

**IMAGE BASED DENTAL IMPRESSION TRAY
SELECTION FROM MAXILLARY ARCHES USING MULTI-
FEATURE WITH ENSEMBLE CLASSIFIER**

MUHAMMAD ASIF HASAN

**FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2021

**IMAGE BASED DENTAL IMPRESSION TRAY
SELECTION FROM MAXILLARY ARCHES USING
MULTI-FEATURE WITH ENSEMBLE CLASSIFIER**

MUHAMMAD ASIF HASAN

**DISSERTATION SUBMITTED IN PARTIAL
FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF COMPUTER SCIENCE
(APPLIED COMPUTING)**

**FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2021

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Muhammad Asif Hasan

Matric No: WOA170005/17036423

Name of Degree: Master of Computer Science (Applied Computing)

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

IMAGE BASED DENTAL IMPRESSION TRAY SELECTION FROM
MAXILLARY ARCHES USING MULTI-FEATURE WITH ENSEMBLE
CLASSIFIER

Field of Study: Image Processing

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature

Date: 14/10/2021

Subscribed and solemnly declared before,

Witness’s Signature

Date: 14/10/2021

Name:

Designation:

**IMAGE BASED DENTAL IMPRESSION TRAY SELECTION FROM
MAXILLARY ARCHES USING MULTI-FEATURE WITH ENSEMBLE
CLASSIFIER**

ABSTRACT

Dental impression tray is frequently used in dentistry to record patient's oral structure for clinical oral diagnosis and treatment planning. Manual procedure of taking impressions is costly, time-consuming, and additionally, no research has been done to select dental impression tray from dental arch images using computer vision in real-life scenarios. In this spirit, an intelligent model is proposed based on computer vision and machine learning to select appropriate dental impression trays from maxillary arch images. A dataset of 52 patients' maxillary arch images that matches one from 4 sizes of Kurten's impression tray have been acquired. Various sets features such as, colors, textures, and shapes of the images were extracted to better characterize the maxillary arch images. Considering the importance of the features in describing the maxillary arch object and to improve the classification performance, a method based on multi-feature with ensemble classifier is proposed. Besides, the performance of a deep-learning based multilayer perceptron neural network is also investigated. The proposed multi-feature with ensemble classifier attained 92.31% precision, 91.75% recall, 91.75% accuracy, respectively, on the dataset. This clearly establishes the feasibility of this study. An illustration of a real-life application of the proposed model is also provided.

Keywords: Dental impression tray, automation in dentistry, computer vision, multi-feature, ensemble classifier.

**IMAGE BASED DENTAL IMPRESSION TRAY SELECTION FROM
MAXILLARY ARCHES USING MULTI-FEATURE WITH ENSEMBLE
CLASSIFIER**

ABSTRAK

Ceper impresi sering digunakan untuk merakamkan struktur mulut pesakit dan digunakan dalam diagnosis oral klinikal dan perancangan rawatan. Prosedur manual untuk merekodkan impresi adalah memakan wang, masa dan tambahan pula, belum ada penyelidikan dijalankan untuk pemilihan ceper impresi dari imej arkus pergigian menggunakan penglihatan computer dalam senario kehidupan sebenar. Maka, model pintar dicadangkan berdasarkan penglihatan komputer dan pembelajaran mesin untuk memilih ceper impresi yang sesuai dari imej arkus pergigian. Sebanyak 52 imej arkus pergigian yang sesuai dengan salah satu daripada 4 saiz ceper impresi jenama Kurten telah diperolehi. Pelbagai set ciri seperti, warna, tekstur dan bentuk dari imej arkus pergigian diekstrak untuk menentukan klasifikasi yang terbaik. Dengan mempertimbangkan ciri-ciri penting dalam mewakili arkus maksila dan mempertingkatkan prestasi klasifikasi, suatu kaedah berdasarkan penggabungan pelbagai ciri dengan pengklasifikasi ensemble dicadangkan. Selain itu, prestasi rangkaian neural perceptron pelbagai lapisan berasaskan pembelajaran mendalam juga disiasat. Penggabungan pelbagai ciri yang dicadangkan dengan pengkelasan ensemble mencapai ketepatan 92.31%, penarikan 91.75%, ketepatan 91.75% pada set data. Ini jelas membuktikan kebolehan kajian ini. Ilustrasi aplikasi sebenar dari model yang dicadangkan juga diberikan.

Keywords: Dental impression tray, automation in dentistry, computer vision, multi-feature, ensemble classifier.

ACKNOWLEDGEMENTS

First of all, my all praises go to Allah (SWT), the most merciful, the most gracious, the compeller who prepared me to complete this dissertation, leading to the partial fulfilment of this Master of Computer Science degree.

Again, I would like to acknowledge and give my warmest thanks to my supervisors, Mr Noorzaily Bin Mohamed Noor and Dr Norli Anida Abdullah. Their guidance and advice carried me through all the stages of writing my dissertation.

I am extremely grateful to my parents for their love, prayers, caring, and sacrifices for educating and preparing me for my future. As a mentor, they were always by my side and supported me morally throughout my life. Also, I would like to thank my wife Tasnia Hassan for her support and encouragement, which really helped me during this difficult time.

Moreover, my heartiest gratitude goes to all the academic staff and friends (specially Tonoy Roy) of the University of Malaya. Finally, I would like to say, all praise is due to Allah (SWT), by whose favors can accomplish all good.

TABLE OF CONTENTS

ABSTRACT	iii
ABSTRAK	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	ix
List OF TABLES	xi
LIST OF SYMBOLS AND ABBREVIATIONS	xii
LIST OF APPENDICES	xiii
CHAPTER 1: INTRODUCTION.....	1
1.1 Overview	1
1.2 Research Problem	4
1.3 Research Aim and Objectives	5
1.4 Research Questions.....	5
1.5 Scope of the Study	6
1.6 Research Significance	7
1.7 Organization of the Thesis	7
CHAPTER 2: LITERATURE REVIEW	9
2.1 Introduction	9
2.2 Prior Works on Dental Impression Tray	9
2.3 Computer Vision.....	11
2.4 Feature Extraction and Selection.....	11
2.4.1 Hand-Crafted Feature Extraction.....	12
2.4.1.1 Grey-Level-Co-Occurrence Matrix	13

2.4.1.2	Local Binary Pattern.....	14
2.4.1.3	Region-based Technique.....	16
2.4.1.4	Contour-based Technique	19
2.4.1.5	Color Features	20
2.4.2	Deep Learning Algorithms.....	24
2.4.2.1	Artificial Neural Network.....	24
2.4.2.2	Multilayer Perceptron Neural Network	25
2.4.2.3	Convolutional Neural Network.....	26
2.5	Classification Algorithms and Benefits of Ensemble Learning.....	28
2.5.1	K-Nearest Neighbor	29
2.5.2	Support Vector Machine	30
2.5.3	Random Forest.....	31
2.5.4	Gradient Boosting Classifier	32
2.5.5	eXtreme Gradient Boosting.....	33
2.6	Ensemble Method	34
2.6.1	Bagging	35
2.6.2	Boosting	36
2.6.3	Stacking.....	36
2.6.4	Voting and Averaging.....	36
2.7	Chapter Summary	37
 CHAPTER 3: RESEARCH METHODOLOGY		38
3.1	Introduction.....	38
3.2	Proposed Method.....	38
3.2.1	Image Acquisition.....	40
3.2.2	Image Augmentation.....	42
3.2.3	Getting Region of Interest	43

3.2.4	Feature Extraction.....	44
3.2.4.1	Color Features	44
3.2.4.2	Texture Features.....	46
3.2.4.3	Local Binary Pattern.....	50
3.2.4.4	Shape Features	50
3.2.5	Data Normalization.....	54
3.2.6	Proposed Ensemble Classifier	54
3.3	Chapter Summary	58
 CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION OF THE PROPOSED MULTI-FEATURE WITH ENSEMBLE CLASSIFIER		59
4.1	Introduction.....	59
4.2	Hardware Setup	59
4.3	Performance of Single Set Feature and Multi-Feature	63
4.4	Performance Based on Ensemble Classifier.....	64
4.5	Results on Multilayer Perceptron Neural Network	66
4.6	Discussion	68
4.7	Chapter Summary	72
 CHAPTER 5: CONCLUSION		73
5.1	Research Summary	73
5.2	Limitation and Future Works	74
	References	75
 LIST OF PUBLICATIONS AND PAPERS PRESENTED		85
 APPENDIX A: Source code for the proposed model.....		86
 APPENDIX B: Dataset		90

LIST OF FIGURES

Figure 1.1: Example of a patient's cast (a); an impression tray (b).....	1
Figure 2.1: GLCM creation process	13
Figure 2.2: Some earlier texture methods and LBP.....	15
Figure 2.3: Three neighborhood examples to describe a texture.....	15
Figure 2.4: RGB color space	21
Figure 2.5: HSV color space	21
Figure 2.6: L*a*b color space	22
Figure 2.7: An artificial neuron (a); An artificial neural network (b).....	25
Figure 2.8: One hidden layer MLP	25
Figure 2.9: Many layers of convolution neural network	27
Figure 2.10: K-NN classification criteria.....	29
Figure 2.11: Maximum margin hyperplane rule.....	30
Figure 2.12: Random Forest model making a prediction.....	31
Figure 2.13: Gradient Boosting Classification criteria	32
Figure 2.14: Weighted aggregation of several models	33
Figure 2.15: Weak learners can be combined to build a better predictive model	35
Figure 3.1: Block representation of the proposed method	39
Figure 3.2: Selected occlusal images for various maxillary arches: (a) crooked teeth; (b) buccal defect; (c) dental braces; (d) dental caries.....	41
Figure 3.2: Finding shapes from an image using contours	51
Figure 3.3: Image translation (a); Image rotation (b); & Image scaling (c).....	54
Figure 3.4: Proposed ensemble classifier.....	55
Figure 4.1: Example of maxillary arches and the trays	59
Figure 4.2: One hidden layer MLP NN.....	67

Figure 4.3: Machine learning classification report (model wise).....	70
Figure 4.4: Illustration of selecting an impression tray from the proposed model in real-life scenario	71
Figure 4.5: Example of identification failures on maxillary arches	72

Universiti Malaya

LIST OF TABLES

Table 3.1: Descriptions of the original images and augmented images	43
Table 3.2: Proposed algorithm for getting ROI	44
Table 4.1: Comparative performance of single set of features and multi-feature	63
Table 4.2: Comparative performance of multi-feature with ensemble classifier	65
Table 4.3: Experimental results of MLP NN on multi-feature.....	68

Universiti Malaya

LIST OF SYMBOLS AND ABBREVIATIONS

CAD	:	Computer-aided design
CNN	:	Convolutional neural network
GB	:	Gradient boosting
GLCM	:	Grey-level-co-occurrence matrix
JAC	:	Jaccard similarity score
KNN	:	K Nearest Neighbor
LBP	:	Local binary pattern
MCC	:	Matthews correlation coefficient
RF	:	Random forest
SVM	:	Support vector machine
XGboost	:	eXtreme gradient boosting

Universiti Malaysia

LIST OF APPENDICES

Appendix A:	86
Source code for creating feature sets	86
Source code for creating feature sets (Cont.)	87
Source code for creating feature sets (Cont.)	88
Source code for creating feature sets (Cont.)	89
Appendix B:	90
Dataset Link	90

Universiti Malaya

CHAPTER 1: INTRODUCTION

1.1 Overview

Dental casts (Figure 1.1-a) are accurate, three-dimensional replicas of a patient's teeth which are made by pouring dental plaster or acrylic into the impressions of the teeth. Dental casts are widely used in dentistry to provide fundamental diagnostic information where most of the advanced diagnosis and treatment planning in prosthodontics, indirect restorations, cosmetic and orthodontic treatments are relying on impression. Specially, for the further casting of dental prosthesis such as bridges, dentures and partial dentures dental casts are mainly used. According to (Gupta et al., 2017), it is necessary to select an appropriate impression tray (Figure 1.1-b) for making an accurate definitive cast.



Figure 1.1: Example of a patient's cast (a); an impression tray (b)

Moreover, improper use of impression tray size may fail in recording the tissue correctly. This kind of error may be corrected by adjusting the tray or choosing a better fitting tray and repeating the impression procedure. This procedure involves manual interactions by a human operator, which makes it burdensome, tedious, user-dependent, and sometimes creates waste of impression materials too, which a professional, well-run dental business should avoid. Furthermore, automation makes it possible to eliminate user dependency and permits evidence-based analysis (Raith et al., 2017). Thus, automatic

identification of dental impression tray has important scientific significance and can be a potential application to enable the automation of workflow in dentistry (Christensen, 2008). Also, it is crucial to facilitate the workflow with minimal cost and time and without automation it is not possible to achieve this task.

Motivated by this, to facilitate the workflow with minimal cost and time, a computer-vision based dental impression tray classification model is proposed in this study. It is worth noting that computer vision and image processing are used interchangeably in many contexts. Anyway, the vision-based impression tray classification model is proposed to select an appropriate impression tray aimed at working in real-life scenario. For instance, an intraoral camera can be used to capture the occlusal view of maxillary arch image and based on the image this classification model will select the appropriate tray. In this proposed study, computer-vision based features such as statistical color features, texture feature, morphological shape features, local binary pattern, color histograms are analyzed to design the vision-based impression tray classification model. These features were chosen because of their potential in describing the object and proven wide application in medical image analysis, e.g., the authors (Chatterjee et al., 2019) used these feature sets for classification of skin lesion types. However, in real-life scenarios, the maxillary arch images can be affected by different light sources, rotation and with the variation of the different types of mouth shapes due to the uncontrolled environment in dentistry. On this account, for this proposed study, the detection and extraction of suitable features is important. It is possible that extracting only one single set of features may not be enough to identify the maxillary arches properly by the classifier. The performance of the classification algorithms greatly relied on suitable features. Inability to detect suitable features may even decrease the classification accuracy as well. In fact, distinct discriminating capabilities are observed in different features. In this case, combination of several features have the high dimensionality which as a result recently drew great

attention in the medical image field because they can solve complex linear and non-linear recognition problems effectively (Yong-lian, 2020).

Therefore, this study focuses on identifying appropriate set of features to improve the classification performance in a limited sample dataset using multi-feature method. A dataset of maxillary arches (52 images) was utilized in this study due to the limited number of participants. Collecting and labelling a large amount of ground medical data is a challenge, especially in the scholarly area. Large-scale datasets or big data collected mainly through grand challenge, taking screening initiatives by countries, crowdsource or application programming interface (Litjens et al., 2017). For this study, it is difficult to fulfil the large-scale datasets requirement, such as the ImageNet dataset which consists of 3.2 million cleanly labelled images (Deng et al., 2009). Literature shows that several studies utilized limited dataset in their study for experimental purposes, e.g. the study by (Yilmaz et al., 2017) used 50 patients' cone beam computed tomography (CBCT) between the years 2013 and 2016 for designing computer-aided diagnosis of periapical cyst; (Akkoc et al., 2016) utilized 40 people's tooth plaster model for gender discrimination study, and (Bozkurt & Karagol, 2020) leveraged 20 images of dental radiographic for human identification. On this ground, through this study, the aim is to sample this maxillary arch dataset as a proof-of-concept. To resolve the overfitting issues which may arise because of this limited sample dataset, artificial image data augmentation has been performed on the dataset such as adding different illumination, scaling, and rotation. After identifying the features, five machine learning classification algorithms were evaluated separately on the set of features. Finally, to suit the problem stated previously and to better categorize the features of maxillary arches properly according to the tray, an attempt to investigate the combination of multiple sets of features with ensemble classifier was done. In addition, a deep-learning based multilayer perceptron (MLP) neural network (NN) was also analyzed as well. The performance of the proposed multi-

features with ensemble classifier method is evaluated in terms of precision, accuracy, recall, F1 score, Jaccard similarity score (JAC), Matthews correlation coefficient (MCC), Zero one loss function, and computational time.

In a nutshell, this study is based on the hypothesis that computer vision and machine learning based classification methods are capable of selecting appropriate dental impression tray from maxillary arch images with sufficient accuracy and can be of potential use in dentistry to automate the digital workflow.

1.2 Research Problem

Stock impression trays are available in various sizes and shapes to accommodate different mouths. It was shown that some commercially available stock impression trays may only be suitable for a particular patient population (Yergin et al., 2001). The development of image-based analysis in dentistry is not as extensive as in biomedical engineering and other associated fields which are inexpensive. Furthermore, in medical applications, computer-aided design (CAD) portrays a significant role such as designing surgical implant analysis, analysis of blood flow, designing prosthesis and many more (Shapi'i et al., 2011). However, leveraging these CAD systems in developing countries (such as Malaysia) is costly and not fully embraced by the dental community compared to other medical fields. The underlying reasons mostly related to initial high setup cost, changing the procedures that practitioners learned in the dental school and became used to it. As an example, the cost of the Dental CAD cam system is 850 USD (~ 3591 RM) (Made, n.d.). On the other hand, the proposed model will leverage a dental intraoral camera to capture the maxillary arch image, costing around RM 49.40 (Lazada.com.my, n.d.). Furthermore, it is worth noting that there is no CAD/CAM system to automatically select dental impression tray to the best of the literature analysis. The aforementioned discussion exhibits that there are still remains some room for improvement. In a nutshell the problem statement of this study is given below:

1. Manual inspection: manually inspecting the dental tray from a patient mouth is time-consuming, laborious, and suffers from the problem of inconsistency and inaccuracy in judgement by different human experts (Yergin et al., 2001).
2. High maintenance CAD systems: although CAD systems showed potential results in the dental community, but it is not fully embraced because of the high maintenance cost (Blatz & Conejo, 2019).
3. Not proper investigation: from the literature analysis (detailed discussion is given in Chapter 2), it is observed that proper investigation has not been performed on the selection of dental impression tray in real-life scenarios. The investigations are limited to small program (Yergin et al., 2001) or on statistical analysis using stone cast images (Rijal et al., 2011; Rijal et al., 2012).

1.3 Research Aim and Objectives

Based on the discussion, this study aims to propose a computer-vision based dental impression tray selection model to work in a real-life scenario. Therefore, the objectives of this study are as follows below:

1. To investigate challenges and problems that exist in previous research.
2. To find the necessary set of computer-vision based features for describing the maxillary arch object properly, according to work in real-life scenario.
3. To propose a novel approach for elevating the classification performance from a limited dataset.
4. To illustrate a computer-vision based dental impression tray selection system for dental practitioners.

1.4 Research Questions

From the defined aim and objectives, the following research questions was formulated:

1. What challenges and problems have been found in the existing dental impression tray research?
2. What kind of feature sets in computer vision are suitable to choose for describing the maxillary arch image properly according to work in real-life scenario?
3. How to evaluate the selected features and the classification algorithms for proposing a novel method to enhance the classification performance from a limited dataset?
4. How to propose a computer-vision based system that can be used in real-life scenarios according to physician's suggestion?

1.5 Scope of the Study

This study emphasizes on proposing a computer vision based dental impression tray selection model according to work in real-life scenario. The dataset consists of maxillary arch images of real-life patients, which is time-consuming to accumulate. It is necessary to take ethical approval (written in consent sheets) through several channels, and the data also needs to be medically validated by the physicians as well. Furthermore, in the case of deploying deep-learning algorithms, a much costly hardware requirement is demanded. Thus, the scope of this study is limited to the following extent:

1. The proposed method implemented on limited dataset due to the restricted number of patients to participate. Furthermore, this current global pandemic (Covid-19) creates hurdles to arrange the maximum number of patients to participate in this experimental purpose.
2. The target patients of this study are adult people ranging from (16-40) belongs to the people of Malaysia; thus, other demographic people of the same age and infant patients should not be taken into account.

3. The proposed proposed approach's feasibility is evaluated based on the rigorous evaluation of several algorithms (five machine learning algorithms together with a deep-learning-based classification algorithm).

1.6 Research Significance

Perhaps the most familiar form of health research is the clinical trial, in which patients volunteer to participate in studies to test the efficacy and safety of new medical inventions. In the healthcare sector, artificial intelligence has already begun to show several promising results. Computer vision (a subfield of artificial intelligence) not only helping medical professionals in saving their valuable time on the daily basic tasks but also saving patient's life as well. Anyway, this study holds that dental community (especially in the developing countries) requires urgent attention for a low-cost automated system that can enhance the quality of physician-patient interactions. This study aims to address and fill up this gap. Consequently, this gap is translated into a computer vision model to select a dental impression tray aims to work in real-life scenarios. Anyway, the contribution of this study is given below:

1. Proposing a first study to select a dental impression tray from the maxillary arch image in a real-life scenario to the best extent of my knowledge.
2. A novel method based on multi-feature with ensemble classifier is proposed to elevate the classification performance from a limited dataset.
3. Illustration of the utilization of the proposed model for dental practitioners.

1.7 Organization of the Thesis

This dissertation encompassed five chapters. Chapter 1 shortly describes the background of the dental impression tray. The potentiality of using computer vision for the automation of the dental impression tray and its advantage in dentistry is also

discussed. The contribution of this study and important highlights that gained from the proposed approach are mentioned at the end of the chapter.

Chapter 2 begins with the brief step-by-step explanation of the prior works on the dental impression tray by various researchers. In the subsequent sections, discussion about computer vision, the effectiveness of hand-crafted features over deep-learning algorithms, discussion of several classification algorithms, and benefits of ensemble methodologies is discussed broadly. In the end, a chapter summary is presented.

Chapter 3 illustrates the research methodology used in this study. A detail block representation of the proposed research methodology, data collection techniques, data augmentation techniques to prevent overfitting, pre-processing of the data, feature extraction, proposed ensemble classification method together with a brief summary is presented at the end of this chapter.

Chapter 4 exhibits the experimental setup, step-by-step experimental evaluation of the proposed methodology, and necessary empirical discussion of the proposed method's convenience and its effectiveness on a limited dataset. Each of the section is presented in detail. A chapter summary is also presented at the end of the chapter.

Chapter 5 concludes the research study and its achievements. It also presents some limitations and recommendations for future work.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter begins with the discussion of the prior works of various studies on dental impression tray selection from dental arch images. In the subsequent steps, description about computer vision, importance of feature selection in computer vision, various feature extraction methods, and the reason for choosing hand-crafted feature engineering over deep-learning is broadly presented. Additionally, various classification algorithms and benefits of ensemble learning methodology were presented as well. Finally, a brief summary is presented at the end of the chapter.

2.2 Prior Works on Dental Impression Tray

According to the literature analysis, the autonomous identification of dental impression trays based on real dental arch images are rarely implemented or described in the literature. As far the studies are limited to small program or chart or statistical analysis using stone cast images. In details the description of the prior works is described below.

- i. The authors (Yergin et al., 2001) proposed a method for finding a tray which best adopts the curvature of the patient automatically for helping practitioners. The casts were taken as test objects and the trays were taken as reference objects for the experimental purpose. A two-dimensional rigid body transformation was used to align the test object with the reference object. In a Euclidean space they represented the objects as point sets and the Euclidean distance transform was used to evaluate the residual distance of the test object to iteratively aligned the reference object. In the conclusion they mentioned that their proposed study can be used as a small program or chart. It can be concluded from this study that to propose a

model or method to work in real-life scenario; a necessary informative description such as, how to deploy the model in real life scenario should be provided. Furthermore, the study should provide some statistical evaluation to describe a conclusive outcome of how their work can be deployed in real-life situation.

- ii. Another two studies have been found in the literature related to stock impression trays. One of the studies by (Rijal et al., 2011) proposed a unique technique to represent dental arch shape in relation to stock tray design. They proposed a hybrid technique comprising dimension selection, clustering, and principal component to represent the shape of the dental arch. In their further study, (Rijal et al., 2012) investigated the homogeneity of 3 groups of arch shape (representation, clustering, and evaluation changes) and suggested 3 groups of multivariate (MV) normal distribution of these groups that may be used to probe the arch shape variation issue. The two studies seem promising, but their analysis is limited to stone cast images rather than utilizing real dental arch images. Besides, the above-mentioned studies did not propose any system or workflow to demonstrate their work according to real-life scenarios.

From the above-mentioned discussion, it is worth noted that although the proposed works seem promising but still several aspects of improvement can be done to further improve the studies. However, in the following sections a brief description of computer vision, importance of suitable features selection and benefits of multi-feature, together with the description of several machine learning and deep-learning algorithms are provided.

2.3 Computer Vision

Computer vision or image processing is an interdisciplinary scientific field that deals with how computers can gain a high-level understanding from digital images or videos. Computer vision-based problems appear simple because apparently human level of vision (even young children) can trivially identify the object even in the worst scenario. For example, identifying a fish in the tank or a fish in the aquarium appears the same although several background objects may exist. Nevertheless, it largely remains an unsolved problem in the computer vision area because of both the limited understanding of biological vision and the complexity of visual perception in a dynamic and nearly infinitely varying physical world. Several studies also mentioned that attaining the capability of human-like recognition remains a challenge in computer vision (Raith et al., 2017). However, progressive advancement in image recognition technology can be seen because of non-stop continuous development in this field. This kind of progress makes this technology open avenues of opportunity for real-world implementation, such as implementing computer vision in the medical image sector. In computer vision, image recognition technology is used to identify the instances of the object in the digital images. To identify the object properly, it is necessary to define appropriate features, and then utilizing the machine learning classification techniques to classify the objects. In simple terms, the features can be described as measurable characteristics or feature vectors of any object and based on their similarity, the classifier categorizes the features according to classes based on their similarity. Therefore, to represent the target object properly, it is necessary to find and select suitable features.

2.4 Feature Extraction and Selection

A comprehensive survey by (Litjens et al., 2017) shows that promising results have been achieved in medical image analysis utilizing computer vision. An example can be the 12th IEEE International Symposium on Biomedical Imaging (ISBI 2015), computer-

automated dental caries detection and X-ray image analysis of cephalometric challenge was defined as one of the grand challenges in the dental community. Research led by the authors (C.-W. Wang et al., 2016) presented a benchmark of dataset, methods, and results of the challenge for further future use of this benchmark. According to the literature, broadly the feature extraction methods in computer-vision mainly divided into two broad categories: (i) hand-crafted based feature extraction and (ii) deep-learning based feature extraction. In the following, the necessary details of the above two methods are provided. It is certain that the target image identification is not only restricted to medical image analysis but also related to other fields of image analysis as well. Therefore, several related areas of computer-vision based studies that can be fit or suitable within this study are included in the following section.

2.4.1 Hand-Crafted Feature Extraction

In computer-vision hand-crafted feature extraction generally refers extracting the significant features from the given image manually. At first, the collected, labelled images are pre-processed. Afterwards, the result-oriented features are extracted in order to form a master feature vector. Finally, the master feature vector is fed to machine learning algorithms to categorize or discriminate into categorizing outputs. In other words, features are also known as visual characteristics or visual features (i.e., to give a computer proper vision characteristic analogous to a human). Several visual features hold a dominant position in various tasks, e.g., bag-of-words is prevalently used in content-based image retrieval task (Ahmed et al., 2019). Anyway, this study focused on medical image-based applications; thus, the visual features that can be or has extensive use in medical sectors or another related field should be scrutinized. In computer-vision based feature extraction methods, especially in the medical image analysis, several authors assured that texture-based feature is considered successful (Mundim et al., 2016; Murugappan & Sabeenian, 2019). To illustrate an example, the study by (Talavera et al.,

2019) finds that while diagnosing skin cancers, essential information can be found while investigating the texture information of melanoma. The texture contains different discriminative information of the melanoma structure, and this provided great help to discriminate the images. The two of the prominent texture-based methods found according to the literature are: (a) Haralick features derived from the grey-level-co-occurrence matrix (GLCM), and (b) local binary pattern (LBP) (Ismail et al., 2019). Necessary description of these two popular texture features is described below.

2.4.1.1 Grey-Level-Co-Occurrence Matrix

To derive textual features, grey-level-co-occurrence matrix (GLCM) is currently the most prevalent statistical method (Zhang et al., 2020). The authors (Haralick et al., 1973) proposed GLCM for texture descriptor, and it has been popular texture descriptor because of its performance till now. According to (Dempere-Marco et al., 2002) GLCM uses a simple tabulation method to show the co-occurrence frequency of the intensity pixel values in an image. This technique first uses a spatial co-occurrence matrix that computes the relationship of pixel values, and then the second-order statistics will be computed by using the matrix (see Figure 2.1).

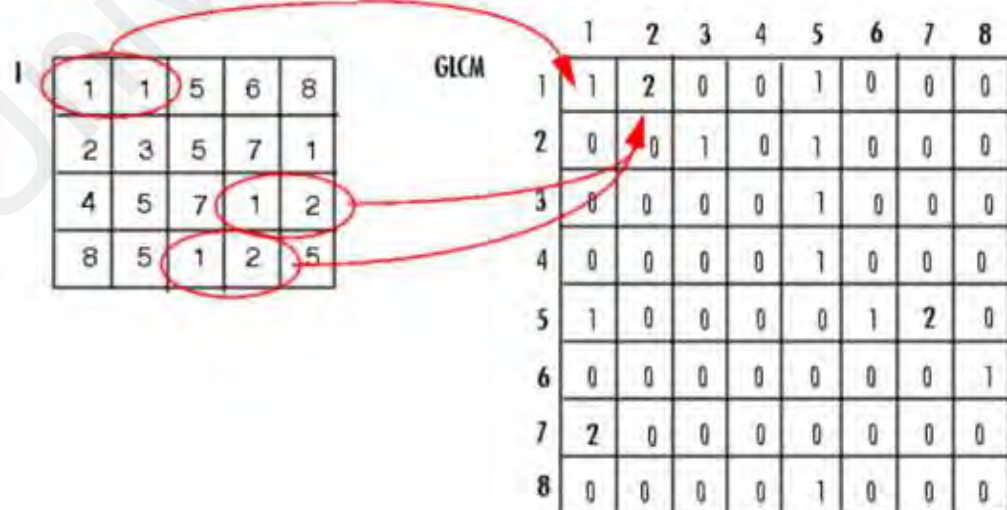


Figure 2.1: GLCM creation process

Typically, the calculation is obtained with the following two parameters: (a) relative distance between the pixel pair and (b) their orientation. For illustration, let $f(x,y)$ be a two-dimensional digital image with a $M \times N$ size and it has N_g grey levels. Then, a particular spatial relationship is satisfied from the GLCM matrix:

$$P(i, j) = \#\{(x_1, y_1), (x_2, y_2) \in M \times N | f(x_1, y_1) = i, f(x_2, y_2) = j\} \quad (2.1)$$

here, the number of collected elements is indicated by #. It should be noted that the size of the matrix P is $(N_h \times N_h)$. Let us assume, d is the Euclidean distance between the points (x_1, y_1) and (x_2, y_2) and the angle between these points is θ , then the various interval distances and angles of GLCM matrix can be determined as $P(i, j, d, \theta)$.

The advantage of GLCM is that it considers the spatial distribution of the intensity of the grey level in the neighborhood from an image. After computing the second-order statistics, different sets of textual features such as Standard deviation, Entropy, Energy, Homogeneity and so on are calculated using the resulting GLCM matrix.

2.4.1.2 Local Binary Pattern

Local Binary Pattern (LBP) is another robust greyscale local feature for texture classification according to (Kas et al., 2020) that firstly proposed by (Ojala et al., 1994; Ojala et al., 1996). It is defined as a greyscale invariant texture operator, derived from a general definition of texture in a local neighborhood. It is a simple but efficient method commonly used for feature extraction in classification tasks of many computer vision applications, including image classification (Nanni et al., 2012), object recognition (Heikkila, Pietikainen, & Schmid, 2009). LBP is related to many well-known texture analysis operators as illustrated in Figure 2.2. The arrows represent the relations between different methods, and the texts beside the arrows summarize the main difference between them. It was adopted in the conception of new frameworks to solve image recognition problem thanks to its ability to effectively represent facial images. In the objective of describing an image based on locality concept, the LBP descriptor labels the pixels value

of an image with decimal numbers and describe the local structure around each pixel. An overlapping block of 3*3 size is used in LBP, and the centre pixel of the 3*3 overlapping block is compared to neighbor pixels, and binary features are extracted (Figure 2.3).

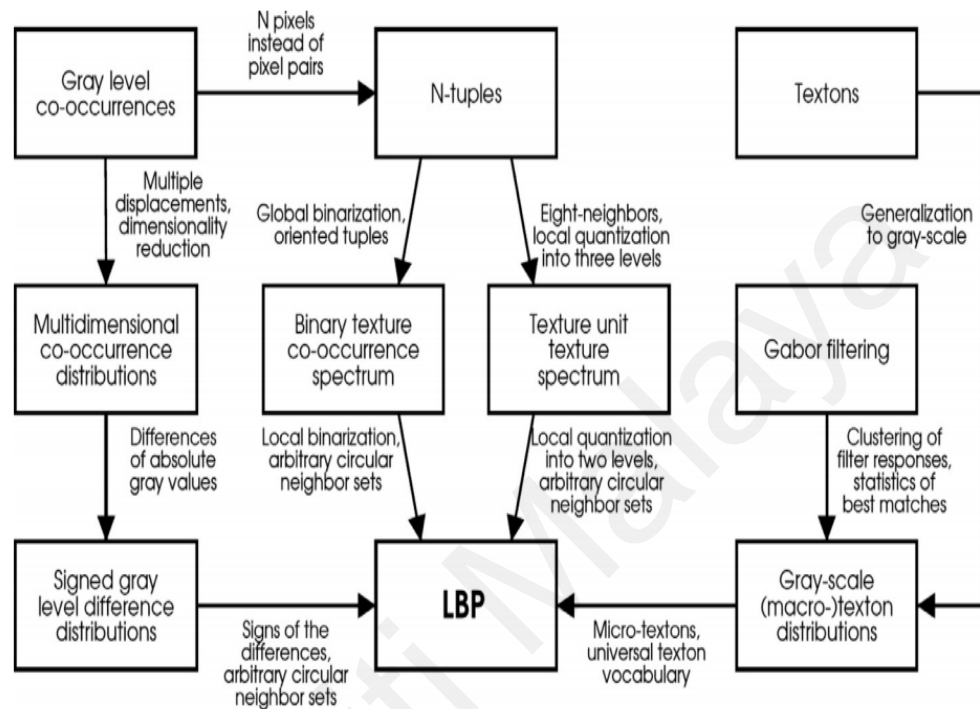


Figure 2.2: Some earlier texture methods and LBP (Hadid et al., 2015)

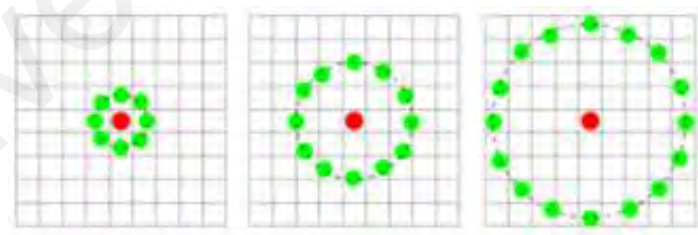


Figure 2.3: Three neighborhood examples to describe a texture

In the above discussion, the importance of the prominent texture feature description is provided. However, limitation exists into these methods such as, to properly distinguish objects leveraging GLCM method necessary texture information is necessary (Li et al., 2015) and LBP is unable to capture dominant feature information in large scale structure (Ojala et al., 2002). So, extracting only texture information from the target image may

not be helpful. Therefore, to describe the meaningful description of an object, other necessary visual features extraction is also necessary.

Investigations of various literature show that shape-based extracted features also showed prominent results in medical image analysis (Chatterjee et al., 2019; Kushwaha et al., 2018). Shape-based detection method focused mainly on edges and corner of an object which is a powerful recipe for accurate object detection. To represent the different types of shapes of various objects generally, two group of techniques employed: (a) region-based, and (b) contour-based technique.

2.4.1.3 Region-based Technique

Region-based methods are applied to the general shapes by representing interior shape information. For the region representation, geometric moments (Ahmed et al., 2017), Hu-moments (Lin et al., 2010), Zernike moments (Shirazi et al., 2016) are leveraged most of the cases. A brief description of these methods is described below.

1. Geometric Moment

Geometric moments have proven to be an efficient tool for image analysis (Xu & Li, 2008). Research on the utilization of moments for object characterization in both invariant and noninvariant tasks has received considerable attention in recent years (Kushwaha et al., 2018). The term “invariant” is often abused in some context, but generally, it refers that whether rotating the target image a little bit the value of the main object image and the reference image will be almost same. Example of using geometric moments are scene matching, image normalization, shape analysis, image retrieval, accurate position detector and many more tasks can be found in (Park & Kwon, 2019; Rizon et al., 2006). It is worth noted that to describe objects using image moments, segmentation of the object image and to

convert it into a binary image is necessary. For a two-dimensional density function $p(x,y)$ the $(p+q)$ th order geometrical moments m_{pq} can be defined as:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q p(x,y) dx dy \quad (2.2)$$

the moments' sequence can be uniquely determined by $p(x,y)$ if only $p(x,y)$ is a piece-wise continuous function and in the x - y plane it has only non-zero values (Hu, 1962).

2. Hu Moment Invariant

Hu moment invariants (or simply Hu moments) proposed by (Hu, 1962) are a set of 7 numbers calculated using central moments that are invariant to image transformations. The first six moments have been proved to be invariant to translation, scale and rotation and the 7th moment's sign changes for image reflection. These seven moments extensively used in several applications such as sign language recognition (Otiniano-Rodriguez et al., 2012), hand gesture recognition (Conseil et al., 2007) and many more. These invariant moments are defined as a special case of complex moments which almost similar to geometric moments. To illustrate an example, let us consider an image as a discrete function $f(x,y)$ with $x = 0,1,\dots,L$ and $y = 0,1,\dots,M$, then Hu moment can be defined as:

$$m_{pq} = \sum_{x=0}^L \sum_{y=0}^M (x + yi)^p (x - yi)^q f(x,y) \quad (2.3)$$

Major drawbacks of these seven moments are that they produce a larger value of geometric moments, which eventually led to noise sensitivity and numerical instabilities (Kotoulas & Andreadis, 2005).

3. Zernike Moments

From the above discussion, it can be noted that geometric moments and Hu moments are pretty much straightforward in their implementation, but they do have a large number of disadvantages. The major one is numerical instabilities because of their large number of dynamic ranges. Also, in the noisy images, high number of inaccurate results also found. To overcome this Teague (Teague, 1980) investigated image moments based on complex radial polynomials which resulted in new types of moments called Zernike moments. These moments are more robust to noise and also rotational invariance. Furthermore, utilizing moment normalization, scale and translation invariance can be implemented too. The Zernike moment of repetition q , and order p is described as follows:

$$Z_{pq} = \frac{p+1}{\pi} \sum_X \sum_Y f(x, y) W_{pq}(r, \theta) \quad (2.4)$$

2.4.1.4 Contour-based Technique

The contour-based technique describes shape properties using object outline (contour). This descriptor can efficiently describe the shape features if the object characteristics are available (Bober, 2001). This contour-based technique has successfully applied in the past in various research sectors, such as describing the popular MPEG shape descriptor (Bober, 2001). The contour-based technique follows some specific properties such as, high variability in an object is expected, which sometimes makes this technique robust to noise. According to (Bober, 2001), some important features of this technique are described below:

- i. It can distinguish between object shapes which have a similar region in pixel distribution.

- ii. Analogous to human semantical, this method searches shapes in an object; however, sometimes little intra-class variability can be ignorable.
- iii. This method is robust to nonrigid deformation in an image.
- iv. Due to perspective transformations, it is robust to the distortions in images and videos.

However, in this study, the geometric moments are leveraged as shape properties because they are more robust to rotation, scaling, which is generally required in real-life dentistry situation. Anyway, under the above discussion, it can be comprehended that to deploy these shape-based methods exploiting the boundary information is important. If the boundary information is distorted or unavailable, then deploying these methods is quite difficult. In some cases, shapes on several objects may be recognized the same; for example, in grey scale image object of different semantic groups may be recognized as the same. For example, the shape of a green color apple and a red color apple will be recognized the same because of the missing color information. Thus, a crucial finding is that sometimes to identifying or distinguishing the objects more specifically exploiting the color information along with shape is also necessary. Therefore, it is required to fuse color with shape to distinguish similar types of objects. In the next section, the importance of color features in computer vision is described.

2.4.1.5 Color Features

The human eye can distinctly recognize the color of an object. In computer vision, to visualize the object for extracting fundamental information and reveal spatial information, proper color analysis is necessary. Furthermore, by adding the spatial color information to the feature vectors can increase the object identification accuracy. Thus, in medical image analysis color features are also analyzed; for example, (Ali et al., 2020) used color features for detecting gastric abnormalities. For accurate object detection,

incorporating the color of an object boundary alongside the shape is a powerful method. At the different axes of an image contains the same color information; therefore, spatial coefficients possess an important object attribute. For color processing, it is important to select an adequate mathematical representation of color, such that all the features of color can be processed independently; basically, the most important features of color are intensity and chromaticity (Gonzales & Woods, 2002). Numerous color spaces are in the literature to represent color; selecting a color space depends on its characteristics. To ease the specification of colors within a three-dimensional coordination system or forming a subspace of the system where a unique point represents every color it is necessary to analyze the color space. In the literature, the color spaces for color image processing are the following: HSV (hue, saturation, value), RGB (red, green, blue), HSI (hue, saturation, intensity), YUV, L^*u^*v , L^*a^*b , YCbCr (Garcia-Lamont et al., 2018). However, in this study, three color space have been utilized, i.e., RGB, HSV and L^*a^*b . Description of each of these three color space is described below.

1. RGB Space

In this color space, every color is represented with the spectral components of red, green and blue. Initially, it was founded for the television technology, however, from digital cameras, scanners, computers to image storage, this color space is widely used. Almost all image processing software and graphics employ this model. The addition of the three individual components of color rays are considered as base of the block color. The values of each component correspond a point within the cube, but usually, the range are normalized to $[0,1]$; hence, space color is represented in a cube as shown in Figure 2.4.

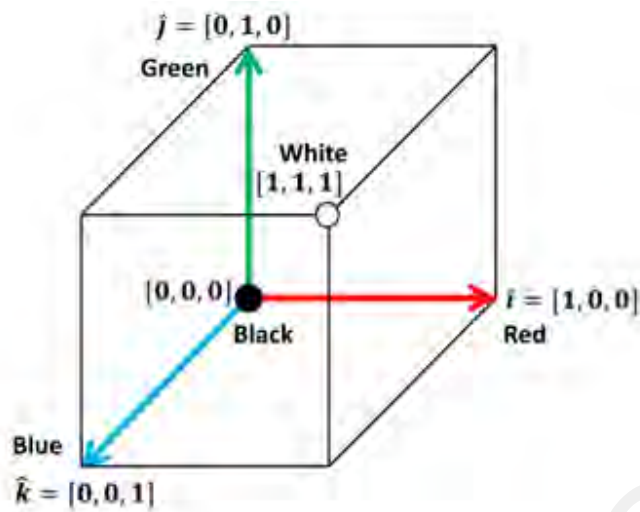


Figure 2.4: RGB color space

2. HSV Space

The HSV (Hue, Saturation, Value) model, also known as HSB (Hue, Saturation, Brightness), is another popular color space. It is a non-linear transformation of the RGB color space and may be used in color progressions. The HSV model is commonly used in computer graphic applications. In this way the HSV often the color wheel is used. In the wheel, the hue is represented by a circular region; a separate triangular region may be used to represent saturation and value (Figure 2.5).

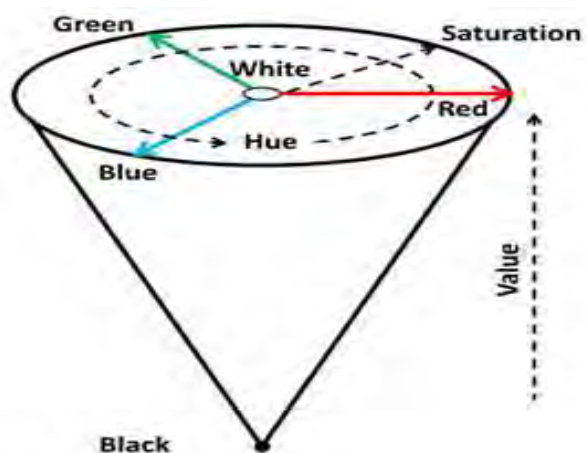


Figure 2.5: HSV color space

It is worth noting that for colors black, white and grey, the hue parameter is undefined. These colors are considered singularities within this color space because they do not have a specific chromaticity.

3. L^*a^*b Space

The CIELAB color space (also known as CIE L^*a^*b) is developed considering linearizing the tonality changes and is designed to approximate human vision. Three variables defined the colors, i.e., intensity represented by L^* and tonality components are represented by a^* and b^* . The location of the colors is different, but the space is similar to the RGB space (see Figure 2.6). The non-linear relations for L^* , a^* , and b^* are intended to mimic the non-linear response of the eye. The relative perceptual differences between any two colors in this space can be approximate by treating each color point in a three-dimensional space and counting the Euclidean distance between them.

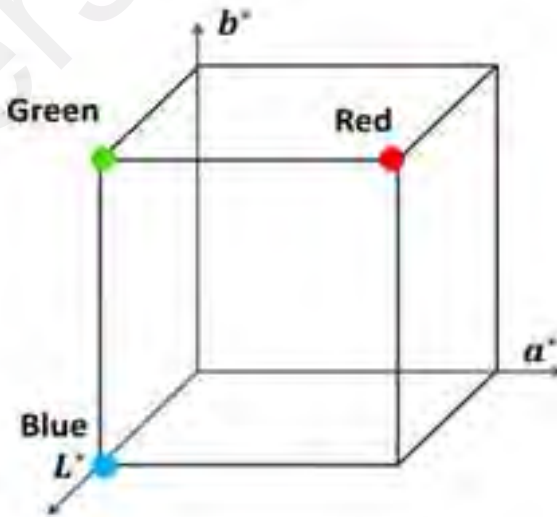


Figure 2.6: L^*a^*b color space

In this study, the above three color spaces were leveraged to describe the color of the maxillary arch object. The spatial color coefficients, along with shape attributes, can

distinguish and recognize objects from a versatile image category (Ahmed et al., 2019). Anyway, the major drawback of leveraging color features is that excessive color information is necessary for the visual perception of any object, and this drawback means sometimes using only, the color features makes it difficult to discriminate the object. An interesting point to note is that in certain scenarios, combination of shape with color features (Ahmed et al., 2019), combination of shape with texture features (Bagri & Johari, 2015), combination of color with texture (Barata et al., 2013) and combination of color, texture alongside shape features proved to be very efficient in elevating the classification performance (Tajeddin & Asl, 2018). For example, (Tyas et al., 2020) used morphological shape, texture and color features with multilayer perceptron neural network for erythrocyte classification in thalassemia cases; (Liu et al., 2019) utilized color and shape features for detection of apple fruits which even outperformed the region-based convolutional neural network. Therefore, it is necessary to find which set of features to choose and combine to increase the classification performance.

The comprehensive discussion above shows that selecting suitable feature sets plays a crucial role in describing an object properly. Only extracting a single set of feature set may not be enough to describe the maxillary arch image. As discussed, color and shape mimic the similar object of different categories, and the texture information distinguishes the object. Therefore, color in spatial coefficients, shape attributes and texture values in integration recognize complex and overlay objects from versatile image groups. Also, it is crucial to state that in real-life dentistry situation, various blurred, rotated and noisy image may be created. Thus, to describe the maxillary arch image properly, multiple set of features is necessary. In the next section, description of deep-learning algorithms and their application in computer vision is described.

2.4.2 Deep Learning Algorithms

In recent years, several scholars have turned to deep-learning algorithms because of their high accuracy and robustness. The primary advantage of deep-learning algorithms is that it incorporates the feature engineering step into its learning process (Dimitropoulos et al., 2017; Kowal et al., 2013; Spanhol et al., 2015). Thus, the burden of the human expert is shifted toward the machine. In addition, least domain knowledge is required in deep-learning approaches as compared to traditional approaches. The popular deep-learning algorithms, according to the literature, is described below.

2.4.2.1 Artificial Neural Network

Artificial neural network has increasingly gained attention in several fields of computational areas where computation is complex or/and intensive computation is required. Artificial neural network has shown promising results in medical image analysis, for example, detection of cardiac imaging (Mannil et al., 2020), diabetic retinopathy detection (Smitha et al., 2018) and so on. Mathematically, artificial neural network can be view as a function $F: u \rightarrow v$; Here, $u = [u_1, u_2, \dots, u_{n_0}]$ is the input and the output is $v = [v_1, v_2, \dots, v_{n_0}]$. In Figure 2.7, a details description is given. It is actually a series of artificial neurons, where a weighted sum operation is performed on the input. Finally, to produce an output, an activation function of a weighted sum is performed. One of the prominent examples of artificial neural network architecture is feed-forward neural network; where only the forward direction information moves (i.e., via N-1 hidden layer, from the input nodes to the output nodes).

2.4.2.2 Multilayer Perceptron Neural Network

Multilayer perceptron (MLP) neural network (NN) is a class of feed-forward neural network or a special case of artificial neural network composed of multiple layers of perceptron (with threshold activation). It is also called second-generation neural network

or shallow neural network and commonly has 1 hidden layer (≤ 2 hidden layers). MLP NN utilize supervised learning technique together with threshold activation. The difference can be observed between an MLP and a linear perceptron because of its multiple layers and non-linear activation. MLP NN is also shown promising results in medical vision sectors, for example, (Avuclu & Basciftci, 2019) used image processing techniques with MLP NN to determine age and gender by examining teeth, and bone structures, (Cervantes-Sanchez et al., 2019) used Gaussian filters and Gabor filters together with MLP NN to automate the segmentation of Coronary arteries etc.

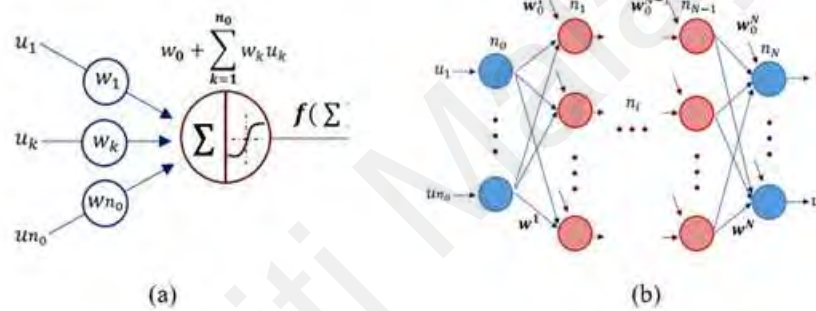


Figure 2.7: An artificial neuron (a); An artificial neural network (b)

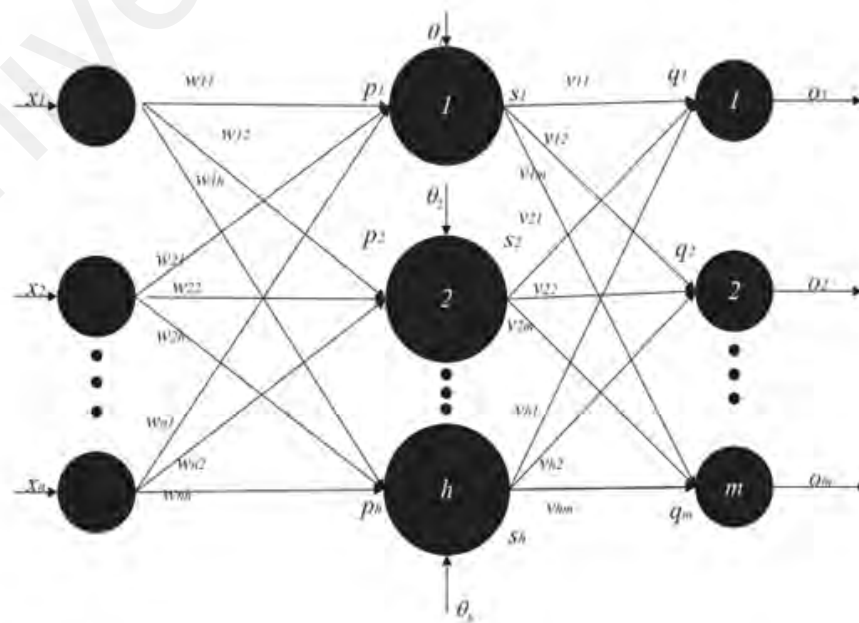


Figure 2.8: One hidden layer MLP (Khishe & Mosavi, 2020)

A three-layer MLP is shown in Figure 2.8. Here, n defined the number of input nodes, h defined the number of hidden layers and m defined the number of output nodes. Trial and error method is usually used to determine the architecture of MLP NN (Avuclu & Basciftci, 2018,2019).

2.4.2.3 Convolutional Neural Network

Convolutional neural network (CNN) is a class of deep neural networks, most commonly used for image segmentation and classification (Shen et al., 2017). From the sample images, CNN can extract features automatically and does not depend on handcrafted feature extraction. CNN used in medical image analysis for automatic classification of peripheral blood cells (Acevedo et al., 2019), region extraction and classification of skin cancer (Saba et al., 2019), detection of medical text semantic similarity (Zheng et al., 2019) and so on. The organization of human visual cortex inspired it and the connectivity pattern is also analogous to the connectivity pattern of human brain. In order to train and test through CNN, each input image will pass through a series of convolutional layers with filters also knows as kernels, pooling, and fully connected layers. Finally, to classify or recognize object according to the category, SoftMax function is used with probabilistic values (either 0 or 1). In the following figure (see Figure 2.9) a complete procedure of CNN is illustrated to show how CNN can process the input image and classifies the images between values. The first convolution layer will filter the fixed-size input images with suitable kernel size and obtain the output.

$$\text{Conv}^k(i, j) = \sum_{u, v} W^{k,l}(u, v) \cdot \text{input}^i(i - u, j - v) + b^{k,l}$$

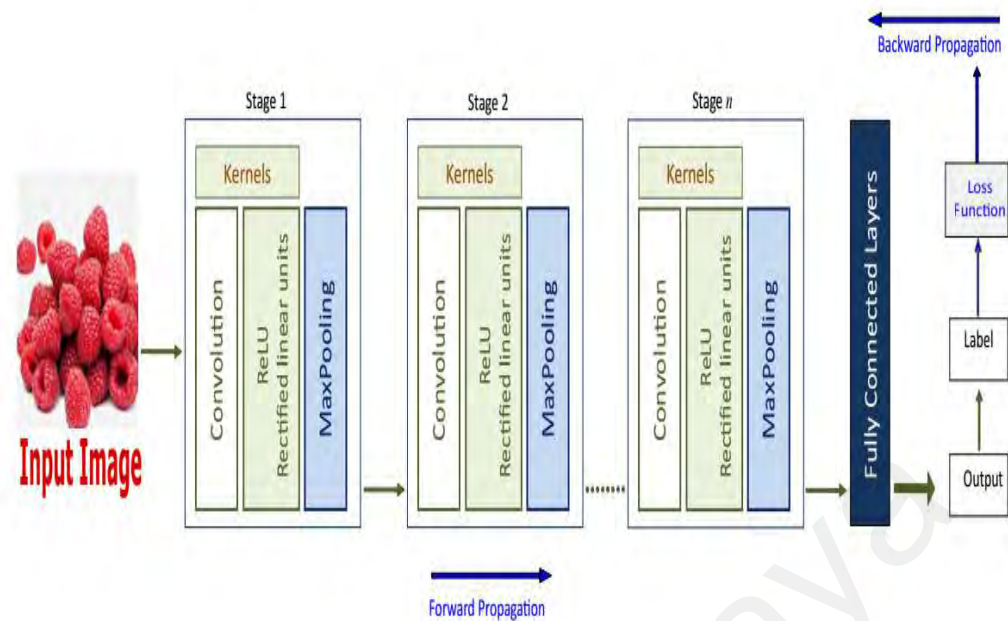


Figure 2.9: Many layers of convolution neural network (Naranjo-Torres et al., 2020)

Here, $W^{k,l}$ denotes the k th kernel and $b^{k,l}$ denotes the bias of the k th kernel. The output of the first convolution layer is connected with a max-pooling layer. Further, the second and third max-pooling/convolutional layer are connected to one another until reaching the desired output size. The fully connected layer neurons are the desired deep features.

From the above discussion, it can be observed that deep-learning algorithms in various medical image sectors showed promising results. Thus, one can conclude that utilizing deep-learning algorithms in medical image analysis is like holy water. However, several drawbacks exist for deploying these algorithms. For example, these algorithms require a massive amount of labelled image data for training. Collecting and labelling these massive data requires a great amount of time and labor cost. To give an instance, the authors (Acevedo et al., 2019) utilized a dataset of 17,092 images for the recognition of peripheral blood cell images. These massive amounts of image collection were done by Biomedic Diagnostic Centre using the CellaVision DM96 analyzer. Again, to train those

massive images, better hardware requirement is needed. For instance, the same authors utilized Nvidia Titan XP Graphics Processing Unit, which almost cost 4,000 MYR according to (Shopee Malaysia, n.d.). In scholarly areas, amassing these amounts of massive data together with the computational resources is quite a challenge. These issues create hurdles to deploy these algorithms. In some cases, traditional methods even outperformed deep learning algorithms too. For example, to detect apple fruits in real-life scenario the authors (Liu et al., 2019) utilized color and shape features which even outperformed the region-based CNN. As a result, scholars still relied on traditional methods that made of hand-crafted image features and machine learning classifiers (Liu et al., 2019; Qiao et al., 2019).

2.5 Classification Algorithms and Benefits of Ensemble Learning

Classification algorithms play a major in compute vision. The intent of the classification process is to categorize all pixels in a digital image into varieties of classes or “themes”. In digital image analysis, perhaps classification is the most important part. In digital image analysis, after extracting the suitable features, perhaps classification is the crucial task. A suitable classification algorithm and a sufficient number of training samples are prerequisites for a successful classification (Lu & Weng, 2007). In machine learning, classification algorithms are the part of supervised learning tasks, i.e., to identify examples of the information classes of interest in the image. Sometimes, these are referred to as “training site”. The opposite part of supervised learning is unsupervised learning which examines a large number of unknown pixels and divides into a number of classes based on natural grouping present in the image values. However, this study focused on the supervised classification task. Several classification algorithms such as k-nearest neighbor (KNN) (Balasubramanian et al., 2018), support vector machine (SVM) (de Sousa Costa et al., 2018), random forest (RF) (He et al., 2019) along with others are leveraged in image classification problems. Anyway, it is not possible to leverage all

aspects of the classification algorithms; thus, in this study based on the literature analysis, five classification algorithms were chosen. These classification algorithms have shown their establish performance and efficacy in image recognition problems tasks. In the following brief description of these classification algorithms is given.

2.5.1 K-Nearest Neighbor

K-Nearest Neighbor (K-NN) is one of the most widely used classifiers in image recognition problems (Ismail et al., 2018). K-NN classifier is an instance-based learning classifier where the hypothesis is constructed directly from training instances. The K-NN classifier is simple and easy to implement for classification. K-NN classifies an object by referring to the feature similarity where the object is assigned to the class based on the majority vote among its k nearest neighbor (Figure 2.10). Here, the k value is typically a small positive integer.

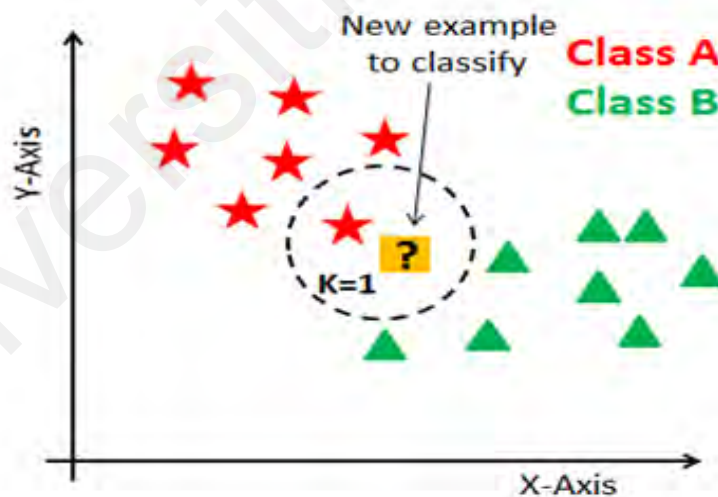


Figure 2.10: K-NN classification criteria

2.5.2 Support Vector Machine

The support vector machine (SVM) is another high-performance classification algorithm and is widely praised in many image classifications tasks (Attamimi et al., 2013; Auria & Moro, 2008; Ozkan et al., 2015). SVM generally follows the maximum

margin hyperplane rule for dividing data and linearly classifies the data (Figure 2.11). However, when the data is not linearly separable, 'kernel trick' is applied to create non-linear classifiers for maximum-margin hyperplanes (1 from reference folder). Some common kernels include:

- i. Homogenous (polynomial kernel)
- ii. Inhomogeneous (polynomial kernel)
- iii. Gaussian radial basis function
- iv. Hyperbolic tangent

The kernel is the transformation by any suitable equations. The effectiveness of SVM depends on the selection of kernel, the kernel's parameters, and soft margin parameter C .

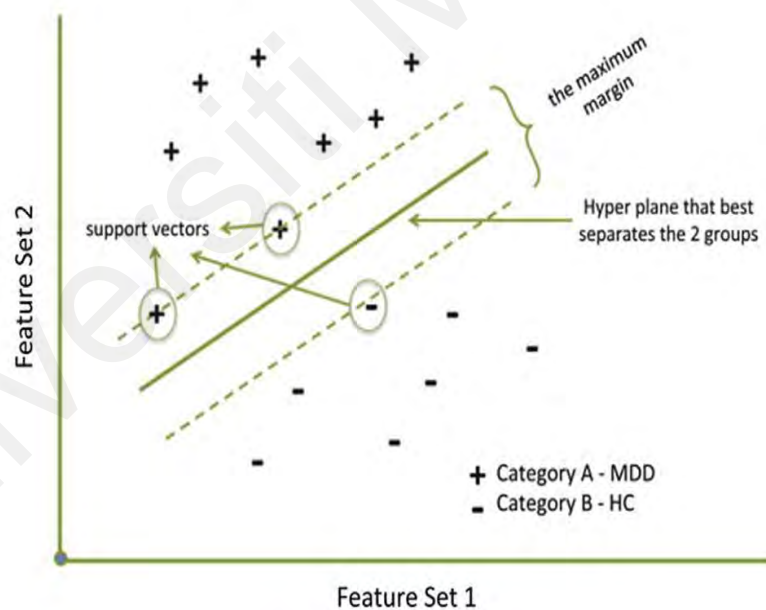


Figure 2.11: Maximum margin hyperplane rule (Pisner & Schnyer, 2020)

2.5.3 Random Forest

Random forest is another prominent classifier, which can predict the class of input, and select the most important features by providing feature importance (i.e. the input vector) (He et al., 2019). As the name implies, Random Forest consists of a large number

of individual decision trees, that operates as an ensemble. Using a bootstrap sample of training data, each tree is finally constructed. This kind of voting mechanism for a large number of decision trees can easily generate training models with a small amount of training data, fast training, less manual intervention, and robustness to noise data and data with missing values, and thus improves the prediction accuracy of the model (Figure 2.12).

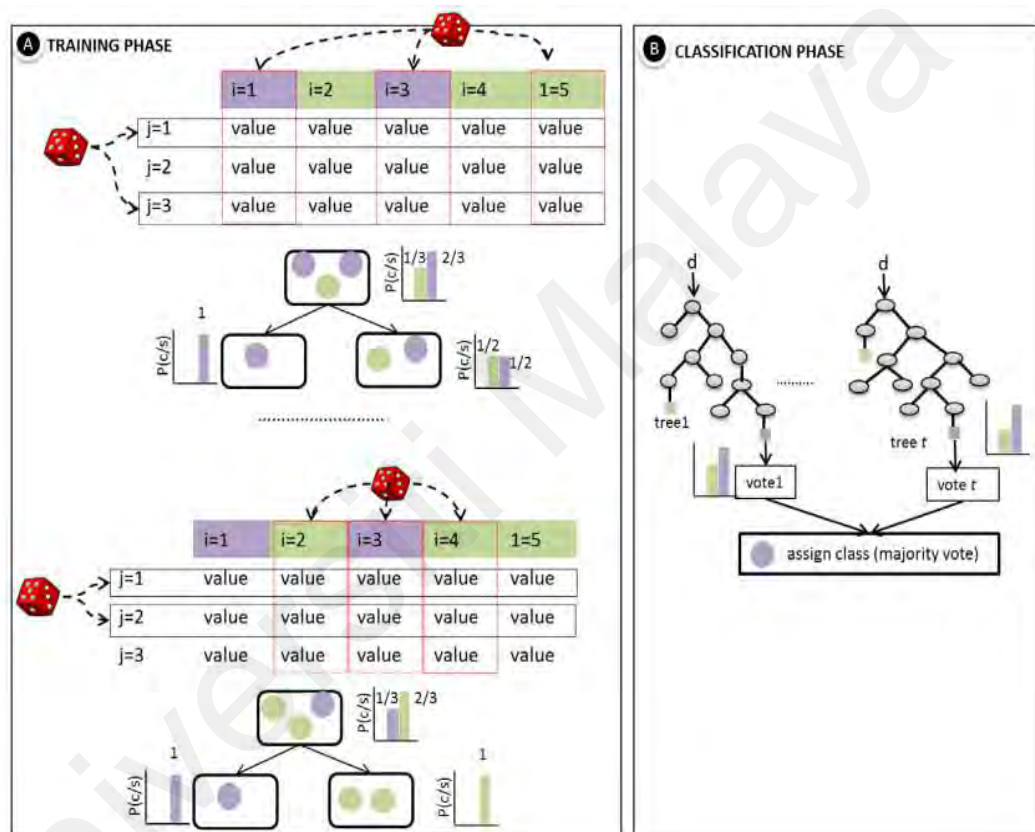


Figure 2.12: Random forest (RF) model making a prediction (Belgiu & Dragut, 2016)

2.5.4 Gradient Boosting Classifier

Boosting is a general method for improving the performance of any given learning algorithm (Borges & Neves, 2020). Boosting classification strategy is the process of combining many weak learning classification algorithms with limited predictive ability into a single more robust classifier capable of producing better prediction of a target.

Gradient boosting (GB) builds an additive model, following a forward stage-wise fashion; and finally allows optimization of arbitrary differentiable loss functions (Hastie et al., 2009) (Figure 2.13). A special case is a binary classification where a single regression tree is included.

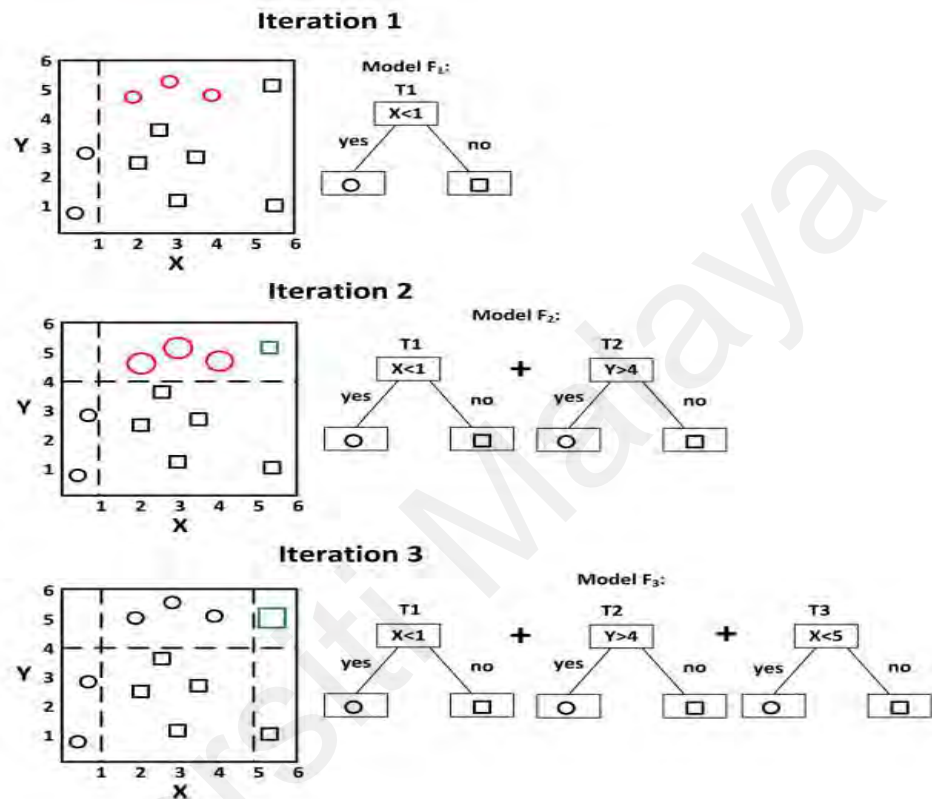


Figure 2.13: Gradient Boosting Classification criteria (Zhang et al., 2018)

2.5.5 eXtreme Gradient Boosting

eXtreme gradient boosting (XGboost) is another implementation of gradient boosting concept but follows a more regularized model formalization to control over-fitting, evidently give it a better performance (Tahmassebi et al., 2019). In a large number of competitions, for example, in Kaggle community it shows predominant results in many machine learning tasks (e.g., classification or regression). XGboost also follows the ensemble learning method where the output result is a single model which eventually came from an aggregated weighted output from several models (Figure 2.14). Contrast to

RF, where trees are grown to their maximum extent; boosting makes use of trees with fewer splits. Generally, researchers agree that compared to other gradient boosting algorithms XGboost works faster (Tahmassebi et al., 2019).

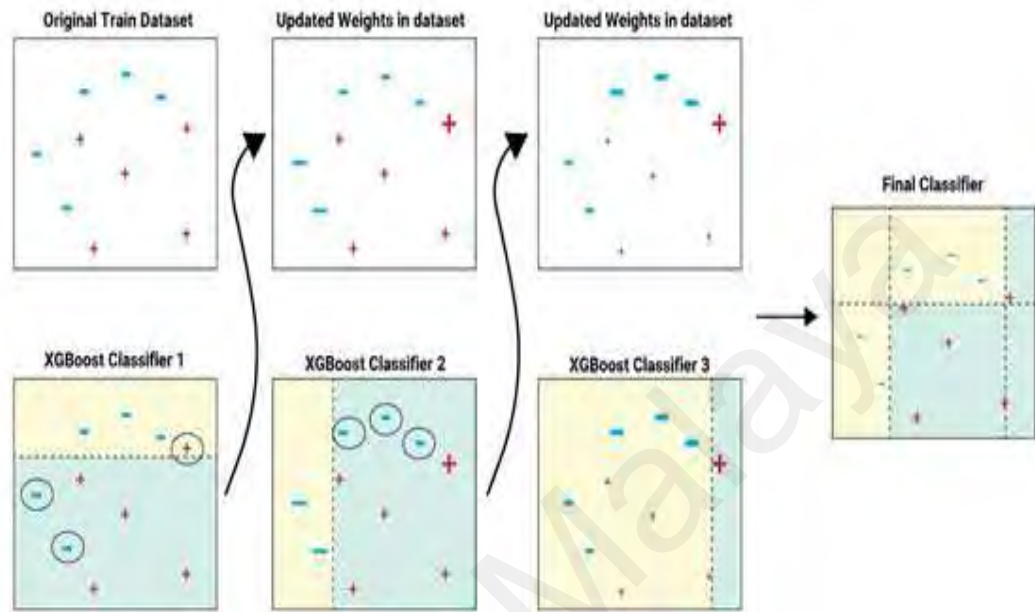


Figure 2.14: Weighted aggregation of several models (Tahmassebi et al., 2019)

The above discussed five classification algorithms have been used for classification tasks in this study. Each classification algorithms have their merits and demerits, for example, SVM could simplify calculations and avoid the problems of overfitting on non-linear and high-dimensional data features (Qiao et al., 2019), whereas K-NN is based on the majority vote that classifies through calculating the distance between data. The distance of the data from different classes may be similar, which may increase misclassification (Duneja & Puyalnithi, 2017). In contrast with the single algorithms, ensemble learning methods attain better predictive performance by utilizing multiple learning algorithms than what could be obtained from any of the individual learning algorithms alone (the learning algorithm can be either classification or regression). Therefore, in this study, ensemble classification methodology is leveraged to enhance the

classification performance. Anyway, in the next section, various ensemble methodology is described in detail.

2.6 Ensemble Method

Ensemble methods use multiple learning algorithms to gain better predictive performance which could be obtained from any single learning algorithm alone (Opitz & Maclin, 1999; Polikar, 2006; Rokach, 2010). The learning algorithms can be either classification or regression-based algorithms. To make a good prediction from a particular problem, supervised learning algorithms perform the task through a hypothesis space (Blockeel, 2011). Ensemble methods also are a class of supervised learning algorithms because they need to be trained and make a prediction from the data. Therefore, a single hypothesis is represented through the trained ensemble. If there is a significant diversity among the models; then ensemble methods tend to yield better predictive results (Kuncheva & Whitaker, 2003; Sollich & Krogh, 1996). To set up ensemble learning methods, it is necessary to set up the base models to be aggregated. The base models sometimes called weak learners. Most of the time, these basic models perform not so well by themselves either because they have a high bias (e.g., low degree of freedom models) or they have too much variance to be robust (e.g., a high degree of freedom models). Then, the idea of ensemble methods is to reduce bias and/or variance of such weak learners by combining several of them together in order to create a strong learner that evidently achieves better predictive performance (Figure 2.15). In the following brief discussion of several ensemble methods are described below.

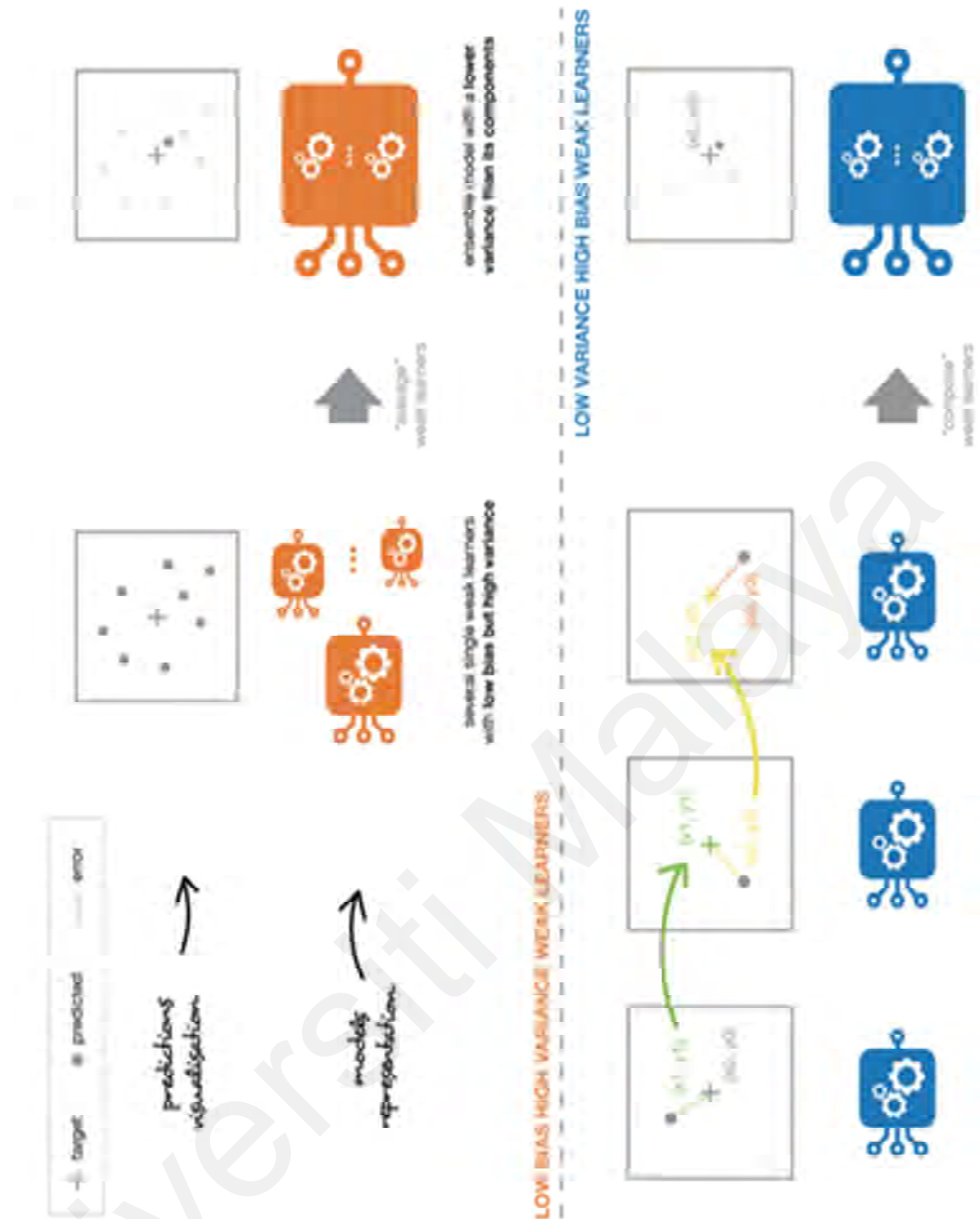


Figure 2.15: Weak learners can be combined to build a better predictive model

2.6.1 Bagging

Bagging (stands for Bootstrap Aggregating) is one of the popular methods of ensemble methods which generate additional data in the training stage and decrease the variance of a learning model by producing random sampling in the original set by replacing the original set. In the new dataset, every element has the probability to appear again. It narrowly tunes the prediction to reach an expected outcome by decreasing the variance

(Bryll et al., 2003). An example can be the decision tree model which tend to have a high variance. Therefore, bagging is applied to it, and the prominent example is the random forest which uses over-bagging (Panov & Džeroski, 2007).

2.6.2 Boosting

Boosting is another general ensemble method that creates a strong classifier from a number of weak classifiers. The term ‘Boosting’ refers to a family of algorithms that convert a number of weak learners to a strong learner. The multiple learner’s data samples create weight at each point of the iteration and create new sets more often. A higher weight is assigned to the learner with good prediction result. A successful boosting algorithm example is AdaBoost for binary classification (Domingo & Watanabe, 2000). With bagging the average of each individual model is creates. Instead in boosting, a new model is created, and the new base-learner model is trained (updated) from the errors of the previous learners.

2.6.3 Stacking

Stacking is an ensemble learning technique that combines multiple classification or regression models via a meta-classifier or a meta-regressor. From the complete training set, the base level learning model (or weak learning) model was trained, and finally, the meta-model is trained on the outputs of the learning model (base model) as features. Heterogenous characteristics often observed in stacking ensemble learning model because the base learning algorithms mostly consists of various learning algorithms (Y. Wang et al., 2019).

2.6.4 Voting and Averaging

Voting and averaging two of the easiest ensemble methods. They are easy to understand and implement and most used by researchers according to (Hosni et al., 2019). For classification tasks voting is used, and averaging is used for regression task. From the

training dataset, multiple learning models (i.e., classification) is created, and like an electoral system prediction on a new data point is made based on various voting scheme.

The above-discussed ensemble methodologies have been widely reported in the literature. As voting ensemble methodologies mostly used by researchers according to (Hosni et al., 2019) thus, the voting ensemble methodology is leveraged in this study.

2.7 Chapter Summary

As presented in the above comprehensive discussion, none of the studies focused on selecting an appropriate dental impression tray in real-life scenario using computer vision to the best extent of my knowledge. Moreover, the selection of impression tray from maxillary arches consists of the objects' changeability, background interference, and shortage of prior knowledge. It is observable that color, texture, and shape features show promising results in different fields of computer vision problems along with medical image analysis, as previously discussed. Thus, in this study, these features are chosen to describe the maxillary arch images. The limited number of features cannot precisely express these maxillary arch images, and the efficiency and stability of the classification algorithms should be preferentially considered as well. Therefore, an investigation of fusing multiple sets of features was done to elevate the classification accuracy for selecting appropriate impression tray. Also, to decrease the variance, bias error and to improve classification rate an ensemble classifier is leveraged combining the best classifiers using soft voting approach. In the next chapter, the proposed research method of this study is presented.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

This chapter portrays the detailed research methodology to design the proposed vision-based dental impression tray, selection model. At first, a block representation of the proposed method is given to briefly summarize the workflow. Subsequently, explanation of each section from the block is discussed in detail. Finally, a summary of this chapter is described as well at the end of the chapter.

3.2 Proposed Method

The complete research methodology to design the proposed vision-based dental impression tray classification model is described in this section. In Figure 3.1, a block representation for the proposed dental impression tray classification model from maxillary arch images is presented. Five main phases generally comprised for the proposed study: (i) acquisition of the image data, (ii) data augmentation, (iii) getting region of interest from the image, (iv) extracting discriminative features and (v) evaluating the classification algorithms and designing the ensemble classifier. For, image data acquisition of the data, 52 patients' have been arranged by taking ethical approval from SEGi University, and their maxillary arch image data have been acquired. After that data augmentation was performed to improve the generalization ability of the training data and to avoid overfitting issue. To get the region of interest (ROI) from the maxillary arch images, several tasks were performed. Through various feature extraction methods, several discriminative and informative features were extracted from the maxillary arch images. In the fifth phase, five classification algorithms were examined and from the classification algorithms results, an ensemble classifier was designed. Detailed working explanation of the all the steps as mentioned above are described in the subsequent sub-sections.

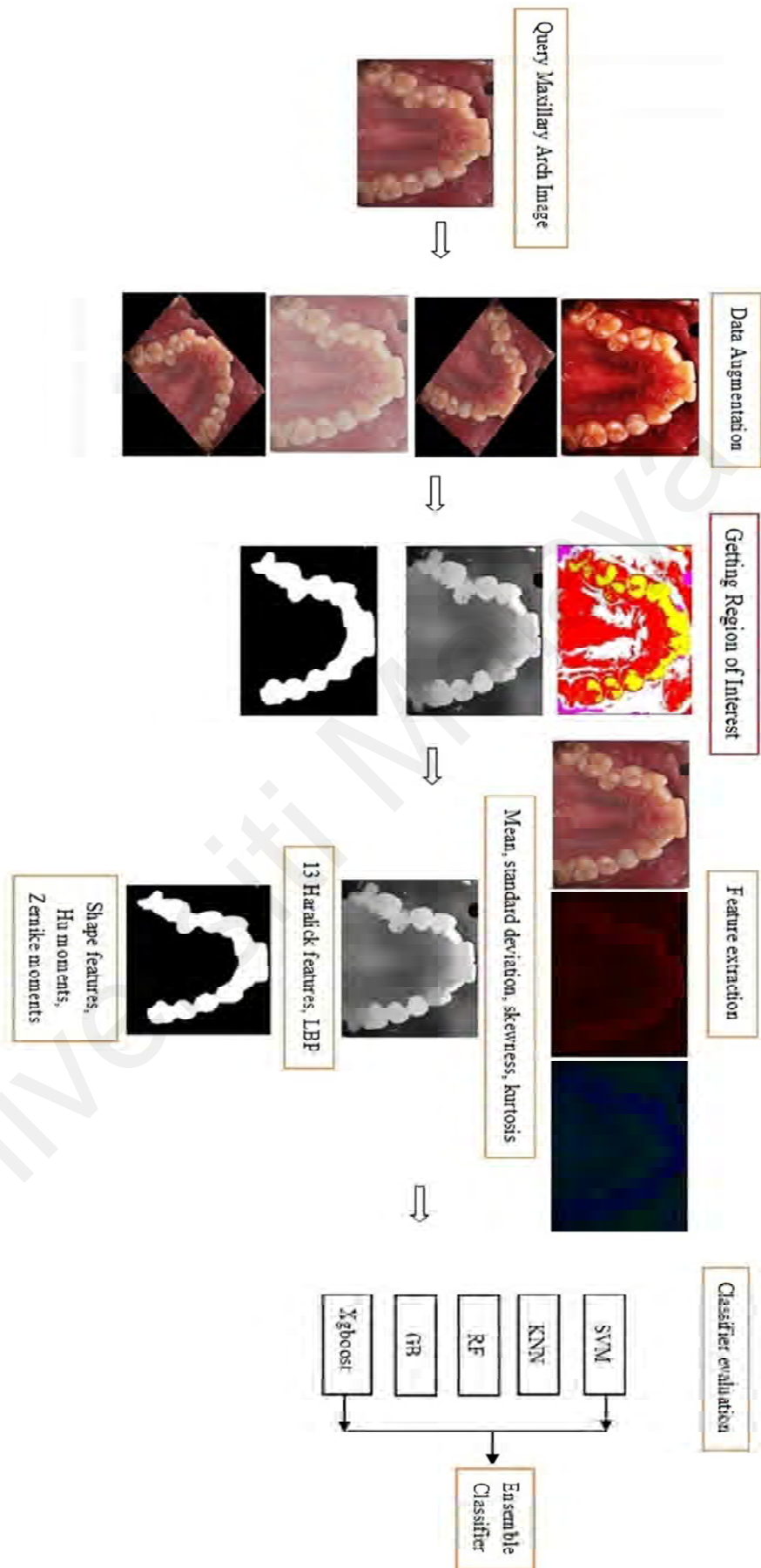


Figure 3.1: Block representation of the proposed method

3.2.1 Image Acquisition

In this study, the input data is the maxillary arch images from real-life patients. To obtain the data, ethical approval was taken by the SEGi University Oral Health Centre Medical Ethics Committee (SEGi EC/SR/FoD/2019-20/12). In the consent sheets, participants received all the necessary information related to the study and signature was taken before the data collection started.

Randomly selected male and female members were selected randomly among walk-in patients of the SEGi oral health centre. The patients' ages generally range from 16-40. Among the patients, some of them had orthodontic treatment history too. The dentist and manufacturer suggested four sizes of impression trays (Daniel Kurten, Germany) which came from four different sizes: small (S), medium (M), large (L) and extra-large (XL). Several brands of stock impression trays were compared in dimensions with numerous stock impression tray brands; but only a slight difference in dimensions were observed. Therefore, four sizes of Kurten brand trays were decided for the experiment according to a dentist's advice. Each of the four impression trays were matched to fit the maxillary image of each of the participants. On the images no modification in dimension or scale were made. Standardize image setting was checked by superimposing the raw images of the both the tray and arch with calibration reference. Following this process, 52 maxillary occlusal images were collected, using intraoral dental photography mirrors (OEM rhodium coated glass photography mirror, China) and DSLR camera kit with dental photography equipment (Nikon D5600, 85mm micro lens and R1C1 speedlight system from Nikon Japan). Among the selected cases several tooth deformations images such as crooked teeth, buccal defect, braces, and dental caries was present. Description of these defects are described below.

i) Crooked Teeth

In children and adult, having misaligned teeth, also known as crooked teeth, are very common. Several reasons are there for having a crooked tooth such as the prolonged habit of sucking thumb, jaw size, which causes the teeth to become crooked (Figure 3.2-a).

ii) Buccal Defect

The buccal defect causes to grow too many teeth's in the mouth. Sometimes these extra teeth are called supernumerary teeth (Figure 3.2-b). In the curved areas of a person's jaw, it can develop quickly at birth.

iii) Dental Braces

Dental braces, also known as orthodontic cases, cases or braces are some kind of devices used in orthodontics that align and straight teeth to help improve patients' oral structure (Figure 3.2-c).

iv) Dental Caries

Dental caries (tooth decay) damage teeth that can happen when bacteria in the mouth make some kind of acids that attack the enamel or surface which eventually leads to a small hole in the teeth (also known as cavity) (Figure 3.2-d). If left untreated, it can cause severe pain, infection, and even tooth loss as well.



Figure 3.2: Selected occlusal images for various maxillary arches: (a) crooked teeth; (b) buccal defect; (c) dental braces; (d) dental caries

Anyway, the selected 52 patients' images are split into training set (70%) and test set (30%) by random sampling method. In this way, 36 images were used for training and 16 images for testing. The images have various resolution, and the format is JPG. The link of the dataset can be found in Appendix B.

3.2.2 Image Augmentation

To create a proper model amassing large amount of data is necessary. It is obvious that building a huge data set for training is beyond the capacity of most global research agencies, especially in the scholarly area. Most of the researchers leveraged popular benchmark datasets to verify their proposed methodology. However, to the best of our knowledge there is no publicly available maxillary arch image dataset found in the literature. For this study, the training dataset consists of 36 images which may not be enough to create a proper model. Possibility lies of overfitting issue (i.e., from the limited training dataset the model will learn too much) or underfitting issue (i.e., the model cannot capture the underlying trend from the limited training data). Furthermore, to train deep-learning algorithms; for example, training CNN from scratch requires a large amount of labelled data to avoid overfitting issue (Krizhevsky et al., 2012). A possible solution is to use image data augmentation to artificially expand the size of the training data by creating modified versions of images in the dataset. This will eventually improve the generalization ability and robustness of the model (Ding et al., 2016). Random combinations of sigmoid correction, blurring, adjusted gamma, different angle rotation was used to augment the training data. It is possible that in the natural dentistry environment, different illumination may create in the images depending on the light sources. Thus, to make to model more robust to this scenario, different illumination of images was added, and rotated images are included considering the rotation of the images taken by dentists. Following (Table 3.1) the description of original images and rotated images were given.


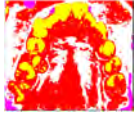



Table 3.1: Descriptions of the original images and augmented images

Tray	Original Images	Augmented Images
Tray 1 (XL)	4	40
Tray 2 (L)	8	80
Tray 3 (M)	14	140
Tray 4 (S)	10	100
Total	36	360

3.2.3 Getting Region of Interest

This stage includes to get the region of interest (ROI) from the maxillary arch image. A thorough investigation has been performed to get the ROI. (Dhivyaa & Vijayakumar, 2019) used K-means clustering algorithm for coarse segmentation which resulted better accuracy. Hence, K-means clustering is utilized for segmentation purpose. The other phases are detailed in Table 3.2. Morphological shape features from contour shows better visual representation of an image (Madian et al., 2019). The shape features such as area, perimeter, aspect ratio and so on are extracted from the final resultant binary images. Anyway, the texture properties are extracted from the greyscale images. Finally, color properties are extracted from the color images. In the case of color and texture feature extraction both the direct input image and the K-means segmented images are used (the code can be found in Appendix A).

Table 3.2: Proposed algorithm for getting ROI

<p>From the file take input color image (rgb)</p>	
<p># Perform image segmentation (K-means clustering) From the image matrix, compute necessary clusters K_i [$i=1$ to n] as long as the convergence criteria met. # $K=9$ is the final cluster Calculate the within sum of squares(wss) wss is the curve plot plot bend is the appropriate number of clusters # it is better to convert into $L*a*b$ color space for better visualization</p>	
<p>Mean shift filtering for image enhancement Initialize $m=1$ and $y_m = x_j$ Compute $y_{m+1} = 1/n_m \sum_{x_i \in S_1(y_m)} x_i$, $m \leftarrow m + 1$ till convergence Assign $x_j = x_j^s, y_{conv}^r$</p>	
<p>Greyscale (gs) conversion from the rgb image $gs = r*0.299 + g*0.587 + b*0.114$</p>	
<p>Obtain the threshold image $< I_T$ # I_T from the Otsu's method determine the best threshold By employing morphological operation dilate and erode the binary image From the binary image select the ROI</p>	

3.2.4 Feature Extraction

In this stage, several sets of discriminative and informative features are extracted. In the following subsections, the details of the feature extraction phases are described.

3.2.4.1 Color Features

The surface of the object in an image can be described by global feature. Many researchers used color features which is a global feature to achieve this purpose (Chen et al., 2019; Souaidi & El Ansari, 2019). The statistical calculations of mean (μ) and standard deviation (σ), kurtosis (γ), and skewness (θ) are extracted from each plane of

the three different color spaces as described in Chapter 2. The statistical formulas are defined as follows:

$$\mu = 1/(LM) \sum_{x=1}^L \sum_{y=1}^M P_{xy} \quad (3.1)$$

$$\sigma = \left(1/(LM) \sum_{x=1}^L \sum_{y=1}^M (P_{xy} - \mu)^2 \right)^{1/2} \quad (3.2)$$

$$\theta = \left(\sum_{x=1}^L \sum_{y=1}^M (P_{xy} - \mu)^3 \right) / (LM\sigma^3) \quad (3.3)$$

$$\gamma = \left(\left(\sum_{x=1}^L \sum_{y=1}^M (P_{xy} - \mu)^4 \right) / (LM\sigma^4) \right) - 3 \quad (3.4)$$

here, P_{xy} is the value of color on row y and column x and, L and M are the dimension of the image.

In addition, color histogram feature descriptors from the three color spaces were also considered. Color histogram is the distribution of colors in an image (Chen et al., 2019). Following a statistical calculation, it represents an estimate of an underlying continuous distribution of the color values. The color histogram may also be represented and displayed as a smooth function defined over the color space that approximates the pixel counts. Color histogram are flexible constructs that can be built from three color spaces such as RGB, HSV, L^*a^*b . A color histogram can be N -dimensional of an image and regarding the spatial location of colors the main focus is only on the proportion of the number of different types of colors. In case of two-dimensional color histogram, a two-dimensional array is followed whereas in three-dimensional color space three-dimensional array is followed. For images represented in the two-dimensional domain, the intensity histogram can be generated in continuous data as follows:

Let, $f \in L^1(\mathbb{R}^n)$ in the Lebesgue space, and the cumulative histogram H can be defined as:

$$H(f)(y) = \mu\{x: f(x) \leq y\} \quad (3.5)$$

here, the Lebesgue measure of sets is μ . $H(f)$ is the real function. The derivative can be defined as the non-cumulative histogram.

$$h(f) = H(f)' \quad (3.6)$$

3.2.4.2 Texture Features

There are many methods for textual feature extraction. However, as described in the literature review, this study is leveraging two of the popular methods: GLCM and LBP. The GLCM matrix of each image is generated and from the matrix 13 textual features such as energy, entropy, contrast, correlation, difference variance, information measures of correlation, as proposed by Haralick (Haralick et al., 1973). For simplicity, only the major set of Haralick features are described below. To learn more details about the other set one can refer to (Dhruv et al., 2019). Anyway, the physical meaning of the above-discussed texture features and the calculation formulas are described below.

i. Energy

The uniformity of the image grey level distribution is measured by energy. Near the main diagonal, the distribution elements of the GLCM are concentrated, and the level distribution in the local area is relatively uniform, and related angular second moment values are relatively large. The angular second moment value is smaller if the values of the symbiosis matrix are equal. It is also an indication that with regular changes, the current texture is a stable texture. The formula of energy defined as:

$$\text{Angular second moment} = \sum_x \sum_y \{p(x, y)\}^2 \quad (3.7)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix.

ii. Entropy

The randomness of the grey distribution and the amount of information of an image is measured by entropy, which represents the complexity of texture in any image. The more complex the texture, the higher the entropy will be. Reciprocally, smaller entropy is seen on the image having more uniform greyscale. The formula is defined by:

$$Entropy = - \sum_x \sum_y p(x,y) \ln(p(x,y)) \quad (3.8)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix.

iii. Variance

In the texture, the degree of regularity reflects by variance. A relatively smaller value of variance is observed if the texture is messy and hard to describe; whereas relatively large value is observed if the regularity is easy to describe and strong. The formula is:

$$Variance = \sum_x \sum_y (x - u)^2 p(x,y) \quad (3.9)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix. The dependence matrix mean is u .

iv. Contrast

The total amount of grey level changes in any part of an image is reflected by contrast. In an image, the contrast will be larger if the local pixel pairs' grey level difference is larger and clearer visual effect of the image is observed. Consequently, the texture's furrow depths also manifested by contrast. The deeper

the groove of the texture, the greater the contrast will be and the sharper the image. However, groove grain shallow will be less, and the image will appear fuzzier if the ratio is small. The contrast formula is given by:

$$Contrast = \sum_x \sum_y p(x, y) (x - y)^2 \quad (3.10)$$

here, the normalized grey-tone is the $(x, y)^{th}$ entry in $p(x, y)$ dependence matrix.

v. Correlation

In the direction of the row and column, the similarity degree of the grey level of an image is measured by correlation. Hence, the size of the values reflects greatly in measuring the local grey level correlation. The biggest value determines the bigger correlation value.

$$Correlation = \sum_x \sum_y p(x, y) \frac{(x - u_i)(y - u_j)}{\sigma_i \sigma_j} \quad (3.11)$$

here, the normalized grey-tone is the $(x, y)^{th}$ entry in $p(x, y)$ dependence matrix. The standard deviation of p_i, p_j are $u_i, u_j, \sigma_i, \sigma_j$.

vi. Sum Entropy

The grey level sum distribution of an image is measured by sum entropy.

$$Sum Entropy = f_8 = - \sum_{i=2}^{2N_g} P_{i+j}(x) \log \{p_{x+y}(i)\} \quad (3.12)$$

here, the normalized grey-tone is the $(x, y)^{th}$ entry in $p(x, y)$ dependence matrix.

vii. Sum Average

In an image, the mean of the grey level sum distribution is measured by sum average.

$$\text{Sum Average} = \sum_{x=2}^{2N_g} x P_{i+j}(x) \quad (3.13)$$

$$P_{x+y}(k) = \sum_{x=1}^{N_g} \sum_{y=1}^{N_g} p(x, y) \quad (3.14)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix.

viii. Sum variance

To measure the dispersion (regarding to the closest mean) of the grey level sum distribution of an image sum variance is used.

$$\text{Sum Variance} = \sum_{i=2}^{2N_g} (x - f_8)^2 P_{x+y}(x) \quad (3.15)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix. The distinct grey levels in the image is N_g .

ix. Difference Variance

It is the measure of heterogeneity which places on the differing intensity level pairs that usually deviate from the mean.

$$\text{Difference variance} = \sum_{x=2}^{2N_g} (x - f_{10})^2 P_{i-j}(x) \quad (3.16)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix. The distinct grey levels in the image is N_g .

x. Difference entropy

The disorder related grey level difference distribution of an image is measured by difference entropy.

$$\text{Difference entropy} = f_{10} = - \sum_{x=0}^{N_g-1} P_{i-j}(x) \log\{p_{i-j}(x)\} \quad (3.17)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix.

xi. Maximal Correlation Coefficient (MCC)

$$\text{Maximal Correlation Coefficient} \quad (3.18)$$

$$= \sqrt{\text{the second largest } P \text{ eigenvalue}}$$

$$P(x, y) = \sum_{k=2} \frac{q(x, k)q(y, k)}{q_x(i)q_y(k)} \quad (3.19)$$

here, the normalized grey-tone is the $(x,y)^{\text{th}}$ entry in $p(x,y)$ dependence matrix.

3.2.4.3 Local Binary Pattern

Another popular technique to describe texture information is local binary pattern (LBP) by estimating the neighborhood pixels of an image (Khan et al., 2019). Let $I_g(x, y)$ be a greyscale image with a dimension of $L \times M$ and (x, y) denotes the position of pixels in the image. Given a central pixel P_c , and the correspondence neighboring pixel P_n ; then, the LBP is calculated as follows (Ojala et al., 2002):

$$\Psi(IF)_{n,r} = \sum_{n=0}^{n-1} i(P_n - P_c)2^n \quad (3.21)$$

here, $i(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$, and n and r denote the neighboring pixel and radius of the neighborhood, respectively.

3.2.4.4 Shape Features

To describe a meaningful insight of any object's shape; contour and morphological shape features are widely used. In medical image analysis, numerous studies can be seen leveraging contour and morphological shape features (Madian et al., 2019). Image contouring is the process of identifying structural outlines of objects in an image which

in turn help to identify shape of the object. In computer vision, contours are designed using edges with an identity, geometrical perimeters and are continuous. In computer vision, an edge in an image is a sharp variation of the intensity function. In greyscale images, this applies to the intensity or brightness of pixels. In color images it can also refer to sharp variations of color. An edge is distinguished from noise by possessing long-range structure. The main properties of an edge are gradient and orientation. However, edge in themselves has no defined shape, identity, and continuity. It is also difficult to differentiate edges from one to another. To overcome this issue, authors relied on contours which is a nice property to describe the shape of an image. An illustration is given in Figure 3.3. From the contour, the image moment was identified which is the certain weighted average of the image pixel intensities.



Figure 3.3: Finding shapes from an image using contours

Let I be a binary image and the pixel intensity at location (x,y) is given by $I(x,y)$:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x,y) \quad (3.22)$$

here, M is the moment which summarizes the shape of that binary image. From the equation we calculated the following:

$$\text{Area } M_{00} = \sum_x \sum_y I(x, y) \quad (3.23)$$

$$\text{Centroid } \bar{x}\bar{y} = \frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \quad (3.24)$$

From the minimum circumscribed rectangular of the sample object, Width (W) and Height (H) is taken. From these, we calculated the following:

$$\text{Aspect ratio } \frac{W}{H} \quad (3.25)$$

$$\text{Rectangular area } W * H \quad (3.26)$$

$$\text{Equivalent diameter } \sqrt{\frac{4 * M_{00}}{\pi}} \quad (3.27)$$

In addition, the Hu invariants moment (Hu, 1962) is used. In this present study, M1-M6 (six orthogonal absolute invariants) and M7 (one skew orthogonal invariant) are used to describe the shape features of the samples. The equations for these moments are given below:

$$J_1 = \mathfrak{N}_{20} + \mathfrak{N}_{02} \quad (3.28)$$

$$J_2 = (\mathfrak{N}_{20} - \mathfrak{N}_{02})^2 + 4\mathfrak{N}_{11}^2 \quad (3.29)$$

$$J_3 = (\mathfrak{N}_{30} - 3\mathfrak{N}_{12})^2 + (3\mathfrak{N}_{21} - \mathfrak{N}_{03})^2 \quad (3.30)$$

$$J_4 = (\mathfrak{N}_{30} + \mathfrak{N}_{12})^2 + (\mathfrak{N}_{21} + \mathfrak{N}_{03})^2 \quad (3.31)$$

$$J_5 = (\mathfrak{N}_{30} - 3\mathfrak{N}_{12})(\mathfrak{N}_{30} + \mathfrak{N}_{12})[(\mathfrak{N}_{30} + \mathfrak{N}_{12})^2 - 3(\mathfrak{N}_{21} + \mathfrak{N}_{03})^2] \quad (3.32)$$

$$+ (3\mathfrak{N}_{21} - \mathfrak{N}_{03})(\mathfrak{N}_{21} + \mathfrak{N}_{03})[3(\mathfrak{N}_{30} + \mathfrak{N}_{12})^2 - (\mathfrak{N}_{21} + \mathfrak{N}_{03})^2]$$

$$J_6 = (\mathfrak{N}_{20} - \mathfrak{N}_{02})[(\mathfrak{N}_{30} + \mathfrak{N}_{12})^2 - (\mathfrak{N}_{21} + \mathfrak{N}_{03})^2] \quad (3.33)$$

$$+ 4\mathfrak{N}_{11}(\mathfrak{N}_{30} + \mathfrak{N}_{12})(\mathfrak{N}_{21} + \mathfrak{N}_{03})$$

$$\begin{aligned}
J_7 = & (3\mathfrak{N}_{21} - \mathfrak{N}_{03})(\mathfrak{N}_{30} + \mathfrak{N}_{12})[(\mathfrak{N}_{30} + \mathfrak{N}_{12})^2 - 3(\mathfrak{N}_{21} + \mathfrak{N}_{03})^2] \\
& - (\mathfrak{N}_{30} - 3\mathfrak{N}_{12})(\mathfrak{N}_{21} + \mathfrak{N}_{03})[3(\mathfrak{N}_{30} + \mathfrak{N}_{12})^2 \\
& - (\mathfrak{N}_{21} + \mathfrak{N}_{03})^2]
\end{aligned} \tag{3.34}$$

These 7 equations are well known Hu moment invariants. The first 6 moments have been proved to be invariant to translation, scale, and rotation and the 7th moment's sign changes for image reflection. A brief description of these image operations is described below.

i. Image translation

The translate operator performs a geometric transformation that maps each picture element's position in an input image into a new position in the output image (Figure 3.4-a). It is not required that the two images' dimensionality need not to be necessary the same. An image element is located at (x_1, y_1) is shifted to a new position (x_2, y_2) in the corresponding output image. A special case of translation is an affine transformation.

ii. Image rotation

In computer vision, image rotation is a common routine application for image matching alignment or other image-based algorithms (Figure 3.4-b). The rotation operation performs a geometric transform which maps the position of (x_1, y_1) of an input image onto a (x_2, y_2) position by rotating it with a degree (θ) . The degree is usually user-defined. It is also a special case of an affine transformation.

iii. Image scaling

Image scaling means resizing the digital input image (Figure 3.4-c). Nyquist sampling theorem viewpoint is mainly interpreted as a form of image resampling or reconstruction. Several algorithms can be leveraged for image scaling such as,

box sampling, nearest-neighbor interpolation, bilinear and bicubic algorithms together with deep convolutional neural network.

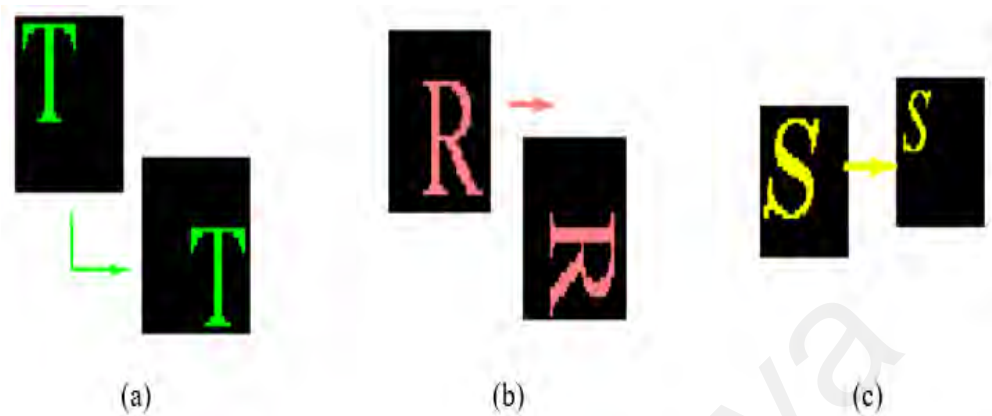


Figure 3.4: (a) Image translation; (b) Image rotation; & (c) Image scaling

3.2.5 Data Normalization

Normalization is the process of scaling individual samples to have unit form. Normalization operation is used to increase the classifier performance. It is required for any learning algorithm (in this case classification) to standardize the dataset for more suitable downstream estimators. Frequent normalization methods are used after feature extraction. However, in this study z-score normalization process is used (Turkoglu & Hanbay, 2019). The standard score of a sample x_i is calculated as:

$$z = \frac{x_i - \mu_i}{\sigma_i} \quad (3.35)$$

here, x_i represents the samples of the data. μ_i and σ_i represents the mean and standard deviation of the data samples.

3.2.6 Proposed Ensemble Classifier

Ensemble learning is a popular paradigm employed to leverage the strength of individual algorithms and mitigate their weakness. To solve a given problem an ensemble

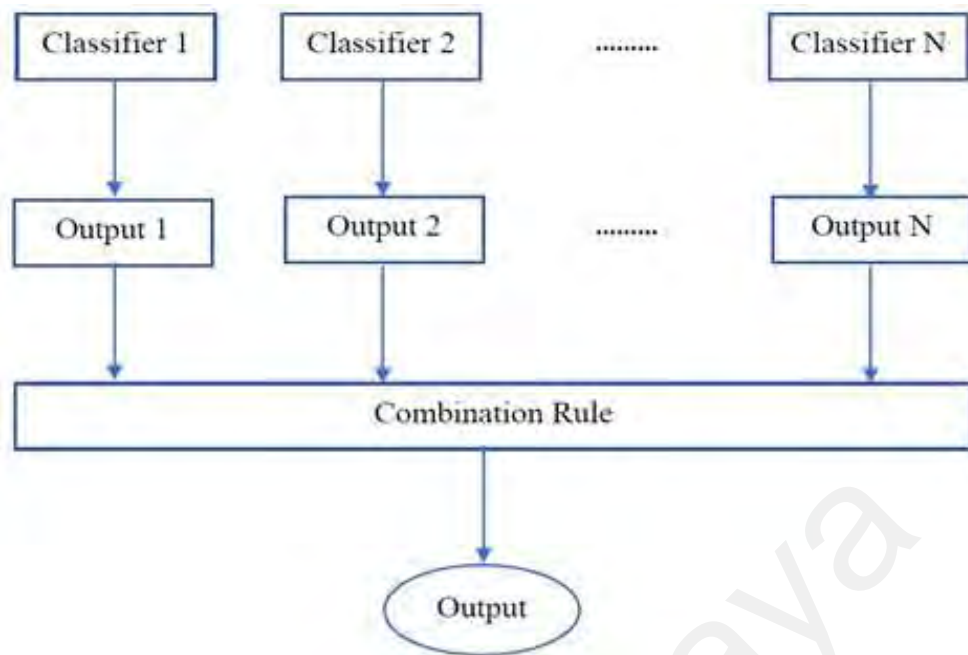


Figure 3.5: Proposed ensemble classifier

techniques combines a set of given single techniques (classification or regression) using an aggregation rule (Idri et al., 2016) as illustrated in Figure 3.5. The main objective is to achieve a high level of performance from this model that at least exceeds the performance accuracy of a single model. As majority voting is effective among other aggregation techniques (Hosni et al., 2019) thus in this study the voting classifier scheme is leveraged. There are two types of voting strategy found according to the literature (Hosni et al., 2019):

i) Hard voting

In hard voting each individual classifier votes for a class and finally the majority one is the winner. According to statistical terms, the mode of distribution of the individually predicted labels is the predicted target label of the hard voting scheme. It is the simplest case of majority voting. To illustrate let us consider 3 binary classifiers to predict either 0 or 1.

Classifier (1) → class 1 (predicted)

Classifier (2) → class 0 (predicted)

Classifier (3) → class 1 (predicted)

Final predicted class = mode {1,0,1} = 1

The final predicted class will be 1 as two of the classifiers predicted 1 as the ensemble decision.

ii) Soft voting

In case of soft voting, the prediction of the class is based on the probabilities p of each classifier. From the previous hard voting discussion, let us assume the probabilities of the classifier as below.

Classifier 1 – [0.9,0.1]

Classifier 2 – [0.8,0.2]

Classifier 3 – [0.4,0.6]

Average probabilities for each class are calculated:

$$P(0) = (0.9 + 0.8 + 0.4)/3 = 0.7$$

$$P(1) = (0.1 + 0.2 + 0.6)/3 = 0.3$$

Since class 0 shows the highest likelihood that is 0.7; thus, based on soft voting, our final class will be '0'.

In contrast with hard voting, soft voting can improve the performance because it considers more information by using each classifiers uncertainty in the final decision (Hosni et al., 2019). Therefore, in this study, soft voting is used to create an ensemble classifier.

Anyway, five machine learning classifiers have been chosen for this study, namely, support vector machine (SVM) (de Sousa Costa et al., 2018), K-nearest neighbor (KNN) (Balasubramanian et al., 2018), random forest (RF) (He et al., 2019), gradient boosting (GB) (Selvapandian & Manivannan, 2018) and eXtreme gradient boosting (XGBoost) (Tahmassebi et al., 2019). These classifiers have been chosen because of their established performance and efficiency. Each classifier has its own advantage and disadvantage; for example, SVM is useful when the dataset is bigger than the features. First, the five classifiers will be evaluated based on their accuracy. After that, the three highest accuracy provided classifiers will be chosen to reduce the memory consumption. To find the optimal parameters for a given classifier, grid search will be employed by performing hyper parameter tuning based on low mean squared error (MSE) rate. For each observation (i) and classifier (n) the MSE is calculated as follows:

$$MSE_{i,n} = \sum_{c=1}^{n_c} (A_c - P_{c,n})^2 / n_c \quad (3.36)$$

here, A_c and P_c are the actual prediction (actual class) and probability of the predicted output (i.e., belonging to each class), where n_c is the total number of classes (two in binary classification).

To create an ensemble model, this study uses the soft voting approach. In soft voting, the class labels are predicted based on the probabilities p for each classifier. Soft voting returns the class label as argmax of the sum of predicted probabilities. The equation for soft voting is given as:

$$\hat{y} = \arg \max \sum_{j=1}^m w_j p_{ij} \quad (3.37)$$

here, for the j_{th} classifier, w_j is the assigned weight.

3.3 Chapter Summary

In this chapter, the outline of the proposed research methodology is presented. It also describes the step-by-step methods for developing the proposed model, as indicated likewise. The next chapter focuses on the experimental implementation and discussion of the proposed methodology.

Universiti Malaya

CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION OF THE PROPOSED MULTI-FEATURE WITH ENSEMBLE CLASSIFIER

4.1 Introduction

In this chapter, at first, the necessary experimental setup of the proposed method is briefly described. Subsequently, the related step-by-step experimental evaluation to validate the proposed method's efficacy together with a deep-learning based algorithm is presented in detail. Furthermore, essential discussion of the proposed method's convenience, its effectiveness on a limited dataset and better performance over a deep-learning algorithm are presented broadly. At the end, the way to use this model in real-life scenarios, some drawbacks and a summary is presented at the end of this chapter.

4.2 Hardware Setup

To perform all the related experiments, Anaconda release on 64-bit Intel® Core™ with 8 GB of RAM was used. A sample of the maxillary dental arch images and impression trays is shown in Figure 4.1.

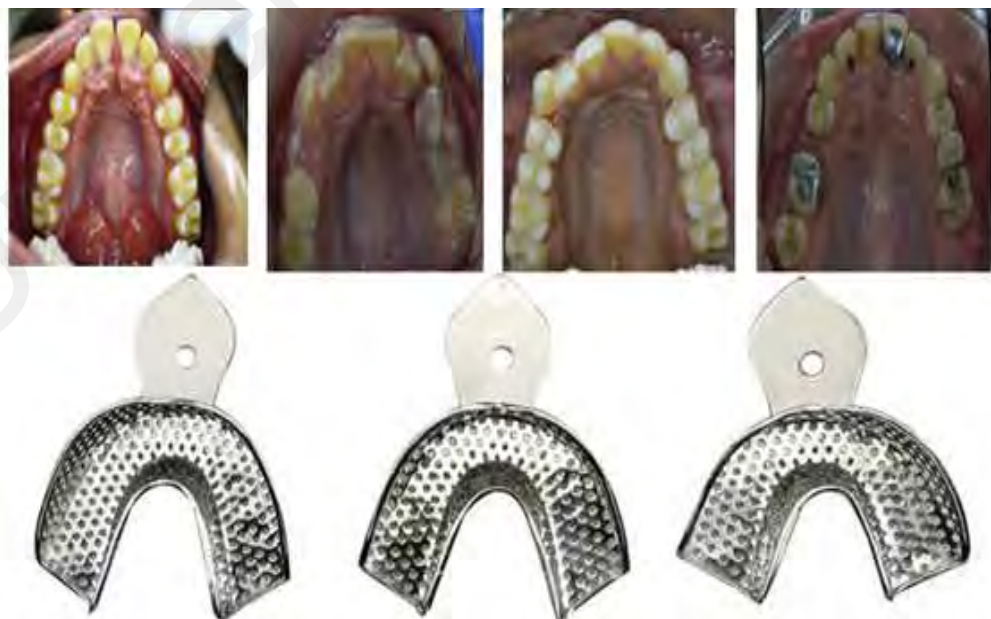


Figure 4.1: Example of maxillary arches and the trays

Generally, the dataset consists of 52 real images collected from different patients with age ranging from 16 to 40 for four different types of trays as suggested by the dentist. For the experimental purpose and for faster calculation, the resolution has been set to 300 * 300. The experimental program is set such that it will automatically convert any image to 300 * 300. Before identifying the appropriate tray from the maxillary arches, several features were extracted as described previously in Chapter 3. In detail, 72 statistical color (SC) features, 1024 color histogram (CH) features, 26 GLCM features, 52 LBP descriptor features and 47 morphological shape (MS) features were extracted. To assess performance of the proposed method, three sets of experiments were carried out. In the first experiment, the performance of different features set is compared separately and by their combination using the five classification algorithms to identify which set of features to combine and to choose from the highest accuracy provider classifiers. In the second experiment, the performance of multi-feature with ensemble classifier is assessed. Finally, a deep-learning based multilayer perceptron neural network classification performance is evaluated in the third experiment. Several statistical tests were performed to investigate the classifier performance. Brief definition of these statistical tests is given below:

i. Recall

In pattern recognition or classification task, recall defines as the number of true positives (TP) divides by the total number of elements that actually belong to the positive class. It actually means the sum of true positives and false negatives (FN) which broadly means that items should belong to the positive class but has not been labelled in this case. A perfect recall score of 1.0 means that for every item found in class B (let us assume) was labelled as belonging to class B. However, it

says nothing about items from other classes were incorrectly labelled as belong to class B.

ii. Precision

Similar to recall in pattern recognition or classification task, the precision for a class is derived from the total number of true positive (TP) case; that is, correctly labelled items belong to the positive class dividing by the total number of cases belong to the positive class (sum of true and false positives (FP) in which some items incorrectly labelled as belong to that class).

iii. Accuracy

In machine learning classification task, accuracy means one straightway to measure how often an algorithm classifies a data point correctly. Out of all data points, accuracy measure the number of correctly predicted data points. To describe in a formal way, it is defined as the number of TP and true negative (TN) divided by the number of TP, TN, FP, and FN. In a nutshell, the data point that any classification algorithm correctly classified as true is the TP, and a data point that the classification algorithm incorrectly classified as false is the FN. On the other hand, if a classification algorithm incorrectly classified a data point then it is called FP or FN e.g., if a classification algorithm classified a false data point as true, it would be FP.

iv. F1 score

The F1 score is the harmonic mean of precision and recall. A perfect precision and recall score indicated by the highest possible value of an F1 score is 1

otherwise lowest possible value of 0 refers either the precision or recall value is 0.

Sometimes, the F1-score also named as Dice similarity coefficient.

v. Jaccard similarity coefficient score

The Jaccard similarity coefficient is used to compare a set of predicted labels for a sample to the corresponding set of labels.

vi. Mathews Correlation Coefficient

The Matthews correlation coefficient (MCC) or phi coefficient is generally regarded as a balanced measure if the classes are of varying different sizes. It returns a value between -1 and +1. MCC is widely used in the field of bioinformatics and machine learning. The MCC does not depend on which class is the positive one and has an advantage over F1 score. Anyway, the equations of the above-mentioned statistical tests are given below:

$$Recall = \frac{TP}{TP + FN} \quad (4.1)$$

$$Precision = \frac{TP}{TP + FP} \quad (4.2)$$

$$Accuracy = \frac{TN + TP}{TP + FN + FP + TN} \quad (4.3)$$

$$F1 = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \quad (4.4)$$

$$JAC = \frac{TP}{TP + FN + FP} \quad (4.5)$$

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \quad (4.6)$$

Anyway, in the following step-by-step rigorous experimental evaluation for validating the efficacy of the proposed model are described in detail.

4.3 Performance of Single Set Feature and Multi-Feature

By utilizing more information, multi-feature can accomplish better identification accuracy and to verify this a comparative experiment is implemented. Since color, texture and shape are different feature descriptors, i.e., have distinct dimensions, so each set of features is evaluated separately. At first, the identification accuracy of the individual set of features are evaluated separately. To further validate the effectiveness, other sets of experiment were done on the combination of best individual features. The identification results (accuracy)% of the individual feature set and their concatenation are shown in Table 4.1.

Table 4.1: Comparative performance of single set of features and multi-feature.

Classifiers	Features				
	SC	GLCM	MS	CH	LBP
SVM	50.00	43.75	75.00	56.25	47.50
KNN (K=3)	68.75	50.00	62.50	62.50	68.75
RF	43.75	68.75	61.50	62.50	68.75
GB	56.25	56.25	56.25	56.25	56.25
Xgboost	68.75	56.25	50.00	68.75	62.50
Classifiers	Features				
	SC + GLCM + MS+ CH + LBP	SC+ MS + CH + LBP	SC+ MS + LBP	MS + LBP	MS + SC
SVM	68.75	84.25	87.50	77.50	81.25
KNN (K=3)	75.00	75.00	75.00	68.75	62.50
RF	52.50	68.75	75.00	68.75	61.50
GB	68.75	68.75	68.75	62.50	56.25
Xgboost	81.25	81.25	68.75	68.75	68.75

*Bolted data indicates the best results

In Table 4.1, the features set shows different performance of recognition under the five classification algorithms. According to the comparative identification accuracy presented in Table 4.1, MS features show the best performance on SVM classifier compared to the other single set of features. GLCM features show the worst identification accuracy compared to others. To further validate the results, five different experiments were performed on the combination of the set of features. It is observed that the combination of SC, MS and LBP achieves 87.50% on SVM classifier. An interesting finding from the experiment is that multiple set of features showed better identification accuracy almost in all cases compared to the single set of features. Thus, it is evident that the multi-feature shows better accuracy compared to the single set of features. Moreover, adding inappropriate features also decrease the classification performance. Two conclusions can be drawn from the analysis: 1) Feature set of SC, MS and LBP shows better identification accuracy than single set of feature accuracy; 2) Among the classifiers SVM, KNN and RF show better performance.

4.4 Performance Based on Ensemble Classifier

In this study a method based on multi-feature with ensemble classifier is proposed based on soft margin technique. SVM, KNN, and RF are chosen for making the final ensemble classifier. Two significant factors have great influence in ensemble classification performance: hyper parameter tuning and assigning weights. As three classification algorithms were finalized, and the algorithms have different parameters. For example, in SVM the penalty factor C and gamma (γ) have important influence in classification performance (Du et al., 2019). To overcome the first issue grid-search method was adopted, which is a classical way of finding optimal parameters of a given classification algorithm (Hsu et al., 2003). A range of predefined parameter values was evaluated using 10-fold cross validation to avoid over-fitting. The low MSE rate (Equation 3.36) of 10-fold cross-validations is used to evaluate classifier accuracy.

At the same time, an exhaustive search was performed to find the optimal weights for each classifier according to the soft voting requirement (Equation 19). The final weights were chosen [0.5, 0.2, 0.3] for SVM, KNN and RF, respectively. Furthermore, the loss function was calculated for each classification algorithm to estimate the loss of the classifier model. A common loss function used with classification is Zero-one loss. It assigns 0 to loss for a correct classification and 1 for an incorrect classification. If \hat{y}_i is the predicted value of the i -th sample and the corresponding true value is y_i , then the 0-1 loss L_{0-1} is defined as:

$$L_{0-1}(y_i, \hat{y}_i) = 1(\hat{y}_i \neq y_i) \quad (4.7)$$

here, the indication function is defined by $1(x)$. It is desired that a proper model should have a smaller Zero-one loss value. To further validate the proposed multi-feature with ensemble classifier various statistical experimental results are provided in Table 4.2.

Table 4.2: Comparative performance of multi-feature with ensemble classifier

Classifier	Combined features (SC + MS + LBP) %				
	Precision	Accuracy	Recall	F1 score	JAC
SVM	90.62	87.50	87.89	87.66	78.13
	MCC	Zero one loss	Training time(s)	Testing time (s)	
	84.04	0.125	0.19	0.03	
	Precision	Accuracy	Recall	F1 score	JAC
KNN (K=3)	79.68	75.00	73.46	73.66	67.45
	MCC	Zero one loss	Training time(s)	Testing time (s)	
	71.01	0.3125	0.11	0.02	
	Precision	Accuracy	Recall	F1 score	JAC

RF	62.20	75.00	76.52	67.59	68.85
	MCC	Zero one loss	Training time(s)	Testing time (s)	
	72.40	0.25	5.10	0.06	
GB	Precision	Accuracy	Recall	F1 score	JAC
	68.43	68.75	69.68	67.51	61.67
	MCC	Zero one loss	Training time(s)	Testing time (s)	
	68.07	0.3252	6.94	0.09	
Xgboost	Precision	Accuracy	Recall	F1 score	JAC
	69.91	68.75	67.73	68.01	63.13
	MCC	Zero one loss	Training time(s)	Testing time (s)	
	66.69	0.25	13.9	0.17	
Ensemble	Precision	Accuracy	Recall	F1 score	JAC
	92.31	91.75	91.75	90.69	88.06
	MCC	Zero one loss	Training time(s)	Testing time (s)	
	88.93	0.082	17.4	0.20	

*Bolded data indicates the best results from the proposed method

4.5 Results on Multilayer Perceptron Neural Network

In addition, deep-learning based MLP NN algorithm was employed, and its performance analyzed on the multi-feature set of (SC, MS, and LBP) of the dataset. As mentioned in the literature review (Chapter 2) that on a limited dataset deep-learning algorithms may not work better. Thus, it is necessary to evaluate with deep-learning algorithm as well to rigorously validate the efficacy of the proposed method. Anyway, the MLP NN was chosen over deep neural network (e.g., CNN) because in case of limited data MLP NN may overcome overfitting issue and showed better classification

performance over CNN (Yun et al., 2019). It is also called second-generation neural network or shallow neural network and commonly has 1 hidden layer (or ≤ 2 hidden layers). In simple terms, an MLP NN can be considered as successively connected series of layers of neurons by weights, which are iteratively adjusted through an optimization process. Back-propagation algorithm is the base for training these neural networks. A three-layer MLP NN is shown in Figure 4.2.

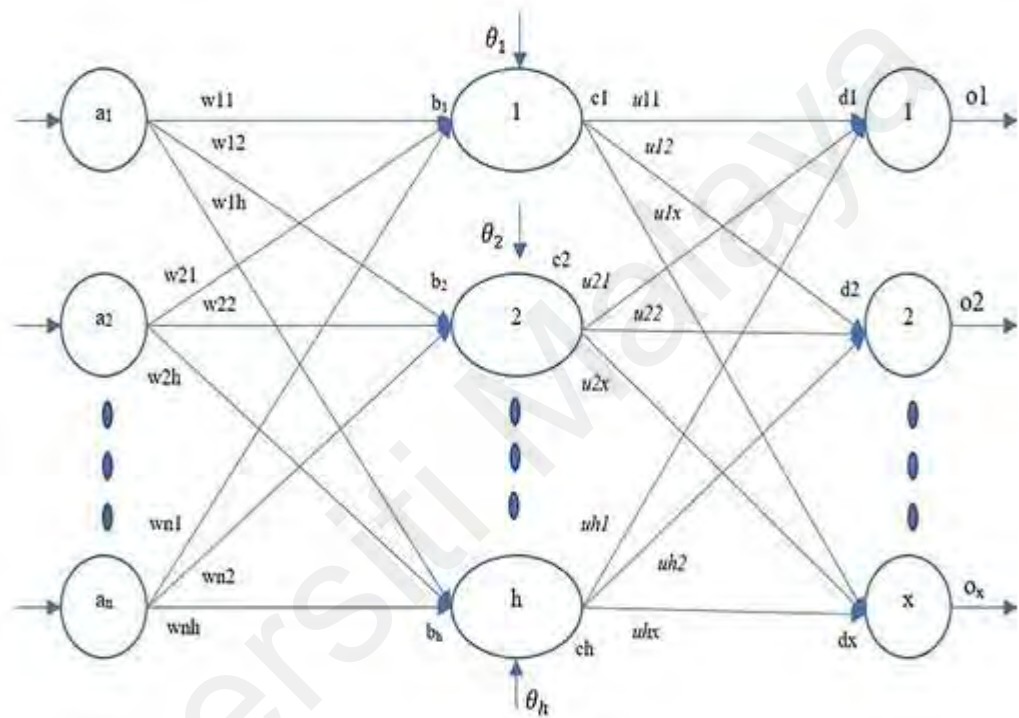


Figure 4.2: One hidden layer MLP NN

From Figure 4.2 it can be seen that one directional connection exists among the nodes which is a special type of the Feed-Forward (FF) neural network family. Here, n defined the number of input nodes, h defined the number of hidden layer and m defined the number of output nodes. The architecture of MLP NN is usually determined by trial and error method (Khishe & Mosavi, 2020). However, in this study, 39 hidden layers, and ‘lbfgs’ optimizer (which belongs to the family of quasi-Newton methods) was used.

Table 4.3: Experimental results of MLP NN on multi-feature.

Precision	Accuracy	Recall	F1 score	JAC
91.75	88.93	88.82	88.56	79.38
MCC	Zero one loss	Training time(s)	Testing time (s)	
83.42	0.113	8.59	0.13	

The MLP NN is able to achieve 88.93% accuracy based on Table 4.3, which is slightly lower than ensemble classification result. It can be inferred that the poor performance of MLP NN is due to the small dataset size. It may be possible that due to the limited samples and with the increment of hidden layer number it got traps to the poor local minimums (Schmidhuber, 2015). However, various researchers leverage different deep-learning algorithms e.g. transfer learning on limited dataset and the results showed better accuracy which outperformed the traditional machine learning classifiers (Zhao et al., 2020). Likewise, it is also possible that the specific MLP NN architecture is unsuitable for the selected dataset. Hence, a conclusive result cannot be stated based on the preceding analysis. In the next section, the overall comparative discussion of the proposed methodology is broadly presented.

4.6 Discussion

After the implementation of the proposed dental impression tray selection model with multi-feature and ensemble classifier and after a rigorous evaluation, in this section experimental discussion of the proposed approach is presented. The convenience of the proposed approach, its usefulness on a limited dataset and better performance over a deep-learning algorithm are described.

In this study, a novel method is proposed based on multi-feature with ensemble classifier to select an appropriate dental impression tray from maxillary arch image. The rigorous evaluation of multi-feature of statistical color, morphological shape, and local

binary pattern with ensemble classifier attained a maximum classification accuracy of 91.75% on this dataset. Single set of color, texture and shape features are unable to describe the arches properly due to the complex changes of lighting setup, absence of texture, and variation of the mouth shapes present in the dataset. In this scenario, the multi-feature with ensemble classifier is able to achieve more than 90% score in precision, recall, and F1 score. The dataset consists of various kinds of low resolution, rotated images but these do not influence the precision, recall, F1 score and accuracy of the proposed method. Also, the Zero-one loss score of ensemble classifier is 0.082, which is lower compared to single machine learning classification algorithms and multilayer perceptron neural network which suggests that the proposed model is worthwhile. This demonstrates that the proposed model based on multi-feature with ensemble classifier plays a significant role in discriminating between the maxillary arch images and can be used in real-time systems. The reason that reduced the classification accuracy of multilayer perceptron neural network is due to the small sample dataset utilized in this experiment and also the possibility of getting trapped into poor local minimums. Generally, deep neural networks perform better if they are fed to large numbers of labelled samples. It is difficult to gather a large number of medical ground truth data which is costly and time-consuming, and the data has to be medically validated by the practitioners as well. However, as previously discussed, using non deep-learning algorithms performs well on limited training samples and is widely reported to be accurate. In various sectors of image classification tasks, it is still arguable how well these deep-learning algorithms perform against machine learning algorithms (Jozdani et al., 2019). Anyway, in this case the ensemble classifier obtained satisfactory performance over multilayer perceptron neural network and the traditional machine learning classification algorithms. A graphical comparison between the classification accuracy is presented in Figure 4.3.

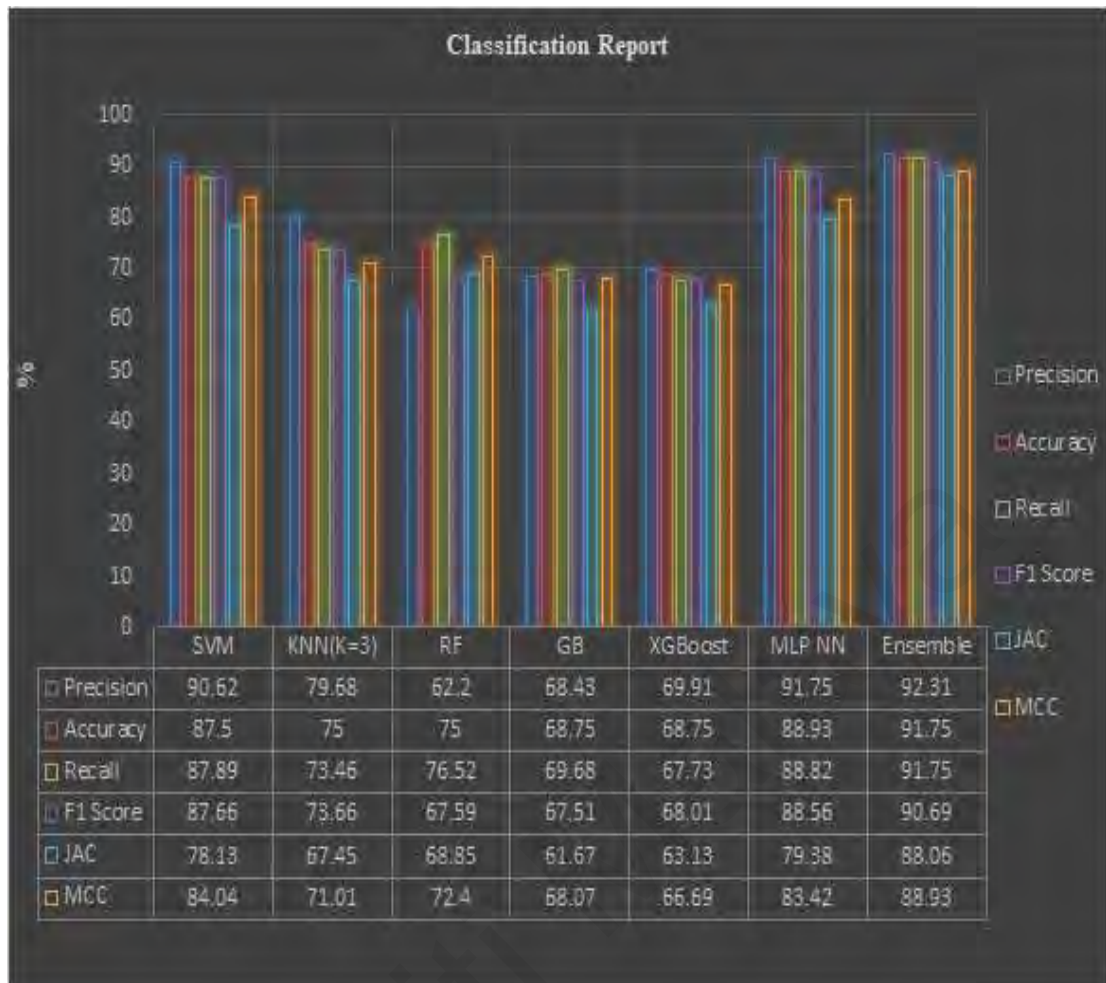


Figure 4.3: Machine learning classification report (model wise)

The efficacy of the proposed method can be validated by the experiments and gave the following conclusion: a) The identification performance of multi-feature is significantly superior compare to single set of features; b) The ensemble classifier method used in this study is more effective for identifying the proper tray compared to the single machine learning classifier.

Additionally, an illustration of the proposed dental tray selection model to use in real-life scenario is shown in Figure 4.4. To illustrate, a dentist can take the occlusal view image of the maxillary arch by an intraoral camera. The proposed model will process the image and determine which tray may best fit the arch without modifications to the tray or use trial and error method even without the need of an expert.

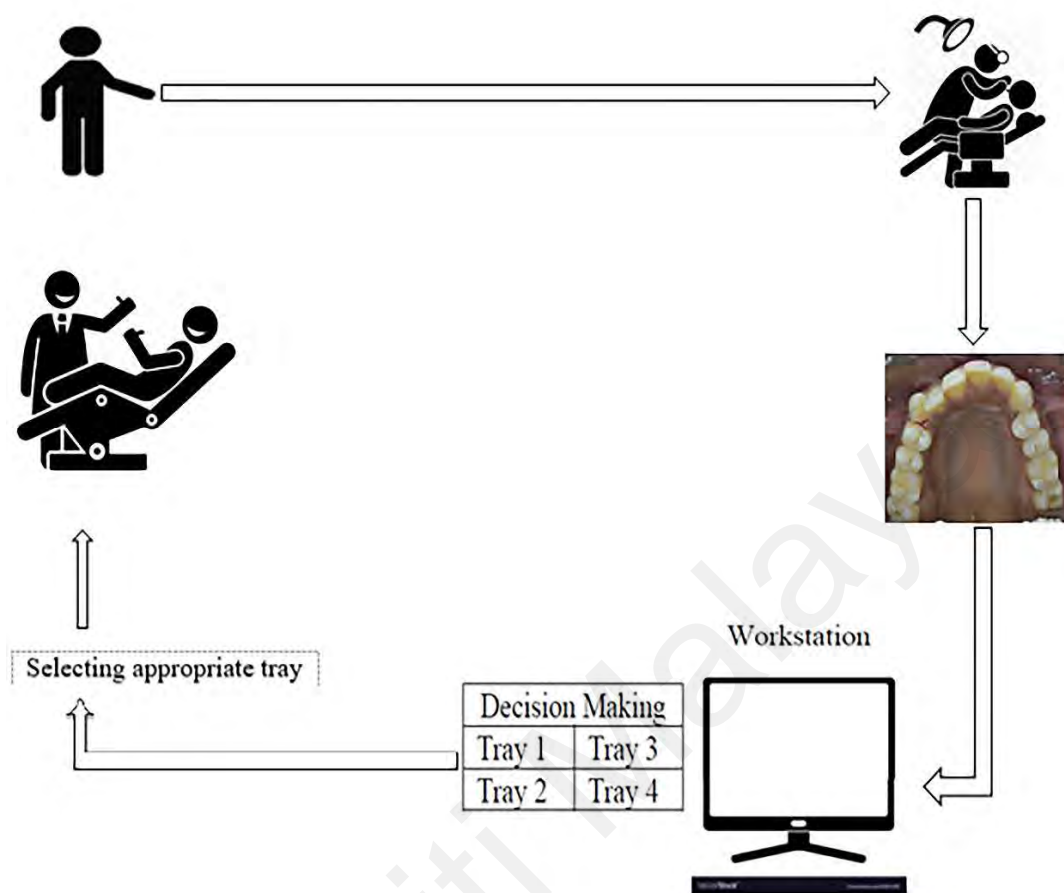


Figure 4.4: Illustration of selecting an impression tray from the proposed model in real-life scenario

It is evident that even expert human observers may oversee the proper tray due to different illumination and low resolution of pixels. However, computational algorithms have the potential to tackle these problems with less error rates, less equipment demand, and less time. This allows instantaneous integration into a clinical computer-assisted diagnostic system. This novel model with computer vision and machine learning presented in this study proved to be capable of selecting appropriate dental impression tray with a potential success rate that signifies its potential for application in clinical practice as part of an automated process chain. Therefore, the hypothesis that computer vision and machine learning can be used to select an appropriate tray from the maxillary

arch is confirmed. Thanks to the automation, it is expected to get a fully user-independent process in the future which will eventually eliminate relying on manual inspection.



Figure 4.5: Example of identification failures on maxillary arches

To have an unbiased view, it is crucial to mention the drawbacks of the proposed method. False classification has often been observed on maxillary images with crooked teeth, braces, and images having less illumination. Examples of these images are shown in Figure 4.5. By employing additional data in the training sets, these faults may be corrected. In addition, in this current study all the participants are adult patients. The data from infant patients have been omitted. The results will differ if people from other demographics of the same and infant's mouth impression images are used. Anyway, in case of adult patients, the proposed method got a satisfactory performance over a limited data which clearly establish the efficacy of this study.

4.7 Chapter Summary

This chapter in detail outlined the necessary computational setup and step-by-step experimental evaluation of the proposed multi-feature with ensemble classifier. The results, according to Table 4.2 clearly indicates that multi-feature with ensemble classifier shows promising performance which establishes the validity of the proposed method. A clear experimental discussion demonstrated that the proposed method could be used in a real-life dentistry scenario. The next chapter concludes the study.

CHAPTER 5: CONCLUSION

In the last chapter of this dissertation, the summary and related findings of this study are presented. Besides, the limitation of the proposed study and potential future work is presented as well.

5.1 Research Summary

In this study, multi-feature with ensemble classifier is proposed to select the appropriate dental impression tray from maxillary arch images. Literature review is scrutinized critically to propose the study and to demonstrate its application in dentistry applications. The first step is to find the essential computer vision-based features to describe the maxillary arch image properly with the aim to work in real-life scenarios. Then, ensemble classifier is leveraged to increase the classification performance. In a nutshell, the proposed method fused statistical color, morphological shape, and local binary pattern features with ensemble classifier to improve its ability in selecting appropriate dental impression tray. Experimental results demonstrate that on a limited dataset, the proposed method attains excellent classification results in identifying appropriate dental impression tray with precision (92.31%), recall (91.75%), F1 score (90.69%) and accuracy (91.75%), jaccard similarity score (88.06)%, matthews correlation coefficient (88.93)% and a Zero-one loss score of 0.082. Thus, the hypothesis that computer vision and machine learning are capable of selecting dental impression tray is confirmed. Therefore, the presented dental impression tray selection procedure, may be leverage in assisting inexperienced dentists and dental laboratory technicians to select the most appropriate impression tray for the Malaysian population and will automate the workflow in dentistry environment.

5.2 Limitation and Future Works

The proposed novel computer vision based dental impression tray selection model shows excellent performance in a limited dataset. However, to have an unbiased view in any study, it is necessary to identify the weakness of the proposed methodology so that other researchers can strengthen the proposed research study. Since the features extracted in this study are very large, it is obvious that large memory space will be occupied. Second, in terms of time, the ensemble classification method takes a little bit of computational time. However, this study prioritized more on classification performance compared to the computational time and space because the experimental study was performed on a CPU with 8 GB of RAM. The hardware requirement is cheap, and anyone can build this model on a personal computer. The instant result is also obtained from the proposed method compared to dental practitioners' timely manual identification. As previously mentioned, deep-learning algorithms (e.g., convolutional neural network) show promising results in medical sectors and have a large potential for further improvement. Through convolutional neural network algorithm pre-processing step is not needed; therefore, this step can be skipped to reduce computational costs. In future work, it would be interesting if other researchers extend this study by doing a deeper investigation utilizing convolutional neural network on a wide-ranging of data to address this study's efficiency in terms of time and space.

REFERENCES

- Abbas, S., Farhan, S., Fahiem, M. A., & Tauseef, H. (2019). Efficient Shape Classification using Zernike Moments and Geometrical Features on MPEG-7 Dataset. *Advances in Electrical and Computer Engineering*, 19(1), 45–51.
- Acevedo, A., Alférez, S., Merino, A., Puigví, L., & Rodellar, J. (2019). Recognition of peripheral blood cell images using convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 180, 105020. <https://doi.org/10.1016/j.cmpb.2019.105020>.
- Ahmed, K. T., Irtaza, A., & Iqbal, M. A. (2017). Fusion of local and global features for effective image extraction. *Applied Intelligence*, 47(2), 526–543.
- Ahmed, K. T., Ummesafi, S., & Iqbal, A. (2019). Content based image retrieval using image features information fusion. *Information Fusion*, 51, 76–99.
- Akkoç, B., Arslan, A., & Kök, H. (2016). Gray level co-occurrence and random forest algorithm-based gender determination with maxillary tooth plaster images. *Computers in Biology and Medicine*, 73, 102–107.
- Ali, H., Sharif, M., Yasmin, M., & Rehmani, M. H. (2020). Color-based template selection for detection of gastric abnormalities in video endoscopy. *Biomedical Signal Processing and Control*, 56, 101668.
- Attamimi, M., Araki, T., Nakamura, T., & Nagai, T. (2013). Visual recognition system for cleaning tasks by humanoid robots. *International Journal of Advanced Robotic Systems*, 10(11), 384.
- Auria, L., & Moro, R. A. (2008). Support vector machines (SVM) as a technique for solvency analysis. (August 1, 2008). DIW Berlin Discussion Paper No. 811, Available at SSRN: <https://ssrn.com/abstract=1424949>.
- Avuçlu, E., & Başçiftçi, F. (2018). New approaches to determine age and gender in image processing techniques using multilayer perceptron neural network. *Applied Soft Computing*, 70, 157–168.
- Avuçlu, E., & Başçiftçi, F. (2019). Novel approaches to determine age and gender from dental x-ray images by using multiplayer perceptron neural networks and image processing techniques. *Chaos, Solitons & Fractals*, 120, 127–138.
- Bagri, N., & Johari, P. K. (2015). A comparative study on feature extraction using texture and shape for content based image retrieval. *International Journal of Advanced Science and Technology*, 80(4), 41–52.

- Balasubramanian, K., Ananthamoorthy, N. P., & Gayathridevi, K. (2018). Automatic Diagnosis and Classification of Glaucoma Using Hybrid Features and k-Nearest Neighbor. *Journal of Medical Imaging and Health Informatics*, 8(8), 1598–1606.
- Barata, C., Ruela, M., Francisco, M., Mendonça, T., & Marques, J. S. (2013). Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Systems Journal*, 8(3), 965–979.
- Belgiu M, Drăguț L (2016) Random forest in remote sensing: A review of applications and future directions. *ISPRS J Photogramm Remote Sens* 114:24–31.
- Blatz, M. B., & Conejo, J. (2019). The Current State of Chairside Digital Dentistry and Materials. *Dental Clinics of North America*, 63(2), 175–197. <https://doi.org/10.1016/j.cden.2018.11.002>
- Blockeel, H. (2011). Hypothesis space. *Encyclopedia of Machine Learning*, 1, 511–513.
- Bober, M. (2001). MPEG-7 visual shape descriptors. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6), 716–719.
- Borges, T. A., & Neves, R. F. (2020). Ensemble of machine learning algorithms for cryptocurrency investment with different data resampling methods. *Applied Soft Computing*, 106187.
- Bozkurt, M. H., & Karagol, S. (2020). Jaw and Teeth Segmentation on the Panoramic X-Ray Images for Dental Human Identification. *Journal of Digital Imaging*, 1–18.
- Bryll, R., Gutierrez-Osuna, R., & Quek, F. (2003). Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets. *Pattern Recognition*, 36(6), 1291–1302.
- Cervantes-Sanchez, F., Cruz-Aceves, I., Hernandez-Aguirre, A., Hernandez-Gonzalez, M. A., & Solorio-Meza, S. E. (2019). Automatic Segmentation of Coronary Arteries in X-ray Angiograms using Multiscale Analysis and Artificial Neural Networks. *Applied Sciences*, 9(24), 5507.
- Chatterjee, S., Dey, D., & Munshi, S. (2019). Integration of morphological preprocessing and fractal based feature extraction with recursive feature elimination for skin lesion types classification. *Computer Methods and Programs in Biomedicine*, 178, 201–218.
- Chen, H., Yang, L., Li, L., Li, M., & Chen, Z. (2019). An efficient cervical disease diagnosis approach using segmented images and cytology reporting. *Cognitive Systems Research*, 58, 265–277.

- Christensen, G. J. (2008). Will digital impressions eliminate the current problems with conventional impressions. *J Am Dent Assoc*, 139(6), 761–763.
- Conseil, S., Bourennane, S., & Martin, L. (2007). Comparison of Fourier descriptors and Hu moments for hand posture recognition. In *2007 15th European Signal Processing Conference* (pp. 1960–1964). IEEE.
- de Sousa Costa, R. W., da Silva, G. L. F., de Carvalho Filho, A. O., Silva, A. C., de Paiva, A. C., & Gattass, M. (2018). Classification of malignant and benign lung nodules using taxonomic diversity index and phylogenetic distance. *Medical & Biological Engineering & Computing*, 56(11), 2125–2136.
- Dempere-Marco, L., Hu, X.-P., MacDonald, S. L. S., Ellis, S. M., Hansell, D. M., & Yang, G.-Z. (2002). The use of visual search for knowledge gathering in image decision support. *IEEE Transactions on Medical Imaging*, 21(7), 741–754.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248–255). Ieee.
- Dhivyaa, C. R., & Vijayakumar, M. (2019). An effective detection mechanism for localizing macular region and grading maculopathy. *Journal of Medical Systems*, 43(3), 53.
- Dhruv, B., Mittal, N., & Modi, M. (2019). Study of Haralick's and GLCM texture analysis on 3D medical images. *International Journal of Neuroscience*, 129(4), 350–362.
- Dimitropoulos, K., Barmpoutis, P., Zioga, C., Kamas, A., Patsiaoura, K., & Grammalidis, N. (2017). Grading of invasive breast carcinoma through Grassmannian VLAD encoding. *PloS One*, 12(9), e0185110.
- Ding, J., Chen, B., Liu, H., & Huang, M. (2016). Convolutional neural network with data augmentation for SAR target recognition. *IEEE Geoscience and Remote Sensing Letters*, 13(3), 364–368.
- Domingo, C., & Watanabe, O. (2000). MadaBoost: A modification of AdaBoost. In *COLT* (pp. 180–189). Citeseer.
- Du, Y., Liu, J., Song, W., He, Q., & Huang, D. (2019). Ocean Eddy Recognition in SAR Images With Adaptive Weighted Feature Fusion. *IEEE Access*, 7, 152023–152033.
- Duneja, A., & Puyalnithi, T. (2017). Enhancing classification accuracy of k-nearest neighbours algorithm using gain ratio. *Int. Res. J. Eng. Technol*, 4, 1385–1388.

- Garcia-Lamont, F., Cervantes, J., López, A., & Rodriguez, L. (2018). Segmentation of images by color features: A survey. *Neurocomputing*, 292, 1–27.
- Gonzales, R. C., & Woods, R. E. (2002). Digital image processing. Prentice hall New Jersey.
- Gupta, S., Narayan, A. I., & Balakrishnan, D. (2017). In Vitro comparative evaluation of different types of impression trays and impression materials on the accuracy of open tray implant impressions: a pilot study. *International Journal of Dentistry*, 2017.
- Hadid, A., Ylioinas, J., Bengherabi, M., Ghahramani, M., & Taleb-Ahmed, A. (2015). Gender and texture classification: A comparative analysis using 13 variants of local binary patterns. *Pattern Recognition Letters*, 68, 231–238.
- Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, (6), 610–621.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- He, B., Ji, T., Zhang, H., Zhu, Y., Shu, R., Zhao, W., & Wang, K. (2019). MRI-based radiomics signature for tumor grading of rectal carcinoma using random forest model. *Journal of Cellular Physiology*, 234(11), 20501–20509.
- Heikkilä, M., Pietikäinen, M., & Schmid, C. (2009). Description of interest regions with local binary patterns. *Pattern Recognition*, 42(3), 425–436.
- Hosni, M., Abnane, I., Idri, A., de Gea, J. M. C., & Alemán, J. L. F. (2019). Reviewing ensemble classification methods in breast cancer. *Computer Methods and Programs in Biomedicine*.
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification. Taipei.
- Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2), 179–187.
- Idri, A., Hosni, M., & Abran, A. (2016). Systematic Mapping Study of Ensemble Effort Estimation. In *ENASE* (pp. 132–139).
- Ismail, A., Idris, M. Y. I., Ayub, M. N., & Yee, L. (2018). Vision-based apple classification for smart manufacturing. *Sensors*, 18(12), 4353.

- Ismail, A., Idris, M. Y. I., Ayub, M. N., & Yee, L. (2019). Investigation of fusion features for apple classification in smart manufacturing. *Symmetry*, *11*(10), 1194.
- Jozdani, S. E., Johnson, B. A., & Chen, D. (2019). Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. *Remote Sensing*, *11*(14), 1713.
- Kas, M., El-merabet, Y., Ruichek, Y., & Messoussi, R. (2020). A comprehensive comparative study of handcrafted methods for face recognition LBP-like and non LBP operators. *Multimedia Tools and Applications*, *79*(1), 375–413.
- Khan, M. A., Lali, M. I. U., Sharif, M., Javed, K., Aurangzeb, K., Haider, S. I., ... Akram, T. (2019). An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection. *IEEE Access*, *7*, 46261–46277.
- Khishe, M., & Mosavi, M. R. (2020). Classification of underwater acoustical dataset using neural network trained by Chimp Optimization Algorithm. *Applied Acoustics*, *157*, 107005.
- Kotoulas, L., & Andreadis, I. (2005). Image analysis using moments. In *5th Int. Conf. on Technology and Automation, Thessaloniki, Greece* (Vol. 360364).
- Kowal, M., Filipczuk, P., Obuchowicz, A., Korbicz, J., & Monczak, R. (2013). Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. *Computers in Biology and Medicine*, *43*(10), 1563–1572.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).
- Kuncheva, L. I., & Whitaker, C. J. (2003). Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Machine Learning*, *51*(2), 181–207.
- Kushwaha A, Singh A, Shrivastav SK (2018) Signature classification using image moments. In: *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications*. Springer, pp 235–244.
- Lazada.com.my. (n.d.). https://www.lazada.com.my/products/intraoral-dental-camera-endoscope-720p-hd-oral-dental-camera-adjustable-6-led-light-real-time-inspect-camera-usb-2mp-tooth-camera-for-pc-laptop-i2269871375-s9639580007.html?spm=a2o4k.searchlist.list.6.5d3d6bf4iZNPd7&search=1&free_shipping=1.

- Li, Y., Wang, S., Tian, Q., & Ding, X. (2015). Feature representation for statistical-learning-based object detection: A review. *Pattern Recognition*, 48(11), 3542–3559.
- Lin, S., Crotty, K. M., & Vazquez, N. (2010, February 23). Shape feature extraction and classification. Google Patents.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciampi, F., Ghafoorian, M., ... Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.
- Liu, X., Zhao, D., Jia, W., Ji, W., & Sun, Y. (2019). A detection method for apple fruits based on color and shape features. *IEEE Access*, 7, 67923–67933.
- Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823–870.
- Made. (n.d.). *brand NEW Gladent Dental CAD Cam systems with high quality*. <https://gladent.en.made-in-china.com/product/nSjJMqulkZVd/China-Brand-New-Gladent-Dental-CAD-Cam-Systems-with-High-Quality.html>.
- Madian, N., Jayanthi, K. B., Somasundaram, D., & Suresh, S. (2019). Identifying Centromere Position of Human Chromosome Images using Contour and Shape based Analysis. *Measurement*, 144, 243–259.
- Mannil, M., Eberhard, M., von Spiczak, J., Heindel, W., Alkadhi, H., & Baessler, B. (2020). Artificial Intelligence and Texture Analysis in Cardiac Imaging. *Current Cardiology Reports*, 22(11), 1–9.
- Mundim, M. B. V, Dias, D. R., Costa, R. M., Leles, C. R., Azevedo-Marques, P. M., & Ribeiro-Rotta, R. F. (2016). Intraoral radiographs texture analysis for dental implant planning. *Computer Methods and Programs in Biomedicine*, 136, 89–96.
- Murugappan, V., & Sabeenian, R. S. (2019). Texture based medical image classification by using multi-scale gabor rotation-invariant local binary pattern (MGRLBP). *Cluster Computing*, 22(5), 10979–10992.
- Nanni, L., Lumini, A., & Brahmam, S. (2012). Survey on LBP based texture descriptors for image classification. *Expert Systems with Applications*, 39(3), 3634–3641.
- Naranjo-Torres J, Mora M, Hernández-García R, et al (2020) A review of convolutional neural network applied to fruit image processing. *Appl Sci* 10:3443.

- Ojala, T., Pietikainen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proceedings of 12th International Conference on Pattern Recognition* (Vol. 1, pp. 582–585). IEEE.
- Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1), 51–59.
- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987.
- Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research*, 11, 169–198.
- Otiniano-Rodríguez, K. C., Cámara-Chávez, G., & Menotti, D. (2012). Hu and Zernike moments for sign language recognition. In *Proceedings of international conference on image processing, computer vision, and pattern recognition* (pp. 1–5).
- Özkan, K., Ergin, S., Işık, Ş., & Işıklı, İ. (2015). A new classification scheme of plastic wastes based upon recycling labels. *Waste Management*, 35, 29–35.
- Panov, P., & Džeroski, S. (2007). Combining bagging and random subspaces to create better ensembles. In *International Symposium on Intelligent Data Analysis* (pp. 118–129). Springer.
- Park, S. W., & Kwon, J. (2019). Sphere generative adversarial network based on geometric moment matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4292–4301).
- Pisner DA, Schnyer DM (2020) Support vector machine. In: Machine Learning. Elsevier, pp 101–121. Academic Press.
- Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, 6(3), 21–45.
- Qiao, X., Bao, J., Zhang, H., Wan, F., & Li, D. (2019a). Underwater sea cucumber identification based on principal component analysis and support vector machine. *Measurement*, 133, 444–455. <https://doi.org/10.1016/j.measurement.2018.10.039>.
- Raith, S., Vogel, E. P., Anees, N., Keul, C., Güth, J. F., Edelhoff, D., & Fischer, H. (2017). Artificial Neural Networks as a powerful numerical tool to classify specific features of a tooth based on 3D scan data. *Computers in Biology and Medicine*, 80(September 2016), 65–76. <https://doi.org/10.1016/j.combiomed.2016.11.013>

- Ramli, R., Idris, M. Y. I., Hasikin, K., Karim, A., Khairiah, N., Abdul Wahab, A. W., ... Arof, H. (2017). Feature-based retinal image registration using D-Saddle feature. *Journal of Healthcare Engineering*, 2017.
- Rijal, O. M., Abdullah, N. A., Isa, Z. M., Davaei, F. A., Noor, N. M., & Tawfiq, O. F. (2011). A novel shape representation of the dental arch and its applications in some dentistry problems. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5092–5095). IEEE.
- Rijal, O. M., Abdullah, N. A., Isa, Z. M., Noor, N. M., & Tawfiq, O. F. (2012). A probability distribution of shape for the dental maxillary arch using digital images. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5420–5423). IEEE.
- Rizon, M., Haniza, Y., Puteh, S., Yeon, A., Shakaff, M., Abdul Rahman, S., ... Karthigayan, M. (2006). Object detection using geometric invariant moment.
- Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1–2), 1–39.
- Saba, T., Khan, M. A., Rehman, A., & Marie-Sainte, S. L. (2019). Region extraction and classification of skin cancer: A heterogeneous framework of deep CNN features fusion and reduction. *Journal of Medical Systems*, 43(9), 289.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117.
- Selvapandian, A., & Manivannan, K. (2018). Performance analysis of meningioma brain tumor classifications based on gradient boosting classifier. *International Journal of Imaging Systems and Technology*, 28(4), 295–301.
- Shapi'i, A., Sulaiman, R., Hasan, M. K., Kassim, A. Y. M., & Hamid, H. A. (2011). Applications of computer aided design (CAD) in medical image technology. *International Journal on Advanced Science, Engineering and Information Technology*, 1(6), 698–701.
- Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221–248.
- Shirazi, S. H., Khan, N., Umar, A., Naz, M. R., & AlHaqbani, B. (2016). Content-based image retrieval using texture color shape and region. *International Journal of Advanced Computer Science and Applications*, 7(1), 418–426.
- Shopee Malaysia. (n.d.). <https://shopee.com.my/Brand-new-Nvidia-Titan-XP-for-sale->

- Smitha, M., Nisa, A. K., & Archana, K. (2018). Diabetic retinopathy detection in fundus image using cross sectional profiles and ann. In *Computational Vision and Bio Inspired Computing* (pp. 982–993). Springer.
- Sollich, P., & Krogh, A. (1996). Learning with ensembles: How overfitting can be useful. In *Advances in neural information processing systems* (pp. 190–196).
- Souaidi, M., & El Ansari, M. (2019). Multi-scale analysis of ulcer disease detection from WCE images. *IET Image Processing*, 13(12), 2233–2244.
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2015). A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7), 1455–1462.
- Tahmasebi, A., Wengert, G. J., Helbich, T. H., Bago-Horvath, Z., Alaei, S., Bartsch, R., ... Kapetas, P. (2019). Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Investigative Radiology*, 54(2), 110–117.
- Tajeddin, N. Z., & Asl, B. M. (2018). Melanoma recognition in dermoscopy images using lesion's peripheral region information. *Computer Methods and Programs in Biomedicine*, 163, 143–153.
- Talavera-Martínez, L., Bibiloni, P., & González-Hidalgo, M. (2019). Computational Texture Features of Dermoscopic Images and Their Link to the Descriptive Terminology-A Survey. *Computer Methods and Programs in Biomedicine*, 105049.
- Teague, M. R. (1980). Image analysis via the general theory of moments. *JOSA*, 70(8), 920–930.
- Turkoglu, M., & Hanbay, D. (2019). Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts. *Applied Mathematics and Computation*, 352, 1–14.
- Tyas, D. A., Hartati, S., Harjoko, A., & Ratnaningsih, T. (2020). Morphological, Texture, and Color Feature Analysis for Erythrocyte Classification in Thalassemia Cases. *IEEE Access*, 8, 69849–69860.
- Wang, C.-W., Huang, C.-T., Lee, J.-H., Li, C.-H., Chang, S.-W., Siao, M.-J., ... Ronneberger, O. (2016). A benchmark for comparison of dental radiography analysis algorithms. *Medical Image Analysis*, 31, 63–76.

- Wang, Y., Wang, D., Geng, N., Wang, Y., Yin, Y., & Jin, Y. (2019). Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection. *Applied Soft Computing*, 77, 188–204.
- Xu, D., & Li, H. (2008). Geometric moment invariants. *Pattern Recognition*, 41(1), 240–249.
- Yergin, E., Ozturk, C., & Sermet, B. (2001). Image processing techniques for assessment of dental trays. In *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 3, pp. 2571–2573). IEEE.
- Yilmaz, E., Kayikcioglu, T., & Kayipmaz, S. (2017). Computer-aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography. *Computer Methods and Programs in Biomedicine*, 146, 91–100.
- Yong-lian, L. (2020). Multi-feature data mining for CT image recognition. *Concurrency and Computation: Practice and Experience*, 32(1), e4885.
- Yun, J., Park, J. E., Lee, H., Ham, S., Kim, N., & Kim, H. S. (2019). Radiomic features and multilayer perceptron network classifier: a robust MRI classification strategy for distinguishing glioblastoma from primary central nervous system lymphoma. *Scientific Reports*, 9(1), 1–10.
- Zhang, C., Huang, C., Li, H., Liu, Q., Li, J., Bridhikitti, A., & Liu, G. (2020). Effect of Textural Features in Remote Sensed Data on Rubber Plantation Extraction at Different Levels of Spatial Resolution. *Forests*, 11(4), 399.
- Zhang Z, Mayer G, Dauvilliers Y, et al (2018) Exploring the clinical features of narcolepsy type 1 versus narcolepsy type 2 from European Narcolepsy Network database with machine learning. *Sci Rep* 8:1–11.
- Zhao, X., Liang, Y., Guo, A. J. X., & Zhu, F. (2020). Classification of small-scale hyperspectral images with multi-source deep transfer learning. *Remote Sensing Letters*, 11(4), 303–312.
- Zheng, T., Gao, Y., Wang, F., Fan, C., Fu, X., Li, M., ... Ma, H. (2019). Detection of medical text semantic similarity based on convolutional neural network. *BMC Medical Informatics and Decision Making*, 19(1), 156.