



TECHNICAL REPORT 0-7139-1
TxDOT PROJECT NUMBER 0-7139

Determination of Pavement Surface Type

Jorge Prozzi
Christian Sabillon-Orellana
Danilo Inoue
Ruohan Li
Anna Hooton
Robing Huang
Hongbin Xu
Joaquin Hernandez

August 2024

Published March 2025

<https://library.ctr.utexas.edu/ctr-publications/0-7139-1.pdf>



Technical Report Documentation Page

1. Report No. FHWA/TX-25/0-7139		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Determination of Pavement Surface Type			5. Report Date Submitted: December 2021		
			6. Performing Organization Code		
7. Author(s) Christian Sabillon, Danilo Inoue, Ruohan Li, Anna Hooton, Robing Huang, Hongbin Xu, Joaquin Hernandez, Jorge A Prozzi			8. Performing Organization Report No. 0-7139		
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3925 W. Braker Lane, 4th Floor Austin, TX 78759			10. Work Unit No. (TRAIS)		
			11. Contract or Grant No. 0-7139		
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Division P.O. Box 5080 Austin, TX 78763-5080			13. Type of Report and Period Covered Technical Report Month Year – June 2024		
			14. Sponsoring Agency Code		
15. Supplementary Notes Project performed in cooperation with the Texas Department of Transportation.					
16. Abstract This study evaluated and developed machine learning models for pavement surface classification to enhance accuracy and reliability using advanced computational techniques. Initially, a detailed literature review on pavement texture and measurement methods provided a foundation for developing a prototype system. This system, integrating laser scanners, high-speed cameras, and lighting systems, marked a significant advancement in capturing high-resolution texture data at highway speeds. Data collection covered 425.5 miles of pavement across Texas, including 313.7 miles of flexible and 111.8 miles of rigid pavements, resulting in over 50,000 high-resolution images and texture profiles from 15 types of flexible pavements and seven types of rigid pavements. This dataset, essential for training and validating machine learning models, was made available for future research. A hierarchical classification method was developed, with picture-based models (PBC) excelling at predicting flexible or rigid pavements (Level 1) and tining orientation (Level 5), while index-based classifiers (IBCs) performed better for intermediate levels of specificity (Levels 2-4). The best models achieved high F1 scores, confirming their effectiveness. Comprehensive validation using six diverse test sets across 20 test sites confirmed the models' robustness and applicability. These findings significantly enhance pavement management systems, enabling more accurate and efficient maintenance strategies. Leveraging advanced machine learning techniques, the developed models can improve road safety, optimize resource allocation, extend pavement lifespan, and support future research and practical applications.					
17. Key Words Pavement surface type, texture, machine learning, supervised learning, unsupervised learning			18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161; www.ntis.gov .		
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages TBD	22. Price		

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized



THE UNIVERSITY OF TEXAS AT AUSTIN
CENTER FOR TRANSPORTATION RESEARCH

Determination of Pavement Surface Type

Christian Sabillon-Orellana
Danilo Inoue
Ruohan Li
Anna Hooton
Robin Huang
Hongbin Xu
Joaquin Hernandez
Jorge A Prozzi

CTR Technical Report:	0-7139
Report Date:	Submitted: June 2024
Project:	0-7139
Project Title:	Determination of Pavement Surface Type
Sponsoring Agency:	Texas Department of Transportation
Performing Agency:	Center for Transportation Research at The University of Texas at Austin

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration

Center for Transportation Research
The University of Texas at Austin
3925 W. Braker Lane, 4th floor
Austin, TX 78759

<http://ctr.utexas.edu>

Disclaimers

Author's Disclaimer: The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation.

Patent Disclaimer: There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine manufacture, design or composition of matter, or any new useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

Engineering Disclaimer

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES.

Research Supervisor: Jorge A Prozzi

Acknowledgments

The authors want to acknowledge the Texas Department of Transportation for providing support during this project, access to the data and for their constant technical advice. In particular, the authors wish to express their appreciation to Dr. Jenny Li, Arturo Perez, Andy Naranjo, Juthika Golder, Marta Stephens, Martin Sotelo, Melissa Benavides, Rachel Cano, Samuel Mendoza, Sarah Tahmoressi, Travis Patton, Dr. Enad Mahmoud, Dr. Hui Wu, and Dr. Andre Smit for participating in the progress meetings, reviewing the technical memoranda, and providing guidance, valuable comments, and feedback. Furthermore, the authors would like to also recognize the contributions of Lacy Peter, Epigemio Gonzalez, Aldo Madrid, Melody Galland, Ed Goebel, and Dr. Andre Smit for providing the vast majority of the data for this project. Our special gratitude also to Shelley Pridgen, RTI Project Manager, for her constant managerial support during the duration of the project.

Table of Contents

Chapter 1. Literature Review.....	22
1.1. Pavement Texture	22
1.1.1. Surface Texture Components.....	24
1.1.2. Measuring Texture	25
1.1.3. Processing of Texture Data	27
1.1.4. Quantifying Pavement Texture	33
1.2. Pavement Mixes and Surfaces	46
1.2.1. Flexible Pavements: Hot mix asphalt (HMA)	46
1.2.2. Flexible Pavements: Surface Treatments.....	48
1.2.3. Rigid Pavement: Textures Constructed in Plastic Concrete	51
1.2.4. Rigid Pavement: Textures Constructed in Hardened Concrete	54
1.3. Machine Learning	56
1.3.1. Types of Learning	56
1.3.2. Supervised Learning (SL): Classification Models.....	57
1.3.3. Unsupervised Learning	74
1.3.4. Assessing Performance	82
Chapter 2. Prototype Development	88
2.1. Past Methods Used by the Performing Agency for Measuring and Determining Pavement Surface Characteristics.....	88
2.1.1. Equipment from Previous Studies	88
2.1.2. Original Prototype for Collecting 2D Surface Profiles.....	90
2.1.3. Original Prototype for Collecting Surface Images.....	91
2.2. Technologies Considered during Prototype Development	92
2.2.1. Laser Scanners	92
2.2.2. Cameras.....	97
2.2.3. Lighting Systems	100
2.3. Current System's Components and Assembly.....	101
2.4. General Specification for a System with Comparable Performance	104
2.5. Target Degree of Prediction Accuracy and Method of Evaluation.....	106
2.5.1. Target Classification Groups based on new Measured Data	108
Chapter 3. Experimental Design and Data Collection.....	109
3.1. Data Collection	110
3.1.1. Locations of Tested Pavements	110

3.1.2. Locations Tested	111
3.1.3. Data Collection Protocol.....	111
3.1.4. Pavement Surfaces Surveyed.....	112
3.2. Data Collection Results.....	113
Chapter 4. Database of Pavement Surface Types.....	116
4.1. The Database.....	116
4.2. Texture Database	116
4.2.1. 2D Profile Data	117
4.2.2. Texture Indexes.....	121
4.2.3. Final Texture Database	126
4.2.4. Inventory of Pavement Texture 0-7139	134
4.3. Picture Database (Image Data Folder).....	136
4.3.1. All Pictures.....	137
4.3.2. Revised Pictures.....	137
4.3.3. Deleted Pictures	138
4.3.4. Inventory of Pavement Pictures 0-7139.....	139
Chapter 5. Model Development, Training, and Testing.....	143
5.1. Data Processing.....	143
5.1.1. Implemented Algorithm.....	143
5.1.2. Computed Indices	145
5.2. Picture Processing.....	147
5.2.1. Quality Control	147
5.3. Machine Learning Implementation.....	148
5.3.1. Index-Based Classification (IBC) Models	149
5.3.2. SL Pictures-Based Classification.....	150
5.4. Results.....	150
5.4.1. Cluster Analysis	150
5.4.2. IBC Models Comparison	163
5.4.3. IBC Classification.....	168
5.4.4. PBC Classification	179
Chapter 6. Critical System Evaluation and Validation.....	186
6.1. Validation data.....	186
6.1.1. Test Set #1 (TS1): Unknown Surfaces.....	187
6.1.2. Test Set #2 (TS2): Previous Classified Pavements by Experts - TxDOT Project No. 0-7031	189

6.1.3. Test Set #3 (TS3): Repeated Measurements	190
6.1.4. Test Set #4 (TS4): Opposite Direction of Traffic	192
6.1.5. Test Set #5 (TS5): Different Lane	193
6.1.6. Test Set #6 (TS6): Different Testing Day	195
6.2. Results.....	195
6.2.1. Test Set #1 (TS1): Unknown Surfaces.....	198
6.2.2. Test Set #2 (TS2): Previous Classified Pavements by Experts - TxDOT Project No. 0-7031	201
6.2.3. Test Set #3 (TS3): Repeated Measurements	202
6.2.4. Test Set #4 (TS4): Opposite Direction of Traffic	205
6.2.5. Test Set #5 (TS5): Different Lane	208
6.2.6. Test Set #6 (TS6): Different Testing Day	210
Chapter 7. Conclusions and Recommendations	212
7.1. Conclusions.....	212
7.2. Recommendations.....	214
Appendix A: Value of Research (VoR)	216
Review of Potential Savings	216
Cost Savings in Maintenance.....	216
Reducing Accident-Related Costs	216
NPV Cost Benefit analysis.....	217
Inputs for the Economic Analysis.....	218
Conclusion	219

List of Figures

Figure 1: Basic terminology: wavelength (1), amplitude (2).....	22
Figure 2: 3D pavement surface recreated from 2D surface profiles (top), 2D transverse pavement profile (bottom), both collected with a laser scanner.	24
Figure 3: Definition of fundamental texture classes, as a function of the wavelengths or spatial frequency (top), and as a function of vertical amplitude (bottom).....	24
Figure 4: Texture wavelength influence on pavement-tire interaction, modified from (19).....	25
Figure 5: Stationary texture measuring devices: (A) SPT, (B) CTM, (C) Laser Texture Scanner, (D) Stationary Linear Profilometer, (E) LS-40, (F) Line Laser Scanner, and (G) multi-line laser sensor attached to the right back side of truck.	27
Figure 6: Effect of measurement environment on height sensors (35).....	28
Figure 7: Common instances of noise extreme spikes (top left), flat signal (top right), mild spike (bottom) (36).....	29
Figure 8: Detrending techniques: regression (left) and integration (right) detrending.	30
Figure 9: 2D pavement surface profile data processing schematic.	32
Figure 10: Complete description of a tribological system (73).	40
Figure 11: Schematic representation of a tribological system (74).	41
Figure 12: Schematic illustration of the form deviations of the first to fourth order (74).	41
Figure 13: Diagrams portraying the computation of mean rough depth (Rz, on the left) and mean roughness value (Ra, on the right) (74)	42
Figure 14: The Moody diagram (82).	43
Figure 15: Definition of the roughness index as defined by (86).	44
Figure 16: Diagram of an ideal pavement surface (102) and a field picture for a dense-graded mix.....	47
Figure 17: Diagram of an ideal pavement surface (102) and a field picture for an open-graded mix	47
Figure 18: Diagram of an ideal pavement surface (102) and a field picture for a gap-graded mix.....	48
Figure 19: Representative diagram (113) and field picture for a single layer seal coat.....	49

Figure 20: Representative diagram (118) and field picture for a high friction surface treatment.	50
Figure 21: Representative diagram (119) and field picture for a thin overlay mix.	50
Figure 22: Representative diagram (120) and field picture for a bonded wearing coarse surface.	51
Figure-23: Field pictures for textures constructed in plastic PCC dragged surfaces burlap (left), artificial turf (right).	52
Figure 24: Field pictures for textures constructed in plastic PCC tined surfaces fixed transverse (left) and random longitudinal tining (right). Direction of travel shown by arrow.	53
Figure 25: Field pictures for textures constructed in plastic PCC exposed aggregate surface.	54
Figure 26: Field picture for texture constructed in hardened PCC using conventional diamond grinding	54
Figure 27: Field picture for texture constructed in hardened PCC using conventional diamond grooving	55
Figure 28: Field picture for texture constructed in hardened PCC using NGCS.	56
Figure 29: Diagram summarizing the traditional machine learning algorithm which can be used to predict pavement surface type based on texture indices.	59
Figure 30: Example of the KNN classification algorithm.	60
Figure 31: Maximum-margin hyperplane and margins for an SVM trained with samples from two classes (136).	62
Figure 32: Basic structure of a DT.	66
Figure 33: Basic structure of a RF.	67
Figure 34: Basic structure of a BA model.	68
Figure 35: Visual depiction of AB using multiple weak learners to create one strong classifier.	69
Figure 36: The architecture of GB (148).	70
Figure 37: Flowchart of XGB algorithm (150).	71
Figure 38: Representation of an ANN Model with 2 hidden layers (151).	72
Figure 39: An input image, followed by a convolutional layer and the output of the convolutional layer (151).	74
Figure 40: A CNN architecture to classify handwritten numbers (153).	74

Figure 41: Comparison between scatterplot of the original data (top) and the reduced data using the first three principal components (bottom).....	76
Figure 42: Density plot overlaid by scatter point of the first and second principal components for texture indices in rigid pavements.	77
Figure 43: Clusters with homogeneity and/or separation (154)	78
Figure 44: Clustering of clustering techniques	79
Figure 45: Sample dendrogram.....	81
Figure 46: Example of a confusion matrix.	83
Figure 47: The ROC space for a “better” and “worse” classifier (165).	87
Figure 48: Original data collection prototype to measure pavement surface profiles at high speed and predict pavement surface.	89
Figure 49: New data collection prototype to measure pavement surface profiles at high speed and predict pavement surface.	90
Figure 50: (A) Prototype magnetically mounted to the passenger side of the vehicle. (B) Close-up of the camera in the middle and two spot-lights focusing the light at the center area of where the images are being captured.	91
Figure 51: AP820 2D laser scanner series (168)	93
Figure 52: 8300 High Speed Inertial Profiler (29) on the left, and Ames AccuTexture 100 single point laser (173) on the right	94
Figure 53: Pathrunner (176) on the top, Macrottexture sensors (177) on the bottom right, and scan obtained from 3-zone texture signature subsystem on the bottom left (177).	95
Figure 54: Images from the LJ-X8000 series sensors from the provider (185) on the top, and images of the sensors mounted on stage for use in laboratory measurements (180) on the bottom.	96
Figure 55: High power laser sensor by itself from a bottom side view (186), shown on the top and the laser sensor mounted on a pushcart and a vehicle to for field use (186) on the bottom.	97
Figure 56: Front and back view of the 3D Machine Vision Ranger3 camera (187).....	98
Figure 57: Front and back view of the Canon 5D Mark IV camera (191)	98
Figure 58: Front and side view of the Arducam 2MP Global shutter OV2311 camera (195).....	99
Figure 59: Front and back view of the FLIR Blackfly S camera (200).....	100
Figure 60: Part (A), camera and lighting system magnetically mounted to the passenger side of the vehicle. Part (B), close-up of the camera in the	

middle and two spot-lights focusing the light at the center area of where the images are being captured.....	100
Figure 61: TX Barlight with a 3-sided T-slot extrusion.	101
Figure 62: Side and rear view of the mounting system.	102
Figure 63: System mounted onto vehicle shown on the top; line laser sensor shown on the bottom-right; and camera and light system (perpendicular to the main beam), shown on the bottom left.....	103
Figure 64: Diagram to wire the camera and light source.....	104
Figure 65: Map showing the location of every pavement section surveyed (54).....	107
Figure 66: Testing sites within TxDOT districts (left), Texas counties (middle) with tested sites in blue, and exact site locations marked by blue circles (right).....	111
Figure 67: Distribution of mileage surveyed broken down by TxDOT district.	114
Figure 68: Distribution of mileage surveyed broken down by pavement type.....	115
Figure 69: Distribution of number of pictures taken broken down by pavement type.	115
Figure 70: Screenshot of the “TxDOT Project 0-7139 Surface Determination” folder view	116
Figure 71: Screenshot of the “(Profiles from Laser)” folder view	116
Figure 72: Screenshot of the “1. 2D Profile Data” folder view	117
Figure 73: Screenshot of the “1. Raw Data (.tsd files)” folder view	117
Figure 74: Screenshot of the “L – Laredo District” subfolder view within the “1.Raw Data (.tsd files)” view, showing the nomenclature used for file naming.....	118
Figure 75: Screenshot of the “2. Raw Profiles (.csv files)” view, showing the that folders have been compressed by district.	120
Figure 76: Screenshot of the “Texture Indexes” folder view, showing the that folders have been compressed by district.	121
Figure 77: Screenshot of the “Per Profile (.csv files)” folder view	122
Figure 78: Screenshot of the “O - Odessa” subfolder view within the “Per Profile (.csv files)” folder.	123
Figure 79: Screenshot from one of the CSV files within the “Per Profile (.csv files)” folder showing all the variables contained within it.	123
Figure 80: Screenshot of the “Aggregated (.csv files)” folder view.....	125

Figure 81: Screenshot of the “A - Lubbock” folder view inside the “Aggregated (.csv files)” folder.....	125
Figure 82: Screenshot the “3. Final Database” folder view.....	127
Figure 83: Diagram indicating the different levels of specificity for the pavements surveyed based on their surface type.	128
Figure 84: Screenshot of the “Summary” tab within the “Inventory of Pavement Texture 0-7139” file.....	135
Figure 85: Screenshot of the “Austin” tab within the “Inventory of Pavement Texture 0-7139” file. This is a representative view of all the other district tabs.	135
Figure 86: Screenshot of the “Image Data (Camera Images)” folder view, showing the that folders have been compressed by district.	136
Figure 87: Screenshot of the “Image Data (Camera Images)” folder view, showing the that folders have been compressed by district.	137
Figure 88: Screenshot of the “Image Data (Camera Images)” folder view, showing the that folders have been compressed by district.	138
Figure 89: Sample images showing what kind of criteria would be considered enough to delete an image from the database: unexpected debris, objects or animals such as a critter’s tail (top left), pavement markings (top right), cracking which can be confused as surface features (bottom left), interface of a pavement change (bottom right).	139
Figure 90: Screenshot of the “3. Deleted Pictures” folder view.....	139
Figure 91: Screenshot of the “Grand Summary” tab within the “Inventory of Pavement Pictures 0-7139” file.	140
Figure 92: Screenshot of the “TxDOT (Ground Truth)” tab within the “Inventory of Pavement Pictures 0-7139” file. This is also a representative view for the “CTR (Extra)” tab. Note that many columns are not shown in this image.	141
Figure 93: Schematic illustrating a detailed description of the data processing algorithm.	143
Figure 94: Example of the SOFA on 2D texture profile for a pavement with new generation diamond grinded surface finish.	144
Figure 95: Pearson correlation coefficient matrix for the final selection of five indices used in flexible pavement clustering.	151
Figure 96: Density plot of the first two principal components for flexible pavements.	152
Figure 97: Dendrogram for flexible pavements with colored rectangles indicating the clustering structure for five groups in the data.	152

Figure 98: Visualization of the five clusters derived from a dendrogram on the scatterplot of the first two principal components.....	153
Figure 99: Representative pictures of each of the five cluster. Members of the same clusters are shown along the row.....	155
Figure 100: Pearson correlation coefficient matrix for the final selection of five indices used in rigid pavement clustering.....	156
Figure 101: Density plot of the first two principal components for flexible pavements highlighting the two clusters that are easily separable by visual inspection.	157
Figure 102: Representative pictures of each of the first two clusters. Members of the same clusters are shown along the row.	157
Figure 103: Visualization of the clustering output from the four clustering algorithms as specified in Table 1. Note that for DBSCAN cluster 0 are noisy data points not belonging to any cluster.....	158
Figure 104: Density plot of the first two principal components for flexible pavements highlighting the clustering structure within the residual cluster.....	159
Figure 105: Representative pictures of each of the five cluster within the residual cluster. Members of the same clusters are shown along the row.	161
Figure 106: Visualization of the clustering output from the four clustering algorithms as specified in Table 16. Note that for DBSCAN cluster 0 are noisy data points not belonging to any cluster.....	162
Figure 107: Summary of F1 scores across all 14 ML models for classification pavements using Level #1 specificity.....	170
Figure 108: Summary of F1 scores across twelve ML models for classification pavements using Level #2 specificity.....	173
Figure 109: Summary of F1 scores across twelve ML models for classification pavements using Level #3 specificity.....	174
Figure 110: Summary of F1 scores across twelve ML models for classification pavements using Level #4 specificity.....	177
Figure 111: Reference map of TS1 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	188
Figure 112:. Reference map of TS1 section #2 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	188
Figure 113: Reference map of TS1 section #3 in Brown County. (Left) Section on Google Maps, and (right) section on TRM map.	188
Figure 114: Reference map of TS1 section #4 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	189
Figure 115: Reference map of TS1 section #5 in Hill County. (Left) Section on Google Maps, and (right) section on TRM map.	189

Figure 116: Reference map of TS2 section #1 in Williamson County. (Left) Section on Google Maps, and (right) section on TRM map.	189
Figure 117: Reference map of TS2 section #2 in Williamson County. (Left) Section on Google Maps, and (right) section on TRM map.	190
Figure 118: Reference map of TS2 section #3 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	190
Figure 119: Reference map of TS3 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	191
Figure 120: Reference map of TS3 section #2 in La Salle County. (Left) Section on Google Maps, and (right) section on TRM map.	191
Figure 121: Reference map of TS3 section #3 in Maverick County. (Left) Section on Google Maps, and (right) section on TRM map.	191
Figure 122: Reference map of TS3 section #4 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	192
Figure 123: Reference map of TS4 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	192
Figure 124: Reference map of TS4 section #2 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	193
Figure 125: Reference map of TS4 section #3 in Webb County. (Left) Section on Google Maps, and (right) section on TRM map.	193
Figure 126: Reference map of TS5 section #1 in La Salle County. (Left) Section on Google Maps, and (right) section on TRM map.	194
Figure 127: Reference map of TS5 section #2 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	194
Figure 128: Reference map of TS5 section #3 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	194
Figure 129: Reference map of TS6 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.	195
Figure 130: Reference map of TS6 section #2 in Hays County. (Left) Section on Google Maps, and (right) section on TRM map.	195
Figure 131: Pavement surface at TS1 Section #1.	198
Figure 132: Pavement surface at TS1 Section #2.	199
Figure 133: Pavement surface at TS1 Section #3.	199
Figure 134: Pavement surface at TS1 Section #4.	200
Figure 135: Pavement surface at TS1 Section #5.	200
Figure 136: Pavement surface at TS2 Section #1 (3).	201
Figure 137: Pavement surface at TS2 Section #2 (3).	202

Figure 138: Pavement surface at TS2 Section #3 (3).	202
Figure 139: Pavement surface at TS3 section #1.	203
Figure 140: Pavement surface at TS3 Section #2.	204
Figure 141: Pavement surface at TS3 Section #3.	204
Figure 142: Pavement surface at TS3 Section #4.	205
Figure 143: Pavement surfaces across both directions of traffic: (top) southbound approach, (bottom) northbound approach in TS4 Section #1.	206
Figure 144: Pavement surfaces across both directions of traffic: (top) southeast approach, (bottom) northwest approach in TS4 Section #2.	207
Figure 145: Pavement surfaces seen across both directions of traffic: (top) southeast approach, (bottom) northwest approach in TS4 Section #3.	208
Figure 146: Pavement surface at TS5 Section #1.	209
Figure 147: Pavement surface at TS5 Section #2.	209
Figure 148: Pavement surface at TS5 Section #3.	210
Figure 149: Pavement surface at TS6 Section #1.	211
Figure 150: Pavement surface at TS6 Section #2.	211
Figure 151: Inputs and outputs of NPV cost-benefit analysis (top), graphical representation of NPV over the course of ten years (bottom).	219

List of Tables

Table 1: List of indexes reviewed and their respective source.	35
Table 2: Aggregate gradation requirements for seal coats in terms of cumulative percent retained at each sieve (114).	49
Table 3. General Specifications for the sensor used in the data collection prototype	105
Table 4: Classification Report Obtained from Decision Tree Model (54)	107
Table 5: TxDOT Districts and personnel who provided the most support and the locations of pavement surfaces to the Performing Agency	110
Table 6: TxDOT Districts and personnel who provided the additional support and the locations of more pavement surfaces to the Performing Agency	111
Table 7: Summary of all the pavement surfaces provided by TxDOT and surveyed by the Performing Agency	113
Table 8: Breakdown of mileage for data collected in terms of pavement type.	113

Table 9: Breakdown of mileage for data collected in terms of dataset.....	113
Table 10: List of letters used to designate TxDOT's district.....	118
Table 11: Testing sites for which the .tsd files are not available.....	119
Table 12: List of indices used in the study	145
Table 13: List of IBC models.....	149
Table 14: Summary of flexible pavement cluster analysis with description of the cluster and median index measurements of the member pavements	154
Table 15: Summary of algorithm specification which enables the identification of clusters RPC1 and RPC2.....	159
Table 16: Summary of algorithm specification which enables the identification of clusters RPC3 through RPC7	160
Table 17: Summary of rigid pavement cluster analysis with description of the cluster and median index measurements of the member pavements	163
Table 18: Summary of Model Sensitivities	165
Table 19: Summary of Model Statistical Assumptions	166
Table 20: Summary of Model Computational Concerns	167
Table 21: Summary of Model Implementation, Interpretability, Transparency and Adaptability	167
Table 22: Confusion matrix for the XGB model in Level #1 specificity	172
Table 23: Classification report for the XGB model in Level #1 specificity	172
Table 24: Confusion matrix for the SNN model in Level #2 specificity	173
Table 25: Classification report for the SNN model in Level #2 specificity	174
Table 26: Confusion matrix for the SVM model in Level #3 specificity	175
Table 27: Classification report for the SVM model in Level #3 specificity	176
Table 28: Classification report for the KNN model in Level #4 specificity	178
Table 29: Confusion matrix for the CNN in Level #1 specificity	180
Table 30: Classification report for the CNN model in Level #1 specificity	180
Table 31: Confusion matrix for the CNN model in Level #2 specificity	181
Table 32: Classification report for the CNN model in Level #2 specificity	181
Table 33: Confusion matrix for the CNN model in Level #3 specificity	182
Table 34: Classification report for the CNN model in Level #3 specificity	183
Table 35: Classification report for the KNN model in Level #4 specificity	184
Table 36: Confusion matrix for the CNN in Level #5 specificity	185
Table 37: Classification report for the CNN model in Level #5 specificity	185

Table 38: Summary of Test Sets and Characteristics	187
Table 39 Abbreviated Nomenclature for Surface Prediction.....	197
Table 40. Prediction made for pavement surfaces in TS1	198
Table 41. Predictions made for pavement surfaces in TS2.....	201
Table 42. Predictions made for pavement surfaces in TS3.....	203
Table 43. Predictions made for pavement surfaces in TS4.....	205
Table 44. Predictions made for pavement surfaces in TS5.....	209
Table 45. Predictions made for pavement surfaces in TS6.....	210

List of Terms

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
A	Accuracy or Astroturf
AB	Adaptive Boosting
AGNES	Agglomerative Nesting
AI	Angularity Index
ANN	Artificial Neural Network
AOI	Area of Interest
AR	Aspect Ratio
ASTM	American Society of Testing Materials
AUC	Area under the Curve
BA	Bootstrap Aggregating
BI	Business Interstate
BOM	Open Mix with the asphalt Binder exudated to the near surface
BPV	British Pendulum Number
BPT	British Pendulum Tester
BWC	Bonded Wearing Coarse
CLL	Complimentary Log-Log Regression
CM	Mean of Cross Width
CNN	Convolutional Neural Network
CS	Chip Seal or Skewness of Cross Width
CSTD	Standard Deviation of Cross Width
CSV	Comma Separated Value
CTM	Circular Track Meter
CTR	Center for Transportation Research
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DG	Diamond Grinding
DMS	Dense Mix Surfaces
DNN	Deep Neural Network
DOT	Department of Transportation
DS1	Ground Truth Dataset
DS2	Holdout Dataset
DT	Decision Tree
FD	Fractal Dimension
FF	Form Factor
FM	Farm to Market
FMI	Fowlkes–Mallows Index
FN	False Negative
FP	False Positive or Floor Distance to Mean Line
FPR	False Positive Rate
FPC	Flexible Pavement Cluster
FR	Feature Roughness
FTT	Fixed Transverse Tining

GB	Gradient Boosting
GLM	Generalized Linear Models
GMM	Gaussian Mixture Models
GNB	Gaussian Naïve Bayes
HCME	Mean of Local Curvature
HCST	Standard Deviation of Local Curvature
HFST	High Friction Surface Treatment
HHT	Hilbert-Huang Transform
HMA	Hot Mix Asphalt
HWMA	Hot and Warm Mix Asphalt
IBC	Index-Based Classification
IH	Interstate Highway
ISO	International Organization for Standardization
JPEG	Joint Photographic Experts Group
KNN	K Nearest Neighbors
LOG	Logistic Regression
MCC	Matthew's Correlation Coefficient
MPD	Mean Profile Depth
MTD	Mean Texture Depth
NASA	National Aeronautics and Space Administration
NGCS	New Generation Concrete Surface
NGDG	New Generation Diamond Grinding
P	Precision
PAM	Partition Around the Medoids
PB	Precoated Crushed Gravel, ..., or limestone rock asphal seal coatt
PBC	Picture Based Classification
PCC	Portland Cement Concrete
PCA	Principal Component Analysis
PFC	Porous Friction Coarse
PRO	Probit Regression
PSD	Power Spectral Density
R	Recall
RA	Mathematical Average Roughness
RAL	Autocorrelation Length
RDT	Maximum Absolute Gradient
RF	Random Forest
RGB	Red Green Blue
RKU	Kurtosis of Height
RLT	Random Longitudinal Tining
RMS	Root Mean Square
ROC	Receiver Operating Curve
RP	Maximum Profile Peak Height
RPC	Rigid Pavement Cluster
RS	Solidity Factor
RSM	Mean Spacing of Profile Irregularities
RTT	Random Transverse Tining

SC	Seal coat of unspecified grade
SC G3	Grade 3 seal coat
SC G3S	Grade 3 seal coat with single size gradation
SC G4	Grade 4 seal coat
SH	State Highway
SL	Supervised Learning
SMA	Stone Matrix Asphalt
SNN	Shallow Neural Network
SOFA	Sabillon-Orellana Filtering Algorithm
SP	SuperPave
SP-C	SuperPave Type C
SP-D	SuperPave Type D
SPT	Sand Patch Test
SV2	Two-Point Slope Variance
SVM	Support Vector Machine
TBWC	Thin Bonded Wearing Coarse
TOM	Thin Overlay Mix
TP	True Positive
TPR	True Positive Rate
TN	True Negative
TRM	Texas Reference Marker
TS#	Test Section #
TxDOT	Texas Department of Transportation
UL	Unsupervised Learning
US	United States Highway
WT	Wavelet Transform
XBG	Extreme Gradient Boosting

Chapter 1. Literature Review

1.1. Pavement Texture

Pavement texture refers to the irregularities on a surface that deviate from an ideally flat pavement. It is a critical characteristic as it influences tire-pavement interactions, affecting noise, friction, rolling resistance, water splash, tire and vehicle wear, and ride quality (1). However, specialized equipment and mathematical tools are necessary for accurate characterization due to its complex nature.

The literature indicates that pavement texture can be represented using two primary approaches: the two-dimensional (2D) approach with linear profiles, and the three-dimensional (3D) approach with pavement surfaces. A linear profile, which is the simplest representation of pavement texture, is a 2D depiction obtained using sensor devices. It is described by two coordinates: distance (measured either longitudinally, along the traffic direction, or transversely, orthogonal to the traffic direction) and height (measured vertically) (2).

These profiles are considered stationary, random functions of the distance along the surface (3). Mathematically, they can be represented as a series of sine and cosine waves with various amplitudes and wavelengths using Fourier analysis. Wavelength, the spatial period of a wave, can be reported in units of length or as spatial frequency (fs), which is the inverse of the wavelength and expressed in cycles per meter. The amplitude is defined as the peak-to-peak height difference (2). **Figure 1** illustrates the main parameters on a surface profile.

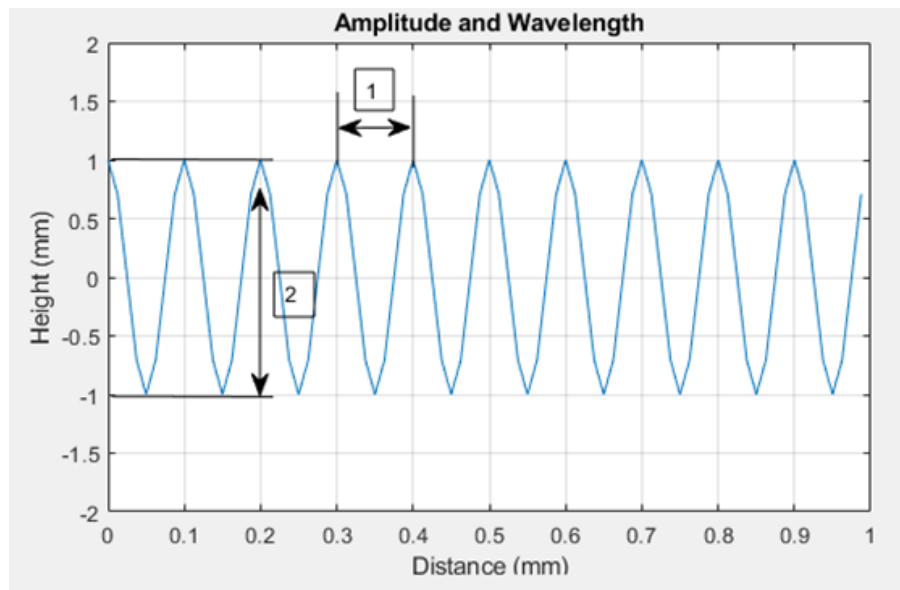


Figure 1: Basic terminology: wavelength (1), amplitude (2).

Pavement texture can also be characterized by using 3D surfaces. This can be achieved by “stitching” together multiple contiguous 2D profiles (4,5), capturing an area image using at least two angled cameras positioned opposite each other (e.g., stereovision systems) (6,7), or combining data from various sensors such as LiDAR, cameras, and lasers to create a comprehensive, high-resolution 3D surface of the pavement (8). These 3D texture surfaces are powerful tools since the data can also be used for identifying and quantifying pavement distress. Moreover, 3D texture summary indices might be more suitable for correlating texture with pavement surface interactions, as they summarize an area of pavement rather than a discrete spot.

Despite the advantages, 3D measurements come with challenges. Wang (7) points out several drawbacks of stereovision systems that are comparable to other 3D data collection systems. First, practical image processing requires very high-performance computing equipment. Using lower-performance equipment compromises data quality, processing speed, or both. Second, image processing is an evolving field, and researchers may encounter unresolved issues while processing data. Third, distinguishing between surface texture features and foreign objects on the pavement, such as oil spills or debris, can be particularly challenging. Lastly, hardware and software compatibility across different vendors’ systems is not necessary but introduces noncomparable survey data from different vendors (7).

Figure 2 illustrates a sample 3D and 2D representation of pavement texture.

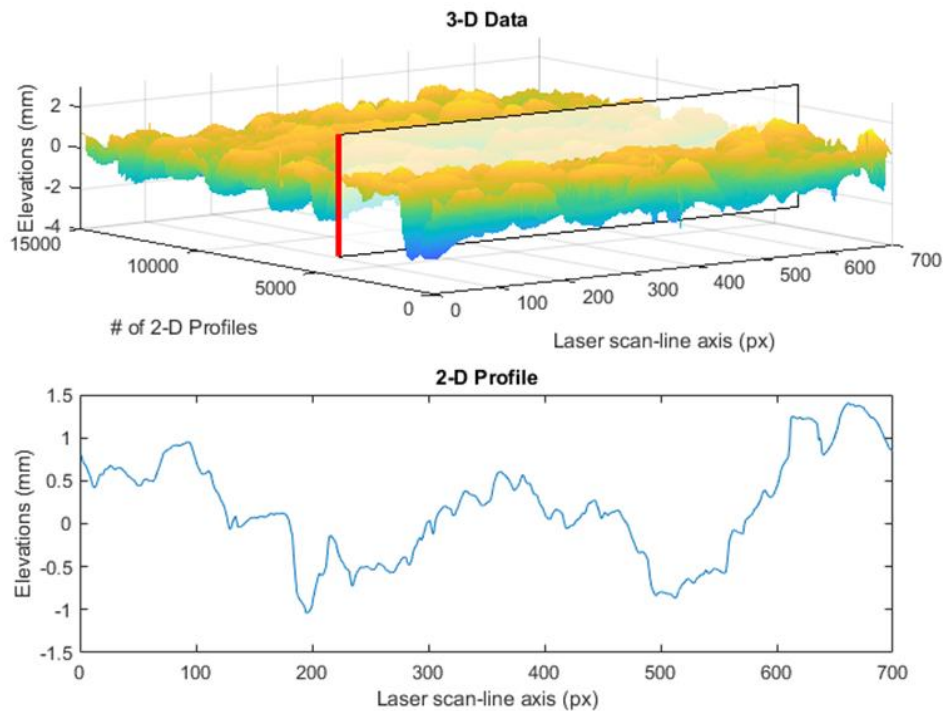


Figure 2: 3D pavement surface recreated from 2D surface profiles (top), 2D transverse pavement profile (bottom), both collected with a laser scanner.

1.1.1. Surface Texture Components

In 1987, the World Road Association, formerly known as the Permanent International Association of Road Congresses (PIARC), established four categories of pavement surface irregularities: unevenness/roughness, megatexture, macrotexture, and microtexture. Each category is defined by the texture's wavelengths or spatial frequency and its vertical amplitude (9). **Figure 3** illustrates the surface texture spectrum, showing the four main texture components along with their respective wavelength and vertical amplitude domains.

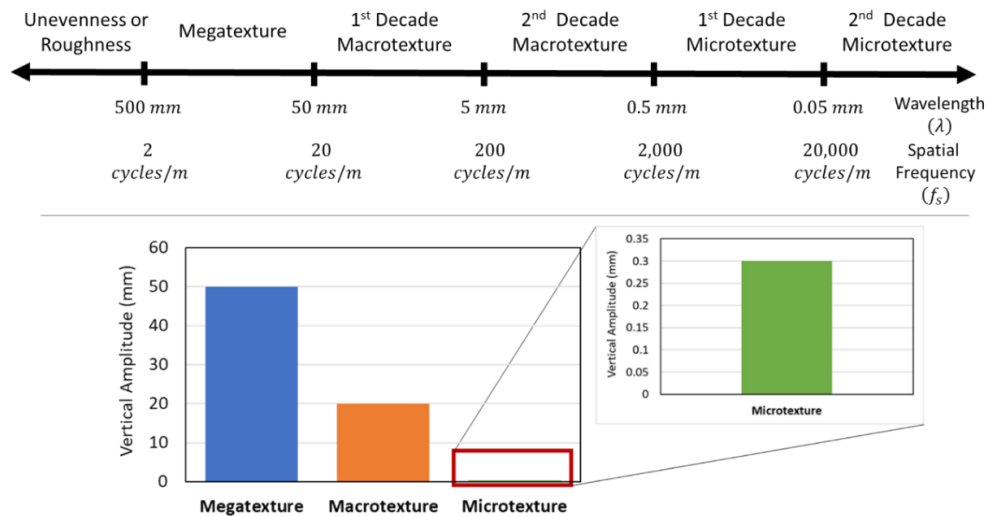


Figure 3: Definition of fundamental texture classes, as a function of the wavelengths or spatial frequency (top), and as a function of vertical amplitude (bottom).

Unevenness or roughness describes the pavement surface irregularities that affect ride quality, smoothness, and serviceability. The reference length for unevenness is equivalent to a short stretch of road (10). These irregularities impact rolling resistance, vehicle dynamics, ride quality, surface drainage, fuel consumption, and maintenance costs (11).

Megatexture refers to the distress, defects, or waviness of the road surface, with wavelengths similar in size to the tire-pavement interface (10). This type of texture is typically the most noticeable to the naked eye, including features like ruts, potholes, major joints, and cracks. Megatexture results from poor construction practices, local settlements, or surface deterioration, adversely affecting ride quality, causing premature vehicle suspension wear, and increasing in-vehicle and external noise due to tire wall vibrations (12).

Macrotexture pertains to the large-scale texture of the pavement surface, influenced by the arrangement and size of aggregates (13). In flexible pavements, the properties of the mixture, such as aggregate shape, size, and gradation, control macrotexture. In rigid pavements, macrotexture is determined by the method of surface finishing, including techniques like dragging, tinning, and grooving (2). Macrotexture plays a crucial role in drainage and wet weather friction characteristics, especially at high speeds, making it essential for pavements with operational speeds of 50 mph or higher to have good macrotexture to prevent hydroplaning (14). Additionally, macrotexture significantly impacts tire-pavement noise (15) and splash and spray (16).

Microtexture refers to the microscopic asperities of the aggregate surface, which control the contact between tire rubber and the pavement surface. Microtexture depends on the mineralogy and petrology of individual aggregate particles, the aggregate source (natural or manufactured), and is influenced by environmental effects and traffic action (17). Microtexture is usually sufficient to provide adequate skid resistance on dry pavements at speeds of 50 mph or higher and on wet, but not flooded, pavements at speeds below 50 mph (14). However, excessive microtexture can be detrimental, as tire wear rates are directly proportional to microtexture levels (18).

Figure 4, inspired by PIARC's work (19), summarizes the significant tire-pavement interactions and which texture components affect them the most.

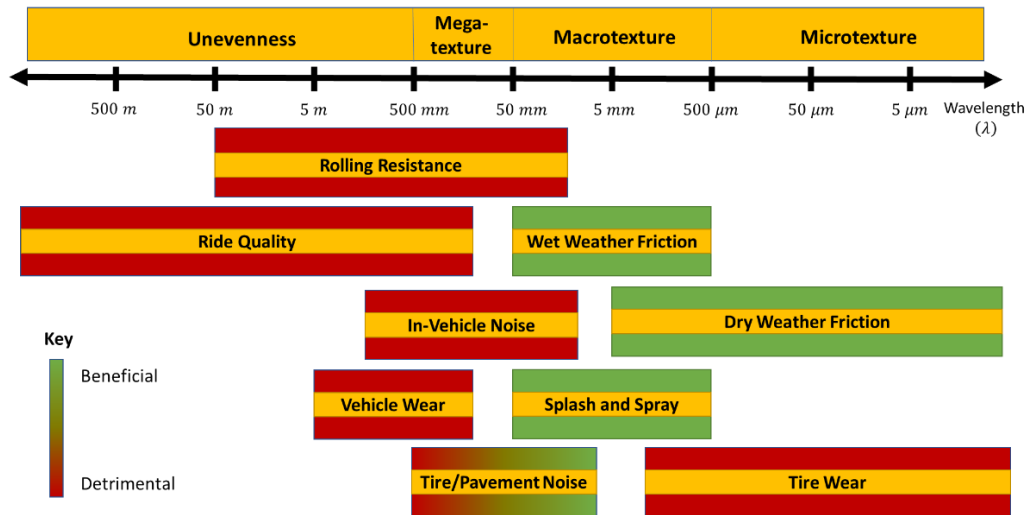


Figure 4: Texture wavelength influence on pavement-tire interaction, modified from (19).

1.1.2. Measuring Texture

Volumetric techniques, such as the outflow meter (20,21) the NASA Grease Smear Method (22), and the Sand Patch Test method (SPT) (21,23), have been used to

evaluate pavement macrotexture in terms of its mean texture depth (MTD). While these tests ingeniously leverage geometry to quantify pavement texture, they have been criticized for their inability to capture microtexture, low repeatability, and high variance across different operators (2,17,22).

Non-contact measurement devices, like the Circular Track Meter (CTM) (24), do not suffer from these limitations. Mean Profile Depth (MPD) measurements from the CTM are consistent across operators at any given location. However, the CTM, and the SPT as well, reduce texture to a single attribute that characterizes only the vertical amplitude of pavement texture at discrete locations.

To overcome these drawbacks, modern approaches prefer the use of sensors that capture a larger pavement area, allowing users to process raw data and compute multiple spatial and spectral indices to summarize the texture of the pavement in terms of its amplitude, spacing, slope, and frequency characteristics, not just MTD or MPD (2). Some of the most used devices in this category include the Ames Laser Texture Scanner 9500 (25), the LS-40 Portable 3D Surface Analyzer (26), the Stationary Laser Profilometer (27), and the line laser scanner (13). As of June 2024, laser sensor technology allows the best static sensors on the market to characterize the full spectrum of macrotexture and the first and second decades of microtexture with high accuracy. However, these devices are stationary and require traffic control for safe field use. Additionally, some devices need shaded conditions during daytime to prevent sunlight from interfering with laser measurements.

To dynamically collect pavement macrotexture data without traffic control, vehicle-mounted laser devices were developed. Laser-based texture profilometry uses technology comparable to road profilers that evaluate pavement smoothness (28). Initially, single-point lasers rated at 1 kHz mounted on vehicles were used to capture 2D longitudinal profiles (29). These vehicles, known as inertial profilers, are not typically used to characterize macro and microtexture. Instead, they are employed to measure pavement roughness. However, since pavement texture is anisotropic, longitudinal profiles do not fully describe all texture characteristics (28). Consequently, a different setup using a line laser sensor rated at 5 kHz was implemented to capture multiple transverse pavement profiles as the vehicle drives (30). This setup is better than point lasers as it captures a 3D image of the pavement, but the resolution in longitudinal and transverse directions differs significantly, and sensors struggle to capture surface finishes on rigid pavements parallel to the laser line direction.

Alternatively, stereovision cameras can reconstruct a 3D pavement surface and avoid some issues associated with laser sensors. However, stereovision systems are best used for low-speed testing as they suffer from blurring due to long exposure

times at high speeds (28). These three configurations are the most widely used for 2D/3D continuous texture measurements. Current research has made slight modifications to these configurations to address certain issues. For instance, Fernando et al. (31) demonstrated that aligning the sensor's laser line at a 45° angle, rather than parallel or orthogonal to the travel direction, is more effective for detecting joints and cracks on rigid pavements.

Figure 5 shows commonly used stationary texture measuring equipment (top) and a laser sensor attached to the back of a truck for continuous texture measurements (bottom). As of June 2024, technology allows for continuous high-speed texture measurements that cover unevenness, megatexture, and macrotexture components, but not microtexture. Some of the best sensors on the market can capture up to 0.186 mm in wavelength during daylight, although such measurements may be noisy at maximum resolution. It is anticipated that in the future, sensors will be powerful enough to capture all four major texture components at network speeds.

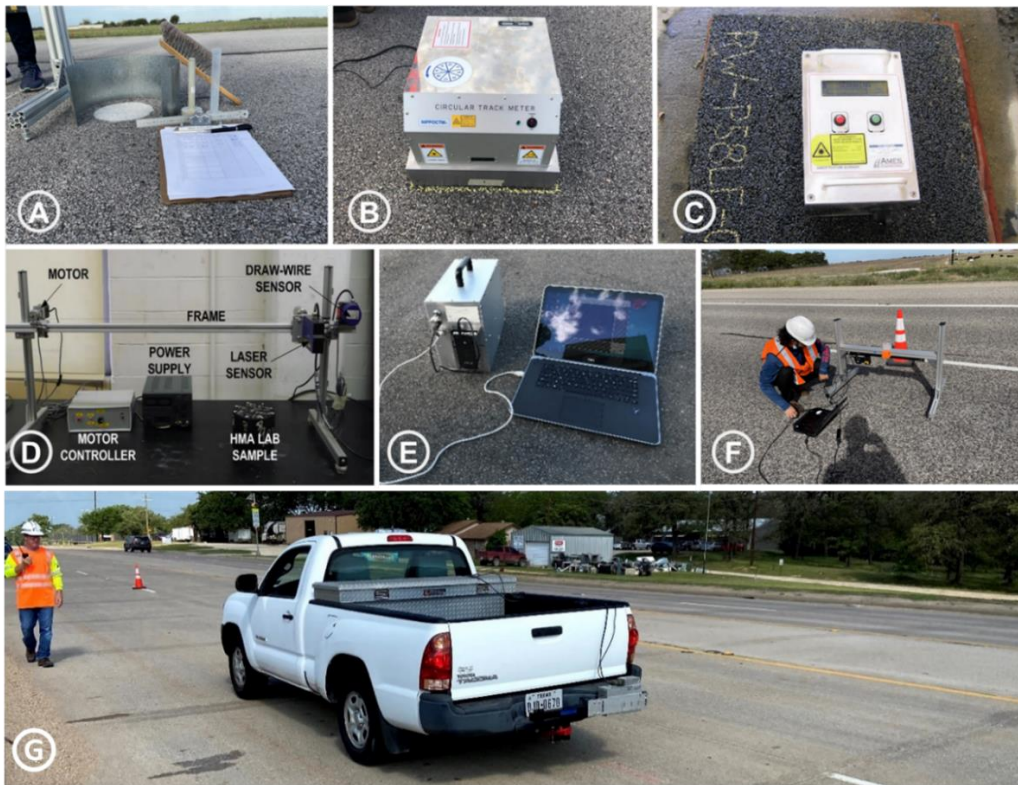


Figure 5: Stationary texture measuring devices: (A) SPT, (B) CTM, (C) Laser Texture Scanner, (D) Stationary Linear Profilometer, (E) LS-40, (F) Line Laser Scanner, and (G) multi-line laser sensor attached to the right back side of truck.

1.1.3. Processing of Texture Data

Despite advancements in data collection equipment, a standardized process for handling these measurements has not yet been established, necessitating further

progress in this area (32). There is a critical need for a standardized methodology to process texture data, as calculating unbiased texture indices requires that the data (i.e., surface profiles or texture surfaces) meet the following criteria: 1) be noise-free, 2) be nearly stationary, and 3) be devoid of missing values.

The presence of white noise, spikes, or flat signals can skew texture indices. Studies like Katicha et al (33), indicate that measurements from vehicle-mounted laser sensors may exhibit spikes, affecting the accuracy of any indices computed from this raw data. This issue is relevant not only at the macrotexture level but also at the unevenness/roughness and megatexture levels, as shown by Sayers and Karamihas (34) in their investigation of spikes' adverse effects on international roughness index (IRI) measurements.

In their study on pavement smoothness, Perera and Kohn (35) identified challenges associated with spikes in profile data collected by Long-Term Pavement Performance profilers. They found that environmental factors influencing height measurements caused biases similar to calibration errors, resulting in spikes in the recorded profiles, as summarized in **Figure 6**.

Factor	Ultrasonic	Laser	Infrared	Optical
Wind	●	○	○	○
Temperature	●	●	●	●
Humidity	○	○	○	○
Surface Moisture	●	●	●	●
Surface Contaminants	●	●	●	●
Pavement Markings	○	○	○	●
Pavement Color	○	○	○	—
Ambient Light	○	○	○	●

● - Strong Effect ● - Effect Under Unusual Circumstances
 ○ - Small or No Effect — - Insufficient Information

Figure 6: Effect of measurement environment on height sensors (35).

1.1.3.1.1. Noise within Pavement Laser Measurements

While spikes are widely recognized as a common form of noise, they are not the only type encountered in signal processing. Noise, generally defined as any interference affecting the measured signal, can stem from various sources, including physical phenomena, environmental conditions, sensors, data acquisition methods, or data transmission. In pavement surface scanning, adverse field conditions frequently introduce noise into the signal. Profiles collected using different laser sensors have demonstrated various noise types, such as white noise,

spikes, and flat signals. The amplitude and frequency characteristics of these noise types can differ across studies due to the diversity of laser devices.

White noise, a random signal with consistent intensity across frequencies, poses a challenge by potentially creating the illusion of capturing microtexture. Therefore, understanding the resolution of the laser in both vertical and lateral directions is essential before assuming accurate microtexture capture. Digital filters, such as low-pass or band-pass filters, can be used to eliminate wavelengths smaller than the sensor’s resolution, thereby reducing the impact of white noise-induced microtexture.

Spikes, characterized by abrupt, short-duration elevation changes, disrupt the pavement profile trend and are visually identifiable. While extreme spikes (i.e., outliers) can be easily detected and removed through automation, mild spikes that do not necessarily represent profile extremes require careful consideration in the detection and removal process (36).

Flat signals result from a combination of low camera exposure time and very dark pavement surfaces. In such instances, the sensor nullifies measured information, producing a flat line where multiple locations share the same elevation as the last “good” point measured by the sensor (36). **Figure 7** provides examples of spikes and flat signals.

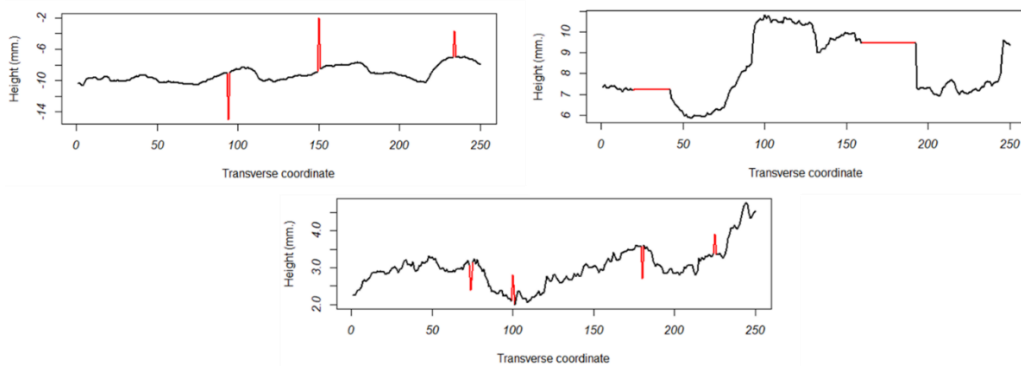


Figure 7: Common instances of noise extreme spikes (top left), flat signal (top right), mild spike (bottom) (36).

1.1.3.1.2. Nearly Stationary Data

The second criterion for processing data highlights the necessity of an almost stationary signal, specifically a time series with constant statistical properties such as mean, variance, and autocorrelation over distance (37). Essentially, this means that pavement elevation data should be centered around zero millimeters with no discernible trends. The goal is to center the data, remove linear trends, and retain seasonality. In this context, seasonality refers to regular and predictable changes

recurring after a specified distance, which is crucial as it reflects the pavement texture due to aggregate gradation in the mix.

Achieving an almost stationary time series requires applying a detrending algorithm to the data. Common detrending methods include integration and regression detrending. In this context, integration involves differencing adjacent observations to obtain an almost stationary time series, while regression detrending uses linear regression to subtract the mean of the regression function at each coordinate, resulting in an approximately stationary time series. Detrending is often described as a two-step process involving the suppression of slope and offset. **Figure 8** illustrates these detrending techniques...

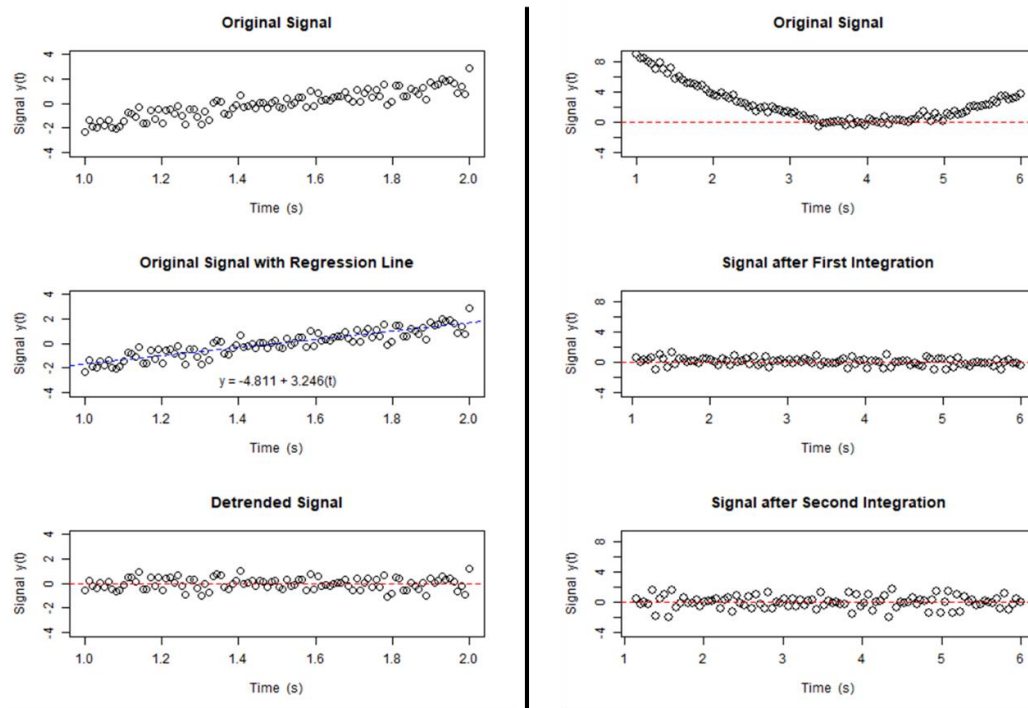


Figure 8: Detrending techniques: regression (left) and integration (right) detrending.

1.1.3.1.3. No Missing Data

The third requirement concerns the absence of missing data, which has significant statistical and practical implications. Missing data, categorized separately due to its unique challenges, often arises from sensor issues, data format conversion, or the removal of noisy data points. Many coding software functions are not inherently designed to handle missing data, leading to potential issues such as algorithm failure or automated imputation without user awareness.

Incomplete data can skew indices, particularly those reliant on sample size, like root mean square (RMS), skewness, and kurtosis. A good rule of thumb is to discard

dynamically collected texture profiles or surfaces where more than 33 percent of the data is noise or missing. Standards such as ISO 13473 (38) and ASTM E1845-15 (39) are more stringent, suggesting thresholds of 10 and 20 percent for stationary texture measurements, respectively.

Researchers should use appropriate imputation techniques to address missing data, where imputation involves replacing missing data with substituted values. While imputation is sometimes criticized, it preserves relationships between variables rather than generating valid estimates for specific case values. It has proven consistent and superior to using only complete data or deleting incomplete cases, which can lead to biased results (40,41). For pavement texture data, if the percentage of missing or removed data exceeds an established threshold, it should be discarded. Otherwise, linear interpolation is recommended as the fastest, simplest, and most accurate imputation method for 2D surface profiles (36). For 3D data, two-dimensional linear interpolation along the lateral directions can be implemented. Ultimately, imputation ensures a complete dataset, facilitates analysis, and eliminates biases that may arise from ignoring missing data.

1.1.3.2. Texture Data Processing Overview

An extensive review of the literature on texture processing and pavement surface characterization reveals multiple methods for processing raw 2D surface profile data, summarized in **Figure 9**. The data processing begins with algorithm calibration, which varies depending on the chosen processing algorithm. This initial step involves estimating thresholds crucial for the later removal of noisy data points from the signal, such as the maximum difference in height between adjacent points before a point is considered a spike. Subsequent steps, both algorithm and analysis dependent, typically include preprocessing tasks such as converting elevation units to millimeters, trimming profile edges, and applying pre-imputation methods to ensure data completeness before further manipulation.

The next phase involves detecting and removing noisy data points from the profile, a process known as denoising. Noise encompasses disturbances such as spikes, flat signals, and white noise that negatively affect the measured signal (36). Methods for addressing noisy data points include the Wavelet Methodology (33), ISO Standard 13473-1 (42), False Discovery Rate (43), ASTM Standard E178-21 (44), and the Sabillon-Orellana Filtering Algorithm (SOFA) (3).

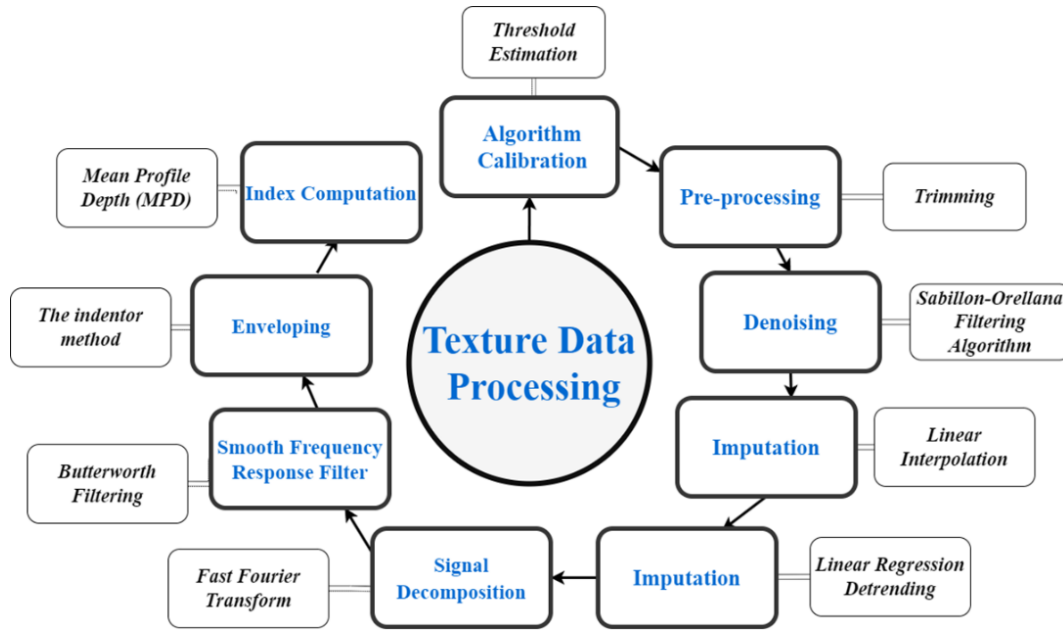


Figure 9: 2D pavement surface profile data processing schematic.

After denoising, an imputation method must be applied to replace any removed data. Linear interpolation is the most common imputation method used in the field and was proven by Sabillon-Orellana (36) to be the simplest, most efficient, and most accurate for 2D surface profile data. Following imputation, the surface profiles must be detrended to eliminate elevation biases and inherent trends, transforming the profiles into a nearly stationary and centered time series of elevation measurements. Both ASTM Standard E1845-15 (39) and ISO Standard 13473-1 (42) recommend using a linear detrending method to center the data.

Subsequent steps, although not always utilized in research studies, may include various mathematical transformations before computing a texture index. These transformations, as documented in ISO Standard 13473-4 (38) and National Cooperative Highway Research Program synthesis report 964 (45), might involve signal decomposition through methods such as the Fast Fourier Transform (46) to break down the profile into wavelength components, the application of smooth frequency response filters like a low-pass Butterworth filter (47) to isolate macrotexture wavelengths, and enveloping methods such as the indenter method (48) to model the contact zone where the tire interacts with the pavement surface.

Depending on the research goals, these transformations may be necessary, but they are not always required before quantifying the processed profiles into a summary index that captures various aspects of the pavement surface, such as MPD, RMS, among others.

1.1.4. Quantifying Pavement Texture

The foundation of pavement texture characterization lies in mathematics, particularly statistics. The 2D texture profiles serve only as a visual representation of a slice of the pavement's surface which needs to be further summarized into a set of *texture indexes* to quantify specific surface characteristic of the profile. Pavement practitioners use well-established and experimental equations to describe texture for all four components. It's important to note that terms like “statistic,” “parameter,” “metric,” and “index” are often used interchangeably in the literature to refer to numerical summaries that characterize features within the profile. In this report, these numerical summaries will consistently be referred to as “indices.”

At the unevenness/roughness level, the most common index is the International Roughness Index. Other indices include the coefficient of planarity, and the bump integrator trailer index, among others (49).

Comprehensive texture indices for the megatexture component are lacking, unlike unevenness/roughness and macrotexture. The primary standardized index for megatexture is the “megatexture level,” defined by ISO 13473-5 (2009). The scarcity of texture indices at this scale is primarily due to the focus on quantifying distresses such as faulting, rutting, and cracking, rather than on texture analysis (50). As of June 2024, no state or federal highway agency explicitly measures megatexture in managing their pavement networks, nor do they specify megatexture for newly paved or resurfaced pavements.

Macrotexture relies on essential indices like mean profile depth (MPD) and mean texture depth (MTD), which are widely recognized in pavement engineering literature. These indices characterize pavement texture and are frequently used in studies correlating surface interactions such as friction (3,51), rolling resistance (52), noise (15,51), and water splash (53). However, relying solely on MPD or MTD provides an incomplete depiction of surface topography, as they only represent the profile's vertical amplitude. Pavements with similar MPDs can still exhibit differing textures, friction levels, and noise levels, as noted by (28). Therefore, a more comprehensive description requires the integration of additional indices.

While many texture indices established at the macrotexture level can be computed at the microtexture level using a high-pass filter to isolate higher frequencies (2), quantifying microtexture remains challenging, particularly during dynamic data collection at speeds of at least 45 mph. Consequently, standardized methods for characterizing this intricate texture level are currently lacking. Past research studies (2,17,54) addressed this issue by using high-precision stationary texture measuring equipment to capture and quantify microtexture. These studies employed

microtexture data to predict skid resistance and noise with varying degrees of success.

Pavement texture quantification involves two primary types of summary indices: spatial and spectral. Spatial indices, computed in the spatial domain, vary with scale, such as MPD at macro and microtexture scales. In contrast, spectral indices, evaluated in the frequency domain, are scale-independent and assessed across various texture wavelengths/frequencies to avoid the complexity of scale-specific definitions (17).

1.1.4.1. Spatial Indices

Spatial texture indices are divided into five categories: amplitude, spacing, hybrid, functional, and feature indices.

Amplitude indices are crucial for characterizing the vertical characteristics of surface deviations on the pavement. The most common in pavement engineering are MPD and MTD. Other examples include the RMS, skewness, kurtosis, solidity factor, and ten-point mean roughness. These amplitude indices are used for various purposes in literature. For more detailed information, refer to the study by Sabillon et al. (3).

Spacing indices measure the horizontal characteristics of surface deviations. These indices are particularly important in fields like manufacturing, where consistent lubrication is necessary to avoid scoring and ensure a smooth surface finish on products like pressed sheet steel (55). In pavement engineering, spacing indices have proven useful in predicting the type of flexible pavement surface using field texture data (54). Common spacing indices include the mean, variance, skewness, and coefficient of variation in the cross width of the profile. The cross width measures the horizontal distance between inflection points along the profile, i.e., the distance between consecutive points crossing the zero-elevation line.

Hybrid indices quantify the profile's slope by combining elements of its spacing and amplitude. Changes in either amplitude or spacing can affect hybrid properties (3). In tribology analysis, surface slope, surface curvature, and developed interfacial area are important factors influencing the tribological properties of surfaces (55). Examples of hybrid indices include the surface area ratio, two-point slope variance, and six-point slope variance.

Functional indices provide information about the surface structure based on a function such as the material bearing ratio curve. This curve is the cumulative probability distribution of the amplitude distribution function, indicating the probability of a texture profile having a certain height (Z) at any position (X) (2).

The automotive industry has sought to optimize and define indices for functional characterization of surface texture, focusing on the contact area between two bodies. The surface bearing ratio index describes the functional characteristics of the pavement surface that may affect tire wear or friction (56). Other functional indices include the areal material ratio and the inverse areal material ratio.

Feature indices are used to describe the surface texture's properties over a predetermined length of the profile, such as 100 mm, as opposed to the entire profile (57). This definition allows these indices to have a dual identity, in the sense that they can be both feature and amplitude like in the case of MPD, or feature and spacing in the case of the Mean Profile Element Spacing. In the pavement engineering literature, these are one of the most commonly used indices to correlate with friction and noise (51) as they encompass MPD, the most commonly used index in the field.

A compilation of 80 spatial texture indexes was created using research articles and standards in transportation engineering, including works from ISO (2021), Zuniga-Garcia (2017), Gadelmawla et al. (2002), and Sabillon-Orellana et al. (2023). This comprehensive list, along with their respective sources, is summarized in Table 1.

Table 1: List of indexes reviewed and their respective source.

Amplitude (Height) Index		
ID	Index	Source
1	Arithmetic Mean Height	ISO 21920-2, 2021
2	Root Mean Square Height	ISO 21920-2, 2021
3	Skewness	ISO 21920-2, 2021
4	Kurtosis	ISO 21920-2, 2021
5	Total Height	ISO 21920-2, 2021
6	Maximum Height per Section	ISO 21920-2, 2021
7	Solidity Factor	Gadelmawla et al, 2002
Spatial Index		
ID	Index	Source
8	Dominant Spatial Wavelength	ISO 21920-2, 2021
9	Mean Spacing of Adjacent Local Peaks	Gadelmawla et al, 2002
10	Mean of Cross Width	Sabillon-Orellana et al 2023
11	Variance of Cross Width	Sabillon-Orellana et al 2023
12	Skewness of Cross Width	Sabillon-Orellana et al 2023
13	Number of Intersections at the Mean Line	Gadelmawla et al, 2002
14	Number of Peaks per Profile Length	Gadelmawla et al, 2002
15	Number of Inflection Points	Gadelmawla et al, 2002
16	Mean Radius of Asperities	Gadelmawla et al, 2002

Hybrid Indexes		
ID	Index	Source
17	Root Mean Square Gradient	ISO 21920-2, 2021
18	Arithmetic Mean of Absolute Gradient	ISO 21920-2, 2021
19	Maximum Absolute Gradient	ISO 21920-2, 2021
20	Developed Length	ISO 21920-2, 2021
21	Developed Length Ratio	ISO 21920-2, 2021
22	Two-Point Slope Variance	Zuniga-Garcia, 2017
23	Six-Point Slope Variance	Zuniga-Garcia, 2017
24	Profile Slope at Mean Line	Gadelmawla et al, 2002
25	Mean Slope of the Profile	Gadelmawla et al, 2002
26	Average Wavelength	Gadelmawla et al, 2002
27	RMS of Wavelength	Gadelmawla et al, 2002
28	Relative Length of the Profile	Gadelmawla et al, 2002
29	Bearing Area Length	Gadelmawla et al, 2002
30	Steepness Factor of the Profile	Gadelmawla et al, 2002
31	Waviness Factor of the Profile	Gadelmawla et al, 2002
32	Roughness of Height Uniformity	Gadelmawla et al, 2002
33	Roughness Height Skewness	Gadelmawla et al, 2002
34	Roughness Pitch Uniformity	Gadelmawla et al, 2002
35	Roughness Pitch Skewness	Gadelmawla et al, 2002
Functional Indexes		
ID	Index	Source
36	Autocorrelation Length	ISO 21920-2, 2021
37	Relative Material Ratio	ISO 21920-2, 2021
38	Material Ratio Height Difference	ISO 21920-2, 2021
39	Core Height	ISO 21920-2, 2021
40	Reduced Peak Height	ISO 21920-2, 2021
41	Reduced Pit Depth	ISO 21920-2, 2021
42	Maximum Peak Height	ISO 21920-2, 2021
43	Maximum Pit Depth	ISO 21920-2, 2021
44	Material Ratio of Hills	ISO 21920-2, 2021
45	Material Ratio of Dales	ISO 21920-2, 2021
46	Area of Hills	ISO 21920-2, 2021
47	Area of Dales	ISO 21920-2, 2021
48	Plateau Root Mean Square Deviation	ISO 21920-2, 2021
49	Dale Root Mean Square Deviation	ISO 21920-2, 2021
50	Material Ratio at Plateau to Dale Transition	ISO 21920-2, 2021
51	Hill Material Volume	ISO 21920-2, 2021
52	Core Material Volume	ISO 21920-2, 2021
53	Core Void Volume	ISO 21920-2, 2021
54	Dale Void Volume	ISO 21920-2, 2021
55	Mean of Local Curvature	ISO 21920-2, 2021
56	Standard Deviation of Local Curvature	ISO 21920-2, 2021

Feature Indexes		
ID	Index	Source
57	Mean Profile Depth	ASTM E1845-15, 2016
58	Maximum Peak Height	ISO 21920-2, 2021
59	Mean Peak Height	ISO 21920-2, 2021
60	Maximum Pit Depth	ISO 21920-2, 2021
61	Mean Pit Depth	ISO 21920-2, 2021
62	Maximum Height	ISO 21920-2, 2021
63	Mean Profile Element Spacing	ISO 21920-2, 2021
64	Maximum Profile Element Spacing	ISO 21920-2, 2021
65	Standard Deviation of Profile Element Spacing	ISO 21920-2, 2021
66	Mean Profile Element Height	ISO 21920-2, 2021
67	Maximum Profile Element Height	ISO 21920-2, 2021
68	Standard Deviation of Profile Element Height	ISO 21920-2, 2021
69	Peak Count Parameter	ISO 21920-2, 2021
70	Density of Peaks	ISO 21920-2, 2021
71	Density of Pits	ISO 21920-2, 2021
72	Arithmetic Mean Peak Curvature	ISO 21920-2, 2021
73	Arithmetic Mean Pit Curvature	ISO 21920-2, 2021
74	Five-point Pit Depth	ISO 21920-2, 2021
75	Five-point Peak Height	ISO 21920-2, 2021
76	Ten-point Height	ISO 21920-2, 2021
77	Mean of Maximum Peak to Valley Height	Gadelmawla et al, 2002
78	Largest Peak to Valley Height	Gadelmawla et al, 2002
79	Third Point Height	Gadelmawla et al, 2002
80	Mean of Third Point Height	Gadelmawla et al, 2002

For further information on these spatial indices refer to the studies by Gadelmawla et al. (55), NASEM (45) , ISO 21920-2 (57), Chen et al. (58), and Sabillon et al. (3).

1.1.4.2. Spectral Indices

Spectral indices refer to those obtained in the frequency domain, requiring Fourier analysis to examine surface texture. A Fourier transform decomposes a texture profile into a function of sinusoidal waves. One common approach to characterize texture in the spectral domain is to determine indices from the texture spectrum. The technical specification ISO/TS 13473-4 (38) describes the procedure for obtaining the texture spectrum expressed in octave or one-third octave bands. The index used to characterize the texture spectrum is known as the texture level, a logarithmic transformation of an amplitude representation for a texture profile with a center wavelength, reported in dB. This approach has mainly been used to find

correlations between tire-pavement noise and road texture, but it has also been used to correlate texture with friction (27,59).

Another commonly used approach is Power Spectral Density (PSD), which analyzes pavement texture in the frequency domain. PSD describes how the energy of a pavement texture profile is distributed over different frequencies. The PSD of a roadway is obtained by applying a Fourier transform to the linear profile of a pavement surface, decomposing it into sinusoidal functions with discrete frequencies. Given the complexity of road profiles, individual amplitudes are almost always small, so the Fourier transform is adjusted to show how the variance of the profiles is distributed over a set of sinusoids. This adjustment is known as the Power Spectral Density (60,61). They used the slope and intercept of the linearized PSD curve to characterize surface macro and microtexture, observing a strong correlation between the log of the PSD and the log of the frequency in most sampled test surfaces.

Wavelet transform (WT) or wavelet analysis is an analytical method developed to overcome the shortcomings of the Fourier transform. Wavelets are functions that satisfy certain mathematical requirements used to represent data. WT decomposes a signal into different frequency components, presenting each component with a resolution matched to its scale (62). The major advantage of WT is its ability to analyze localized areas of a larger signal, capturing aspects like trends, breakdown points, and discontinuities that other signal analysis techniques might miss. In road roughness analysis, WT can reveal localized surface irregularities such as surface depressions, potholes, surface heaving, and bumps (63).

Other studies have used fractal and multi-fractal theory to characterize texture. A fractal is a mathematical set with a fractal dimension that usually exceeds its topological dimension and may be non-integral (64,65). Fractals are typically self-similar patterns, meaning they appear the same from near and far. The fractal theory for pavement assumes that texture irregularities follow similar patterns at different scales. To appreciate the similarity of texture patterns at different levels, the patterns need to be scaled by different factors known as the fractal dimension in the coordinate axis (66). Villani et al. (67) evaluated pavement surface friction properties using fractal analysis and developed a methodology to select aggregate type and mix design for optimized skid resistance performance.

Although the spectral indices described above are based on solid physical principles, they are often too complex for road engineers to use due to the significant time required to compute the accompanying indices (68). To address this issue, simpler approaches based on modern signal processing techniques have been developed. Among these, the simplest and most promising is the Hilbert-

Huang Transform (HHT) (69). Unlike the theoretical approaches of Fourier or Laplace transforms, HHT is an empirical method that decomposes a signal into a set of functions to obtain instantaneous amplitudes and frequencies.

1.1.4.3. Surface Characteristics in other Fields

The spatial and spectral indices reviewed in the previous section are indices used in pavement engineering to describe roadway surface characteristics. However, pavement engineering is just one of many fields that uses surface engineering concepts to understand and quantify the morphological characteristics of surfaces to enhance desired properties. This section explores how other disciplines quantify and control the surface roughness of their materials.

1.1.4.3.1. Materials Science: Surface Engineering

Surface engineering, a sub-discipline of materials science, deals with the surface of solid matter. It is an interdisciplinary field focused on manipulating the chemical, physical, and mechanical properties of materials to enhance surface properties independently from those of the underlying substrate. Enhancements may target visual appearance, tactile properties, optical properties, wettability, corrosion resistance, or tribological behavior (70).

Surface engineering processes can be divided into three basic groups. The first group includes processes that modify the existing surface without changing its composition, such as transformation hardening, surface melting, and surface texturing. The second group involves processes that change the surface composition, such as altering the crystal structure or creating surface precipitates through chemical reactions. The third group applies a material to the surface, generally referred to as coatings. This includes practices like painting, weld hardfacing, and electroplating, where a surface-engineered layer is applied over the original material to achieve a desired property (70).

Many processes used in surface engineering are also observed in pavement engineering. For example, surface texturing is commonly done on rigid pavements during curing or after hardening to reduce pavement/tire noise, improve drainage, or enhance skid resistance (71,72). Additionally, most surface treatments applied to pavements today are examples of surface engineering coatings used to enhance skid resistance without modifying the original flexible or rigid pavement underneath.

1.1.4.3.2. Tribology

Tribology is an interdisciplinary field drawing on principles from physics, chemistry, materials science, and mechanical engineering to study the interactions between surfaces in relative motion, including friction, wear, and lubrication. A tribological system has functional characteristics and performs technical tasks related to energy, material, or signal transmission. The tribological properties depend on system function, load complexity, and structure. Understanding the complete tribological system (**Figure 10**) is essential for evaluating and designing many products. Friction and wear are considered system properties rather than material characteristics, meaning that even slight modifications to the system can significantly alter these properties (73,74).

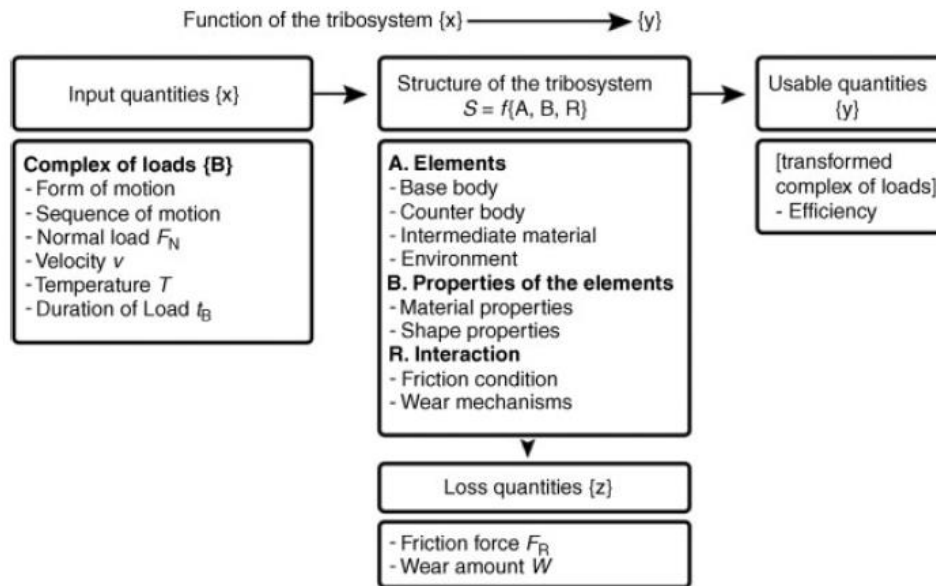


Figure 10: Complete description of a tribological system (73).

The tribological system involves frictional and wear contact between two bodies, the base body and the counter-body, which move relative to each other, with or without an intermediate medium (**Figure 11**). Interactions on the surface of the solid arise from this relative movement. The load complexity and structure determine the wear behavior of the tribological system (74). Tribological studies characterize the interacting surfaces' material properties at both physical and chemical levels, although this study focuses on the physical level.

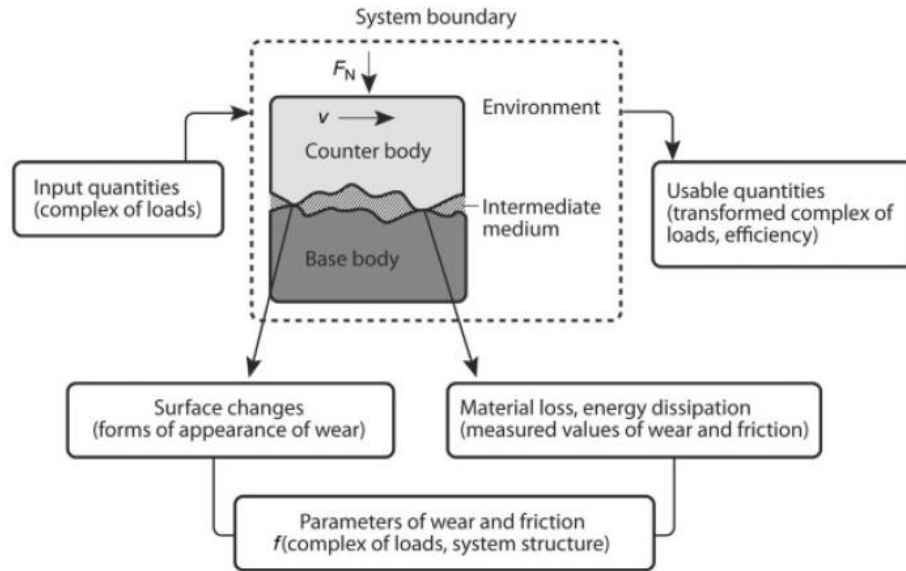


Figure 11: Schematic representation of a tribological system (74).

Physically, tribologists assert that all technical surfaces deviate from their ideal geometric form due to manufacturing processes. According to DIN 4760 (75), deviations can be categorized from the first to the sixth order, analogous to the texture components in pavement engineering (e.g., roughness, megatexture). The first order has the longest wavelength, while the sixth has the shortest. Form deviations from the first to the fourth order are typically superimposed on the real surface profile (74), as shown in **Figure 12**.

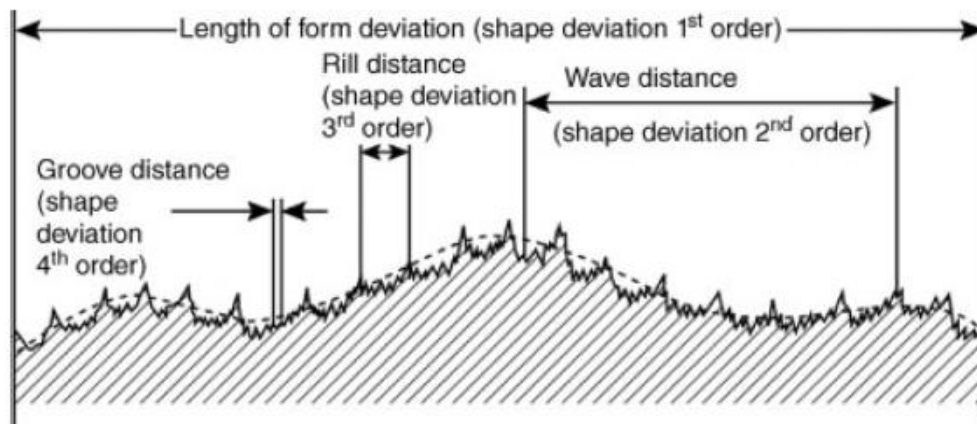


Figure 12: Schematic illustration of the form deviations of the first to fourth order (74).

The microscopic form deviations of the third and fourth orders are typically determined using a contact stylus and described as roughness values. Measurement procedures for these orders are standardized according to ISO 21920-2 (57). The most frequently cited roughness values in these orders are the mean roughness depth (R_z) and the mean roughness value (R_a), both illustrated in **Figure 13**. These

indices are commonly used by tribologists to perform quality assurance and control on materials during experimentation.

Form deviations of the fifth and sixth orders refer to deviations in the microstructure and lattice structure of materials. Due to the manufacturing process, the microstructure of material surfaces almost always differs from that within the material (74). While tribologists have indices to quantify the fifth and sixth orders, analogous to the first and second decades of microtexture, their characterization is too complex for practical purposes in pavement surface identification.

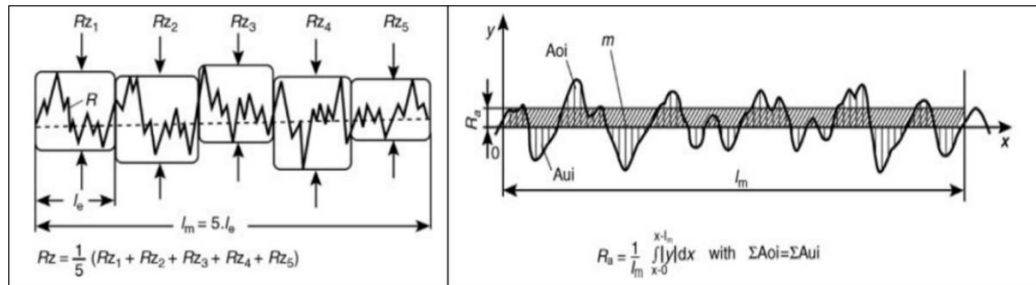


Figure 13: Diagrams portraying the computation of mean rough depth (Rz , on the left) and mean roughness value (Ra , on the right) (74)

1.1.4.3.3. Fluid Dynamics

Fluid dynamics, a subdiscipline of fluid mechanics, describes the flow of fluids, including liquids and gases. It encompasses several subdisciplines such as aerodynamics (the study of air and other gases in motion) and hydrodynamics (the study of liquids in motion) (76). Surface engineering plays a crucial role in fluid dynamics as engineers seek to reduce turbulence in airflow or control pressure drops during internal flow in fluid systems (77).

Surface roughness was identified as an important index in fluid flow as early as the nineteenth century by Darcy (78), who conducted experiments with pipes of varying roughness. Subsequent work by Fanning (79), Nikuradse (80), and Colebrook (81) further expanded the understanding of how surface roughness influences pressure drops. In 1944, Moody presented Colebrook's results graphically in the well-known Moody diagram, where the Darcy friction factor is represented as a function of Reynolds number over a relative roughness (82). This plot covered both laminar and turbulent regions, showing that while roughness has little effect in the laminar region, it plays a major role in the turbulent region. In the turbulent region, the friction factor increased with Reynolds number and asymptotically reached a constant value at higher Reynolds numbers. This constant asymptotic value of the friction factor increased with increasing relative roughness (77), as shown in **Figure 14**.

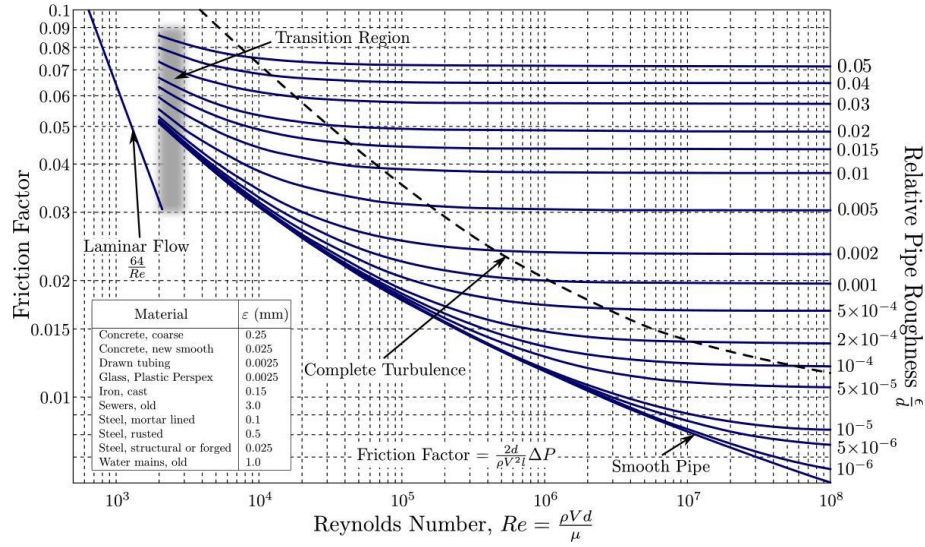


Figure 14: The Moody diagram (82).

In fluid dynamics, surface texture is defined as the composite profile of certain deviations typical of an ideal real surface. Since Nikuradse's work in 1937, a series of 2D indices aimed at describing amplitude, spacing, waviness, and shape of material surfaces have been developed and documented in several standards (57,83–85). Despite the breadth of available surface indices, common industry practice is to specify and control surfaces with a single indicator, generally the mathematical average roughness (Ra). However, experience has shown that this single index is insufficient to describe the complete functionality of a surface or its interface. A surface with sharp spikes, deep pits, or general isotropy may all yield the same Ra but perform very differently. While Ra remains useful as a general guideline on surface texture, it has proven too general to describe the functional nature of surfaces in today's increasingly complex applications (77). This overreliance on Ra in fluid dynamics mirrors the overuse and limitations of MPD in pavement engineering.

Over the years, numerous indices have been developed to address the limitations of using a single index. These have been broadly classified into categories such as roughness, waviness height, spacing and amplitude, shape, and areal indices. Advances in instrumentation have provided information on surface texture in various formats and resolutions (77). Work by Kandlikar et al. (86) proposed six new roughness indices for studying systems with relative roughness values up to 14% for single-phase fluid flow applications. These include three indices for roughness characterization (maximum profile peak height, Rp; mean spacing of profile irregularities, Rsm; and floor distance to mean line, Fp) and three indices for localized hydraulic diameter variation (maximum, minimum, and average). The roughness is then defined as $Rp + Fp$, as shown in **Figure 15**. This index, along

with a more accurate representation of pipe diameter, improved modeling performance and extended the model into areas not previously covered, as demonstrated in a study by Schmitt and Kandlikar (87).

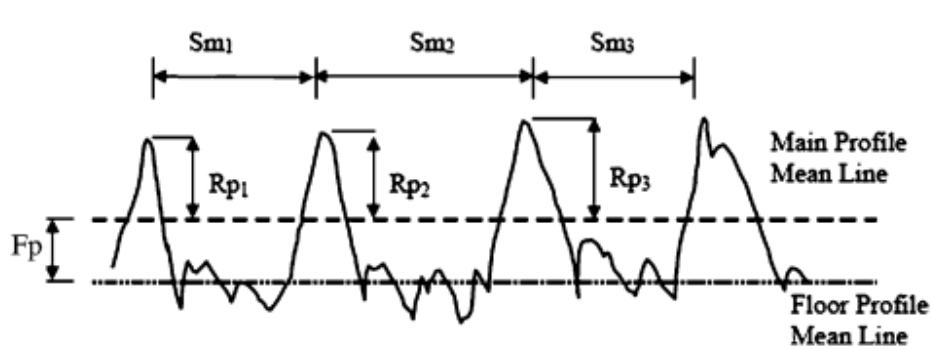


Figure 15: Definition of the roughness index as defined by (86).

1.1.4.3.4. Geological Sciences

Geological sciences have made significant contributions to pavement engineering, particularly in mineralogy and morphological characterization of aggregates. Within geological sciences, essential particle morphology is expressed in terms of contour shape, angularity, and surface texture (88).

Contour shape reflects the overall change characteristics of particles at a large spatial scale, distinguishing between flat, elongated, or flat and elongated particles (89). It is often defined in terms of aspect ratio (AR) and form factor (FF), as shown in **equations (1)** and **(2)**, respectively. Where a and b represent the lengths of the major and minor axes of the equivalent ellipse, respectively, A as the particle area, and P as the particle perimeter.

$$AR = \frac{a}{b} \quad (1)$$

$$FF = \frac{4\pi A}{P^2} \quad (2)$$

Angularity refers to the sharpness of the edges and corners of a particle's 2D image, reflecting its spatial characteristics at the mesoscopic scale (90). The angularity index (AI) is a common index used to describe angularity. This index measures the degree of change between the particle's perimeter or outer boundary and the perimeter of its equivalent ellipse, as indicated by **equations (3)** and **(4)**. A higher AI value indicates a more angular particle. Particles with more edges and corners tend to form stronger internal friction, preventing movement and maintaining structural stability under external forces (89). The formulas for AI are:

$$AI_1 = \left(\frac{P}{P_e} \right)^2 \quad (3)$$

$$AI_2 = \left(\frac{P_c}{P_e} \right)^2 \quad (4)$$

Where P is the perimeter, Pc is the outer boundary perimeter, and Pe is the equivalent ellipse perimeter of each particle.

Surface texture reflects the complexity of a particle's uneven surface at the microscopic scale, characterizing its microstructure to some extent (89). Fractal geometry theory and its evaluation method are widely applied, with fractal dimension (FD) and feature roughness (FR) considered reliable indicators of a particle's geometric complexity (91). These indices are defined as per **equations (5) and (6)**.

$$N(c_i > d) = B * d^{-FD} \quad (5)$$

$$FR = K^{\frac{1}{FD}} * 100 \quad (6)$$

Where ci is the mean dimension, d is the aperture size, B is a dimension index, and K is the fractal intercept.

These three morphological properties—contour shape, angularity, and surface texture—are not mutually exclusive. Accurate prediction of particle behavior requires understanding the origin and evolution of fragmental particles from minerals (92,93). Characterizing particle morphology is key to determining their formation mechanisms, defining fragmented particle properties (94), and predicting behaviors such as fluid-particle interaction and physical resistance (95).

Several other indices describe particle morphology. Morphological indices, which determine geometrical variables like length, radius, perimeter, area, and volume, are categorized into 1D, 2D, and 3D images based on associated methods and variables (96). Typically, 1D morphological indices, such as aspect ratio, elongation, and flatness, are based on particle length in three dimensions and can be measured quickly with a ruler or caliper. 2D morphological indices are measured using image analysis of particle projections, including projection perimeter, area, and circumscribing circles (96,97). For 3D visualization and quantitative analysis of particle microstructures, Fourier transform interferometry and X-ray computed

tomography are very useful. These methods analyze 3D morphological indices such as sphericity and 3D angularity (98).

Despite the development of numerous indices, there is a lack of consensus on which particle morphological indices are most effective in describing morphological characteristics (89).

1.2. Pavement Mixes and Surfaces

The previous sections have discussed the general characterizations of surface roughness and texture in pavement engineering and other disciplines. This section will introduce the most common pavement surfaces found on Texas highways, providing a brief description of what makes each pavement mix or surface unique.

A roadway pavement surface can fall into two types: flexible or rigid. A flexible pavement consists of a relatively thin asphalt wearing course over layers of granular base and subbase, protecting the subgrade from overstresses (99). In contrast, a rigid pavement comprises portland cement concrete slabs poured over a base course, designed to be much stiffer than the foundation material, providing a strong and durable surface, particularly for heavy traffic and loads (99).

1.2.1. Flexible Pavements: Hot mix asphalt (HMA)

The HMA is an asphalt concrete that combines heated asphalt binder and aggregates with air voids, that are then laid down and compacted at specific temperatures to form a durable road surface. This process results in a flexible and relatively impermeable surface for diverse engineering applications such as highways, taxiways, and parking lots. HMA can be further subdivided based on their aggregate matrix: dense-graded, open-graded, and gap-graded mixes.

1.2.1.1. Dense-Graded Mixes

A dense-graded mix is a commonly used mixture for structural surface courses, characterized by a high concentration of well-graded aggregate particles, ensuring a dense arrangement, as shown in **Figure 16**, that stabilizes and impermeabilizes the surface. The percentage of air voids in this mix depends on aggregate gradation, with fuller curves resting close to the density line (45 degrees) (100). These mixes, classified by nominal maximum aggregate particle size, are versatile for various pavement layers and traffic conditions, serving structural, friction, leveling, and patching purposes (101).

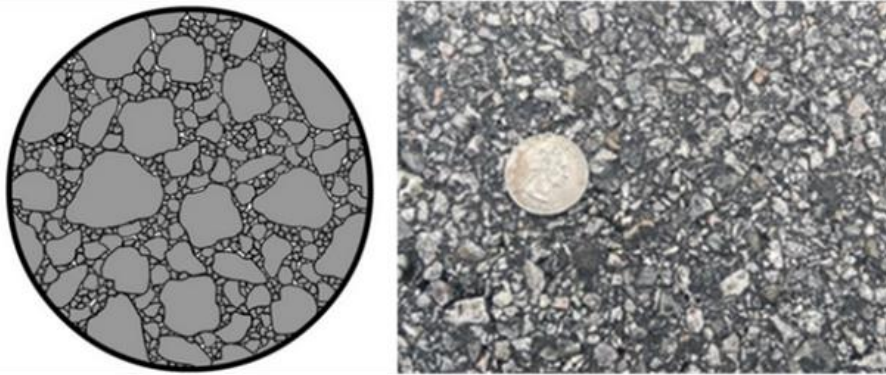


Figure 16: Diagram of an ideal pavement surface (102) and a field picture for a dense-graded mix

1.2.1.2. Open-Graded Mixes

An open-graded mix, often referred to as: “open-graded friction course” (OGFC) or porous friction course (PFC), is a permeable asphalt layer with a skeleton of uniformly sized aggregates and a minimal amount of fines (103), as shown in **Figure 17**. The high percentage of internal air voids is achieved through uniform aggregate grading, typically consisting of 50-60 percent particles of similar size. OGFCs are effective in improving surface friction, draining water, reducing hydroplaning, and lowering noise levels (103), but they face challenges such as clogging of voids and structural issues leading to a shorter lifespan compared to dense-graded mixes. Nonetheless, the application of modified binders and preventive maintenance can significantly extend their service life (104).

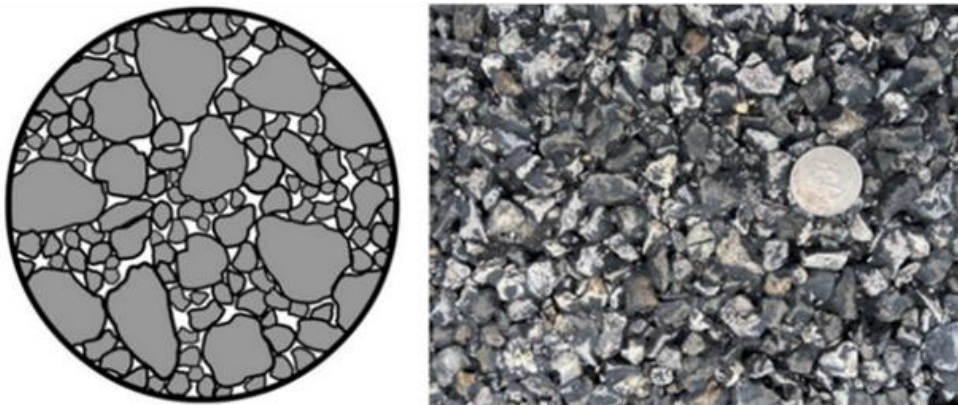
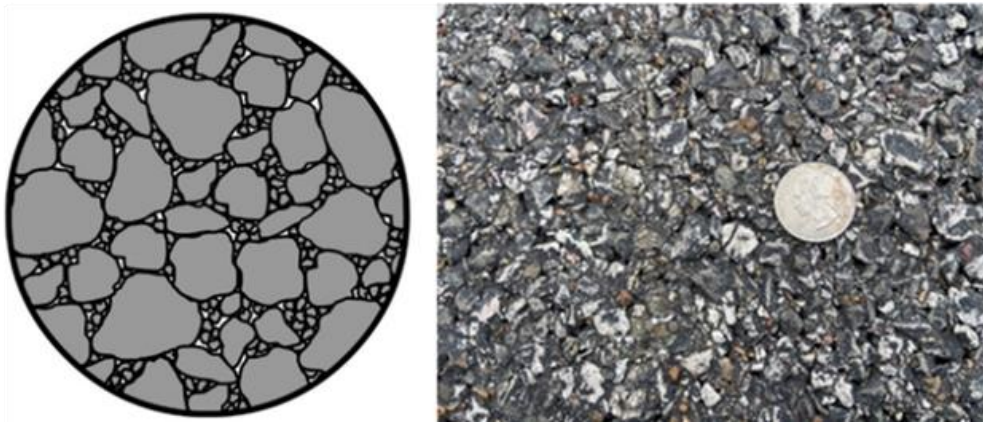


Figure 17: Diagram of an ideal pavement surface (102) and a field picture for an open-graded mix

1.2.1.3. Gap-Graded Mixes: Stone Matrix Asphalt (SMA)

An SMA is a gap-graded mix known for its stable stone-to-stone interlock, developed initially in Europe for studded tire wear and rutting prevention through

the use of a stable stone-on-stone skeleton held together by a rich mixture of asphalt cement, as shown in **Figure 18**, along with stabilizing agents such as fibers and/or asphalt modifiers (105). SMAs gained popularity in the United States since 1991, featuring approximately 70 percent coarse aggregate, 3-4 percent air voids, and a high filler content (106). Despite higher upfront costs, SMAs offer enhanced durability, fatigue and rutting resistance, reduced tire noise, and improved wet weather friction (105). However, the rough texture of an SMA can lead to a higher number of internal air voids, which are associated with a decline in performance, even if the total volume of air voids matches that of a conventional asphalt pavement (107).



***Figure 18:** Diagram of an ideal pavement surface (102) and a field picture for a gap-graded mix.*

1.2.2. Flexible Pavements: Surface Treatments

A surface treatment involves strategically applying binder and aggregates combined to the current pavement surface, aiming to reduce the deterioration rate while enhancing performance in terms of durability, and safety. The selection of a surface treatment is influenced by factors such as pavement condition, traffic volume, and desired performance outcomes. This study categorizes surface treatments into four groups: seal coats, high-friction surface treatments, thin overlay mixes, and bonded wearing courses. However, note that the names for these treatments may vary across countries and different states within the U.S.

1.2.2.1. Seal coats

Seal coats, also known as chip seals or surface dressings, are a cost-effective pavement preservation method applied to existing paved surfaces (108). The process involves applying an asphalt binder layer followed by uniformly-graded aggregate, as shown in **Figure 19**, to create a weather-resistant seal, preventing air and water intrusion, and extending pavement life (108,109). Seal coating is

effective for roads with light traffic, providing a skid-resistant surface that can last six to ten years when applied correctly (110,111). This method encompasses various treatments, including crack seal, fog seal, slurry seal, and many more (109), where the main difference lies in the type and gradation of aggregate used (112). The study specifically focuses on seal coats in Texas, where the Texas Department of Transportation has established standards for different grades based on permissible aggregate sizes, as shown in **Table 2**.

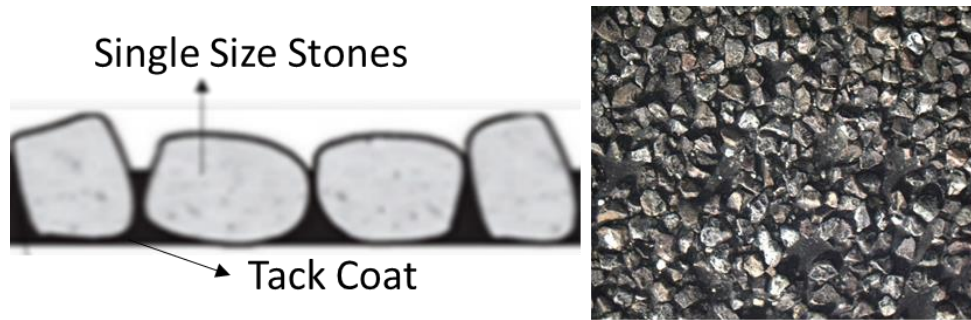


Figure 19: Representative diagram (113) and field picture for a single layer seal coat.

Table 2: Aggregate gradation requirements for seal coats in terms of cumulative percent retained at each sieve (114).

Sieve	Grade								
	1	2	3S ¹	3		4S ¹	4	5S ¹	5
				NL ²	L ²				
1"	~	~	~	~	~	~	~	~	~
7/8"	0-2	0	~	~	~	~	~	~	~
3/4"	20-35	0-2	0	0	0	~	~	~	~
5/8"	85-100	20-40	0-5	0-5	0-2	0	0	~	~
1/2"	~	80-100	55-85	20-40	10-25	0-5	0-5	0	0
3/8"	95-100	95-100	95-100	80-100	60-80	60-85	20-40	0-5	0-5
1/4"	~	~	~	95-100	95-100	~	~	65-85	~
#4	~	~	~	~	~	95-100	95-100	95-100	50-80
#8	99-100	99-100	99-100	99-100	98-100	98-100	98-100	98-100	98-100

¹ Single size gradation

² NL – Non-lightweight aggregates, L – Lightweight Aggregates

1.2.2.2. High Friction Surface Treatments (HFST)

HFSTs differ from seal coats in that they use high-quality aggregates like calcined bauxite with polymer binders (e.g., Epoxy-resin, polyester-resin, or polyurethane-resin), as shown in **Figure 20**, to endure extreme shear forces and maintain or restore pavement friction. This treatment is applied in areas prone to high-crash incidents, providing enhanced control for motorists in dry and wet conditions (115,116). The choice of aggregate material considers factors like cost, availability, and performance. HFSTs have proven effectiveness in reducing accidents,

particularly in wet conditions, offering durability, long-lasting results, and cost-effectiveness compared to geometric design improvements (116,117). They present a resilient and environmentally friendly solution with minimal traffic impacts where insufficient friction contributes to crashes, and pavement quality is good (115–117).

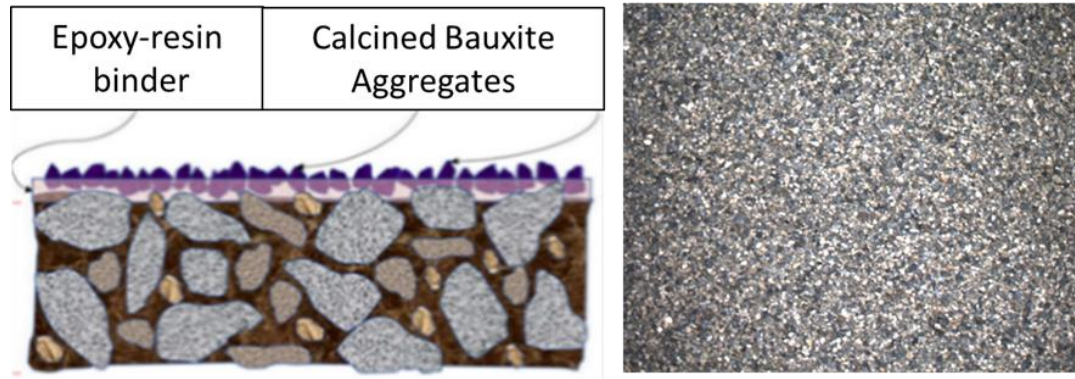


Figure 20: Representative diagram (118) and field picture for a high friction surface treatment.

1.2.2.3. Thin-Overlay Mix (TOM)

A TOM is a type of asphalt mixture used for pavement overlays, typically with a relatively thin layer and thus, smaller nominal aggregate size, as shown in **Figure 21**. This mix is designed to preserve the pavement structure by providing a smooth and durable surface to improve the ride quality and extend the life of existing pavements (119). Thin overlays may include modified binders and fine aggregates to enhance performance and resist common distresses such as cracking. The specific mix design and materials may vary based on project requirements and standards of the jurisdiction, ensuring the overlay meets performance and durability expectations.

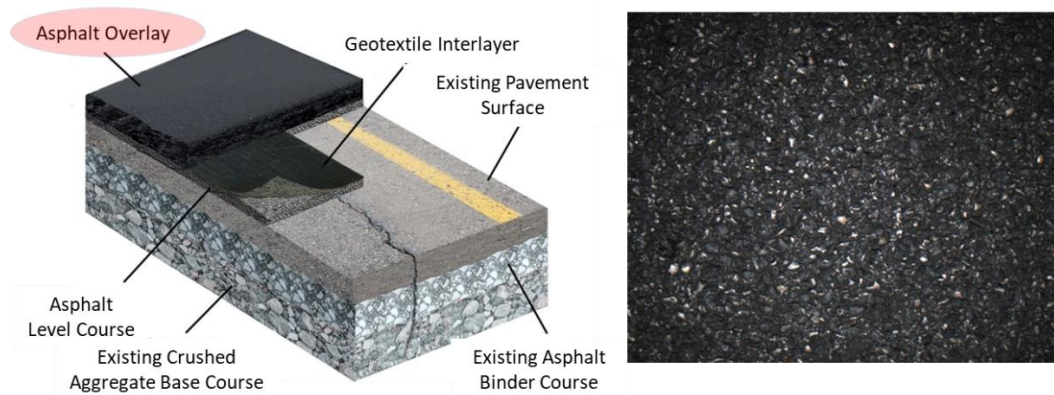


Figure 21: Representative diagram (119) and field picture for a thin overlay mix.

1.2.2.4. Bonded Wearing Coarse (BWC)

A BWC refers to a specialized asphalt overlay applied to an existing pavement surface. The application of a bonded wearing course not only aims to enhance the surface characteristics, ride quality, and overall performance of the road but also ensures a secure and lasting bond between the layers, designed to create a strong connection with the underlying pavement, which enhances durability and prevents delamination and crack propagation (120). It typically involves the use of high-quality materials and asphalt binders to ensure a reliable connection between the new overlay and the existing pavement, as shown in **Figure 22**. Additionally, there are variations such as thin bonded wearing courses (TBWC) and ultra-thin bonded wearing courses, (UTBWC) each varying levels of thickness is tailored to the specific requirements of pavement enhancement needed for the road.

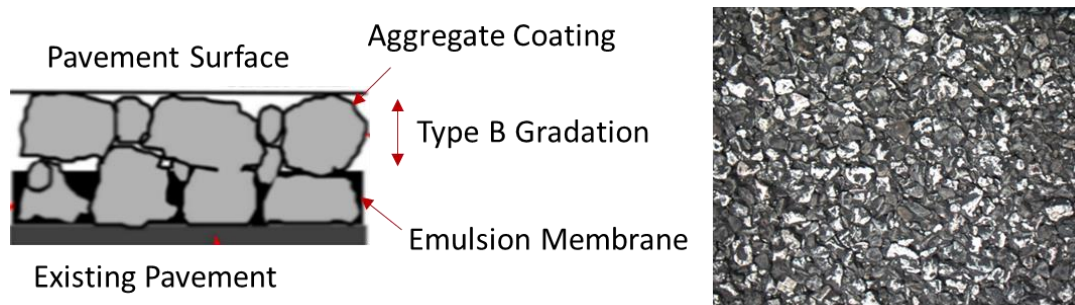


Figure 22: Representative diagram (120) and field picture for a bonded wearing coarse surface.

1.2.3. Rigid Pavement: Textures Constructed in Plastic Concrete

Various techniques are employed to texture the surface of portland Cement Concrete (PCC) pavements in their plastic state (i.e., when the concrete is still fresh and unhardened), aiming to create a safe, high-friction surface while considering noise reduction from tire-pavement interaction. Nonetheless, some texture types and orientations are inherently more prone to noise production than others (121).

1.2.3.1. Dragged Texturing: Burlap, Carpet, Brooming and Artificial Turf

Dragged texturing techniques for plastic concrete pavement involve using objects like moistened coarse burlap, brooms, carpet, or artificial turf, usually dragged along the direction of traffic. Burlap, carpet, and broom dragging create shallow textures (approximately 0.008 inches) with longitudinal striations and low macrotexture. Burlap drag, prevalent until the mid-1960s, provides a relatively quiet surface but may lack wet-weather friction at high speeds and can decrease in friction over time (122).

Artificial turf drag, pioneered by the Minnesota Department of Transportation, produces textures ranging between 0.06 to 0.12 in. in depth. The depth varies depending on turf factors and the concrete mix design (72). Noise and friction data suggest values comparable to durable asphalt pavements (123), highlighting the importance of a high-quality, low water/cement ratio concrete mixture for a lasting turf drag texture (124). These two PCC dragged surface texturing techniques are shown in **Figure-23**.

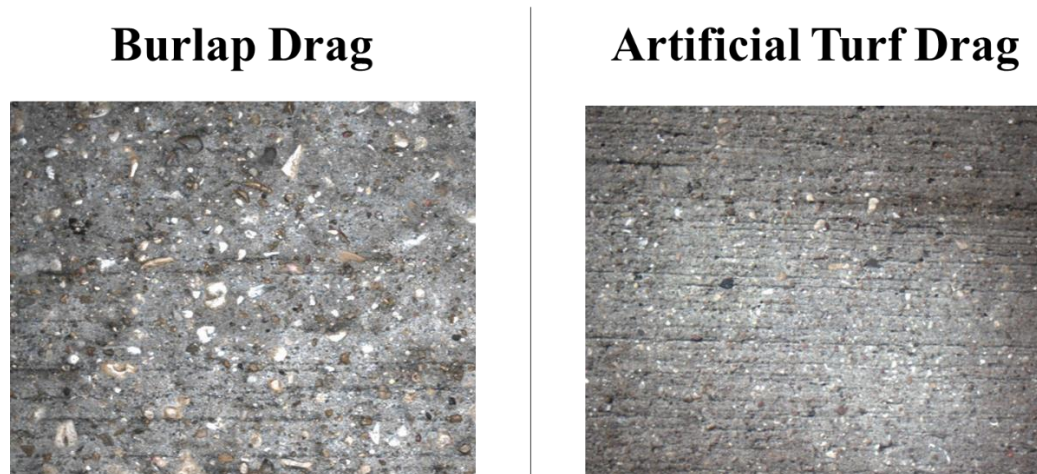


Figure-23: Field pictures for textures constructed in plastic PCC dragged surfaces burlap (left), artificial turf (right).

1.2.3.2. Tining

Tining is a texturing method often used in conjunction with burlap drag which creates grooves on the pavement with specified dimensions, where groove depths typically vary between 0.06 and 0.25 inches. Actual groove depth is influenced by concrete mix properties, tine down-pressure, and timing of the tining operation (72). Transverse tining, prevalent until the late 2010s, yields durable and hydroplaning-resistant surfaces for concrete pavements but tended to generate highly objectionable noise (72,124).

Longitudinal tining, the current standard, tends to produce pavement surfaces with reduced tire-pavement noise and a reduced variability of the noise itself (125). Friction testing shows minimal differences in skid resistance across pavements textured with transverse and longitudinal tining (71,125). Advantages of using longitudinal instead of transverse tining include reduced noise, earlier curing compound application, and less variability. Challenges involve ensuring grooves are straight and control for the uniformity of shallow grooves (125). Additionally, tine spacing can be fixed or random, with fixed spacing being regular and predetermined, and random spacing being irregular (71). These two PCC surface tined texturing techniques are shown in

Figure 24.

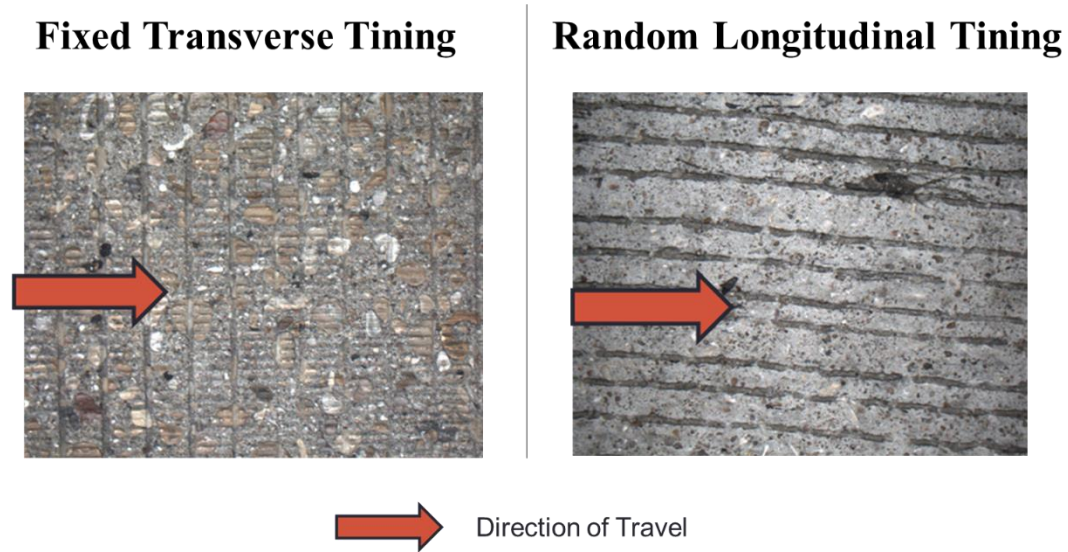


Figure 24: Field pictures for textures constructed in plastic PCC tined surfaces fixed transverse (left) and random longitudinal tining (right). Direction of travel shown by arrow.

1.2.3.3. Exposed Aggregate

Exposed aggregate texture is achieved by incorporating hard, angular, polish-resistant coarse aggregate into the surface concrete mixture and exposing it after placement using water or a surface-applied set retarder with mechanical brushing (72), as shown in **Figure 25**. While not commonly used in the U.S. (124), properly designed exposed aggregate surfaces are reported to have reduced noise, improved friction, and good durability (122,126). The construction cost of an exposed aggregate surface typically adds around 10 percent to the paving cost, although some short demonstration projects have reported higher costs (12).

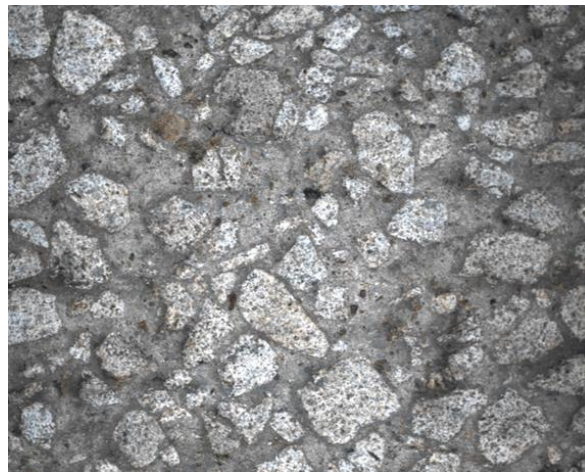


Figure 25: *Field pictures for textures constructed in plastic PCC exposed aggregate surface.*

1.2.4. Rigid Pavement: Textures Constructed in Hardened Concrete

Various techniques are employed to texture the surface of hardened concrete pavements, aiming to enhance tire-pavement noise characteristics and friction properties. Common methods include diamond grinding, diamond grooving, and shotblasting. Shotblasting will not be covered because it is not commonly used in Texas; nor will surface milling with carbide teeth be included, since it is deemed unsuitable for creating concrete pavement texture due to the potential damage to joints and its generation of unacceptable tire-pavement noise levels (72).

1.2.4.1. Conventional Diamond Grinding

Diamond grinding is a technique that involves removing a thin layer (typically 0.1 to 0.8 inches) from the surface of hardened concrete using closely spaced diamond saw blades (around 50-60 blades per foot) mounted on a rotating shaft, as shown in **Figure 26**. The resulting texture depends on factors such as aggregate hardness, blade spacing, and post-grind treatments (72). Since its inception in 1956, diamond grinding is now recognized as a highly effective method for pavement texturing and surface profiling. It is beneficial in pavement rehabilitation and restoration programs, improving ride quality, restoring surface friction, and reducing tire-pavement noise (72). The friction-increment and noise-reducing can be lost with time and wear, depending on traffic levels and aggregate hardness, but can be restored through additional grinding operations. Any slight pavement thickness reductions associated with grinding are generally compensated by increases in concrete elastic modulus (stiffness) over time (127).



Figure 26: *Field picture for texture constructed in hardened PCC using conventional diamond grinding*

1.2.4.2. Diamond Grooving

Diamond grooving is a technique that adds macrotexture to hardened concrete by cutting grooves in the pavement surface, resembling tining, as shown in **Figure 27**. The grooves are typically longitudinally cut with a blade spacing (center-to-center) of 0.75 inches and a depth ranging from 0.12 to 0.25 inches. On occasion, grooves are cut transversely at intersections or on airfield runways (72). This method proves highly effective in preventing hydroplaning and enhancing wet weather pavement friction. Longitudinal grooving also offers increased resistance to lateral skidding in curved sections (124).

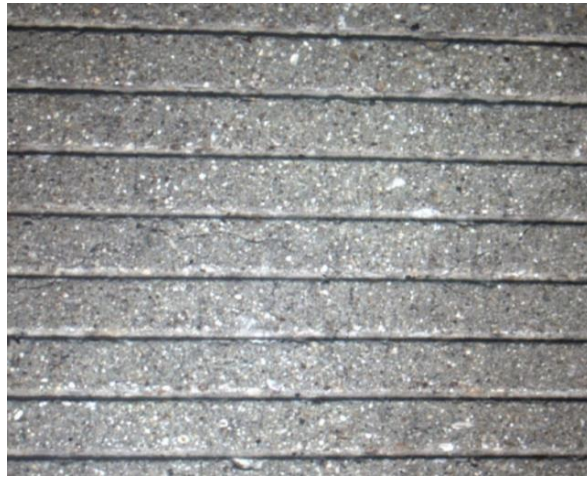


Figure 27: Field picture for texture constructed in hardened PCC using conventional diamond grooving

1.2.4.3. Next-Generation Concrete Surfaces (NGCS)

NGCS is a diamond saw-cut concrete pavement texturing technique developed around 2006, creating a texture resembling a hybrid of diamond grinding and grooving (72), as shown in **Figure 28**. It is typically produced in two passes, utilizing a flush grind followed by a grooving operation with closely spaced blades (72). The resulting texture is predominantly “negative,” featuring grooves with minimal upward-oriented texture, providing effective dry- and wet-weather friction while being one of the quietest concrete pavement textures. NGCS is applied to reasonably smooth pavement surfaces, either in new construction or after diamond grinding (128). Although construction costs are higher than conventional diamond grinding due to the need for up to three equipment passes, NGCS is suitable for both pavement restoration and new construction projects (72).



Figure 28: Field picture for texture constructed in hardened PCC using NGCS.

1.3. Machine Learning

This section introduces machine learning and its potential for determining pavement surface types using field texture or picture data. Before proceeding, it is crucial to use proper machine learning terminology: the dependent variable is referred to as the target, and the independent variables as features.

Machine learning is a subfield of artificial intelligence focused on developing algorithms that can automatically learn from data and make predictions or decisions without being explicitly programmed (129,130). Machine learning algorithms can be used for both regression and classification problems. Regression involves predicting a continuous numerical value given a set of input features (130). For instance, a regression model might predict the price of a house based on its size, location, and other features.

Classification, on the other hand, involves assigning a class label to an input sample based on a set of features (130). Examples include predicting whether an email is spam or categorizing images into different classes such as “dog,” “cat,” or “person.” For this particular study, the Performing Agency is interested in predicting pavement surface type (a discrete outcome) given a set of features about its surface texture. Thus, classification algorithms will be the main focus of this literature review.

1.3.1. Types of Learning

Given the focus on “learning” in machine learning, there are various types of learning, often referred to as problems, encountered by practitioners. Some types describe entire subfields like “supervised learning,” while others, such as “transfer

learning,” leverage existing models. Although there are approximately 14 recognized types, this study will focus on the two main learning problems.

The first learning problem is supervised learning, where the algorithm is trained on a labeled dataset, with correct outputs provided for each input sample. The goal is to learn the mapping from inputs to outputs, enabling the algorithm to generalize to new, unseen data and make accurate predictions (129).

The second type is unsupervised learning, where the algorithm is not given labeled data and must find structure in the data independently. The goal is to identify patterns or relationships for tasks such as dimensionality reduction, clustering, or association rule learning (129). Examples include clustering analysis, principal component analysis, and autoencoders.

Each type of learning has its advantages and disadvantages, and the choice of algorithm depends on the specific problem and data available. This study emphasizes supervised and unsupervised learning algorithms, as these have been most used in determining pavement surface types.

1.3.2. Supervised Learning (SL): Classification Models

Supervised learning (SL) methods instruct models to map a set of features to their corresponding labels, utilizing numerous frameworks for this mapping. These frameworks are:

- ***Distance-based methods***, including k-nearest neighbors (KNN) and Support Vector Machines (SVMs), classify instances based on their proximity to other labeled instances. KNN utilizes a simple distance measurement to predict a class based on the nearest neighbors, while SVMs create a hyperplane that maximally separates the classes in the feature space.
- ***Probability-based models***, such as Naive Bayes and Generalized Linear Models (GLMs), rely on statistical techniques to estimate the probabilities of data points belonging to various classes based on their input features. These models are particularly effective when there is a clear statistical link between the input features and the output labels and are most useful when prior probabilities are well understood.
- ***Tree-based methods*** utilize hierarchical models to segment data and make predictions, with Decision Trees being a foundational technique within this category. These methods work by breaking down the dataset into smaller subsets through a series of feature-based decisions, forming a tree structure

where each node represents a decision point, and each leaf represents a prediction.

- **Ensemble methods** enhance predictive accuracy in machine learning by combining multiple models, termed “weak learners,” into a more robust “strong learner.” These techniques operate on the principle that a group of weak models can collectively outperform individual ones. Common techniques include *bootstrap aggregating*, *boosting*, and *stacking*. Ensemble methods are valued for their ability to reduce errors and provide more reliable and consistent predictions.
- **Network-based** models, namely neural networks, are distinguished by their ability to model complex, nonlinear relationships through layers of interconnected nodes, or neurons. Each neuron processes inputs from previous layers and passes its output to subsequent layers, enabling the network to learn deep representations of the data through a process called backpropagation, which adjusts the weights of connections to minimize prediction errors.

Figure 29 illustrates five foundational frameworks that employ unique principles to address diverse predictive challenges. It is important to note that this is not an exhaustive list, as more models are being developed as the field evolves. The subsequent section provides an overview of the traditional SL models under each of these frameworks.

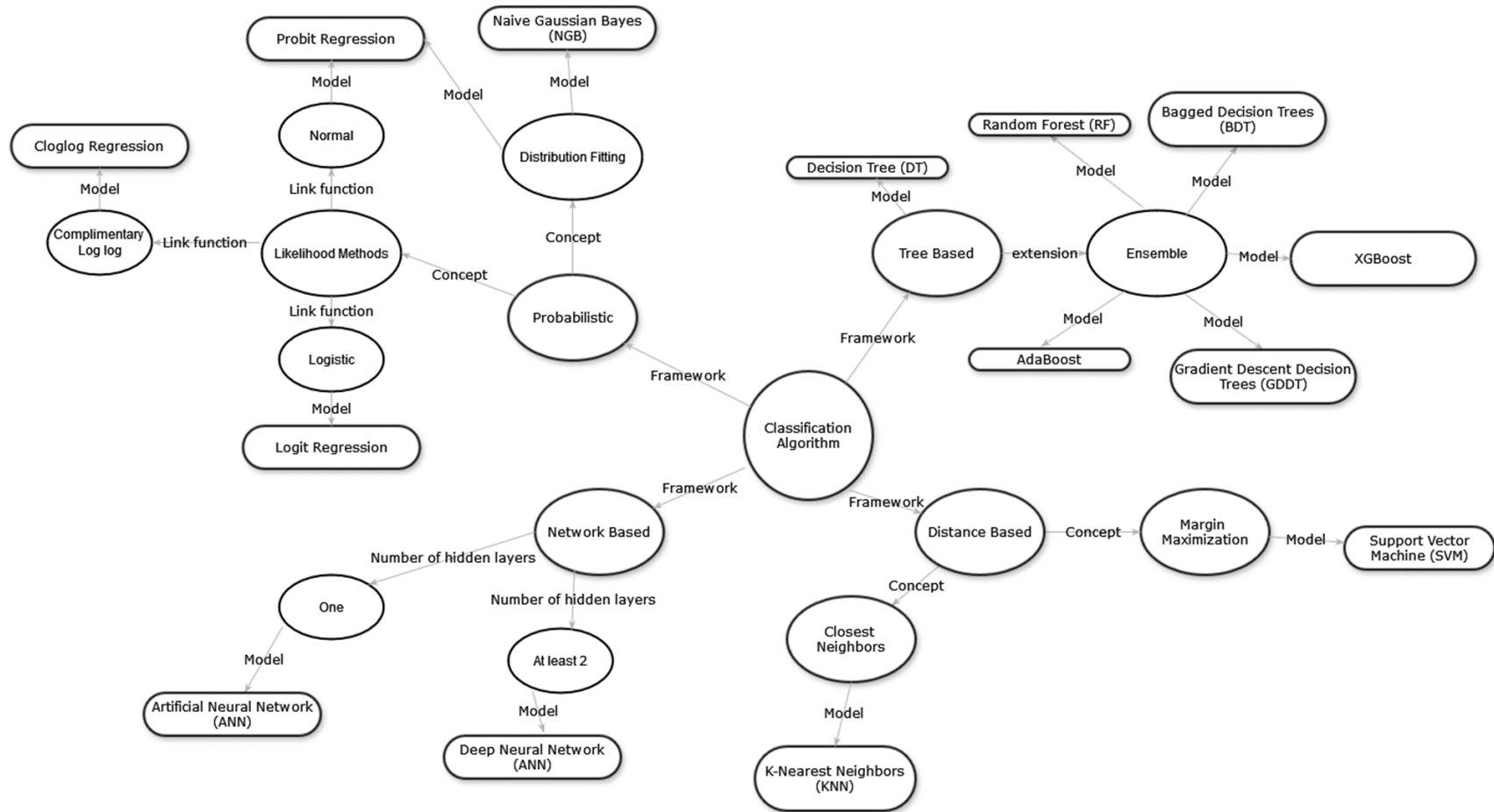


Figure 29: Diagram summarizing the traditional machine learning algorithm which can be used to predict pavement surface type based on texture indices.

1.3.2.1. K-Nearest Neighbor (KNN)

KNN is a simple and popular non-parametric machine learning algorithm used for classification and regression problems, although it is mostly applied to classification problems. Non-parametric models do not make assumptions about the data distribution (131,132). The KNN algorithm is based on the idea that an instance's behavior or characteristics can be determined by the behavior or characteristics of its neighbors (133). It employs instance-based learning to find patterns in features that map to labels. Instance-based learning implies that the model memorizes every feature vector for all observations in the training dataset and uses that knowledge to make predictions on unseen data. Because this model relies on memorization, its learning phase is considered instantaneous and straightforward (132).

The KNN algorithm works by computing the distance between feature vectors for each observation and creating a distance matrix. The algorithm then selects a number “K” of the closest neighbors to the observation being predicted. The closest neighbors are those with the smallest entries in the distance matrix. Once the neighbors are selected, the model uses their labels to make a prediction. In regression-KNN, the prediction is made by taking the mean of the labels for all nearest neighbors. In classification-KNN, inference is made using a majority voting system (133), as shown in **Figure 30**.

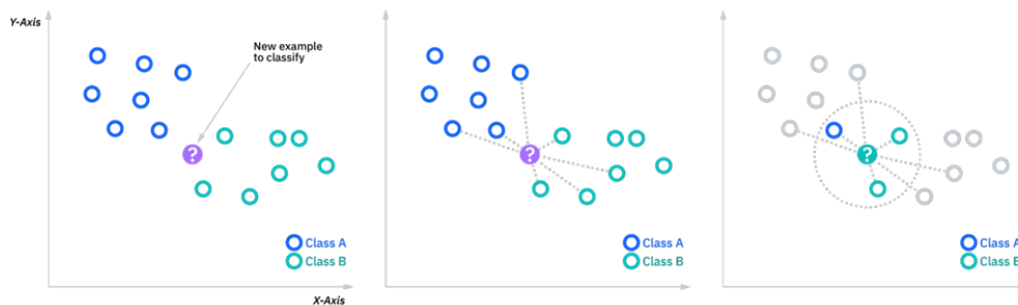


Figure 30: Example of the KNN classification algorithm.

The key parameters of the KNN algorithm are the number of nearest neighbors (K) and the distance metric used to calculate similarity between instances. Common distance metrics include Euclidean distance, Manhattan distance, and Cosine similarity (132). KNN assumes that objects or subjects in close proximity are similar, whereas those far apart are dissimilar. This assumption indirectly implies that the algorithm assumes all features provided to the model are independent of each other and highly correlated with the label.

The KNN model offers several advantages. First, it has an instantaneous training phase because the algorithm simply memorizes distances between data points.

Secondly, the model requires only simple distance calculations to explain its predictions, making it straightforward to understand. Thirdly, KNN is a “white box” algorithm, meaning it is easy to interpret. Lastly, it can predict complex behaviors using simple mathematics.

However, the KNN model also has some limitations. Since it memorizes data, this can be a limiting factor if the dataset is large, and the value of K is high. Additionally, the model is highly susceptible to the curse of dimensionality, making it unsuitable for datasets with many features. There is no agreed-upon method to choose the optimal K , which can affect the model’s performance. KNN struggles to make accurate predictions for outlier data points and, even if the model provides high accuracy, it offers no insight into how the data is generated. Finally, its reliance on the notion of distance can be influenced by the scale and distribution of features (132).

1.3.2.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression problems. It aims to find a hyperplane, or decision boundary, that best separates different classes of data points. The optimal hyperplane is determined by finding the values of the SVM’s parameters that control the margin and kernel function. The margin is the distance between the hyperplane and the nearest data points from each class, and the goal of the algorithm is to maximize this margin while minimizing classification errors (134).

The kernel function maps the input data into a higher-dimensional space where it can be more easily separated by the hyperplane. In a binary classification problem, the SVM algorithm finds the hyperplane with the largest margin, as shown in **Figure 31**. The maximum margin is the distance between the closest data points from each class and the hyperplane. These closest points, called support vectors, define the hyperplane (130).

For non-linearly separable data, the SVM algorithm can transform the data into a higher-dimensional space using kernel functions, making it possible to find a linear hyperplane that separates the data (134,135). This approach allows SVM to handle complex classification tasks effectively.

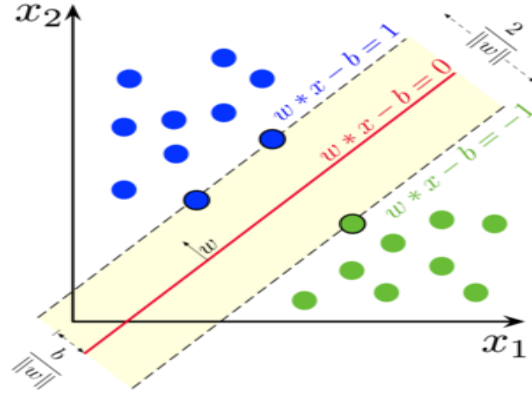


Figure 31: Maximum-margin hyperplane and margins for an SVM trained with samples from two classes (136).

One of the main advantages of SVM is its ability to handle high-dimensional data effectively. SVM maximizes the margin between classes, resulting in more accurate and robust classification, even when the number of features exceeds the number of samples (130,135). Additionally, SVM is versatile and supports different kernel functions, such as linear, polynomial, and radial basis function, allowing it to fit various types of data. Another advantage of SVM is its robustness to outliers and noise. By maximizing the margin between classes, outliers are less likely to fall within this margin, making SVM perform better on noisy data compared to other algorithms (134).

However, SVM has several disadvantages. One major drawback is its computational complexity. SVM can be slow when working with large datasets and requires significant memory, as it needs to solve a quadratic optimization problem for every sample in the dataset (135). This makes SVM computationally expensive and impractical for large datasets. Additionally, SVM is sensitive to parameter tuning, requiring careful selection of the kernel function and tuning of parameters such as the regularization parameter and kernel coefficient. This can be time-consuming and challenging, especially for non-experts. Another disadvantage is that SVM is inherently a binary classification algorithm, capable of classifying data into only two classes. To extend it to multi-class classification, techniques like one-vs-one or one-vs-all are used, but these can be computationally expensive and may not scale well for large datasets (130). Lastly, SVM's interpretability is limited because it produces a black-box model, making it challenging to understand how predictions are made. This can be a significant disadvantage in applications where interpretability is essential. However, techniques such as feature importance measures can help understand the contribution of each feature to the final prediction (130).

1.3.2.3. Gaussian Naive Bayes (GNB)

Naive Bayes is a probabilistic machine learning algorithm based on Bayes' theorem, primarily used for classification problems. It is termed "naive" due to its strong independence assumption between features, assuming that features are independent and have no impact on each other, despite often being correlated in practice (130). The GNB model extends the Naive Bayes model by assuming that each class within the labels has a unique Gaussian distribution, hence the name.

In the GNB classifier, the goal is to estimate the class-conditional probabilities, assuming that the features for each class follow a Gaussian distribution. Given a new instance with a set of features, the algorithm calculates the posterior probability of each class label using Bayes' theorem and assigns the label with the highest posterior probability to the instance. Gaussian Naive Bayes is one of several variants of the Naive Bayes algorithm, with others including Multinomial Naive Bayes and Bernoulli Naive Bayes, each suited to different types of data and feature distributions (130,137).

One of the main advantages of GNB is its ease of implementation and computational efficiency, making it well-suited for large datasets. It performs well with high-dimensional data and is particularly effective when the number of features is much greater than the number of observations. GNB also provides probabilistic predictions, which can be useful in decision-making (137,138).

However, GNB has some limitations. Its assumption that all features are normally distributed and independent is not always true in real-world data, leading to suboptimal performance when features are correlated. The algorithm also assumes that all features contribute equally to the outcome, which may not be the case in some datasets, resulting in suboptimal performance. Moreover, GNB is prone to the "zero-frequency problem," where it assigns a probability of zero to a class when a feature is not present in the training set. This issue can be addressed through smoothing techniques. Lastly, GNB may not be the best choice for tasks where maximizing accuracy is the goal, as more complex models like decision trees or neural networks may outperform it (130,137).

1.3.2.4. Generalized Linear Models (GLM)

A GLM is an extension of traditional linear regression models that allows for response variables to have error distributions other than a normal distribution. GLMs are defined by three components: a linear predictor, a link function, and a probability distribution from the exponential family. The linear predictor is a combination of model parameters and explanatory variables. The link function connects the means of the distribution of the response variable to the linear

predictor. The exponential family encompasses a wide range of distributions, including normal, binomial, Poisson, and gamma distributions, making GLMs highly versatile for various types of data.

This report will focus specifically on logistic, Probit and complementary log-log regression models as these are commonly used in discrete choice modeling, survival analysis and ML classification problems

1.3.2.4.1. *Logistic Regression (LOG)*

LOG is a statistical method used for binary classification problems, though it can be extended to multi-class classification. The goal is to model the relationship between independent features and a binary target (139). It uses the logistic function, or sigmoid function, to map independent variables to a probability range of 0 to 1, representing the likelihood of the input data belonging to a certain class. The model is trained using maximum likelihood estimation to find the coefficients that best fit the observed data (139).

After training, the model can make predictions for new instances by inputting features and computing the predicted class probability. A classification threshold, often set at 0.5, determines the final class decision. LOG is popular due to its simplicity, interpretability, and ability to handle both linear and non-linear relationships. It is computationally efficient and easy to implement. However, it has limitations, such as sensitivity to outliers, the requirement for feature independence, and difficulty in capturing complex non-linear relationships (139).

1.3.2.4.2. *Probit Regression (PRO)*

PRO is a statistical method for binary classification problems similar to LOG, with the primary difference being that it uses the standard normal cumulative distribution function (Probit function) to model the relationship between independent variables and the binary dependent variable. In PRO, the relationship between input variables and the binary outcome is modeled as a linear combination of the inputs, with coefficients estimated through maximum likelihood estimation (140). The estimated coefficients are used to calculate the predicted value of the cumulative distribution function of the standard normal distribution, which is then transformed into a probability value between 0 and 1.

One advantage of the Probit model is its ability to estimate thresholds or cut-off points at which the binary outcome changes, which is useful when the relationship between inputs and the binary outcome is not linear. PRO also provides a natural probabilistic interpretation of the predicted probabilities due to the Probit function (140).

However, PRO is less widely used than LOG because it is computationally more intensive and less interpretable. Additionally, it assumes that errors are normally distributed, which may not be suitable for all datasets (140).

1.3.2.4.3. *Complimentary Log-Log Regression (CLL)*

CLL is a type of regression analysis used in binary classification problems. It is a variant of LOG, where the logistic transformation is replaced by the complimentary Log-Log function. This function maps predicted probabilities to a space that better models the probabilities of extreme events (i.e., events close to 0 or 1), providing a better fit for datasets with rare target variables or many outliers (140).

The coefficients of the independent variables in CLL are estimated using maximum likelihood estimation. Once the model is trained, it can make predictions for new instances by inputting the features and computing the predicted CLL transformed probabilities. A threshold, such as 0.5, is set to convert these probabilities into binary class predictions (140).

Although less widely used than LOG, CLL is useful when the target variable is rare or when there are many outliers. However, it is more computationally intensive and can be more difficult to interpret compared to LOG or PRO (140).

1.3.2.5. **Decision Tree (DT)**

A DT is a popular and simple machine learning algorithm used for both classification and regression problems. It is a tree-like model representing a series of decisions and their possible consequences. Each node in the tree represents a test on an input feature, each branch represents the outcome of the test, and the leaves represent the final prediction or class, as shown in **Figure 32**. The decisions in the tree are based on the most important feature that splits the data into homogeneous groups (130). The algorithm recursively splits the data into smaller groups based on the feature that provides the highest information gain. This process continues until all data in each group belongs to the same class, or until the maximum tree depth or the minimum number of samples per leaf is reached. The final result is a tree with branches representing decisions based on features and leaves representing the final prediction or class. To make a prediction for a new instance, the instance is passed down the tree, following the branches corresponding to the feature values until a leaf is reached, providing the final prediction (130).

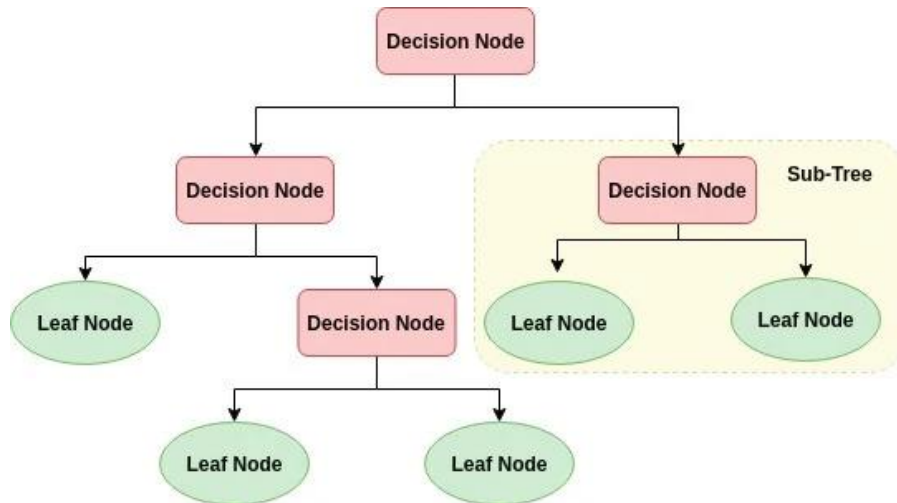


Figure 32: Basic structure of a DT.

DTs are widely used due to their interpretability and ease of use. One of their biggest advantages is their interpretability, as the tree structure makes it easy to explain results to a non-technical audience. They can handle missing values and irrelevant features, as not all features need to be relevant for the prediction task. Additionally, DTs are non-parametric, meaning they do not make any assumptions about the underlying data distribution (141). They are also fast to train and make predictions, making them suitable for real-time applications ((130).

However, DTs also have some disadvantages. One major challenge is overfitting, especially when the tree grows very deep, resulting in a complex structure that may not generalize well to unseen data. DTs can also be unstable, as small changes in the data can lead to a completely different tree structure and different predictions. They can be biased towards features with many levels or towards classes with more samples in the training data. Additionally, DTs are limited in their expressiveness, as they can only capture linear decision boundaries. Finally, the algorithm used to build DTs is a greedy approach, which may not always result in the optimal tree structure (141).

1.3.2.6. Random Forest (RF)

RF are an ensemble learning method that constructs multiple decision trees during training and outputs the modal class for classification or the mean prediction for regression. The random component arises from training each decision tree with a different, randomly selected subset of the original data, with each subset containing a random set of variables to ensure low correlation among the trees (142,143). A diagram showing the general structure of the RF is shown in **Figure 33**.

The fundamental concept behind a RF model is the “wisdom of crowds.” This approach works well because many relatively uncorrelated models operating as a committee will outperform any individual constituent model. The low correlation between models is crucial. Similar to how a diversified investment portfolio with low-correlated assets (e.g., stocks and bonds) is more robust, uncorrelated models in a RF produce ensemble predictions that are more accurate than individual predictions. This is because the trees protect each other from their individual errors. While some trees may be wrong, many others will be correct, allowing the forest to collectively make more accurate predictions (142,143).

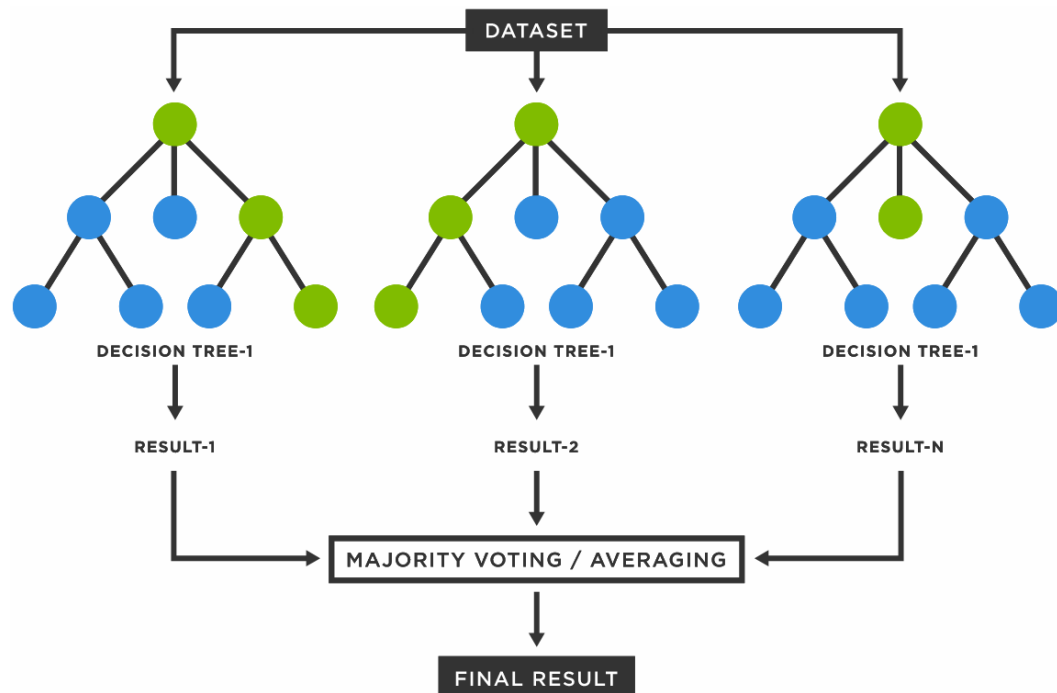


Figure 33: Basic structure of a RF.

RF models offer several advantages over single DTs, including improved accuracy, reduced overfitting, better handling of noisy data and outliers, and the ability to estimate feature importance. It is also relatively fast and easy to implement, making it a popular choice for many applications. However, RF can be computationally expensive for large datasets, and its predictions may be slower than those of other algorithms because it needs to make a prediction for each tree in the forest (142,143).

1.3.2.7. Bootstrap Aggregating (BA)

BA, also known as bagged decision trees, involves creating multiple DTs using random subsets of the training data. For each tree in the ensemble, a new subset of the training data is created by randomly selecting examples from the original

training set with replacement, meaning some examples may appear multiple times in the new subset while others may be left out. Each DT in the BA model is then trained using this new subset of training data, and a different subset of input features is used for each node split in the tree to increase diversity and prevent overfitting (144).

During prediction, each DT in the BA model independently evaluates the input features and generates a prediction. The final prediction is determined by aggregating the predictions of all individual trees, typically by averaging in regression problems or taking the majority vote in classification problems. By combining the predictions of multiple DTs, BA reduces the variance of the model and improves its accuracy. Additionally, because each tree is trained on a different subset of the training data, BA helps reduce the impact of outliers and noise in the data (144). **Figure 34** shows the basic structure of a bagged decision tree.

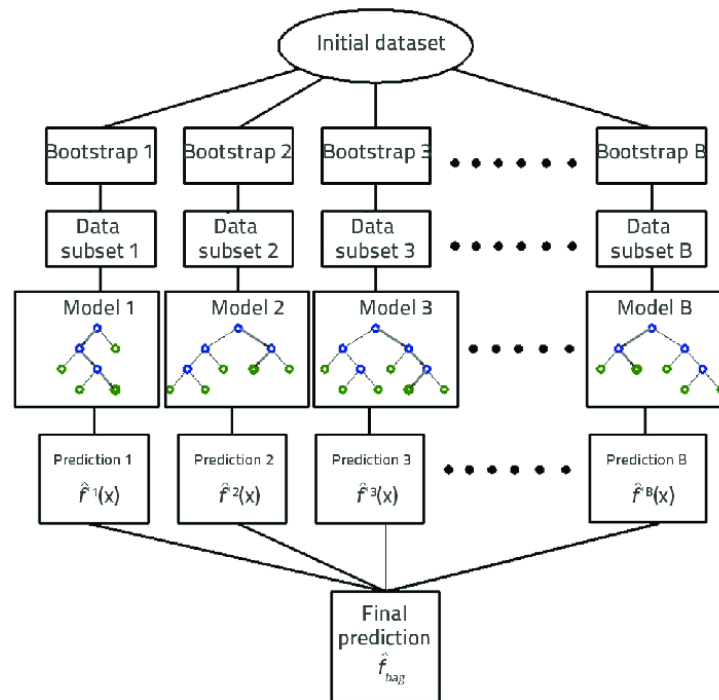


Figure 34: Basic structure of a BA model.

BA models offer several advantages over single DT, such as improved accuracy, reduced overfitting, better handling of noisy data and outliers, and the ability to estimate feature importance. They are also relatively fast and easy to implement, making them a popular choice for many applications. However, BA models can be computationally expensive for large datasets, and their predictions may be slower than those of other algorithms because a prediction must be made for each tree in the bag (144).

1.3.2.8. Adaptive Boosting (AB)

AB, also known as AdaBoost, is a machine learning algorithm used for binary classification problems. It is a meta-algorithm that combines several weak learners, such as DTs with small depth, to form a strong learner. The algorithm iteratively fits weak learners on weighted versions of the training data, updating the weights of instances in each iteration to give more importance to the misclassified instances. The final prediction is obtained by weighted voting of the weak learners (145,146).

In each iteration, AB fits a weak learner on the weighted training data and assigns a weight to the weak learner based on its accuracy. The weights of the instances are updated so that misclassified instances have higher weights in the next iteration. This process continues until a stopping criterion is met, such as a maximum number of iterations or a minimum error rate. The final prediction is made by weighted voting of the weak learners, where the weight of each weak learner is proportional to its accuracy, giving more importance to those that perform well on the training data (146). **Figure 35** shows a visual example of how AB works.

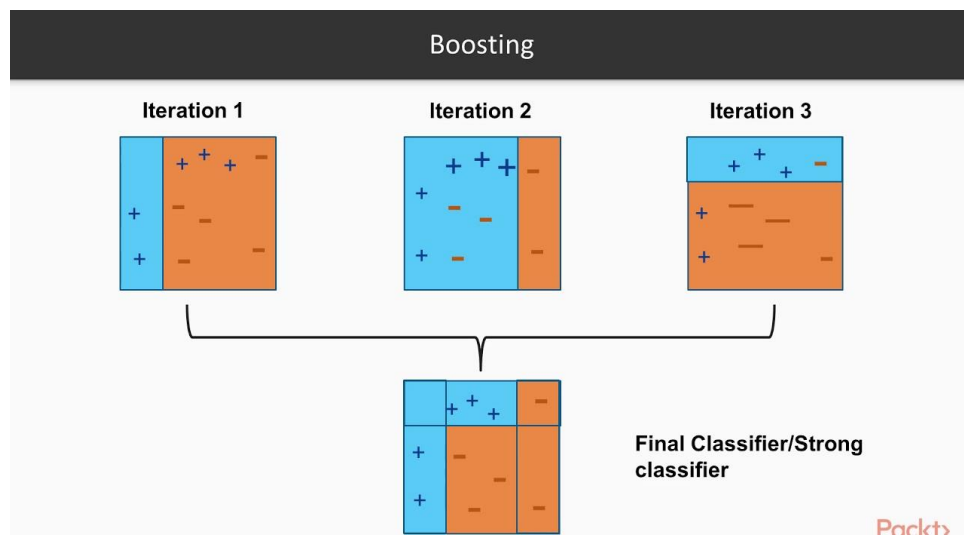


Figure 35: Visual depiction of AB using multiple weak learners to create one strong classifier.

One of the main advantages of AB is its ability to handle both linear and non-linear decision boundaries, making it suitable for a wide range of data distributions and types. Additionally, AB is not prone to overfitting, can handle noisy data and outliers effectively, and is computationally efficient, making it suitable for large-scale applications (146).

However, AB has some disadvantages. It is sensitive to the choice of weak classifiers and the number of iterations; if weak classifiers are not chosen carefully, or if too many iterations are used, the algorithm can become overly complex and

perform poorly (146). AB is also sensitive to class imbalance, which can lead to poor performance if one class has many more samples than the other. Additionally, AB requires significant hyperparameter tuning, which can be time-consuming and requires expertise. Finally, AB can be sensitive to outliers, as they can significantly impact the algorithm's performance (145).

1.3.2.9. Gradient Boosting (GB)

GB is a machine learning algorithm used for both classification and regression problems. It is an ensemble learning algorithm that combines several simple weak learners, such as SVMs or DTs, to form a strong learner. The algorithm works by iteratively fitting weak learners on the training data and adjusting their weights based on the errors made in previous iterations. The final prediction is obtained by weighted voting of the weak learners, where the weight of each weak learner is proportional to its accuracy (147). GB can achieve higher accuracy than single weak learners and other machine learning algorithms and is widely used in applications such as credit scoring, customer segmentation, and disease diagnosis (146). **Figure 36** shows the basic architecture of GB using DT as weak learners.

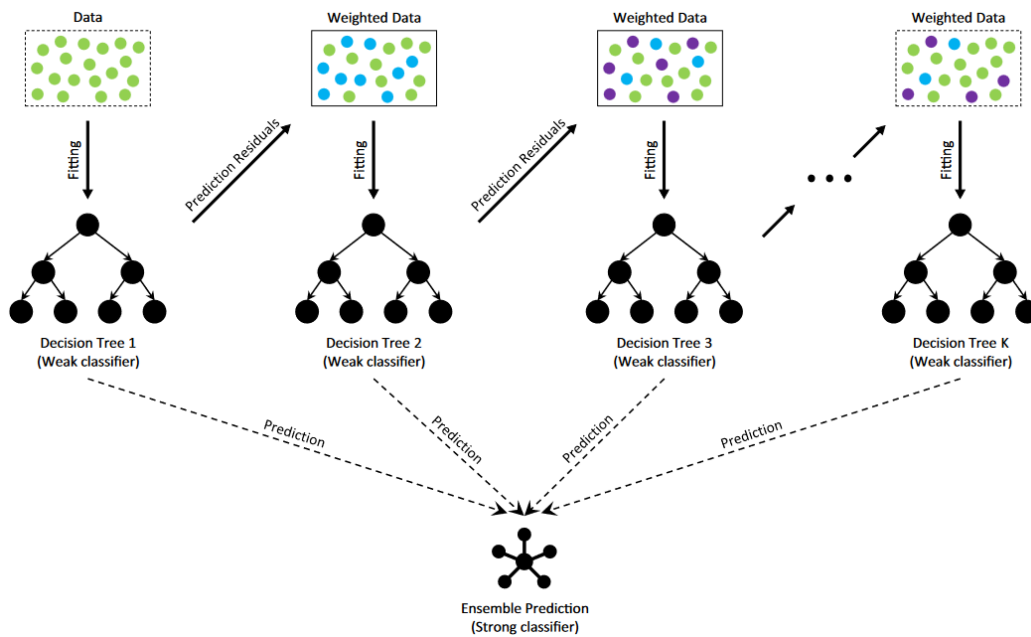


Figure 36: The architecture of GB (148).

One of the main advantages of GB is its ability to handle large datasets with many features and instances, as well as their capability to manage noisy and complex data. The algorithm is relatively fast and easy to implement and can be parallelized for even faster computation. However, GB can be sensitive to overfitting, especially if the number of iterations is too large or if the weak learners are too complex. It is

also computationally expensive, and its predictions may be slower than those of other algorithms (146,147).

1.3.2.10. Extreme Gradient Boosting (XGB)

XGB, also known as XGBoost, is a popular gradient boosting library widely used for both regression and classification problems. It is highly customizable, with many parameters that can be tuned to optimize performance for different datasets and problems. XGB builds an ensemble of decision trees sequentially, each tree correcting the errors of the previous one. During training, the algorithm calculates the gradient of the residuals and uses this information to adjust the weights of each instance in the dataset (149).

XGB uses parallel processing and memory optimization to make the training process efficient and scalable. It implements advanced techniques for handling overfitting and improving prediction accuracy, such as regularization, early stopping, and tree pruning. It also supports GPU acceleration, making it much faster than other gradient boosting libraries (149). **Figure 37** shows a schematic for how XGB works.

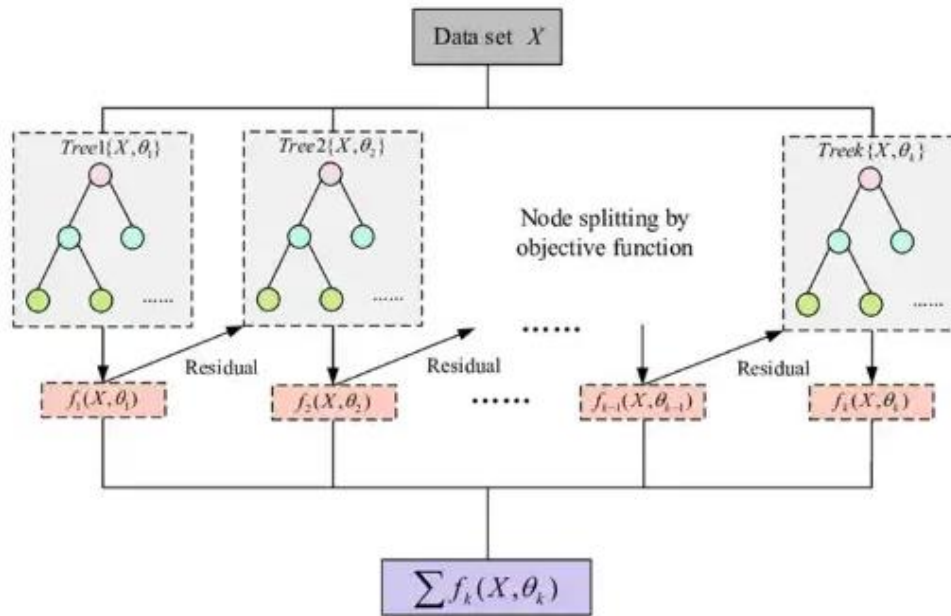


Figure 37: Flowchart of XGB algorithm (150).

In addition to its efficiency and scalability, XGBoost is known for its interpretability, providing feature importance scores that help understand each feature's contribution to the prediction. However, XGB can be sensitive to noisy

data and requires careful tuning of hyperparameters, such as the learning rate and the number of trees, to achieve good results (149).

1.3.2.11. Neural Networks

Neural networks are a subset of machine learning that mimics how the human brain learns by using a network of connected nodes (neurons) to output a prediction. The general architecture of an artificial neural network (ANN) comprises three types of layers: the input layer, hidden layers, and the output layer (130) as shown in **Figure 38**. The input layer consists of the data fed into the model, such as images of handwritten numbers for a digit recognition task. The hidden layers, which can range from a single layer to many layers depending on data complexity, contain neurons that learn specific features of the data. Neurons perform mathematical calculations optimized with each feature to reduce the error between the model's prediction and label.

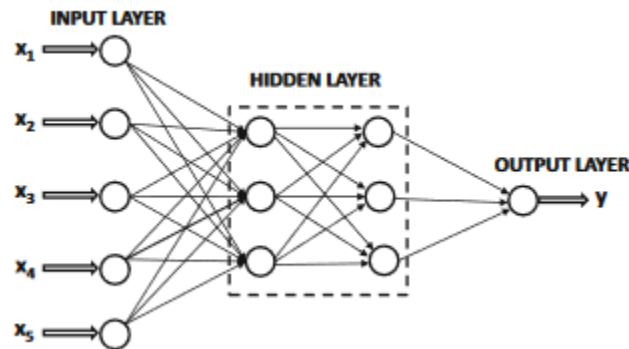


Figure 38: Representation of an ANN Model with 2 hidden layers (151).

Typically, neural network models are supervised, meaning they are trained on labeled data. The data is usually split into training, validation, and testing sets in a 70%, 20%, and 10% ratio, respectively. The training data adjusts the numerical values in the neurons, while the validation data monitors the model's performance during training. The output layer, connected to the last hidden layer, makes the final decision about the input data. It produces a vector where each value represents the confidence that the input belongs to a particular class, with the highest value indicating the predicted class (130).

When an ANN has a single hidden layer, it is considered a shallow neural network (SNN). In contrast, when there are two or more hidden layers, it is referred to as a deep neural network (DNN). DNNs are capable of capturing more complex patterns in the data, making them suitable for tasks requiring high-level abstractions, such as image and speech recognition. The depth of a neural network allows it to learn

hierarchical representations, improving its ability to generalize from training data to unseen instances.

However, even with the improved predictive power provided by DNNs relative to SNN. It is important to know that SNNs are widely used due to the universal approximation theorem, which states that a neural network can approximate any function with at least one hidden layer (152). This implies that any computational function can be approximated by a SNN. However, the theorem does not specify the optimal architecture needed to best approximate the function (152), leaving it up to the data analyst to determine the best structure. Despite their versatility, ANNs are often considered “black boxes” because their internal workings do not provide insights into the structure of the approximated function. Nevertheless, for many high-complexity tasks, the ability to predict outputs accurately is more valuable than understanding the underlying structure.

1.3.2.11.1. *Convolutud Neural Networks*

Convolutud Neural Networks (CNN) are a specialized type of neural network model that excel in image classification for computer vision and image analysis. In fact, picture-based classification is largely dominated by neural networks, like the CNN, as other methods like RF, KNN, and GNB achieve poorer accuracy.

CNNs extract distinct features from images, enabling the classification of various classes within the data. Unlike traditional ANNs, CNNs are particularly effective for pattern recognition in images. For example, a color image of a cat might be 32 x 32 pixels with three channels (R, G, B). While an ANN architecture could require neurons for each pixel in each channel, resulting in high computational time, a CNN can condense the input data into smaller sizes as it progresses through the layers, making it more efficient (153). **Figure 39** shows how the convolutional layer fits into the ANN’s architecture.

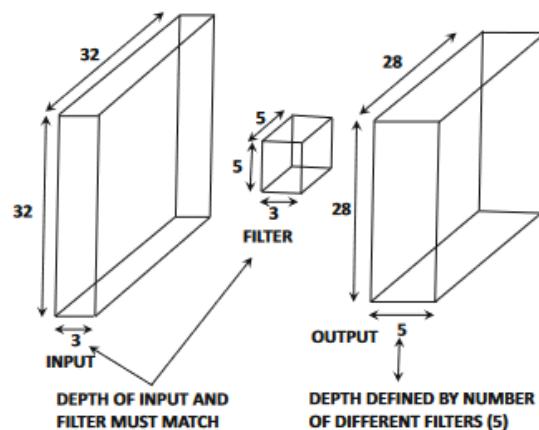


Figure 39: An input image, followed by a convolutional layer and the output of the convolutional layer (151).

The overall architecture of a CNN is similar to an ANN, but the hidden layers in ANNs are replaced by three types of layers in CNNs: convolutional layers, pooling layers, and fully connected layers (**Figure 40**). The input and output layers function the same as in ANNs. The convolutional layer takes smaller segments from the image and applies filters/kernels learned during training. to extract low-level features, such as edges, used for classification. The pooling layer then undersamples the data, reducing the number of calculations needed. Finally, the fully connected layer produces a confidence score for the image class, similar to the output in traditional ANNs. These layers can be repeated multiple times within the model (153).

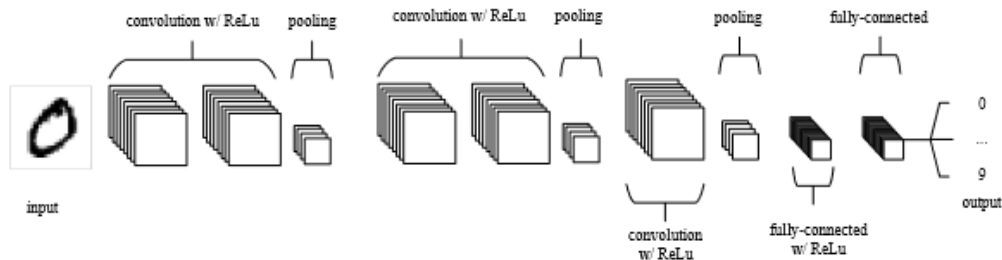


Figure 40: A CNN architecture to classify handwritten numbers (153).

CNN models involve fewer neuron calculations compared to ANNs for image data. However, CNNs have many parameters that require tuning, often making the process trial-and-error to find the optimal architecture. As the model becomes more complex with more layers, the computational power required increases significantly, sometimes beyond the capacity of a regular personal computer. Training time for CNNs is also dependent on computational power. Overfitting is a common issue, where the model starts to memorize the training data rather than learning generalizable patterns. This can be mitigated through various tools and methods designed to prevent overfitting. Balancing the input complexity and the number of layers is crucial to maintaining a model's effectiveness without overfitting.

1.3.3. Unsupervised Learning

Unsupervised Learning (UL) is a branch of machine learning that trains algorithms using data without predefined labels or outcomes determined by human input. UL is categorized into three primary types: clustering, association rule learning, and dimensionality reduction. *Clustering* groups objects so that those within the same group are more similar to each other than to those in other groups, enhancing pattern

recognition and classification within datasets. *Association rule learning* identifies significant relationships between variables in large databases, uncovering strong rules using various measures of interestingness. *Dimensionality reduction* simplifies data by focusing on principal variables, reducing the number of variables considered without substantial information loss, which supports data compression and feature extraction.

1.3.3.1. Dimensionality Reduction

PCA is a dimensionality reduction technique designed to simplify the complexity of high-dimensional data while preserving essential trends and patterns. It achieves this by forming linear combinations of the original data to create a new set of variables, known as principal components. The first principal component captures the maximum amount of variance from the original variables, with each subsequent component accounting for the maximum amount of remaining unexplained variance. These components are uncorrelated and ordered, ensuring that the initial few components retain the majority of the variation found in all the original variables.

In the context of clustering, PCA facilitates the projection of high-dimensional data into a lower-dimensional space with minimal information loss, enhancing data interpretability. One effective method to visualize PCA's efficacy is by comparing scatterplot matrices of the first three principal components against those of all the original variables. To illustrate the benefits of PCA, consider an example involving high-dimensional data. **Figure 41** displays a density scatterplot of the original data containing six variables (top) and the first three principal components (bottom).

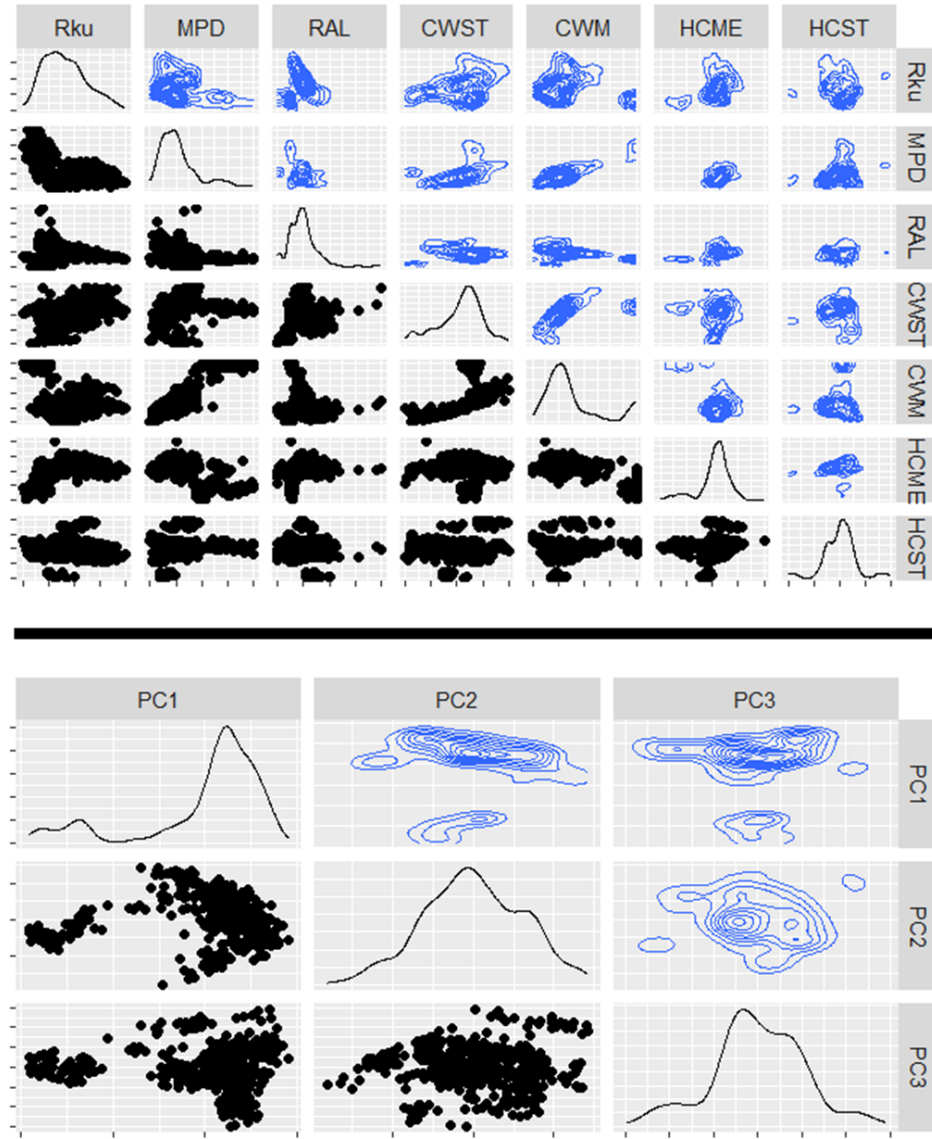


Figure 41: Comparison between scatterplot of the original data (top) and the reduced data using the first three principal components (bottom).²

The scatterplot matrices are divided into three distinct regions. The main diagonal illustrates the distribution of each variable, providing a straightforward visual representation of individual data trends. The off-diagonal elements on the bottom left consist of simple scatterplots that showcase relationships between two variables. In contrast, the off-diagonal elements on the top right consist of density plots, which are useful for highlighting areas of high density and identifying clusters.

² Note that the numerical values on both axes were suppressed due to excessive overlap at this resolution which distracts from cluster visualization.

While both the original data scatterplot matrix (top) and the principal component scatterplot matrix (bottom) represent the same data, the complexity of the top section, which includes a larger number of variables, renders it more difficult to interpret. The PCA scatterplot matrix, on the other hand, provides a clearer and more interpretable visual due to its reduced dimensionality, featuring only three variables instead of six. This simplification greatly enhances the ease of understanding the underlying data structure and the relationships between variables.

Reducing the number of variables from six to three often suffices to reveal the clustering structure within the data, particularly by examining the density plot of the first and second principal components. These components typically capture most of the variance from the original variables. **Figure 42** zooms into the top middle element of the PCA scatterplot matrix to showcase the density plot for these two principal components. Visually, at least three distinct groups are easily identifiable: a large cluster at the top, a smaller cluster in the mid-right, and another cluster in the bottom middle. The clarity of these groupings in the density plot provides valuable insights that inform and guide the subsequent application of clustering algorithms, effectively leveraging the streamlined data structure achieved through PCA to improve clustering accuracy.

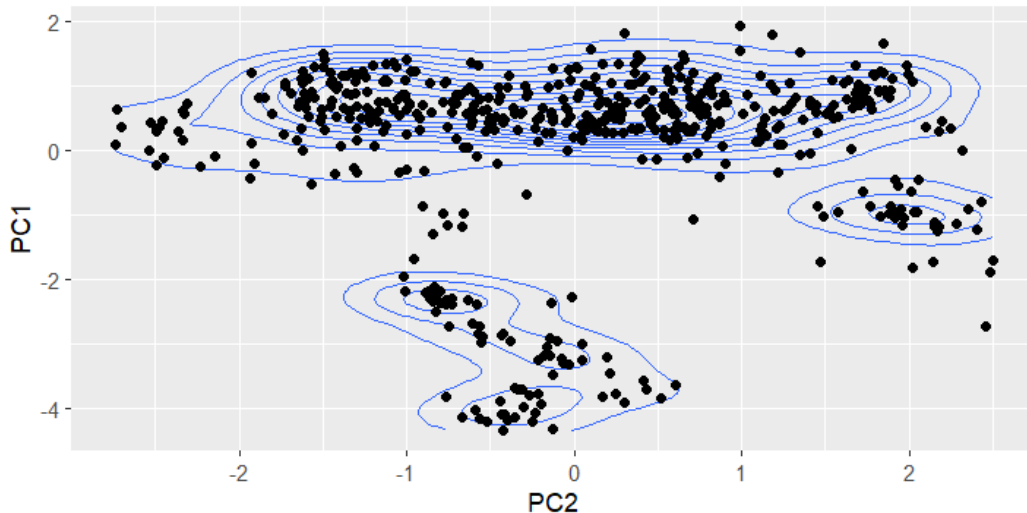


Figure 42: Density plot overlaid by scatter point of the first and second principal components for texture indices in rigid pavements.

1.3.3.2. Clustering

Clustering methods are distinct from discrimination or assignment methods—types of supervised machine learning used for classification—where the groups are predefined. Cluster analysis focuses on identifying and revealing naturally

occurring groups within the data, often referred to as clusters. A cluster is generally characterized by the homogeneity of its members (i.e., internal cohesion) and their distinct separation from members of other clusters (i.e., external isolation). Such properties are illustrated informally with a diagram in **Figure 43**. The clusters present in this figure will be clear to most observers without attempting an explicit formal definition.

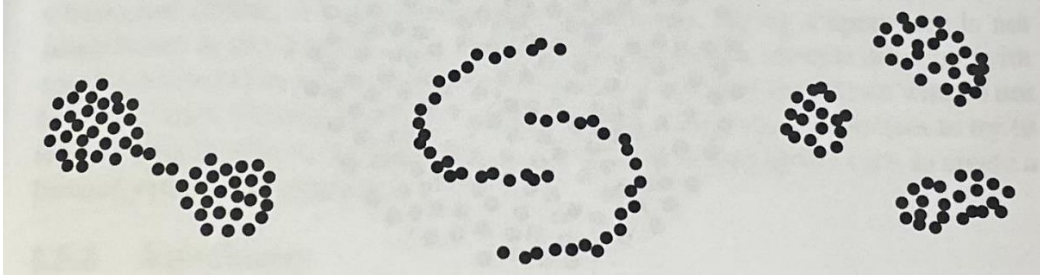


Figure 43: Clusters with homogeneity and/or separation (154)

Identifying clusters mathematically is significantly more complex than merely visualizing them. The earlier definition of clusters was intentionally qualitative to highlight the challenges inherent in rigorously defining concepts such as homogeneity and separation. The difficulty in precisely defining these concepts, and the uncertainty around how the human brain visually isolates clusters, as illustrated in **Figure 43**, have led to a proliferation of clustering algorithms. The vast array of documented methods—each adapting the definition of *what* constitutes a cluster and *how* to measure homogeneity and separation—continues to grow as new algorithms are developed. **Figure 44** provides a diagram that offers a broad overview of just a few of the many clustering techniques and frameworks employed by researchers across various fields of study.

Given the sheer number of clustering algorithms available, the Performing Agency could not perform an exhaustive analysis of all techniques. Instead, the focus was placed on identifying a clustering method that not only mirrors the clusters visually discernible through PCA but also produces reproducible results when tested with an alternative method. To this end, the Performing Agency evaluated several algorithms, including:

- Partition Around Medoids (PAM),
- Agglomerative Nesting (AGNES),
- Gaussian Mixture Models (GMM), and
- Density-Based Spatial Clustering of Applications with Noise (DBSCAN).

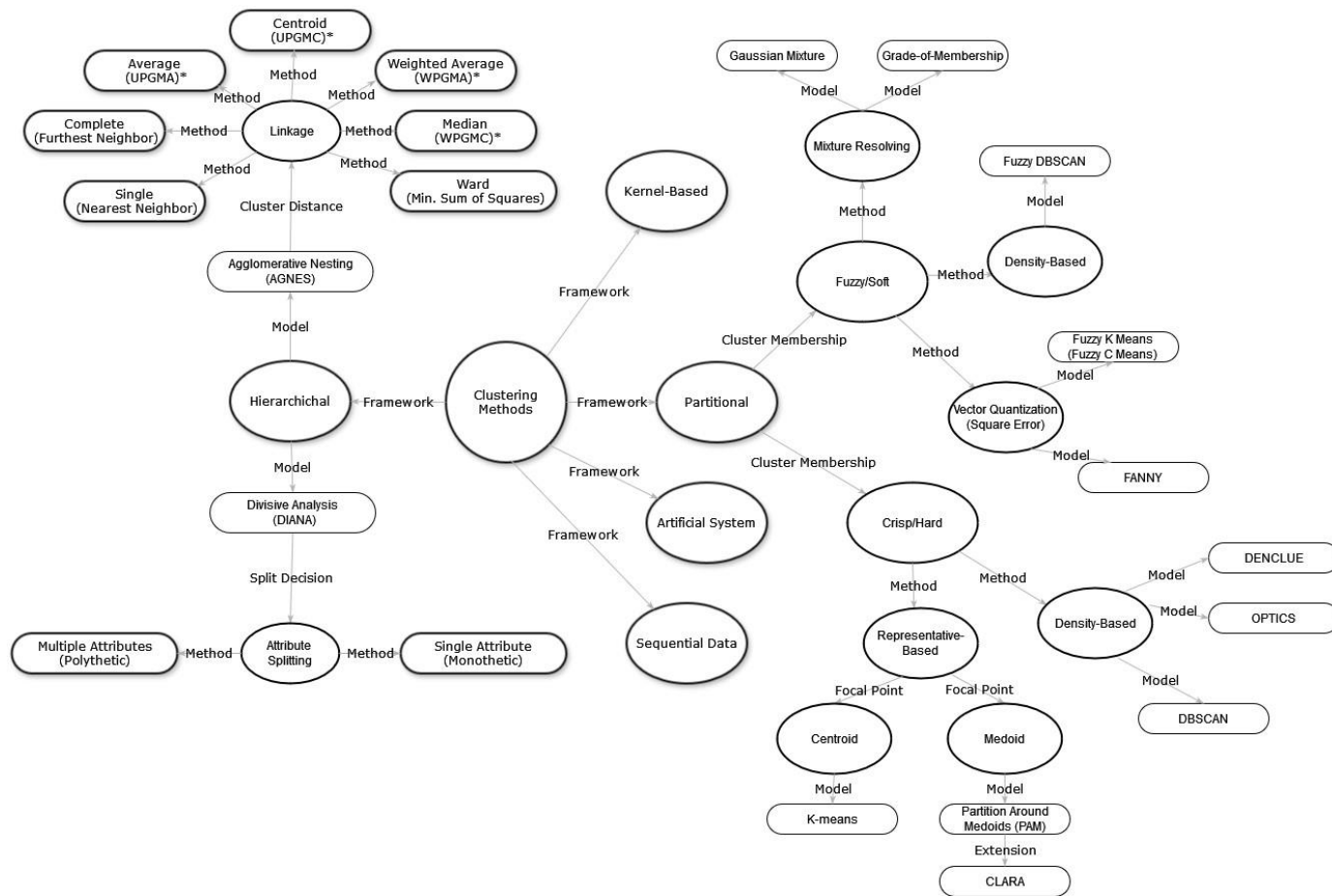


Figure 44: Clustering of clustering techniques

1.3.3.2.1. *Partition Around Medoids (PAM)*

PAM is an unsupervised clustering algorithm that organizes a dataset into clusters using actual data points as central elements, known as medoids. Unlike k-means clustering, which minimizes variance within clusters, PAM reduces the sum of dissimilarities between points and their respective medoids, making it more robust to noise and outliers. During the clustering process, PAM iteratively refines its clusters by evaluating potential medoids and selecting those that yield the most coherent and compact groupings. This method enhances the robustness of the results and provides an intuitive understanding of each cluster through real, representative examples (155).

PAM is robust to noise and outliers because it uses medoids instead of means, minimizing dissimilarity rather than variance. This makes PAM effective for datasets with irregular distributions, providing interpretable and realistic cluster representations. However, PAM can be computationally intensive, especially with large datasets, due to the need to evaluate all possible medoid candidates, resulting in higher time complexity compared to k-means. It can also be sensitive to the initial selection of medoids. Despite these challenges, PAM is valuable for datasets requiring robustness to outliers and clear cluster interpretability (156).

1.3.3.2.2. *Agglomerative Nesting (AGNES)*

AGNES is a hierarchical clustering technique that groups data points based on their similarities using an agglomerative approach. Each data point starts as its own cluster, and clusters are gradually merged based on a specified linkage criterion, such as minimum, maximum, or average distance between clusters. This process results in a dendrogram (**Figure 45**), a tree-like diagram that displays the nested grouping of points and clusters, which is invaluable for analysts seeking to understand the data structure at different levels of granularity and to identify the natural hierarchy in data. AGNES is particularly useful when the number of clusters is unknown, allowing analysts to explore the hierarchical structure and determine an appropriate number of clusters by interpreting the dendrogram. To decide which clusters should be combined, a measure of dissimilarity between sets of observations is required, typically using metrics like Euclidean distance and linkage criteria such as Ward's method, which minimizes the sum of squared differences within all clusters.

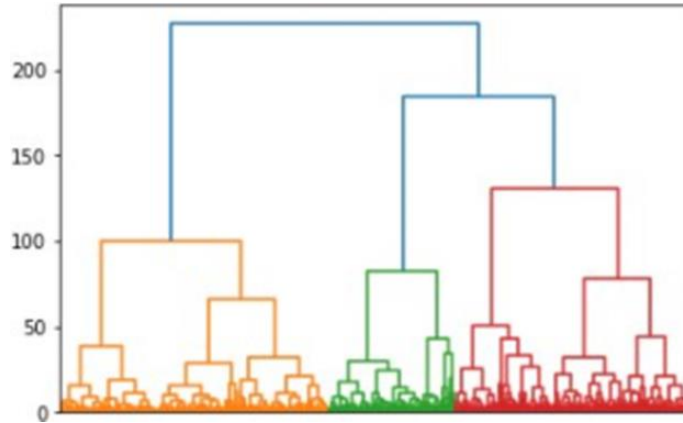


Figure 45: Sample dendrogram.

The mergers and splits in AGNES are determined using a greedy strategy, making locally optimal choices at each stage. While this approach may not always yield the global optimum, it often produces satisfactory results efficiently (157). However, the main drawback of AGNES is its computational time, with a time complexity of $O(N^3)$. Despite this, agglomerative hierarchical clustering with a specified metric and linkage criterion is guaranteed to converge to the same unique solution every time (158).

1.3.3.2.3. *Density-Based Spatial Clustering Application with Noise (DBSCAN)*

DBSCAN is a clustering algorithm that defines clusters as areas of high density separated by areas of low density. Unlike methods requiring a predetermined number of clusters, DBSCAN operates based on two parameters: *epsilon*, which specifies the neighborhood radius in terms of distance, and *minPts*, the minimum number of points required to form a dense region. This flexibility makes DBSCAN suitable for datasets with irregularly shaped and sized clusters and effectively handles noise. DBSCAN also excels at excluding outlier data points that do not belong to any cluster, providing a robust solution for real-world data clustering where outliers are prevalent (155). However, DBSCAN has disadvantages, such as struggling with datasets of varying density, as it relies on parameters like *epsilon* and *MinPts* that need careful tuning. Additionally, it can be computationally intensive for large datasets. (159).

1.3.3.2.4. *Gaussian Mixture Models (GMMs)*

GMMs are probabilistic models that posit all data points are generated from a finite number of Gaussian distributions with unknown parameters. Part of the mixture models family, GMMs offer a soft-clustering approach, calculating the probability of each data point belonging to each distribution rather than assigning each point to a single cluster. This flexibility allows GMMs to accommodate clusters of varying

sizes and correlation structures, making them particularly useful in fields such as image processing, pattern recognition, and bioinformatics, where they help elucidate underlying parameters in data exhibiting multiple modes of variation (155).

Nonetheless, GMMs have several disadvantages, including sensitivity to initial parameter values, which can lead to convergence at local optima. They assume data is generated from Gaussian distributions, making them unsuitable for non-Gaussian data or datasets with significant outliers. Moreover, GMMs are computationally intensive, especially for large or high-dimensional datasets, and require specifying the number of components in advance, which can lead to overfitting or underfitting (160).

1.3.4. Assessing Performance

Goodness-of-fit metrics are statistical measures used to evaluate how well a model fits the data it is meant to represent. These metrics compare observed values to predicted values, providing a quantitative assessment of the model's fit. The choice of metric depends on the specific problem and desired properties of the metric. It is crucial to use appropriate metrics and interpret results carefully to ensure accurate quality assessment.

In classification problems, numerous goodness-of-fit metrics exist, with no consensus on a single preferred metric. For example, logistic regression models often use deviance or log-likelihood, while neural networks or decision tree frameworks may use confusion matrices or the F1 score. This section briefly describes commonly used goodness-of-fit statistics in machine learning classification models.

1.3.4.1. Deviance or Maximum Likelihood

Deviance is a goodness-of-fit statistic used in statistical hypothesis testing, generalizing the sum of squares of residuals in ordinary least squares to cases where model-fitting is achieved by maximum likelihood (161). It measures the difference between observed and predicted outcomes in a discrete choice regression model, with lower deviance indicating a better fit. Deviance is formulated as:

$$D = -2[\log(\mathcal{L}_R) + \log(\mathcal{L}_U)] \quad (7)$$

Where D is the deviance, LR is the likelihood under the restricted model, and LU is the likelihood under the unrestricted model. Deviance can compare logistic regression models and test hypotheses about predictor-response relationships (161).

1.3.4.2. Confusion Matrix

A confusion matrix evaluates the performance of a classification algorithm (162,163). It is a table with true class values as rows and predicted class values as columns. Diagonal entries represent correctly classified samples (true positives and true negatives), while off-diagonal entries represent misclassified samples (false positives and false negatives), as shown in **Figure 46**.

By examining the confusion matrix, various performance metrics such as accuracy, precision, recall, and F1 score can be calculated, providing different perspectives on classifier performance. The confusion matrix is a valuable tool for analyzing and comparing classification algorithms, helping identify areas for improvement and determining the best algorithm for a particular task (163).

		Predicted	
		Negative (N) -	Positive (P) +
Actual	Negative -	True Negative (TN)	False Positive (FP) Type I Error
	Positive +	False Negative (FN) Type II Error	True Positive (TP)

Figure 46: Example of a confusion matrix.

Understanding the definitions of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) is crucial as these terms form part of the confusion matrix and are foundational for many other goodness-of-fit metrics. TPs refer to instances where the classifier correctly predicts the positive class, such as correctly identifying a patient with a disease. TNs are instances where the classifier correctly predicts the negative class, like correctly identifying a patient without a disease. FPs are instances where the classifier incorrectly predicts the positive class, mistaking a healthy patient for having the disease. FNs occur when the classifier incorrectly predicts the negative class, failing to identify a disease in a patient who actually has it (162).

1.3.4.3. Accuracy

Accuracy (A) measures the overall correctness of a classification algorithm in predicting the class of instances. It is defined as the number of correct predictions (both true positives and true negatives) divided by the total number of predictions:

$$A = \frac{TP + TN}{TP + FP + FN + TN} \quad (8)$$

Accuracy provides a general idea of how well a classifier performs but can be misleading when the class distribution is imbalanced. In such cases, a classifier always predicting the majority class may show high accuracy without making meaningful predictions (130,162).

1.3.4.4. Precision (Specificity)

Precision (P), also known as specificity, is the probability of correctly predicting a given class. It measures the proportion of correctly classified groups over the total number of observations for each group and the classifier's ability to avoid false positives:

$$P = \frac{TP}{TP + FP} \quad (9)$$

Precision ranges from 0% to 100%, with 0% indicating all predictions are false and 100% indicating perfect precision with no false positives.

1.3.4.5. Recall (Sensitivity)

Recall (R), also known as sensitivity or true positive rate, measures the probability of correct predictions given a class prediction and is computed as per **equation (10)**. It measures the proportion of correctly classified groups over the total number of predictions for each group and the classifier's ability to avoid false negatives:

$$R = \frac{TP}{TP + FN} \quad (10)$$

Recall also ranges from 0% to 100%, with 0% indicating all predictions are false and 100% indicating perfect recall with no false negatives. Recall and precision are similar but incomplete on their own; a model can have perfect precision but poor recall in imbalanced datasets. These metrics should be analyzed together or through a combined statistic (130,162).

1.3.4.6. F1 Score

The F1 Score, also known as the F-measure or F1 value, combines the precision and recall of a classifier into a single metric by taking their harmonic mean, as shown in **equation (11)**.

$$F1\text{-score} = \frac{2PR}{P + R} \quad (11)$$

The F1 score provides a single value representing the overall performance of a classifier, considering both its ability to avoid false positives (precision) and to identify all relevant instances (recall). This makes the F1 score particularly useful for evaluating classifier performance when the class distribution is imbalanced, as it balances two complementary metrics (162).

The F1 score ranges from 0% to 100%, with 100% indicating perfect balance and 0% indicating the worst possible balance. It can also be adjusted to prioritize recall over precision, depending on the goal. This adjustment is shown in **equation (12)**.

$$F1_{\beta} = \frac{(1 + \beta)PR}{\beta P + R} \quad (12)$$

Where β is the weighting factor. Common β values are 0.5 and 2.0. Choosing $\beta=0.5$ gives higher weight to recall relative to precision (making false positives more concerning), while $\beta=2.0$ gives higher weight to precision (making false negatives more concerning). Setting $\beta=1$ reverts to the original F1 score, which equally weights precision and recall.

1.3.4.7. Matthew's Correlation Coefficient (MCC)

MCC is a performance metric for binary classification problems, considering all elements of the confusion matrix (162). It is a balanced metric that summarizes overall classifier performance by means of **equation (13)**.

$$MCC = \frac{([TP][TN]) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (13)$$

MCC ranges from -1.0 to +1.0, with +1.0 indicating perfect prediction, 0 indicating random prediction, and -1.0 indicating perfectly incorrect prediction. MCC can be more robust for imbalanced datasets compared to other metrics. It accounts for all parts of the confusion matrix and can be a stronger indicator of performance than the F1 score. However, this robustness comes at the cost of interpretability. The F1 score is easier to understand and calculate, focusing solely on precision and recall. In most cases, the F1 score is sufficient to distinguish high-performing classifiers (164).

MCC is particularly useful for highly unbalanced datasets where the positive case is less important. In such scenarios, MCC provides a balanced evaluation, while the

F1 score can yield high values only when the classifier accurately identifies the positive (minority) class (162).

1.3.4.8. Fowlkes–Mallows index

The Fowlkes–Mallows Index (FMI) is a performance metric used to determine the similarity between two clusters obtained after running a clustering algorithm. It can also be used as a goodness-of-fit metric for a binary classification algorithm. FMI measures similarity between either two hierarchical clustering models or a clustering and a benchmark classification (162). It is calculated as the geometric mean of precision and recall as per **equation (14)**:

$$FMI = \sqrt{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}} \quad (14)$$

FMI ranges from 0% to 100%, with 100% indicating perfect similarity between the two sets of predictions and 0% indicating complete dissimilarity. This metric is most often used when comparing unsupervised machine learning implementations and is rarely seen as a goodness-of-fit metric for supervised learning classification models.

1.3.4.9. Receiver Operating Curve (ROC) and Area under the ROC (AUC)

The Receiver Operating Curve (ROC) visually assesses a model's performance using a graph with the false positive rate (FPR or 1-specificity) on the x-axis and the true positive rate (TPR or recall) on the y-axis at various classification thresholds or hyperparameter values, as shown in **Figure 47**. The ROC curve illustrates the trade-off between TPR and FPR, helping to select the optimal threshold or hyperparameter value that balances the risk of false positive and false negative predictions (162,163). A classifier with perfect performance has an ROC curve that passes through the top left corner of the plot (TPR = 1, FPR = 0), while a classifier with random performance has an ROC curve close to the diagonal (TPR = FPR). The closer a point is to the top left of the ROC curve (0, 1.0), the better the classifier's performance.

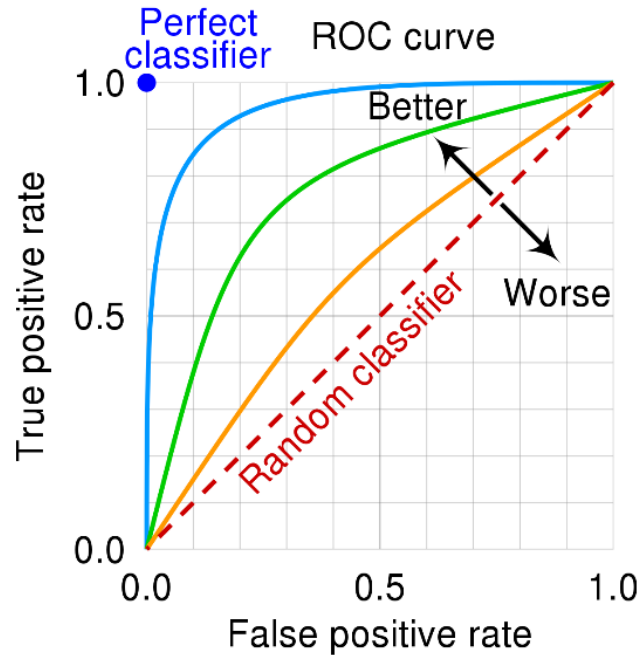


Figure 47: The ROC space for a “better” and “worse” classifier (165).

While the ROC curve is useful for tuning hyperparameters and finding the best performing model within a framework, it is not used to compare different models. Instead, the Area under the ROC Curve (AUC) is used. The AUC is a numerical measure summarizing the overall performance of a classifier by evaluating its ability to distinguish between positive and negative instances across all possible classification thresholds or hyperparameter values (146,166).

The AUC ranges from 0% to 100%. An AUC of 0.0 indicates a classifier that perfectly predicts negative input data as positive and vice versa. An AUC of 0.5 indicates a classifier equivalent to random guesses, while an AUC of 1.0 indicates a classifier that perfectly predicts all data without error. The area under the ROC curve is typically estimated using the trapezoidal rule (162).

This section highlights commonly cited performance metrics in the literature. For a more comprehensive list, refer to studies by Fawcett (166), Powers (162), Pirayonesi and El-Diraby (147), and Chicco et al. (164).

1.3.4.10. Aggregation Methods for Goodness-of-fit Metrics

When discussing goodness-of-fit statistics, particularly those using elements within the confusion matrix, it’s important to note that these metrics usually quantify the model’s performance in classifying a single category of the target variable. Thus, for each category being classified, there will be separate values for accuracy, precision, recall, F1 score, or MCC. To quantify the overall performance of the

model across all categories, these metrics must be aggregated in one of three ways (167).

The first method is to compute the model performance as a macro or unweighted performance metric. This means the goodness-of-fit metric for the model is the arithmetic mean of each class's performance metric, assigning equal weight to the performance of each class, regardless of the number of observations in each class. Macro statistics are only relevant when the class distribution is well-balanced (167).

The second method is to compute performance as a weighted average of the performance observed for each class. This method accounts for the proportion of the sample allocated to each class when summarizing the model's performance, making it more robust to imbalanced datasets (167).

The third method is to compute a micro average of the performance metrics. The micro average is the global average for the goodness-of-fit metric of interest. Instead of computing an F1 score for each category and averaging those, the micro average counts the total number of TP, FP, and FN observed across all categories and computes a single F1 score based on that (167).

Micro averages tend to perform worse on imbalanced datasets compared to macro and weighted average F1 scores. This is because micro averages give equal importance to each observation, meaning classes with more observations will have a larger impact on the final score, potentially hiding the performance of minority classes and amplifying the majority. In contrast, macro F1 scores give equal importance to each class, ensuring that a majority class contributes equally along with the minority, thus returning more objective results on imbalanced datasets (167).

Chapter 2. Prototype Development

2.1. Past Methods Used by the Performing Agency for Measuring and Determining Pavement Surface Characteristics

2.1.1. Equipment from Previous Studies

The Performing Agency conducted friction and texture measurement tests in northeast Austin, Texas. The test circuit included Farm-to-Market roads, State Highways, and US highways with various flexible pavements, such as dense-graded, open-graded, gap-graded mixes, and chip seals. Pavement sections of 800

meters were measured for both friction and texture, with data collected along the lane centerline and the left wheel path.

The equipment used included a prototype developed in-house, featuring continuous friction measurement equipment, the GripTester MK2, and a line laser scanner. Although the GripTester information is not relevant to this study, it was retrofitted with a line-laser scanner to capture the texture profile before wetting the surface. The line-laser was positioned in front of the water nozzle and test tire, centered on the test tire's path, and oriented transversely to the direction of travel, covering the 100 mm contact patch of the tire. The line-laser scanner measured the entire macrotexture range (50 to 0.5 mm) and a portion of the first decade of microtexture (>0.322 mm). A profile was captured every 40 mm using the GripTester's internal distance measuring unit, allowing synchronized data collection from both pieces of equipment. Figure 48 shows an image of the data collection prototype. **Figure 48** shows an image of the data collection prototype.



***Figure 48:** Original data collection prototype to measure pavement surface profiles at high speed and predict pavement surface.*

During this study, cluster analysis revealed that texture statistics could identify six distinct pavement surfaces. Additionally, a decision tree classifier using four texture statistics achieved an F1 Score of approximately 90%. This strong predictive power was hypothesized to extend beyond pavement friction to models predicting rolling resistance, splash and spray, tire/pavement noise, and other pavement surface interactions.

2.1.2. Original Prototype for Collecting 2D Surface Profiles

The previous pavement texture measurement equipment had limitations, such as low laser power, which resulted in noise and missing data due to sunlight and dark-colored pavements. Additionally, it relied on the GripTester for mounting and triggering the sensor. To address these issues, the Performing Agency developed a standalone system that eliminates the need for the GripTester and incorporates a high-power line-laser scanner capable of measuring pavement texture at higher rates. The system's flexible mounting mechanism allows for versatile configurations of the line-laser.

To further enhance measurement accuracy, the Performing Agency plans to integrate a wheel encoder on the test vehicle's rear wheel. This encoder will trigger the sensor based on the distance traveled, ensuring evenly spaced measurements instead of varying profiles due to changes in vehicle speed.

Preliminary testing positioned the laser's line at a 45-degree angle relative to the direction of travel, enhancing the algorithm's ability to classify pavement surfaces by capturing both transverse and longitudinal features. The sensor's position can be adjusted to measure the right or left wheel path or the center of the lane. The initial test involved installing the system on the rear hitch receiver of a full-size truck, with the line-laser positioned on the left wheel path, as shown in **Figure 49**. The data showed reduced noise and the ability to perform testing under full sunlight and at highway speeds. Future tests will include the addition of a wheel encoder to this setup.



***Figure 49:** New data collection prototype to measure pavement surface profiles at high speed and predict pavement surface.*

2.1.3. Original Prototype for Collecting Surface Images

In parallel to the aforementioned setup, the Performing Agency tested a proof of concept by using images of pavement surfaces to classify the type using supervised machine learning. A prototype consisting of a machine vision camera installed on a vehicle, positioned 610 mm above the surface, covering an area of 127 x 100 mm, was tested. Initial tests revealed challenges due to changing lighting conditions, such as the position of the sun and overcast skies, which caused irregular lighting. To address this, two high-power spotlights with wavelengths simulating midday sunlight were added. While these spotlights improved image quality in shaded conditions, they were not as powerful as natural sunlight, leading to variations in image contrast, as shown in (Figure 50).



Figure 50: (A) Prototype magnetically mounted to the passenger side of the vehicle. (B) Close-up of the camera in the middle and two spot-lights focusing the light at the center area of where the images are being captured.

The prototype was tested on pavement near the city of Austin, previously evaluated with friction and texture equipment, in the same sections with homogeneous surfaces. The vehicle, equipped with the prototype, traveled at 60 mph while capturing images at 170 frames per second, collecting approximately 10,000 images per section under varying lighting conditions. Six flexible pavements were analyzed, reduced from prior testing due to repaved roads.

To expand the study, various concrete surface finishes were sampled, including carpet drag, diamond grinding, longitudinal tining, transverse tining, and combinations of these finishes. Samples were collected from cities across Texas, including Austin, Abbott, Corsicana, Dallas, and Houston. Quality control eliminated images that were unclassifiable due to heavy distress, external features, or contaminants. Approximately 60,000 images underwent manual classification, resulting in a small database of four distinct concrete types..

2.2. Technologies Considered during Prototype Development

This section highlights the most promising sensor devices and hardware considered by the Performing Agency in designing the data collection prototype. The final prototype consists of three major components: 1) a laser scanner, 2) a high-speed camera, and 3) a lighting system to illuminate the pavement at night or when shadows affect the camera's field of view.

2.2.1. Laser Scanners

The purpose of the laser scanner is to measure the surface elevations of the pavement with high resolution, allowing for the quantification of texture indexes to characterize its morphological features. The laser sensors discussed in this section were candidates considered for the final prototype.

2.2.1.1. AccuProfile 820 Laser Scanners

These sensors are part of the Acuity Laser product line by Schmitt Measurement Systems, Inc. Commonly used in automation and quality control, these sensors measure and verify dimensions, detect flaws, and identify surface variations (168). The laser measurements create two-dimensional and three-dimensional representations of the scanned item and perform well on shiny or jagged targets, in environments with high ambient light, and are robust against vibration and shocks (169). Typical applications for a 2D laser line scanner include steel and aluminum production, automotive, aerospace technology, electronics, robotics, and welding (169–172). The AccuProfile 820 Laser Scanners offer various case sizes, measuring ranges, and diode/accessory options for a wide range of measurement applications (168), as shown in **Figure 51**. Although popular in other engineering and scientific fields, there are no peer-reviewed papers in pavement engineering literature that have used these sensors.

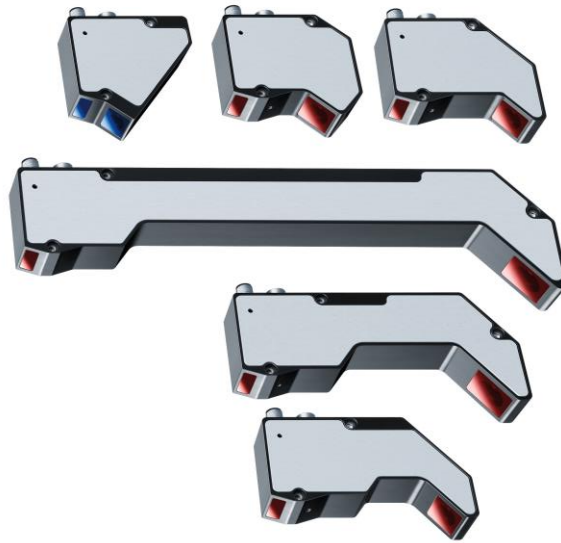


Figure 51: AP820 2D laser scanner series (168)

2.2.1.2. 8300 High Speed Inertial Profiler - Ames AccuTexture 100

The Model 8300 High Speed Inertial Profiler (left side of **Figure 52**) is an inertial profiling system designed by AMES Engineering to enhance efficiency for paving contractors, testing agencies, and State and Federal highway departments (29). The Model 8300 offers multiple add-ons and configurations, including the AMES AccuTexture 100 single point laser sensor (right side of **Figure 52**). This sensor uses a single point laser rated at 100 kHz to measure the longitudinal profile of roadways at highway speeds (at least 45 mph) and is touted as “the new standard in accurate texture measurements” (173). The system also features a Profiler GPS-DMI, which replaces the need for a wheel-mounted encoder, improving data collection efficiency and operator safety (29).

Tests conducted by AMES Engineering with the AccuTexture on flexible and rigid pavements demonstrated that this sensor significantly reduces variances in elevation measurements due to speed or surface type, proving superior to the previously used Optocator sensor (173). The sensor has been utilized in studies on the quality control of seal coats (174) and the integration of network-level macrotexture measurements into pavement management systems (175).



Figure 52: 8300 High Speed Inertial Profiler (29) on the left, and Ames AccuTexture 100 single point laser (173) on the right

2.2.1.3. Pathway Services Laser Sensor System

Pathway Services specializes in automated pavement condition data collection in North America. Their data collection vehicle, the Pathrunner (**Figure 53**), is equipped with advanced equipment for various pavement condition assessments, including image collection, distress data collection, and skid and texture measurements (176). The texture measuring equipment includes three-line laser sensors positioned along the two wheel paths and the center line of the lane, all oriented at a 10-degree angle from the transverse direction (177). These sensors feed data to the “3-zone texture signature subsystem” for microtexture characterization at high speeds and to their macrotexture system for pavement macrotexture analysis (177).

Replicating their system is currently not possible due to the novelty of the technology, and most technical information about their system is not publicly available. So far, the system has been referenced in one technical report (178) and one conference presentation (179), where their texture characterization approach is explained.

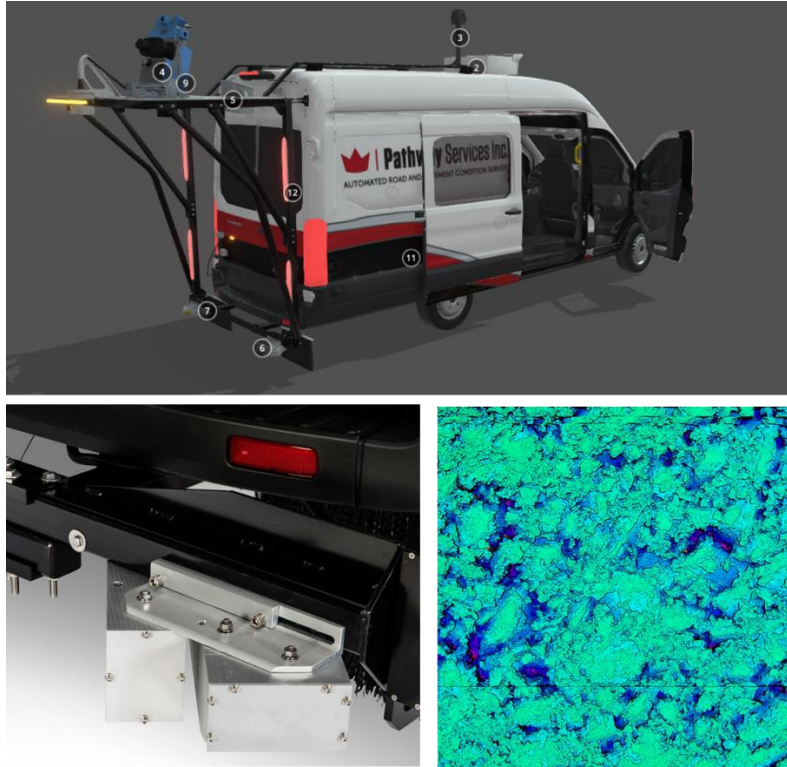


Figure 53: Pathrunner (176) on the top, Macrotexture sensors (177) on the bottom right, and scan obtained from 3-zone texture signature subsystem on the bottom left (177).

2.2.1.4. Keyence LJ-X8000 series

The LJ-X8000 series of laser sensors (**Figure 54**) manufactured by Keyence is a widely used 2D/3D line laser scanner available in the market. These sensors are employed in diverse applications such as pavement engineering (180,181), materials engineering (182), precision manufacturing (183), and quality control (184). They measure up to 3200 points per profile, ensuring accurate measurements of any target shape or material (185). The scanners are compatible with four different controller options, allowing users to align system capabilities with application requirements. With a maximum laser linewidth of over 720mm (28”), multiple sensors can be paired together to scan larger targets. Programming these high-performance sensors involves three simple steps, making accurate in-line 3D measurement accessible for users of any experience level(185). Additionally, the sensors use a distance-correction algorithm to maintain consistent point spacing within operating limits.

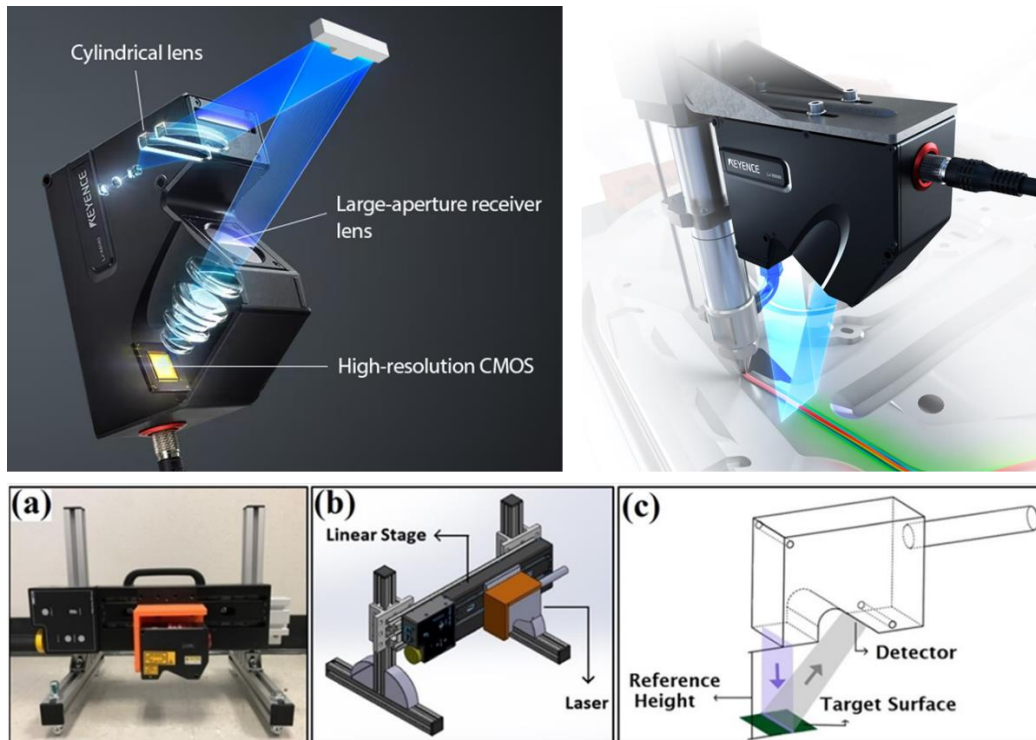


Figure 54: Images from the LJ-X8000 series sensors from the provider (185) on the top, and images of the sensors mounted on stage for use in laboratory measurements (180) on the bottom.

2.2.1.5. HyMIT Sensors

HyMIT, based in Austin, Texas, specializes in manufacturing laser sensors for pavement engineering applications. These sensors (**Figure 55**) are primarily used for high-speed data collection of texture, roughness, or distress quantification at speeds up to 70 mph (186). Known for their high-power lasers, these sensors allow for short exposure times, minimize moving average error, and provide good quality data at speeds below 70 mph (186). Typically mounted on adjustable braces on the back of a vehicle, these sensors can be positioned to measure different parts of the lane (wheel path versus center), set at various heights, and installed in different orientations (parallel, perpendicular, or angled with traffic direction). The Texas Department of Transportation (TxDOT) Maintenance Division currently uses these sensors for quality control measurements across their network.

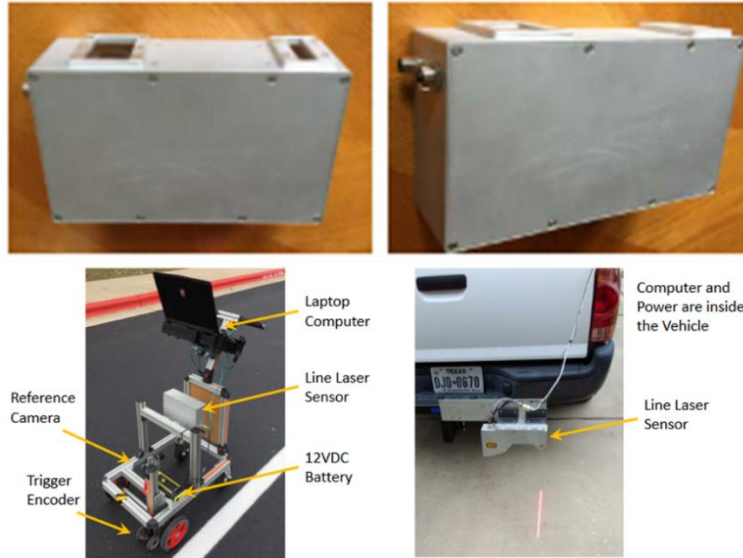


Figure 55: High power laser sensor by itself from a bottom side view (186), shown on the top and the laser sensor mounted on a pushcart and a vehicle to for field use (186) on the bottom.

2.2.2. Cameras

The camera's purpose is to capture high-resolution images of the pavement surface, complementing the texture data collected by the laser sensor. These images are used in image detection algorithms to determine the pavement type.

2.2.2.1. 3D Machine Vision Ranger3

The 3D Machine Vision Ranger3 camera (**Figure 56**) is a high-speed, high-resolution camera manufactured by SICK Sensor Intelligence. This advanced camera features a distinctive CMOS sensor that enables rapid 3D measurement, providing precise assessments of shape, volume, and positioning across various objects. It ensures reliable precision on both dark and bright surfaces without requiring increased laser power (187). The Ranger3 integrates with standard software and is easy to incorporate mechanically, making it versatile for modern industrial needs (187). In research applications, the camera has been used in pavement engineering studies to estimate pavement texture depth (188), in industrial engineering for automating quality control pipelines (189), and in structural engineering studies on crack detection in concrete bridges (190).



Figure 56: Front and back view of the 3D Machine Vision Ranger3 camera (187)

2.2.2.2. Canon 5D Mark IV

The Canon 5D Mark IV (**Figure 57**) is an advanced camera system featuring a 30.4 Megapixel full-frame CMOS sensor, designed for versatile imaging in various lighting conditions (191). It offers 4K Motion JPEG video capabilities at 30p and 24p frame rates and can achieve a resolution of 8.8 Megapixels (191). Its optical architecture includes a 150,000-pixel RGB+IR metering sensor, which ensures precise exposure quantification, sensitivity to changing light conditions, and enhanced scene recognition and facial detection (191). These features collectively enhance its imaging capabilities. The camera has been used in studies on computer vision and pattern recognition (192), dentistry for exploring oral rehabilitations using a digital protocol (193), and has been featured in Australian Photography publications (194). However, there is no scientific literature indicating its use in pavement engineering applications.



Figure 57: Front and back view of the Canon 5D Mark IV camera (191)

2.2.2.3. Arducam 2MP Global shutter OV2311

The Arducam 2MP Global shutter OV2311 camera's module (**Figure 58**) is primarily designed for Raspberry Pi boards and connects directly to the RPi's CSI-2 camera interface without additional hardware (195). This camera, based on a 1/2.9-inch Omnivision OV2311 image sensor, provides full-frame, sub-sampled, and windowed 8/10-bit MIPI images, operating at up to 60fps in full resolution on

a 2-lane MIPI bus, with complete user control over image quality (195). The OV2311 offers highly accurate gaze- and eye-tracking capabilities and high near-infrared quantum efficiency to minimize active illumination power and reduce system power requirements (195). However, this camera model was discontinued between December 2022 and August 2023. Its successor, the Arducam PiVariety 2MP Global Shutter OV2311, offers the same quality with additional features. As of August 2023, the Arducam OV2311 has been used in one peer-reviewed geotechnical engineering study on 3D deformation measurements during triaxial testing (196). The Arducam series has also been used in smart surveillance (197), crop disease detection (198), and biology and physiology (199).

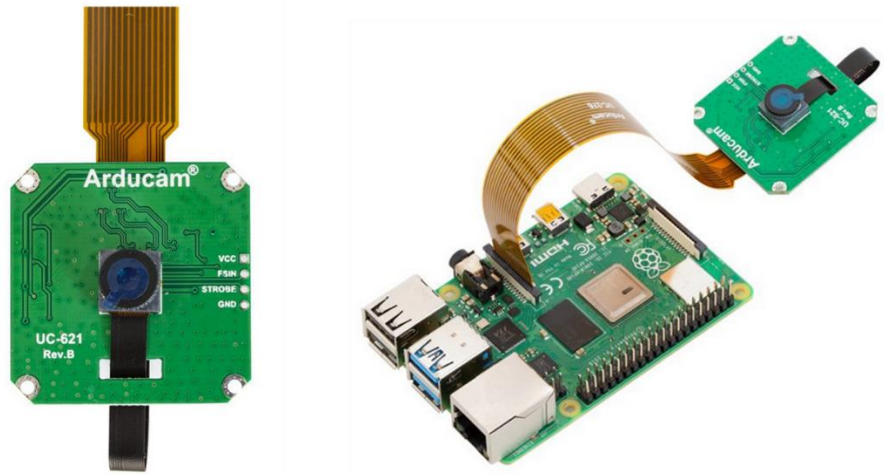


Figure 58: Front and side view of the Arducam 2MP Global shutter OV2311 camera (195)

2.2.2.4. Blackfly S USB 1.3 MP

The Blackfly S camera (**Figure 59**) features advanced sensors in a compact ice-cube form factor, enabling the capture of high-resolution images even at high speeds (200). It offers both automatic and precise manual control over image capture and on-camera pre-processing. Additionally, it is compatible with third-party software and hardware, supports the development of custom applications with rich sample code and descriptive API logging, and works with a wide range of operating systems and host system hardware architectures (200). While the Blackfly S has not been cited as a sensor used in pavement engineering, it has been utilized in various scientific fields such as geographic information science (201), computer and electrical engineering (202), and mechanical engineering (203).



Figure 59: Front and back view of the FLIR Blackfly S camera (200)

2.2.3. Lighting Systems

The purpose of the lighting system is to ensure that the camera always captures images with sufficient light, regardless of environmental and lighting conditions at high speeds. This is particularly important when driving at night, under cloudy skies, or when the vehicle's shadow falls over the area being photographed. It is also crucial when driving over terrain that casts shadows, such as tree lines, buildings, or bridges.

2.2.3.1. Dual Magnetic Spotlights

In the first iteration of the prototype, researchers attached two generic spotlights to a test vehicle. The spotlights were mounted on magnetic stands, allowing for easy orientation towards the area where the high-speed camera was focused, as shown in **Figure 60**. This versatile setup allows the placement of both the camera and lights anywhere on the vehicle with a solid, flat metal surface. The lights can draw 12 V power either from the vehicle itself or from a separate battery inside the test vehicle. These lights remain on continuously during testing, regardless of the time of day or terrain conditions.



Figure 60: Part (A), camera and lighting system magnetically mounted to the passenger side of the vehicle. Part (B), close-up of the camera in the middle and two spot-lights focusing the light at the center area of where the images are being captured.

2.2.3.2. Metaphase – Line lights

Metaphase, with over two decades of expertise in advanced LED lighting, has a history of working with clients in various fields, including road and railroad track inspection. Ultimately, Metaphase’s TX Barlight 6000K (**Figure 61**) was chosen, offering over 100,000 lux in continuous mode and up to eight times more power when configured as a strobe light. This setup covers an area of 5 x 5 inches (127 x 127 mm) from a height of 2 feet (0.6 m).

The engineers at Metaphase collaborated with the Performing Agency to synchronize the camera and lightbar by carefully wiring the camera’s strobe-out signal to the light trigger input, ensuring proper synchronization between the light flashing and camera exposure. Although the current configuration is non-shrouded, plans are underway to install a second lightbar and add shrouding for the camera and lights to ensure that all illumination in the inspection area is solely from the designated light source.



Figure 61: TX Barlight with a 3-sided T-slot extrusion.

2.3. Current System’s Components and Assembly

In the initial phases of this study, passenger vehicles, such as sedans and small-size pickup trucks, served as the primary means to transport the prototype system. However, these vehicles exhibited significant vertical displacement on highways due to their suspension systems bouncing more than expected, causing the sensors to frequently go outside their working range. This led to blurry images, abnormal readings, and unwanted vibrations in the laser signal. To address this issue, A heavy-duty pickup truck with robust build, strong suspension, and reliable towing capacity became the chosen host vehicle for field testing to address this challenge. This selection was based on the stiff suspension of these trucks, which proved

resilient against large vertical bounces. Nevertheless, it's worth noting that these trucks had the drawback of being more than 2 ft above the ground, measured from the rear hitch receiver. In contrast, smaller trucks could be as low as 1.4 ft, and sedans as low as 1 ft.

To accommodate different vehicle heights, a fully customizable mounting system was developed. This system, made from 1.57 x 1.57 in (40 x 40 mm) aluminum profiles, allows for quick assembly, removal, and modification. It enables vertical and lateral positioning of the system's components with respect to the vehicle. The system is fixed to the test vehicle via the rear hitch, using 2 x 2 in (50.8 x 50.8 mm) steel square tubing inserted into the hitch and secured by a heavy-duty bolt. The other end of the steel piece has a vertical 2 x 2 in (50.8 x 50.8 mm) square tube with a T-shape, where the horizontal part of the T fastens the main body of the aluminum profile.



Figure 62: Side and rear view of the mounting system.

The horizontal aluminum profile serves as a rail on which the sensors can be mounted, allowing for independent adjustments of each component, such as the line laser and the camera/light system **Figure 63**. These profiles can be positioned anywhere within the vehicle's width (e.g., left wheel path, center wheel path, right wheel path) and permit height adjustments of up to 3 feet, addressing the issue of varying vehicle heights. When mounting the sensors, it is crucial to measure their heights accurately to set them at the correct stand-off distances. Specifically, the laser sensor requires a height of 18 in (457 mm) from the camera window to the surface, while the camera/light needs to be 24 in (610 mm) from the camera lens to the surface. This configuration accommodates vertical movement in both sedans and pickup trucks. Measurements should be taken on a level surface, such as concrete, and account for the weight of passengers and/or extra load in the vehicle.



***Figure 63:** System mounted onto vehicle shown on the top; line laser sensor shown on the bottom-right; and camera and light system (perpendicular to the main beam), shown on the bottom left.*

For electrical wiring, each sensor has specific power and data communication requirements. The line laser requires a 12 V power supply from the vehicle's cigarette lighter, consuming 0.7 amps. Data transmission is handled by a 20-foot (6 m) Cat 6E Ethernet cable connecting the sensor to a laptop in the front passenger seat. The area camera operates on a 5 V supply and connects to the laptop via a 20-foot (6 m) USB 3.1 cable, which provides both power and data transmission. The lighting system uses a 24 V power source from a battery pack. Wiring for the camera and light source, including synchronization connections, is shown in **Figure**

64, with their respective pins and color-coded wires from each pigtail. For safety, a waterproof enclosure houses the wires, reinforced with special wire connectors.

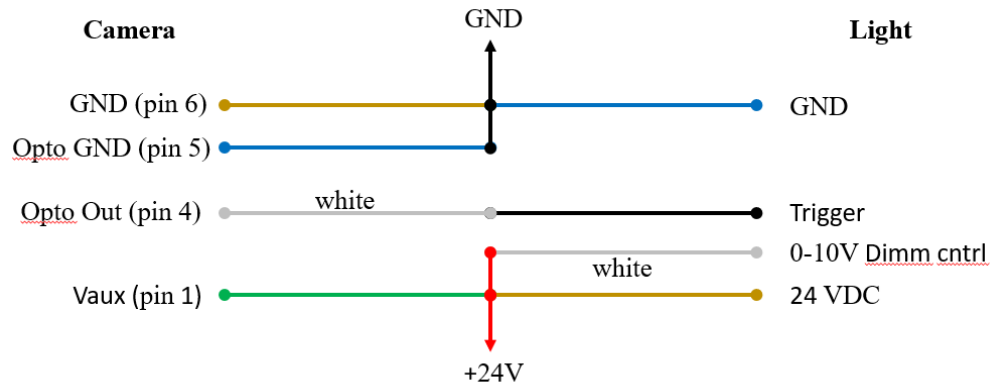


Figure 64: Diagram to wire the camera and light source.

Synchronization between the area camera/light and the line laser sensor is achieved through a wheel encoder connected to the vehicle's wheel. This encoder is programmed to activate the laser and the camera/light sequentially every 40 mm of forward travel.

2.4. General Specification for a System with Comparable Performance

This section outlines the specifications for the three main components used in this project: the line-laser sensor, the area camera, and the light source for the area camera. The requirements for the line laser sensor, as proposed by the Performing Agency, are presented in **Table 3**. and should be considered the minimum standard for future data collection of the same quality. These parameters were established through extensive testing in real-world scenarios under various conditions, including different vehicle heights, ambient conditions, and road surface types.

Table 3. General Specifications for the sensor used in the data collection prototype

Line Laser	
Working Distance	400-450 mm
Z-Range (+/-)	75 mm
Resolution Height	< 5 micrometers
Field of View	142 - 190 mm
Wavelength	810 nm
Optical Power	2.3 W
Data Points	> 2000
Sampling Cycle	≥ 600 Hz
Resolution Lateral	< 100 micrometers
Trigger	TTL
AOI	608 pixels
Area Camera	
Resolution	$\geq 1280 \times 1024$
Frame Rate	≥ 50
Megapixels	≥ 1.3
Chroma	Color
Operating Temperature	0° to 50°C
Readout Method	Global shutter
ADC	10-bit
Exposure Range	$\leq 10.0 \mu\text{s}$
Interface	USB 3.1 or Ethernet
Light Source	
Color	White, 6000K
Operating Temperature	0° to 40°C
IP Rating	IP64
Intensity	$\geq 100,000$ Lux

The first eleven parameters under the “Line Laser” bracket are crucial for developing a system capable of operating effectively at speeds up to 80 mph, in both extremely bright and dark conditions, and on various pavement surfaces. The working distance was set at 400 mm with a tolerance of ± 75 mm to accommodate vehicle movements during travel. A stiff suspension is ideal to stay within this range. With the working distance and Z-Range set, the height resolution was 3.8 micrometers (anything less than 5 micrometers is acceptable). The field of view

should ideally be between 142 to 190 mm, providing sufficient coverage to capture a 100 mm length (transverse) at a 45-degree angle relative to the direction of travel.

The laser's wavelength should be at least 810 nm to minimize interference from the solar spectrum, which peaks around 500 nm. A minimum optical power of 2.3 W is necessary to ensure visible profiles on very dark pavements, allowing the laser sensor's camera to operate with exposure times as low as 3-5 microseconds. While the number of data points per profile is optional, it is recommended to have at least 2000 data points for adequate lateral resolution.

An external trigger source is advisable, although the hardware can function without one. The trigger collects profiles at predefined intervals of distance traveled, independent of vehicle speed. In this project, profiles were collected approximately every 44 mm, with the TTL trigger sending a pulse to the line laser. The trigger distance can be adjusted based on the camera's sampling cycle. The laser sensor can operate at speeds up to 80 mph, collecting data every 44 mm at a sampling rate of 812 Hz. For this project, data was collected at 50 mph, equating to a sampling rate of 508 Hz. The Area of Interest (AOI) is a defined window within the camera image where the software focuses on the laser line, enhancing sampling cycles and Z-Range. The AOI should have at least 608 pixels.

For the "Area Camera," there are nine main parameters to consider. The camera must have a high frame rate (≥ 50 FPS), low exposure time (≤ 10 microseconds), and a global shutter. At 50 mph, the image blur should be no more than 6 mils (0.15 mm). The camera should have a lens capable of capturing an area of approximately 15 x 15 cm from a height of 50 cm, such as a 2/3" lens with a 12 mm focal length and various iris stops (F1.4/F4/F8/F16).

Finally, the light source should provide sufficient illumination as the main lighting for the area camera. This can be a continuous light source or a synchronized light burst with the area camera.

2.5. Target Degree of Prediction Accuracy and Method of Evaluation

In a surface prediction study conducted for TxDOT project 0-7031, results indicated that it was possible to distinguish six distinct pavement surfaces on flexible pavements (54). These were chip seals with a high macrotexture, chip seals with a low to medium macrotexture, open-graded mixes, gap-graded mixes or SMAs, dense course-graded mixes, and dense fine-graded mixes

The dataset in this study included 21 pavement sections, each at least 0.5 miles long, located on the east side of Travis County near Austin, as shown in **Figure 65**.

Texture measurements were collected along the centerline of the outer lane using a line laser sensor with a resolution of 0.161 mm in the transverse direction and 5 microns in the vertical direction.

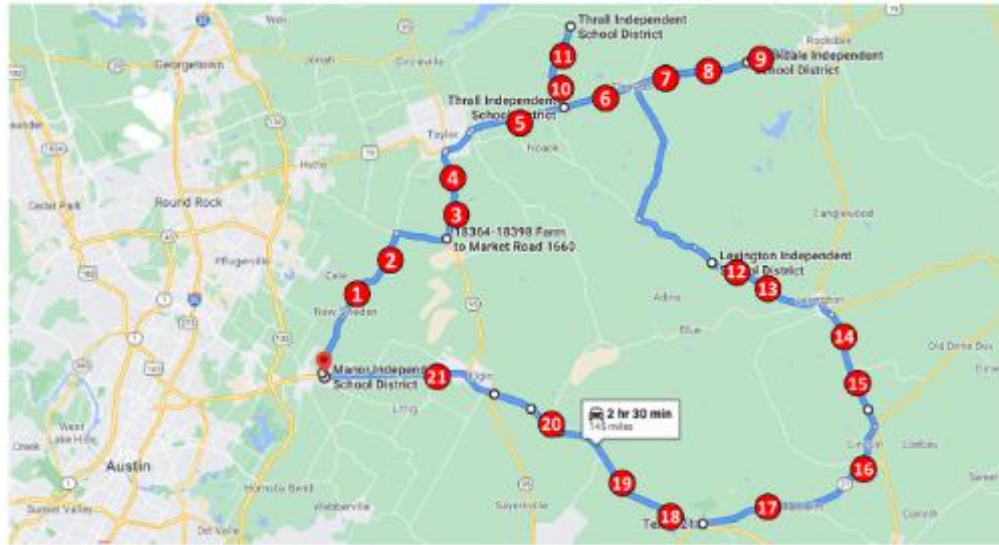


Figure 65: Map showing the location of every pavement section surveyed (54)

The researchers reported a prediction accuracy, measured in terms of the F1 score, that ranged between 80% to 98% using a decision tree classification model, as shown in Table 4. Where the easiest pavements to classify were the open and gap-graded mixes and the relatively harder ones were the chip seals and dense graded mixes (54).

This preliminary study demonstrated that pavement surface identification in flexible pavements is possible and feasible using the technology available to the Performing Agency. However, the study did not address how pavement age or the presence of distresses such as rutting, cracking, or flushing (bleeding) would affect prediction accuracy. These factors could potentially reduce the accuracy for each specific surface. Alternatively, pavements that are severely cracked, rutted, or flushed might be classified as distinct surfaces.

Table 4: Classification Report Obtained from Decision Tree Model (54)

	Precision	Recall	F1 Score	Sample Size
CS1	0.82	0.89	0.85	523
CS2	0.85	0.91	0.88	471
OM	0.98	0.98	0.98	484
BOM	0.96	0.98	0.97	484
DCM	0.88	0.73	0.80	582
DFM	0.78	0.81	0.80	236
Overall Model Accuracy				0.89

Note: CS = chip seal, OM = open-graded mixes, BOM = open mix with the asphalt binder exudated to the near surface, DCM = dense coarse mixes, and DFM = dense fine mixes and microsurfacing

2.5.1. Target Classification Groups based on new Measured Data

Based on previous results, the Performing Agency aims to develop prediction models for pavement surface classification. Initially, a model will be created to distinguish rigid pavements from flexible pavements with the highest possible accuracy. Flexible pavements will then be categorized into the following broad groups based on their surface profile and images:

- Surface Treatments: likely to include all pavements that resemble a seal coat,
- Dense Mixes: likely to include all pavements resembling dense graded mixes (note that gap-graded mixtures may appear similar), and
- Open Mixes: likely to include all pavements resembling open graded mixes.

After classifying these surfaces, more specific classifications can be made. For instance, surface treatments can be broken down by the gradation of the chip seal, while dense mixes can be further divided into coarse mixes, fine mixes, and gap-graded mixes. Rigid pavements will be categorized based on their surface profile and images into the following groups:

- Transverse and Longitudinal Tining
- Carpet or Astroturf Drag
- Diamond Grinding
- Exposed Aggregates

Although no published articles or peer-reviewed papers have classified rigid pavements, a preliminary analysis on a few rigid surfaces around Austin found that transverse tinning, carpet drag, and diamond grinding could be distinguished with F1 scores above 70%. However, the sample size was small, with only 10 pavement sections, each 0.5 miles long. The Performing Agency expects that with the aid of image recognition, the classification accuracy for different types of rigid pavement surface finishing/texturing can be improved to at least 80%.

Chapter 3. Experimental Design and Data Collection

The goal of this project was to create a comprehensive database of pavement surfaces across Texas, accounting for variability in pavement type (flexible vs. rigid), texture (positive vs. negative), surface (different mixes and texturing techniques), materials (different aggregate types), and age (pavements at different service life stages). To achieve this, the Performing Agency collaborated with multiple TxDOT district offices in North, East, West, South, and Central Texas, requesting locations of at least 10 highway sections with diverse surfaces, including: *Hot and Warm Mix Asphalt Surfaces*: dense, gap, and open-graded mixes of varying gradations; *Portland Cement Concrete Surfaces*: carpet/burlap drag, longitudinal/transverse tining, diamond grinding, grooving, etc.; and *Surface Treatments*: seal coats of different grades, microsurfacing, fog seals, etc.

The Performing Agency prioritized newer pavements for their certainty regarding pavement type. However, some agencies provided locations of pavements over ten years old without any surface-altering maintenance. While most variability sources are accounted for, pavement age and aggregate source are not always known with certainty. Additionally, variations in asphalt binder content, water-cement ratio, or chemical admixtures are not considered due to data collection difficulties. All data was collected using the prototype system detailed in Chapter 2.

Before data collection, two data subsets were established: one where the pavement surface type is known, based on TxDOT personnel information, and another where the surface type is unknown. The known subset serves as the training, testing, and validation dataset for developing classification models. The unknown subset is a holdout dataset, where the best classification algorithm will predict pavement types based on surface texture and images. These predictions will be compared to actual images to assess the algorithm's accuracy.

For consistency, a “testing site” refers to a highway section longer than 0.4 miles where equipment was run continuously, while a “pavement section” refers to a 0.1-mile-long section within the testing site. Each pavement section is homogeneous, containing only one type of pavement surface.

The Performing Agency collected a substantial amount of data to ensure a robust dataset for developing machine learning algorithms. They gathered 400 miles of pavement texture data and over 100,000 pavement surface images, with at least 300 miles on flexible pavements and 100 miles on rigid pavements. Of the 400 miles, 235 are part of the ground truth dataset for model development, and 165 are part of the holdout dataset for model visualization.

3.1. Data Collection

3.1.1. Locations of Tested Pavements

The Performing Agency contacted personnel in the materials or maintenance divisions of multiple TxDOT districts to obtain reliable ground truth information on the locations of pavements with varying surface types. Initially, the contract specified that the Performing Agency would collect “60 different sections, 20 at each of three districts: Austin, Dallas, and Houston.” However, the Performing Agency exceeded this requirement by collecting more than 60 sections across more than three TxDOT districts. This success was due to the enthusiasm, willingness, and proactiveness of several TxDOT districts in aiding the Performing Agency.

Table 5 summarizes the six TxDOT districts that contributed the most information and resources, along with the key personnel in those districts who assisted in obtaining the pavement sections. The main contact persons in these districts were contacted exclusively by email. Table 6 summarizes additional supporting districts that provided aid either with the data collection process or by suggesting locations of other pavement surfaces of interest to the Performing Agency.

Table 5: TxDOT Districts and personnel who provided the most support and the locations of pavement surfaces to the Performing Agency

District	Main Contact	Supporting Personnel
Atlanta	Lacy Peters	Cody Fuller
Austin	Andre Smit	Andy Naranjo John Wirth
El Paso	Aldo Madrid	Mauricio Esquivel Lara Monica Ruiz Anthony Marquez Christopher J. Weber
Houston	Melody Galland	Viet Pham Juan Fuentes John Zientek III
Laredo	Epigemio Gonzalez	Jesus Saavedra
Lubbock	Ed Goebel	Mike Stroope

Some TxDOT districts in the original experimental design could not be surveyed as planned. Thus, to prevent prolonged delays in data collection, the Performing Agency reached out to all other adjacent TxDOT districts in an effort to obtain reliable ground truth to move forward.

Table 6: TxDOT Districts and personnel who provided the additional support and the locations of more pavement surfaces to the Performing Agency

District	Main Contact	Supporting Personnel
Amarillo	Cody Harris	~
Bryan	Joel Withem	~
Fort Worth	Tom Brown	~
Odessa	Zane Honeyfield	~
San Angelo	Evan Jones	Lonnie Green III
San Antonio	Alejandro Miramontes	Jose Emilio Ramos
Tyler	Dustin Morgan	~
Waco	Jerrod Swift	~
Wichita Falls	Brian Moore	~
Yoakum	Bradley Polasek	~

3.1.2. Locations Tested

The Performing Agency surveyed numerous pavement surfaces as discussed in the previous section. All the testing sites are portrayed in **Figure 66**, which contains three maps of Texas, highlighting TxDOT districts, Texan counties, and exact locations for each of the testing sites used in this study.

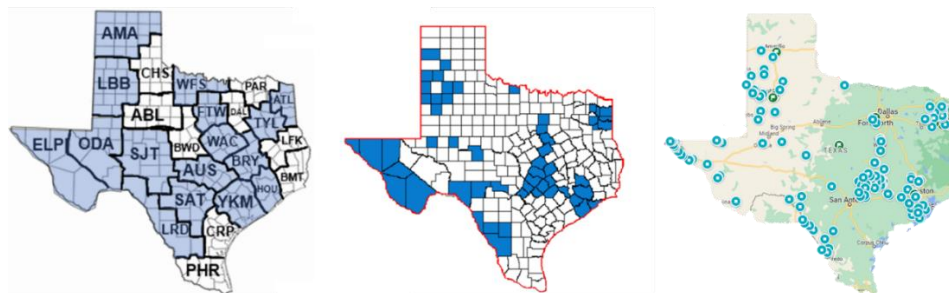


Figure 66: Testing sites within TxDOT districts (left), Texas counties (middle) with tested sites in blue, and exact site locations marked by blue circles (right).

3.1.3. Data Collection Protocol

The Performing Agency made a remote survey of the locations provided by the Receiving Agency's personnel using a combination of Google Street map view and video recording from Pathweb (in house software from Pathway) prior to going to the field. This survey allowed the Performing Agency to pinpoint the exact locations where a pavement surface started and ended. These resources were also used to determine if the pavement had visible distresses, such as raveling, bleeding, cracking, patching, rutting, and punchouts, among others that could potentially affect the quality of the data. The Performing Agency later planned which routes would be taken to travel from one testing site to another and coordinated with

district personnel to let them know that data collection was going to occur during certain dates, in case the Performing Agency required field assistance.

Unlike pavement sections used to evaluate friction, the sections in this project were not necessarily flat or straight. Data was collected at grade, along turns, under varying lighting conditions, and without traffic control to mimic real-world conditions faced by the Receiving Agency's contractors. All texture measurements were collected along the inner wheel path of the rightmost lane. The laser sensor was oriented at a 45-degree angle from the transverse direction to detect patterns of tining, grinding, or grooving on rigid pavement. The average spacing between adjacent points within a 2D profile was 0.18 mm, which varied based on the height of the laser sensor. The laser sensor was installed at a height of 17.5 inches (44.45 cm) from the ground, with an average spacing of 44 mm between adjacent profiles. The vertical resolution of the laser sensor was 5 microns, and each raw profile collected 2046 data points. Pavement surface images were taken as close as possible to the inner wheel path, approximately 30 cm to the right of its center. The camera and laser sensor were synchronized to trigger simultaneously.

Before data collection, the Performing Agency parked the surveying vehicle about half a mile from the start of the pavement site. Researchers mounted the equipment, performed quality control checks, activated the overhead amber lights, and recorded the pavement's surface temperature. The vehicle then traveled in the rightmost lane, accelerating to a speed of 50 mph, at which point cruise control was activated to maintain this speed. Researchers then began collecting data and monitored the equipment to ensure proper operation. Data collection ended a few yards from the end of the pavement surface of interest or if there was an unexpected obstacle or pavement change. Afterward, the researchers parked the vehicle in a safe spot and dismounted the sensors for safe transport. At the end of the day, all raw data files were backed up to the cloud.

3.1.4. Pavement Surfaces Surveyed

The Performing Agency prioritized obtaining a wide variety of distinct pavement surfaces, as previously mentioned. Table 7 summarizes all the types of pavement surfaces provided by the Receiving Agency's personnel. These labels will generally serve as the ground truth for developing machine learning classification models. In a few instances, the surface type information provided by the Receiving Agency did not match the actual site conditions. For example, a location expected to have a high friction surface course turned out to be a rigid pavement with transverse tining. These mismatches constitute less than 5% of the total dataset. Instead of discarding these sections, they will be added to the holdout dataset for proper labeling later on.

Table 7: Summary of all the pavement surfaces provided by TxDOT and surveyed by the Performing Agency.

Flexible Pavements	Rigid Pavements
Dense Graded Mixes (C, D)	AstroTurf or Carpet Drag
Thin Overlay Mix	Longitudinal and Transverse Tining
Superpave (C,D)	Fixed and Random Tining
Stone Mastic Asphalt (Rubber, D)	Conventional Diamond Grinding
Permeable Friction Coarse	New Generation Diamond Grinding
Thin Bonded Wearing Coarse	
Chip Seals (Grades 3, 3S, and 4)	
Fog Seals	
High Friction Surface Treatments	
Microsurfacing	

3.2. Data Collection Results

This subsection provides a summary of the data collection in terms of mileage and the number of pictures taken. Table 8 summarizes the mileage collected by pavement type, showing that the Performing Agency met the requirement of measuring at least 300 miles of flexible pavement and 100 miles of rigid pavement. Table 9 categorizes the mileage collected into the ground truth and holdout datasets. The ground truth dataset includes pavement sections where the surface type is known and matches field observations. The holdout dataset includes sections where the surface type is unknown or where field observations did not match the provided information.

Table 8: Breakdown of mileage for data collected in terms of pavement type.

Type	Mileage (mi.)
Flexible Pavement	313.7 (73.73%)
Rigid Pavement	111.8 (26.27%)
Total	425.5 (100.00%)

Table 9: Breakdown of mileage for data collected in terms of dataset.

Type	Mileage (mi.)
Ground Truth	238.4 (56.03%)
Hold-out	187.1 (43.97%)
Total	425.5 (100.0%)

Figure 67 summarizes the mileage collected, broken down by TxDOT district. El Paso had the most data collected, with almost 80 miles, while Tyler had the least, with three miles. **Figure 68** shows the mileage broken down by the type of pavement surface reported by each district. Some districts provided detailed information, such as the grade of the aggregates, their source, and whether the

aggregate was lightweight. Others simply specified that the pavement was a seal coat.

Figure 69 shows the number of pictures taken, broken down by the type of pavement surface reported by each district. The abbreviations used are as follows: HFST stands for High Friction Surface Course, A represents AstroTurf, SC G3S denotes a grade 3 seal coat with single size gradation, DG refers to conventional diamond grinding, FTT is fixed transverse tining, NGDG indicates new generation diamond grinding, SC G3 signifies a grade 3 seal coat, RLT stands for Random Longitudinal Tining, TOM is a thin overlay mix, TBWC is a thin bonded wearing course, RTT means random transverse tining, SC G4 indicates a grade 4 chip seal, SMA refers to stone matrix asphalt, SC stands for a seal coat of unspecified grade, SP represents a Superpave mix, and PFC denotes a permeable friction course.

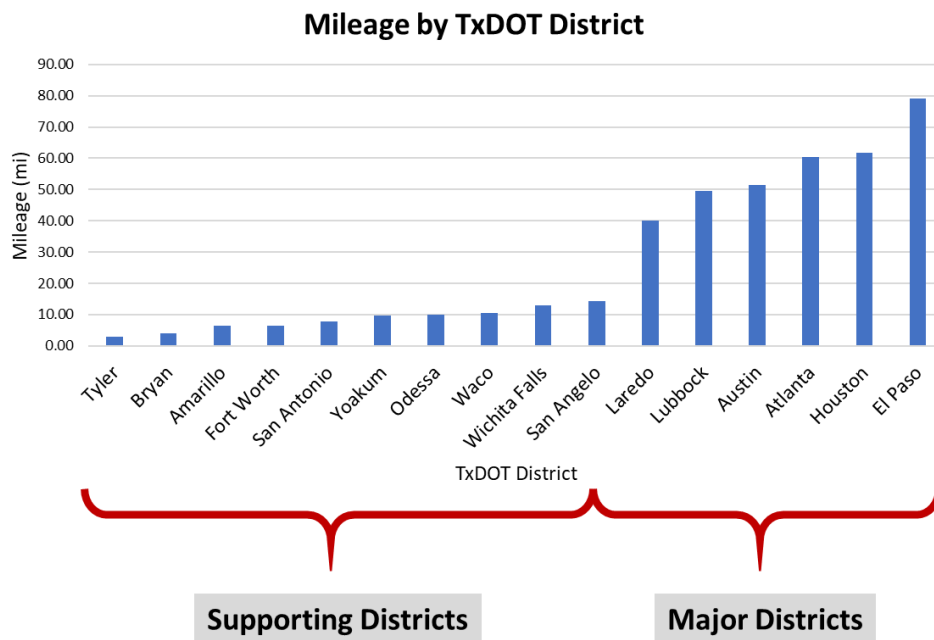


Figure 67: Distribution of mileage surveyed broken down by TxDOT district.

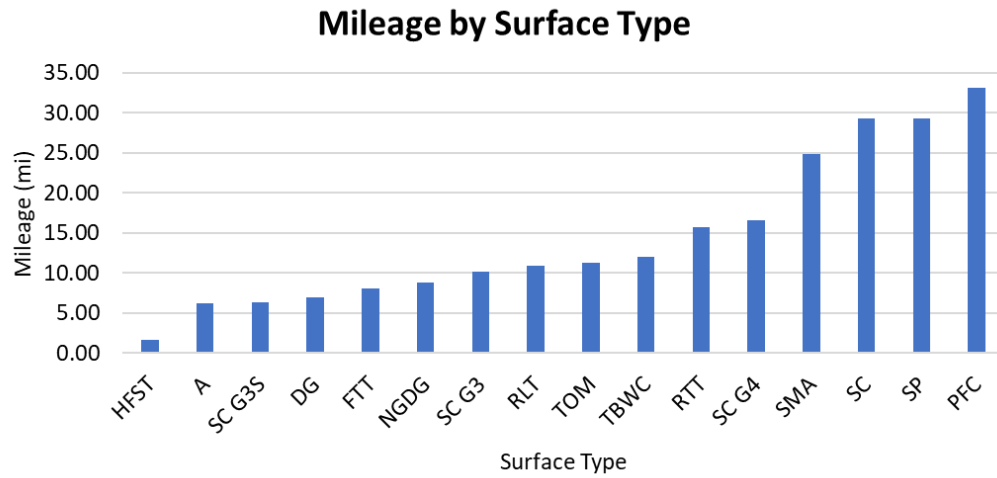


Figure 68: Distribution of mileage surveyed broken down by pavement type.

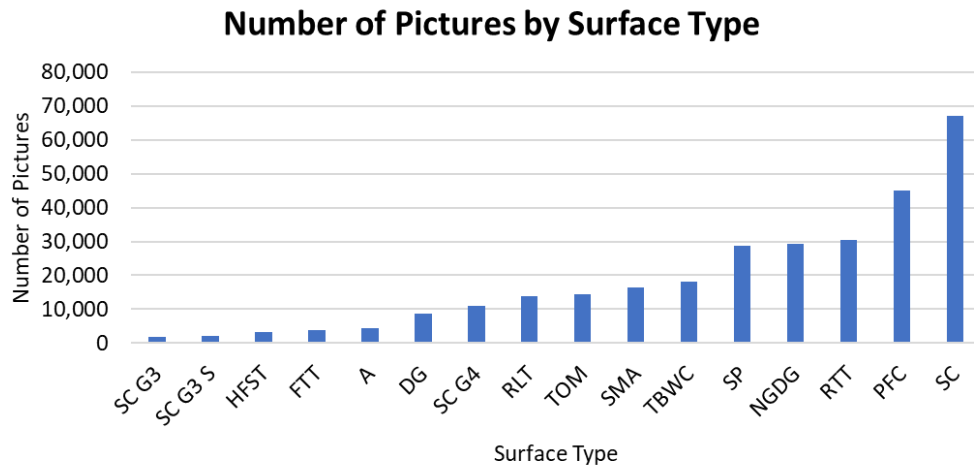


Figure 69: Distribution of number of pictures taken broken down by pavement type.

Chapter 4. Database of Pavement Surface Types

4.1. The Database

The database for Product #1 (P1) will be delivered to the Receiving Agency in the form of an external hard drive disk named “0-7139 Data” which comprises pavement texture data coming from two-dimensional (2D) profiles and area pictures of the pavement surface from a high-speed camera. These data can be found inside the folder “*TxDOT Project 0-7139 Surface Determination*,” following the folders and files shown in **Figure 70**:

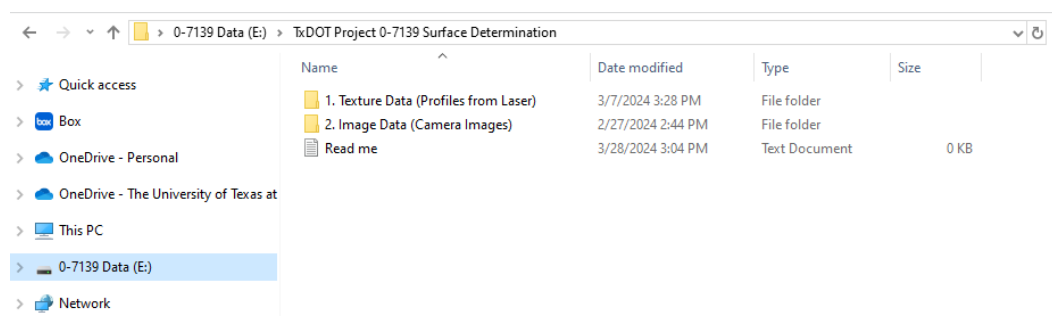


Figure 70: Screenshot of the “*TxDOT Project 0-7139 Surface Determination*” folder view

4.2. Texture Database

The “*Texture Data (Profiles from Laser)*” folder is organized into three subfolders and one Excel file, as shown in **Figure 71**.

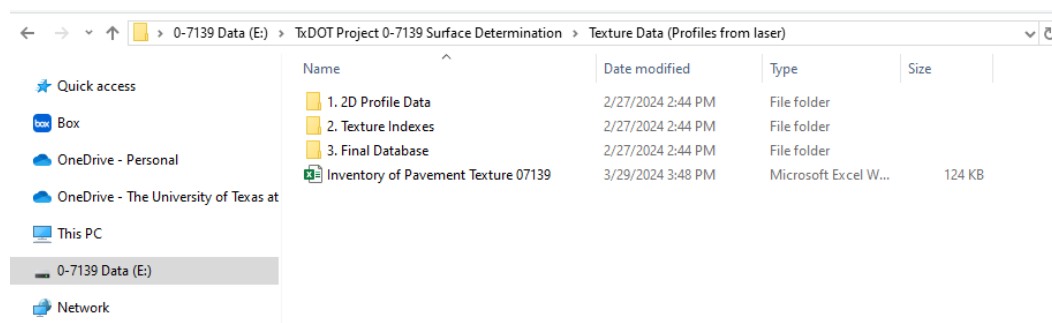


Figure 71: Screenshot of the “*(Profiles from Laser)*” folder view

The folder entitled “*1. 2D Profile Data*” holds both the raw and processed profile data. The “*2. Texture Indexes*” folder contains the computed texture indexes for each profile and the same indexes aggregated into 0.1-mile pavement sections. Lastly, the “*3. Final Database*” subfolder consolidates texture index data for all sites, while also adding details about testing locations such as district, county,

approximate GPS locations, among others. Further explanations of these directories are provided in subsequent sections. The Excel file in this folder holds the inventory information for each site.

4.2.1. 2D Profile Data

The folder entitled “1. 2D Profile Data” contains the raw data, raw profiles, and the processed profiles, as shown in **Figure 72**. Each of these folders is then subdivided by TxDOT District, as shown in **Figure 73**.

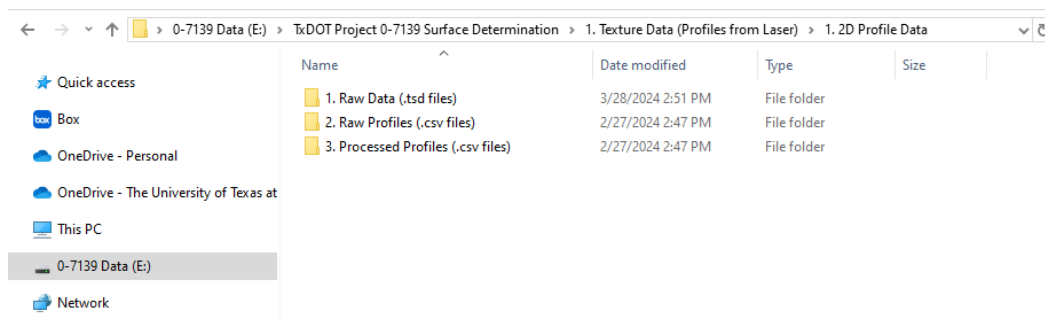


Figure 72: Screenshot of the “1. 2D Profile Data” folder view

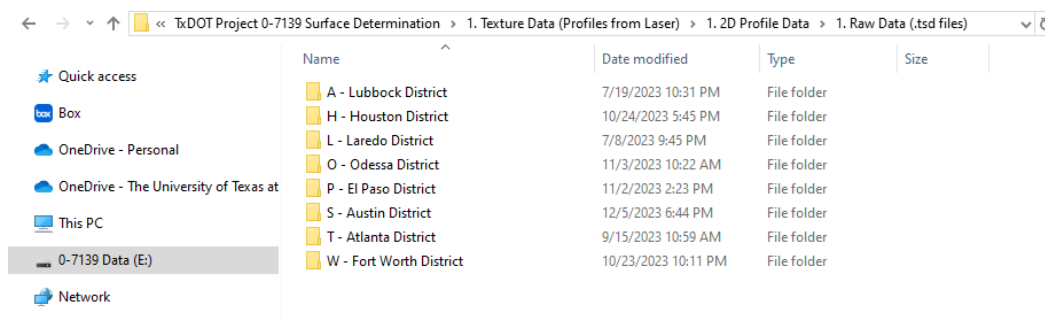


Figure 73: Screenshot of the “1. Raw Data (.tsd files)” folder view

4.2.1.1. Raw Data

The “1. Raw Data (.tsd files)” folder contains the initial output from the data collection software, featuring both range and intensity measurements for each surveyed site. This data remains in its original, unaltered form, meaning it has not undergone trimming, detrending, denoising, imputation or any other data manipulation technique. The naming of these files adheres to a specific convention determined by the Performing Agency, structured as follows:

X#_SiteName_TestType_#

The naming convention for most files in the “Raw Data” folder begins with an uppercase letter which represents the primary location where data was collected. Each letter is associated with a specific TxDOT district, as detailed in Table 10.

Table 10: List of letters used to designate TxDOT’s district.

Letter	District	Letter	District	Letter	District
A	Lubbock	O	Odessa	T	Atlanta
H	Houston	P	El Paso	W	Fort Worth
L	Laredo	S	Austin		

Following this initial character there is a number that represents the sequence in which the sites were surveyed. An underscore then precedes a designated name for the site, chosen to reflect the nearest city, town, county, or notable landmark, though the selection is largely arbitrary. Another underscore introduces the test type whose nomenclature will be explained using examples in the next paragraph. The final segment, separated by an underscore, is a numerical batch identifier for the site, starting at 0 and incrementing by one for each batch. A “batch” refers to a collection of 15,360 profiles, which represents the maximum number of profiles per file that the data collection software allows. An example of the naming convention using real data from the project is shown in **Figure 74**.

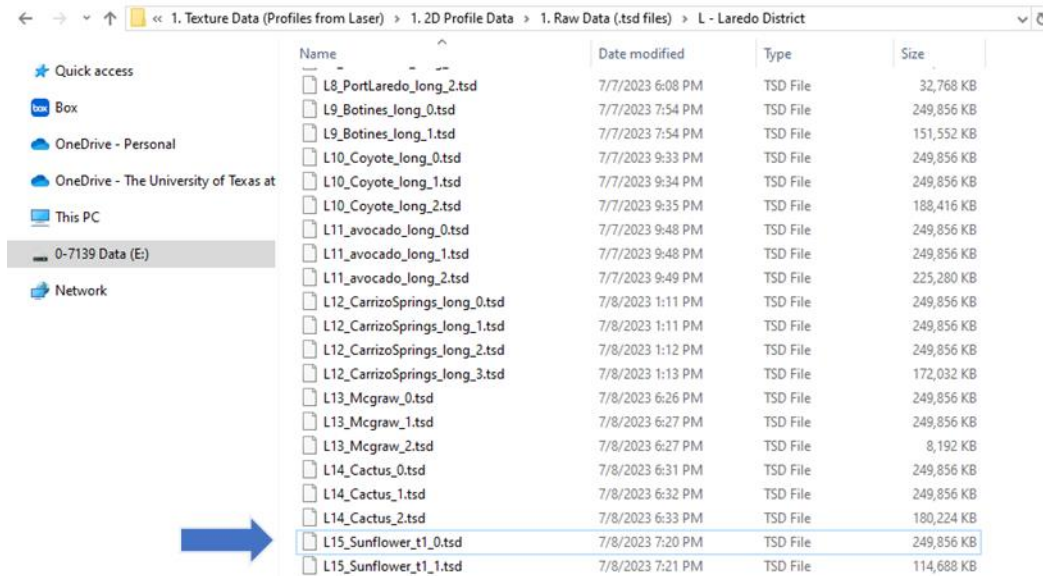


Figure 74: Screenshot of the “L – Laredo District” subfolder view within the “1.Raw Data (.tsd files)” view, showing the nomenclature used for file naming.

The file name “L15_Sunflower_t1_0.tsd” starts with the letter L, meaning that the data was collected in or near the Laredo District. It is then followed by the number 15, meaning that the site was the 15th in the survey sequence. The proximity of a sunflower field is noted in the site’s descriptive name “Sunflower.” Since the test conducted was the first of multiple runs (repeated run), the term “t1” was employed. Finally, because this file pertains to the very first batch of data collected, it was designated by the number “0” at the end of the file name.

Conversely, a file named “A4_Cricket_long_3” or “A4_Cricket_3” indicates that the site was the fourth site surveyed in the vicinity of the Lubbock District as indicated by the term, A4. This site was labelled “Cricket” arbitrarily due to the notable presence of crickets at the location and not because there was a town named Cricket. The term “long” at the end of the filename (or its omission) signifies the execution of a single long run test. Finally, the batch identifier “3” specifies that this file contains the fourth batch of collected data.

The range (texture profile) data contained within the .tsd files was subsequently transformed into .csv format to facilitate manipulation via statistical coding software. Nonetheless, a discrepancy can be observed for testing sites in the Austin, Laredo, and Lubbock Districts, shown in Table 11. This happened because the original .tsd files for a few locations in these districts were not successfully uploaded to cloud storage following their conversion to .csv format, resulting in their loss. Despite the issue, all the range data within these files was successfully saved as .csv and stored within this database. This implies that the intensity data for the sites shown in Table 11 is lost. Nonetheless, because no practical application was found for the intensity data, the Performing Agency did not utilize it.

Table 11: Testing sites for which the .tsd files are not available

Austin District	Laredo District	Lubbock District
S1 – MopacRiver	L20 – Rattlesnake	A13 – Lockney
S2 – DripSpring	L21 – Comstock	A14 – Dead
S3 – BeeCaves	L22 – Cattle	A15 – HaleCounty
S4 – LagoVista	L23 – MountainHome	A16 – Silo
	L24 - Picnic	A17 – Peter

4.2.1.2. Raw Profiles

The folder entitled “2. Raw Profiles (.csv files)” contains the unprocessed range data from the .tsd files, now reformatted into CSV (Comma-Separated Values) format. This transformation enhanced the ease of data processing and manipulation, particularly when utilizing statistical and database management tools such as RStudio, Jupyter Notebook, or MATLAB. The naming convention for these files mirrors that of the .tsd files. Structurally, the raw files consist of 2,048 columns (points comprising a texture profile), with the number of rows varying according to the length of each test site. Each row represents a distinct texture profile. The average spacing between points along the transverse direction (from one column to the next) varies with the laser sensor’s height during measurement. The spacing information is documented in the “Texture_Database.csv” file. The standard spacing between adjacent profiles (from row to row) is approximately 0.44 mm.

Each cell entry in these files represents a “biased and trended” elevation of the pavement surface, measured in decimeters (dm). The term “biased and trended” means that a systematic positive bias is intentionally added to the elevation measurements. This bias helps in storing most data points as large integers, simplifying data storage compared to storing numbers with decimal points. Moreover, the elevation measurements might show a positive or negative trend depending on the road’s grade or lateral slope, which influences the angle of the laser sensor during measurements.

For analysis, these measurements need to be converted into millimeters (mm) using a conversion factor of 0.01. To correct the biases and trends in each profile, a linear regression detrending method can be applied. This process effectively centers (detrends) the data, preparing it for further analysis. While there is no universally accepted standard for these data processing steps, the Performing Agency recommends the following routine: (1) convert the data to millimeters; (2) remove noisy data points; (3) impute the removed data; and (4) detrend the data.

Considering the significant size of the data, the Performing Agency recommends employing advanced database management tools such as RStudio, Oracle, or MATLAB for efficient data handling and analysis. Utilizing simpler programs like Microsoft Excel or Access to open these files might lead to partial data retrieval due to memory limitations. To enhance storage efficiency, all eight district subfolders within the “2. Raw Profiles” directory have been compressed, as shown in **Figure 75**.

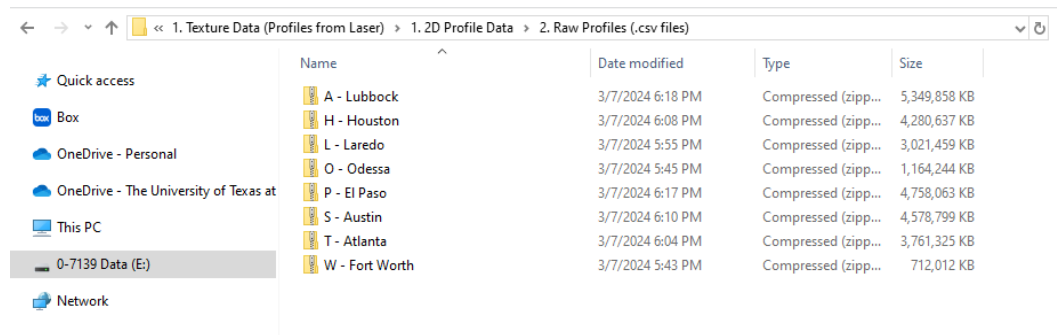


Figure 75: Screenshot of the “2. Raw Profiles (.csv files)” view, showing the that folders have been compressed by district.

4.2.1.3. Processed Profiles

The folder was labeled “3. Processed Profiles (.csv files)” contains profiles that have undergone comprehensive processing using a texture processing algorithm. The Performing Agency conducted rigorous quality control on the raw data to ensure the cleanest data possible was used. Initially, the data was converted from decimeters to millimeters. Subsequently, any data points outside the central 150

mm length of each profile were removed. The number of points within this central segment varies between sites, depending on the point-to-point spacing, ranging from 587 points (with a spacing of 0.256 mm) to 933 points (with a spacing of 0.161 mm). This trimming step reduces the data size to a manageable level and eliminates potentially inaccurate measurements typically found near the profile edges. The selected 150 mm length is optimal for calculating key texture indexes such as the mean profile depth (MPD)

Following the trimming, the profiles underwent further scrutiny via a specialized texture processing algorithm, namely the Sabillon-Orellana Filtering Algorithm (SOFA). For more details on this algorithm, please see Sabillon-Orellana et al (3). In brief, SOFA identifies and removes both spike and flat signal noise from the profile data. It also excludes any profiles with over 33% missing or noisy data, fills gaps using linear interpolation, and applies linear regression detrending to centralize the data. The processed profiles provide an accurate representation of the pavement surface in millimeters, characterized by minimal noise.

The nomenclature of the files in this folder is similar to that of the raw data with the exception that now the files have the suffixes “*DN_trim*,” indicating that the files were first trimmed and then denoised. Lastly, the eight district subfolders inside “3. *Processed Profiles*” have all been compressed similar to how they were in the “Raw Profiles” folder.

4.2.2. Texture Indexes

There are two folders in the “2. *Texture Indexes*” folder (**Figure 76**) that are subdivided into eight folders named after a TxDOT District.

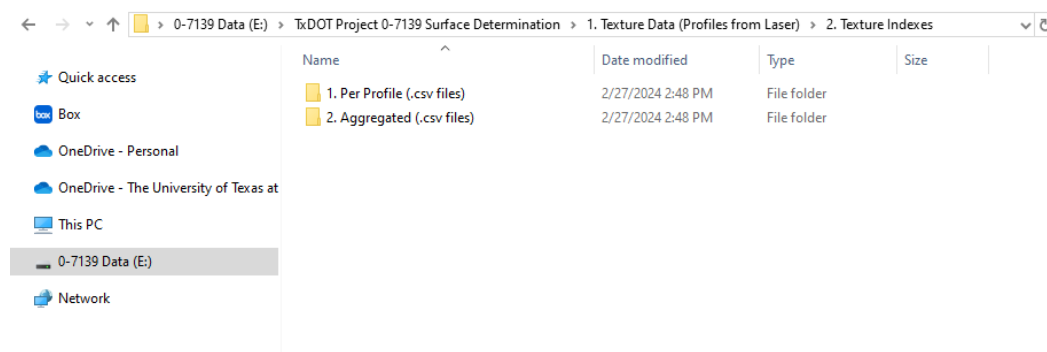


Figure 76: Screenshot of the “Texture Indexes” folder view, showing the that folders have been compressed by district.

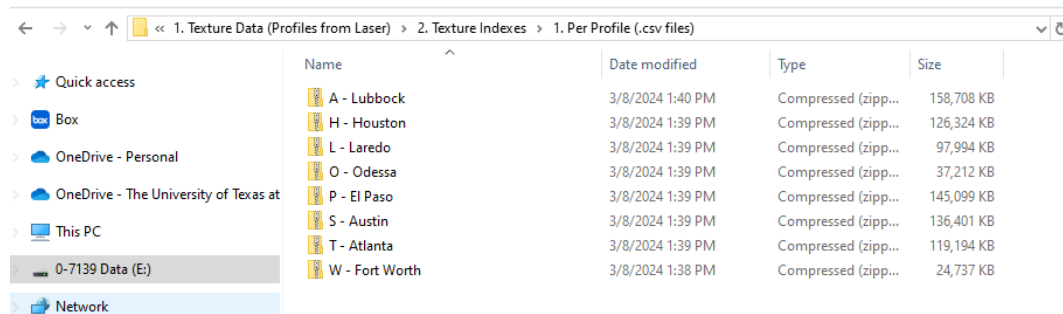
Careful analysis of these indexes revealed significant linear correlations among them. This finding is helpful since it implies that indexes that are difficult to understand or compute can be estimated with a high level of goodness-of-fit using

a more familiar or easier to compute index. Indexes exhibiting a correlation higher than 60% were deemed redundant and subsequently excluded from the study, with preference given to those that were simpler to both compute and understand. Exceptions were made for those variables with established relevance in pavement engineering such as mean profile depth, the root mean square of height, and the two-point slope variance. Also, the continued inclusion of these particular indexes is justified by their widespread recognition and application in the field, rendering them invaluable to industry professionals despite the redundancy.

This correlation criteria reduced the effective list of indexes from 80 to 19 given the correlation criteria. Despite that, feasible computation time proved key when choosing which variables to use in the final database. Even though a few indexes were uncorrelated with others, their computation time was still excessively lengthy for practical application of texture processing at a network-level scale. Consequently, this consideration required a further reduction of the list, resulting in a selection of 12 indexes: Root Mean Square Height (RMS), Kurtosis (Rk), Solidity Factor (Rs), Mean Cross Width (Cm), Standard Deviation of Cross Width (Cstd), Skewness of Cross Width (Cs), Maximum Absolute Gradient (Rdt), Two-Point Slope Variance (Sv2), Autocorrelation Length (Ral), Mean Local Curvature (Hcme), Standard Deviation of Local Curvature (Hcst), and Mean Profile Depth (MPD)

4.2.2.1. Texture Indices per Profile

The “1. Per Profile (.csv files)” folder (**Figure 77**) contains eight compressed folders for each of the TxDOT districts. Within each district folder, there are CSV files for test sites, as shown in **Figure 78**. All CSV files contain the set of indexes described in Section 4.2.2, as depicted in **Figure 77**. Notice that the nomenclature of the CSV files is similar to that within the “3. Processed Profiles” folder, with the addition of a “IND_” suffix standing for the word “index” after the profiles were trimmed and denoised.



Name	Date modified	Type	Size
A - Lubbock	3/8/2024 1:40 PM	Compressed (zip...	158,708 KB
H - Houston	3/8/2024 1:39 PM	Compressed (zip...	126,324 KB
L - Laredo	3/8/2024 1:39 PM	Compressed (zip...	97,994 KB
O - Odessa	3/8/2024 1:39 PM	Compressed (zip...	37,212 KB
P - El Paso	3/8/2024 1:39 PM	Compressed (zip...	145,099 KB
S - Austin	3/8/2024 1:39 PM	Compressed (zip...	136,401 KB
T - Atlanta	3/8/2024 1:39 PM	Compressed (zip...	119,194 KB
W - Fort Worth	3/8/2024 1:38 PM	Compressed (zip...	24,737 KB

Figure 77: Screenshot of the “Per Profile (.csv files)” folder view

Name	Date modified	Type	Size
Ind_DN_trim_01_Hammett	11/11/2023 12:42 PM	Microsoft Excel C...	4,588 KB
Ind_DN_trim_02_Averitt	11/11/2023 1:59 PM	Microsoft Excel C...	16,196 KB
Ind_DN_trim_03_Baiza	11/11/2023 3:02 PM	Microsoft Excel C...	13,332 KB
Ind_DN_trim_04_Odessa	11/11/2023 3:47 PM	Microsoft Excel C...	9,787 KB
Ind_DN_trim_05_GardenCity	11/11/2023 4:43 PM	Microsoft Excel C...	12,418 KB
Ind_DN_trim_06_ConchoValley	11/11/2023 5:02 PM	Microsoft Excel C...	4,131 KB
Ind_DN_trim_07_SanAngelo	11/11/2023 7:04 PM	Microsoft Excel C...	27,329 KB

Figure 78: Screenshot of the “O - Odessa” subfolder view within the “Per Profile (.csv files)” folder.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	distance	na_counter	RMS	Rku	Rso	MPD	Cwm	Cwst	Cws	Ral	Sv2	Rdt	Hcme	Hcst	
2	1	16.01362862	0.132108	-0.30983	-0.80632	0.32747	3.248883	4.230685	1.607992	13.02885	0.136662	0.645336	-0.01601	0.825179	
3	6	14.13969336	0.168398	1.544349	-0.53798	0.317798	3.32108	4.240302	2.001187	6.64216	0.124113	1.138938	-0.00142	0.742534	
4	8	14.82112436	0.135247	0.678842	-0.80918	0.31419	3.049972	4.525219	2.587308	22.48116	0.142399	0.998845	0.00675	0.750503	
5	9	13.9693356	0.145254	0.591326	-1.18685	0.457088	3.396559	4.073192	2.10694	12.51792	0.138424	0.470653	-0.01237	0.741055	
6	10	15.84327087	0.12277	0.660373	-1.17102	0.331492	3.32108	3.840135	1.834445	9.707773	0.132025	0.803491	-0.0058	0.685644	
7	13	16.69505963	0.162291	0.518116	-0.57558	0.390168	4.395547	4.429199	1.196108	7.919499	0.144577	1.884199	0.002971	0.737116	
8	14	18.22827939	0.266484	3.476776	-0.5996	0.476426	3.832016	5.815823	2.552219	9.707773	0.159349	0.501229	-0.00155	0.748118	
9	15	18.56899489	0.18356	-0.75574	-0.76333	0.335956	4.26996	7.461121	2.41292	18.64914	0.134319	0.841745	-0.00274	0.814309	
10	18	14.65076661	0.153206	-0.10983	-0.98	0.251059	3.736215	4.963721	2.075843	14.81713	0.144118	0.812422	0.001607	0.787176	
11	19	17.37649063	0.20126	-0.02488	-0.72371	0.425082	5.33745	8.931859	3.171935	19.92648	0.149541	0.591511	0.006517	0.75663	
12	23	15.84327087	0.217751	1.635921	-0.90036	0.522143	3.475549	5.427065	3.424571	9.96324	0.166088	0.784616	-0.00757	0.798425	
13	24	19.9318569	0.178376	0.927594	-1.03752	0.55973	3.32108	3.547399	2.114006	3.832016	0.166169	0.588243	0.00119	0.797014	
14	25	16.35434412	0.193645	-0.10721	-1.02356	0.411484	3.736215	5.958953	3.182592	14.05072	0.145197	0.542064	0.007157	0.777488	

Figure 79: Screenshot from one of the CSV files within the “Per Profile (.csv files)” folder showing all the variables contained within it.

Each CSV file has 14 columns output (two general properties and the 12 indexes from Section 4.2.2 within it). The following is a data dictionary to understand them:

PER PROFILE (.CSV FILES) DATA DICTIONARY

1. **“distance”**: This is a numerical variable. It *DOES NOT* actually represent Distance; instead, it is an index indicating the order in which the profiles were collected. It is possible for some indexes to be missing. For instance, in **Figure 79** the values 2, 3, 4, and 5 are missing. That means that those profiles had more than 20% noise or missing data which is the threshold level of uncertainty for this study.
2. **“na_counter”**: This is a numerical variable: It ranges from 0 to 100% and indicates the perctange of data points within the profile that were either missing or removed by the SOFA for being considered noise. The value of this variable will never exceed 20%.
3. **“RMS”**: This is a numerical variable. It represents the amplitude index “root mean square height” (ID# 2)³. RMS is measured in units of mm.
4. **“Rku”**: This is a numerical variable. It represents the amplitude index

³ ID refers to the ID displayed in Table 1: List of indexes reviewed and their respective source.

“kurtosis” (ID# 10)³. Rku is a unitless index.

5. **“Rso”**: This is a numerical variable. It represents the amplitude index “solidity factor” (ID# 7)³. Rso is a unitless index.
6. **“MPD”**: This is a numerical variable. It represents the feature index “mean profile depth” (ID# 57)³. MPD is measured in units of mm.
7. **“Cwm”**: This is a numerical variable. It represents the spacing index “mean of cross width” (ID# 10)³. Cwm is measured in units of mm.
8. **“Cwst”**: This is a numerical variable. It represents the spacing index “standard deviation of cross width” (ID# 11)³. Cwst is measured in units of mm.
9. **“Cws”**: This is a numerical variable. It represents the spacing index “skewness of cross width” (ID# 12)³. Cws is a unitless index.
10. **“Ral”**: This is a numerical variable. It represents the functional index “autocorrelation length” (ID# 36)³. Ral is measured in units of mm.
11. **“Sv2”**: This is a numerical variable. It represents the hybrid index “two-point slope variance” (ID# 22)³. Sv2 is measured in units of mm.
12. **“Rdt”**: This is a numerical variable. It represents the hybrid index “maximum absolute gradient” (ID# 19)³. Rdt is measured in units of mm..
13. **“Hcme”**: This is a numerical variable. It represents the functional index “mean of local curvature” (ID# 55)³. Hcme is measured in units of mm^{1/2}.
14. **“Hcst”**: This is a numerical variable. It represents the functional index “standard deviation of local curvature” (ID# 56)³. Hcst is measured in units of mm^{1/2}.

4.2.2.2. Aggregated Texture Indexes

The “2. Aggregated (.csv files)” folder contains eight uncompressed folders for each of the TxDOT districts, as shown in **Figure 80**. The information contained within these folders has now been aggregated or summarized every 0.1 miles similar to the pavement management systems used by the Receiving Agency. The aggregation involves taking the mean, median, and standard deviation for the 12 texture indexes shown in Section 4.2.2

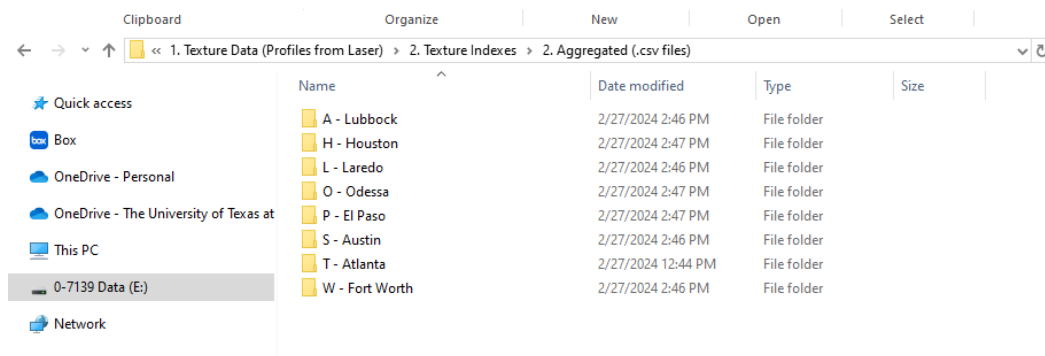


Figure 80: Screenshot of the “Aggregated (.csv files)” folder view

Inside each of the eight district folders, there will be a single CSV file, as opposed to one file per test site. That file will have the name of the district followed by the terms “_texture_collection,” as shown in **Figure 81**.

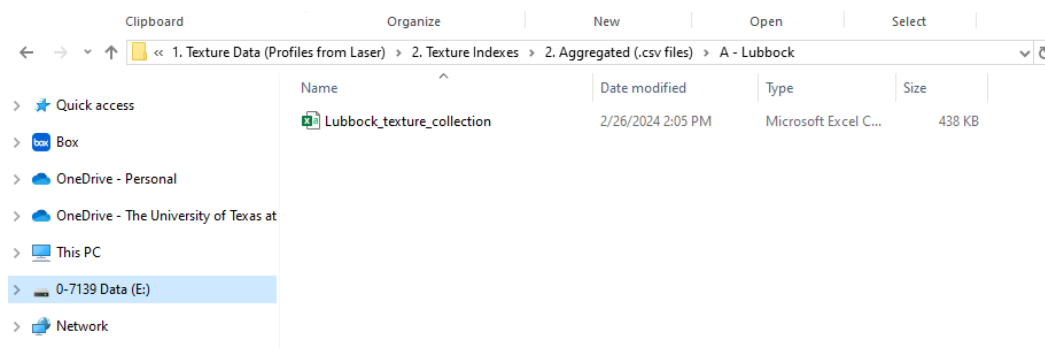


Figure 81: Screenshot of the “A - Lubbock” folder view inside the “Aggregated (.csv files)” folder.

The following is a data dictionary that explains the variables found inside the CSV files exemplified in **Figure 81**:

AGGREGATED (.CSV FILES) DATA DICTIONARY

1. **“Site_Number”:** This is a categorical variable. This variable is composed by combining of the uppercase characters shown in Table 10 and the number which indicates the order in which this section was surveyed.
2. **“Type_of_Test”:** This is a categorical variable: It indicates the type of test conducted on the test site. “Long” indicates the test site was surveyed a single time, and “Repeat” indicates it has been tested at least twice.
3. **“Run”:** This is a numerical variable. It identifies the chronological order of

data collection, especially on sites with repeated measurements. A value of 1 indicates the data comes from the first run, 2 for the second, and so on. For testing sites where the “Type_of_Test” is “Long” this variable will be strictly marked as 1.

4. **“n”**: This is a numerical variable. The number of rows or profiles that were summarized (by using of a median, mean, or standard deviation) for the 0.1 mile long pavement section in consideration.
5. **“Mile_group”**: This is a numerical variable. This variable starts at 0.1 and increases in units of 0.1 miles. It represents a pavement section that is 0.1 miles long, as it would in the Pavement Management Systems used by the Receiving Agency.
6. **“###_med”**: The “###” term stands for any of the texture indexes discussed in Section 4.2.2. All the variables ending in “_med” represent the median for the index in consideration over the 0.1 mile long pavement section.
7. **“###_mean”**: The “###” term stands for any of the texture indexes discussed in Section 4.2.2. All the variables ending in “_mean” represent the arithmetic average for the index in consideration over the 0.1 mile long pavement section.
8. **“###_sd”**: The “###” term stands for any of the texture indexes discussed in Section 4.2.2. All the variables ending in “_sd” represent the sample standard deviation for the index in consideration over the 0.1 mile long pavement section.

4.2.3. Final Texture Database

The Performing Agency compiled all the texture data collected during Task 4 alongside other information about the testing site, such as GPS location, direction of travel of the vehicle, county, and district locations, among others, onto a single CSV file entitled “*Texture_Database.csv*.” **Figure 82** shows the “3. *Final Database*” folder view.

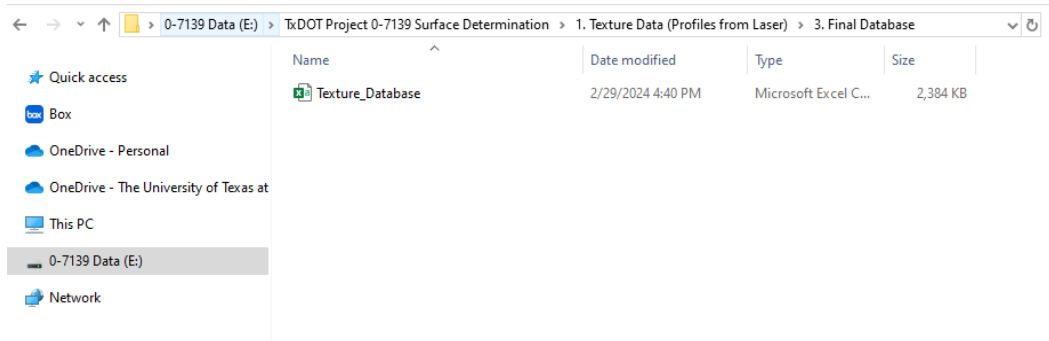


Figure 82: Screenshot the “3. Final Database” folder view.

The following is a data dictionary describing all the relevant information found inside the “*Texture_Database.csv*” file. Note that since this database is used to classify a vast array of pavement surfaces, some of the variables will reference a so-called *level of specificity*, which will be explained in more detail in Chapter 5. In summary, the *level of specificity* is an integer ranging from 1 to 5, where one is the most basic and generalized grouping of pavements (pavement type: flexible vs rigid), and the higher the number goes, the more specific the classification is. For instance, an arbitrary portland cement concrete pavement in Texas could have the following levels of specificity:

- Level 1: Rigid Pavement
- Level 2: Plastic Surface Finished (textured while concrete was fresh)
- Level 3: Tinned Surface Texturing
- Level 4: Transverse Tinning
- Level 5: Fixed Transverse Tinning

Figure 83 depicts the tree-diagram for the specificity levels used for this project.

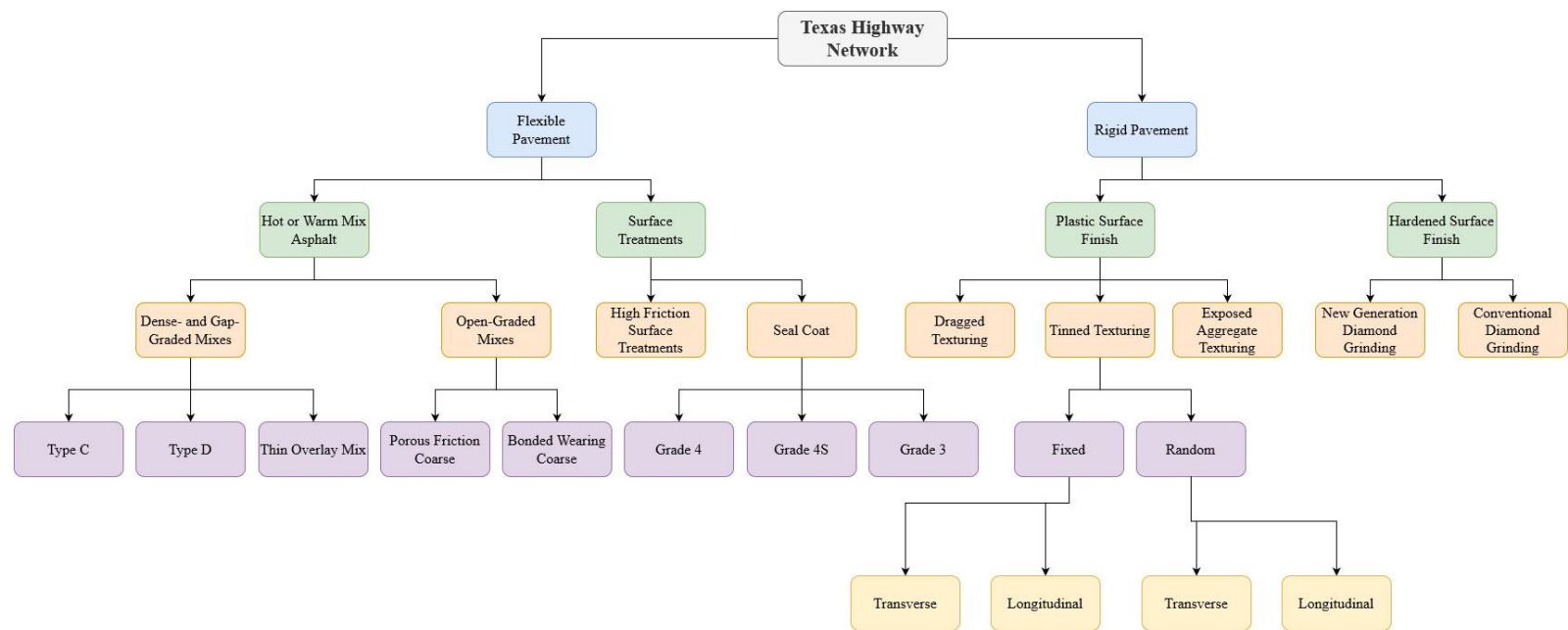


Figure 83: Diagram indicating the different levels of specificity for the pavements surveyed based on their surface type.

TEXTURE_DATABASE DATA DICTIONARY

1. ***“Folder Name”***: This is a categorical variable. It represents one of the eight TxDOT district folder names which have been used in this study.
2. ***“District”***: This is a categorical variable: It represents the TxDOT district where the section of interest can be located. Note that there are two sections that were collected in the state of New Mexico due to their proximity to a pavement section of interest. Those sections will have a suffix (NM) to indicate they come from New Mexico followed by the corresponding district name that the New Mexico DOT has assigned to that region.
3. ***“County”***: This is a categorical variable. It represents the Texas county where the section of interest can be located. Note that there are two sections that were collected in the state of New Mexico due to their proximity to a pavement section of interest. Those sections will have a suffix (NM) to indicate they come from New Mexico followed by the corresponding New Mexico county name.
4. ***“Site_Number”***: This is a categorical variable. This variable is composed by combining of the uppercase characters shown in Table 10 and the number which indicates the order in which this section was surveyed.
5. ***“Site_Name”***: This is a categorical variable. It shows the name that was assigned to the section of interest. As mentioned earlier, this name tends to be related to the closest town, city, or landmark, but it is ultimately arbitrary.
6. ***“Highway”***: This is the name of the highway ID using the Receiving Agency’s nomenclature. This variable does not include roadbed information.
7. ***“Direction”***: This is a categorical variable. It gives a rough estimate of the vehicle’s direction of travel, possible values include:

EB – Eastbound	NW – Northwest	SW – Southwest
NB – Northbound	SB – Southbound	WB – Westbound
NE – Northeast	SE – Southeast	
8. ***“Approx_GPS_Location”***: This is a set of numerical variables. It contains the longitude and latitude coordinates for the testing site using the decimal degree system.

9. ***“Pavmt_Type”***: This is a categorical variable. It indicates if the pavement is “Flexible” or “Rigid.” This is also the first level of specificity within **Figure 83**.

10. ***“General Surface Type”***: This is a categorical variable. It represents a general category for pavement surface as determined by the Performing Agency. Note that this variable is only meaningfully defined for testing sites where the Receiving Agency provided information about the surface type. This also represents the second level of specificity shown in **Figure 83**. Possible categories include:

HWMA	Plastic	Unknown
Surface Treatment	Hardened	

HWMA stands for hot and warm mix asphalt pavements. Surface treatment encompasses all flexible pavements consisting of seal coats, microsurfacing, high friction surface treatment and thin overlay mixes. Plastic represents rigid pavement whose surface finish was done while the portland Cement Concrete (PCC) was in its “plastic” or fresh state. Hardened represents rigid pavement whose surface finish was done after the PCC had already fully hardened. Finally, Unknown indicates that the section surface type is not fully known.

11. ***“Specific Surface Type”***: This is a categorical variable. It represents a more specific classification for pavement surface as determined by the Performing Agency. Note that this variable is only meaningfully defined for testing sites where the Receiving Agency provided information about the surface type. Furthermore, this is the third level of specificity within **Figure 83**. Possible categories include:

Dense- and Gap-Graded	Dragged	NGDG
Open-Graded	Tining	Unknown
Seal Coat	Exposed Aggregate	
High Friction Surface Treatment	CDG	

Dense- and Gap-Graded mixes cover the coarse and fine graded mixes as well as Superpave (SP) and stone matrix asphalt (SMA) mixes. Open-Graded encompasses permeable friction and bonded wearing courses that were surveyed. Seal Coat covers all the gradations of seal coats as well as fog seals and microsurfacing. High Friction Surface Treatment (HFST) are strictly HFST, or seal

coats built with volcanic rock also known to provide high friction levels. Dragged surfaces include Astroturf, carpet and burlap drag. Exposed Aggregates only encompasses that surface type. Tinning accounts for transverse and longitudinal tining as well as fixed and random tining. CDG stands for conventional diamond grinding. NGDG stands for new generation diamond grinding. Finally, “Unknown” indicates that the site’s surface type is not fully known.

12. “Specific Surface Type II”: This is a categorical variable. It represents a more specific classification for pavement surface as determined by the Performing Agency. Note that this variable is only meaningfully defined for testing sites where the Receiving Agency provided information about the surface type. Furthermore, this the fourth level of specificity within **Figure 83**. Possible categories include:

Type C	SC Grade 4	Longitudinal
Type D	Fog Seal	Transverse
TOM	Microsurfacing	Unknown
SC Grade 3	PFC	
SC Grade 3S	BWC	

Type C, Type D refer to hot mix asphalt pavements following the mix design options within Tex-204-F (204). Gradations type C for SMAs, SPs, TOMs, and dense-graded mixes all look identical for surface texture point of view (even though they are different from the structural and design aspect) thus were group together across the mix designs. The same was done for Type D. TOM covers all dense- and gap-graded mixes whose gradation is composed mostly of fine aggregates. The term SC stands for seal coat and the gradation numbers correspond to those found in Item 302: Aggregates for Surface Treatments within the Standard Specifications for Construction and Maintenance of Highways, Streets, and Bridges (114). Fog seal and microsurfacing correspond their respective surface types as defined in Item 315 and Item 350, respectively, of the aforementioned standard (114). PFC and BWC stand for permeable friction course a bonded-wearing course, respectively and are defined as per Tex-204-F (204). In this project only thin-bonded wearing courses were evaluated. The terms longitudinal and transverse refer to the directionality of the tining within a rigid pavement. Finally, “Unknown” indicates that the section surface type is not fully known or not applicable.

13. “Specific Surface Type III”: This is a categorical variable. It represents a more specific classification for pavement surface as determined by the Performing Agency. Note that this variable is only meaningfully defined for testing sites where the Receiving Agency provided information about the surface type. Furthermore, this is the fifth level of specificity within **Figure 83** and it is only applicable to rigid pavements. Possible categories include:

Fixed

Random

Unknown

Fixed indicates that the spacing within the tines of the pavements are consistent, whereas random indicates the spacing between tines is randomly distributed. “Unknown” indicates that the section surface type is not fully known or not applicable.

14. “Surface”: This is a categorical variable. It represents the true surface type of the pavement as indicated by the Receiving Agency.

Astroturf	Recycled Plastic Mix
Burlap	Seal Coat
Burlap and Transverse Tining	Seal Coat GR 3 Lightweight
Dense Type C	Seal Coat GR 4 hard rock
Exposed Aggregates	Seal Coat GR 4 Lightweight
Fixed Transverse Tining	Seal Coat Volcanic Rock
Fog Seal	SMA
Grade 4 / Hot-Rubber Seal	SMA-HMA
HFST	SP-C
Longitudinal Tining	SP-D
Microsurfacing	TBWC
NGCS	TOM
NGDG	TOM-C
PB 3S (Gradation 3S)	Transverse Tining
PD 3	Transverse Tining and Carpet Drag
PFC	Unknown
Random Transverse Tining	

Where, **HFST** stands for High Friction Surface Treatment; **NGCS**, Next Generation Concrete Surface; **NGDG**, Next Generation Diamond Grinding; **PB 3S**, Precoated Crushed Gravel, Crushed Slag, Crushed Stone or limestone rock asphalt with a “3S” aggregate gradation; **PD 3**, Precoated Crushed Gravel, Crushed Slag, Crushed Stone with a “3” aggregate gradation; **PFC**, Permeable friction course; **SMA**, Stone Matrix Asphalt; **HMA**, Hot Mix Asphalt; **SP-C**, SuperPave Type C; **SP-D**, SuperPave Type D; **TBWC**, Thin

Bonded Wearing Coarse; **TOM**, Thin Overlay Mix; and **TOM-C**, Thin Overlay Mix Type C.

15. ***“Type_of_Test”***: This is a categorical variable. It indicates the type of test conducted on the test site. “Long” indicates the test site was surveyed a single time, and “Repeat” indicates it has been tested at least twice.
16. ***“Run”***: This is a numerical variable. It identifies the chronological order of data collection, especially on sites with repeated measurements. A value of 1 indicates the data comes from the first run, 2 for the second, and so on. For testing sites where the “Type_of_Test” is “Long” this variable will be strictly marked as 1.
17. ***“Provider”***: This is a categorical variable: This variable specifies the source of the surface information and the test section’s location. When set to “TXDOT,” it indicates that the surface details and site location were provided by the Receiving Agency. If set to “CTR,” then the Performing Agency found the site, but the surface information is considered unknown.
18. ***“Texture_Point_Spacing”***: This is a numerical variable. It represents the point-to-point spacing within each profile for a given test site. The spacing is measured in units of mm.
19. ***“Mile_group”***: This is a numerical variable. This variable starts at 0.1 and increases in units of 0.1 miles. It represents a pavement section that is 0.1 miles long, as it would in the Pavement Management Systems used by the Receiving Agency.
20. ***“n_texture”***: This is a numerical variable. The number of rows or profiles that were summarized (by using of a median, mean, or standard deviation) for the 0.1 mile long pavement section in consideration.
21. ***“###_med”***: The “###” term stands for any of the texture indexes discussed in Section 4.2.2. All the variables ending in “_med” represent the median for the index in consideration over the 0.1 mile long pavement section.
22. ***“###_mean”***: The “###” term stands for any of the texture indexes discussed in Section 4.2.2. All the variables ending in “_med” represent the median for the index in consideration over the 0.1 mile long pavement section.
23. ***“###_sd”***: The “###” term stands for any of the texture indexes discussed In Section 4.2.2. All the variables ending in “_med” represent the median for the index in consideration over the 0.1 mile long pavement section.

4.2.4. Inventory of Pavement Texture 0-7139

The Microsoft Excel file “*Inventory of Pavement Texture 0-7139*” served as document to catalog and summarize the information contained within the “texture” portion of the database in P1. Researchers from the Performing Agency would record information in this file immediately after collecting the data. **Figure 84** shows the view within the “*Summary*” tab of the document.

Part A (**Figure 84**) presents a summary of the districts that were surveyed. The districts’ names that are bolded contributed a significant amount of data to this project. Conversely, if not bolded, the district contributed with some sections to the database. Lastly, if the district name is almost white that implies that no data was collected in that district.

Part B summarizes the mileage that the Performing Agency agreed to survey and the mileage that was truly surveyed by the end of Task #4.

Part C shows a collection of tabs, of which the first tab is the *summary* of all districts, and the subsequent tabs are named after each district where data was collected. **Figure 85** shows a sample view of the spreadsheet for the tabs in Part C.

	A	B	C	D	E
1	#	District	Pavement Types	Miles Tested	
2	1	Abilene			
3	2	Amarillo	Flexible and Rigid	6.40	
4	3	Atlanta	Flexible and Rigid	60.48	
5	4	Austin	Flexible and Rigid	51.37	
6	5	Beaumont			
7	6	Brownwood			
8	7	Bryan	Flexible and Rigid	4.11	
9	8	Childress			
10	9	Corpus Christi			
11	10	Dallas			
12	11	El Paso	Flexible and Rigid	79.09	
13	12	Fort Worth		6.48	
14	13	Houston	Flexible and Rigid	61.87	
15	14	Laredo	Flexible Only	40.18	
16	15	Lubbock	Flexible Only	49.57	
17	16	Lufkin			
18	17	Odessa	Flexible Only	9.91	
19	18	Paris			
20	19	Pharr			
21	20	San Angelo	Flexible Only	14.18	
22	21	San Antonio	Flexible and Rigid	7.88	
23	22	Tyler	Flexible Only	2.82	
24	23	Waco	Rigid Only	10.60	
25	24	Wichita Falls	Flexible and Rigid	12.85	
26	25	Yoakum	Flexible and Rigid	9.70	
27					
28					
29	Total Prome	400	Total Mileage Measured	427.48	
30			Total Mileage Remaining	(27)	
31					
32	Total Rigid	100	Rigid Mileage Measured	110.34	
33			Rigid Mileage Remaining	(10)	
34					
35	Total Flexible	300	Flexible Mileage Measured	317.15	
36			Flexible Mileage Remaining	(17)	
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					
101					
102					
103					
104					
105					
106					
107					
108					
109					
110					
111					
112					
113					
114					
115					
116					
117					
118					
119					
120					
121					
122					
123					
124					
125					
126					
127					
128					
129					
130					
131					
132					
133					
134					
135					
136					
137					
138					
139					
140					
141					
142					
143					
144					
145					
146					
147					
148					
149					
150					
151					
152					
153					
154					
155					
156					
157					
158					
159					
160					
161					
162					
163					
164					
165					
166					
167					
168					
169					
170					
171					
172					
173					
174					
175					
176					
177					
178					
179					
180					
181					
182					
183					
184					
185					
186					
187					
188					
189					
190					
191					
192					
193					
194					
195					
196					
197					
198					
199					
200					
201					
202					
203					
204					
205					
206					
207					
208					
209					
210					
211					
212					
213					
214					
215					
216					
217					
218					
219					
220					
221					
222					
223					
224					
225					
226					
227					
228					
229					
230					
231					
232					
233					
234					
235					
236					
237					
238					
239					
240					
241					
242					
243					
244					
245					
246					
247					
248					
249					
250					
251					
252					
253					
254					
255					
256					
257					
258					
259					
260					
261					
262					
263					
264					
265					
266					
267					
268					
269					
270					
271					
272					
273					
274					
275					
276					
277					
278					
279					
280					
281					
282					
283					
284					
285					
286					
287					
288					
289					
290					
291					
292					
293					
294					
295					
296					
297					
298					
299					
300					
301					
302					
303					
304					
305					
306					
307					
308					
309					
310					
311					
312					
313					
314					
315					
316					
317					
318					
319					
320					
321					
322					
323					
324					
325					
326					
327					
328					
329					
330					
331					
332					
333					
334					
335					
336					
337					
338					
339					
340					
341					
342					
343					
344					
345					
346					
347					
348					
349					
350					
351					
352					
353					
354					
355					

“Pavmt_Type,” and “Surface” are the same as those described in the data dictionaries for earlier files. “Location Description” provides a description for the location of the site. “Age” indicates the best guess of how many years have passed since the site was constructed or received its most current maintenance treatment. “Ground_truth” is equal to “1” if the surface information was provided by the Receiving Agency and “0” otherwise. “Coordinates” are the site’s GPS coordinates written using the decimal degree system. “Miles_Tested” indicates the mileage tested with the equipment, and the “Provider” indicates specifically which person provided the surface information for the site and on what date the information was provided. If the surface information is unknown, then the term “CTR” is used. Section B, in **Figure 85**, summarizes the mileage collected in the districts in terms of the pavement type and in terms of who provided the surface information.

4.3. Picture Database (Image Data Folder)

The 2D pictures taken with a high-speed camera were used in conjunction with the texture measurements to complement the classification of pavement surfaces. All the picture data has been stored in a folder entitled “2. Image Data (Camera Images)” inside the main directory of the hard drive. This folder is broken down into three subfolders: “1. All Pictures,” that contains all images collected during this study, “2. Revised Pictures Only” that contains only pictures who passed a quality control inspection, and “3. Deleted Pictures Only” that contains images deemed noisy, misleading, or useless for the purposes of training a surface detection model. A spreadsheet named “Inventory of Pavement Pictures 0-7139” that contains bookkeeping records for all the images collected across the eight district folders is attached to the pictures database. Each of the folders displayed in **Figure 86** will be discussed in more detail in the upcoming sections.

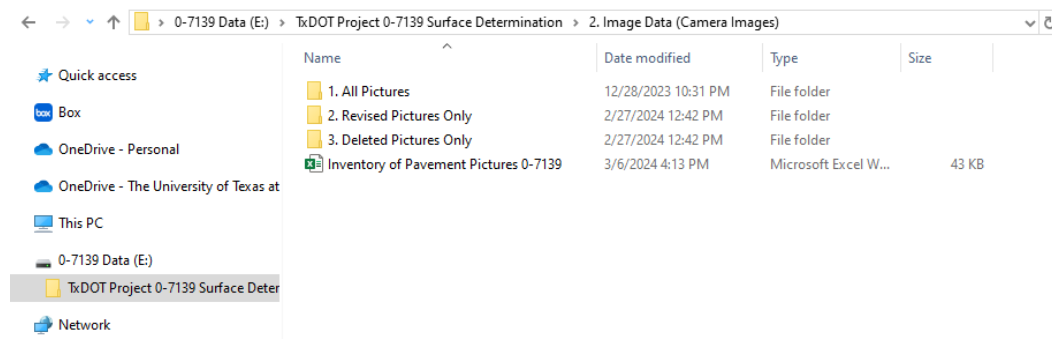


Figure 86: Screenshot of the “Image Data (Camera Images)” folder view, showing the that folders have been compressed by district.

4.3.1. All Pictures

The “1. All Pictures” folder is broken up into nine folders (all compressed due to significant file size). Eight of these folders correspond to the same districts which have been previously discussed in the sections for texture of this document. A ninth folder, named “*Bleeded Pavmts in Brownwood*,” was included as additional data, and it contains pictures of multiple seal coats collected from the Brownwood district. This dataset can be used in the future to detect the presence and degree of severity of bleeding in surface treatments. **Figure 87** displays the folder view for the “1. All Pictures” folder.

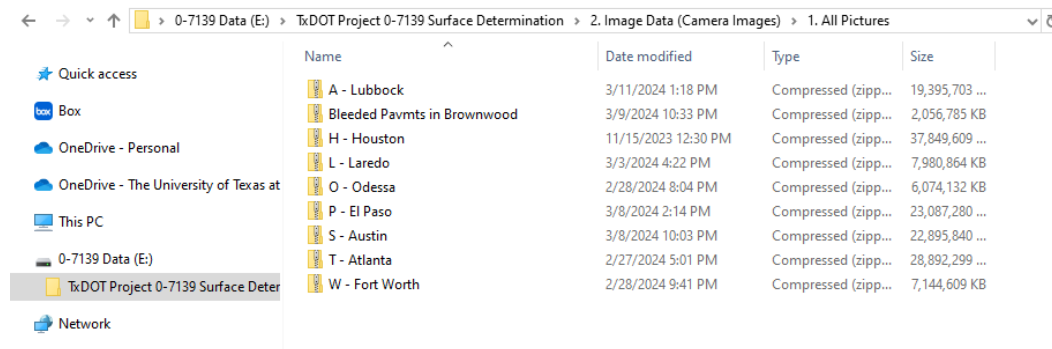


Figure 87: Screenshot of the “Image Data (Camera Images)” folder view, showing the that folders have been compressed by district.

The “1. All Pictures” folder contains images that have not been subjected to quality control. As a result, some pictures may include obstructions such as cables across the image, road debris, or suffer from issues with extreme contrast or brightness. Additionally, sections labeled as “Ground Truth” might contain significant anomalies like large pavement patches or bridges with different surface types in the middle of the test site. These inconsistencies could potentially mislead machine learning algorithms during training. Therefore, the Performing Agency advises using the images in this folder to develop an automated method for distinguishing between high and low-quality pictures, rather than for classifying pavement surface types.

4.3.2. Revised Pictures

The folder “2. Revised Pictures” contains only those pictures that passed a rigorous manual visual inspection conducted by the Performing Agency. There were two standards of quality control criteria applied to the image data: one for test sites that comprised the “Ground Truth” data subset, and another for the “Holdout” data subset.

The “Ground Truth” sites are those for which the surface information was known in advance and provided by the Receiving Agency. This data will be used for training and testing the machine learning models. Any picture that did not comply with the information provided by the Receiving Agency was moved to the “Deleted Pictures” folder. Additionally, if an image was of poor quality or had features that would obstruct or complicate surface classification (e.g., excessive cracking), it was also moved to the “3. Deleted Pictures” folder.

The “Holdout” data subset comprises pavement sites where the surface type is not known, thus it follows a different quality control criterion. In this subset, a picture was only deleted if it possessed features that would obstruct or complicate surface classification. All pictures passing this criterion will be used to assess the accuracy of the algorithm in detecting pavement changes. Data from the “Holdout” subset will not be used for training or testing the pavement detection models but will serve solely as a validation dataset. The “2. Revised Pictures” is subdivided into eight compressed folders as shown in **Figure 88**. The additional file for bled pavements was not part of this quality control process.

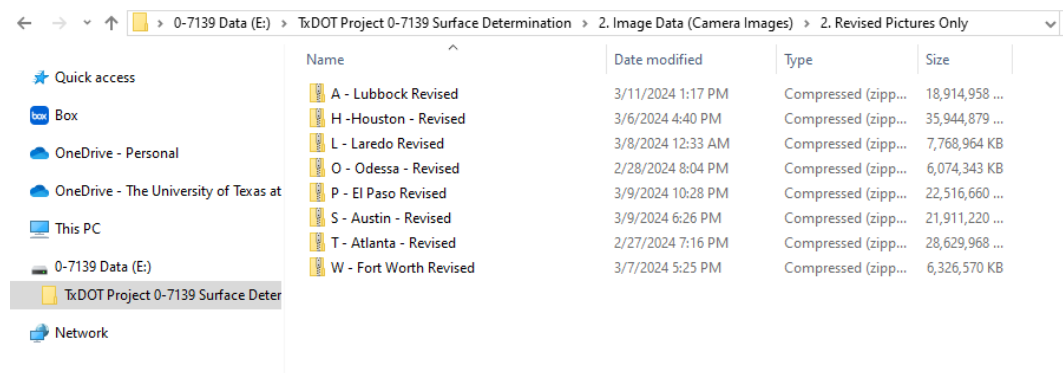


Figure 88: Screenshot of the “Image Data (Camera Images)” folder view, showing the that folders have been compressed by district.

4.3.3. Deleted Pictures

The “3. Deleted Pictures” folder contains all of the pictures that were removed during the quality control stage. **Figure 89** depicts some sample images with features that would be not representative of an ideal homogenous pavement surface, or whose inclusion in the database could result in detrimental performance. The file sizes of these compressed folders are significantly smaller compared to the ones of “2. Revised Pictures,” as shown in **Figure 90**. The “3. Deleted Pictures” folder is broken up into eight district subfolders. Notice that once a folder is extracted and opened, not all the test sites for that specific district will have a folder. If a test site is missing, that implies that the pictures taken from that site were all homogenous.

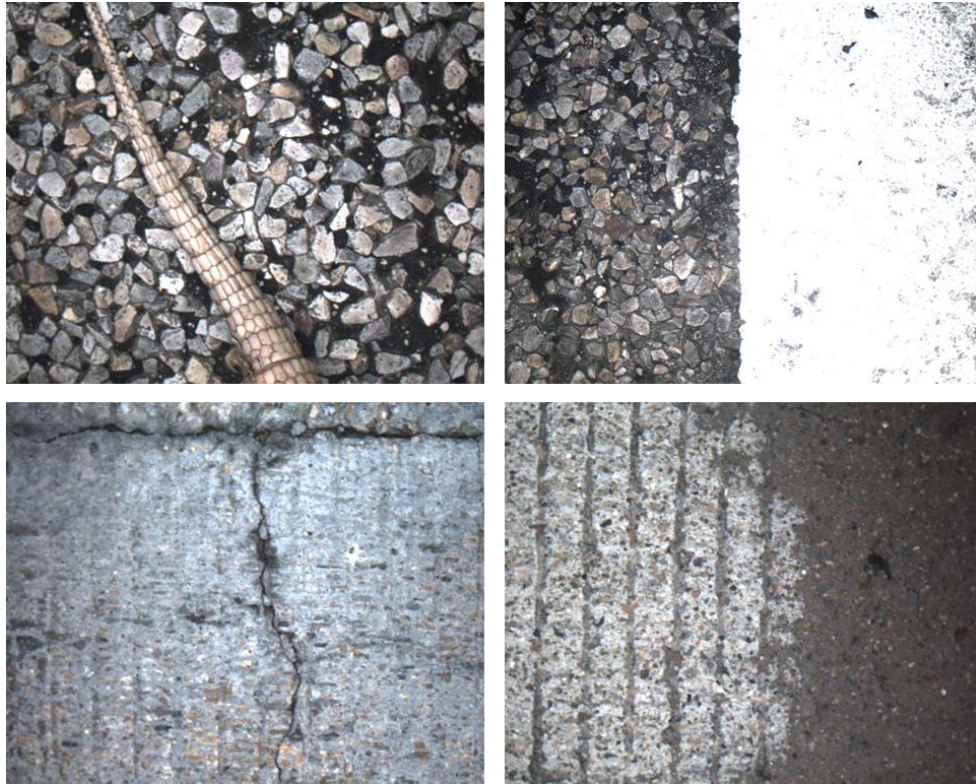


Figure 89: Sample images showing what kind of criteria would be considered enough to delete an image from the database: unexpected debris, objects or animals such as a critter's tail (top left), pavement markings (top right), cracking which can be confused as surface features (bottom left), interface of a pavement change (bottom right).

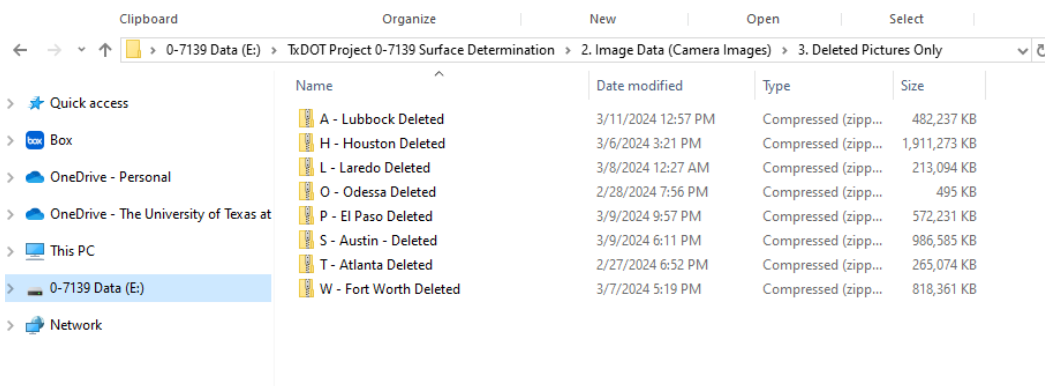


Figure 90: Screenshot of the “3. Deleted Pictures” folder view.

4.3.4. Inventory of Pavement Pictures 0-7139

The Microsoft Excel file “*Inventory of Pavement Pictures 0-7139*” served as a document to catalog and summarize the information contained within the “pictures” portion of the database in P1. Researchers from the Performing Agency would record information in this file immediately after collecting the data in the field.

Figure 91 shows the view within the “*Grand Summary*” tab of the document.

Ground Truth Data							
Folder Initial	District	N_TxDOT_sites	N_pictures_TxDOT	N_pictures_revised_TxDOT	N_pictures_removed_TxDOT	N_Homogenous_sites	N_Heterogenous_sites
T	Atlanta	10	60,072	53,029	1,043	10	0
S	Austin	20	66,275	61,853	4,422	13	7
P	El Paso	18	44,522	42,754	1,768	12	6
W	Fort Worth	4	23,960	20,395	2,965	0	4
H	Houston	20	176,657	170,262	6,395	16	4
L	Laredo	13	72,161	71,101	1,060	10	3
A	Lubbock	14	55,051	54,203	842	12	2
O	Odessa	0	-	-	-	-	0

Holdout Data							
Folder Initial	District	N_CTR_sites	N_pictures_CTR	N_pictures_revised_CTR	N_pictures_removed_CTR	N_Homogenous_sites	N_Heterogenous_sites
T	Atlanta	16	81,232	81,083	143	8	8
S	Austin	11	37,725	37,666	59	8	3
P	El Paso	12	53,123	52,273	850	5	7
W	Fort Worth	1	2,502	2,502	0	1	0
H	Houston	9	52,882	51,000	1882	5	4
L	Laredo	11	13,581	13,513	68	9	2
A	Lubbock	14	42,530	41,147	1443	10	4
O	Odessa	7	45,176	45,172	4	6	1

Summary			
Folder Initial	District	N_sites	Total_Number_Pictures
T	Atlanta	26	141,304
S	Austin	31	104,000
P	El Paso	30	97,645
W	Fort Worth	5	26,462
H	Houston	29	229,539
L	Laredo	24	85,742
A	Lubbock	28	97,641
O	Odessa	7	45,176
Total		180	827,509

Grand Summary	TxDOT (Ground Truth)	CTR (Extra)	+
---------------	----------------------	-------------	---

Figure 91: Screenshot of the “Grand Summary” tab within the “Inventory of Pavement Pictures 0-7139” file.

Part A (**Figure 91**) summarizes the picture information for all the sites which composed the Ground Truth dataset.

Part B does the same for those sites within the *Holdout* dataset.

Part C summarizes the relevant information across both datasets. The columns, which can be found within the “Grand Summary” tab, are:

- “Folder Initial” covers the initials as discussed in Table 10.
- “District” is the name of the district where most data was collected.
- “N_dataset_sites” is the number of sites within a specific dataset. “N_pictures_dataset” is the total number of pictures taken within that district. “N_pictures_revised_dataset” stands for the number of pictures that passed the quality control within a given district.
- “N_pictures_deleted_dataset” is the number of pictures removed from a district because they did not pass the quality control.

- “N_Homogenous_sites” is the number of sites for which the surface type was homogenous across all pictures.
- “N_Heterogenous_sites” is the number of sites for which there is at least one pavement change or different surface texturing across all the pictures.

Lastly, Part D shows the tabs that can be found in this file. The other two tabs hold more specific information about the pictures in each site based on which dataset they belong to. **Figure 92** depicts a sample view of the spreadsheet for the other tabs in Part D.

	A	B	C	D	E	F	G	H	I
1	Folder	District	County	Site_Nu	Site_Name	Highway	Directio	Approx_	Pavmt_Type
2	T	Atlanta	Upshur	T1	Gilmer	FM0049	EB	32.7241	Flexible
3	T	Atlanta	Cass	T12	Junk	SH0008	SB	33.1742	Flexible
4	T	Atlanta	Cass	T16	Danger	US0059	SW	33.0943	Flexible
5	T	Atlanta	Marrion	T22	Scott	US0059	SB	32.7908	Flexible
6	T	Atlanta	Harrison	T23	Marshall	IH0020	WB	32.4799	Flexible
7	T	Atlanta	Harrison	T24	Pine	US0080	NW	32.5034	Flexible
8	T	Tyler	Gregg	T26	Longview	IH0020	WB	32.4549	Flexible
9	T	Atlanta	Upshur	T5	Ore	US0259	NB	32.7610	Flexible
10	T	Atlanta	Morris	T6	Lonestar	US0259	NB	32.9703	Flexible
11	T	Atlanta	Titus	T7	Family	SH0049	EB	33.1445	Flexible
12									
13	S	Austin	Hays	S10	Comal	FM0032	WB	29.9421	Flexible

Figure 92: Screenshot of the “TxDOT (Ground Truth)” tab within the “Inventory of Pavement Pictures 0-7139” file. This is also a representative view for the “CTR (Extra)” tab. Note that many columns are not shown in this image.

Both the “TxDOT (Ground Truth)” and the “CTR (Extra)” hold the same information for different sites. In each of these tabs, the following information can be found:

- The columns, “Folder,” “District,” “County,” “Site_Number,” “Site_Name,” “Highway,” “Direction_of_Travel,” “Approx_GPS_location,” “Type_of_Test,” and “Run,” have the same meanings as the ones given in the data dictionary for the “Final Texture Database.”
- “TXDOT Surface” is the surface information as provided by the Receiving Agency.
- “N_Pictures” is the total number of pictures taken at a given site.
- “N_kept” is the number of pictures that passed the quality control.

- “N_deleted” is the number of pictures removed for not passing the quality control.
- “Homogeneity” has a value of “1” if the site was homogenous across all pictures and 0 otherwise.
- “Approx_n_pvmt_changes” is an educated guess based on the pictures for the total number of pavement changes seen across the pictures. A pavement change constitutes going from one pavement type to another (flexible to rigid), from one type of flexible pavement to another (HMA to surface treatment), or from one PCC surface finish to another (NGDG to tining).
- “Excessive_Bleeded sections” counts the number of pictures with too much bleeding to determine surface type.
- “Combined_surfacing_sections” counts the number of pictures where two types of PCC texturing techniques are in equal proportions, such as longitudinal and transverse tining creating a sort of tining mesh.
- “Notes” are comments on why each picture was removed, or comments of the pictures which remained in the database but have interesting characteristics such as a thin layer of dirt on the pavement’s surface.

Chapter 5. Model Development, Training, and Testing

5.1. Data Processing

This section outlines the Performing Agency's approach to processing the data. It was divided into two subsections, each dedicated to a different type of data collected: 1) 2D surface profiles, and 2) 2D RGB area images of the pavement surface.

5.1.1. Implemented Algorithm

The Performing Agency developed **Figure 93** to illustrate the specific steps used to process the texture data, following the structure highlighted in **Figure 9**. The Performing Agency calibrated their processing algorithm by assessing at least 2,000 profiles to estimate five thresholds for identifying mild and extreme spikes in the data. Next, during the preprocessing stage, the algorithm converted the units of both the profile measurements and the thresholds to millimeters, trimmed all measurements not within the middle 150 mm, and pre-imputed any missing data using the median height. For the denoising stage, the Performing Agency implemented SOFA⁴ (3). Linear interpolation imputation was then performed on profiles where the total number of removed data did not exceed 20%. Profiles where removed noise exceeded 20% were discarded. Once imputed, a linear regression detrending subroutine was implemented to center the profile at an elevation of 0 mm.

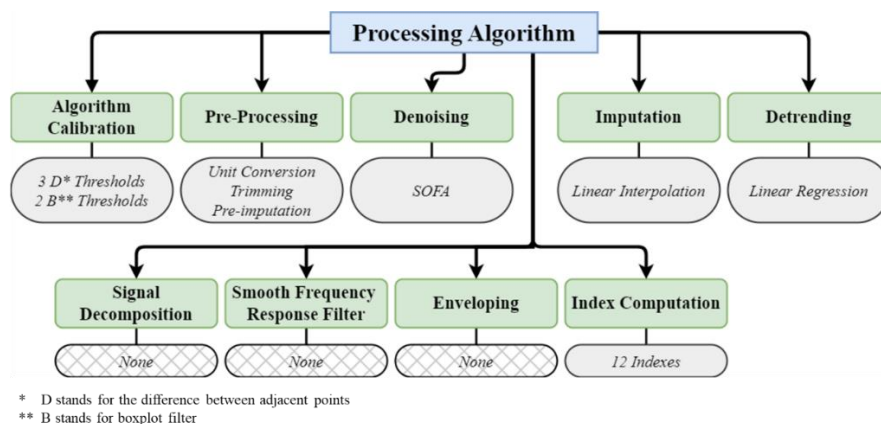


Figure 93: Schematic illustrating a detailed description of the data processing algorithm.

⁴ SOFA is an algorithm developed by The University of Texas at Austin which uses boxplot filters and the difference between adjacent points alongside the previously determined thresholds to detect and remove spikes or flat signals from the profile.

The Performing Agency did not apply any signal decomposition, smooth frequency response filters, or enveloping methods to the profiles, as these methods require significant computational resources when processing large quantities of texture data, particularly at a network level. After full processing, a large representative sample of profiles from each tested site was visually inspected to ensure the algorithm performed as expected. **Figure 94** depicts a comparison between a raw profile and a fully processed profile. Fully processed profiles were then used to compute a set of twelve texture summary indices.

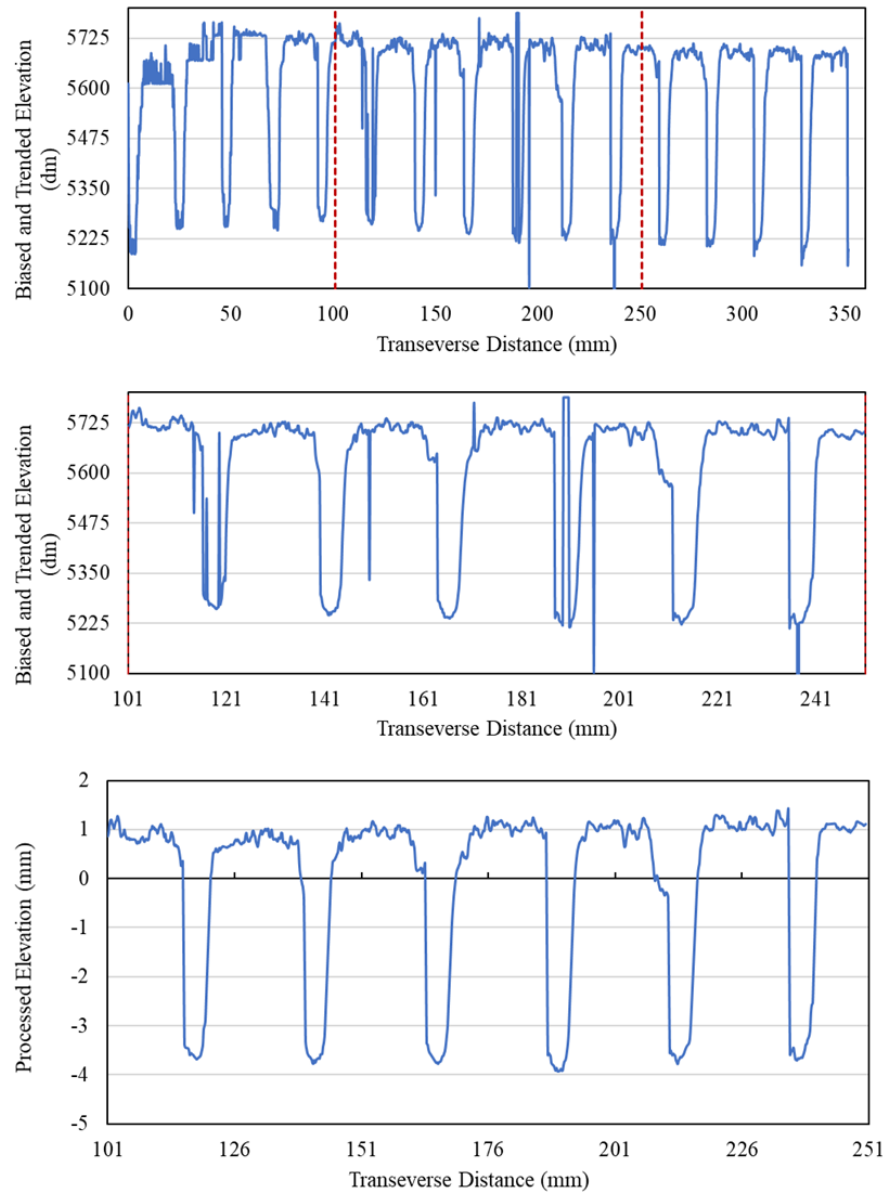


Figure 94: Example of the SOFA on 2D texture profile for a pavement with new generation diamond grinded surface finish.

5.1.2. Computed Indices

The Performing Agency sourced material from Gadelmawla et al. (55), ISO 21920-2 (57), and Sabillon et al. (3). Putting together these sources yielded a list of 80 spatial texture indices. However, due to the complexity and lengthy computation time of some indices, particularly spectral ones, the list was refined by excluding those that were impractical for efficient processing.

Spatial texture indices were organized into five categories:

- *Amplitude indices*: to characterize vertical deviations.
- *Spacing indices*: to quantify horizontal characteristics.
- *Hybrid indices*: to describe characteristics regarding the profile's slope.
- *Functional indices*: derived more complex textural feature by using auxiliary functions like the material ratio curve.
- *Feature indices*: typically include amplitude, spacing, hybrid, and functional indices, calculated over specific segments—such as 100 mm—rather than across the entire profile.

Significant linear correlations identified among these indices suggested that simpler, well-understood indices could approximate more complex ones effectively. Therefore, indices exhibiting over 60% correlation were excluded, except for essential ones widely recognized in pavement engineering, such as MPD and RMS of height. This correlation-based selection process reduced the list from 80 to 19 indices. Further reduction prioritized indices requiring minimal computation time, refining the selection to those most feasible for network-level analysis. This process resulted in 12 essential indices, ensuring their practicality and familiarity to industry professionals. These indices are detailed in Table 12.

Table 12: List of indices used in the study

Amplitude (Height) Index	
Index	Equation
Root Mean Square Height	$RMS = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (h_i - \bar{h})^2}$
Kurtosis	$R_k = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \frac{(h_i - \bar{h})^4}{RMS^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$
Solidity Factor	$R_s = \frac{\min(h_i)}{\max(h_i)}$

Spacing Index	
Index	Equation
Mean of Cross Width	$C_m = \frac{1}{n} \sum_{i=1}^n x_i$
Standard Deviation of Cross Width	$C_{std} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$
Skewness of Cross Width	$C_s = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \frac{(x_i - \bar{x})^3}{C_{std}^3}$
Hybrid Indices	
Index	Equation
Maximum Absolute Gradient	$R_{dt} = \max \left(\left \frac{d h(x)}{dx} \right \right)$
Two-Point Slope Variance	$SV_2 = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{h_i - h_{i+1}}{\Delta x} \right)^2}$
Functional Indices	
Index	Equation
Autocorrelation Function	$f_{ACF} = \frac{\frac{1}{l_e - t_x } \int_0^{l_e - t_x } (h(x) - \bar{h})(h(x + t_x) - \bar{h}) dx}{\frac{1}{l_e} \int_0^{l_e} (h(x) - \bar{h})^2 dx}$
Autocorrelation Length	$R_{al} = \min(t_x)$
Local Curvature	$\kappa(x) = \frac{d^2 h(x)/dx^2}{(1 + (d h(x)/dx)^2)^{3/2}}$
Mean of Local Curvature	$Hcme = \frac{1}{n} \sum_{i=1}^n \kappa(x)_i$
Standard Deviation of Local Curvature	$Hcst = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\kappa(x)_i - \overline{\kappa(x)})^2}$
Feature Indices	
Index	Equation
Mean Profile Depth	$MPD = 0.5 [\max(h_1, \dots, h_{m/2}) + \max(h_{m/2}, \dots, h_m)]$

Where h_i and $h(x)$ is the elevation at point i or x , respectively, \bar{h} is the mean elevation, n is the number of datapoints, $h_{m/2}$ is the elevation value midway through segment, h_m is the elevation value at the end of segment, x_i is the

horizontal spacing between inflection points along the profile, \bar{x} is the mean spacing between inflection points, t_x is a spatial shift in the horizontal direction, l_e is the evaluation length, and l_0 is the overlap interval, $d h(x)/dx$ and $d^2 h(x)/dx^2$ are the first and second derivatives, respectively, of the scale-limited profile h with respect to the position x .

5.2. Picture Processing

5.2.1. Quality Control

The high-speed pictures underwent a rigorous quality control process before any manipulation was conducted, adhering to two distinct standards tailored for each data subset. The “Ground Truth” data subset (DS1) comprises all test sites with known surface information provided by the Receiving Agency. Pictures in DS1 were removed if they met any of the following criteria:

- 1) The surface type in the images did not match the information provided by the Receiving Agency.
- 2) The images exhibited excessively high or low contrast or brightness, impeding accurate pavement type determination.
- 3) The images displayed significant pavement distress, such as cracking or bleeding, which could affect classification accuracy.
- 4) The images contained foreign objects, like debris, cables, or critters, unrepresentative of a pavement surface.

Pictures from DS1 that passed the quality control process were subsequently utilized for training and validation of the image classification algorithm. In contrast, the ‘Holdout’ data subset (DS2) consists of pavement sites with unidentified surface types. Thus, these pictures adhere to a slightly different quality control standard. Images from DS2 were excluded based on criteria #2, #3, or #4 of the DS1 standards. Criterion #1 cannot be implemented as the “ground truth” is not known. Successfully vetted images from DS2 were designated as part of the testing dataset to heuristically evaluate the accuracy of the model developed using DS1.

5.2.1.1. Data Augmentation

The practice of manipulating data prior to model training is generally referred to as “pre-processing,” “data processing,” or “feature engineering,” in various contexts. However, within the context of image classification, this step is specifically termed “data augmentation.” Data augmentation ensures that the model receives input

conducive to learning genuine patterns rather than merely memorizing noise. Data augmentation derives its name from its capacity to artificially augment both the volume and variety of the training dataset, proving particularly valuable in scenarios where data accessibility is limited.

The initial phase of data augmentation involved resizing the images to a uniform resolution of 224 x 224 pixels, a significant reduction from the original 1280 x 1024 pixels. This adjustment serves two purposes: it diminishes the computational resources required and simplifies the extraction of effective contextual information crucial for classification, which becomes more challenging with larger images. Subsequently, the Performing Agency introduced variations in lighting conditions by adding noise through over- or under-saturation of the images. When executed correctly, this strategy enhances model robustness by simulating a wider array of environmental factors, akin to those encountered in field conditions where lighting varies drastically when traveling at 50 mph.

Further data augmentation techniques, such as rotations, cropping, shearing, and zoom adjustments, contribute significantly to the model's ability to generalize. Rotations simulate scenarios where the camera is physically rotated, cropping mimics situations where the image resolution is smaller than expected, shearing represents the perspective shift seen when objects are viewed at an angle, and zoom adjustments replicate varying distances between the camera and the pavement surface. Collectively, these techniques aim to make the classification algorithm invariant to translation and rotation, ensuring minimal impact on classification accuracy regardless of camera movement or orientation.

5.3. Machine Learning Implementation

The Performing Agency implemented two unsupervised learning techniques: dimensionality reduction and clustering. Employing a higher number of variables (i.e., higher dimensionality) generally enhances the performance of machine learning algorithms due to the richer data environment provided. However, when the dimensionality exceeds three variables, visualizing data effectively becomes a challenge. To overcome this issue, the Performing Agency utilized a dimensionality reduction technique known as Principal Component Analysis (PCA). This method was instrumental in revealing potential clustering patterns within the complex data. Building on this, a series of clustering techniques was subsequently applied to objectively identify and analyze the clusters that were visualized through PCA.

Then, the Performing Agency utilized traditional classification supervised learning algorithms on texture indices derived from 2D profiles and trained nuanced

supervised learning image recognition models to predict pavement surface types from 2D RGB images.

5.3.1. Index-Based Classification (IBC) Models

The Performing Agency trained a comprehensive suite of 14 traditional supervised learning classification models which utilize texture indices as features to predict pavement surface type. The models have been summarized in Table 13 and encompass a diverse range of methodologies, each tailored to specific aspects of the classification challenge:

Table 13: List of IBC models

Model name (Abbreviation)	Model name (Abbreviation)
K-Nearest Neighbor (KNN)	Bootstrap Aggregating (BA)
Gaussian Naïve Bayes (GNB)	Adaptive Boosting (AB)
Logistic Regression (LOG)	Gradient Boosting (GB)
Probit Regression (PRO)	Extreme Gradient Boosting (XGB)
Complementary Log-log Regression (CLL)	Support Vector Machine (SVM)
Decision Tree (DT).	Shallow Neural Network (SNN)
Random Forest (RF)	Deep Neural Network (DNN)

These IBC models performed classification tasks in hierarchical manner, according to the hierarchy shown in **Figure 83**. The hierarchy of pavement specificity begins with the broad differentiation between flexible and rigid pavements. Subsequent levels refine this categorization: the second level splits flexible pavements into hot and warm mix asphalt and various surface treatments, while rigid pavements are categorized by the timing of surface texturing—either “Plastic” for fresh concrete texturing or “Hardened” for post-set texturing. Further distinctions are made at the third specificity level based on the types of asphalt mixes and seal coats, and the numerous texturing technique for PCC texturing. The highest levels of specificity detail the gradations within seal coats and asphalt mixes and differentiates rigid pavement tining into longitudinal or transverse patterns, with an additional distinction in the latter for the spacing of tining lines. This structured approach ensures detailed and nuanced classification, crucial for accurate pavement management and maintenance strategies.

5.3.2. SL Pictures-Based Classification

The Performing Agency utilized TensorFlow, an open-source machine learning framework developed by Google to leverage the capabilities of CNNs for classifying images of pavement surfaces. This framework is specifically designed to facilitate the building, training, and deployment of machine learning models, including those for deep learning. The CNN model implemented for image classification uses *transfer learning*, a technique where a model developed for a specific task is repurposed as the starting point for a model on a second task. This approach leverages pre-trained models, which have already learned a substantial number of relevant features from large and diverse datasets, to improve learning efficiency and performance on related but distinct tasks.

In this particular application, the model of choice was ResNet50, a 50-layer CNN developed by Microsoft Research Asia (205), pre-trained on a dataset comprising over a million images depicting approximately 1,000 different objects. This pre-trained model, known for its deep architecture and robustness in feature detection across various visual domains, was adapted for the specific task of pavement surface classification.

5.4. Results

This section is broken down into four subsections: cluster analysis, a comparison of index-based classification (IBC) model characteristics, IBC model results, and picture-based classification (PBC) model results

5.4.1. Cluster Analysis

Cluster analysis was performed to determine whether the data naturally forms groups that correlate with the surface types specified by their mix design. If the UL analysis produces clusters that logically correspond to pavement surface types, it would suggest that the classification algorithms are capable of identifying different surfaces using texture indices alone. However, no similar inference can be drawn for image data, as no UL analysis has been conducted in that domain yet. Given the variability in conditions and ages of the tested pavement surfaces, it is probable that these clusters will reflect differences in surface conditions more than surface types. The Performing Agency carried out this analysis focusing on both flexible and rigid pavements within the Ground Truth (DS1) dataset.

5.4.1.1. Flexible Pavements

The first step of cluster analysis is to find uncorrelated variables. Thus, a pairwise correlation matrix was computed to isolate those pavement texture indices whose

Pearson correlation coefficient was lower than 60%. **Figure 95** depicts the five indices used for the flexible pavements subset of DS1: height kurtosis (RKU), two-point slope variance (SV2), autocorrelation length (RAL), mean of local curvature (HCME), and the standard deviation of local curvature (HCST).

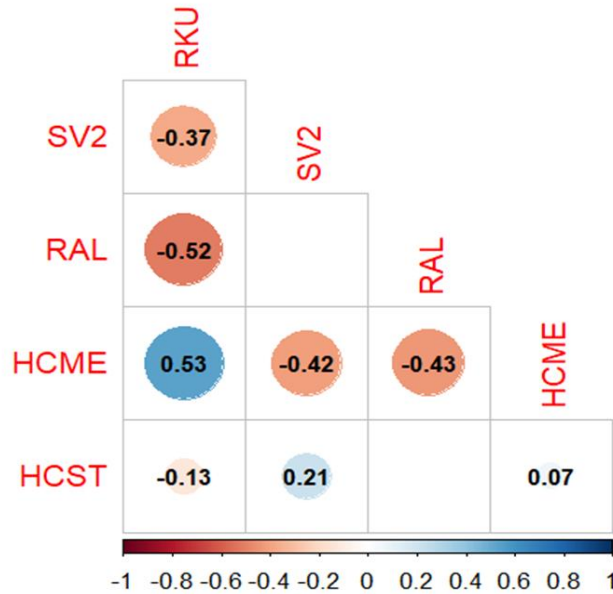


Figure 95: Pearson correlation coefficient matrix for the final selection of five indices used in flexible pavement clustering.

Next, the Performing Agency plotted a density plot of the first two principal components, as shown in **Figure 96**.

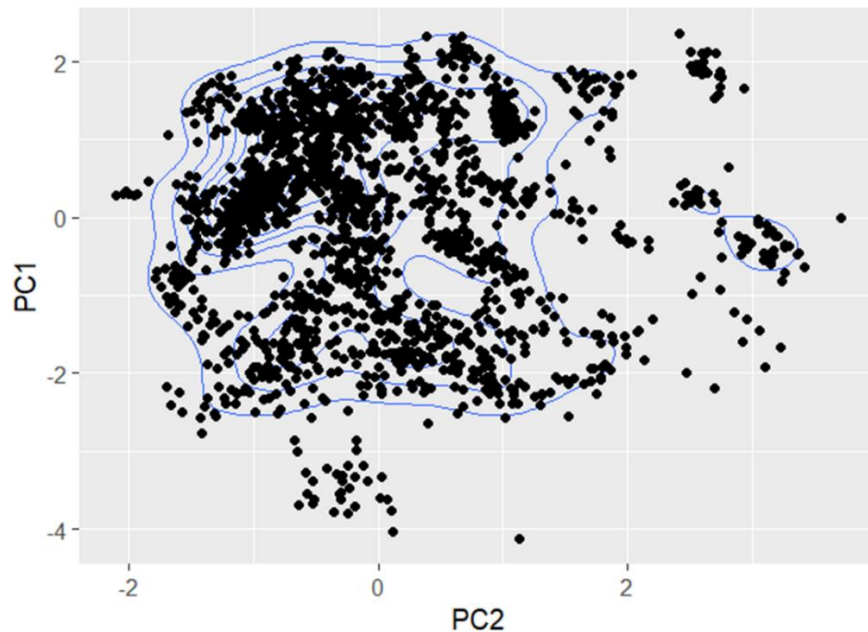


Figure 96: Density plot of the first two principal components for flexible pavements.

Although it is challenging to discern a clear clustering structure through visual inspection alone, a dendrogram (**Figure 97**) was utilized to enhance the visualization of the clustering arrangement. A dendrogram is a tree-like diagram that displays the relationships formed by hierarchical clustering. Each leaf on the dendrogram represents an individual data point, while branches merge at different heights, indicating the level of similarity between clusters. The height at which branches converge serves as a crucial indicator for estimating the optimal number of clusters in the data.

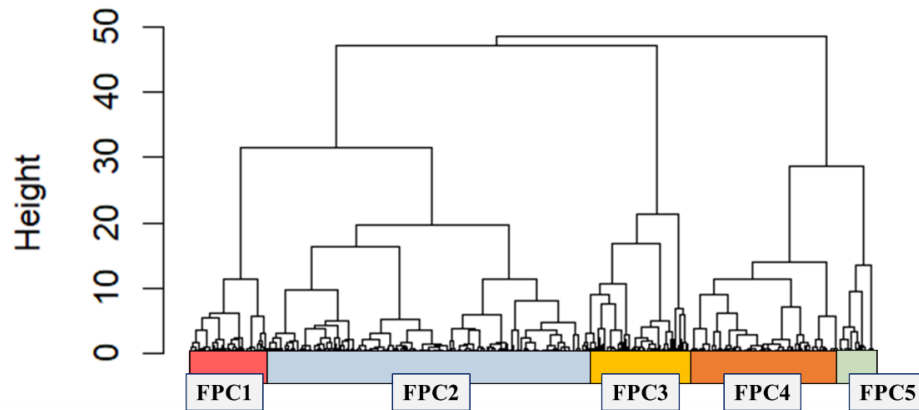


Figure 97: Dendrogram for flexible pavements with colored rectangles indicating the clustering structure for five groups in the data.

The dendrogram depicted in **Figure 97** has been enhanced by the addition of colored rectangles along the bottom of the x-axis, which represent the size of the clusters, assuming the data contains five clusters⁵. This assumption is based on the height at which branches join, serving as the cutoff for cluster membership. The selection of exactly five clusters is informed by the distinct textural properties of pavements within these groups. Moreover, these clusters demonstrate a significant degree of separation when analyzed in the principal component domain. To further illustrate this, the density plot in **Figure 98** has been adjusted to display how the clusters, identified in the dendrogram of **Figure 97** overlap on the scatterplot for the first two principal components.

The clusters, as depicted in the dendrogram, were identified using the AGNES model. This method's successful application in delineating these clusters necessitated the scaling of the indices from **Figure 95** using robust scaling, as

⁵ The clusters in **Figure 98** are labelled FPC#, where FPC stands for flexible pavement cluster followed by an index used to distinguish the clusters.

outlined by **equation (15)**. This precise scaling process was essential to achieve the cluster formations observed in the dendrogram.

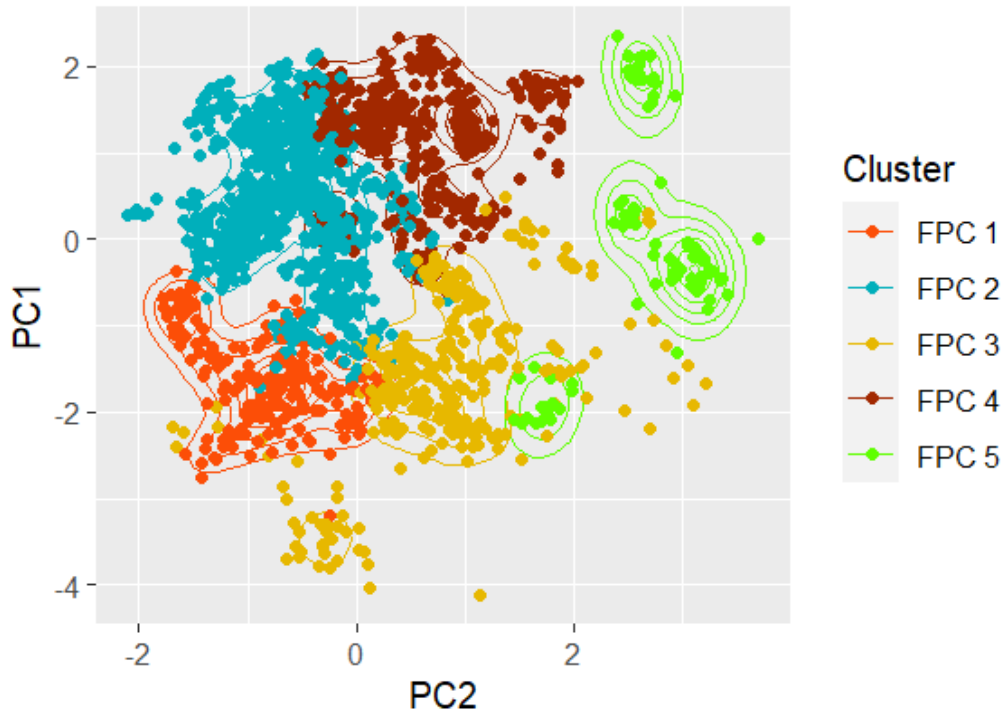


Figure 98: Visualization of the five clusters derived from a dendrogram on the scatterplot of the first two principal components.

$$X_R = \frac{x_i - \tilde{x}}{IQR} \quad (15)$$

Where, X_R is a robust scaled index, x_i is the original value of index i , \tilde{x} is the median value of the index, and IQR is the interquartile range of the index or the difference between the third and first quartiles of the index.

Once the indices are scaled, they are clustered using the *Ward* linkage method, a method of measuring the distance between clusters that minimizes within-cluster variance. This method systematically identifies the pair of clusters that, when merged, results in the smallest increase in total within-cluster variance, thereby ensuring that the most similar groups are combined to form more cohesive cluster structures.

It should be noted that none of the other three methods tested were able to replicate these five clusters precisely, due to differences in cluster assignment techniques. Consequently, only the AGNES model was utilized for clustering flexible pavements, owing to the effective classification it achieves. Detailed descriptions of the pavements comprising each cluster, along with representative median values

for six texture indices of the pavements within those clusters, are presented in Table 14. The number of pavement section in each cluster and their corresponding percentage out of the total, shown in parenthesis, is shown the right side of the table.

Table 14: Summary of flexible pavement cluster analysis with description of the cluster and median index measurements of the member pavements

FPC1: Highest macrotexture pavements					n: 198 (11%)
MPD	RKU	SV2	RAL	p HCME	HCST
2.49	-0.36	1.08	1.05	-5.65	2.86
FPC2: Course well-graded surfaces resembling a dense mix					n: 816 (46%)
MPD	RKU	SV2	RAL	p HCME	HCST
1.27	1.37	0.68	0.89	-4.97	3.12
FPC3: High macrotexture and uniform gradation					n: 279 (16%)
MPD	RKU	SV2	RAL	p HCME	HCST
1.97	1.97	0.67	1.48	-5.40	3.11
FPC4: Lowest macrotexture pavements					n: 391 (22%)
MPD	RKU	SV2	RAL	p HCME	HCST
0.92	0.92	0.54	0.96	-5.00	2.23
FPC5: Coarse surfaces with uniform gradation					n: 92 (5%)
MPD	RKU	SV2	RAL	p HCME	HCST
1.42	0.82	0.57	1.46	-6.93	1.33

Table 14 indicates that all pavements in FPC1 exhibit the highest macrotexture, as evidenced by their elevated median MPD and SV2 values. FPC2, the largest of the five clusters, primarily consists of two pavement types: (1) dense and gap-graded mixes with higher than usual macrotexture, and (2) open-graded mixes and seal coats showing signs of deterioration and polishing, making them resemble coarse dense-graded mixes. FPC3 predominantly includes seal coats of Gradation 3 and similar pavement surfaces, characterized by high MPDs and significant variation in profile heights, as denoted by RKU values and the standard deviation of local curvature. Additionally, the substantial RAL suggests that these surfaces have uniform gradations, with profile patterns recurring every 1.48 mm. FPC4 comprises mixes, primarily dense and gap-graded, which display minimal macrotexture, as indicated by low median MPD and SV2. FPC5 mainly consists of seal coats and PFCs undergoing polishing. These surfaces maintain a macrotexture slightly above average, as reflected by their MPDs and SV2, yet they do not exhibit the

smoothness typical of a polished dense-graded mix. Representative pavement types for each cluster are depicted in **Figure 99**.

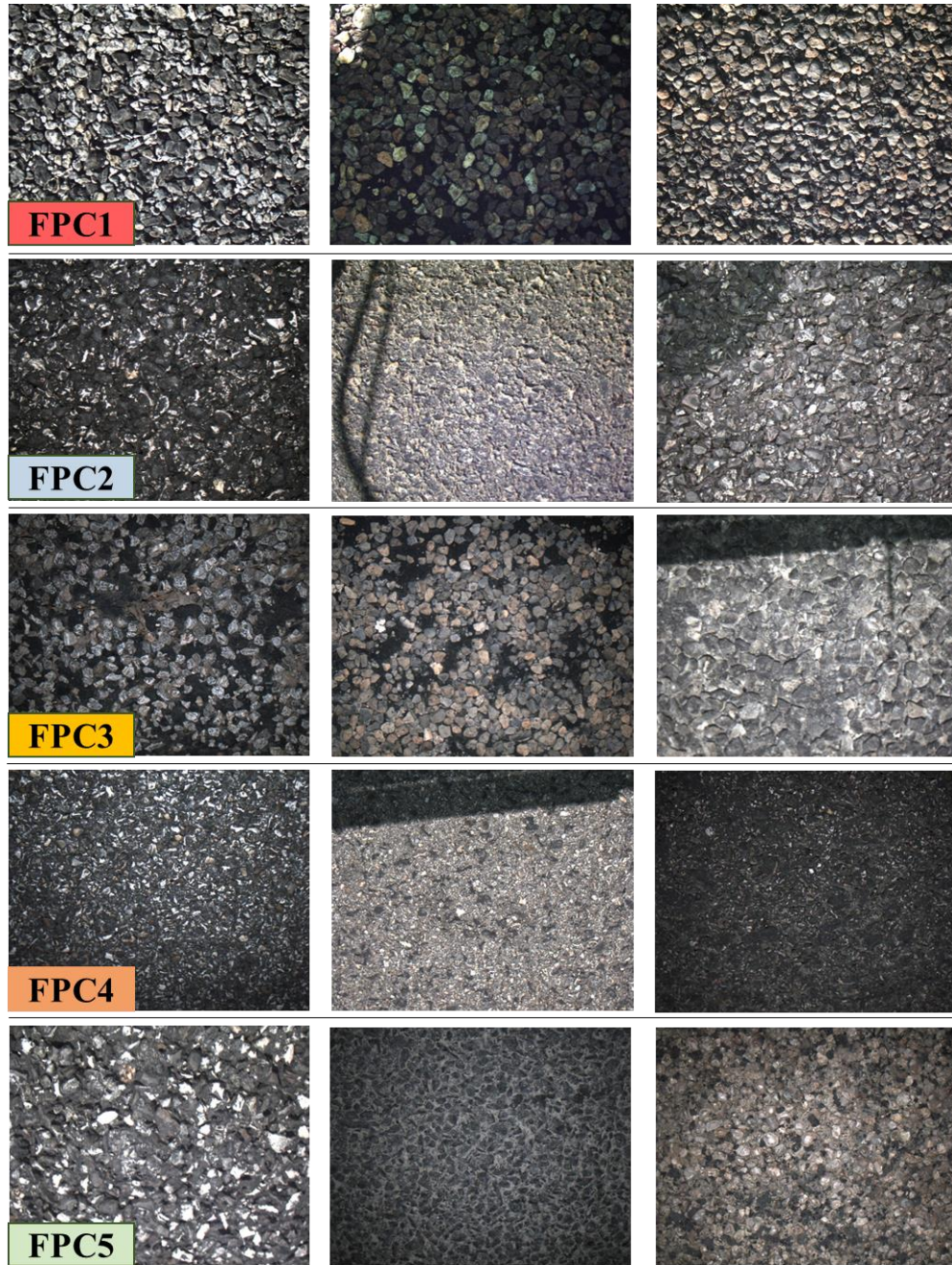


Figure 99: Representative pictures of each of the five cluster. Members of the same clusters are shown along the row.

5.4.1.2. Rigid Pavements

The texture indices calculated for flexible and rigid pavements exhibit different correlational patterns owing to the distinct textural properties inherent in each

pavement type. Consequently, it is essential to conduct separate correlation analyses before advancing to the cluster analysis. **Figure 100** presents the five indices utilized for the rigid pavement subset of DS1: Mean Profile Depth (MPD), Kurtosis of Height (RKU), Standard Deviation of Cross Width (CWST), Autocorrelation Length (RAL), and Mean of Local Curvature (HCME).

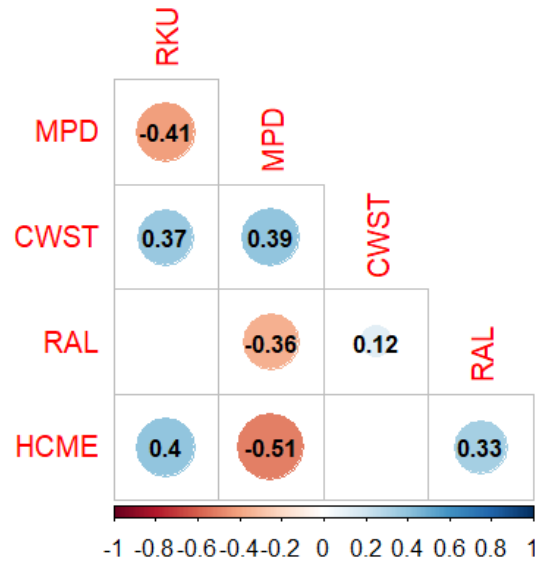


Figure 100: Pearson correlation coefficient matrix for the final selection of five indices used in rigid pavement clustering.

Subsequent analysis by the Performing Agency revealed two distinct and easy separable clusters in the density plot of the first two principal components (**Figure 101**), a contrast to the findings with flexible pavements. These clusters were identified as new generation (RPC1) and conventional (RPC2) diamond grinding pavement surfaces, showcased in **Figure 102**.

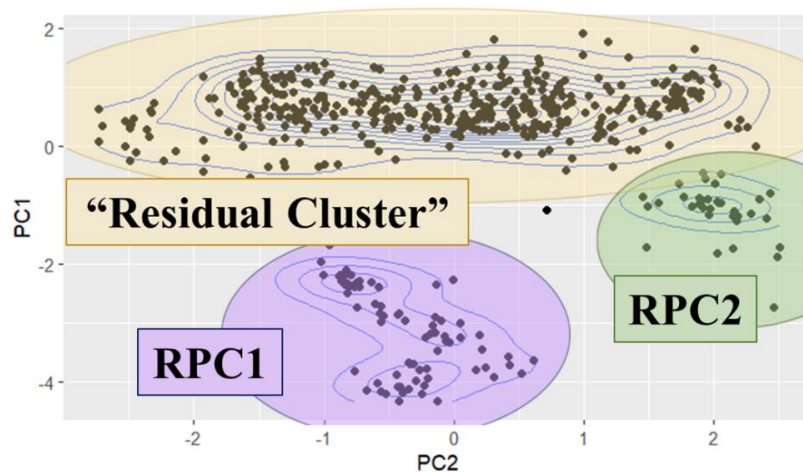


Figure 101: Density plot of the first two principal components for flexible pavements highlighting the two clusters that are easily separable by visual inspection⁶.

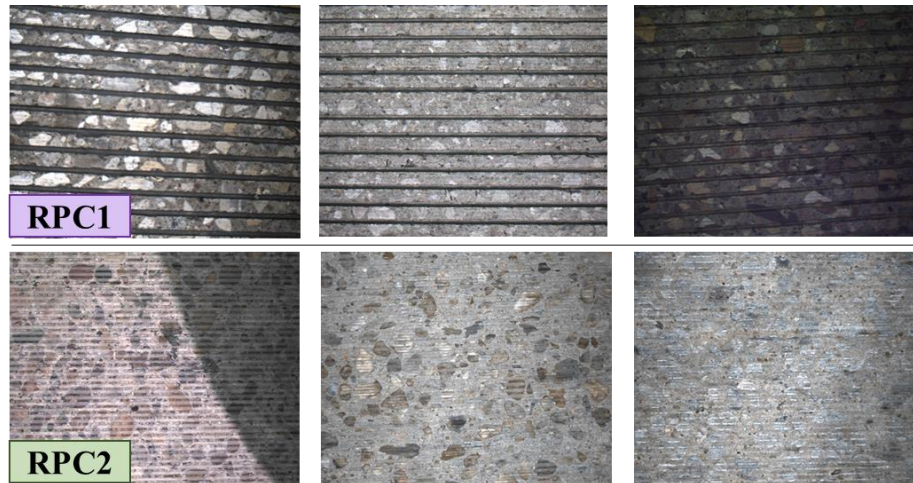


Figure 102: Representative pictures of each of the first two clusters. Members of the same clusters are shown along the row.

To confirm the detectability of these clusters, the Performing Agency applied all four clustering algorithms discussed in Section 2, all of which successfully identified the two clusters. AGNES and DBSCAN recognized three clusters, while PAM and GMM initially divided the larger, topmost cluster (labeled as the residual cluster in **Figure 101**) into two before identifying RPC2. This division is attributed to their limitation in detecting clusters of equal sizes. **Figure 103** illustrates the clustering outcomes for each algorithm, with a summary of the specific algorithm parameters used to identify clusters RPC1 and RPC2 detailed in Table 15.

⁶ The clusters in **Figure 101** are labelled RPC#, where RPC stands for rigid pavement cluster followed by an index used to distinguish the clusters.

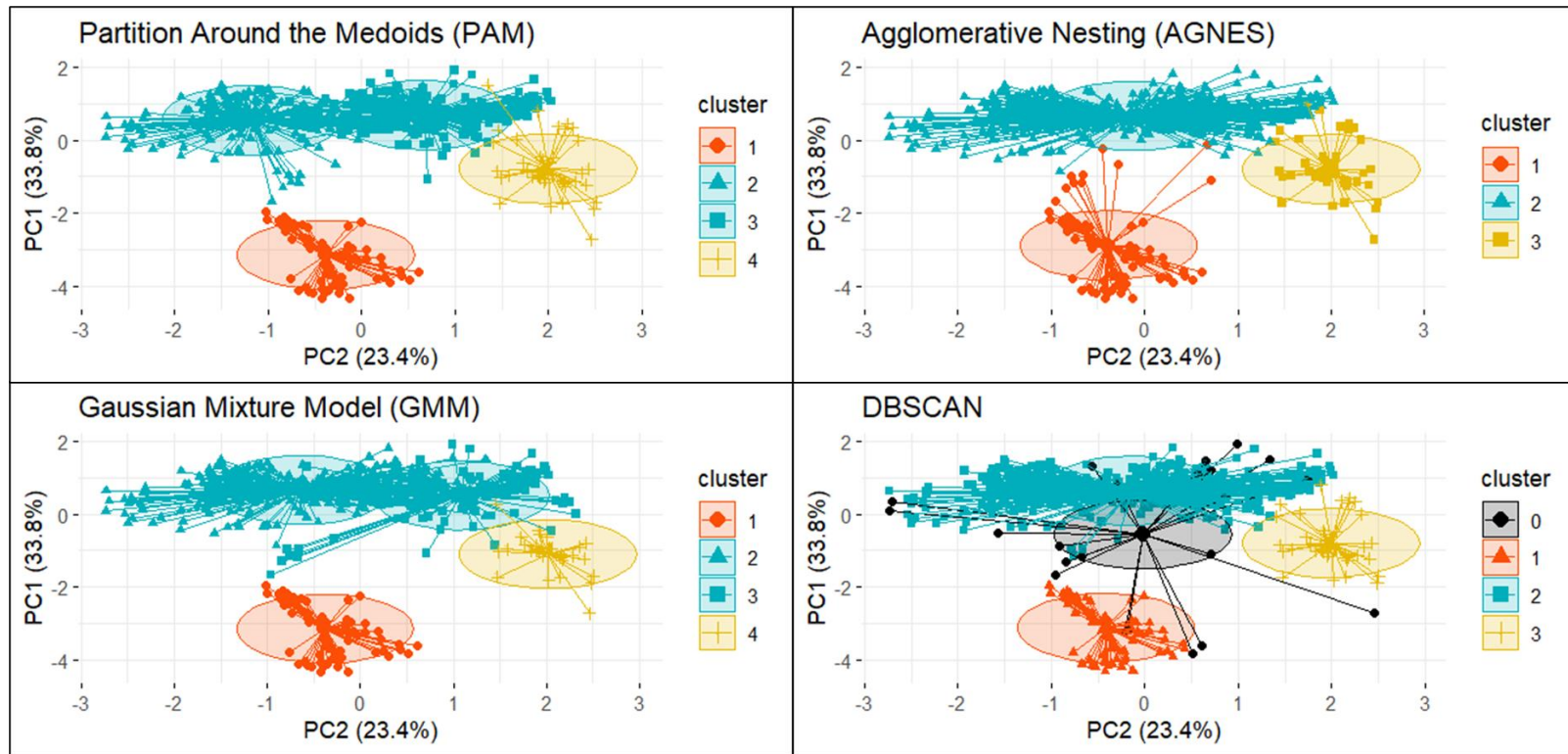


Figure 103: Visualization of the clustering output from the four clustering algorithms as specified in Table 1. Note that for DBSCAN cluster 0 are noisy data points not belonging to any cluster.

Table 15: Summary of algorithm specification which enables the identification of clusters RPC1 and RPC2

Algorithm	Specification	
PAM	<i>Feature Scaling:</i>	Robust
	<i>Number of Clusters:</i>	4
	<i>Distance Metric:</i>	Manhattan
AGNES	<i>Feature Scaling:</i>	Robust
	<i>Number of Clusters:</i>	3
	<i>Linkage Method:</i>	Ward
GMM	<i>Feature Scaling:</i>	Robust
	<i>Number of Clusters:</i>	4
DBSCAN	<i>Feature Scaling:</i>	Robust
	<i>Epsilon:</i>	1
	<i>Minimum Points:</i>	17

The data within **Figure 101** not included in RPC1 and RPC2 form a substantial cloud, termed the residual cluster. Although data points within this cluster were separable, the dominance of the two more distinct clusters obscured their internal structure. Consequently, data points from the two diamond grinding clusters were removed to clarify the residual cluster's structure. A density plot for the residual cluster is displayed in **Figure 104**, highlighting five areas of high density with colored circles.

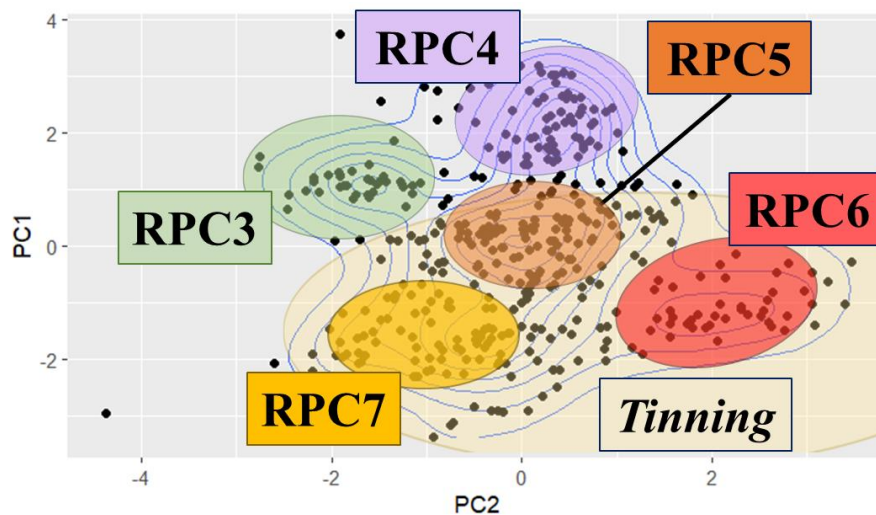


Figure 104: Density plot of the first two principal components for flexible pavements highlighting the clustering structure within the residual cluster.

RPC3 consists solely of exposed aggregate surfaces, while RPC4 includes pavement surfaces with dragged texturing, such as AstroTurf, and others surfaces so polished they appear dragged. The remaining clusters involve various types of tining: RPC5 features shallow grooves with short landing spacing, RPC6 exhibits deeper grooves and wider spacing, and RPC7 consists of heavily polished tined surfaces. Examples of these clusters are shown in **Figure 105**. Table 16 provides a summary of the hyperparameter specifications for the four clustering algorithms used.

Table 16: Summary of algorithm specification which enables the identification of clusters RPC3 through RPC7

Algorithm	Specification	
PAM	<i>Feature Scaling:</i>	Robust
	<i>Number of Clusters:</i>	5
	<i>Distance Metric:</i>	Manhattan
AGNES	<i>Feature Scaling:</i>	Robust
	<i>Number of Clusters:</i>	6
	<i>Linkage Method:</i>	Ward
GMM	<i>Feature Scaling:</i>	Robust
	<i>Number of Clusters:</i>	7
DBSCAN	<i>Feature Scaling:</i>	Robust
	<i>Epsilon:</i>	0.615
	<i>Minimum Points:</i>	13

While visually identifying these five clusters may seem straightforward, mathematically detecting them presents a unique challenge. Nonetheless, all four clustering techniques implemented were capable of identifying these five clusters, though with varying degrees of efficiency, as illustrated in **Figure 105**. PAM proved to be the most efficient, successfully discerning the clusters with a specification of only five groups, whereas GMM was the least efficient, requiring detection of seven clusters to identify the five groups depicted in **Figure 106**. DBSCAN managed to identify subsets of each cluster, but adjusting its hyperparameters was a time-consuming process.

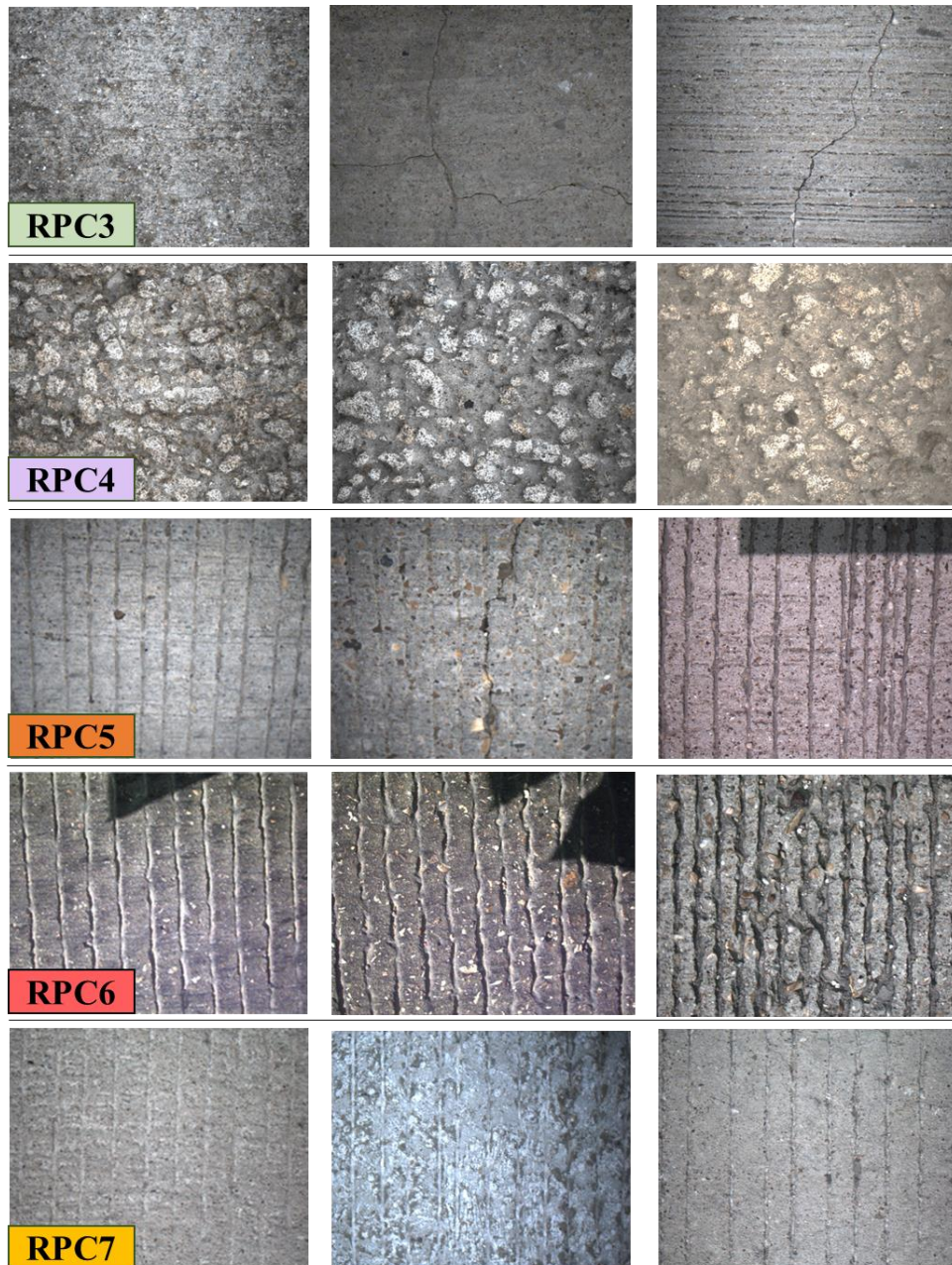


Figure 105: Representative pictures of each of the five cluster within the residual cluster. Members of the same clusters are shown along the row.

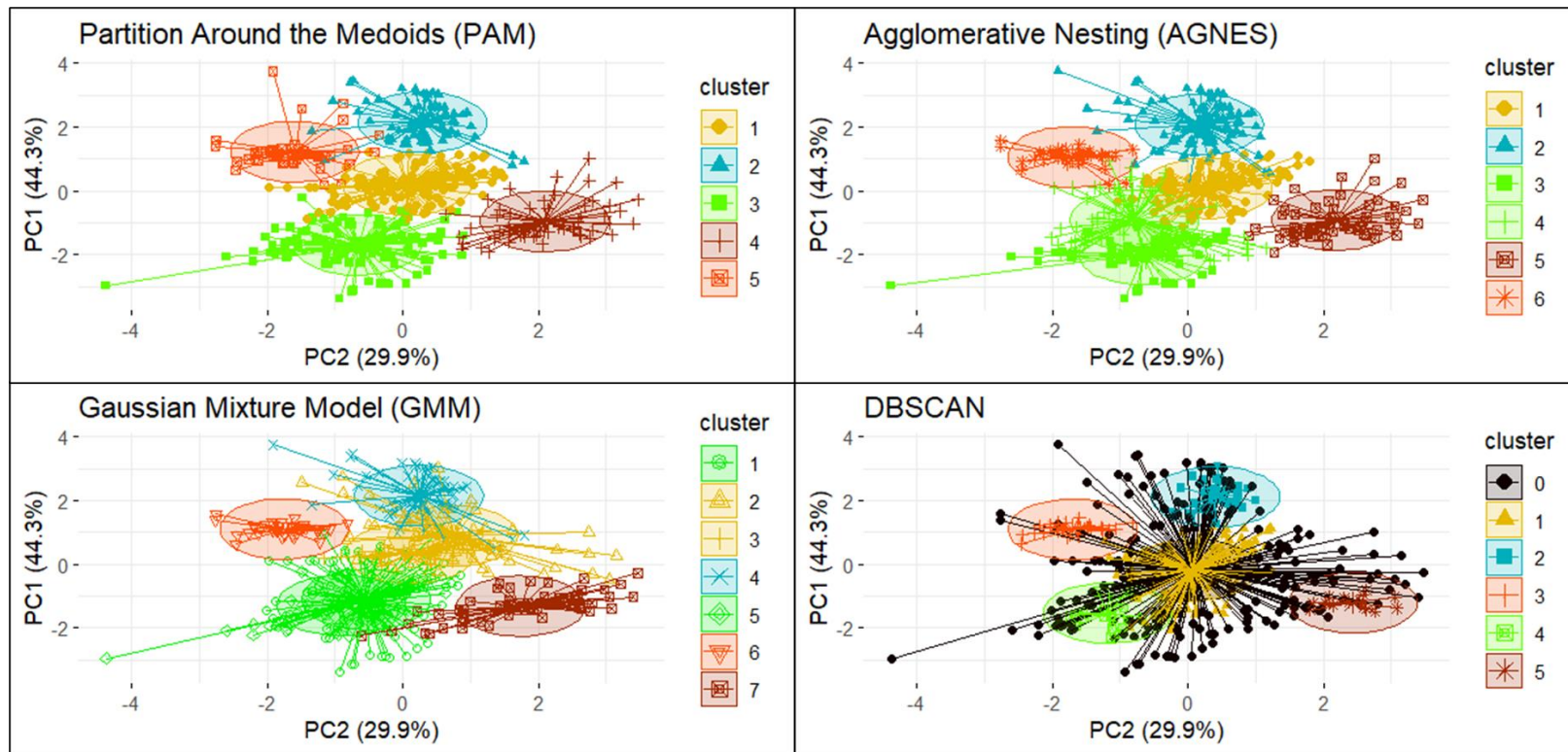


Figure 106: Visualization of the clustering output from the four clustering algorithms as specified in Table 16. Note that for DBSCAN cluster 0 are noisy data points not belonging to any cluster.

The results of the cluster analysis for rigid pavements are summarized in Table 17. Similar to the approach for flexible pavements, this table includes a description of each cluster, a set of six texture indices with their corresponding median values within the same cluster, the number of pavement sections in each cluster, and their percentage of the total. This comprehensive overview allows for a detailed comparison and understanding of the textural characteristics of each cluster.

Table 17: Summary of rigid pavement cluster analysis with description of the cluster and median index measurements of the member pavements

RPC1: New Generation Diamond Grinding						n: 76 (14%)
MPD	Rku	Cwst	Cws	Ral	p Hcme	
1.46	-0.20	5.89	0.21	0.69	-5.66	
RPC2: Conventional Diamond Grinding						n: 47 (8%)
MPD	Rku	Cwst	Cws	Ral	p Hcme	
0.79	0.11	1.96	0.84	0.40	-6.25	
RPC3: Exposed aggregates						n: 50 (9%)
MPD	Rku	Cwst	Cws	Ral	p Hcme	
0.52	0.43	3.46	2.38	1.34	-6.03	
RPC4: Dragged surfaces						n: 67 (12%)
MPD	Rku	Cwst	Cws	Ral	p Hcme	
0.77	-0.14	5.40	1.92	1.54	-5.63	
RPC5: Tining, shallow grooves, short landing spacing						n: 141 (26%)
MPD	Rku	Cwst	Cws	Ral	p Hcme	
0.67	0.89	5.05	1.90	1.11	-5.68	
RPC6: Tining, deep grooves, wide landing spacing						n: 126 (23%)
MPD	Rku	Cwst	Cws	Ral	p Hcme	
0.92	1.41	6.46	1.52	1.00	-5.44	
RPC7: Tining, significantly polished						n: 53 (10%)
MPD	Rku	Cwst	Cws	Ral	p Hcme	
0.59	2.47	5.71	2.10	0.82	-6.18	

5.4.2. IBC Models Comparison

The Performing Agency conducted a qualitative analysis to comprehensively evaluate the characteristics of the SL models before proceeding with their training

and validation. This analysis involved formulating a series of yes/no questions across four categories: model sensitivities (Set S), statistical assumptions (Set A), computational concerns (Set C), and implementation feasibility (Set I). These questions aimed to evaluate the adequacy, limitations, transparency, interpretability, and adaptability of the models. The responses to these questions were systematically organized in tables, employing a color-coded system where each color has distinct meanings depending on the question set.

5.4.2.1. Model Sensitivities (Set S)

The Performing agency explored several dimensions of model sensitivity, including:

- S1: Is the model sensitive to outliers?
- S2: Is the model sensitive to the scale of the features?
- S3: Is the model sensitive to multicollinearity?
- S4: Is the model sensitive to irrelevant features?
- S5: Is the model sensitive to imbalanced data?
- S6: Is the model sensitive to high dimensionality?
- S7: Is the model sensitive to missing data?
- S8: Is the model sensitive to autocorrelation?
- S9: Is the model sensitive to nonlinearity in the data?

The sensitivity of each model was depicted using a color-coded table (Table 18). Cells indicating a significant decrease in performance, as measured by the F1 score, were colored *red*. *Yellow* cells indicated performance changes contingent upon specific model attributes, such as the choice of base learner in ensemble methods. Conversely, cells unaffected by the criterion, demonstrating model robustness, were colored *green*.

Table 18: Summary of Model Sensitivities

	KNN	GNB	LOG	PRO	CLL	DT	RF	BA	AB	GB	XGB	SVM	SNN	DNN
S1	red	red	red	red	red	green	green	yellow	red	red	red	yellow	red	red
S2	red	red	red	red	red	green	green	yellow	yellow	yellow	yellow	red	red	red
S3	yellow	red	red	red	red	red	green	yellow	yellow	yellow	green	yellow	yellow	yellow
S4	red	red	red	red	red	red	yellow	green	red	yellow	yellow	red	red	red
S5	red	red	red	red	yellow	red	yellow	yellow	red	yellow	green	red	red	red
S6	red	green	red	red	red	yellow	yellow	yellow	red	green	yellow	yellow	green	green
S7	red	red	red	red	red	yellow	yellow	yellow	yellow	yellow	green	red	red	yellow
S8	yellow	red	red	red	red	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
S9	green	red	yellow	yellow	yellow	green	green	green	green	green	green	green	green	green

The sensitivity analysis reveals differing levels of responsiveness among various models to factors such as outliers, multicollinearity, and feature scale. Notably, models such as Decision Trees (DT) and Random Forests (RF) exhibit robustness (*green* cells) against challenges like high dimensionality and irrelevant features, making them well-suited for complex data landscapes. Conversely, models like Logistic Regression may exhibit sensitivity (*red* cells) to multicollinearity, affecting their performance when predictor variables are highly correlated. This underscores the versatility of certain models across diverse data scenarios, while emphasizing the need for meticulous data preprocessing to ensure precise outcomes with others.

5.4.2.2. Model Assumptions (Set A)

The following set of questions assessed the models based on their statistical assumptions about the data:

- A1: Does the model require homoskedasticity?
- A2: Does the model require linearity between features and labels?
- A3: Does the model require the residuals to follow a given distribution?
- A4: Does the model require the features to follow a given distribution?
- A5: Does the model require the observations to be independent?
- A6: Does the model require the constituent models to be independent?
- A7: Is there another statistical assumption(s) the model makes?

For each assumption, a *red* cell in Table 19 indicated a requirement by the model, potentially limiting its applicability without fulfilling these conditions. A *green* cell signaled no such requirement, enhancing the model's flexibility. Additional

statistical assumptions were specifically noted with a superscript detailing the nature of the assumption. This detailed assessment helps clarify the suitability of various models under different conditions, ensuring a more informed selection and application of the appropriate modeling technique. Nonetheless, some models operate on assumptions beyond statistical principles. For instance, KNN relies on the notion that proximity in distance signifies similarity, whereas greater distance indicates dissimilarity. These non-parametric assumptions were not considered in this analysis.

Table 19: Summary of Model Statistical Assumptions

A1														
A2														
A3														
A4														
A5														
A6														
A7														

¹ Assumption of Independence of Irrelevant Alternatives (only applicable to multinomial logit)

² Latent variable (label in the probit model) is assumed to come from a normal distribution and the assumption of proportional odds.

³ Assumption of proportional hazards

The adherence to statistical assumptions significantly affects the suitability of models for different data types. For example, models that require a linear relationship between features and labels, such as Logistic Regression (LOG), are marked in *red* when this assumption is unmet, highlighting their limitations in managing nonlinear relationships. The presence of additional assumptions, indicated with a red cell and a superscript (e.g., Probit Regression -PRO), necessitates a thorough understanding of the data distribution prior to deployment. This emphasizes the critical role of preliminary data analysis in ensuring that the conditions for optimal model performance are satisfied.

5.4.2.3. Model Computational Concerns (Set C)

The following set of questions assessed the models' performance in terms of computational resources:

- C1: Does the model take relatively long to train?
- C2: Does the model take relatively long to make predictions?
- C3: Does the model have high computer memory requirements?

In the corresponding table (Table 20), a *red* cell indicates undesirable traits, *yellow* suggests conditional desirability depending on data quantity or hyperparameter settings, and *green* denotes favorable characteristics.

Table 20: Summary of Model Computational Concerns

	KNN	GNB	LOG	PRO	CLL	DT	RF	BA	AB	GB	XGB	SVM	SNN	DNN
C1	green	green	green	green	yellow	yellow	red	red	red	red	yellow	red	red	red
C2	red	green	green	green	green	green	red	red	red	yellow	green	yellow	yellow	yellow
C3	red	green	yellow	yellow	yellow	yellow	red	red	red	red	yellow	red	red	red

The results for the question in Set C highlight the practical considerations in model deployment, particularly in terms of computational efficiency. For instance, models like Deep Neural Networks or most of the ensemble methods, with high demands for computational resources, may not be practical in time-sensitive or hardware-constrained environments. Conversely, simpler models such as Gaussian Naïve Bayes may be more appropriate for quick deployment or on systems with limited computational power due to their lower demands.

5.4.2.4. Model Complexity and Interpretability (Set I)

The final set of questions seeks to explore the models' transparency, complexity, interpretability, and adaptability:

- I1: Does the model have a “black box” framework?
- I2: Is the model's tuning complex?
- I3: Is the model easy to interpret?
- I4: Can the model handle multinomial data?

In the corresponding table, (Table 21) cells are colored *red* if the answer is no, *yellow* if the outcome is dependent on data volume or hyperparameter configuration, and *green* if the answer is yes.

Table 21: Summary of Model Implementation, Interpretability, Transparency and Adaptability

	KNN	GNB	LOG	PRO	CLL	DT	RF	BA	AB	GB	XGB	SVM	SNN	DNN
I1	green	green	green	green	green	green	yellow	yellow	yellow	yellow	yellow	yellow	red	red
I2	green	green	yellow	yellow	yellow	green	red	red	yellow	red	red	red	red	red
I3	green	green	green	green	green	green	yellow	red	yellow	yellow	yellow	red	red	red
I4	green	green	green	green	red	green	green	green	green	green	green	green	green	green

Interpretability is a critical aspect, especially in sectors where transparency is essential, such as healthcare, finance, and civil engineering. The analysis shows that models like K-Nearest Neighbors (KNN) and Decision Trees (DT) are more interpretable (indicated by green cells), allowing for clearer insight into the decision-making process. Conversely, models falling under the “black box” category, such as Deep Neural Networks (DNN), provide less transparency in how decisions are formulated, which can pose regulatory or operational challenges.

In summary, these tables offer crucial insights for selecting the most suitable models for specific tasks, by weighing both their technical capabilities and practical implications. This careful selection process ensures that the chosen machine learning models not only achieve optimal performance but also comply with deployment feasibilities and industry regulations.

5.4.3. IBC Classification

The classification of pavement surfaces based on their texture indices was executed using all fourteen ML algorithms outlined in Section 3.1.1. However, some algorithms were restricted and unable to perform multi-class classification due either to inherent limitations (the CLL model) or the absence of readily available code libraries that would make the model adaptable to those circumstances (the PRO model).

The performance of these ML algorithms, including those for image classification, was evaluated based on the weighted average model F1 Score—a metric that calculates the harmonic mean of precision (the proportion of *properly* classified pavements over all classifications performed) and recall (the ratio of *properly* classified pavements over the total number pavements predicted for that class). The F1 score ranges from 0 to 1, where values close to 1 signify near-perfect classification and scores near 0 indicate pervasive misclassification. The weighted average F1 score was chosen over the standard F1 score to account for varying sample sizes across different pavement types, but for simplicity, it is referred hereafter as merely “F1 score.”

In machine learning, data is typically divided into three distinct datasets before model development: *training*, *validation*, and *testing*. The *training* dataset is used to teach the model to recognize patterns and learn from them. The *validation* dataset is used to finetune the model’s parameters and select the best version by providing feedback on its performance during training. The *testing* dataset evaluates the final model’s performance in an unbiased manner after training and validation. This structured division ensures that the model can generalize effectively to new, unseen

data, rather than merely memorizing patterns within the training data. Note that *all F1 scores reported in this report are derived from the testing dataset.*

This section is segmented into four subsections, each focusing on a specific level of prediction accuracy achieved by the algorithm. Notably, the fifth level of specificity is omitted from this section of the report due to insufficient performance observed when using texture indices.

5.4.3.1. Level of Specificity #1: Flexible Versus Rigid

The dataset used to classify between rigid and flexible pavements comprised the ground truth dataset (DS1) and the holdout dataset (DS2). DS1, with a total of 2,346 0.1-mile-long pavement sections, employed an unconventional 30%-70% training-validation split. This specific division was chosen because pavement sections from the same site might be included in both training and validation datasets, and such double inclusions could negatively affect model performance in the testing dataset. By limiting the training data to 30%, the algorithms were compelled to generalize patterns effectively, preventing overfitting to overly site-specific patterns.

In contrast, DS2, consisting of 2,019 samples, was exclusively used to evaluate the performance of the trained models (the testing dataset). The distinction between rigid and flexible pavements, designated as Level of Specificity #1, was the only category where training, validation and testing data was distributed between DS1 and DS2. This arrangement was possible because accurate classification of pavement type was based on visual inspections conducted by the Performing Agency during data collection. For higher levels of specificity, only data from DS1 was used as the details for those levels could not be ascertained with complete certainty from DS2.

The hyperparameters for all ML classification models, where applicable, were optimized to maximize the testing F1 score while ensuring that the training and validation F1 scores did not significantly differ from the testing F1 score. This approach helps prevent overfitting of the models by keeping the differences between these scores negligible for practical purposes. Once trained for level #1 specificity, the models were tested against the data in DS2, and their testing F1 scores were recorded, as illustrated in **Figure 107**.

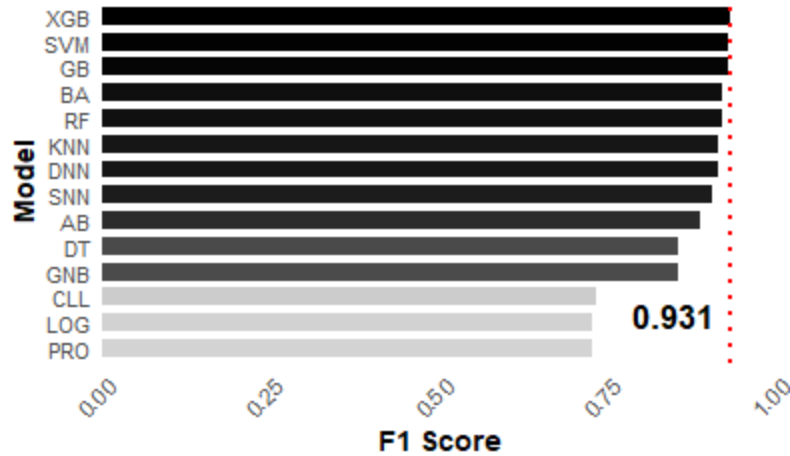


Figure 107: Summary of F1 scores across all 14 ML models for classification pavements using Level #1 specificity.

The bar chart suggests that ensemble methods, particularly gradient boosting (XGB and GB), were highly effective in classifying whether a pavement is rigid or flexible, achieving a maximum testing F1 score of 0.931. Conversely, at the lower end of the spectrum, probabilistic models such as LOG, PRO, CLL, and GNB were observed. Among them, LOG performed the poorest with an F1 score of 0.73. Nevertheless, it's worth noting that even the least performing algorithm attained a predictive power close to 0.800, a commonly accepted threshold for practical implementation of algorithms. This underscores the ease of achieving a high F1 score within the first level of specificity using only texture indices, as anticipated.

Table 22 and Table 23 present the confusion matrix and classification report for the best model, XGB, at classifying the first level of specificity.

Table 22: Confusion matrix for the XGB model in Level #1 specificity

		<i>Predicted Class</i>	
		<i>Rigid</i>	<i>Flexible</i>
<i>True Class</i>	<i>Rigid</i>	840	39
	<i>Flexible</i>	44	288

Table 23: Classification report for the XGB model in Level #1 specificity

Class	Testing Sample Size	Testing F1 Score	Weighted Average F1 Score
Flexible	884	0.959	0.931
Rigid	327	0.874	

5.4.3.2. Level of Specificity #2: General Mix/Surface Type

The second level of specificity involves differentiating flexible pavements into surface treatments and hot and warm mix asphalts and categorizing rigid pavements into plastic and hardened surface texturing. The dataset for training, validating, and testing these models exclusively comprises DS1, with a data split of 30%-20%-50%, respectively. A crucial aspect of these and all subsequent models is the inclusion of pavement type (flexible or rigid) as one of the input variables. This adjustment prevents the misidentification of smoother flexible pavement surfaces as their rigid counterparts, such as mistaking a thin overlay mix for a polished dragged concrete surface.

Ground truth data regarding pavement type was utilized as an input variable. However, predictions from the best-performing model at Level #1 will be tested to assess the effectiveness of this approach.

Due to their limitations in predicting categorical outcomes and the absence of readily available libraries for implementation, two models, namely CLL and PRO, were excluded from the analysis. All remaining models were finely tuned to maximize the testing F1 score while ensuring that both training and validation scores remained high to prevent overfitting. The testing F1 scores for these models are presented in **Figure 108**.

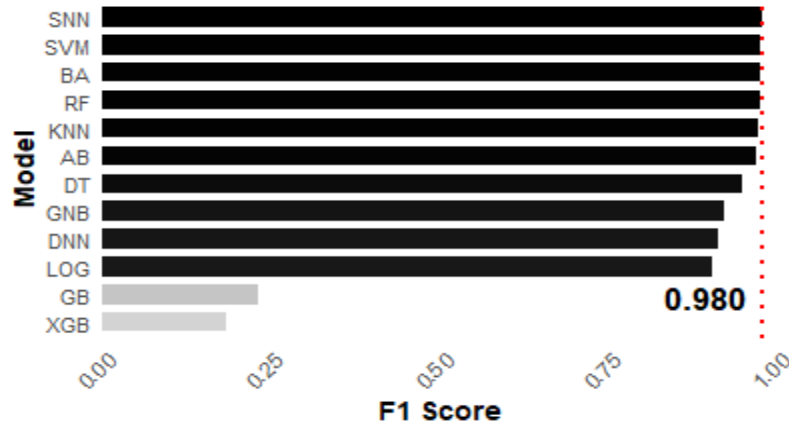


Figure 108: Summary of F1 scores across twelve ML models for classification pavements using Level #2 specificity.

The bar chart reveals that the shallow neural network (SNN) achieved the highest performance, with an F1 score of 0.980, closely matched by the next five highest scores. Interestingly, gradient boosting models populated the lower end of the chart, with the XGB model recording a notably low F1 score of 0.185. This poor performance is attributable to the inherent limitations of the gradient boosting framework, where an error by a weak learner, given the structure of the provided data, proves difficult for subsequent weak learners to correct.

It's important to note that these high results were achieved with the inclusion of pavement type information as an input variable. Without this data, the maximum F1 score for the model dropped to around 0.750, highlighting the challenges in accurately identifying specific pavement surfaces using only texture indices without prior knowledge of the pavement type. Table 24 and Table 25 provide a detailed look at the confusion matrix and the classification report for the best performing SNN model at this second level of specificity, respectively.

Table 24: Confusion matrix for the SNN model in Level #2 specificity

		<i>Predicted Class</i>			
		<i>HWMA</i>	<i>ST</i>	<i>PCT</i>	<i>HCT</i>
<i>True Class</i>	<i>HWMA</i>	557	10	0	1
	<i>ST</i>	9	321	0	0
	<i>PCT</i>	0	0	205	2
	<i>HCT</i>	0	0	3	61

Where HWMA stands for hot and warm mix asphalt, *ST* stands for surface treatment, *PCT* stands for plastic concrete texturing, and *HCT* stands for hardened concrete texturing.

Table 25: Classification report for the SNN model in Level #2 specificity

Class	Testing Sample Size	Testing F1 Score	Weighted Average F1 Score
Hot and Warm Mix Asphalt	567	0.961	0.980
Surface Treatment	321	0.984	
Plastic Concrete Texturing	208	0.987	
Hardened Concrete Texturing	64	0.973	

5.4.3.3. Level of Specificity #3: Specific Mix/Surface Type

At the third level of specificity, the general surface types are further subdivided into specific mixes or types of PCC surface finishing techniques, as detailed in **Figure 109**. The dataset utilized for training, validation, and testing the model once again comprised DS1, with the same data split as observed in the second level of specificity (30%-20%-50%). Notably, the pavement type was incorporated as one of the inputs to predict the third level of specificity, enhancing the accuracy of surface prediction, as depicted in **Figure 109**.

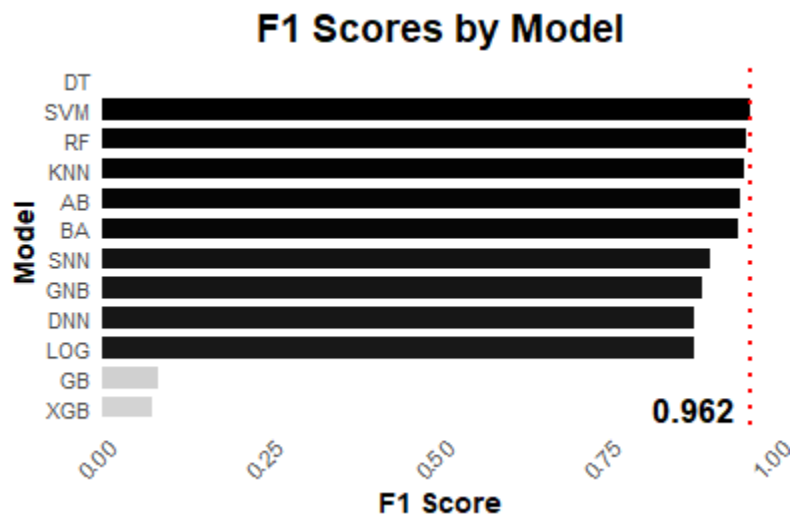


Figure 109: Summary of F1 scores across twelve ML models for classification pavements using Level #3 specificity.

Figure 109 shows that the support vector machine (SVM) model exhibited the strongest performance with an F1 score of 0.962, closely followed by the next four highest scores. Despite a slight decrease in F1 score due to the complexity of classifying more pavement surfaces, the results remained highly satisfactory. However, the gradient boosting models continued to struggle, underfitting the data due to inherent limitations in their framework that compound prediction errors from earlier weak learners during the development process. Gradient boosting models aside, the least effective model was the logistic regression (LOG), achieving a still respectable F1 score of 0.870, which is still considered adequate for scientific applications.

Notably, in **Figure 109**, the decision tree model (DT) did not register an F1 score. This absence resulted from the model's inability to accurately predict any instance of the high friction surface treatment (HFST) pavement surface, a limitation due to its simplistic framework and insufficient support for that particular surface type in the dataset. This outcome was anticipated, given the decision tree model's inadequacy for handling complex classification tasks like those required here. Table 26 and Table 27 provide a detailed confusion matrix and classification report for the best-performing SVM model at the third level of specificity, respectively.

Table 26: Confusion matrix for the SVM model in Level #3 specificity

		<i>Predicted Class</i>								
		<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>S5</i>	<i>S6</i>	<i>S7</i>	<i>S8</i>	<i>S9</i>
<i>True Class</i>	<i>S1*</i>	325	14	5	1	0	0	0	0	0
	<i>S2</i>	2	217	0	0	0	0	0	0	0
	<i>S3</i>	7	1	317	0	0	0	0	0	0
	<i>S4</i>	0	0	0	6	0	0	0	0	0
	<i>S5</i>	0	0	0	0	28	3	1	0	0
	<i>S6</i>	0	0	0	0	2	151	0	2	1
	<i>S7</i>	0	0	0	0	0	0	16	0	0
	<i>S8</i>	0	0	0	0	0	5	0	39	0
	<i>S9</i>	0	0	0	0	0	0	0	0	21

**S1* stands for dense- and gap-graded mixes, *S2* stands for open-graded mixes, *S3* stands for seal coat, *S4* stands for high friction surface treatment, *S5* dragged texturing, *S6* stands for tined texturing, *S7* stands for exposed aggregate texturing,

S8 stands for next generation diamond grinding, and S9 stands for conventional diamond grinding.

Table 27: Classification report for the SVM model in Level #3 specificity

Class	Validation Sample Size	Validation F1 Score	Weighted Average F1 Score
Dense- and Gap-Graded Mixes	334	0.957	0.962
Open-Graded Mixes	232	0.962	
Seal Coat	322	0.980	
High Friction Surface Treatment	7	0.923	
Dragged Texturing	30	0.903	
Tinned Texturing	159	0.959	
Exposed Aggregate Texturing	17	0.970	
Next Generation Diamond Grinding	41	0.918	
Conventional Diamond Grinding	22	0.977	

5.4.3.4. Level of Specificity #4: Specific Mix/Surface Type II

The fourth level of specificity categorizes pavement surfaces into highly detailed classifications based on the information provided by the Receiving Agency. At this level, dense and gap-graded mixes, as well as seal coats, are differentiated by their gradation, while open mixes are subdivided into Porous Friction Courses (PFCs) and bonded wearing courses. Similarly, tined concrete pavements are classified into fixed and random tining. However, certain surfaces such as exposed aggregates, HFSTs, all diamond grinding types, and dragged concrete surfaces do not lend themselves to further specification and are thus excluded from the dataset used for training these models. Additionally, some sites lacked sufficiently detailed surface information to achieve this level of specificity; for example, many surface treatments were simply labelled as seal coats without detailed gradation information, leading to their exclusion from the dataset. Consequently, the final subset of data from DS1 used for model training comprised 1,604 observations, a reduction from the original 2,346.

The data split for training, validation, and testing was adjusted to 30%-40%-30%. This change was necessary as tuning these models with only 20% for validation previously proved to be inadequate. Pavement type continued to be used as an input variable to enhance the accuracy of surface predictions at this advanced level of specificity, as shown in **Figure 110**.

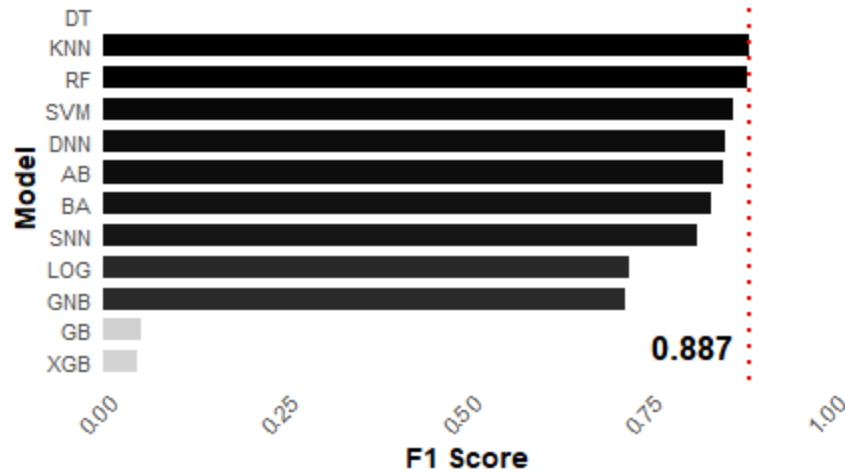


Figure 110: Summary of F1 scores across twelve ML models for classification pavements using Level #4 specificity.

Figure 110 shows that the K-nearest neighbor (KNN) model outperformed others with an F1 score of 0.887, closely followed by the Random Forest (RF) model. Although there was a significant reduction in the F1 score due to an increase in the number of pavement surfaces to classify, it remained above the previously mentioned threshold of 0.80.

However, the Decision Tree (DT) model continued to struggle, failing to classify at least three of the pavement surfaces, and the gradient boosting models persistently underperformed despite multiple optimization efforts. The next lowest F1 scores, around 0.70, were recorded by the probabilistic models, Gaussian Naive Bayes (GNB) and Logistic Regression (LOG). Despite these scores seeming relatively high, they were not practically useful. Preliminary tests on the holdout dataset DS2 showed that GNB and LOG failed to consistently assign a surface type to sites with homogeneous surfaces, instead inaccurately suggesting multiple surface types for subsequent sites despite photographic evidence and change point analysis indicating no such variation.

Due to the large number of surface types involved in the fourth level of specificity, the confusion matrix for the best model is not displayed; only the classification report for the top-performing KNN model is provided in Table 28. Notably, even the best model achieved low F1 scores for certain pavement types such as random

ting, bonded wearing course, and fog seals, which were too low to be considered effective. These low scores suggest either a lack of sufficient data to accurately classify these surfaces, as was the case with random tining where only one site was available, or a minimal textural variance that made distinguishing these surfaces challenging, as seen with bonded wearing courses compared to PFCs and fog seals compared to different gradation seal coats.

Table 28: Classification report for the KNN model in Level #4 specificity

Class	Validation Sample Size	Validation F1 Score	Weighted Average F1 Score
Type C	43	0.882	0.887
Type D	25	0.808	
Thin Overlay Mix	29	0.909	
Porous Friction Course	101	0.915	
Bonded Wearing Course	36	0.677	
Seal Coat Grade 4	40	0.914	
Seal Coat Grade 3S	19	0.905	
Seal Coat Grade 3	27	0.880	
Fog Seal	25	0.792	
Microsurfacing	8	0.933	
Fixed Tining	24	0.958	
Random Tining	6	0.667	

Note that the IBC models do not provide a Level #5 specificity, which differentiates transverse from longitudinal tining. This omission is due to the manner in which the texture data was collected. Since the laser sensor was positioned at a 45-degree angle to the direction of traffic, both longitudinal and transverse tining exhibit identical appearances at the profile level. Distinguishing between these two surfaces would necessitate a more intricate 3D analysis of the texture data.

Consequently, the Performing Agency chose to classify this level of specificity using picture classification models instead.

5.4.4. PBC Classification

The Performing Agency enhanced the ResNet50 architecture for pavement surface classification by adding an extra hidden layer with 512 neurons and making the entire network trainable. This modification allowed the model to better discern unique pavement characteristics such as aggregate shapes, texture patterns, and material colors through the dynamic updating of weights and biases in all layers during training.

For optimization, the model utilized the Adam optimization algorithm, known for its effectiveness with large datasets and complex architectures, offering advantages over traditional methods like stochastic gradient descent. Each neuron within the hidden layers employed a rectified linear unit (ReLU) activation function to introduce non-linearity, enabling the capture of complex data patterns and improving learning capabilities.

The model's output layer featured a Softmax activation function, ideal for multi-class classification by normalizing the output values between 0 and 1, representing the likelihood of each class, with all probabilities summing to one. To enhance model reliability and prevent overtraining, checkpoints were used to save the most effective model iterations, and training could be halted early if no significant improvements were observed. Additionally, a dropout layer was incorporated to randomly deactivate a portion of the neurons during training, reducing overfitting by introducing noise into the system and promoting better generalization.

The finalized CNN model was trained and validated using a dataset of 48,328 manually inspected images representing all surveyed surface types, with a data split of 70%-30% for training and validation, respectively. After achieving satisfactory F1 scores, the model was tested on a new set of images. Detailed results for each level of specificity are provided in the subsequent sections.

5.4.4.1. Level of Specificity #1: Flexible Versus Rigid

The testing dataset utilized for the CNN model comprised a total of 1,300 rigid and 1,599 flexible pavements, all sourced from testing sites distinct from those used in the model's training and validation phases. A summary of the confusion matrix and classification report can be found in Table 29 and Table 30, respectively. The results unequivocally indicate that images are the most effective means of distinguishing between flexible and rigid pavement, boasting a remarkable predictive power with an F1 score of 0.995. The few misclassifications observed

predominantly involved heavily polished rigid pavements erroneously classified as flexible pavement. Given sufficient time for proper training and validation of a CNN model, this approach likely supersedes the use of indices when predicting pavement surface at the first level of specificity.

Table 29: Confusion matrix for the CNN in Level #1 specificity

		<i>Predicted Class</i>	
		<i>Flexible</i>	<i>Rigid</i>
<i>True Class</i>	<i>Flexible</i>	1,599	0
	<i>Rigid</i>	12	1,288

Table 30: Classification report for the CNN model in Level #1 specificity

Class	Testing Sample Size	Testing F1 Score	Weighted Average F1 Score
Flexible	1,599	1.000	0.995
Rigid	1,300	0.989	

5.4.4.2. Level of Specificity #2: General Mix/Surface Type

The same dataset utilized for level one predictions was employed to forecast pavement surface at the second level of specificity. Summaries of the confusion matrix and classification report can be found in Table 31 and Table 32 respectively. Similar to the IBC models, the CNN incorporated the ground truth pavement type as an additional feature to enhance prediction accuracy at the second level of specificity. However, despite parameter optimization efforts, the CNN's performance fell short, yielding an F1 score of 0.904. In comparison, the best-performing IBC model (the SNN, also a neural network) achieved a higher F1 score of 0.980. Notably, CNN encountered the greatest challenges when classifying surface treatments compared to other surface types.

Table 31: Confusion matrix for the CNN model in Level #2 specificity

		<i>Predicted Class</i>			
		<i>HWMA</i>	<i>ST</i>	<i>PCT</i>	<i>HCT</i>
<i>True Class</i>	<i>HWMA</i>	1,286	139	0	0
	<i>ST</i>	114	681	0	0
	<i>PCT</i>	0	0	507	12
	<i>HCT</i>	0	0	12	148

Where HWMA stands for hot and warm mix asphalt, *ST* stands for surface treatment, *PCT* stands for plastic concrete texturing, and *HCT* stands for hardened concrete texturing.

Table 32: Classification report for the CNN model in Level #2 specificity

Class	Testing Sample Size	Testing F1 Score	Weighted Average F1 Score
Hot and Warm Mix Asphalt	1,400	0.911	0.904
Surface Treatment	820	0.844	
Plastic Concrete Texturing	519	0.976	
Hardened Concrete Texturing	161	0.922	

5.4.4.3. Level of Specificity #3: Specific Mix/Surface Type

Utilizing the same database employed for levels one and two, predictions for pavement surface at the third level of specificity were made. A summary of the confusion matrix and classification report is provided in Table 33 and Table 34, respectively. Once again, the ground truth pavement type was included as an additional feature in the CNN model. However, as observed previously, the CNN's performance diminished with increased specificity. The model achieved an F1 score of 0.876, notably lower than the 0.962 achieved by the SVM model when utilizing texture indices. This decline in performance was primarily attributed to two surfaces: conventional diamond grinding and high friction surface treatment, while most other surfaces maintained sufficiently high F1 scores.

Table 33: Confusion matrix for the CNN model in Level #3 specificity

		<i>Predicted Class</i>								
		<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>S5</i>	<i>S6</i>	<i>S7</i>	<i>S8</i>	<i>S9</i>
<i>True Class</i>	<i>S1</i>	810	34	12	3	0	0	0	0	0
	<i>S2</i>	5	541	0	0	0	0	0	0	0
	<i>S3</i>	17	2	790	0	0	0	0	0	0
	<i>S4</i>	0	0	0	15	0	0	0	0	0
	<i>S5</i>	0	0	0	0	70	8	2	0	0
	<i>S6</i>	0	0	0	0	5	376	0	5	2
	<i>S7</i>	0	0	0	0	0	0	40	0	0
	<i>S8</i>	0	0	0	0	0	12	0	97	0
	<i>S9</i>	0	0	0	0	0	0	0	0	53

Where *S1* stands for dense- and gap-graded mixes, *S2* stands for open-graded mixes, *S3* stands for seal coat, *S4* stands for high friction surface treatment, *S5* dragged texturing, *S6* stands for tined texturing, *S7* stands for exposed aggregate texturing, *S8* stands for next generation diamond grinding, and *S9* stands for conventional diamond grinding.

Table 34: Classification report for the CNN model in Level #3 specificity

Class	Validation Sample Size	Validation F1 Score	Weighted Average F1 Score
Dense- and Gap-Graded Mixes	832	0.859	0.876
Open-Graded Mixes	578	0.869	
Seal Coat	801	0.881	
High Friction Surface Treatment	18	0.444	
Dragged Texturing	75	0.787	
Tinned Texturing	396	0.944	
Exposed Aggregate Texturing	42	0.850	
Next Generation Diamond Grinding	102	0.942	
Conventional Diamond Grinding	55	0.650	

5.4.4.4. Level of Specificity #4: General Mix/Surface Type II

A slightly smaller dataset was employed to evaluate the CNN model's performance in predicting pavement surface at the fourth level of specificity. This reduction was necessitated by the same constraint observed with the IBC models: certain pavement sites lacked the requisite detail to reach this level of pavement surface granularity. The classification report for the fourth level of specificity using images is succinctly outlined in Table 35.

Similar to previous levels of specificity, the ground truth pavement type was incorporated into the CNN as an additional feature. However, this time, there was a notable decline in the F1 score, plummeting to as low as 0.717. With this decline, the model's efficacy diminishes significantly, falling below the 0.800 threshold deemed acceptable for practical application. Consequently, the Performing Agency strongly advises against utilizing this particular CNN model for predicting pavement surface at the fourth level of specificity.

Table 35: Classification report for the KNN model in Level #4 specificity

Class	Validation Sample Size	Validation F1 Score	Weighted Average F1 Score
Type C	275	0.483	0.717
Type D	160	0.556	
Thin Overlay Mix	186	0.676	
Porous Friction Course	646	0.828	
Bonded Wearing Course	230	0.298	
Seal Coat Grade 4	256	0.769	
Seal Coat Grade 3S	122	0.533	
Seal Coat Grade 3	173	0.653	
Fog Seal	160	0.679	
Microsurfacing	51	0.933	
Fixed Tining	154	0.920	
Random Tining	38	0.727	

5.4.4.5. Level of Specificity #5: Specific Mix/Surface Type III

For the fifth and final level of specificity, which distinguishes between transverse and longitudinal tining, a distinct dataset was utilized for testing purposes. This new data set was expanded to include numerous instances of both longitudinal and transverse tining, in contrast to the less than 200 instances present in the previous testing dataset. Notably, in this iteration, the pavement type was not included as an additional input to the model.

The confusion matrix and classification report for the fifth level of specificity are summarized in Table 36 and Table 37, respectively. The results are indicative of the CNN model's exceptional performance in classifying transverse and longitudinal tining pictures of pavements, achieving near-perfection with only one misclassification. This outcome aligns with expectations, given that the tining

patterns in the pictures are orthogonal to each other, resulting in a testing F1 score of 0.999.

Table 36: Confusion matrix for the CNN in Level #5 specificity

		<i>Predicted Class</i>	
		<i>Transverse</i>	<i>Longitudinal</i>
<i>True Class</i>	<i>Transverse</i>	600	1
	<i>Longitudinal</i>	0	599

Table 37: Classification report for the CNN model in Level #5 specificity

Class	Testing Sample Size	Testing F1 Score	Weighted Average F1 Score
Transverse	600	1.000	0.999
Longitudinal	600	0.999	

Based on these results, the Performing Agency recommends a tiered approach to predicting pavement surface type. For initial pavement type classification at the first level of specificity, a picture-based CNN model is advised due to its superior performance. For subsequent levels of specificity—second through fourth—it is recommended to employ an index-based model, which has demonstrated the best fit, followed by another CNN model to classify the fifth and final level of specificity.

The agency has thoroughly evaluated each model’s statistical assumptions, sensitivities, computational demands, implementation feasibility, transparency, and predictive accuracy. This comprehensive analysis allows the Receiving Agency to select the most suitable model based on their specific accuracy requirements and operational needs.

Currently, these recommendations are provisional. The Performing Agency will proceed to validate these findings against the holdout dataset and additional testing scenarios, including varied pavement conditions across different traffic directions and repeated runs over the same pavement.

Chapter 6. Critical System Evaluation and Validation

6.1. Validation data

The Performing Agency has collected a set of data to validate the performance of the models developed in Chapter 5. This data is a combination of new data, some even unknown, that was never used at any point in the previous study), new and repeated measurements of the same sites of the original study, and even pavement sections used in other TxDOT projects. All these pavement sections were broken down into seven test sets which are described as follows:

- **Test Set #1 (TS1):** Comprised of pavements sections that were not included as part of the database delivered during Task #5. Therefore, they were not used to train and validate the algorithms developed during Task #6. These sites are approximately 1 to 3 miles long and homogeneous, and their surface type is *unknown*.
- **Test Set #2 (TS2):** Comprised of pavements sections that were originally collected for TxDOT Project No. 0-7031 “Towards Efficient Prediction of Highway Friction on an Annual Basis on Texas Network.” The researchers in that study used expert knowledge to define their pavement surface information. Their classification will be compared against the predictions produced by this algorithm.
- **Test Set #3 (TS3):** This set is comprised of four sites that were used for the development of these algorithms. These sites were surveyed on the same day, along the same direction of traffic and along the same lane. Essentially, they are *perfect repeats* of the data collected for the study. The goal of this test set is to assess consistency of the final prediction.
- **Test Set #4 (TS4):** This set is comprised of three sites that were used for the development of these algorithms. These sites were surveyed on the same day, along the outer lane, but in *different directions* of traffic (i.e., in opposite directions). The goal of this test set is to assess consistency of prediction across opposite lanes of the same sections and the impact of generalizing the pavement surface classification using data from only one direction.
- **Test Set #5 (TS5):** This set is comprised of three sites that were used for the development of these algorithms. These sites were surveyed on the same day, along the same direction of traffic, but in *different lanes* (either inner

or outer). The goal of this test set is to assess consistency of prediction across different lanes in a road.

- **Test Set #6 (TS6):** This set is comprised of one site that was used for the development of these algorithms. These sites were surveyed along the same direction of traffic, on the same lane, but on *different days*. While the data for this test set is limited, the goal was to assess if the prediction would change if the data was collected within a span of a week.

The summary of these test sites is presented in **Table 38**.

Table 38: Summary of Test Sets and Characteristics

Test Set	Description	Characteristics	Purpose
TS1	New pavement sections with unknown pavement surface information.	Completely homogeneous pavements, ranging from 1 to 3 miles not used during model training or validation.	Evaluate model performance on new, previously unused pavement data.
TS2	Pavement section from other studies where surface information is partially known.	Pavement sections from TxDOT project 0-7031, defined by expert opinions.	Compare expert opinions with algorithm predictions.
TS3	Pavement sections with repeated measurements along the same lane and same day.	Four sites surveyed on the same day, in the same lane, in the same direction of traffic.	Assess consistency of predictions and texture data with repeated measurements.
TS4	Pavement sections with repeated measurements in opposite directions of traffic.	Three sites surveyed in different directions of traffic on the same day.	Evaluate consistency of predictions across different traffic directions.
TS5	Pavement sections with repeated measurements in different lanes.	Three sites surveyed in different lanes (inner vs outer) on the same day.	Assess consistency of predictions across different lanes on a road.
TS6	Pavement sections with repeated measurements on different days.	One site surveyed in the same lane and direction of traffic on different days within a week.	Determine if predictions change with data collected on different days.

6.1.1. Test Set #1 (TS1): Unknown Surfaces

The Performing Agency collected data on five completely homogenous sites, whose information was unknown and not utilized in the development of models for this study.

The first site (**Figure 111**) is located in the Austin District, within Travis County, in the City of Austin along SH 0071 between Texas Reference Markers (TRM) 580 and 582.

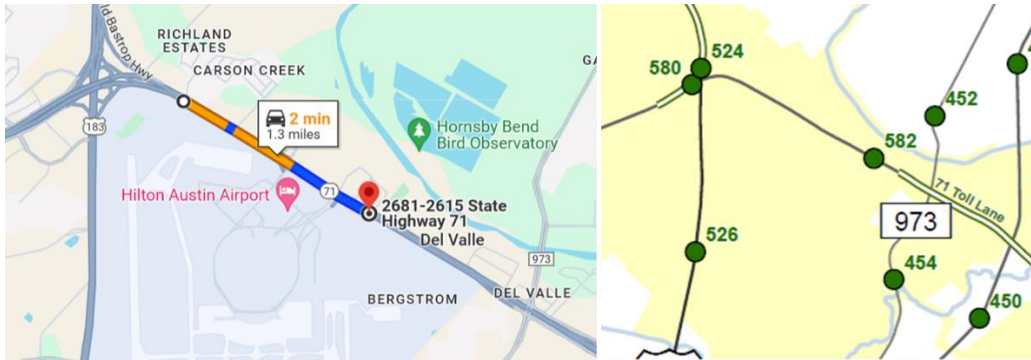


Figure 111: Reference map of TS1 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The second site (**Figure 112**) is located in the Austin District, within Travis County, in the City of Austin along US 0290 between TRMs 572 and 574.

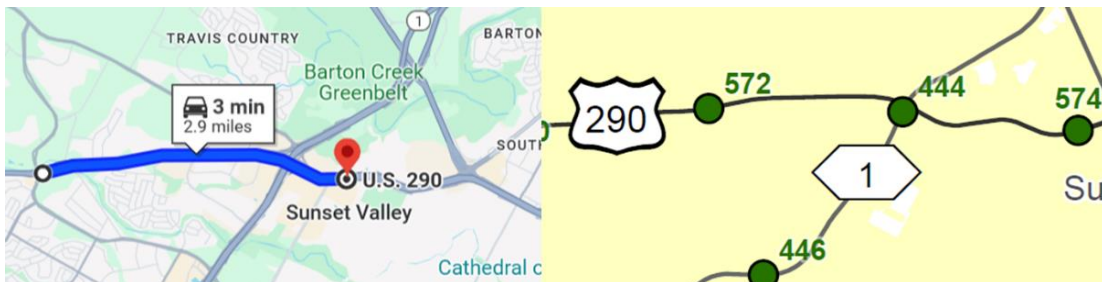


Figure 112: Reference map of TS1 section #2 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The third site (**Figure 113**) is located in the Brownwood District, within Brown County, southeast of the City of Early along US 0084 between TRMs 566 and 570.

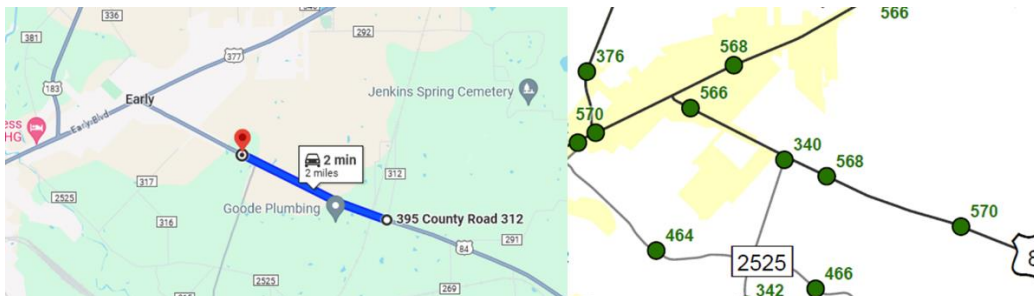


Figure 113: Reference map of TS1 section #3 in Brown County. (Left) Section on Google Maps, and (right) section on TRM map.

The fourth site (**Figure 114**) is located in the Austin District, within Travis County, in the City of Austin along SH 0071 between TRMs 576 and 580.

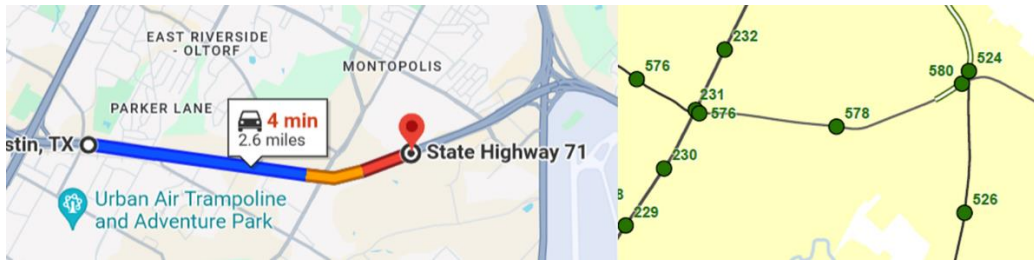


Figure 114: Reference map of TS1 section #4 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The fifth site (**Figure 115**) is located in the Waco District, within Hill County, north of the City of Abbott along IH 0035 between TRMs 362 and 364.

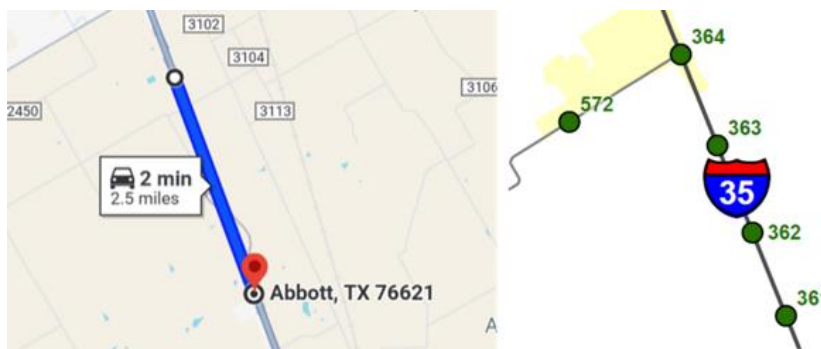


Figure 115: Reference map of TS1 section #5 in Hill County. (Left) Section on Google Maps, and (right) section on TRM map.

6.1.2. Test Set #2 (TS2): Previous Classified Pavements by Experts - TxDOT Project No. 0-7031

The sixth site (**Figure 116**) is located in the Austin District, within Williamson County, between the cities of Thrall and Rockdale along FM 1063 between TRMs 416 and 418.

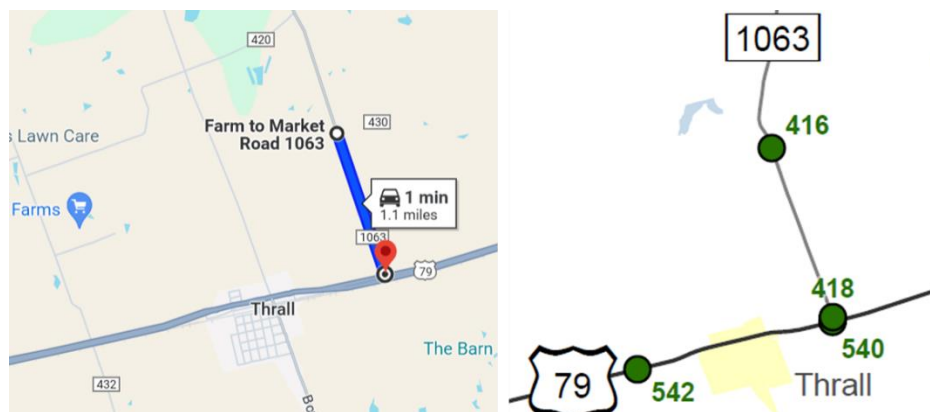


Figure 116: Reference map of TS2 section #1 in Williamson County. (Left) Section on Google Maps, and (right) section on TRM map.

The seventh site **Figure 117** is located in the Austin District, within Williamson County, between the cities of Coupland and Taylor along SH 0095 between TRMs 428 and 430.

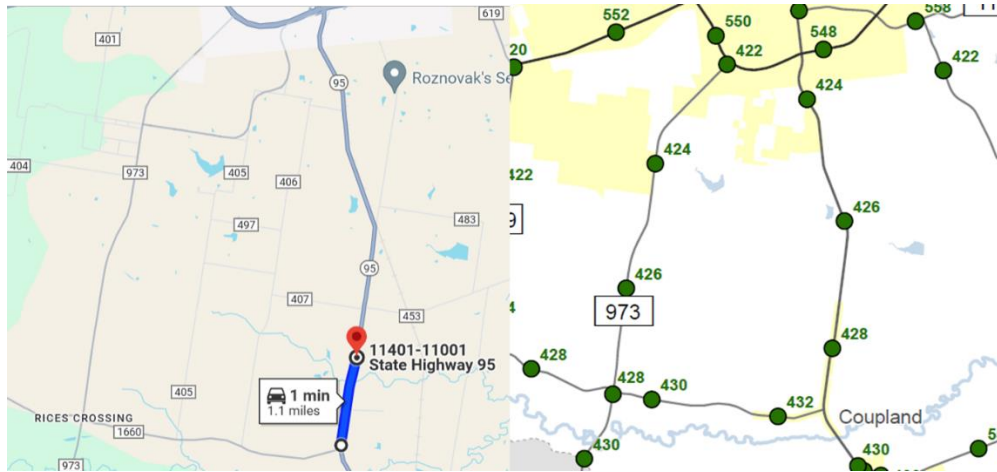


Figure 117: Reference map of TS2 section #2 in Williamson County. (Left) Section on Google Maps, and (right) section on TRM map.

The eighth site (**Figure 118**) is located in the Austin District, within Bastrop County, northwest of the unincorporated community of Paige along US 0290 between TRMs 626 and 622.



Figure 118: Reference map of TS2 section #3 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

6.1.3. Test Set #3 (TS3): Repeated Measurements

The ninth site (**Figure 119**) is located in the Austin District, within Travis County, in the City of Austin along SH 0001 between Texas Reference Markers (TRM) 440 and 438.



Figure 119: Reference map of TS3 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The tenth site (**Figure 120**) is located in the Laredo District within the La Salle County east of the city of Cotulla along SH 0097 between TRMs 488 and 450.

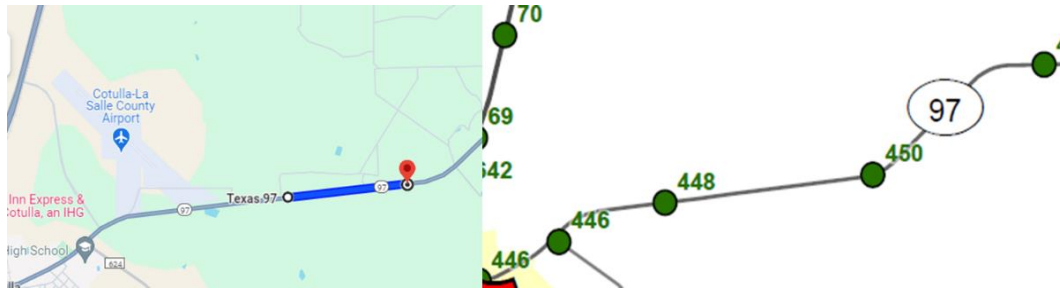


Figure 120: Reference map of TS3 section #2 in La Salle County. (Left) Section on Google Maps, and (right) section on TRM map.

The 11th site (**Figure 121**) is located in the Laredo district within Maverick County north of the city of Eagle Pass along FM 1589 between TRMs 548 and 546.

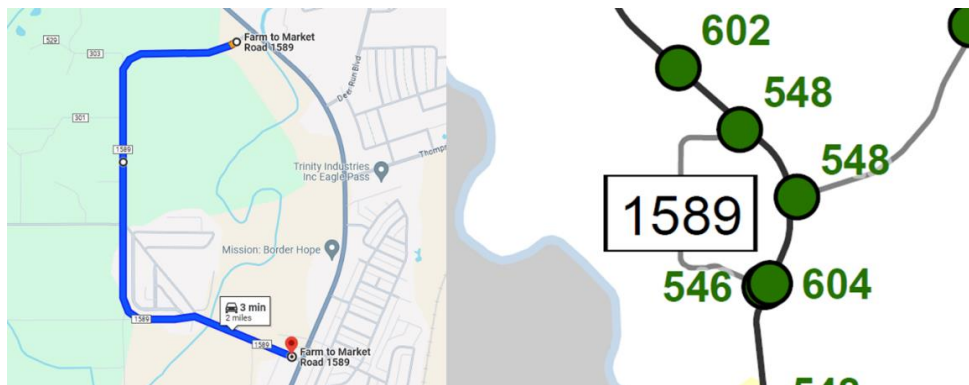


Figure 121: Reference map of TS3 section #3 in Maverick County. (Left) Section on Google Maps, and (right) section on TRM map.

The 12th site (**Figure 122**), and final repeated measurement site, is located in the Austin district within Travis County west of the city of Cedar Park along FM 1431 between TRMs 550 and 552.

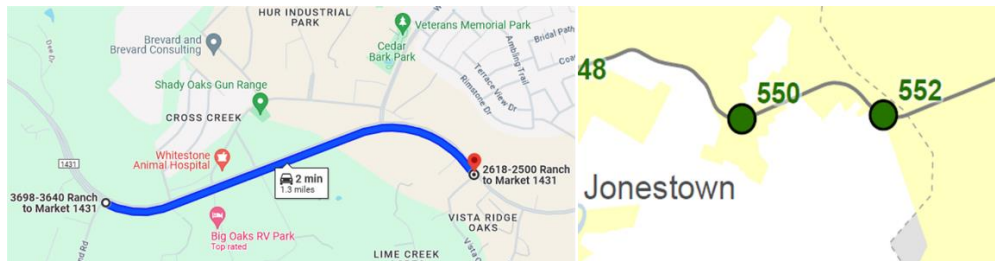


Figure 122: Reference map of TS3 section #4 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

6.1.4. Test Set #4 (TS4): Opposite Direction of Traffic

The 13th site (**Figure 123**) is located in the Austin District, within Travis County, in the City of Austin along IH 0035 between Texas Reference Markers (TRM) 234 and 235.

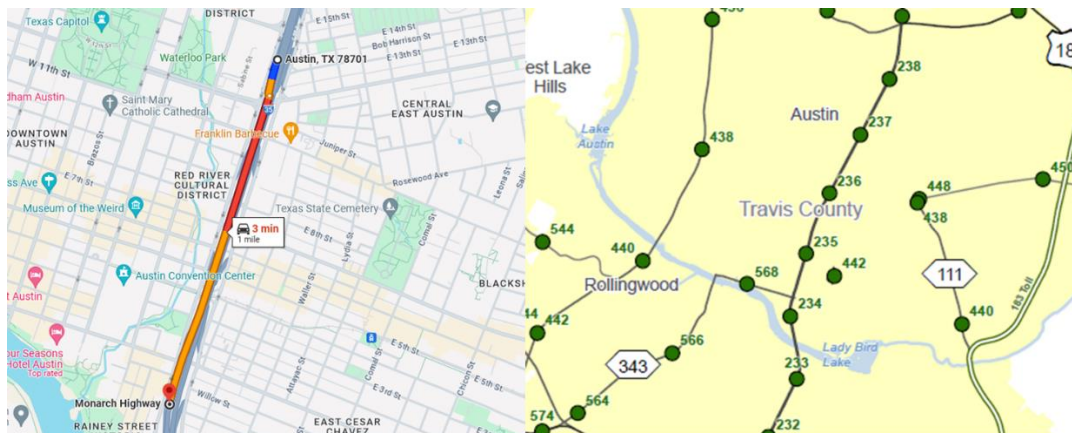


Figure 123: Reference map of TS4 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The 14th site (**Figure 124**) is also located in the Austin District within Travis County between the unincorporated communities of Fentress and Prairie Lea along SH 0080 between TRMs 478 and 480.

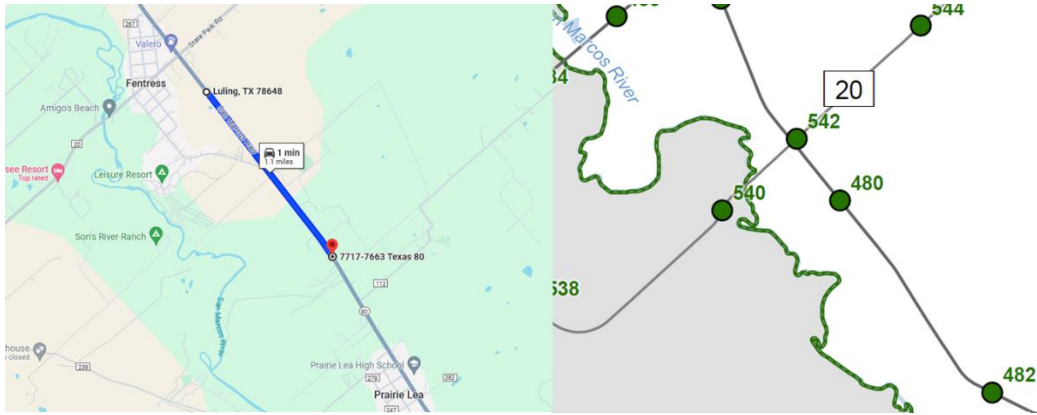


Figure 124: Reference map of TS4 section #2 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The 15th site (**Figure 125**) is located in the Laredo District, within the Webb County north of the Laredo Colombia Solidarity International Bridge along FM 1472 between TRMs 422 and 424.

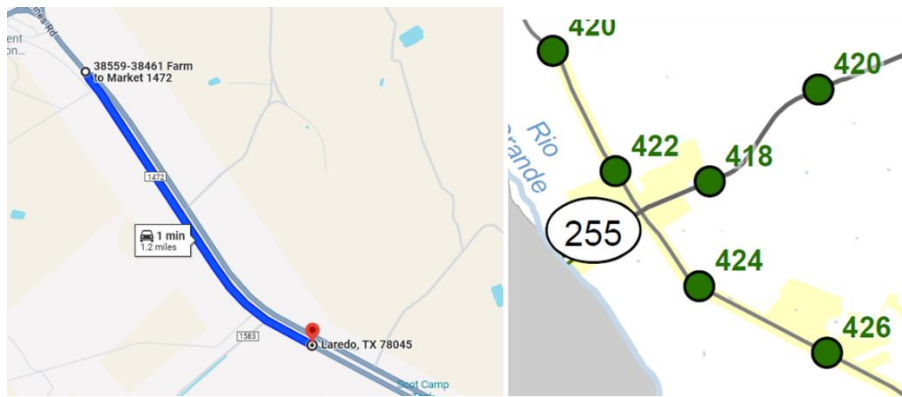


Figure 125: Reference map of TS4 section #3 in Webb County. (Left) Section on Google Maps, and (right) section on TRM map.

6.1.5. Test Set #5 (TS5): Different Lane

The 16th site (**Figure 126**) is located in the Laredo District, within the La Salle County, in the City of Cotulla along BI 0035C between TRMs 642 and 644.

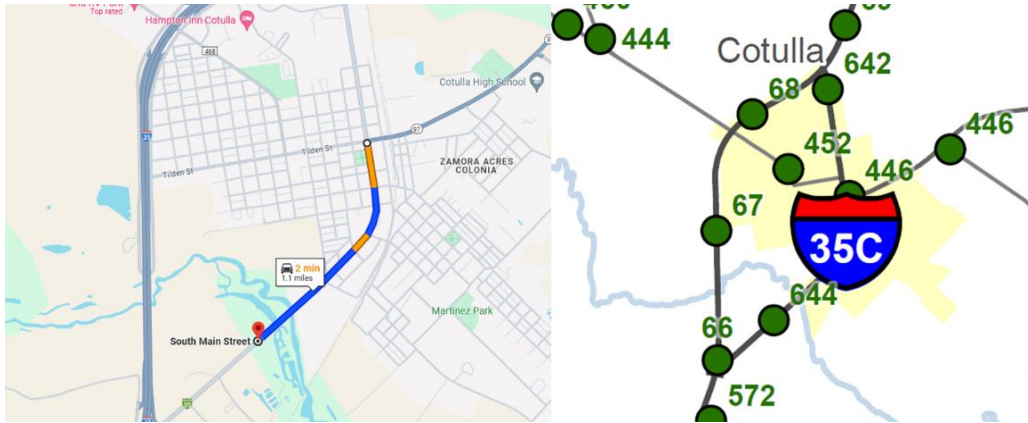


Figure 126: Reference map of TS5 section #1 in La Salle County. (Left) Section on Google Maps, and (right) section on TRM map.

The 17th site (**Figure 127**) is located in the Austin District, within the Travis County, southwest of the village of Briarcliff along SH 0071 between TRMs 550 and 552.



Figure 127: Reference map of TS5 section #2 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The 18th site (**Figure 128**) is located in the Austin District, within the Hays County, northwest of the village of Bear Creek along US 0290 between TRMs 560 and 562.

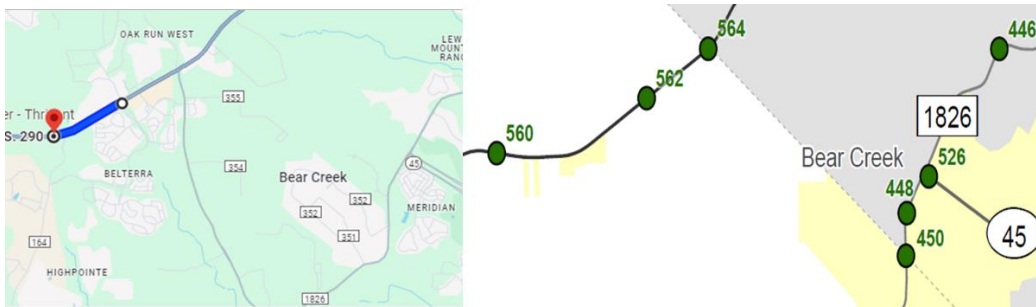


Figure 128: Reference map of TS5 section #3 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

6.1.6. Test Set #6 (TS6): Different Testing Day

The 19th site (**Figure 129**) is located in the Austin District, within Travis County, west of the city of Creedmoor along FM 1327 between TRMs 532 and 534.



Figure 129: Reference map of TS6 section #1 in Travis County. (Left) Section on Google Maps, and (right) section on TRM map.

The 20th and last site (**Figure 130**) is located in the Austin District, within Hays County, south of the city of Wimberley along FM 0032 between TRMs 520 and 522.



Figure 130: Reference map of TS6 section #2 in Hays County. (Left) Section on Google Maps, and (right) section on TRM map.

6.2. Results

This section presents the results obtained after applying the algorithms developed during Task #6 to all six testing sets (20 test sites) aforementioned. As per the recommendation given in Chapter 5, a picture classification model was used to predict Level 1 specificity of pavement surface type (Flexible vs Rigid), and an index-based model was used for all subsequent levels of specificity. The output of these algorithms is provided in a hierarchical manner, i.e., the first level of specificity first, followed by the second level, followed by third, and so on until the fifth level is reached. All predictions are separated by a hyphen such that the final prediction complies with the following notation:

LVL 1 Prediction – ... – LVL 5 Prediction

Recall that all levels of prediction, after the first level, used the predicted pavement type information (LVL 1 Prediction) to make subsequent predictions, thus is it highly unlikely that a pavement firstly classified as “Rigid” will be proceeded by a LVL 2 Prediction of “Hot Mix Asphalt.” However, that is the only nested dependency built in these models. It is possible for the algorithm to predict a “Hot Mix Asphalt” in LVL 2, and then predict a “Seal Coat” in LVL 3.

The Performing Agency has developed an abbreviated nomenclature for the predictions made at each level of specificity to simplify the prediction notation, summarized in Table 39.

Table 39 Abbreviated Nomenclature for Surface Prediction

Level 1	
F – Flexible	R - Rigid
Level 2	
HMA – Hot mix asphalt	PT – Plastic (Fresh) Concrete Texturing
ST – Surface treatment	HT – Hardened Concrete Texturing
Level 3	
DGM – Dense and Gap Graded Mixes	DT – Dragged Texturing
OGM – Open Graded Mixes	TT – Tined Texturing
HFST – High Friction Surface Treatment	EAT – Exposed Aggregate Texturing
SC – Seal Coat	NGDG – New Generation Diamond Grinding
	CDG – Conventional Diamond Grinding
Level 4	
TC – Gradation Type C	TRT – Transverse Tining
TD – Gradation Type D	LT – Longitudinal Tining
TOM – Thin Overlay Mixt	Ø - Prediction is not applicable*
PFC – Porous Friction Coarse	
BWC – Bonded Wearing Coarse	
G4 – Grade 4 Aggregates	
G4S – Grade 4S Aggregates	
G3 – Grade 3 Aggregates	
Level 5	
Ø - Prediction is not applicable*	RT - Random
	FT – Fixed Transform
	Ø - Prediction is not applicable*

* Certain surfaces cannot be specified beyond Level 3. Thus, their Level 4 and 5 predictions were assigned the null set symbol.

6.2.1. Test Set #1 (TS1): Unknown Surfaces

A summary of the prediction provided for all five sections in TS1 is presented in **Table 40**.

Table 40. Prediction made for pavement surfaces in TS1

TS1 Section	Classification Prediction
1	F-HMA-DGM-TD-Ø
2	R-PT-TT-DS-Ø
3	F-ST-SC-G3-Ø
4	R-PT-TT-TRT-RTT
5	R-PT-TT-TRT-FTT

The pavement in TS1 Section #1 appears to be a HMA based on a visual inspection (**Figure 131**). Specifically, a dense mix that has been cold milled in preparation for the application of a surface treatment. The age of the pavement is unknown. The prediction algorithm accurately predicted that this was a flexible pavement, and it is very likely that is, indeed, a hot mix asphalt. Furthermore, the algorithm proceeded to predict that the pavement is a dense or gap-graded mix with a gradation of Type D.

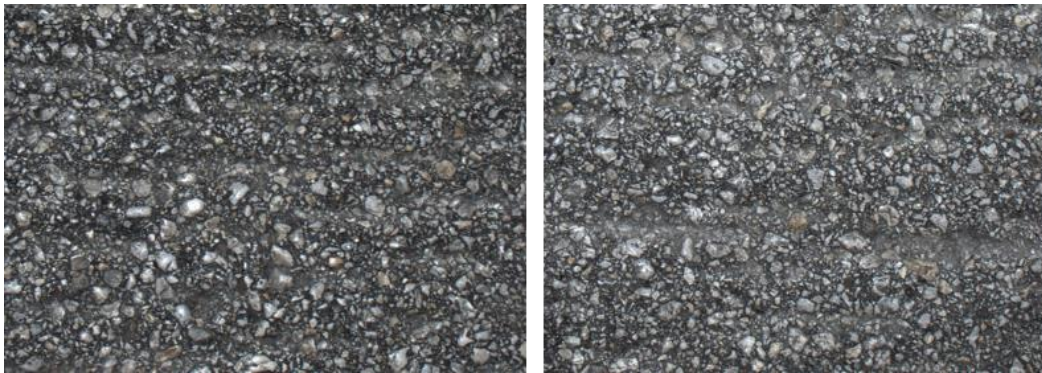


Figure 131: Pavement surface at TS1 Section #1.

The next pavement is TS1 Section #2 which appears to be a rigid pavement (**Figure 132**), of unknown age, with transverse tining (direction of traffic is orthogonal to the tining). The algorithm correctly predicted that the pavement was rigid, with a plastic texturing, but failed at specifying the type of tinning, instead the algorithms claims that the pavement looks more like a carpet drag pavement than a transverse tinned surface. This may be due to the fact that the pavement surface is already deteriorated, but it is still a misclassification at the Level 4 of specificity.

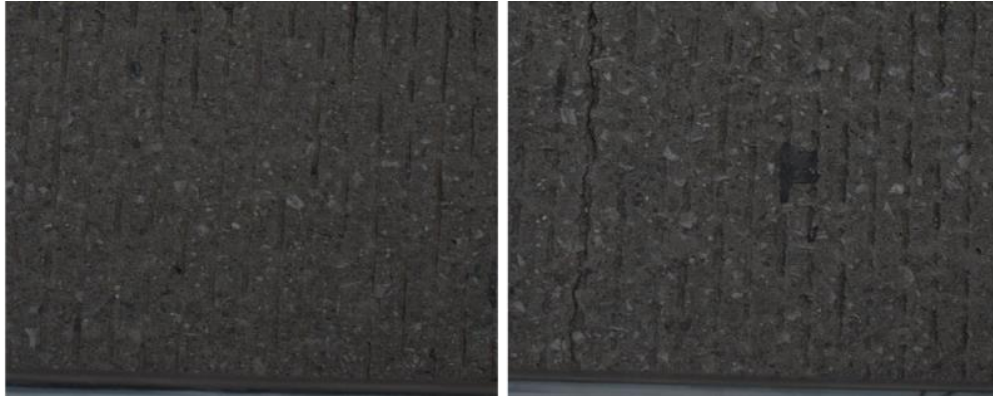


Figure 132: Pavement surface at TS1 Section #2.

The next pavement is TS1 Section #3 which appears to a surface treatment with signs of bleeding and unknown age (**Figure 133**). Note that in the hierarchy of specificity, a “bleeded asphalt pavement” is *not* one of the options. Thus, the algorithm correctly predicted that this pavement is a flexible pavement, and then it proceeded to infer that it was a surface treatment, specifically a grade 3 seal coat.

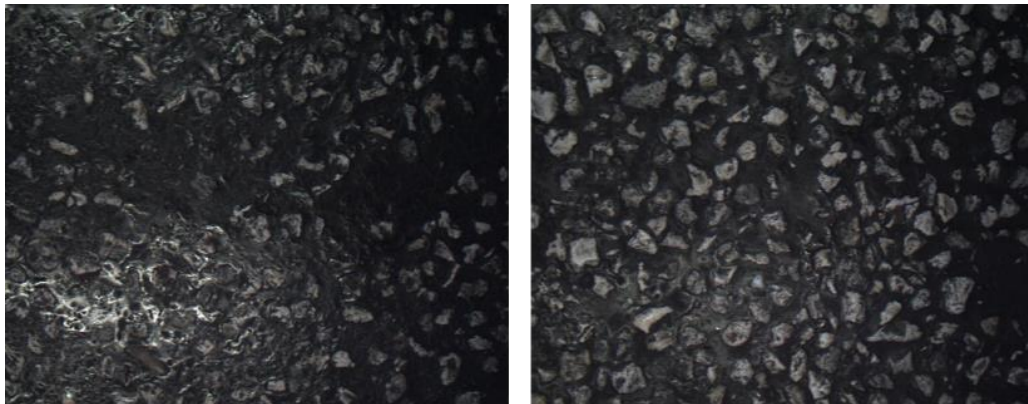


Figure 133: Pavement surface at TS1 Section #3.

The next site in this test set is TS1 Section #4, which appears to be a rigid pavement of unknown age with transverse tining (**Figure 134**). In this instance, the algorithm was able to accurately predict the surface type up to the Level 4 of specificity with complete accuracy. However, neither the Performing nor Receiving Agency had any records that indicated whether this pavement was originally designed as a fixed or random transverse tining. Thus, the classification on the last level of specificity could not be confirmed, but the algorithm prediction was random tining.

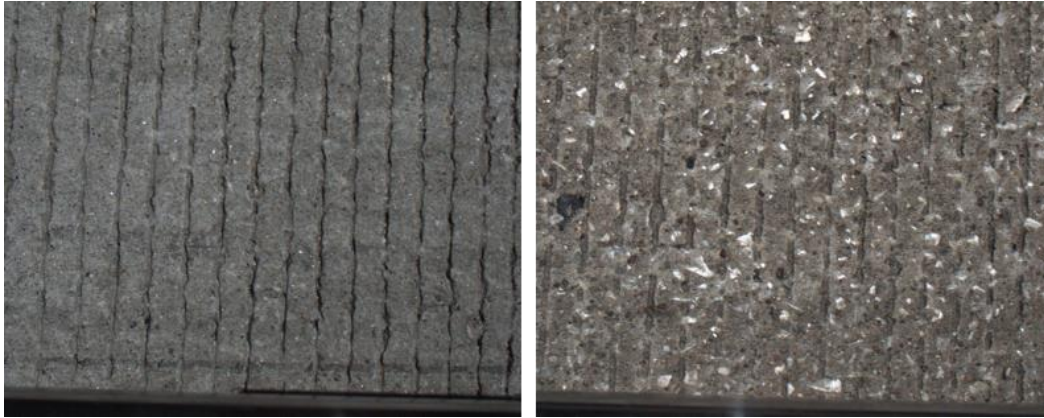


Figure 134: Pavement surface at TS1 Section #4.

The final test site was TS1 Section #5 which appears to be a rigid pavement of unknown age with a combination of carpet drag and transverse tining (**Figure 135**). Sites like this one were eliminated for the sake of simplicity and accuracy within the Ground Truth (refer to TM#5) data used in the model training phase. The algorithm was able to output a prediction up until the fifth level of specificity. It predicted that the pavement was rigid, with plastic texturing, transverse tining and that the tining was fixed. From **Figure 135**, it appears to indicate that for this surface, even though both carpet dragging and tining can be seen in the pictures, the transverse tining is more prevalent texture wise.

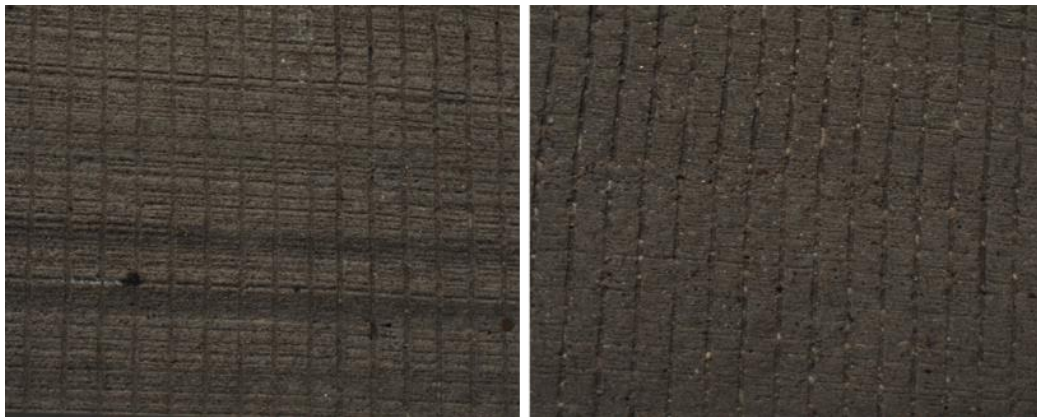


Figure 135: Pavement surface at TS1 Section #5.

In summary, the algorithm was able to learn patterns in the pavement surface that would agree with visual inspection up until the Level 3 of specificity. Once Level 4 specificity prediction is required, there is a chance that the algorithm prediction would not match the expected pavement surface type.

6.2.2. Test Set #2 (TS2): Previous Classified Pavements by Experts - TxDOT Project No. 0-7031

A summary of the prediction provided for all three sections in TS2 is presented in **Table 41**.

Table 41. Predictions made for pavement surfaces in TS2

TS2 Section	Classification Prediction	Expert Classification
1	F-ST-SC-G4-Ø	High Macrotexture Chip Seal
2	F-HMA-DGM-TC-Ø	Dense Coarse Mix
3	F- HMA -OGM-PFC-Ø	Porous Friction Coarse

The pavement at TS2 Section #1 appears to have a surface treatment and the researchers claimed that the surface type was a high macrotexture chip seal (**Figure 136**). The prediction made by algorithm appears to partially agree with the assessment that the surface is indeed a seal coat, but in this instance the algorithm predicts the gradation is of Grade 4. It is unknown whether that final prediction about the gradation of the seal coat is correct, but the prediction complied with the experts' classification up until Level 2 specificity.

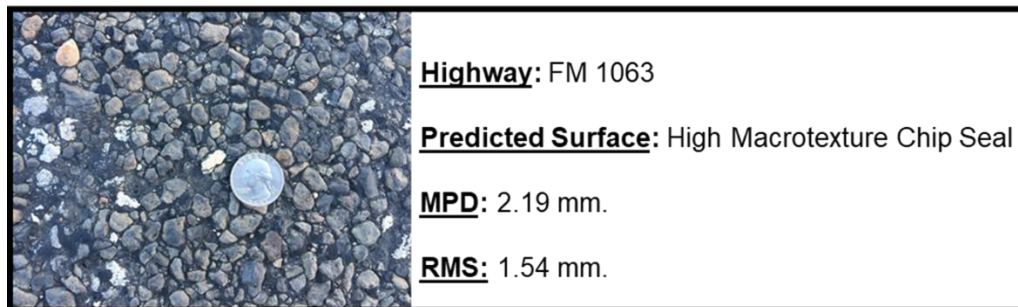


Figure 136: Pavement surface at TS2 Section #1 (3).

The next pavement is TS2 section #2. This site appears to be a HMA dense-graded mix, and it was previously classified as a dense coarse mix (**Figure 137**). The prediction by the algorithm agreed with the experts' opinion as it predicted that the surface was an HMA, specifically of dense or gap-graded mix with a Type C gradation, which is the coarsest gradation (in terms of nominal maximum aggregate size) that can be predicted for dense- and gap-graded mixes given the training data used to develop the model.

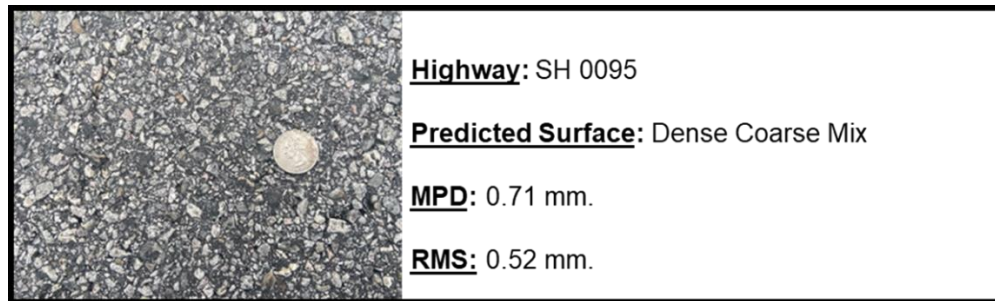


Figure 137: Pavement surface at TS2 Section #2 (3).

The final pavement site is TS2 Section #3 which appears to be an HMA open-graded mix, and it was previously classified as a PFC (**Figure 138**). Once again, the algorithm agreed perfectly with the experts' classification as it predicted HMA, specifically an open-graded mix and a PFC. This accuracy in prediction was expected since the picture shows a well-maintained PFC surface, and those pavement surfaces have a very characteristic texture.



Figure 138: Pavement surface at TS2 Section #3 (3).

In conclusion, the algorithm effectively replicates the experts' classification on pavement surface type for the three samples collected from TxDOT Project 0-7031. This promising outcome highlights that for pavements near Austin, where extensive data has been gathered, the algorithm's predictions closely align with those of experts.

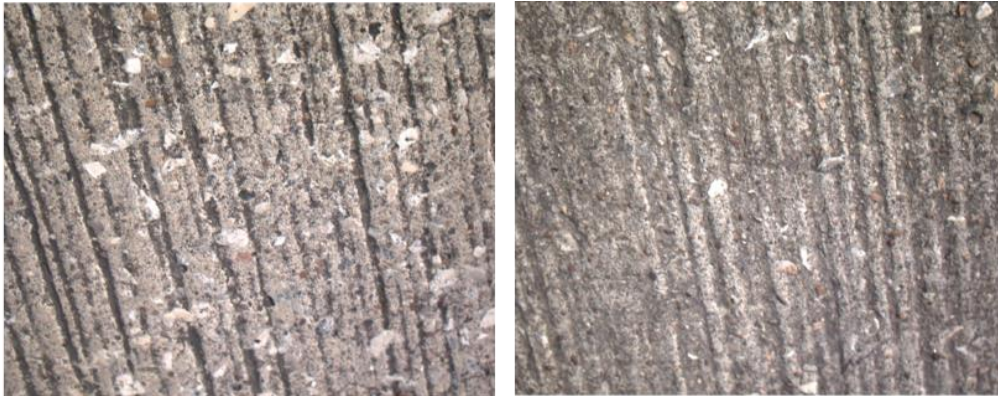
6.2.3. Test Set #3 (TS3): Repeated Measurements

A summary of the prediction provided for all four sections in TS3 is presented in **Table 42**.

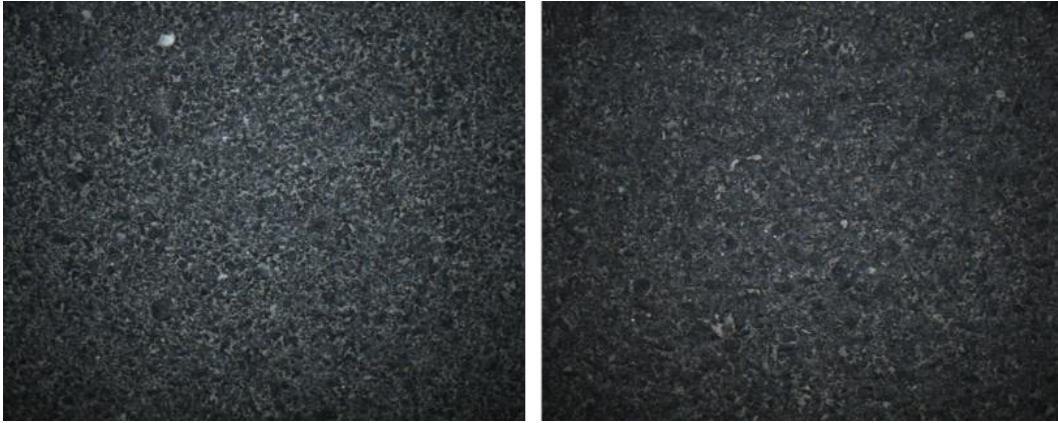
Table 42. Predictions made for pavement surfaces in TS3

TS3 Sections	Run	Classification Prediction	True Surface
1	Run #1	R-PT-TT-DS-Ø	<i>Longitudinal Tining</i>
	Run #2	R-PT-TT-DS-Ø	
2	Run #1	F-HMA-DGM-TOM-Ø	<i>SMA</i>
	Run #2	F-HMA-DGM-TOM-Ø	
3	Run #1	F-ST-SC-G3-Ø	<i>Seal Coat</i>
	Run #2	F-ST-SC-G3-Ø	
4	Run #1	F-HMA-OGM-PFC-Ø	<i>PFC</i>
	Run #2	R-HMA-OGM-PFC-Ø	

The pavement at site TS3 #1 was identified as a rigid pavement with longitudinal tining of an unknown age (**Figure 139**). The prediction across both runs were identical; however, the algorithm was not able to distinguish between a carpet drag texturing and a longitudinal tining. The mistake in prediction was made at the fourth level of specificity.

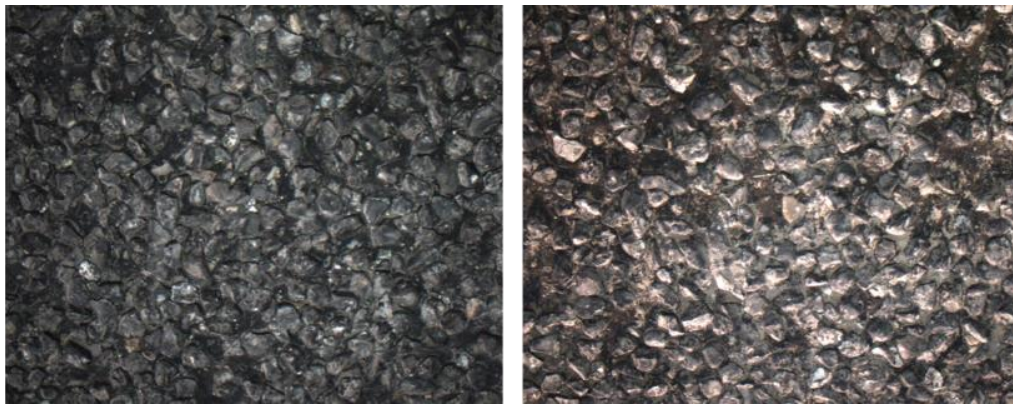
**Figure 139: Pavement surface at TS3 section #1.**

The pavement at site TS3 #2 was denoted as a stone matrix asphalt (SMA) of at most four years (**Figure 140**). The prediction across both runs were identical and the algorithm was able to correctly predict that the pavement was a dense- or gap-graded mix. However, given that it is difficult to distinguish between the different types of HMA using only surface information about the pavement, the algorithm was not able to explicitly specify it as a SMA. Instead, it did its best to assess the mix's gradation and estimates that the pavement has a gradation equivalent to that of a TOM, which are usually within the dense gradation domain.



***Figure 140:** Pavement surface at TS3 Section #2.*

The pavement at TS3 #3 was denoted as a seal coat made out of crushed stone with unknown age (**Figure 141**). The prediction across both runs were identical, and the algorithm was able to correctly predict that the pavement was a seal coat. Moreover, the algorithm estimated that the was a Grade 3, which is the gradation with largest aggregates used to train the algorithm.



***Figure 141:** Pavement surface at TS3 Section #3.*

The final site, TS3 #4, was denoted as a PFC (**Figure 142**) and three years old, according to its work history. The prediction across both runs were identical and the algorithm was able to correctly predict all levels of specificity up to the fourth level. Once again, the PFCs in good condition were relatively straightforward to classify.

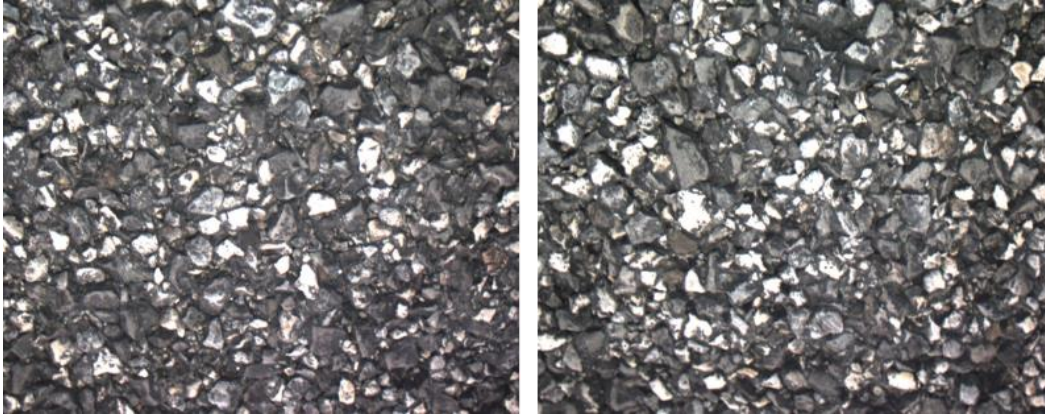


Figure 142: Pavement surface at TS3 Section #4.

In summary, the algorithm provided consistent predictions whenever the pavement section was measured multiple times along the same lane and direction of traffic.

6.2.4. Test Set #4 (TS4): Opposite Direction of Traffic

A summary of the prediction provided for all three sections in TS4 is presented in **Table 43**.

Table 43. Predictions made for pavement surfaces in TS4

TS4 Sections	Direction of Traffic	Classification Prediction	True Surface
1	Northbound	F-HMA-DGM-PFC*-Ø	<i>SMA</i>
	Southbound	F-HMA-DGM-PFC**-Ø	
2	Southeast	F-HMA-DGM-TOM-Ø	<i>TOM</i>
	Northwest	F-HMA-DGM-TOM-Ø	
3	Southeast	F-ST-SC***-G3S ^ψ -Ø	<i>Seal Coat</i>
	Northwest	F-ST-SC-G4-Ø	

*The majority of 0.1-mile sections were predicted as PFC and a few were predicted to be TOMs.

**The majority of 0.1-mile sections were predicted as PFC and a few were predicted to be BWCs.

*** The majority of 0.1-mile sections were predicted as SC but a few were predicted as DGM.

^ψ The majority of 0.1-mile sections were predicted as G3S but a few were predicted as G4.

The pavement at site TS4 #1 was denoted as a SMA (**Figure 143**) pavement with an approximate age of six years, along both directions of traffic. The algorithm was able to provide consistent predictions across both directions of traffic, but only because a majority voting system was implemented to decide what type of pavement was prevalent across the mile-long sections that were tested. On the northbound approach, three out of the ten 0.1-mile-long sections were classified as TOM at the fourth level of specificity, potentially indicating the presence of patches or deteriorated pavement. Along the southbound approach, four out of the ten pavement sections were classified as BWCs. This type of mix is similar to a PFC

as they are both open-graded, but in general, BWCs tend to have lower MPDs. Again, this could be an indication of deterioration or clogging of the PFC.

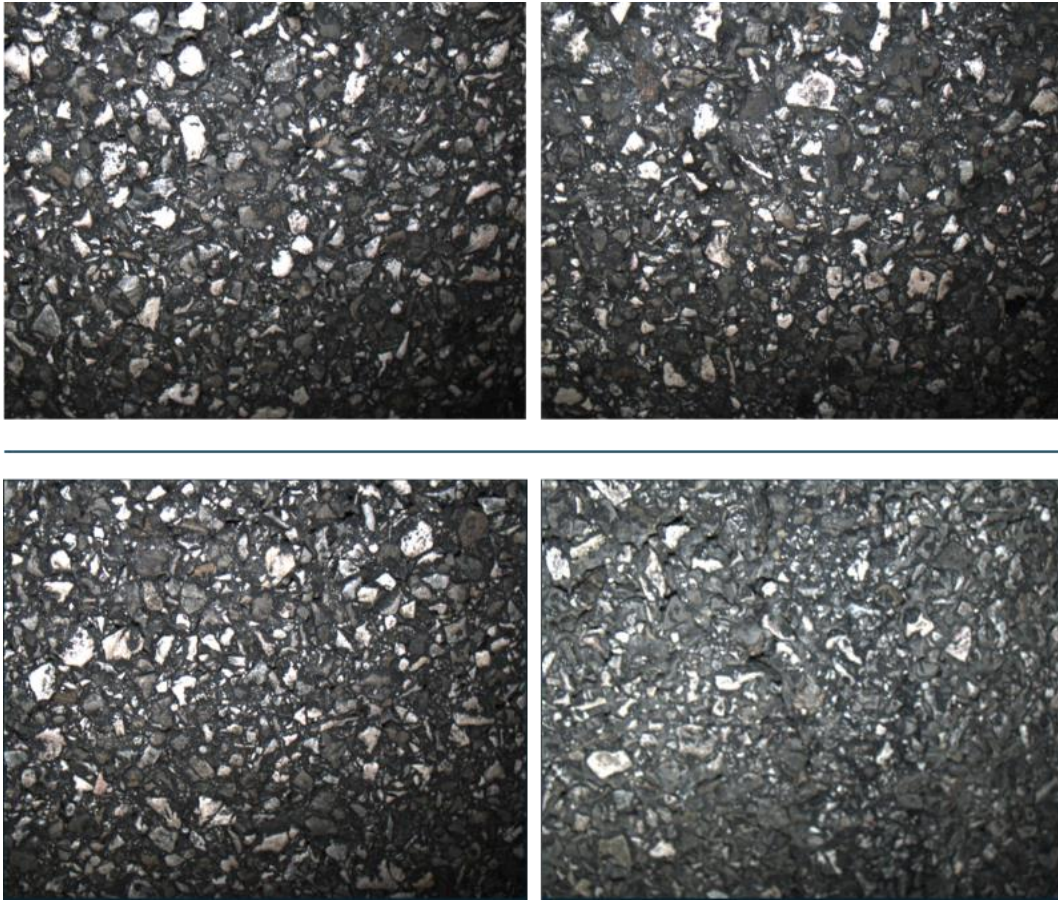


Figure 143: *Pavement surfaces across both directions of traffic: (top) southbound approach, (bottom) northbound approach in TS4 Section #1.*

The pavement at site TS4 #2 was denoted as a TOM (**Figure 144**) with an approximate age of five years, along both directions of traffic. Here the predictions made across both directions of traffic were identical and completely correct. This is a good indication given that a visual inspection of the pavement of the site would give the impression that the pavement in one direction of traffic was older than the other, due to the grayer color that resembled asphalt binder oxidation aging.

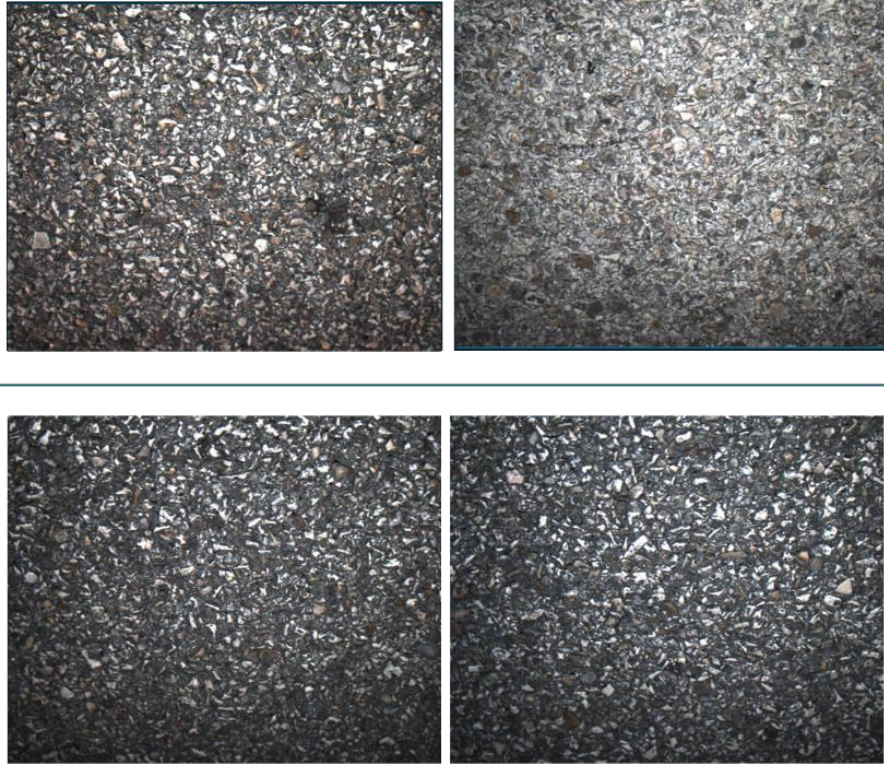


Figure 144: Pavement surfaces across both directions of traffic: (top) southeast approach, (bottom) northwest approach in TS4 Section #2.

Finally, the pavement at site TS4 #3 was denoted as a grade 3 seal coat (**Figure 145**) with an approximate age of three years, along both directions of traffic. Prior to discussing the results, it should be noted that the southeast approach had patches along the inner wheel path where the measurements were taken, as can be seen in one of the pictures in **Figure 145**. The predictions made across both directions of traffic were only similar up to the third level of specificity and, even then, there were differences in the results.

The southeast approach had four out of the 11 pavement sections classified as dense- or gap-graded mixes, whereas the northwest approach classified consistently as a seal coat. This difference can be an indication of the presence of multiple asphalt patches along the wheel patch, but also of the chip seal degradation. Furthermore, at the fourth level of specificity, the southeast approach was classified as a grade 3S seal coat (with a few sections classifying as grade 4), whereas the northwest approach was classified as a grade 4 seal coat consistently. Unfortunately, no proper assessment can be made about the fourth level of specificity, as the exact gradation at this site is unknown to the Performing Agency.

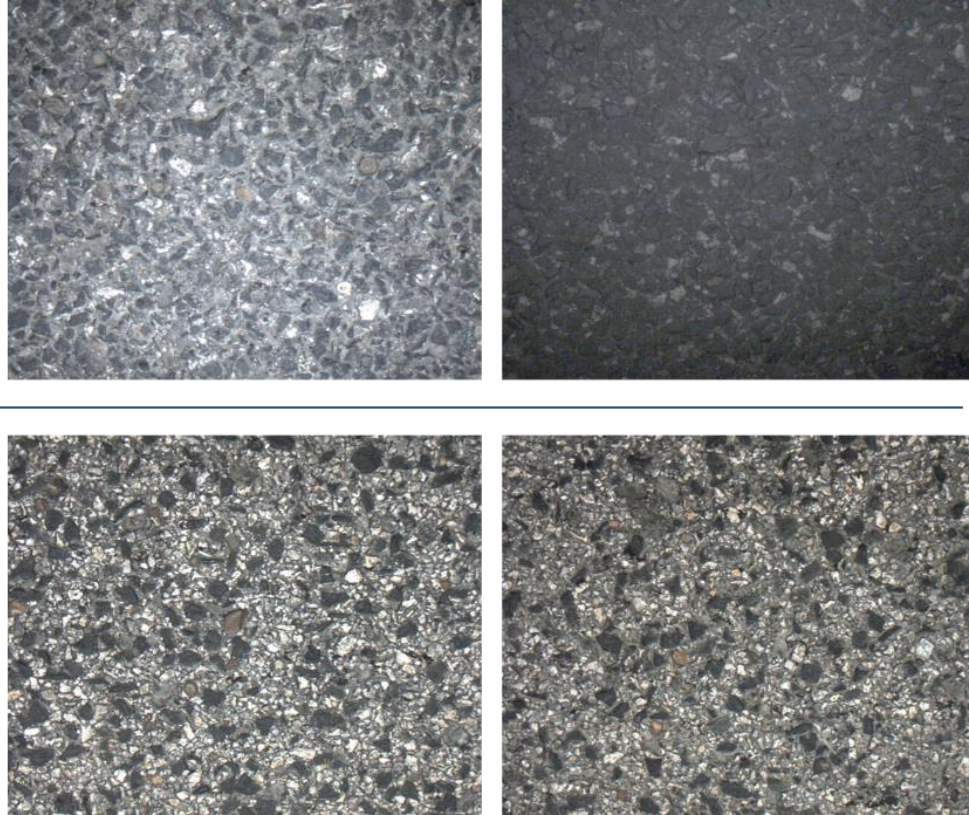


Figure 145: Pavement surfaces seen across both directions of traffic: (top) southeast approach, (bottom) northwest approach in TS4 Section #3.

In summary, the Performing Agency concludes that it is unlikely that surface predictions made across different directions of traffic will always be identical. There were instances in the field survey where one direction of traffic has one type of pavement and the other direction has different one, but even in cases where the surface type was the same, it is still possible for one direction to be more deteriorated relative to the other. Hence the generalization of pavement surface made by assessing only one of the directions should be done carefully, and always rely on other available information.

6.2.5. Test Set #5 (TS5): Different Lane

A summary of the prediction provided for all three sections in TS5 is presented in **Table 44**.

Table 44. Predictions made for pavement surfaces in TS5

TS5 Sections	Lane Measured	Classification Prediction	True Surface
1	Inner	F-HMA-DGM-PFC-Ø	<i>SMA</i>
	Outer	F-HMA-DGM-PFC-Ø	
2	Inner	F-HMA-OGM-PFC-Ø	<i>PFC</i>
	Outer	F-HMA-OGM-PFC-Ø	
3	Inner	F-HMA-OGM-PFC-Ø	<i>PFC</i>
	Outer	F-HMA-OGM-PFC-Ø	

The pavement at site TS5 #1 was denoted as a HMA (**Figure 146**) of unknown age. The algorithm was able to provide consistent predictions along the inner and outer lanes. However, those predictions were only accurate up until the third level of specificity. At the fourth level of specificity the algorithm predicted that the surface was a PFC, instead of attempting to predict the gradation of the surface.

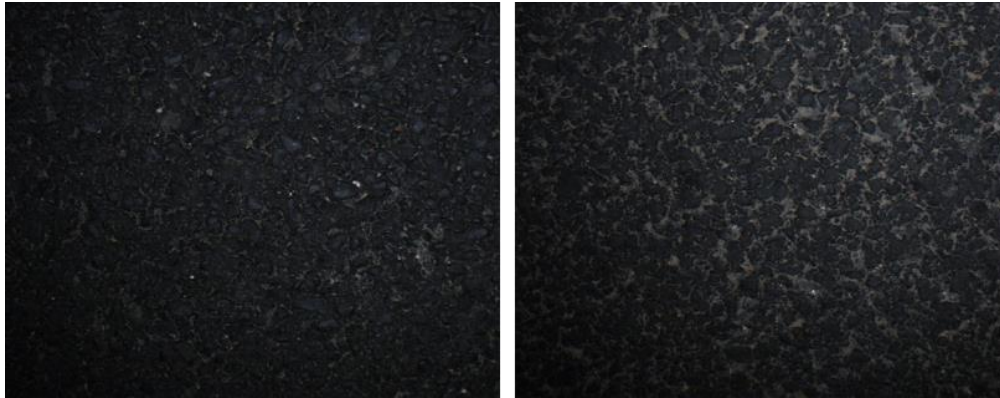


Figure 146: Pavement surface at TS5 Section #1.

The pavement at site TS5 #2 was denoted as a PFC (**Figure 147**) with an approximate age of five. The algorithm was able to correctly classify this pavement surface and also provide consistent predictions along both lanes in the site.

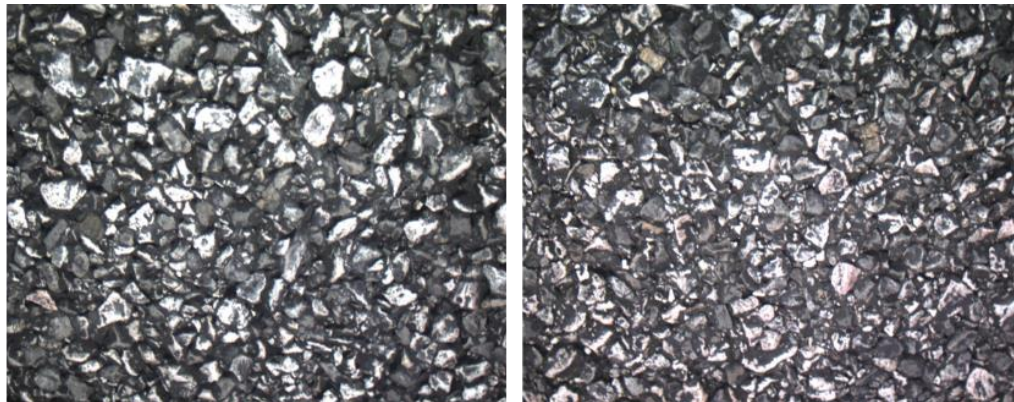


Figure 147: Pavement surface at TS5 Section #2.

The pavement at site TS5 #3 was also denoted as a PFC (**Figure 148**) with an approximate age of five years. Similar to the previous site, the algorithm was able to correctly classify this pavement surface and provide consistent predictions along both lanes in the site. Again, PFCs appeared to be the easiest surface to classify.

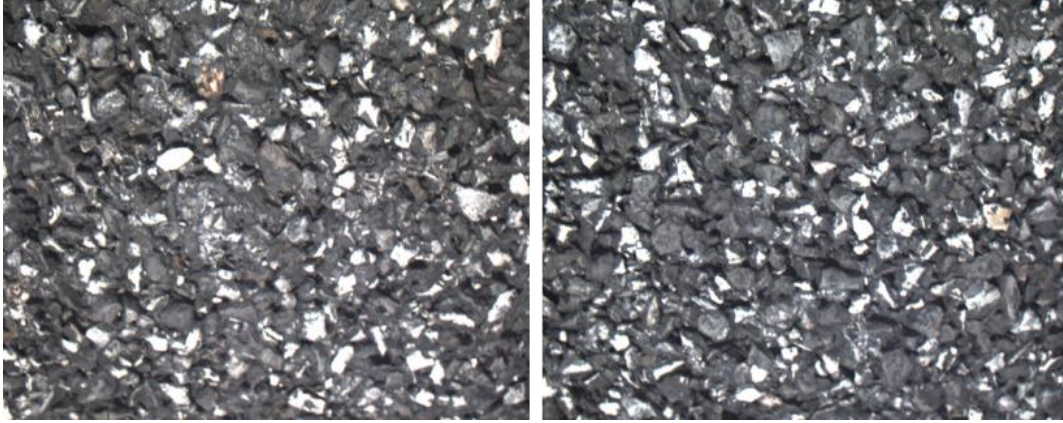


Figure 148: Pavement surface at TS5 Section #3.

In summary, it appears that the algorithm is accurate in predicting pavement surface type up to the third level of specificity in this testing set, while also being always consistent. That is in part a result of the sampling method conducted by the Performing Agency, as these sites were chosen precisely because the pavement was homogenous throughout their length. Thus, provided that the pavement is the same across the lanes, predictions will not be different between different lanes.

6.2.6. Test Set #6 (TS6): Different Testing Day

A summary of the prediction provided for all two sections in TS6 is presented in **Table 45**.

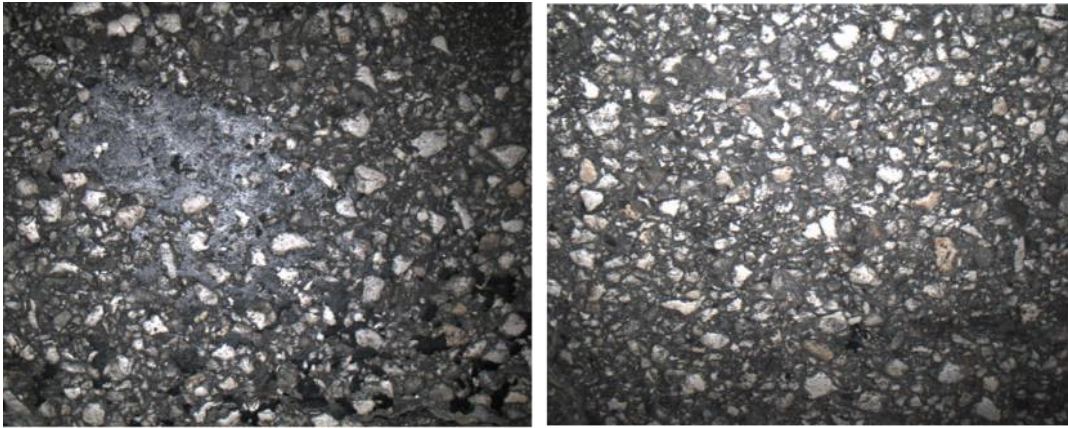
Table 45. Predictions made for pavement surfaces in TS6

TS6 Sections	Time Measured	Classification Prediction	True Surface
1	Initial Day	F-HMA-OGM- BWC*-Ø	<i>TBWC</i>
	One Day After	F-HMA-OGM-BWC-Ø	
2	Initial Day	F-HMA-DGM-TOM-Ø	<i>TOM</i>
	Three Days After	F-HMA-DGM-TOM-Ø	

*The majority of 0.1-mile sections were predicted as BWC and a few were predicted to be G3S.

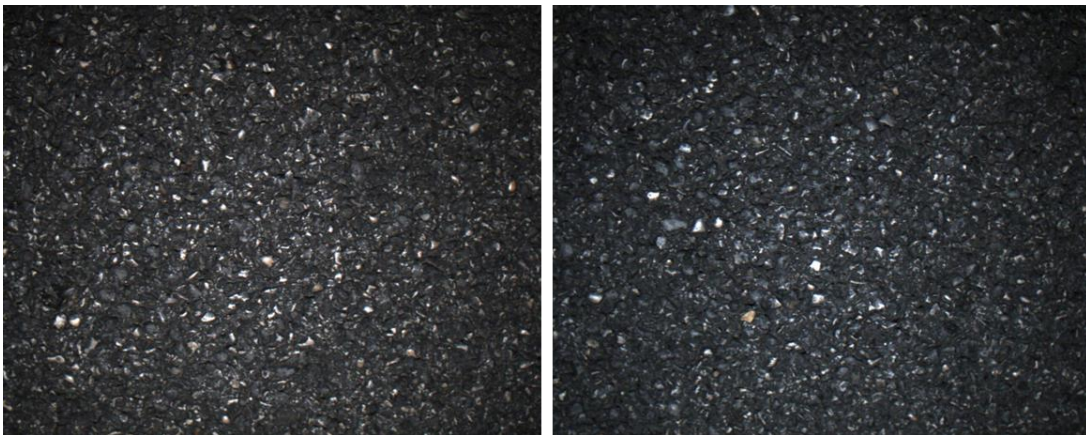
The pavement at this location was denoted as a TBWC (**Figure 149**) with an approximate age of three years. The algorithm was able to produce similar prediction results on this site after a full day elapsed. The only difference occurred at the section level, where two of the ten pavement sections were classified as a seal coat of gradation 3S at the fourth level of specificity. Other than that, the prediction

was consistent and correct according to information received from personnel in the Receiving Agency.



***Figure 149:** Pavement surface at TS6 Section #1.*

The pavement at this location was denoted as a TOM (**Figure 150**) with an approximate of three years. The algorithm was able correctly and consistently predict the pavement surface type for this site when testing was spaced three days apart. Further, it appears that TOM surfaces were easy to classify in the sample collected for testing purposes, even though when the model was being calibrated these surfaces did not have an F1 score above 90%.



***Figure 150:** Pavement surface at TS6 Section #2.*

In summary, it appears that even when measurements are taken a few days apart, the prediction of the surface type remains consistent. This is explained by the fact that, unless an extreme weather event happens or a specific usage of the road such as the case of one pass of a super-heavy truck, it is unlikely that a pavement surface will deteriorate enough over the span of a few days to completely change its surface texture characteristics to the point where it would be classified differently.

Chapter 7. Conclusions and Recommendations

7.1. Conclusions

During this study, a comprehensive evaluation and development of machine learning models for pavement surface classification were undertaken to enhance accuracy and reliability using advanced computational techniques. An initial detailed literature review on pavement texture and measurement methods provided a solid foundation for developing a prototype system. This prototype marked a significant advancement in capturing high-resolution texture data at highway speeds, integrating innovative technologies like laser scanners, high-speed cameras, and lighting systems to meet rigorous evaluation conditions.

The data collection phase of this project covered 425.5 miles of pavement across Texas, including 313.7 miles of flexible and 111.8 miles of rigid pavements. Over 50,000 high-resolution images and detailed texture profiles were gathered from 15 distinct types of flexible pavements and seven types of rigid pavements. This diverse dataset, crucial for training and validating machine learning models, has also been made available to the Receiving Agency via Product #1 (P1), ensuring future research and continuous advancements in pavement surface classification.

In terms of the unsupervised machine learning analysis conducted, the clustering results demonstrated that texture indices could effectively classify various pavement surfaces. Moreover, the Performing Agency identified Partition Around Medoids (PAM) as the best clustering method due to its robustness to noise and outliers. For flexible pavements, five distinct clusters were found, each representing unique textural properties. These included pavements with the highest macrotexture, coarse well-graded surfaces, high macrotexture with uniform gradation, lowest macrotexture pavements, and coarse surfaces with uniform gradation. For rigid pavements, seven clusters were identified: new generation diamond grinding, conventional diamond grinding, exposed aggregates, dragged surfaces, tining with shallow grooves and short landing spacing, tining with deep grooves and wide landing spacing, and significantly polished tined surfaces .

In terms of classification, the Performing Agency performed a comprehensive qualitative analysis of each machine learning algorithm to be tested, and developed a hierarchical classification method for pavements, organized by levels of specificity. At the broadest level (Level 1), the distinction is made between flexible and rigid pavements. Subsequent levels provide more detailed classifications, such as mix/surface type, gradation, and texturing orientation. Comprehensive testing and model tuning revealed that picture-based models (PBC) excelled at predicting

whether a pavement is flexible or rigid, achieving an accuracy of 0.995, compared to 0.935 for index-based models (IBM).

For Levels 2 through 4, the IBMs demonstrated superior performance, provided the predicted pavement type was also included as an input. At Level 2, the best model was the Stochastic Neural Network (SNN) with an F1 score of 0.980. At Level 3, the Support Vector Machine (SVM) model achieved an F1 score of 0.962. At Level 4, the K-Nearest Neighbor (KNN) model led with an F1 score of 0.887. The study also noted that Gradient Boosting (GB) and Extreme Gradient Boosting (XBG) models were less effective beyond the first level of specificity, and Decision Trees (DT) were inadequate past the second level.

At Level 5, which involved determining the orientation of tining (transverse or longitudinal), PBC models vastly outperformed IBC models, achieving an exceptional F1 score of 0.999. This hierarchical approach and the detailed performance metrics of various models underscore the efficacy of combining texture indices and image data to achieve precise and specific pavement surface classification, thereby enhancing pavement management and maintenance strategies.

The classification results indicated that the best joint classifier uses PBC to predict levels 1 and 5 of specificity, while levels 2 through 4 are best predicted using IBCs. A rigorous evaluation and validation of the joint classifier ensued, involving six diverse test sets across 20 test sites. The unknown surfaces (TS1) test set, comprising pavement sections not included in the training database, demonstrated the model's ability to generalize to new data with high accuracy and minimal errors. The previously classified pavements (TS2) test set confirmed the model's reliability by comparing expert classifications with model predictions. Repeated measurements (TS3) assessed the model's consistency, indicating robustness in repeated evaluations. The opposite direction of traffic (TS4) test set showed the model's capability to generalize across different directional data, achieving consistent performance with minimal variance. The different lane (TS5) test set demonstrated the model's ability to handle lane-specific variations, consistently classifying surfaces accurately across different lanes. Lastly, the different testing day (TS6) test set assessed temporal consistency, with the model's predictions remaining stable over short time spans (one to three days). Overall, the validation process confirmed the models' robustness and effectiveness in various real-world conditions.

The research demonstrated that integrating picture-based models and index-based classifiers significantly enhances the accuracy and reliability of pavement surface classification. Picture-based models excelled in broad and fine-grained

classifications, while index-based classifiers effectively managed intermediate specificity levels. Rigorous validation across diverse test sets confirmed the models' robustness and real-world applicability, ensuring consistent performance under various conditions. These findings have important implications for pavement management systems, enabling more accurate and efficient maintenance strategies. Leveraging advanced machine learning techniques, the developed models can improve road safety by using surface classification in empirical models to predict skid resistance, optimize resource allocation by better understanding pavement types across the network, and enhance pavement surface analysis by providing detailed information of pavement surfaces. This supports future research and practical applications in the field.

7.2. Recommendations

The research highlighted several areas for future work to further enhance pavement surface classification and management. One key area is developing systems capable of real-time data acquisition and processing, which will provide immediate insights and feedback during pavement inspections, allowing for timely maintenance decisions. Additionally, investing in advanced laser technologies with higher resolution and sampling rates is also crucial, as these improvements will enable the capture of finer surface details, including microtexture, thus enhancing the predictive power of the models.

On the modeling side, exploring the use of 3D texture statistics can significantly enhance classification accuracy, especially for complex surfaces that demand detailed analysis. Additionally, expanding the study to encompass a wider variety of pavement types, such as concrete finishes like sand blasting, newly developed pavement materials, and obtaining larger samples of surfaces with limited observations, such as high friction surface treatments, will increase the models' versatility. Future research should also involve alternative classification approaches beyond the hierarchical method used in this study, or more innovative machine learning classifiers, allowing for a comparative analysis of results to identify the most effective techniques.

Finally, the Performing Agency strongly encourages the application of these pavement classification techniques in practical scenarios, such as the segmentation of highways. This involves allowing data collection equipment to operate continuously over long distances, enabling precise identification of where pavement changes occur. These changes can include transitions from one type of pavement mix to another or shifts from a high-quality pavement mix to a heavily deteriorated one, even if the mix remains the same. Implementing these techniques

will enhance the ability to monitor and manage pavement conditions more effectively.

Appendix A: Value of Research (VoR)

An analysis of the economic benefits brought about by TxDOT Research Project 0-7139 is explained in this appendix. Investing in the research and development of advanced machine learning models for pavement surface classification can yield significant financial benefits for the Receiving Agency. By enhancing the accuracy and reliability of pavement surface classification, this research project has the potential to save millions of dollars annually in maintenance costs, extend the lifespan of pavement infrastructure, and improve road safety, thereby reducing accident-related expenses.

Review of Potential Savings

Cost Savings in Maintenance

One of the primary financial benefits of this research is the significant reduction in maintenance costs. Traditional methods of pavement inspection and maintenance are labor-intensive and often require frequent inspections to identify issues. By implementing machine learning models that accurately classify pavement surfaces and predict deterioration, the Receiving Agency can transition to a more proactive maintenance approach. This shift from reactive to proactive maintenance can result in substantial cost savings.

For example, a study by Michigan State University estimates that every dollar spent on preventive maintenance can save between \$4 and \$10 in future repair costs (206). Applying this ratio to the Receiving Agency's 2024-25 annual maintenance and replacement budget (207) of \$13.9 billion (which includes preventive, routine, reactive maintenance, and rehabilitation), the potential savings become evident. Using a conservative estimate of \$4 in savings for every dollar spent, the Receiving Agency could save approximately \$55.6 billion in future deferred maintenance costs. Even if only 0.01% of these total savings is attributed to preventive maintenance guided by accurate pavement classification, this would still result in annual savings of approximately \$5.5 million.

Reducing Accident-Related Costs

Improved pavement conditions directly contribute to enhanced road safety, reducing the number of accidents caused by poor road conditions. According to the National Center for Statistics and Analysis, approximately 2.49 million injuries and 0.42 million deaths were reported in 2021 alone (208). The economic impact of these motor vehicle crashes is profound, costing \$340 billion in direct losses in

2019 and the broader societal harm reaching \$1.4 trillion when accounting for quality-of-life valuations (208,209). Moreover, longitudinal trends reveal a concerning escalation in both crashes and fatalities, with a notable 10% spike in 2020 and 2021 compared to previous years (210). In 2006, poor road conditions were reported to contribute to 31.4% of all traffic crashes nationwide. The influence of road conditions on crash frequency and severity was significant, with road-related factors becoming more prominent as crash severity increased. Specifically, road-related crashes accounted for 38.2% of non-fatal injuries, translating to approximately 2.2 million cases, and 52.7% of fatalities, amounting to 22,455 deaths.(211).

Investing in machine learning models for pavement classification and enhancing the predictive power of empirical models for skid resistance can significantly reduce accident-related costs. According to a macroeconomic analysis by Sabillon et al.(3) accurate network-level predictions for skid resistance, of which the surface information is a key input variable to improve goodness-of-fit, can yield a conservative net savings of approximately \$20 million annually. Since the outcomes of this study provide essential inputs rather than the models themselves, it is reasonable to attribute only 10% of these savings to the enhanced pavement surface classification. This results in an estimated annual savings of \$2 million directly guided by improved pavement classification.

Additionally, there are other potential sources of savings and benefits for the Receiving Agency. For instance, the ability to track the performance of pavement mixes over time can extend pavement life and enable more efficient resource allocation. However, for the purposes of this Value of Research (VoR) analysis, it suffices to focus on the primary cost savings derived from improved maintenance (\$5.5 million per year) and reduced accident-related costs (\$2 million per year).

NPV Cost Benefit analysis

The final economic analysis was conducted by using the Excel template provided by TxDOT. This template performs a net present value (NPV) cost-benefit analysis by considering:

1. Project budget: the total amount of money allocated to finance this research project, measured in dollars,
2. Project duration: the agreed upon timeframe for project completion, measured in years,
3. Expected value per year: An estimation of the annual savings incurred by TxDOT after implementing the project's products, measured in dollars,

4. Expected value duration: the timeframe over which this economic analysis is conducted, measured in years, and
5. Discount rate: the interest rate used in discounted cash flow analysis to determine the present value of future cash flows, measured as a percentage.

Many of the inputs were dictated by TxDOT or could not be varied as they were based on values from the contract; however, there are two terms, Exp. Value (per Yr.) and Expected Value Duration (Yrs.), which the Performing Agency had full freedom to vary. The values associated with those two terms (highlighted in yellow at the top of Figure C.2) governed the outputs of the economic analysis.

Inputs for the Economic Analysis

The project budget was set at \$414,534.68. This value was the agreed upon budget as stipulated in the project's contract team. The input project duration was 2.00274 years. The University Handbook (212) states that the project duration is not rounded. The project commenced on September 1, 2022, and the termination date is August 31, 2023. There are 731 days from the start date of the project to the end date (with the end date included), which equates to 2.00274 years. The expected value per year was \$7.5 million. This input is the total savings that were computed for the macro analysis. As mentioned in the previous sections, most of the values used in both analyses were as conservative as possible. The expected value duration of the project was assumed to be 10 years to reflect a potential timeframe between the inception of this project and the time it takes for it to be fully implemented. Finally, the input for the discount was 5% as recommended by the University Handbook (212). The inputs and outputs of this economic analysis can be seen at the top of **Figure 151** and a graphical representation of the NPV measured in millions of dollars over the timeframe of the economic analysis can be seen at the bottom of **Figure 151**.

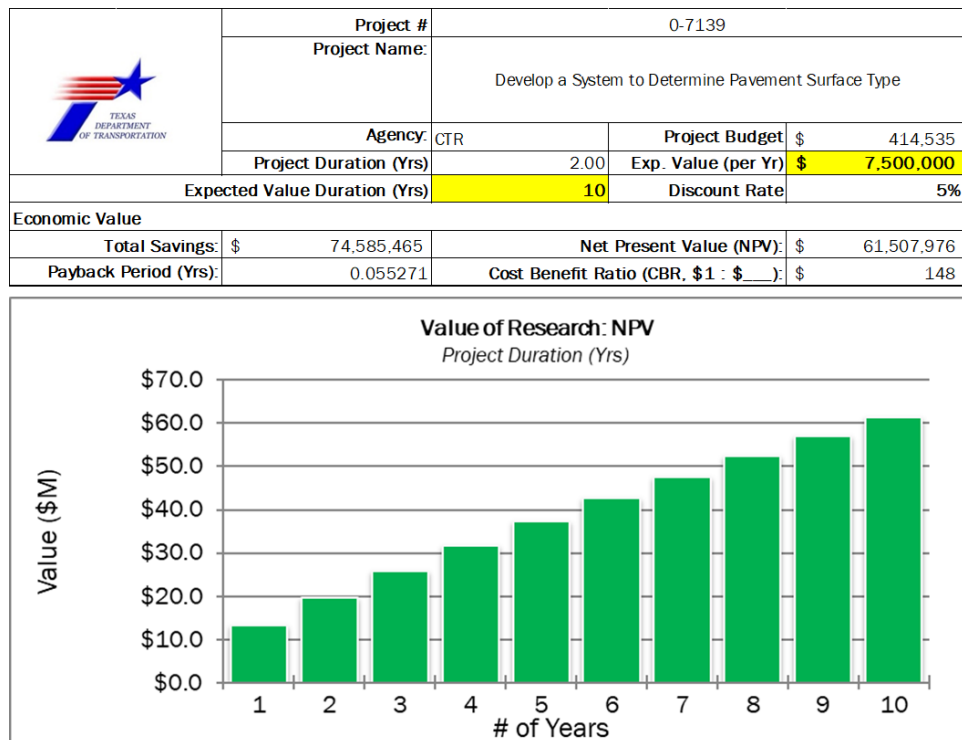


Figure 151: Inputs and outputs of NPV cost-benefit analysis (top), graphical representation of NPV over the course of ten years (bottom).

Conclusion

The projected economic savings from this project for the Receiving Agency are estimated to be \$70 million over a decade. Mathematically, this project pays for itself in a mere 0.055 years and boasts a cost-benefit ratio of 148. Such impressive figures are typical for projects focused on driver safety and infrastructure maintenance.

Investing in advanced machine learning models for pavement surface classification offers a compelling value proposition for the Receiving Agency. The potential financial benefits include reduced maintenance costs, extended pavement lifespan, optimized resource allocation, and decreased accident-related expenses, collectively leading to substantial cost savings. By embracing these innovative approaches, the Receiving Agency can enhance operational efficiency, improve road safety, and ensure the longevity of its infrastructure, ultimately delivering significant economic and societal benefits.

References

1. Henry J. Evaluation of Pavement Friction Characteristics. Washington D.C.: National Cooperative Highway Research Program; 2000.
2. Zuniga-Garcia N, Prozzi JA. High-Definition Field Texture Measurements for Predicting Pavement Friction. Washington D.C.: Transportation Research Record; 2019.
3. Sabillon C, Hernandez J, Li R, Huang R, Prozzi J. Efficient Model for Predicting Friction on Texas Highway Network. Austin: Texas Department of Transportation; 2023.
4. Luo W, Liu L, Li L. Measuring rutting dimension and lateral position using 3D line scanning laser and inertial measuring unit. *Autom Constr.* 2020 Mar;111:103056.
5. Zhang D, Xu X, Lin H, Gui R, Cao M, He L. Automatic road-marking detection and measurement from laser-scanning 3D profile data. *Autom Constr.* 2019 Dec;108:102957.
6. Ahmed M, Haas CT, Haas R. Toward Low-Cost 3D Automatic Pavement Distress Surveying: The Close Range Photogrammetry. *Can J Civ Eng.* 2011;38(12):1301–13.
7. Wang KCP. Elements of automated survey of pavements and a 3D methodology. *J Mod Transp.* 2011 Mar;19(1):51–7.
8. Inzerillo L, Di Mino G, Roberts R. Image-based 3D reconstruction using traditional and UAV datasets for analysis of road pavement distress. *Autom Constr.* 2018 Dec;96:457–69.
9. PIARC. PIARC Technical Committee on Surface Characteristics. Brussels: Permanent International Association of Road Congresses; 1987.
10. AASHTO. AASHTO Guide for Pavement Friction. Washington D.C.: American Association of State Highway and Transportation Officials; 2008.
11. Robbins M, Tran N. Literature Review: The Impact of Pavement Roughness on Vehicle Operating Costs. Auburn: National Center for Asphalt Technology; 2015.
12. Wu CL, Nagi MA. Optimizing Surface Texture of Concrete Pavement. Skokie: Portland Cement Association (PCA); 1995.
13. Zuniga-Garcia N. Prediction Friction with Improved Texture Characterization. (Master Thesis). Austin: The University of Texas at Austin; 2017.

14. Hibbs B, Larson R. Tire Pavement Noise and Safety Performance, PCC Surface Texture Technical Working Group. Washington D.C.: Federal Highway Administration (FHWA); 1996.
15. Sabillon-Orellana CA, Li R, Hernandez JB, Prozzi JA. Evaluation of the Relationship between Friction, Texture, and Noise. Okemos: United States Department of Transportation; 2021.
16. Rungruangvirojn P, Kanitpong K. Measurement of Visibility Loss due to Splash and Spray: Porous, SMA, and Conventional Asphalt Pavements. *Int J Pavement Eng*. 2010;11:499–510.
17. Serigos PA, Buddhavarapu P, Gorman GM, Hong F, Prozzi JA. The Contribution of Micro- and Macro-texture to the Skid Resistance of Flexible Pavement. Austin: U.S. Department of Transportation; 2016.
18. Gunaratne M, Bandara N, Medzorian J, Chawla M, Ulrich P. Correlation of Tire Wear and Friction to Texture of Concrete Pavements. *J Mater Civ Eng*. 2000;12:46–54.
19. PIARC. International PIARC Experiment to Compare and Harmonize Texture and. Paris: Permanent International Association of Road Congresses; 1995.
20. ASTM. ASTM E2380/2380M-15 Test Method for Measuring Pavement Texture Drainage Using an Outflow Meter [Internet]. ASTM International; 2019 [cited 2024 Jun 13]. Available from: <http://www.astm.org/cgi-bin/resolver.cgi?E2380E2380M-15R19>
21. Doty R. Study of the Sand Patch and Outflow Meter Methods of Pavement Surface Texture Measurement. In: *Surface Texture Versus Skidding: Measurements, Frictional Aspects, and Safety Features of Tire-Pavement Interactions* [Internet]. ASTM International 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959; 1975 [cited 2024 Jun 13]. p. 42–61. Available from: <https://asmedigitalcollection.asme.org/astm-ebooks/book/1211/chapter/27784314/Study-of-the-Sand-Patch-and-Outflow-Meter-Methods>
22. FAA. AC 150/5320-12C Measurement, Construction, and Maintenance of Skid Resistant Airport Pavement Surfaces. Washington D.C.: Federal Aviation Administration; 1997. Report No.: 150/5320-12C.
23. Chamberlin W, Amsler D. Measuring Surface Texture by the Sand-Patch Method. In: *Pavement Surface Characteristics and Materials* [Internet]. ASTM International 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959; 1982 [cited 2024 Jun 13]. p. 3–15. Available from: <https://asmedigitalcollection.asme.org/astm->

ebooks/book/1398/chapter/27801446/Measuring-Surface-Texture-by-the-Sand-Patch-Method

24. ASTM. ASTM E2157-15 Test Method for Measuring Pavement Macrotexture Properties Using the Circular Track Meter [Internet]. ASTM International; 2019 [cited 2024 Jun 13]. Available from: <http://www.astm.org/cgi-bin/resolver.cgi?E2157-15R19>
25. Lin Y, Dong C, Wu D, Jiang S, Xiang H, Weng Z. Study of Pavement Macro- and Micro-Texture Evolution Law during Compaction Using 3D Laser Scanning Technology. *Appl Sci*. 2023 May 6;13(9):5736.
26. Peng, Li, Zhan, Wang, Yang. Finite Element Method-Based Skid Resistance Simulation Using In-Situ 3D Pavement Surface Texture and Friction Data. *Materials*. 2019 Nov 21;12(23):3821.
27. Miller T, Swiertz D, Tashman L, Tabatabaee N, Bahia HU. Characterization of Asphalt Pavement Surface Texture. *Transp Res Rec*. 2012;2295:19–26.
28. El Gendy A, Shalaby A, Saleh M, Flintsch GW. Stereo-Vision Applications to Reconstruct the 3D Texture of Pavement Surface. *Int J Pavement Eng*. 2011;12:263–73.
29. AMES Engineering. High Speed Inertial Road Profiler [Internet]. 2020 [cited 2024 Jun 13]. Available from: <https://amesengineering.com/products/8300-high-speed-profiler/>
30. Oracheff A, Wang G, Fletcher JH. Macrotexture Assessments Using Line-Laser Technology. *Transp Res Rec J Transp Res Board*. 2022 Aug;2676(8):1–9.
31. Fernando EG, Hu S, Crockford W. Comparative Evaluation of Pavement Macrotexture Measurements from Different Devices. *Transp Res Rec*. 2022;
32. Perera RW, Orthmeyer R, De Leon Izeppi E. AFD 90, Group III Texture Measurements and Use. Brussels: World Road Association; 2013.
33. Katicha S, Mogrovejo DE, Flintsch GW, De Leon Izeppi E, Carrasco DM. Latest Development in the Processing of Pavement Macrotexture Measurements of High Speed Laser Devices. In Alexandria: 9th International Conference on Managing Pavement Assets; 2015.
34. Sayers MW, Karamihas SM. Interpretation of Road Roughness Profile Data. Ann Arbor: Federal Highway Administration; 1996.
35. Perera RW, Kohn SD. Issues in Pavement Smoothness: A Summary Report. Plymouth: National Cooperative Highway Research Program; 2002.

36. Sabillon-Orellana. Evaluation of Data Imputation Techniques in Pavement Texture Processing. (Master Thesis). Austin: The University of Texas at Austin; 2020.
37. Nason GP. Stationary and Non-stationary Time Series. In: Statistics in Volcanology. London: Geological Society; 2006. p. 129–42.
38. ISO. ISO 13473-4: Characterization of Pavement Texture by Use of Surface Profiles – Part 4: One Third Octave Band Spectral Analysis of Surface Profiles. Geneva: International Organisation for Standardisation; 2024.
39. ASTM. ASTM E1845-15: Standard Practice for Calculating Pavement Macrotexture Mean Profile Depth. West Conshohocken: American Society for Testing and Materials International; 2015.
40. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol. 2013;64(5):402.
41. Shadbahr T, Roberts M, Stanczuk J, Gilbey J, Teare P, Dittmer S, et al. The impact of imputation quality on machine learning classifiers for datasets with missing values. Commun Med. 2023 Oct 6;3(1):139.
42. ISO. ISO 13473-1: Characterization of Pavement Texture by Use of Surface Profiles - Part 1: Determination of mean profile. Geneva: International Organisation for Standardisation; 2009.
43. Katicha SW, Mogrovejo DE, Flintsch GW, De Leon Izeppi ED. Adaptive Spike Removal Method for High-Speed Pavement Macrotexture Measurements by Controlling the False Discovery Rate. Transp Res Rec J Transp Res Board. 2019;2525:100–10.
44. ASTM. ASTM E178-21: Standard Practice for Dealing with Outlying Observations. West Conshohocken: American Society for Testing and Materials International; 2021.
45. NASEM. Protocols for Network-Level Macrotexture. Washington, DC: The National Academies Press; 2021.
46. Cooley J, Tukey J. An Algorithm for the Machine Calculation of Complex Fourier Series. Math Comput. 1965;297–301.
47. Butterworth S. On the Theory of Filter Amplifiers. Exp Wirel Eng. 1930;536–41.
48. Gottaut C, Goubert L. Texture-based Descriptors for Road Surface Properties and How They can be Used in the Appropriate Standards. Vienna: Rosanne; 2016. (Innovative, cost-effective construction and maintenance for safer, greener and climate resilient roads). Report No.: P7-SST-2013-RTD-1.

49. Paterson WD. International Roughness Index: Relationship to Other Measures of. *Transp Res Rec.* 1986;1084.
50. Sanberg U. Road Macro- and Megatexture Influence on Fuel Consumption,. *ASTM Spec Tech Publ Am Soc Test Mater.* 1990;1031:460–79.
51. Ongel A, Harvey JT, Kohler E, Lu Q, Steven B. Investigation of Noise, Durability, Permeability, and Friction Performance Trends for Asphaltic Pavement Surface Types: First- and Second-Year Results. Berkeley: California Department of Transportation; 2008.
52. Kane M, Riahi E, Do MT. Tire/Road Rolling Resistance Modeling: Discussing the Surface Macrotexture Effect. *Coatings.* 2021 May 2;11(5):538.
53. Flintsch GW, Tang L, Katicha SW, de Leon Izeppi E, Viner H, Dunford H, et al. Splash and Spray Assessment Tool Development Program. Washington D.C.: Federal Highway Administration; 2014.
54. Sabillon-Orellana CA, Hernandez JB, Bassil M, Prozzi JA. Classification of Flexible Pavements Based on Texture Data. *Transp Res Rec.* 2022;2676:140–51.
55. Gadelmawla ES, Koura MM, Maksoud TMA, Elewa IM, Soliman HH. Roughness Parameters. *J Mater Process Technol.* 2002;123:133–45.
56. Underwood SB, Castorena C, Goenaga B, Rogers P. Evolution of Pavement Friction and Macrotexture After. Raleigh: North Carolina Department of Transportation; 2022.
57. ISO. ISO 21920-2: Geometrical Product Specifications (GPS) - Surface Texture: Profile - Part 2: Terms, Definitions and Surface Texture Parameters. Geneve: International Organisation for Standardisation; 2021.
58. Chen S, Liu X, Luo H, Yu J, Chen F, Zhang Y, et al. A State-of-the-Art Review of Asphalt Pavement Surface Texture and its Measurement Techniques. *J Road Eng.* 2022;2:156–80.
59. Sanberg U, Descornet G. Road Surface Influence on Tire/Road Noise. In Miami: Proceedings of the International Conference Noise Control Engineering, Noise Control for the 80's; 1980.
60. Sayers M, Karamihas S. The Little Book of Profiling. New York City: Regent of the University of Michigan; 1998.
61. Serigos PA, Smit ADF, Prozzi JA. Incorporating Surface Microtexture in the Prediction of Skid Resistance of Flexible Pavements. *Transp Res Rec.* 2014;2457:105–13.

62. Addison PS. The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine, and Finance. 2nd ed. Boca Raton: CRC Press; 2017.
63. Wei L, Fwa TF, Zhe Z. Wavelet Analysis and Interpretation of Road Roughness. *J Transp Eng*. 2005;131:120–30.
64. Diaspro A. Review of “Fractals and Chaos: The Mandelbrot Set and Beyond”, by B. Mandelbrot. *Biomed Eng OnLine*. 2005 Dec;4(1):30, 1475-925X-4–30.
65. Mandelbrot BB. The Fractal Geometry of Nature. 1st ed. W.H. Freeman and Co.; 1982. 286–287 p.
66. Rajaei S, Chatti K, Dargazany R. A Review: Pavement Surface Micro-texture and Its Contribution to Surface Friction. *TRB 96th Annual Meeting Compendium of Papers*; 2017; Washington D.C.
67. Villani MM, Scarpas A, de Bondt A, Khedoe R, Artamendi I. Application of Fractal Analysis for Measuring the Effects of Rubber Polishing on the Friction of Asphalt Concrete Mixtures. *Wear*. 2014;320:179–88.
68. Kane M, Rado Z, Timmons A. Exploring the Texture–Friction Relationship: From Texture Empirical Decomposition to Pavement Friction. *Int J Pavement Eng*. 2014;16:919–28.
69. Huang H, Pan J. Speech Pitch Determination Based on Hilbert-Huang Transform. *Signal Process*. 2006;86:792–803.
70. Hutchings I, Shipway P. Surface Topography and Surfaces in Contact. In: *Tribology: Friction and wear of engineering materials*. Second Editions. Cambridge, MA: Elsevier; 2017. p. 7–35.
71. FHWA. Surface Texture for Asphalt and Concrete Pavements. Washington D.C.: Federal Highway Administration; 2005.
72. Snyder M. Concrete Pavement Texturing. Washington D.C.: Federal Highway Administration; 2019.
73. Habig KH, editor. *Tribologisches Verhalten keramischer Werkstoffe: Grundlagen und Anwendungen*. Ehningen bei Böblingen: expert-Verl; 1993. 222 p. (Kontakt & Studium Werkstoffe).
74. Mang T, Bobzin K, Bartels T. *Industrial Tribology: Tribosystems, Friction, Wear and Surface Engineering, Lubrication* [Internet]. 1st ed. Wiley; 2010 [cited 2024 Jun 13]. Available from: <https://onlinelibrary.wiley.com/doi/book/10.1002/9783527632572>

75. DIN. DIN 4760:1982-06, Gestaltabweichungen; Begriffe, Ordnungssystem [Internet]. Berlin: Beuth Verlag GmbH; 1982 [cited 2024 Jun 13]. Available from: <https://www.beuth.de/de/-/-/962490>
76. Acheson DJ. Elementary fluid dynamics. Oxford : New York: Clarendon Press ; Oxford University Press; 1990. 397 p. (Oxford applied mathematics and computing science series).
77. Taylor JB, Carrano AL, Kandlikar SG. Characterization of the effect of surface roughness and texture on fluid flow—past, present, and future. *Int J Therm Sci*. 2006 Oct;45(10):962–8.
78. Darcy H. Recherches expérimentales relatives au Mouvement de l'eau dans les tuyaux. Paris: Creative Media Partners; 1857. 302 p.
79. Fanning JT. A practical treatise on hydraulic and water-supply engineering: relating to the hydrology, hydrodynamics, and practical construction of water-works, in North America. New York: Van Nostrand; 1887.
80. Nikuradse J. Laws of flow in rough pipes. Washington D.C.: National Advisory Committee for Aeronautics; 1937.
81. Colebrook CF. Turbulent Flow in Pipes, with Particular Reference to the Transition Region between the Smooth and Rough Pipe Laws. *J Inst Civ Eng*. 1939 Feb;11(4):133–56.
82. Moody LF. Friction Factors for Pipe Flow. *J Fluids Eng*. 1944 Nov 1;66(8):671–8.
83. ASME. Surface texture: (surface roughness, waviness, and lay): an international standard. New York, NY: The American Society of Mechanical Engineers; 2019. 126 p. (ASME).
84. Dagnall H. Exploring surface texture. Leicester: Rank Taylor Hobson; 1980. 170 p.
85. Whitehouse DJ. Handbook of Surface Metrology [Internet]. 1st ed. Boca Raton: Routledge; 2023 [cited 2024 Jun 13]. Available from: <https://www.taylorfrancis.com/books/9780203752609>
86. Kandlikar SG, Schmitt D, Carrano AL, Taylor JB. Characterization of surface roughness effects on pressure drop in single-phase flow in minichannels. *Phys Fluids*. 2005 Oct 1;17(10):100606.
87. Schmitt DJ, Kandlikar SG. Effects of Repeating Microstructures on Pressure Drop in Rectangular Minichannels. In: ASME 3rd International Conference on Microchannels and Minichannels, Parts A and B [Internet]. Toronto, Ontario, Canada: ASME; 2005 [cited 2024 Jun 13]. p. 281–9. Available

from:

<http://proceedings.asmedigitalcollection.asme.org/proceeding.aspx?articleid=1573694>

88. NCHRP. Test Methods for Characterizing Aggregate Shape, Texture, and Angularity [Internet]. Washington, D.C.: Transportation Research Board; 2007 [cited 2024 Jun 13]. Available from: <http://www.nap.edu/catalog/14017>
89. Xing B, Fan W, Lyu Y, Sun H, Che J. Influence of particle mineralogy and size on the morphological characteristics of mineral fillers. *J Mater Res Technol*. 2021 Nov;15:3995–4009.
90. Mora CF, Kwan AKH. Sphericity, shape factor, and convexity measurement of coarse aggregate for concrete using digital image processing. *Cem Concr Res*. 2000 Mar;30(3):351–8.
91. Xu Y. Fractal dimension of demolition waste fragmentation and its implication of compactness. *Powder Technol*. 2018 Nov;339:922–9.
92. Liu EJ, Cashman KV, Rust AC. Optimising shape analysis to quantify volcanic ash morphology. *GeoResJ*. 2015 Dec;8:14–30.
93. Stachowiak GP, Stachowiak GW, Podsiadlo P. Automated classification of wear particles based on their surface texture and shape features. *Tribol Int*. 2008 Jan;41(1):34–43.
94. Dürig T, Ross PS, Dellino P, White JDL, Mele D, Comida PP. A review of statistical tools for morphometric analysis of juvenile pyroclasts. *Bull Volcanol*. 2021 Nov;83(11):79.
95. Yang Y, Wang JF, Cheng YM. Quantified evaluation of particle shape effects from micro-to-macro scales for non-convex grains. *Particuology*. 2016 Apr;25:23–35.
96. Bagheri GH, Bonadonna C, Manzella I, Vonlanthen P. On the characterization of size and shape of irregular particles. *Powder Technol*. 2015 Jan;270:141–53.
97. Đuriš M, Arsenijević Z, Jaćimovski D, Kaluđerović Radoičić T. Optimal pixel resolution for sand particles size and shape analysis. *Powder Technol*. 2016 Nov;302:177–86.
98. Wang H, Wang D, Liu P, Hu J, Schulze C, Oeser M. Development of morphological properties of road surfacing aggregates during the polishing process. *Int J Pavement Eng*. 2017 Apr 3;18(4):367–80.

99. Yoder EJ, Witczak MW. Principles of pavement design. 2d ed. New York: Wiley; 2010.
100. Sutton W. Sizes and Grading of Aggregates for Road Maintenance and Construction. Indianapolis: Indiana State Highway Commission; 2016.
101. McLeod NW. Design of dense graded asphalt concrete pavements. Can Tech Asph Assoc. 1987;32:352.
102. Mix Types – Pavement Interactive [Internet]. 2003 [cited 2024 May 21]. Available from: <https://pavementinteractive.org/reference-desk/pavement-types-and-history/pavement-types/pavement-typesmix-types/>
103. Taylor G. Open-Graded Friction Courses (OGFC). Continuing Education and Development; 2014. Report No.: C02-058.
104. Allen Cooley L Jr, Ray Brown E, Watson DE. Evaluation of Open-Graded Friction Course Mixtures Containing Cellulose Fibers. Transp Res Rec. 2000;1723:19–25.
105. Garcia J, Hansen K. Hot Mix Asphalt Pavement Mix Types and Selection Guide. 4th ed. Washington D.C.: Federal Highways Administration; 2001.
106. Brown ER. Experience with Stone Matrix Asphalt in the United States. Auburn: National Center for Asphalt Technology; 1993.
107. Xie H, Watson DE. Determining Air Voids Content of Compacted Stone Matrix Asphalt Mixtures. Transp Res Rec. 2004;1891:203–11.
108. Hanson FM. Bituminous Surface Treatment of Rural Highways. Wellington: New Zealand Society of Civil Engineers; 1935.
109. El Hachem Y. The Uteas Seal Coat Design Method using 3D Laser Technology. Master Thesis. Austin: The University of Texas at Austin; 2019.
110. Bateman D. Design Guide for Road Surface Dressing. Wokingham: Transport Research Laboratory; 2016.
111. Epps JA, Gallaway BM, Hughes CH. Field Manual on Design and Construction of Seal Coats. College Station: Texas State Department of Highways and Public Transportation; 1981.
112. Mouaket I, Sinha K, White T. Guidelines for Management of Chip and Sand Seal Coating Activity in Indiana. Transp Res Rec. 1992;1507:81–90.

113. Sculer S, Epps-Martin A. Best Practices for Emulsified Asphalt Chip Seals. Washington D.C.: American Association of State Highway and Transportation Officials; 2021.
114. TxDOT. Standard Specifications for Construction and Maintenance of Highways, Streets, and Bridges. Austin: Texas Department of Transportation; 2014.
115. FHWA. High Friction Surface Treatments (HFST). Washington D.C.: Federal Highway Administration; 2022.
116. Holzschuher C. High Friction Surface Treatment Guidelines: Project Selection, Materials, and Construction. Gainesville: Florida Department of Transportation; 2017.
117. Sheehan T. High Friction Surface Treatment (HFST) I-57 and I-74 Interchange. In Champaign: 101st IL Transportation and Highway Engineering (T.H.E.) Conference; 2015.
118. Pranav C, Do MT, Tsai YC. Analysis of High-Friction Surface Texture with Respect to Friction and Wear. Coatings. 2021;11:758.
119. Brazer. Benchmark. 2023 [cited 2024 May 22]. Asphalt Leveling Course. Available from: <https://www.benchmark-inc.com/resources/results/2018/05/22/asphalt-leveling-course-installation-utilization-and-long-term-benefits>
120. Wegman D, Sabouri M. Ultra-Thin Bonded Wearing Course (UTBWC) Snow, Ice, and Wind Effects: Transportation Research Synthesis. St. Paul: Minnesota. Department of Transportation; 2018.
121. Rasmussen R, Sohaney R, Wiegand P. How to Reduce Tire-Pavement Noise: Better Practices for Constructing and Texturing Concrete Pavement Surfaces. Ames: National Concrete Pavement Technology Center; 2011.
122. Rasmussen R, Mun E, Ferragut T, Wiegand P. A Comprehensive Field Study on Concrete Pavement Solutions for Reducing Tire-Pavement Noise. In Honolulu: InterNoise 2006; 2006.
123. Hanson D, Waller B. Evaluation of the Noise Characteristics of Minnesota Pavements. St. Paul: Minnesota Department of Transportation; 2005.
124. Snyder M. Pavement Surface Characteristics: A Synthesis and Guide. Skokie: American Concrete Pavement Association; 2006.
125. FHWA. Texture of Concrete Pavements Observations from the FHWA Mobile Concrete Trailer (MCT). Washington D.C.: Federal Highway Administration; 2018.

126. Sandberg U, Ejsmont JA. Tyre/Road Noise: Reference Book. 1st ed. Kisa: Informex Ejsmont & Sandberg Handelsbolag; 2002.
127. Rao S, Yu HT, Darter MI. Longevity of Diamond-Ground Concrete Pavements. *Transp Res Rec.* 1999;1684:128–36.
128. Scofield L. Development and Implementation of the Next Generation Concrete Surface. Skokie: American Concrete Pavement Association; 2012.
129. Mitchell TM. Machine learning. Nachdr. New York: McGraw-Hill; 2013. 414 p. (McGraw-Hill series in Computer Science).
130. Russell SJ, Norvig P. Artificial intelligence: a modern approach. Third edition, Global edition. Boston Columbus Indianapolis New York San Francisco Upper Saddle River Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo: Pearson; 2016. 1132 p. (Prentice Hall series in artificial intelligence).
131. Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory.* 1967 Jan;13(1):21–7.
132. Harrison O. Medium. 2019 [cited 2024 Jun 13]. Machine Learning Basics with the K-Nearest Neighbors Algorithm. Available from: <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>
133. Fix E, Hodges JL. Discriminatory analysis: small sample performance. Randolph Field: USAF School of Aviation Medicine; 1952. Report No.: Tex., Project 21-49-004, Rept. 11.
134. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995 Sep;20(3):273–97.
135. Hastie T, Tibshirani R, Friedman J. Support Vector Machines and Flexible Discriminants. In: *The Elements of Statistical Learning* [Internet]. New York, NY: Springer New York; 2009 [cited 2024 Jun 13]. p. 417–55. (Springer Series in Statistics). Available from: <http://link.springer.com/10.1007/978-0-387-84858-7>
136. Larhmam. Support vector machine. In: *Wikipedia* [Internet]. 2024 [cited 2024 Jun 13]. Available from: https://en.wikipedia.org/w/index.php?title=Support_vector_machine&oldid=1225371653
137. Hastie T, Tibshirani R, Friedman J. Kernel Smoothing Methods. In: *The Elements of Statistical Learning* [Internet]. New York, NY: Springer New

- York; 2009 [cited 2024 Jun 13]. p. 210–2. (Springer Series in Statistics). Available from: <http://link.springer.com/10.1007/978-0-387-84858-7>
138. Hand DJ, Yu K. Idiot’s Bayes—Not So Stupid After All? *Int Stat Rev*. 2001 Dec;69(3):385–98.
 139. Tolles J, Meurer WJ. Logistic Regression: Relating Patient Characteristics to Outcomes. *JAMA*. 2016 Aug 2;316(5):533.
 140. Agresti A. Foundations of linear and generalized linear models. Hoboken, New Jersey: John Wiley & Sons Inc; 2015. 1 p. (Wiley series in probability and statistics).
 141. Wu X, Kumar V, Ross Quinlan J, Ghosh J, Yang Q, Motoda H, et al. Top 10 algorithms in data mining. *Knowl Inf Syst*. 2008 Jan;14(1):1–37.
 142. Ho TK. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell*. 1998 Aug;20(8):832–44.
 143. Ho TK. Random decision forests. In: *Proceedings of 3rd International Conference on Document Analysis and Recognition* [Internet]. Montreal, Que., Canada: IEEE Comput. Soc. Press; 1995 [cited 2024 Jun 13]. p. 278–82. Available from: <http://ieeexplore.ieee.org/document/598994/>
 144. Breiman L. Bagging predictors. *Mach Learn*. 1996 Aug;24(2):123–40.
 145. Freund Y, Schapire RE. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J Comput Syst Sci*. 1997 Aug;55(1):119–39.
 146. Hastie T, Tibshirani R, Friedman J. Boosting and Additive Trees. In: *The Elements of Statistical Learning* [Internet]. New York, NY: Springer New York; 2009 [cited 2024 Jun 13]. p. 343–5. (Springer Series in Statistics). Available from: <http://link.springer.com/10.1007/978-0-387-84858-7>
 147. Pirayonesi SM, El-Diraby TE. Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index. *J Infrastruct Syst*. 2020 Mar;26(1):04019036.
 148. Deng H, Zhou Y, Wang L, Zhang C. Ensemble learning for the early prediction of neonatal jaundice with genetic features. *BMC Med Inform Decis Mak*. 2021 Dec;21(1):338.
 149. Sagi O, Rokach L. Approximating XGBoost with an interpretable decision tree. *Inf Sci*. 2021 Sep;572:522–42.

150. Guo R, Zhao Z, Wang T, Liu G, Zhao J, Gao D. Degradation State Recognition of Piston Pump Based on ICEEMDAN and XGBoost. *Appl Sci*. 2020 Sep 21;10(18):6593.
151. Aggarwal CC. *Neural Networks and Deep Learning: A Textbook* [Internet]. Cham: Springer International Publishing; 2023 [cited 2024 Jun 13]. Available from: <https://link.springer.com/10.1007/978-3-031-29642-0>
152. Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Netw*. 1989 Jan;2(5):359–66.
153. O'Shea K, Nash R. *An Introduction to Convolutional Neural Networks* [Internet]. arXiv; 2015 [cited 2024 Jun 13]. Available from: <https://arxiv.org/abs/1511.08458>
154. Gordon AD. *Classification*. 1st ed. London: Chapman and Hall CRC; 1980.
155. Everitt B, Landau S, Leese M, Stahl D. *Cluster analysis*. 5th ed. Chichester, West Sussex, U.K: Wiley; 2011. 330 p. (Wiley series in probability and statistics).
156. Kaufman L, Rousseeuw PJ. *Finding Groups in Data: An Introduction to Cluster Analysis* [Internet]. 1st ed. Wiley; 1990 [cited 2024 Jun 13]. (Wiley Series in Probability and Statistics). Available from: <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316801>
157. Nielsen F. *Introduction to HPC with MPI for Data Science* [Internet]. Cham: Springer International Publishing; 2016 [cited 2024 Jun 14]. (Undergraduate Topics in Computer Science). Available from: <http://link.springer.com/10.1007/978-3-319-21903-5>
158. Sibson R. SLINK: An optimally efficient algorithm for the single-link cluster method. *Comput J*. 1973 Jan 1;16(1):30–4.
159. Ester M, Kriegel H, Sander, Xu X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Knowl Discov Data Min*. 1996;
160. Bishop CM. *Pattern recognition and machine learning*. New York: Springer; 2006. 738 p. (Information science and statistics).
161. Jørgensen B. *The theory of dispersion models*. London: Chapman & Hall; 1997. 237 p. (Monographs on statistics and applied probability).
162. Powers DMW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. 2020 [cited 2024 Jun 13]; Available from: <https://arxiv.org/abs/2010.16061>

163. Stehman SV. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens Environ.* 1997 Oct;62(1):77–89.
164. Chicco D, Tötsch N, Jurman G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Min.* 2021 Feb 4;14(1):13.
165. Martin T. Receiver Operating Characteristic. In: Wikipedia [Internet]. 2024 [cited 2024 Jun 13]. Available from: https://en.wikipedia.org/wiki/Receiver_operating_characteristic
166. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett.* 2006 Jun;27(8):861–74.
167. Leung K. Medium. 2022 [cited 2024 Jun 13]. Micro, Macro & Weighted Averages of F1 Score, Clearly Explained. Available from: <https://towardsdatascience.com/micro-macro-weighted-averages-of-f1-score-clearly-explained-b603420b292f>
168. Bjorn. 2D-3D Laser Scanners - Line Profile Scanner [Internet]. Acuity Laser. 2020 [cited 2024 Jun 13]. Available from: <https://www.acuitylaser.com/2d-3d-laser-scanners/>
169. Bjorn. AccuProfile 820 Laser Scanner - 2D-3D Laser Profiler [Internet]. Acuity Laser. 2022 [cited 2024 Jun 13]. Available from: <https://www.acuitylaser.com/product/laser-sensors/2d-3d-laser-scanners/accuprofile-820-laser-scanner/>
170. Al Khawli T, Anwar M, Gan D, Islam S. Integrating laser profile sensor to an industrial robotic arm for improving quality inspection in manufacturing processes. *Proc Inst Mech Eng Part C J Mech Eng Sci.* 2021 Jan;235(1):4–17.
171. Al Khawli T, Anwar M, Alzaabi A, Sunda-Meya A, Islam S. Machine Learning for Robot-Assisted Industrial Automation of Aerospace Applications. In: 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC) [Internet]. Miyazaki, Japan: IEEE; 2018 [cited 2024 Jun 13]. p. 3695–8. Available from: <https://ieeexplore.ieee.org/document/8616622/>
172. Plachinda P, Ritter C, Sabharwall P. Digital Engineering Sensor Architectures for Future Microreactor Builds [Internet]. 2021 Oct [cited 2024 Jun 13] p. INL/EXT-21-63631-Rev000, 1827623. Report No.: INL/EXT-21-63631-Rev000, 1827623. Available from: <https://www.osti.gov/servlets/purl/1827623/>

173. AMES Engineering. Texture and Profile Single Point Laser Sensor [Internet]. 2019 [cited 2024 Jun 13]. Available from: <https://amesengineering.com/products/ames-accutexture-100/>
174. Zhao G, Li S, Jiang Y, Lee J. Quality Assurance Procedures for Chip Seal Operations Using Macrottexture Metrics [Internet]. Purdue University; 2019 Jan [cited 2024 Jun 13]. Available from: <https://docs.lib.purdue.edu/jtrp/1671/>
175. Maeger K. Enhancement of Network Level Macrottexture Measurement Practices Through Deterioration Modeling and Comparison of Measurement Devices for Integration into Pavement Management Systems [Masters Thesis]. [Blacksburg]: Virginia Tech University; 2018.
176. Pathway Services. Pathway Services Inc. 2023 [cited 2024 Jun 13]. Services. Available from: <https://www.pathwayservices.com/services/>
177. Pathway Services. Pathway Services Inc. 2023 [cited 2024 Jun 13]. Technology. Available from: <https://www.pathwayservices.com/technology/>
178. Blanco R. Texture Signature: A Statistical Analysis of Pavement Texture Data [Internet]. Washington DC; 2022 [cited 2024 Jun 13]. Available from: <https://www.pathwayservices.com/news/13468/>
179. Blanco R, Constable P. The Pavement Texture Signature: An Alternative Analysis of the Texture Spectrum. RPUG Session 4.2; 2022.
180. Dong N, Prozzi J, Ni F. Reconstruction of 3D Pavement Texture on Handling Dropouts and Spikes Using Multiple Data Processing Methods. Sensors. 2019 Jan 11;19(2):278.
181. Kouchaki S, Roshani H, Prozzi JA, Garcia NZ, Hernandez JB. Field Investigation of Relationship between Pavement Surface Texture and Friction. Transp Res Rec J Transp Res Board. 2018 Dec;2672(40):395–407.
182. Li BQ, Li XC, Lu X. Microstructure and compressive properties of porous Ti–Nb–Ta–Zr alloy for orthopedic applications. J Mater Res. 2019 Dec 30;34(24):4045–55.
183. Zhao J, Huang J, Xiang Y, Wang R, Xu X, Ji S, et al. Effect of a protective coating on the surface integrity of a microchannel produced by microultrasonic machining. J Manuf Process. 2021 Jan;61:280–95.
184. Ma ZX, Ning J, Yu B, Zhang LJ, Fan JH, Yuan LG. Effects of process parameters and scanning patterns on quality of thin-walled copper flanges manufactured by selective laser melting. J Manuf Process. 2021 Dec;72:419–30.

185. Keyence. Keyence. 2023 [cited 2024 Jun 13]. 2D/3D Laser Profiler - LJ-X8000 series | KEYENCE America. Available from: <https://www.keyence.com/products/measure/laser-2d/lj-x8000/>
186. Huang Y. Moving Reference Road Profiler. The 30th Annual Road Profile Users' Group; 2018; Rapid City.
187. SICK. Sick. 2023 [cited 2024 Jun 13]. Ranger3 - Machine vision | SICK. Available from: <https://www.sick.com/us/en/catalog/products/machine-vision-and-identification/machine-vision/ranger3/c/g448354>
188. Weng Z, Xiang H, Lin Y, Liu C, Wu D, Du Y. Pavement texture depth estimation using image-based multiscale features. *Autom Constr.* 2022 Sep;141:104404.
189. Ravničan J, Marinko A, Noveski G, Kalabakov S, Jovanović M, Gazvoda S, et al. A Prestudy of Machine Learning in Industrial Quality Control Pipelines. *Informatika* [Internet]. 2022 Jun 15 [cited 2024 Jun 13];46(2). Available from: <https://www.informatika.si/index.php/informatika/article/view/3938>
190. Merkle D, Schmitt A, Reiterer A. SENSOR EVALUATION FOR CRACK DETECTION IN CONCRETE BRIDGES. *Int Arch Photogramm Remote Sens Spat Inf Sci.* 2020 Aug 14;XLIII-B2-2020:1107–14.
191. Canon. Canon. 2023 [cited 2024 Jun 13]. EOS 5D Mark IV Body. Available from: <https://www.usa.canon.com/shop/p/eos-5d-mark-iv?color=Black&type=New>
192. Ignatov A, Van Gool L, Timofte R. Replacing Mobile Camera ISP with a Single Deep Learning Model [Internet]. *arXiv*; 2020 [cited 2024 Jun 13]. Available from: <https://arxiv.org/abs/2002.05509>
193. Josipović N, Stolica D. Oral rehabilitation of a patient using a digital protocol. *J Dent.* 2022 Jun;121:104025.
194. Graham M. Have camera, will travel. *Aust Photogr.* 2020;36–43.
195. Arducam. Arducam 2MP OV2311 Global Shutter Mono Visible Light Camera Modules for Raspberry Pi [Internet]. Arducam. 2022 [cited 2024 Jun 13]. Available from: <https://www.arducam.com/product/arducam-ov2311-mipi-2mp-mono-global-shutter-camera-module-for-raspberry-pi/>
196. Wang Z. Development of an Image-Based 3D Deformation Measurement System for Triaxial Testing [Doctoral]. [Perth]: The University of Western Australia; 2022.

197. Rai P, Rehman M. ESP32 Based Smart Surveillance System. In: 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) [Internet]. Sukkur, Pakistan: IEEE; 2019 [cited 2024 Jun 13]. p. 1–3. Available from: <https://ieeexplore.ieee.org/document/8673463/>
198. Abu Bakar MN, Muhammad Audah LH. The Crop Disease Detection on Vertical Farming using Image Processing. *Evolution. Evol Electr Electron Eng.* 2021;2(2):989–98.
199. Kaniewska MM, Chvalová D, Dolezel D. Impact of photoperiod and functional clock on male diapause in cryptochrome and pdf mutants in the linden bug *Pyrrhocoris apterus*. *J Comp Physiol A* [Internet]. 2023 Jun 11 [cited 2024 Jun 13]; Available from: <https://link.springer.com/10.1007/s00359-023-01647-5>
200. Teledyne Flir. Teledyne Flir. 2023 [cited 2024 Jun 13]. Blackfly S USB3 | Teledyne FLIR. Available from: <https://www.flir.eu/products/blackfly-s-usb3?vertical=machine+vision&segment=iis>
201. Frentzos E, Tournas E, Skarlatos D. DEVELOPING AN IMAGE BASED LOW-COST MOBILE MAPPING SYSTEM FOR GIS DATA ACQUISITION. *Int Arch Photogramm Remote Sens Spat Inf Sci.* 2020 Aug 6;XLIII-B1-2020:235–42.
202. Gomez GAO, Orozco-Lopez JO, Casillas-Rodriguez FJ, Mora-Gonzalez M. 2D automatic alignment of optical spatial filtering system by PID control. In: 2022 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC) [Internet]. Ixtapa, Mexico: IEEE; 2022 [cited 2024 Jun 13]. p. 1–6. Available from: <https://ieeexplore.ieee.org/document/10018654/>
203. Pizzagalli SL, Bondarenko Y, Baykara BC, Niidas A, Kuts V, Kerm M, et al. Forestry Crane Immersive User Interface for Control and Teleoperation. In: Volume 2B: Advanced Manufacturing [Internet]. Columbus, Ohio, USA: American Society of Mechanical Engineers; 2022 [cited 2024 Jun 13]. p. V02BT02A021. Available from: <https://asmedigitalcollection.asme.org/IMECE/proceedings/IMECE2022/86649/V02BT02A021/1156840>
204. TxDOT. Pavement Manual: Section 6 Hot-Mix Asphalt Pavement Mixtures. Austin: Texas Department of Transportation; 2021.
205. Lenyk Z, Park J. Microsoft Vision Model: A state-of-the-art pretrained vision model [Internet]. Microsoft Research. 2021 [cited 2024 Jun 13]. Available from: <https://www.microsoft.com/en-us/research/blog/microsoft-vision-model-resnet-50-combines-web-scale-data-and-multi-task-learning-to-achieve-state-of-the-art/>

206. Hicks G, Seeds S, Peshkin D. Selecting a preventive maintenance treatment for flexible pavements [Internet]. Foundation for Pavement Preservation; 2000. Available from: <https://rosap.ntl.bts.gov/view/dot/14254>
207. TxDOT. Transportation Funding in Texas Fiscal Years 2024-25 [Internet]. Texas Department of Transportation; 2024. Available from: <https://ftp.txdot.gov/pub/txdot-info/fin/funding-brochure.pdf>
208. NCSA. Summary of Motor Vehicle Traffic Crashes: 2021 Data. Washington D.C.: National Center for Statistics and Analysis; 2023.
209. Blincoe L, Wang T, Swedler JS, Coughlin T, Lawrence B, Guo F, et al. The Economic and Societal Impact of Motor Vehicle Crashes,. Washington D.C.: National Highway Traffic. Safety Administration; 2023.
210. IIHS. Fatality facts 2021: Yearly snapshot. Arlington: Insurance Institute for Highway Safety; 2023.
211. Zaloshnja E, Miller TR. Cost of Crashes Related to Road Conditions, United States, 2006. *Ann Adv Automot Med Annu Sci Conf.* 2009 Oct 5;53:141–53.
212. TxDOT. University Handbook: The Research and Technology Implementation Division [Internet]. Austin: Texas Department of Transportation; 2022. Available from: <https://ftp.txdot.gov/pub/txdot-info/rti/university-handbook.pdf>

