

Analytical study of computer vision-based pavement crack quantification using machine learning techniques

2015

Soroush Mokhtari

University of Central Florida

Find similar works at: <http://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

STARS Citation

Mokhtari, Soroush, "Analytical study of computer vision-based pavement crack quantification using machine learning techniques" (2015). *Electronic Theses and Dissertations*. Paper 1156.

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.

ANALYTICAL STUDY OF COMPUTER VISION-BASED PAVEMENT CRACK QUANTIFICATION USING MACHINE LEARNING TECHNIQUES

by

SOROUSH MOKHTARI

B.S. University of Tabriz, 2005

M.Sc. K.N. Toosi University of Technology, 2007

M.Sc. University of Central Florida (UCF), 2012

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Civil, Environmental and Construction Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2015

Major Professor: Hae-Bum Yun

ABSTRACT

Image-based techniques are a promising non-destructive approach for road pavement condition evaluation. The main objective of this study is to extract, quantify and evaluate important surface defects, such as cracks, using an automated computer vision-based system to provide a better understanding of the pavement deterioration process. To achieve this objective, an automated crack-recognition software was developed, employing a series of image processing algorithms of crack extraction, crack grouping, and crack detection. Bottom-hat morphological technique was used to remove the random background of pavement images and extract cracks, selectively based on their shapes, sizes, and intensities using a relatively small number of user-defined parameters. A technical challenge with crack extraction algorithms, including the Bottom-hat transform, is that extracted crack pixels are usually fragmented along crack paths. For de-fragmenting those crack pixels, a novel crack-grouping algorithm is proposed as an image segmentation method, so called MorphLink-C. Statistical validation of this method using flexible pavement images indicated that MorphLink-C not only improves crack-detection accuracy but also reduces crack detection time.

Crack characterization was performed by analysing imagerial features of the extracted crack image components. A comprehensive statistical analysis was conducted using filter feature subset selection (FSS) methods, including Fischer score, Gini index, information gain, ReliefF, mRmR, and FCBF to understand the statistical characteristics of cracks in different deterioration stages. Statistical significance of crack features was ranked based on their relevancy and redundancy. The statistical method used in this study can be employed to avoid subjective crack

rating based on human visual inspection. Moreover, the statistical information can be used as fundamental data to justify rehabilitation policies in pavement maintenance.

Finally, the application of four classification algorithms, including Artificial Neural Network (ANN), Decision Tree (DT), k-Nearest Neighbours (kNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) is investigated for the crack detection framework. The classifiers were evaluated in the following five criteria: 1) prediction performance, 2) computation time, 3) stability of results for highly imbalanced datasets in which, the number of crack objects are significantly smaller than the number of non-crack objects, 4) stability of the classifiers performance for pavements in different deterioration stages, and 5) interpretability of results and clarity of the procedure. Comparison results indicate the advantages of white-box classification methods for computer vision based pavement evaluation. Although black-box methods, such as ANN provide superior classification performance, white-box methods, such as ANFIS, provide useful information about the logic of classification and the effect of feature values on detection results. Such information can provide further insight for the image-based pavement crack detection application.

***This work is dedicated to my parents “Ahmad” and “Eftekhar”, and my sisters
“Setareh” and “Saba”***

Without their support, this work would have not been possible.

ACKNOWLEDGMENTS

I would like to offer my sincerest gratitude to Dr. Hae-Bum Yun, my advisor, for his patient guidance, continuous encouragement, and abundant kindness. I would also like to offer thanks to the members of our research group specially Namgyu and Toni for their friendship and support.

I would also acknowledge the Department of Civil, Environmental and Construction Engineering for providing technical and financial support for this work.

TABLE OF CONTENTS

LIST OF FIGURES	x
LIST OF TABLES	xiv
CHAPTER 1: INTRODUCTION	1
1.1 Research Problem Statement	1
1.2 Aims and Objectives	3
1.3 Organization of Dissertation	7
CHAPTER 2: IMPROVEMENT OF CRACK DETECTION ACCURACY USING A NOVEL CRACK DE-FRAGMENTATION TECHNIQUE IN IMAGE-BASED ROAD ASSESSMENT.....	10
2.1 Introduction.....	10
2.2 Background	17
2.2.1 Pavement Crack Rating Manuals in U.S.A.....	17
2.2.2 Road Rating Practice for Flexible Pavement in Florida State	20
2.3 Collection of Pavement Surface Images	22
2.4 Crack Extraction Using Bottom-Hat Transform.....	23
2.5 Crack Grouping Using MorphLink-C.....	29
2.6 Discussion	34
2.6.1 Effects of MorphLink-C on Crack-Detection Accuracy.....	34
2.6.2 Effects of MorphLink-C on Crack Feature Subset Selection	41

2.6.3 Adaptive Property of MorphLink-C	47
2.7 Conclusions.....	51
CHAPTER 3: A NOVEL ADAPTIVE IMAGE-PROCESSING TECHNIQUE TO CONNECT FRAGMENTED CRACK PIXELS FOR FLEXIBLE ROAD PAVEMENT	54
3.1 Introduction.....	54
3.2 Brief Overview of Morphological Technique.....	60
3.3 Removal of Random Backgrounds Using Bottom-Hat Transform.....	63
3.4 Data Preparation.....	65
3.4.1 Data Preparation Using Road Surveying Images.....	65
3.5 Grouping of Fragmented Crack Pixels	70
3.5.1 Procedures of MorphLink-C	70
3.5.2 Feature Extraction Before and After MorphLink-C	77
3.6 Discussion	84
3.6.1 Statistics of Crack Lengths and Widths for Different Pavement Conditions ..	84
3.6.2 Effects on Crack Detection Accuracy Using ANN Classifier	91
3.7 Conclusions.....	100
CHAPTER 4: STATISTICAL CHARACTERIZATION OF CRACK FEATURE FOR AGING ROAD PAVEMENT USING COMPUTER-VISION TECHNIQUE	102
4.1 Introduction.....	102
4.2 Preparation of Crack Feature Dataset	107

4.2.1 FDOT's Road Surface Images of Flexible Pavement.....	107
4.2.2 Preliminary Image Processing Using MorphLink-C	108
4.2.3 Manual Separation of Crack and No-crack Objects	113
4.3 Feature Extraction of Crack and No-Crack Objects	116
4.3.1 Methods of Feature Extraction.....	116
4.3.2 Feature Extraction Results	123
4.4 Discussion	124
4.4.1 Statistical Distributions of the Feature Data	124
4.4.2 Statistical Correlations between Crack Features.....	127
4.4.3 Ranking of Crack Features for Optimal FSS	129
4.5 Conclusions.....	141
CHAPTER 5: COMPARATIVE STUDY OF CLASSIFICATION METHODS FOR IMAGE-BASED PAVEMENT CRACK DETECTION	145
5.1 Introduction.....	145
5.2 Morphological Image Processing and MorphLink-C for Crack Detection	148
5.3 Database Preparation and Feature Extraction	151
5.4 Review of the Selected Classification Methods.....	155
5.4.1 Neural Networks	158
5.4.2 K-Nearest Neighbors	161
5.4.3 Decision Trees	163

5.4.4 Adaptive Neuro-Fuzzy Inference System	165
5.5 Pavement Crack Detection.....	167
5.5.1 Application of Classification Methods to Pavement Data	167
5.5.2 Artificial Neural Networks	169
5.5.3 Decision Trees	171
5.5.4 K-Nearest Neighbors	173
5.5.5 Adaptive Neuro-Fuzzy Inference System	174
5.6 Discussion of the Results	176
5.6.1 Performance of Classifiers.....	176
5.6.2 Stability of Classification Results for the Imbalanced Training Dataset.....	178
5.6.3 Computation Time	180
5.6.4 Stability of Classification Results for Different Pavement Conditions	180
5.6.5 Clarity of Classification Procedure and Logic.....	181
5.7 Conclusions.....	188
CHAPTER 6: CONCLUSIONS	191
APPENDIX: APPROVAL LETTERS	197
REFERENCES.....	199

LIST OF FIGURES

Figure 1-1. Schematic illustration of study	4
Figure 2-1. Levels of crack recognition in road rating applications.	13
Figure 2-2. Proposed crack-recognition method for flexible road pavement.	15
Figure 2-3. Wrapper FSS method (Gutierrez-Osuna, 2014).....	16
Figure 2-4. Schematic designation of wheel path.....	21
Figure 2-5. Sample flexible pavement images collected with the Laser Road Imaging System (LRIS).	23
Figure 2-6. Morphological erosion transform of a grayscale image.....	25
Figure 2-7. Morphological operation procedures designed to detect black cracks from a gray-scale image.....	27
Figure 2-8. Binary images using different intensity-thresholding methods.	29
Figure 2-9. Proposed crack-grouping method.	33
Figure 2-10. A schematic of the ANN configuration.	37
Figure 2-11. The error analysis results of the ANN classifier for crack length with and without the MorphLink-C algorithm.....	41
Figure 2-12. The mean-square-errors for different feature subset combinations.	43
Figure 2-13. Sample classification results using all the six features after MorphLink-C.....	46
Figure 2-14. Sample classification results using three features of the length, texture and position after MorphLink-C.....	46
Figure 2-15. Sample crack-classification results using the ANN classifier.....	49

Figure 2-16. Adaptive process of the proposed fragment connection approach for different SD sizes.....	50
Figure 3-1. The overview of the proposed study of MorphLink-C	58
Figure 3-2. Diamond-shaped structuring element with size 7×7 pixels.....	61
Figure 3-3. Morphological erosion transform of a grayscale image.....	62
Figure 3-4. FDOT multi-purpose survey vehicle with laser road imaging system.....	66
Figure 3-5. Sample flexible-pavement images used in this study.	67
Figure 3-6. Sample crack-extraction results using the bottom-hat transform.....	69
Figure 3-7. The procedures of MorphLink-C algorithm for different crack types.	73
Figure 3-8. Sample results of the proposed MorphLink algorithm.	76
Figure 3-9. Sample result of the proposed crack grouping method.....	77
Figure 3-10. The features extracted from crack image components.....	78
Figure 3-11. Sample features of crack image components.	83
Figure 3-12. Wheel-path designation (FDOT, 2012a).....	84
Figure 3-13. The percentages of crack components for different pavement conditions.....	87
Figure 3-14. Box plots of crack length and width before and after MorphLink-C.....	91
Figure 3-15. A schematic of the ANN configuration.	92
Figure 3-16. Apparently successful classification in number and length.	94
Figure 3-17. Hypothesis test.	96
Figure 3-18. Type 1 and 2 errors before and after MorphLink-C.....	98
Figure 3-19. Sample classification results for different threshold values.....	99
Figure 4-1. Statistical characterization of crack features using computer-vision techniques....	106

Figure 4-2. Grey-scale images of flexible pavement in different surface conditions.....	108
Figure 4-3. Results of the MorphLink-C method.	112
Figure 4-4. The software for manual selection of crack and no-crack image components.	114
Figure 4-5. A sample result of manual selection of crack components.	115
Figure 4-6. Features extracted from flexible pavement images for crack detection.....	117
Figure 4-7. Measurement of the orientation feature.	119
Figure 4-8. Wheel path position feature.	122
Figure 4-9. Statistical distributions of crack and no-crack features.....	126
Figure 4-10. Pair-wise scatter plots of the feature correlation.....	128
Figure 4-11. Correlation coefficients of the cracks feature dataset.	129
Figure 4-12. Filter and wrapper FSS algorithms.	131
Figure 4-13. Terminology and procedure of the filter FSS.	132
Figure 4-14. The mean-square-errors of ANN classifier for different feature subset combinations (Wu et al., 2014).	140
Figure 4-15. Percentages of Class-1B, Class-2, and Class-3 cracks in different stages of pavement deterioration.	144
Figure 5-1. General procedure of this chapter.	148
Figure 5-2. Application of bottom-hat transformation and MorphLink-C methods for detecting cracks from pavement images.....	150
Figure 5-3. Wheel path designation defined by Florida Department of Transportation.....	155
Figure 5-4. One of the randomly generated datasets in the example problem.	158
Figure 5-5. Classification of example problem using artificial neural network classifier.....	160

Figure 5-6. Classification of sample problem using k-Nearest Neighbor method.	162
Figure 5-7. Classification of sample problem using decision tree method.	164
Figure 5-8. Classification results using ANFIS method.	166
Figure 5-9. Successful classification rates of ANN for training, validation and test datasets....	170
Figure 5-10. Successful classification, missed cracks and false alarms of ANN for test dataset.	171
Figure 5-11. Successful classification rates of DT for training, validation and test datasets.	172
Figure 5-12. Successful classification, missed cracks and false alarms of DT for test dataset. .	172
Figure 5-13. Successful classification rates of kNN for training, validation and test datasets... 173	
Figure 5-14. Successful classification, missed cracks and false alarms of kNN for test dataset.174	
Figure 5-15. Successful classification rates of ANFIS for training, validation and test datasets.	175
Figure 5-16. Successful classification, missed cracks and false alarms of ANFIS for test dataset.	175
Figure 5-17. Outputs of synthetic problem using ANN classifier.	183
Figure 5-18. Outputs of synthetic problem using DT classifier.....	184
Figure 5-19. Outputs of synthetic problem using kNN classifier.	185
Figure 5-20. Outputs of synthetic problem using ANFIS classifier.	186
Figure 5-21. Trained membership functions of area, length, intensity and location.	187
Figure A-1. Permission from ASCE Journal of Computing in Civil Engineering for using the paper entitled, ‘Improvement of crack-detection accuracy using a novel crack defragmentation technique in image-based road assessment’ as the first chapter of this study.	198

LIST OF TABLES

Table 2-1. A summary of crack-rating manuals for flexible and rigid pavements in U.S.A.....	18
Table 2-2. Description of the feature-extraction procedures before and after MorphLink-C.	34
Table 2-3. Training datasets for the ANN classification before and after MorphLink-C algorithm.	
All objects include both crack and non-crack image components.....	39
Table 2-4. Averaged MSE and training time of the ANN classification before and after MorphLink-C	44
Table 2-5. Computational times for the proposed morphological crack quantification method. .	45
Table 3-1. Averaged crack width estimated with Equation 3-5.....	74
Table 3-2. Dataset preparation of crack image components.....	86
Table 3-3. Statistics of the lengths of crack and non-crack image components.....	88
Table 3-4. Statistics of the widths of crack and non-crack image components.....	89
Table 4-1. The numbers and length of crack image components in manual selection.	116
Table 4-2. The feature dataset of the crack image components.....	123
Table 4-3. Results of statistical feature evaluation for different pavement conditions.	137
Table 4-4. Combined selected features for different pavement conditions.	139
Table 5-1. The numbers and length of crack and non-crack image components in training dataset.	152
Table 5-2. Statistical properties of all classes of data for the example problem.	157
Table 5-3. Error quantification of all classification algorithms for the example problem.	167
Table 5-4. Quantification of classification error based on the number of image components..	176

Table 5-5. Quantification of classification error based on the length of image components. 177

Table 5-6. Quantification of classification error for pavement in different stages of deterioration.

..... 181

CHAPTER 1: INTRODUCTION

1.1 Research Problem Statement

Roadway system as a main component of infrastructure, plays a critical socio-economic role by providing transportation for people and commodities. In 2011, four million miles of roadways carried more than three trillion vehicles in the United States. The backbone of the U.S. interstate highway system was constructed after the Federal-Aid Highway Act of 1956 signed by President Eisenhower, and the maintenance of rapidly aging highway system is considered as a significant engineering challenge. Deficiencies were partly due to the designed life of pavements. However, increasing traffic load which exceeded the predicted design values also contributed to early wear and tear of the pavements. According to American Society of Civil Engineers (ASCE) Report Card for America's Infrastructure (2013), more than 30% of major roads are in poor or mediocre condition. Aging roadway increases the vehicle repair costs by 67 billion dollar every year. Urban roads are deteriorated faster than rural roads. It is well known that since the cost of reconstructing road is extremely expensive, it is much more cost-effective to intervene in the early stages of pavement deterioration. Thus, the goal of preventive maintenance is "*applying the right treatment to the right pavement at the right time*" by the AASHTO Lead State Team on Pavement Preservation. In addition, pavement condition plays an important role in safety of roadways and nearly one-third of U.S. traffic fatalities are caused by poor pavement condition (ASCE, 2013). Traffic fatalities decreased annually over the past decade, however they still cost US economy around two hundred and thirty billion dollars each year.

It should be noted that most of funding for roadways and transportation is supplied through Federal Highway Trust Fund (HTF) from gas tax revenue. However, the gas tax revenue has not changed since 1993 and new fuel economy standards will result in less revenue in future. Consequently, Highways Trust Fund is expected to decrease fifty seven billion dollars between 2012 and 2022 (ASCE, 2013). It has been estimated that maintaining all US highways in their current condition would cost more than one hundred billion dollars, annually, and improving the highways needs an additional hundred seventy billion dollars each year. However, total available annual fund for this purpose is only ninety billion dollars, meaning that roadways are deteriorating further and further every year.

By 1980, it has been well established that developing Pavement Management Systems (PMS) are necessary to predict the long term performance of roadways and prioritize maintenance policies. Nowadays, a pavement management systems seek to identify the best strategy to maintain pavements serviceability (Finn, 2011). Over the past decades, all 50 states have developed pavement management programs that describe required physical measurements for rating pavements serviceability. Surveying distress data is a major task in every PMS. Using pavement distress data, agencies need to evaluate and interpret the data to determine the pavement condition and predict the long-term performance of pavement. Manual survey is still considered as a major method in pavement condition evaluation. In this method, inspectors identify and record different types of surface distress by walking or driving over the target segments of pavement. However, this approach is highly subjective, qualitative, and sometimes inaccurate. Evaluation results may vary due to personal judgment, distress type or severity (Smith, Freeman, & Pendleton, 1998).

Manual inspection can also prolong the procedure, cause traffic interruption and impose safety issues especially in high-volume highways.

Some automated pavement data collection approaches have been developed to overcome the shortcomings of manual inspection. In image-based systems, pavement images will be used to identify different distress types. An advantage of this approach is that quantitative analysis of pavement images can reduce the subjectivity of data interpretation, which is a significant limitation in the manual approach. To minimize the subjectivity, the pavement evaluation system should be fully automated based on mathematical procedures using computer vision and machine learning algorithms. An automated system should be able to identify different types of distress from the pavement images and quantify the severity and extend of each defect. However, each distress type may appear various in shape, size, color, orientation, etc. Moreover, different pavement textures will cause considerable variation in image background. The random background along with the presence of potholes, oil stains, lane-marks, wet areas and other objects make the crack identification procedure extremely difficult. Therefore, the identification of pavement defects requires to elaborate computer vision techniques and machine learning algorithms.

1.2 Aims and Objectives

Cracking is a major cause of pavement deterioration that determines the short-term and long-term performance of pavements. Statistical characteristics of pavement cracks are essential for pavement condition assessment which is a performance measure for evaluating roadway system. The main objective of this study is to extract useful information for pavement evaluation by mining the crack data from an automated computer vision based system. The contribution of

pavement crack detection and quantification to performance-based evaluation of roadways is illustrated in Figure 1-1a. This research aims to use the knowledge, discovered using the computer vision based system, as accurate quantitative information to explain the crack progression and pavement deterioration which can be used to justify pavement maintenance policies.

For this purpose, an automated computer vision based method, integrating crack-extraction, crack-grouping, and crack-detection processes is employed in this study. Schematic illustration of the computer vision-based method is presented in Figure 1-1b.

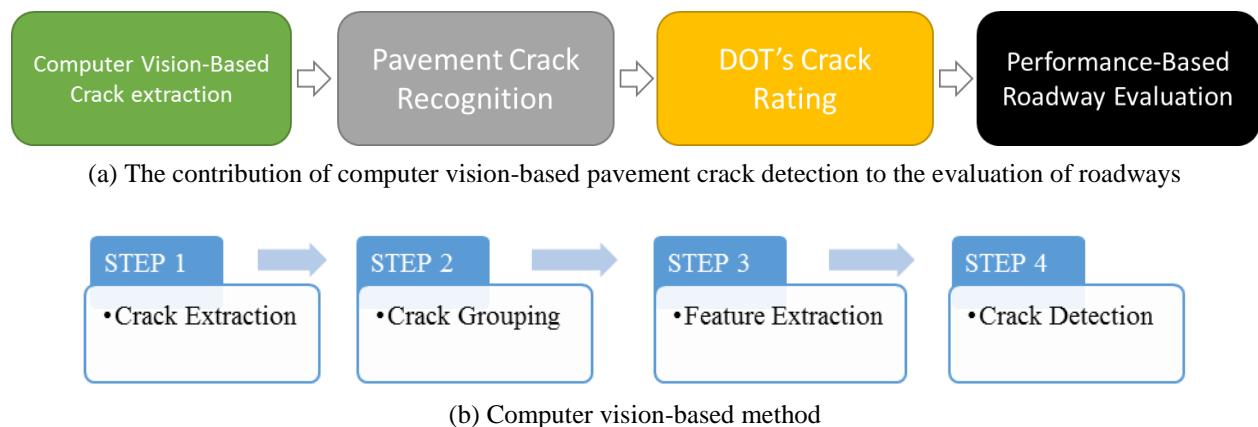


Figure 1-1. Schematic illustration of study

Morphological bottom-hat transformation is employed in this research for extracting cracks from pavement images (Step 1), because this method is capable of detecting cracks, selectively based on their size, shape and intensity with relatively small number of parameters.

A problem with the results of most crack-extraction methods, including morphological bottom-hat transformation is that detected cracks are fragmented and have many disjoints along the crack path. Several crack-grouping methods have been proposed to solve the fragmentation

problem, including: dilation and thinning transforms (He, Qiu, Wang, Zhang, & Xie, 2011; W. Huang & Zhang, 2012), seed-growing method (Q. Li et al., 2011), crack-tree method (Zou, Cao, Li, Mao, & Wang, 2012) and tensor-voting method (J. Huang et al., 2014). However, the width of cracks cannot be measured using most of the existing crack-segmentation methods. Crack width is an important measure for pavement evaluation and most state and federal agencies use crack width to evaluate type and severity of cracks. Moreover, the effect of segmentation technique on crack detection accuracy should be evaluated. Therefore, an adaptive crack-grouping algorithm, so called MorphLink-C, is studied in Chapter 3. This method can be used for any pixels level crack detection method and is not limited to a certain type of crack. MorphLink-C also provides an accurate way of computing averaged crack width. In order to quantify the effect of MorphLink-C on crack detection procedure, the performance of crack detection method before and after the segmentation is evaluated to show that MorphLink-C not only improves crack-detection accuracy but also reduces crack detection time.

The image processing techniques can efficiently remove random background from the pavement images, but, the results still include non-crack image components. Machine learning (classification) methods are commonly used to separate cracks from non-crack image components (Step 3). Different features have been employed in similar computer-vision based studies. Intensity-based feature, area and texture roughness measures are the most widely used features for separating crack image components from non-cracks (Kaseko & Ritchie, 1993; Kirschke & Velinsky, 1992; Xu, Huang & Chiang, 2003; Cheng, 1996; and Hu & Zhao, 2010). However, the common practice of feature selection is mostly intuitive and based on human observations. To cover this problem, six crack features including area, length, orientation, intensity, texture

roughness, and wheel path position, which are commonly used in pavement applications, are extracted from pavement in different deterioration stages. A comprehensive statistical analysis of the extracted features, using wrapper exhaustive search with ANN classifiers and filter feature subset selection (FSS) methods, including Fischer score, Gini index, information gain, ReliefF, mRmR, and FCBF, was conducted in Chapter 4, to avoid subjective and intuitive feature selection based on human observations. The optimal features subset that can describe important crack characteristics is determined by avoiding irrelevant or redundant features. The Knowledge and information that can be inferred from each feature and the importance of feature values for crack evaluation is also rarely discussed in pavement assessment literature. Characteristics of pavement cracks, in different deterioration stages, are quantified and interpreted to link the outputs of vision-based method to pavement evaluation measures. Special emphasis is placed to discover knowledge from the statistical analysis of features and provide useful information to characterize cracking process and pattern in different stages of aging flexible pavement.

Finally, application of machine learning methods for detecting cracks from pavement images is studies in Chapter 5. Classification techniques that are commonly employed for this purpose include: artificial neural networks (Lee & Lee, 2004; Moghadas Nejad & Zakeri, 2011b; Nguyen, Avila, & Begot, 2009; Saar & Talvik, 2010; Siddharth, Ramakrishnan, Krishnamurthy, & Santhi, 2012), support vector machine (Evdorides, Schlotjes, Henning, & Burrow, 2014; Gavilán et al., 2011; Moussa & Hussain, 2011; Salari & Ouyang, 2012), decision trees (Ho, Chou, & Lin, 2012; Moghadas Nejad & Zakeri, 2011a; Zhou & Wang, 2012), and k-nearest neighbors (Jahanshahi, Masri, Padgett, & Sukhatme, 2011). However, limited research has been carried out to compare different classification algorithms and discuss their advantages and drawbacks for the

image-based pavement crack detection problem. Moreover, the available studies, often focus on classification performance of the methods, disregarding the information and knowledge that can be inferred from the classification procedure. To address these issues, different classification algorithms in terms of computational complexity and clarity of procedure (i.e. being white or black box) are selected for evaluation in this study. The selected methods include: Artificial Neural Network (ANN), Decision Tree (DT), k-Nearest Neighbours (kNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS). Classification methods are evaluated for: 1) prediction performance, 2) computation time, 3) stability of results for highly imbalanced datasets in which, the number of crack objects are significantly smaller than the number of non-crack objects, 4) stability of the classifiers performance for pavements in different deterioration stages, and 5) interpretability of results and clarity of the procedure, to provide a comprehensive comparison of classification methods and signify their advantages and drawbacks for the pavement crack detection.

1.3 Organization of Dissertation

The application of morphological bottom-hat transformation for extracting the cracks from pavement images and MorphLink-C for segmentation of crack image components are presented in Chapter 2. This chapter that can be considered as the pilot study for this research is also published in ASCE Computing in Civil Engineering (Wu, L., Mokhtari, S., Nazef, A., Nam, B., & Yun, H.-B., 2014). Contributions of the current study in this paper include: 1) A comprehensive study of pavement crack rating manuals in the US with emphasis on Florida Department of Transportation guidelines , 2) preparing validation data by applying the bottom-hat transform and Morphlink-C to pavement surface images which have been collected with a Laser Road Imaging

System (LRIS) by the Florida Department of Transportation (FDOT) and extracting the designed features, 3) quantifying the effect of MorphLink-C on optimal feature subset using wrapper exhaustive search with ANN classifiers before and after the MorphLink-C and 4) evaluating the effect of MorphLink-C on crack detection by quantifying the detection performance before and after the MorphLink-C using the optimal feature subset.

A detailed explanation of the MorphLink-C method and experimental validation of its results is presented in Chapter 3. A comprehensive statistical analysis of the MorphLink-C effects on crack features before and after segmentation is conducted in this chapter.

Feature extraction and evaluation using the wrapper and filter feature subset section methods is presented in Chapter 4. The main contributions of this chapter to the subject matter include: 1) data preparation by manual selection of crack image component from 26 pavement images in different deterioration stages. 2) Extraction of six features that are commonly used in pavement crack detection applications including area, length, intensity, texture roughness, location and orientation. 3) A comprehensive statistical analysis of pavement cracks, in different deterioration stages using statistical distribution and correlations between crack features and 4) ranking crack feature for optimal feature subset selection. The finding of this chapter is under review for publication as journal paper (Mokhtari, S., Yun, H.-B., & Wu, L., 2015a). Chapter 5 aimed to contribute to the subject matter by providing comprehensive study of four well-known classification algorithms that have been widely used in pavement crack detection applications and compare the methods based on their performance, computation time, stability for highly imbalanced datasets, stability for pavements in different deterioration stages and interpretability

of results and clarity of the procedure, to provide a further insight into the advantages and drawbacks of each method for the pavement crack detection problem.

Finally, a summary of this research along with conclusions are provided in Chapter 6.

CHAPTER 2: IMPROVEMENT OF CRACK DETECTION ACCURACY USING A NOVEL CRACK DE-FRAGMENTATION TECHNIQUE IN IMAGE-BASED ROAD ASSESSMENT¹

2.1 Introduction

Effective maintenance of aging road pavement is a great engineering challenge to road maintenance authorities. The road network in U.S.A. includes more than 4 million miles of public roadways. According to the American Society of Civil Engineers (2013), 32% of America's major roads are in poor or mediocre condition, costing U.S. motorists who travel on deficient pavement \$67 billion per year in additional repairs and operating costs. It is estimated that \$101 billion in annual capital investment is needed to maintain the current pavement condition between 2008 and 2028, and \$170 billion per year to improve the current mediocre condition. However, federal, state and local governments are spending only \$91 billion per year due to budget limitation. It is established that current pavement maintenance approaches are not sustainable with rapid decrease of pavement condition and performance due to underfunding maintenance. Pavement preventive maintenance has received increasing attentions by many road maintenance agencies as an effort to improve road maintenance efficiency in current under-budget conditions (Federal Highway Administration, 2001). The pavement preventive maintenance is defined by the AASHTO's Lead State Team on Pavement Preservation as "*applying the right treatment to the right pavement at*

¹ The content of this chapter also appeared in:

Wu, L., Mokhtari, S., Nazef, A., Nam, B., & Yun, H.-B. (2014). Improvement of crack detection accuracy with novel crack grouping technique for flexible pavement. *ASCE Computing in Civil Engineering*. Using the paper as a chapter of this study is **with permission from ASCE** (please see the appendix). The author contributed to all sections except 2.4, 2.5, 2.8, and related parts in literature review and conclusions.

the right time." Therefore, accurate pavement condition assessment is vital for effective pavement prevention maintenance. Accurate condition assessment capability is also necessary to predict future deterioration rates and to establish rehabilitation strategies and budget.

Image-processing techniques to assess road condition are considered as a promising non-destructive method to quantify pavement distresses by analyzing pavement surface images. Crack in computer vision can be defined as follows: *a group of low-intensity pixels compared to neighboring pixels, which forms into an arbitrary line shape with length, width and directions (crack-pixel level). One or more crack pixels can further form into a complicated crack network of a single line, branched line or polygon which are not necessarily continued in its line paths due to the fragmentation of the crack pixels (crack-network level).* To deal with the multi-level topological shapes of crack images, an integrated image-processing approach needs to be employed for computer-aided crack recognition: 1) crack extraction, 2) crack grouping, 3) crack detection, and 4) crack classification. The description of each image-processing level is shown in Figure 2-1.

For the crack-extraction process (Level 1), many crack-extraction algorithms have been developed. In some early research, most crack detection approaches were *statistic thresholding methods*. For example, Koutsopoulos & Downey (1993) compared four intensity thresholding methods: Otsu's method, regression-based histogram method, relaxation method and Kittler's method. They observed that the regression-based histogram method provides the best results for pavement crack detection. Oliveira & Correia (2009) calculated the difference between a modified Otsu method (Dong, Yu, Ogunbona, & Li, 2008) and half of the standard deviation of all image pixel intensities as the threshold to separate pixels into non-cracks and potential cracks. In Nguyen,

Begot, Duculty, & Avila (2011), those connected pixels in four orientations ($0^\circ, 45^\circ, 90^\circ, 135^\circ$) with the smallest sum up intensity constructed four minimal paths. Pavement surface texture and crack shapes as well as pixel intensity have been used as crack features to improve crack detection accuracy. Song, Petrou, & Kittler (1995), Petrou, Kittler, & Song (1996), and Hu & Zhao (2010) used the texture analysis to abstract the crack information. Yan, Bo, Xu, & He (2007), and Sun & Qiu (2007) used the shape information and morphological filter to identify crack locations. Another popular approach for crack extraction is *edge detection methods*. Tsao, Kehtarnavaz, Chan, & Lytton, (1994), Li, (2003) and Ayenu-Prah & Attoh-Okine (2008) applied the traditional Sobel edge detection method to identify pavement crack locations. Tsai, Kaul, & Mersereau (2010) found that the dynamic optimization method outperformed in pavement crack detection among six different methods, including statistical thresholding, Canny edge detection, multi-scale wavelet, crack seed identification, iterative clipping, and dynamic optimization methods. *Seed-growing methods* have gained attentions recently for pavement crack extraction. In this approach, seed pixels are usually selected for local minimum intensity pixels. When a seed pixel is connected to neighboring low-intensity pixels, the connected lines form a crack shape. Crack Seed Verification (Y. Huang & Xu, 2006), F*Seed growing (Li, Zou, Zhang, & Mao, 2011), CrackTree (Zou, Cao, Li, Mao, & Wang, 2012), and tensor voting (Huang, Liu, & Sun, 2014) are some extensions of the seed growing-based algorithms. *Morphological techniques* that process images using structural elements have been employed for pavement crack extraction. The structuring element is a binary image component having a certain shape and size used to manipulate images. In Yan, Bo, Xu, & He (2007), crack edges could obtained by applying morphological dilation transform subtracted with morphological erosion transform. Sun, Salari, & Chou (2009) used dilation and erosion

transforms to fill gaps between crack disjoints. Then two crack objects were connected together if their endpoints were in the same neighbor area (i.e. 4 rows and 20 columns for horizontal crack). Jing & Aiqin (2010) used morphological opening transform to remove isolated crack noise. Mancini, Frontoni, & Zingaretti (2013) extracted crack components with top-hat transform. Different pavement distress types, such as potholes, were studied by Koch & Brilakis (2011), Koch, Jog, & Brilakis (2013). Golparvar-Fard, Balali, & Garza (2012) extended their image-based studies for the recognition of various highway assets using 3-D laser scanning point cloud data combined with semantic Textron forests approach.

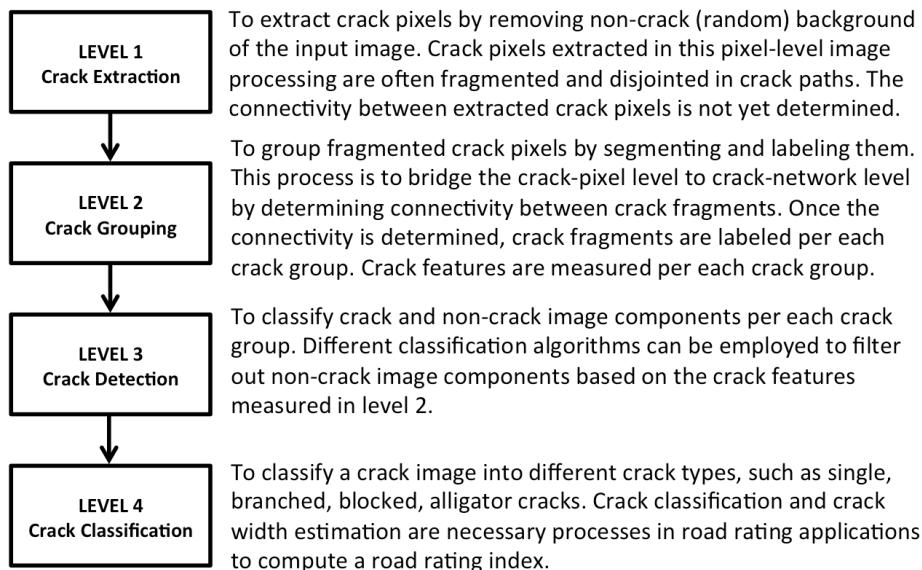


Figure 2-1. Levels of crack recognition in road rating applications.

A technical challenge in the crack-extraction process (Level 1) is that cracks should be extracted from random pavement background due to large variations of texture, roughness and intensity, spots and stains, oil and water spilling, and road markings. In addition, pavement crack

varies greatly in their shapes, sizes and widths, and multiple cracks can form a more complicated crack network, such as branched, block and alligator cracks. Resulting crack images usually include many non-crack objects due to random background noises. Therefore, effective crack-detection process (Level 3) is necessary to remove non-crack objects to increase crack recognition accuracy. Another problem in Level 1 is that crack image components are usually fragmented having (multiple) disjoints in their crack paths. Hence, crack-grouping process (Level 2) should be employed since crack fragmentation can be misleading in characterizing crack features, which results in lowering crack recognition accuracy.

The first objective of this study is to propose an automated image-processing method, integrating the crack-extraction, crack-grouping, and crack-detection processes. To validate the proposed method, flexible pavement surface images are used, which have been collected with a Laser Road Imaging System (LRIS) by the Florida Department of Transportation (FDOT). The morphological bottom-hat transform (Salembier, 1990) is used to extract crack images. To be shown in Section 2.4, the bottom-hat transform can effectively extract crack image components by removing various random pavement backgrounds, including pixel-level intensity noise due to rough pavement texture as well as region-level noise due to partial surface wetness. Although the bottom-hat transform can extract crack components from flexible pavement images, they are usually fragmented in their crack paths. Thus, it is necessary to employ a crack-grouping algorithm. In this study, a novel crack-grouping algorithm, called MorphLink-C, is proposed for de-fragmentation after the crack-extraction process. This algorithm is based on the morphological technique, which consists of 1) fragment grouping using dilation transform, and 2) fragment

connection using thinning transform. The advantages of the proposed MorphLink-C algorithm include:

- It can be used with arbitrary crack-extraction method.
- The connection of crack fragments is adaptive without involving complicated calculation of crack orientation, length, intensity, etc. for arbitrary crack types, such as single, branched, block and alligator cracks.
- It provides a simple and accurate way to measure crack width that is an important measure in road rating applications.

After MorphLink-C, the processed image still contains both crack and non-crack objects.

To filter non-crack objects, an artificial neural network (ANN) classifier is used as a crack-detection method. The proposed integrated crack-recognition approach is shown in Figure 2-2.

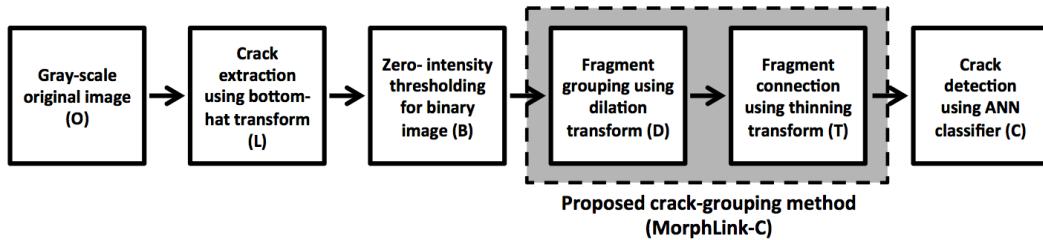


Figure 2-2. Proposed crack-recognition method for flexible road pavement.

The second objective of this study is to validate the effectiveness of the proposed crack-grouping algorithm. In this study, it is hypothesized that crack-detection accuracy can be improved with MorphLink-C by better representation of real crack feature characteristics. To validate this hypothesis, the accuracy of the ANN classifier is analyzed through feature subset selection (FSS)

method. FSS is defined as the process of selecting a subset of relevant features for use in the crack classification. FSS methods can be categorized into different methods depending on the training data types, objective function, and search methods. Among numerous FSS methods, since the wrapper FSS method is generally associated with a classifier to evaluate feature subsets by their predictive accuracy on test data by statistical resampling or cross-validation, the crack-detection accuracy is analyzed with the wrapper FSS method in this study. A wrapper FSS process is illustrated in Figure 2-3.

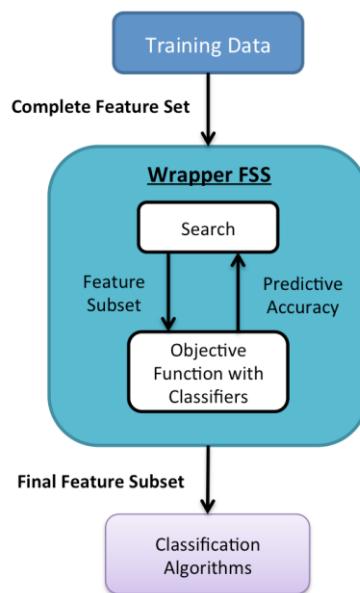


Figure 2-3. Wrapper FSS method (Gutierrez-Osuna, 2014).

A total of six crack features are measured before and after MorphLink-C, including length, area, orientation, texture, intensity, and position that are commonly used in pavement applications. In general, involving more features improves classification accuracy, while the computation time

increases. Since image processing-based road rating involves a large number of pavement image data often surveyed annually, finding the optimal feature subset is critical to develop an effective crack-recognition method. Hence, using the wrapper FSS the selection of the optimal feature subset is investigated based on the accuracy of ‘crack’ and ‘no-crack’ classification and computation time.

This paper is presented as follows. Section 2.2 provides the background of road rating manuals and practices in U.S.A. Section 2.3 describes the hardware specifications of the FDOT’s LRIS used to collected road surface images. In Sections 2.4 and 2.5, the procedures of crack extraction using the bottom-hat transform and crack grouping using MorphLink-C are described, respectively. Section 2.6 discusses the effectiveness of the proposed MorphLink-C as a de-fragmentation algorithm through the wrapper FSS method using an ANN classifier.

2.2 Background

2.2.1 Pavement Crack Rating Manuals in U.S.A.

Crack is a major pavement distress, and federal and state governments have developed standardized crack rating procedures to assess pavement conditions. Table 2-1 summarizes some of pavement crack-rating manuals developed by federal and state governments in U.S.A. The table includes crack classes, measure of extent, and crack features in assessment for both flexible and rigid pavements.

Table 2-1. A summary of crack-rating manuals for flexible and rigid pavements in U.S.A.

Agents	Pavement Type	Crack classes and extent		Crack features in assessment
American Association of State Highway and Transportation Officials (1993)	Flexible	Alligator or fatigue crack	Severity: low, medium, high	Pattern, width & spalling condition
			Extent: area	Area
		Block crack	Severity: low, medium, high	Pattern, width, spalling & sealing condition
			Extent: area	Area
		Joint reflection crack from PCC slab	Severity: low, medium, high	Spalling & sealing condition, width, bump occurrence & vicinity to random cracks
			Extent: length	Length
		Longitudinal and transverse crack	Severity: low, medium, high	Spalling & sealing condition, width, vicinity to random cracks & bump occurrence
			Extent: length	Length
	Rigid (JCP or CRCP)	Slippage crack	Severity: not defined	-
			Extent: area	Area
		Corner break	Severity: low, medium, high	Spalling, faulting, break-up & width
			Extent: number	Number
		D-crack	Severity: low, medium, high	Width of affected area, pattern, spalling & patching
			Extent: number	Number
		Longitudinal crack	Severity: low, medium, high	Width, faulting & spalling
			Extent: length	Length
		Transverse and diagonal crack	Severity: low, medium, high	Width, spalling, faulting & sealing condition
			Extent: number	Number
Federal Highway Administration (2003)	Flexible	Fatigue crack	Severity: low, moderate, high	Physical characteristics
			Extent: area	Area
		Block crack	Severity: low, moderate, high	Width, size of blocks & visual characteristics
			Extent: area	Area
		Edge crack	Severity: low, moderate, high	Physical properties
			Extent: length	Length of cracks
		Longitudinal crack	Severity: low, moderate, high	Width & vicinity to other cracks
			Extent: length	Length of cracks
		Transverse crack	Severity: low, moderate, high	Width, area & vicinity to other cracks
			Extent: length & number	Length & number of cracks

Agents	Pavement Type	Crack classes and extent		Crack features in assessment
Federal Highway Administration (2003)	Rigid (RC)	D-Crack	Severity: low, moderate, high Extent: number	Visual characteristics Number & area of D-Cracking
		Longitudinal crack	Severity: low, moderate, high	Width, spalling condition & visual characteristics
			Extent: length	Length of cracks
		Transverse crack	Severity: low, moderate, high	Width, spalling condition, faulting depth
			Extent: number & length	Length & number of cracks
		Corner breaking	Severity: low, moderate, high	Length of spalling, depth of faulting & patching condition
			Extent: number	Number at each segment
Florida Department of Transportation (2012a, 2012b)	Flexible	Crack	Type: Class IB, Class II, Class III	Width, depth & pattern
			Extent: %	Percentage of affected area
	Rigid	Transverse crack	Severity: light, moderate, severe	Width & physical properties
			Extent: number	Number of cracks
		Longitudinal crack	Severity: light, moderate, severe	Width & physical properties
			Extent: number	Number of cracks
		Corner crack	Severity: light, moderate, severe	Width & physical properties
			Extent: number	Number of cracks
Illinois Department of Transportation (2004)	Flexible & Rigid	Condition Rating Survey		Comparison with reference pictures
Kansas Department of Transportation (2013)	Flexible	Fatigue crack	Severity: Fc1, Fc2, Fc3, Fc4 Extent: length	Physical characteristics Length of cracking
		Transverse crack	Severity: T0, T1, T2, T3	Width, length & visual properties
			Extent: 1 or 2 digit number	Number of full width cracks
		Block crack	Severity: 1,2,3, 4	Block size & presence of secondary cracking
		General condition: scale of 1-10		Comparison with reference pictures
New York State Department of Transportation (2010)	Flexible & Rigid	Faulting: presence		Visual inspection
		Spalling: isolated or general		Percentage of the total length of the segment
		Alligator crack: isolated or general		Percentage of the total length
		Widening drop-off: low or high		Visual inspection
		Block crack: %		Pattern, area & length
Texas Department of Transportation (2010)	Flexible	Alligator crack: %		Pattern & length
		Longitudinal crack: %		Orientation, width, pattern & length
		Transverse crack: %		Orientation, width, pattern & length

Agents	Pavement Type	Crack classes and extent		Crack features in assessment
Texas Department of Transportation (2010)	Rigid (CRCP)	Average Crack spacing: number		Average crack spacing (transverse)
		Spalled crack: number		Length & space of cracks
	Rigid (JCP)	Corner breaks: number		Length & visual inspection
		D-Cracking: number		Crack spacing & physical pattern
		Spalls: number		Length & width
Washington State Department of Transportation (1992)	Flexible	Alligator crack	Severity: low, medium, high	Width & pattern
			Extent: %	Percentage of wheel paths length
		Longitudinal crack	Severity: low, medium, high	Width & pattern
			Extent: %	Percentage of segment length
		Transverse crack	Severity: low, medium, high	Width & pattern
			Extent: freq. per 100 ft	Frequency per 100ft
		Block crack	Severity: low, medium, high	Size of blocks & width of cracks
			Extent: length	Full length of segment
	Rigid (RCP)	Crack	Severity: low, medium, high	Number per panel
			Extent: %	Area percentage

2.2.2 Road Rating Practice for Flexible Pavement in Florida State

To establish standards for road condition rating using collected pavement images, FDOT developed Pavement Condition Survey Handbook for Flexible Pavement (Florida Department of Transportation, 2012a) and Pavement Condition Survey Handbook for Rigid Pavement (Florida Department of Transportation, 2012c). In the FDOT manual for flexible pavements, cracks will be classified into Classes IB, II or III based on their widths, general patterns (forming block or alligator cracks) and severities of branching and spalling. The percent affected area of distress for each class will be calculated within and outside of the wheel path (CW and CO, respectively as presented in Figure 2-4). According to this manual, the effect of raveling and patching should be added to the area of Class III cracks (their effects will be considered together); therefore the total percentage of affected area can be calculated as presented in Equation 2-1.

$$\text{Total Percent affected area} = \text{Class IB} + \text{Class II} + \text{Class III} + \text{Raveling} + \text{Patching} \quad (2-1)$$

In order to determine the predominant crack type, percent affected area of the three classes should be compared. The predominant class has the largest percentage of area. The percent area of each crack type will be used to calculate deduction values from provided tables and crack rating can be calculated using equation 2-2.

$$\text{Crack Rating (Flexible Pavement)} = 10 - \text{CW} + \text{CO} \quad (2-2)$$

where CW and CO are the percent affected area of all crack types inside and outside of the wheel path, respectively. The overall defect rating of rigid pavements can be calculated using Equation 2-3.

$$\text{Defect Rating (Rigid Pavement)} = 100 - \text{summation of deduction values} \quad (2-3)$$

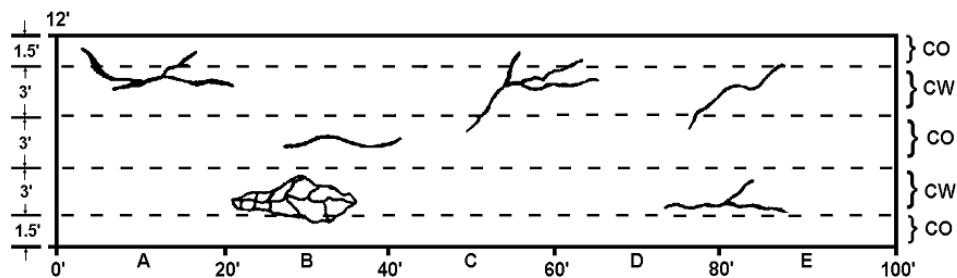


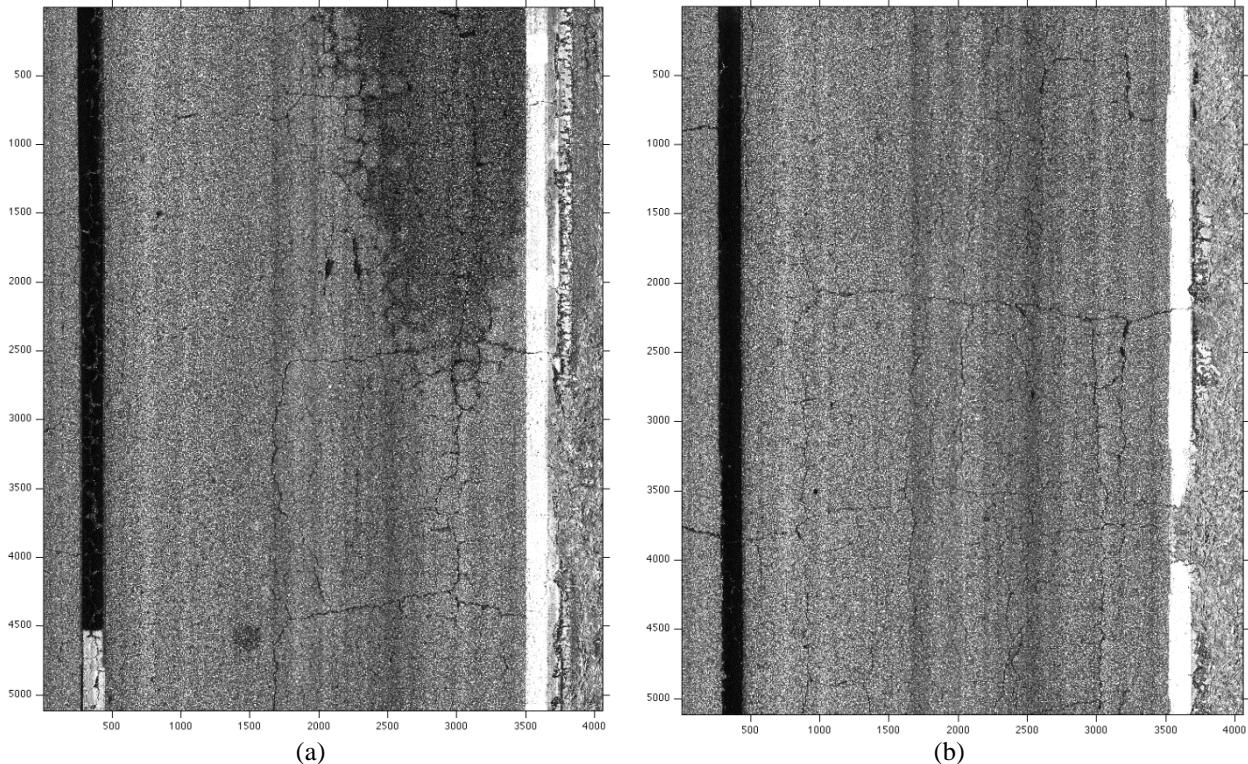
Figure 2-4. Schematic designation of wheel path

2.3 Collection of Pavement Surface Images

The Florida Department of Transportation (FDOT) conducts annual pavement condition surveys as a part of the Pavement Management System (Florida Department of Transportation, 2013). In 2006, the FDOT acquired a Multi-Purpose Survey Vehicle (MPSV) consisting of a self-contained van equipped with an Inertial Profiler System, an Inertial Navigation System, a Laser Rut Measurement System (LRMS) and a Laser Road Imaging System (LRIS) to capture pavement images at highway speed. The LRIS is composed of two high-resolution linescan cameras and laser illuminators that are configured to image up to 4-m transverse road section with about 1-mm resolution at speeds of 60 mile per hour. The camera is mounted above 1,960 mm from road surface, which has a 20-mm focal length. The image-sensor pixel size is 0.01 mm. Thus, the image resolution of the LRIS used in this study is 0.98 mm per pixel. Sample flexible pavement images are shown in Figure 2-5.

While ride quality and rut depth are collected automatically, surface distress evaluation like cracking is based on “manual” windshield survey. Such manual distress evaluation is currently used for network as well as project level pavement condition surveys. However, manual distress surveys could involve exposure to hazardous conditions and involve subjectivity and bias in the rating procedure. To process a large number of pavement image data surveyed annually, FDOT needed a computer application that could accurately and efficiently detect and quantify cracks from pavement images with minimal human intervention. FDOT sponsored a research project to evaluate potential computer-based applications for crack detection and quantification from images. The results from the study identified two commercially available software packages which were compatible with the MPSV LRIS. Of the two software packages evaluated, the results indicated

the crack quantification accuracy of both applications was limited and was satisfactory only when computer-based manual evaluation was used (Gunaratne, Amarasiri, & Nasseri, 2008).



* The numbers in the x- and y-axis are pixel numbers.

Figure 2-5. Sample flexible pavement images collected with the Laser Road Imaging System (LRIS).

2.4 Crack Extraction Using Bottom-Hat Transform

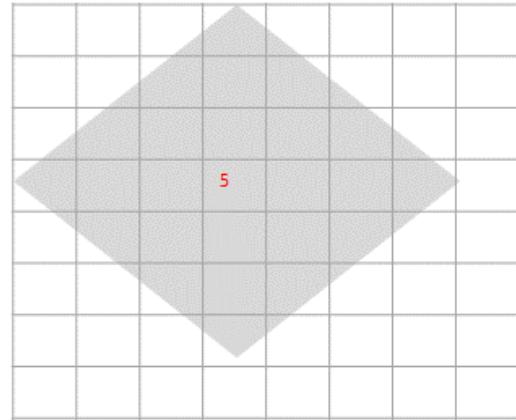
The morphological image processing technique uses mathematical morphology as a tool for extracting image components that are useful in the representation and description of various region shapes, such as boundaries, skeletons, and convex hull (Gonzalez, Woods, & Eddins, 2009).

The mathematical foundation of the morphological technique is based on the theory of set algebra by Minkowski (1903) and on theory of topology by Matheron (1975) (Jahanshahi, Kelly, Masri, & Sukhatme, 2009; Pratt, 2001). A general description about the morphological image processing technique can be found in Dougherty & Lotufo (2003).

Let $\{p\}$ be the set of pixels in a two-dimensional digital image. Thus, $\{p\}$ partitions the xy -plane into a grid, with the coordinates of the center of each grid being a pair of elements (x, y) from the Cartesian coordinate. A function $f(x, y)$ is said to be a digital image if (x, y) are integers from $\{p\}$ and f is mapping that assigns an intensity value to each distinct pair of (x, y) . The morphological technique applies a shape of binary image referred to as a *structuring element* on an input image. The structuring element is a binary image component whose shape can be designed for different purposes of image manipulation. The center of the structuring element is called the pixel of interest (POI). The pixels within the structuring element boundary have one, and the rest have zero. Figure 2-6a shows an example of grayscale input image with a diamond structuring element. In the morphological technique, dilation and erosion transforms are two basic operators to manipulate $f(x, y)$. The *dilation* is an operation that “grows” or “thickens” objects in an image, while the *erosion* is an operation that “shrinks” or “thins” in an image. It is a common convention in image processing that the dilation and erosion of the original image (O) with the structuring element (S) are expressed as $O \oplus S$ and $O \ominus S$, respectively. An advantage of the morphological technique is that multiple basic operators can form another operation for more sophisticated image processing. For example, two commonly-used operations are opening and closing as $O \circ S = (O \ominus S) \oplus S$ and $O \bullet S = (O \oplus S) \ominus S$, respectively.

60	68	126	159	15	21	53	122
62	64	5	146	19	15	21	122
15	21	53	135	122	15	19	57
17	19	15	167	122	25	122	57
253	122	15	148	62	147	14	60
5	77	56	125	11	156	62	62
16	45	46	136	126	159	15	15
13	85	123	149	19	15	167	165

(a) Input image and structuring element



(b) Output image

Figure 2-6. Morphological erosion transform of a grayscale image.

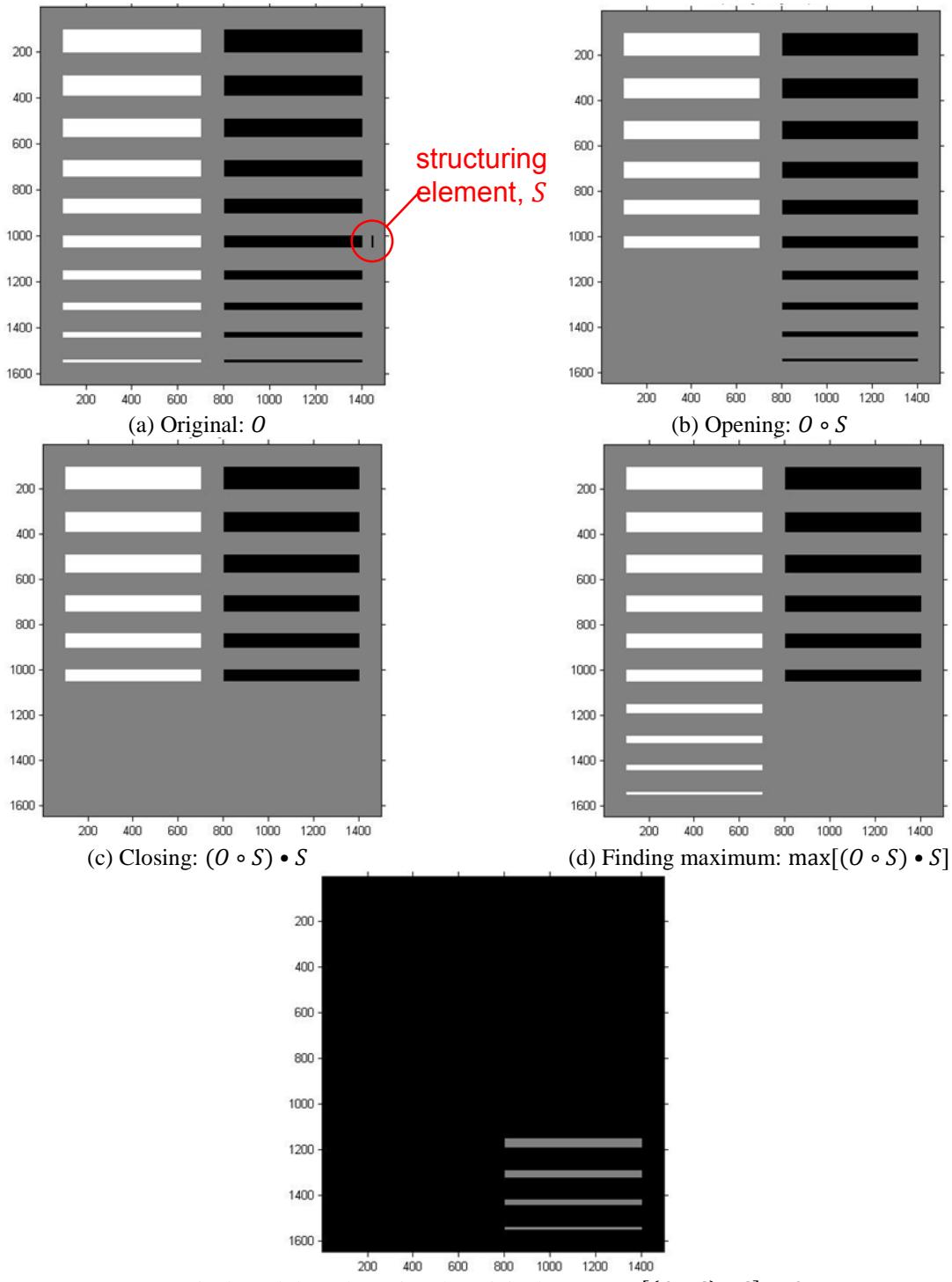
Crack extraction is a pixel-level operation to detect crack-like objects from a pavement image by removing random background. For flexible pavement, the random background usually includes asphalt types, pavement coating, surface texture and roughness, pavement aging and raveling, spots, stains and scratches, oil and water spilling, and road markings. Salembier (1990) and Jahanshahi et al. (2009) proposed the modified bottom-hat and top-hat transforms to detect black and white cracks as presented in Equations 2-4 and 2-5, respectively.

$$L_b = \max \left[\left(O \circ S_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} \right) \bullet S_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}, O \right] - O \quad (2-4)$$

$$L_t = O - \min \left[\left(O \bullet S_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} \right) \circ S_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}, O \right] \quad (2-5)$$

where L is a gray-scale image as the output of the morphological transforms; O is the original gray-scale image; $S_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}$ is the line-shape structuring element rotating 0° , 45° , 90° and 135° ; \circ is the opening transform; and \bullet is the closing transform. To demonstrate how a crack

can be detected using the bottom-hat transform, an example is shown in Figure 2-7. Figure 2-7a shows an original gray-scale image with white and black horizontal cracks of 10 different crack widths from 10 to 100 pixels with a 10-pixel increment. The goal in this example is to detect only black cracks that are less than 50 pixels wide from the gray background. For it, a structuring element was designed to be a vertical line with the size of 50×1 . Figures 2-7b, 2-7c and 2-7d illustrate the step-by-step procedures of the bottom-hat transform. The resulting image of the bottom-hat transform is shown in Figure 2-7e. Only black cracks with less than 50-pixel crack widths were detected as expected. It should be noted that the detected cracks, which were originally black, are converted into gray, and the background, which was originally gray, is now converted into black after the bottom-hat transform. However, the absolute intensity difference of the detected crack and background remained the same before and after the transform. The above example shows several advantages of the bottom-hat transform for pavement crack extraction. By designing the structuring element properly of its shape (i.e., vertical line for horizontal cracks), size (i.e., 50 pixels), and combination of morphological operators (i.e., bottom-hat transform), one can detect pavement crack *selectively* based on their shapes (e.g., horizontal line-shape crack), sizes (e.g., crack width), and intensities (e.g., black crack). Another advantage is that the bottom-hat transform uses a relatively small number of user-defined parameters. In this example, only *one* parameter is required of the structuring element length to detect black cracks with less than 50-pixel crack widths by employing the structuring element of a vertical line. Minimizing user-intervention with a small number of user-defined parameters is important to develop a fully automated system.



* The x- and y-axes are pixels.

Figure 2-7. Morphological operation procedures designed to detect black cracks from a grayscale image.

The bottom-hat transform is also advantageous to recognize dark cracks from random pavement background. To validate it, the bottom-hat transform was applied to the flexible pavement images collected by FDOT. The FDOT manual (Florida Department of Transportation, 2012b) specifies that the Class-III crack width be 25.4 mm (1.0 in) or larger. In this study, the line-shape structuring element with 100-pixel length was used in the analysis, which is equivalent to 98.0 mm (3.86 in). Therefore, the structuring element was designed to be about 4 times larger than the minimum Class-III crack width to ensure to detect cracks in all classes that are specified in the manual. By rotating the structural element for 0°, 45°, 90° and 135°, cracks having arbitrary line orientation could be detected.

As shown in Figure 2-8a, the original gray-scale pavement image (**O**) has severe alligator cracks with some spallings. The cracks are seen in random asphalt background which is due to aged pavement surface texture and ununiform partial wetness in the top-right corner. The bottom-hat transform result is shown in Figure 2-8b. To demonstrate its effectiveness of background removal, the result (**B**) is compared with two histogram-based intensity thresholding methods for pavement crack detection, Otsu's intensity thresholding (Otsu, 1979) and neighboring difference histogram methods (Q. Li & Liu, 2008) in Figures 2-8c and 2-8d, respectively. In the original pavement image (**O**) in Figure 2-8a, the surface roughness and partial wetness cause different random backgrounds. The surface roughness causes pixel-level intensity variation due to the randomness of pavement surface texture. The partial wetness causes regional intensity variation due to low intensity in the wet area in the top-right corner. Therefore, being a global intensity thresholding method, the Otsu's method was affected by both the surface roughness and wetness as shown in Figure 2-8c. Using the neighboring difference histogram method, background removal

was improved by reducing the effect of the roughness; however, the wetness was not removed effectively as shown in Figure 2-8d. Figure 2-8b shows that both the surface roughness and wetness are effectively removed using the bottom-hat transform.

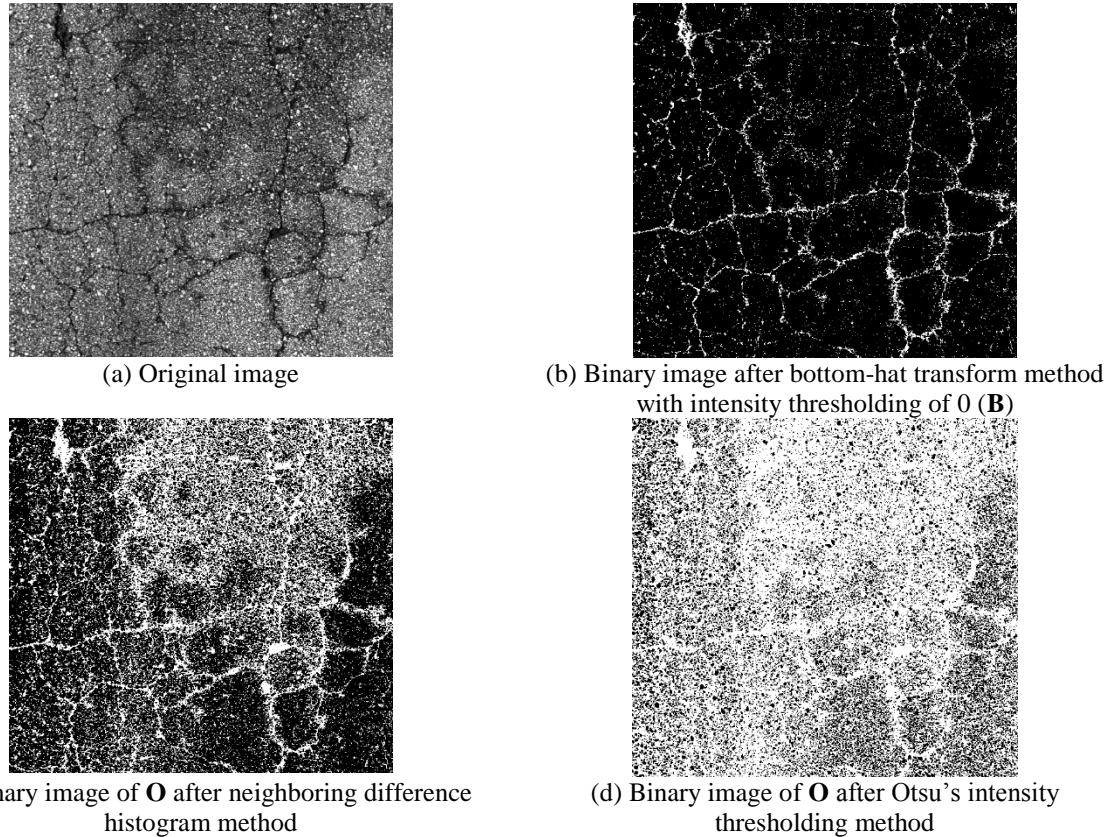


Figure 2-8. Binary images using different intensity-thresholding methods.

2.5 Crack Grouping Using MorphLink-C

In this study, the MorphLink-C is proposed as a novel crack-grouping method to segment crack fragments after crack extraction. It consists of two-step processes:

1. Apply the morphological dilation transform to the binary image of \mathbf{B} that contains crack fragments as

$$\mathbf{D} = \mathbf{B} \oplus \mathbf{S}_D \quad (2-6)$$

where \mathbf{S}_D is the structuring element of the dilation transform; \oplus is the morphological dilation operator; and \mathbf{D} is the resulting binary image after the dilation transform. The dilation is an image operation that “grows” or “thickens” image components, fragmented crack pixels in this case. \mathbf{S}_D is selected to be a square structuring element.

2. Apply the morphological thinning transform to \mathbf{D} to connect the fragments as a continuous crack line within a dilation boundary as

$$\mathbf{T} = \mathbf{D} - \text{hit-and-miss}(\mathbf{D}, \mathbf{S}_T) \quad (2-7)$$

where ‘hit-and-miss’ is the morphological hit-and-miss operator; \mathbf{S}_T is the structuring element for skeletonization; and \mathbf{T} is the resulting binary image after the thinning transform.

After MorphLink-C, grouped crack fragments are labeled per each dilation boundary, and the corresponding crack features can be measured, which represent different characterizations of real crack. An advantage of the proposed MorphLink-C is that it provides a simple and accurate way to measure “averaged” crack width per grouped fragments as

$$(\text{Averaged crack width})_k = \frac{\sum_i (\text{Area of fragmented crack pixels})_k}{\sum_i (\text{Length of the connected crack line})_k} \quad (2-8)$$

where ' $\sum_i (\text{Area of fragmented crack pixels})_k$ ' is the summation of the total areas of the fragmented crack pixels within the k -th dilation boundary; i is the index of the fragmented objects in the k -th dilation boundary; and ' $\sum_i (\text{Length of the connected crack line})_k$ ' is the summation of the total lengths of connected crack line within the k -th dilation boundary. Crack width is an important factor in pavement road rating since the severity of crack deterioration is often controlled by the maximum crack width. For example, FDOT classifies pavement conditions with crack width in road rating: hairline cracks less than or equal to 3.18 mm (1/8 in) for Class 1B, cracks greater than 3.18 mm (1/8 in) for Class 2, cracks greater than 6.35 mm (1/4 in) for Class 3 (Florida Department of Transportation, 2012b).

The proposed MorphLink-C algorithm was used for the crack-grouping process using FDOT's flexible pavement images. A sample result of a single vertical crack is shown in Figure 2-9. The binary image after the bottom-hat transform (**B**) is shown in Figure 2-8b. The numbers shown in the figure are the pixel count of 8-connected neighborhood pixel clusters. Although the bottom-hat transform effectively extract crack components from a flexible pavement image, two problems can be observed. First, the extracted crack objects are fragmented with (multiple) disjoints in their crack paths. Although the image **B** is representational to crack shapes with great local details, the vertical crack has multiple disjoints in its crack paths. Therefore, the crack areas of the 8-connected pixels measured in Figure 2-9b are not a good representation of the area of real vertical crack. Second, the resulting image **B** still contains non-crack objects after removing pavement background. One can recognize from Figure 2-8b that a vertical crack exists in the

middle surrounded by false cracks. That is, the pixel clusters along the vertical crack approximately in $75 < x < 175$ and $45 < y < 350$ range can be considered as “crack” objects. Although the pixel clusters are discontinued at several locations in the crack’s vertical path, one could judge that the crack is a single vertical crack. The rest pixel clusters around the vertical crack can be judged to be “non-crack” objects.

Figure 2-9c shows the result of the fragment grouping using the dilation transform. The boundary range is determined based on relative proximity between the fragments, which can be specified with the \mathbf{S}_D size. In this study, a square structuring element, \mathbf{S}_D , was used, which has the size of 10×10 pixels ($= 9.8 \times 9.8 \text{ mm}^2$). The numbers shown in Figure 2-9c are the summation of the pixel counts of pixel clusters in each boundary. After MorphLink-C, one can observe that the de-fragmented area in Figure 2-9c represents the real vertical crack more accurately than the fragmented areas in Figure 2-9b. For example, the largest boundary containing the vertical crack includes 12 pixel clusters. The pixel count in the largest boundary is 1842 pixels that is the summation of pixel counts of the 12 clusters. The smallest true crack in the largest boundary has 10 pixels at (95, 260). Without the boundary, this crack would be easily misclassified as a non-crack object since the largest non-crack object in Figure 2-9b has 44 pixels at (160, 20). Hence, the accuracy of the crack-classification result would be improved with MorphLink-C.

Figure 2-9d shows the result of the fragment connection using the thinning transform. The numbers shown in the figure are the length of the thinned line that has the width of one pixel. The connection of crack fragments is *adaptive* without involving complicated calculation of crack orientation, length, intensity, etc. for arbitrary crack types, such as single, branched, block and alligator cracks.

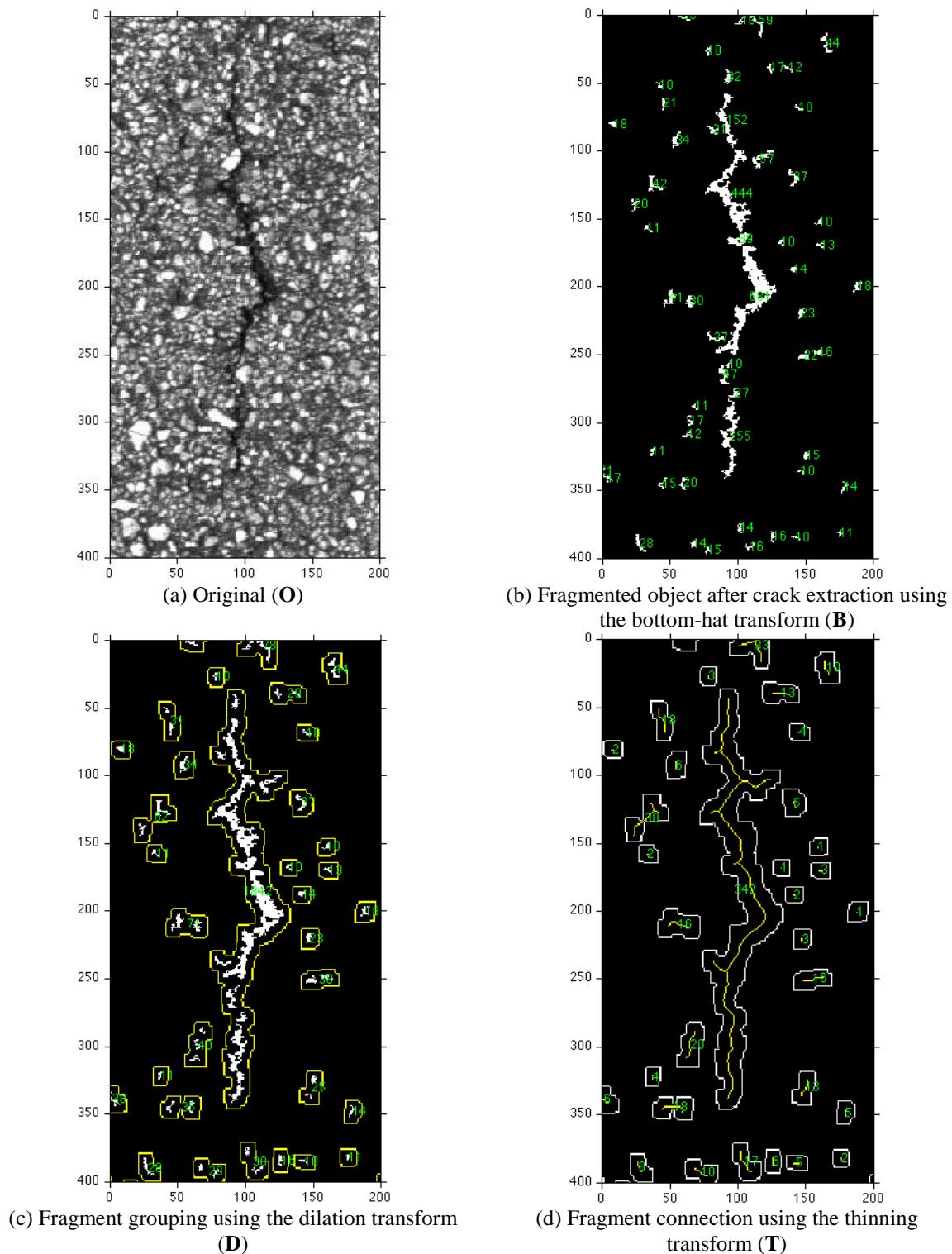


Figure 2-9. Proposed crack-grouping method.

2.6 Discussion

2.6.1 Effects of MorphLink-C on Crack-Detection Accuracy

To evaluate the effect of MorphLink-C on classification accuracy, six features were extracted, which have been commonly used in literatures including area, length, orientation, texture, intensity, and location. For a comparison, the six features were measured before and after MorphLink-C. Before MorphLink-C, the six features were measured for the fragments of 8-connected pixels in image **B**. After MorphLink-C, the six features were measured for the grouped fragments of the dilation boundaries in images **D** and **T**. The feature extraction procedures of the six features are described in Table 2-2.

Table 2-2. Description of the feature-extraction procedures before and after MorphLink-C.

Feature	Before MorphLink-C	After MorphLink-C
Area (mm²)	The area is measured for the fragments of 8-connected pixels in image B .	The area is measured by summing fragment areas within the same dilation boundary in image D .
Length (mm)	The length is measured per the thinned line of fragments, which is obtained with the thinning transform of image B .	The length is measured for the thinned line of the dilation boundary in image T .
Orientation (degree)	The orientation is measured for the fragments of 8-connected pixels in image B . The angle of the orientation is measured between the x-axis and the major axis of the ellipse that contains the 8-connected fragments. The orientation angle ranges between -90 degree and 90 degree, and the counterclockwise is the positive direction.	The orientation is measured per the dilated segment in image D . The angle of the orientation is measured between the x-axis and the major axis of the ellipse that contains the dilated segment.
Texture (co-occurrence index)	The texture is measured using the co-occurrence index of sub-image tile of 200×200 pixels using image B .	The texture is measured using the co-occurrence index of sub-image tile of 200×200 pixels using image D .
Intensity (8-bit scale)	The intensity is calculated by averaging the intensities of the original grayscale image O per the 8-connected pixels.	The intensity is calculated by the summation of the intensity of the original grayscale image O weighted by fragment area in each dilation boundary.
Wheel path designation (0, 1 and 2)	CW = 2, CO = 1, and outside of lane markers = 0 in image B .	CW = 2, CO = 1, and outside of lane markers = 0 in image D .

Among them, the texture feature was considered since rough pavement surface has higher potential for developing cracks. It was calculated using the co-occurrence matrix which can quantify the surface roughness in different regions of the pavement surface. The co-occurrence matrix, first defined by Haralick (1975), is based on the vicinity (offset) of pixels with certain intensities. Mathematical expression of the co-occurrence matrix is as follows:

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1 & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0 & \text{otherwise} \end{cases} \quad (2-9)$$

where C is the co-occurrence matrix; i and j are the grey levels; Δx and Δy describe the offset between the pixels; and $I(p, q)$ is the intensity of the image at (p, q) in a $(m \times n)$ pixel segment of the image. In simple words, the co-occurrence matrix presents the occurrence frequency of two intensity values at a given offset. In this study, the co-occurrence matrix was calculated using a binary image, which is a symmetrical 2×2 matrix in which element $C(0,0)$ representing the co-occurrence frequency of two non-crack pixels next to each other and $C(1,0)$ or $C(0,1)$ representing the co-occurrence frequency of neighboring crack and non-crack pixels. The value of $C(1,1)$, which represents the co-occurrence frequency of neighboring crack pixels, is considered as *the co-occurrence index* of the region of interest. A region with a higher number of the co-occurrence index can be considered to be rougher than a region with a lower index. Detailed background and mathematical formulation of the co-occurrence matrix is presented in Kaseko & Ritchie (1993). The co-occurrence matrix method is a region-based method as opposed to a pixel-based. Thus, to obtain the co-occurrence matrix in this study, first each image was

divided into regions with 200×200 pixels, and the co-occurrence matrix was calculated for each region. Then the co-occurrence index of each region was normalized by dividing to the number of pixels in that region.

The position feature was employed to incorporate the wheel-path designation for road rating. According to FDOT (2012a), flexible pavement regions confined to wheel path (CW) are subjected to more frequent traffic loading than the regions outside of wheel path (CO) as shown in Figure 2-4. Consequently, CW has higher potential for developing fatigue cracks. To obtain the position features before and after MorphLink-C, the lane markers were detected from the road image, and the CW and CO regions were divided based on the ratios of the distance between the left and right lane markers: $CO:CW:CO:CW:CO = 1:2:2:2:1$. Then, image components in CW were assigned to two, image components in CO were assigned to one, and image components on and outside the lane markers were assigned to zero. When only one lane marker can be observed in the road image, CW and CO were determined based on the distances from inside the lane marker ($CO:CW:CO:CW:CO = 0.46 \text{ m}: 0.91 \text{ m}: 0.91 \text{ m}: 0.91 \text{ m}: 0.46 \text{ m}$) since the distance can be measured based on the pixel resolution (e.g., 0.98 mm per pixel in this study). When both lane markers are not available in the pavement image, CW and CO were determined based on the distances from the centerline of the image by assuming the center of the lane was located at the image centerline.

In order to find the optimal crack feature subset, an ANN classifier was employed, which might be one of most popularly used supervised classification methods for crack detection. The ANN is a network of simple processing units called neurons which are connected using a set of weights and biases. The ANN classifier that was used in this study consisted of a seven-layer feed-

forward network, including one input layer and one output layer. Each hidden layer contained 10 neurons. Tangent-sigmoid and linear transfer functions were used in the hidden and outputs layers, respectively. The input of the classifier was the six features. Having the linear function in the output layer, the output of the classifier was a real number ranged between 0 and 1. Since the output was a non-crack object toward 0 and a crack object toward 1. Therefore, a threshold should be set between 0 and 1 to classify an image object into non-crack when its output is smaller than the threshold or crack when its output is greater than equal to the threshold. A schematic of the ANN configuration is shown in Figure 2-10.

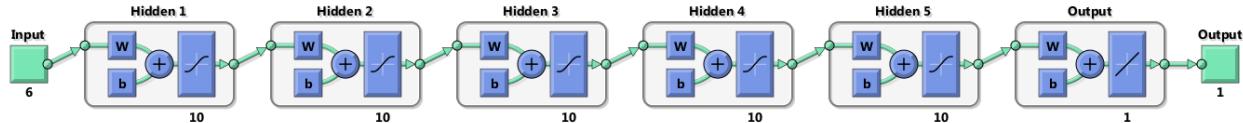


Figure 2-10. A schematic of the ANN configuration.

Being a supervised machine learning method, training the ANN classifier is the process of adjusting these weights and biases, so that the classifier produces desirable network outputs. For the network optimization, the Levenberg-Marquardt learning method was used, associated with a mean-square-error (MSE) function to minimize the error between the network output and target output:

$$\text{MSE}(O) = \frac{1}{n} \sum_{i=1}^n (T - O)^2 \quad (2-10)$$

where O is the $(n \times 1)$ output vector of neural network; T is the $(n \times 1)$ vector of target values; and n is the number of data.

Being a supervised machine learning method, the ANN classifier requires a training dataset for optimizing the values of weights and biases. Training datasets were prepared through manual selection of true cracks in a total of 26 flexible pavement images; thus unselected pixels were categorized into false cracks. Correspondingly, target outputs for ‘crack’ and ‘non-crack’ objects were considered to be 1 and 0, respectively. A summary of selected images along with the total number and length of image components is shown in Table 2-3. Before MorphLink-C, a total of 512731 objects were extracted, including 26656 (5.20%) crack and 486075 (94.80%) non-crack objects in number. The total length of the objects was 6499.8 m, including 487.2 m (7.50%) for crack and 6012.6 m (92.50%) for non-crack objects. After MorphLink-C, a total of 264981 objects were extracted, including 3090 (1.17%) crack and 261891 (98.83%) non-crack objects. The total length of the object was 3915.1 m, including 444.3 m (11.35%) for crack and 3470.8 m (88.65%) for non-crack objects. Therefore, one can observe that the ratio of the crack and non-crack objects is unbalanced both in the number and length.

Table 2-3. Training datasets for the ANN classification before and after MorphLink-C algorithm. All objects include both crack and non-crack image components.

Image #	Before the proposed crack grouping method (Without MorphLink-C)				After the proposed crack grouping method (With MorphLink-C)			
	Cracks (#)	Cracks (mm)	All objects (#)	All objects (mm)	Cracks (#)	Cracks (mm)	All objects (#)	All objects (mm)
1	22	325	15780	255569	3	273	7519	149979
2	64	849	15486	268241	13	905	7345	149701
3	481	8980	16802	212659	60	7718	8498	128759
4	536	12715	18504	239567	70	10876	9138	138994
5	190	2766	19978	219915	15	2984	10991	129565
6	317	4069	19742	209313	50	4625	10865	129312
7	476	5700	18673	202897	94	6280	10685	118494
8	402	6876	20072	225010	42	7244	10800	132737
9	465	10403	20571	253828	57	8888	10429	141403
10	185	2360	21063	258457	38	2586	10108	150859
11	0	0	20269	239926	0	0	10008	143116
12	1062	23217	19967	274102	141	17584	9934	162640
13	1410	27231	19310	276034	151	20726	10396	151451
14	1455	28456	19716	259098	116	23428	10477	151652
15	81	1140	21339	233896	2	1214	12184	153605
16	11	173	20689	229400	2	203	12035	145013
17	2448	41169	21711	257300	310	40253	10937	170581
18	1489	30906	21852	257122	144	29242	11122	166114
19	1434	23436	21906	260571	212	22422	10705	166817
20	1779	31985	19770	277335	183	26962	10230	157482
21	2257	39468	19637	303304	213	33843	9843	173981
22	2020	34204	19818	291783	252	27005	10156	167410
23	2014	39459	21672	259987	219	38955	10984	168922
24	1876	32760	21508	260848	239	34580	11072	166833
25	3006	58015	20959	276174	320	55307	10219	179091
26	1176	20542	15937	197481	144	20203	8301	120591
Sum	26656	487204	512731	6499817	3090	444306	264981	3915102
Mean	1025.2	18738.6	19720.4	249993.0	118.8	17088.7	10191.6	150580.8
Stdv	890.2	16481.4	1870.9	27345.4	99.6	15288.2	1189.5	17061.2

After crack extraction before and after MorphLink-C, all the six features were measured from the image objects. Since the datasets were extracted from 26 images, the entire image objects were randomly shuffled to avoid the biasness per image, and partitioned into 60% for training, 20% for validation and 20% for testing to compare two ANN classifier models trained before and after MorphLink-C.

The crack-detection accuracy of the ANN classification is shown in Figure 2-11. Since the classifier output is a real number between 0 and 1, the classification result depends on the threshold value: the output less than the threshold is classified into a non-crack object, and the output greater than or equal to the threshold is classified into a crack object. Figure 2-11a shows the apparent success of the ANN classification before and after MorphLink-C. It was calculated the summed length of correctly classified objects divided by the summed length of total objects. The result shows that the apparent success was improved after MorphLink-C for all thresholds: the apparent success after MorphLink-C is higher than 98% for all threshold, while the apparent success before MorphLink-C is higher than 85% for all threshold.

Since the datasets are unbalanced between the crack and non-crack objects, more rigorous analysis would be needed to evaluate the crack-detection accuracy. A statistical hypothesis test was conducted to analyze the crack-classification errors before and after MorphLink-C. The null hypothesis (H_0) was defined as “*the object is NOT crack*”, and the alternative hypothesis (H_a) was defined as “*the object is crack*”. Hence, the false positive (or Type-I) error happens when the classifier says that the pixel object is a crack, but the object is not crack in truth. The false negative (Type-II) error happens when the classifier says that the pixel object is not a crack, but the object is crack in truth. Figure 2-11b compares the results of false positive before and after MorphLink-C. Overall, the false positive increases as the threshold increases. The false positive increases without MorphLink-C when the threshold is less than or equal to 0.2, while the false positive increases with MorphLink-C when the threshold is greater than 0.2. In many applications, false positive is considered acceptable when its percentage ($p = a$) is less than 5%. Hence, the level of

false positive is acceptable both before and after MorphLink-C when the threshold is greater than or equal to 0.3.

Figure 2-11c compares the percentages of the false negative before and after MorphLink-C. Overall, the false negative increases as the threshold increases. One can also observe that the false negative decreases after MorphLink-C for all thresholds. The percentage of the false negative ($p = b$) after MorphLink-C is lower than 20% when the threshold is smaller than or equal to 0.3. Since a hypothesis test is usually considered as a good test when the power of test ($p = 1 - b$) is higher than 80%, the threshold of 0.3 after MorphLink-C was chosen to be the optimal threshold of the ANN classifier in this study.

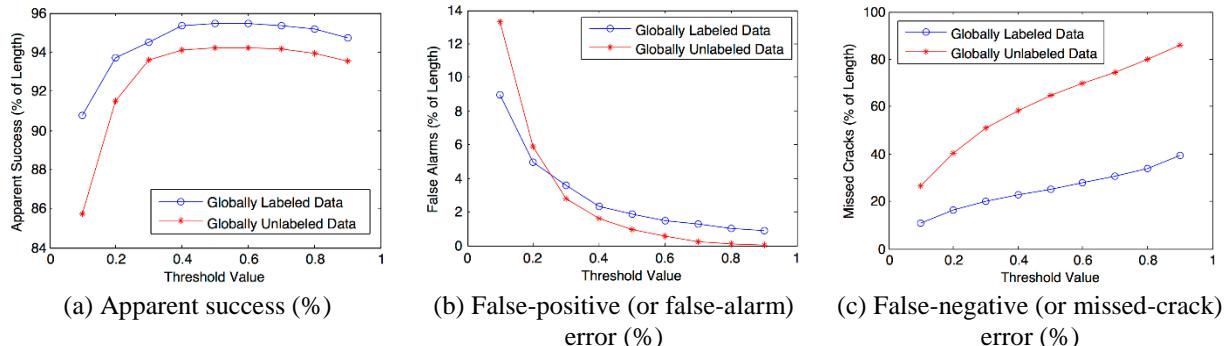


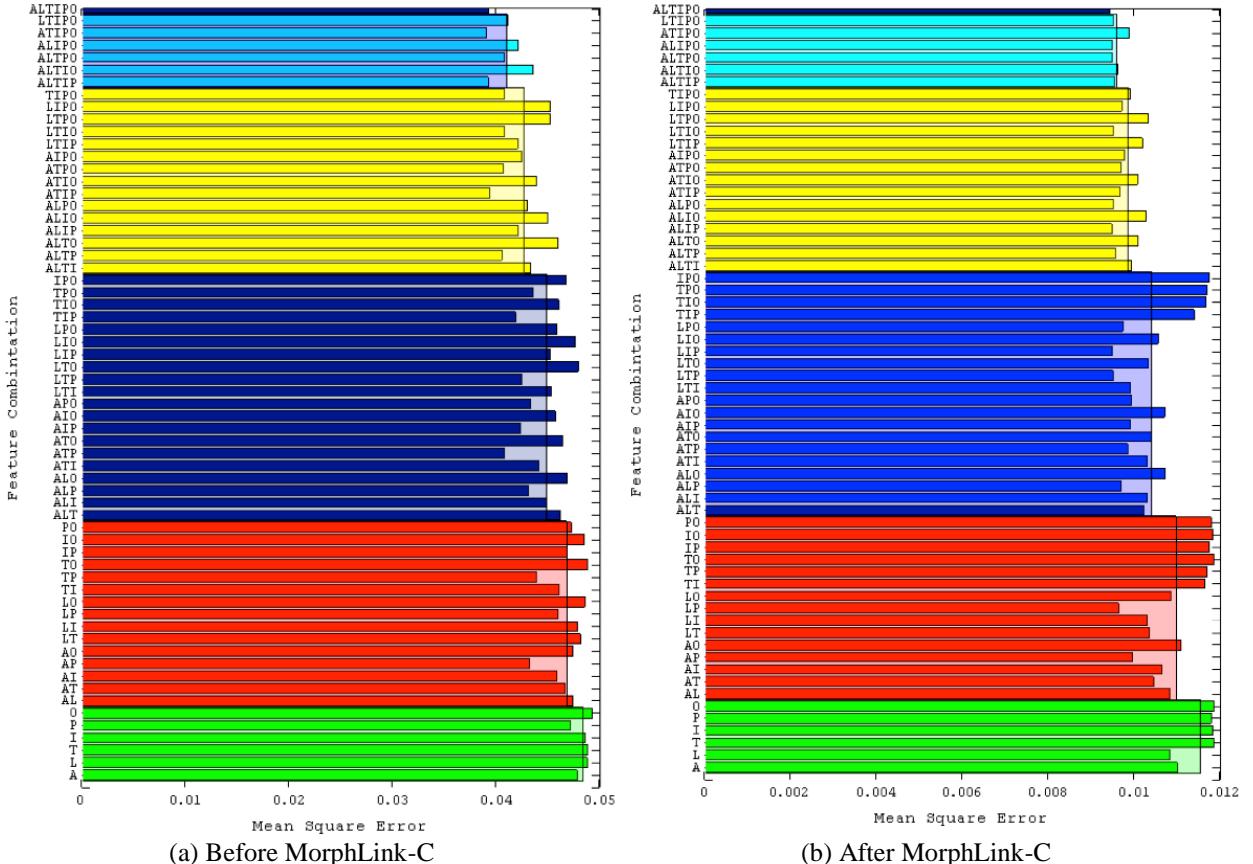
Figure 2-11. The error analysis results of the ANN classifier for crack length with and without the MorphLink-C algorithm.

2.6.2 Effects of MorphLink-C on Crack Feature Subset Selection

In general, when more features are used in classification process, classification accuracy increases while training time increases. Since a large number of pavement image data are usually involved in road-rating applications, however, finding the optimal feature subset is important to

develop an effective crack-recognition method. In the previous section, it was shown that the crack-detection accuracy could be improved after the proposed MorphLink-C. In this section, the effects of MorphLink-C on FSS are discussed to find the optimal crack feature subset, considering both the crack-detection accuracy and computation time for classifier training. Hence, a wrapper feature evaluation is used since the ANN classifier is used in this study.

Considering the all six features, a total of $2^6 - 1 = 63$ combinations of feature subsets were considered. The data were partitioned into 60% in the classifier training, 20% in validation and 20% in testing. The ANN classifier was trained and validated, and the crack-detection accuracy was calculated with the MSE in Equation 2-10 using the testing data. This procedure was repeated 5 times to ensure that the network is trained properly avoiding the local minima. The results of the MSE calculation for different feature combinations are shown in Figure 2-12. The results show that overall MSE decreases significantly after MorphLink-C about 2.5 times. This result confirms that the proposed MorphLink-C improves the crack-detection accuracy in terms of MSE for different combinations of crack feature subsets. It is also observed that MSE decreases when more features are used in the classifier training both before and after MorphLink-C. When one feature is used, the reduction of MSE is more sensitive after MorphLink-C to the area (0.0110) and length (0.0108) than the other features, the texture (0.0118), intensity (0.0118), position (0.0118) and orientation (0.0118), while no significant difference is observed before MorphLink-C, in which the numbers in the parentheses show the corresponding MSE. Thus, it can be postulated that the area and length features were directly affected by grouping crack fragments after MorphLink-C, which resulted in improving the crack detection accuracy.



* The y-axis shows the crack area (A), length (L), texture (T), intensity (I), wheel-path location (L), and orientation (O).

Figure 2-12. The mean-square-errors for different feature subset combinations.

Table 2-4 summarizes the averaged MSE and training times of the ANN classification for different numbers of crack feature combinations. In the analysis, the resolution of the image was 5940×4044 , and the image size was about 5.75 MB. The computer was equipped with an Intel® Core™ i7 – 2600 3.40 GHz CPU and 16 GB RAM. Overall, as more features are used in training, the training time increases while the MSE decreases both before and after MorphLink-C. However, the extrema of the training time shows fluctuations while the MSE is relatively stable. Thus, larger training time does not necessarily yield a more accurate crack detection result. Fluctuation in

training time would be largely affected by the initial biases and weights of the ANN classifier. Using MorphLink-C, both the MSE and training time are reduced for all combinations of feature subsets: overall 69.3 % for MSE and 76.52%. Therefore, the results show the proposed MorphLink-C can improve the detection accuracy in terms of the MSE. Reduced training time with MorphLink-C is practically important when a large number of pavement images need to be trained, which is true in many road rating applications.

Table 2-4. Averaged MSE and training time of the ANN classification before and after MorphLink-C.

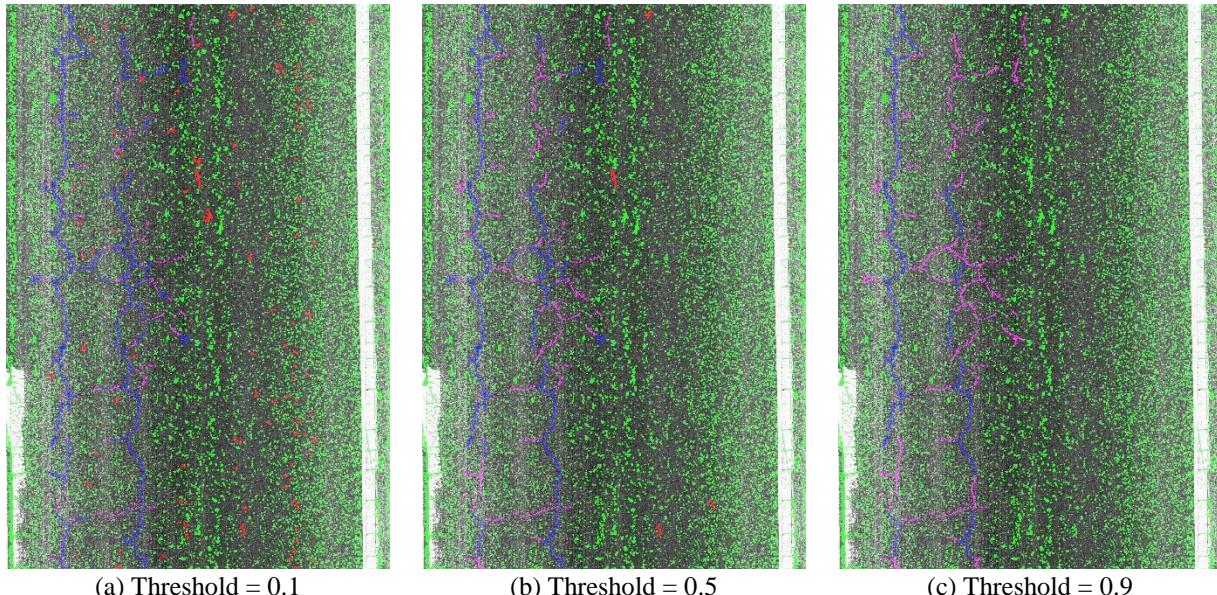
# of features	Before MorphLink-C						After MorphLink-C					
	Avg. MSE	Max. MSE	Min. MSE	Avg. time (min)	Max. time (min)	Min. time (min)	Avg. MSE	Max. MSE	Min. MSE	Avg. time (min)	Max. time (min)	Min. time (min)
1	0.0483	0.0492	0.0471	63.1	209.9	4.8	0.0115	0.0118	0.0108	16.8	50.9	2.1
2	0.0467	0.0488	0.0431	209.2	466.9	16.8	0.0110	0.0118	0.0096	57.3	158.2	7.5
3	0.0447	0.0478	0.0407	316.4	852.2	2.0	0.0104	0.0117	0.0095	90.3	260.1	11.9
4	0.0426	0.0459	0.0393	382.2	608.3	70.5	0.0098	0.0114	0.0093	105.6	247.8	25.8
5	0.0414	0.0434	0.0390	289.4	428.3	213.3	0.0096	0.0099	0.0095	83.2	116.7	37.2
6	0.0391	0.0391	0.0391	371.0	371.0	371.0	0.0094	0.0094	0.0094	167.1	167.1	167.1

Computation times of the proposed crack-recognition approach were also measured using 100 flexible images, and the computation times for different levels of crack recognition are summarized in Table 2-5. The average computation times for fragment grouping and fragment connection are 0.04 and 0.76 seconds, respectively. Therefore, the computation time of the proposed MorphLink-C is significantly shorter than that of crack extraction using the bottom-hat transform. The computation time could be further reduced by implementing the algorithms using compiled language, such as C or C++, since the tested algorithm was coded using a script language using Matlab®.

Table 2-5. Computational times for the proposed morphological crack quantification method.

Level of crack recognition	Avg. time (sec)	Max. time (sec)	Min. time (sec)	No. of images
Crack extraction using bottom-hat transform	33.37	34.15	32.63	100
Fragment grouping using dilation transform	0.04	0.05	0.03	
Fragment connection using thinning transform	0.76	1.42	0.41	

To demonstrate the effects of FSS, the classification results using the all six features are compared with the results using three features. First, the classification results using the all six features for different thresholds are shown in Figure 2-13. It is shown that the falsely detected objects (red) increases when the threshold is close to zero, while the falsely filtered object (pink) increases when the threshold is close to one. Figure 2-14 shows the classification result using the three-feature subset of the length, texture and position. When the threshold is 0.1, the result is similar to the result with the six features in Figure 2-13a, except the falsely detected objects on the lane marker in the lower left corner. When the threshold is 0.5, the vertical crack in the left wheel path is falsely filtered, compared to the result with the six features in Figure 2-13b. When the threshold is 0.9, more falsely filtered objects can be observed, including the alligator crack in the mid-left of the image. The MSE of the six-feature case is 0.0094, while the MSE of the three-feature case is 0.0105.



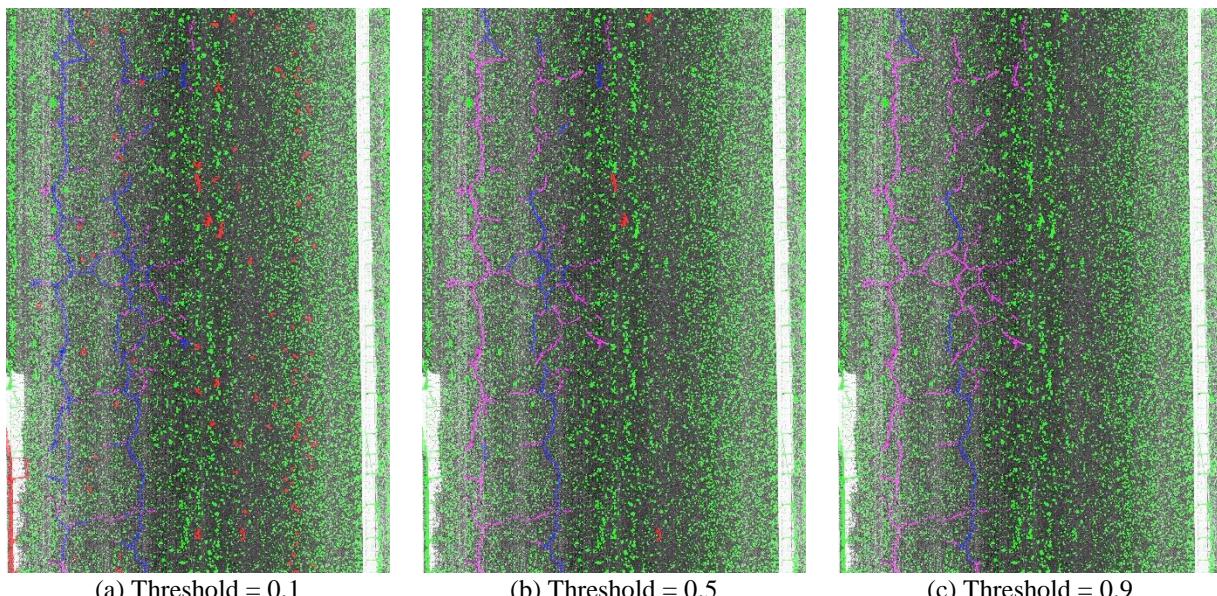
(a) Threshold = 0.1

(b) Threshold = 0.5

(c) Threshold = 0.9

* blue: correctly detected objects, green: correctly filtered objects, red: falsely detected objects, and pink: falsely filtered objects.

Figure 2-13. Sample classification results using all the six features after MorphLink-C.



(a) Threshold = 0.1

(b) Threshold = 0.5

(c) Threshold = 0.9

* blue: correctly detected objects, green: correctly filtered objects, red: falsely detected objects, and pink: falsely filtered objects.

Figure 2-14. Sample classification results using three features of the length, texture and position after MorphLink-C.

Finally, Figure 2-15 shows sample results of the crack extraction using the bottom-hat transform, and the crack detection using the ANN classifier combined with MorphLink-C. The results show that the proposed crack-recognition method not only can detect crack accurately, but also can successfully reject non-crack pavement background.

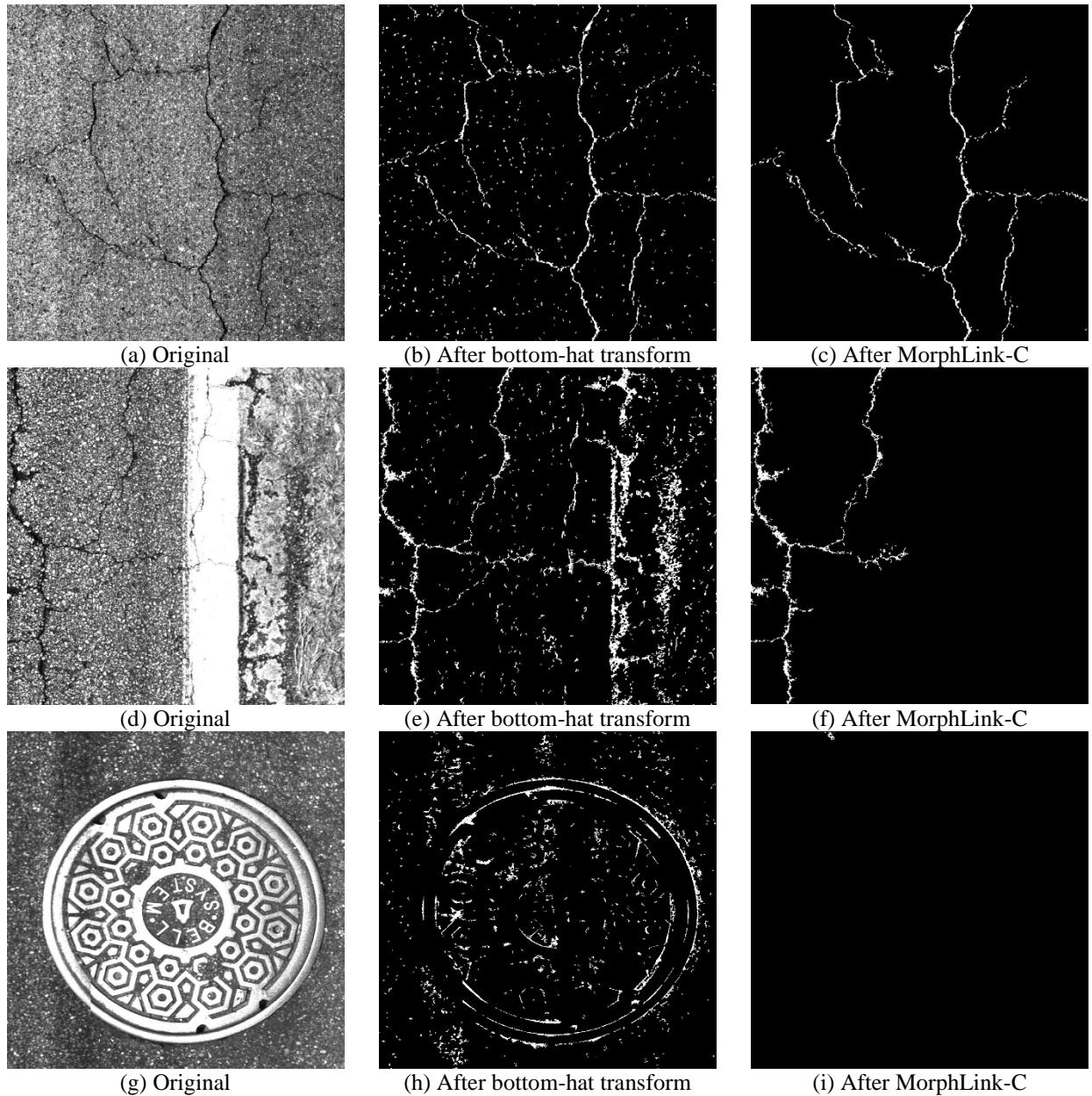
2.6.3 Adaptive Property of MorphLink-C

Figure 2-16 demonstrates the adaptive property of the proposed MorphLink-C. The effects of the S_D size are investigated, which determines the level of the crack-line abstraction after the thinning transform. An original gray-scale pavement image is shown in Figure 2-16a. The image shows severe alligator cracks as well as some local spallings. The pavement surface was smeared due to partial wetness in top-left corner and along some parts of the cracks and spallings. The image T was obtained through the crack-recognition procedures proposed in Figure 2-2. The results of crack-line abstraction are shown in Figures 2-16b to 2-16g for an increasing S_D sizes from 1×1 pixel to 20×20 pixels.

Some horizontal cracks in wet areas are not detected with a structuring element smaller than 5×5 pixels. One can see in the original image that the horizontal cracks in the wet areas are less distinguishable than the cracks in the dry areas; thus the ANN-classifier misses those cracks when a small dilation boundary is used. When the S_D size increases, the small cracks that have been missed earlier began to be detected, which improved the crack classification accuracy. It is recommended that the square structuring element size be selected between 10 mm and 20 mm for flexible pavement. For a smaller structuring element, the crack lines are discontinued and more representational to local details (e.g., the local spallings). When the S_D size increases, the disjointed

crack lines in Figure 2-11b get connected adaptively without involving complicated computation of crack-path orientations, lengths, and intensities. In addition, the level of crack-line abstraction increases with a larger $\mathbf{S_D}$ size.

The $\mathbf{S_D}$ size can be determined based on the proximity limit to neighboring cracks. The $\mathbf{S_D}$ size determines the proximity limit of neighboring cracks. For example, with a 15-pixel (14.7-mm) square structuring element, neighboring cracks within a 29.4-mm range will be considered as a part of the same crack. The selection of the $\mathbf{S_D}$ size depends on the application. In this study, the $\mathbf{S_D}$ with the size of 10×10 pixels ($= 9.8 \times 9.8 \text{ mm}^2$) was used.



* (a)-(c) branched cracks, (d)-(f) branched cracks with white lane marker, and (g)-(i) Manhole.

Figure 2-15. Sample crack-classification results using the ANN classifier.

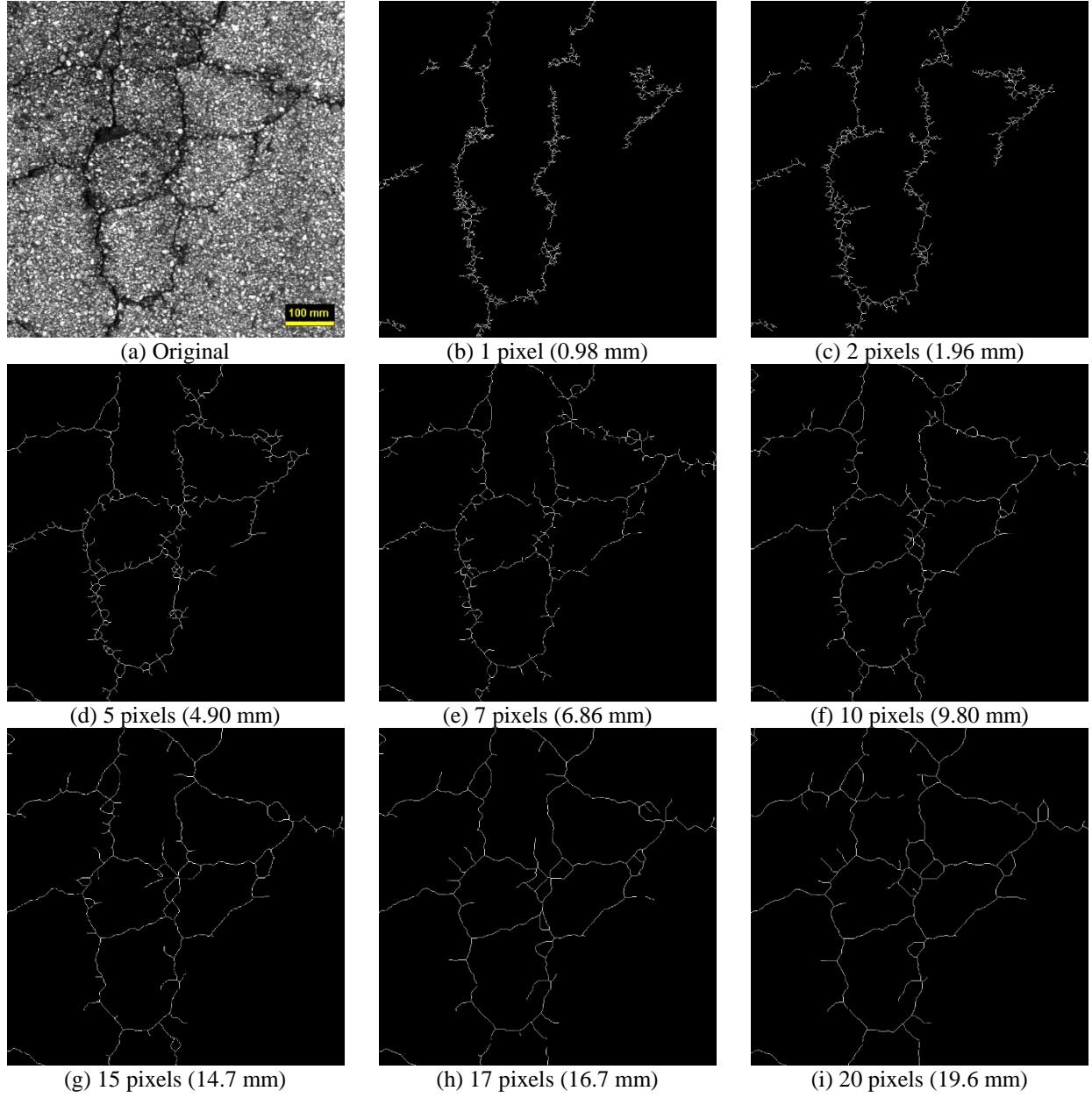


Figure 2-16. Adaptive process of the proposed fragment connection approach for different S_D sizes.

2.7 Conclusions

An automated crack-recognition method was proposed, which consists of the crack-extraction, crack-grouping, and crack-detection processes to address multi-level image processing goals in road rating applications. The MorphLink-C algorithm was proposed as a novel crack-grouping method. It is composed with two sub-processes, including fragment grouping using the dilation transform, and fragment connection using the thinning transform. The proposed crack-grouping method is different from He, Qiu, Wang, Zhang, & Xie (2011) and W. Huang & Zhang (2012) that also utilized the dilation and thinning transforms. In these studies, the dilation transform was applied for multiple times to crack fragments until they were connected to each other. Then, the thinning transform was applied to obtain crack skeleton. This method has two drawbacks. First, the stopping criterion of multiple dilation transforms is not clear, particularly when a noise is closely located to crack. Second, the method did not provide the way to measure crack width that is an important measure in road rating applications.

The effects of the crack-grouping process were evaluated on the crack-detection accuracy and feature subset selection, which have been rarely discussed in existing literatures. Using the ANN classifier, the followings can be concluded from the results:

- A common problem of crack extraction algorithms is that extracted crack image components are usually fragmented, having (multiple) disjoints in their crack paths. The proposed MorphLink-C improves overall crack-recognition performance by grouping crack fragments, which enables better characterization of crack features. It can be used with any crack-extraction algorithm as a crack de-fragmentation technique.

- MorphLink-C increases the crack-detection accuracy by reducing the false negative error. Lowering the false negative error is important in pavement crack recognition since the image after crack extraction usually includes relatively a small number of crack objects compared to no-crack objects.
- Using MorphLink-C, the crack-detection accuracy increases for all tested 63 combinations of crack feature subsets as well as the classification training time decreases. These are advantageous in road rating applications to reduce the training time with reduced feature subsets, which involves a large number of pavement images.
- The proposed fragment-connection method is adaptive for different crack types, including single, branched, blocked and alligator cracks, without involving time-consuming computation of crack orientation, length, and intensity.
- MorphLink-C provides an accurate way of computing “averaged” crack width that is an important measure in road rating applications. Using existing crack-segmentation algorithms, such as a seed-growing method (Q. Li et al., 2011) and crack-tree method (Zou, Cao, Li, Mao, & Wang, 2012), and tensor-voting method (J. Huang et al., 2014), crack width cannot be measured although they have been effectively used to connect adjacent crack fragments.
- Although the wrapper FSS method was used in this study to evaluate crack-detection accuracy associated with a classifier for different combinations of crack feature subsets, the relevance, irrelevance and redundancy of the features were not known. Since correlational statistics of crack features are important to understand road deterioration severity and patterns, future study would be necessary. For this, filter FSS method

could be used to evaluate feature subsets by their statistical information contents, such as interclass distance, statistical dependence, entropies or information-theoretic measures.

CHAPTER 3: A NOVEL ADAPTIVE IMAGE-PROCESSING TECHNIQUE TO CONNECT FRAGMENTED CRACK PIXELS FOR FLEXIBLE ROAD PAVEMENT

3.1 Introduction

Crack is an important distress type that influences short-term and long-term road pavement performance. Effective pavement rehabilitation polices can only be established with reliable prediction of future pavement cracking rates based on quantitative assessment of past and present pavement conditions. Many image-based crack-recognition techniques have been employed to provide necessary quantitative measures of cracks in pavement surface images.

Crack recognition for flexible road pavement is a challenging application for two main reasons. First, crack patterns are highly arbitrary, such as single, branched, block and alligator cracks that are observed at different stages of road deterioration. Second, it is difficult due to noisy random backgrounds of pavement aging, coating types, surface roughness, spots and stains, oil and water spills, road markings, or manholes. Numerous crack-extraction techniques have been developed for arbitrary crack objects from random pavement background. Some of many crack-extraction techniques include intensity thresholding, edge detection, texture analysis, seed growing, wavelet transform, empirical mode decomposition, artificial neural network, and fuzzy set theory (Wu, Mokhtari, Nazef, Nam, & Yun, 2014). A common problem of crack images after extraction is that they are fragmented due to (multiple) disjoints in their crack paths, which can be misleading in characterizing crack features. Another problem is that resulting crack images include

non-crack objects due to random background noises, which could lower crack recognition accuracy.

Crack in computer vision can be defined as *a group of low-intensity pixels compared to neighboring pixels, which forms into an arbitrary line shape in length, width and directions (crack-pixel level)*. One or more crack pixels can further form into a complicated crack network of a single line, branched line or polygon, which are not necessarily continued in its line paths due to the fragmentation of the crack pixels (crack-network level). In order to deal with the multi-level topological shapes of crack images, different image-processing levels need to be employed for computer-aided crack recognition:

Crack extraction (Level 1) is defined to extract crack pixels by removing non-crack (random) background of the input image. Crack pixels extracted in this pixel-level image processing are often fragmented and discontinued in crack-line paths. The connectivity between extracted crack pixels is not yet determined in this level.

Crack grouping (Level 2) is defined to group fragmented crack pixels extracted in level 1 by image segmentation. This process is to bridge the crack-pixel level to crack-network level by determining the connectivity between fragmented crack pixels. Once the connectivity is determined, crack pixels are labeled with a crack-identification number per each crack group. Different crack features, such as crack length, area, width, orientation, intensity, and locations, can be also measured per each crack group.

Crack detection (Level 3) is defined to classify crack and non-crack image components per each crack group. Different classification algorithms can be employed to filter out non-crack image components based on the crack features measured in level 2.

Crack classification (Level 4) is defined to classify a crack image into single, branched, blocked, alligator cracks, etc.

Researchers have developed some crack-segmentation techniques to link disjointed crack image components. As the first class of the methods, Sun, Salari, and Chou (2009) developed a simple method to connect the end points of two adjacent crack fragments with a straight line when the endpoints are within a certain window size (e.g., 4 rows and 20 columns for horizontal crack). In W. Huang and Zhang (2012), two fragments were connected at the nearest points between them with a straight line instead of two ends points when the distance was less than a certain threshold. In Liu, Xu, Yang, Niu, and Pan (2008), two fragments were connected at their end points with a straight line. However, the connectivity was determined based on crack length, orientation and gap distance in their study. Another class of crack-segmentation methods is to determine the connectivity based on crack seeds. Li, Zou, Zhang, and Mao (2011) developed a seed-growing method, called FoSA, associated with linking algorithm to connect adjacent crack elements and pruning algorithm to remove short branches of the linear structures and extract the main crack globally. Zou, Cao, Li, Mao, and Wang (2012) also developed a seed-based method, called CrackTree. In this method, crack position and shape were determined using the minimum spanning tree (MST) that could be calculated from the crack probability map based on tensor-voting technique. J. Huang, Liu, and Sun (2014) also applied the tensor-voting technique in their pavement crack detection application combining a 2D image with 3D information based on Dempster-Shafer theory. A drawback of these approaches is that crack width cannot be measured, which is an important measurement to control the level of crack distress in many maintenance applications. Morphological technique has also been used to link disjointed crack components.

Chu (2010) developed a crack-linking technique for bituminous pavement by combining a series of morphological operations. After obtaining crack skeleton with thinning transform, dilation transform was applied to fill the gaps between crack disjoints that are smaller than the size of the applied structuring element. Then erosion transform was used to restore original crack skeleton after filling the gaps. He, Qiu, Wang, Zhang, and Xie (2011) and W. Huang and Zhang (2012) also used the morphological thinning and dilation transforms but in different procedures. First an edge detection technique was applied to extract crack object from a pavement image. Since resulting crack images are usually fragmented, they applied the dilation transform to fragmented crack images for multiple times to expand each fragment area until they were connected each other. Then the thinning transform was applied to obtain crack skeleton. Being morphological approaches, the above techniques have an advantage to selectively connect gaps that are randomly located in a pavement image. However, crack width cannot be measured using these techniques, and the stopping criterion of the last two approaches cannot be clearly determined particularly when crack fragments are closely located to background noises. Moreover, the effects of most crack-segmentation techniques have been rarely evaluated on the accuracy of crack detection through statistical error analysis.

In this study, it is hypothesized that an effective crack-grouping method can provide better representation of real crack features, which is important for reliable road deterioration assessment. It is also hypothesized that an effective crack-grouping method can improve crack-detection accuracy by grouping crack fragments that are extracted with a crack-extraction algorithm. Therefore, the objectives of this study are two-folded. The first objective is to develop a novel crack-grouping algorithm, called MorphLink-C, to deal with crack fragmentation problems. The

second objective is to investigate the effect of the proposed algorithm on crack-detection accuracy. Hence, this study mainly focuses on Levels 2 and 3 in the above crack-recognition processes. The proposed algorithm consists of two sub-processes: i) fragments grouping using morphological dilation transform, and ii) connection of fragments using morphological thinning transform. The proposed MorphLink-C algorithm can be used with any crack-extraction method in pixel level to connect fragmented crack pixels that are placed in crack-line paths. This algorithm is not limited to a simple crack shape but can be used for complicated crack shapes in network level, such as single cracks, branched cracks, block cracks, and alligator cracks, whose classification is important to evaluate the severity of crack-induced damages in road rating applications. In road rating applications, it is important to crack width since crack-induced pavement distress is usually controlled based on crack width by most road maintenance authorities. The proposed method provides a simple but accurate way to estimate an “averaged” crack width that varies in its complicated crack paths. The procedures of the proposed MorphLink-C algorithm are shown in the second gray box in Figure 3-1.

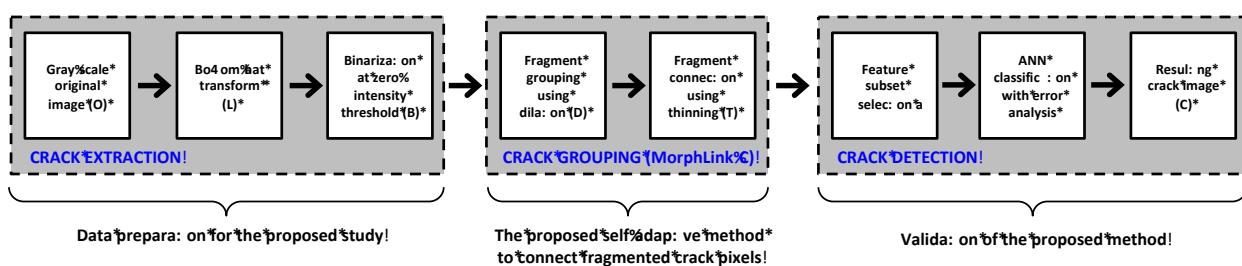


Figure 3-1. The overview of the proposed study of MorphLink-C.

To validate the proposed crack-grouping algorithm, first a crack-extraction method has to be employed to prepare necessary crack (and non-crack) image objects extracted from realistic

crack images. Flexible pavement images were collected using a Multi-Purpose Survey Vehicle (MPSV) equipped with a Laser Road Imaging System (LRIS) by Florida Department of Transportation (FDOT) for the state road surveying. Once road surface images are obtained, the morphological bottom-hat transform originally developed by Salembier (1990) is used to extract crack pixels that are necessary to validate the proposed MorphLink-C algorithm (Jahanshahi, Kelly, Masri, & Sukhatme, 2009; Pratt, 2001). The morphological image-processing technique has been applied in road crack extraction applications. It utilizes mathematical morphology as a tool for extracting image components that are useful in the representation and description of various region shapes, such as crack-like shapes in this study. The mathematical foundation of the morphological technique is based on the theory of set algebra by Minkowski (1903) and on theory of topology by Matheron (1975). In the morphological technique, an input image is processed based on spatial relationship between pixels with pixels and an image with a structuring element have been employed in pavement crack detection. Here, the structuring element is a structure with certain shape and size used to define object pixels and neighbor pixels. More detailed description about the morphological technique will be presented in the subsequent section. The necessary data preparation procedures are shown in the first gray box in Figure 3-1.

In order to achieve the second objective of this study, the effect of MorphLink-C is investigated on the accuracy of crack detection. The Artificial Neural Network (ANN) classifier that might be one of most commonly used classification methods in crack-detection applications is employed to evaluate how the proposed algorithm improves the accuracy of crack detection associated with the ANN classifier. A total of six features that are commonly used in literatures are extracted from crack (and non-crack) image objects. Then the classifier is trained using the

features before and after MorphLink-C. Finally, their false positive and negative errors in crack detection are compared through statistical hypothesis tests. The validation process is shown in the third gray box of Figure 3-1.

The outline of this study is as follows. A brief overview of morphological image-processing technique is presented in the Section 3.2 to provide background for the understanding of crack extraction and grouping techniques used in this study. The procedures of data preparation are described to obtain pavement surface images and fragmented crack (and non-crack) image components that are necessary to evaluate the proposed crack-grouping algorithm are presented in Section 3.4. Section 3.5 of this study describes the proposed MorphLink-C algorithm and the procedures of image processing. The procedures of the six-feature extraction, including area, length, orientation, texture, intensity, and position are also described in this section. Finally, Section 3.5 discusses the effects of MorphLink-C on crack-detection results using ANN classifier to evaluate how the proposed crack-grouping approach improves the accuracy of crack detection.

3.2 Brief Overview of Morphological Technique

The morphological image processing technique uses mathematical morphology as a tool for extracting image components that are useful in the representation and description of various region shapes, such as boundaries, skeletons, and convex hull (Gonzalez, Woods, & Eddins, 2009). The mathematical foundation of the morphological technique is based on the theory of set algebra by Minkowski (1903) and on theory of topology by Matheron (1975) (Jahanshahi et al., 2009; Pratt, 2001). A general description about the morphological image-processing technique can be found in Dougherty and Lotufo (2003).

Let $\{p\}$ be the set of pixels in a two-dimensional digital image. Thus, $\{p\}$ partitions the XY -plane into a grid, with the coordinates of the center of each grid being a pair of elements (x, y) from the Cartesian coordinate. A function $f(x, y)$ is said to be a digital image if (x, y) are integers from $\{p\}$ and f is mapping that assigns an intensity value to each distinct pair of (x, y) . The morphological technique applies a shape of binary image referred to as a *structuring element* on an input image. The structuring element is a binary image component whose shape can be designed for different purposes of image manipulation. The center of the structuring element is called the pixel of interest (POI). The pixels within the structuring element boundary have one, and the rest have zero. Figure 3-2 illustrates an example of diamond-shaped structuring element.

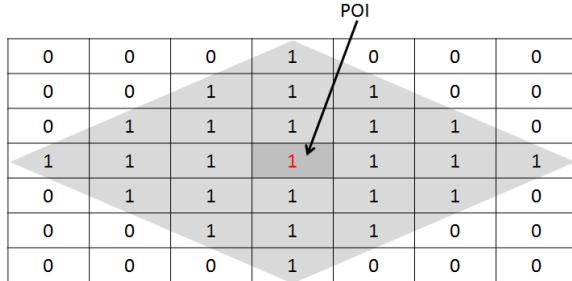


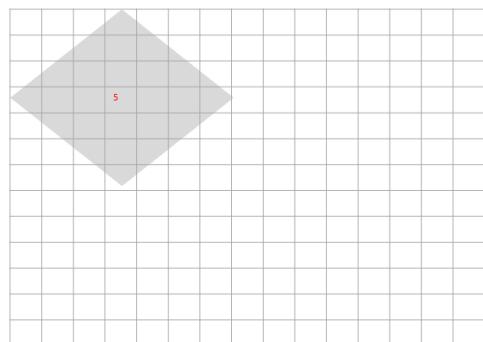
Figure 3-2. Diamond-shaped structuring element with size 7×7 pixels.

In a morphological operation, the structuring element moves on the original image from the left-top corner to the right-bottom corner to create an output image of the same size. The pixel intensity of the output image is determined by the comparison of the input image and the neighboring pixels to the POI, whose boundary is determined by the structuring element shape. The comparison rule is defined by different operators of image-manipulation function $f(x, y)$. For example, the dilation and erosion transforms are two basic morphological operators. In the dilation

transform, the output POI value is determined to be the maximum intensity value of all neighboring pixels. The maximum value in binary image is 1. Thus, the output image “grows” or “thickens” after dilation, if any neighboring pixel has 1. In contrast, the output POI value of the erosion transform is determined to be the minimum value of all neighboring pixels. Thus, the output image “shrinks” or “thins” after erosion if any neighboring pixel has 0. Figure 3-3 illustrates the erosion transform of a grayscale image with the diamond-shaped structuring element. In the illustration, the minimum value of the POI neighborhoods is 5. Therefore, the POI in the output image is 5.

60	68	126	159	15	21	53	122	69	25	36	136	19	253	148
62	64	5	146	19	15	21	122	156	64	5	19	122	123	156
15	21	53	135	122	15	19	57	61	62	64	145	122	56	19
17	19	15	167	122	25	122	57	61	1	0	136	57	59	122
253	122	15	148	62	147	14	60	68	126	128	60	57	126	122
5	77	56	125	11	156	62	62	64	5	19	59	65	147	57
16	45	46	136	126	159	15	15	21	53	122	62	64	156	57
13	85	123	149	19	15	167	165	77	57	61	1	0	159	62
16	14	14	17	19	15	21	53	56	60	68	126	128	146	15
159	156	46	12	165	60	68	126	122	62	64	5	19	135	165
147	42	63	15	125	62	64	5	19	15	21	53	122	167	53
156	45	57	16	165	15	21	53	122	26	56	145	122	148	19
168	48	78	18	148	46	78	45	122	39	25	25	156	253	122

(a) Input image and structuring element



(b) Output image

Figure 3-3. Morphological erosion transform of a grayscale image.

An advantage of the morphological method is that multiple basic operators can form another operation for more sophisticated image processing. For example, although the erosion transform is effective to eliminate small objects of undesirable foreground pixels, “salt noise” that are random in their sizes and positions, it has a disadvantage that all regions of the foreground pixels will be affected indiscriminately. In this case, the opening transform can be used. In contrast, the dilation transform is used to remove random “pepper noise”. It is a common convention in image processing that the dilation and erosion of the input image (I) with the structuring element

(S) are expressed as $I \oplus S$ and $I \ominus S$, respectively. Using the basic operations of the dilation and erosion transforms, the opening and closing transforms can be defined as $I \circ S = (I \ominus S) \oplus S$ and $I \bullet S = (I \oplus S) \ominus S$, respectively.

3.3 Removal of Random Backgrounds Using Bottom-Hat Transform

The morphological method has been used in many crack-extraction applications, and some of literatures are introduced in this section. In Yan, Bo, Xu, and He (2007), the morphological operation was applied to obtain crack edges through the dilation transform subtracted with the erosion transform. Then the closing transform was used to fill the gaps of cracks. Sun, Salari, and Chou (2009) filtered out image noise by a mean intensity and variance based nonlinear filter and a threshold. The dilation and erosion transforms were operated to fill gaps between cracks. Then two cracks were connected together if their endpoints were in the same neighbor area (i.e. 4 rows and 20 columns for horizontal crack). Jing and Aiqin (2010) also converted lower intensity pixels into potential crack pixels by threshold method. Pixels surrounded by seven or eight potential crack pixels were selected as crack pixels. Then morphological opening operation was operated to remove isolated noise. Mancini, Frontoni, and Zingaretti (2013) removed the noise of pavement image by morphological top-hat algorithm first. After enhancing the contrast, median filter and Wiener filter were applied for further removal of image noise. Otsu threshold filter was used to convert the image into binary image. Then a Snake-GVF algorithm is developed to find the final crack boundary.

Salembier (1990) and Jahanshahi et al. (2009) proposed the modified bottom-hat and top-hat transforms to detect black and white cracks as presented in Equations 3-1 and 3-2, respectively.

$$\mathbf{L} = \max \left[\left(\mathbf{I} \circ \mathbf{S}_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} \right) \bullet \mathbf{S}_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}, \mathbf{I} \right] - \mathbf{I} \quad (3-1)$$

$$\mathbf{L} = \mathbf{I} - \min \left[\left(\mathbf{I} \bullet \mathbf{S}_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} \right) \circ \mathbf{S}_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}, \mathbf{I} \right] \quad (3-2)$$

where \mathbf{L} is a gray-scale image as the output of the morphological transforms after the bottom-hat or top-hat transforms; \mathbf{I} is the input gray-scale image; $\mathbf{S}_{\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}$ is the line-shape structuring element rotating 0° , 45° , 90° and 135° ; \circ is the opening transform; and \bullet is the closing transform.

The morphological technique using the bottom-hat transform has several advantages over the other crack-extraction algorithms. First, it is advantageous to extract low-intensity pixels of arbitrary cracks from random pavement backgrounds with a small number of user-defined parameters. The bottom-hat transform is also advantageous to recognize dark cracks from random pavement background. Random background is usually due to intensity variations, caused by asphalt and concrete pavement types, pavement coating, surface texture and roughness, pavement aging and raveling, spots, stains and scratches, oil and water spilling, and road markings. Wu et al. (2014) showed that the bottom-hat transform could effectively remove various random background noises in both pixel level (e.g., surface roughness due to granular aggregates and bitumen binder) and regional level (e.g., partial wetness due to oil or water spills). In the bottom-hat transform, one has to specify only the length of structuring element, which requires a very small number of user-defined parameters. In Wu et al. (2014), the length of the structuring element was set to be 98.0 mm (3.86 in) that is about 4 times greater than the minimum crack width of 25.4 mm (1.0 in) in Crack-III condition, the worst class of pavement condition in FDOT road rating

(Florida Department of Transportation, 2012). Therefore, the bottom-hat transform is employed as the crack-extraction method in this proposed study.

3.4 Data Preparation

3.4.1 Data Preparation Using Road Surveying Images

FDOT conducts annual Pavement Condition Surveys (PCS) as part of the Pavement Management System (Florida Department of Transportation, 2013). In 2006, FDOT acquired a Multi-Purpose Survey Vehicle (MPSV) consisting of a self-contained van equipped with an Inertial Profiler System, an Inertial Navigation System, a Laser Rut Measurement System (LRMS) and a Laser Road Imaging System (LRIS) to capture pavement images at highway speed. The LRIS is composed of two high-resolution linescan cameras and laser illuminators that are configured to image up to 4-m transverse road section with about 1-mm resolution at speeds of 60 mile per hour. The camera is mounted above 1,960 mm from road surface, which has a 20-mm focal length. The image-sensor pixel size is 0.01 mm. Thus, the image resolution of the LRIS used in this study is 0.98 mm per pixel. The FDOT's surveying vehicle equipped with LRIS is shown in Figure 3-4.



Figure 3-4. FDOT multi-purpose survey vehicle with laser road imaging system.

A total of 26 flexible pavement images were selected in data preparation. Based on cracking severity and background randomness, the pavement images were selected from the road sections in 11 good, 7 intermediate, and 8 poor conditions. The good condition images had minor or no single cracks (longitudinal and transverse) and low background randomness (oil and water stains, eroded lane marks, shadows and etc.). The intermediate condition images had single cracks and slight block cracks on a moderate random background. The poor condition images represented pavements with severe single cracks and/or block and alligator cracks on a highly random background.

Sample pavement images used in this study are shown in Figure 3-5. The figures show that the pavement image in good condition (Figure 3-5a) has the lowest intensity; however the right side of the poor pavement (Figure 3-5c) has lower intensity than the intermediate pavement (Figure 3-5b). Therefore, it is not necessarily true that lower intensity represents better pavement condition. The pavement images also contains non-crack random backgrounds, including surface texture patterns, oil stains, scratches, cat's eyes, road markings, and road shoulders. These non-

crack backgrounds are highly arbitrary in their shapes, intensities and locations, which makes successful crack extraction become very challenging.

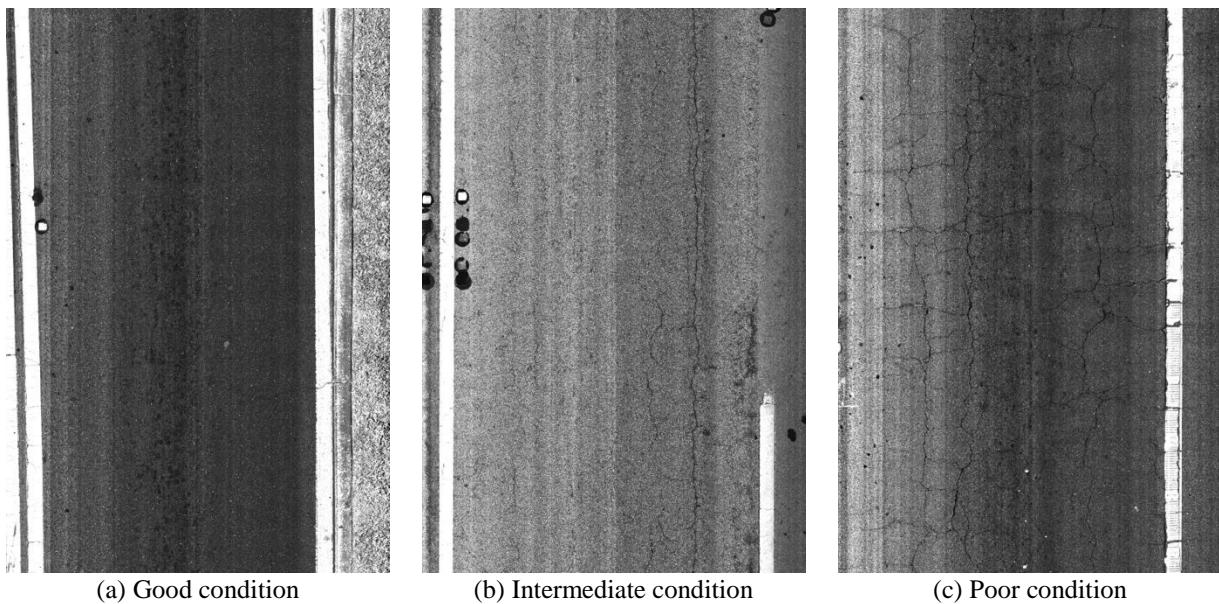
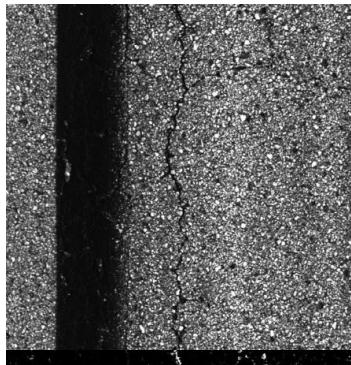


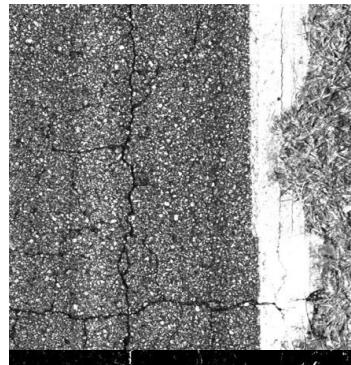
Figure 3-5. Sample flexible-pavement images used in this study.

Once road surface images were collected, the bottom-hat transform was applied for crack extraction, which was shown in Equation 3-1. The Florida Department of Transportation (2012) specifies that the Class-III crack width be 25.4 mm (1.0 in) or larger. In this study, the line-shape structuring element in Equation 3-1 was set to have 100-pixel length, which is equivalent to 98.0 mm (3.86 in). Therefore, the structuring element was designed to be about 4 times larger than the minimum Class-III crack width to ensure to detect cracks in all classes that are specified in the manual. By rotating the structural element for 0° , 45° , 90° and 135° , cracks in arbitrary orientations could be extracted. The output image of the bottom-hat transform is a grayscale image. Thus, to

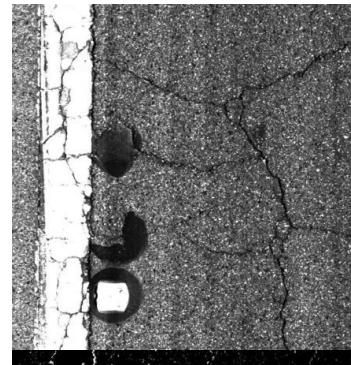
convert the grayscale image into a binary image, an intensity threshold has to be determined: a pixel with an intensity higher than the threshold is to be as a crack pixel, while a pixel with an intensity less than or equal to the threshold is to be as a non-crack pixel. In this study, the threshold was set to be zero intensity to minimize the effects of the intensity threshold; hence the analysis results would be solely affected by the proposed crack-grouping algorithm. More detailed procedures of the bottom-hat transformation in crack extraction for flexible pavement can be found in Wu et al. (2014). Sample results of crack extraction using the bottom-hat transform for different crack types with random background noises are shown in Figure 3-6.



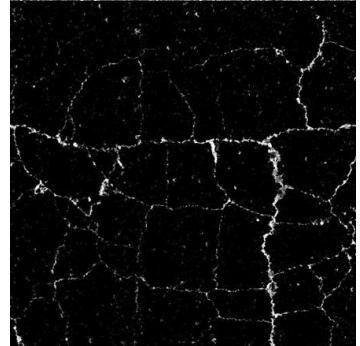
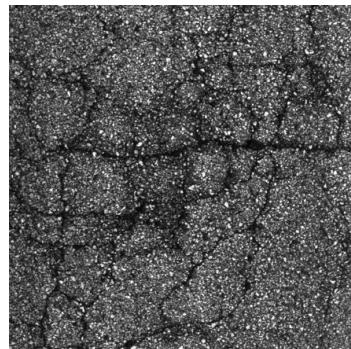
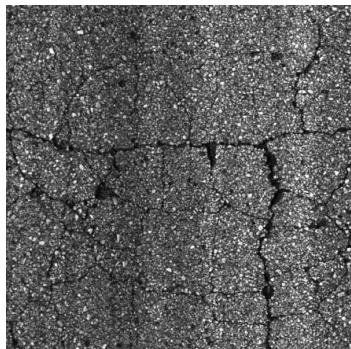
(a) Branched crack & black lane marker



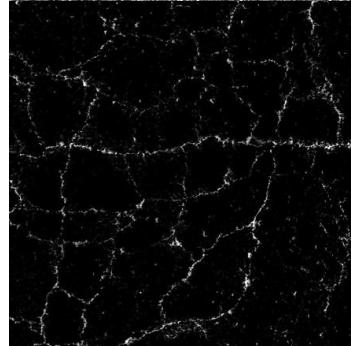
(b) Branched crack, white lane marker & road shoulder



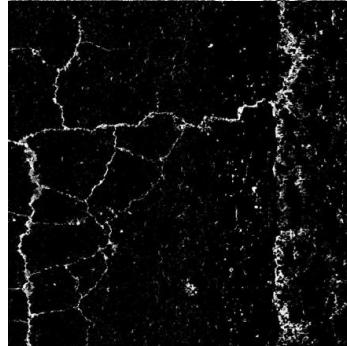
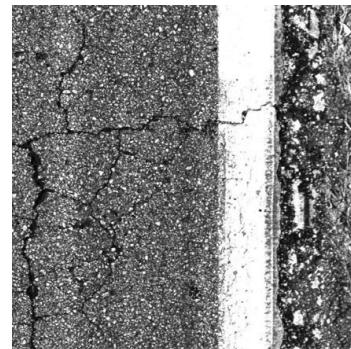
(c) Branched crack, white lane marker & cat's eyes



(d) Block crack



(e) Alligator crack & partial water wetness



(f) Alligator crack, white lane marker & road shoulder

Figure 3-6. Sample crack-extraction results using the bottom-hat transform.

3.5 Grouping of Fragmented Crack Pixels

3.5.1 Procedures of MorphLink-C

Fragmented crack pixels extracted using an arbitrary crack-extraction algorithm—the bottom-hat transform was used in this study—have a great variety in their sizes, shapes, orientations and spacing. Once crack fragments are extracted, the goal of the proposed MorphLink-C is two-folded: i) to group fragmented crack pixels by determining their connectivity to neighboring ones, and ii) to connect fragmented crack pixels of the same group as a continuous line with one-pixel crack width. The procedures of the proposed MorphLink-C algorithm to connect fragmented crack pixels are shown as follows:

1. Extract crack pixels by removing random backgrounds using a crack-pixel extraction method. The binary image after crack pixel extraction is **L**.
2. Apply the morphological dilation to **L** as:

$$\mathbf{D} = \mathbf{L} \oplus \mathbf{S}_D \quad (3-3)$$

where **S_D** is the structuring element of the dilation transform; \oplus is the morphological dilation operator; and **D** is the resulting binary image after the dilation transform. The dilation is an image operation that “grows” or “thickens” image components, fragmented crack pixels in this case. **S_D** is selected to be a square structuring element.

3. Apply the morphological thinning transform to D to connect the fragmented crack pixels as a continuous crack line within a dilation boundary as

$$\mathbf{T} = \mathbf{D} - \text{hit-and-miss}(\mathbf{D}, \mathbf{S}_T) \quad (3-4)$$

where ‘hit-and-miss’ is the morphological hit-and-miss operator; \mathbf{S}_T is the structuring element for skeletonization; and \mathbf{T} is the resulting binary image after the thinning transform.

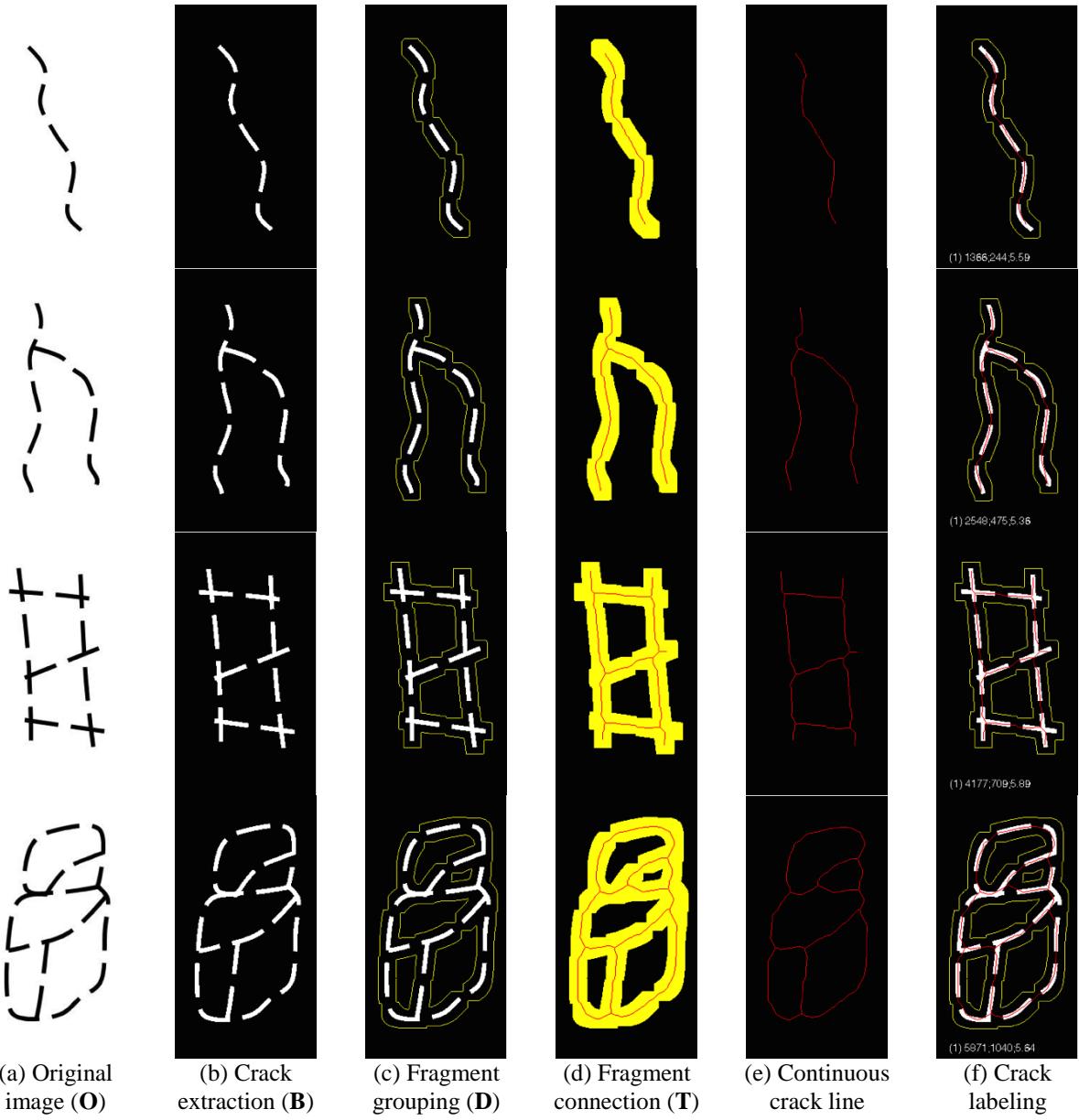
4. Label the fragmented crack pixels and connected crack line within the same dilation boundary with the crack-network identification number of k .
5. Since crack width varies along the crack-line path, one needs to develop a way to measure representative crack width. The averaged crack width of the k -th dilation boundary can be determined as

$$(\text{Averaged crack width})_k = \frac{\sum_i (\text{Area of fragmented crack pixels})_k}{\sum_i (\text{Length of the connected crack line})_k} \quad (3-5)$$

where ‘ $\sum_i (\text{Area of fragmented crack pixels})_k$ ’ is the summation of the total areas of the fragmented crack pixels within the k -th dilation boundary; i is the index of the fragmented objects in the k -th dilation boundary; and ‘ $\sum_i (\text{Length of the connected crack line})_k$ ’ is the summation of the total lengths of connected crack line within the k -th dilation boundary.

The above procedures are applied to different crack types in Figure 3-7 for illustration purpose. For Step 1, the binary image (Figure 3-7b) converted from the original image (Figure 3-7a), which is disjointed along the crack paths. The fragment width is 6 pixels, and the gaps between the disjoints are 16 pixels. The orientations of the fragmented pixels are arbitrary depending on

crack types. For Step 2, the dilation transform is used for fragment grouping by determining the connectivity between crack fragments based on their relative proximity. The proximity limit can be determined only by the size of \mathbf{S}_D for arbitrary crack shapes and orientations. Since the disjoint gap is 16 pixels in this example, the size of \mathbf{S}_D should be set to be greater than the gap size for the fragments to be grouped into the same crack network. In this example, the size of \mathbf{S}_D is set to be 19×19 pixels. The resulting crack-grouping boundary after the dilation transform is shown in Figure 3-7c. The crack boundary includes all the crack fragments in this case since the \mathbf{S}_D size is greater than the (largest) gap size. For Step 3, the thinning transform is applied to connect the crack fragments within the same crack boundary (Figure 3-7d). The crack fragments within the dilation boundary are now connected with a continuous line with one-pixel crack width. It should be noted that the proposed algorithm does not computationally intensive computation for linking and pruning that is required in other methods, such as seed-growing algorithms. Instead, it is able to connect fragments, simply by skeletonizing the dilated region for arbitrary crack shapes (Figure 3-7e). For Step 4, once fragment grouping is conducted, the crack fragments within the same boundary are labeled with a same crack-network identification number (Figure 3-7f). The averaged crack width can be calculated using Equation 3-5 for the grouped crack fragments as summarized in Table 3-1: 5.59 pixels for the single crack, 5.36 pixels for the branched crack, 5.89 pixels for the block crack, and 5.64 pixels for the alligator crack. The estimated crack widths are smaller than the initial crack width of 6 pixels since it takes the average of crack fragments (i.e., 6 pixels) and disjoints (i.e., 0 pixels) in their crack paths.



* From the top, single crack, branched crack, block crack and alligator crack.

Figure 3-7. The procedures of MorphLink-C algorithm for different crack types.

Table 3-1. Averaged crack width estimated with Equation 3-5.

Crack type	Single	Branched	Block	Alligator
Area	1366	2548	4177	5871
Length	244	475	709	1040
Avg. width	5.59	5.36	5.89	5.64

Figure 3-8 illustrates the adaptive characteristic of MorphLink-C in grouping crack fragments. The crack fragments augmented to the real flexible pavement image in Figure 3-8a are highly arbitrary in their sizes, shapes, orientations, and spacing. The crack fragments—extracted using the bottom-hat transform in this study—get connected automatically as the S_D size increases from 1×1 pixels (Figure 3-8b) to 15×15 pixels (Figure 3-8f). The dilation boundary of each fragment works like a “stepping stone” to automatically link the dilation boundaries of neighboring fragments when the S_D size increases. The crack lines are also connected to each other when their dilation boundaries are merged together. The lengths and orientations of crack line connection are determined through thinning transform.

The proposed MorphLink-C algorithm has the following properties in fragment grouping for crack area and fragment connection for crack line:

- The size of S_D is recommended to be chosen an odd pixel number for determining the center position of the square structuring element.
- When the S_D size is smaller than or equal to the shortest distance between two neighboring fragments, their dilation boundaries are isolated, and their crack lines are not connected.
- When the S_D size is larger than the shortest distance between two neighboring fragments, their dilation boundaries are merged into one. The merged boundary can be both an open

form (e.g., single crack and branched crack) and a closed form (e.g., block crack and alligator crack).

- When the dilation boundaries are merged, the crack lines of the neighboring fragments are also connected. The length and orientation to connect two crack lines are determined automatically through thinning transform without involving user-specified parameters. The connected crack lines can also be both an open form (e.g., single crack and branched crack) and a closed form (e.g., block crack and alligator crack). The crack line becomes smoother when S_D size is larger.
- The shapes and locations of fragments work like pivots or “stepping stones” in grouping fragmented pixels with a dilation boundary. Therefore, the accuracy of crack line identification is affected by the accuracy of associated crack extraction algorithms. The identified crack line becomes more accurate when fragments are more densely located.

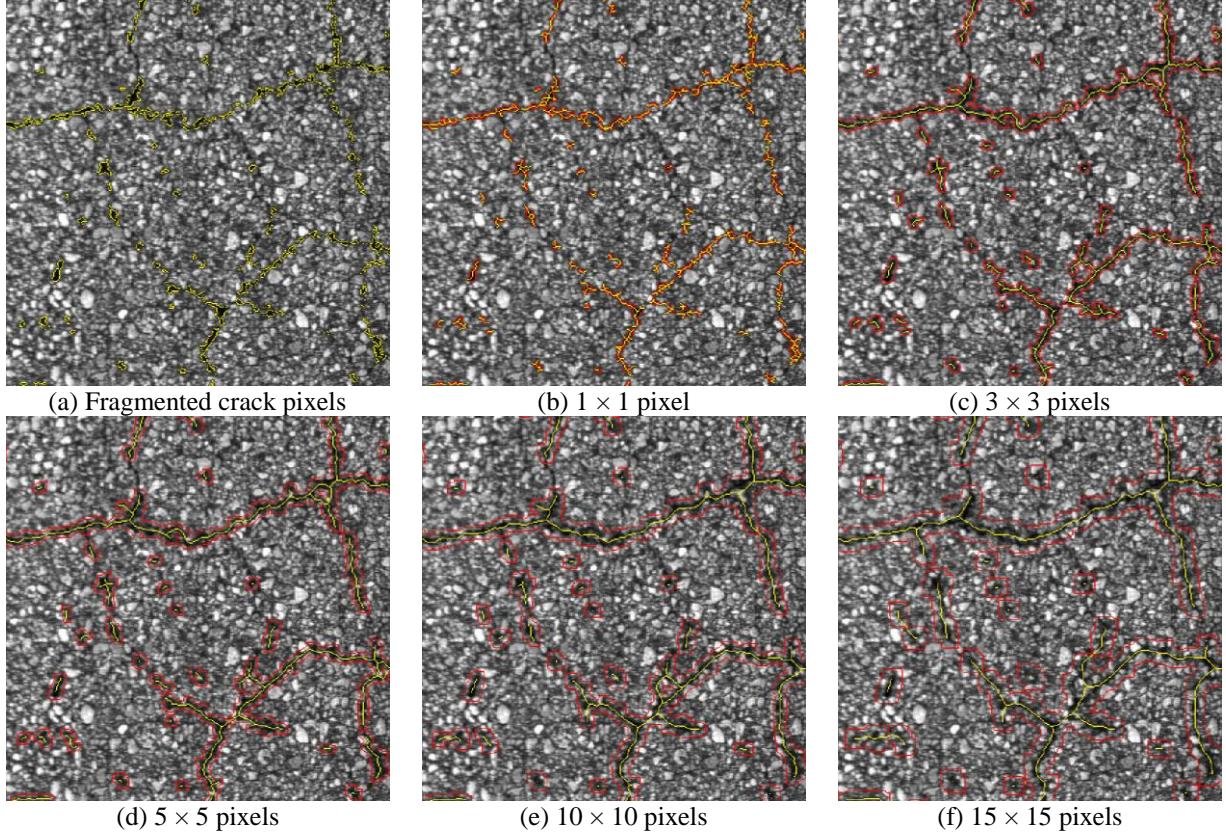
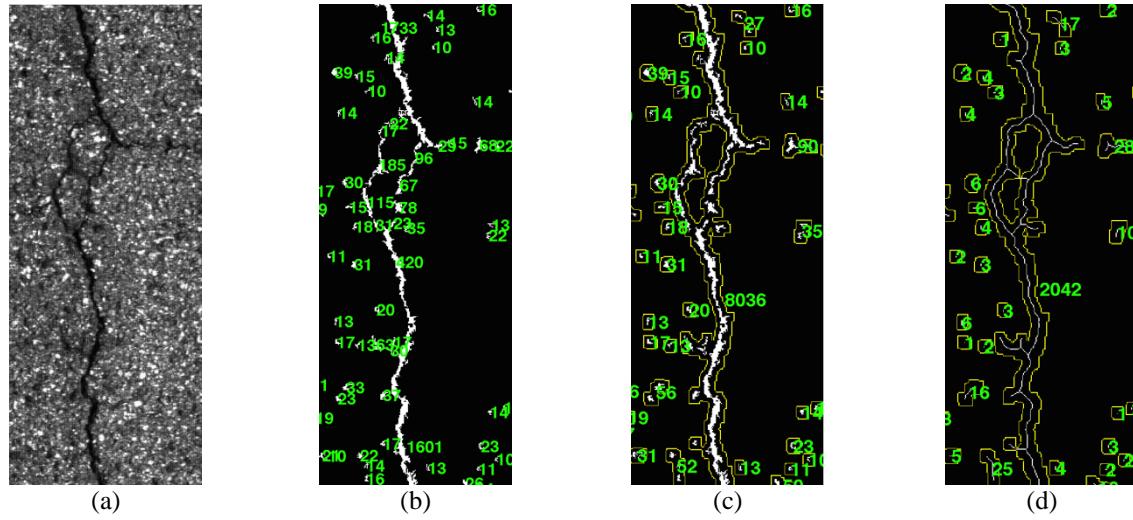


Figure 3-8. Sample results of the proposed MorphLink algorithm.

Figure 3-9 illustrates crack labeling before and after the proposed MorphLink-C algorithm.

The original crack image in Figure 3-9a is processed using the bottom-hat transform to extract crack pixels. The resulting image is shown in Figure 3-9b. Although the crack is accurately extracted by removing pavement background, the extracted crack objects are fragmented. Since the connectivity of these fragments has not been determined yet, they are grouped per 8-connected pixels. The areas of the 8-connected pixels are shown in pixel square in Figure 3-9b. Figure 3-9c shows that the fragments are grouped per dilation boundary with the S_D size of 11×11 pixels. Then, the fragments are labeled per dilation boundary. The figure shows that the summation of the

areas of fragments within each boundary. The crack lines in Figure 3-9d are also labeled per dilation boundary since all the fragments in the same boundary are connected with a continuous line. The figure shows the lengths of the crack lines in pixel.



* (a) Original gray-scale crack image; (b) Fragmented cracks after the bottom-hat transform. The number is the area of the image components of 8-connected pixels; (c) Fragment grouping using the dilation transform. The number is the area of the sum of crack-pixel clusters within each dilation boundary; (d) Fragment connection using the thinning transform. The number is the length of the connected crack line in each dilation boundary.

Figure 3-9. Sample result of the proposed crack grouping method.

3.5.2 Feature Extraction Before and After MorphLink-C

In order to evaluate the effect of the proposed crack grouping approach on crack detection results (Level 3), crack and non-crack datasets should be prepared before and after MorphLink-C. For necessary data preparation, the crack and non-crack image components were manually selected before and after MorphLink-C using the 26 pavement images described before. Once crack and

non-crack image components were selected, six features were extracted from crack image components, including area, length, orientation, texture, intensity, and location. These features can be divided into four categories, geometry, texture, brightness and position as shown in Figure 3-10, which have been widely employed in literatures. The image “before MorphLink-C”, used in the feature extraction was the resulting image after the binarization at zero-intensity threshold (**B**) in Figure 3-9b. The image “after MorphLink-C” used in the feature extraction was the image after the fragment grouping using dilation (**D**) or the fragment connection using thinning (**T**) in Figure 3-9c.

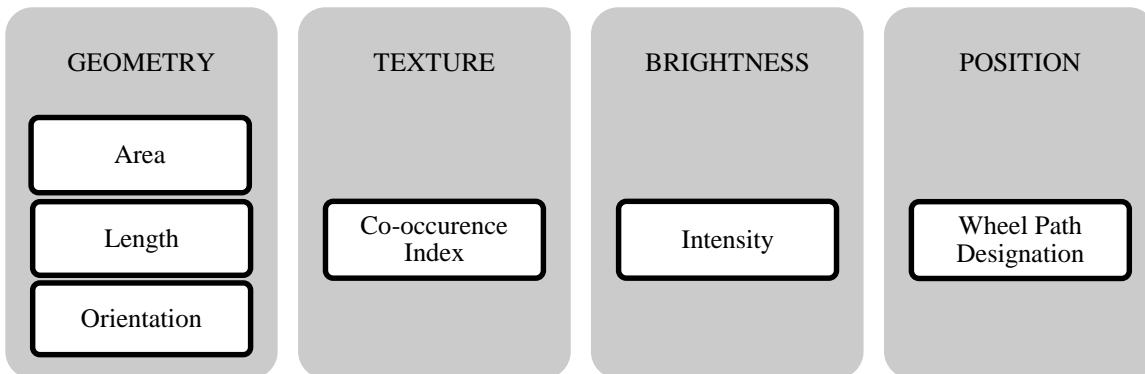


Figure 3-10. The features extracted from crack image components.

Area: The area feature *before* MorphLink-C was extracted using the image **B**. The area of the image object of 8-connected pixels was measured, and sample results are shown in Figure 3-11a. The numbers in the figure are shown in mm^2 . The area feature *after* MorphLink-C was extracted using the image **D**. The area was calculated by summing the areas of the fragmented image components within the same dilation boundary in Figure 3-11d. For example, the two

fragmented image components have the areas of 9.6 mm^2 and 42.3 mm^2 in the bottom-right corner of Figure 3-11a, while the corresponding area after MorphLink-C is 51.9 mm^2 .

Length: To extract the length feature *before* MorphLink-C, first the morphological thinning transform was applied to the image **B** as

$$\mathbf{T}' = \mathbf{B} - \text{hit-and-miss}(\mathbf{B}, \mathbf{S}_{\mathbf{T}'}) \quad (3-6)$$

Then the lengths of the thinned lines of the fragmented image components were measured by multiplying the number of pixels in the thinned line by the pixel resolution in mm per pixel as shown in Figure 3-11b since the thinned line always has one-pixel line width. The numbers in the figure are shown in mm. The length feature *after* MorphLink-C was extracted from the image **T**. Figure 3-11e shows that the fragmented image components within the same dilation boundary are converted into a single connected line. The length of the connected line was measured by multiplying the number of pixels in the line by the pixel resolution in mm per pixel.

Orientation: The orientation feature *before* MorphLink-C was extracted from the image **B**, while the orientation feature *after* MorphLink-C was extracted from the image **D**. The angle of the orientation was measured between the x-axis and the major axis of the ellipse that contains the fragmented or grouped image components before and after MorphLink-C, respectively. The orientation angle ranges between -90 degree and 90 degree, and the counterclockwise is the positive direction. The angles shown in Figures 3-11c and 3-11g are in degree.

Texture: The texture feature was considered since rough pavement surface has higher potential for developing cracks. The texture feature was calculated using the co-occurrence matrix

which can quantify the surface roughness in different regions of the pavement surface. The co-occurrence matrix, first defined by Haralick (1975), is based on the vicinity (offset) of pixels with certain intensities. Mathematical expression of the co-occurrence matrix is as follows:

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1 & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0 & \text{otherwise} \end{cases} \quad (3-7)$$

where C is the co-occurrence matrix; i and j are the grey levels; Δx and Δy describe the offset between the pixels; and $I(p, q)$ is the intensity of the image at (p, q) in a $(m \times n)$ pixel segment of the image. In simple words, the co-occurrence matrix presents the occurrence frequency of two intensity values at a given offset. Defined offsets in this study are the 8-neighboring pixels to identify weather crack pixels are next to each other or dispersed over the entire region. Therefore, the co-occurrence matrix can be simplified to (Ahuja & Rosenfeld, 1978; Kaseko & Ritchie, 1993; Sahoo et al., 1988):

$$C = C_{0,1} + C_{1,1} + C_{1,0} + C_{1,-1} + C_{0,-1} + C_{-1,-1} + C_{-1,0} + C_{-1,1} \quad (3-8)$$

For an 8-bit grey-scale image with the intensity between 0 and 255, the co-occurrence matrix has 256×256 entries where the element (i, j) represents how many times grey levels i and j are adjacent in the image. Although the co-occurrence matrix was originally developed for a grey-scale image, the concept can be extended to any type of image, such as a binary image. In

this study, the co-occurrence matrix was calculated using a binary image. Therefore, the co-occurrence matrix is a symmetrical 2×2 matrix in which element $C(0,0)$ representing the co-occurrence frequency of two non-crack pixels next to each other and $C(1,0)$ or $C(0,1)$ representing the co-occurrence frequency of neighboring crack and non-crack pixels. The value of $C(1,1)$, which represents the co-occurrence frequency of neighboring crack pixels, is considered as *the co-occurrence index* of the region of interest. A region with a higher number of the co-occurrence index can be considered to be rougher than a region with a lower index. Detailed background and mathematical formulation of the co-occurrence matrix is presented in Kaseko and Ritchie (1993). The co-occurrence matrix method is a region-based method as opposed to a pixel-based. Thus, to obtain the co-occurrence matrix in this study, first each image was divided into regions with 200×200 pixels, and the co-occurrence matrix was calculated for each region. Then the co-occurrence index of each region was normalized by dividing to the number of pixels in that region.

Intensity: The intensity feature *before* MorphLink-C was calculated by averaging the intensities of the original grayscale image (**O**) for the 8-connected pixels as shown in Figure 3-11d. The intensity feature *after* MorphLink-C was calculated by averaging the intensity of the original grayscale image (**O**) as

$$(\text{Averaged crack intensity})_k = \frac{\sum_i (\text{Area of framed crack pixels} \times \text{grayscale intensities in image } \mathbf{O})_k}{\sum_i (\text{Area of framed crack pixels})_k} \quad (3-9)$$

A sample result of the intensity feature after MorphLink-C is shown in Figure 3-11h. The intensity values shown in the figure are in 8-bit grayscale.

Position: The position feature was employed to incorporate the wheel-path designation for road rating. According to FDOT (2012a), flexible pavement regions confined to wheel path (CW) are subjected to more frequent traffic loading than the regions outsides of wheel path (CO) as shown in Figure 3-12. Consequently, CW has higher potential for developing fatigue cracks. To obtain the position features before and after MorphLink-C, the lane markers were detected from the road image, and the CW and CO regions were divided based on the ratios of the distance between the left and right lane markers shown in Figure 3-12 (i.e., CO:CW:CO:CW:CO = 1:2:2:2:1). Then, image components in CW were assigned to two, image components in CO were assigned to one, and image components on and outside the lane markers were assigned to zero. When only one lane marker can be observed in the road image as shown in Figure 3-12, CW and CO were determined based on the distances from inside the lane marker (CO:CW:CO:CW:CO = 0.46 m: 0.91 m: 0.91 m: 0.91 m: 0.46 m) since the distance can be measured based on the pixel resolution (e.g., 0.98 mm per pixel in this study). When both lane markers are not available in the pavement image, CW and CO were determined based on the distances from the centerline of the image by assuming the center of the lane was located at the image centerline.

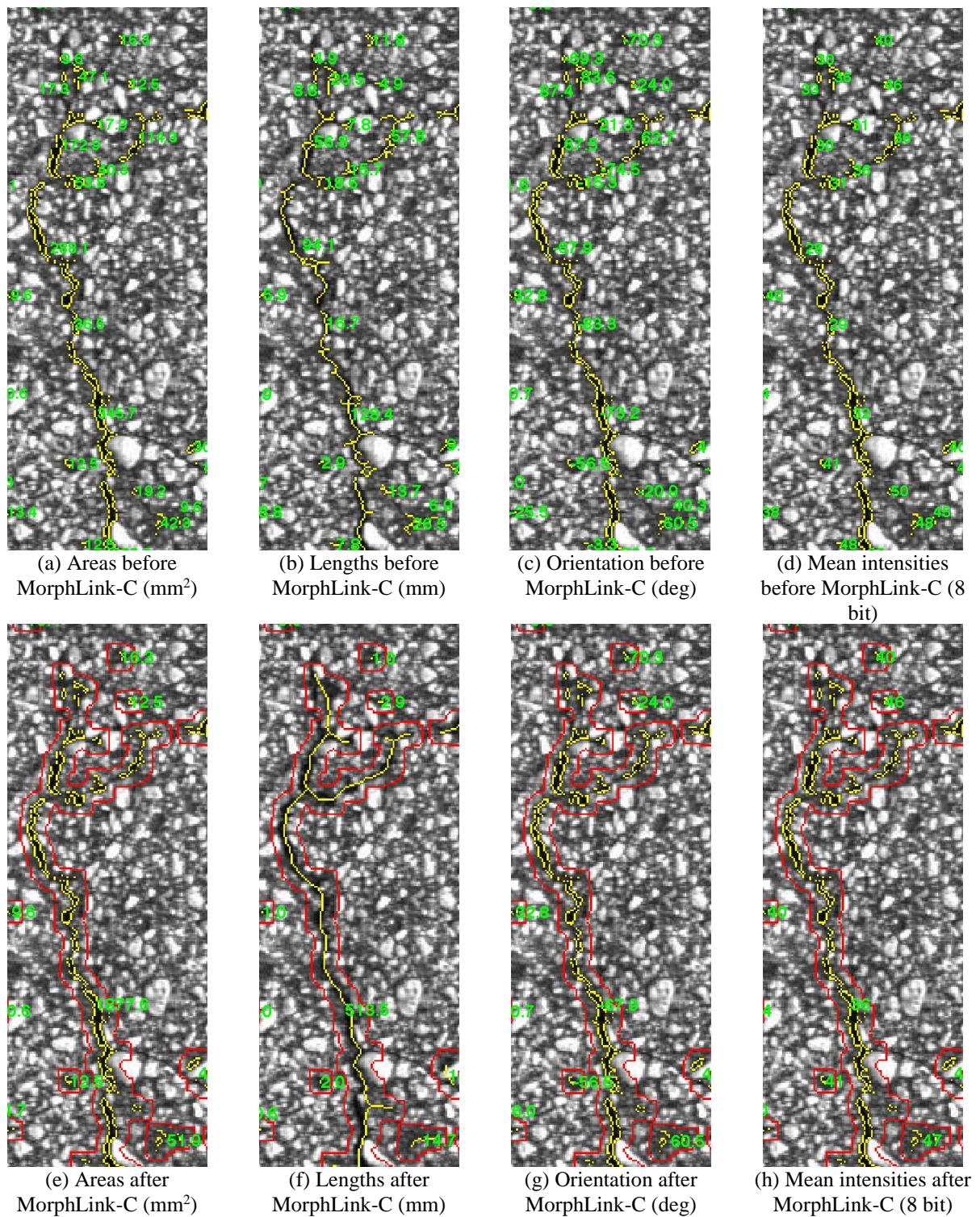


Figure 3-11. Sample features of crack image components.

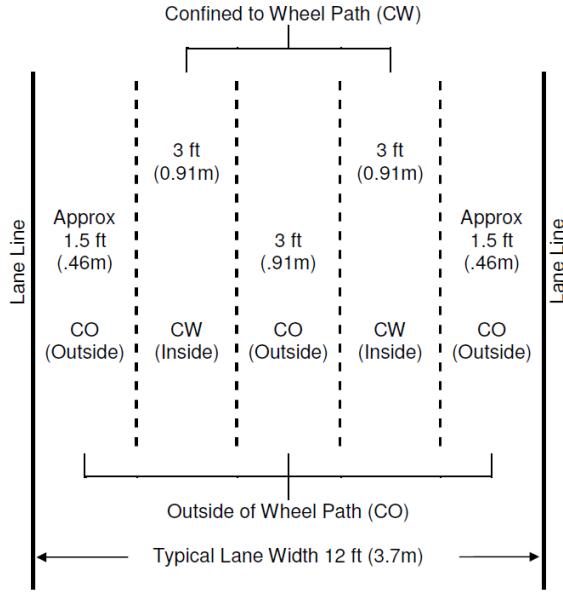


Figure 3-12. Wheel-path designation (FDOT, 2012a).

3.6 Discussion

3.6.1 Statistics of Crack Lengths and Widths for Different Pavement Conditions

The above six features were automatically extracted and measured for the crack and non-crack image objects that were manually selected before and after MorphLink-C using the 26 pavement images. The statistics in terms of the numbers and lengths of crack and non-crack objects for good, intermediate and poor-condition pavement images are summarized in Table 3-2.

For the good condition, a total of 206940 crack and non-crack objects were extracted before MorphLink-C, which had the length summation of 2585.4 m. Among them, cracks were 3138, which had the length summation of 55.0 m. Thus, before MorphLink-C, the cracks out of the total image components were 1.52% in number and 2.13 % in length. After MorphLink-C, a

total of 106386 crack and non-crack objects were extracted, which had the length summation of 1512.9 m. Among them, cracks were 442, which had the length summation of 52.4 m. Thus, after MorphLink-C, cracks out of the total image components were 0.42% in number and 3.46 % in length.

For the intermediate condition, a total of 144584 crack and non-crack objects were extracted before MorphLink-C, which had the length summation of 1787.0 m. Among them, cracks were 7956, which had the length summation of 152.3 m. Thus, before MorphLink-C, cracks out of the total image components were 5.50% in number and 8.52 % in length. After MorphLink-C, a total of 77085 crack and non-crack objects were extracted, which had the length summation of 1101.1 m. Among them, cracks were 866, which had the length summation of 132.7 m. Thus, after MorphLink-C, cracks out of the total image components were 1.12% in number and 12.0 % in length.

For the poor condition, a total of 161207 crack and non-crack objects were extracted before MorphLink-C, which had the length summation of 2127.5 m. Among them, cracks were 15562, which had the length summation of 279.9 m. Thus, before MorphLink-C, cracks out of the total image components were 9.65% in number and 13.15 % in length. After MorphLink-C, a total of 81510 crack and non-crack objects were extracted, which had the length summation of 1301.1 m. Among them, cracks were 1782, which had the length summation of 259.3 m. Thus, after MorphLink-C, cracks out of the total image components were 2.19% in number and 19.93% in length.

Table 3-2. Dataset preparation of crack image components.

Image #		Before the proposed crack grouping method (Without MorphLink-C)				After the proposed crack grouping method (With MorphLink-C)			
		Cracks (#)	Cracks (mm)	All objects (#)	All objects (mm)	Cracks (#)	Cracks (mm)	All objects (#)	All objects (mm)
Good condition (11)	1	22	325	15780	255569	3	273	7519	149979
	2	64	849	15486	268241	13	905	7345	149701
	3	481	8980	16802	212659	60	7718	8498	128759
	4	536	12715	18504	239567	70	10876	9138	138994
	5	190	2766	19978	219915	15	2984	10991	129565
	6	317	4069	19742	209313	50	4625	10865	129312
	7	476	5700	18673	202897	94	6280	10685	118494
	8	402	6876	20072	225010	42	7244	10800	132737
	9	465	10403	20571	253828	57	8888	10429	141403
	10	185	2360	21063	258457	38	2586	10108	150859
	11	0	0	20269	239926	0	0	10008	143116
Intermediate condition (7)	12	1062	23217	19967	274102	141	17584	9934	162640
	13	1410	27231	19310	276034	151	20726	10396	151451
	14	1455	28456	19716	259098	116	23428	10477	151652
	15	81	1140	21339	233896	2	1214	12184	153605
	16	11	173	20689	229400	2	203	12035	145013
	17	2448	41169	21711	257300	310	40253	10937	170581
	18	1489	30906	21852	257122	144	29242	11122	166114
Poor condition (8)	19	1434	23436	21906	260571	212	22422	10705	166817
	20	1779	31985	19770	277335	183	26962	10230	157482
	21	2257	39468	19637	303304	213	33843	9843	173981
	22	2020	34204	19818	291783	252	27005	10156	167410
	23	2014	39459	21672	259987	219	38955	10984	168922
	24	1876	32760	21508	260848	239	34580	11072	166833
	25	3006	58015	20959	276174	320	55307	10219	179091
	26	1176	20542	15937	197481	144	20203	8301	120591
All	Sum	26656	487204	512731	6499817	3090	444306	264981	3915102
	Mean	1025.2	18738.6	19720.4	249993.0	118.8	17088.7	10191.6	150580.8
	Stdv	890.2	16481.4	1870.9	27345.4	99.6	15288.2	1189.5	17061.2

The above statistics in number and length are summarized in Figure 3-13. The figure shows that the percentages of crack objects increase for both before and after MorphLink-C when pavement condition is worse. This result shows that the percentage of crack objects extracted using the bottom-hat transform increases for deteriorated pavement surfaces. For the number of cracks in Figure 3-13a, the crack percentages before MorphLink-C are greater than those after

MorphLink-C for all pavement conditions. This is due to the crack grouping effect after MorphLink-C since multiple fragmented pixels in the vicinity of crack boundaries are labeled as the same crack group. In contrast, for the length of cracks in Figure 3-13b the crack percentages after MorphLink-C are greater than those before MorphLink-C for all pavement conditions. This result indicates that the fragmentation of crack images after crack extraction using bottom-hat transform can be greatly improved by crack grouping using the proposed MorphLink-C algorithm.

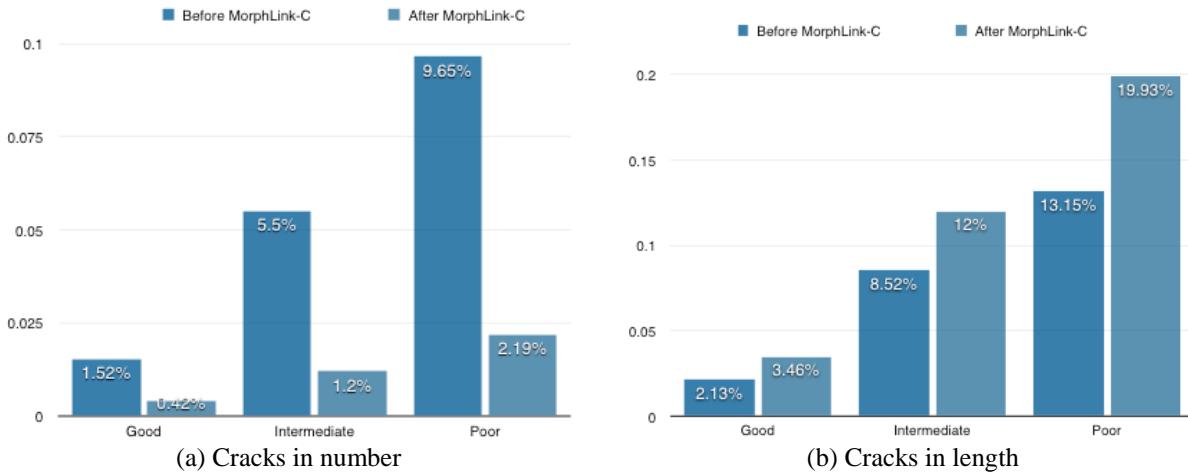


Figure 3-13. The percentages of crack components for different pavement conditions.

Crack length is an important feature to characterize the severity of pavement deterioration. Table 3-3 shows the statistics of the lengths of crack and non-crack objects before and after MorphLink-C. The mean of crack lengths significantly increases after MorphLink-C, while the mean of non-crack lengths remains relatively constant. This indicates that the proposed grouping algorithm is effective selectively on crack objects by removing disjoints between crack fragments. The standard deviation of crack lengths increases after MorphLink-C as the pavement condition is

poorer: 281.6 mm for good, 386.9 mm for intermediate, and 430.4 mm for poor conditions. Thus, it can be postulated that the range of crack length increases as pavement has severe cracks, such as box or alligator cracks.

Table 3-3. Statistics of the lengths of crack and non-crack image components.

Type	MorphLink-C	Good condition				Intermediate condition				Poor condition			
		Mean	Stdv	Max	Min	Mean	Stdv	Max	Min	Mean	Stdv	Max	Min
Crack	Before	17.19	28.64	504.7	0.98	18.76	36.99	1099	0.98	17.62	36.84	1142	0.98
	After	116.1	281.6	3744	0.98	150.1	386.9	4571	0.98	142.6	430.4	6378	0.98
Non-crack	Before	12.18	35.46	504.7	0.98	11.73	120.5	1099	0.98	12.43	100.7	1142	0.98
	After	13.51	103.3	21004	0.98	12.45	78.2	14158	0.98	12.81	114.7	19919	0.98

* The numbers are in millimeter.

Crack width is another important feature that should be measured since the deterioration level of flexible pavement is often controlled by the maximum crack width. For example, like many road maintenance authorities, FDOT classifies pavement conditions with crack width in road rating: hairline cracks less than or equal to 3.18 mm (1/8 in) for Class 1B, cracks greater than 3.18 mm (1/8 in) for Class 2, cracks greater than 6.35 mm (1/4 in) for Class 3 (Florida Department of Transportation, 2012). The proposed MorphLink-C algorithm provides an effective way to measure “averaged” crack width after crack grouping. The averaged crack width measurement after MorphLink-C shown in Equation 3-10 is compared with the averaged crack width measurement before MorphLink-C as

$$(\text{Averaged crack width})_l = \frac{(\text{Area of 8-connected crack pixels})_l}{(\text{Length of 8-connected crack line})_l} \quad (3-10)$$

where l is the index of fragmented 8-connected crack pixels after bottom-hat transform in image \mathbf{L} ; $(\text{Area of 8-connected crack pixels})_l$ is the area of the l -th fragmented 8-connected crack pixels as shown in Figure 3-8a; and $(\text{Length of 8-connected crack line})_l$ is the crack length of the l -th fragmented 8-connected crack pixels, which is obtained using the thinning transform as shown in Figure 3-8b.

Table 3-4 shows the statistics of the widths of crack and non-crack objects before and after MorphLink-C. It shows that as pavement condition is poorer, the means of crack widths increase both before and after MorphLink-C. The standard deviation of crack widths after MorphLink-C monotonically increases as 3.06 mm for good, 3.54 mm for intermediate and 3.55 mm for poor pavement conditions. The standard deviations of crack width before MorphLink-C are significantly larger and more fluctuate than those after MorphLink-C as 9.70 mm for good, 13.05 mm for intermediate, and 18.30 mm for poor conditions. This result indicates that the proposed MorphLink-C improves statistical stability and accuracy in the estimation of averaged crack width.

Table 3-4. Statistics of the widths of crack and non-crack image components.

Type	MorphLink .C	Good condition				Intermediate condition				Poor condition			
		Mean	Stdv	Max	Min	Mean	Stdv	Max	Min	Mean	Stdv	Max	Min
Crack	Before	4.20	9.70	276.8	0.03	4.88	13.05	421.4	0.03	5.40	18.30	748.5	0.01
	After	3.62	3.06	32.34	1.10	3.98	3.54	43.12	1.10	4.06	3.55	34.79	1.07
Non-crack	Before	3.14	15.80	2805	0.00	3.65	109.5	39517	0.00	3.51	44.34	11908	0.00
	After	6.14	5.72	260.7	0.88	6.21	6.05	252.8	0.85	6.36	6.06	154.8	0.82

* The widths are shown in millimeter.

The box plots are shown in Figure 3-14 to illustrate the imbalance of data in crack length and width. On each box, the central mark indicates the median, and the top and bottom edges of

the box are the 25th and 75th percentiles. The whiskers in the dashed lines show the lower limit (LL) and upper limit (UL) as

$$LL = \hat{q}_{0.25} - 1.5 IQR \quad UL = \hat{q}_{0.75} + 1.5 IQR \quad (3-11)$$

where $\hat{q}_{0.25}$ is the 25th percentile; $\hat{q}_{0.75}$ is the 75th percentile; and the interquartile range $IQR = \hat{q}_{0.75} - \hat{q}_{0.25}$. The idea of the whiskers is that observations lying outside these limits are possible outliers that could represent extreme points that arise naturally according to the distribution. The horizontal dotted line marks the limit if any points are outside it, and two gray lines outside the dotted line mark the compression region if any points are compressed. The maximum and minimum values of the crack length and width are shown in Tables 3-3 and 3-4. Figures 3-14a and 3-14c show that before MorphLink-C, most of crack objects have small lengths as shown within the UL of the whiskers less than 3 cm for all pavement conditions. This result is due to the disjoints of crack fragments. Figure 3-14b shows that the median, LL and UL of crack length significantly increase after MorphLink-C. The median, LL and UL of averaged crack width also increase after MorphLink-C as shown in Figure 3-14d. The above results show that the imbalancedness of data in crack length and width can be mitigated by removing disjoints between crack fragments, which can improve crack-detection accuracy to be discussed in the subsequent section.

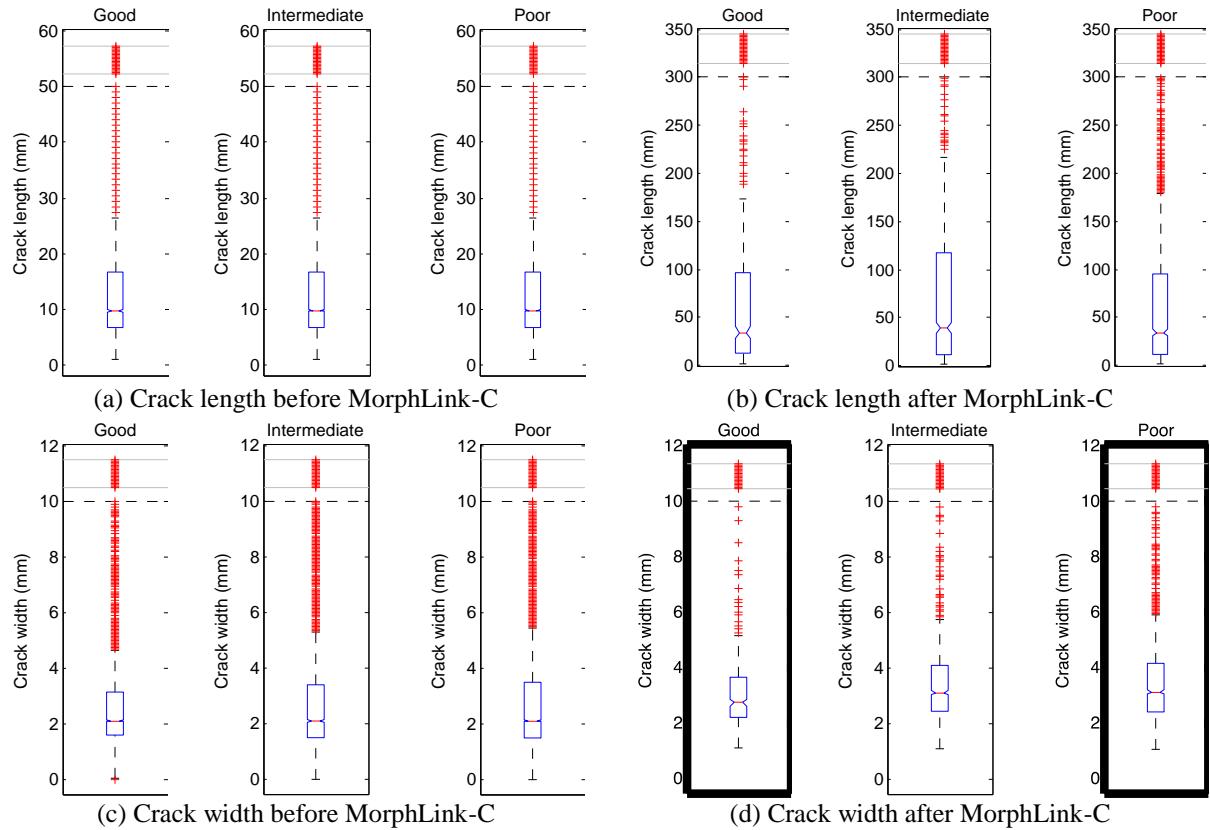


Figure 3-14. Box plots of crack length and width before and after MorphLink-C.

3.6.2 Effects on Crack Detection Accuracy Using ANN Classifier

In order to evaluate the effect of the proposed crack grouping approach on crack detection results (Level 3), the classification results before and after MorphLink-C was compared. The ANN classifier might be one of most commonly used supervised classification methods for crack detection. The ANN is a network of simple processing units called neurons which are connected using a set of weights and biases. The ANN classifier that was used in this study consisted of a seven-layer feed-forward network, including one input layer and one output layer. Each hidden

layer contained 10 neurons. Tangent-sigmoid and linear transfer functions were used in the hidden and outputs layers, respectively. The input of the classifier was the six features that were extracted from pavement images, including the area, length, orientation, intensity, texture and position. Having the linear function in the output layer, the output of the classifier was a real number ranged between 0 and 1. Since the output was a non-crack object toward 0 and a crack object toward 1. Therefore, a threshold should be set between 0 and 1 to classify an image object into non-crack when its output is smaller than the threshold or crack when its output is greater than equal to the threshold. A schematic of the ANN configuration is shown in Figure 3-15.

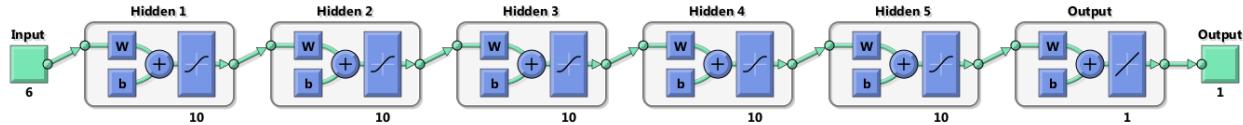


Figure 3-15. A schematic of the ANN configuration.

Being a supervised machine learning method, training the ANN classifier is the process of adjusting these weights and biases, so that the classifier produces desirable network outputs. For the network optimization, the Levenberg-Marquardt learning method was used, associated with a mean-square-error (MSE) function to minimize the error between the network output and target output:

$$\text{MSE } (O) = \frac{1}{n} \sum_{i=1}^n (T - O)^2 \quad (3-12)$$

where O is the $(n \times 1)$ output vector of neural network; T is the $(n \times 1)$ vector of target values that were manually selected; and n is the number of training samples.

The proposed crack grouping approach was evaluated using the above ANN classifier. Necessary datasets extracted before and after MorphLink-C were prepared by manual selection of crack and non-crack objects using the 26 pavement images in Table 3-2. Before MorphLink-C, a total of 512731 objects were extracted, including 26656 (5.20%) crack and 486075 (94.80%) non-crack objects in number. The total length of the objects was 6499.8 m, including 487.2 m (7.50%) for crack and 6012.6 m (92.50%) for non-crack objects. After MorphLink-C, a total of 264981 objects were extracted, including 3090 (1.17%) crack and 261891 (98.83%) non-crack objects. The total length of the object was 3915.1 m, including 444.3 m (11.35%) for crack and 3470.8 m (88.65%) for non-crack objects.

After crack extraction before and after MorphLink-C, all the six features were measured from the image objects. Since the datasets were extracted from 26 images, the entire image objects were randomly shuffled to avoid the biasness per image, and partitioned into 60% for training, 20% for validation and 20% for testing to compare two ANN classifier models (i.e., before and after MorphLink-C). The training of the ANN classifier was performed in the following procedures:

- 1.** The classifier was trained using the training dataset.
- 2.** The MSE of classification result was calculated using the validation dataset.
- 3.** Steps 1 and 2 were repeated five times to avoid training at local minima, and the classifier with the lowest MSE using the validation dataset was selected as the best-trained classifier.
- 4.** Steps 1, 2 and 3 were performed using the datasets before and after MorphLink-C.

Once the ANN classifier was trained, the classification accuracy was evaluated using the testing dataset. First, the accuracy was calculated by the number of truly classified crack and non-crack objects divided by the number of all objects, and the percentage of apparent success in number is shown in Figure 3-16a. In the figure, the x-axis is the threshold of the classifier output. Overall, the classification results in *number* after MorphLink-C are more accurate for all thresholds than the classification results before MorphLink-C. The apparent success after MorphLink-C is higher than 98% for all thresholds, while the apparent success before MorphLink-C is higher than 85% for all thresholds. Classification result in *length* is another important way to measure the performance of the trained classifiers. The apparent success was also calculated by the length of truly classified crack and non-crack objects divided by the total length of all objects. The apparent success in length is shown in Figure 3-16b. Similar to Figure 3-16a, the classification results after MorphLink-C are more accurate than the classification results before MorphLink-C. The apparent success after MorphLink-C is higher than 90% for all thresholds, while the apparent success before MorphLink-C is higher than 84% for all thresholds.

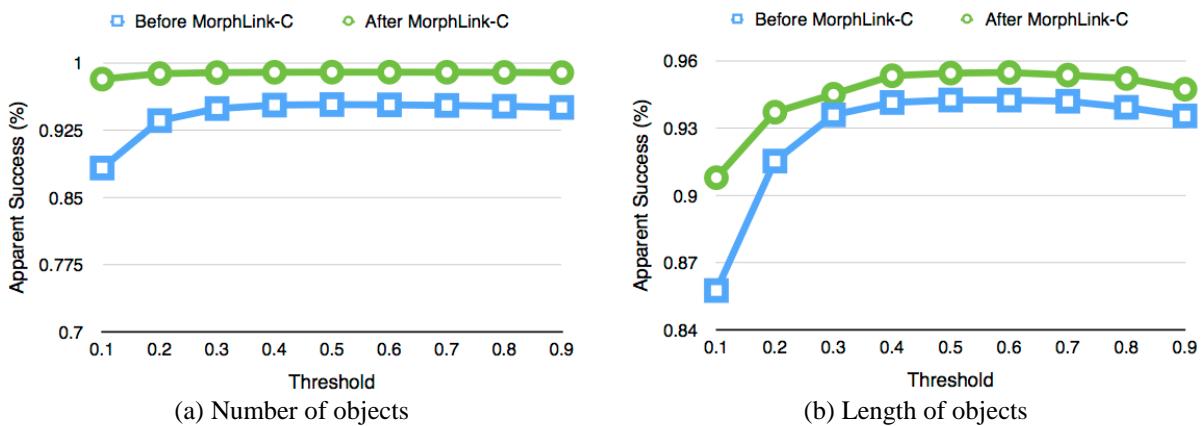


Figure 3-16. Apparently successful classification in number and length.

Although the above apparent success results show the proposed MorphLink-C can improve crack-detection accuracy, they should be carefully interpreted for two reasons. First, the ratio of crack and non-crack objects contained in the datasets is highly unbalanced. As described earlier, the ratios of the crack objects in *number* are less than 5.20% before MorphLink-C and 1.17% after MorphLink-C, and the ratios of the crack objects in *length* are less than 7.50% before MorphLink-C and 11.35% after MorphLink-C. Consequently, the training results for crack and non-crack object could be also unbalanced. Second, the ratio of small and large crack objects is also unbalanced as many large crack objects are outliers as lying outside the *ULs* of the box plots in Figure 3-14. For these reasons, having a large crack object is statistically an extreme event in the datasets, which could affect classification results. Thus, more rigorous investigation was conducted to analyze classification errors through statistical hypothesis tests.

In the hypothesis test, the null hypothesis (H_0) was that “the extracted object is *not* crack”, and the alternative hypothesis (H_a) was “the extracted object is crack”. In the top-left quadrant of Figure 3-17, the probability to accept H_0 when it is true ($p = 1 - a$) is called confidence level. In the bottom-left quadrant, rejecting H_0 when it is true is called Type 1 error (or false positive), and its probability ($p = a$) is called significance level. In the top-right quadrant, accepting H_0 when in fact it is false is called Type 2 error (or false negative), and its probability is $p = b$. In the bottom-right quadrant, rejecting H_0 when it is true is what a good classifier should make. Its probability ($p = 1 - b$) is called power of test, which measures how good a statistical test is.

A good classifier model reduces both Type 1 and 2 errors simultaneously for given datasets. For the ANN classifier, Type 1 and 2 errors are determined with the threshold. In general, when the threshold increases, Type 1 error decreases while Type 2 error increases, which have an inverse

relation. Therefore, the goal of the hypothesis test is to find the threshold for acceptable levels of Type 1 and 2 errors for given application. For Type 1 error, a significant level is set, usually $\alpha = 0.05$. Hence, once Type 1 error (or confidence level $p = 1 - \alpha$) is set, the power of test measures the goodness of hypothesis test. Conventionally a test with a power ($p = 1 - \beta$) of 0.8 or higher is considered a good test.

		<u>DECISIONS</u>	
		Accepting H_0 when it is true; good decision ($p = 1 - \alpha$ or confidence level)	Accepting H_0 when it is false; Type 2 error ($p = \beta$)
H_0 is true	Accepting H_0 when it is true; good decision ($p = 1 - \alpha$ or confidence level)	Accepting H_0 when it is false; Type 2 error ($p = \beta$)	
	Rejecting H_0 when it is true; Type 1 error ($p = \alpha$ or significance level)	Rejecting H_0 when it is not true; good decision ($p = 1 - \beta$ or power of test)	
H_a is true	Rejecting H_0 when it is true; Type 1 error ($p = \alpha$ or significance level)	Accepting H_0 when it is false; Type 2 error ($p = \beta$)	

Figure 3-17. Hypothesis test.

The results of the statistical hypothesis tests are shown in Figure 3-18. Based on the *number* of objects in Figure 3-18a, Type 1 errors ($p = \alpha$) for both before and after MorphLink-C decrease when the threshold increases as expected. It is also observed that the error is reduced after MorphLink-C for all thresholds. Overall, Type 1 error ($p = \alpha$) decreases both before and after MorphLink-C as the threshold increases toward 1 as expected, and they are smaller than 5% except the error before MorphLink-C at the threshold of 0.1. In Figure 3-18b, Type 2 errors ($p = \beta$) for both before and after MorphLink-C increase when the threshold increases as expected. Overall, Type 2 errors are greater than Type 1 errors. This might be due to the unbalance data between the numbers of crack and non-crack objects. Type 2 error after MorphLink-C is smaller than the error

before MorphLink-C when the threshold is greater than 0.5, while Type 2 error after MorphLink-C is greater than the error before MorphLink-C when the threshold is less than and equal to 0.5. Type 2 error after MorphLink-C is 39.3% when the threshold is 0.1, which results in the power of test ($p = 1 - b$) of 60.7%.

The Type 1 and 2 errors based on object *length* in Figures 3-18c and 3-18d can be thought as the errors based on object numbers weighted with their lengths. Hence, if a large image object is falsely accepted or rejected, the error will be considered heavier than for a small object. Therefore, the errors in length would be a more accurate measure than the errors in number to evaluate pavement condition in road rating applications. Figure 3-18c shows that Type 1 errors for both before and after MorphLink-C decrease as the threshold increases, which is a similar trend to Figure 3-18a. The error levels, however, are slightly higher than those in Figure 3-18a. In Figure 3-18d, Type 2 error increases as the threshold increases, which is also similar to Figure 3-18b. Type 2 error is significantly reduced after MorphLink-C for all thresholds. This result is different from Figure 3-18b particularly when the threshold is less than 0.5. Although Type 2 error before MorphLink-C is smaller than the error after MorphLink-C in number in Figure 3-18b, the error relation becomes opposite in length in Figure 3-18d. Thus, it can be postulated that relatively large crack objects are falsely rejected before MorphLink-C when the threshold is less than 0.5. At the threshold of 0.1, the Type 2 error is 11.02%, equivalent to the power of test of 88.98% that can be considered as a good test. However, the corresponding Type 1 error is 8.99% that is greater than the conventionally accepted error limit of 5%. At the threshold of 0.2, the Type 2 error is 16.38% that can be still considered as a good test since the equivalent power of test is 83.62% that is higher than 80%. The corresponding Type 1 error is 4.99%, which is an acceptable error level being

smaller than 5%. Therefore, 0.2 can be selected the optimal threshold value for the trained ANN classifier after MorphLink-C.

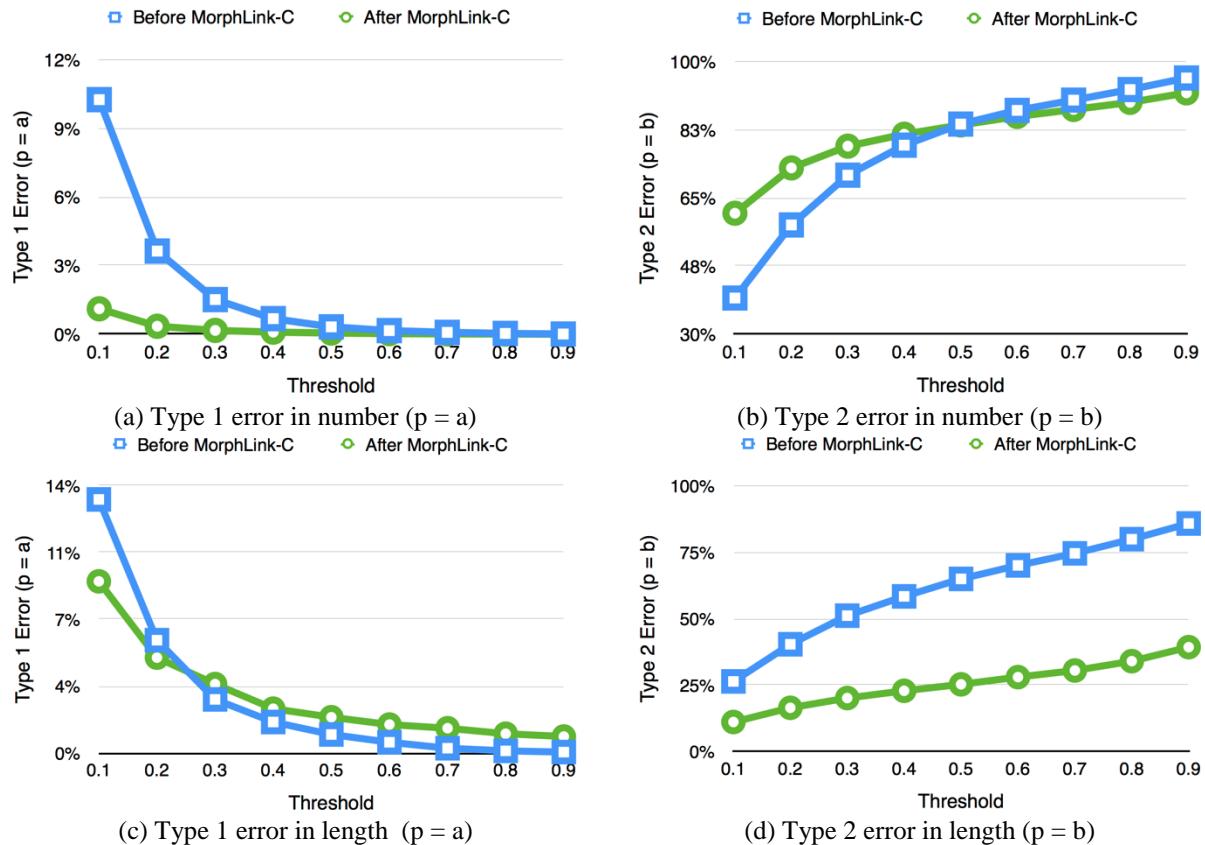


Figure 3-18. Type 1 and 2 errors before and after MorphLink-C.

Figure 3-19 shows sample classification results before and after MorphLink-C for the thresholds of 0.2, 0.5 and 0.8. In the figure, blue shows correctly accepted classification for crack objects, green shows correctly rejected classification for non-crack objects, red show falsely accepted classification for non-crack objects (i.e., Type 1 error), and pink shows falsely rejected classification for crack objects (i.e., Type 2 error). For both before and after MorphLink-C when the threshold is 0.2, the pink decreases while red increases; and when the threshold is 0.8, the pink

increases while red decreases. In the right-side wheel path, the blue shows truly accepted crack objects that are highly fragmented before MorphLink-C while they are connected after MorphLink-C.

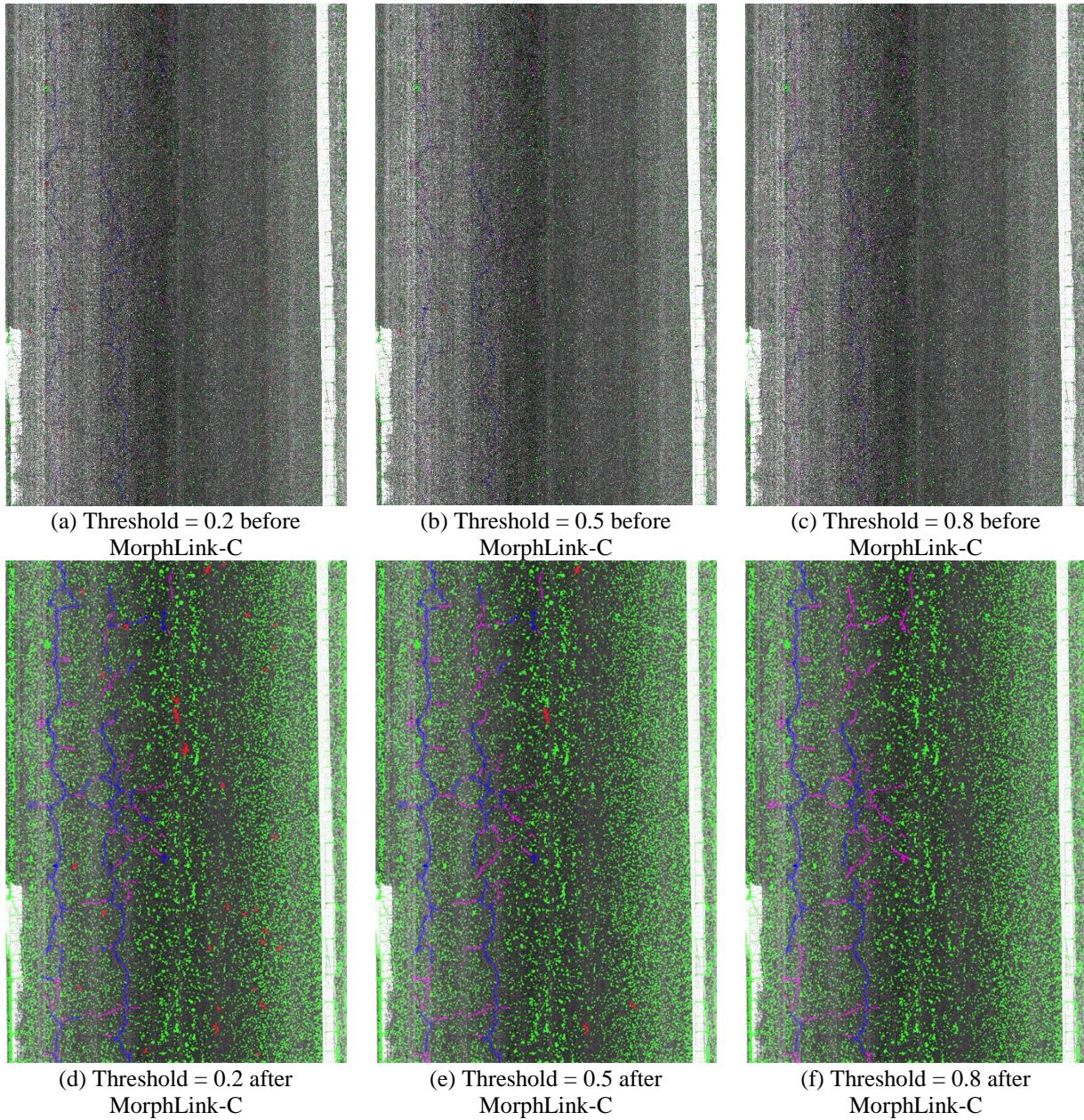


Figure 3-19. Sample classification results for different threshold values.

3.7 Conclusions

The MorphLink-C algorithm was proposed in this study as a novel crack-grouping approach. It is mainly based on morphological image processing techniques, associated with the dilation and thinning transforms to remove disjoints between crack fragments. The proposed grouping approach is important to characterize cracks in flexible pavement and to improve crack detection accuracy for different levels of pavement deterioration. There are several advantages of the proposed MorphLink-C over traditional crack-segmentation methods:

- It is an adaptive method to determine the connectivity of crack fragments, which can be associated with any crack-extraction algorithm.
- In the determination of connectivity, only one user-defined parameter, the size of \mathbf{S}_D , is needed for arbitrary crack types of flexible pavement.
- The \mathbf{S}_D size can be determined based on the largest proximity between two adjacent crack fragments. For example, in this study \mathbf{S}_D was set as a square structuring element with the size of 11×11 pixels, which are equivalent to $10.78 \times 10.78 \text{ mm}^2$. Hence, when the gap between two fragments is less than 10.78 mm , they are considered as the same crack. Specifying the \mathbf{S}_D size allows more precise control on the spatial proximity than specifying the number of dilation transform that was proposed in He et al. (2011) and Huang and Zhang (2012).
- The proposed fragment connection method based on the thinning transform allows representing fragmented crack image components as a continuous crack line with the one-pixel width. No specific user-defined parameter is required for an arbitrary crack type, including single, branched, block and alligator cracks.

- Crack width is an important feature that should be measured since the deterioration level of flexible pavement is often controlled by the maximum crack width. The proposed crack width estimation method in Equation 3-5 provides an effective way to characterize flexible pavement cracks.

In this study, it was hypothesized that the proposed approach could improve crack-detection accuracy by grouping crack fragments that were extracted with a given crack-extraction algorithm—the morphological bottom-hat transform in this study. It was validated through statistical analyses of crack-detection results as follows:

- The statistical imbalancedness between crack and no-crack objects could be mitigated with the proposed crack-grouping approach by connecting disjoints of crack fragments.
- The proposed crack-grouping algorithm improved the accuracy of crack detection for flexible pavement by finding the optimal ANN threshold of 0.2 for less than 5% Type 1 (or false positive) error and 20% Type 2 (or false negative) error, which are commonly accepted error levels in many classification applications.

CHAPTER 4: STATISTICAL CHARACTERIZATION OF CRACK FEATURE FOR AGING ROAD PAVEMENT USING COMPUTER-VISION TECHNIQUE²

4.1 Introduction

Crack is a major cause of pavement deterioration that determines short-term and long-term performance of road pavement. Transportation maintenance authorities have developed standard crack rating procedures to assess pavement conditions based on surveying results of road surface images. A summary of crack rating procedures by federal and state departments of transportation is presented in Wu et al. (2014). Road surveying vehicles are often employed to periodically collect road surface images. Road surface surveying is usually conducted to determine the present condition, to compare present with past condition, to predict future deterioration rates, and to provide justification for the policies of rehabilitation priority and budget allocation. Effective pavement rehabilitation policies can only be established with reliable quantitative assessment of past and present pavement conditions. However, current crack rating techniques are primarily based on qualitative human visual inspection. Drawbacks of the human visual inspection are well established. Human visual inspection results are subjective, and its prediction of future deterioration that is based on qualitative inspection results is consequently inaccurate.

² The contents of this chapter is in preparation to be published as journal paper.

Mokhtari, S., Yun, H.-B., & Wu, L. (2015a). *Statistical characterization of crack features for aging road pavement using computer vision technique*. Unpublished manuscript.
The author contributed to all sections except 4.2.2.

Computer-vision technique has been used as a promising non-destructive method to assess conditions of aging road pavement by analyzing pavement surface images. Crack in computer vision can be defined as a group of low-intensity pixels compared to neighboring pixels, which forms into an arbitrary line shape in length, width and directions (crack-pixel level). One or more crack pixels can further form into a complicated crack network of a single line, branched line or polygon, which are not necessarily continued in its line paths due to the fragmentation of crack pixels (crack-network level). Hence, a challenge in computer vision applications is to recognize crack image components having multi-level topological shapes from random noisy background of pavement surface texture, spots and stains, oil and water spills, road markings, manholes, etc. Many computer-vision techniques have been applied to extract crack objects from road surface images. Some of numerous applications include intensity thresholding methods (Dong, Yu, Ogunbona, & Li, 2008; Hu & Zhao, 2010; Koutsopoulos & Downey, 1993; Nguyen, Avila, & Begot, 2009; Oliveira & Correia, 2009; Petrou, Kittler, & Song, 1996; Song, Petrou, & Kittler, 1995; Sun & Qiu, 2007), edge detection methods (Ayenu-Prah & Attoh-Okine, 2008; J. Li, 2003; Tsai, Kaul, & Mersereau, 2010; Tsao, Kehtarnavaz, Chan, & Lytton, 1994), crack seeds (J. Huang, Liu, & Sun, 2014; Q. Li, Zou, Zhang, & Mao, 2011; Zou, Cao, Li, Mao, & Wang, 2012) and morphological techniques (Chu, 2010; He, Qiu, Wang, Zhang, & Xie, 2011; W. Huang & Zhang, 2012). Another challenge is that crack image components extracted using computer-vision techniques often have many disjoints in crack paths, which prohibits accurate measurement of crack shape. Yun, Mokhtari, and Wu (2015) proposed a novel crack grouping technique, so called the MorphLink-C, to connect fragmented crack image components. They showed that effective crack-grouping technique could improve representation of real crack features by defragmenting

disjoints of crack image components. In their experimental study, MorphLink-C could significantly improve crack detection accuracy by combining classifiers, such as the Artificial Neural Network (ANN).

Although the above computer-vision techniques have advantages over human visual inspection to obtain quantitative data of crack, it has been rarely studied how they can be used in quantifying the severity of pavement deterioration. The imagery information of crack extracted using computer-vision techniques can be measured as different crack features, such as area, length, width, orientation, texture roughness, intensity, and wheel-path position. Thus, characteristics of pavement crack in different deterioration stages can be quantified with the statistics of crack features.

The objective of this study is to understand the statistics of crack features for different surface conditions of flexible pavement, which can be measured using computer-vision techniques. The statistics of crack features are important to understand for effective road maintenance in several dimensions. First, feature statistics can be used to assess the severity of pavement deterioration by measuring various imagery characteristics of surface crack. Hence, the statistical information of crack features in different deterioration stages can be used as fundamental data to justify rehabilitation policies. Second, the relevance of features to be crack can be measured, and the significance of crack features can be ranked based on relevance measures. Third, the optimal number of crack features that can describe important crack characteristics can be determined by avoiding irrelevant or redundant features. Here, the relevance is defined as the characteristic of a feature that has an influence on the output (i.e., crack in this study) and the role cannot be assumed by the rest of features. On the other hand, irrelevance is defined as those features not having any

influence on the output. Redundancy exists whenever a feature can take the role of another feature (Ladha & Deepa, 2011). More formal definition of the relevance and redundancy are presented in (Molina, Belanche, & Nebot, 2002).

To achieve the proposed study objective, statistical characterization of crack features is conducted, associated with computer-vision techniques, which is summarized in Figure 4-1. First, to obtain necessary pavement surface images for the proposed study, gray-scale images of flexible pavement collected with a Multi-Purpose Survey Vehicle (MPSV) by Florida Department of Transportation (FDOT) are used in this study. Using the road surface images, computer-vision techniques developed by Wu et al. (2014) and Yun et al. (2015) are used to extract crack image components (Step 1) and to group crack fragments for better representation of actual crack (Step 2). The computer-vision technique used in this study will be described in Section 4.2. Once the preliminary image processing is completed, the results usually include both crack and no-crack image components. Therefore, crack components have to be separated from no-crack components. In this study, crack components are manually selected (Step 3) to ensure to exclude no-crack components in the statistical characterization that will be followed in the subsequent steps. Seven crack features of area, length, width, orientation, texture roughness, orientation and wheel path position that are commonly used in road rating applications will be extracted for crack and no-crack components (Step 4). Detailed procedures and results of feature extraction will be described in Section 4.3. Finally, crack features will be evaluated for different pavement conditions using their statistical distributions, pair-wise scatter plots, and filter feature subset selection methods (Step 5). Statistical feature subset selection (FSS) is defined as the process of selecting the optimal subset of features. For the type of objective function, FSS can be categorized into the filter and

wrapper FSS methods. In this study, the filter FSS method will be used since the wrapper FSS method can only be used with a predetermined classifier to evaluate feature subsets by predictive accuracy on test data. Wu et al. (2014) shows an application of the wrapper FSS using an artificial neural network (ANN) classifier. The filter FSS methods generally evaluate feature subsets by their information contents, such as interclass distance, statistical dependence, entropies or information-theoretic measures to quantify the relevance and redundancy of crack features. Fischer score, Gini index, information gain, ReliefF, minimum-redundancy-maximum-relevance (mRMR), and fast correlation based filter (FCBF) are employed in this study.

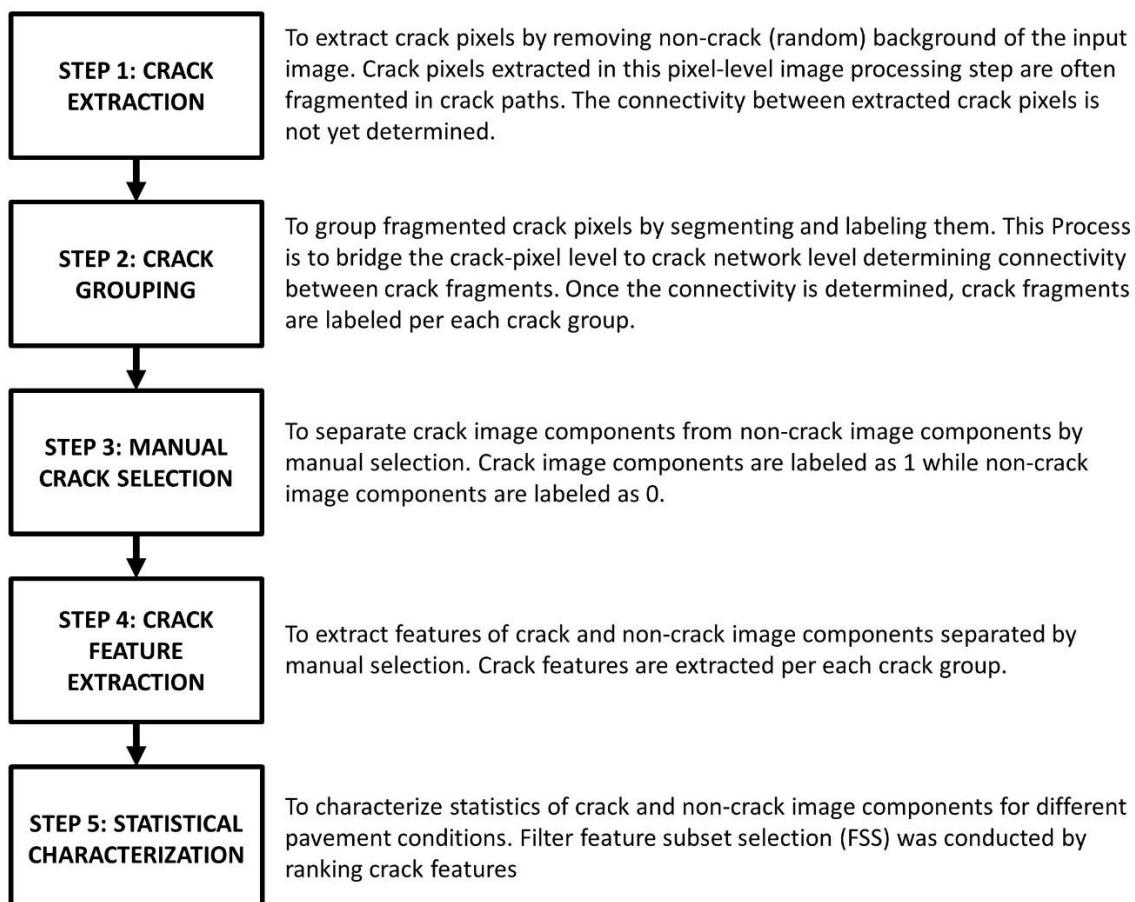


Figure 4-1. Statistical characterization of crack features using computer-vision techniques.

4.2 Preparation of Crack Feature Dataset

4.2.1 FDOT's Road Surface Images of Flexible Pavement

To prepare necessary crack feature datasets, first the flexible pavement images were acquired using the Multi-Purpose Survey Vehicle (MPSV) operated by the Florida Department of Transportation (FDOT). The MPSV is a self-contained van equipped with an Inertial Profiler System, an Inertial Navigation System, a Laser Rut Measurement System (LRMS), and a Laser Road Imaging System (LRIS) to capture pavement images at highway speed. LRIS is composed of two high-resolution lines-can cameras and laser illuminators that are configured to image up to 4-m transverse road section with about 1-mm resolution at speeds of 60 miles per hour. The camera is mounted above 1,960 mm from road surface, which has a 20-mm focal length. The image-sensor pixel size is 0.01 mm. Therefore, the image resolution of LRIS used in this study is 0.98 mm per pixel. Once acquired, the pavement images were classified into three categories of good, intermediate and poor depending on the severity of surface cracking. Sample grey-scale images of flexible pavement are shown in Figure 4-2.

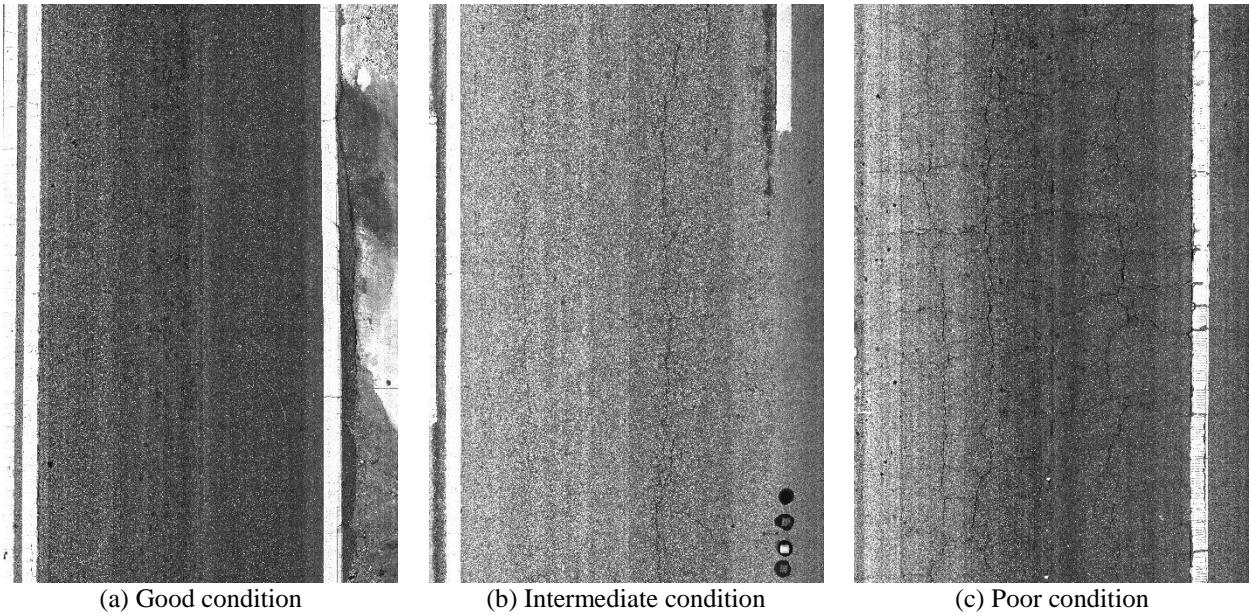


Figure 4-2. Grey-scale images of flexible pavement in different surface conditions.

4.2.2 Preliminary Image Processing Using MorphLink-C

Once necessary pavement images are collected, computer-vision technique was used to extract crack image components from original pavement images. In this study, computer-vision techniques developed by Yun et al. (2015) are be used to extract crack image components by removing random background and to group crack fragments for better representation of pavement cracks. The method of crack grouping is called the MorphLink-C. The advantage of this defragmentation method is to improve the crack detection accuracy and reduce the computation time. The detailed image-processing procedures of the pre-processing techniques are presented in Yun et al. (2015). A brief description of the image-processing technique used in this study is presented in this section.

The morphological image processing technique uses mathematical morphology as a tool for extracting image components that are useful in the representation and description of various region shapes, such as boundaries, skeletons, and convex hull (Gonzalez, Woods, & Eddins, 2009). The morphological technique is advantageous to extract specific image components from random background by manipulating two image-processing controls: the combination of morphological transformations and the shape of the structuring element. In the morphological technique, dilation and erosion are two basic operators to manipulate the digital image of $f(x, y)$. The dilation is an operation that “grows” or “thickens” objects in an image, while the erosion is an operation that “shrinks” or “thins” objects. Specific manners and extent of these operations are controlled by a shape of binary image, called a *structuring element*. The common conventions of the morphological dilation and erosion operations are $I \oplus S$ and $I \ominus S$ respectively, where I is the gray-scale image and S is the structuring element. In crack detection applications, the bottom-hat transform is commonly employed for dark cracks (Jahanshahi, Kelly, Masri, & Sukhatme, 2009; Salembier, 1990):

$$L = \max[(O \circ S_{L\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}) \bullet S_{L\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}, O] - O \quad (4-1)$$

where L is a gray-scale image as the output of the morphological transformations; O is the original gray-scale pavement image; $S_{L\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}$ is the line-shape structuring element with the length of l pixels, which rotates at 0° , 45° , 90° and 135° to detect cracks in arbitrary orientations; \circ is the opening transformation; \bullet is the closing transformation; and the subtraction, ‘ $-$ ’ is a logical

subtraction defined by $X - Y = X \cap NOT Y$. Wu et al. (2014) demonstrated that the morphological bottom-hat transform could effectively detect crack pixels by removing both pixel-level and regional random background from the original pavement images. The length of the structuring element of $l = 100$ pixels, equivalent to 98.0 mm (3.86 in), were used in the analysis to detect cracks whose widths are smaller than the length of the structuring element. However, the bottom-hat transform results as shown in Figure 4-3b have two practical problems in the pavement crack detection. The first is that due to the significant variations of the intensity, the resulting pixels after the bottom-hat transform include both crack and no-crack image components. For example, the rough pavement surface of the lower right corner is falsely detected as crack in Figure 4-3b. The second is that the detected crack image components are fragmented, and therefore a single crack can be detected in many smaller segments as shown on the left of Figure 4-3b. Consequently, the crack fragmentation can significantly bias the statistics of crack features. For example, the length of the fragmented crack is measured smaller than the length of the actual crack.

MorphLink-C can be used to deal with the crack fragmentation problem. To segment the fragmented crack pixels in Figure 4-3b, the extended boundaries were found around the detected pixels in \mathbf{L} as:

$$\mathbf{D} = \mathbf{L} \oplus \mathbf{S}_D \quad (4-2)$$

where \mathbf{L} is the image after the bottom-hat transform; \mathbf{S}_D is a square structuring element with the size of $d \times d$ pixels; and \oplus is the morphological dilation transformation. Once the dilation

boundaries were determined, the fragmented pixels in \mathbf{L} were segmented by the dilation boundary. In this analysis, the size of the structuring element was set to $d = 10$ pixels, which is equivalent to 9.8 mm (0.39 in). That is, two fragmented crack pixels apart up to 19.6 mm in all directions are considered as a single crack in the same segment. A sample result of the crack segmentation is shown in Figure 4-3c. The result shows that the fragmented cracks in the left are segmented within the same dilation boundary. It is observed that the crack pixels within the dilation boundary are still discontinued in Figure 4-3c. Thus, it will be convenient to represent many fragmented pixels within the same boundary as a single continuous line. Such crack-line abstraction can be conducted using the morphological thinning transform as

$$\mathbf{T} = \mathbf{D} - \text{hit-and-miss}(\mathbf{D}, S_T) \quad (4-3)$$

where \mathbf{D} is the binary image after the dilation transform in Equation 4-2; ‘hit-and-miss’ is the morphological hit-and-miss transformation; and S_T is the structuring element for skeletonization by the morphological thinning. The resulting image after the crack-line abstraction is shown in Figure 4-3d. It shows that the fragmented pixels in a boundary are transformed into a continuous single or branched line which represents the region shapes of the actual crack better than with the fragmented pixels. The resulting image after the preliminary image processing includes both crack and no-crack components. Hence, crack image components should be separated from no-crack image components to determine the statistics of crack features.

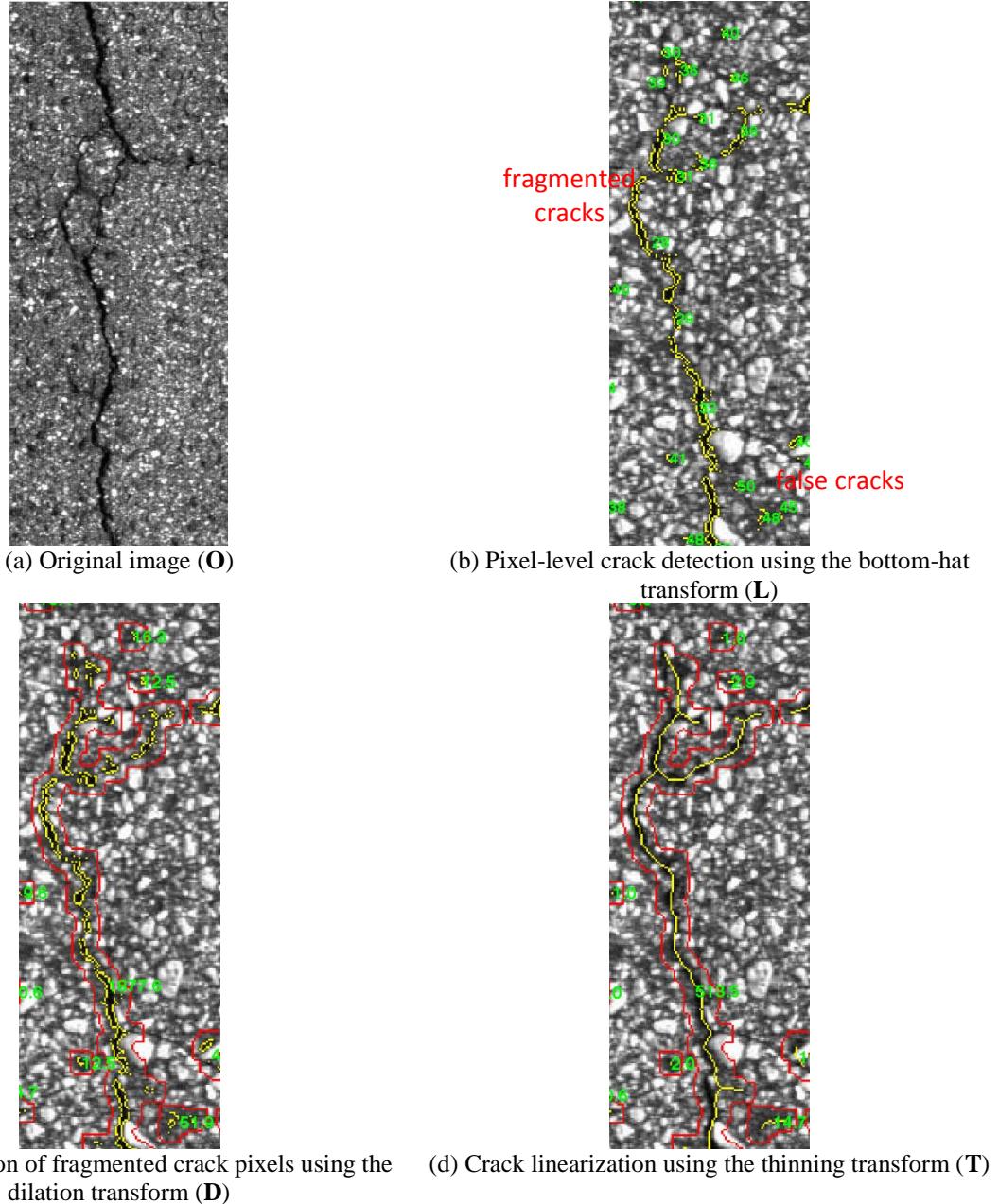


Figure 4-3. Results of the MorphLink-C method.

4.2.3 Manual Separation of Crack and No-crack Objects

To prepare the necessary feature datasets, a total of 26 pavement images were chosen, including 11 images in good, 7 images in intermediate, and 8 images in poor pavement condition. A more number of images were chosen from good conditions since these images usually have a less number of cracks than the images in poor condition. The pavement images were then pre-processed using the image-processing method described before to extract crack and no-crack components. In this study, crack components were manually selected to ensure not to include no-crack components. Then, selected crack components were labeled as one while unselected no-crack components were labeled as zero. Since the manual selection involves a large number of the image components, a software program was developed for accurate labeling with pixel-level accuracy as shown in Figure 4-4. In the figure, selected crack components are shown as blue pixels surrounded in the yellow boundary in the top-right pane. Unselected no-crack components are shown as the yellow boundary only without blue pixels. The software was associated with the image-processing techniques described in earlier to have the crack selection process be semi-automated. The yellow boundary is the dilation boundary, \mathbf{D} , in Equation 4-3.

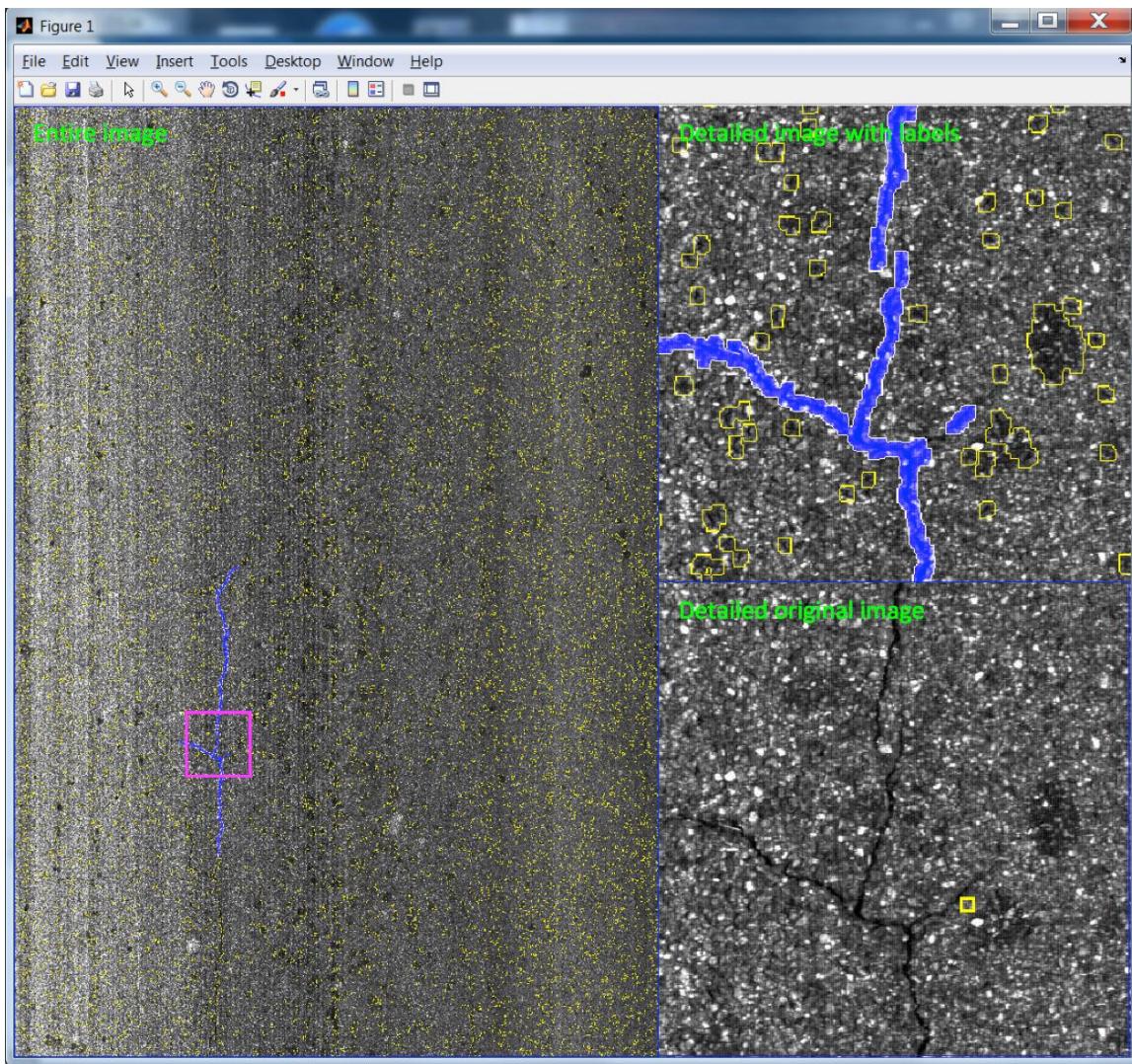
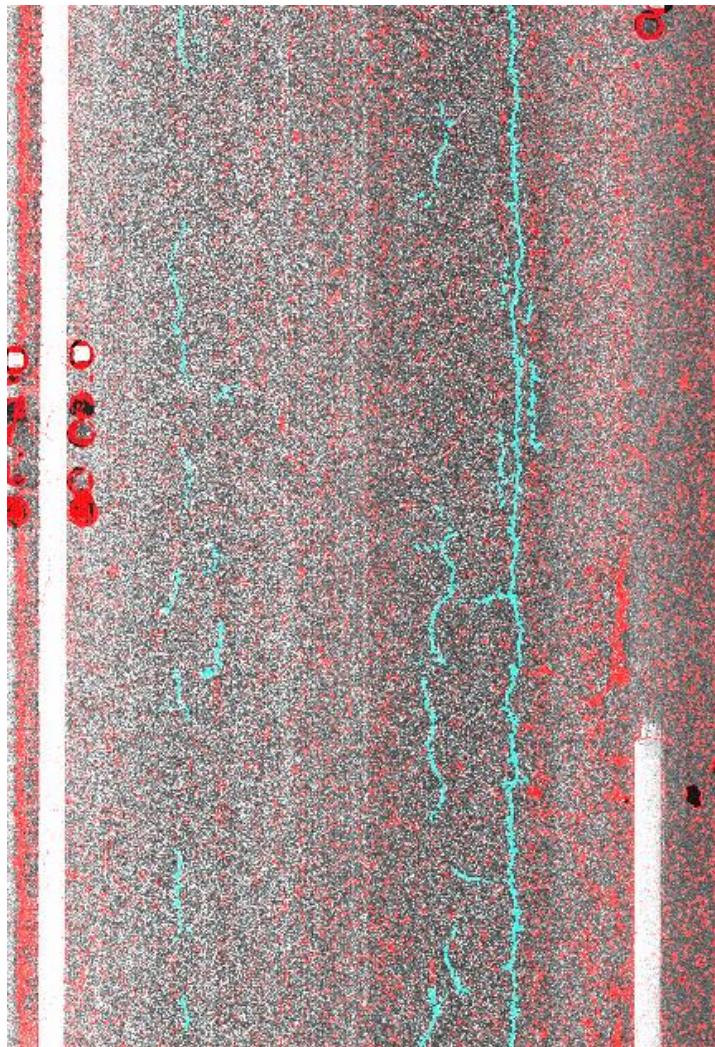


Figure 4-4. The software for manual selection of crack and no-crack image components.

A sample result of manual selection of crack and no-crack image components is shown in Figure 4-5. The numbers and lengths of crack and no-crack image components manually selected are summarized in Table 4-1. Out of 264,981 crack and no-crack components, the total number of crack components used in this study was 3090 whose total length was 444.3 m.



* The crack components are shown in cyan, and the no-crack components are shown in red.

Figure 4-5. A sample result of manual selection of crack components.

Table 4-1. The numbers and length of crack image components in manual selection.

Image #	Crack objects (#)	Crack objects (mm)	All objects (#)	All objects (mm)
Good (11)	1	3	273	7519
	2	13	905	7345
	3	60	7718	8498
	4	70	10876	9138
	5	15	2984	10991
	6	50	4625	10865
	7	94	6280	10685
	8	42	7244	10800
	9	57	8888	10429
	10	38	2586	10108
	11	0	0	10008
Intermediate (7)	12	141	17584	9934
	13	151	20726	10396
	14	116	23428	10477
	15	2	1214	12184
	16	2	203	12035
	17	310	40253	10937
	18	144	29242	11122
	19	212	22422	10705
Poor (8)	20	183	26962	10230
	21	213	33843	9843
	22	252	27005	10156
	23	219	38955	10984
	24	239	34580	11072
	25	320	55307	10219
	26	144	20203	8301
	Sum	3090	444306	264981
All	Mean	118.8	17088.7	10191.6
	Stdv	99.6	15288.2	1189.5
				3915102
				150580.8
				17061.2

4.3 Feature Extraction of Crack and No-Crack Objects

4.3.1 Methods of Feature Extraction

Once crack and no-crack image components were manually selected, different features were extracted, which characterized the imagery information of crack. In this study, a total of seven features were used, which have been commonly employed in road rating applications, including

area, length, orientation, texture roughness, intensity and wheel path position. The extraction method of crack and no-crack features are described in this section.

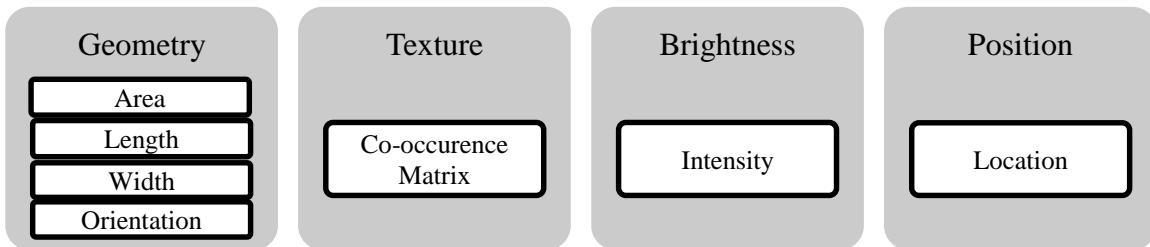


Figure 4-6. Features extracted from flexible pavement images for crack detection.

Area: The area feature measures the area of the crack and no-crack image components obtained after the top-hat transform (L), which are within the same dilation boundary. It can be calculated by counting the number of pixels within the same dilation boundary as

$$Area_D = \sum_{L=1}^{Number\ of\ L\ in\ D} (Area_L)_D \quad (4-4)$$

where D is the segment index determined after the dilation transformation; L is the index of the fragmented pixel clusters after the top-hat transformation, which are within the same dilation boundary of D ; and $(Area_L)_D$ is the area of the L -th image component in the D -th segment.

Length: The length feature measures the length of the ‘crack’ and ‘no-crack’ thinned lines after the thinning transform (T), which are surrounded with the same dilation boundary. For a single crack, the length feature should measure the length along the crack path particularly when

the crack is not straight. For a branched crack, the length feature should measure the summed length of all crack branches. The length feature can be calculated as

$$Length_D = \text{Number of pixels on the thinned line in } D \quad (4-5)$$

Width: In flexible pavement maintenance, the capability of measuring crack width is important since crack width is a key measure to determine the severity of pavement deterioration in road rating applications. Crack width cannot be measured using many conventional image processing-based crack segmentation methods, such as seed-growing method by Li, Zou, Zhang, and Mao (2011) and CrackTree by Zou, Cao, Li, Mao, and Wang (2012). Using the proposed method, MorphLink-C, crack width can be accurately measured without involving heavy computation (Yun et al., 2015). Since crack width varies along the crack-line path, the averaged crack width can be determined as

$$(\text{Averaged crack width})_k = \frac{\sum_i (\text{Area of fragmented crack pixels})_k}{\sum_i (\text{Length of the connected crack line})_k} \quad (4-6)$$

where ' $\sum_i (\text{Area of fragmented crack pixels})_k$ ' is the summation of the total areas of the fragmented crack pixels within the k -th dilation boundary; i is the index of the fragmented objects in the k -th dilation boundary; and ' $\sum_i (\text{Length of the connected crack line})_k$ ' is the summation of the total lengths of connected crack line within the k -th dilation boundary. Therefore, the crack

width was not directly measured from image components but was calculated using the area and length features.

Orientation: Finally, the orientation feature measures the angle in radian between the horizontal direction and the major axis of the ellipse that encloses the dilation boundary after the dilation transform (**D**). The measurement of the orientation feature is illustrated in Figure 4-7.

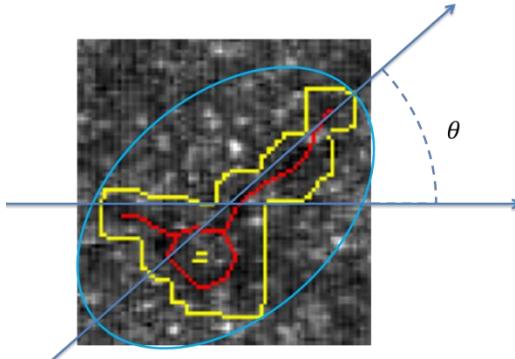


Figure 4-7. Measurement of the orientation feature.

Texture: The texture feature measures the roughness of the pavement surface. The main logic behind defining a texture index is that rough parts of the pavement have higher potential for developing cracks. Co-occurrence matrix, first introduced by Haralick, Shanmugam, & Dinstein, (1973), has been used to quantify the roughness of surfaces from 2D images. The mathematical expression of co-occurrence matrix for a $m \times n$ grey scale image can be written as:

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1 & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0 & \text{otherwise} \end{cases} \quad (4-7)$$

where C is the co-occurrence matrix, i and j are the grey levels, Δx and Δy describe the offset between the pixels and $I(p, q)$ is the grey level of the image at (p, q) . By this definition, co-occurrence matrix presents the relative frequency of two intensity values (i, j) at a given offset $(\Delta x, \Delta y)$. In other words, each entry of the co-occurrence matrix represents the number of times, two pixels with grey levels i and j occurred at distance $(\Delta x, \Delta y)$. In image-based pavement evaluation, finding neighboring dark pixels is important because pavement defects usually appear as blocks of pixels with low grey-levels. Therefore, offset is usually defined as the 8-neighboring pixels and the co-occurrence index is described as (Ahuja & Rosenfeld, 1978; Kaseko & Ritchie, 1993; Sahoo, Soltani, & Wong, 1988)

$$C = C_{0,1} + C_{1,1} + C_{1,0} + C_{1,-1} + C_{0,-1} + C_{-1,-1} + C_{-1,0} + C_{-1,1} \quad (4-8)$$

For a grey-scale image with 256 levels of grey, the co-occurrence matrix will have 256×256 entries where element (i, j) represents the number of times, grey levels i and j were adjacent in the image. The main assumption in using the co-occurrence matrix is that the pixels inside the objects are homogeneous; therefore they will contribute to near-diagonal entries while the gradient of grey-level change is high for pixels around the boundaries (or edges) and they will mainly contribute to off-diagonal elements (Sahoo et al., 1988). Using this argument, a high number of near-diagonal elements for low grey levels (darker pixels) can be an indication of highly cracked and rough surface. Although the co-occurrence matrix was originally defined for grey-scale images, the concept can be extended to any type of image. In this study, the co-occurrence matrixes

of binarized images are employed. For this purpose, each binary image is divided into 200×200 pixels regions (tiles) first. Then, the co-occurrence index of each tile is calculated using Equations 4-7 and 4-8. To normalize the index, values are divided by the total number of pixels in each tile (200×200). Finally the normalized frequency of neighboring crack pixels is calculated from the 2×2 normalized co-occurrence matrix.

Intensity: The intensity feature measures the pixel intensity of the ‘crack’ and ‘no-crack’ image components obtained after the top-hat transform (L), which are surrounded with the same dilation boundary. Since the image components in the boundary can have different areas, the intensity is weighted by the area as

$$Intensity_D = \frac{\sum_{L=1}^{Number\ of\ L\ in\ D} (Intensity_L)_D \times (Area_L)_D}{\sum_{L=1}^{Number\ of\ L\ in\ D} (Area_L)_D} \quad (4-9)$$

where $(Intensity_L)_D$ is the averaged pixel intensity of the L -th image component in the D -th segment.

Location: The location of wheel-path designation feature determines whether the ‘crack’ and ‘no-crack’ image components after the bottom-hat transform (L) are located inside or outside of the wheel path. This feature may be useful since the pavement crack is usually more severe along the wheel path due to traffic loads; consequently the regions inside the wheel path have higher potential for developing cracks. Therefore, some highway management authorities begin to consider the wheel-path effect in their pavement condition survey. For example, in *Flexible Pavement Condition Survey Handbook* by the Florida Department of Transportation, (2012a), the

pavement surface is divided into the inside of the wheel path (CW) and the outside of the wheel path (CO) as shown in Figure 4-8a. To implement this concept, first the lane markers of the road are identified, and then the region between the insides of the left and right lane markers are divided into five smaller regions for the proportions of $CO:CW:CO:CW:CO = 1:2:2:2:1$ based on their widths. Once the regions are divided, the image components within the same dilation boundary are labeled as follows: two for the components in CO, one for the components in CW, and zero for the components outsides of the lane markers. A sample result of determining the location feature is shown in Figure 4-8b.

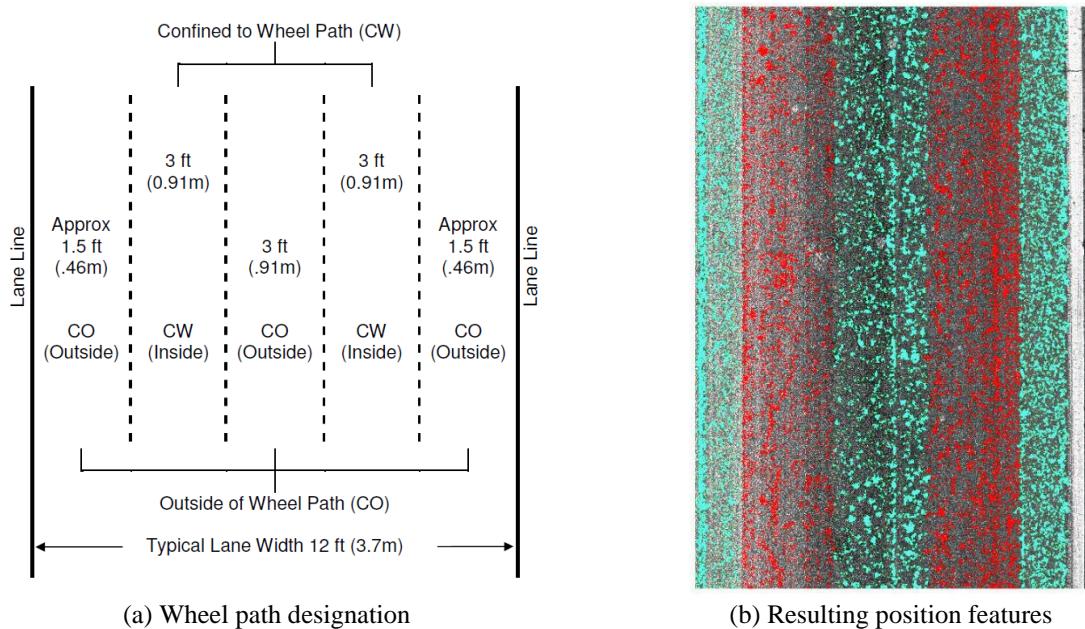


Figure 4-8. Wheel path position feature.

4.3.2 Feature Extraction Results

The results of feature extraction for crack image components for good, intermediate, and poor pavement conditions are summarized in Table 4-2. The table shows the statistics of ranges, means, modes, medians, and standard deviations of the seven crack features.

Table 4-2. The feature dataset of the crack image components.

	Features	Data types	Units	Range	Mean	Mode	Median	Standard deviation
Good	Length	Continuous	mm	[0.98,3743.6]	116.13	1.96	34.30	281.60
	Area	Continuous	mm ²	[9.60,15478.0]	362.01	9.60	84.04	1080.50
	Intensity	Continuous	0~255 (8-bit)	[16.22,49.43]	30.71	26.5	30.64	6.13
	Texture	Continuous	0~1	[0.016, 0.347]	0.0908	0.038	0.083	0.0446
	Location	Discrete	0, 1, 2	{0,1,2}	-	2	2	-
	Orientation	Continuous	deg	[0.069,179.63]	92.2316	0.069	96.26	26.9462
	Width	Continuous	mm	[1.10,32.34]	3.62	1.96	2.74	3.06
Intermediate	Length	Continuous	mm	[0.98, 4570.7]	150.11	2.94	39.69	386.90
	Area	Continuous	mm ²	[9.60, 20489.0]	524.61	9.60	105.64	1589.60
	Intensity	Continuous	0~255 (8-bit)	[16.50,76.00]	40.97	30.00	40.92	9.90
	Texture	Continuous	0~1	[0.023, 0.537]	0.12	0.09	0.10	0.07
	Location	Discrete	0, 1, 2	{0,1,2}	-	2	2	-
	Orientation	Continuous	deg	[0,179.93]	89.99	0.00	88.43	30.42
	Width	Continuous	mm	[1.09,43.12]	3.98	2.94	3.15	3.54
Poor	Length	Continuous	mm	[0.98, 6377.8]	142.59	1.96	34.30	430.37
	Area	Continuous	mm ²	[9.60, 27222.0]	512.19	9.60	94.12	1803.90
	Intensity	Continuous	0~255 (8-bit)	[11.23, 108.25]	39.20	28.00	39.18	12.09
	Texture	Continuous	0~1	[0.023, 0.852]	0.1349	0.13	0.12	0.0773
	Location	Discrete	0, 1, 2	{0,1,2}	-	2	2	-
	Orientation	Continuous	deg	[0,180]	88.5172	0	85.46	35.14
	Width	Continuous	mm	[1.07,34.79]	4.06	4.90	3.13	3.55

4.4 Discussion

4.4.1 Statistical Distributions of the Feature Data

Figure 4-9 shows the statistical distributions of crack and no-crack features for good, intermediate and poor pavement conditions. Some observations about the statistical characteristics of crack features are discussed in this section.

Area and length: The ranges of area and length are wide for all pavement conditions. The upper boundary of the ranges increases as pavement condition is poorer, while the lower bounds remained the same. Based on this result, good pavement has less severe cracks, such as single or branched cracks, while poor pavement has all kinds of cracks including more severe cracks, such as alligator cracks. The standard deviations also increase as pavement condition is poorer. It indicates that the dispersion of crack size increases when pavement deteriorates. Poor pavement of course has more cracks in number. In addition, those distributions are non-gaussian in the majority of small cracks. The frequency of large cracks is very low although large cracks are more interesting in maintenance practices. Therefore, false negative error needs to be minimized not to miss large cracks in automated crack detection applications.

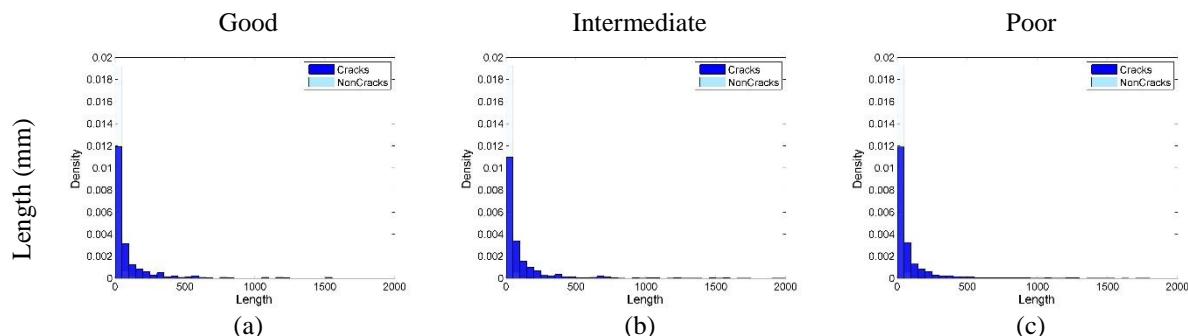
Intensity: The distribution becomes wider when pavement deteriorates further. The standard deviation increases from 6.13 for good condition to 12.09 for poor condition. The mean also tends to increase for poorer condition: 30.71 for good, 40.97 for intermediate, and 39.20 for poor conditions. Since higher intensity means a brighter pixel in a gray-scale image, one can observe that crack image components become brighter as pavement deteriorates. It is due to eroded

surface of aging pavement. The distribution of no-crack component is wider and its mean is higher than those of crack components since crack has usually lower intensity than random background.

Texture: The distributions are uni-modal and skewed toward a lower co-occurrence index. The mean increases from 0.0908 for good pavement to 0.1349 for poor pavement. Since a higher co-occurrence index indicates more surface cracking, the feature statistics also reflects the severity of deterioration. The standard deviation also increases as pavement is in poor condition: 6.13 for good, 9.90 for intermediate and 12.09 for poor conditions.

Orientation: The distribution for good condition in Figure 4-9n has two modes near the angles of -90° and 90° . It should be noted that the angle of -90° is equal to the angle of 90° , which indicates the longitudinal direction. Therefore, one can observe that the majority of crack in good pavement condition is longitudinal cracks. As pavement condition is poorer, the frequency of 0° becomes higher. It indicates that transverse crack is developed in later stages of pavement deterioration than longitudinal crack.

Width: The means of crack width in Equation 4-6 were 3.62 mm for good, 3.98 mm for intermediate, and 4.06 mm for poor condition in mean. The standard deviations were 3.06 mm for good, 3.54 mm for intermediate, and 3.55 mm for poor condition. The statistics show that crack width increases when pavement is more deteriorated.



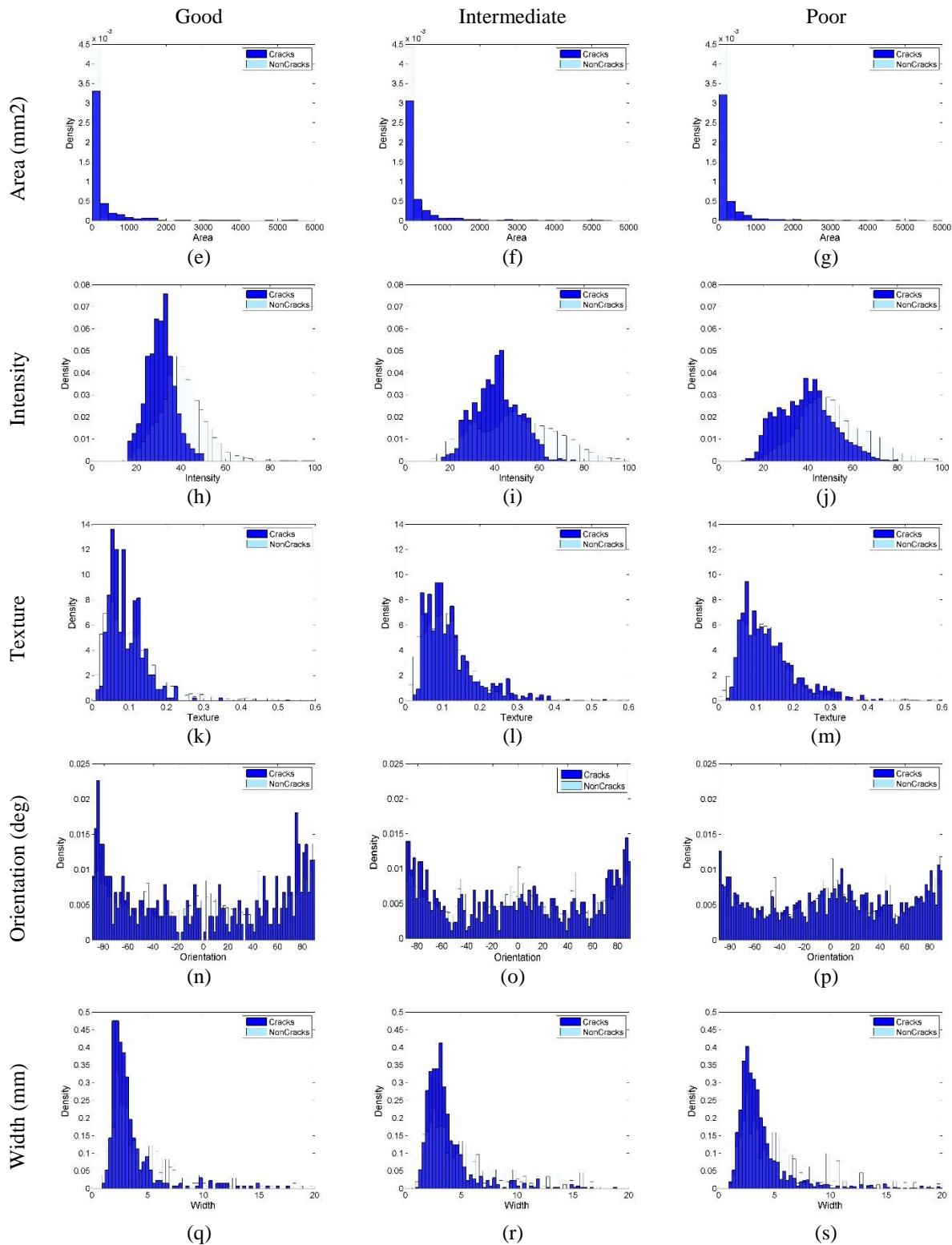


Figure 4-9. Statistical distributions of crack and no-crack features.

4.4.2 Statistical Correlations between Crack Features

Statistical correlations between the six crack features used in this study are compared. Pairwise scatter plots of the features for all pavement conditions are shown in Figure 4-10, and correlation coefficient plots for all, good, intermediate and poor pavement conditions are shown in Figure 4-11. The correlation coefficient was calculated as

$$\rho_{ij} = \frac{E[F_i - \mu_{F_i}]E[F_j - \mu_{F_j}]}{\sigma_{F_i}\sigma_{F_j}} \quad (4-10)$$

where F_i is the i -th crack feature; μ_{F_i} is the mean of F_i ; σ_{F_i} is the standard deviation of F_i ; and $0 \leq \rho_{ij} \leq 1$. Some observations of statistical correlation are discussed in this section. First, a strong correlation is observed between area and length in Figure 4-10, and the corresponding correlation coefficient is 0.99 for all four pavement conditions in Figure 4-11. One can easily postulate that crack area is large for a long crack. Hence, the area and length are redundant with a large correlation coefficient. Having both features should be avoided for the optimal FSS. Second, notable correlation is observed between the area and texture, and the length and texture. The correlation coefficients of the area and texture are between 0.41 and 0.48 and those of the length and texture are between 0.40 and 0.48. This result is why the surface roughness increases (or the co-occurrence index increases) when more cracks exist in the corresponding texture tile. Third, no significant difference is observed in the correlation coefficients for different pavement conditions. This result indicates that redundancy between the features is rarely affected by pavement

condition. Finally, although the pair-wise scatter plots and correlation coefficient plots provides useful graphical representation to understand statistical relations between features, they have some limitations. Since these plots compare only two features at a time, statistical relations between more than two features can be hardly understood. In addition, since these plots provide only qualitative information of feature statistics, more advanced methods are needed for quantitative measures of the relevance and redundancy to rank crack features for the optimal FSS, considering all features simultaneously.

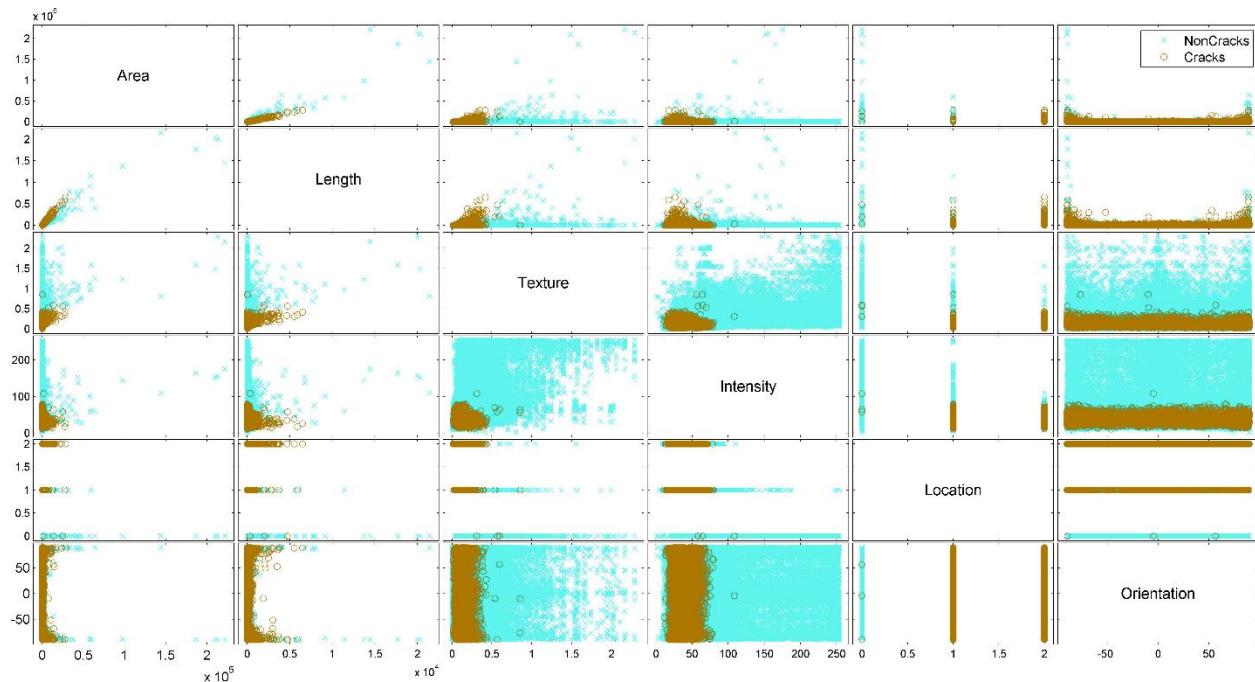


Figure 4-10. Pair-wise scatter plots of the feature correlation.

Orientation	-0.03	-0.03	-0.03	0.01	-0.01	1.00
Location	0.00	0.01	-0.06	0.14	1.00	-0.01
Intensity	-0.13	-0.13	-0.10	1.00	0.14	0.01
Texture	0.43	0.42	1.00	-0.10	-0.06	-0.03
Length	0.99	1.00	0.42	-0.13	0.01	-0.03
Area	1.00	0.99	0.43	-0.13	0.00	-0.03

(a) All conditions

Orientation	-0.01	-0.02	-0.01	0.03	-0.04	1.00
Location	0.04	0.03	-0.18	-0.30	1.00	-0.04
Intensity	-0.13	-0.15	-0.05	1.00	-0.30	0.03
Texture	0.46	0.48	1.00	-0.05	-0.18	-0.01
Length	0.99	1.00	0.48	-0.15	0.03	-0.02
Area	1.00	0.99	0.46	-0.13	0.04	-0.01

(b) Good condition

Orientation	-0.04	-0.04	-0.03	-0.03	0.00	1.00
Location	0.08	0.08	0.09	0.36	1.00	0.00
Intensity	-0.14	-0.14	-0.11	1.00	0.36	-0.03
Texture	0.48	0.48	1.00	-0.11	0.09	-0.03
Length	0.99	1.00	0.48	-0.14	0.08	-0.04
Area	1.00	0.99	0.48	-0.14	0.08	-0.04

(c) Intermediate condition

Orientation	-0.03	-0.03	-0.04	0.02	-0.00	1.00
Location	-0.03	-0.02	-0.07	0.23	1.00	-0.00
Intensity	-0.13	-0.13	-0.13	1.00	0.23	0.02
Texture	0.41	0.40	1.00	-0.13	-0.07	-0.04
Length	0.99	1.00	0.40	-0.13	-0.02	-0.03
Area	1.00	0.99	0.41	-0.13	-0.03	-0.03

(d) Poor condition

Figure 4-11. Correlation coefficients of the cracks feature dataset.

4.4.3 Ranking of Crack Features for Optimal FSS

Relevance and *redundancy* of crack features are important measures for feature evaluation.

By selecting the optimal feature subset, one can reduce the dimension of feature data space by including only relevant features and excluding redundant features. Consequently, efficiency of statistical analysis can be improved using optimal feature subsets. Mathematical definition of relevance and redundancy can be expressed as:

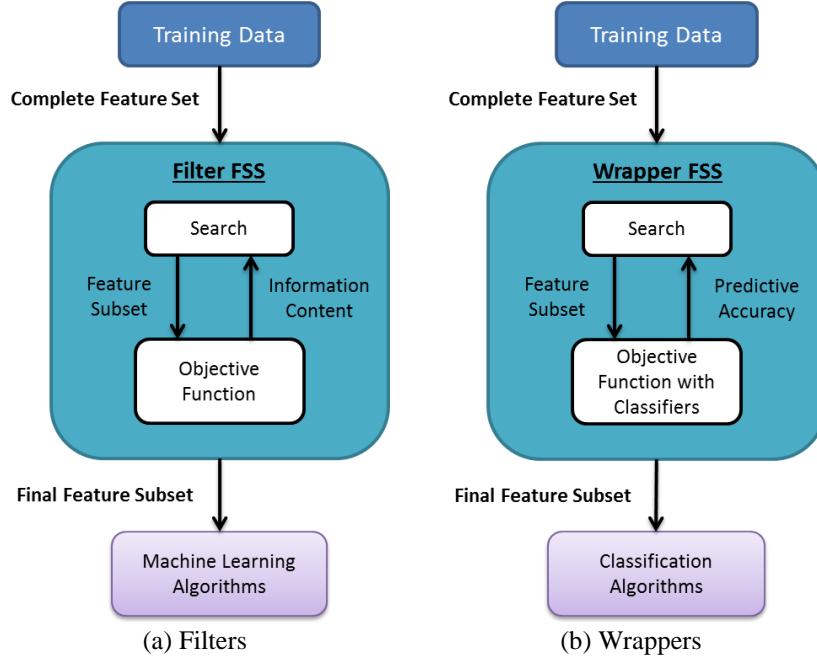
Feature F_i is relevant iff

$$\exists S'_i \subseteq S_i \text{ such that } P(C|G_i, S'_i) \neq P(C|S'_i) \quad (4-11)$$

Feature F_i is redundant iff

$$P(C|F_i, S_i) = P(C|S_i) \text{ but } \exists S'_i \subseteq S_i \text{ such that } P(C|F_i, S'_i) \neq P(C|S'_i) \quad (4-12)$$

where F is the set of all features; F_i is the i -th feature; $S_i = F - \{F_i\}$; C is the class label; and P is the conditional probability of the class label. Statistical FSS methods have been developed for feature evaluation, which can be categorized into filter, wrapper, and hybrid FSS methods. The *filter* FSS method generally evaluates feature subsets by their information contents, such as interclass distance, statistical dependence, entropies or information-theoretic measures. The *wrapper* FSS method is generally associated with a classifier to evaluate feature subsets by their predictive accuracy on test data. The *hybrid* FSS method is the combination of the filter and wrapper methods. This study focuses on using the filter FSS method since the objective of this study is for statistical characterization of crack features by evaluating information contents in the feature datasets rather than for evaluating prediction accuracy of associated classifiers. Wu et al. (2014) discussed results of the wrapper FSS method using an ANN classifier for pavement crack detection. A comparison of the filter and wrapper FSS process are illustrated in Figure 4-12.



*Courtesy of Dr. Gutierrez-Osuna, University of Texas A&M.

Figure 4-12. Filter and wrapper FSS algorithms.

In this study, six different filter FSS models are used, including Fisher score, Gini index, information gain, ReliefF, fast correlation based filter (FCBF), and minimum-redundancy-maximum-relevance (mRMR). The selected methods are widely accepted filter models that cover a wide range of statistical criteria on information gained, entropy, correlation and distance for feature evaluation. Description of each model is presented below. In this discussion, features are the attributes of image components, including the area, length, intensity, texture, location and orientation. Therefore, each image component has a set of six feature instances describing its imagery information. On the other hand, the image components belong to one of two classes, crack or no-crack, which were manually selected. Therefore, all the set of the six feature instances can be assigned to those two class labels. The goal of the filter FSS is to find the optimal feature subset

through the statistical characterization of the relations between the feature instances and the class labels. This concept is illustrated in Figure 4-13.

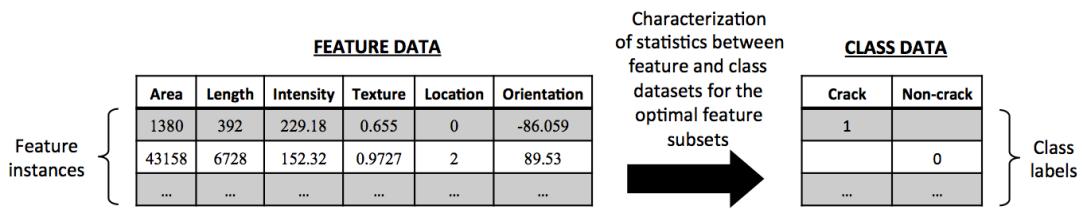


Figure 4-13. Terminology and procedure of the filter FSS.

Fisher Score: The objective is to find the feature which has the largest distance between data points from different classes as well as the smallest distance between the data points of the same class. Thus, if all the instances of data are plotted in feature spaces, the mean of feature data from each class should be distinctly different from others, and data points in each class should be located around its mean value (Gu, Li, & Han, 2012). The mathematical formulation of Fisher score can be described as

$$S_i = \frac{\sum_{j=1}^J n_j (\mu_{ij} - \mu_i)^2}{\sum_{j=1}^J n_j \rho_{ij}^2} \quad (4-13)$$

where S_i is the fisher score of the i -th feature; μ_{ij} and ρ_{ij}^2 are the mean and the variance of the i -th feature in the j -th class respectively; n_j is the number of instances in the j -th class; μ_i

is the mean of instances of the i -th feature; and J is the number of classes. The limitation is that the redundancy cannot be measured using Fisher score since each feature is evaluated separately.

Gini index: Using the Gini index, the number of class labels (C_j) for a given feature (f_i) is calculated. The feature is *impure* if it contributes to different class labels (i.e., instances of the feature belong to different classes). The features can be ranked using the calculated impurity. By this means, a better feature mostly contributes to a single class label (Teoh, Nguwi, & Cho, 2009).

The Gini index is defined as:

$$G_i = 1 - \sum_{j=1}^J [p(C_j|f_i)]^2 \quad (4-14)$$

where G_i is the Gini index of the i -th feature; J is the number of classes; f_i is the i -th feature; and C_j is the j -th class. The $p(C_j|f_i)$ indicates the probability of having different class labels given the i -th feature instances. The maximum Gini index is $(1 - 1/K)$ when features are equally distributed in different classes. The minimum Gini index is zero when all instances of a feature belong to a single class (Zhao et al., 2010). Similar to Fisher score, redundancy of features cannot be evaluated using the Gini index.

Information Gain: Features are ranked based on dependence between features and class (Tang, Alelyani, & Liu, 2013). The dependence can be determined by measuring uncertainty associated with each feature, as

$$IG(f_i, C) = H(f_i) - H(f_i|C) \quad (4-15)$$

where IG is the information gain of the i -th feature to separate different class labels; $H(f_i)$ is the entropy of the i -th feature; f_i is the i -th feature; and $H(f_i|C)$ is the entropy of f_i given all class labels (C). $H(f_i)$ and $H(f_i|C)$ can be calculated as

$$H(f_i) = -\sum_k^K p(x_k) \log_2(p(x_k)) \quad (4-16)$$

$$H(f_i|C) = -\sum_j^J p(C_j) \sum_k^K p(x_k|C_j) \log_2(p(x_k, C_j)) \quad (4-17)$$

where x_k is the k -th instance of the i -th feature; f_i is the i -th feature; C_j is the j -th class; K is the number of instances of f_i ; and J is the number of classes. According to Equation 4-15, the information gain is the difference between the entropy of a feature and the entropy of a feature given class labels as additional information (Forman, 2003). Thus, the maximum value of $IG(f_i, C)$ is one when features have the same number of instances (i.e., $H(f_i) = 1$) but all the instances belong to a unique class label (i.e., $H(f_i|C) = 0$). The information gain can be used to measure the relevance of features. The redundancy of features cannot be measured using this method.

ReliefF: The objective is to select features having most their instances belong to a single class label. The relative distance of features from different class labels can be calculated as (Tang et al., 2013):

$$RF_{f_i} = \frac{1}{2} \sum_{k=1}^K d(X_{ik} - X_{iM_k}) - d(X_{ik} - X_{iH_k}) \quad (4-18)$$

where RF_{f_i} is the ReliefF weight, assigned to the i -th feature; K is the number of instances randomly selected from each feature; X_{ik} is the value of the k -th instance in the i -th feature; X_{iM_k} is the value of the nearest instances to X_{ik} with the same class label; X_{iH_k} is the values of the nearest instances to X_{ik} with different class labels; and d is a distance measure. ReliefF finds the K -nearest hits and misses for given instances of a feature. Using the nearest hits and misses, the contribution of each feature for separating class labels is calculated (Hall, 1999).

Minimum-Redundancy-Maximum-Relevance (mRMR): mRMR selects the subset of features that has the minimum redundancy between its members, and the maximum relevance with class labels (Tang et al., 2013). The minimum redundancy is defined as:

$$\min W = \frac{1}{|S|^2} \sum_{i,j \in S} |Cor(f_i, f_k)| \quad (4-19)$$

where W is the measure of redundancy between the i -th feature and k -th feature of the selected feature subset; f_i is the i -th feature; f_k is the k -th feature of the selected feature subset; $Cor(f_i, f_k)$ is the correlation between instances of f_i and f_k ; and $|S|$ is the number of elements of the selected feature subset. The maximum relevance can be measured using the F-statistic of data as:

$$\max V = \frac{1}{|S|^2} \sum_i F(f_i, C_j) \quad (4-20)$$

where V is the measure of relevancy between i -th feature of the selected feature subset and j -th class label; C_j is the j -th class; $F(f_i, C_j)$ is the F-statistics. Therefore, this method is advantageous over the previous methods since it ranks features based on both the relevance and redundancy. A limitation of mRMR is that it is a ranking-based method, so that quantitative measurements, such as interclass distance, statistical dependence, entropies or information-theoretic measures, are not available.

Fast Correlation Based Filter (FCBF): In this method, symmetrical uncertainty (SU) is employed as a measure of correlation between features and classes, is employed. SU can be calculated as (Zhao et al., 2010):

$$SU(f_i, C) = 2 \frac{IG(f_i, C)}{H(f_i) + H(C)} \quad (4-21)$$

where, IG is the information gain in Equation 4-15; and H is the entropy in Equation 4-16. FCBF method selects features that have not only the highest correlation with class labels (i.e., the largest $SU(f_i, C)$ value) but also the least correlation with other features (i.e., the least $SU(f_i, f_j)$ value).

Using the above six filter FSS methods, the crack features were evaluated. The results of statistical feature evaluation for different pavement conditions are summarized in Table 4-3. The numbers in the parentheses in the table indicate the ranks of the features with a higher relevance. For FCBF, the X mark indicates that the corresponding feature is selected as the optimal feature

subset, while the - mark indicates that the corresponding feature is not selected as the optimal feature subset.

Table 4-3. Results of statistical feature evaluation for different pavement conditions.

All Conditions

	Area	Length	Texture	Intensity	Location	Orientation
Fisher Score	0.0687 (4)	0.0994 (3)	0.0024 (5)	1.0036 (2)	1.2945 (1)	0.0011 (6)
Gini Index	0.0228 (2)	0.0225 (1)	0.0236 (5)	0.0235 (4)	0.0236 (3)	0.0237 (6)
Info. Gain	0.0160 (2)	0.0172 (1)	0.0020 (5)	0.0033 (4)	0.0055 (3)	0.0002 (6)
ReliefF	0.0009 (6)	0.0024 (5)	0.0164 (3)	0.0295 (2)	0.0515 (1)	0.0147 (4)
mRMR	(6)	(1)	(5)	(4)	(3)	(2)
FCBF	-	X	-	-	X	-

Good Condition

	Area	Length	Texture	Intensity	Location	Orientation
Fisher Score	0.0007 (3)	0.0043 (1)	0.0003 (5)	0.0006 (4)	0.0038 (2)	0.0001 (6)
Gini Index	0.0108 (2)	0.0107 (1)	0.0110 (5)	0.0109 (4)	0.0109 (3)	0.0110 (6)
Info. Gain	0.0067 (2)	0.0078 (1)	0.0031 (5)	0.0025 (3)	0.0035 (4)	0.0004 (6)
ReliefF	0.0006 (6)	0.0023 (4)	0.0211 (5)	0.0416 (2)	0.0468 (1)	0.0305 (3)
mRMR	(3)	(1)	(6)	(5)	(4)	(2)
FCBF	-	X	-	X	-	-

Intermediate Condition

	Area	Length	Texture	Intensity	Location	Orientation
Fisher Score	0.0032 (3)	0.0272 (1)	0.0001 (5)	0.0020 (4)	0.0056 (2)	0.0000 (6)
Gini Index	0.0211 (2)	0.0208 (1)	0.0222 (5)	0.0211 (3)	0.0211 (4)	0.0222 (6)
Info. Gain	0.0172 (2)	0.0177 (1)	0.0016 (5)	0.0071 (3)	0.0050 (4)	0.0000 (6)
ReliefF	0.0006 (6)	0.0029 (5)	0.0200 (3)	0.0604 (1)	0.0551 (2)	0.0180 (4)
mRMR	(5)	(1)	(6)	(4)	(3)	(2)
FCBF	-	X	-	-	-	-

Bad Condition

	Area	Length	Texture	Intensity	Location	Orientation
Fisher Score	0.0029 (4)	0.0197 (1)	0.0000 (5)	0.0068 (3)	0.0128 (2)	0.0000 (6)
Gini Index	0.0386 (2)	0.0382 (1)	0.0411 (5)	0.0406 (3)	0.0406 (4)	0.0412 (6)
Info. Gain	0.0279 (2)	0.0295 (1)	0.0029 (5)	0.0137 (3)	0.0113 (4)	0.0000 (6)
ReliefF	0.0011 (6)	0.0028 (5)	0.0191 (3)	0.0405 (2)	0.0737 (1)	0.0189 (4)
mRMR	(4)	(1)	(6)	(5)	(3)	(2)
FCBF	-	X	-	-	X	-

It is observed that the ranks using the Gini index and information gain methods are almost identical for all pavement conditions. Since the Gini index and information gain measure similar quantities (i.e., the number of instances of a feature that belong to a single class are measured in both methods), the observed rankings are justifiable. Since the first four methods in Table 4-3, the Fisher score, Gini index, information gain, and ReliefF, evaluate features only based on the relevance, the ranks of these methods can be combined using Borda scoring method. Borda scoring is a simple multi-criteria decision making (MCDM) method. In this method, rankings of each feature under different criteria (i.e., the objective functions of those four FSS methods in this study) are added up, and then features can be ranked using the aggregated rank (Mokhtari, Madani, & Chang, 2012; Madani et al., 2014; Read et. al, 2012). Borda scoring (BS) procedure can be expressed as:

$$BS_{f_i} = R_{f_i|Fisher\ Score} + R_{f_i|Gini\ Index} + R_{f_i|Info.\ Gain} + R_{f_i|ReliefF} \quad (4-22)$$

where BS_{f_i} is the Borda score of i -th features and $R_{f_i|FSS\ method}$ is the rank of f_i under specified FSS method (Mokhtari, 2012). A better feature should have the smaller BS value. The combined results of the Fisher score, Gini index, information gain and ReliefF using the Borda scoring are summarized in Table 4-4. The numbers in the parentheses are the ranking of crack features based on the relevance. From the results, the length, wheel-path location and intensity are top three features that have high relevance to be pavement crack. On the other hand, the orientation is ranked the lowest. It should be noted that the ranking of ReliefF is very different from the

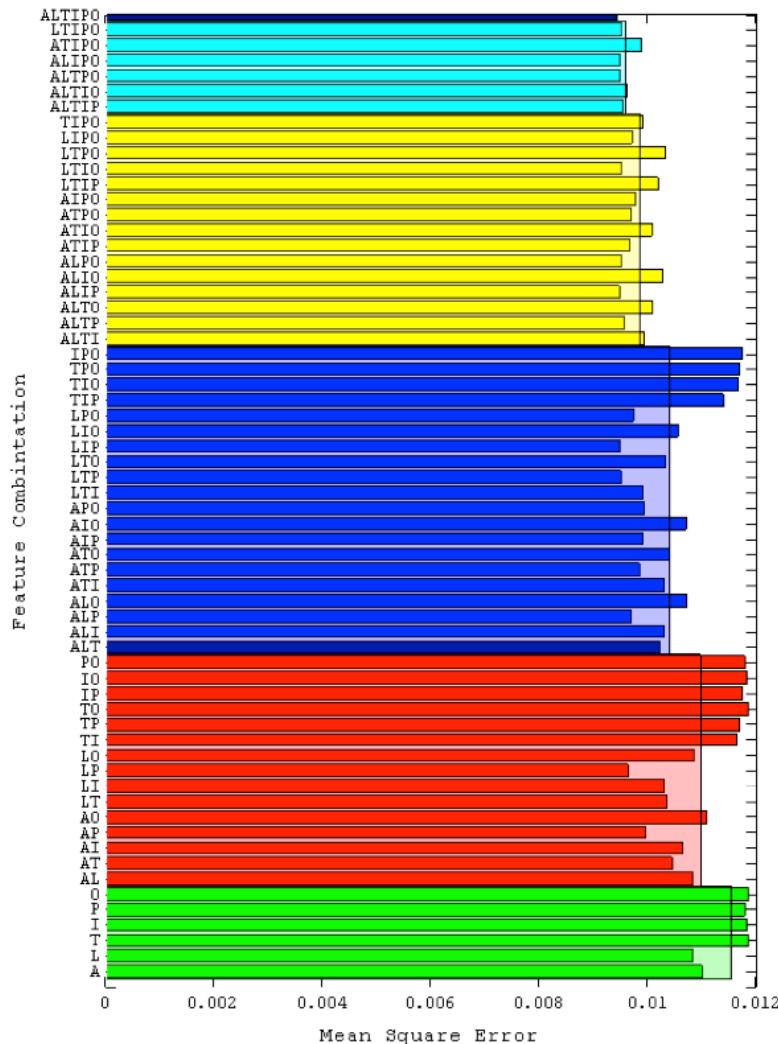
rankings of the other three methods. Unlike the other methods, the objective function of ReliefF is only based on the distances of feature instances without the uncertainty evaluation of data. Thus, the result of ReliefF would be unreliable when dispersion of feature dataset is large.

Table 4-4. Combined selected features for different pavement conditions.

	Area	Length	Texture	Intensity	Location	Orientation
Good	13 (3)	7 (1)	20 (5)	13 (3)	10 (2)	21 (6)
Intermediate	13 (3)	8 (1)	18 (5)	13 (3)	10 (2)	21 (6)
Poor	14 (4)	8 (1)	18 (5)	11 (2)	11 (2)	22 (6)
All	14 (4)	10 (2)	18 (5)	12 (3)	8 (1)	22 (6)

The above rankings based on the relevance only are compared with the rankings with mRMR and FCBF based on both the relevance and redundancy. FCBF selected the length and location as the optimal feature subsets. mRMR selected the length, orientation and location are the top three features. Thus, the orientation was ranked at the second with mRMR, while it was not selected with FCBF. In addition, the result of the orientation with mRMR does not agree with the rankings in Table 4-3. In the last column of Figure 4-10, one can observe that there is no clear distinction between the ranges of crack and no-crack components in x-axis. Finally, from the observation of pavement surface images, one can find that the orientation of pavement crack can be in any direction, particularly when pavement condition is poor with alligator cracks. Therefore, the mRMR result of the orientation would not be reliable. In summary of the above results, the length, location and intensity are selected as the optimal feature subset. This result agrees with the wrapper FSS result by Wu et al. (2014). In Figure 4-14, the classification error of crack and no-crack were significantly reduced when ANN classifier was with the length, location and intensity

features (LIP). The feature combination of the length and position (LP) also produced a low classification error, which agrees with the FCBF result in Table 4-3.



* The y-axis shows the crack features of the area (A), length (L), texture (T), intensity (I), wheel-path location (L), and orientation (O).

Figure 4-14. The mean-square-errors of ANN classifier for different feature subset combinations (Wu et al., 2014).

4.5 Conclusions

This study showed how crack features, extracted using computer-vision techniques can be used to evaluate cracks in flexible pavement in different damage stages. Using the novel computer-vision techniques proposed in this study, crack image components could be accurately extracted from surface images of flexible pavement. Multiple features were extracted from manually selected crack components, which have been commonly used in road pavement literatures. These features contained imagery information that could characterize pavement crack. Therefore, this study was aimed to investigate if the statistics of crack image components identified using currently available computer-vision techniques could provide useful information to characterize crack patterns in different stages of aging flexible pavement. The information of crack patterns could be used as fundamental data to provide justification for rehabilitation policies. Using the computer-vision techniques proposed in this study, a total of 264981 image components were identified from 26 surface images of pavement in good, intermediate, and poor conditions. Then, six different features were extracted from the image components, including the area, length, texture, intensity, location, and orientation. The image components were also manually separated into two classes and labeled as 3090 cracks and 261891 no-cracks. Statistical analyses were conducted to understand statistics of the crack feature data and to find the optimal feature subset by characterizing the relations between the feature instances and the class labels. Based on the statistical characterization results, some conclusions can be made as follows:

- The filter FSS methods used in this study could find the optimal feature sets by considering both relevance and redundancy. Finding the optimal feature subsets are

important for effective statistical characterization by eliminating irrelevant and redundant features.

- The filter FSS results showed that the length, location and intensity were the top-three ranked features as the optimal feature set. That is, these features are the three most important features characterizing image components to be crack for the given images.
- Crack in wheel paths (CW) usually has more weights in crack rating calculation than crack outside wheel paths (CO) (Florida Department of Transportation, 2012b). The filter FSS result supports the validity of having different weighting factors based on the wheel path designation in crack rating calculation since statistically more cracks were detected in wheel paths through image processing.
- The distributions of crack orientation showed that poor pavement had similar percentages of longitudinal and transverse cracks, while good pavement had mostly longitudinal crack. This fact reflects the progress of structural crack caused by traffic loads, such as wheel-path longitudinal crack and alligator crack. Alligator cracking is a major structural distress, which is caused by fatigue damage in flexible pavement with granular and weakly stabilized bases. Alligator crack first appears as parallel longitudinal cracks in wheel paths, and then it progresses into a network of interconnecting (transverse) cracks resembling the skin of an alligator (K. T. Hall, Correa, Carpenter, & Elliot, 2001).
- Crack width is an important factor in crack rating. Using the proposed method in Equation 4-6, the mean of crack width was measured at 3.62 mm for good, 3.98

mm for intermediate, and 4.06 mm for poor condition as pavement is more damaged. The standard deviation was 3.06 mm for good, 3.54 mm for intermediate, and 3.55 mm for poor condition. Florida Department of Transportation (2012) classifies crack into Class-1B crack has a width less than 3.18 mm, Class-2 crack has a width between 3.18 mm and 6.35 mm, and Class-3 crack has a width greater than 6.35 mm. Therefore, the percentages of Class-1B, Class-2 and Class-3 cracks for different pavement conditions can be calculated as shown in Figure 4-15. Good pavement has 65.4% Class-1B cracks, 26.5% Class-2 cracks, and 8.1% Class-3 cracks. Intermediate pavement has 51.4% Class-1B cracks, 39.0% Class-2 cracks, and 9.6% Class-3 cracks. Poor pavement has 51.7% Class-1B cracks, 37.6% Class-2 cracks, and 10.7% Class-3 cracks. These results show that the percentage of Class-3 crack increases as pavement is more damaged. The percentage of Class-1B crack decreases while the percentage of Class-2 crack increases when pavement is in intermediate or poor conditions. In comparison of the intermediate and poor conditions, the percentages of Class-1B and Class-2 cracks are almost the same, while poor pavement has more Class-3 cracks than intermediate pavement.

- The above results validate that the statistical characterization methods associated with the computer-vision techniques used in this study can extract useful information of pavement cracks in different deterioration stages, which can be accurate quantitative information to understand past and current pavement conditions and to justify pavement maintenance policy (Noori et al., 2014).

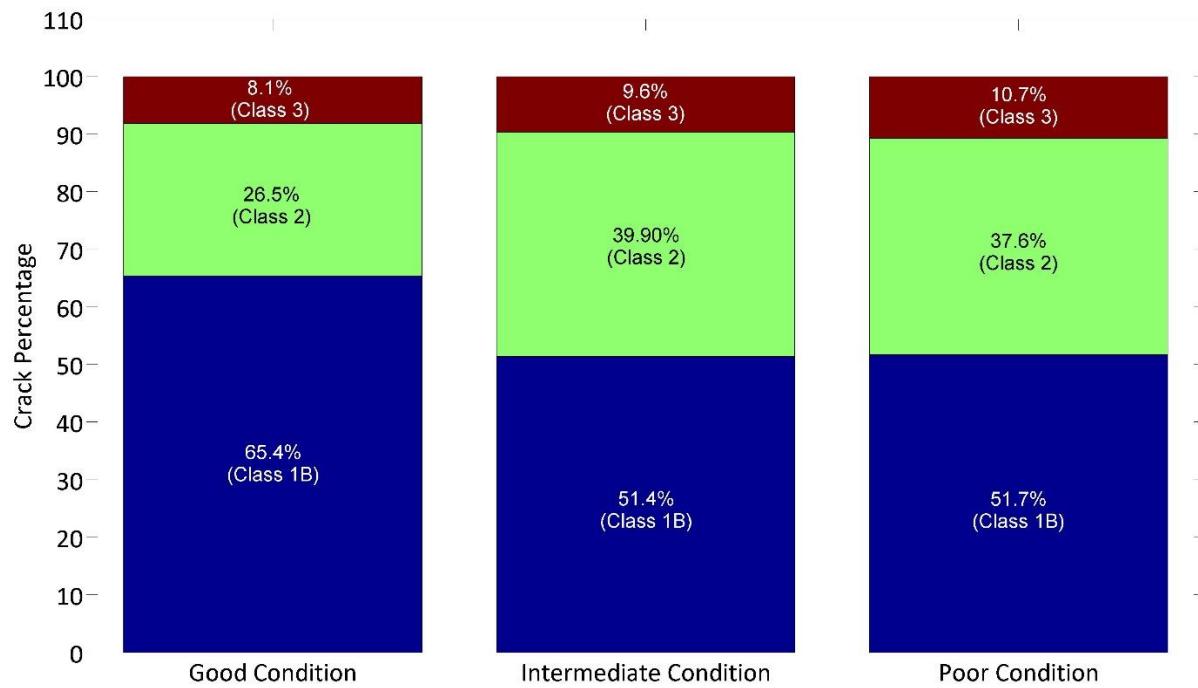


Figure 4-15. Percentages of Class-1B, Class-2, and Class-3 cracks in different stages of pavement deterioration.

CHAPTER 5: COMPARATIVE STUDY OF CLASSIFICATION METHODS FOR IMAGE-BASED PAVEMENT CRACK DETECTION³

5.1 Introduction

Roadways as a main component of infrastructure, play an important socio-economic role by providing transportation for people and commodity. Therefore, long term performance prediction, maintenance and development of roadways is a necessity. In the US, all 50 states have developed pavement management programs that describe required physical measurements for rating pavements serviceability. Correspondingly, collection of distress data is a major task in every pavement management program. Current data collection and evaluation methods are mostly manual or semi-automated. However such methods can prolong the evaluations procedure, cause traffic interruption and impose safety issues especially in high-volume highways. Moreover, the results of manual pavement evaluation methods may vary due to personal judgment.

Image-based techniques are considered as a promising non-destructive method for evaluating pavement surface condition. Especially, for cracking which is a major cause of pavement deterioration and determines the short-term and long-term performance of pavements. However, developing a fully automated image-based pavement crack detection method is challenging. Firstly, cracks should be identified from the random background of pavement images with large variations of texture, roughness, intensity, spots and stains, oil and water spilling, and

³ The contents of this chapter is in preparation to be published as journal paper.

Mokhtari, S., Yun, H.-B., & Wu, L. (2015b). *Comparative Study of classification methods for image-based pavement crack detection*. Unpublished manuscript.
The author contributed to all sections except 5.2.

road markings. Secondly, crack geometry (e.g. length and direction) and pattern (e.g. longitudinal, transverse, block and alligator) can also vary significantly. In this study, computer-vision framework developed by Wu et al. (2014), is employed to detect crack image components by removing the random background of pavement images. Their method is based on morphological bottom-hat transformation which is capable of detecting cracks, selectively based on their size, shape and intensity with relatively small number of parameters.

However, the problem with the results of morphological bottom-hat transformation is that detected cracks are fragmented and have many disjoints along the crack path. To solve this problem, Yun et al. (2015), developed a crack grouping method based on morphological dilation and thinning transforms, so called MorphLink-C. this method can be used with any pixel-level crack-extraction method. The details of computer-vision techniques are out of the scope of this study; however, a brief description of these method are provided in Section 5.2.

Although the image processing techniques can efficiently remove random background from the pavement images, but, the results still include non-crack objects. Machine learning methods are commonly used to classify image components into cracks and non-cracks. Different classification methods have been employed to detect cracks from pavement images including: artificial neural networks (ANN) (Lee & Lee, 2004; Moghadas Nejad & Zakeri, 2011b; Nguyen, Avila, & Begot, 2009; Saar & Talvik, 2010; Siddharth, Ramakrishnan, Krishnamurthy, & Santhi, 2012), adaptive neuro fuzzy inference system (ANFIS) (Bianchini & Bandini, 2010; Terzi, 2013), support vector machine (SVM) (Evdorides, Schlotjes, Henning, & Burrow, 2014; Gavilán et al., 2011; Moussa & Hussain, 2011; Salari & Ouyang, 2012), decision trees (Ho, Chou, & Lin, 2012; Moghadas Nejad & Zakeri, 2011a; Zhou & Wang, 2012), and k-nearest neighbours (Jahanshahi,

Masri, Padgett, & Sukhatme, 2011). However, limited research has been carried out to compare different classification algorithms and discuss their advantages and drawbacks for the image-based pavement crack detection problem. The available studies, often focus on classification performance of the methods, disregarding the information and knowledge that can be inferred from the classification procedure. State-of-the-art classification methods are getting more complex in order to provide superior classification performance (Freitas, Wieser, & Apweiler, 2010). However, it is even more difficult to follow the decision making procedure or interpret the logic of classification for such methods (Hammer, Mokbel, Schleif, & Zhu, 2012). From this points of view, classification techniques can be divided into ‘black-box’ and ‘white-box’ methods.

The objective of this study is to evaluate and compare different classification methods for detecting cracks from pavement images. Although high performance and low error rates are essential for a suitable classifiers but, the information and knowledge that can be inferred from the classification procedure can provide further insight into the cracking procedure. Such information can be used by agencies and decision makers for better policy making and pavement management planning.

To achieve the proposed study objective, a four-step procedure is followed as presented in Figure 5-1. In the first step, gray-scale images of flexible pavement collected with a Multi-Purpose Survey Vehicle (MPSV) by Florida Department of Transportation (FDOT), are analysed using the computer-vision technique developed by Wu et al. (2014). In the second step, the cracks grouping method (MorphLink-C), developed by Yun et al. (2015) is employed to deal with fragmentation of crack images components. A brief description of computer-vision techniques are presented in Section 5.2. In order to prepare the training dataset for classification, 27 sample images of

pavements in different deterioration stages are selected and six features including: area, length, orientation, texture, orientation and wheel path position, that are commonly used in image-based pavement evaluation is extracted for crack and no-crack components (Step 3). Details of data preparation and feature extraction is described in Section 5.3. Different classification algorithms in terms of computational complexity and clarity of procedure (i.e. being white or black box) are selected for evaluation (Step 4) in this study. The selected methods include: Artificial Neural Network (ANN), Decision Tree (DT), k-Nearest Neighbours (kNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS). A comprehensive review of the classification methods along with their application procedure using an example problem is presented in Section 5.4. Finally, selected classification algorithms are applied to the pavement dataset, prepared in Step 2. Application procedure and discussion of the results are presented in Section 5.5 and 5.6, respectively.

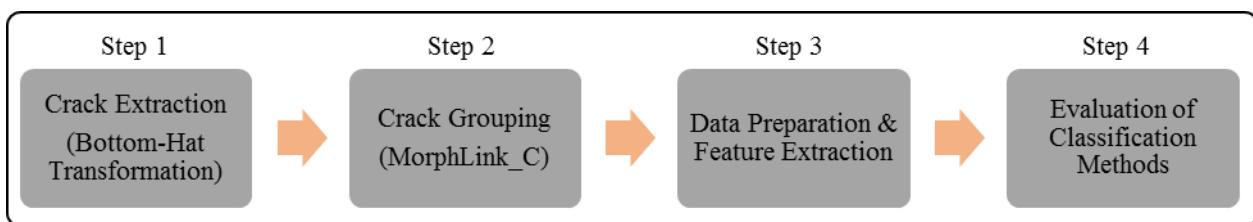


Figure 5-1. General procedure of this chapter.

5.2 Morphological Image Processing and MorphLink-C for Crack Detection

The morphological image processing technique is based on mathematical morphology. In this technique, a binary image, so called structuring element is applied to input image. Different region shapes can be extracted from input image by controlling the shape of structuring element

and combining morphological operators. Dilation and erosion are two basic morphological operators that can be used to manipulate the input image. The dilation “grows” or “thickens” objects in an image, while the erosion “shrinks” or “thins” objects. The common mathematical expressions of dilation and erosion are $\mathbf{O} \oplus S$ and $\mathbf{O} \ominus S$, respectively. Where \mathbf{O} , is the grey-scale input image and S is the structuring element.

Wu et al. (2014) used morphological bottom-hat transformation to detect dark cracks in pavement images.

$$\mathbf{L} = \max[(\mathbf{O} \circ S_{l\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}) \bullet S_{l\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}, \mathbf{O}] - \mathbf{O} \quad (5-1)$$

where \mathbf{L} is a gray-scale outputs image; \mathbf{O} is the gray-scale input pavement image; $S_{l\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}}$ is a line-shape structuring element with the length of l pixels, which rotates at 0° , 45° , 90° and 135° to detect cracks in arbitrary orientations; the subtraction, ‘ $-$ ’ is a logical subtraction defined by $X - Y = X \cap NOT Y$; \circ and \bullet are the opening and closing transformations that are defined using basic dilation and erosion operators as:

$$\mathbf{O} \circ S = (\mathbf{O} \ominus S) \oplus S \quad (5-2)$$

$$\mathbf{O} \bullet S = (\mathbf{O} \oplus S) \ominus S \quad (5-3)$$

Wu et al. (2014), demonstrated that bottom-hat transformation can effectively remove the random background of pavement images with limited number of user-defined parameters. In their study, the optimum length of the structuring element is found to be 98.0 mm (3.86 in). Application of the bottom-hat transformation to detect cracks from a sample pavement image is presented in Figure 5-2.

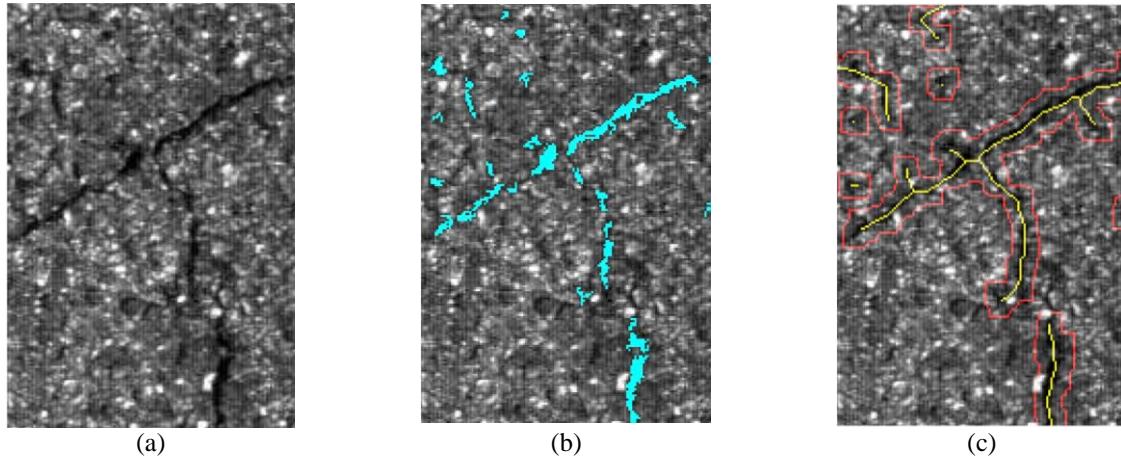


Figure 5-2. Application of bottom-hat transformation and MorphLink-C methods for detecting cracks from pavement images.

However, the bottom-hat transform results (Figure 5-2b) are fragmented and cracks have several disjoints along their path. MorphLink-C is proposed by Yun et al. (2015) as a segmentation technique to deal with the crack fragmentation problem. In this method, crack boundaries are extended using morphological dilation and crack fragments are connected using the extended boundaries. Mathematical expression of the procedure is presented in Equation 5-4.

$$\mathbf{D} = \mathbf{L} \oplus \mathbf{S}_D \quad (5-4)$$

where \mathbf{L} is the output image of the bottom-hat transform; S_D is a square structuring element with the size of $d \times d$ pixels; and \oplus is the morphological dilation transformation. Yun et al. (2015), recommended $d = 10$ pixels for the size of square structuring element, which is equivalent to 9.8 mm (0.39 in). Extended boundaries of crack are presented in red in Figure 5-2c. Based on the results, crack fragments are grouped into a single boundary however, to demonstrate the crack path and have a crack-line abstraction morphological thinning transform can be employed as:

$$\mathbf{T} = \mathbf{D} - \text{hit-and-miss}(\mathbf{D}, S_T) \quad (5-5)$$

where \mathbf{D} is the binary image after the dilation transform; ‘hit-and-miss’ is the morphological hit-and-miss transformation; and S_T is the structuring element for skeletonization by the morphological thinning. Thinning results are shown in Figure 5-2c in yellow.

5.3 Database Preparation and Feature Extraction

For this study a total of 26 pavement images including 11 images in good, 7 images in intermediate, and 8 images in poor pavement condition, have been analyzed using the suggested morphological framework. Sample Images are selected based on the randomness of background and severity of cracking to make sure that the dataset can represent the characteristics of cracks in different deterioration stages of pavement. More pavement images are selected from good conditions since these images usually have relatively smaller number of cracks.

As mentioned earlier, the results of computer-vision techniques includes crack and non-crack image components; therefore, to prepare a dataset for training and validation of the classifiers, crack and non-crack components are manually separated. The number of crack and non-crack components along with the equivalent length of detected objects are presented in Table 5-1.

Table 5-1. The numbers and length of crack and non-crack image components in training dataset.

Image #	Crack objects (#)	Crack objects (mm)	Non-Crack objects (#)	Non-Crack objects (mm)
Good (11)	1	3	273	7516
	2	13	905	7332
	3	60	7718	8438
	4	70	10876	9068
	5	15	2984	10976
	6	50	4625	10815
	7	94	6280	10591
	8	42	7244	10758
	9	57	8888	10372
	10	38	2586	10070
	11	0	0	10008
Intermediate (7)	12	141	17584	9793
	13	151	20726	10245
	14	116	23428	10361
	15	2	1214	12182
	16	2	203	12033
	17	310	40253	10627
	18	144	29242	10978
Poor (8)	19	212	22422	10493
	20	183	26962	10047
	21	213	33843	9630
	22	252	27005	9904
	23	219	38955	10765
	24	239	34580	10833
	25	320	55307	9899
	26	144	20203	8157
All	Sum	3090	444306	261891
	Mean	118.8	17088.7	10072.8
	Stdv	99.6	15288.2	1089.9
				3470796
				133492.1
				1773

It should be emphasized that there is a significant difference between the number of crack and non-crack components (3090 crack components and 261891 non-cracks components). Imbalanced datasets may negatively impact the performance of the classification methods. Partitioning the dataset for training, validation and testing the performance of classifiers will be explained in Section 5.5.

Six features including: area, length, texture, intensity, location and orientation that are commonly used in image-based pavement rating were extracted in this study to represent the characteristics of crack and non-crack image components for the classification algorithms. Mokhtari, Yun, & Wu (2015a), conducted a comprehensive statistical study to evaluate these features for road rating applications. A brief description of these features is presented here.

Area: This feature measures the area of the crack and no-crack image components. Area can be calculated by counting the number of pixels within the same dilation boundary as:

$$Area_D = \sum_{L=1}^{Number\ of\ L\ in\ D} (Area_L)_D \quad (5-6)$$

where D is the segment index determined after the dilation transformation; L is the index of the fragmented pixel clusters after the top-hat transformation, which are within the same dilation boundary of D ; and $(Area_L)_D$ is the area of the L -th image component in the D -th segment.

Length: length of crack and non-crack components can be measure by counting the number of pixels of thinned lines after the thinning transform (T).

$$Length_D = \text{Number of pixels on the thinned line in } D \quad (5-7)$$

Orientation: the orientation feature measures the angle in degree, between the horizontal direction and the major axis of the ellipse that encloses the dilation boundary after the dilation transform (D).

Texture: This feature represents the roughness of the pavement surface using Co-occurrence matrix. The main logic behind defining a texture index is that rough parts of the pavement have higher potential for developing cracks. Co-occurrence matrix is a measure of intensity and the offset between the pixels of image components. A complete description of texture index is provided by Mokhtari et al. (2015a). In this research, inputs images after MorphLink-C are divided into 200×200 image tiles. Co-matrix of each tile is calculated and normalized by the total number of pixels in that tile.

Intensity: The intensity feature measures the weighted average pixel intensity of the crack and non-crack image components as:

$$Intensity_D = \frac{\sum_{L=1}^{Number\ of\ L\ in\ D} (Intensity_L)_D \times (Area_L)_D}{\sum_{L=1}^{Number\ of\ L\ in\ D} (Area_L)_D} \quad (5-8)$$

where $(Intensity_L)_D$ is the averaged pixel intensity of the L -th image component in the D -th segment.

Location: The location feature determines whether the crack and non-crack image component is located inside or outside of the wheel path. The wheel path designation is defined by some pavement management agencies (Florida Department of Transportation, 2012) to consider the higher cracking potential of wheel path which is directly affected by moving traffic.

FDOT (2012a) is divided the pavement surface into the inside of the wheel path (CW) and the outside of the wheel path (CO) as shown in Figure 5-3. The image components within the CO are assigned feature value of two, one is assigned to the components in CW, and zero for the components outsides of the lane markers.

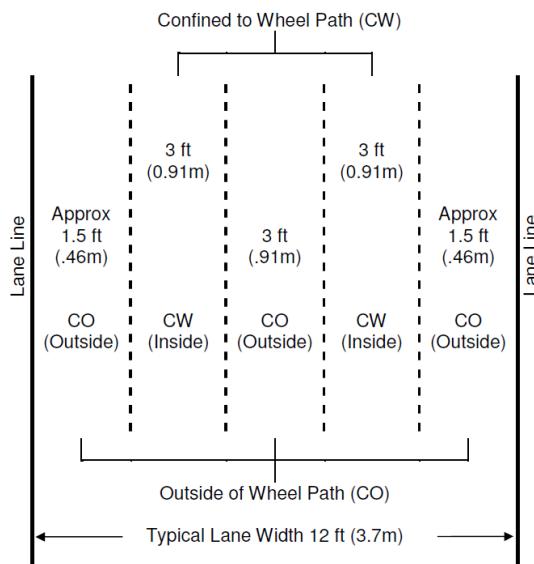


Figure 5-3. Wheel path designation defined by Florida Department of Transportation.

5.4 Review of the Selected Classification Methods

Four classification algorithms namely, Artificial Neural Network (ANN), Decision Tree (DT), k-Nearest Neighbours (kNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) is considered for detection cracks from pavement images. In order to demonstrate the classification procedure for each method, a sample problem with two features and four classes of data is

employed in this section. Each data point has two characteristics, quantified in feature values and the classifier uses these values to assign one of the four possible class labels to each data point. However, generating enough sample datasets for training, validation and testing each classification method, requires the statistical properties of data in each class to be constant; otherwise, the difference between classifiers' performance over training dataset and test dataset will be significant, since the data is not following a constant pattern and is completely random. For this purpose, data samples in each class are generated from a predefined bivariate distribution function. In this way, arbitrary number of data samples with constant statistical characteristics can be generated using a Monte-Carlo simulation for each class of data. The first class of data follows a bivariate normal distribution with mean vector of $\mu = [4 \ 4]$ and variance matrix of $\sigma = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$. Off-diagonal entries of the variance matrix are not zero; therefore, features in this class are correlated and distribution is not symmetrical. Second class follows a symmetrical normal distribution with mean vector of $\mu = [4 \ 6.5]$ and variance matrix of $\sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The log-normal distribution of the third class has a mean of $\mu_{log} = [2.2 \ 1]$ and variance of $\sigma_{log} = \begin{bmatrix} 0.2 & 0.25 \\ 0.25 & .2 \end{bmatrix}$ and the samples of the forth class are uniformly distributed around $\mu = [11 \ 6.5]$ with $\sigma = [5.33 \ 1.33]$. Finally, Class labels 1 through 4 are assigned to the classes as target values for training classifiers. The summary of statistical properties of sample data is presented in Table 5-2.

Table 5-2. Statistical properties of all classes of data for the example problem.

Class Labels	Distribution	Statistical Parameters	# of Samples/Dataset	Target Value
Class 1	Normal	$\mu = [4 \ 4]$ $\sigma = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$	250	1
Class 2	Normal	$\mu = [4 \ 6.5]$ $\sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	250	2
Class 3	Log-Normal	$\mu_{log} = [2.2 \ 1]$ $\sigma_{log} = \begin{bmatrix} 0.2 & 0.25 \\ 0.25 & 2 \end{bmatrix}$	250	3
Class 4	Uniform	$[11 \pm 4 \ 6.5 \pm 2]$	250	4

In this study, 100 datasets are generated for training, validation and testing the classification methods. Each dataset is consist of 1000 data points equally divided among 4 possible classes (250 data points in each class) adding up to a total of 100,000 data points. 60 datasets are used for training to tune and optimize the adaptive parameters of the classifiers. 20 datasets are used for validation to evaluate the generalization ability of the classifier and prevent the over-fitting problem. Finally, the remaining 20 datasets are used as test data which was not used during the training. A sample dataset (1000 data points) is presented in Figure 5-4. Mathematical background of the classifiers along with their application procedure for the example dataset is presented in subsequent sections.

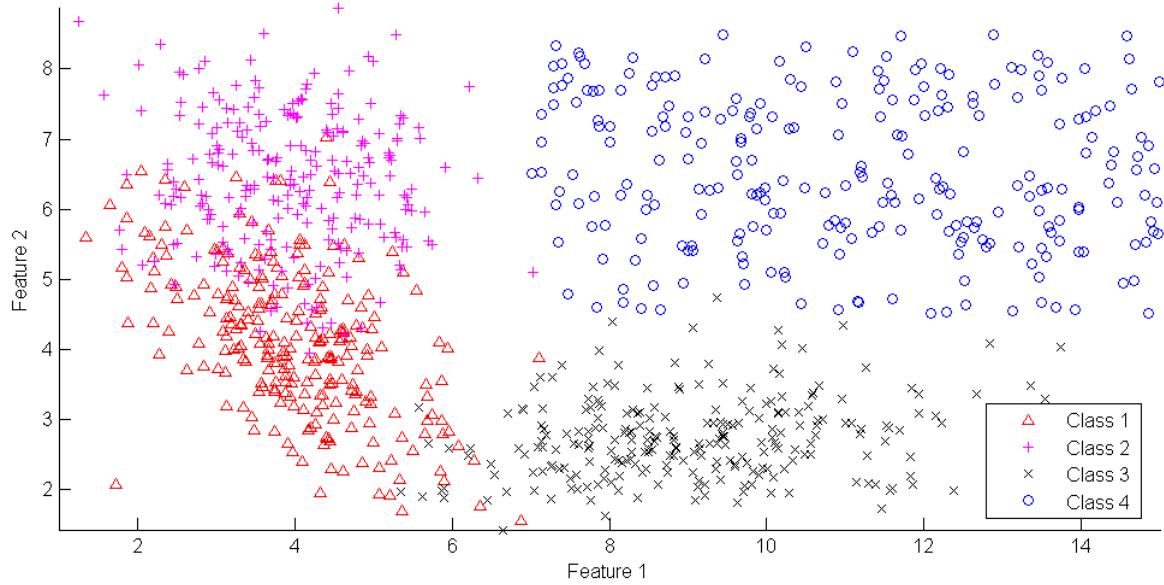


Figure 5-4. One of the randomly generated datasets in the example problem.

5.4.1 Neural Networks

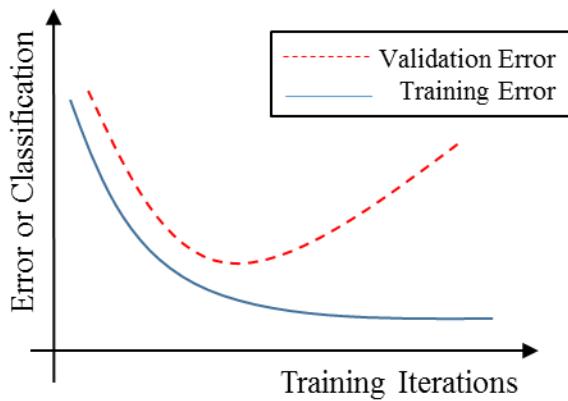
Artificial Neural Network (ANN) is a data driven technique that has been widely studied and applied to various engineering applications. Neural networks can be applied to classification problems without any assumption about the underlying statistics and they can estimate any function with arbitrary accuracy. (Hornik, Stinchcombe, & White, 1989; Hornik, 1991). Among different types of neural networks (e.g. radial basis function networks, recurrent neural networks, modular neural networks and etc.), feed-forward multilayer networks have been considered in this study not only because they are the most widely used neural network classifiers but also due to considerable study that has been carried out on these networks. Learning and generalization are the most important concepts in neural networks (Amirikian & Nishimura, 1994; Baum & Haussler,

1989; Sietsma & Dow, 1991; Mirakhorli, Farahani, & Ramtin, 2009). Learning is the process of tuning the network based on training data while generalization is the predictive ability of classifier for other datasets. Network size and model along with selected features and sample size are the main parameters that can affect the learning and generalization of a network (Zhang, 2000). Empirical studies show that sample size effect on learning is much more than number of hidden nodes (Hung, Hu, Shanker, & Patuwo, 1996; Richard & Lippmann, 1991) and changing the training data can affect the efficiency of networks, drastically (Breiman, 1996). Imbalanced training datasets can also reduce the predictive performance of neural networks (Wilson & Sharda, 1994). In an imbalanced training dataset, number of samples from classes is significantly different.

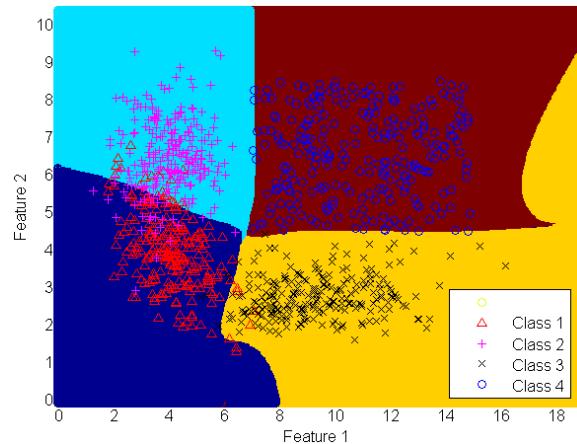
Neural networks are powerful approximators; correspondingly, over-fitting is common a problem of ANN classifiers. An over-fitted network has very low error for training data (bias) but show poor results when applied to the test sample (variance). Numerous studies have been carried out to prevent the over-fitting problem and the proposed methods include cross validation, training with penalty terms, weight decay and node pruning method (Zhang, 2000; Ramtin & Pazour, 2014).

In application of ANN classifier to the sample problem a feed-forward network with 5 layers (1 input layer, 3 hidden layers and 1 output layer), and 10 nodes in each hidden layer was employed. The network is oversize, considering the number of features and samples, to avoid required analysis for finding the optimum network size. Tangent-sigmoid transfer functions in hidden nodes map the inputs into the output layer where a linear transfer functions magnifies the results to any desired range. A Mean Square Error (MSE) function is employed as the cost function to evaluate the difference between network outputs and classification targets. Finally, Levenberg-

Marquardt method is used to optimize the weights and biases of the network to minimize the training error. In early stages, the error for training and validation datasets decrease. When the networks over-fits the training data the validation error starts to increase while training error continues to decrease. A schematic trends of training and validation errors are presented in Figure 5-5a.



(a) Trend of classification error



(b) Classification results using ANN

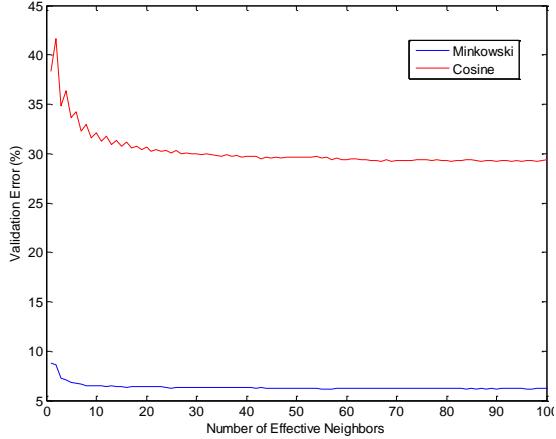
Figure 5-5. Classification of example problem using artificial neural network classifier.

To prevent the over-fitting problem, the network is trained for 60 percent of the data samples (60,000 data points) and in each iteration, the validation error is calculated for a different set of 20,000 data points. The training is stopped, if the validation error was not decreasing for 10 consecutive iterations. The results are presented in Figure 5-6b. Statistical error analysis is performed for the classification results of the test set (20,000 data points) and the results are presented in Table 5-3.

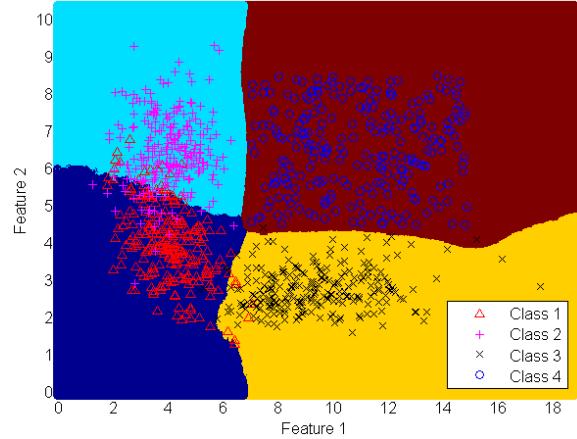
5.4.2 K-Nearest Neighbors

k-Nearest Neighbor (kNN) method is a simple yet widely used classification method. In this method, each data point will be assigned the same class label as its k-nearest neighboring points. It is important to normalize the features so that the difference in ranges won't affect the results; otherwise, predictions are more biased toward the feature with largest range of variation. One major disadvantage of using kNN is that the entire training dataset should be stored in the memory of computer; therefore, classification of large problems may be computationally expensive.

Several modifications have been proposed to improve the performance of kNN classifiers including distance-based weighting for the nearest neighbors. These weighting methods include Minkowski distance metric and Earth Mover Distance (EMD) (Cunningham & Delany, 2007). However, calculating the weight matrixes for each point can increase the computation time. Alternative searching strategies such as, Case-Retrieval Nets, Footprint-Based Retrieval, Fish and Shrink and Cover Trees have also been proposed to substitute the exhaustive search, used in standard kNN. For the standard kNN method, type of distance measure and the number of effective neighboring points are the most important factors that can influence the classification performance. To analyze the effect of these parameters, classification performance with different number of effective neighbors and two distance measures is considered. For this purpose, the classifier is trained for 60% percent of the sample data and then, the validation error is calculated for 20% of samples. Using an exhaustive search, the validation error for different distance measures and number of neighbors is calculated and the results are presented in figure 5-6a.



(a) Error of validation



(b) Classification results using kNN

Figure 5-6. Classification of sample problem using k-Nearest Neighbor method.

It can be observed that in early stages of analysis, classifier's performance improves as the number of neighboring points increases; however, no significant trend can be observed for classifiers that consider more than 20 neighboring points. The analysis also compares the performance of kNN classifiers for different distance measures. The results depict that, Minkowski distance measures provided lower validation error compared to cosine measure. This result was expected since cosine measure is usually more suitable when the angular proximity of data should be considered for classifying the samples.

Considering the above mentioned analysis, a 30-nearest neighbor classifier with Minkowski distance measure is trained and validated for 60 and 20 percent of the sample data, respectively. Larger number of effective neighbors will increase the computation time while a smaller value will cause over-fitting problem. Classification results of feature space using the kNN classifier is presented in Figure 5-6b.

5.4.3 Decision Trees

Decision trees perform the classification by repetitive division of features space. A decision tree is formed by a set of test nodes (internals) and decision nodes (leaves). Each test node divides the feature space into two or more subspaces using a splitting rule (function) and at the bottom of the tree, decision nodes will assign output values to the subspaces. In most studies, splitting rules are univariate discrete functions. In simple words, they use a thresholds values for a single feature to divide the feature space. Various types of univariate splitting measures have been introduced among which, Impurity Based Criteria, Normalized Impurity Based Criteria, Distance Measure, Binary Criteria, and Towing Criteria are the most widely studied (Rokach & Maimon, 2005). Although some splitting rules may outperform others for specific problems, but it has been reported that in many cases, the type of rules will not have a significant effect on performance of decision trees (Rokach & Maimon, 2005). Division of the features space will continue until no more splitting is possible or a stopping criterion is met. Therefore, stopping criteria play an important role in controlling the complexity and performance of decision tress (Breiman, Friedman, Stone, & Olshen, 1984). Choosing stringent stopping criteria will cause very small and under-fitted trees while loose criteria will cause large and over-fitted classifiers. Breiman et al. (1984), proposed a methodology, so called pruning techniques, to control the complexity of decision trees. In these methods, a loose stopping criterion will be selected first, so that tree completely over-fits the training dataset. In the next step, tree branches will be pruned considering their effect on validation error of the tree (generalization). The result is a smaller (less complicated) tree and all the branches have acceptable contribution to generalization accuracy. Widely used pruning techniques include, Cost-Complexity Pruning, Reduced-Error Pruning, Minimum-Error

Pruning (MEP), Pessimistic Pruning, Error-Based Pruning (EBP), Optimal Pruning, and Minimum Description Length Pruning. However, several studies suggested that there is no optimal pruning technique and selection of the method is case-dependent. In this study pruning is applied by controlling the minimum number of data samples per leafs. Increasing the minimum number of samples in each leaf forces the tree to include more data samples in each feature subspace and the tree depth will decrease. Error analysis of decision trees with different number of samples per leaf is presented in figure 5-7.

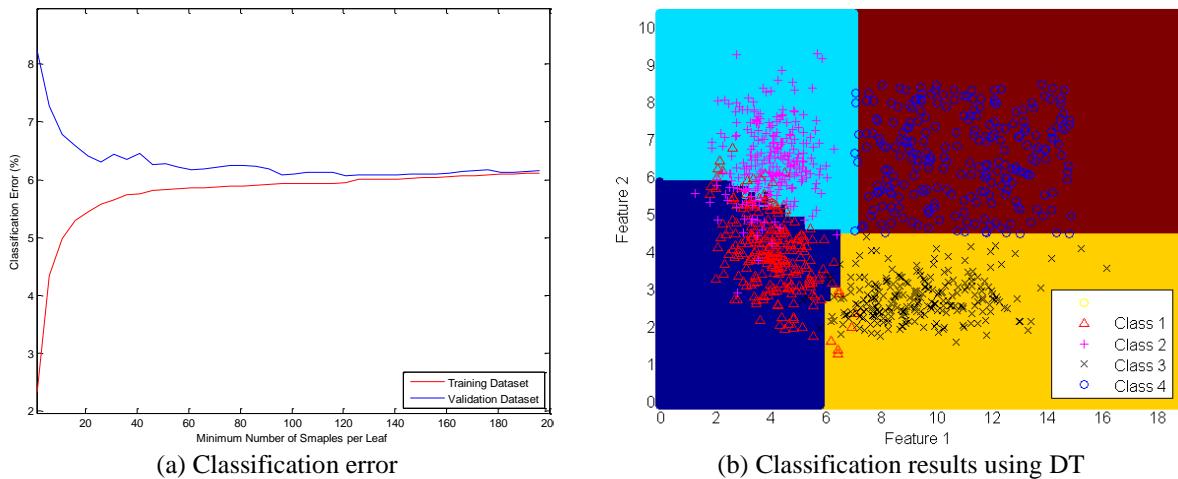


Figure 5-7. Classification of sample problem using decision tree method.

At the beginning of the analysis, the tree with at least one sample per leaf has the lowest error for training data because the tree can be deep enough to over-fit the training data, completely. However, the error of validation is at its peak due to low generalization capability. For higher number of samples per leaf (the depth of tree decreases), the training error increases but the generalization of classifier improves. The trend continues until it reaches a steady state at 120 samples per leaf. Considering these results, a decision tree will have a minimum of 120 samples per leaf

is trained and applied to the samples problem. Classification results are presented in figure 5-7b.

Error of classification is quantified for the test data and the results are presented in Table 5-3.

5.4.4 Adaptive Neuro-Fuzzy Inference System

Fuzzy inference systems use a set of if-then rules to simulate human reasoning in vaguely defined feature space (Jang, 1993; Rafinejad, Ramtin, & Arabani, 2009). However, tuning the membership functions of fuzzy variables can be challenging. In neuro-fuzzy inference systems, artificial neural network is employed as a learning tool to adjust membership functions and minimize the classification error. Fusion of fuzzy inference systems and neural networks can provide better results than each individual technique since these methods can compensate each other's shortcomings. Unlike neural networks, fuzzy inference systems are able to handle inherent uncertainties and vagueness in input data. Moreover, they are easy to interpret since they represent the logical procedure of decision making through their if-then rules but neural networks do not provide knowledge about their internal workings. On the other hand, neural networks have better learning and generalization abilities. A fuzzy inference system consists of five fundamental components: fuzzification block, rule base, data base, decision-making unit and defuzzification block. The rule base includes all the fuzzy if-then rules and database contains the membership function of features. Together, rule base and the database are usually referred to as the knowledge base. In decision-making unit the membership degree of variables are combined through fuzzy operators to calculate a weight for each rule. Proper consequences of the rules can be evaluated using these weights. Aggregation of the consequences is performed in defuzzification block which provides a crisp output. Fuzzy systems and neural networks can be combined into three different

neuro-fuzzy classes (Vieira, Dias, & Mota, 2004). In cooperative and concurrent neuro-fuzzy systems, neural networks are employed in pre-processing or post-processing phases to prepare data or interpret the outcomes of fuzzy system components. However, in a hybrid neuro-fuzzy system the first hidden layer of the neural networks performs the fuzzification of input variables. In the second layer, fuzzy operators are applied to calculate the weight of rules. The third layer normalizes the weights and in forth layer, consequences of rules are determined. Finally, defuzzification in final layer provides the crisp results.

Application of ANFIS method is very similar to ANN. For this method, 60 percent of the sample data is considered for training and in each iteration, the FIS is validated for 20 percent of the data. To prevent the FIS from over-fitting the data, the procedure is stopped when the validation error reaches a steady state or starts to increase. The results of classification using ANFIS method are presented in Figure 5-8.

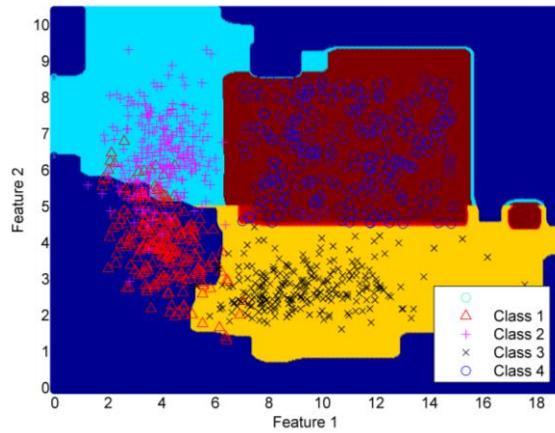


Figure 5-8. Classification results using ANFIS method.

Two trapezoidal fuzzy membership function is used to describe the range of each feature.

Quantified error of each classification method for the test data (20,000 data points) are presented in Table 5-3.

Table 5-3. Error quantification of all classification algorithms for the example problem.

Classification Method	Data Class	False Alarms (%)	Missed Cracks (%)	Successful Classification (%)
Artificial Neural Networks	Class 1	11.18	14.68	85.32
	Class 2	16.18	9.92	90.08
	Class 3	1.99	5.30	94.70
	Class 4	1.26	0.71	99.29
	Overall	7.6525	7.6525	92.3475
k-Nearest Neighbors	Class 1	21.73	7.91	92.09
	Class 2	6.53	19.18	80.82
	Class 3	2.40	4.30	95.70
	Class 4	1.74	1.01	98.99
	Overall	8.10	8.10	91.90
Decision Tree	Class 1	16.38	13.90	86.10
	Class 2	12.29	13.66	86.34
	Class 3	3.31	3.48	96.52
	Class 4	0.74	1.67	98.33
	Overall	8.18	8.18	91.82
Adaptive Neuro-Fuzzy Inference System	Class 1	12.99	13.37	86.63
	Class 2	17.54	13.55	86.45
	Class 3	2.02	6.73	93.27
	Class 4	1.12	0.0	99.98
	Overall	8.41	8.41	91.58

5.5 Pavement Crack Detection

5.5.1 Application of Classification Methods to Pavement Data

In this section, aforementioned classification methods are applied to the data that was extracted from pavement images. Each classifier is trained, validated and tested using the same procedure, explained using the sample problem. However, there are significant differences

between the datasets. The example problem has equal number of samples from each class of data while in the pavement data, there is a significant difference between the number of cracks and non-crack objects. Moreover, each features in the sample problem has predefined and constant statistics while the statistical distribution of features for the pavement data is unknown. Therefore, random partitioning of the dataset might cause several problems, including: 1) training dataset might contain very small number of crack data because the partitioning is random and number of crack samples are far less than the number of non-cracks. In such case, there is not enough information about the crack data for the classifier to learn the difference between the class labels; 2) the randomly selected training dataset might not reflect the characteristics of crack and non-crack objects. In other words statistical properties of the training and testing datasets might be very different. In this case, the classifier will yield poor results because it is trained for a very different data.

In order to avoid the former condition, the crack and non-crack datasets are partitioned, separately. In simple words, 60 percent of both datasets (60% of crack data and 60% of non-crack data) are combined and allocated for training, 20 percent for validation and 20 percent for testing. Using this procedure, that number of crack and non-crack objects for training, validation and testing stay the same. To address the second issue, this random partitioning method is repeated 10 times and each classifier is trained, validated and tested (10 times) using each dataset. The results of all iterations are then, compared to make sure that all the results are within an acceptable range and there is no significant variation in the performance of classifiers over different randomly generated datasets. Explanation of the training procedures is provided using the example problem;

therefore, only a brief description of the application procedure for pavement data is provided in this section.

5.5.2 Artificial Neural Networks

A feed-forward network with 7 layers (1 input layer, 5 hidden layers and 1 output layer) is used to classify the pavement data into crack and non-crack objects. The network input is a vector of 6 features, extracted from each object and the output should be 1 for crack objects and 0 for non-cracks. Each hidden layer contains 10 nodes with tangent-sigmoid transfer function. The transfer function of the output layer is linear to magnify the results to the target values (one for cracks and zero for non-cracks). The performance is measured using the Mean Square Error (MSE) of network outputs. This cost function measures the difference between network outputs and classification targets. Levenberg-Marquardt optimization technique is adopted to adjust the weights and biases of the network in order to minimize MSE over the training dataset. Finally, the stopping criteria is defined using the validation error. In simple words, validation error is calculated in each iteration of training and the procedure ends when the validation error starts to increase or remains unchanged for 10 consecutive iterations.

It should be emphasized that the target values of training are 1 and 0 for cracks and non-cracks, respectively. However, the output layer of the networks have linear transfer function; correspondingly, network outputs are not necessarily 0 or 1. This problem could be avoided by selecting a step function instead of linear transfer function for the output layer of the networks. In this study however, a threshold value between 0 and 1 is defined so that output values more or less than the threshold can be considered as cracks or non-cracks, respectively. This threshold value

can be used for tuning the results and controlling classifier's type I and type II errors (false alarms and missed cracks, respectively).

Performance of the neural network classifier over training, validation and test datasets are presented in Figure 5-9. Horizontal axis depicts different threshold values for network outputs and the box plots for each threshold values illustrates the successfully classified samples during the entire 10 iterations of the random partitioning (10 iterations).

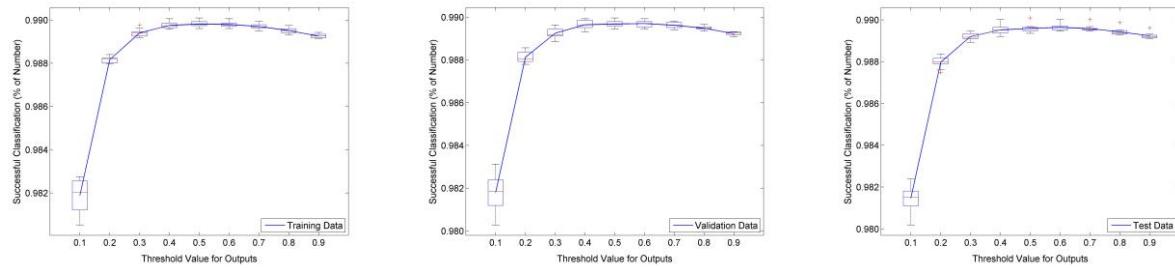


Figure 5-9. Successful classification rates of ANN for training, validation and test datasets.

As mentioned earlier, random partitioning of the dataset is repeated 10 times to ensure that the training dataset can represent the entire characteristics of the crack and non-crack objects. Considering the results (Figure 5-9), no significant variation in performance of classifiers can be observed.

In order to select the threshold value for the outputs of the network, the percentages of type I and type II errors need to be considered. For this purpose, total length of successfully detected cracks along with and the length of missed cracks (Type II error) and falsely detected cracks (Type I error) are calculated for test dataset. The results are presented in Figure 5-10.

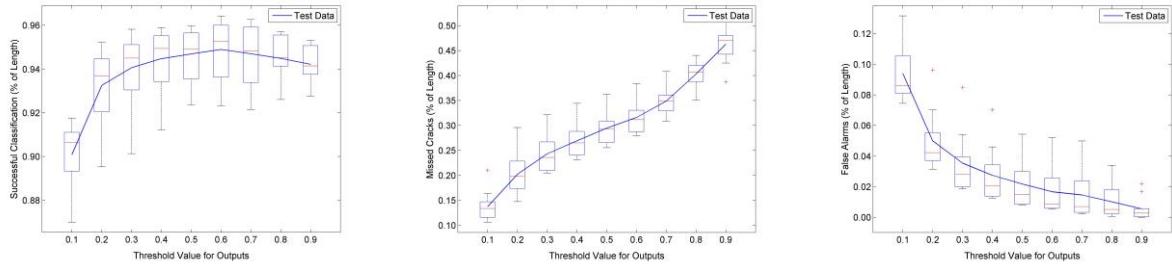


Figure 5-10. Successful classification, missed cracks and false alarms of ANN for test dataset.

Considering the results (Figure 5-10), a threshold value of 0.1 is selected for this application. Correspondingly, output values, less than 0.1 are identified as non-cracks and samples with an output larger than 0.1 are cracks.

5.5.3 Decision Trees

Decision tree classifier was also trained using the same features and data partitioning method. As mentioned earlier, the depth of decision tree has significant impact on generalization ability and performance of the tree. For the example problem, the depth of tree was controlled by increasing the minimum number of data samples per leafs. The same procedure is followed for pavement data and the error of classification for training, validation and test data is evaluated to select the optimum number of samples per leaf (correspondingly, depth of tree). The percentage of successfully classified sample for training, validation and test datasets are presented in Figure 5-11.

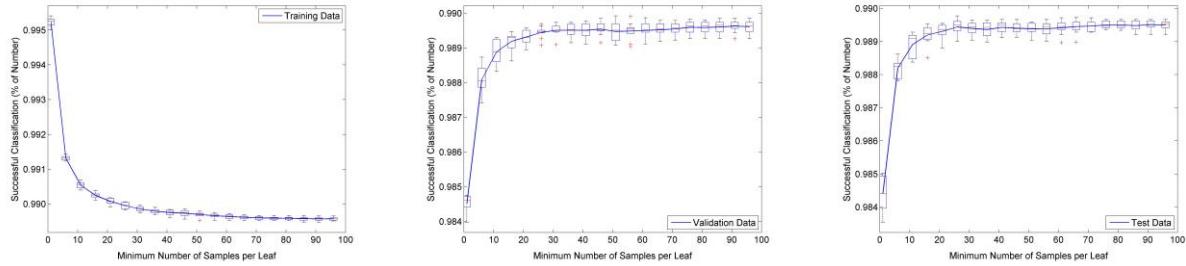


Figure 5-11. Successful classification rates of DT for training, validation and test datasets.

Considering the results, decision tree with 30 sample per leaf is selected for pavement crack detection. Decision trees with less than 30 samples per leaf seem to over-fit the training dataset since they have smaller training error but the validation error is relatively high. It should be noted that the random partitioning does not have significant effect on the classifier's performance since the variations of performance are not significant.

In order to quantify the error of classification, the length of successfully classified objects, missed cracks and falsely detected objects are calculated and the percentages of type I and type II error (in terms of length of objects) are presented in Figure 5-12.

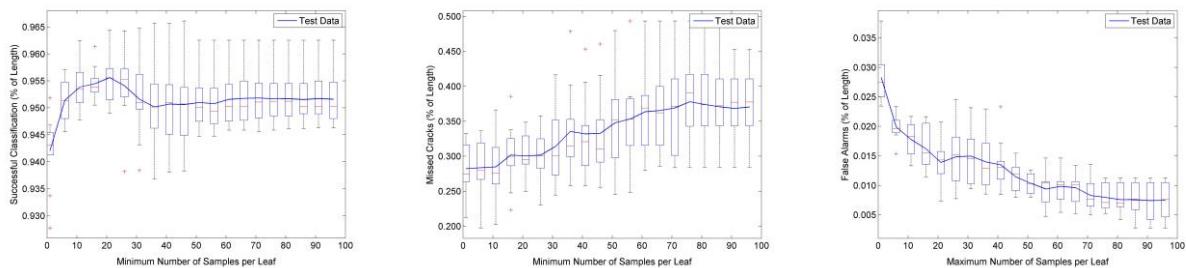


Figure 5-12. Successful classification, missed cracks and false alarms of DT for test dataset.

5.5.4 K-Nearest Neighbors

As mentioned earlier, K-nearest neighbors method follows a relatively simple logic for classifying the objects. For the pavement dataset a nearest neighbor model with Minkowski distance measure is considered. However, a sensitivity analysis is required to determine the optimum number of nearest neighbors. Figure 5-13 presents the percentage of successfully classified samples using kNN classifiers with different number of effective nearest neighbors. The analysis is carried out using the aforementioned features and data partitioning method.

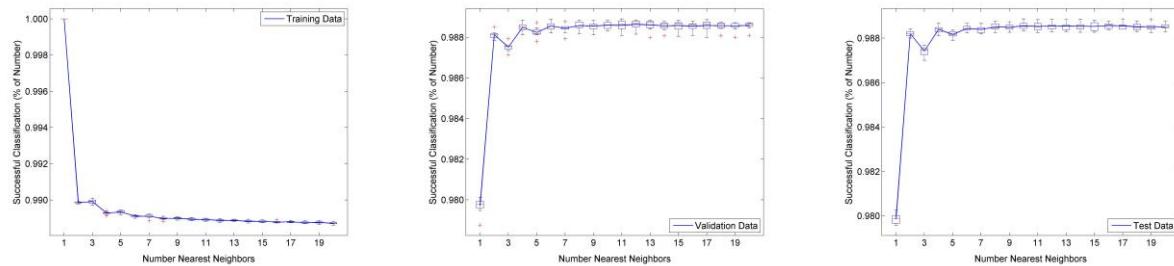


Figure 5-13. Successful classification rates of kNN for training, validation and test datasets.

Considering the results, a 30-nearest neighbor classifier is selected for pavement crack detection. The error of classification using this type of classifier is quantified in terms of length of cracks. The results are presented in Figure 5-14.

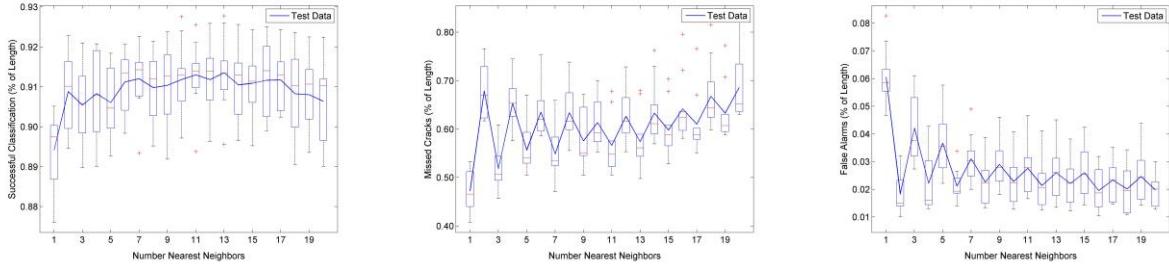


Figure 5-14. Successful classification, missed cracks and false alarms of kNN for test dataset.

5.5.5 Adaptive Neuro-Fuzzy Inference System

The type and number of membership functions can have significant effect on performance of ANFIS classifier. To select the type of membership functions, statistical distributions of features, studied by Mokhtari et al. (2015a) is considered and trapezoidal, Gaussian and generalized bell membership functions are studies for the pavement dataset. Generalized bell membership function provides faster convergence and better results compared to the other membership function and is selected for the pavement crack detection problem.

Classification performance increased marginally (less than 3 percent) when the number of membership functions for each feature increased from two to four and six while the computation time increased drastically. Therefore, two generalized bell membership function is assigned to each feature. The application procedure is similar to ANN classifier and the training is continued until error of validation reaches a steady-state or starts to increase. It should be noted that output of ANFIS is continuous and a threshold value should be selected to divide the outputs into crack and non-cracks. Similar procedure is conducted for ANN classifier. The classification results of ANFIS method for different threshold values is presented in figure 5-15.

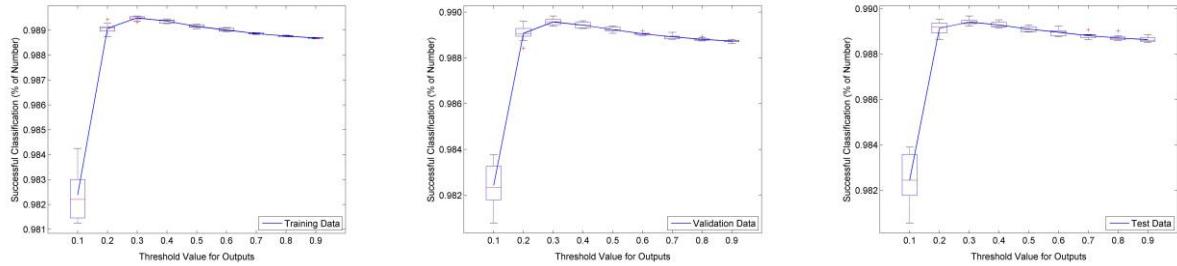


Figure 5-15. Successful classification rates of ANFIS for training, validation and test datasets.

Considering the results, no significant variation is observed during 10 iterations of classification for training, validation and test dataset and the data partitioning method provides satisfactory results for ANFIS method. In order to select the output threshold value, length of missed cracks and false alarms are calculated for the test dataset and the results are presented in Figure 5-16.

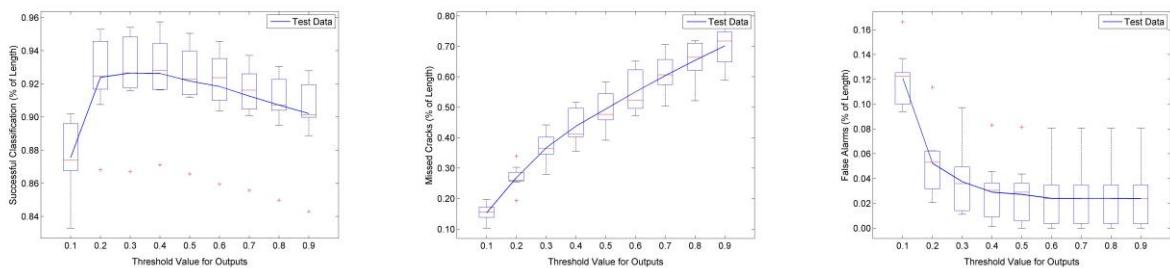


Figure 5-16. Successful classification, missed cracks and false alarms of ANFIS for test dataset.

Based on the results, threshold value of 0.1 is selected for output values. Quantification of error for all the classification methods along with discussion of the results are presented in the next section.

5.6 Discussion of the Results

5.6.1 Performance of Classifiers

In order to evaluate the classifiers' performance, the percentages of successfully classified objects, missed cracks and falsely detected objects should be calculated. Each classifier is evaluated 10 times due to the random partitioning method and the statistics of performance for training, validation and test datasets are presented in Table 5-4.

Table 5-4. Quantification of classification error based on the number of image components.

Method	Data Type	Successful Classification			Missed Cracks			False Alarms		
		Mean	Stdev	Range	Mean	Stdev	Range	Mean	Stdev	Range
DT	Training	98.98	0.0040	[98.98, 98.99]	80.52	1.3882	[78.15, 81.88]	0.08	0.0166	[0.06, 0.10]
	Validation	98.93	0.0168	[98.91, 98.95]	82.68	1.0733	[81.39, 84.62]	0.10	0.0197	[0.08, 0.14]
	Test	98.94	0.0177	[98.91, 98.97]	82.45	2.0649	[79.15, 86.11]	0.09	0.0253	[0.06, 0.14]
kNN	Training	98.90	0.0100	[98.89, 98.92]	89.25	0.6159	[88.76, 90.40]	0.05	0.0054	[0.04, 0.06]
	Validation	98.84	0.0247	[98.79, 98.88]	92.01	0.9893	[90.61, 93.85]	0.08	0.0161	[0.06, 0.11]
	Test	98.84	0.0153	[98.82, 98.87]	92.49	1.1302	[89.98, 94.18]	0.08	0.0073	[0.07, 0.10]
ANN	Training	98.19	0.0752	[98.05, 98.28]	60.73	2.0182	[59.17, 65.91]	1.12	0.0905	[0.97, 1.27]
	Validation	98.18	0.0855	[98.03, 98.31]	61.95	2.1269	[59.70, 66.99]	1.11	0.1038	[0.92, 1.29]
	Test	98.15	0.0648	[98.02, 98.24]	61.35	2.2331	[57.19, 66.07]	1.15	0.0806	[1.00, 1.28]
ANFIS	Training	98.23	0.0973	[98.13, 98.43]	65.49	0.9527	[64.19, 66.88]	1.01	0.1081	[0.80, 1.13]
	Validation	98.24	0.0917	[98.07, 98.37]	65.91	1.8979	[63.43, 69.42]	1.00	0.1035	[0.82, 1.17]
	Test	98.25	0.1167	[98.05, 98.39]	65.44	2.0112	[61.71, 68.66]	1.00	0.1287	[0.84, 1.19]

The cost function of all classification methods are evaluating the performance based on the number of successfully classified samples. However, from pavement management point of view, measuring the performance in terms of length of cracks is more advantageous. For one thing, codes and pavement maintenance manuals often quantify the extent of cracking based on the length of cracks. And for another, the number of miss classified objects is not a good measure of crack

detection accuracy. For instance, missing a small crack has the same effect on the performance of classifier as missing a large crack; however, from crack-detection point of view, missing a large crack is more important. Following this logic, the percentage of successfully classified, missed cracks and false alarms in terms of the length of objects are presented in Table 5-5.

Table 5-5. Quantification of classification error based on the length of image components.

Method	Data Type	Successful Classification			Missed Cracks			False Alarms		
		Mean	Stdev	Range	Mean	Stdev	Range	Mean	Stdev	Range
DT	Training	96.00	0.2665	[95.41, 96.33]	26.49	2.1060	[23.31, 29.56]	1.16	0.3961	[0.84, 2.15]
	Validation	95.04	0.7227	[93.48, 96.04]	29.33	3.5792	[21.57, 33.43]	1.79	0.8835	[1.08, 4.11]
	Test	95.20	0.4161	[94.40, 95.68]	29.98	4.6609	[25.12, 39.57]	1.50	0.4471	[0.94, 2.31]
kNN	Training	92.79	0.3919	[91.93, 93.46]	45.93	2.6665	[42.72, 50.51]	2.27	0.2697	[1.83, 2.78]
	Validation	92.72	0.9296	[89.06, 92.23]	54.70	6.6972	[43.25, 68.13]	3.27	0.4370	[2.48, 4.20]
	Test	91.20	0.7964	[89.34, 92.27]	54.85	5.2160	[47.03, 65.94]	3.10	0.8534	[2.00, 4.90]
ANN	Training	91.24	1.4423	[87.78, 92.70]	12.04	3.9890	[9.67, 23.02]	8.33	1.6774	[6.85, 12.62]
	Validation	90.55	2.0758	[85.49, 92.63]	14.94	4.1697	[10.38, 23.24]	8.77	2.44	[5.59, 14.61]
	Test	90.09	1.5528	[87.00, 91.76]	13.77	3.0953	[10.65, 21.07]	9.39	1.8985	[7.44, 13.16]
ANFIS	Training	89.54	0.3840	[88.84, 89.94]	13.40	0.9306	[12.64, 15.76]	10.09	0.4556	[9.57, 10.96]
	Validation	89.22	1.3920	[87.06, 91.57]	13.28	2.1806	[10.84, 16.84]	10.44	1.6268	[7.77, 12.95]
	Test	87.56	2.0383	[83.30, 90.18]	15.34	2.5388	[10.29, 19.69]	12.09	2.1323	[9.39, 16.63]

Considering the results (Table 5-5), all the classifiers provide relatively similar successful classification result. The percentage of successfully classified objects using decision tree and K-nearest neighbor method are marginally higher than those of artificial neural networks and adaptive neuro-fuzzy inference system. However, percentage of successfully classified samples is not an adequate measure of classifiers performance because total length of non-crack objects are significantly larger than the accumulated length of cracks (imbalanced data); correspondingly missed cracks have less impact on successful classification rate compared to false alarms. For this reason, percentage of missed cracks and false alarms should also be considered for evaluating the

performance of different classification algorithms. Based on the results, kNN is not an acceptable methods for the current crack detection framework since it missed more than 50 percent of the cracks (in terms of length). Significant percentage of missed cracks can alter the results of pavement evaluation drastically. Other methods produced more acceptable results: DT have significantly higher rate of missed crack and slightly lower rate of false alarms compared to ANN and ANFIS. Decision tree method failed to detect around 30 percent of crack length and falsely detected less than 2 percent of the non-crack objects while the type 1 and 2 error for ANN and ANFIS was around 10 and 15 percent, respectively. ANN slightly outperformed ANFIS with lower percentage of both, missed cracks and false alarms. It should be emphasized that the error rates of ANN and ANFIS can be tuned using the threshold value as explained earlier. Choosing a small threshold value reduces the number of missed cracks but increases the number of false alarms; while a large threshold causes less false alarms but more missed cracks. The output threshold value for ANN and ANFIS makes them more flexible and compatible for the pavement crack detection compared to DT and kNN which have categorical outputs (zero or one).

5.6.2 Stability of Classification Results for the Imbalanced Training Dataset

It has been mentioned earlier that there is a significant difference between the number of crack and non-crack objects and the dataset is highly imbalanced. This problem can have negative impact on the performance of classifiers because most classification methods, seek to minimize the number of misclassified data, disregarding the difference in cost of misclassification errors (Ganganwar, 2012). In many real-world applications (e.g. medical diagnosis, fraud detection and detection of oil spills), available data is highly imbalanced and several studies attempt to deal this

class of problems (Japkowicz & Stephen, 2002; Weiss, 2004). Proposed methods can be classified into data-level or algorithmic-level solutions (Ganganwar, 2012). Data-level solutions usually introduce re-sampling techniques that aim to balance the number of data points by under-sampling the more frequent instances or over-sampling the small class. Under-sampling methods ignore samples from the majority class and by this means, balance the training set. However, the main disadvantage of the under-sampling techniques is that potential useful information of the discarded samples will be neglected. Over-sampling methods aim to balance the training data by modify the minority class. In random or directed over-sampling, no new example is created; instead existing samples are repeated to increase the number of data points in smaller class. Nevertheless, this approach causes the over-fitting problem in many cases (Sun, Wong, & Kamel, 2009). In algorithmic-level solutions, the effect of imbalanced data is countered by modifying the classification algorithm (usually modifying the cost function). However, algorithmic-level are not available for all classification techniques and determination of misclassification costs can be a challenging task (Sun et al., 2009). Further study is required to select a data-level or an algorithmic-level solution for the pavement crack detection problem but, the data partitioning is modified in this study, to prevent some of the effects of imbalanced data on training data preparation. In order to evaluate the efficiency of the data partitioning method, standard deviation and rang of classifiers performance for all the randomly generated datasets is evaluated (Table 5-5). Considering the results, no significant trend or clear difference is observed among the results. In simple words, using the proposed random partitioning technique, classifiers produce similar results in every iteration and the main objectives of this method are satisfied.

5.6.3 Computation Time

Computations time is another important factor that should be considered for adopting a classifier. The time, required for all iterations of training, validation and testing is calculated for each classifier. One complete iteration include: 1) partitioning the entire pavement dataset; 2) training and validation of the classifier, and; 2) applying the classifier to all data partitions. The procedure is conducted on a computer with Intel® Core™ i7 – 2600, 3.40 GHz CPU and 16 GB of RAM. The average computation time for training, validation and testing a decision tree was 26.7 seconds which was faster than all other methods. Similar procedure for a k-nearest neighbors algorithm took 680.1 seconds. Longer computation times were expected for ANN and ANFIS methods since the number of adaptive parameters (e.g. weights and biases) are significantly higher than those of DT or kNN. Similar procedure took 927.1 and 8766.1 seconds for ANN and ANFIS, respectively.

5.6.4 Stability of Classification Results for Different Pavement Conditions

The Consistency of classifiers performance for pavements in different conditions is another important factor that should be considered. Mokhtari et al. (2015a), conducted a statistical study of different features (characteristics) of cracks for pavements in different conditions. According to their results, crack features can change significantly as the pavements ages. For instance, the length and area of cracks in good condition pavements are usually smaller than those of poor condition pavements. In the current study, classifiers are trained for features, extracted from pavement images in good, intermediate and poor conditions. Therefore, to make sure that the classifier

maintain its performance for all pavement types, the classification errors are quantified for pavements in good, intermediate and bad conditions, separately. The results are presented in Table 5-6.

Table 5-6. Quantification of classification error for pavement in different stages of deterioration.

Method	Good Condition			Intermediate Condition			Bad Condition		
	Successful	Missed Cracks	False Alarms	Successful	Missed Cracks	False Alarms	Successful	Missed Cracks	False Alarms
DT	98.06	29.11	0.97	95.11	26.14	1.98	93.82	28.29	0.68
kNN	95.05	50.52	3.32	91.64	46.26	3.17	87.97	49.65	2.67
ANN	89.45	12.21	10.49	91.52	10.31	8.23	92.24	13.27	6.39
ANFIS	88.25	11.39	11.76	88.98	10.96	11.02	91.48	14.11	7.13

Based on the results, there are slight changes in performance and error rates of all classifiers for pavement in different conditions. However, no significant fluctuation or drastic trend was observed for any of the methods and all classifiers seem to maintain their performance for pavements in different conditions.

5.6.5 Clarity of Classification Procedure and Logic

Comparison of performances and error rates is essential for finding the most suitable classification algorithm for the crack detection framework, but maximizing the performance is not the only objective of the current study. State-of-the-art classification methods are getting more complex to provide better classification performance (Freitas et al., 2010). As the methods grow more complex, it is even more difficult to follow the decision making procedure and inspect or interpret the logic of classification (Hammer et al., 2012). From this points of view, classification methods such as Artificial Neural Networks (ANN) or Support Vector Machines (SVM) that do

not provide an interpretable logic for their results are called ‘black-box’ methods. On the other hand, classification methods such as Decision Tree (DT) and K-Nearest Neighbors (kNN) are have interpretable procedures and their logic is easy to visualize are called ‘white-box’ methods. The acquired knowledge from interpreting the logic and the procedure of decision making can provide further insight into the pavement cracking problem. Such information can be used by agencies and decision makers for better policy making and pavement management planning.

In order to have a better understanding of the classification procedure and logic of decision making for each classifier, a synthetic dataset containing all possible feature values is generated. For this purpose the range of values for each continuous crack feature (i.e. area, length, texture, intensity and orientation) is divided into 10 intervals and the center of each interval is included into the synthetic dataset. Location is the only categorical feature that has 3 discrete possible values (0 for beyond the lane marker, 1 for outside of the wheel path and 2 for inside the wheel path) and all 3 values are included in the synthetic dataset. Considering the number of features, a total of 3×10^5 combination of features (synthetic samples) was possible. Using the synthetic data as input for pre-trained classifiers (classifiers are trained using the actual pavement dataset, as described in Section 5.5) can help understanding the decision making process of each classifier for different ranges of feature values. The outputs of ANN classifier for the synthetic dataset is evaluated and the results are presented in Figure 5-17. The horizontal axes depict different ranges of feature values and the vertical axes show the percentage of cracks and non-cracks for synthetic sample with the corresponding feature value.

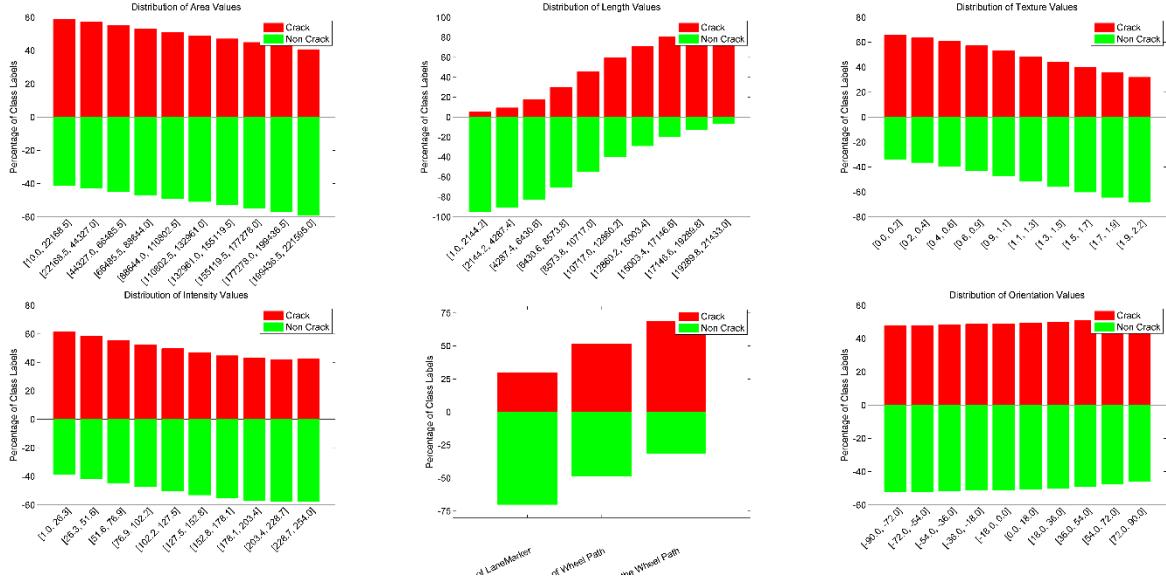


Figure 5-17. Outputs of synthetic problem using ANN classifier.

Mokhtari et al. (2015a), conducted a statistical study of the features of crack and non-crack objects in pavement images. Based on their results, cracks are larger in area, longer in length, darker in intensity and more frequent within the wheel path compared to non-crack objects. However, there are no meaningful differences between the texture indices and orientations of cracks and non-cracks.

Based on the results of synthetic dataset (Figure 5-17), the perception of ANN classifier form length, intensity, location and orientation features are compatible with the statistics of the actual pavement data. Synthetic samples with higher values of length are more frequently classified as cracks and the probability of being classified as crack decreases as the intensity value increases (object becomes brighter). The percentage of crack objects increases when the location feature indicates a sample within the wheel zone and finally, the results are unchanged for different orientation values. However, the behavior of the ANN classifier for area and texture values are not

following the logic of crack detection. The percentage of cracks decreased when the area and the roughness index of the synthetic sample increased. The results also indicate that, length and location are the most important feature for the ANN classifier because the value of these features have the most significant effect on the output of the method (i.e. outputs changed 100 percent over the entire range of length feature and around 50 percent for over location values). Area, texture and intensity are less effective for an ANN classification and the method is insensitive to the orientation of the sample.

The results of synthetic dataset does not provide an explanation for the behavior of the classifier but can be useful to demonstrate the knowledge, learned from the collected pavement images. The results of similar analyses for decision tree and K-nearest neighbors methods are presented in Figures 5-18 and 5-19, respectively.

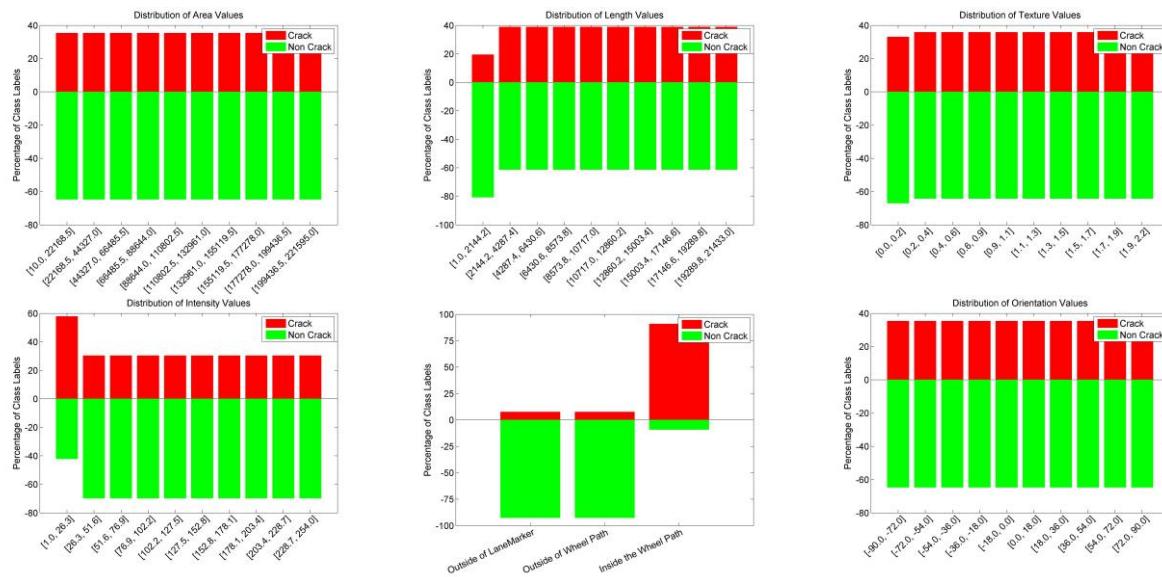


Figure 5-18. Outputs of synthetic problem using DT classifier.

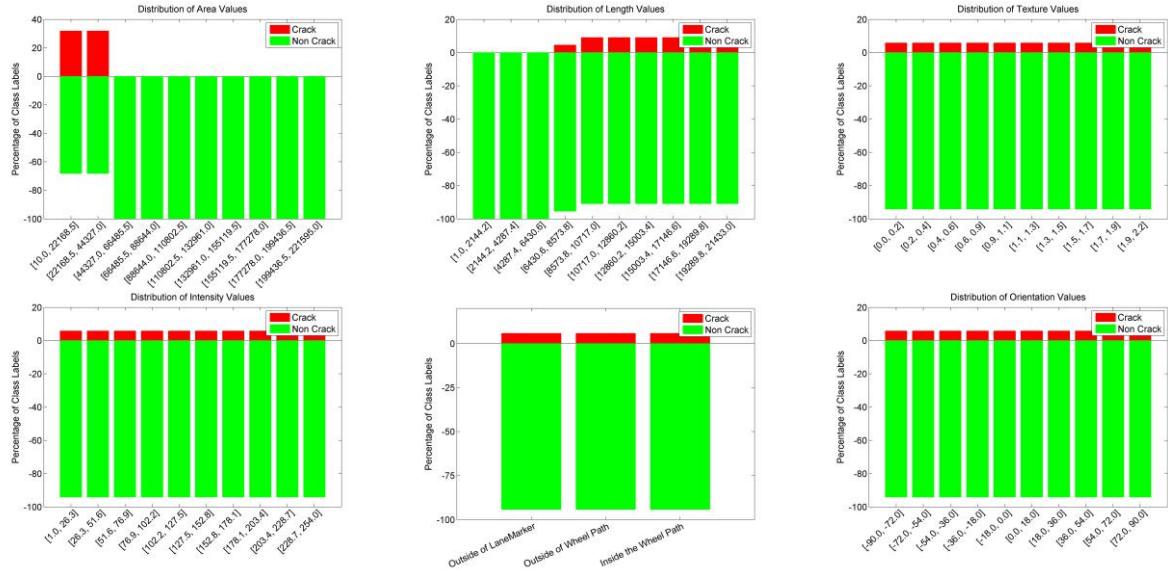


Figure 5-19. Outputs of synthetic problem using kNN classifier.

The knowledge, learned using DT and kNN methods are very different compared to the ANN method. DT algorithm is not sensitive to the area, texture and orientation features but the behavior of this method for length, intensity and location feature is compatible with human perception of the crack detection procedure. As mentioned earlier, DT perform the classification by repetitive division of features space. This logic can also be observed in the results of synthetic dataset. The change in percentage of cracks and non-cracks are not as gradual as the results of ANN method.

kNN method, on the other hand is not sensitive to most feature (e.g. length, texture, intensity, location and orientation) and the perception of this method from the area feature is against the statistics of the actual pavement data because smaller objects are more frequently classified as cracks compared to larger objects. This behavior might explain the poor results of the kNN method for detecting cracks in actual pavement images.

The outputs of ANFIS method for the synthetic dataset is also evaluated and the results are presented in Figure 5-20.

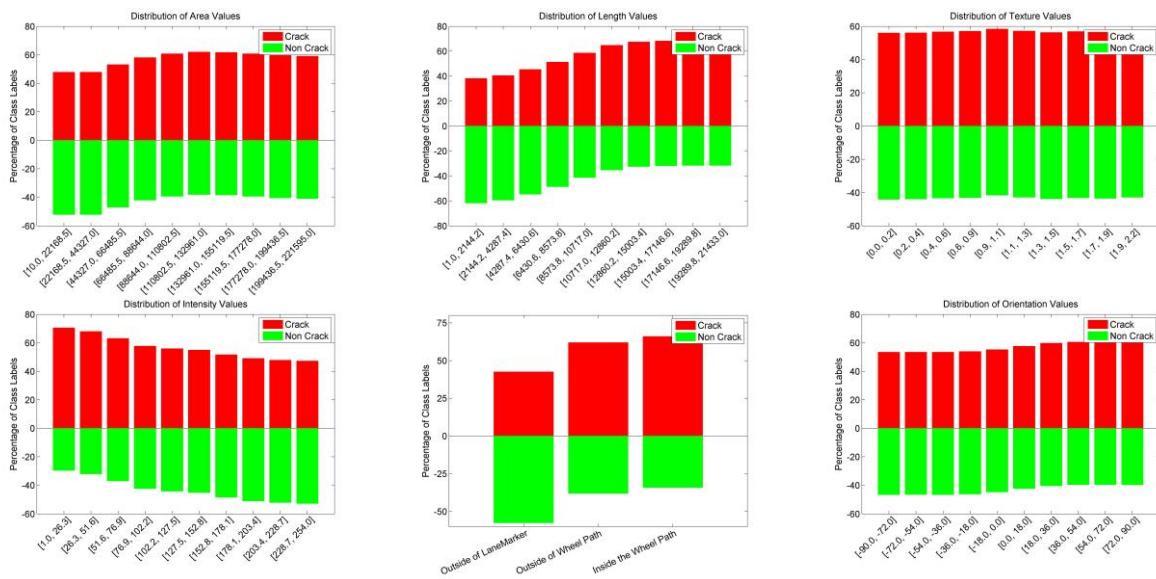


Figure 5-20. Outputs of synthetic problem using ANFIS classifier.

The perception of ANFIS classifier form all the features are compatible with the statistics of the actual pavement data. Synthetic samples with larger area, longer length, darker intensity that are located within the wheel path are more frequently classified as cracks while the effect of texture and orientation features are not significant. ANFIS method is also a white-box method and its decision making process can be visualized using the shape of the trained membership functions. Membership functions of ANFIS for the effective features (i.e. area, length, intensity and location) are presented in Figure 5-21.

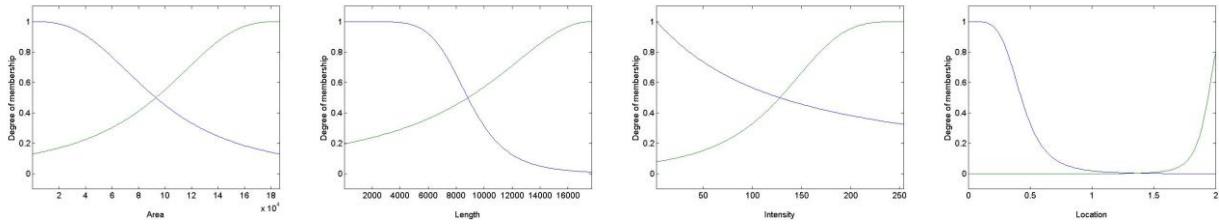


Figure 5-21. Trained membership functions of area, length, intensity and location.

Apart from the membership function of area feature that did not change significantly during the training process, other membership functions can provide further insight into the pavement crack evaluation and quantification. Using the results each feature can be divided, verbally, into two categories. For instance, objects can be short or long based on the membership functions of the length feature. The intensity membership function can also be used to divide object into bright or dark. It should be noted that further division of each feature is possible, using more membership functions for each feature but it will increase the computation time drastically. Based on the membership functions of length, three ranges of length can be roughly identified. Considering the degree of membership, objects with length of 0 to 5000 are short. From 5000 to 12000, objects have average length and objects more than 12000 pixels are long. The same definitions can be used for intensity as well. Intensity of less than 50 corresponds to dark objects while an intensity of more than 200 is bright. Objects with intensity feature of 50 to 200 are within the intermediate region. Although similar interpretation for location feature is not a meaningful but, the shape of the membership functions indicate that the wheel path designation is an effective feature for crack detection. Based on the results, objects that are located outside of the lane marks and inside the wheel path have high membership degree on different functions while the objects, outside of the wheel path belong to neither functions.

ANFIS is the only method among others that provides information about the classification procedure and the ranges of feature values through the shape of its membership functions. Following the logic of this method is even possible without the results of synthetic problem which might not be convenient to conduct in practical pavement evaluations.

5.7 Conclusions

In this study, application of four classification algorithms including Artificial Neural Network (ANN), Decision Tree (DT), k-Nearest Neighbours (kNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) is investigated to detect cracks from pavement images. In order to do so, computer vision framework, developed by We et al. (2014) and Yun et al. (2015) is employed for extraction and segmentation of crack components from pavement images that were collected with a Multi-Purpose Survey Vehicle (MPSV) by Florida Department of Transportation (FDOT). The results of the computer vision technique include crack and non-crack components. Therefore, six imagerial features including area, length, texture roughness, intensity, location and orientation are extracted from crack and non-crack components of 26 pavement at different deterioration stages. These features are used to train classifiers and evaluate their advantages and drawbacks for the image-based pavement crack detection framework. However, the extracted database is highly imbalanced and the number of cracks are significantly smaller than the number of non-crack components. Moreover, the training pavement images are selected from different stages of deterioration. Therefore, classification methods are evaluated for: 1) prediction performance, 2) computation time, 3) stability of results for highly imbalanced datasets, 4) stability of the

classifiers performance for pavements in different deterioration stages, and 5) interpretability of results and clarity of the procedure.

Comparative studies of classification algorithms often focus on the rate or percentage of successfully classified samples but, it is not an adequate measure of classifiers performance for pavement evaluation applications. Firstly, pavement management manuals usually quantify the extent of cracking based on the length of cracks. Secondly, missed cracks have less impact on successful classification rate compared to false alarms due to imbalanced dataset. Therefore, success rate of classification along with the rate of negative and positive errors in terms of the length of objects is employed to evaluate the performance of classifiers. Based on the results, ANN and ANFIS are more flexible and compatible for the pavement crack detection application. These methods not only provide superior performance but also are more flexible and compatible for the pavement crack detection because their outputs are continuous and their error rates (false positive and negative) can be tuned using an output threshold value based on the application. However computation times for ANN and ANFIS are significantly longer than those of kNN and DT due to their large number of adaptive training parameters. Especially ANFIS which took more than two hours to train, validate and test for the pavement dataset.

This study was also concerned with the effects of imbalanced data and pavement image selection on performance of classifiers. To avoid the effects of imbalanced data on training data preparation, crack and non-crack datasets are partitioned separately and the classification results are evaluated for 10 training iterations. This partitioning is considered satisfactory since the classifiers produce similar results in every iteration and no significant trend or clear difference is observed. Since the training dataset that was extracted from pavement images in different stages

of deterioration, the classifiers were applied to good, intermediate and bad condition pavements separately. The results indicate that the current method of pavement sample-image selection is satisfactory since the classifiers maintained their performance for pavements in different conditions. These findings can be used as a guideline for data preparation and validation procedure for other image-based pavement crack detection methods.

Classification results of a synthetic problem were also evaluated to provide further insight into the logic of classification and perception of each classifier from the crack detection problem. Based on the results ANFIS is the only method that completely follows the human perception from the statistics of the actual pavement data. In this method, samples with larger area, longer length, darker intensity that are located within the wheel path are more frequently classified as cracks. Being a white box method, ANFIS is also the only method that provides further information about the feature values and their effect of classification outcome. The inferred knowledge from the membership functions of this method can be used to characterize the imagerial properties of detected image components.

CHAPTER 6: CONCLUSIONS

This study aimed to extract useful information for pavement condition assessment by mining the data, acquired from an automated computer vision based system. The knowledge, discovered from evaluating the results and interpreting the detection procedure is used to explain the cracking progression and pavement deterioration process. The automated image-based system, used in this study, performs a sequence of computer vision and machine learning techniques to detect cracks from pavement images and quantify the imagery characteristics (features) of cracks, automatically, to avoid subjective crack rating based on human visual inspection.

In this study, the random background of pavement images is removed using a bottom-hat morphological image processing technique. Experimental validation of this technique using the pavement data, collected with a Multi-Purpose Survey Vehicle (MPSV) by the Florida Department of Transportation (FDOT), is presented in the first chapter. Based on the results, the main advantage of the bottom-hat transformation is that cracks can be detected, selectively based on their shapes, sizes, and intensities using a relatively small number of user-defined parameters.

The problem of most image-based crack detection methods, including bottom-hat transformation, is that the extracted crack-image components are usually fragmented along the crack paths; therefore, a novel crack de-fragmentation technique, so called MorphLink-C, is proposed to connect crack fragments. This method consists of two sub-processes: 1) fragments grouping using morphological dilation transform, and 2) adaptive connection of fragments using morphological thinning transform. This method can be used for any pixels level crack detection method and is not limited to a certain type of crack. MorphLink-C can be used for complicated

crack networks, such as single cracks, branched cracks, block cracks, and alligator cracks. Experimental validation results using the pavement dataset indicate that the MorphLink-C increases the crack-detection accuracy by reducing the false negative error. For the collected dataset, MorphLink-C increased the percentage of successfully classified objects by less than 5 percent and decreased the false negative error (missed cracks) up to 50 percent. MorphLink-C also reduces the computation time for crack detection by almost 50 percent.

Although the image processing techniques are efficient for removing random background pixels from pavement images, the results still include non-crack objects. Therefore, application of four classification algorithms including Artificial Neural Network (ANN), Decision Tree (DT), k-Nearest Neighbours (kNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) is investigated for the crack detection framework. Classification methods are evaluated for different criteria including: 1) prediction performance, 2) computation time, 3) stability of results for highly imbalanced datasets in which, the number of crack objects are significantly smaller than the number of non-crack objects, 4) stability of the classifiers performance for pavements in different deterioration stages, and 5) interpretability of results and clarity of the procedure. Special emphasis is placed on white box methods (e.g. Decision Tree and Adaptive Neuro-Fuzzy Inference System) which are more interpretable because the acquired knowledge from interpreting the logic and the procedure of classification can provide further insight into the pavement cracking procedure.

The results of this study indicate that the percentage of successfully classified samples, which is commonly used in most studies, is not an adequate measure of classifiers performance for pavement evaluation applications and the percentage of negative and positive errors in terms of the length of objects should be employed to compare the performance of classifiers. The

percentage of successfully classified samples for all classifiers were above 98 percent when applied to the collected dataset (26 images). However, comparison of classifiers based on the percentage of missed cracks and false alarms in terms of the length of cracks indicate that DT and KNN methods that have simple classification logics are not capable of detecting cracks from pavement images and ANN and ANFIS methods provide superior performances. The percentage of missed cracks for DT and KNN is 50 and 30 percent, respectively. ANN and ANFIS missed less than 10 percent of the crack length and produced around 10 percent false alarms for the extracted data. A different configuration of ANN and ANFIS is also proposed in this research to make these methods more flexible and compatible for the pavement crack detection. These methods are configured to produce continuous outputs so that their error rates (false positive and negative) can be tuned using an output threshold value. However, ANN acts as a ‘black box’ and understanding its classification logic is hardly possible. ANFIS, on the other hand, is a ‘white box’ method and provides significant information about the logic of classification and the effect of feature values on crack detection results. ANFIS is also the only method that completely follows the human perception from the statistics of the actual pavement data (samples with larger area, longer length, darker intensity that are located within the wheel path are more frequently classified as cracks). Another advantage of ANFIS is that the inferred knowledge from the membership functions of this method can be used to quantify the qualitative properties of detected image components. For the collected data from 26 pavement images, qualitative characteristics such as long, medium and short cracks is quantified using the input membership function of the trained ANFIS classifier. Based on the results, short objects are less than 6000 millimeter while long image components are more than 14000 millimeters and medium objects are within the range of 6000 to

12000 millimeter. Similar information can be extracted for other features of extracted image components. These quantitative measure are essential for pavement condition assessment.

The results also signifies the importance of training data preparation and validation procedure. Training dataset should be extracted from pavement images in different stages of deterioration so that the dataset can represent the characteristics of cracks in good, intermediate and bad condition pavements. Training datasets were prepared through manual selection of true cracks in a total of 26 flexible pavement images including 11 images in good, 7 images in intermediate, and 8 images in poor pavement condition. A more number of images were chosen from good conditions since these images usually have a less number of cracks than the images in poor condition. A total of 264981 image components were extracted, including 3090 (1.17%) crack and 261891 (98.83%) non-crack objects. The total length of the object was 3915.1 m, including 444.3 m (11.35%) for crack and 3470.8 m (88.65%) for non-crack objects.

The characteristics of pavement cracks are extracted and quantified as imagery features using the proposed computer vision method. Six crack features including area, length, orientation, intensity, texture roughness, and wheel path position, which are commonly used in pavement applications, are extracted from surface images of flexible pavement in different deterioration stages. A comprehensive statistical analysis of the extracted features, using wrapper exhaustive search with ANN classifiers and filter feature subset selection (FSS) methods, including Fischer score, Gini index, information gain, ReliefF, mRmR, and FCBF, was conducted to avoid subjective and intuitive feature selection based on human observations. Finding the optimal feature subset is also important for effective statistical characterization and eliminating irrelevant and redundant information. Based on the results, length, location and intensity are the most important features,

characterizing image components to be crack for the given images. Based on the results of the wrapper feature selection method using ANN classifier, the feature subset that includes these three features provided a mean square error or 0.0095 which is comparable to the performance of the full feature subset (6 features). These feature also ranks as the top three relevant features according to filter feature evaluation methods.

Special emphasis is placed on statistics of crack image components, identified using the computer-vision method, in correspondence with the main objective of this study. Knowledge, discovered from statistical analysis of the features can provide useful information to characterize cracking process and pattern in different stages of aging flexible pavement. The information of crack patterns could be used as fundamental data to provide justification for rehabilitation policies. The statistical analysis indicates that poor pavement segments had similar percentages of longitudinal and transverse cracks, while good pavement had mostly longitudinal crack. This fact reflects the progress of structural cracks, such as longitudinal and alligator cracks which are usually caused by traffic loads. Alligator cracking is a major structural distress, which is caused by fatigue damage in flexible pavement with granular and weakly stabilized bases. Based on the results, alligator crack first appears as parallel longitudinal cracks in wheel paths, and then progresses into a network of interconnecting (transverse) cracks. The results also support the validity of having different weighting factors based on the wheel path designation in crack rating calculation, proposed by Florida Department of Transportation (FDOT, 2012b) since statistically, more cracks were detected in wheel paths through image processing. The statistical analysis also provides quantified information about the types of cracks based on their width in different stages of pavement deterioration. For the collected data from 26 pavement images, the mean of crack width

was measured at 3.62 mm for good, 3.98 mm for intermediate, and 4.06 mm for poor condition as pavement is more damaged. Considering the guidelines of the Florida Department of Transportation (2012) the percentages of Class-1B, Class-2 and Class-3 cracks for different pavement conditions can be calculated using the width of cracks. For the collected data, the percentage of Class-3 crack increases as pavement ages. The percentage of Class-1B crack decreases while the percentage of Class-2 crack increases when pavement is in intermediate or poor conditions. In comparison of the intermediate and poor conditions, the percentages of Class-1B and Class-2 cracks are almost the same, while poor pavement has more Class-3 cracks than intermediate pavement.

The above results validate that the statistical characterization methods associated with the computer-vision techniques used in this study can extract useful information of pavement cracks in different deterioration stages, which can be accurate quantitative information to understand past and current pavement conditions and to justify pavement maintenance policy. Moreover, the findings of this study and the comprehensive study of different feature subset selection methods and classification algorithms, can also be used in other similar studies to identify the optimal set of features and select the best classification method considering the objectives of the automated image-based pavement evaluation method. However, further study is recommended for other aspects of vision based pavement evaluation. The proposed should be applied for larger datasets in order to generalize the results and provide overall condition assessment of the roadways. The method can also be expanded to be able to detect other type of pavement defects such as potholes and patches. Further study is also required to select a data-level or an algorithmic-level solution to reduce the effect of imbalance dataset on performance of classifiers.

APPENDIX: APPROVAL LETTERS



RightsLink®

Home

Create Account

Help



AMERICAN SOCIETY OF CIVIL ENGINEERS

Title:

Improvement of Crack-Detection Accuracy Using a Novel Crack Defragmentation Technique in Image-Based Road Assessment

Author:

Liuliu Wu, Soroush Mokhtari, Abdenour Nazef, et al

Publication:

Journal of Computing in Civil Engineering

Publisher:

American Society of Civil Engineers

Date:

11/14/2014

Copyright © 2014, ASCE. All rights reserved.

LOGIN

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials. Already a RightsLink user or want to learn more?

Permissions Request

As an ASCE author, you are permitted to reuse your own content for another ASCE or non-ASCE publication.

Please add the full credit line "With permission from ASCE" to your source citation. Please print this page for your records.

Type of use: Dissertation/Thesis

Portion: full article

Format: electronic

Use of this content will make up more than 25% of the new work: no

Author of this ASCE work or ASCE will publish the new work: yes

BACK**CLOSE WINDOW**

Copyright © 2015 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#), [Terms and Conditions](#).
Comments? We would like to hear from you. E-mail us at customercare@copyright.com.

Figure A-1. Permission from ASCE Journal of Computing in Civil Engineering for using the paper entitled, 'Improvement of crack-detection accuracy using a novel crack defragmentation technique in image-based road assessment' as the first chapter of this study.

REFERENCES

- Ahuja, N., & Rosenfeld, A. (1977). *A Note on the Use of Second-Order Gray Level Statistics for Threshold Selection* (No. TR-572). Maryland University College Park Computer Science Center.
- American Association of State Highway and Transportation Officials. (1993). *AASHTO Guide for Design of Pavement Structures*. Washington D.C.
- American Society of Civil Engineers. (2013). *2013 Report Card for America's Infrastructure* (pp. 1–74). Washington, D.C., USA.
- Amirikian, B., & Nishimura, H. (1994). What size network is good for generalization of a specific task of interest? *Neural Networks*, 7(2), 321–329. [http://doi.org/10.1016/0893-6080\(94\)90026-4](http://doi.org/10.1016/0893-6080(94)90026-4)
- Ayenu-Prah, A., & Attoh-Okine, N. (2008). Evaluating pavement cracks with bidimensional empirical mode decomposition. *EURASIP Journal on Advances in Signal Processing*, 2008(1), 861701. doi:10.1155/2008/861701
- Baum, E. B., & Haussler, D. (1989). What size net gives valid generalization? *Neural Computation*, 1(1), 151–160. <http://doi.org/10.1162/neco.1989.1.1.151>
- Bianchini, A., & Bandini, P. (2010). Prediction of pavement performance through neuro-fuzzy reasoning. *Computer-Aided Civil and Infrastructure Engineering*, 25(1), 39–54. <http://doi.org/10.1111/j.1467-8667.2009.00615.x>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140. <http://doi.org/10.1007/BF00058655>

- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees*. Chapman and Hall (1 edition). Retrieved from <http://www.amazon.com/Classification-Regression-Wadsworth-Statistics-Probability/dp/0412048418>
- Cheng, H. D. (1996). Automated real-time pavement distress detection using fuzzy logic and neural network. In S. B. Chase (Ed.), *Nondestructive Evaluation Techniques for Aging Bridges and Highways* (Vol. 2946, pp. 140–151). doi:10.1117/12.259131
- Cheng, H. D., Chen, J.-R., Glazier, C., & Hu, Y. G. (1999). Novel approach to pavement cracking detection based on fuzzy set theory. *Journal of Computing in Civil Engineering*, 13(4), 270–280. doi:10.1061/(ASCE)0887-3801(1999)13:4(270)
- Cheng, H. D., Shi, X. J., & Glazier, C. (2003). Real-time Image thresholding based on sample space reduction and interpolation approach. *Journal of Computing in Civil Engineering*, 17(4), 264–272. doi:10.1061/(ASCE)0887-3801(2003)17:4(264)
- Cheng, H. D., Wang, J., Hu, Y. G., Glazier, C., Shi, X., & Chen, X. (2001). Novel approach to pavement cracking detection based on neural network. *Transportation Research Record*, 1764(1), 119–127. doi:10.3141/1764-13
- Chou, J., O'Neill, W. A., & Cheng, H. D. (1995). Pavement distress evaluation using fuzzy logic and moment invariants. In *Transportation Research Record* (pp. 39–46). Washington, DC: Transportation Research Board. Retrieved from <http://pubsindex.trb.org/view.aspx?id=453036>
- Chou, J., O'Neill, W., & Cheng, H. (1994). Pavement distress classification using neural networks. In *Proceedings of IEEE International Conference on Systems, Man and*

Cybernetics (Vol. 1, pp. 397–401). San Antonio, TX: IEEE.

doi:10.1109/ICSMC.1994.399871

Chu, Y. (2010). Algorithm study on thinning and keeping connectivity of bituminous pavement crack images. In *2010 International Conference of Information Science and Management Engineering* (Vol. 265, pp. 550–553). Xi'an China: Ieee. doi:10.1109/ISME.2010.265

Cuhadar, A., Shalaby, K., & Tasdoken, S. (2002). Automatic segmentation of pavement condition data using wavelet transform. In *IEEE CCECE2002. Canadian Conference on Electrical and Computer Engineering*, (Vol. 2, pp. 1009–1014). IEEE.
doi:10.1109/CCECE.2002.1013082

Cunningham, P., & Delany, S. J. (2007). *k -Nearest Neighbour Classifiers* (pp. 1–17). Dublin.

Dong, L., Yu, G., Ogunbona, P., & Li, W. (2008). An efficient iterative algorithm for image thresholding. *Pattern Recognition Letters*, 29(9), 1311–1316.
doi:10.1016/j.patrec.2008.02.001

Dougherty, E. R., & Lotufo, R. A. (2003). *Hands-on Morphological Image Processing*. The Society of Photo-Optical Instrumentation Engineers (SPIE).

Evdorides, H. T., Schlotjes, M. R., Henning, T. F. P., & Burrow, M. P. N. (2014). Using support vector machines to predict the probability of pavement failure. *Proceedings of the ICE - Transport*, 1–11. <http://doi.org/10.1680/tran.12.00084>

Federal Highway Administration. (2001). *Pavement Preventive Maintenance Guidelines*.

Retrieved from

<http://www.mdt.mt.gov/publications/docs/brochures/research/toolbox/FHWA/PavPrevM ainGuides.pdf>

Federal Highway Administration. (2003). *Distress Identification Manual for the Long-Term Pavement Performance Program*. McLean, VA.

Finn, F. (1998). Pavement management systems--Past, present, and future. *Public Roads*, 62(1).

Florida Department of Transportation. (2012a). *Flexible Pavement Condition Survey*.

Tallahassee, Florida. Retrieved from

<http://www.dot.state.fl.us/statematerialsoffice/administration/resources/library/publications/researchreports/pavement/flexiblehandbook.pdf>

Florida Department of Transportation. (2012b). *Rigid Pavement Condition Survey*. Tallahassee, Florida. Retrieved from

<http://www.dot.state.fl.us/statematerialsoffice/pavement/performance/pcs/>

Florida Department of Transportation. (2013). *History of Florida Pavement Condition Survey: 1973-2013*. Tallahassee, Florida. Retrieved from

<http://www.dot.state.fl.us/statematerialsoffice/pavement/performance/pcs/>

Florida Department of Transportation. (2014). *2014 Performance Report*. Tallahassee, Florida. Retrieved from <http://www.dot.state.fl.us/planning/performance/2014/>

Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *The Journal of Machine Learning Research*, 3, 1289–1305. Retrieved from <http://dl.acm.org/citation.cfm?id=944919.944974>

Freitas, A. A., Wieser, D. C., & Apweiler, R. (2010). On the importance of comprehensible classification models for protein function prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 7(1), 172–82.

<http://doi.org/10.1109/TCBB.2008.47>

- Ganganwar, V. (2012). An overview of classification algorithms for imbalanced datasets. *Int. J. Emerg. Technol. Adv. Eng*, 2(4), 42–47. Retrieved from http://www.ijetae.com/files/Volume2Issue4/IJETAE_0412_07.pdf
- Gavilán, M., Balcones, D., Marcos, O., Llorca, D. F., Sotelo, M. A., Parra, I., Amírola, A. (2011). Adaptive road crack detection system by pavement classification. *Sensors*, 11(10), 9628–9657. <http://doi.org/10.3390/s111009628>
- Golparvar-Fard, M., Balali, V., & Garza, de la J. M. (2012). Segmentation and recognition of highway assets using image-based 3D point clouds and semantic texton forests. *Journal of Computing in Civil Engineering*, (540). doi:10.1061/(ASCE)CP.1943-5487.0000283
- Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2009). *Digital Image Processing Using Matlab* (2nd ed.). Gatesmark Publishing.
- Gu, Q., Li, Z., & Han, J. (2012). Generalized Fisher score for feature selection. *arXiv*. Retrieved from <http://arxiv.org/abs/1202.3725>
- Gunaratne, M., Amarasinghe, S., & Nasseri, S. (2008). *Investigation of Automated and Interactive Crack Measurement Systems* (No. BD544-36). Florida Department of Transportation.
- Hall, K. T., Correa, C. E., Carpenter, S. H., & Elliot, R. P. (2001). *Rehabilitation Strategies for Highway Pavements* (Vol. 35). Transportation Research Board.
- Hall, M. (1999). *Correlation-Based Feature Selection for Machine Learning*. The University of Waikato. Retrieved from <https://www.iri.fr/~pierres/donn%20es/save/these/articles/lpr-queue/hall99correlationbased.pdf>
- Hammer, B., Mokbel, B., Schleif, F. M., & Zhu, X. (2012). White box classification of dissimilarity data. In Lecture Notes in Computer Science. *Including subseries Lecture*

Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, Vol. 7208 LNAI, pp.

309–321. Retrieved from http://doi.org/10.1007/978-3-642-28942-2_28

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 610–621.
doi:10.1109/TSMC.1973.4309314

He, Y., Qiu, H., Wang, J., Zhang, W., & Xie, J. (2011). Studying of road crack image detection method based on the mathematical morphology. In *4th International Congress on Image and Signal Processing*, pp. 967–969.

Ho, T.-W., Chou, C.-C., & Lin, J.-D. (2012). Study on pavement distress image recognition using clustering and classification algorithms and spatial database techniques. In *Transportation Research Board 91st Annual Meeting*. Retrieved from <http://trid.trb.org/view.aspx?id=1130740>

Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2), 251–257. [http://doi.org/10.1016/0893-6080\(91\)90009-T](http://doi.org/10.1016/0893-6080(91)90009-T)

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366. [http://doi.org/10.1016/0893-6080\(89\)90020-8](http://doi.org/10.1016/0893-6080(89)90020-8)

Hu, Y., & Zhao, C. (2010). A local binary pattern based methods for pavement crack detection. *Journal of Pattern Recognition Research* 1, 1(20103), 140–147.

Hu, Y., & Zhao, C. (2010b). A novel LBP based methods for pavement crack detection. *Journal of Pattern Recognition Research*, 5(1), 140–147. doi:10.13176/11.167

- Huang, J., Liu, W., & Sun, X. (2014). A pavement crack detection method combining 2D with 3D information based on Dempster-Shafer theory. *Computer-Aided Civil and Infrastructure Engineering*, 29(4), 299–313. doi:10.1111/mice.12041
- Huang, W., & Zhang, N. (2012). A novel road crack detection and identification method using digital image processing techniques. In *7th International Conference on Computing and Convergence Technology (ICCCT)*, pp. 397–400. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6530365
- Huang, Y., & Xu, B. (2006). Automatic inspection of pavement cracking distress. *Journal of Electronic Imaging*, 15(1), 013017. doi:10.1117/1.2177650
- Hung, M. S., Hu, M. Y., Shanker, M. S., & Patuwo, B. E. (1996). Estimating posterior probabilities in classification problems with neural networks. *International Journal of Computational Intelligence and Organizations*, 1(1), 49–60.
- Illinois Department of Transportation. (2004). *Condition Rating Survey Manual*. Springfield, IL.
- Jahanshahi, M. R., Kelly, J. S., Masri, S. F., & Sukhatme, G. S. (2009). A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. *Structure and Infrastructure Engineering*, 5(6), 455–486.
doi:10.1080/15732470801945930
- Jahanshahi, M. R., Masri, S. F., Padgett, C. W., & Sukhatme, G. S. (2011). An innovative methodology for detection and quantification of cracks through incorporation of depth perception. *Machine Vision and Applications*, 24(2), 227–241.
<http://doi.org/10.1007/s00138-011-0394-0>

- Jang, J. (1993). ANFIS: adaptive-network-based fuzzy inference system. *Systems. Man and Cybernetics, IEEE Transactions*, 23(3). Retrieved from
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=256541
- Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis*, 6(5), 429–449.
- Javidi, B., Stephens, J., Kishk, S., Naughton, T., McDonald, J., & Isaac, A. (2003). *Pilot for Automated Detection and Classification of Road Surface Degradation Features*. State Publications. Rocky Hill, CT. Retrieved from
<http://cslib.cdmhost.com/cdm/singleitem/collection/p128501coll2/id/34215>
- Jing, L., & Aiqin, Z. (2010). Pavement crack distress detection based on image analysis. *2010 International Conference on Machine Vision and Human-Machine Interface*, (c), 576–579. doi:10.1109/MVHI.2010.10
- Kansas Department of Transportation. (2013). *Pavement Management Information System*. Topeka, KS.
- Kaseko, M. S., & Ritchie, S. G. (1993). A neural network-based methodology for pavement crack detection and classification. *Transportation Research Part C: Emerging Technologies*, 1(4), 275–291. doi:10.1016/0968-090X(93)90002-W
- Kirschke, K. R., & Velinsky, S. A. (1992). Histogram based approach for automated pavement crack sensing. *Journal of Transportation Engineering*, 118(5), 700–710. doi:10.1061/(ASCE)0733-947X(1992)118:5(700)
- Koch, C., & Brilakis, I. (2011). Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, 25(3), 507–515. doi:10.1016/j.aei.2011.01.002

- Koch, C., Jog, G. M., & Brilakis, I. (2013). Automated pothole distress assessment using asphalt pavement video data. *Journal of Computing in Civil Engineering*, 27 July/August, 370–378. doi:10.1061/(ASCE)CP.1943-5487.0000232.
- Koutsopoulos, H., & Downey, A. (1993). Primitive-based classification of pavement cracking images. *Journal of Transportation*, 119(3), 402–418. doi:10.1061/(ASCE)0733-947X(1993)119:3(402)
- Ladha, L., & Deepa, T. (2011). Feature selection methods and algorithms. *International Journal on Computer Science and Engineering*, 3(5), 1787–1797.
- Lee, B. J., & Lee, H. D. (2004). Position-invariant neural network for digital pavement crack analysis. *Computer-Aided Civil and Infrastructure Engineering*, 19(2), 105–118.
<http://doi.org/10.1111/j.1467-8667.2004.00341.x>
- Li, J. (2003). Image processing algorithm for detecting the pavement crack diseases. *Computer Engineering and Applications*, 35, 211–213. Retrieved from
http://en.cnki.com.cn/Article_en/CJFDTOTAL-JSGG200335067.htm
- Li, Q., & Liu, X. (2008). Novel approach to pavement image segmentation based on neighboring difference histogram method. *2008 Congress on Image and Signal Processing*, 792–796.
doi:10.1109/CISP.2008.13
- Li, Q., Zou, Q., Zhang, D., & Mao, Q. (2011). FoSA: F* Seed-growing approach for crack-line detection from pavement images. *Image and Vision Computing*, 29(12), 861–872.
doi:10.1016/j.imavis.2011.10.003

- Liu, F., Xu, G., Yang, Y., Niu, X., & Pan, Y. (2008). Novel approach to pavement cracking automatic detection based on segment extending. *2008 International Symposium on Knowledge Acquisition and Modeling*, 610–614. doi:10.1109/KAM.2008.29
- Madani, K., Sheikhmohammady, M., Mokhtari, S., Moradi, M., & Xanthopoulos, P. (2014). Social Planner's Solution for the Caspian Sea Conflict. *Group Decision and Negotiation*, 23(3), 579–596. doi:10.1007/s10726-013-9345-7
- Mallat, S., & Zhong, S. (1992). Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7), 710–732. doi:10.1109/34.142909
- Mancini, A., Frontoni, E., & Zingaretti, P. (2013). Road pavement crack automatic detection by MMS images. In *21st Mediterranean Conference on Control & Automation* (pp. 1589–1596).
- Maode, Y., Shaobo, B., Kun, X., & Yuyao, H. (2007). Pavement crack detection and analysis for high-grade highway median-filter algorithm with 3 morphological detection edge. In *2007 8th International Conference on Electronic Measurement and Instruments* (pp. 4–548–4–552). IEEE. doi:10.1109/ICEMI.2007.4351202
- Matheron, G. (1975). *Random Sets and Integral Geometry*. New York, New York, USA: Wiley.
- Meignen, D., Bernadet, M., & Briand, H. (1997). One application of neural networks for detection of defects using video data bases: identification of road distresses. In *Database and Expert Systems Applications. 8th International Conference, DEXA '97* (pp. 459–464). IEEE Comput. Soc. doi:10.1109/DEXA.1997.617332
- Minkowski, H. (1903). Volumen und oberfläche. *Mathematische Annalen*, 57(4), 447–495.

- Mirakhori, A., Farahani, M. H., & Ramtin, F. (2009). New approach in supplier selection using linear physical programming. In *Service Operations, Logistics and Informatics*, SOLI'09. IEEE/INFORMS International Conference (pp. 47-51). IEEE.
<http://doi.org/10.1109/SOLI.2009.5203902>
- Moghadas Nejad, F., & Zakeri, H. (2011a). A comparison of multi-resolution methods for detection and isolation of pavement distress. *Expert Systems with Applications*, 38(3), 2857–2872. <http://doi.org/10.1016/j.eswa.2010.08.079>
- Moghadas Nejad, F., & Zakeri, H. (2011b). An expert system based on wavelet transform and radon neural network for pavement distress classification. *Expert Systems with Applications*, 38(6), 7088–7101. <http://doi.org/10.1016/j.eswa.2010.12.060>
- Mokhtari, S. (2012). *Developing a Group Decision Support System (GDSS) for Decision Making Under Uncertainty* (Master's thesis, University of Central Florida Orlando, Florida).
- Mokhtari, S., Madani, K., & Chang, N.-B. (2012). Multi-criteria decision making under uncertainty: application to California's Sacramento-San Joaquin Delta problem. In *World Environmental and Water Resources Congress 2012* (pp. 2339–2348). Reston, VA: American Society of Civil Engineers. doi:10.1061/9780784412312.236
- Mokhtari, S., Yun, H.-B., & Wu, L. (2015a) *Statistical characterization of crack feature for aging road pavement using computer-vision techniques*. Unpublished manuscript.
- Mokhtari, S., Yun, H.-B., & Wu, L. (2015b) *Comparative Study of classification methods for image-based pavement crack detection*. Unpublished manuscript.

- Molina, L. C., Belanche, L., & Nebot, À. (2002). Feature selection algorithms : a survey and experimental evaluation. In *2002 IEEE International Conference on Data Mining* (pp. 306–313). IEEE. doi:10.1109/ICDM.2002.1183917
- Moussa, G., & Hussain, H. (2011). A new technique for automatic detection and parameters estimation of pavement crack. In *4th International Multi-Conference on Engineering and Technological Innovation* (pp. 1–6). Orlando, Florida. Retrieved from http://www.iiis.org/CDs2011/CD2011SCI/IMETI_2011/PapersPdf/FA884ZF.pdf
- New York State Department of Transportation. (2010). *Pavement Condition Assessment*. Albany, New York.
- Nguyen, T. S. Y., Avila, M., & Begot, S. (2009). Automatic detection and classification of defect on road pavement using anisotropy measure. In *European Signal Processing Conference* (pp. 617–621). Glasgow. Retrieved from <http://hal.archives-ouvertes.fr/hal-00666919>
- Nguyen, T. S., Begot, S., Duculty, F., & Avila, M. (2011). Free-Form anisotropy: a new method for crack detection on pavement surface images. In *18th IEEE International conference on Image processing* (pp. 1093–1096). Bruxelles, Belgium.
- Noori, M., Tatari, O., Nam, B., Golestani, B., & Greene, J. (2014). A stochastic optimization approach for the selection of reflective cracking mitigation techniques. *Transportation Research Part A: Policy and Practice*, 69, 367-378.
<http://doi.org/10.1016/j.tra.2014.09.006>
- Oh, H., Achenie, L., & Garrick, N. (1997). *Automated Pavement Evaluation System for Pavement Distress Assessment*. University of Connecticut, Department of Civil and

Environmental Engineering. Retrieved from http://www.netc.uconn.edu/pdf/jhr97-261_94-3.pdf

Oliveira, H., & Correia, P. L. (2008). Supervised strategies for crack detection in images of road pavement flexible surfaces. In *European Signal Processing Conference (EUSIPCO'08)* (pp. 25–29). Retrieved from <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2008/papers/1569105104.pdf>

Oliveira, H., & Correia, P. L. (2009). Automatic road crack segmentation using entropy and image dynamic thresholding. In *European Signal Processing Conference (EUSIPCO'09)* (pp. 622–626). Retrieved from <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2009/contents/papers/1569192726.pdf>

Oliveira, H., & Lobato Correia, P. (2008). Identifying and retrieving distress images from road pavement surveys. In *2008 15th IEEE International Conference on Image Processing* (pp. 57–60). IEEE. doi:10.1109/ICIP.2008.4711690

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, And Cybernetics*, SMC-9(1), 62–66. doi:10.1109/TSMC.1979.4310076

Petrou, M., Kittler, J., & Song, K. Y. (1996). Materials processing technology. *Journal of Materials Processing Technology*, 56, 158–167.

Pratt, W. K. (2001). *Digital Image Processing* (3rd ed.). New York, New York, USA: Wiley.

Rafinejad, S. N., Ramtin, F., & Arabani, A. B. (2009). A new approach to generate rules in genetic algorithm solution to a job shop schedule by fuzzy clustering. In *Proceedings of*

the World Congress on Engineering and Computer Science, USA.

<http://doi.org/10.1109/SOLI.2009.5203902>

Ramtin, F., & Pazour, J. A. (2014). Analytical models for an automated storage and retrieval system with multiple in-the-aisle pick positions. *IIE Transactions*, 46(9), 968-986.

<http://doi.org/10.1080/0740817X.2014.882037>

Read, L., Mokhtari, S., Madani, K., Maimoun, M., & Hanks, C. (2013). A multi-participant, multi-criteria analysis of energy supply sources for Fairbanks, Alaska. In *Proceedings of the 2013 World Environmental and Water Resources Congress* (pp. 1247-1257).

<http://dx.doi.org/10.1061/9780784412947.123>

Richard, M. D., & Lippmann, R. P. (1991). Neural network classifiers estimate bayesian a posteriori probabilities. *Neural Computation*, 3(4), 461–483.

<http://doi.org/10.1162/neco.1991.3.4.461>

Rokach, L., & Maimon, O. (2005). Top-down induction of decision trees classifiers – a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 35(4), 476–487.

Saar, T., & Talvik, O. (2010). Automatic Asphalt pavement crack detection and classification using neural networks. In *2010 12th Biennial Baltic Electronics Conference* (pp. 345–348). IEEE. <http://doi.org/10.1109/BEC.2010.5630750>

Sahoo, P., Soltani, S., & Wong, A. K. . (1988). A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41(2), 233–260. doi:10.1016/0734-189X(88)90022-9

- Salari, E., & Ouyang, D. (2012). An image-based pavement distress detection and classification. *In 2012 IEEE International Conference on Electro/Information Technology* (pp. 1–6). IEEE. <http://doi.org/10.1109/EIT.2012.6220706>
- Salembier, P. (1990). Comparison of some morphological segmentation algorithms based on contrast enhancement. application to automatic defect detection. *In Proceedings of 5th European signal processing conference* (pp. 833–836).
- Siddharth, S., Ramakrishnan, P. K., Krishnamurthy, G., & Santhi, B. (2012). Pavement crack classifiers: A comparative study. *Research Journal of Applied Sciences, Engineering and Technology*, 4(24), 5434–5437. Retrieved from http://www.jourlib.org/paper/2560485#.VQhNaY7F_wo
- Sietsma, J., & Dow, R. J. F. (1991). Creating artificial neural networks that generalize. *Neural Networks*, 4(1), 67–79. [http://doi.org/10.1016/0893-6080\(91\)90033-2](http://doi.org/10.1016/0893-6080(91)90033-2)
- Smith, R. E., Freeman, T. J., & Pendleton, O. J. (1996). Evaluation of automated pavement distress data collection procedures for local agency pavement management. *Research Project GC10470, Washington State Department of Transportation, Olympia, Washington.*
- Song, K. Y., Petrou, M., & Kittler, J. (1995). Texture crack detection. *Machine Vision and Applications*, 8(1), 63–75. doi:10.1007/BF01213639
- Subirats, P., Dumoulin, J., Legeay, V., & Barba, D. (2006). Automation of pavement surface crack detection using the continuous wavelet transform. *In 2006 International Conference on Image Processing* (Vol. 1, pp. 3037–3040). IEEE. doi:10.1109/ICIP.2006.313007

- Sun, B., & Qiu, Y. (2007). Automatic identification of pavement cracks using mathematic morphology. In *International Conference on Transportation Engineering 2007* (ICTE2007) (pp. 1783–1788). Reston, Virginia, USA.
doi:[http://dx.doi.org/10.1061/40932\(246\)292](http://dx.doi.org/10.1061/40932(246)292)
- Sun, Y., Salari, E., & Chou, E. (2009). Automated pavement distress detection using advanced image processing techniques. *2009 IEEE International Conference on Electro/Information Technology*, 373–377. doi:10.1109/EIT.2009.5189645
- Sun, Y., Wong, A. K. C., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(04), 687–719.
<http://doi.org/10.1142/S0218001409007326>
- Tang, J., Alelyani, S., & Liu, H. (2013). Feature Selection for Classification: A Review. *Data Classification: Algorithms and Applications*. Editor: Charu Aggarwal, CRC Press In Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. Retrieved from http://www.public.asu.edu/~jtang20/publication/feature_selection_for_classification.pdf
- Teoh, T.-T., Nguwi, Y.-Y., & Cho, S.-Y. (2009). Towards a portable intelligent facial expression recognizer. *Intelligent Decision Technologies*, 3(3), 181–191. doi:10.3233/IDT-2009-0062
- Terzi, S. (2013). Modeling for pavement roughness using the ANFIS approach. *Advances in Engineering Software*, 57, 59–64. <http://doi.org/10.1016/j.advengsoft.2012.11.013>
- Texas Department of Transportation. (2010). *Pavement Management Information System*. Austin, Texas.

- Tsai, Y.-C., Kaul, V., & Mersereau, R. M. (2010). Critical assessment of pavement distress segmentation methods. *Journal of Transportation Engineering*, 136(1), 11–19. doi:10.1061/(ASCE)TE.1943-5436.0000051
- Tsao, S., Kehtarnavaz, N., Chan, P., & Lytton, R. (1994). Image-based expert-system approach to distress detection on CRC pavement. *Journal of Transportation Engineering*, 120(1), 52–64.
- Vieira, J., Dias, F. M., & Mota, A. (2004). Neuro-Fuzzy Systems : A Survey Neuro Fuzzy Systems Types of Neuro-Fuzzy Systems. In *5th WSEAS NNA International Conference on Neural Networks and Applications*. Udine, Italia.
- Wang, K. C. P., Li, Q., & Gong, W. (2008). Wavelet-based pavement distress image edge detection with à Trous algorithm. *Transportation Research Record: Journal of the Transportation Research Board*, 2024(-1), 73–81. doi:10.3141/2024-09
- Washington State Department of Transportation. (1992). *Pavement Surface Condition Rating Manual*. Olympia, WA.
- Weiss, G. M. (2004). Mining with rarity. *ACM SIGKDD Explorations Newsletter*, 6(1), 7. <http://doi.org/10.1145/1007730.1007734>
- Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, 11(5), 545–557. [http://doi.org/10.1016/0167-9236\(94\)90024-8](http://doi.org/10.1016/0167-9236(94)90024-8)
- Wu, L., Mokhtari, S., Nazef, A., Nam, B., & Yun, H.-B. (2014). Improvement of crack detection accuracy with novel crack grouping technique for flexible pavement. *ASCE Computing in Civil Engineering*.

- Yan, M., Bo, S., Xu, K., & He, Y. (2007). Pavement crack detection and analysis for high-grade highway median-filter algorithm with 3 morphological detection edge. *In The Eighth International Conference on Electronic Measurement and Instruments (ICEMI'2007)*.
- Yun, H.-B., Mokhtari, S., & Wu, L. (2015). Crack Recognition and Segmentation Using Morphological Image-Processing Techniques for Flexible Pavement. *Journal of the Transportation Research Board*.
- Zhang, G. P. (2000). Neural networks for classification: a survey. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 30(4), 451–462.
<http://doi.org/10.1109/5326.897072>
- Zhao, Z., Morstatter, F., Sharma, S., Alelyani, S., Anand, A., & Liu, H. (2010). *Advancing feature selection research*. ASU feature selection repository.
- Zhou, G., & Wang, L. (2012). Co-location decision tree for enhancing decision-making of pavement maintenance and rehabilitation. *Transportation Research Part C: Emerging Technologies*, 21(1), 287–305. <http://doi.org/10.1016/j.trc.2011.10.007>
- Zhou, J., Huang, P. S., & Chiang, F.-P. (2003). Wavelet-aided pavement distress image processing. In *Optical Science and Technology, SPIE's 48th Annual Meeting* (pp. 728-739). International Society for Optics and Photonics.
- Zhou, J., Huang, P. S., & Chiang, F.-P. (2006). Wavelet-based pavement distress detection and evaluation. *Optical Engineering*, 45(2), 027007. doi:10.1117/1.2172917
- Zou, Q., Cao, Y., Li, Q., Mao, Q., & Wang, S. (2012). CrackTree: Automatic crack detection from pavement images. *Pattern Recognition Letters*, 33(3), 227–238.
doi:10.1016/j.patrec.2011.11.004