

This is to certify that the

dissertation entitled

APPLE SORTING USING NEURAL NETWORKS, STATISTICAL
CLASSIFIERS AND SPECTRAL REFLECTANCE IMAGING

presented by

Ismail Kavdir

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Agricultural Engineering



Major professor

Daniel E. Guyer

Date August 4, 2000

MSU is an Affirmative Action/Equal Opportunity Institution

O-12771

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
SEP 21 6 04 2006		

**APPLE SORTING USING NEURAL NETWORKS, STATISTICAL
CLASSIFIERS AND SPECTRAL REFLECTANCE IMAGING**

By

Ismail Kavdir

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Agricultural Engineering

2000

ABSTRACT

APPLE SORTING USING NEURAL NETWORKS, STATISTICAL CLASSIFIERS AND SPECTRAL REFLECTANCE IMAGING

By

Ismail Kavdir

Apple sorting was performed using artificial backpropagation neural network classifiers (single and multiple) and statistical classifiers including k-nearest neighbor, decision tree, and Bayesian. Pixel gray values and texture features obtained from the entire apple images were used as input to artificial neural network classifiers; these features were used either alone or in combination in neural network classifiers. Statistical classifiers, on the other hand, were used only with texture features.

Two types of classification experiments were applied; 2-class classification which included a defective apple group and a non-defective apple group, and 5-class classification which included all the defective and good apple groups.

In general, backpropagation neural network classifiers provided superior classifications using pixel intensity features in both categories of 2-class and 5-class classifications compared to statistical classifiers, which used texture features. Using the combination of pixel gray values from images or the combination of pixel gray values and texture features, the backpropagation neural network provided further improvements in 5-class classification. Using wavelengths beyond 1000 nm improved the identification success of the defects such as bruise, russet, and bitter pit on Golden Delicious apples and leaf roller defect on Empire apples.

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my major professor, Dr. Daniel E. Guyer, for the time, effort and guidance he devoted to me. I thank Dr. Guyer for encouraging me throughout my research. I also would like to thank Dr. Ajit K. Srivastava for his time and guidance in the first two years of my study at Michigan State University. I appreciate the research assistantships given by Michigan State University through Dr. Guyer and Dr. Srivastava.

I would like to thank Dr. George C. Stockman, my committee member from the Computer Science Department, for his insightful ideas and guidance. I would also like to thank to Dr. Randolph M. Beaudry, my committee member from the Horticulture Department, for his inputs and guidance.

I would like to thank Trakya University in Turkey for giving me the scholarship to study a PhD program at Michigan State University.

I thank Dr. Clarence Hansen, retired professor from Department of Agricultural Engineering, and his family for their moral support since we came to United States. Special thanks to Mario Fusco, Adnan Degirmencioglu, Fei Ni Tuang and Jim Shaper, my friends at MSU, for their support and encouragement.

My deep appreciation goes to my wife Yasemin who has just successfully defended her PhD thesis, for her patience and sacrifices. My special thanks to my little patient daughter “Selin” for her understanding and great supporting smiles. Finally I like to thank my parents for their insights, sacrifices and dedications for my education.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER 1. INTRODUCTION	1
1.1 Background of the Study	1
1.2 Associated Problems	4
1.3 The Significance of the Study	5
1.4 Overview of the Methodology	7
1.5 Objectives and Hypothesis	8
CHAPTER 2. LITERATURE REVIEW	10
2.1 Theoretical Background	10
2.1.1 Machine Vision	10
Texture Features	12
2.1.2 Spectral Reflectance	15
Interaction Between Light and Biological Material	15
2.1.3 Classification	19
2.1.3.1 Backpropagation Artificial Neural Network Classifier	20
Network Structure	20
The Multi-layer Neural Network	22
Training	23
Adjusting the Weights in the Output Layer	24
Adjusting the Weights in the Hidden Layer	26
Adding Bias to a Neuron	27
Learning Rate and Momentum	27
2.1.3.2 K-Nearest Neighbor Classifier	28
2.1.3.3 Decision Tree Classifier	31
2.1.3.4 Bayesian Classifier	32

2.2 Empirical Studies	35
2.2.1 Machine Vision Applications.....	35
2.2.2 Spectral Reflectance Applications.....	37
2.2.3 Applications of Artificial Neural Networks	40
2.2.4 Comparison of Statistical and Neural Network Classifiers.....	44
CHAPTER 3. MATERIALS AND METHODS.....	48
3.1 Data Collection.....	49
3.1.1 Apple Defects Used in the Study	50
3.1.1.1 Bruise	51
3.1.1.2 Puncture.....	52
3.1.1.3 Leaf Roller.....	52
3.1.1.4 Russet (Netting).....	52
3.1.1.5 Bitter Pit	53
3.2 Image Acquisition	54
3.3 Image Processing.....	56
3.3.1 Reducing the Resolution of Images	56
3.3.2 Background Segmentation	57
3.3.3 Elimination of Specular Reflectance from Images.....	58
3.4 Preparation of Image Data for Artificial Classifiers	60
3.4.1 Pixel Gray Values as Features from an Image	60
3.4.2 Extraction of Texture Features from an Image	61
3.5 Designing the Artificial Classifiers Based on the Feature Types and the Number of Output Classes	62
Feature Types.....	63
3.5.1 Classification with Two-Output Classes	64
3.5.1.1 Design of Single Backpropagation Neural Network Classifier for Classification and Wavelength Selection Using Pixel Intensity Values	64
Performance Evaluation.....	67
Wavelength Selection	69
3.5.1.2 Classification Using Texture Features from Images	69
3.5.1.2.1 Single Backpropagation Neural Network Classifier.....	70

3.5.1.2.2 Statistical Classifiers.....	70
K-Nearest Neighbor Classifier	71
Decision Tree Classifier	71
Bayesian Classifier	72
3.5.1.3 Classification Using Combined Features from Images	72
3.5.1.3.1 Using Pixel Intensities From Two Images at Different Wavelengths.....	73
3.5.1.3.2 Using Pixel Intensity and Texture Features	73
3.5.1.3.3 Combining Two Images by Averaging.....	73
3.5.2 Classification with Five Output Classes.....	74
Design of the Multiple Backpropagation Neural Network Classifier.....	76
CHAPTER 4. RESULTS AND DISCUSSION	79
4.1 Image Processing.....	79
4.1.1 Reducing the Resolution of Images	79
4.1.2 Background Segmentation	81
4.1.3 Elimination of Specular Reflectance from Images.....	86
4.2 Classification with Two-Output Classes	87
4.2.1 Classification and Wavelength Selection Using Single Backpropagation Neural Network Classifier and Pixel Intensity Values.....	87
4.2.2 Classification Using Texture Features	93
4.2.3 Classification Using Combined Features	97
4.2.3.1 Using Pixel Intensities From Two Images at Different Wavelengths	97
4.2.3.2 Using Pixel Intensities and Texture Features	98
4.2.3.3 Combining Two Images by Averaging.....	98
4.3 Classification with Five-Output Classes	99
4.3.1 Classification and Wavelength Selection Using Single Backpropagation Neural Network Classifier and Pixel Intensity Values.....	99
4.3.2 Classification Using Multiple Backpropagation Neural Network Classifier	101
4.3.3 Classification Using Texture Features	106
4.3.4 Classification Using Combined Features	107

4.3.4.1 Using Pixel Intensities From Two Images at Different Wavelengths	107
4.3.4.2 Using Pixel Intensities and Texture Features	109
4.3.4.3 Combining Two Images by Averaging.....	109
4.3.5 Statistical Tests of the Results from 5-Class Classification	110
4.4 Classification of Stem and Calyx	114
4.5 Summary and Conclusions	114
APPENDIX A: Graphs of Classification Success and Accuracy Rate versus Spectral Reflectance For Defects on Empire and Golden Delicious Apples.....	121
APPENDIX B: Confusion Matrices for the Varieties and Classification Methods Used in 5-Class Classification.....	139
BIBLIOGRAPHY	143

LIST OF TABLES

Table 3.1 Classifiers, types of input features and the number of classes used.....	63
Table 3.2 Groups of defective and good apples included in the study for each variety ...	64
Table 3.3 Codes for the output targets for each class used in training for Empire variety	66
Table 3.4 Numbers of apples in training and in testing sets used in classifications	67
Table 4.1 Classification results obtained from using images at different resolutions.....	80
Table 4.2 Training and testing times for two-class classification with neural network....	81
Table 4.3 Comparisons of classification results obtained using images with and without background at the most effective wavelengths	82
Table 4.4 Three effective wavelengths for individual surface characteristics	89
Table 4.5 The best two classifiers in two-class classification.....	95
Table 4.6 Training and testing times for five-class classification with neural network..	101
Table 4.7 Classification results of an Empire apple with Leaf Roller at 3 different wavelengths through 4 different classifiers	102
Table 4.8 Selected classification results from using all of the classifiers and features used for Empire and Golden Delicious varieties (5-Class Classification).....	104
Table 4.9 Significant test results when all the defect groups are considered as one class (Empire, $P < 0.05$)	111
Table 4.10 Significant test results when each defect group is considered separately (Empire, $P < 0.05$).....	112
Table 4.11 Significant test results when all the defect groups are considered as one class (Golden Delicious, $P < 0.05$).....	112
Table 4.12 Significant test results when each defect group is considered separately (Golden Delicious, $P < 0.05$).....	113
Table A.1 Classification results of defective apples at single and double effective wavelengths (Two-Classes).....	132
Table A.2 Classification results obtained using combined features with the single backpropagation neural network (Two-Classes).....	133
Table A.3 Classification results from two-class classification using pixel intensity and texture features at effective wavelengths (Empire).....	134

Table A.4 Classification results from two-class classification using pixel intensity and texture features at effective wavelengths (Golden Delicious)	135
Table A.5 Classification results from five-class classification using single backpropagation neural network classifier (Empire)	136
Table A.6 Classification results of five-class classification using single backpropagation neural network classifier (Golden Delicious).....	137
Table A.7 Results obtained from multiple backpropagation neural network classifier ..	138
Table B.1 Using pixel intensities in single neural network (Empire)	139
Table B.2 Using pixel intensities in multiple neural networks (Empire).....	139
Table B.3 Using pixel intensities from two images in single neural network (Empire) .	139
Table B.4 Using pixel intensity and texture features in single neural network (Empire)	139
Table B.5 Using texture features in single neural network (Empire).....	139
Table B.6 Using texture features in 1-nearest neighbor classifier (Empire)	140
Table B.7 Using texture features in 3-nearest neighbor classifier (Empire).....	140
Table B.8 Using texture features in decision tree classifier (Empire).....	140
Table B.9 Using texture features in Bayesian classifier (Empire)	140
Table B.10 Using pixel intensities in single neural network (Golden D.)	140
Table B.11 Using pixel intensities in multiple neural networks (Golden D.).....	141
Table B.12 Using pixel intensities from two images in single neural network (Golden D.)	141
Table B.13 Using pixel intensity and texture features in single neural network (Golden D.)	141
Table B.14 Using texture features in single neural network (Golden D.).....	141
Table B.15 Using texture features in 1-nearest neighbor classifier (Golden D.)	141
Table B.16 Using texture features in 3-nearest neighbor classifier (Golden D.)	142
Table B.17 Using texture features in decision tree classifier (Golden D.).....	142
Table B.18 Using texture features in Bayesian classifier (Golden D.)	142

LIST OF FIGURES

Figure 2.1 Comparison of the gray-levels in an image to form spatial dependence matrix (Tomita and Tsuji, 1990).....	14
Figure 2.2 Sample images with different texture features (normalized).....	15
Figure 2.3 A simplified schematic figure of a fruit interacting with light (Reflectance values given are for fruits in general, Chen, 1978).	16
Figure 2.4 Artificial neuron with an activation function (Wasserman, 1989)	21
Figure 2.5 Sigmoidal activation function (Wasserman, 1989).....	21
Figure 2.6 Two-layer backpropagation artificial neural network (Wasserman, 1989)	23
Figure 2.7 1-nearest neighbor classifier in the case of two-output classes	29
Figure 2.8 Modules of decision-making procedure in a k-nearest neighbor classifier	30
Figure 2.9 Modules of decision-making procedure in a decision tree classifier.....	32
Figure 2.10 Modules of decision-making procedure in a Bayesian classifier.....	35
Figure 3.1 A schematic diagram of imaging configuration.....	55
Figure 3.2 Schematic of the backpropagation neural network in classifying an apple image	65
Figure 3.3 A confusion matrix for two classes (James, 1985)	68
Figure 3.4 A schematic description of the multiple backpropagation neural networks	78
Figure 4.1 Classification of bruised apples against good apples (Empire)	83
Figure 4.2 Classification of calyx view images against non-calyx (good) images (Empire)	84
Figure 4.3 Classification of images with russet against good apple images (Golden Delicious)	85
Figure 4.4 Examples of apple images with calyx and good tissue views in Figure A.4 ...	86
Figure 4.5 Appearances of defects at different wavelengths	90
Figure 4.6 Classification of bruised apples against good apples (Golden Delicious).....	94
Figure A.1 Classification of bruised apples against good apples (Empire)	121
Figure A.2 Classification of punctured apples against good apples (Empire)	122
Figure A.3 Classification of apples with leaf roller against good apples (Empire)	123

Figure A.4 Classification of calyx view images against non-calyx (good) apple images (Empire)	124
Figure A.5 Classification of apples with stem against good apples (Empire)	125
Figure A.6 Classification of apples considering all the defects and natural features against good apples (5 Classes, Empire)	126
Figure A.7 Classification of apples with bitter pit against good apples (Golden Delicious)	127
Figure A.8 Classification of calyx view images against non-calyx (good apple) images (Golden Delicious)	128
Figure A.9 Classification of stem view images against non-stem (good apple) images (Golden Delicious)	129
Figure A.10 Classification of apples with russet against good apples (Golden Delicious)	130
Figure A.11 Classification of apples considering all the defects and natural features against good apples (5 Classes, Golden Delicious)	131

Chapter 1. INTRODUCTION

This dissertation reports a study of apple sorting based on surface quality using spectral imaging, machine vision and classifiers including backpropagation neural networks, Bayesian, k-nearest neighbor and decision tree. Chapter one of the dissertation presents the background of the study, specifies the problem, describes its significance, and presents an overview of the methodology used. In Chapter 2, theoretical and empirical backgrounds of the methodology applied are reviewed. Chapter 3 presents the application details of the methodology used and establishes an introduction to the results. Finally, Chapter 4 presents the results and discussion.

1.1 Background of the Study

Apples are susceptible to defects by the nature of their variety, growing conditions and operations applied to them during growing, harvesting and post harvest handling. Defects, which develop during the growing season, are caused by bacteria, fungus, larvae, virus, insect sting, weather conditions (dry, cold, intense sun light, rain, frost, hail etc) and competition between fruits and leaves for water and nutrients. Harvesting and post harvest handling defects are usually caused by storage conditions and the impacts to which apples are exposed during operations of loading, unloading, washing, dewatering, waxing, drying and transporting on the inspection belts. Consequently, in their final destination after harvest, apples may have different types of defects caused by one or more of the mostly external effects. Having different defects coming from different causes makes automated defect detection and classification

challenging. This is because the severity, depth and form of a defect may vary drastically depending on the cause.

Defect detection was done manually in early fruit sorting applications by the help of people trained on the standards of the quality grades of apples. Later, technical advancements, especially in the areas of machine vision and spectral imaging, helped the automation of the grading process of fruits according to shape, size, color and recently blemishes.

Machine vision was widely applied in many sectors of industry, especially in electronic and automotive, before being applied in agricultural areas. It is primarily used for inspection and quality control. Application of machine vision has been increasing rapidly in recent years. The following are primary steps followed in image processing; image generation, product location, scrutiny, measurement of features and classification (Yang, 1992).

Product location and scrutiny are two important steps in image processing. First, the product, which is under investigation, is separated from the background and from other objects, if there are any in the scene. Second, all pixels in the segmented parts are examined to find features. If it is necessary, images are preprocessed to enhance contrast between features, to remove noise or to eliminate problems caused by lighting.

Spectral reflectance imaging originated from the areas of chemistry and remote sensing (Muir et al., 1989). Ability of light to detect the elements of a matter by activating molecules of the related element at a specific wavelength has drawn attractions from many areas including agriculture. Activation levels of the molecules depend on the energy of the incoming light. For instance, while long wavelength radiations such as

radio and microwaves excite gases, shorter wavelengths affect liquids and solids.

Absorption of light takes place in the visible and ultraviolet as the electrons reach higher energy levels. On the other hand, molecular absorption in the infrared region develops by means of vibrations and rotations of molecules. Although rotational absorption bands are mainly in the far infrared, vibrational absorption bands are in the near infrared and have been used extensively in analysis of food and agricultural products.

After interacting with the absorbing molecules, light is re-emitted from the surface of the material. This re-emitting happens in the form of transmission or reflection. If the reflected light is measured, the information it carries about the structure of the elements from which it was reflected can be obtained. At this point machine vision plays an important role from acquiring the information coming from the sample examined, to processing it.

Another development that led the studies of automatic sorting of agricultural produce has been in artificial neural networks. Artificial neural networks have been widely used in many areas, including agriculture, since they regained their popularity in the early 1980s. They provided a massive parallelism, learning and generalization ability, representation and computation power and fault tolerance to the solution of automatic sorting problems.

Detecting blemishes automatically relies on machine vision, the relation between the light and the matter, and an artificial classifier that will be robust and fault tolerant enough to interpret the linear or non-linear relation between the input data and targeted output.

1.2 Associated Problems

Some challenging problems in automating the sorting process of fruits can be expressed as follows:

In a two dimensional image, blemishes such as punctures and bruises can easily be confused with natural features of apples such as calyx and stem, with natural color variations over a product surface or with the pixels at the edges of the apple image, which are usually darker in color compared to the central area of the image. This situation may sometimes require additional feature extraction algorithms such as using structured lighting (Yang, 1993 and Campins et al., 1997) to identify stem and calyx. This approach requires using different algorithms and acquiring multiple images to differentiate the calyx or stem from the blemishes.

Another problem in image processing for fruit sorting is that different types of defects may exist on the surface of the object being classified and it is usually difficult to predict their locations, orientations, sizes and shapes. Also, the contrast between a defective tissue, such as a bruised area, and a good tissue is often very low making it difficult to develop an algorithm to segment the defective area.

Consistent orientation of apples is a difficult task considering shape and size variations within and between cultivars. It is reasonable to expect that the materials handling operation of a sorting line would present the apple to the detection operation in any orientation. Therefore, a robust image processing and classification algorithm is required to discriminate stem or calyx from any other defect without being dependent on the orientation.

1.3 The Significance of the Study

Quality classification of fruits is an important procedure in marketing and processing. Efforts into more capable, efficient and accurate automated fruit classification systems continue as industry priorities as manual fruit grading has the drawbacks such as subjectivity, tediousness, cost, availability, and inconsistency.

However, applying automatic sorting or quality control in agriculture is not as straightforward as applying them in other industries such as the electronic and automotive. There are two main differences. First, the working agricultural environment is highly variable (weather, soil, etc). Second, agricultural produce are highly variable due to their inherent morphological diversity (Blackmore and Steinhouse, 1993).

Automatic sorting of some types of produce according to color, size and shape has been commercially possible. More work is needed for sorting automatically according to blemishes to improve the classification success and to develop a more generalized algorithm that is capable of recognizing many types of defects and natural features on a fruit surface. Automating defect detection is a key step to the full automation of the sorting procedure.

Research to date has focused usually on using one type of feature in the classification of agricultural commodities. They either used only texture, shape features or pixel gray values in the spectral images. However, using all of the features available such as, texture, shape features and pixel gray values in images at different wavelengths together could help to increase the classification success, as was suggested by Campins et al. (1997).

Spectral reflectance has potential to bring out the specific information from a defect on a fruit. The effective wavelength for highlighting a defect from its surroundings may vary with the defect being investigated. Therefore, use of spectral reflectance for defect detection can be very useful. Use of multiple wavelengths from different regions of the spectrum in the existence of multiple defects on the fruit surface could bring different and complementary information together benefiting the classification system with improved accuracy.

Shapes of defects can be categorized as, for instance, round or long etc. although they are not consistent. While most of the bruises tend to be round, cuts and punctures tend to be long. On the other hand, russet (netting) does not have a consistent shape. Shapes of calyx and stem areas are unique, although they may not preserve their uniqueness in different orientations.

In conventional methods of automatic fruit grading, determining the shape of a defective area would require segmentation of the defect, which may often be difficult due to low contrast between the blemish and good tissue. In this situation where feature extraction is applied in image processing and the blemishes are segmented, shape features of defects or calyx-stem area would help in recognizing the objects on the apple surface together with other features. However in this study, as the gray values in all of the pixels of an image were input in a backpropagation neural network classifier, shape features belonging to the defects or stem-calyx were expected to be seen and extracted by this classifier, eliminating most of the feature extraction or preprocessing beforehand.

Texture may also be an important feature in discriminating different parts of an apple surface. Throop and Aneshansley (1993) used texture features to recognize apple

characteristics. In texture calculations, distribution of the pixels in the defective area is examined and the spatial relation between the pixels is extracted. There is no requirement for defect segmentation to calculate the textural features, as the entire image, i.e. all pixels is used in the calculations.

Nevertheless, extracting and using textural information requires extra time and effort. Thus, if the classification success and error rate are within error limits by using only the gray value information from the pixels at effective wavelengths, texture features may be eliminated from the list of features. Classifying an apple image by means of an artificial neural network using only pixel gray values is much faster than extracting the texture features and then using them in classification.

1.4 Overview of the Methodology

Brief information on the methodology used in the experiment is given here. Theoretical detail on methodology is given in Chapter 2, while the information on the application of the methodology is presented in Chapter 3 of the dissertation.

Images were acquired using a black and white vidicon camera, a lighting system connected to a monochromator that adjusts the light according to wavelength, a personal computer and software.

Two types of features were used in the experiment, pixel gray-level values and texture from the same image at the same or different wavelengths. Forty wavelength bands from 540 nm to 1320 nm with increments of 20 nm were examined to find the effective wavelength(s) for a specific defect. A backpropagation artificial neural network classifier was used to select the effective wavelength(s) among all of the 40 wavelength

bands examined. Images of good apples and defective apples taken at the effective wavelength(s) were used in two classification applications; 2-class classification where only the good apple group and one defective apple group were used and 5-class classification where the good group including stem and calyx, and each of the defective groups were included in the classification. Pixel gray values and texture features of angular second moment, correlation and contrast calculated from the image were input in the classifier either alone or combined depending on the classifier used.

In this research, machine vision, spectral imaging and backpropagation neural networks were used together to separate defective apples from good ones. In addition to neural network classifier, statistical classifiers, such as Bayesian, k-nearest neighbor and decision tree were also studied. In the case of 5-class classification, another artificial neural network application was used in addition to the one mentioned above. This was a combination of the individual backpropagation neural networks used in each good-defect classification. Two types of information, pixel gray values and texture features, from all pixels in the apple images were used as features in the classification applications.

1.5 Objectives and Hypothesis

The overall objective of this study was to develop a classification system for automatic nondestructive apple sorting according to surface quality conditions using machine vision, spectral imaging and artificial classifiers. To obtain this overall target, the following sub-objectives were established;

- 1) Determine effective wavelength(s) and classifier for separating good tissue from defective tissue.

- 2) Develop a classification system with backpropagation neural network classifier, which will accept the pixels in an entire apple image as features eliminating feature extraction tasks.
- 3) Evaluate the use of texture features that would enhance the classification success.
- 4) Evaluate image resolution in terms of classification speed and accuracy.
- 5) Minimize the image enhancement and processing operations.
- 6) Automatically recognize and differentiate the natural features of apples such as, calyx and stem.

The hypothesis of this research was that multiple image-based features such as surface gray level values and texture features embedded in an image could be measured and used either separately or combined in one decision process using the techniques of multiple spectral imaging, machine vision and classifiers such as neural networks and statistical to have an enhanced classification success. The aim here is to eliminate defect segmentation in apple sorting; the task of detecting the blemishes or stem-calyx is done automatically by the artificial classifiers. In this sense, application of automated apple sorting here would look like apple sorting performed by human sorters as the artificial classifiers have the ability to tune themselves according to kind and severity of the defects.

Chapter 2. LITERATURE REVIEW

Chapter 2 of the dissertation presents information on literature review. There are two main sections; the first one explains the theoretical background of the methods used throughout the research. The second part of the literature review includes information on some practical applications of the methodology and their relation to current study.

2.1 Theoretical Background

In this section, theoretical information is given on the areas of machine vision, spectral reflectance, artificial neural networks and statistical classifiers that were used in this study. Methodology explained here is closely related with application details of the experiment presented in Chapter 3.

2.1.1 Machine Vision

A machine vision system consists of a camera, a computer that has a plug-in image processing card, a display monitor and a lighting system (Figure 3.1). The camera generates the image in the form of analog signals. The analog signal is transferred to the card where it is converted into digital form. Once it is digitized, an image can be stored, processed, displayed and interpreted through computer programming (Yang, 1992).

Digital image processing is a very important part of machine vision. Digital image processing refers to processing a two dimensional picture by a digital computer. A digital image can be expressed as a function $f(x, y)$ in two dimensions. The smallest unit of a

digitized image is the picture element (pixel), which can have gray level between 0 and 255.

Following are the main interest areas for image processing tasks:

- 1) Image digitization and coding for the optimization of storage and transmission capabilities,
- 2) Image enhancement and restoration for improving the quality of images,
- 3) Image analysis including feature extracting, segmentation and classification operations on images for automated machine vision applications (Graf, 1982).

Noise and distortion may occur in images in the processes of image acquisition and transmission. Thus, image enhancement operations are applied to images to remove noise, to correct distortions, and to minimize the effects of errors such as motion blur or similar errors during image acquisition (Jahne, 1995).

Many filtering methods have been used for image enhancement. Spatial averaging is one of the important image filtering operations. In this filtering method, each pixel is replaced by a weighted average of its neighborhood pixels (Jain, 1989). Similar to spatial averaging, image resolution can be reduced by averaging the pixels in the input image using the equation below:

$$x(i, j) = \frac{1}{N} \sum_i^n \sum_j^n y(i, j) \quad (2.1)$$

where $y(i, j)$ and $x(i, j)$ are input and output images respectively and n is the size of the window used for averaging (filtering). N is the number of pixels in the selected window.

In reducing resolution, gray values in the pixels of the reduced sized image are replaced by the average of the gray values of the pixels in the selected window in the

original image. The amount of reduction in the size of the image is determined by the size of the window selected for averaging.

Image analysis maintains the tools to extract useful information from images. In this process, objects are first located within an image. Then, features characterizing the objects are calculated and finally the objects in the images are classified based on the features extracted. Texture, explained below, is an important part of image analysis as it can provide an image recognition system with the required information about the object in the scene.

Texture Features

Texture has been used for the analysis of different types of images, such as microscopic, aerial and satellite. Textural features represent the spatial distribution of tonal variations in an image at various wavelengths such as the visible and infrared portions of the spectrum. Texture and tone are two concepts that are closely related. While tone is related with the varying shades of the gray, texture deals with the spatial distribution of gray tones in the image (Haralick et al., 1973).

In processing images, the pictorial information is represented as a function of two variables (x, y) . An image is stored as a two-dimensional array in its digital form. If $K_x = (1, 2, \dots, N_x)$ and $K_y = (1, 2, \dots, N_y)$ are the X and Y spatial domains, then $K_x \times K_y$ is the set of resolution cells and the digital image I is a function which assigns some gray-tone value $G = \{1, 2, \dots, N_g\}$ to each and every resolution cell (Haralick et al., 1973):

$$I : K_x \times K_y \rightarrow G . \quad (2.2)$$

Methodology of spatial gray-level dependence matrices was used by Haralick et al. (1973) to classify aerial photographs and satellite images. In this method, the reoccurrence probability of a pixel, which has a particular gray-level at a determined distance and orientation from any given pixel that has the same gray level, is calculated.

Function $F(i, j, d, \theta)$ forms the spatial gray-level dependence matrices; where i and j are the (x, y) coordinates of the matrix, d is the distance, and θ is the orientation between two pixels that have the identical gray-level value (Figure 2.1-a). Distance d may have any value (Figure 2.1-b). On the other hand, θ can have the values of 0° , 45° , 90° , 135° etc. (Tomita and Tsuji, 1990).

Spatial dependence matrices are used to calculate statistical texture features such as Angular Second Moment (ASM), Contrast (C) and Correlation (CR) as follows:

$$ASM = \sum_i \sum_j \{F(i, j)\}^2 \quad (2.3)$$

$$C = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} F(i, j) \right\} \quad (2.4)$$

$$CR = \frac{\sum_i \sum_j (i, j) F(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (2.5)$$

where

$$\mu_x = \sum_{i=1}^{N_g} i \sum_{j=1}^{N_g} F(i, j) \quad (2.6)$$

$$\mu_y = \sum_{j=1}^{N_g} j \sum_{i=1}^{N_g} F(i, j) \quad (2.7)$$

$$\sigma_x = \sum_{i=1}^{N_g} (i - \mu_x)^2 \sum_{j=1}^{N_g} F(i, j) \quad (2.8)$$

$$\sigma_y = \sum_{j=1}^{N_g} (j - \mu_y)^2 \sum_{i=1}^{N_g} F(i, j). \quad (2.9)$$

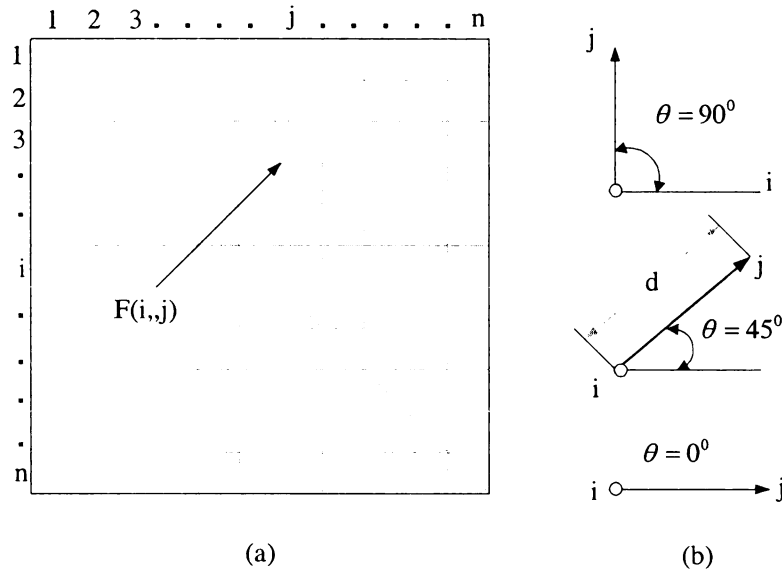


Figure 2.1 Comparison of the gray-levels in an image to form spatial dependence matrix (Tomita and Tsuji, 1990)

Angular second moment is a measure of homogeneity in the image. For a homogeneous image with minimal number of entries of large magnitude in the co-occurrence matrix, the angular second moment will be large. The contrast measures the amount of local variations in the image; an increase in the occurrences of similar intensities in the spatial dependence matrix shows a low contrast, while the frequency of different intensities occurring together is a sign of increased contrast. The correlation on the other hand shows the existence of linearity in the image. The correlation of images with large areas of similar gray values will be higher than the correlation of images with high variance in gray values (Shearer and Holmes, 1990).

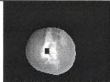
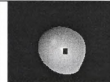
Texture Features	Good Apple	Leaf Roller	Bitter Pit
(at 1320 nm)			
Angular Second Moment	0.83	0.03	0.15
Contrast	0.00	1.00	0.37
Correlation	1.00	0.01	0.01

Figure 2.2 Sample images with different texture features (normalized)

Three sample images with their three texture features are presented in Figure 2.2. Magnitudes of the features, which were normalized, are in agreement with the definition of each feature. For instance, the good apple has the most homogeneous texture so it has the highest angular second moment, while the apple with leaf roller has the minimum homogeneity. The apple with bitter pit on the other hand is in between the other two apples in each texture measurement.

2.1.2 Spectral Reflectance

Optical reflectance can be used to evaluate tissue characteristics near the surface of a biological material. Some such characteristics are color, surface blemishes, and separation of undesired materials (Chen, 1978). Examples of studies on blemish detection using spectral reflectance can be seen in the empirical studies section later in this Chapter.

Interaction Between Light and Biological Material

The interaction between light and biological material is a complex physical phenomenon. To simplify the complexity of molecular absorption of light, it is assumed

that molecules vibrate only at fixed frequencies when interacting with light and thus, absorb only light of that particular frequency or wavelength (Muir et al., 1989).

When a light beam is projected on an object, part of the incident beam is reflected by the surface and the rest is transmitted into the object. There are three possible pathways for the transmitted light into the object; it is absorbed, reflected back to the surface, or transmitted through the object. Radiation absorbed during the interactions between light and matter may be transformed into another form of radiation, such as fluorescence and delayed-light emission (light emitted from the sample after removing the light source). The amount of radiant energy in the forms of reflectance, transmittance, absorption, or emission may vary depending on the properties of the biological material or incoming radiation (Chen, 1978). Light entering such materials is scattered in various directions as elements in most biological materials are optically non-homogeneous (Figure 2.3).

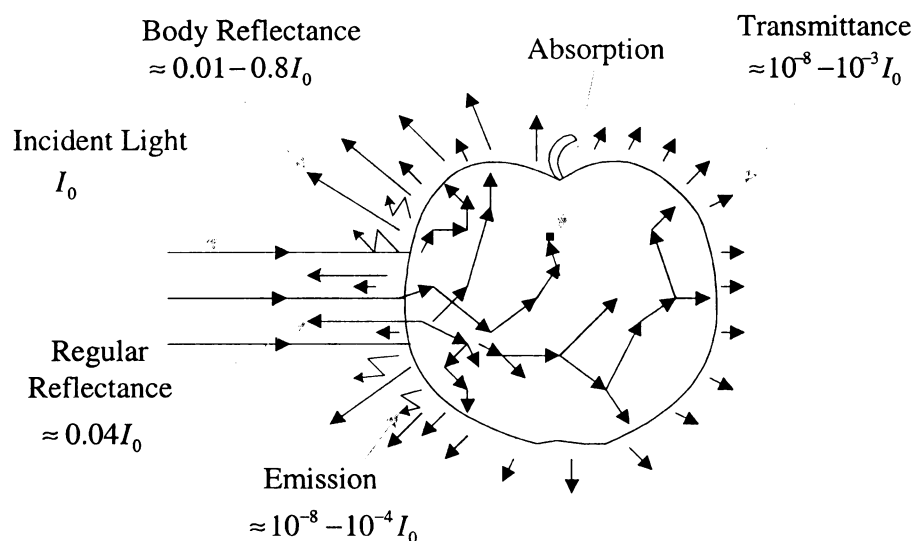


Figure 2.3 A simplified schematic figure of a fruit interacting with light (Reflectance values given are for fruits in general, Chen, 1978).

As an average over all fruit, about 4% of the incident radiation is reflected back (regular reflectance) after the first contact between the light beam and a fruit. The rest of the light entering the object encounters small particles inside the cellular structure and scatters in all directions. Much of the light returns back to the surface where the light beam initially enters the fruit and leaves the fruit from there. This type of reflection is termed body reflectance. The remaining dispersed light continues its travel through the fruit; some of the light may finally exit from the fruit at a distance from the point of incidence. On the other hand, some light is absorbed by the molecules in various elements in the fruit. The amount of absorption varies with the type of the element, wavelength, and the distance the light travels (Chen, 1978).

Body reflectance may contain information on the characteristics of the region near the surface of the point of incidence. If the reflected light is measured, quality information on the subject such as color and surface defects can be obtained. Also, information on the internal quality of the fruit may be obtained (Chen, 1978).

Two factors affect the intensity of light coming from an object; its optical density and the distance between the light source and the object. For most produce, intensity from reflectance, which is 1-80% of the incident energy depending on the absorption, is much higher than transmittance or emission. Therefore, reflectance is more commonly used in quality evaluations.

Spectral reflectance is usually measured by comparing the reflectance from an object with the reflection from a reference surface. Absorption in biological materials is generally expressed using optical density, which is expressed as follows

$$\text{Optical Density} = \log_{10}(E_1/E_2) \quad (2.10)$$

where E_1 is the incident radiant energy and,

E_2 is the radiant energy transmitted through the object.

Due to the complexity in the relations between light and biological material, it is difficult to formulize the interactions such as absorption, transmittance, and emission.

However, transmittance (T) and absorption (A) can be roughly expressed as

$$T = \frac{E_2}{E_1} \quad (2.11)$$

and absorbency as,

$$A = \log_{10}(1/T) = \log_{10}(E_1/E_2) \quad (2.12)$$

which is the same with optical density. Absorbency includes the losses from absorption, reflection and scatter of light (Chen, 1978).

Spectral reflectance properties of agricultural produce can be determined by measuring the radiant energy from the samples over a range of wavelengths. Different spectral reflectance responses are obtained from different biological materials as the constituents of materials vary depending on the variety, kind and conditions. For instance, color pigments and water absorb light in specific wavelength regions. Depending on the variety, other elements of produce may affect light scattering characteristics resulting in changes in the path length of the light and in the amount of light absorbed.

In the measurement system of the optical properties of produce, three components are critical; light source, light dispersing tool (monochrometer) and the detector of the reflected radiant light. In measuring the targeted radiation, such as reflectance or transmittance within a known wavelength band, the light source is assumed to be

powerful enough to emit the adequate energy. Also, the sensitivity of the detector, for example a camera in spectral reflectance imaging, must be compatible with the intensity and the wavelength range of reflected radiation. Obtaining light within a desired wavelength can be accomplished by using a prism or a grating to disperse a light beam into a spectrum and over a narrow slit to pass only the desired wavelength band.

2.1.3 Classification

Classification accuracy obtained from a classification algorithm that is parametric or non-parametric is usually dependent on the data. Parametric classifiers such as Bayesian usually result in lower classification accuracy when the number of features increases, as accuracies in estimating the parameters decreases with increasing number of features (Raudys and Jain, 1991). In addition to the low estimation rates, non-linearity between the input patterns and the output classes increases with high numbers of features. However, statistical classifiers based on parameter estimation perform as well as the non-parametric classifiers when the number of features is low.

In many pattern recognition problems, the pattern classes are multi-modal. So, it is difficult to use conventional partitioning applications such as Bayesian Rule (in Bayesian classifier) for classifier design. In these situations, the non-parametric partitioning of the feature space is usually preferred (Sethi and Sarvarayudu, 1982). Classifiers such as, backpropagation artificial neural network, k-nearest neighbor and decision tree are nonparametric classifiers.

2.1.3.1 Backpropagation Artificial Neural Network Classifier

Artificial neural networks, which were originally inspired by biological neural networks, are parallel-distributed massive computing systems consisting of a large number of simple highly interconnected processing units to process the information through its travel from the input nodes to the output nodes.

Backpropagation learning rule is commonly used in multilayer neural networks. In this learning rule, the network is presented with pairs of input and target patterns. Weights between the simple processing units of neurons are adjusted by iterating input patterns throughout the network until the error between the network output and the targeted output is minimized. Details of the backpropagation learning rule are given below.

Network Structure

The smallest processing unit of a backpropagation neural network is the neuron that is shown in Figure 2.4. Two operations take place in these neurons during learning; first, a set of inputs coming either from outside or a previous layer are multiplied with their assigned weights and all the results are summed. The result from this calculation can be named as *SUM*:

$$SUM = \sum_{i=1}^n x_i w_i \quad (2.13)$$

where x_i is the i^{th} input to the neuron and w_i is the weight associated with it. Calculation of *SUM* is repeated for each neuron in the layer. Second, *SUM* calculated in the first step, is entered in a modifying function to produce the final output *OUT* from the neuron (Wasserman, 1989) as follows:

$$OUT = \frac{1}{1 + \exp^{(-SUM)}} \quad (2.14)$$

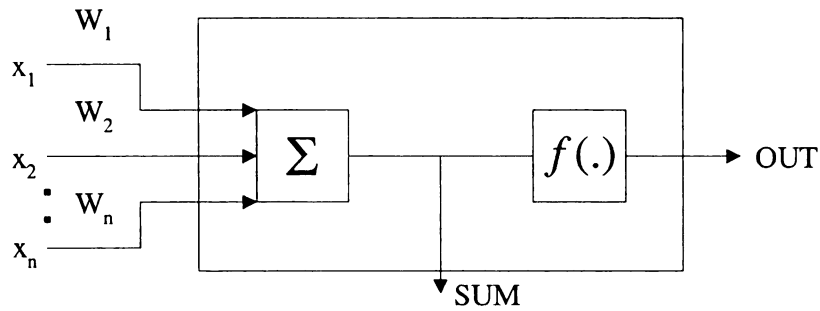


Figure 2.4 Artificial neuron with an activation function (Wasserman, 1989)

The most common function used in the neurons of a backpropagation network is sigmoidal activation given in Equation 2.14 and shown in Figure 2.5. Using this function in backpropagation neural networks is advantageous as it has a simple derivative as given in Equation 2.15:

$$\frac{\partial OUT}{\partial SUM} = OUT(1 - OUT). \quad (2.15)$$

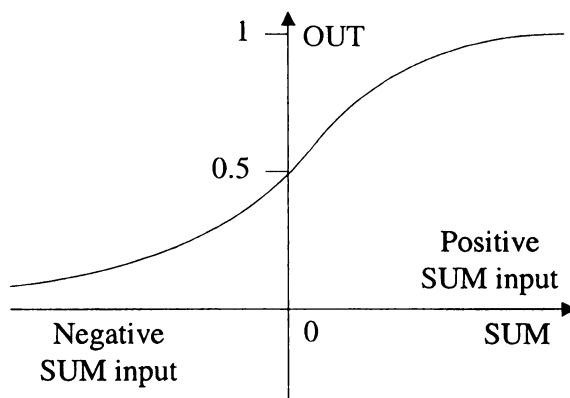


Figure 2.5 Sigmoidal activation function (Wasserman, 1989)

The sigmoidal function is continuous and smooth. Thus, it is possible to differentiate the sigmoid function in any step of the backpropagation-learning algorithm. After calculating the error rate at the end of a forward pass during learning, updating the weights by propagating backward in the network requires using the differentiated form of the sigmoid function as explained below. This function outputs between two boundaries (for example; $[0,1]$, $[-1,1]$) enabling normalization. Another advantage of using the sigmoid function is that it increases the representation power of a multi-layer neural network by introducing the nonlinearity into the system.

The Multi-Layer Neural Network

A typical multi-layer neural network that uses a backpropagation learning rule is shown in Figure 2.6. The first layer that symbolizes the inputs coming from outside is called the input layer. The number of the nodes in the input layer is equal to the number of features that are available to use. No mathematical operation takes place in this layer. Next, inputs from neurons in the input layer and the weights that connect the input nodes and the hidden layer are multiplied and summed forming the *SUM* explained above. Finally, *OUT* given in Equation 2.14 is calculated in each neuron in the hidden layer as an output signal. Then, output (*OUT*) from each neuron in the hidden layer becomes an input to the output layer after being multiplied by the weights assigned previously. The number of layers is not limited in backpropagation neural networks. However, the more layers used in the network, the more time is needed in learning and consequently in testing. The flow of the information, i.e. calculations from one layer to the other, is the same for any multi-layer network (Wasserman, 1989).

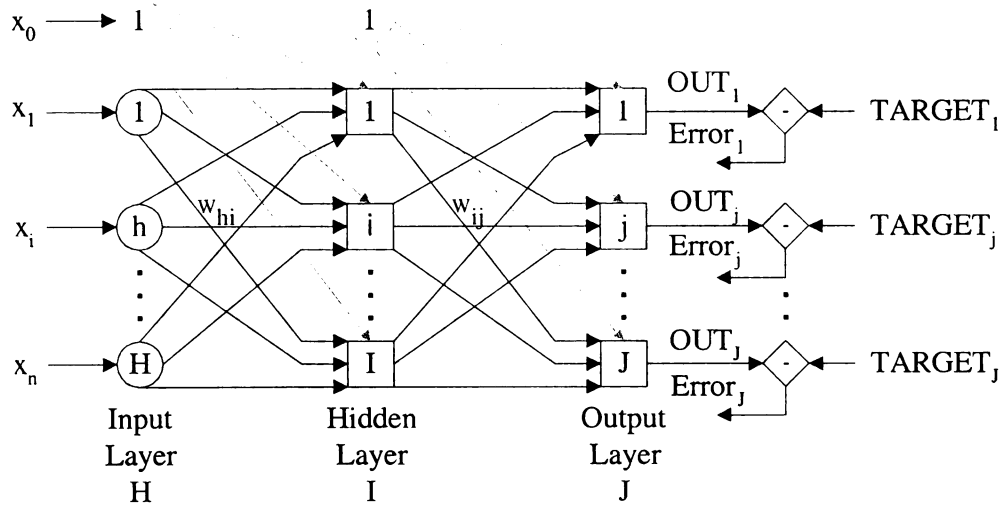


Figure 2.6 Two-layer backpropagation artificial neural network (Wasserman, 1989)

Training

Training is aimed at adjusting the weights throughout the network by processing a set of training data to obtain the desired set of outputs. This type of training is called supervised learning, which means that there is a targeted output to approach for each training sample to learn.

All weights in the network are randomly initialized to small numbers to prevent the saturation of the network. Otherwise the network may not learn.

The following steps are taken in training the backpropagation network:

- 1) Input the selected training sample and its associated target, which in supervised learning tells which class the particular training data belongs, in the network and start the process, assuming the initial weights are already selected and ready to use.

- 2) Proceed with calculations throughout the network and determine the final output from the network.
- 3) Calculate the error between the output from the network and the targeted output.
- 4) Using the error calculated in the previous step as feedback, adjust the weights in the network to minimize the error.
- 5) Repeat steps 1 through 4 until obtaining the desired error rate (Wasserman, 1989).

In Figure 2.6, outputs of the neurons in the hidden layer I are calculated first and then these outputs are used as inputs to the output layer J . Outputs from the layer J are the outputs of the entire network. These calculations refer to the steps 1 and 2 explained above, which take place in the forwarding pass in the network. Next, error rate is calculated and then weights are adjusted as mentioned in steps 3 and 4, which represent the backward pass. When the targeted error rate is reached, training the network is ended and the weights representing the memory of the network are stored. Later, this weight set obtained from training is used in classification of the testing data. The training set should represent the population well enough to develop an efficient classifier for maximum classification performance.

Adjusting the Weights in the Output Layer

In Figure 2.6, if a training process for a single weight from neuron i in the hidden layer I to neuron j in the output layer J is considered, the output of neuron j in the output

layer J is subtracted from its assigned target value to find the error. Error signal at the j^{th} neuron of the output layer for n^{th} training pattern is defined by

$$e_j(n) = Target_j(n) - OUT_j(n) \quad (2.16)$$

where $Target_j$ is the real target assigned to the training sample before the learning process has started and OUT_j is the calculated output.

The energy function for the entire output layer for the n^{th} training pattern, on the other hand, is obtained as follows:

$$E(n) = \frac{1}{2} \sum_{j=1}^J e_j^2(n) \quad (2.17)$$

where j is the number of the nodes in the output layer. Energy function, E , of the network, which is a function of error signals, is minimized over a training set to adjust the weights in the network (Hassoun, 1995):

$$\delta_{j\mu} = -\frac{\delta E(n)}{\delta W_{ij}} = \frac{\delta(1/2 \sum_{j=1}^J e_j^2(n))}{\delta OUT_j / \delta SUM} \quad (2.18)$$

where, $\delta_{j\mu}$ is the local gradient that minimizes the energy function over the weights between output and hidden layers. The same expression can simply be written as follows:

$$\delta_{j\mu} = OUT_j(1 - OUT_j)(Target_j - OUT_j). \quad (2.19)$$

Using Equation 2.19, weights of the network can be adjusted with the equations below;

$$\Delta W_{ij} = \eta \delta_{j\mu} OUT_{i\mu} \quad (2.20)$$

$$W_{ij}(n+1) = W_{ij}(n) + \Delta W_{ij} \quad (2.21)$$

where,

η is a learning rate coefficient and its value is usually between 0 and 1,

$W_{ij}(n)$ is the weight from neuron i in the hidden layer to the neuron j in the output layer at n^{th} iteration before any adjustment,

$W_{ij}(n+1)$ is value of the weight at the next iteration $(n+1)$, after adjustment,

OUT_{ii} is the output from neuron i in the hidden layer I .

Adjusting the Weights in the Hidden Layer

Equations 2.20 and 2.21 are also used to adjust the weights for the hidden layer.

However, local gradient (δ_{ii}) needed for neuron i in the hidden layer is calculated by summing all the products between weights and the local gradient (δ_{jj}) values that belong to the neurons in the output layer, as there is no target for the hidden layer and by multiplying this sum with the derivative of the sigmoid function including the output from neuron i in the hidden layer as given below:

$$\delta_{ii} = OUT_{ii} (1 - OUT_{ii}) \left(\sum_j \delta_{jj} W_{ij} \right). \quad (2.22)$$

After finding the δ_{ii} for neuron i in the hidden layer using the Equation 2.22, weights coming to the hidden layer from the input layer can be adjusted using Equations 2.20 and 2.21 explained above. The two backward-pass procedures explained so far to adjust the weights connecting hidden and output layers and input and hidden layers completes the weight adjustments of a two-layer backpropagation neural network.

Adding Bias to a Neuron

A bias with a trainable weight is added to each neuron in hidden and output layers to speed the process of learning. A bias that is shown in Figure 2.6 as $x_0 = 1$ for hidden and output layer nodes adjusts the origin of the sigmoid function to speed the convergence of training (Wasserman, 1989).

Learning Rate and Momentum

Two parameters, learning rate (η) and momentum (α) are used in the training process to rapidly and efficiently adjust the weights. Learning rate regulates the amount of change in weight adjustment. A high learning rate requires less number of epochs for training. However, there is a possibility that using high learning rate may cause the weights to oscillate between positive and negative values without reaching optimality. On the other hand, momentum allows using a high learning rate with controlled oscillation in the weights. Momentum coefficients determine the direction of the movement in the error space maintaining a rapid convergence during training and a stable learning based on the previous weight change. Addition of the momentum coefficient modifies the weight adjustment equations for the network as follows (Orchard and Phillips, 1991):

$$\Delta W_{ij}(n+1) = \eta(\delta_j OUT_{ii}) + \alpha[\Delta W_{ij}(n)] \quad (2.23)$$

$$W_{ij}(n+1) = W_{ij}(n) + \Delta W_{ij}(n+1). \quad (2.24)$$

Application of momentum in adjusting the weights throughout the network helps the energy function, which is actually the mean square of the error signals, to reach the globally optimum solution.

Despite all the information about neural networks, there is not a general model of them to use for any type of classification (Timmermans and Hulzebosch, 1995). So, to find the right neural network for a specific sorting problem, one should try different settings such as the number of neurons in hidden layer(s), the number of hidden layers, values of learning rate and momentum, or the type of transfer function in the neurons of hidden and output layers.

Learning from input-output relationships is the most important feature of artificial neural networks. However, when new patterns (features) are added to the training set, artificial neural networks need to be retrained. A drawback of artificial neural networks is that how it learns is not clearly understood. Sometimes long training periods can also be considered as a drawback.

2.1.3.2 K-Nearest Neighbor Classifier

K-nearest neighbor is a nonparametric classifier that does not assume the form of conditional density of a class. It directly assigns a test pattern to one of the classes. To be able to classify a given test pattern, all the training data is used (Young and Calvert, 1974). Thus, the rules for k-nearest neighbor classifiers are data driven. For instance, to apply 1-nearest neighbor decision rule, all distances between the given test pattern and all available training data must be measured (Figure 2.7). Then, the test pattern is assigned to the class that has the member with minimum distance from the pattern being tested.

The only parameter that should be determined is “ K ”, the number of the nearest neighbors to consider in calculations. The value of K depends on the number of training

data (n); so, K is a function of n , $K \propto f(n)$. With a larger number of samples, larger numbers of K can be chosen.

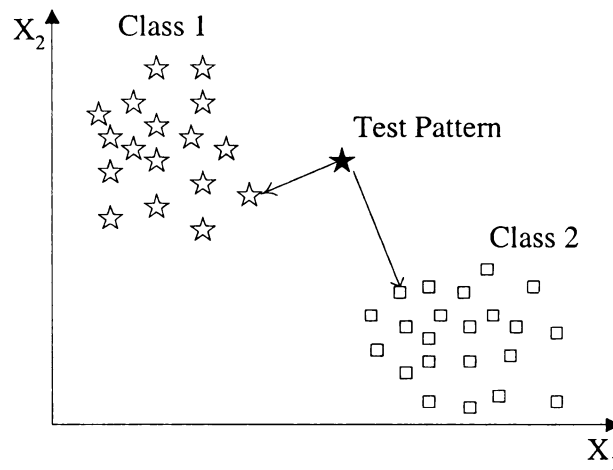


Figure 2.7 1-nearest neighbor classifier in the case of two-output classes

The best way to find the optimum value for K is to classify testing data using different k -nearest neighbor classifiers and compare the results in terms of classification success.

Two of the distance metrics that can be used in a k -nearest neighbor classifier are:

Euclidean Distance;

$$d(i, k) = \left[(x_i - x_k)^T (x_i - x_k) \right]^{1/2} \quad (2.25)$$

and Mahalanobis Distance,

$$d(i, k) = (x_i - x_k)^T S^{-1} (x_i - x_k) \quad (2.26)$$

where S is the covariance matrix of the features and, i and k are the two points distance between which is measured.

Basically, there are two modules in this classifier; testing set and training set.

Distance measurement can be considered as a third module, and also when $K > 1$, finding

the class that has the majority of members closest to the pattern being tested may be considered as a fourth module (Figure 2.8). The test pattern is assigned the label of the class most frequently represented in the k-nearest neighbor training samples when $K > 1$.

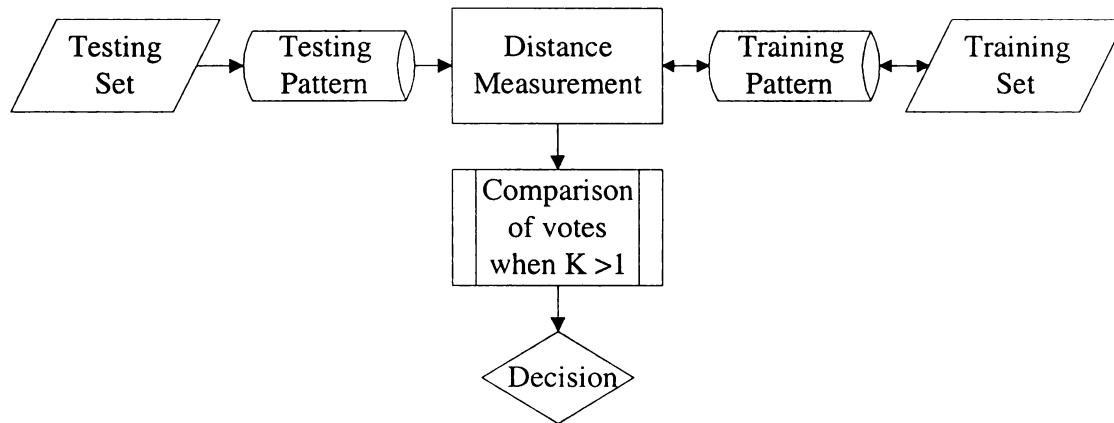


Figure 2.8 Modules of decision-making procedure in a k-nearest neighbor classifier

No training is applied in a k-nearest neighbor classifier as the training set already represents the features of the samples to be learned. In another words, the training set is the memory of the training samples. Memory increases along with an increase in training data.

K-nearest neighbor classifiers are easy to build as there is not a training process. When a new set of training data or a new feature is added to the existing training data set, there is no need for a new training process for this addition. However, there are some disadvantages of using k-nearest neighbor classifiers; they occupy large memory and require intensive computations.

2.1.3.3 Decision Tree Classifier

Decision trees are hierarchical and nonparametric classifiers. They are constructed by repeated splitting of subsets of training data into two descendant subsets. If every terminal has patterns that belong to only one class, then the error is zero. In designing a tree classifier, a single feature or a subset of features can be used. Each terminal in the decision tree has a class label. It is possible to use two or more terminal subsets with the same class label.

The simplest decision tree is the binary tree. In a binary tree, each node is split into two subsets. Secondly, a single feature should be used in each node. Two decisions should be made:

- 1) Which feature should be used in each node? and,
- 2) What threshold should be used for each feature?

Not all of the available features measured are necessarily used in building the decision tree. Selection of the features is done by a tree-partitioning algorithm. Terminal nodes, which are also known as leaves, are determined by partitioning the decision tree. At each node t of the classification tree, there is a probability distribution p_{tk} for the classes. Each pattern in the training set is assigned to a terminal node. So at each terminal node there is a random sample n_{tk} from a multinomial distribution of p_{tk} .

Based on the descriptions made above, the conditional likelihood is expressed as follows,

$$\prod_{\text{patterns}, i} P[i]_{x_i} = \prod_{\text{leaves}, t} \prod_{\text{classes}, k} P_{tk}^{n_{tk}} \quad (2.27)$$

where $[i]$ is the leaf to which the pattern i is assigned. Detailed information on tree partitioning methodology can be found in Chapter 13 of Venables and Ripney (1994).

The first module in a decision tree classifier is the training set (Figure 2.9) that contains all the features. Designing the decision tree using the features (training set) available is the next step. Nodes in the tree are split and the tree is formed with the target of having one class in each terminal. Output of this step is a constructed decision tree classifier. Later, a test pattern can be classified using this decision tree. The output class for the test pattern is determined based on the features and their associated threshold values in the decision tree.

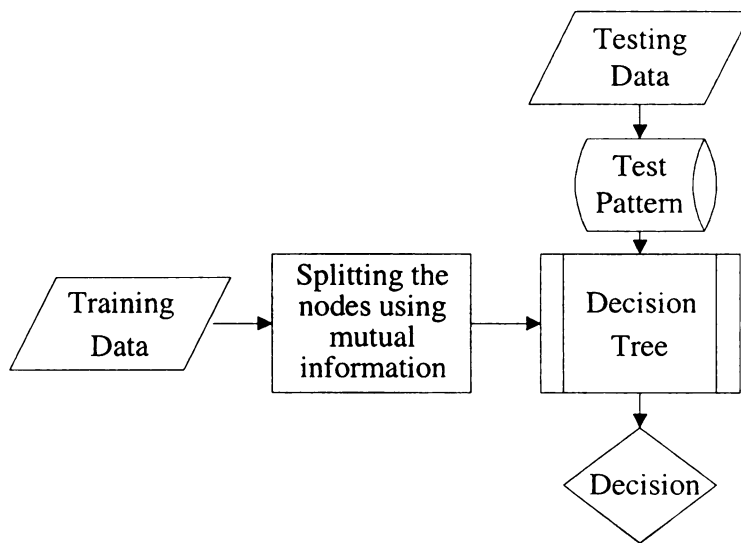


Figure 2.9 Modules of decision-making procedure in a decision tree classifier

The memory of a decision tree classifier is the tree constructed using the training data. Unlike the k-nearest neighbor classifier, when an addition is made to the existing training data, a new training (constructing a new decision tree) must be performed.

2.1.3.4 Bayesian Classifier

In this type of classifier, a given test pattern is assigned to one of M ($w_1, w_2, \dots, w_M$) classes based on its feature vector $x = (x_1, x_2, \dots, x_N)$. It is assumed that

feature vector x has a probability density function based on the pattern class (Jain, 1988).

In this study, conditional densities are assumed to be multivariate Gaussian (normal distribution). Gaussian densities are unimodal and require only two parameters, data mean and covariance.

If it is assumed that occurrences of a vector x and a class w_i are random, the likelihood of x being in w_i is given as (Precetti et al., 1993):

$$P(w_i | x) = \frac{P(x | w_i)P(w_i)}{P(x)} \quad (\text{Bayes Rule}) \quad (2.28)$$

where $P(w_i | x)$ is posteriori probability of w_i given x ,

$P(w_i)$ is a *priori* probability for class i ,

$P(x)$ is the probability that x occurs,

$P(x | w_i)$ is the conditional density (exists for all classes), which is assumed

Gaussian distributed and given as below:

$$p(x | w_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}[(x - \mu_i)^t \Sigma_i^{-1} (x - \mu_i)]\right\} \quad (2.29)$$

where the expression $\exp\left\{-\frac{1}{2}[(x - \mu_i)^t \Sigma_i^{-1} (x - \mu_i)]\right\}$ is the Mahalanobis distance

measurement that together with the exponential component in the Gaussian distribution function makes the probability of a pattern small if it is far from the mean.

In this parametric approach, prior probabilities of the classes are determined using training samples from each class. Unknown parameters of μ and Σ , which are the mean and covariance matrix of a class, are estimated from training samples using the technique of Maximum Likelihood Estimation given as follows:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.30)$$

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2. \quad (2.31)$$

Using these estimated parameters, discriminant functions for each class are calculated and the same functions are later used to classify the testing patterns. Discriminant functions under the assumptions made above can be developed as follows:

$$g_i(x) = \ln P(w_i | x) = \ln \left[\frac{P(x | w_i) P(w_i)}{P(x)} \right], \quad (2.32)$$

$$g_i(x) = \log_e [P(x | w_i)] + \log_e [P(w_i)], \quad (2.33)$$

$$g_i(x) = -\frac{1}{2} (x - \mu_i)' \Sigma_i^{-1} (x - \mu_i) - \frac{1}{2} \ln |\Sigma_i| + \ln P(w_i). \quad (2.34)$$

Equation 2.34 is calculated for each class and the decision rule given below is used to assign a testing pattern to the class whose discrimination function gives the highest value.

The discriminant function $g_k(x)$, for class w_k is given as:

$$\text{Assign } x \text{ to class } w_k, \text{ if } g_k(x) > g_l(x) \text{ for all } l \neq k. \quad (2.35)$$

Stages of decision making in a Bayesian classifier are the training data used in estimating the parameters of the discriminant functions, testing data from which the testing pattern is selected, evaluation of the testing pattern through all the discriminant functions, comparison of the outputs from the discriminant functions and decision (Figure 2.10).

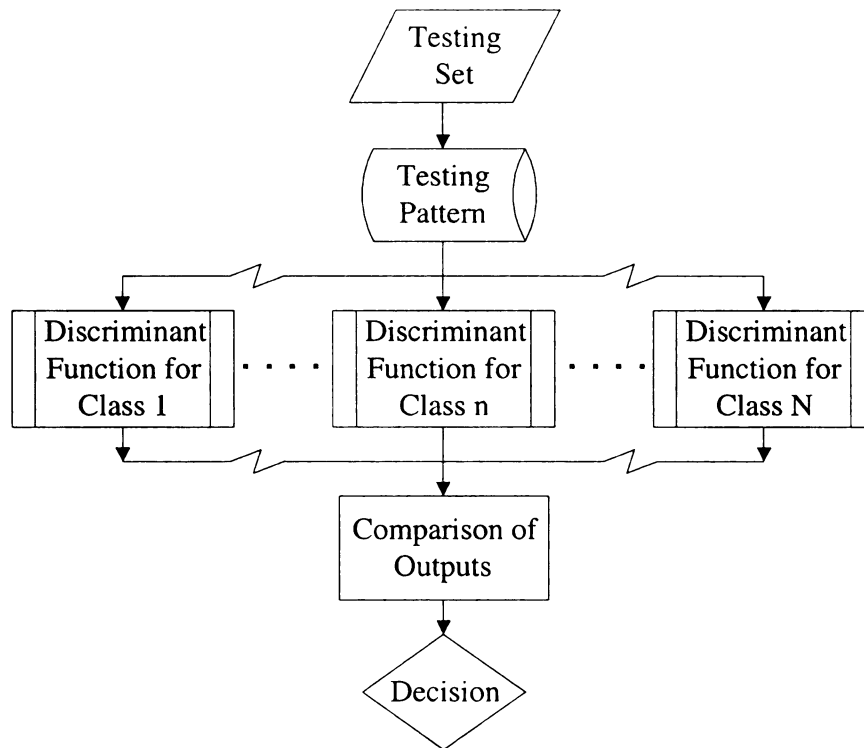


Figure 2.10 Modules of decision-making procedure in a Bayesian classifier

2.2 Empirical Studies

In this section, practical applications in the areas of machine vision, spectral reflectance, artificial neural networks and statistical classifiers as related to this study are reviewed.

2.2.1 Machine Vision Applications

Classification using machine vision and artificial classifiers has an expanding application in agriculture; bell pepper sorting, plant classification, apple classification, peach classification, and surface defect detection on fruit are some of the applications. In the following applications of machine vision, object segmentation and feature extraction in the images were applied.

Eight different plant species were classified using shape features which were obtained from binary images (Guyer et al., 1986). Four shape features of complexity, elongatedness, central moment, and principal axis moment were extracted from the overall image and singulated objects (leaves) in the image resulting in a total of eight features. In the classification procedure, corn plants formed one class and the rest of the plants belonging to the other seven species were accepted as the second class. Separation of corn plants from other species was performed eight times using two features in the beginning and each time increasing the number of features up to 9. The trial using 8 features performed the most successful classification resulting in an error rate of 9%.

In studying texture features of apple bruises, Throop and Aneshanesley (1993) used fifteen texture features; ten of them were calculated from co-occurrence matrix and five were calculated from gray level run length matrix. Bruises on apple varieties such as Golden Delicious and Red Delicious were studied. Two sub-images with different resolutions such as 50x50 pixels and 20x20 pixels were used in texture calculations. The image of 50x50 pixels included the bruise area and some good tissue in it, while the image of 20x20 pixels included only either the bruised tissue or the good tissue. In the results, images of 50x50 pixels provided better discrimination of bruised tissue than the images of 20x20 pixels. Bruises on Red Delicious apples were detected with higher accuracies compared to the bruises on Golden Delicious apples.

However, using a selected region for texture calculations is a drawback for automated classification algorithms. There is a requirement for the segmentation of the related region when a sub-region is used to calculate the texture features. This would increase the image processing. Moreover, as was mentioned, when there is a low contrast

between the defective region and the good tissue, it would be difficult to locate the region with the defect.

On the other hand, when an entire image is used for the calculation of texture features, there would be no search for a target region that contains the defective sub-regions. In this research, the entire image was used for texture calculations without any segmentation of the defective areas.

2.2.2 Spectral Reflectance Applications

Spectral reflectance has been widely used for assessing the quality aspects of agricultural commodities. Application of spectral reflectance varies depending on the equipment being used. In some studies measurement of spectral reflectance has been performed in sub-areas on a commodity using a spectrophotometer where regional measurements are made from defective areas and from non-defective areas separately. As another way of measuring the spectral reflectance, some researchers used spectral reflectance imaging that takes into account a selected part of an image or the whole image of the commodity to detect the flawed parts in an object.

In a study using the spectrophotometer a small portion of a whole apple was used in measuring the reflectance values (Geoola et al., 1994). Three apple groups were used in the experiment: Good apples; bruised apples that were kept at room temperature for 90 min; and bruised apples that were kept at room temperature for 24 h. Spectra of diffuse reflectance in the wavelength range of 400-840 nm was studied. Sample slices of 15 mm thickness were used, as the aperture of the spectrophotometer was not appropriate to use the whole apple. Classification criterion of the average reflectance values at n selected

wavelengths in the range of 750–800 nm was found as the most effective feature in discriminating the bruised tissue on Golden Delicious apples. Classification results of 96.1% for unbruised apples, 88.4% for bruised apples left at room temperature for 90 min and 93.7% for bruised apples left at room temperature for 24 h were obtained respectively.

In a similar research, spectral reflectance measurements were performed using a spectrophotometer from the bruised and unbruised sub-regions on Red Delicious apples (Upchurch et al., 1990). Before the measurements bruised apples were kept at room temperature for 20-25 h. Linear regression wavelength analysis was used to select the most effective independent variable that included one, two or three wavelengths in the range of 400-1000 nm. In the results, the lowest correlation was obtained by using single wavelength (50% misclassification), while the ratio of normalized difference and the derivative models, which included two and three wavelengths respectively, resulted in better correlations (2.5% and 3.5% misclassifications).

Using the same equipment in the wavelength range of 350-1200 nm, spectral analysis of peach defects, such as bruise, scar, scale, brown rot, and wormhole was performed. The Mahalanobis distance method was used to measure the classification accuracies of the defects. The criterion that used three wavelengths at 650, 720, and 815 nm had the highest Mahalanobis distance for most of the defect types on the peaches (Miller and Delwiche, 1991).

In the second method of reflectance measurement, which applies spectral reflectance imaging at different wavelengths, either all of the pixels in the image or some extracted features are input into a classifier. One example to this method can be multi-

spectral image analysis for inspection of poultry carcasses (Park et al., 1998). In this study, pixel gray values and also extracted Fast Fourier Transform intensity values were used as features to differentiate wholesome carcasses from unwholesome carcasses using an artificial neural network classifier. While combined features of pixel gray values at 540 nm and 700 nm worked best for an efficient classification (93.3%), Fast Fourier Transform intensity features at 700 nm worked better than other wavelengths (90%).

In analyzing a large number of defects on five different varieties of apples, such as Red Delicious, Golden Delicious, Crispin, McIntosh and Empire, gray level intensities of defective and non-defective regions were obtained by averaging the gray values in the selected windows in the spectral reflectance images (Aneshansley et al., 1997). The percentage of classification error was determined using Mahalanobis distance in the wavelength range of 460-1030 nm. High classification error was found for defects such as russet and sunburn at almost all of the wavelengths used. The greatest contrast between the defective areas such as bitter pit, leaf roller, scab etc. and unblemished areas was found at wavelength 750 nm. Defects such as bitter pit, leaf roller, stored bruises, cork and some early season insect stings showed reflectance values higher than undamaged tissue at wavelength 1030 nm. For defects such as scald and new bruise reflectance values from the damaged tissue was lower than the undamaged tissue at wavelength 970 nm.

Spectral reflectance information from a whole image was used in detecting bruises on peaches and apricots (Zwiggelaar et al., 1996). In this research, a pre-study was conducted to determine the effective wavelengths that would increase the contrast between bruised and non-bruised areas. Reflectance values measured by means of a

monochrometer were combined using Ratio, Derivative and Normalized difference methods, which were also used in the works done by Upchurch et al. (1990) and Miller and Delwiche (1991). Results were evaluated using a maximum distance method. Images of the fruits were acquired at selected effective wavelengths to be used for the bruise detection procedure. A flooding algorithm, considering the whole image, was used to detect bruised areas on the fruit surface (Yang, 1994). Classification success of 65% was obtained using images from a single wavelength for apricots and the ratio of two images for peaches.

Regardless of method used, either a spectrophotometer in a sub-region of an object or spectral reflectance imaging of a whole image or a portion of an image, the ultimate goal has always been the same, being able to detect the flawed regions on the surface of a matter. However, not using a whole image presents a limitation in applying automatic classification, as it is not practical or sometimes possible to select the sub-regions for defect segmentation in automatic sorting applications.

2.2.3 Applications of Artificial Neural Networks

There are different types of neural networks based on their connection types and architectures. Backpropagation neural network is one of them. In recent years, the backpropagation learning rule has been used very commonly in artificial neural network applications including multi-layer neural networks. According to Hassoun (1995), development of backpropagation increased popularity of artificial neural networks. Some of the problems backpropagation artificial neural networks have been used to solve are

pattern classification, clustering, function approximation, prediction, optimization and control (Jain et al., 1996).

Artificial neural networks are very robust in learning even non-linear relations between input features and output categories, such as in sorting biological materials. Thus, they have been effectively used in classification problems in many agricultural areas. Multi-layer feed forward neural networks are the most commonly used artificial neural network model in the classification of agricultural produce.

Apple surface features were classified using machine vision and neural networks (Yang, 1993). In this study, two images of an apple, one under diffused lighting and the other under structured lighting were used. In the primary investigation, selected sub-regions in the images, which included the blemished area, with structured light were used as input features to neural networks. All the pixels in the selected sub-images were used as features to the neural network after applying minimum processing. However, a high unacceptable classification error was obtained due to the complexity in the images with inconsistent stripes in them. In the second approach, features such as average positive and negative curvatures, average number of points with zero curvatures and average curve length were extracted from the striped images used in the first approach. In addition to these features, area, compactness, a slope feature, average gray level intensity and variance of gray-level intensity were extracted from the image under the diffuse lighting. A total of nine features were extracted from two different images of an apple. No exhaustive search was applied to find the effective number of nodes in the hidden layer or effective number of hidden layers. The resulting number of nodes in the input layer and the hidden layer were obtained based on the trial and error. Successful results were

obtained with an average classification success of 96.6% in classifying images into three blemish output categories.

Apples were classified according to their colors using back propagation neural networks (Nakano, 1997). Nine color features including color gradients, variances and chromatic coordinates were extracted from apple images. Two separate sets of backpropagation neural networks were used. The first neural network was used to decide whether the red color of an apple was normal or abnormal. Second neural network was used to grade the apples into five output categories based on the quality of the color. Satisfactory results were obtained except for one quality group that was confused either with one upper class or one lower class.

Golden Delicious apples were graded using features extracted by machine vision and classification algorithms based on rule-bases and neural network (Heinemann et al., 1997). Four features including mean hue values for color, two dimensional moment for shape, diameter for size, and percentage of average surface area affected by defects for defect were used in the classification applications. In the results, 94% and 84% agreements were obtained for color and shape classifications respectively between human and machine classification applications. While 94% agreement was found for russet identification, 75% agreement for identification of defects such as cuts and diseases was obtained. The system could not classify fly-speck defect.

Another backpropagation neural network was used to separate broken corn kernels from whole kernels (Liao et al., 1993). Eight morphological features, such as local maximum curvatures, symmetry ratio, aspect ratio and tip cap variation ratio were extracted for shape description of the corn kernel profile from the images under diffuse

reflected light. Images were taken basically at two positions; with the germ side up and down. The backpropagation neural network used in discriminating the broken kernels from the whole ones had eight nodes in the input layer, two hidden layers with thirty-two nodes in the first layer and eight nodes in the second, and an output layer with two nodes. The neural network classifier provided accuracy rates of 89% and 94% for the whole and broken round kernels with the germ side up, and 94% and 96% for the whole and broken flat kernels with the germ side down.

Cracks in eggs were detected with backpropagation neural networks using histogram features from color images (Patel et al., 1995). Red, green and blue histograms from the image of an egg were extracted and concatenated to form the features that were input to a neural network. Concatenating gray values from three histograms of red, green and blue resulted in total of 768 input features (3×256). Averaging two adjacent gray values in the combined histogram with 768 units reduced the number of input features to 384. Twenty-four nodes that were found based on trial and error were used in the hidden layer. The system provided an accuracy of 97.78%.

In almost all of the applications either using sub-images or the entire images mentioned so far, both image enhancement and feature extraction procedures were performed. These applications usually required a dense computation and significant amount of time. For instance, in an apple sorting system, two different types of images were acquired under different lighting conditions to extract two different sets of features, which increased image processing time and complexity of the system (Yang, 1993). In addition, calyx and stem were not included or evaluated, instead separate features and algorithms were used to classify them. All of these limitations create a disadvantageous

situation for automatic classification. However, a classification algorithm that minimizes image enhancement and processing tasks and is capable of recognizing multiple defects and calyx/stem ends on a fruit would be more acceptable in an automatic classification system.

2.2.4 Comparison of Statistical and Neural Network Classifiers

Prior to artificial neural networks, and to date, statistical classifiers have been commonly used in classification studies in agriculture. However, success of applying statistical classifiers in classification problems has been dependent on the data used. Problems with traditional statistical classifiers are declared as follows (Yang, 1992):

- 1) Some of the selected features can be statistically dependent,
- 2) The training set of samples is often not large enough to represent all the classes,
- 3) The class-conditional probability densities are not well known.

An artificial neural network classifier is usually not affected by the first and third problems mentioned, as it does not make any assumption about the input data. The second problem, however may still affect a neural network classifier.

A good comparison of statistical and artificial backpropagation neural network classifiers was done in classifying cereal grains using machine vision (Luo et al., 1999). A total of twenty-eight morphological and color features were used as input to the artificial classifiers. These features were extracted from the images that contained twenty-five kernels. Segmentation and enhancement were performed to improve the images. Three types of classifiers were compared in two cases of classification problems. In the first one, cereal grain kernels including two types of wheat varieties, barley, rye and oats

were classified using a parametric statistical classifier, k-nearest neighbor classifier and a multi-layer neural network classifier. The results demonstrated that neural network and k-nearest neighbor classifiers gave similar results while the parametric statistical classifier performed significantly lower than the other two. In the second study where healthy and damaged kernels from the same groups were classified, similar classification performances were obtained from the artificial classifiers. In summary, k-nearest neighbor and multi-layer neural network classifiers were superior to the parametric statistical classifier in each application. However, application simplicity and speed of statistical classifiers were mentioned as possible advantages that made them attractive to use.

A Bayesian classification method was used in segmenting the defects on a bi-colored apple such as Jonagold (Leemans et al., 1999). Pixel gray values from color images were used as features. Probabilities of having a defective apple and a healthy apple were assumed to be equal. Parameters such as the probability of a pixel belonging to class healthy (or defect) and of having a specific color were estimated using gray values from pixels. Then, the Bayesian classifier was tested using a test set of pixel gray values. Although pixels in a transition area, between the ground color and blush, and in some russet areas were misclassified or poorly classified, good results were obtained for other defects such as bitter pit, fungi damage, scar, frost damage, bruise, insect damage and scab.

In recognition of handwritten characters, Bayesian and neural network classifiers were compared. Two types of input features, pixel gray value and texture, were used. Gray value features from pixels were obtained from a binary and size normalized image

of 16x16. Texture features were uniform length feature vectors of 636 binary components obtained from the images. In evaluating the performances of the two classifiers using two types of features, backpropagation neural network was found more successful than the Bayesian classifier using the pixel features. This was explained as a result of the linear discrimination function the Bayesian classifier had, contrary to the non-linear discrimination function used in the backpropagation network. However, using texture features, Bayesian classifier performed slightly better classification compared to the backpropagation neural network. It was also mentioned that for high dimensional finite training samples such as pixel gray values in an image, calculation of probability density functions was a drawback for a statistical classifier so that lower performances would result due to the assumptions made (Lee et al., 1991).

Deck et al. (1995) compared a backpropagation neural network classifier and a statistical (Fisher Discriminant) classifier in sorting potatoes according to defects of greening, shatter bruise and shape. Two images, one from top and the other from side, were used in inspecting the defects of greening and bruise. In sorting according to greening defect 6 hue histogram bin values corresponding to the color green from both views were summed together and used as features in training and testing the classifiers. In bruise detection, 20 manually selected hue values in bruised or non-bruised areas were used as features to classifiers. Only the top view image was used in inspecting shape defect. Features of 1st through 10th radial Fourier harmonics except the 2nd one were used as features to the classifiers. Backpropagation neural network classifier had the highest accuracy for greening (74%) and shape (73.3%) detections. On the other hand, Fisher method had the highest accuracy for shatter bruise detection (76.7%).

Artificial neural networks are preferred in many research areas with their robustness rooted from the massive parallelism and ability to learn even non-linear relations between input data and output categories. However, despite the facts about the advantages of artificial neural networks, traditional statistical classifiers still preserve their popularity with their easy to use and clear structure. Depending on the data used, they may result in better classification by being more efficient in terms of time spent in training and testing.

Chapter 3. MATERIALS AND METHODS

The following steps and experiments were taken in conducting the research:

- 1) Data collection
- 2) Image acquisition
- 3) Image processing
 - a. Reducing the resolution of images
 - b. Background segmentation
 - c. Elimination of specular reflectance from the images
- 4) Preparation of image data for artificial classifiers
 - a. Preparation of pixel gray values as features from an image
 - b. Extraction of texture features from an image
- 5) Designing the artificial classifiers based on the classification applications
 - a. Classification with two-output classes
 - i. Classification using pixel intensity values from images
 1. Design of single backpropagation neural network
 - ii. Classification using texture features from images
 1. Single backpropagation neural network classifier
 2. K-nearest neighbor classifier
 3. Decision tree classifier
 4. Bayesian classifier

iii. Classification using combined features from images using single backpropagation neural network

1. Using pixel intensities from two images at different wavelengths
2. Using pixel intensities and texture features together
3. Combining two images by averaging

b. Classification with five-output classes

Same steps were repeated here as in the case of using two-output classes (a). The only difference was that five classes included all defective and non-defective apple groups including (stem and calyx view image groups). Also a second application of a backpropagation neural network was used- a multiple backpropagation neural network that was a combination of single neural networks developed for 2-class classification.

3.1 Data Collection

Two apple varieties Golden Delicious and Empire, one with light and the other with dark skin colors respectively were included in this study. Two apple varieties of different colors were selected to increase the variability of the experimental data and to challenge the classification system. The contrast between the reflectance values from defective and healthy tissues may be strong, for instance, for a variety with light ground color, while it may be weak for a variety with dark ground color.

Apples from the Empire data set belonged to the same strain and were all harvested on the same day of 16th of September 1998. The Golden Delicious data set contained several random strains and was harvested on two different dates, 21st of September and 9th of October of 1998. All the apples were harvested manually in the experimentation orchards of Michigan State University.

Selection of the specific two varieties was based on their availability in the orchard and the availability of defects on these varieties. After harvest, apples were kept in a cooler at 37° F until image acquisition. They were held at room temperature for an hour before image acquisition.

3.1.1 Apple Defects Used in the Study

Although more defects were selected and collected for each variety, leaf roller was included in this study for Empire variety, as the number of apples for other defects was low. Similarly, russet (netting) and bitter pit were chosen for Golden Delicious variety. Other defects such as bruise and puncture, which usually occur during post harvest operations due to external impacts, were created by applying force to the apples as explained below.

Apple samples were kept at room temperature for an hour before bruising and puncturing. After applying forces to them, apples were held at room temperature for another 24 h before image acquisition.

3.1.1.1 Bruise

Apples were released for a free fall from a height of 20 cm onto a concrete floor. Selection of the 20 cm drop height was based on the commonly used drop height range of 5-30 cm in similar applications in the literature when studying the reflectance of the bruise such as by Brown et al. (1974), Upchurch et al. (1990, 1991 and 1994), Crowe and Delwiche (1996), and Miller and Delwiche (1991). It was declared that bruises on Golden Delicious apples that resulted from falls of less than 10 cm height could not be detected after 90 min (Geoola et al., 1994).

In the bruising procedure, apples were held in such a way that their stem and calyx axis would be parallel to the floor expecting a bruise formation on the cheek. It is believed that using a fixed drop height with varying sizes of apples created somewhat variable bruises on apple samples. Geometric location of the bruise created affected severity of damage on the apple surface. It is likely that more severe bruises would form on the peaked areas of apples than the areas on the flat surfaces. This situation may explain the differences in the appearances of the bruises in different locations of apple fruit. All of the variations in the structure of the bruises made it more challenging to detect them in automatic classification applications.

Skin color of the bruised area is often normal or slightly discolored. Tissue under the skin in the bruised area is brown and filled with the fluid from broken cells. After prolonged storage this fluid evaporates and the tissue in the bruised area becomes dry and spongy (Aneshansley et al., 1997).

3.1.1.2 Puncture

Apples were impacted with various objects to create punctures. In most of the punctured areas the skin was broken. The amount of force applied to each apple was different to provide puncture severity and appearance variability.

3.1.1.3 Leaf Roller

This is a defect caused by an insect called Leaf Roller. Adult moths lay eggs on the upper branches of apple trees in early spring. Later generations lay eggs on the leaves. As the larvae feed on the leaves, they roll and web the leaves together. Damage usually occurs in the areas of skin and outer flesh of apples that touch the infested leaves or hang in clusters. Russet scars develop in the following maturing period in the damaged areas (Anonymous, 1998).

3.1.1.4 Russet (Netting)

Russetting affects some apple cultivars including Golden Delicious. Some early sprays to prevent the scab fungus and the infection of Powdery Mildew may cause russetting. Russet can also occur due to frost just after bloom or cool spring weather between the drop of the blossom and formation of apples (Manhart, 1995).

There are 5 morphological classes of russet that are effective on apples: stem-bowl, calyx-end, netted, lenticular, and solid (USDA-ARS, 1999). There are separate and multiple reasons behind each of the russet types mentioned. It is believed that russet is initiated during the rapid period of cell enlargement. Due to the first heat, stressed cell expansion beneath the epidermis exceeds the stretching capability of epidermal and

cuticle layers and therefore, the surface cracks and cells become separated and exposed. Later, exposed cells begin producing brown, corky cells to protect against further injury or dehydration. Thus, the degree of injury and ability of the epidermal layer to recover determine the amount of russetting that will form. Russet is usually skin deep and the flesh is not affected.

3.1.1.5 Bitter Pit

A primary reason for bitter pit formation is shortage of calcium in the apple, not necessarily in the tree or in the soil. Larger apples are more susceptible than small-sized apples. Excessive growth in tree branches causes the calcium to be consumed primarily by the tree or the leaves. Insufficient water during summer causes the tree to take some water, and consequently calcium, from the apple fruit. Low pH of the soil may also cause the formation of bitter pit (Manhart, 1995).

Small brown lesions of bitter pit defects that develop in the flesh of the fruit are, depending on the cultivar, 2-10 mm in diameter. The tissue below the skin is dark and corky. Depressed spots on the surface develop at or near harvest or after a period of cold storage. These spots usually become darker and get more sunken compared to the healthy surrounding skin. After one to two months in storage, they reach their full development (Andris et al., 1999).

Golden Delicious is one of the cultivars that are susceptible to bitter pit. After the tissues are affected, higher rates of respiration and ethylene are produced and more protein and pectin are synthesized with increased migration of organic ions into the affected areas. More starch gain is seen in the affected areas than in the healthy tissue.

Moreover, a mineral imbalance in the apple flesh occurs with low levels of calcium and relatively high concentrations of potassium and magnesium. Due to the low levels of calcium, selective permeability of cell membranes decreases causing a cell injury and necrosis (Andris et al., 1999).

3.2 Image Acquisition

As previously noted, the main objective of the study was to separate defective apples from good apples by using the reflected appearance features obtained through spectral reflectance imaging and artificial classifiers.

A laboratory machine vision system was used to obtain images of apples illuminated at desired bands of light energy by a monochromator controlled halogen light source (Oriel Instruments model 77250) having a slit opening of 40 nm. A black and white vidicon camera (Hamamatsu C2741-03) with an enhanced NIR range and a sensitivity range of 400-2000 nm, a Pentium 200 MHz personal computer and QuantIm (Zedec Inc.) software were used for image acquisition. Problems with background drift and lag that are common with vidicon cameras were decreased by turning on the camera at least 2-4 h beforehand for it to warm and stabilize. A special cut-off filter was used to prevent visible range harmonics from illuminating the apple in the NIR range. A white pad standard was used to calibrate the system. Images were acquired and saved in TIFF format. A schematic diagram of the image acquisition system can be seen in Figure 3.1.

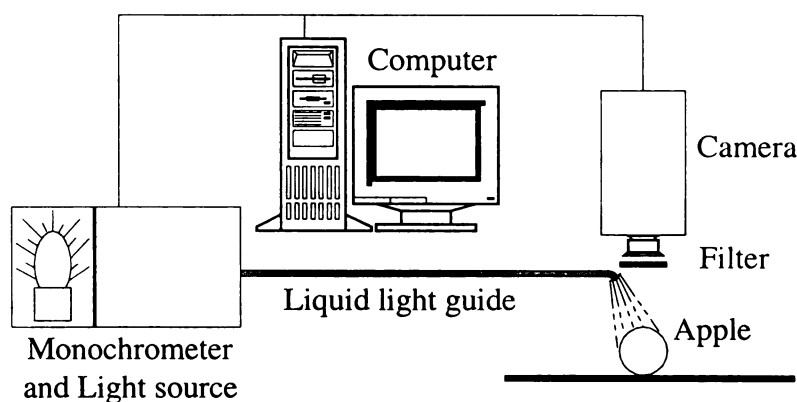


Figure 3.1 A schematic diagram of imaging configuration

Images were taken in a dark room with only the light source, coming through a liquid light guide from the monochromator, cast on the apple. The wavelength of the light emitted changed automatically with desired increments and within the desired range. Distance between the lens of the camera and the background level where the apple to be imaged stood was 40 cm. Distance between the light source and the background level, on the other hand, was about 10 cm.

Presentation of the apples to the camera was done by hand for this developmental study. Once an apple was positioned in its desired orientation on a black matte background, forty images in the resolution of 480x640 pixels were acquired from wavelengths 540 to 1320 nm with a step size of 20 nm. By extending the wavelength to 1320 nm from around 1000 nm, which has not been commonly used in fruit assessing studies, it was expected to get clearer distinctive reflectance from defective and non-defective areas on the apple surface. Selection of the wavelength range of 540-1320 nm was based on the visual inspection of the images between wavelengths from 400 to 2000 nm. Above 1320 nm and below 540 nm no valuable information (which might enhance defect discrimination) was seen in the images.

The purpose of taking forty images was to determine the effective wavelength(s) for detecting a defect. Artificial classifiers would use images at the most effective wavelength to separate the defective apples from good ones. When there is more than one effective wavelength for a defect, possible combinations of them would be used for improved classification success.

3.3 Image Processing

A goal and potential uniqueness of this study was that minimal computationally intensive image enhancement or processing operations be applied. Three major image-processing operations applied were reduction of image resolution, background segmentation, and specular area deletion. For comparison purposes, images with background were also used in classification of some defects. The three major image-processing operations applied in this study are explained below.

3.3.1 Reducing the Resolution of Images

Decrease in the resolution of the images was performed as the original image resolution of 480x640 pixels (307,200 pixels) was not feasible in terms of time to train the artificial neural networks used in selecting the effective wavelengths and in classification. By averaging the pixel gray values in a selected window, resolution of an image was reduced into 60x80 pixels. Equation 2.1 was used for this operation. All classification applications were performed using images at resolution of 60x80 pixels.

However, for comparison purposes additional classification applications were performed using four different resolutions of images such as 480x640 (original image

size), 240x320, 120x160, and 60x80 pixels. These image resolutions were tested on classifications of three different groups of apple images: bruised area view apple images against good tissue view apple images (Empire), calyx view apple images against good tissue view apple images (Empire), and bitter pit view apple images against good tissue view apple images (Golden Delicious). Classification accuracies obtained from using images at different resolutions were compared and the results are presented in the following Chapter.

3.3.2 Background Segmentation

Background segmentation in the images was applied to minimize the confusion of the classification process with background pixels. With no background in the images, it was expected that the artificial classifiers would be more sensitive to the reflectance difference between defective and non-defective areas. The following procedure was applied to segment the apple from its background.

The image of an apple at wavelength 760 nm was selected among forty images taken at forty different wavelength bands based on the visual assessment of the maximum contrast between the apple and the background. A threshold pixel gray value, which was determined manually from the images, was used to segment the background pixels from the apple pixels. A few of the pixels in the edges of the apples were lost in this operation. After this operation, a binary master image was obtained for each apple. Then, the binary image and the forty original images with background of the same apple were used in an AND operation separately. In the end, an apple image with no background was obtained for each wavelength band from 540 to 1320 nm for all the apples in each group.

However, images with background were also used in the classification of some defective apple groups against good apple group in the Empire data set in a pre-study. This application was performed only to compare the classification accuracies obtained from using images with and without background.

3.3.3 Elimination of Specular Reflectance from Images

A just harvested apple has a thin layer of wax that has a naturally rough surface. Therefore, regular reflectance from the apple surface is diffuse and the body reflectance, which is the reflectance of the incident light after being reflected from internal parts of the apple, transmits the characteristic color of the apple to the detector. However, in automatic sorting operations apples are washed, dried, and usually waxed to clean and to improve their appearance. During these operations natural wax of apples is either lost or smoothed. A smooth surface, on the other hand does not reflect an incoming light diffusely. Instead, regular reflectance produces a glared area, which is also referred as the specular component of the reflectance (Birth and Zachariah, 1976).

To simulate the cleaning operation of apples in sorting lines and also to get rid of dirt, apples were cleaned with a piece of dry cloth before collecting the images. As a result of this action, images of apples used in the experiment did have the specular areas that had very high and abnormal gray values. Sizes of these specular regions were generally quite small compared to surface of the apple in the images; however, they varied with the apple geometry.

As the shapes of these specular regions might have been correlated with the types of defects on surfaces of apples, it was decided to eliminate them from the images to prevent a biased classification.

Despite efforts to eliminate the unwanted small regions of specular reflectance during the image acquisition process using filters etc., limited success was obtained as extra filters would decrease the amount of light that reaches the camera.

Some image processing efforts were also made to clear the images of the specular reflectance regions. One attempt was to average the pixels of a selected window just outside of the glared region and to replace the pixel gray values inside the glared region with the calculated average. However, after processing the images with this approach, tracks of the specular reflectance regions in the images still seemed to exist. Therefore, specular reflectance regions were automatically deleted in the images using a fixed sized window centered on the pixel that had the maximum gray level in the image. The pixel with maximum gray value was always in the central area of the specular reflectance region. The final processed image used in the wavelength and classification studies contained only the gray values from apple pixels after extracting the background and the specular regions.

It should be noted here that images of apples were taken in a period of two-three weeks, due to the time to acquire forty images per apple. During this period of time there was a slight possibility that adjustments of the image acquisition system, such as time (2-4 h) between turning on the camera and starting the image acquisition, might have been changed from time to time. However, in classification procedures apples were divided in training and testing groups according to their numbers as odds and evens, which

eventually would solve any biasing in images and yet challenge the classification systems further requiring a more robust classifier.

3.4 Preparation of Image Data for Artificial Classifiers

Two types of features were used from images; one feature group was the complete set of pixel gray values from the entire image and the second group was the texture features extracted from the same image. Extraction and preparation of these features are explained below.

3.4.1 Pixel Gray Values as Features from an Image

Following the three image-processing operations explained above, pixel gray values of images were written into text files as long vectors. Later, this vector of gray values was input to the artificial neural network. The artificial neural network was the only classifier used with this type of feature set.

A preliminary study on finding an efficient way of using pixel gray values in the classification process was performed. Two types of gray values were tested in this approach. In the first one, original gray values from 0 to 255 were used. In the second approach, normalized pixel gray values were used. Normalization enabled standardizing the input data eliminating the disadvantage of using different ranges of gray values with large gaps between them. At the end of the normalization, data is represented in the intervals such as [0, 1] and [-1, 1]. Linear normalization, which is described below, was used in this research:

$$x_{kl,norm} = \frac{(x_{kl} - x_{k,min})}{(x_{k,max} - x_{k,min})}, \quad (3.1)$$

where, x_{kl} is a real value,

$x_{kl,norm}$ is a normalized value,

$x_{k,min}$ and $x_{k,max}$ are the minimum and the maximum values for the variable x_k respectively (Kasabov, 1996).

In this pre-study, improved classification results were obtained by using the normalized features, thus, during all of the wavelength selection and classification applications normalized pixel gray values were used.

3.4.2 Extraction of Texture Features from an Image

The same images used for the artificial neural network classifier study described in the previous section were used in the calculation of texture features. First, a spatial dependence matrix and then three texture features of angular second moment, contrast and correlation were calculated as explained in section 2.1. Three texture features were calculated at 0 degree (θ) and considering 1 unit (d) distance between the pixels (Figure 2.1) in the image.

Similar to the case using pixel gray values, a preliminary study was performed to compare the normalized and non-normalized texture features, as the magnitudes of three non-normalized texture features were quite different from one another. Linear normalization described in the previous section was applied to calculated texture features. In normalization procedure, the same texture features for different classes were normalized in one group. This was applied to all three-texture feature groups to eliminate the differences between their ranges (Ebert and Dobbins, 1990).

Results obtained from using texture features with and without normalization showed that using normalized features improved the classification success. Thus, normalized texture features were used in all of the classification studies using texture features. Texture features were used either alone or together with the pixel gray values in the classifiers as explained in the following sections.

In this study, instead of using a selected window on a defective area, the entire image was used for the calculation of the texture features. However, pixels with zero gray value from the background and the deleted specular regions of the image were excluded from the calculations, as they did not have a direct effect on the texture.

3.5 Designing the Artificial Classifiers Based on the Feature Types and the Number of Output Classes

Five different classifiers were used in classifying the images into desired output groups. These were nonparametric Single and Multiple Backpropagation Neural Networks, K-Nearest Neighbor and Decision Tree classifiers and a parametric Bayesian classifier. Two types of features and their combinations as explained below were used in the classifications. Two classifications were performed with two different numbers, 2 and 5 of output classes (Table 3.1).

Before explaining the application and design procedures of the classifiers, types of the features and the classification applications with different number of output classes are explained. Later, classifiers are explained first based on the number of output classes and second based on the feature type. Each apple variety was considered separately in each classification study.

Table 3.1 Classifiers, types of input features and the number of classes used

Input Features	Classifiers				
	SBNN†	MBNN††	K-NN‡	Decision Tree	Bayesian
Pixel Intensity	X	X			
Texture	X		X	X	X
Combined	X				
Number of Classes					
2	X		X	X	X
5	X	X	X	X	X

† Single backpropagation neural network, †† Multiple backpropagation neural network,
‡ K-Nearest Neighbor

Feature Types

1) Intensity values in the pixels of an image: this feature was used only in the (single and multiple) backpropagation neural network classifiers, as this type of classifier has a unique ability to learn nonlinear relations between large numbers of input features and output categories. However, this feature was not used in the Bayesian classifier as the expected classification error of a Bayesian classifier using a finite training sample can increase as the number of features increases due to the inaccuracies in estimating the parameters of the classifier (Raudys and Jain, 1991). Similarly, using pixel gray values as features in the classifiers of k-nearest neighbor and decision tree was not practical as it would take a long time to test an apple image with so many input features. For instance, in testing an image in a k-nearest neighbor classifier, distance measurements between all of the pixel gray values in the testing image and in the training images would require excessive computations.

2) Textural features calculated from the intensity values of an image: these features were used in every classifier to compare performances of the classifiers. Contrary to the situation with pixel gray values, there were only three features used in this application.

Thus, parameter estimation for the Bayesian classifier was expected to be reliable. Also,

distance measurement in the k-nearest neighbor classifier and tree splitting operations in the decision tree classifier were expected to be fast.

Considering the number of output classes, there were two situations; the first was the case with two-output classes and the second one was the case with five-output classes. Explanation of the classification applications and the descriptions of the classifier designs were done based on these two classification applications.

3.5.1 Classification with Two-Output Classes

In this application, groups of defective area view apple images were separately classified against the non-defective view apple images as shown in Table 3.2. Determining the effective wavelengths in discriminating a specific defect from good tissue was the aim of this application. Classifiers and the features used in this category are explained below.

Table 3.2 Groups of defective and good apples included in the study for each variety

Apple Variety	Apple Tissue Characteristics					
Empire	Good	Bruise	Stem	Calyx	Leaf Roller	Puncture
Golden Delicious	Good	Bruise	Stem	Calyx	Russet	Bitter Pit

3.5.1.1 Design of Single Backpropagation Neural Network Classifier for Classification and Wavelength Selection Using Pixel Intensity Values

This type of classifier is actually the backpropagation neural network explained in section 2.3. The number of input nodes is equal to the number of pixels in an image (60x80 pixels). Pixel gray values were concatenated in a vector form (4800x1) as can be seen in Figure 3.2 below for a two-class classification problem. Although there are

suggested methods of choosing the number of neurons in hidden layer, such as finding the number of clusters available in the training data set (Kasabov, 1996), ten neurons were found to be efficient based on trial and error. Similar previous studies by Yang (1993), Nakano et al. (1995), and Patel et al. (1995) used this rationale of choosing the number of neurons in hidden layer in classification of agricultural commodities. Using two or more hidden layers was avoided, as it would increase the training time. The number of output nodes was two or five based on the number of desired classification classes.

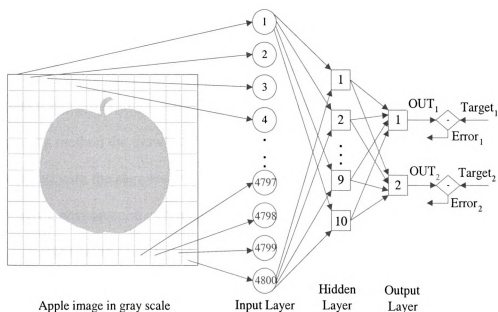


Figure 3.2 Schematic of the backpropagation neural network in classifying an apple image

As described in section 2.3, a sigmoid function was used in the nodes of hidden and output layers. Selection of values for the coefficients of learning rate (η) and momentum (α) were again found based on trial and error. The maximum number of iterations in training was set to 4000. However, convergence was established generally

around 400-2000. Error rate convergence criterion for the neural network to stop learning was 0.002. Output targets that were used in the training procedure (supervised learning) were coded as shown in Table 3.3.

Table 3.3 Codes for the output targets for each class used in training for Empire variety

Number of Classes	Target Codes in Training					Name of the Target Class
	Class1	Class2	Class3	Class4	Class5	
2	1.0	0.0				Good Apple
	0.0	1.0				Defective Apple
5	1.0	0.0	0.0	0.0	0.0	Good Apple
	0.0	1.0	0.0	0.0	0.0	Bruised Apple
	0.0	0.0	1.0	0.0	0.0	Stem + Calyx
	0.0	0.0	0.0	1.0	0.0	Leaf Roller
	0.0	0.0	0.0	0.0	1.0	Puncture

An alternative training method that basically feeds the training samples from each class respectively to the neural network was used in this study. By using this alternative training method the drawback of feeding the network with all the samples from one class first and with the samples from another class later was prevented as this type of training causes a phenomenon called catastrophic forgetting (Kasabov, 1996).

The same coding used in training in Table 3.3 was used for the testing data set. However, for testing the code for one pattern was repeated for all patterns in a quality group in a sequence, as there was no catastrophic forgetting problem in testing. Next, the codes for the second quality group were repeated and so on.

Initial weights for each neural network application were randomly selected between -0.3 and 0.3 based on the number of nodes in the input, hidden and output layers of the network. Numbers of apples used in classification studies are given in Table 3.4.

Table 3.4 Numbers of apples in training and in testing sets used in classifications

Apple Variety	Classification	Number of apples in training (†)		Number of apples in testing (†)	
Empire	Good vs Bruised	52	(26 ^g + 26 ^b)	52	(26 + 26 ^b)
	Good vs Stem	46	(26 ^g + 20 ^s)	36	(26 + 10 ^s)
	Good vs Calyx	46	(26 ^g + 20 ^c)	36	(26 + 10 ^c)
	Good vs Leaf Roller	46	(26 ^g + 20 ^l)	36	(26 + 10 ^l)
	Good vs Puncture	41	(26 ^g + 15 ^p)	31	(26 + 05 ^p)
	All 5 groups (Stem and calyx as one group)	117	(26 ^g +26 ^b +30 ^{sc} +20 ^l +15 ^p)	97	(26 ^g +26 ^b +30 ^{sc} +10 ^l +05 ^p)
Golden D.	Good vs Bruised	50	(24 ^g + 26 ^b)	50	(24 ^g + 26 ^b)
	Good vs Stem	39	(24 ^g + 15 ^s)	39	(24 ^g + 15 ^s)
	Good vs Calyx	44	(24 ^g + 20 ^c)	34	(24 ^g + 10 ^c)
	Good vs Russet	48	(24 ^g + 24 ^r)	37	(24 ^g + 13 ^r)
	Good vs Bitter Pit	44	(24 ^g + 20 ^{bp})	38	(24 ^g + 14 ^{bp})
	All 5 groups (Stem and calyx as one group)	134	(24 ^g +26 ^b +40 ^c +24 ^r +20 ^{bp})	97	(24 ^g +26 ^b +20 ^{sc} +13 ^r +14 ^{bp})

† Number of apples for each group of apples: (^g) Good, (^b) Bruise, (^s) Stem, (^c) Calyx, (^l) Leaf Roller, (^p) Puncture, (^r) Russet, (^{bp}) Bitter Pit, (^{sc}) Stem-Calyx

Performance Evaluation

In practice, a classifier is first designed using training samples; then, test samples are fed into the classifier one at a time. Finally, the percentage of misclassified samples is taken as the estimate of probability of misclassification. To get accurate error estimation, training and testing samples should be statistically independent, and the number of samples used to estimate the error rate should be large.

After converging at the end of training period, the neural network classifier was validated by using the training data set. Later, the neural network was evaluated by classifying the testing data. In each application, performance of the neural network classifier was measured by calculating error rate and classification accuracy. Calculation of error rate in the neural network classifier was given in Equations 2.16 and 2.17.

Error Rate to find the classification accuracy was calculated using

$$E = \frac{e}{N} \quad (3.2)$$

where, E is estimate of the true error rate, and e is the number of misclassified patterns in N samples (James, 1985).

In this research, output obtained after testing a pattern in the neural network was compared with the target output of the related pattern. If there was no match between a calculated output and its targeted output, the related test pattern was declared as a misclassification.

Error estimation is needed to learn the probability of making a mistaken classification for a future randomly chosen sample. Error rate predicts the performance of a recognition system. More detailed evaluation of a classification system can be obtained through confusion matrices; a confusion matrix is a way of displaying a full breakdown of error rates among the groups in the classification. In another words, confusion matrix is an $n \times n$ contingency table of actual group to classified group. For a two-class classification a confusion matrix shown in Figure 3.3 can be established:

		Actual		
		1	2	
Classified	1	E_{11}	E_{21}	N_{1c}
	2	E_{12}	E_{22}	N_{2c}
		N_{1a}	N_{2a}	

Figure 3.3 A confusion matrix for two classes (James, 1985)

Where E_{ij} is the number from group i classified as group j , N_{ia} is the number actually in group i and, N_{ic} is the group classified to group i (James, 1985). Confusion matrices formed for each classification application in this study are given in Appendix B. C-programming language in Microsoft C++ was used in coding the algorithm for the neural networks.

Wavelength Selection

Reflectance values from defective and non-defective tissues of an apple are expected to be different. Therefore, every 20 nm increment within the range of 540 - 1320 nm was investigated for an improvement in defective versus good tissue discrimination. Forty separate runs were performed using the single backpropagation neural network classifier for each classification application shown in Table 3.4. Effective wavelengths that would enhance the classification between the defective and non-defective apples were selected using the criterion of classification accuracy and error rate.

3.5.1.2 Classification Using Texture Features from Images

Three textural features calculated from the images at effective wavelengths were used. Information on how these features were obtained is given in sections 2.1 and 3.4.2. Four artificial classifiers were included in the textural feature study. Single backpropagation artificial neural network is one classifier. Information on its design procedure is given in section 3.5.1.1. Other classifiers, which may also be called as statistical classifiers, such as k-nearest neighbor, decision tree, and Bayesian are

explained below. Coding of the algorithms for statistical classifiers were done in Matlab programming language.

3.5.1.2.1 Single Backpropagation Neural Network Classifier

The same methodology used to construct the backpropagation neural network with the pixel intensities in section 3.5.1.1 was used here. However, as the number of textural features was three, only three input nodes were used in the input layer of the network. The number of nodes for the hidden layer was selected by experimenting with different numbers of nodes from one to ten. Five neurons in the hidden layer resulted in the maximum classification accuracy and minimum error rate. A neural network with no hidden layer was also used, but no convergence was obtained. Final settings of the network were 3-5-2 as the number of nodes in input layer, hidden layer and output layer respectively. Similar to the previous neural network model, learning rate and momentum coefficients were both selected as 0.25 based on the trial and error method. All other design parameters were the same as with the previously explained neural network.

3.5.1.2.2 Statistical Classifiers

The same three texture features described for the neural network classifier in previous section were used in the following statistical classifiers. Classification error was estimated by calculating the probability of misclassification, as it was described in the section for neural network above.

K-Nearest Neighbor Classifier

Detailed information on this classifier is given in section 2.3.3. The only parameter that had to be chosen for this classifier was the value of K, which was the number of the neighbor members included in the distance measurements. Values of K were chosen as one and three in this study. Test pattern in the 1-nearest neighbor classifier was assigned to the closest class based on the Mahalanobis distance (section 2.3.3). However, in the 3-nearest neighbor classifier, the output class was chosen if it had the two closest members to the test pattern.

Decision Tree Classifier

Some theoretical background on construction of decision trees is given in section 2.3.4 in the previous Chapter. The probability model was used in partitioning the nodes in the decision tree.

Selection of the features to be used in classification and the thresholds required to split the nodes were done using S-plus software (Venables and Ripney, 1994). S-plus software constructs the decision tree by processing the training data based on the tree-partitioning algorithm used. The outputs from this process are the features selected for nodes and their thresholds. In another explanation, three texture features are further evaluated to find whether there is an effective subset of features to split the tree. If all the features are found effective, then only the threshold values are determined in this process. Later, knowing the effective features and the associated threshold values with them, a decision tree is constructed for classification of testing data by means of the “if then” rules.

Bayesian Classifier

In this parametric approach, which was explained in detail in section 2.1.3.4, data was assumed to have a multivariate Gaussian distribution that was given in Equations 2.28 through 2.35.

Prior probabilities of $P(w_i)$ in Equation 2.28, where $i=1, 2$ for a two-class classification problem and $i=1, \dots, 5$ for a five-class classification problem, were assumed to be equal for each class. Unknown parameters of μ and Σ , which are mean and covariance matrix for each class, were estimated from training samples using Equations 2.30 and 2.31. Using these parameters, discriminant functions for each class were calculated and later used to separate the testing patterns.

3.5.1.3 Classification Using Combined Features from Images

In this experiment, to attempt to improve classification accuracy, features of pixel intensities and texture from one or more images were combined in different ways. Only single backpropagation neural network classifier was used in each classification application with combined features because other classifiers cannot efficiently handle the high number of inputs which occur when all pixels are used.

Coefficients of learning rate and momentum used in the neural network were both 0.25. In constructing and using the backpropagation neural network for this application, sometimes the number of nodes in both input and hidden layers and sometimes only in the input layer were changed based on the experiment. All other operations were the same as with the applications of previous single neural network classifier.

3.5.1.3.1 Using Pixel Intensities From Two Images at Different Wavelengths

The two most effective wavelengths in identifying a defect were determined based on the highest classification success and lowest error rate obtained in the classification application with single backpropagation neural network. Later, pixel gray values from the two selected images were concatenated resulting in a total of 2×4800 features. This application doubled the number of input features to the neural network increasing the time for training. The increase in the time for training depended on the application. Nodes in the hidden layer were doubled (twenty nodes). Using twenty nodes in the hidden layer enabled the network to learn better and to test more accurately compared to using ten nodes as in the case of using pixel intensities from only one image.

3.5.1.3.2 Using Pixel Intensity and Texture Features

Pixel gray values and texture features were combined as one feature set although the source of features used was the same image (not necessarily the same wavelength). The total number of features was 4803, 4800 belonging to pixel gray values and 3 being the textural features.

Similar to the previous application of combining the features, the number of nodes in the input layer was changed to 4803. However, the number of nodes (ten) in the hidden layer was not changed as increase in the number of input nodes was only three.

3.5.1.3.3 Combining Two Images by Averaging

In this experiment, two images at different effective wavelengths were combined in one image by averaging them. By this operation it was expected to include information

from different wavebands into one image. Similar approach was used by Geoola et al. (1994) for bruise detection using a spectrophotometer method. Selection of the two images at two different wavelengths for averaging was based on the best effective wavelengths found for the defective group of apples at hand. Different combinations of the wavelength pairs among the best wavelength group for a specific defect were tried for possible further improvement in the classification success.

3.5.2 Classification with Five Output Classes

In this classification category all defects were considered together in the classification problem. A total of five output classes were used for each variety; these five classes for Empire variety included apple images with good tissue, bruise, leaf roller, puncture, and as the fifth group, the combination of stem and calyx in the view. For the Golden Delicious variety they were similarly apple images with good tissue, bruise, russet, bitter pit and as the last group the combination of stem and calyx in the view. Although stem and calyx were considered separately in the two-class (good-stem, or good-calyx) classification, for efficiency, they were combined into one group in five-class classification as they were similar to each other in nature and had very close effective wavelengths in both varieties. Wavelength 740 nm was chosen as the effective wavelength for this combined class.

Similar tasks that were performed for the two-class classification category in section 3.5.1 were repeated for the five-class classification category. No major change took place in the structures of the classifiers except the portions related with the output classes. While the number of the nodes in the input layer and hidden layer of the single

backpropagation neural network did not change, the number of nodes in the output layer was increased to five. Other settings were the same as with the previous application of the neural network.

A Logistic Regression in SAS software was used in statistical analysis of the classification results from 5-class classification. Outputs from classifiers were assumed to have a binomial distribution. On the other hand, Miller and Delwiche (1991) did not suggest to use a F-test analysis of variance associated with Mahalanobis distance to test for a significant difference in class means considering the number of degrees of freedom. This conclusion was drawn, as the significance levels would only be valid when the groups are multivariate normal with equal covariance matrices.

Statistical classifiers required same modifications in their structures to use them in five-class classification. In the k-nearest neighbor classifier, the number of distance measurements between the testing pattern and all the available training patterns increased with an increased number of classes. A similar modification was made in the Bayesian classifier. In the decision tree classifier, the number of final nodes increased with an increased number of classes. In short, more computations were needed in each classifier with increased number of classes.

A different classifier, which is the multiple backpropagation neural networks, was used in 5-class classification in addition to the other classifiers. Information on the application of this classifier is given below.

Design of the Multiple Backpropagation Neural Network Classifier

Single backpropagation neural network classifiers developed earlier for detection of individual defects were combined in a special way to form the multiple backpropagation neural network classifier. In this classifier, more than one single neural network classified a test pattern at the same time. Then, decision for the output class of the test pattern was made based on the maximum output obtained from the single classifiers.

The single backpropagation neural networks used in this classifier are the classifiers that were constructed for the classification of the specific defects against the non-defective apples at the best effective wavelengths. For instance for Empire, four single neural network classifiers were used in the multiple backpropagation neural network classifier. Each of these classifiers was obtained from the training operations between the good apple image group and the apple image groups with leaf roller, puncture, stem and calyx, and bruise sides in the view respectively (Figure 3.4). Each individual classifier in the multiple neural network classifier was trained using one specific defect against the good apple group. Therefore, all the classifiers have a memory of good apples although there is no separate classifier for good apples.

The test pattern at the four effective wavelengths of the four single neural network classifiers (four images of the same testing apple) is input into the multiple neural networks and classified by each single classifier separately. A total of sixteen outputs are obtained from the multiple neural network classifier, four outputs per single classifier. In the next step, the first task is to find whether the test image is a defective or a non-defective apple based on the majority of decisions made as defective or good by

individual neural network classifiers. If the majority of the decisions are for good apple then it is decided as a good apple. However, if the majority of decisions are for defective apples then, another stage of decision-making starts to assign the testing image into a specific defect group. In this category, the decision and consequently the classifier with maximum output (weight) is determined to find the defect type.

This approach is depicted in Figure 3.4: Four images of the same testing image at four different wavelengths are represented as I_1 , I_2 , I_3 , and I_4 . Sixteen outputs obtained by processing the four inputs are represented as $O_1, O_2, \dots, O_{16}$. For simplicity, only one image (I_1) and its regarding outputs are used in explaining the classification procedure.

First, image I_1 is input into the multiple neural networks. It is classified by each single neural network classifier in the multiple neural networks. A total of four outputs are obtained; one from each single neural network as O_1, O_5, O_9 , and O_{13} . First, these four decisions are checked to find if they are for good apple group or a defective apple group by comparing the weights they were assigned.

For instance, if the weight set assigned to O_1 is (0.00078 0.98870) then, this decision is for a leaf roller defect; or if the weight set is (0.99784 0.00541) then, the output O_1 is a decision for a good apple. In the result, if the numbers of decisions for good apple group and the defective apple group are equal or the number of decisions for defective group is more, image I_1 is classified as a defective apple. Otherwise it is classified as a good apple. However, if it is desired to find the specific defect type then, one more step is applied comparing the weight pairs for the defect decisions.

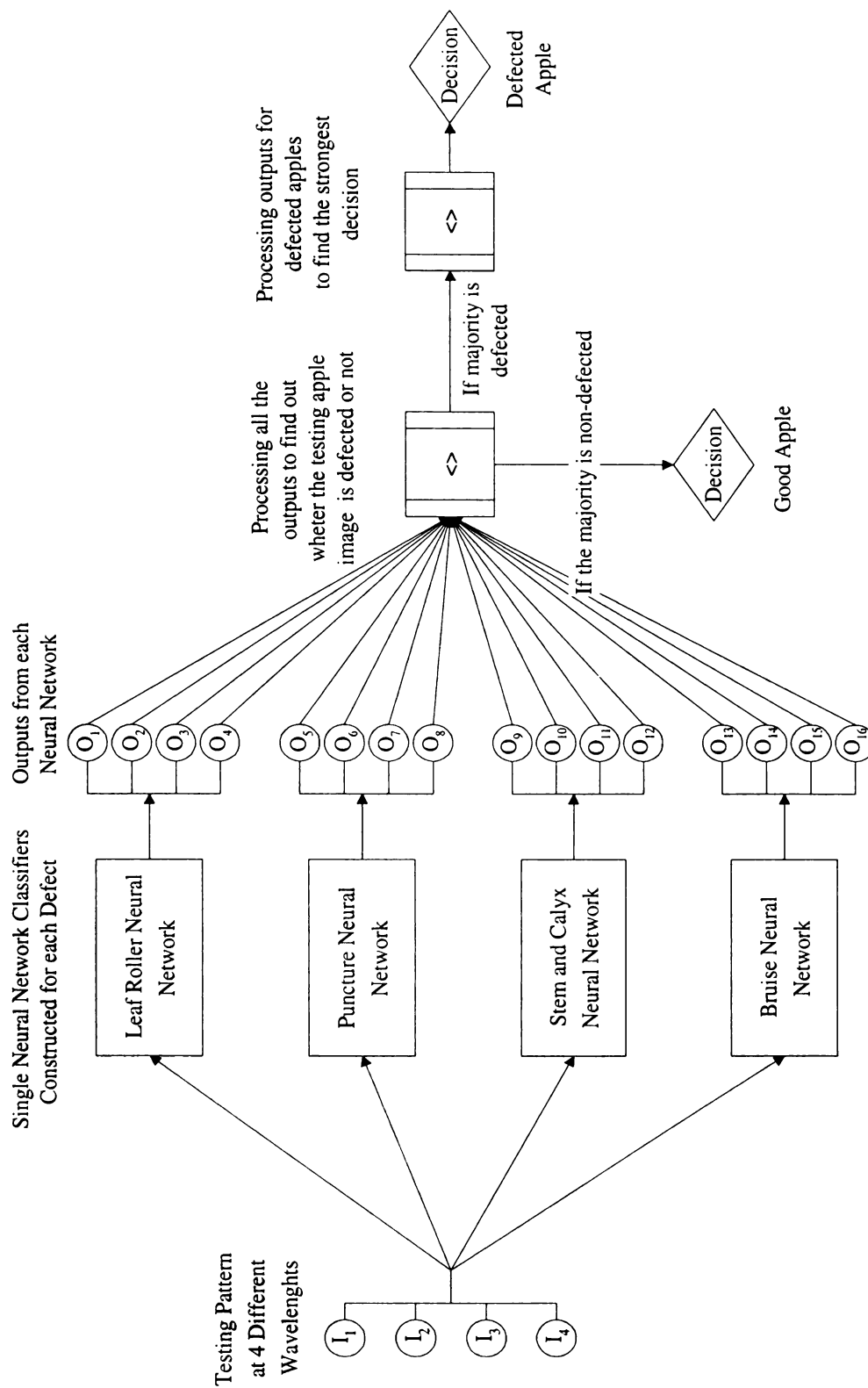


Figure 3.4 A schematic description of the multiple backpropagation neural networks

Chapter 4. RESULTS AND DISCUSSION

This study used neural networks and statistical classifiers to classify multispectral image information of apples with defective, healthy, or stem-calyx sides in the view. Uniqueness of the study was to use the pixel intensity and texture information either alone or combined from all the pixels in an image as features, minimizing image processing tasks.

4.1 Image Processing

Image processing tasks of reducing resolution, background segmentation, and deletion of the specular reflectance areas were used to increase the efficiency of the classification system, to enhance the images and to remove any biased information (specular reflectance) from the images.

4.1.1 Reducing the Resolution of Images

Four different image resolutions were tested as shown in Table 4.1. Decreasing the resolution did not decrease classification success or increase error rate. Classifications of images of defective apples against images of healthy apples were done using the single backpropagation neural network classifier. Successful results with images at reduced resolutions imply that spectral and textural information in the images were preserved even if the image resolution was reduced. Moreover, better classification accuracy was obtained using the resolution of 60x80 pixels for bitter pit defect on Golden Delicious apples compared to using the original image resolution. This may be interpreted as a

result of an improvement in the images introduced by averaging (filtering) during the image resizing operations. Based on this result, for all further applications in this study, images at the resolution of 60x80 pixels were used.

Table 4.1 Classification results obtained from using images at different resolutions

Surface Characteristics	Wavelength (nm)	Resolution (Pixels) *	Classification Success (%)	Accuracy Rate (%)†
Bruise (Empire)	880	480x640	98.1	98.7
		240x320	96.2	97.2
		120x160	96.2	97.3
		60x80	98.1	97.5
Calyx (Empire)	700	480x640	88.9	92.4
		240x320	91.7	92.8
		120x160	88.9	92.0
		60x80	88.9	93.3
Bitter Pit (Golden D.)	1320	480x640	98.0	97.4
		240x320	100.0	99.7
		120x160	100.0	99.4
		60x80	100.0	99.7
Russet (Golden D.)	1200	480x640††	81.1††	85.5††
		240x320	86.5	89.0
		120x160	89.2	89.8
		60x80	89.2	92.5

† Calculated using the operation of $(100 - (100 \times \text{Error Rate}))$, †† Results were obtained at higher error rate than the targeted one (0.002): iteration was stopped at 6800 as the convergence was too slow, * Number of apples used in this experiment are given in Table 3.4.

Times spent in training and testing the single neural network classifier at different resolutions for the selected defects on Golden Delicious and Empire apples are presented in Table 4.2. As it can be seen from the table time spent, especially in training, decreases significantly as the resolution decreases. It should be noted that high similarity between defective and non-defective tissues increases the time period in training and also in testing, as can be seen in classification of apples with russet defect against apples with good tissue. Despite the variations in time periods for training, time spent in testing an

apple image did not vary much among the defect types (between 0.02 and 0.05 at 60x80 pixels resolution)

Table 4.2 Training and testing times for two-class classification with neural network

Surface Characteristic (Wavelength)	Resolution (Pixels)	Training Time per Apple (s)	Testing Time per Apple (s)
Bitter Pit, (Golden D., 1320 nm)	480x640	490.41	4.09
	240x320	81.82	0.78
	120x160	9.55	0.11
	60x80	2.18	0.02
Russet, (Golden D., 1200 nm)	480x640	-†	-†
	240x320	543.75	0.51
	120x160	26.25	0.16
	60x80	5.00	0.05
Bruise, (Empire, 880 nm)	60x80	2.31	0.04
Calyx, (Empire, 700 nm)	60x80	3.57	0.03
Leaf Roller, (Empire, 600 nm)	60x80	3.48	0.03
Puncture, (Empire, 680 nm)	60x80	3.46	0.03

† Iteration was stopped at 6800 due to very slow convergence.

4.1.2 Background Segmentation

Background subtraction was applied to improve the images by eliminating any possible confusion between the apple in the scene and the background. However, classification operations using the original images that included background also resulted in successful classification accuracies. In Table 4.3 classification results at the most effective wavelengths for bruise and calyx on Empire and for russet on Golden Delicious are given as examples. Classification results at other wavelengths obtained from using images with and without background can be seen in Figures 4.1, 4.2 and 4.3 for the defects in Table 4.3.

In general better results were obtained using images with background compared to using images without background. However excluding the background was still

Table 4.3 Comparisons of classification results obtained using images with and without background at most effective wavelengths

Surface Tissue		With Background		Without Background	
(Empire)	Wavelength (nm)	Classification Success (%)	Accuracy Rate (%)†	Classification Success (%)	Accuracy Rate (%)†
Bruise	880	98.1	97.8	98.1	97.5
Calyx	700	94.4	94.7	88.9	93.3
(Golden D.)					
Russet	1200	91.9	91.9	89.2	92.5

† Calculated using the operation of $(100 - (100 * \text{Error Rate}))$

important in this research to make sure the classifiers see only the radiance from the apple as there might be some variations in the reflectance from the background in the wavelength range used.

Adding noise to the training data was suggested as a method of improving learning ability of the backpropagation neural network (Kasabov, 1996). Local minima, which is a very common problem in backpropagation learning, can be overcome by adding noise to the input data. Images may have contained some noise coming from the light source or camera etc. With the existence of background this noise would be more than that of the images with no background as the noise would exist in the pixels in background too. This may be an explanation for slightly better classification results obtained using the images with background.

In conclusion, finding the same most effective wavelengths for the defects given in Table 4.3 using images either with background or without background suggests that background subtraction is not necessary for defect detection. Thus, background segmentation can be eliminated from future image processing tasks.

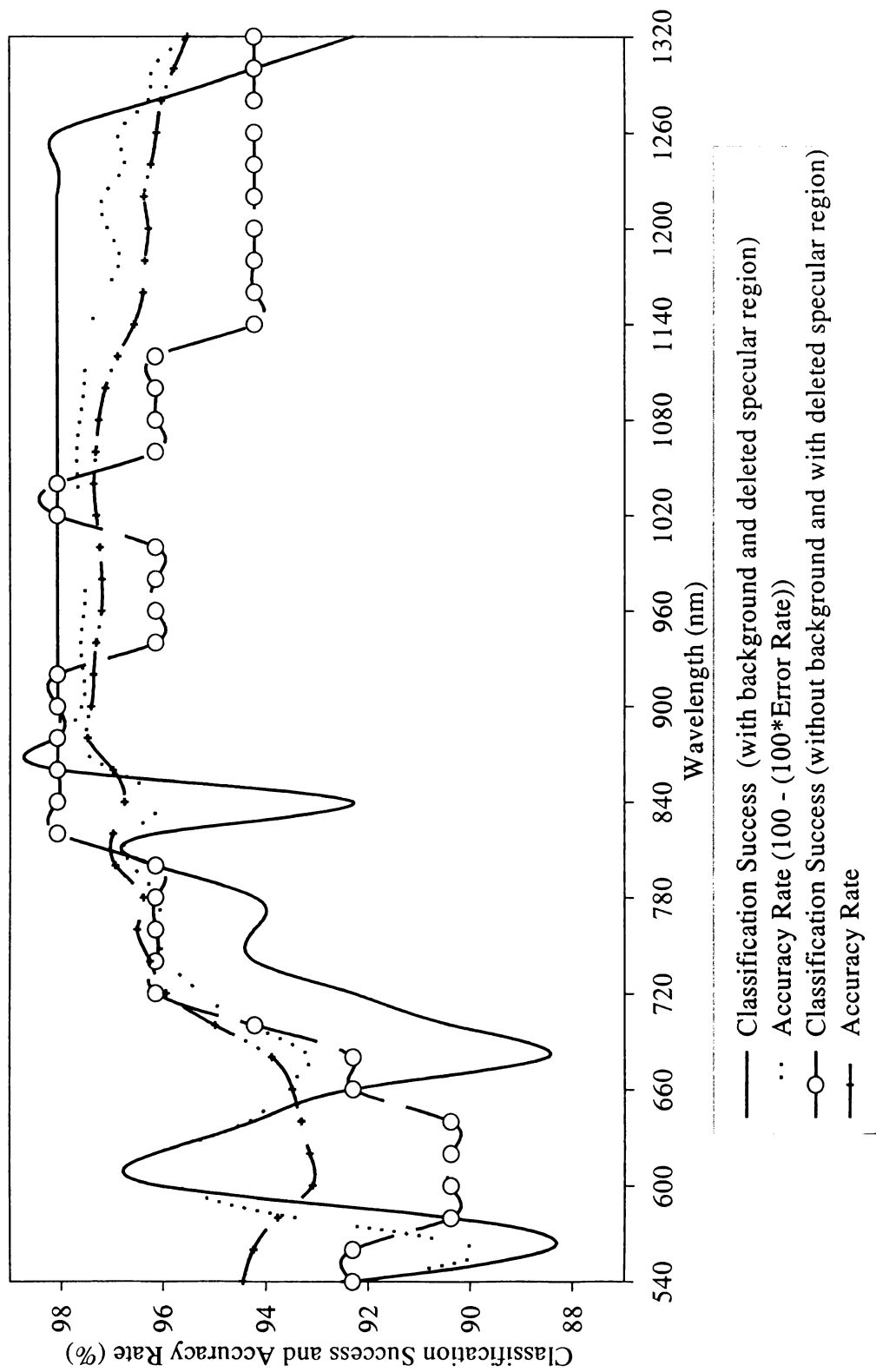


Figure 4.1 Classification of bruised apples against good apples (Empire)

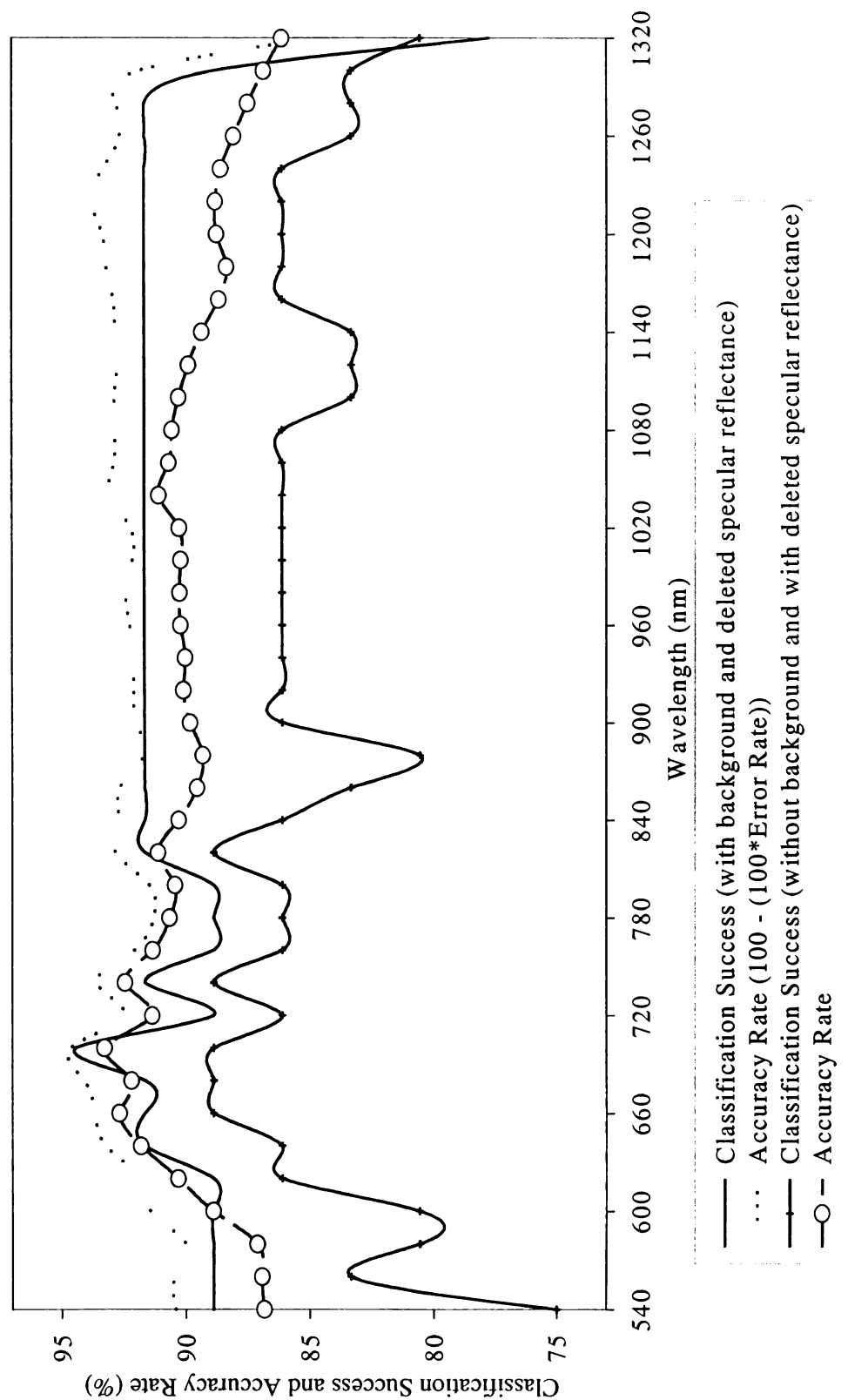


Figure 4.2 Classification of calyx view images against non-calyx (good) images (Empire)

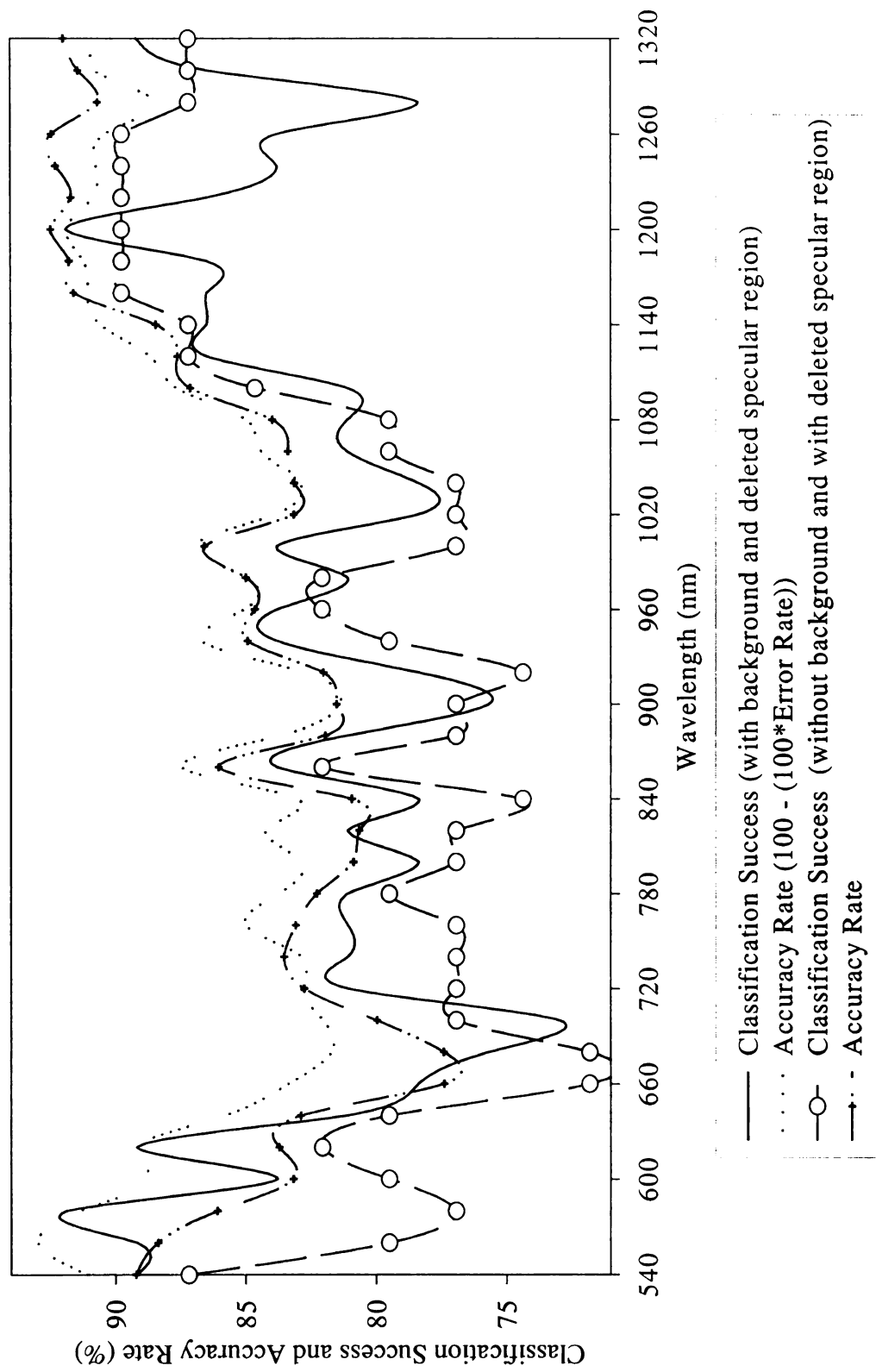


Figure 4.3 Classification of images with russet against good apple images (Golden Delicious)

4.1.3 Elimination of Specular Reflectance from Images

Specular reflectance regions were removed from the images as they were shown to bring a biased feature to the images with their specific shapes and reflectance values, which appeared to depend on the defect type (section 3.3.2). Examples of images with specular areas can be seen in Figure 4.4 where, for instance, the shapes of specular reflectance regions in calyx or stem view images tend to be long while appearing rounder in a non-calyx view (good tissue) images.

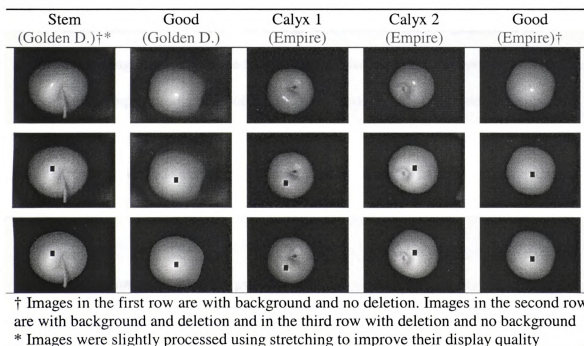


Figure 4.4 Examples of apple images with calyx and good tissue views in Figure A.4

Results of the classifications of calyx view images against the non-calyx (good tissue) images with deleted specular regions are shown in Figure A.4 in Appendix A. Classification accuracies obtained using the original images with background and no deletion of the areas with specular reflectance were 100% at and after the wavelength 620 nm. For instance, in calyx identification at the most effective wavelength of 700 nm

classification success using the images with background and no deletion was 100% while results from other classification applications that used the images with deleted specular reflectance areas were lower, 94.4% for the images with background and deletion and 88.9% for images with no background and with deletion.

Having a high classification success using images with the specular reflectance areas in classification of calyx view images against the non-calyx images shows that the existence of the specular regions biases the classification. By deleting the specular reflectance regions, images were cleared of the biased features.

It was shown that images with background could be used in classification applications using pixel gray value features. If the image acquisition system is carefully designed, image-processing tasks of reducing the resolution and deletion of the specular points could be eliminated in the classification procedure; for instance, images can be taken directly at desired smaller resolutions. Formation of the specular reflectance areas may be prevented with improvements in lighting and camera configurations.

4.2 Classification with Two-Output Classes

Each defect or calyx/stem view image was studied individually against apple images with healthy tissue view under lighting from visible and NIR regions to determine the effective wavelengths in recognizing the defective (or stem/calyx) apple images.

4.2.1 Classification and Wavelength Selection Using Single Backpropagation Neural Network Classifier and Pixel Intensity Values

It should be noted here that classification and wavelength selection were conducted simultaneously as part of the single backpropagation neural network experiment.

Classification results at effective wavelengths are reported for specific defects as a result of classification procedure.

Effective wavelengths selected for identification of defects can be categorized in two groups based on classification accuracies obtained for classification of the selected defect (stem/calyx). Effective wavelengths in the first group are the most efficient ones as they enabled the system to classify a defect (stem/calyx) with minimum error rate and the highest classification accuracy. It should be noted here that when the same classification success was obtained using the images at two different wavelengths, the wavelength that is associated with the lower error rate is selected as the most effective wavelength. Effective wavelengths in the second group provide the next highest classification success after the wavelengths in the first group. Based on these definitions the selected effective wavelengths are given in Table 4.4 for each defect and stem-calyx on the two apple varieties.

Wavelength 680 nm was the most effective for discrimination of puncture on Empire apples. Wavelengths 660 and 700 nm having the same classification success with wavelength 680 nm were in the first wavelength group although they had slightly higher error rates compared to wavelength 680 nm. Wavelengths 1020 and 1100 nm on the other hand were the most effective wavelengths in the second group of wavelengths that resulted in lower classification successes and higher error rates compared to the effective wavelengths in the first group mentioned above. These observations can also be seen in Figure A.2 of Appendix A. It should be noted that reflectance for puncture does not change much as wavelength changes. This may be due to the loss of water, which was

released from the damaged tissue through the broken skin leaving the punctured area dry (Figures 4.5 and A.2).

Table 4.4 Three effective wavelengths for individual surface characteristics

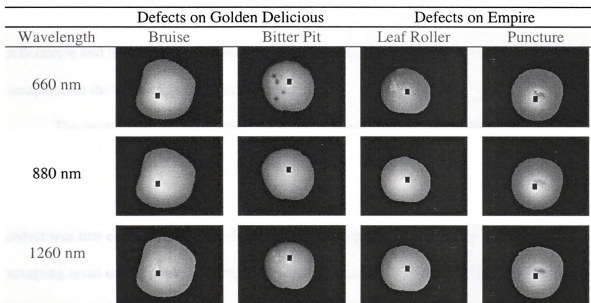
Surface Characteristic (Empire)	Wavelengths in the First Group (nm) †	Wavelengths in the Second Group (nm) ††
Bruise	880 , 900, 920	1060, 940, 1080
Calyx	700 , 660, 740	640, 720, 760
Stem	740	760, 780
Puncture	680 , 700, 660	1020, 1100, 880
Leaf Roller	600 , 580, 1320	1000, 940, 1020
(Golden D.)		
Bruise	1260 , 1300, 1280	1240, 1220, 1180
Calyx	660 , 680, 620	640, 700, 560
Stem	760	900, 920, 720
Bitter Pit	1320 , 1300, 1280	1240, 1200, 1180
Russet	1200 , 1260, 1240	1320, 1300, 1280

† Wavelengths that provided the highest classification success. †† Wavelengths that provided the second highest classification success. Wavelengths in two groups are in the order from the lower to the higher error rate. Maximum three wavelengths are shown.

Leaf Roller defect on Empire apples had the most effective wavelengths both in the visible and NIR regions. The best three wavelength bands were 600, 580, and 1320 nm in the order from the lower to higher error rate. At these wavelengths all apples were classified correctly. Variation in the classification accuracy depending on the wavelength of this defect is shown in Figure A.3 of Appendix A.

The most effective wavelength for bitter pit defect on Golden Delicious apples was 1320 nm. Two other effective wavelengths that were in the same group with wavelength 1320 nm with slightly higher error rates were 1300 and 1280 nm. The change in the appearance of an apple with bitter pit at different wavelengths in parallel with the change in classification results from the single neural network classifier may suggest the following interesting conclusion; although the bitter pit defects on the apple at

wavelength 660 nm in Figure 4.5 looked much clearer than those at wavelength 1260 nm, better classification results were obtained using the images at wavelength 1260 nm. Difference in the appearances of the defects at two wavelengths was mainly in their gray values. While the gray values of the defective areas were quite low (dark) at 660 nm, they turned into much higher almost opposite gray values (light) at 1260 nm.



† Images were slightly processed using stretching to improve their display quality

Figure 4.5 Appearances of defects at different wavelengths

The same relation can be seen for the apple with leaf roller in Figure 4.5. However, for leaf roller the higher gray values were obtained in the beginning of the spectral range (540-1320 nm) contrary to the bitter pit. In short, for both defects improved classification results were obtained at wavelengths where the defective areas had higher gray values (light color) than the surrounding healthy regions. Having the similar results from using the images with undeleted specular regions that had extremely

high gray values, it may be suggested that the backpropagation neural network gives more value to the pixels with high gray values.

Dark reflection from the bitter pit defects at 660 nm may be due to absorption of the light by some elements in the defective region. On the other hand, high reflectance from the same blemished areas at wavelength 1260 nm can be a sign of some existing constituents that have less absorbing ability than the normal tissue. Another possibility may be that light reflected from the blemished area may contain transmitted light, body reflectance and the regular reflectance together making the region lighter in color compared to the surrounding healthy region.

The most effective wavelength for another important defect, russet on Golden Delicious apples, was 1200 nm. Two other effective wavelengths in the order from lower error rate to the higher were 1260 and 1240 nm. Overall classification success of this defect was low compared to other defects. This may be due to the special formation of russetting areas on the apple surface. There did not seem to be a clear distinction between the russetting areas and the healthy tissue on the apple surface. Three of the four misclassified apples were good ones, which were falsely rejected as russet at wavelength 1200 nm. This may suggest that any slight russet patches on the good apples, which might have been ignored in selecting the apples for training and testing before image acquisition, might have caused the good apples mentioned above to be falsely rejected. This shows that separation of good apples from apples with the russet defects was a challenge for the neural network. Similar findings were reported in classifying Golden Delicious apples with russet defect by Aneshansley et al. (1997). High classification error

was reported for classifying apples with russet against good apples using Mahalanobis Distance classification method.

Successive wavelengths from 840 to 920 nm were the most effective wavelengths for discrimination of bruised apples from good apples in the Empire variety. The wavelength having the lowest error rate among them was 880 nm. Having the effective wavelengths around the 900 nm region of the NIR can be explained by the interaction between the released water from the corrupted cells and light (Weisskopf, 1968). At this wavelength water absorbs light more than the surrounding good tissue, causing the bruised area to look darker depending on the severity of the bruise.

The most effective wavelength to differentiate bruise on Golden Delicious apples was found to be wavelength 1260 nm as it gave the highest classification accuracy and yet the lowest error rate. As can be seen in Figure 4.6, bruise was classified more accurately at the wavelengths beyond 1000 nm, especially beyond 1200 nm. This suggests an increase in the absorption of the light at longer wavelengths by the constituents in the defective area and, thus, greater reflectance differences between good and defective tissue. Furthermore, depending on the location of the bruise, a texture that formed inside the bruise area created another feature for the classifiers. The change in the reflectance from one wavelength to another in the bruised and other defect areas can also be seen in the images in Figure 4.5 as a change in the contrast between defective and good tissues. Changes in contrast in the defective areas matches the changes in the classification successes and error rates in the respective graphs given for each defect as follows: Figures A.2 and A.3 in Appendix A for puncture and leaf roller respectively on

Empire apples; Figures 4.6 and A.7 for bruise and bitter pit respectively on Golden Delicious apples.

Although the best wavelengths for detecting the bruised regions on Golden Delicious apples were longer than 1000 nm, wavelength 880 nm was efficient for discriminating bruise on Golden Delicious apples. Thus, wavelength 880 nm was demonstrated to be a common wavelength for bruise detection on both Golden Delicious and Empire varieties.

For both varieties the effective wavelengths for calyx and stem were in the visible range between 660 and 740 nm as was expected. Unlike the defects mentioned above, there was no corruption in the apple cells that would affect the reflection of light, thus, the clearest appearance of the calyx and stem were obtained in the visible range of the spectra (Figures A.4 and A.5 for Empire and Figures A.8 and A.9 for Golden Delicious in Appendix A).

4.2.2 Classification Using Texture Features

After determining the effective wavelength(s) for each specific defect for both varieties in the previous section, three texture features of angular second moment, contrast and correlation explained in sections 2.1.1 and 3.4.2 were extracted from the images at the selected wavelengths. These texture features were used in the classifiers of single backpropagation neural network, k-nearest neighbor, decision tree and Bayesian.

The best two classification results for each variety and defect using the classifiers mentioned are shown in Table 4.5. Detailed information on the same results is shown in Tables A.3 and A.4 in Appendix A.

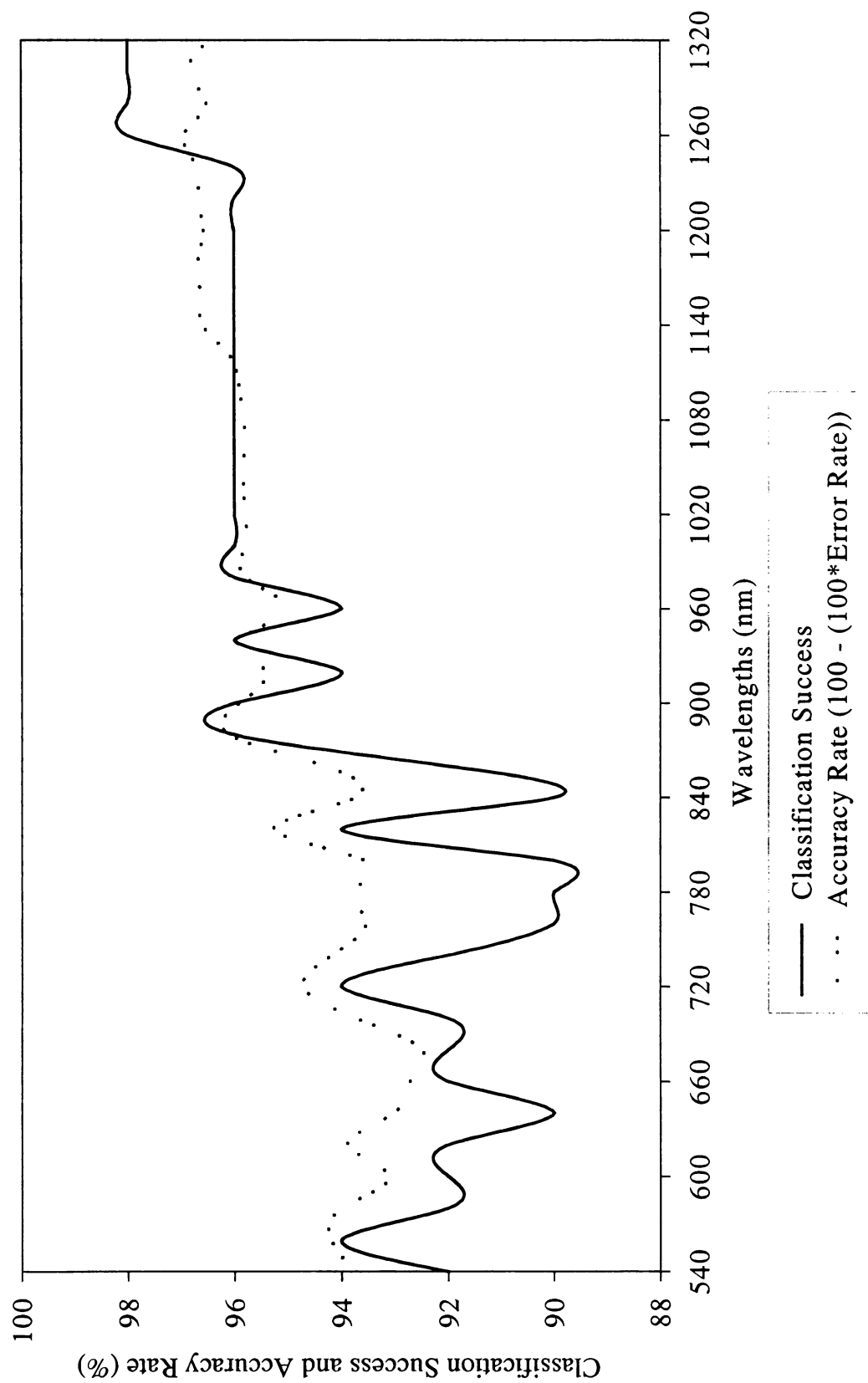


Figure 4.6 Classification of bruised apples against good apples (Golden Delicious)

Table 4.5 The best two classifiers in two-class classification

Surface Characteristic and Wavelength (nm)	Classifier	Classification Success (%)	Accuracy Rate (%)†	Number of Mis classified Apples
(Empire)				
Bruise, 880	SBPNN-TeX ^b	100.0	99.9	0
	1-NN ^c	100.0	100.0	0
Leaf R., 600	SBPNN-Pixel ^a	100.0	99.6	0
	Decision Tree ^d	100.0	100.0	0
Stem, 740	SBPNN-Pixel ^a	100.0	99.3	0
	SBPNN-TeX ^b	91.7	92.6	3
Calyx, 700	SBPNN-TeX ^b	100.0	99.1	0
	SBPNN-Pixel ^a	88.9	93.3	4
Puncture, 680	SBPNN-Pixel ^a	96.7	97.5	1
	Decision Tree ^d	93.6	93.6	2
(Golden D.)				
Stem, 760	SBPNN-Pixel ^a	100.0	98.5	0
	SBPNN-TeX ^b	100.0	99.8	0
Calyx, 660	3-NN ^c	100.0	100.0	0
	SBPNN-TeX ^b	97.1	98.7	1
Bitter Pit, 1320	SBPNN-Pixel ^a	100.0	99.7	0
	SBPNN-TeX ^b	94.7	95.1	2
Bruise, 1260	SBPNN-Pixel ^a	98.0	96.9	1
	Bayesian ^d	98.0	98.0	1
Russet, 1200	Bayesian ^d	91.9	91.9	3
	SBPNN-Pixel ^a	89.2	92.5	4

† 100-(100*Error Rate), (^a) Single Backpropagation Neural Network Classifier using Pixel features, (^b) Single Backpropagation Neural Network Classifier using Texture Features, (^c) Nearest Neighbor Classifier using texture features, (^d) Using texture features

For bruise defect on Empire there was no significant difference between the classification results obtained from different classifiers or input features. However, the best classification results were obtained using texture features with the classifiers of single backpropagation neural network and 1-nearest neighbor (SBPNN-TeX and 1-NN in Table 4.5). Both classifiers classified all the apples correctly. On the other hand, for bruise defect on Golden Delicious, which had the most effective wavelength at 1260 nm, improved classification results were obtained using pixel intensity features with single backpropagation neural network classifier (SBPNN-Pixel in Table 4.5) and using texture

features with Bayesian classifier (98%, classification success). Classification results obtained from any classifier with either pixel or texture features was 90% or higher for this defect on both varieties.

For identification of the apple stem end images, single backpropagation neural network classification either using pixel gray values or texture features produced better results compared to other applications (Tables 4.5) for both varieties. This classifier recognized all the stem images successfully for both varieties. For Golden Delicious apples using only texture features with the same neural network classifier produced the same results recognizing all the images successfully. Results of stem recognition from other classifiers for Golden Delicious were also successful, with the classification success rates of 94.9% and above.

Similar results were obtained in calyx recognition on Empire apples. The neural network strategy performed better classification, (recognizing all the images successfully), especially using texture features. Identification of images of Golden Delicious apples with calyx in view was as successful as it was with stem identification for most of the classifiers and features used except decision tree classifier. The 3-nearest neighbor classifier recognized all the testing images successfully. Neural network classifiers and the Bayesian classifier followed this classifier in performance.

All classifiers identified the leaf roller defect on Empire with classification successes ranging from 97.2% to 100%. The neural network using pixel intensity values and the decision tree classifier using texture features recognized all the apple images with leaf roller successfully.

For the puncture defect on Empire, the best classification result of 96.7% was obtained using the single neural network classifier that used the pixel gray values as features. The second best results were obtained using the texture features with single neural network and decision tree classifiers.

Russet on Golden Delicious apples was best classified by the Bayesian classifier with a classification success rate of 91.9% while the classification successes from other classifiers ranged between 78.4% and 89.2%. Severity of russet defect was variable causing confusion between apple images with russet defect and good tissue in the view.

Bitter pit defect on Golden Delicious apples was classified at success rates from 92.1% to 100%, having the best result from the single neural network classifier using pixel gray values as features.

4.2.3 Classification Using Combined Features

Different sets of features were used together in the single backpropagation neural network classifier to attempt further improvement in the classification performance.

4.2.3.1 Using Pixel Intensities From Two Images at Different Wavelengths

In the wavelength selection procedure for individual defects in the two-class classification category, it was seen that most of the defects had effective wavelength(s) from close regions of the spectra. Therefore, combining two images using the pixel gray values at two neighbor effective wavelengths did not improve the classification accuracy further compared to using pixel gray values from a single image at the best effective wavelength. This was probably because there was not any additional information

introduced to the classifiers by using an extra image which was very similar to the first one.

However, for bruise and leaf roller (Figures A.1 and A.3 in Appendix A) defects on Empire apples and for stem on Golden Delicious apples (Figure A.9 in Appendix A) combining two images improved the classifications resulting in slightly lower error rates, although the classification success was the same. Contrary to other defects, bruise and leaf roller defects on Empire and stem on Golden Delicious had effective wavelengths from two regions which were to some extent apart from each other in the spectra. For the rest of the defects on the two varieties, combining images negatively affected the classifications causing a decrease in the classification success in general. Examples of the results of combining pixel gray values from two images using the backpropagation neural network are given in Table A.1 in Appendix A.

4.2.3.2 Using Pixel Intensities and Texture Features

Combining pixel gray values and texture features together produced similar results as were obtained by using pixel gray values from two images (Table A.2 Appendix A). Similar to using pixel gray values from two images some slight decreases in the error rates in the classifications of leaf roller on Empire and stem on Golden Delicious were seen while no improvement in classification success was obtained.

4.2.3.3 Combining Two Images by Averaging

Using the feature combination method of averaging two images at different wavelengths into one image helped to reduce the error rates in classifying bruise, stem

and russet on Golden Delicious apples although no improvement was obtained in the classification success.

Despite drops in the error rates, classification success did not improve in any of the feature combination methods used in two-class classification. However, decrease in the error rate of a classification is still an improvement as error rate is the only criterion used in weight adjustment during training process of the neural network. Thus, having a lower error rate at the end of the classification is a sign of having a more robust classifier.

4.3 Classification with Five-Output Classes

In addition to the classifiers described in section 3.5.2 for two-class classification, a multiple backpropagation neural network classifier was used to classify all tissue types (5) in a single classifier. The only change applied to the neural network was to increase the number of nodes from 2 to 5 in the output layer. Also the classification tasks applied in the statistical classifiers were repeated five times instead of two.

4.3.1 Classification and Wavelength Selection Using Single Backpropagation Neural Network Classifier and Pixel Intensity Values

It should be noted that classification and wavelength selection are the same applications. Classification accuracy and error rate at the effective wavelengths are reported as the optimum classification results obtained for 5 apple groups considered.

Similar to the two-class classification, 40 apple images collected in the range of 540-1320 nm were used in a single backpropagation neural network classifier. Forty classification results, obtained one at each bandpass increments, were compared to find the effective wavelength(s) for the classification of a group of apples containing five

subgroups in it. Results from this experiment are presented in Figure A.6 in Appendix A for Empire variety and in Figure A.11 for Golden Delicious. Also some selected results, which include the most optimum classification results, from this application can be seen in Tables A.5 and A.6 in the same Appendix.

If the confusions among the defective groups and among the good apple groups such as good apple, stem and calyx are not considered, there is no misclassification among the 5 subgroups at wavelengths 740, 1000, 1080 and 1100 nm for Empire (Table A.5). The classification and accuracy rates were 100% and 96.2% respectively using the images at the most effective wavelength (740 nm):

$$\text{Where the Accuracy Rate (\%)} = 100 - (100 * \text{Error Rate}) \quad (4.1)$$

If the consideration is not to tolerate any confusion between any output classes, no matter if it is between the good apples (stem, calyx and good tissue view apple images) or between the defective apples, then the classification success is 86.6% at the same wavelength (740 nm). If the misclassification between stem and calyx, and good apples is not considered, but misclassification is considered between defects then the classification success is 91.8% at wavelength 740 nm.

For Golden Delicious apples, the most effective wavelength was found to be 1260 nm at which the classification success and accuracy rate disregarding any misclassification among the defective groups and among the good apple groups were 90.7% and 93.8% respectively (Table A.6 in Appendix A). On the other hand, classification success was 84.5% while still considering the confusions between the defective groups but disregarding the confusion between the good apple groups. Lower

classification success rates were obtained for Golden Delicious apples, which may be due to the randomly selected multiple strains and two different harvest dates for this variety.

Not having the same effective wavelengths for Empire and Golden Delicious varieties might be because of different defect types used for each variety with the exception of bruise. Also the slightly different number of samples that were used either in training or in testing might have influenced effective wavelengths. Difference in skin color might also have had an effect on the resulted different wavelengths, as it could affect the appearances of defects.

Time periods spent in training and testing the single backpropagation neural network classifier in 5-class classification application are given in Table 4.6 for two varieties at the selected effective wavelengths. Although the testing time did not change compared to 2-class classification, training time increased significantly in 5-class classification as 5 different groups were used.

Table 4.6 Training and testing times for five-class classification with neural networks

Variety and Wavelength	Resolution (Pixels)	Training Time per Apple (s)	Testing Time per Apple (s)
Empire (740 nm)	60x80	369.23	0.03
Golden Delicious (1260 nm)	60x80	591.04	0.03

4.3.2 Classification Using Multiple Backpropagation Neural Network Classifier

Multiple backpropagation neural networks, as described earlier in section 3.5.1.1.2, are simply a combination of single backpropagation neural network classifiers that were established and selected for each defect (two-class classification) at the most effective wavelength using one defective (or stem/calyx) group of apples and the group of good apples.

For Empire variety, four single backpropagation neural network classifiers were selected from the training operations between good apple group and the defect groups of leaf roller at wavelength 680 nm, puncture at 680 nm, stem-calyx group at 700 nm, and bruise at 880 nm. Wavelength 680 nm was chosen as the common wavelength for leaf roller and puncture classifiers to decrease the number of input images from four to three.

The concept for the classification here was to classify apples in each group at three different wavelengths using four classifiers explained above. For example, an apple with its images at wavelengths 680, 700 and 880 nm is input into 4 different classifiers (leaf roller classifier @ 680 nm, puncture classifier @ 680 nm, stem-calyx classifier @ 700 nm and bruise classifier @ 880 nm) and it is classified into the class whose classifier outputs the lowest error rate.

In classification of a testing apple through each defect classifier, the classifier checks the input defect for its similarity to the defect on which the classifier was originally trained; if there is a similarity then it is highly possible that the testing apple will be classified as the defect the classifier knows. If there is no similarity, the defective apple being tested is classified as good apple as it is the only other choice available in a two-class classification.

Table 4.7 Classification results of an Empire apple with leaf roller at 3 different wavelengths through 4 different classifiers

Wavelength (nm) †	Bruise C.‡ @ 880 nm		Stem-Calyx C. @ 700 nm		Leaf Roller C. @ 680 nm		Puncture C. @ 680 nm	
	Good	Defect	Good	Defect	Good	Defect	Good	Defect
680	0.105	0.898	0.986	0.015	0.068	0.936	0.079	0.924
700	0.081	0.921	0.987	0.014	0.056	0.948	0.064	0.939
880	0.094	0.909	0.988	0.013	0.064	0.938	0.035	0.967

‡Classifier, † An apple image at 3 different wavelengths (680, 700, 880 nm) was inputted into 4 different classifiers; 12 results in the columns of “Defect” were evaluated.

Weights assigned by each defect classifier in the classification of an Empire apple with leaf roller defect are presented in “Defect” columns in Table 4.7 as an example of the multiple neural networks approach. A testing Empire apple image with leaf roller at three different wavelengths was classified by 4 defect classifiers of bruise, stem-calyx, leaf roller and puncture. Total of 12 outputs (weights) are evaluated for the strongest decision. In this example the test apple with leaf roller was classified as a puncture by the puncture classifier and by the combination criteria as the weight assigned by this classifier was the strongest among all (0.967 for image at 880 nm).

It can also be seen in the same table that the stem-calyx classifier classified the same testing apple as a good apple as the leaf roller does not look like a stem or a calyx in nature. Also, the weights assigned by the bruise classifier for the same apple are not as strong as the ones assigned by the puncture classifier or leaf roller classifier as the similarity between bruise and leaf roller defects is weak.

Although the answer to the question of why the testing apple was not classified as leaf roller is not clear, it may be that the puncture classifier trained on the puncture defects, which had in general lower gray values than that of the leaf roller defect, weighted the testing image with leaf roller more than did the leaf roller classifier. The leaf roller classifier was already trained on the leaf roller apples and so the testing pattern with leaf roller would not be too different from the training data.

However, for the puncture classifier a testing pattern that is already similar to the patterns the classifier was trained on and with higher gray values in the defective areas may boost the decision weight produced by the classifier. One reason for this action of the puncture classifier may be that neural networks tend to be driven more by the spectral

features with higher gray values, as was discussed earlier. Or, the reason for a leaf roller defect to be detected better by a puncture classifier instead of a leaf roller classifier may also be due to the specific wavelength selected.

The advantage of a multiple neural network classifier is that it may provide a more robust defect detection system when the misclassification among the defects is not

Table 4.8 Selected classification results from using all of the classifiers and features used for Empire and Golden Delicious varieties (5-Class Classification)

Classifiers‡†	Bad-Good (%)	Bad-Good MC‡	Separate (%)†	Separate MC
(Empire)				
SBPNN-Pixel^a	100.0	0	91.8	8
MBPNN-Pixel^b	99.0	1	84.5	15
SBPNN-2Pixel^d	99.0	1	88.7	11
SBPNN-Px+Tx^c	99.0	1	93.8	6
SBPNN-TeX^c	93.8	6	88.7	11
3-NN^f	87.6	12	81.4	18
1-NN^f	86.6	13	79.4	20
Decision Tree^g	84.5	15	72.3	24
Bayesian^g	76.3	23	63.9	35
(Golden Delicious)				
SBPNN-2Pixel^d	94.9	5	89.7	10
MBPNN-Pixel^b	92.8	7	82.5	17
SBPNN-Pixel^a	90.7	9	84.5	15
SBPNN-Px+Tx^c	89.7	10	85.6	14
Bayesian^g	89.7	10	83.5	16
SBPNN-TeX^c	89.7	10	82.5	17
1-NN^f	89.7	10	79.4	20
Decision Tree^g	86.6	13	81.4	18
3-NN^f	85.6	14	76.3	23

‡† At the most effective wavelengths; 740 nm for Empire, 1260 nm for Golden Delicious, ‡ Number of misclassified apples, † When all defects are considered separately although stem, calyx and good are assumed as one good group, (^a) Single Backpropagation Neural Network Classifier with Pixel features from one image, (^b) Multiple Backpropagation Neural Network Classifier with Pixel features from one image, (^c) Using Texture features only, (^d) Using Pixel features from two images at two most effective wavelengths; 740 and 760 nm for Empire, 620 and 1260 nm for Golden Delicious, (^e) Using Pixel features from one image and Texture features from another (at effective wavelengths)- 740 and 620 nm for Empire, and 1260 and 760 nm for Golden Delicious, (^f) Nearest Neighbor Classifier using texture features, (^g) Using texture features

considered as can be seen in the classification of Golden Delicious apple group containing 5-output groups in this study (Table 4.8). The possibility of missing a defective apple by a classification system may be reduced by using the multiple neural network classifier.

Classification results of an apple group including 5 different classes obtained using the multiple neural network classifier are given in Table A.7 in Appendix A (also in Table 4.8). Table A.7 shows the classification results obtained from different grouping of wavelengths used in the multiple neural network classifier. In some applications, fewer effective wavelengths were used in an attempt to reduce the number of input images to the classifier for more efficient classification. For instance, using three wavelengths (680, 700, and 800 nm) for Empire variety resulted in the same or better classification results compared to using four effective wavelengths. On the other hand for Golden Delicious apples, using wavelengths of 660 nm (calyx/stem classifier), 1200 nm (russet), 1260 nm (bruise), and 1320 nm (bitter pit) resulted in a better classification compared to using input images only at two effective wavelengths (660 and 1200 nm). Confusion matrices of the classifications using the multiple neural networks (Table B.2 for Empire and B.11 for Golden Delicious) and other classifiers are given in Appendix B.

Classification success (99%) obtained by using a multiple neural network classifier for Empire variety was close to the classification success (100%) obtained using the single neural network classifier when no discrimination was considered among the defect groups. Classification successes obtained from single and multiple neural networks were 91.8% and 84.5% respectively when the confusion between defect groups is considered.

However, for Golden Delicious better results were obtained with the multiple neural network classifier (92.8%) compared to using the single neural network classifier (90.7%) in classifying the apples as bad and good in general without considering the defect types separately. On the other hand, classification success (84.5%) of the single neural network was slightly higher than that (82.5%) of the multiple neural networks when no misclassification among the defects was allowed. The classification success in the second case was calculated by considering the misclassifications between the defective groups but not within the good groups, i.e. misclassification between stem-calyx and good apples was accepted. The disadvantage of the multiple neural network classifier is that it requires longer testing time as it evaluates multiple inputs and outputs.

4.3.3 Classification Using Texture Features

Similar to the classification application with two outputs, only three texture features such as angular second moment, contrast and correlation were used in the classifiers. Images at selected effective wavelengths determined in the previous sections were used in each five-class classification for both varieties.

The highest classification success in classifying Empire apples using texture features only was obtained using single neural network classifier (93.8%) when defects were considered as one group and 88.7% classification success was obtained when each defect was considered separately (Table 4.8). Following the neural network classifier the best results were obtained using 3-nearest neighbor, 1-nearest neighbor, decision tree and Bayesian classifiers respectively.

In classifying Golden Delicious apples using texture features only, single backpropagation neural network, Bayesian and 1-nearest neighbor classifiers produced the highest classification success of 89.7% when considering all the defects as one group. If the defects are considered separately, classification success for Bayesian classifier becomes 83.5%, for single neural network classifier, 82.5% and for 1-nearest neighbor classifier, 79.4%.

For either variety, using texture features alone did not provide a better classification compared to using pixel gray values in the back propagation neural network classifier. Also, using texture features is associated with increased time requirements for both training and testing. Extraction of texture features is a computationally intensive task that takes a certain amount of time for each image. On the other hand, the testing procedure using pixel intensity features in the application of a neural network is quite fast compared to testing using texture features. However, time spent in training using either texture features or pixel intensity features are similar.

4.3.4 Classification Using Combined Features

Only the single neural network classifier was used in the classification applications that used combined features. In this stage, three feature combination methods were studied to further improve the classification accuracy.

4.3.4.1 Using Pixel Intensities From Two Images at Different Wavelengths

Combining pixel gray values from two images at different wavelengths was one method. The application of combining images worked well for Golden Delicious variety

increasing the classification accuracy and decreasing the error rate significantly in classifying five different classes at the same time. Not every two pair of wavelengths worked the same way. To find the most effective pair of wavelengths, images at single most effective wavelengths that resulted in the highest classification successes were combined and tested to find the optimum result.

For Golden Delicious, images at wavelengths 1260 and 620 nm, which were the best two wavelengths respectively in classification of five classes using a single image (single wavelength), were used together. This combination gave the highest classification success of 94.9% among all of the classification applications when no discrimination between the defects was considered. Likewise in the case when defects were evaluated separately the highest classification success of 89.7% was obtained using the combined images explained above (Table 4.8). Apparently combining two images helped to improve the classification accuracy, as the complementary spectral information from two images at effective wavelengths from separate regions of the spectra was brought together (effective wavelength range for individual defects on Golden Delicious apples is 660-1320 nm, Table 4.4).

However, combining images did not result in the same improvement for the Empire variety as can be seen in Table 4.8 (effective wavelength range for individual defects on Empire apples is 600-880 nm). This was probably because the effective wavelengths for this variety were closer to each other compared to effective wavelengths for Golden Delicious apples. So, adding a new image from a second wavelength, which was close to the first one, did not bring any new information to the neural network.

4.3.4.2 Using Pixel Intensities and Texture Features

In the second method of combining features, pixel gray values and texture features from images were used together. For both varieties, positive improvements in the classification accuracies were obtained compared to using only pixel information from one image when defects were considered as one class.

Using pixel gray values at 740 nm and texture features at 620 nm for Empire apples provided the best classification accuracy (93.8%) of all classification methods in the case when each defect was evaluated separately (Table 4.8). Also, combining pixel intensity values and texture features for Empire variety brought the classification accuracies for the majority of defect groups close to or above 90% except for the puncture (Table A.5 in Appendix A). Using the texture features from wavelength 620 nm especially improved the results of punctured and leaf roller defects, which might not be visible enough at wavelength 740 nm.

Slight improvement was obtained by combining pixel gray values and texture features for Golden Delicious apples when no misclassification was allowed between the defect groups. Classification success for this case was improved from 84.5% to 85.6% by adding texture features to the pixel intensity values. Addition of texture features to pixel gray value features increased the identification success of the individual surface characteristics for both varieties.

4.3.4.3 Combining Two Images by Averaging

In the third feature combination method, two images of an apple at two different wavelengths were combined by averaging. However, effects of this method were not as

well as those of the two combinations explained above. In the case of classifying 2 classes, averaging 2 images resulted in almost the same classification success and error rate with other applications such as using only the pixel gray values or using two images together. However in the case of classifying 5 classes together, averaging images decreased the classification success and increased the error rate compared to the case using pixel gray values only. This was due probably to the loss of information on defects resulted from the averaging.

4.3.5 Statistical Tests of the Results from 5-Class Classification

There seemed to be more variations among the classification results from 5-class classification compared to 2-class classification. Thus, statistical analysis was performed among the 5-class classification results to see if there was any significant difference between the classifiers. In statistical tests of the classification methods, Logistic Regression in SAS software was used assuming the classification outputs from the classifiers had a binomial distribution.

Among the classification results for Empire, when the defects are considered as one class, there were significant differences between statistical and neural network classifiers at the 95% confidence level. There was no significant difference between the statistical classifiers except between 1-NN (86.6%) and Bayesian (76.3%), and between 3-NN (87.6%) and Bayesian (76.3%) classifiers (Table 4.9).

Similarly, at the 95% confidence level in the case when each defect is considered as one class, there were significant differences between the statistical classifiers and

Table 4.9 Significant test results when all the defect groups are considered as one class (Empire, $P < 0.05$)*

	1-NN	3-NN	Bayesian	D.Tree	MBNN-P	SBNN-2P	SBNN-PT	SBNN-T
1-NN ¹		0.739	0.017	0.593	0.011	0.011	0.011	0.036
3-NN ²			0.007	0.467	0.014	0.014	0.014	0.059
Bayesian				0.087	0.001	0.001	0.001	0.001
D.Tree ³					0.001	0.007	0.007	0.022
MBNN-P ⁴						1	1	0.094
SBNN-2P ⁵							1	0.094
SBNN-PT ⁶								0.094
SBNN-T ⁷								

(¹) 1-nearest neighbor, (²) 3-nearest neighbor, (³) D.Tree: decision tree, (⁴) MBNN-P: multiple backpropagation neural network using pixel gray values, (⁵) SBNN-2P: single backpropagation neural network using pixel gray values from two images, (⁶) SBNN-PT: single backpropagation neural network using pixel gray values and texture features, (⁷) SBNN-T: single backpropagation neural network using texture features only, * SBNN-P single backpropagation neural network using pixel gray values, which classified all the testing apples correctly, was excluded from the analysis as a drawback result of the software used.

especially the single neural network classifier using pixel gray values from one image, using combined features of pixel gray values and texture features, and using texture features only (Table 10). Among the statistical classifiers, there were significant differences between 1-NN (79.4%) and Bayesian (63.9%) classifiers, and between 3-NN (81.4%) and Bayesian (63.9%) classifiers. Also, among neural network classifiers, there were significant differences between the multiple neural networks classifier and single neural network classifier using combined features of pixel gray values and texture (Table 10).

Among the classification results for Golden Delicious, when the defects are considered as one class, only significant differences were between 3-nearest neighbor classifier (85.6%) and single neural network classifier (94.9%) using pixel gray values from two images as features and between decision tree classifier (86.6%) and single

Table 4.10 Significant test results when each defect group is considered separately (Empire, $P < 0.05$)

	1-NN	3-NN	Bayesian	D.Tree	MBNN-P	SBNN-2P	SBNN-P	SBNN-PT	SBNN-T
1-NN ¹		0.527	0.001	0.414	0.336	0.073	0.016	0.008	0.028
3-NN ²			<0.001	0.200	0.590	0.163	0.035	0.018	0.051
Bayesian				0.054	0.002	<0.001	<0.001	<0.001	<0.001
D.Tree ³					0.106	0.009	0.001	<0.001	0.007
MBNN-P ⁴						0.285	0.111	0.022	0.415
SBNN-2P ⁵							0.256	0.059	1
SBNN-P ⁶								0.415	0.440
SBNN-PT ⁷									0.201
SBNN-T ⁸									

(¹) 1-nearest neighbor, (²) 3-nearest neighbor, (³) D.Tree: decision tree, (⁴) MBNN-P: multiple backpropagation neural network using pixel gray values, (⁵) SBNN-2P: single backpropagation neural network using pixel gray values from two images, (⁶) SBNN-P: single backpropagation neural network using pixel gray values from one images, (⁷) SBNN-PT: single backpropagation neural network using pixel gray values and texture features, (⁸) SBNN-T: single backpropagation neural network using texture features only

Table 4.11 Significant test results when all the defect groups are considered as one class (Golden Delicious, $P < 0.05$)

	1-NN	3-NN	Bayesian	D.Tree	MBNN-P	SBNN-2P	SBNN-P	SBNN-PT	SBNN-T
1-NN ¹		0.099	1	0.317	0.440	0.171	0.819	1	1
3-NN ²			0.285	0.705	0.073	0.007	0.277	0.347	0.156
Bayesian				0.317	0.440	0.136	0.808	1	1
D.Tree ³					0.137	0.023	0.373	0.467	0.255
MBNN-P ⁴						0.481	0.564	0.407	0.440
SBNN-2P ⁵							0.253	0.099	0.171
SBNN-P ⁶								0.705	0.808
SBNN-PT ⁷									1
SBNN-T ⁸									

(¹) 1-nearest neighbor, (²) 3-nearest neighbor, (³) D.Tree: decision tree, (⁴) MBNN-P: multiple backpropagation neural network using pixel gray values, (⁵) SBNN-2P: single backpropagation neural network using pixel gray values from two images, (⁶) SBNN-P: single backpropagation neural network using pixel gray values from one images, (⁷) SBNN-PT: single backpropagation neural network using pixel gray values and texture features, (⁸) SBNN-T: single backpropagation neural network using texture features only

neural network classifier using pixel gray values from two images as features at 95% confidence level (Table 4.11).

In the case when each defect is considered as one class for Golden Delicious at 95% confidence level, the only significant difference was between 3-nearest neighbor classifier (76.3%) and single neural network classifier (89.7%) using pixel gray values from two images as features (Table 4.12).

Table 4.12 Significant test results when each defect group is considered separately (Golden Delicious, $P < 0.05$)

	1-NN	3-NN	Bayesian	D.Tree	MBNN-P	SBNN-2P	SBNN-P	SBNN-PT	SBNN-T
1-NN ¹		0.404	0.371	0.637	0.564	0.052	0.370	0.240	0.491
3-NN ²			0.125	0.224	0.257	0.007	0.158	0.071	0.156
Bayesian				0.593	0.827	0.203	0.835	0.670	0.819
D.Tree ³					0.842	0.089	0.590	0.433	0.796
MBNN-P ⁴						0.070	0.655	0.491	1
SBNN-2P ⁵							0.226	0.205	0.146
SBNN-P ⁶								0.763	0.695
SBNN-PT ⁷									0.532
SBNN-T ⁸									

(¹) 1-nearest neighbor, (²) 3-nearest neighbor, (³) D.Tree: decision tree, (⁴) MBNN-P: multiple backpropagation neural network using pixel gray values, (⁵) SBNN-2P: single backpropagation neural network using pixel gray values from two images, (⁶) SBNN-P: single backpropagation neural network using pixel gray values from one images, (⁷) SBNN-PT: single backpropagation neural network using pixel gray values and texture features, (⁸) SBNN-T: single backpropagation neural network using texture features only

In general, there were significant differences between almost all of the statistical and neural network classifiers when considering defects as one class or as separate classes for Empire variety. However, for Golden Delicious apples, the only significant differences were between statistical classifiers of 3-nearest neighbor and decision tree, compared against the single neural network classifier using pixel gray value features from two images.

4.4 Classification of Stem and Calyx

As it can be seen from Tables A.5 and A.6 in Appendix A, stem and calyx were classified 85-100% correctly in five-class classification studies depending on the wavelength used. Misclassification occurred usually between good class and stem or calyx, which was an acceptable error as can be seen in most of the confusion matrices in Appendix B.

In this study, no extra algorithm or procedure was applied to discriminate the stem and calyx from normal or defective parts. The backpropagation neural network was able to recognize these natural parts of apples to the extent of not classifying them as defects without performing any further processing.

4.5 Summary and Conclusions

The following conclusions can be drawn from the results obtained in this study:

Image Processing:

- Reducing original image resolution from 480x640 pixels into 60x80 pixels did not change the classification accuracy. Thus, the resolution of 60x80 pixels was used in all the classification applications throughout the research. Using reduced resolution provided significant decrease in time to train and test the artificial classifiers.

- In all of the classification applications, background in the apple images was subtracted to make sure that the classifiers sense the reflectance from defective areas more efficiently. However, in a pre-study to find the effect of existence of background on classification success, it was found that existence of background did not affect the

classification success of the classifiers. Moreover, better classification performance was obtained in general with the existence of background.

- In a pre-study to test the effect of small specular regions in the images on classification success, it was found that specular regions highly biased the classifiers as they provided biased features to classifiers with their specific shapes dependent on the apple surface characteristics such as defects, calyx, stem, and healthy tissue. Therefore, specular regions were removed from images by automatically deleting the pixels in the specular region with a fixed sized window. Images with deleted specular regions were used in all of the classification applications.

- It was concluded that in the future applications, all three image processing applications used in the this study could be eliminated by acquiring images directly at reduced resolution, eliminating specular regions from images with adjustments in image acquisition system and using images with backgrounds.

Classification with two-classes:

- Using the single backpropagation neural network classifier and pixel gray value features, the most effective wavelengths for optimum identifications of apple surface characteristics such as bruise, leaf roller, puncture, bitter pit, russet, calyx and stem separately against healthy (good) apple tissue (two-class classification) were determined.

- For Empire variety, the single backpropagation neural network classifier using pixel gray values as features provided the most optimum classification results for the following surface characteristics against good apple tissue; stem (100%), calyx (88.9%),

leaf roller (100%), and puncture (96.7%). Bruise defect on this variety was classified most effectively by the same classifier using the texture features (100%).

- For Golden Delicious, the single backpropagation neural network classifier using pixel gray values resulted in following optimum classification success results for surface characteristics of; bruise (98%), stem (100%), russet (89.2%) and bitter pit (100%). Calyx was best classified by 3-nearest neighbor classifier using texture features (100%).

- The single backpropagation neural network classifier using pixel gray values provided better classification performance compared to the statistical classifiers that used texture features.

- Combining features such as pixel gray values from two images at different wavelengths, pixel gray values from one image and texture features from another, and averaging two images at two different wavelengths into one image did not provide improved classification compared to using one type of feature. This was probably due to the effective wavelengths that were close to each other.

- It was observed that the backpropagation neural network gives more value to high gray values in the images assuming that high gray values are important features. This conclusion was based on the classification of defects such as bitter pit and leaf roller and also classification of images with specular regions (no deletion).

Classification with Five-Classes:

- Similar to the two-class classification category, effective wavelengths for an apple group containing three defect types, a combined stem and calyx group and a good

(healthy) apple group were found using single backpropagation neural network and intensity values from image pixels.

- Wavelengths 740 and 1260 nm were found as the most effective for Empire and Golden Delicious apple groups respectively.

- For Empire variety, the most successful classifications were obtained by the classifiers of single backpropagation neural network using pixel gray values from one image (100%), using combined pixel gray values from two images (99%), using combined pixel and texture features from two images (99%), and multiple backpropagation neural networks (99%) using pixel gray values from one image as features when confusions between two defect groups and two good apple groups (healthy tissue, stem and calyx) were ignored. However, when no confusion between the defect groups is accepted, the highest classification success (93.8%) was obtained by single backpropagation neural network using combined features of pixel gray values and texture features. In this case, when no confusion was allowed between the two defect groups, the second best classification result (91.8%) was obtained by using single backpropagation neural network classifier with pixel gray value features.

- For Golden Delicious, in the case when defect groups are considered as one class, the most successful classification was obtained by using the single backpropagation neural network classifier with combined pixel gray values from two images (94.9%). The second best classification in this category was obtained from multiple backpropagation neural network classifier using pixel gray values from one image (92.8%). In the case when no confusion is allowed between two defect groups, the best classification (89.7%) was again obtained by using the single back propagation neural network using combined

pixel gray values from two images at the two most effective wavelengths. Second best result (85.6%) in this category was obtained from the single backpropagation neural network classifier using the combined features of pixel gray values and texture features.

- Combining pixel gray value features from two images at most effective wavelengths (Golden D.) and combining pixel gray values from one image and texture features from another image at effective wavelengths increased the classification success for both varieties of apples using the single backpropagation neural network classifier. Successful classification results with combined features are believed to be due to the availability of effective wavelengths in different regions of the spectra, contrary to the two-class classification category where the effective wavelengths were usually from the same region of the spectra. Also, having different surface characteristics with different effective wavelengths together in this classification category may be another reason for improved classification results with combined features from different wavelengths.

- A different application of a backpropagation neural network was introduced by the multiple backpropagation neural network classifier that provided improved classification results for Golden Delicious apples when all the defects are assumed to be one class.

- In classification of Empire apples, neural network classifiers performed significantly higher classification results compared to statistical classifiers. Although the neural network classifiers performed better for classification of Golden Delicious apples compared to statistical classifiers, the only significant differences were between statistical classifiers of 3-nearest neighbor and decision tree, compared against the single neural network classifier using pixel gray value features from two images.



Overall Conclusions:

- Backpropagation neural network classifiers were superior to statistical classifiers in most of the classification applications studied. In general neural network classifiers performed better using intensity features compared to using texture features.

- Superiority of neural network classifiers especially in five-class classification can be explained with the ability of this type of classifier to learn the non-linear relation between the input features and output classes.

- Using wavelengths beyond 1000 nm improved the identification success of the defects such as bruise, russet, and bitter pit on Golden Delicious apples, and leaf roller on Empire apples. This broad range of wavelengths also provided improved classification success especially for Golden Delicious in five-class classification allowing the combination of images at the wavelengths from different regions of the spectra.

- In conclusion, the backpropagation neural networks provided robust and effective classifications using pixel gray values in both categories of two-class and five-class classifications. Using the combination of pixel gray values from images or the combination of pixel gray values and texture features, the backpropagation neural network provided further improvement in five-class classification.

APPENDICES

APPENDIX A

Graphs of Classification Success and Accuracy Rate versus Spectral Reflectance For Defects on Empire and Golden Delicious Apples

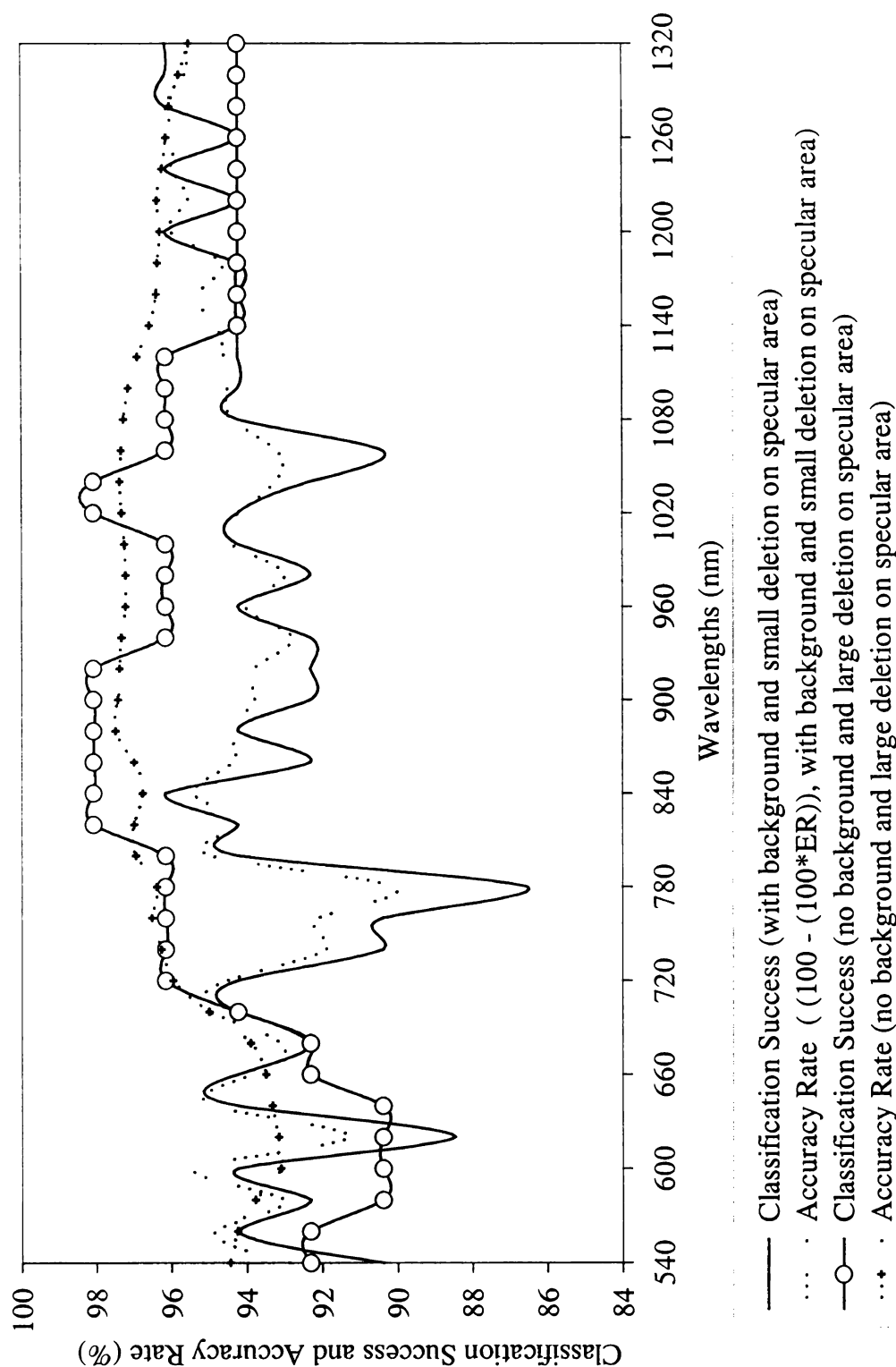


Figure A.1 Classification of bruised apples against good apples (Empire)

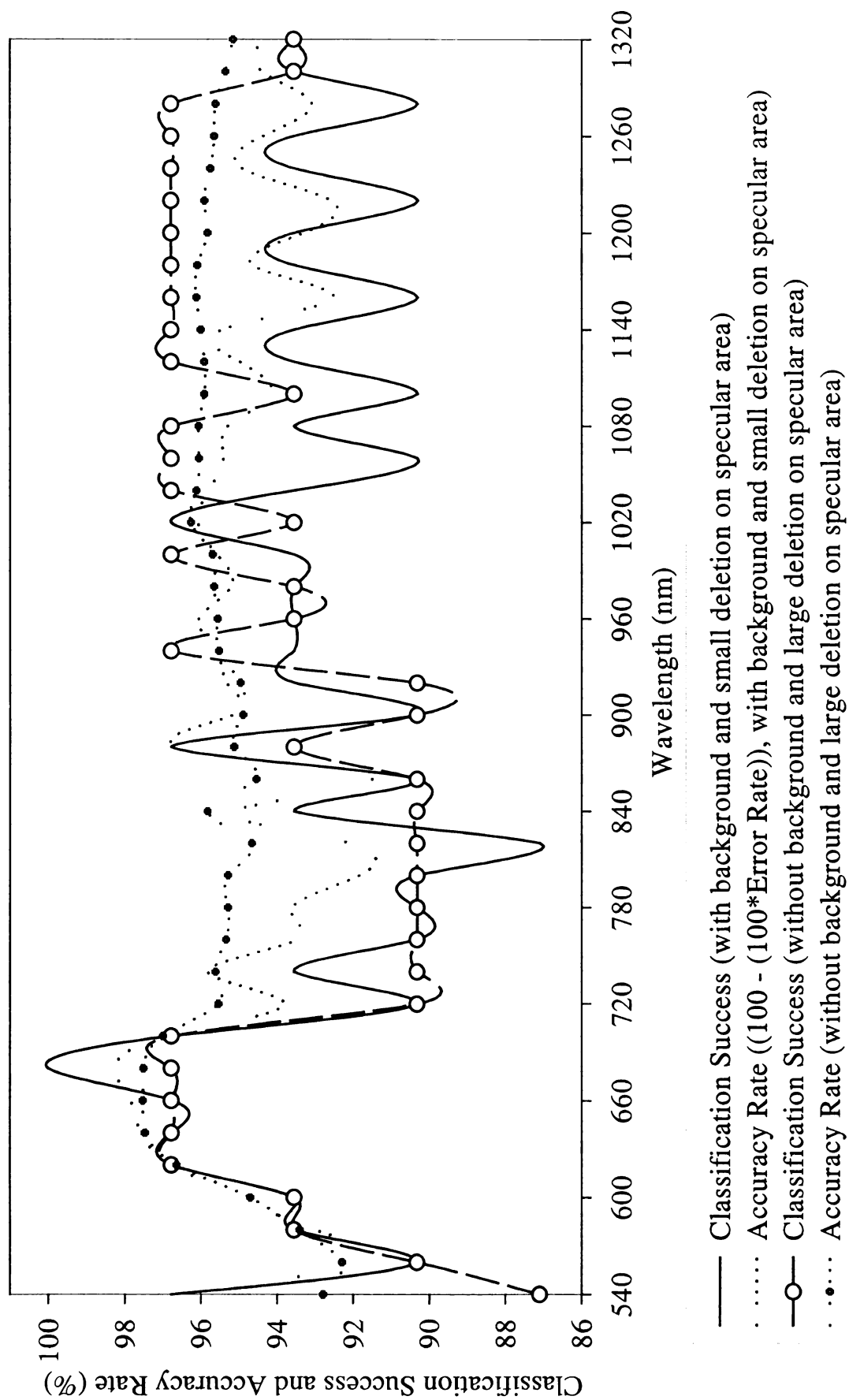


Figure A.2 Classification of punctured apples against good apples (Empire)

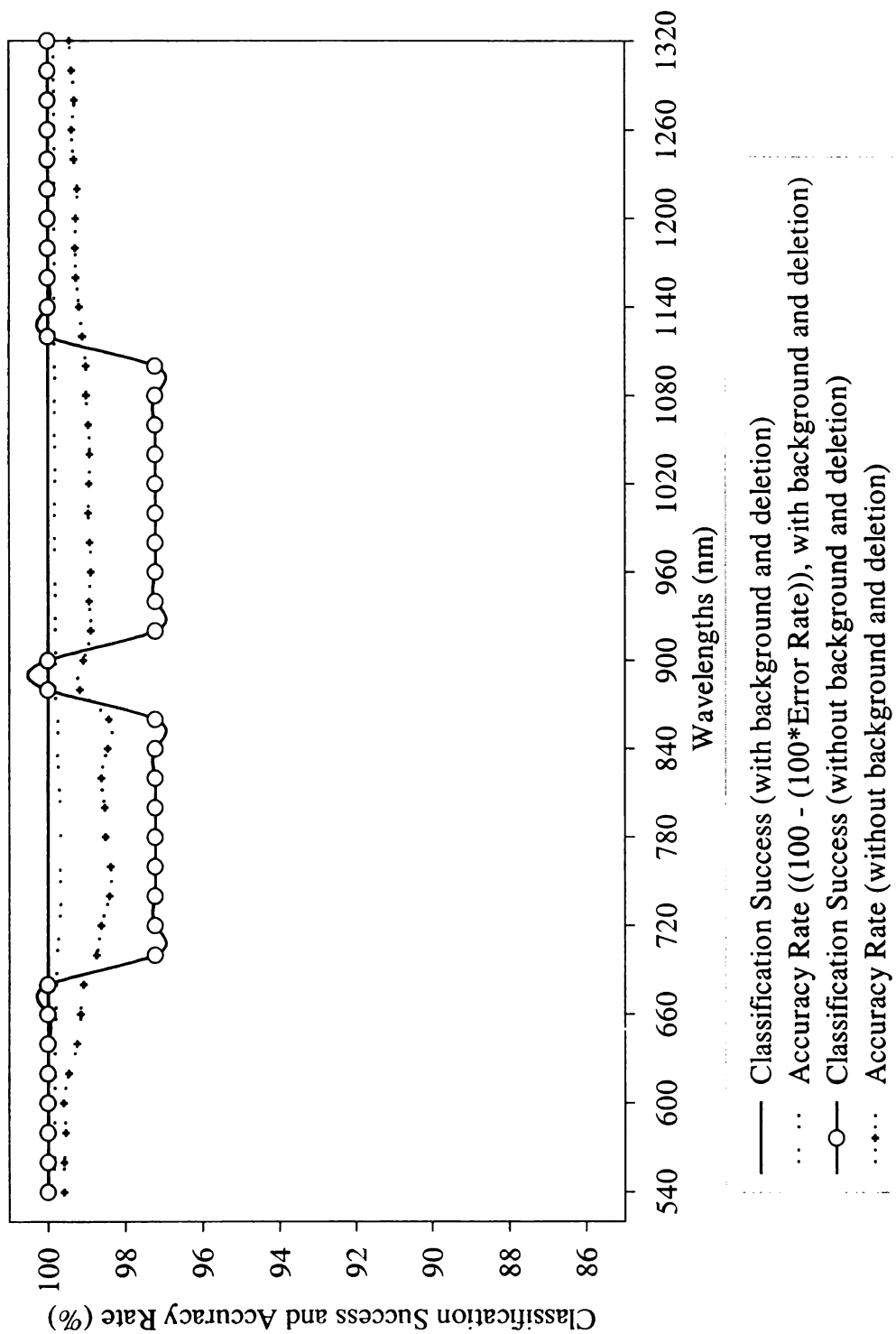


Figure A.3 Classification of apples with leaf roller against good apples (Empire)

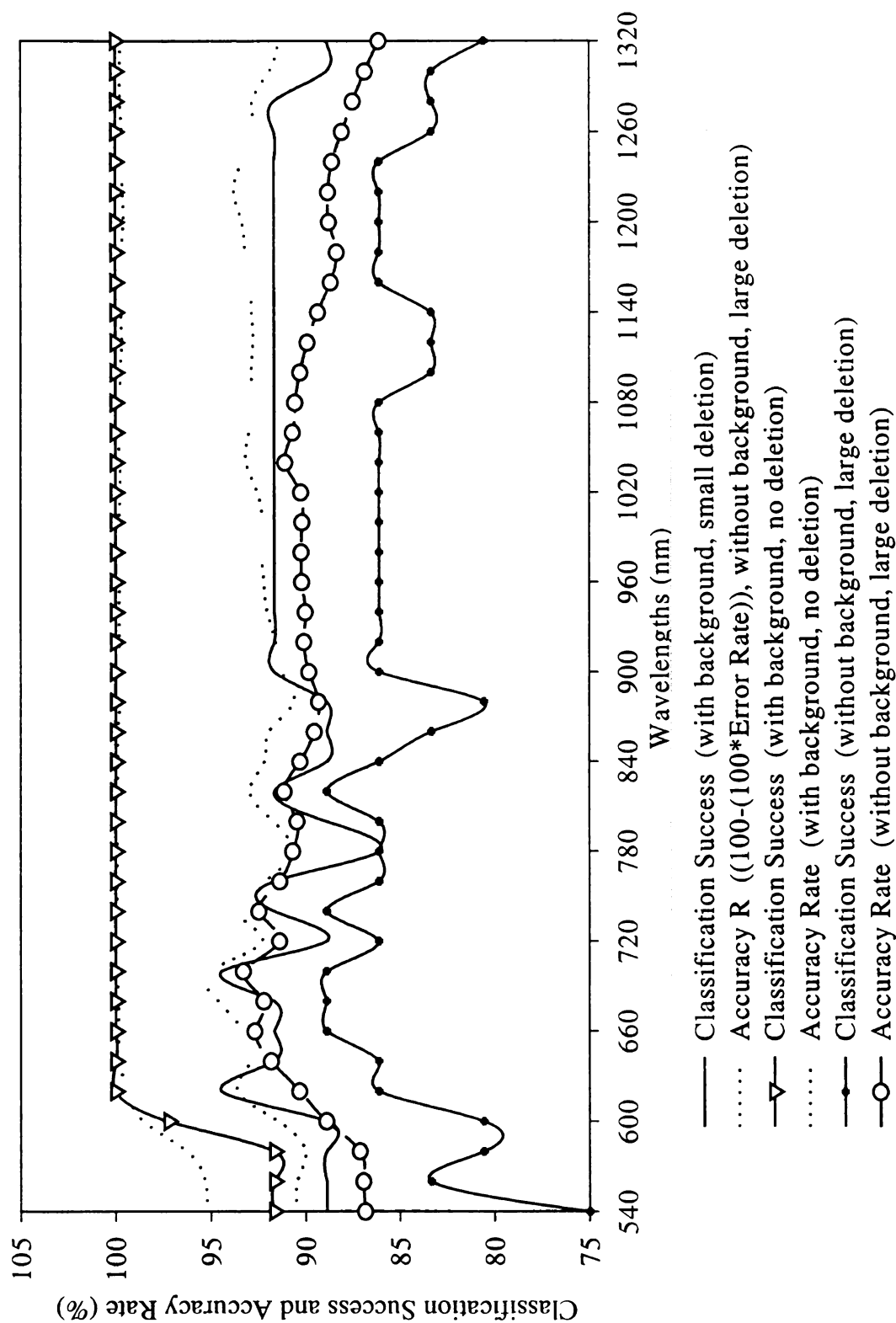


Figure A.4 Classification of calyx view images against non-calyx (good) apple images (Empire)

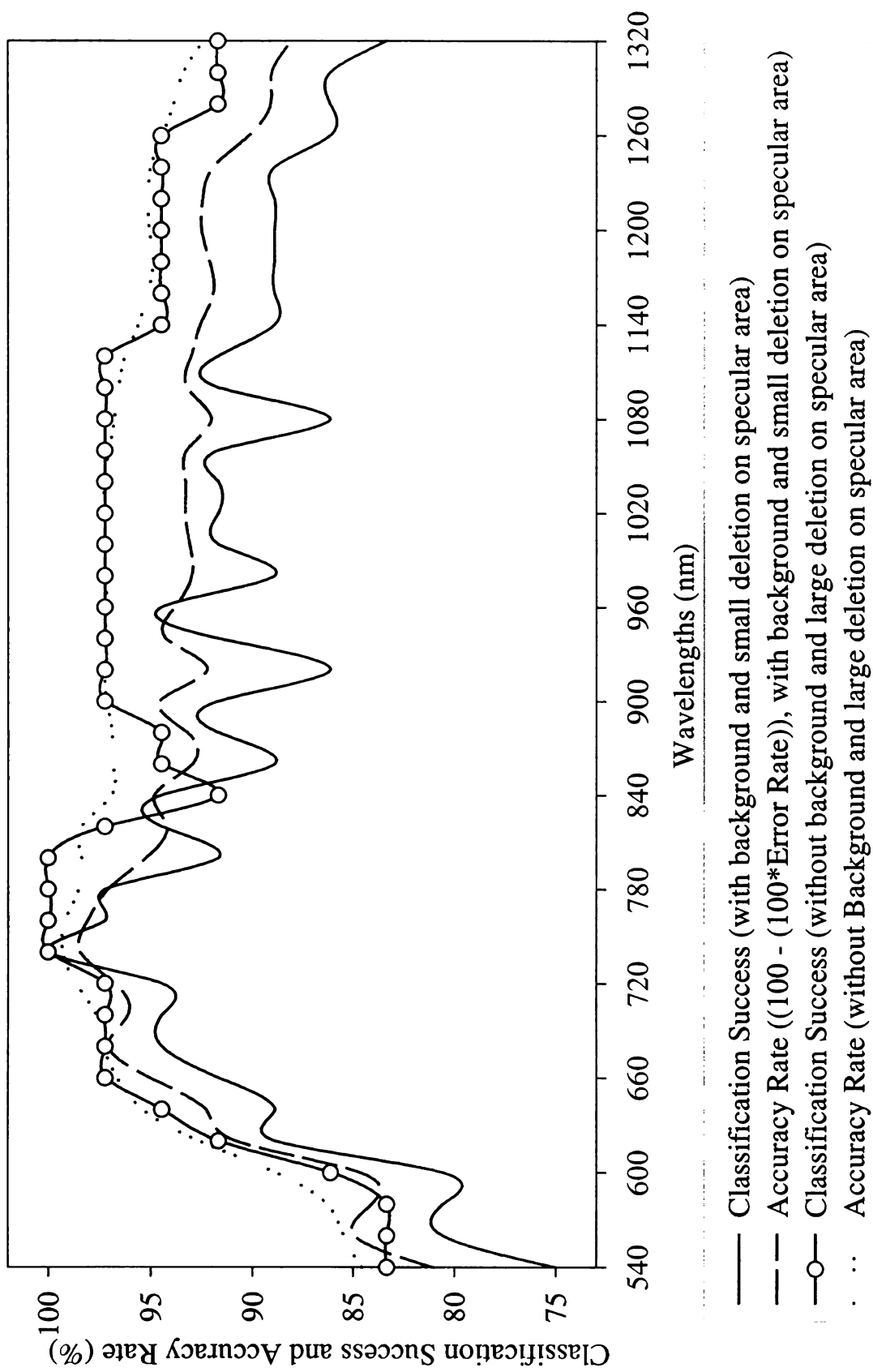


Figure A.5 Classification of apples with stem against good apples (Empire)

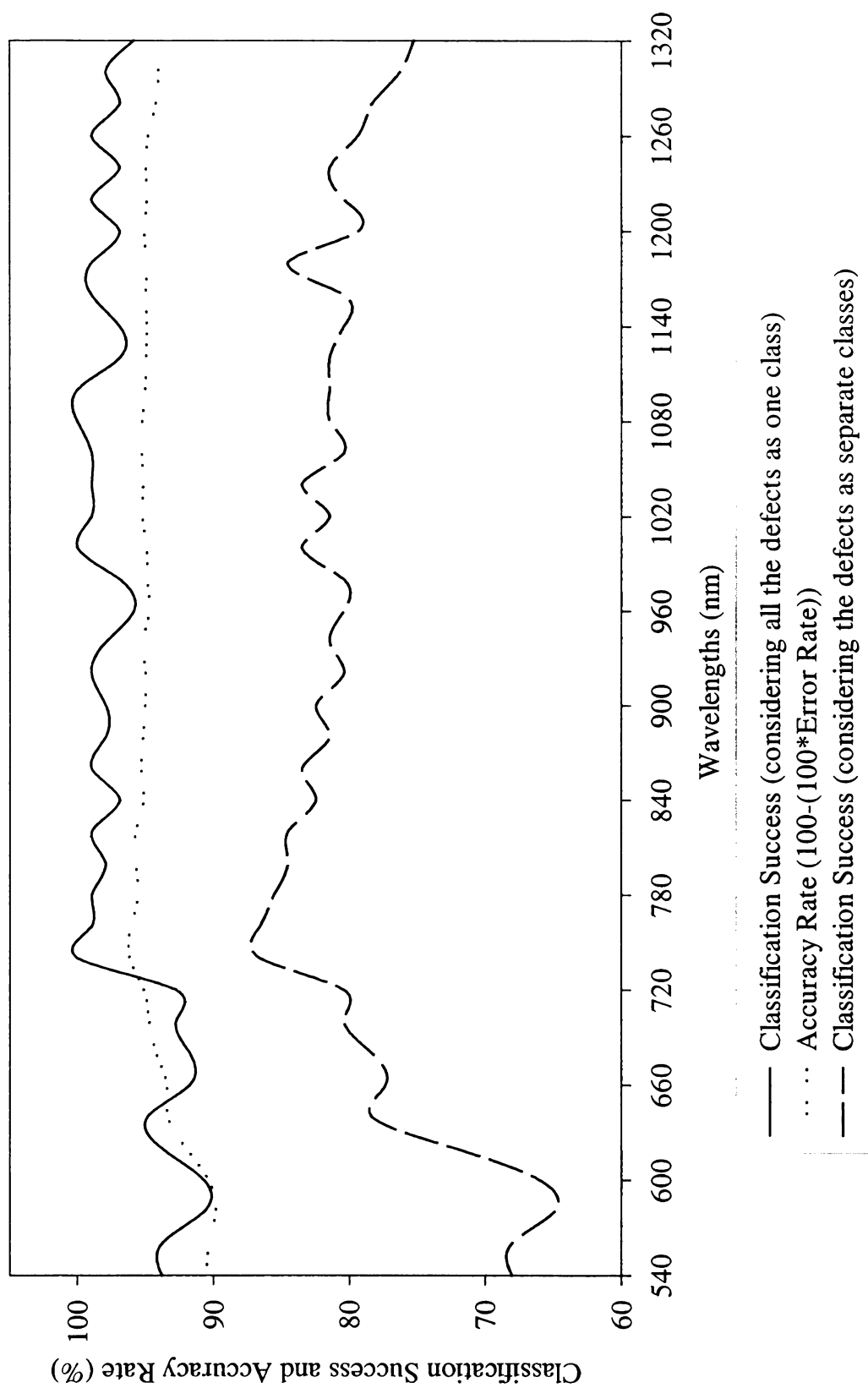


Figure A.6 Classification of apples considering all the defects and natural features against good apples (5 Classes, Empire)

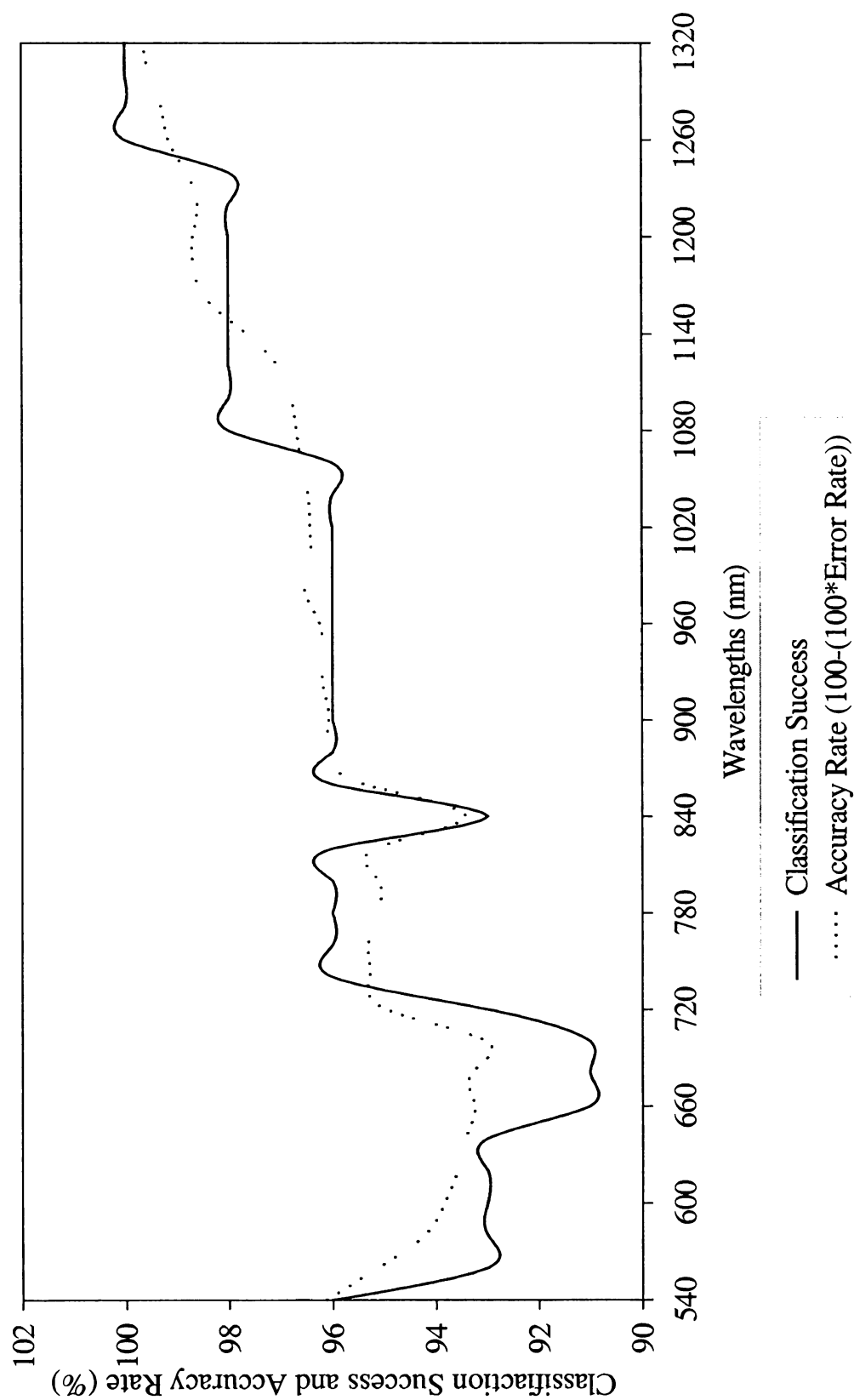


Figure A.7 Classification of apples with bitter pit against good apples (Golden Delicious)

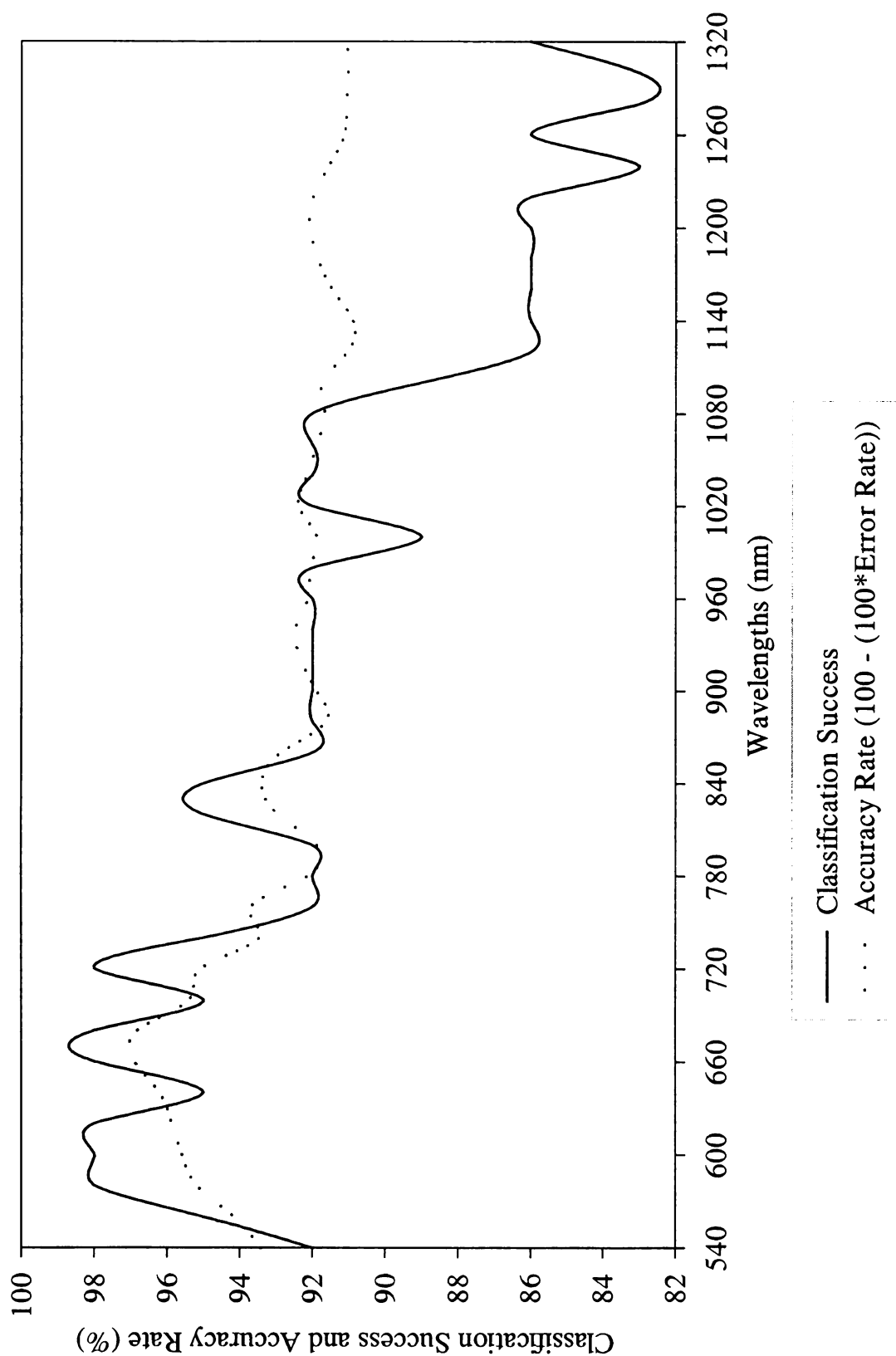


Figure A.8 Classification of calyx view images against non-calyx (good apple) images (Golden Delicious)

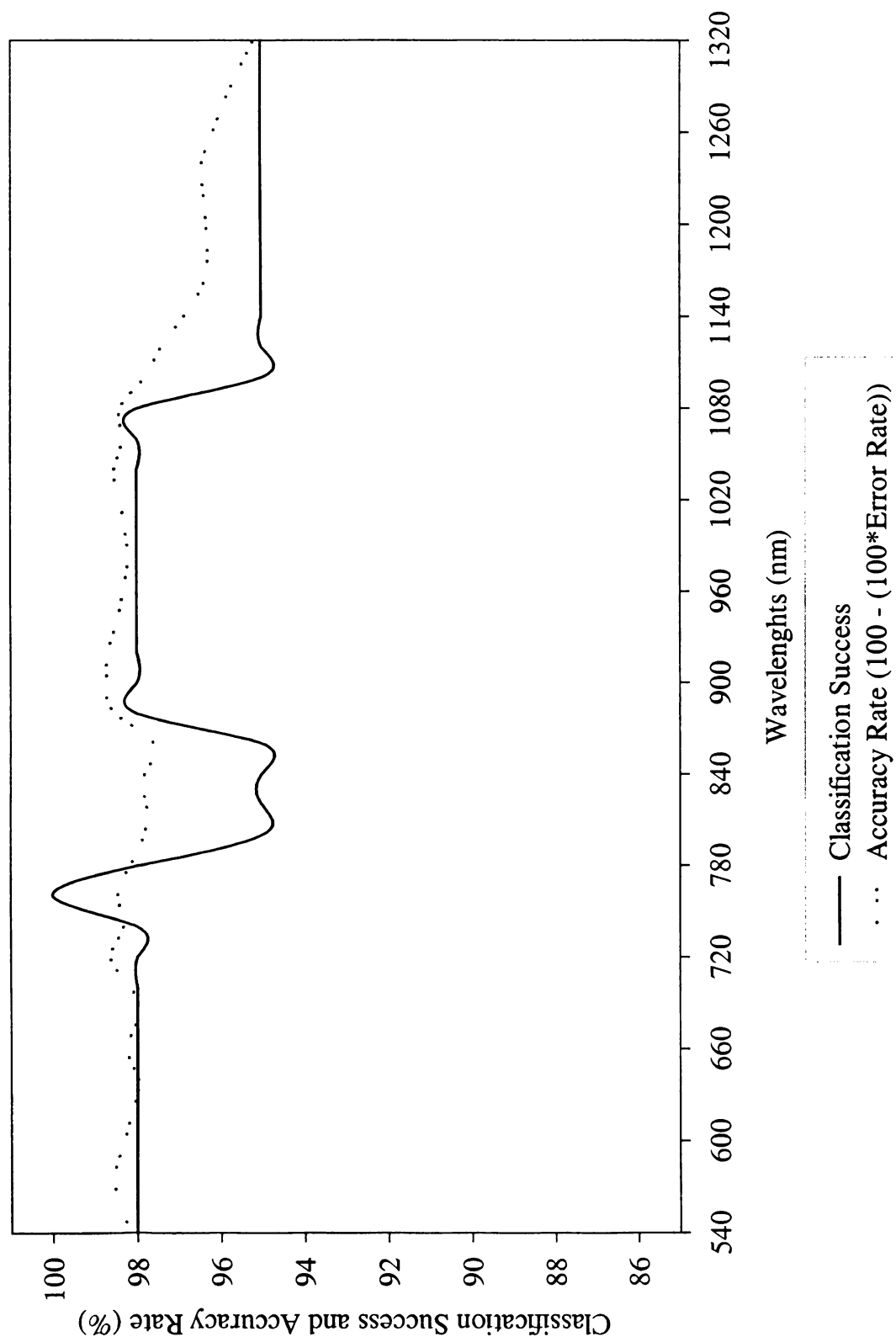


Figure A.9 Classification of stem view images against non-stem (good apple) images (Golden Delicious)

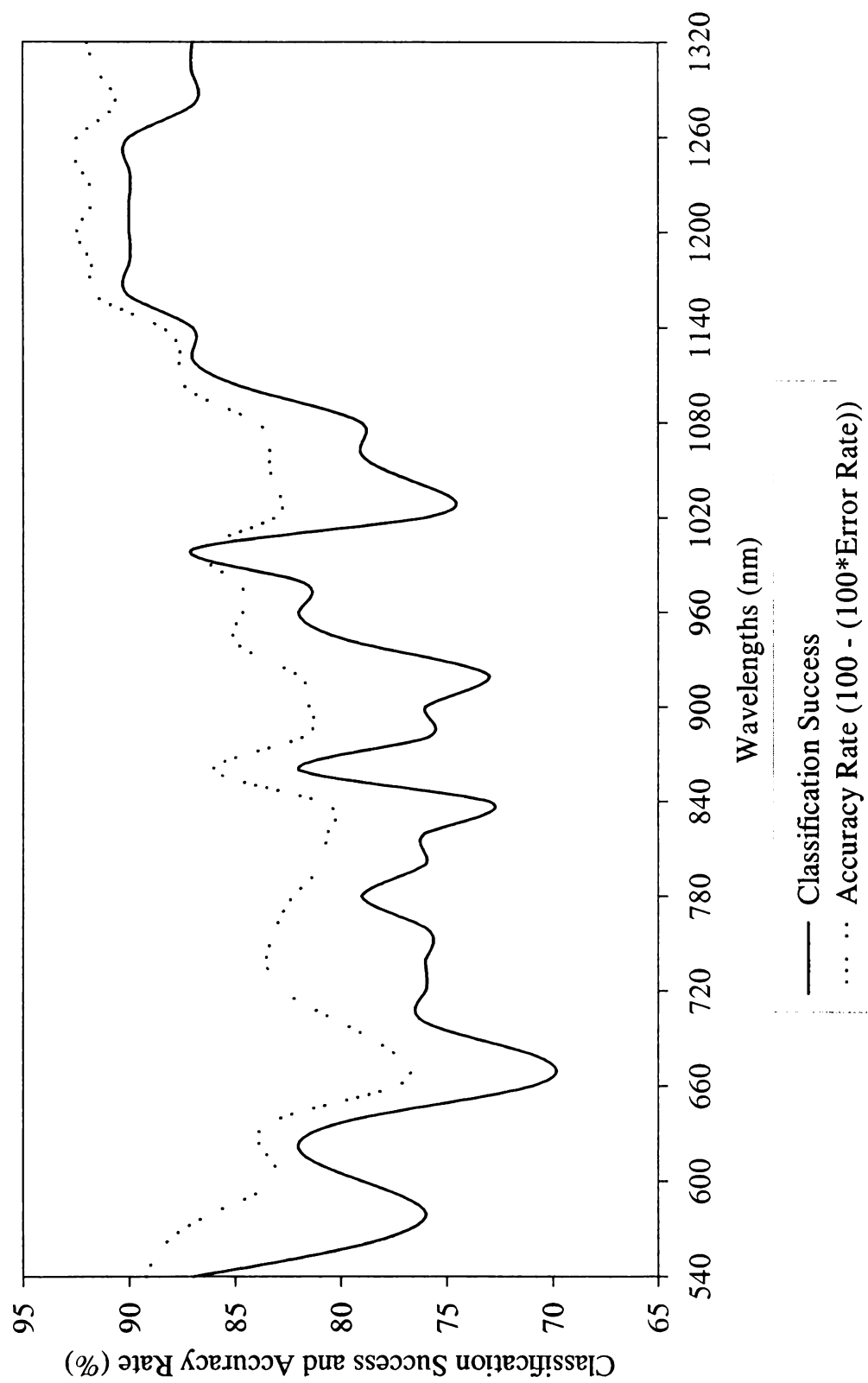


Figure A.10 Classification of apples with russet against good apples (Golden Delicious)

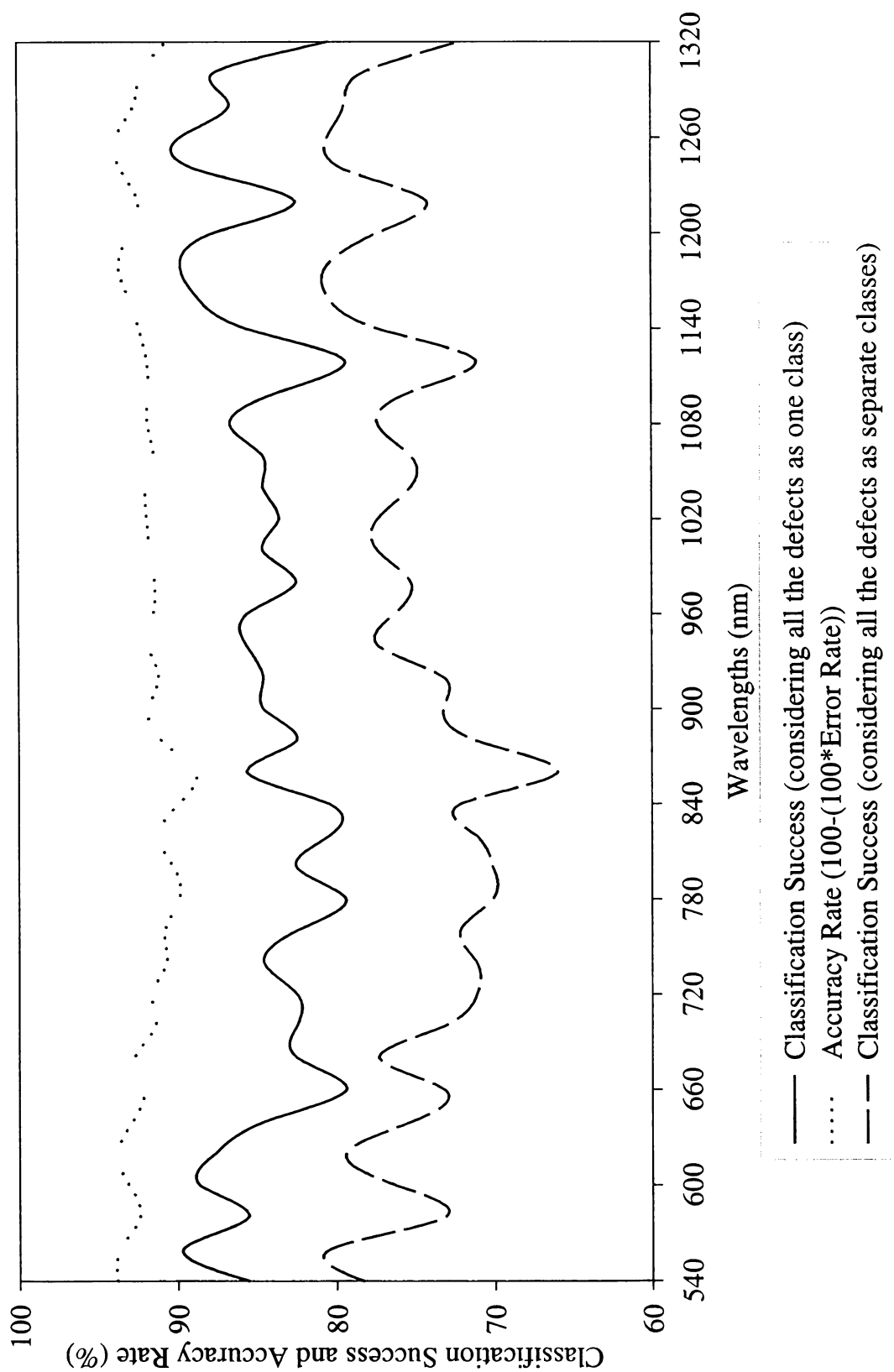


Figure A.11 Classification of apples considering all the defects and natural features against good apples (5 Classes, Golden Delicious)

Table A.1 Classification results of defective apples at single and double effective wavelengths (Two-Classes)

Surface Characteristics	Single Wavelength (nm)	Classification Success (%)	Accuracy Rate (%)‡	Two Wavelengths (nm)	Classification Success (%)	Accuracy Rate (%)‡
Empire						
Bruise	880	98.1	97.5	880+1040	98.1	97.7
Stem	740	100.0	98.3	740+760	91.7	95.5
Calyx	700	88.9	93.3	700+740	88.9	92.1
Leaf Roller	600	100.0	99.6	600+1320	100.0	99.6
Puncture	680	96.8	97.5	680+700	90.3	90.9
Golden D.						
Bruise	1260	98.0	96.9	1260+1280	96.0	96.3
Stem	760	100.0	98.5	760+900	100.0	98.6
Calyx	660	97.1	96.9	660+680	97.1	96.3
Russet	1200	89.2	92.5	1200+1320	89.2	90.8
Bitter Pit	1320	100.0	99.7	1300+1320	100.0	99.0

‡ Accuracy rate was calculated using the operation of $(100-(100*\text{Error Rate}))$

Table A.2 Classification results obtained using combined features with the single backpropagation neural network (Two-Classes)

Surface Ch*	Pixel Gray Values ^b				Pixel Gray Values + Texture ^c				Averaging 2 Images ^d			
	WL ^a (nm)	CS (%)‡	AR (%)†	WL (nm) ^e	CS (%)‡	AR (%)†	WL (nm) ^e	CS (%)‡	AR (%)†	WL (nm) ^e	CS (%)‡	AR (%)†
Empire												
Bruise	880	98.1	97.5	920	96.2	97.4	920	98.1	97.4	920	98.1	97.4
Stem	740	100.0	99.3	760	100.0	97.6	760	94.4	95.8	760	94.4	95.8
Calyx	700	88.9	93.3	740	86.1	93.2	740	86.1	91.4	740	86.1	91.4
Leaf R.	600	100.0	99.6	1320	100.0	99.6	1320	100.0	99.6	1320	100.0	99.6
Puncture	680	96.8	97.5	700	87.1	90.9	700	87.1	90.7	700	87.1	90.7
Golden D.												
Bruise	1260	98.0	96.9	1260	98.0	97.2	1260	98.0	97.0	1260	98.0	97.0
Stem	760	100.0	98.5	760	100.0	98.5	760	100.0	98.9	760	100.0	98.9
Calyx	660	97.1	96.9	660	97.1	96.6	660	91.2	94.7	660	91.2	94.7
Russet	1200	89.2	92.5	1200	89.2	92.2	1200	89.2	92.5	1200	89.2	92.5
Bitter Pit	1320	100.0	99.7	1300	100.0	99.4	1300	100.0	99.3	1300	100.0	99.3

*Surface Characteristic, (^a) Selected effective wavelength for the first wavelength, (^b) Only pixel gray values as features, (^c) Pixel gray values were used with texture features together. (^d) Using gray values from the image that is the average of two selected images, (^e) Wavelength of second image used for texture or averaging,

† Accuracy Rate (100-(100*Error Rate)). ‡ Classification Success

Table A.3 Classification results from two-class classification using pixel intensity and texture features at effective wavelengths (Empire)

Surface Characteristics	Classifier	Classification Success, (%)	Accuracy Rate† (%)	Number of Misclassified Apples
Bruise, 880 nm	SBPNN-TeX ^b	100.0	99.9	0
	1-NN ^c	100.0	100.0	0
	SBPNN-Pixel ^a	98.1	97.5	1
	Bayesian ^d	98.1	98.1	1
	3-NN ^c	98.1	98.1	1
	Decision Tree ^d	98.1	98.1	1
Stem, 740 nm	SBPNN-Pixel ^a	100.0	99.3	0
	SBPNN-TeX ^b	91.7	92.6	3
	3-NN ^c	88.9	88.9	4
	1-NN ^d	86.1	86.1	5
	Bayesian ^d	75.0	75.0	9
	Decision Tree ^d	72.2	72.2	10
Calyx, 700 nm	SBPNN-TeX ^b	100.0	99.1	0
	SBPNN-Pixel ^a	88.9	93.3	4
	Decision Tree ^d	83.3	83.3	6
	Bayesian ^d	77.8	77.8	8
	1-NN ^c	75.0	75.0	9
	3-NN ^c	72.2	72.2	10
Leaf R., 600 nm	SBPNN-Pixel ^a	100.0	99.6	0
	Decision Tree ^d	100.0	100.0	0
	SBPNN-TeX ^b	97.2	97.8	1
	Bayesian ^d	97.2	97.2	1
	1-NN ^c	97.2	97.2	1
	3-NN ^c	97.2	97.2	1
Puncture, 680 nm	SBPNN-Pixel ^a	96.8	97.5	1
	SBPNN-TeX ^b	93.6	93.6	2
	Decision Tree ^d	93.6	93.6	2
	1-NN ^c	90.3	90.3	3
	3-NN ^c	90.3	90.3	3
	Bayesian ^d	83.9	83.9	5

† 100-(100*Error Rate), (^a) Single Backpropagation Neural Network Classifier using pixel features, (^b) Single Backpropagation Neural Network Classifier using texture features, (^c) Nearest Neighbor Classifier using texture features, (^d) Using texture features

Table A.4 Classification results from two-class classification using pixel intensity and texture features at effective wavelengths (Golden Delicious)

Surface Characteristics	Classifier	Classification Success (%)	Accuracy Rate (%)†	Number of Misclassified Apples
Bruise, 1260 nm	SBPNN-Pixel ^a	98.0	96.9	1
	Bayesian ^d	98.0	98.0	1
	Decision Tree ^d	94.0	94.0	3
	1-NN ^c	92.0	92.0	4
	3-NN ^c	92.0	92.0	4
	SBPNN-TeX ^b	90.0	90.7	5
Stem, 760 nm	SBPNN-Pixel ^a	100.0	98.5	0
	SBPNN-TeX ^b	100.0	99.8	0
	Bayesian ^d	97.4	97.4	1
	3-NN ^c	97.4	97.4	1
	1-NN ^c	94.9	94.9	2
	Decision Tree ^d	94.9	94.9	2
Calyx, 660 nm	3-NN ^c	100.0	100.0	0
	SBPNN-Pixel ^a	97.1	96.9	1
	SBPNN-TeX ^b	97.1	98.9	1
	Bayesian ^d	97.1	97.1	1
	1-NN ^c	91.2	91.2	3
	Decision Tree ^d	76.5	76.5	8
Russet, 1200 nm	Bayesian ^d	91.9	91.9	3
	SBPNN-Pixel ^a	89.2	92.5	4
	3-NN ^c	89.2	89.2	4
	Decision Tree ^d	89.2	89.2	4
	1-NN ^c	86.5	86.5	5
	SBPNN-TeX ^b	78.4	81.2	7
Bitter Pit, 1320 nm	SBPNN-Pixel ^a	100.0	99.7	0
	SBPNN-TeX ^b	94.7	95.1	2
	Bayesian ^d	94.7	94.7	2
	1-NN ^c	92.1	92.1	3
	3-NN ^c	92.1	92.1	3
	Decision Tree ^d	92.1	92.1	3

† 100-(100*Error Rate), (^a) Single Backpropagation Neural Network Classifier using pixel features, (^b) Single Backpropagation Neural Network Classifier using texture features, (^c) Nearest Neighbor Classifier using texture features, (^d) Using texture features

Table A.5 Classification results from five-class classification using single backpropagation neural network classifier (Empire)

WL ^{††} (nm)	CS (%) [‡]	AR (%) [†]	CS (%)	All ²	B-G ³	CS (%)	Good ⁴	Bruise ⁵	StCl ⁶	Leaf R. ⁷	CS (%)	Puncture ⁸
740	100.0	96.2	86.6	91.8	88.5	92.3	93.3	90.0	90.0	0.0	0.0	0.0
1000	100.0	94.9	83.5	90.7	88.5	92.3	86.7	70.0	70.0	20.0	20.0	20.0
1080	100.0	95.3	81.4	90.7	80.8	92.3	86.7	80.0	80.0	0.0	0.0	0.0
1100	100.0	95.1	81.4	90.7	80.8	92.3	86.7	80.0	80.0	0.0	0.0	0.0
760	99.0	96.1	86.6	90.7	88.5	96.2	93.3	80.0	80.0	0.0	0.0	0.0
740+760	99.0	95.6	85.6	88.7	88.5	96.2	96.7	50.0	50.0	20.0	20.0	20.0
740+1080t	99.0	96.7	89.7	90.7	96.2	88.5	96.7	90.0	90.0	20.0	20.0	20.0
740+1260t	99.0	96.7	90.7	91.8	96.2	88.5	96.7	90.0	90.0	40.0	40.0	40.0
740+620t	99.0	96.7	92.8	93.8	96.2	96.2	96.7	80.0	80.0	60.0	60.0	60.0
1080+740t	96.9	96.3	88.7	92.8	88.5	96.2	86.7	100.0	100.0	40.0	40.0	40.0
660+740	94.9	95.8	85.6	87.6	80.8	96.2	93.3	70.0	70.0	40.0	40.0	40.0

^{††} Wavelength, [‡] Classification Success, [†] Accuracy Rate (100-(100*Error Rate)),

(¹) Any defect considered as Bad and, stem, calyx, and good considered as Good, (²) Every group is considered separately,

(³) Confusion between defects is not allowed, and, stem, calyx, and good considered as Good, (⁴) Good Apples, (⁵) Bruised Apples,

(⁶) Stem and Calyx considered together, (⁷) Apples with Leaf Roller, (⁸) Punctured Apples,

(t) Symbolizes wavelength at which texture features are calculated; pixel gray values from the first wavelength and texture from the second

Table A.6 Classification results of five-class classification using single backpropagation neural network classifier (Golden Delicious)

WL ^{††} (nm)	CS (%) [‡] B-G ¹	AR (%) [†] B-G ¹	CS (%) All ²	CS (%) B-G ³	CS (%) Good ⁴	Bruise ⁵	StemCalyx ⁶	CS (%) Russet ⁷	CS (%) Bitter Pit ⁸
620+1260	94.9	95.8	88.7	89.7	79.2	100.0	95.0	61.5	100.0
1260	90.7	93.8	80.4	84.5	83.3	73.1	85.0	61.5	100.0
560	89.7	93.7	80.4	82.5	79.2	76.9	95.0	53.9	92.9
1180	89.7	93.7	80.4	83.5	83.3	69.2	95.0	53.9	100.0
1260+760t	89.7	94.4	83.5	85.6	79.2	80.8	95.0	61.5	100.0
620	87.6	93.8	79.4	79.4	75.0	73.1	100.0	53.9	92.9
540	85.6	93.8	78.4	79.4	66.7	80.8	100.0	46.2	92.9
1260+620t	82.5	92.6	75.3	77.3	75.0	53.9	95.0	61.5	100.0
760+1260t	81.4	92.8	77.3	77.3	75.0	73.1	85.0	53.9	100.0

^{††} Wavelength, [‡] Classification Success, [†] Accuracy Rate (100-(100*Error Rate)),

⁽¹⁾ Any defect considered as Bad and, stem, calyx, and good considered as Good, ⁽²⁾ Every group is considered separately,

⁽³⁾ Confusion between defects is not allowed, and, stem, calyx, and good considered as Good, ⁽⁴⁾ Good Apples, ⁽⁵⁾ Bruised Apples,

⁽⁶⁾ Stem and Calyx considered together, ⁽⁷⁾ Apples with Russet, ⁽⁸⁾ Apples with Bitter Pit,

(t) Symbolizes the texture features from the wavelength τ is adjacent; pixel gray values from the first wavelength and texture from the second

Table A.7 Results obtained from multiple backpropagation neural network classifier

Empire										
Effective Wavelength	CS (%) [‡]	CS (%)	CS (%)	CS (%)	CS (%)	CS (%)	CS (%)	CS (%)	CS (%)	CS (%)
Combinations (nm)	B-G ¹	All ²	B-G ³	Good ⁴	Bruise ⁵	StCl ⁶	Leaf R. ⁷	Puncture ⁸		
660, 740, 880	97.9	82.5	82.5	96.2	80.8	96.7	10.0	80.0		
680, 700, 880	99.0	84.5	84.5	92.3	92.3	100.0	10.0	100.0		
680, 700, 880, 600	97.9	83.5	85.6	92.3	92.3	100.0	20.0	80.0		
660, 700, 880, 600	99.0	82.3	83.5	96.2	92.3	100.0	10.0	80.0		
660, 700, 1280, 600	96.9	82.3	83.5	96.2	92.3	100.0	0.0	60.0		
Golden Delicious										
660, 1200, 1260, 1320	92.8	82.5	82.5	92.3	92.3	100.0	10.0	100.0		
660, 1260	91.8	82.5	82.5	96.2	80.8	96.7	10.0	80.0		

‡ Classification Success, (1) Any defect considered as Bad and, stem, calyx, and good considered as Good, (2) Every group is considered separately, (3) Confusion between defects is not allowed, and, stem, calyx, and good considered as Good, (4) Good Apples, (5) Bruised Apples, (6) Stem and Calyx considered together, (7) Apples with Leaf Roller (Russet for Golden D.), (8) Apples with Puncture (Bitter Pit for Golden D.)

APPENDIX B
Confusion Matrices for the Varieties and Classification Methods Used in 5-Class Classification

Table B.1 Using pixel intensities in single neural network (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	23	0	3	0	0
Bruise	0	24	0	0	2
Stem-Calyx	2	0	28	0	0
Leaf Roller	0	0	0	9	1
Puncture	0	2	0	3	0

Table B.2 Using pixel intensities in multiple neural networks (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	24	0	1	0	1
Bruise	0	21	0	0	5
Stem-Calyx	0	0	30	0	0
Leaf Roller	0	2	0	1	7
Puncture	0	0	0	0	5

Table B.3 Using pixel intensities from two images in single neural network (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	23	0	2	0	1
Bruise	0	25	0	0	1
Stem-Calyx	1	0	29	0	0
Leaf Roller	0	2	0	5	3
Puncture	0	1	0	3	1

Table B.4 Using pixel intensity and texture features in single neural network (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	25	0	0	0	1
Bruise	0	25	0	0	1
Stem-Calyx	1	0	29	0	0
Leaf Roller	0	1	0	8	1
Puncture	0	1	0	1	3

Table B.5 Using texture features in single neural network (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	21	1	2	1	1
Bruise	0	22	2	2	0
Stem-Calyx	3	0	26	0	1
Leaf Roller	0	0	0	9	1
Puncture	0	2	0	0	3

Table B.6 Using texture features in 1-nearest neighbor classifier (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	15	2	2	2	5
Bruise	1	21	0	2	2
Stem-Calyx	3	0	26	0	1
Leaf Roller	1	0	1	7	1
Puncture	0	2	0	0	3

Table B.7 Using texture features in 3-nearest neighbor classifier (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	16	1	2	3	4
Bruise	1	21	1	3	0
Stem-Calyx	4	0	24	0	2
Leaf Roller	0	0	0	9	1
Puncture	0	2	0	0	3

Table B.8 Using texture features in decision tree classifier (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	17	3	1	1	4
Bruise	1	23	1	0	1
Stem-Calyx	11	0	15	0	5
Leaf Roller	3	0	0	6	1
Puncture	1	2	1	0	1

Table B.9 Using texture features in Bayesian classifier (Empire)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	7	2	1	2	14
Bruise	2	15	0	5	4
Stem-Calyx	13	0	14	1	2
Leaf Roller	0	0	0	9	1
Puncture	0	2	0	0	3

Table B.10 Using pixel intensities in single neural network (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	20	0	2	2	0
Bruise	1	19	4	2	0
Stem-Calyx	2	0	17	1	0
Russet	1	4	0	8	0
Bitter Pit	0	0	0	0	14

Table B.11 Using pixel intensities in multiple neural networks (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	23	0	0	1	0
Bruise	0	22	2	1	1
Stem-Calyx	0	1	16	3	0
Russet	0	4	0	9	0
Bitter Pit	0	4	0	0	10

Table B.12 Using pixel intensities from two images in single neural network (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	19	1	1	3	0
Bruise	0	26	0	0	0
Stem-Calyx	0	0	19	1	0
Russet	0	3	0	8	2
Bitter Pit	0	0	0	0	14

Table B.13 Using pixel intensity and texture features in single neural network (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	19	0	1	4	0
Bruise	2	21	3	0	0
Stem-Calyx	1	0	19	0	0
Russet	1	4	0	8	0
Bitter Pit	0	0	0	0	14

Table B.14 Using texture features in single neural network (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	16	0	1	7	0
Bruise	0	20	0	6	0
Stem-Calyx	0	0	20	0	0
Russet	1	1	0	11	0
Bitter Pit	0	0	2	0	12

Table B.15 Using texture features in 1-nearest neighbor classifier (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	20	2	0	2	0
Bruise	0	18	0	8	0
Stem-Calyx	0	1	17	0	2
Russet	1	2	0	10	0
Bitter Pit	0	0	2	0	12

Table B.16 Using texture features in 3-nearest neighbor classifier (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	16	3	0	3	2
Bruise	0	20	0	6	0
Stem-Calyx	0	1	16	0	3
Russet	0	3	0	10	0
Bitter Pit	0	0	2	0	12

Table B.17 Using texture features in decision tree classifier (Golden D.)

	Good	Bruise	Stem-Calyx	Russet	Bitter Pit
Good	13	5	3	2	1
Bruise	0	23	0	3	0
Stem-Calyx	2	1	17	0	0
Russet	0	2	0	11	0
Bitter Pit	0	0	4	0	10

Table B.18 Using texture features in Bayesian classifier (Golden D.)

	Good	Bruise	Stem-Calyx	Leaf Roller	Puncture
Good	9	0	12	3	0
Bruise	0	18	2	6	0
Stem-Calyx	0	0	20	0	0
Russet	0	0	0	13	0
Bitter Pit	1	0	4	0	9

BIBLIOGRAPHY

BIBLIOGRAPHY

- Andris, H., B. Mitcham and C.H. Crisosto. 1999. Fruit physiological disorders (Apples). [Online] Available.
<http://postharvest.ucdavis.edu/Produce/Disorders/apple/pdapbit.html>
- Aneshansley, D.J., J.A. Throop and B.L. Upchurch. 1997. Reflectance spectra of surface defects on apples. Proceedings of the Sensors for nondestructive Testing International Conference, Orlando, Florida. P.143:160.
- Anonymous. 1998. Common fruit insects. Cooperative Extension work in Agriculture and Home Economics, State of Indiana, Purdue University and U.S. Department of Agriculture cooperating. H.A. Wadsworth, Director. [Online] Available
<http://www.entm.purdue.edu/Entomology/ext/targets/e-series/EseriesPDF/E-89.pdf>
- Birth, G.S. 1976. How light interacts with foods. Quality Detection in Foods. Publication of ASAE, St. Joseph, MI. P. 6-11.
- Birth, G.S. and G.L. Zachariah. 1976. Spectrophotometry of agricultural products. Quality Detection in Foods. Publication of ASAE, St. Joseph, MI. P. 1-5.
- Blackmore, B.S. and T. Steinhauser. 1993. Intelligent sensing and self-organizing fuzzy logic techniques used in agricultural automation. ASAE Paper No. 931048, St. Joseph, MI.
- Brown, G.K., L.J. Segerlind and R. Summit. 1974. Near-infrared reflectance of bruised apples. Transactions of the ASAE 17(1):534-536.
- Campins, J., A.J. Throop and D.J. Aneshansley. 1997. Apple stem and calyx identification for automatic sorting. ASAE paper no. 97-3079, St. Joseph, MI.
- Chen, P. 1978. Use of optical properties of food materials in quality evaluation of materials sorting. Journal of Food Process Engineering: (2) 307-322.

- Crowe, T.G. and M.J. Delwiche. 1996. Real-time defect detection in fruit-Part I: Design concepts and development of prototype hardware. *Journal of Food Process Engineering*: (2) 307-322.
- Deck, S.H., C.T. Morrow, P.H. Heinemann and H.J. Sommer. 1995. Comparison of a neural network and traditional classifier for machine vision inspection of potatoes. *Applied Engineering in Agriculture*. ASAE: 11(2): 319-326.
- Eberthart, R.C. and R.W. Dobbins. 1990. *Neural Network PC Tools*. Academic Press Inc. London.
- Geoola, F., F. Geoola and U.M. Peiper. 1994. A spectrophotometric method for detecting surface bruises on Golden Delicious apples. *Journal of Agricultural Engineering Research*: (58) 47-51.
- Graf, G.L. 1986. Automatic detection of surface blemishes on apples using digital image processing. PhD Thesis, Cornell University: 21-24.
- Guyer, D.E., G.E. Miles, M.M. Schreiber, O.R. Mitchell and V.C. Vanderbilt. 1986. Machine vision and image processing for plant identification. *Transactions of the ASAE* 29(6):1500-1507.
- Heinemann, P.H., C.T. Morrow, R.M. Crassweller, H.J. Sommer, W.N. Marrazzo and B. He. 1997. Machine vision based station for grading of Golden Delicious apples. *Proceedings of the Sensors for nondestructive Testing International Conference*, Orlando, Florida. P.239: 248.
- Haralick, R.M., K. Shanmugam and I. Dinstein. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*. SMC-3(6), November:610-621.
- Hassoun, M.H. 1995. *Fundamentals of Artificial Neural Networks*. A Brandford Book. The MIT Press, Cambridge, Massachusetts. P. 197.
- Howard, D. and M. Meale. 1992. *Neural network toolbox*. The math work inc.
- Jahne, B. 1995. *Digital image processing*. Springer-Verlag Berlin Heidelberg: 100.

- Jain, A.K. 1988. Pattern recognition. International Encyclopedia of Robotics Applications and Automation. John Wiley and Sons, Inc. P 1052:1055.
- Jain, A.K. 1989. Fundamentals of digital image processing. Prentice-Hall, Inc. Simon and Shuster. P 244.
- Jain, A.K. 1997. Pattern recognition and analysis course notes. CPS802, Michigan State University, USA.
- Jain, A.K., J. Mao, and K. M. Mohiuddin. 1996. Artificial neural networks: A tutorial. IEEE Transactions. 31-44.
- James, M. 1985. Classification algorithms. London: William Collins Sons and Co.
- Kasabov, N.K. 1996. Foundation of neural networks, fuzzy systems and knowledge engineering. A Bradford Book. The MIT Press, Cambridge, Massachusetts. P. 335.
- Lee, D.S., S.N. Srihari and R. Gaborski. 1991. Bayesian and neural network pattern recognition: a theoretical connection and empirical results with handwritten characters. Artificial neural networks and statistical pattern recognition (Sethi, I. K. and A. K. Jain, 1991): 89-108.
- Leemans, V., H. Magein and M.F. Destain. 1999. Defect segmentation on Jonagold apples using color vision and a Bayesian classification method. Computers and Electronics in Agriculture, 23(1999):43-53.
- Liao, K., M.R. Paulsen, J.F. Reid, B.C. Ni and E.P. Bonifacio-Maghirang. 1993. Corn kernel breakage classification by machine vision using a neural network classifier. Transactions of the ASAE 36(6):1949-1953.
- Luo, X., D.S. Jayas and S.J. Symons. 1999. Comparison of statistical and neural network methods for classifying cereal grains using machine vision. Transactions of the ASAE 42(2):413-419.
- Manhart, W. 1995. Apples for the twenty-first century. North American Tree Company. P. 243-246.

- Miller, B.K. and M.J. Delwiche. 1991. Peach defect detection with machine vision. Transactions of the ASAE 34(6):2588-2597.
- Muir, A.Y., I.D.G. Shirlaw and D.C. McRae. 1989. Machine vision using spectral imaging techniques. Agricultural Engineer, Autumn, 1989:78-80.
- Nakano, K. 1997. Application of neural networks to the color grading of apples. Computers and Electronics in Agriculture, 18(1997):105-116.
- Nakano, K., K. Takizaya and Y. Ohtsuka. 1995. The application of neural network systems to the evaluation of quality of fruits. IFAC Artificial Intelligence in Agriculture, Wageningen, the Netherlands: 105-200.
- Ni, B., M.R. Paulsen and J.F. Reid. 1997. Corn kernel crown shape identification using image processing. Transactions of the ASAE 40(3):833-838.
- Orchard, G.A. and W.A. Phillips. 1991. Neural Computation. Lawrence Erlbaum Associates Ltd., Publishers, East Sussex, U.K.83:84.
- Park, B., Y.R. Chen and M. Nguyen. 1996. Multispectral image analysis using neural network algorithm. ASAE Paper No. 96-3034, St. Joseph, MI.
- Patel, V., C., R.W. McClendon and J.W. Goodrum. 1995. Detection of cracks in eggs using color computer vision and artificial neural networks. ASAE Paper No. 95-3258, St. Joseph, MI.
- Precetti, C., J. and G.W. Krutz. 1993. Building a color classifier. ASAE Paper No. 93-3003, St. Joseph, MI.
- Raudys, S. and A.K. Jain. 1991. Small sample size problem in designing artificial neural networks. Artificial neural networks and statistical pattern recognition (Sethi, I. K. and A. K. Jain, 1991): 33-50.
- Sethi, I.K. and G.P.R. Sarvarayudu. 1982. Hierarchical classifier design using mutual information. IEEE Transactions of Pattern Recognition and Machine Intelligence. Vol PAMI-4, NO.4: 441-445.

- Shearer, S.A. 1990. Color and defect sorting of bell peppers using machine vision. Transactions of the ASAE 33(6): 2045-2050.
- Shearer, S.A. and R.G. Holmes. 1990. Plant identification using color co-occurrence matrices. Transactions of the ASAE 33(6): 2037-2044.
- Throop, A.J. and D.J. Aneshanesley. 1993. Investigation of texture analysis features to apple bruises. ASAE paper no. 93-3527, St. Joseph, MI.
- Timmermans, A.J.M. and A.A. Hulzebosch. 1995. Computer vision system for on-line sorting of pot plants using an artificial neural network classifier. Computers and Electronics in Agriculture 15(1996) 41-55.
- Tomita, F. and S. Tsuji. 1990. Computer analysis of visual textures. Kluwer Academic Publisher. London.
- Upchurch, B.L., H.A. Affeldt, W.R. Hruschka and J.A. Throop. 1991. Optical detection of bruises and early frost damage on apples. Transactions of ASAE 34(3): 1004-1009.
- Upchurch, B.L., H.A. Affeldt, W.R. Hruschka, K.H. Norris and J.A. Throop. 1990. Spectrophotometric study of buises on whole 'Red Delicious' aples. Transactions of ASAE 33(2):385-389.
- Upchurch, B.L., J.A. Throop and D.J. Aneshansley. 1994. Influence of time, bruise-type, and severity on near-infrared reflectance from apple surfaces for automatic bruise detection. Transactions of ASAE 37(5): 1571-1575.
- USDA-ARS Tree Fruit Research Laboratory. 1999. Physiological fruit disorders in the Pacific Northwest. 1104 N. Western Ave. Wenatchee, WA 98801 [Online] Available <http://www.tfrrl.ars.usda.gov/Disorders/applephysio.html#stembowl>
- Venables, W.N. and B.D. Ripley. 1994. Modern applied statistics with S-plus. 329-348.
- Wasserman, P.D. 1989. Neural Computing. Van Nostrand Reinhold, New York. 44-54.
- Weisskopf, V.F. 1968. How light interacts with matter. Scientific American, September. 60-71.

- Wooley, J.T. 1971. Reflectance and transmittance of light by leaves. *Plant Physiology*. 47:656-662.
- Yang, Q. 1992. The potential for applying machine vision to defect detection in fruit and vegetable grading. *Agricultural Engineering*, autumn 1992, 74-79.
- Yang, Q. 1993. Classification of apple surface features using machine vision and neural networks. *Computers and Electronics in Agriculture*, 9 (1993) 1-12.
- Yang, Q. 1993. Finding stalk and calyx of apples using structured lighting. *Computers and Electronics in Agriculture*, 8 (1993) 31-42.
- Yang, Q. 1994. An approach to apple surface feature detection by machine vision. *Computers and Electronics in Agriculture*, 11 (1994) 249-264.
- Yonekawa, S., N. Sakai and O. Kitani. 1996. Identification of idealized leaf types using simple dimensionless shape factors by image analysis. *Transactions of the ASAE*. 39(4):1525-1533.
- Young, T.Y. and T.W. Calvert. 1974. Classification estimation and pattern recognition. American Elsevier Publishing Company, inc. 95-100.
- Zwiggelaar, R., Q. Yang, E. Garcia-Pardo and C.R. Bull. 1996. Use of spectral information and machine vision for bruise detection on peaches and apricots. *Journal of Agricultural Engineering Research*: (63) 323-332.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 02092 6980