

DEVELOPMENT OF FACIAL EXPRESSION DETECTION MODEL FOR STROKE PATIENTS



A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Computer Science and Information Technology

Department of Mathematics and Computer Science

FACULTY OF SCIENCE

Chulalongkorn University

Academic Year 2019

Copyright of Chulalongkorn University

การพัฒนาตัวแบบการตรวจจับการแสดงสีหน้าสำหรับผู้ป่วยโรคหลอดเลือดสมอง



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ภาควิชาคณิตศาสตร์และวิทยาการ

คอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2562

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

5773102423 : MAJOR COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

KEYWORD: Facial expression Decision tree Stroke patients

Rawinan Praditsangthong : DEVELOPMENT OF FACIAL EXPRESSION
DETECTION MODEL FOR STROKE PATIENTS. Advisor: Assoc. Prof.
PATTARASINEE BHATTARAKOSOL, Ph.D.

A stroke patient should be cared and treated closely since the patient cannot speak, communicate, and move the body when needed. In addition, the number of stroke patients is increasing in Thailand and around the world. Unfortunately, the number of medical staffs does not vary by the number of stroke patient. Thus, the aim of this research is to develop a facial expression detection model for stroke patients during their treatments. This research proposes the facial expression detection model for stroke patients from facial features, such as Interpalpebral Fissure (IPF), Palpebral Fissure Length (PFL), Palpebral Fissure Region (PFR), Inner Brow Raisers (IBR), Brow Lower (BL), Inner and Outer Lid Raiser (LR), and Lip Part (LP). These features are applied to develop the facial expression detection model using the decision tree algorithm. Furthermore, there are two factors that can determine the facial expression detection, which are gender and age. From the experiment, the proposed the facial expression detection model can identify normal situations and abnormal situations for stroke patients with 95% accuracy; the value of precision is 91% and the value of recall is 100%. This model can be assisted for raising patient's safety.

Field of Study: Computer Science and
Information Technology

Student's Signature

Academic Year: 2019

Advisor's Signature

ACKNOWLEDGEMENTS

First of all, I would like to give my gratitude to my advisor, Associate Professor Dr. Pattarasinee Bhattarakosol, for guidance and support research funding (ASAHI GLASS FOUNDATION) in this research, including the patent of this research. All of her advices help me go through all difficulties and make it possible.

I would like to thank Assistant Professor Aurauma Chutinet and Wasan Akarathanawat from Chulalongkorn Comprehensive Stroke Center of Excellence, King Chulalongkorn Memorial Hospital for include and exclude participants for data collection.

I would like to thank Associate Professor Dr. Nagul Cooharajanone, Associate Professor Dr. Panjai Tantatsanawong, Associate Professor Dr. Chatchawit Aporntewan, and Associate Professor Dr. Kanitha Patarakul, chairman and committee of my dissertation, for assistance.

I would like to thank Rangsit University for the scholarships since August 10, 2014 – August 11, 2019. Moreover, I would like to thank my bachelor's advisor, Assistant professor Weerawat Liemmanee and Mr. Ekapong Nopawong, for guarantee scholarships from Rangsit University.

I would like to give my sincere thanks to Miss Tipaporn Juengchareonpoon for her encouragement and being my accompany during data collection in the night shift.

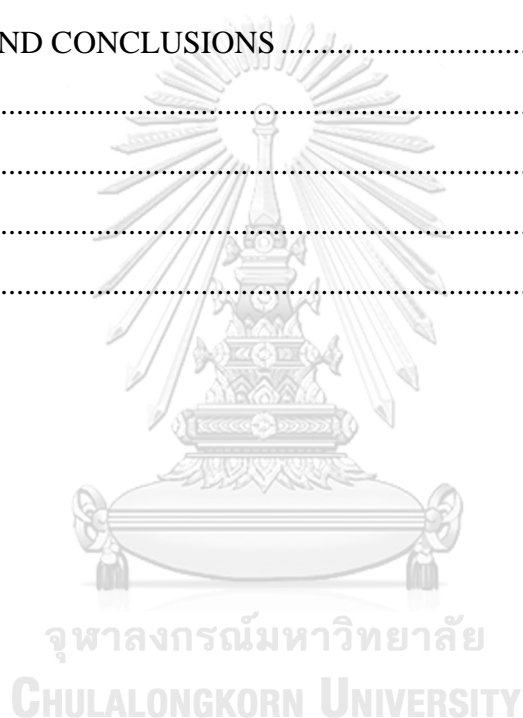
Finally, I would like to express my gratitude to my beloved parents for their supportive whenever I felt down.

Rawinan Praditsangthong

TABLE OF CONTENTS

	Page
.....	iii
ABSTRACT (THAI)	iii
.....	iv
ABSTRACT (ENGLISH).....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS	vi
LIST OF TABLES.....	viii
LIST OF FIGURES.....	x
INTRODUCTION	1
1.1 Rationale.....	1
1.2 Research Objectives.....	4
1.3 Research Scopes	4
1.4 Benefits	4
1.5 Structure of this Thesis	5
CHAPTER II.....	6
RELATED WORKS.....	6
2.1 Stroke Patient.....	6
2.2 Facial Action Coding System (FACs)	8
2.3 Machine Learning	11
2.4 Tools and Libraries	26
2.5 Techniques	28
CHAPTER III	31
METHODOLOGY AND ANALYSIS	31
3.1 Research Procedure	31
3.2 Material.....	32
3.3 Preliminary Study	34

3.4 Experiment.....	39
3.5 Analysis Method.....	40
CHAPTER IV	42
RESULT	42
4.1 Parameter Selection: First Phase	42
4.2 Parameter Selection: Second Phase	53
4.3 Parameter Selection: Third Phase	65
CHAPTER V	84
DISCUSSIONS AND CONCLUSIONS	84
5.1 Discussions	84
5.2 Conclusions.....	85
REFERENCES	87
VITA.....	94



LIST OF TABLES

	Page
Table 1 Key FAST of Stroke Description.....	7
Table 2 Key FASTER of Stroke Description	7
Table 3 Features in 68 Facial landmarks	11
Table 4 Attributes of Training Data.....	12
Table 5 Customer Behavior Dataset	20
Table 6 Clustered dataset	21
Table 7 New centroid of clustered dataset.....	22
Table 8 Show the smallest value.....	23
Table 9 Combine rows and columns for Cluster 1	24
Table 10 Combine rows and columns for Cluster 2	24
Table 11 Combine rows and columns for Cluster 3	25
Table 12 Combine rows and columns for Cluster 4	25
Table 13 Comparison.....	29
Table 14 Structure and Sub-Structure of the Face	37
Table 15 The IPF means with standard deviations (millimeters)	42
Table 16 The PFL means with standard deviations (millimeters).....	43
Table 17 The PFR means with standard deviations (millimeters).....	43
Table 18 Values of variables in this experiment.....	44
Table 19 Confusion Matrix of the Emotion Classification Decision Tree	51
Table 20 Comparison of the results between Decision Tree and Random Forest Model	52
Table 21 The average with standard deviation of the IBR (millimeters)	54
Table 22 The average with standard deviation of the BL (millimeters).....	55
Table 23 The average with standard deviation of the Inner Lid Raiser (LR) (millimeters).....	56
Table 24 The average with standard deviation of the Outer Lid Raiser (LR) (millimeters).....	56

Table 25 The Distance Link's Lip Part (LP) average with standard deviation (millimeters).....	57
Table 26 The Values of All Parameters	58
Table 27 Confusion Matrix of the Emotion Classification Decision Tree	64
Table 28 Comparison of the results between Decision Tree and Random Forest Model	65
Table 29 The IPF means with standard deviations (Millimeters).....	67
Table 30 The PFL means with standard deviations (Millimeters).....	68
Table 31 The PFR means with standard deviations (Millimeters)	69
Table 32 The IBR means with standard deviations (Millimeters).....	69
Table 33 The average with standard deviation of the BL (millimeters)	70
Table 34 The average with standard deviation of the Inner Lid Raiser (LR) (millimeters).....	70
Table 35 The average with standard deviation of the Outer Lid Raiser (LR) (millimeters).....	71
Table 36 The Distance Link's Lip Part (LP) average with standard deviation (millimeters).....	71
Table 37 The Values of All Parameters	72
Table 38 Confusion Matrix of the Status Classification Decision Tree Model.....	82
Table 39 Comparison of the results between Decision Tree and Random Forest Model	83
Table 40 Facial features and accuracy of each phase	85

LIST OF FIGURES

	Page
Figure 1 Action Units in Basic Emotions	9
Figure 2 68 facial landmarks (Sagonas, 2016).....	10
Figure 3 Decision Tree.....	13
Figure 4 Random Forest.....	14
Figure 5 Gradient Boosted Trees	15
Figure 6 Support Vector Machine.....	16
Figure 7 Rectangle area and Rectangle features	17
Figure 8 Viola Jones Object Detection	18
Figure 9 Cell	19
Figure 10 Histogram of Oriented Gradients	19
Figure 11 k-Means clustering	20
Figure 12 Scatter plot of customer behavior dataset.....	21
Figure 13 Scatter plot of cluster one and cluster two.....	23
Figure 14 Reinforcement Learning concept (Sutton & Barto, 2018)	26
Figure 15 Infrastructure of Facial Recognition Alert System (FRAS)	33
Figure 16 The External Elements	34
Figure 17 Example of Mark point with Facial Landmark Detection (FLD) of Phase I	35
Figure 18 Diagram of the Emotion Classification of the External Elements	36
Figure 19 Sub-Structures of the Face.....	37
Figure 20 Example of Mark point with Facial Landmark Detection (FLD) of Phase II	38
Figure 21 Diagram of the Emotion Classification of the Sub-Structures	38
Figure 22 Example of Mark point with Facial Landmark Detection (FLD) of Phase III	40
Figure 23 Relation among Gender, Emotion, IPF, PFL, and PFR.....	44
Figure 24 Expected relationships among independent and dependent variables.....	45

Figure 25 Relationships among all variables	49
Figure 26 Decision Tree Algorithm ID3 of Emotion Classification.....	50
Figure 27 Refinement of Emotion Classification Decision Tree Algorithm ID3	51
Figure 28 Comparison of classification performance and runtime.....	52
Figure 29 Comparison of Receiver Operating Characteristics (ROC) between Decision Tree and Random Forest.....	53
Figure 30 Inner Brow Raiser (IBR)	54
Figure 31 Brow Lower (BL).....	55
Figure 32 The Inner and Outer Lid Raiser (LR)	55
Figure 33 The Inner Lip Part	56
Figure 34 Relationships among independent and dependent variables	58
Figure 35 Relationships of all variables.....	63
Figure 36 Decision Tree Algorithm for Emotion Classification	64
Figure 37 Comparison of Receiver Operating Characteristics (ROC) between Decision Tree and Random Forest.....	65
Figure 38 Interpalpebral Fissure	67
Figure 39 Palpebral Fissure Length (PFL).....	68
Figure 40 Relationships among independent and dependent variables	73
Figure 41 Relationships among variables	81
Figure 42 Decision Tree Algorithm for Status Classification	82
Figure 43 Comparison of Receiver Operating Characteristics (ROC) between Decision Tree and Random Forest.....	83

CHAPTER I

INTRODUCTION

This dissertation focuses on the extraction of facial features to detect normal situations and abnormal situations through the facial expression of stroke patients. Thus, the main topics in this thesis consist of exploring the facial features, creating a pattern of the features, generating a classification model when the faces are changed.

This chapter is organized into four sections. Section 1.1 is the rationale and problem statement, following by the research objectives in Section 1.2. Then, the scopes and limitations are determined in Section 1.3. Finally, the organization of the dissertation is described in Section 1.4.

1.1 Rationale

Currently, the number of stroke patients is continuously increasing around the world, including Thailand (Avan et al., 2019 and Areechokchai et al., 2017). Generally, the stroke patients were admitted by ischemic stroke or hemorrhagic stroke. Most stroke patients have caused by thrombosis in the cardiovascular system; it made the ischemic stroke (Hao et al., 2017). About 10-15 percent of all stroke patients have the fragile blood vessels and the high blood pressure; this is the cause of the hemorrhagic stroke (Wintermark & Rizvi, 2018). All stroke patients cannot perform any vocal communication, and they cannot move their bodies. Moreover, the patients are always confused in the thinking process; so, they cannot understand any conversations from other persons. However, a severe level of symptoms depends on the level of the disease, or the damage level of the stroke. In some cases, the patients cannot respond to the external stimuli because they have a severely high level of the disease (Cooksley, Rose & Holland, 2018). Fortunately, some patients can express emotion from incentives or painful. Although these patients cannot communicate and move, they can express their feelings via faces, lips, or eyes' components, such as brow raisers, brow lower, eyes (Ancillao et al., 2019).

Since stroke patients usually cannot move themselves, the medical monitoring for medical treatments after the surgery must be reacted properly. Besides, these patients stay usually unconscious for few days after their operations; moreover, when they wake up from this unconsciousness, some of them may be in paralysis and speechless. Thus, they cannot push any button for help when they really need. In the worst case, these patients must wait the medical staffs to come for their visit at the time of the ward schedule. In such case, this brings a risk of the patient's life though the blood pressure and heart rate monitor are implemented in a patient bed; the medical staffs may not assist the patients in time. Furthermore, in the real lives, the number of doctors, medical staffs, and caregivers for stroke patients is very limited

when comparing with the number of the stroke patients which are increasing (Meretoja et al., 2017 and Sugii et al., 2019). Therefore, this limitation of these medical specialists and caregivers can affect to safety of the stroke patients because the patients should be monitored for all times.

For the reason of patient safety, it motivates many researchers to study and design for patient monitoring processes or patient monitoring systems of stroke patients at all the times. Most researchers study facial expression of the stroke patients because of limitation in bodies' movement. For example, the facial expression was applied to the paralysis classification of paralyzed patients; the facial landmarks are detected and the distances between two landmarks are measured afterwards. Then, the values of the distances are used to define the facial features (Barbosa et al., 2019). Moreover, some researchers (Umirzakova et al., 2018) studied to detect stroke symptoms using facial features with wrinkles on the forehead area, eye moving, mouth drooping, and cheek line. This research uses facial images from the iBUG 300-w dataset, removes noises from the images, and marks each image with the 68 facial feature coordinates, finally, computes the distances between two points, such as the left mouth point to the right mouth point, etc. The research of (Umirzakova et al., 2018) was applied to create a stroke detection algorithm with 91% accuracy. Later, these researchers (Umirzakova et al., 2019) studied only the cheek wrinkle in normal people and stroke people; the outcome of this study was used to detect stroke symptom according to the facial images for automatic stroke symptom detection system using Active Appearance Model (AAM). This model refers to a match with a statistic model of object's shape and creates a new image for wrinkle detection; this new algorithm has accuracy as 92%. Though this new algorithm does not apply any machine learning techniques, its accuracy is quite high. Therefore, if the researchers adopt the machine learning mechanism, accuracy of the facial detection and classification should be improved.

The method that has been mentioned in the previous paragraph can be called as pattern recognition since its' originals has objective in detecting an object based on its detected patterns. Furthermore, this proposed pattern recognition system for facial detection in several domains, such as security or healthcare. For example, pattern recognition was applied in the music therapies program to identify the feelings of patients after listening to music (Liu, Lu & Hong-Jiang, 2003). In the healthcare domain, the automatic detection system is employed to detect an anomalous situation for patients in the hospital called the anomaly detection system (Kadri et al. 2016). Moreover, the facial stroke recognition system is applied to classify patients between no-stroke and stroke (Chang et al. 2018). These researchers used the concept of pattern recognition to consists of three stages: preprocessing, feature extraction, and classification. The preprocessing stage refers to the preparation of a raw dataset from non-computational data format to be a computational format with data cleaning and data transforming. This preprocessing stage is important because it filters improper data, such as incomplete or inconsistent data, from the dataset before sending to the next step. Thus, the outcome from this stage is the suitable dataset that is ready to be

extracted as features. A feature extraction stage refers to selection of some features from data input like RGB value in the color images. The feature extraction step composes of three processes: feature construction (FC), feature selection (FS), and feature reduction (FR). The FC process aims to disclose a relationship of the features from the input dataset obtained from the preprocessing stage. The result from FC is the features and their relationships that can be identified from the input dataset. Then, these identified features are sent to the FS process, and subsets of these features are defined with a combination of the flexible format (Rudnicki et al. 2015). Furthermore, the FR process also reduces noisy features according to a specific criterion (Sondhi, 2009). Consequently, the feature extraction stage optimizes the number of features for creating a classification model. Then, the classification class in the experiment has to be applied, such as no-stroke, and stroke, to the training dataset. The product from the training process is the classification model that is used to predict the value of the class when the values of features are known. However, this classification model must be evaluated by the testing process using testing dataset before used. Furthermore, there is some mechanisms to assist identification feature extraction obviously called Facial Action Coding System (FACS) (Ekman & Rosenberg, 2005). The mechanisms of FACS use the AUs not only for facial detection, emotion detection, classification, and recognition, but also for the machine learning process to improve the classification and recognition model. For example, a decision tree algorithm was applied to recognize emotions on social networking sites using the sentence level classification (Dixit, Kumar-Pal et al. 2017), or to detect emotions on the faces from images (Salmam et al. 2016).

Although there are many studies of facial expression and investigate facial features for stroke patients, dissimilar variables on the face have been selected for each research. As mentioned above, wrinkle on the face, such as forehead area, eye moving, mouth drooping, and cheek line, had been applied to detect symptom of stroke. Unfortunately, the variables of wrinkle cannot discover the symptom of patient strokes completely. So, the stroke symptom detection, finally, uses a moving hand combining with the cheek line wrinkle to obtain a better detection result (Umirzakova et al., 2018 and Umirzakova et al., 2019). Another interesting method that had been implemented to detect the stroke symptom was the use of regions around left and right eyes, including mouth. This method can identify both stroke and no-stroke situations, but not include the critical circumstance of stroke (Chang et al., 2018). Thus, monitoring for the critical state of the stroke patients must be performed manually. Therefore, whenever the caregiver is not available when the patients need caring according to their serious sickness, there can be a risk for patients' lives. According to such condition, monitoring of stroke patients to improve patient safety and quality in the hospital when an emergency occurred is vital.

For the problem mentioned above, this research aims to apply the feature extraction mechanism based on the extracted features from the variables on the face, such as inner brow, brow, eyelid, and lip, to create a classification model for

distinguishing between normal and abnormal situation, or critical situation, of stroke patients. Then, the alarm will be sent afterwards to a caregiver when needed.

1.2 Research Objectives

To develop a facial expression detection model using facial features.

1.3 Research Scopes

This dissertation mainly concerns about the feature extraction techniques on the face for creating the classification model and developing the automatic alert system. Thus, recruitment criteria of volunteers in this research is listed below.

1. Samples must be Thai stroke-patients and non-Thai actors from the movies.
2. Samples should have age ranges between 18 to 80 years old, both male and female.
3. Samples must not be blind, cataract, or have a mental health problem, such as psychotic disorder and mood disorder.
4. Samples must be able to perform facial expressions, such as brow raiser, lip part, open eyes, and close eyes.

1.4 Benefits

The automatic facial expression detection system has been developed and implemented in various areas, especially in the healthcare systems. Moreover, some automatic facial expression detection techniques are depended on high performance devices and may not be worth or flexible to be implemented in the real lives. Otherwise, the techniques might not support the critical care system for critical stroke patients or other sickness. Therefore, this research proposes an efficiency facial expression detection model in the form of a classification model that can classify stroke patients who are in the abnormal status out off the normal stroke patients. Thus, the caregiver will have automatic assistants to monitor patients for all times as needs and patients' relatives will feel comfortable that the patient is in the good care for all times. Besides, the devices that are implemented in this research are common camera that can be found easily in the market. So, this proposed method is not only easily to be implemented, but also worthwhile to be used for monitoring any patients who require a close caring.

1.5 Structure of this Thesis

The organization of the remainder of this thesis is as follows. Chapter II provides the definitions and the related works about the stroke patients symptoms, facial extraction, and machine learning algorithms. Then, methodology and analysis are described in Chapter III while the results are explained in Chapter IV. Finally, the discussions and the conclusions are presented in Chapter V.



CHAPTER II

RELATED WORKS

This chapter contains details of some literatures that related to this research. In Section 2.1, literatures of stroke patients monitoring symptom of the stroke patients have been surveyed, following with the Facial Action Coding System (FACs) technique in Section 2.2, including the facial landmark on the face. The Machine learning has been elaborated afterwards in Section 2.3. Then, tools and libraries are shown in Section 2.4 where Section 2.5 describes techniques related to the researches.

2.1 Stroke Patient

Stroke or cerebrovascular accident is the symptom of the lack of oxygen and blood in the brain since the artery is constricted, blocked, or broke. When blood cannot flow to the brain, the brain tissue will be destroyed and lost control of the body's movement. The severity of this symptom is depending on the damage of the tissues and the strength of each person; so it can be classified to various levels. Some patients may express the weakness of one side body or face immediately, while some patients may take a long time to reveal the symptoms. However, stroke patients should meet a doctor at once when the symptoms occurred.

Stroke is divided into two types: ischemic stroke, and hemorrhagic stroke. The first type of the stroke is the ischemic stroke caused by the artery blocking. In such case, the oxygen blood in the artery cannot flow to the brain (Summers et al., 2009 and Mozzafarian et al., 2016 and National Center for Chronic Disease Prevention and Health Promotion, 2020). Most stroke patients, including Thailand and foreign countries, often are ischemic stroke (Summers et al., 2009 and Hao et al., 2017 and Areechokchai et al., 2017).

The second type of the stroke is a hemorrhagic stroke; this type can be separated into two subtypes: intracerebral hemorrhage, and subarachnoid hemorrhage. For the intracerebral hemorrhage, it occurs when the artery cracks and causes flooding blood to the brain. On the other hand, the subarachnoid hemorrhage refers to the situation that there is bleeding surrounding the brain and causes blood-blend with tissue in the brain. The researches had found that the intracerebral hemorrhage occurs more than the subarachnoid hemorrhage (Summers et al., 2009 and National Center for Chronic Disease Prevention and Health Promotion, 2020). Moreover, stroke can occur in both males and females in a similar proportion. However, the risk of stroke often happens in the elderly population more than young adults (Zöllner et al., 2020).

Though ischemic stroke and hemorrhagic stroke are caused by different artery's damage, the stroke's signs can be observed in the same from for their initial symptoms. The signs of the stroke occurrence, both males and females, are facial

weakness, one side of the body weakness, speak troubling and confusing, see disturbing, walking difficulty, and headache severity (National Center for Chronic Disease Prevention and Health Promotion, 2020). Since the stroke patient needs to have medical treatment within three hours after the first symptom is noticeable, every second during the stroke may damage the artery and the brain that can cause to death (National Center for Chronic Disease Prevention and Health Promotion, 2020).

Therefore, observation of stroke symptoms in the risk group such as the elderly can be assisted and be diagnosed in time. In 1998, United Kingdom had defined a crucial key for initial stroke testing, called as FAST which comes from Face, Arms, Speech, and Time. Later, this method was implemented by other healthcare organizations, including educate to the public on detecting symptoms of a stroke. Each acronym can be described in Table 1.

Table 1 Key FAST of Stroke Description

Acronym	Stand for	Description
F	Face	Test by smiling, drooping on one side of the face
A	Arms	Test by raise both arms, drifting downward on one arm
S	Speech	Test by repeat a simple phrase, speaking difficult
T	Time	Call emergency when occur these signs

Next, some stroke experts at Beaumont Health created an acronym as FASTER to identify the stroke's key. FASTER will be the importance of recognition of signs and symptoms of the stroke. Thus, there are added S (Stability) and E (Eyes), every acronym is displayed in Table 2 (Beaumont).

Table 2 Key FASTER of Stroke Description

Acronym	Stand for	Description
F	Face	Test by smiling, drooping on one side of the face
A	Arms	Test by raise both arms, drifting downward on one arm
S	Stability	Test by standing, trouble walking or difficulty balancing
T	Talking	Test by repeat a simple phrase, talking difficult
E	Eyes	Test by seeing, loss of vision in one eye or both eyes
R	React	Call emergency when occur these signs

After the stroke patients were taken away to the hospital, the doctor will diagnose the cause of stroke from blood tests, angiograms, CT scans, MRI scans, and electrocardiograms (EKG). The treatment of ischemic stroke will use anticoagulant; it helps resist clot-busting blood until the artery blocked. Moreover, antiplatelet may be treated to the patient for solving the clot-busting blood platelet.

In the case of hemorrhagic stroke, the artery in the brain cracked, a doctor may have to surgery the brain. However, situation of stroke patients can be complicated during the treatment period, such as high blood pressure, dyslipidemia, pulmonary embolism, or hydrocephalus. Thus, patient monitoring is very important for patient safety.

2.2 Facial Action Coding System (FACS)

Facial Action Coding System (FACS) refers to measure the movement of anatomy system on the human face. FACS was announced in 1978 by Ekman and Friesen, and it was revised in 2002 (Ekman & Friesen, 1976 and Ekman & Rosenberg., 2005: 12-13). Ekman and Friesen (Ekman et al., 1976) developed FACS under six basic emotions, such as happiness, anger, disgust, surprise, sadness, and fear. FACS consists of action units (AUs), the facial action code (FAC) name, and muscular basis (Zhang et al., 2018). Each AU is a number that represents a basic action of a particular muscle. Moreover, each FAC name is the title of the activity that describes the response on the face related to the AU. Additionally, the muscular basis explains the motions of muscles. Thus, each fundamental emotion can be represented by AUs; for example, happiness contains the AU6 and the AU12, while surprise comprises the AU1, AU2, AU5, AU25, and AU26. The AU6 refers to the muscle under an eyelid when the cheek raises, and the AU12 describes the corner of the mouth when smiling, as displayed in Figure 1.



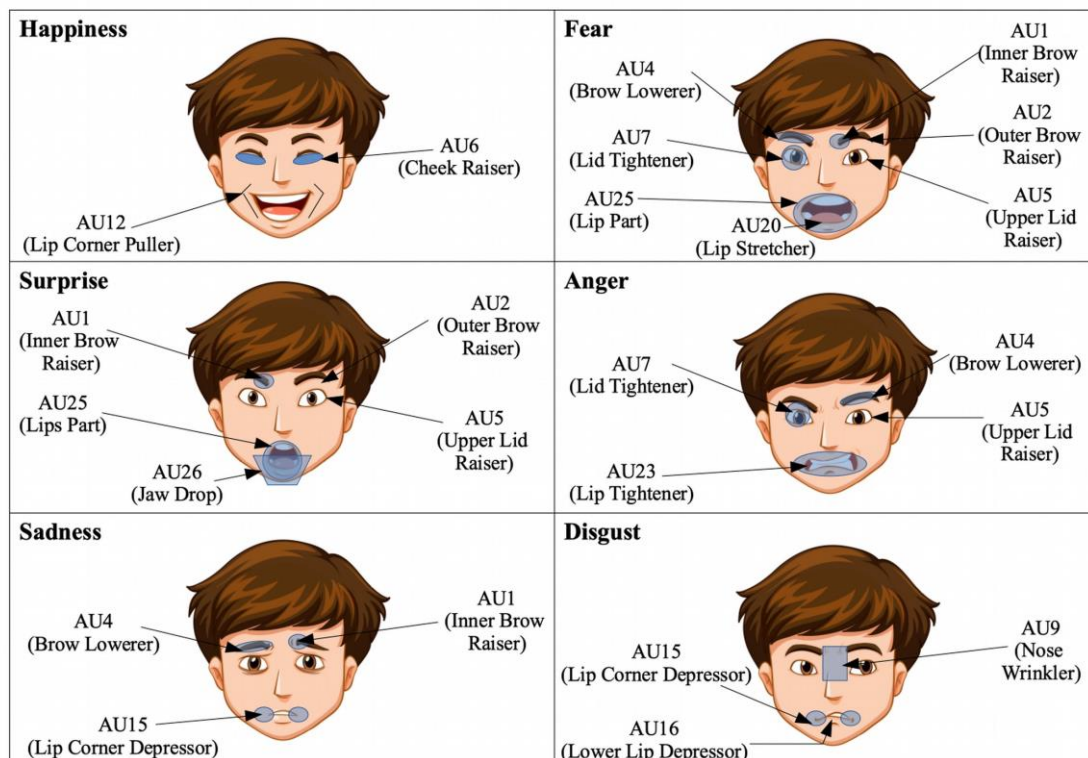


Figure 1 Action Units in Basic Emotions

In the past decade, various studies have investigated facial detection and recognition using the Facial Action Coding System (FACS) based on Paul Ekman's theory (Moeini, Moeini et al., 2017 and Wolf, 2015 and Ekman et al., 1976 and Benitez-Quiroz et al., 2015). Furthermore, FACS has been applied in the areas of neuroscience and psychology to detect sentiments of mental illness (Wolf, 2015), including autism disorder (Manfredonia et al., 2018) through facial expression. Therefore, FACS is relevant to the structure of facial movement. The measurement of facial emotional expression using encoding on the face is composed of many action units (AUs).

Although each AUs will present facial muscle movement, clue muscle will display in the form of an area. It is not exhaustive for the changing face of detection and recognition. Thus, point determination will show a mark of the changing on the face obviously as explained in the following section.

2.2.1 Facial Landmark

Facial landmarks refer to the position's identifying of points or feature points on the human face through muscle movement of organs, such as eyebrow, eye, nose, and mouth, called as Facial Landmark Detection (FLD). The FLD is essential to analyze emotion status via facial expression. There are various researchers those

applied the FLD for facial feature detection in computer vision task such as facial verification (Amato et al., 2018), or head pose estimation (Xu & Kakadiaris, 2017). Then, the FLD is used to develop an automatic detection system in the healthcare area research. For example, there is to investigate facial palsy in Parkinson disease and stroke with 68 facial landmarks (Guarin et al., 2020), as displayed in Figure 2.

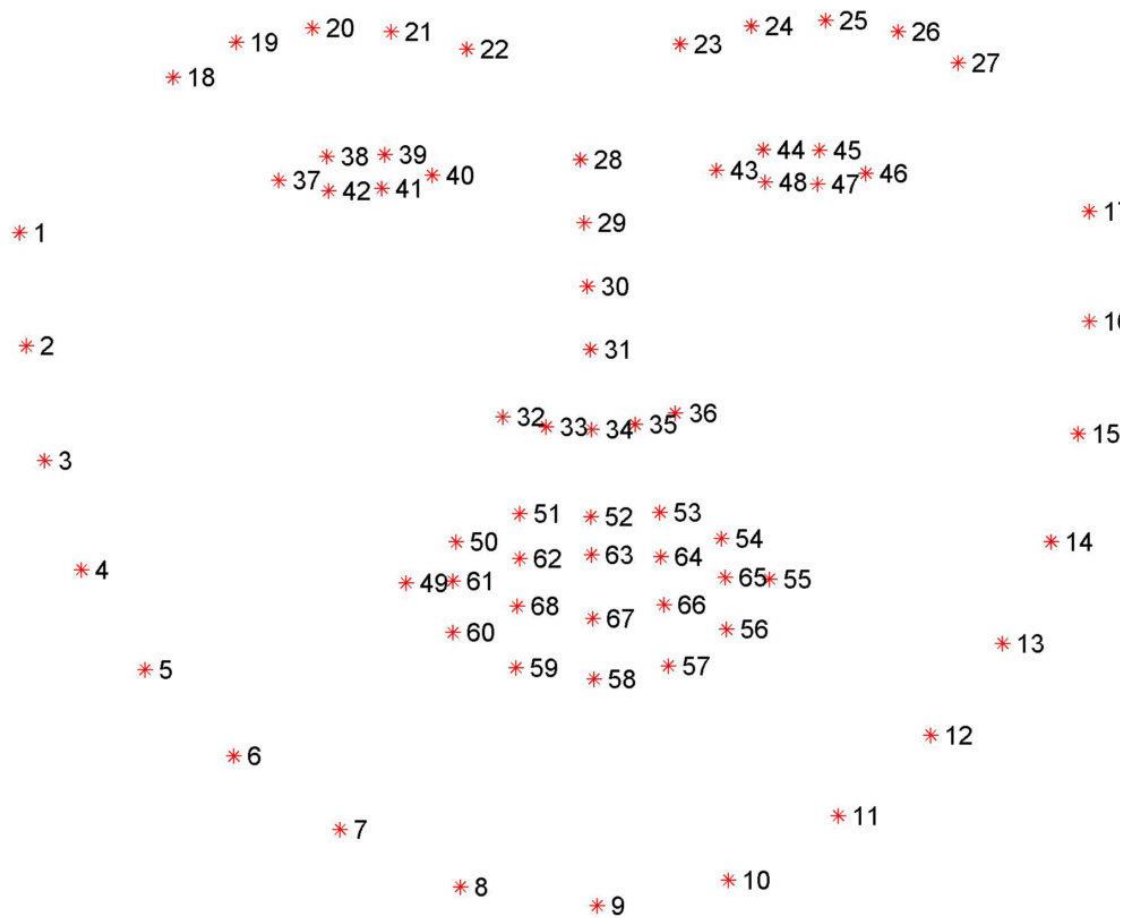


Figure 2 68 facial landmarks (Sagonas, 2016)

The points of 68 facial landmarks can be implemented using opensource Machine Learning (ML) library such as Dlib. The Dlib library can extract the coordinate of facial feature points on the face from an image. The 68 facial landmarks can be divided into eleven areas as left jawline, right jawline, chin, the inner edge of lips, the outer edge of lips, bottom of the nose, bridge of the nose, left eye, right eye, left eyebrow, and right eyebrow, as presented in Table 3.

Table 3 Features in 68 Facial landmarks

No.	Feature	Point
1	Left jawline	1-8
2	Chin	9
3	Right jawline	10-17
4	Left eyebrow	18-22
5	Right eyebrow	23-27
6	Bridge of nose	28-31
7	Bottom of nose	32-36
8	Left eye	37-42
9	Right eye	43-48
10	Outer edge of lips	49-60
11	Inner edge of lips	61-68

After obtaining the facial landmark, facial features can be investigated from a distance changing between two points; for example, the position between the top of the eyelid (point 38) and the bottom of the eyelid (point 42) has changed from distance measurement using Euclidean distance. Then, these features are the training dataset in the ML model for specifying a pattern.

2.3 Machine Learning

In the last few decades, ML has been an extensively interesting and favorite study of computer scientists and developers. Based on its efficiency, ML has been applied to develop various systems, such as customer decision support systems, computer vision systems, cybersecurity systems, and pattern recognition systems, etc.

It is a fact that humans have been trying to invent a perfect neural network that can work similar to the human brain. Therefore, the first neural network system was invented in the year 1943 by a neurophysiologist and a young mathematician (Spitzer, 1998). Their invention led to the implementation of artificial intelligence which is a simulation system of the human intelligence process by a computer or neural network system (Nasser & Abu-Saser, 2019 and Nasser et al., 2019). After the implementation of AI, computers are able to analyze and provide the optimal solution for any complex problems from human knowledge. Nevertheless, human knowledge can also be obtained from a large amount of data that are stored in databases of organizations. Therefore, a machine learning system was introduced to extract new knowledges and new solutions from all available data in the organizations. In such a manner, it can be said that ML can perform many more complicated tasks than AI; and, ML is a subset of AI (Handelman et al., 2019). Considering the analysis process of ML, the data are analyzed through a data structure that can be seen as one layer of analysis. Thus, the extracted content can be limited. Therefore, to obtain deep and clarified information, deep learning (DL) was implemented to analyze data in multiple layers. The result

from data analysis of multiple layers in DL is totally different from the result from ML because data are analyzed in various dimensions. As a result, it can be said that DL is a subset of ML (Hutter et al., 2019).

ML is associated with artificial intelligence (AI); and, deep learning (DL) is inseparable. AI is superset of ML; AI can be embedded in hardware, robots, sensors, and the Internet of Things (IoT). Therefore, AI means the development of computer programmes that are capable of imitating humans (Lu et al., 2018). Afterwards, the abilities to classify and predict patterns by computer program have been improved and called machine learning (ML). Thus, ML relates to exploring computational algorithms with learning from the surrounding environment. ML is composed of supervised learning, unsupervised learning, and reinforcement learning.

2.3.1 Supervised learning

Supervised learning is training to make comparison according to a learning condition set by a human, and these comparisons are called training process. The training process will map an input to an output data; the output data will be marked with a label and be used for learning. After learning the training data, the learning system can explore an answer to a problem without human intervention. In general, the supervised learning can be employed to solve two issues: regression, and classification problems. A regression model is used to predict continuous values from either single or multiple features, such as salary or weight. In a classification problem, it is used to classify the binary class, such as “0” or “1”, “black” or “white”, and “animal” or “plant”, including multi-class problems, such as documentaries, action movies, and comedy movies. The fundamental algorithms of supervised learning are such as a decision tree, a random forest, and gradient boosted trees.

The decision tree is deployed when classifying a group of data using data features, called attributes. For example, the marketing department requests for types of payment methods: credit card, debit card, or cash, from customers when they buy a product at the store. The customer data are used to establish a training dataset based on some attributes’ specifications, such as age, income, occupation, and payment methods. Within the training dataset, the payment method is the output of the classification method based on data of age, income, and occupation. Thus, the output of the classification is labelled as the payment method, as shown in Table 4.

Table 4 Attributes of Training Data

Age	Income	Occupation	Payment Methods
30	High	Officer	Credit Card
20	Low	Student	Cash
43	Medium	Analyst	Debit Card
8	Medium	Accounting	Credit Card
56	Medium	Housekeeper	Cash

According to the data in Table 4, some attributes will be chosen for the classification, based decision tree model that comprises an internal node, branch, and leaf node. In this study, the internal nodes represent age, income, and occupation. The links refer to the selection conditions of the internal node's value. For each internal node, there can be binary links or multiple links. On the other hand, the leaf node is applied to display the classification's result namely the payment methods. These descriptions are drawn in Figure 3.

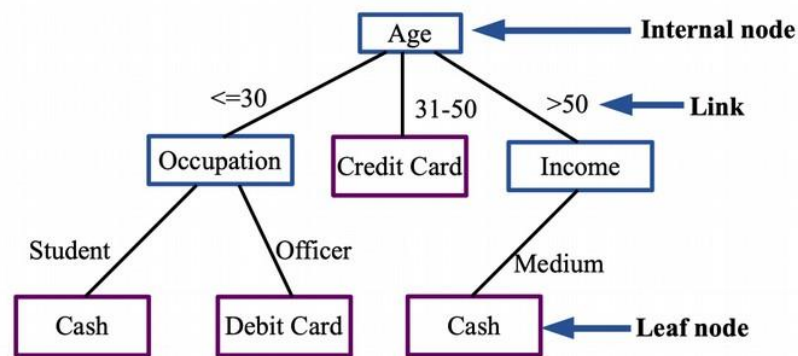


Figure 3 Decision Tree

The next fundamental algorithm is a random forest where many sub-models from the decision trees are deployed to create the required model. Each sub-model refers to an individual decision tree with an independent dataset; among different datasets may or may not have overlapping attributes. When data are arranged to many sub-datasets, as shown in Figure 4, each dataset can derive to a specific decision tree. All these trees are passing through the voting procedure where a new dataset is input to all trees and the prediction accuracy from every tree will be compared. Only one best decision tree, the tree with smallest prediction errors, out of all trees will be selected as the decision classification tree afterwards.

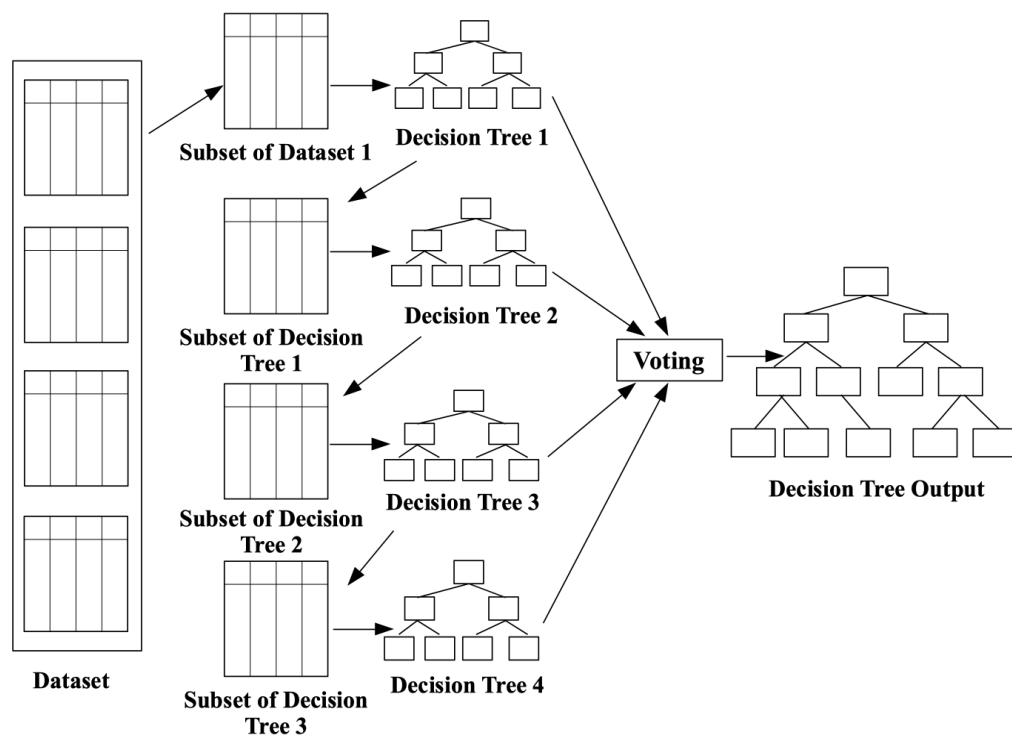


Figure 4 Random Forest

As mentioned above, one of the fundamental algorithms of supervised learning is gradient boosted trees. This algorithm uses a boosting technique presented in Figure 5, to create sequence decision trees as the prediction models. Consider the processes in Figure 5, when the dataset#1 was derived to the first tree, only some data in the dataset#1 were selected to create the tree#1. Thus, the data in the tree#1 is named as dataset#2 and it is subset of dataset#1. However, if the accuracy of the classification tree#1 is low, the process to create the tree# i is repeated using dataset# i until the classification model has high accuracy in the acceptable score. The entire process is called an ensemble method. The advantage of this boosting technique is to help increasing the efficiency of the trees and obtain the optimized prediction model.

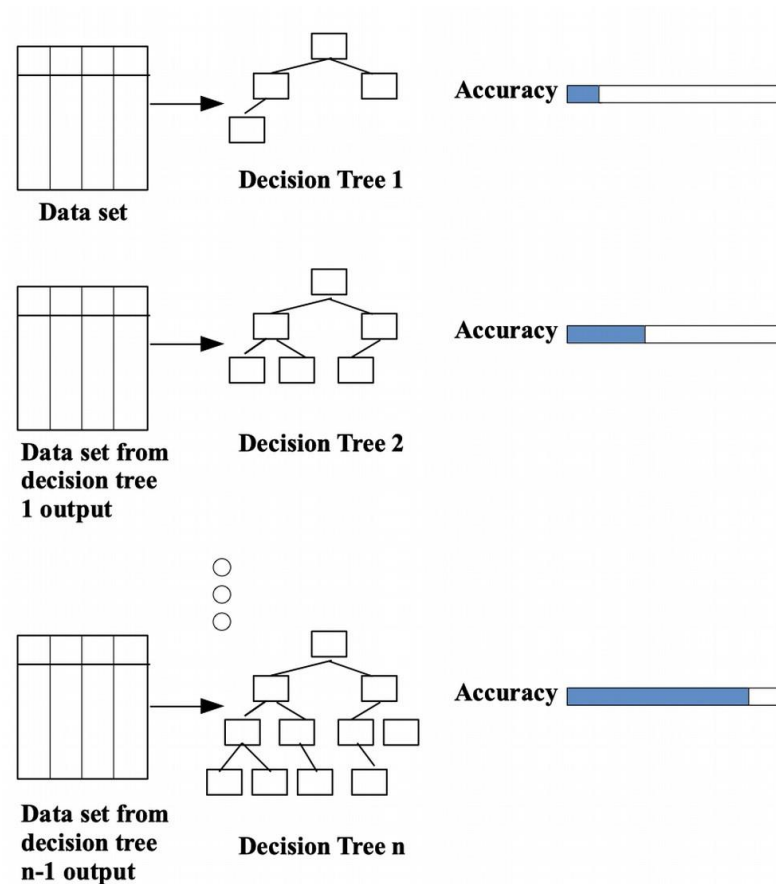


Figure 5 Gradient Boosted Trees

Furthermore, there are other supervised learning mechanisms of ML, such as naive Bays, nearest neighbor, and support vector machines. Naive Bayesian learning is a method developed by Thomas Bayes; it is a classification model that uses the probability based on Bayes theorem for an independent event. For example, assume $P(x)$ represents the probability of event x . $P(x/A)$ describes the probability of event x when event A occurred. Thus, the event will be predicted to occur from Eq (1). However, the events used in the prediction must be consistent with the events to be predicted

$$P(x/A) = [P(A/x) * P(x)] / P(A) \text{ -----(2.1)}$$

The next algorithm is the k-nearest neighbor (k -NN) that is used to predict data from the attribute label. One record of the dataset is represented with a point. A new point must find the closest point for classification. However, the classification cannot be concluded from only one nearest point because of lack of credibility. Thus, the trusted outcome can be predicted from the multiple points, called k .

The last algorithm is the support vector machine (SVM) that is applied to classify two-group datasets based on a straight line or a hyperplane. Since many hyperplanes can be drawn between two-group datasets, only one of the hyperplanes that has the farthest distance from both groups is the most important line and is named as the best hyperplane. From Figure 6, three straight lines can separate the data into groups. Unfortunately, the distance between the L1 band to the closest point is very short while the distance between L3 to its nearest point is also short. Therefore, the gap between L2 and its closest point is the largest. Thus, the optimized hyperplane that has the maximum margin from both datasets is L2.

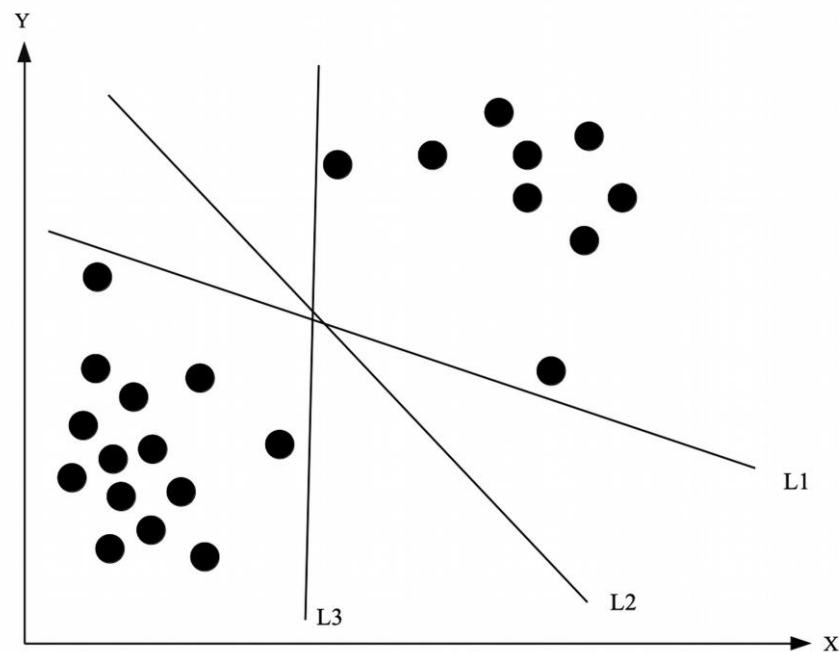


Figure 6 Support Vector Machine

All of these supervised learning algorithms mentioned above can be solved for both regression and classification. The algorithms are applied to classify data both small and medium-sized datasets, plus creating a classification model or a prediction model. However, for any complex datasets, parallel data processing such as the deep learning (DL) method, is required. DL is usually applied to process numerous data for improving a result of the prediction model. The concept of DL is like the artificial neural networks (ANN) that the DL's learning layers can be implemented as multilayer networks.

A learning model, either created from the ML or DL mechanisms, needs to execute the feature extraction for detecting objects from an input source, such as numbers, texts, and images. Commonly, the feature extraction techniques of ML consist of the Viola-Jones framework, which is based on Haar-like features (Goyal et al., 2017, Elmer et al., 2005, Kadir et al., 2014, and Akash et al., 2016) and the histogram of oriented gradients (HOG). This Viola-Jones framework was firstly

introduced for facial detection; the mechanism of this method is to distinguish between positive pixels and negative pixels from the face images.

As mentioned above, the Viola-Jones framework using the Haar-like features converts a face image from a colour image to the grey-scale image. Then, the grey-scale image is used to generate an integral image by calculating the rectangle area, as displayed in Figure 7(a).

The integral area (IA) from a rectangle is calculated by pixel coordinates of point A, B, C, and D, as shown in Eq (2). The IA represents the integral area values of a rectangle to set a window and sliding windows corresponding to both vertical and horizontal features of the images. The windows are arranged from the weighted sum of 2D integral rectangle features, presented in Figure 7(b).

$$IA = IA(x_A, y_A) - IA(x_B, y_B) - IA(x_C, y_C) + IA(x_D, y_D) \text{-----}(2.2)$$

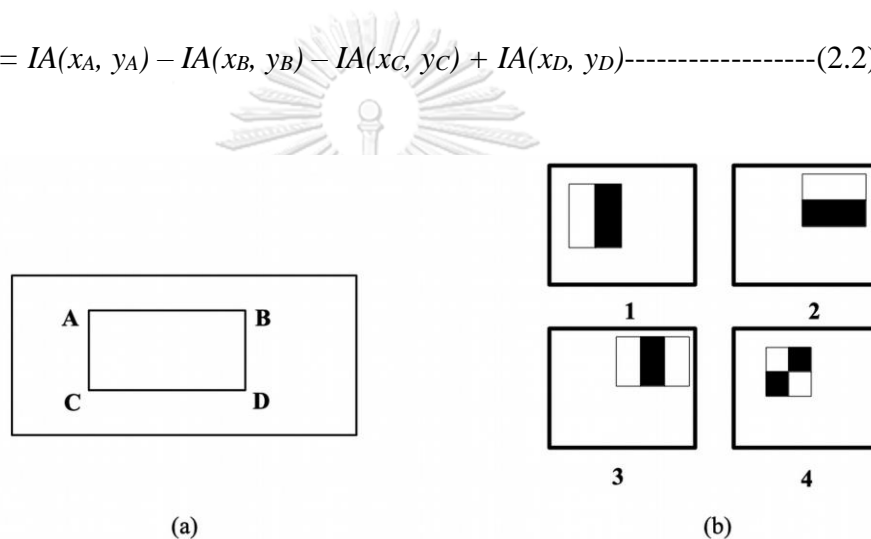


Figure 7 Rectangle area and Rectangle features

Black areas are positive weight, and white areas are a negative weight, and this representation is called Haar-like Features. Each step is set up as N -face recognition filters, and the positive weight is identified as the face throughout the repeated process, as shown in Figure 8.

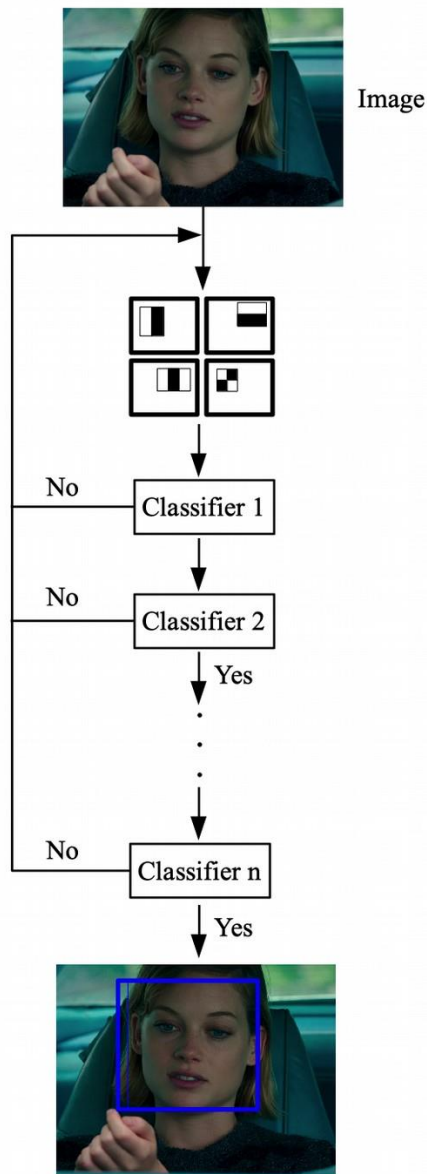


Figure 8 Viola Jones Object Detection

The next object detection method is the histogram of oriented gradients (HOG) which focuses on the extraction of shape for human and non-human classification. The first step is to explore the edge of the object; then, the image is divided into many small regions, each of $n \times n$ size each, called as cells, as displayed in Figure 9. Next, calculation on each cell are performed to find the size and directions for both the horizontal (x-axis) and vertical (y-axis), called gradient histograms, as presented in Eq (2.3), Eq (2.4), Figure 10(a), and Figure 10(b).

$$G_x = f(x + 1, y) - f(x - 1, y) \text{-----}(2.3)$$

$$G_y = f(x, y + 1) - f(x, y - 1) \text{ -----(2.4)}$$

Then, the gradient histograms are computed by the gradient frequency using Eq (2.5). The results are normalized from the histogram of each cell and the specification feature is drawn, called the HOG descriptor, as shown in Figure 2.10(b).

$$\theta(x,y) = \text{atan} (G_x^2 / G_y^2) \text{ -----(2.5)}$$

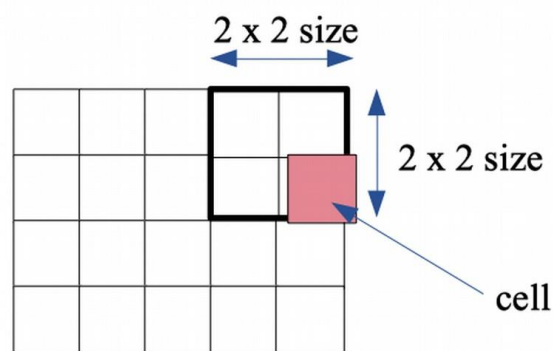


Figure 9 Cell



a. Input Image



b. Histogram of Oriented Gradients (HOG)

Figure 10 Histogram of Oriented Gradients

2.3.2 Unsupervised learning

The second of machine learning is unsupervised learning that is learning no teacher and unknown output. Thus, unsupervised learning will extract knowledge or features from input data. Unsupervised learning can be separated into two kinds: clustering and dimensionality reduction. Clustering is a grouping of data from similar items, for example, image from uploading photos to social media may be grouped with the same face. Though, the social media does not know which images as whom, clustering will be extracted from all items on the face and grouping similar face. The dimensionality reduction refers to decrease dimension of data from many dimensions

to two dimensions. This type of unsupervised learning no need to store the data completely, it can be clustered of data.

The favorite clustering of unsupervised learning is k -Means clustering; it is partition data into k groups from similar features of input data, as displayed in Figure 11(a). The alphabet k is a number of groups to identify before training.

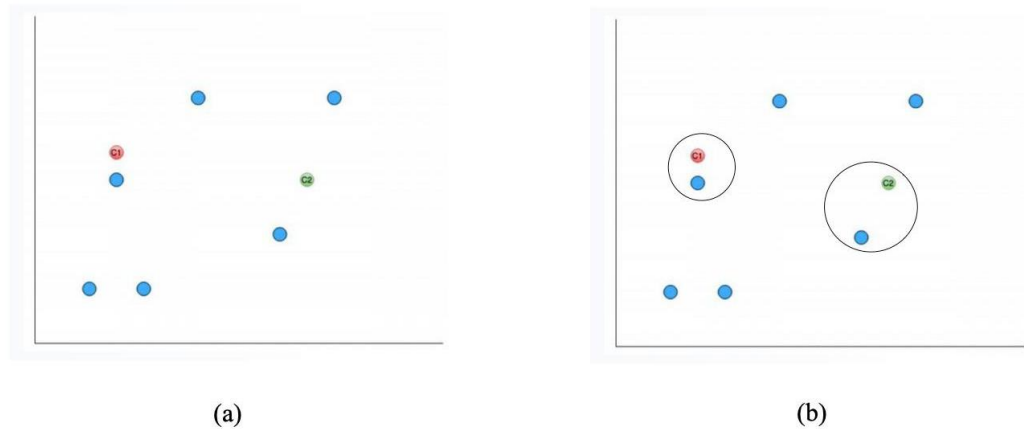


Figure 11 k -Means clustering

From Figure 11 has the input data as red point and green point for grouping. Consider distancing of red point; it will be grouped from the shortest distance between red point and that point including green point, as shown in Figure 11(b). For example, a company have to analyze and cluster customer behavior from product purchase, data followed in Table 5.

Table 5 Customer Behavior Dataset

No.	Customer ID	Amount of Purchase (Bath)	Number of Visit Shop
1	ID112	64	22
2	ID115	72	74
3	ID221	40	22
4	ID226	130	99
5	ID775	105	70

Assume that this dataset has to cluster into two groups, so k value is defined as 2. The dataset in Table 5 can be displayed in the scatter plot in Figure 12(a). Then, the k -Means algorithm will try grouping from the calculation of the center point of each data group. From Figure 12(b), point one and point 5 are identified as center points called the centroid. The next step is distance calculation between the centroid and other points; data will be grouped with the nearest centroid using formula Eq (2.6).

$$Distance = \sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2} \text{ ----- (2.6)}$$

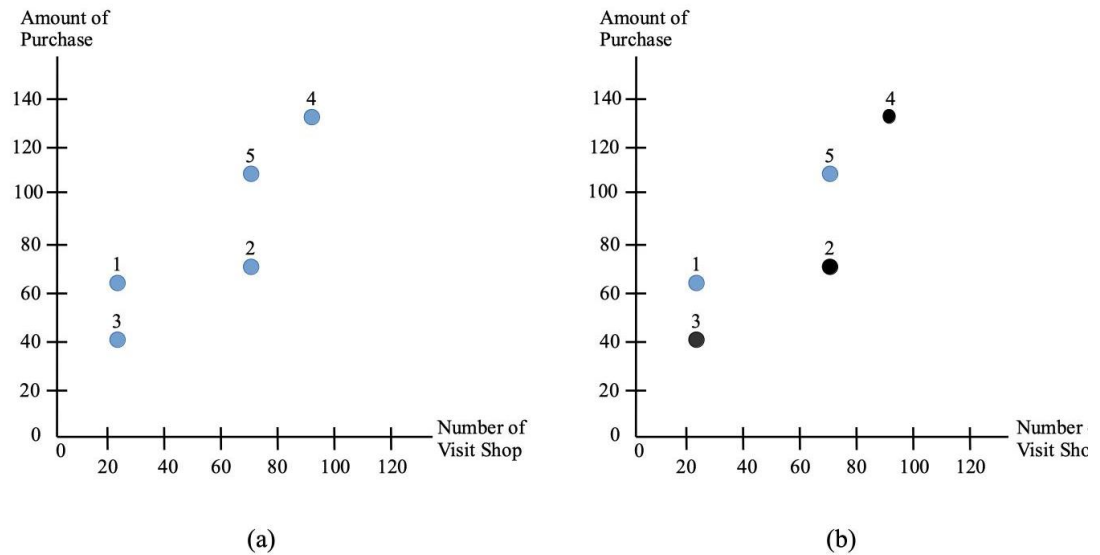


Figure 12 Scatter plot of customer behavior dataset

After distance calculated, the dataset can be clustered as shown in Table 6. Point one and point three are grouped into cluster one while point two, point four, and point five are grouped into cluster two.

Table 6 Clustered dataset

Points	(x, y)	Distance of Centroid 1 (22, 64)	Distance of Centroid 5 (70, 105)	Cluster
1	(22, 64)	0.00	63.13	1
2	(74, 72)	52.61	33.24	2
3	(22, 40)	24.00	80.80	1
4	(99, 130)	101.41	38.29	2
5	(70, 105)	63.13	0.00	2

Then, each cluster have to calculate new mean of group for new centroid, as shown in Eq (2.7) and Eq (2.8). According to data in Table 6, cluster one and cluster two have to calculate new centroid, as displayed in Table 7.

$$\text{Cluster 1} = ((22+22)/2, (64+40)/2) = (22, 52) \text{ -----(2.7)}$$

$$\text{Cluster 2} = ((74+99+70)/3, (72+130+105)/3) = (81, 102.33) \text{ -----(2.8)}$$

Table 7 New centroid of clustered dataset

Points	(x, y)	Distance of Centroid 1 (22, 52)	Distance of Centroid 5 (81, 102.33)	Cluster
1	(22, 64)	12.00	70.36	1
2	(74, 72)	55.71	31.13	2
3	(22, 40)	12.00	65.17	1
4	(99, 130)	109.60	33.01	2
5	(70, 105)	71.51	11.32	2

After calculating the distance of new centroid, cluster one and cluster two have to calculate the new mean of each group again. The new centroid must not be different from the previous centroid; otherwise, the calculation must be repeated until the new centroid has the same value as the first centroid.

In the case of the changed centroid, it will be calculated until the centroid value does not change. Thus, the centroid calculation of the second time is presented in Eq (2.9) and Eq (3.0).

$$\text{Cluster 1} = ((22+22)/2, (64+40)/2) = (22, 52) \text{ -----(2.9)}$$

$$\text{Cluster 2} = ((74+99+70)/3, (72+130+105)/3) = (81, 102.33) \text{ -----(3.0)}$$

CHULALONGKORN UNIVERSITY

Consider the outcomes of Eq (2.7) and Eq (2.9), these outcomes are centroids of cluster one; Moreover, they are equivalent to one another although the computing performed on different tables. Similarly, it can be concluded that the centroid of cluster two can be obtained from Eq (2.8) and Eq (3.0). Grouping of the cluster one and the cluster two are displayed in Figure 13.

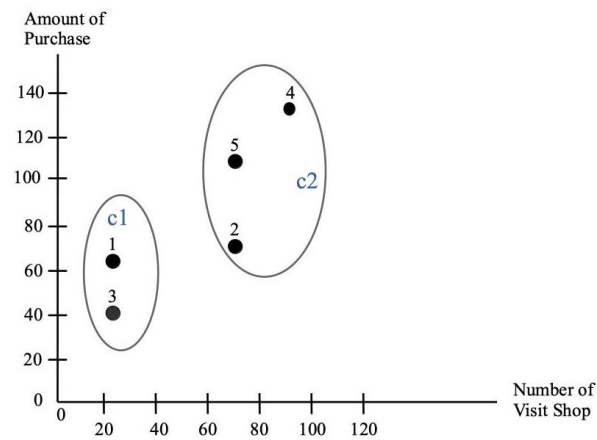


Figure 13 Scatter plot of cluster one and cluster two

The next clustering algorithm is the hierarchical clustering; the data are clustered based on distance calculation like the k -Means clustering's concept. Example of this method can be considered from Table 8, the same cluster is selected from the smallest value, except zero, which is 24.00 in the row#1 and column#3 or vice versa.

Table 8 Show the smallest value

No. (x, y)		(x ₂ , y ₂)				
		1 (22, 64)	2 (74, 72)	3 (22, 40)	4 (99, 130)	5 (70, 105)
(x ₁ , y ₁)	1 (22, 64)	0.00	52.61	24.00	101.41	63.13
	2 (74, 72)	52.61	0.00	61.06	63.16	33.24
	3 (22, 40)	24.00	61.06	0.00	118.44	80.80
	4 (99, 130)	101.41	63.16	118.44	0.00	38.29
	5 (70, 105)	63.13	33.24	80.80	38.29	0.00

From Table 8, rows of number one and number three are combined and columns of number one and number three are merged as cluster one. The results of these can be shown in Table 9.

Table 9 Combine rows and columns for Cluster 1

No. (x, y)	Cluster 1	2 (74, 72)	4 (99, 130)	5 (70, 105)
Cluster 1	0.00	52.61	101.41	63.13
2 (74, 72)	52.61	0.00	63.16	33.24
4 (99, 130)	101.41	63.16	0.00	38.29
5 (70, 105)	63.13	33.24	38.29	0.00

This step will be repeated until all data are combined. Thus, Table 9 is used to group as cluster two, the smallest value is 33.24 of the distance between data number two and data number five. It can be displayed in Table 10.

Table 10 Combine rows and columns for Cluster 2

No. (x, y)	Cluster 2	4 (99, 130)	5 (70, 105)
Cluster 2	0.00	63.16	33.24
4 (99, 130)	63.16	0.00	38.29
5 (70, 105)	33.24	38.29	0.00

From Table 10, the smallest value is 38.29 in row#5 and column#4. It can be combined rows and columns for cluster three as displayed in Table 11.

Table 11 Combine rows and columns for Cluster 3

No. (x, y)	Cluster 2	Cluster 3
Cluster 2	0.00	33.24
Cluster 3	33.24	0.00

From Table 11, the smallest value between cluster two and cluster three is 33.24, so cluster four has to combine rows and columns as presented in Table 12.

Table 12 Combine rows and columns for Cluster 4

No. (x, y)	Cluster 4
Cluster 4	0.00

From Table 8 to Table 12, all data are grouped, and it can be clustered as four clusters. The hierarchical clustering shows a relationship of the dataset; this clustering is usually applied for blind grouping, or the number of groups is undefined before hands. Consequently, this method is time consuming and it does not always fit for all data types.

2.3.3 Reinforcement learning

Reinforcement learning is different from supervised learning and unsupervised learning; it does not use any data repositories for learning and pattern identification. The concept of reinforcement learning is to deliver cumulative rewards and punishments from the measurement of performance under defined situations. Thus, reinforcement learning is applied in the game industries. For example, Tower defense game aims to attack objects or items of an enemy and protects the tower from any intruders. Thus, the character needs to choose the right movement direction to prevent their tower, including attack the enemy's tower. In the event of choosing the right direction, the characters will receive a reward, and the characters have to learn more than one round for cumulative rewards. On the other hand, the first round of learning may have incorrect movement; the characters will receive a punishment, such as a score deduction (Michailidis et al., 2019 and Phillipa et al., 2014). Reinforcement learning concept is presented in Figure 14.

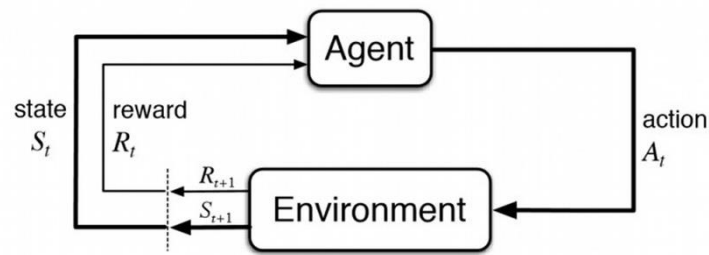


Figure 14 Reinforcement Learning concept (Sutton & Barto, 2018)

2.4 Tools and Libraries

This research deploys tools and libraries for the development of an algorithm, including testing of the algorithm. These tools and libraries are described below.

2.4.1 PyCharm Community Edition

In this research uses Python language to develop a machine learning algorithm because of its suitability to an integrated development environment (IDE). There are many IDEs, either free or licensed versions, such as Visual Studio Code, Visual Studio, Atom, Sublime Text, Spyder, and PyCharm. Most IDEs in the market are available for Linux, Windows, and Mac OS.

Microsoft created Visual Studio Code or VS Code in 2015; VS Code is an open-source, and free version. Additionally, it supports many languages, such as Python, C#, C++, JavaScript, HTML, and CSS. VS Code is a fast and lightweight while Visual Studio (VS) is a heavy-weight IDE and it does not support for Linux. Furthermore, VS has both unpaid versions and paid versions; the unpaid version normally supports only small teams with open-source projects. Atom, Sublime Text, and Spyder are equivalent to VS Code because they are free, fast, and lightweight.

Besides the software mentioned above, this research uses the PyCharm Community Edition (CE) version for the development of the algorithm and system. Further, PyCharm CE version is an open-source and free version like the IDEs above. Moreover, PyCharm CE also has the Jupyter Notebook as the IDE to combine all libraries for development. The characteristics of the Jupyter Notebook is interactive computing that can display output step by step. Consequently, programmers can check the results line-by-line.

2.4.2 Libraries

This research uses two libraries for facial detection as OpenCV and Dlib. OpenCV summarizes from Open Source Computer Vision Library, is an open-source for image detection such as facial recognition, or object detection. It can use C++ or Python language for development with all libraries from a pipeline for extraction feature. Dlib is a modern toolkit for image processing library with Python or C++ language. It is applied in robotics, embedded devices, or smartphone. The primary function provides features for detecting point on the face and extracting a feature from the landmark. Moreover, it supports the development of ML or DL algorithm. Thus, this research employs Dlib to mark spots on the face for facial detection in a normal situation and abnormal situation.

2.4.3 Python libraries

Development of machine learning algorithms contains many modules for ML and DL; the favorite modules is Keras, TensorFlow, and Scikit-learn. Keras is one of Python library to support neural network model and to prepare vital portable models. TensorFlow was developed by the Brain Team from Google TensorFlow. TensorFlow is suitable for neural network model like Keras; this can train dataset as N -dimensional matrices. Thus, TensorFlow can process a large amount of data with the TensorFlow GPU core.

The last favorite module is Scikit-learn, that is suitable for complex data. Moreover, Scikit-learn has several features to support many ML algorithms, such as nearest neighbors, logistic regression, random forest, decision tree, and k -Means clustering. Furthermore, the new version of Scikit-learn was added with new functions for cross-validation features and N -dimensional matrices in the case of a complexed dataset. Thus, this research selects Scikit-learn python library to develop ML algorithm based the decision tree model.

2.4.4 Software tool for data science and machine learning

There are various software tools for data science and ML in the market. The two interesting tools are Weka, and RapidMiner Studio because both Weka and RapidMiner Studio are easy to use. After defining the input data for software, users can select a model for training and testing. However, Weka has limitation in displaying the result and model optimization while RapidMiner Studio is much powerful. RapidMiner Studio offers a visual workflow designer, output, and graph, including various optimization techniques. Thus, the researchers mostly choose the RapidMiner Studio for investigation ML model.

Since there are various tools have been mentioned for creating a classification model, this research starts using the ML model to derive a classification model, then validate the obtained model by comparing with a model obtained from the execution of Scikit-learn. The expected outcome for this process is that these two models should be the same, otherwise, an error checking must be performed at the Scikit-learn process.

2.5 Techniques

In a decade, there are many researchers who applied the facial action coding system (FACs) and ML in several areas, such as security, behavior analysis, animation, and healthcare. Some research uses facial expression recognition for identification and verification from an image in the field of security (Shreve et al., 2017). Moreover, facial expression recognition is one solution of a smart city project (Sajjad et al., 2019). In the behavior analysis, an automatic real-time emotion recognition system with facial image analysis to monitor human behavior was adopted to combat autism and reduce traffic fatalities (Kundu & Saravanan, 2017). Furthermore, facial expression using FACs was applied to use in an Avatar by marking points on the face, such as eyes, lips, and eyebrows. This research is to study the six basics emotion like angry, happy, sad, scared, and surprise for define of emotion expression on the face of the Avatar (Puklavage et al., 2010).

In the healthcare area, there are many researchers who studied stroke symptom detection methods or facial palsy detections, including the development of an automatic stroke detection system. Since most stroke patients have speechless and paralysis, treatment monitoring is essential for patient safety. Therefore, these researchers tend to investigate detection methods using facial features according to the facial expression because the stroke patients can still express emotions via their faces. For example, some researchers investigate facial features about facial movement from facial paralysis to nerve damage using Active Shape Models Plus Local Binary Patterns (ASMLB). The ASMLB refers to splitting and shaping on the face as eyebrows, eyes, nose, cheek, and mouth. Then, all items on the face is used to seek for pattern movement. This research uses the facial images from the patients with the equipment set for photography. Nonetheless, the patients need to take a seat while photography, this might uncomfortable for stroke patents. The accuracy for facial classification between paralyzed face and normal face as 93.95% and this algorithm cannot diagnosis of facial paralysis (Wang et al., 2014).

In 2018, Hsu et al., are to study detection of facial palsy syndrome from eyes and mouth using Hierarchical Detection Network (HDN). The HDN composes of three components: facial detection, facial landmark detection, and palsy region detection. The first component is the facial detection, it will create rectangle area on the face. Then, the facial landmark detection is used to be the references to implement

the grid of 8 x 8 cells, each cell is applied to detect eyes and mouth called class eyes and class mouth. The last component is palsy region detection from class eyes and class mouth for classification of paralyzed face and normal face. The accuracy is 90.50% (Hsu et al., 2018). Next, Hsu et al., is to investigate the analysis of facial palsy or bell's palsy with facial detection in 2019. The research proposed Deep Hierarchical Network (DHN) based on CNN. This technique will be divided into two sub-networks, Line Segment Network (LSN) and Double Dropout Network (DDN) after facial landmark detection. LSN will be estimated the line segments to connect facial landmarks and DDN will be located the facial landmarks. In this research, Hsu et al., increased one variable as nose from the previous research in 2018 and proposed new techniques. The accuracy in this research is 91.20% (Hsu et al., 2019).

Umirzakova and Whangbo had been studying stroke symptom detection using facial features in 2018. This research uses to detect stroke symptoms from forehead wrinkle, eye pupil, cheek area, and mouth using Active Appearance Model (AAM). This algorithm will detect object shape from color or intensity of the images, then the algorithm will create the new images for training. The accuracy of the experiment is 91.00 % (Umirzakova and Whangbo, 2018). Next, these researchers improved the detection accuracy for predicting the stroke symptoms. They selected only one facial feature from the previous research; this selection is the cheek area. In addition, they also use hand moving to assist the detection. This research still applied to use the AAM for cheek wrinkle line detection, and motion detection algorithm and skin model are used to detect hand moving. The accuracy is 92.00% (Umirzakova and Whangbo, 2019). The research in 2019 is not use forehead wrinkle, eye pupil, and mouth because these features cannot express stroke symptoms obviously. Although they were selected cheek area, it still is not classifying the symptoms precisely.

Table 13 Comparison

No.	Related Works	Techniques	Features	Outputs	Accuracy
1	Wang et al., 2014	ASMLBP	Upper and Lower Eyebrows, Interpalpebral Fissure, Nose Region, Cheek Lines, Lip Part	Classification between paralyzed face and normal face	93.95 %
2	Hsu et al., 2018	HDN	Edge of Eyes, Edge of Lip	Classification between paralyzed face and normal face	90.50 %
3	Hsu et al., 2019	DHN	Edge of Eyes, Edge of Lip	Classification between paralyzed face and normal face	91.20 %
4	Umirzakova et	AAM	Eye pupil	Classification	91.00 %

No.	Related Works	Techniques	Features	Outputs	Accuracy
	al., 2018		distance between right to left, Cheek Lines, Lip Part	between normal person and stroke person	
5	Umirzakova et al., 2019	AAM	Cheek Lines,	Classification between normal person and stroke person	92.00 %

From the list of research mentioned above, most researchers were intended to study the facial features for classifying between a normal person and a stroke person. Nevertheless, the stroke person is usually unable to move their body. Thus, the body movement cannot use for detecting the stroke patients since most stroke patients have trouble in speaking and moving their bodies. Though, there are researchers using the severity level detection of stroke patients from the facial images via the equipment set of photography, these researchers cannot classify each severity level clearly. Moreover, some researchers have to detect facial images of the patients from the photography equipment while sit; unlikely that this posture is uncomfortable for the strokes.

Some facial features in these researches are not be adequate for classification severity level symptoms or some facial features do not need for detection symptoms. Therefore, the researcher of this dissertation aims to investigate essential facial features for situation between normal situation and abnormal situation of the stroke patients. Thus, the expected result of this dissertation can be assisted stroke patients to receive safety while treatment.

CHAPTER III

METHODOLOGY AND ANALYSIS

This chapter explains all methodologies and analysis of the research; Section 3.1 describes the research procedures, including the preliminary study process. Then, the materials are described in Section 3.2 which consist of hardware components and software components. Furthermore, the preliminary study and the experiment are described in Section 3.3 and Section 3.4, respectively. Finally, the analysis method is presented in Section 3.5.

3.1 Research Procedure

This research is designed using the prospective analytic study. The research procedure consists of twelve steps, listed below.

1. Review literatures related to facial feature extraction of the eye detection and the facial detection mechanisms.
2. The ethics review committee for research involving human research, Chulalongkorn University and research affairs of faculty of medicine, Chulalongkorn University.
3. Classify types of facial features from the feature extraction technique using eyes, brow, nose, and lip.
4. Determine detection variables using types of facial features related to the defined types.
5. This research will randomly select 60 samples for preliminary study from the Internet Movies Database (IMDb) websites to obtain the expression patterns from palpebral region features and facial features. Then, the experiment will randomly select 10 samples from Chulalongkorn Comprehensive Stroke Center of Excellence, King Chulalongkorn Memorial Hospital.
6. Establish an expression-classification model from the collected samples.
7. Measure the accuracy of the model using true positive (TP), true negative (TN), false positive (FP), false negative (FN), accuracy, classification error, and running time.
8. Develop a facial feature detection software prototype, so called as the expression classification software, based on the expression-classification model.
9. Verify the prototype with the previous volunteer group.
10. Refining the prototype in order to obtain higher accuracy of the detection.
11. Validate the expression classification software prototype using the learning system, PyCharm Community edition with Scikit-learn, for acceptance checking.
12. Compare the classification performance, and classification error of the proposed technique with other existing techniques using the receiver operating characteristics (ROCs).

3.2 Material

The environment of the expression classification software consists of hardware and software components as showed in Figure 15. Details of each component are elaborated below.

1. Hardware Components

The emotional classification system composes of the Raspberry Pi 3 Model B board, Raspberry Pi camera module, a personal computer, and connecting with an intranet. Specifications of each component are described as follow.

1.1 Raspberry Pi 3 Model B board

- 1.1.1 Qu Quad Core 1.2GHz Broadcom BCM2837 64bit CPU
- 1.1.2 1 GB RAM
- 1.1.3 BCM43438 wireless LAN and Bluetooth Low Energy (BLE) on board
- 1.1.4 100 Base Ethernet
- 1.1.5 40-pin extended GPIO
- 1.1.6 4 USB 2 ports
- 1.1.7 4 Pole stereo output and composite video port
- 1.1.8 Full size HDMI
- 1.1.9 CSI camera port for connecting a Raspberry Pi camera
- 1.1.10 DSI display port for connecting a Raspberry Pi touchscreen display
- 1.1.11 Micro SD port for loading your operating system and storing data
- 1.1.12 Upgraded switched Micro USB power source up to 2.5A

1.2 Raspberry Pi camera module

- 1.2.1 Night vision camera
- 1.2.2 Support up to 2 infrared LED and fill flash
- 1.2.3 5 mega pixel OV5647 sensor
- 1.2.4 Camera dimension: 25 mm x 24 mm
 - 1.2.4.1 CCD size: ¼ inch
 - 1.2.4.2 Aperture (F): 1.8
 - 1.2.4.3 Focal length: 3.6 mm (adjustable)
 - 1.2.4.4 Diagonal: 75.7 degree
 - 1.2.4.5 Sensor best resolution: 1080p

1.3 Personal Computer

- 1.3.1 Quad Core Intel
- 1.3.2 ECC main memory supports to error checking code at least 8 Giga bytes
- 1.3.3 SSD hard drive at least 4 Tera bytes
- 1.3.4 Support network interface card dual port Gigabit Ethernet

1.4 Intranet System

- 1.4.1 The intranet supports 10/100/1000 Mbps communication connection speeds from all local places.

2. Software Components

This software will be developed using PyCharm Community Edition (CE) version 2018.2 based on OS X 10.15.2, with Python language. In addition, an open source library called DLIB and OpenCV library is embedded to the PyCharm Community Edition (CE) so that the software can communicate with the RPi3 Model B with the RPi camera module in order to detect the face.

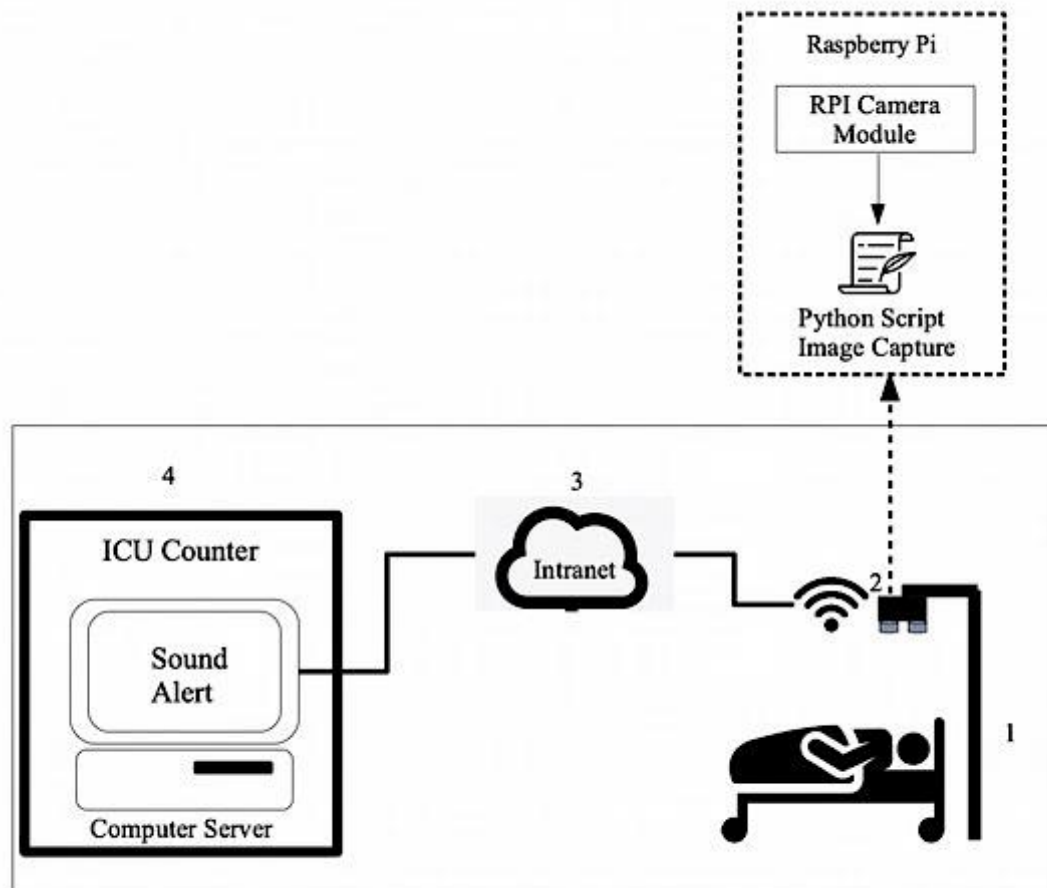


Figure 15 Infrastructure of Facial Recognition Alert System (FRAS)

From Figure 15, a camera-detection set consists of the camera's stand (#1) and a Raspberry Pi 3 (RPi3) Model B with Raspberry Pi (RPi) camera module (#2). The RPi3 Model B with the RPi camera module is responsible for capturing images from the face of patients in every movement. Moreover, the RPi3 Model B board embedded python script for reading 68 facial landmark points. Then, the landmark point values will be sent to a computer server (#4) for state analysis via the intranet system (#3). The alarm will be activated when an abnormal situation is detected. Moreover, this infrastructure was published by the Department of Intellectual Property, Thailand (Bhattarakosol et al., 2019). However, the alarm will be developed in the future.

3.3 Preliminary Study

This section presents the external elements of eyes and the sub-structures of face that are used as the proposed classification model with the decision tree algorithm.

3.3.1. External Elements of Eyes

Elements of eyes can be classified to the internal elements and the external elements. The internal elements refer to components in the eyes, such as the iris, cornea, pupil, retina, and lens. Most researches (Hansen et al., 2010 & Zafeiriou et al., 2015) applied the internal elements with specific equipment to measure their changes. In contrast, the preliminary study of this research is focused on the change of the external elements of eyes, which are the interpalpebral fissure (IPF), the palpebral fissure length (PFL), and the palpebral fissure region (PFR). From Figure 16, the IPF is the vertical distance between the palpebral superius point (PS) and the palpebral inferius point (PI); and, the PFL is the horizontal distance between the endocanthion (EN) and the exocanthion (EX) as is presented in Figure 16. Additionally, the IPF and the PFL are applied to calculate the PFR; the formula is presented in Eq.(3.1). Furthermore, these elements were performed by Facial Landmark Detection (FLD) with the Dlib library for marking positions on the image as displayed in Figure 17.

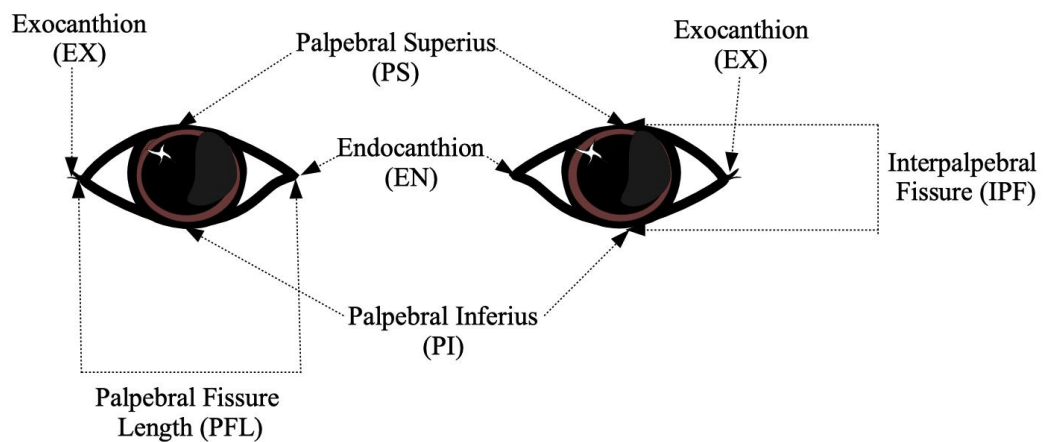


Figure 16 The External Elements

$$PFR = (\pi/4) * IPF * PFL \quad \text{-----} \quad (3.1)$$

As mentioned above, in the preliminary study period, all images are captured from horror-thriller-murder movies, and the frontal face is the only object to be used.

After the fear images are captured, these images are inserted to the developed application to detect and extract the face. The application was developed using PyCharm version CE editor, with Python language. In addition, an open source library, called DLIB library and OpenCV library, was embedded into the PyCharm CE. Consider an image, one image normally consists of one frontal face, background, and other objects. Among these components, the most important object is the frontal face; and, this will be captured by the DLIB library and OpenCV library, displayed in Figure 17. However, the size of this captured image has no standard size or the real size of the human face; the real size of a face is defined in (Poston, 2000). So, the size of the frontal face is adjusted to be the real size and this resized face is added with 68 facial mark-points on the face. As a result, the new face-image is normalized and adjusted to 800x800 pixels. After marking the face with 68 mark-points, the PI, PS, EX and EN points are defined. Then, the IPF and the PFL can be calculated, following with the PFR value. Then, the PFR in the normal state is compared with the PFR in the fear state using paired t -test with significant value equal to 0.05. Finally, the emotion either neutral or fear can be identified according to the value of PFR. Processes of as presented in Figure 18. Diagram in Figure 18 shows steps of the External Elements of Eyes to be detected, measured, and classified.

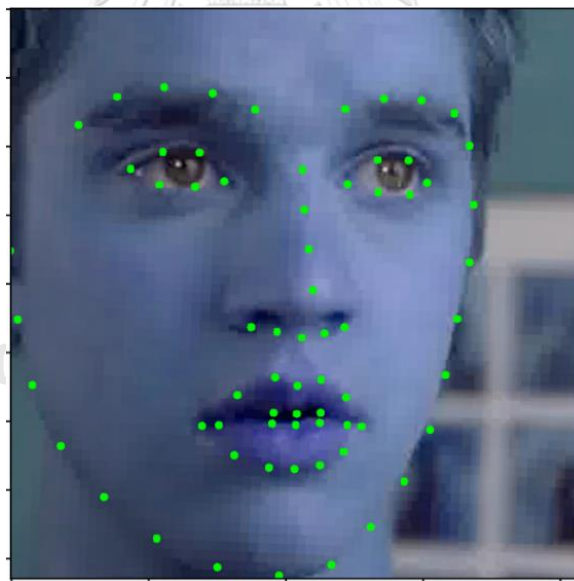


Figure 17 Example of Mark point with Facial Landmark Detection (FLD) of Phase I

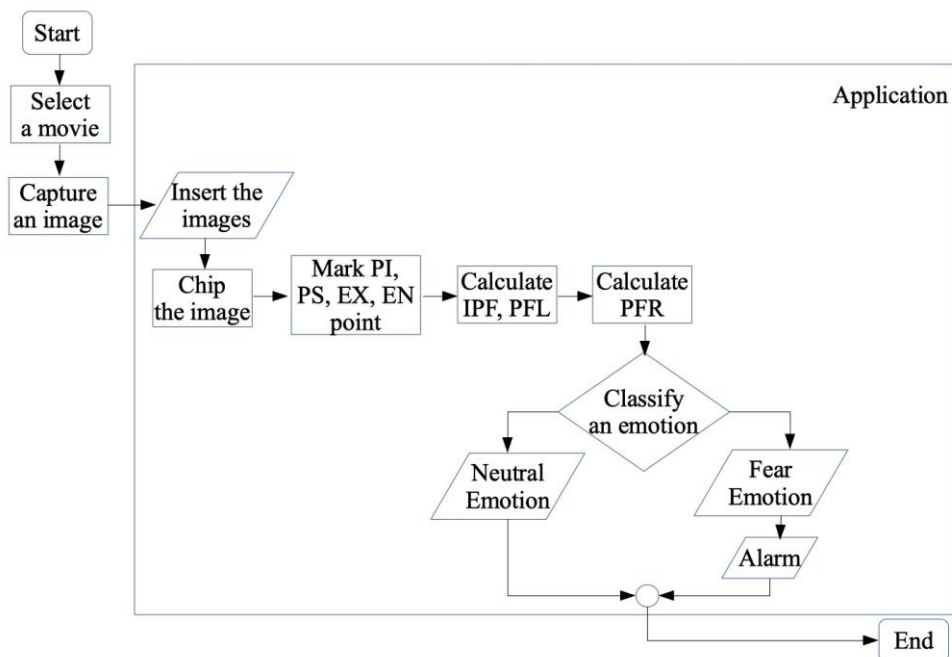


Figure 18 Diagram of the Emotion Classification of the External Elements

3.3.1.1 Data Collection

From 33 selected videos, there are only two states to be chosen from a performer: neutral face state, and fear face state. The movies are collected from the IMDb website, and they are ranked by experts and leading critics. The collected movies must have high rankings, be in high definition (HD) and have qualified players. This experiment used 60 samples with the age range of 18-52 years. Each sample is selected six scenes from the movies, three scenes of the neutral emotion and three scenes of the fear emotion. Thus, the total number of images is 360 pieces, and the duration of data collection is two months.

3.3.2. Sub-Structures of Faces

Considering a human face as a structure of the face which consists of forehead, two eyes, one nose shape, one cheek, one chin, and one mouth (Takalkar et al., 2018). In addition, each structure contains sub-structures as shown in Table 14.

From Table 14, when the structures are altered, the sub-structures are also altered. Thus, the second step of this preliminary study is to find the differences among sub-structures, such as the inner brow raiser (IBR), brow lower (BL), lid raiser (LR), and lip part (LP). To achieve this objective, this research applied 68 facial landmarks to detect these sub-structures of the face, as presented Figure 19.

Table 14 Structure and Sub-Structure of the Face

Structures	Sub-Structures
Forehead	Eyebrows, Ear
Eyes	Eyelid, Eyelash
Nose Shape	Nose
Cheek	Cheek, Chin
Mouth	Inner Lip, Outer Lip

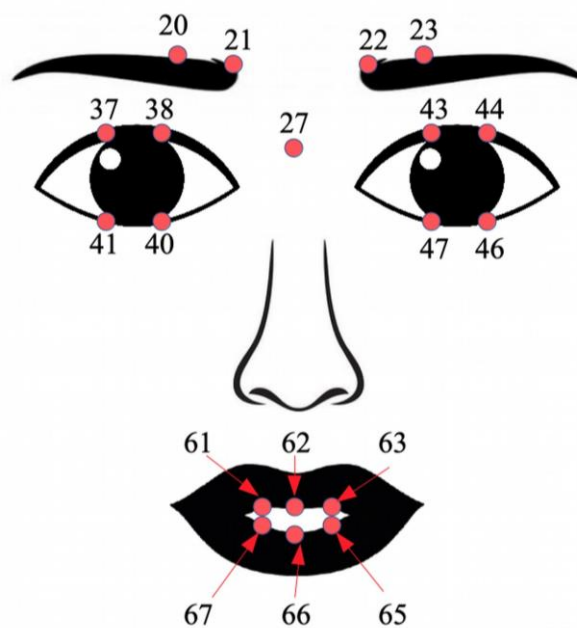


Figure 19 Sub-Structures of the Face

Similarly, to the external element of eyes, all frontal face images were captured from the horror-thriller movies, and these images must be reformed size to the particular 800x800 pixels or called as the real scale size of the face images (Poston, 2000). After re-sizing, 68 facial landmarks were pointed over a piece of the facial image followed by the sub-structure of the face, displayed in Figure 20. Moreover, DLIB library and OpenCV library are also used to locate the facial landmarks on the IBR, BL, LR, and LP. Then, the IBR is calculated using the number pair (20, 27) of the right eyebrow and the number pair (23, 27) of the left eyebrow. The BL is the distance between the position number 21 to the position number 22. This LR contains 4 distance-values; those are computed from the position number 37 to the position number 41, and the position number 38 to the position number 40 of the right eye, including the position numbers 43 to 47, and the position number 44 to the position number 46 of the left eye. Like the LR, the LP value, an essential structure for the change of the lip's shape, composed of six positions: 61, 62, 63, 64, 65, and 66. Then, all six positions are paired and the length of each pair is calculated;

the order pairs of these positions are (61, 65), (62, 66), and (63, 67). Finally, the emotion classification, both normal and fear, can be identified according to the value of the IBR, BL, LR, and LP as presented in Figure 21.

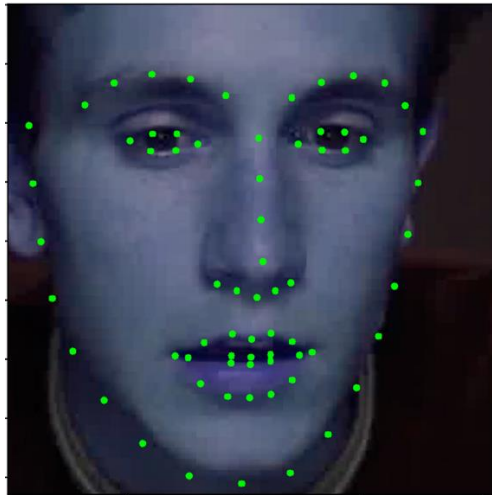


Figure 20 Example of Mark point with Facial Landmark Detection (FLD) of Phase II

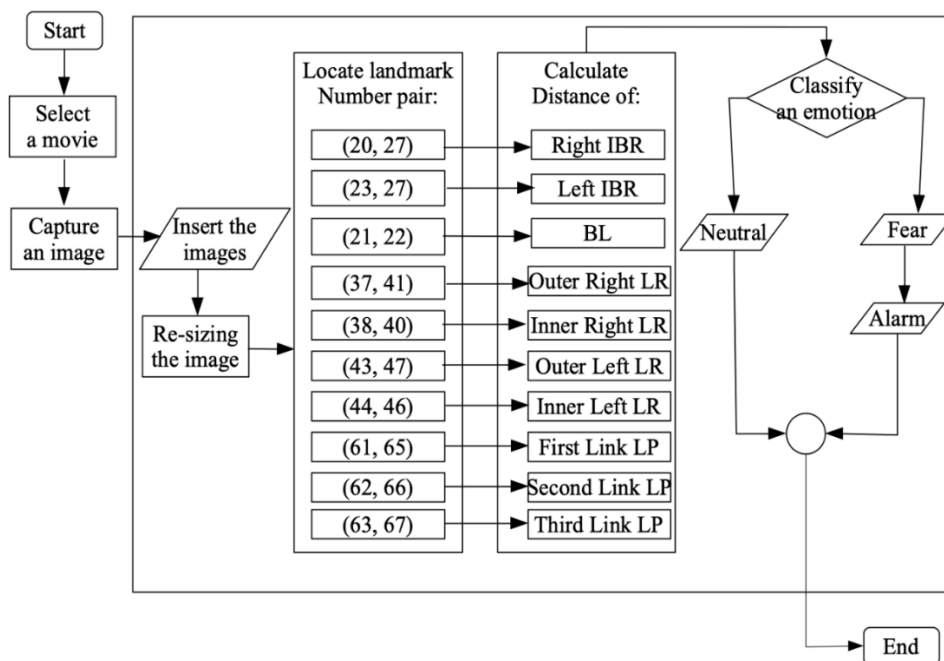


Figure 21 Diagram of the Emotion Classification of the Sub-Structures

3.3.2.1 Data Collection

From 23 selected films, there are two situations from 20 samples as neutral situation and fear situation. The movies are collected from the Internet Movie Database (IMDb), an online database movie website. The experiment for this preliminary study used 360 images from 60 samples with the age range of 18-61 years. As same as the previous preliminary study, each sample uses six images from the cinemas, three scenes of the normal situation and three scenes of the fear situation. Thus, the total number of images is 360 pictures, and the duration of data collection is two months.

3.4 Experiment

From the preliminary study, the external elements of eyes and the sub-structures of the face are applied to detect a facial expression, both normal situation and abnormal situation.

3.4.1 Data Collection

From the preliminary study, the external elements of eyes and the sub-structures of the face are applied to detect a facial expression, both normal situation and abnormal situation.

Similarly, to the first phase and the second phase, all frontal face images were captures from the participations as stroke patients from Chulalongkorn Comprehensive Stroke Center of Excellent, King Chulalongkorn Memorial Hospital. Then, DLIB library and OpenCV library are also used to locate the facial landmarks on the IPF, PFL, PFR, IBR, BL, LR, and LP as displayed in Figure 22.



Figure 22 Example of Mark point with Facial Landmark Detection (FLD) of Phase III

3.5 Analysis Method

Since there are the preliminary study and the experiment, the external elements of the preliminary study are to collect all information such as gender, age, and images of 60 samples from 33 movies. Moreover, the sub-structures of the face are to accumulate gender, age, and pictures of 60 samples from 23 films. All films selected images from the internet movie database (IMDb) website; it is high rankings and high definition (HD). In the experiment, there is collect 10 participants Chulalongkorn Comprehensive Stroke Center of Excellence, King Chulalongkorn Memorial Hospital, through a stroke expert. Furthermore, this collection has requested an ethics review committee for research involving human research, Chulalongkorn University, and research affairs of faculty of medicine, Chulalongkorn University.

3.5.1 Statistical Method

According to the external elements of the preliminary study, all retrieved images have calculated the distance of the IPF, PFL, and PFR. All external elements led to determine means including gender and age. The mean values of all data are compared to the difference between neutral emotion and fear emotion using a paired *t*-test with a 95% confidence level.

The sub-structures of the face are used to identify the normal emotion and fear emotion based on the average values of IBR, BL, LR, and LP. However, all average values were tested with the bell curve as a non-normal distribution. Thus, the non-parametric Wilcoxon signed ranks test, is deployed to compare the differences

between the neutral emotion and the fearful emotion. The Wilcoxon signed ranks test is performed, employing a 95% confidence level or the 0.05 significance level.

The external elements and the sub-structures of the face are used to identify the normal situation and abnormal situation based on IPF, PFL, PFR, IBR, BL, LR, and LP. However, all average values were tested with the bell curve as a non-normal distribution. Thus, the non-parametric Wilcoxon signed ranks test, is deployed to compare the differences between the normal situation and the abnormal situation. The Wilcoxon signed ranks test is performed, employing a 95% confidence level or the 0.05 significance level.

3.5.2 Machine Learning Mechanism

After the statistical method, all data are divided into two sets: the training set, and the testing set. The training set is used in the pattern creation, and the testing set is applied for the classification and analysis processes. These sets were splitting by a simple random sampling method using the proportion of 7:3. Then, there are select the parameters and create the emotion classification model. Then, this model led to creating a decision tree algorithm. After that, the performance of the emotion classification model is measure depends on four values: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). These values are used to find Precision and Recall as follow in Eq (3.2), Eq (3.3), and Eq (3.4).

$$Precision = \frac{TP}{TP+FP} \text{-----}(3.2)$$

$$Recall = \frac{TP}{TP+FN} \text{-----}(3.3)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \text{-----}(3.4)$$

CHAPTER IV

RESULT

This chapter elaborates the parameter selection results and model derivation using both statistical methods and machine learning mechanisms. In Section 4.1 to Section 4.2 describes the outcome of parameter selection techniques, then Section 4.3 is the final model of this study after applying all parameters to identify the fear emotion.

4.1 Parameter Selection: First Phase

This phase is to study and find parameters of the external elements of eyes, such as the Interpalpebral Fissure (IPF), Palpebral Fissure Length (PFL), and Palpebral Fissure Region (PFR), including gender and age. From the training set, 40 samples are 47.5% male and 52.5% female. Additionally, the highest age range is between 18-24 years old, 32.5% of the entire samples. The statistical testing is the paired samples t -test to determine the difference of mean values and standard deviation values of neutral emotion and fear emotion, under 0.05 significant level or 95% confidence level, as detailed below.

4.1.1 Statistical Results

These results consist of determining the difference of mean values and standard deviation of the IPF, PFL, and PFR.

4.1.1.1. Test of the IPF means and standard deviations differences

The IPF is tested by pair t -test, with a 95% confidence level. The result shows at least one significant difference between the IPF means of the right eyes under the neutral and the fear emotions, p -value = $0.003 < 0.05$. Moreover, the left eyes also have significant means different under two emotions with p -value = $0.001 < 0.05$. Besides, the values of the IPF in the fear situation are bigger than the values of the IPF in the neutral situation in both eyes. Furthermore, the percentage of alter in the IPF means in females is lower than the percentage of change in the IPF means in males. The IPF means with standard deviations of both eyes are displayed in Table 15.

Table 15 The IPF means with standard deviations (millimeters)

Emotion	IPF	Gender		Total Mean
		Male	Female	
Normal	Right	8.83 ± 1.22	10.21 ± 0.99	9.52 ± 0.98
	Left	8.76 ± 1.24	10.30 ± 1.09	9.53 ± 1.09
Fear	Right	10.45 ± 1.24	11.35 ± 1.05	10.90 ± 0.64
	Left	10.50 ± 1.33	11.34 ± 1.16	10.92 ± 0.60

4.1.1.2. Test of the PFL means and standard deviations differences

The result of the t -test value when compared between the PFL means of the right eyes under neutral and fear emotions has p -value = $0.046 < 0.05$. Furthermore, the percentage of change in the PFL means in the right eyes is higher than the percentage of change in the PFL means in the left eyes. The PFL means with standard deviations of both eyes are shown in Table 16.

Table 16 The PFL means with standard deviations (millimeters)

Emotion	PFL	Gender		Total Mean
		Male	Female	
Normal	Right	29.80 ± 1.88	30.98 ± 2.26	30.39 ± 0.84
	Left	30.75 ± 2.48	31.21 ± 1.18	30.98 ± 0.28
Fear	Right	30.50 ± 2.27	30.90 ± 1.53	30.70 ± 0.64
	Left	30.37 ± 1.54	30.16 ± 2.21	30.26 ± 0.15

4.1.1.3. Test of the PFR means and standard deviations differences

The result of the t -test value when compared between the PFR means are the same as the result obtained in the IPF means of differences. Furthermore, the PFR means in the fear state are more significant than the PFR means in the neutral state for both genders. Additionally, the percentage of change in males' eyes is much bigger than the percentage of change in females' eyes. The PFR means with standard deviations are drawn in Table 17.

Table 17 The PFR means with standard deviations (millimeters)

Emotion	PFR	Gender		Total Mean
		Male	Female	
Normal	Right	207.06 ± 33.79	248.84 ± 33.12	227.95 ± 29.54
	Left	212.17 ± 37.95	252.64 ± 30.47	232.41 ± 28.62
Fear	Right	251.39 ± 41.85	276.12 ± 34.62	263.76 ± 17.49
	Left	250.80 ± 37.01	269.68 ± 41.65	260.24 ± 13.35

According to three simple t -test, it can be summarized that the sizes of eyes in the fear state are larger or wider than the sizes of eyes in the normal state. In addition, the males' eyes usually change more than the females' eyes when they are scared. Therefore, it could say that genders relate to emotions and also influence to the changes of the mean values of the IPF, the PFL, and the PFR. Thus, when the mean values of any indicators have been altered to be larger than normal, there is a possibility that the person is in the scary mode. The relationship among these 3 parameters can be drawn in Figure 23.

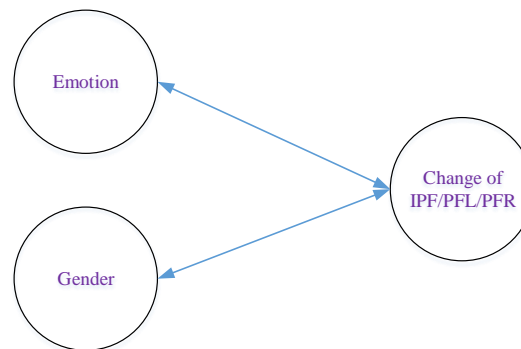


Figure 23 Relation among Gender, Emotion, IPF, PFL, and PFR

4.1.2 Choosing Parameters

Based on the statistical results, there are three important parameters to be considered as classifier factors: gender, IPF, and PFR. However, the values of IPF and PFR are depended on each sample that has individual characteristics, such as gender, age-range, status of emotion, and both eyes' sizes. Thus, the variable IPF and PFR can be counted as the dependent variables while the sample's characteristics can be specified as independent variables. The emotion classification model that is derived from some of these variables should be optimized, i.e. uses as small number as possible in the model but provides high accuracy in the classification result. Since there are various variables obtained from every samples: gender, age-range, status of emotion, and eye's location. The values of these variables are displayed in Table 18. Furthermore, Figure 24 shows relationships among these parameters; the red dash line represents the expected effect of factors towards an emotion. In addition, the relationship between gender and emotion, and the relationship between age and emotion can be determined by calculating the correlation.

Table 18 Values of variables in this experiment

Independent Variable	Possible values
gender	Male, Female
age	18-24, 25-31, 32-38, 39-45, > 45
location	Right, Left
IPF	[8.83, 11.34]
PFL	[29.80, 30.98]
PFR	[212.17, 276.12]
Dependent Variable	Possible values
status	Neutral, Fear

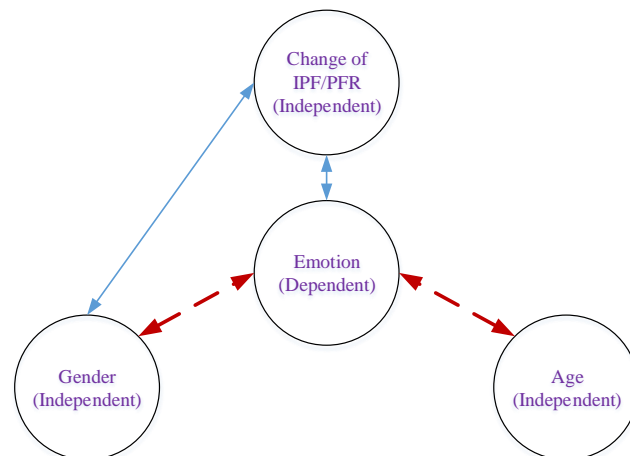


Figure 24 Expected relationships among independent and dependent variables

Since this phase classified IPF and PFR as the independent variables, the selected dependent variables should be able to identify the dissimilarity between neutral and fear emotion precisely. To achieve this requirement, statistical testing must be performed to determine the significant difference of IPF or PFR between neutral and fear using all these independent variables. Although there are various factors that are related to both dependent variables, the created classification model must be optimized and implemented by the smallest numbers of factors to maintain simplicity of the classification modeling. Therefore, statistical analysis to identify factors and their impacts towards the values of IPF or PFR was performed. Nevertheless, the distributions of IPF and PFR were determined and found out as multimodal distribution. Therefore, the non-parametric analysis, Kruskal-Wallis test, is applied using 95% confident level.

4.1.2.1 IPF consideration

The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effect to the IPF mean values.
- Different genders, male and female, have different effects to the IPF mean values.
- Different emotions, neutral and fear, have different effects to the IPF mean values.
- Different eye-sides, right eye and left eye, have different effects to the IPF mean values.

The results of these basic testing found that there are two variables that do not have affect to the mean of IPF: the variables of age-range and eye-sides, with p -values $> 0.05 = \alpha$. On the other hand, the variables of genders and emotions can categorize the differences of the IPF mean values, with p -values $< 0.05 = \alpha$.

The next step is to identify the impacts of interactions between these factors. The interactions between factors are considered based on combinations of levels among different factors. For example, the interactions between emotions and genders can be derived into 4 combinations of their levels: male with neutral emotion, female with neutral emotion, male with fear, and female with fear. These interactions can be marked as different groups of treatments, then, Kruskal-Wallis test is applied. As four factors are considered, assumptions to determine the interactions among factors can be derived to 11 assumptions as listed below.

- At least one combination between emotions and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination between emotions and gender has significant different from others towards the mean values of the IPF.
- At least one combination between eye-sides and genders has significant different from others towards the mean values of the IPF.
- At least one combination between emotions and ages has significant different from others towards the mean values of the IPF.
- At least one combination between genders and ages has significant different from others towards the mean values of the IPF.
- At least one combination between eye-sides and ages has significant different from others towards the mean values of the IPF.
- At least one combination among genders, ages, and emotions has significant different from others towards the mean values of the IPF.
- At least one combination among genders, ages, and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination among emotion, ages, and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination among genders, emotions, and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination among genders, ages, emotions, and eye-sides has significant different from others towards the mean values of the IPF.

From the assumption above, the testing results discover that only the combination between ages, and eye-sides has no significant difference in the mean values of the IPF, with $p\text{-value} = 0.643 > 0.05 = \alpha$; or, the combination between ages, and eye-sides has no effect to the mean values of the IPF.

As a consequence of the statistical testing shown above, there are two variables that cannot influence to the change of the IPF values, i.e. ages, and eye-sides. Although these two factors cause effects when combining with other factors, they can be eliminated from the classification model in order to obtain the model simplification.

4.1.2.2 PFL Consideration

Although the PFL is not the appropriate variable for the emotion classification from the PFL-test of differences, the value of the PFL is used to compute the PFR

values that are used to indicate an emotion. So, it can say that the PFL values are used to identify an emotion with any other values. Thus, the next process is to analyze the value of the PFL using the non-parametric analysis, Kruskal-Wallis test, with 95% confident level.

The basic testing assumptions are listed as follows.

- At least one level of age-range has different effects to the PFL mean values.
- Different genders, male and female, have different effects to the PFL mean values.
- Different emotions, neutral and fear, have different effects to the PFL mean values.
- Different eye-sides, right and left eyes, have different effects to the PFL mean values.

The results of these fundamental testing found that there are three variables that do not have effect to the PFL: age-range, eye-sides, and emotions, with $p\text{-value} > 0.05 = \alpha$. On the contrary, another variable that can categorize the differences of the PFL mean values is genders, with $p\text{-values} < 0.05 = \alpha$. Therefore, genders are used to combine with others where the interactions among factors are listed below.

- At least one combination between genders and emotions has significant different from others towards the mean values of the PFL.
- At least one combination among genders, ages, and emotions has significant different from others towards the mean values of the PFL.
- At least one combination among genders, ages, and eye-sides has significant different from others towards the mean values of the PFL.
- At least one combination among genders, ages, emotions, and eye-sides has significant different from others towards the mean values of the PFL.

According to all assumptions above, the statistical results indicated that there is no significant difference among the mean values of PFL under combination of genders, and eye-sides, with $p\text{-value} = 0.106 > 0.05 = \alpha$; and there is no significant difference among the mean values of PFL under combination of genders, emotions, and eye-sides, with $p\text{-value} = 0.135 > 0.05 = \alpha$. So, the combination of genders, and eye-sides and the combination of genders, emotions, and eye-sides have no effect to the mean values of the PFL.

Though, there is no significant difference among mean values of PFL under combination of genders, emotions, and eye-sides, as shown above, these combinations still need to be applied to the next step for the PFR calculation. Thus, the second classification model also contain variables, genders, ages, emotions, and eye-sides.

4.1.2.3 PFR Consideration

The last process is to analyze the PFR value using non-parametric analysis, Kruskal-Wallis test, with 95% confident level.

The basic testing assumptions are listed as follows.

- At least one level of age-range has different effects to the PFR mean values.
- Different genders, male and female, have different effects to the PFR mean values.
- Different emotions, neutral and fear, have different effects to the PFR mean values.
- Different eye-sides, right eye and left eye, have different effects to the PFR mean values.

The results of these basic testing found that the eye-side is the only one variable that does not have effect to the mean of PFR, with p -values $> 0.05 = \alpha$. On the other hand, combination among genders, ages, and emotions can differentiate the PFR mean values, with p -values $< 0.05 = \alpha$.

The assumption can be derived to 10 assumptions as listed below.

- At least one combination between genders and ages has significant different from others towards the mean values of the PFR.
- At least one combination between genders and emotions has significant different from others towards the mean values of the PFR.
- At least one combination between genders and eye-sides has significant different from others towards the mean values of the PFR.
- At least one combination between ages and emotions has significant different from others towards the mean values of the PFR.
- At least one combination between emotions and eye-sides has significant different from others towards the mean values of the PFR.
- At least one combination among genders, ages, and emotions has significant different from others towards the mean values of the PFR.
- At least one combination among genders, ages, and eye-sides has significant different from others towards the mean values of the PFR.
- At least one combination among emotion, ages, and eye-sides has significant different from others towards the mean values of the PFR.
- At least one combination among genders, ages, and emotions has significant different from others towards the mean values of the PFR.
- At least one combination among genders, ages, emotions, and eye-sides has significant different from others towards the mean values of the PFR.

The results of these interaction testing discover that all combinations between ages, and eye-sides has no effect to the mean values of the PFR, or there is no significant different of mean values of the PFR, with p -value $= 0.234 > 0.05 = \alpha$.

Although the eye-side variable cannot influence the change of the PFR values, this variable is still important to use with other factors. Thus, the third classification model includes genders, ages, emotions, and eye-sides.

Referring to all statistical results mention previously, it is clear that there are relationships among parameters in Table 18; the relationship diagram can be illustrated in Figure 25.

According to the relationship diagram in Figure 25, there is only 2 direct links toward emotion. One link is the link of correlation between change of the IPF/ PFR and emotion; the second link is the link of correlation between emotion and interaction of gender and age. However, there are direct links of correlations between some interactions with gender, or eye-side, or age. So, it can imply that there are some indirect correlations between gender with emotion, eye-side with emotion, and age with emotion by passing through the existing links of correlations. Thus, it is the right to say that every independent parameter has relationship with emotion.

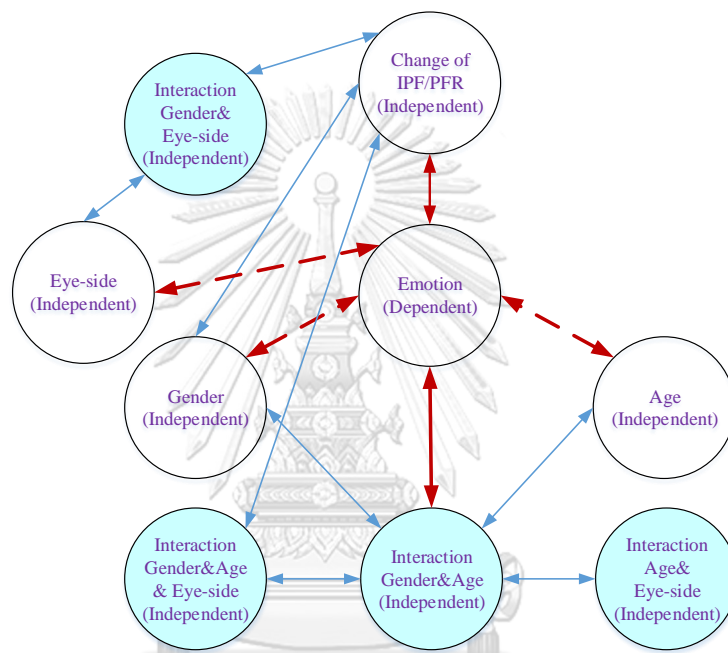


Figure 25 Relationships among all variables

4.1.3 Machine Learning Results

4.1.3.1 Creating Model

Based on the conclusion from the previous section, all parameters led to compile with a data science software tool, called RapidMiner Studio version 8.1. The changed pattern from the parameters is customized to a general model as an emotional classification model, a many-to-one relationship among variables as shown in Eq (4.1).

$$Emotion = M(\text{gender, age, eye-side, IPF, PFR}) \text{ -----(4.1)}$$

Eq (4.1) refers to “*Emotion is the model of gender, age, eye-side, IPF, and PFR with possible interactions*”; the values of all parameters are defined in Table 18.

Since the emotion classification model is many-to-one relation and the solution of the Emotion model can be either neutral or fear, a decision tree can be implemented to express all possible paths for emotion's derivation. In this research, the decision tree algorithm ID3 is chosen to create a tree-like graph or model. Thus, the data set was imported to RapidMiner version 8.1 for creating a classification model of the emotion expressions from a value of a target attribute called data labelling.

Since there are 40 training data sets, attributes of the training dataset consist of the results from calculation the average of the IPF and the PFR, both on the right eye and the left eye. Furthermore, the structure of the decision tree algorithm ID3 is shown in Figure 26.

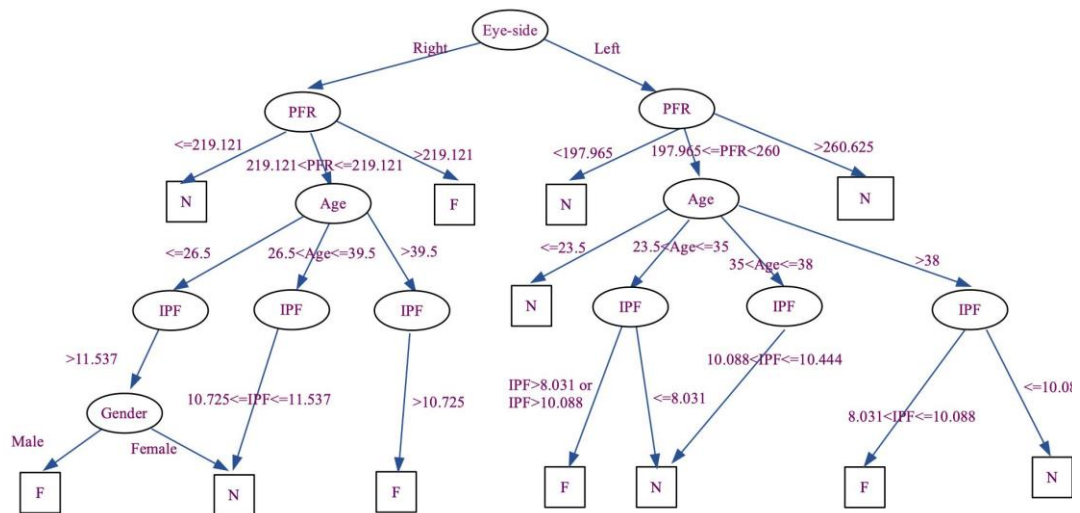


Figure 26 Decision Tree Algorithm ID3 of Emotion Classification

According to the multiple links tree in Figure 26, a simplified decision tree is derived as shown in Figure 27. Based on the decision tree algorithm ID3 of emotion classification in Figure 27, the links of nodes to explore emotions between neutral and fearful emotion is refined, including the depth of tree. Consequently, the tree length is increased from the root node for precise and correct the emotion classification model.

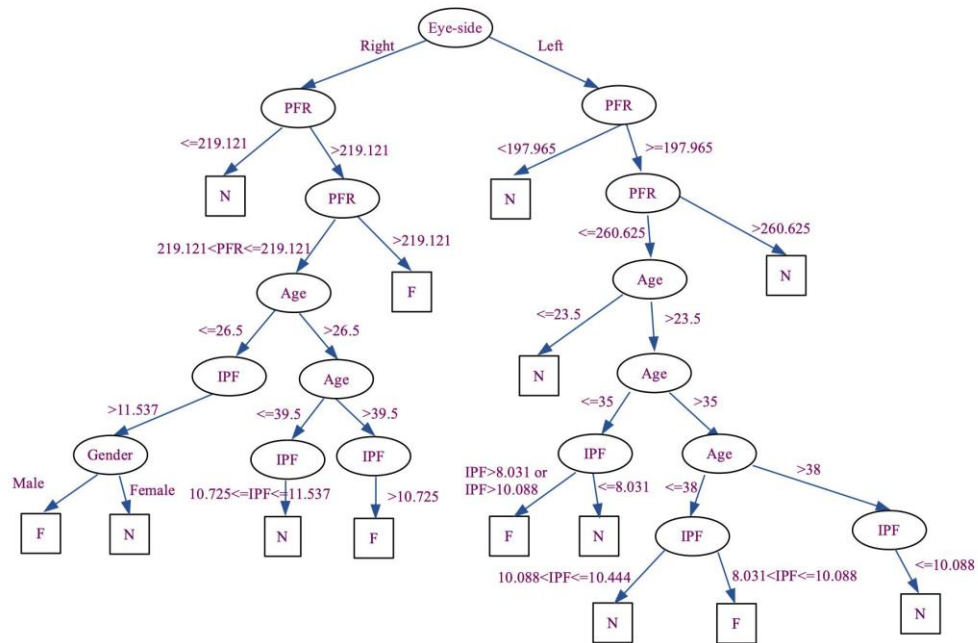


Figure 27 Refinement of Emotion Classification Decision Tree Algorithm ID3

4.1.3.2 Simulation base on the emotion classification decision tree

The training led to learn for the emotional classification model and the testing set was used to display performance of the emotional classification model using confusion matrix table as shown in Table 19.

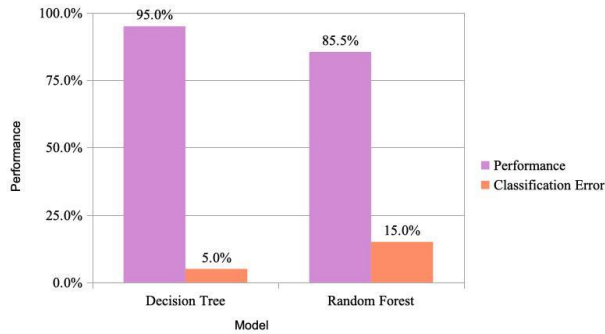
Table 19 Confusion Matrix of the Emotion Classification Decision Tree

Predict/Actual	Neutral	Fear	Class Precision
Neutral	18	2	90.00 %
Fear	0	20	100.00 %
Class Recall	100.00 %	90.00 %	

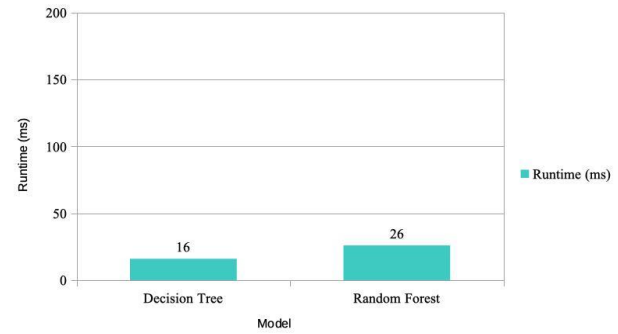
From the testing results, TP is 18, FP is 0, FN is 2, and TN is 20. Consequently, Precision value is 100%, the probability that the model can present the right answer is 1. Additionally, the Recall value is 90%, the correctness probability of the solution when uses the model is 0.90. Finally, the performance of this model is 95%; these can be interpreted that there is a small chance that this model can predict the wrong emotion. Moreover, the results of emotion classification and classification error, including training runtime, are compared using the same datasets with the Random Forest Model which can be seen that these results are much better than Random Forest as displayed in Table 20, Figure 28 and Figure 29.

Table 20 Comparison of the results between Decision Tree and Random Forest Model

Model	Performance	Classification Error	Runtime
Decision Tree	95.00 %	5.00 %	16 ms
Random Forest	85.50 %	15.00 %	26 ms



a. Comparison of classification performance



b. Comparison of runtime (ms)

Figure 28 Comparison of classification performance and runtime

Figure 29 displayed performance comparison between the decision tree method and the random forest method with receiver operating characteristics (ROC). The red line represents the decision tree, and the blue line represents the random forest. This indicates that the performance of the decision tree method is better than the random forest method.

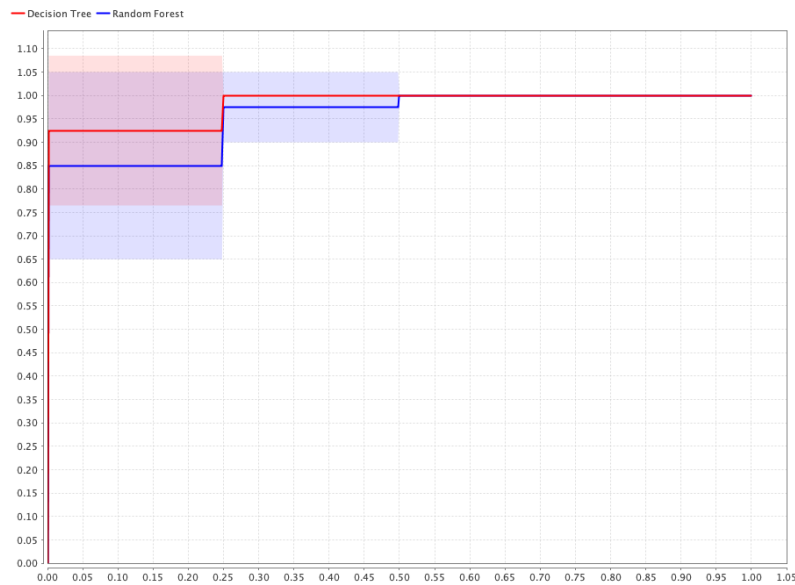


Figure 29 Comparison of Receiver Operating Characteristics (ROC) between Decision Tree and Random Forest

4.2 Parameter Selection: Second Phase

This phase is to explore parameters of the sub-structure of face such as the Inner Brow Raiser (IBR), Brow Lower (BL), Lid Raiser (LR), and Lip Part (LP), including gender and age. From the training set, 240 images, 40 are separated to comprise groups that are represented as 50.0%, male and female. Additionally, the highest age range is between 18-25 years old, 35.0% of the entire sample group. In the statistical testing, the non-parametric Wilcoxon signed ranks test, is deployed to compare the differences between the neutral emotion and the fearful emotion. The Wilcoxon signed ranks test is performed employing a 95% confident level as detailed below.

4.2.1 Statistical Results

The part contains the results of the determination on the differences of average and standard deviation values of the IBR, BL, LR, and LP.

4.2.1.1. Test of the IBR means and standard deviations differences

Considering the distribution of the data, the distribution of the inner brow raisers is not normal as displayed in Figure 30. Therefore, the non-parametric analysis is applied to analyze the differences between neutral and fear situations. As a result, using the Wilcoxon signed ranks test with a 95% confidence level, the statistical values indicated that, for both males and females, there are no significant different of the means between the IBRs under the neutral and fear emotions, with p -value = $0.417 > 0.05$ and p -value = $0.454 > 0.05$, respectively. Moreover, the calculation also points out that in the fear state, the mean value of the right IBR of males is significantly different from females, but not the left side. Nevertheless, the mean of

the IBRs for both sides in males under the normal mode shows no significant difference from the mean of the IBRs in females.

Considering the differences of means of the IBRs under the neutral state and fear state, the non-parametric Wilcoxon signed ranks test, it can determine that there is a significant difference between the means of the right IBRs of males when the emotion has been changed from neutral to fear, with $p\text{-value} = 0.00 < 0.05$, similar to the left IBRs. Like males, there is also a significant difference between the means of the right IBRs of females, which is the same as the left IBRs, when their temperament was altered from normal to the scary state, with $p\text{-value} = 0.00 < 0.05$. The distance average and standard deviation of the IBRs on all states of both males and females are displayed in Table 21.

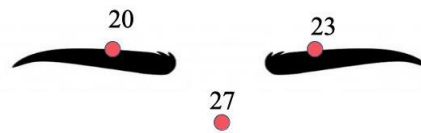


Figure 30 Inner Brow Raiser (IBR)

Table 21 The average with standard deviation of the IBR (millimeters)

Emotion	IBR	Gender		Total Mean
		Male	Female	
Normal	Right	37.18 ± 3.58	39.42 ± 3.12	38.30 ± 3.35
	Left	37.45 ± 3.13	38.42 ± 2.58	37.94 ± 2.86
Fear	Right	38.14 ± 3.39	39.32 ± 4.08	38.73 ± 3.74
	Left	39.85 ± 3.64	39.81 ± 2.59	39.83 ± 3.12

4.2.1.2. Test of the BL means and standard deviations differences

Considering the BL on Figure 31, which is the distance between the right end of the left eyebrow to the left end of another eyebrow, the means of this BL in both gender under two emotions were calculated and compared for differences, as shown in Table 22. With the same reason in the previous section, running the Wilcoxon signed ranks test indicates that there is no significant difference between the mean values of the BLs of males and females, with $p\text{-value} = 0.514 > 0.05$. In the neutral situation, the BL's mean in males is significantly different from the BL's mean in females, with $p\text{-value} = 0.000 < 0.05$. Furthermore, under the fear situation, there is a significant difference between the mean values of the BL in males and females, with $p\text{-value} = 0.000 < 0.05$. Furthermore, in both males and females, the mean value of the BL in a normal mode is significantly different from the mean value of the BL in a fear mode, with $p\text{-value} = 0.000 < 0.05$.



Figure 31 Brow Lower (BL)

Table 22 The average with standard deviation of the BL (millimeters)

Emotion	Gender		Total Mean
	Male	Female	
Neutral	25.42 ± 3.04	29.15 ± 2.66	27.29 ± 2.85
Fear	26.27 ± 3.30	27.17 ± 3.32	26.72 ± 3.31

4.2.1.3. Test of the LR means and standard deviations differences

This section focuses on the positions of the LRs that can be classified as the inner lid and the outer lid for each eye. Based on Figure 32, the outer LRs are the number pairs (37, 41) of the right eye and (44, 46) of the left eye while the inner LRs are the number pairs (38, 40) of the right eye and (43, 47) of the left eye.

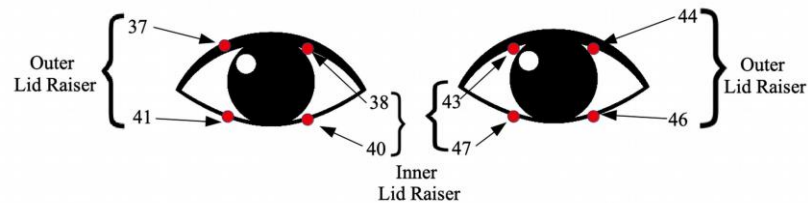


Figure 32 The Inner and Outer Lid Raiser (LR)

Since the distribution of the LRs of both eyes from all samples is not normal, the Wilcoxon signed ranks test, which is a non-parametric method, is applied. From computation of the Wilcoxon signed ranks test, the result shows that there is a significant difference of the LR mean values under the neutral state of both eyes, with $p\text{-value} = 0.001 < 0.05$; this is unlikely to the fear state where $p\text{-value} = 0.105 > 0.05$. In addition, this analysis also pointed out that the average sizes of the LRs from both eyes of both genders are significantly different when emotions have been changed from neutral to fear, with $p\text{-value} = 0.000 < 0.05$. The average and standard deviation of the inner lid raiser and the average and standard deviation of the outer lid raiser are shown in Table 23 and Table 24, respectively.

Table 23 The average with standard deviation of the Inner Lid Raiser (LR) (millimeters)

Emotion	Inner LR	Gender		Total Mean
		Male	Female	
Normal	Right	8.87 ± 1.72	9.93 ± 1.85	9.40 ± 1.79
	Left	8.99 ± 1.79	10.08 ± 1.79	9.54 ± 1.79
Fear	Right	9.23 ± 1.58	10.70 ± 2.02	9.97 ± 1.80
	Left	9.32 ± 2.01	10.77 ± 2.01	10.05 ± 2.01

Table 24 The average with standard deviation of the Outer Lid Raiser (LR) (millimeters)

Emotion	Outer LR	Gender		Total Mean
		Male	Female	
Normal	Right	8.88 ± 1.87	8.83 ± 1.87	9.45 ± 1.87
	Left	8.83 ± 1.85	10.02 ± 1.88	9.43 ± 1.87
Fear	Right	9.22 ± 2.03	10.72 ± 2.05	9.97 ± 2.04
	Left	9.16 ± 2.06	10.69 ± 2.17	9.93 ± 2.12

4.2.1.4. Test of the LP means and standard deviations differences

This section describes the average and standard deviation of the inner LP where six positions on the lip were defined, as presented in Figure 33. To study the change of the lip's shape, three ordered pairs are taking in the consideration: (61, 65), (62, 66), (63, 67). Moreover, the link between points in each pair is created and measured. Once all links were defined, the normality test of the length of all links is performed using the Shapiro-Wilk test. Unfortunately, the result of the test indicated that the distribution of the length of all links is not normal; the average and standard deviation of all LP links are displayed in Table 25. Since the distribution of links is not normal, the comparisons among links' means must rely on the non-parametric method. Using Wilcoxon signed ranks test with 95% confident level, it can determine that the means of all lengths for order pairs (61, 65), (62, 66), and (63, 67) under the neutral state are significantly different from those of fears, p -value = 0.001 < 0.05, p -value = 0.000 < 0.05, and p -value = 0.003 < 0.05.

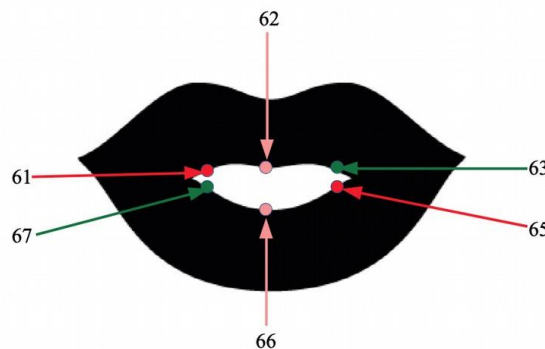


Figure 33 The Inner Lip Part

Table 25 The Distance Link's Lip Part (LP) average with standard deviation (millimeters)

The First Link			
Emotion	Gender		Total Mean
	Male	Female	
Neutral	17.27 ± 2.55	16.88 ± 3.39	17.08 ± 2.97
Fear	18.03 ± 3.73	18.55 ± 4.38	18.29 ± 4.06
The Second Link			
Emotion	Gender		Total Mean
	Male	Female	
Neutral	4.68 ± 3.80	4.71 ± 4.38	4.70 ± 4.09
Fear	6.91 ± 5.59	8.53 ± 6.27	7.72 ± 5.93
The Third Link			
Emotion	Gender		Total Mean
	Male	Female	
Neutral	17.25 ± 2.56	16.80 ± 3.40	17.03 ± 2.98
Fear	17.80 ± 3.66	18.58 ± 4.01	18.19 ± 3.84

4.2.2. Choosing Parameters

Referring to the statistical results above, it can be concluded that the efficiency of emotion detection can be achieved by five parameters: gender, the inner BR, the BL, the LR, and the LP. Nonetheless, there are small details for each parameter that have to be concerned, such as defining whether the inner BR and the LR are on the right or the left side on the face, or the whether the LR is associated with the outer lid or the inner lid, the including of age, and the links over the LP. Therefore, to obtain a precise classification of emotion expression from a face, various parameters must be applied; these parameters are gender, age, side, position, the inner BR, the BL, the LR, and the LP. Table 26 presents all values of parameters. Moreover, Figure 34 displays possible relationships between independent variables and dependent variable; the red dot line represents the effect of factors an emotion. Furthermore, the relationship between gender and emotion, and the relationship between age and emotion can be determined by calculating the correlation.

Table 26 The Values of All Parameters

Independent Variable	Values
Gender	Male, Female
Age-Range	18-24, 25-31, 32-38, 39-45, 46-53, 54-61
Side	Right, Left
Position	Outer, Inner
Link	First, Second, Third
Inner Brow Raiser (IBR)	[37.18, 39.85]
Brow Lower (BL)	[25.42, 29.15]
Lid Raiser (LR)	[8.83, 10.77]
Lip Part (LP)	[4.68, 18.58]
Dependent Variable	Values
State of Emotion	Normal, Fear

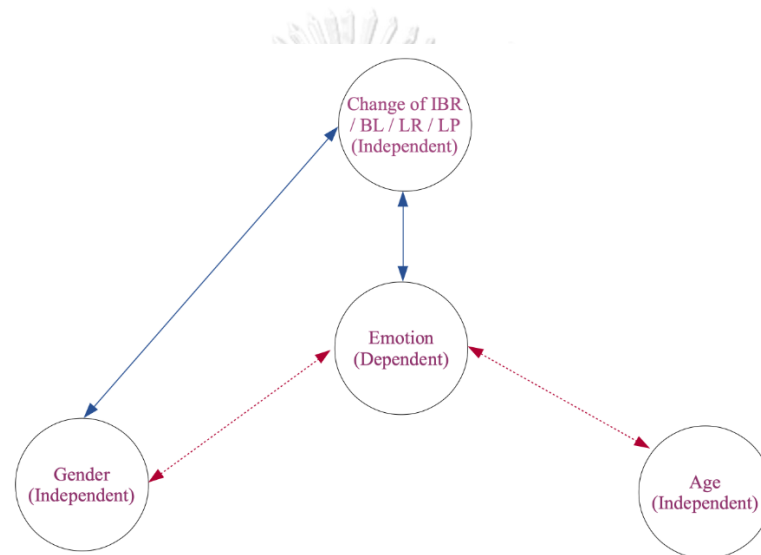


Figure 34 Relationships among independent and dependent variables

Since the second phase classified IBR, BL, LR, and LP as the independent variable, the selected dependent variable stores either the value of “neutral” or “fear”. Therefore, statistical testing must be performed to determine the significant differences of IBR, BL, LR, or LP both two emotions using all these independent variables. Although there are various factors that are related to the dependent variable, the created classification model must be optimized and implemented by the smallest numbers of factors to maintain simplicity of the classification modeling. Thus, statistical analysis to identify factors and their impacts towards the values of IBR, BL, LR, or LP was performed. However, the distributions of IBR, BL, LR, and LP were determined as non-normal distribution. Therefore, the non-parametric analysis, namely Kruskal-Willis test, is applied using 95% to be the confident level.

4.2.2.1 IBR Consideration

The basic testing assumptions are listed as follows.

- At least one level of age-range has different effect to the IBR mean values.
- Different genders, male and female, have different effects to the IBR mean values.
- Different emotions, neutral and fear, have different effects to the IBR mean values.
- Different brow-sides, right brow and left brow, have different effects to the IBR mean values.

To prove those assumptions, the Wilcoxon signed rank test with 95% confident level is deployed. The results of these testing found that there is only one variable that has no effect to the IBR mean values, with p -values $> 0.05 = \alpha$; this variable is the age-range. On the other hand, other variables that can categorize the differences of the IBR mean values are genders, emotions, and brow-sides, with p -value $< 0.05 = \alpha$. Thus, these variables are used to combine with others where the interactions among factors can be derived to 10 assumptions as listed below.

- At least one combination between genders and age-ranges has significant different from others towards the IBR mean values.
- At least one combination between genders and emotions has significant different from others towards the IBR mean values.
- At least one combination between genders and brow-sides has significant different from others towards the IBR mean values.
- At least one combination between brow-sides and emotions has significant different from others towards the IBR mean values.
- At least one combination between age-ranges and emotions has significant different from others towards the IBR mean values.
- At least one combination between age-ranges and brow-sides has significant different from others towards the IBR mean values.
- At least one combination between age-ranges, genders, and emotions has significant different from others towards the IBR mean values.
- At least one combination between age-ranges, genders, and brow-sides has significant different from others towards the IBR mean values.
- At least one combination between genders, emotions, and brow-sides has significant different from others towards the IBR mean values.
- At least one combination between age-ranges, genders, emotions, and brow-sides has significant different from others towards the IBR mean values.

According to the assumptions above with the use of the Wilcoxon signed rank test with 95% confident level, the statistical results indicated that there is no significant differences among the IBR mean values under combinations of genders, ages, and emotions with p -value = 0.456 $> 0.05 = \alpha$; and there is no significant difference among the IBR mean values under combinations of ages and emotions, with p -value = 0.781 $> 0.05 = \alpha$. In addition, there is no significant difference among the IBR mean values under combinations of ages and brow-sides, both the right side

and the left side, with $p\text{-value} = 0.916 > 0.05 = \alpha$ and with $p\text{-value} = 0.238 > 0.05 = \alpha$ respectively. Moreover, the combinations of ages, genders, and brow-sides are no significant difference among the IBR mean values in two brow-sides, with $p\text{-value} = 0.299 > 0.05 = \alpha$ and with $p\text{-value} = 0.916 > 0.05 = \alpha$. Thus, it can be seen that all combinations of ages have no effect to the IBR mean values.

Although the age-ranges variable has no impact to the IBR mean values, this variable is still essential to combine with genders, emotions, and brow-sides for creating the classification model.

4.2.2.2 BL Consideration

This process is similar to the IBR analysis where the non-parametric namely Wilcoxon signed ranks test with 95% confident level is applied. The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effects to the BL mean values.
- Different genders, male and female, have different effects to the BL mean values.
- Different emotions, neutral and fear, have different effects to the BL mean values.

The results of these basic testing found that there are two variables that do not have effect to the BL mean values, with $p\text{-value} > 0.05 = \alpha$; these variables are age-ranges and genders. On the other hand, the variables of genders and emotions can categorize the differences of the BL mean values, with $p\text{-values} < 0.05 = \alpha$. Thus, four assumptions can be derived as listed below.

- At least one combination between ages and genders has significant different from others towards the BL mean values.
- At least one combination between ages and emotions has significant different from others towards the BL mean values.
- At least one combination between genders and emotions has significant different from others towards the BL mean values.
- At least one combination between ages, genders, and emotions has significant different from others towards the BL mean values.

Based on the assumptions above, the testing results discover that only the combination between genders and emotions has at least one significant difference in the BL mean values, with $p\text{-value} = 0.000 < 0.05 = \alpha$. However, there is no significant difference among the BL mean values under combinations ages, genders, and emotions, with $p\text{-value} = 0.514 > 0.05 = \alpha$.

4.2.2.3 LR Consideration

This process is similar to the IBR analysis where the non-parametric namely Wilcoxon signed ranks test with 95% confident level is applied. The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effect to the LR mean values.
- Different genders, male and female, have different effects to the LR mean values.
- Different emotions, neutral and fear, have different effects to the LR mean values.
- Different eye-sides, right and left eyes, have different effects to the LR mean values.
- Different positions, in and out positions, have different effects to the LR mean values.

The results of these fundamental testing found that there are two variables: age-range and emotions, that do not have effect to the LR, with $p\text{-value} > 0.005 = \alpha$. On the other hand, other variable that can categorize the differences of the LR values are genders, eye-sides, and positions, with $p\text{-values} < 0.005 = \alpha$. Thus, genders, eye-sides, and positions are used to combine with others where the interactions among factors are listed below.

- At least one combination between genders and ages has significant different from others towards the LR mean values.
- At least one combination between genders and emotions has significant different from others towards the LR mean values.
- At least one combination between ages and emotions has significant different from others towards the LR mean values.
- At least one combination between genders and eye-sides has significant different from others towards the LR mean values.
- At least one combination between genders and positions has significant different from others towards the LR mean values.
- At least one combination between emotions and eye-sides has significant different from others towards the LR mean values.
- At least one combination between emotions and positions has significant different from others towards the LR mean values.
- At least one combination between genders, eye-sides and positions has significant different from others towards the LR mean values.
- At least one combination between ages, eye-sides and positions has significant different from others towards the LR mean values.

According to the assumptions above, the testing results discover that only the combination between ages and emotions has no significant difference in the LR mean values, both position inner and position outer in the fear mode, with $p\text{-value} = 0.105 > 0.005 = \alpha$ and with $p\text{-value} = 0.139 > 0.005 = \alpha$.

Though, the ages and emotions variables cannot influence the change of the LR mean values, these variables are still needed to use with other factors. Thus, the third classification model includes genders, ages, emotions, eye-sides, and positions.

4.2.2.4 LP Consideration

The last process is to analyze the LP value using non-parametric analysis, Wilcoxon signed ranks test, with 95% confident level. The basic testing assumptions are listed as follows.

- At least one level of age-range has different effect to the LP mean values.
- Different genders, male and female, have different effects to the LP mean values.
- Different emotions, neutral and fear, have different effects to the LP mean values.
- Different links, the first link, the second link, and the third link, have different effects to the LP mean values.

The results of these rudimentary testing found that all variables have effect to the LP mean values of the first link, the second link, and the third link, with $p\text{-value} < 0.005 = \alpha$. Thus, three assumptions can be derived as listed below.

- At least one combination between genders and emotions has significant different from others towards the LP mean values.
- At least one combination between ages and emotions has significant different from others towards the LP mean values.
- At least one combination between genders, ages and links has significant different from others towards the LP mean values.

The results of these interactions testing indicate that all combinations between genders and emotions have at least one significant difference between the LP mean values, with $p\text{-value} = 0.000 < 0.05 = \alpha$; and there is at least one significant difference among the LP mean values under the combination of ages and emotions, with $p\text{-value} = 0.000 < 0.05 = \alpha$. Therefore, all variables: ages, genders, emotions, and links are applied to create the fourth classification model.

Referring to all statistical results mention previously, there are relationships among parameters in Table 26. So, the relationship diagram that illustrates the relationships between the dependent variable and the independent variables is drawn in Figure 35. From this figure, the direct effect to the change of Emotion is the Change of IBR/BL/LR/LP while the indirect effects occur from Age, Side, Position, Link, and Gender. These indirect effects are caused by all interactions of variables: Age, Side, Position, Link, and Gender, influence on the Change of IBR/BL/LR/LP which has impact to the Emotion.

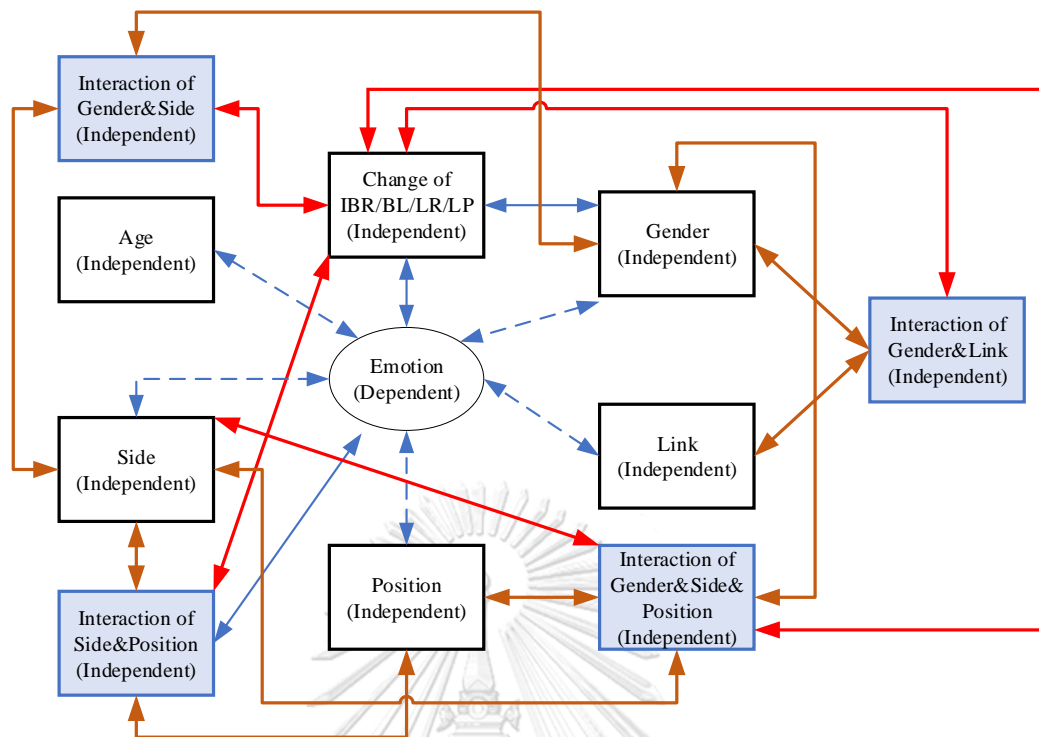


Figure 35 Relationships of all variables

4.2.3 Machine Learning Results

4.2.3.1. Creating Model

After all parameters are defined, the next step is the model implementation based on the discovered parameters. To accomplish the objective, this phase deploys the Scikit-Learn, a machine learning library to interpret all parameters via the PyCharm Community Edition version 2018.2 with Python language. The pattern was modified to a general classification model, as presented in Eq (4.2),

$$Emotion = M(\text{gender}, \text{age}, \text{side}, \text{position}, \text{link}, \text{IBR}, \text{BL}, \text{LR}, \text{LP}) \text{ -----(4.2)}$$

From Eq (4.2), Emotion is obtained from the model M which M refers to the model with parameters gender, age, side, position, IBR, BL, LR, and LP; the possible values of all parameters are defined in Table 26.

The emotion classification model was applied to classify emotion as either neutral or fear using a decision tree algorithm. Thus, 360 images of the dataset were imported to PyCharm Community Edition with Scikit-Learn. Then, this dataset is separated into the training dataset and the testing dataset, with a ratio of 70:30. Therefore, 240 images of data are the training dataset that contains the cluster attributes, for which the expected result after the learning process, namely, a label, is obtained. Furthermore, the decision tree algorithm received as another outcome of the learning process is displayed in Figure 36.

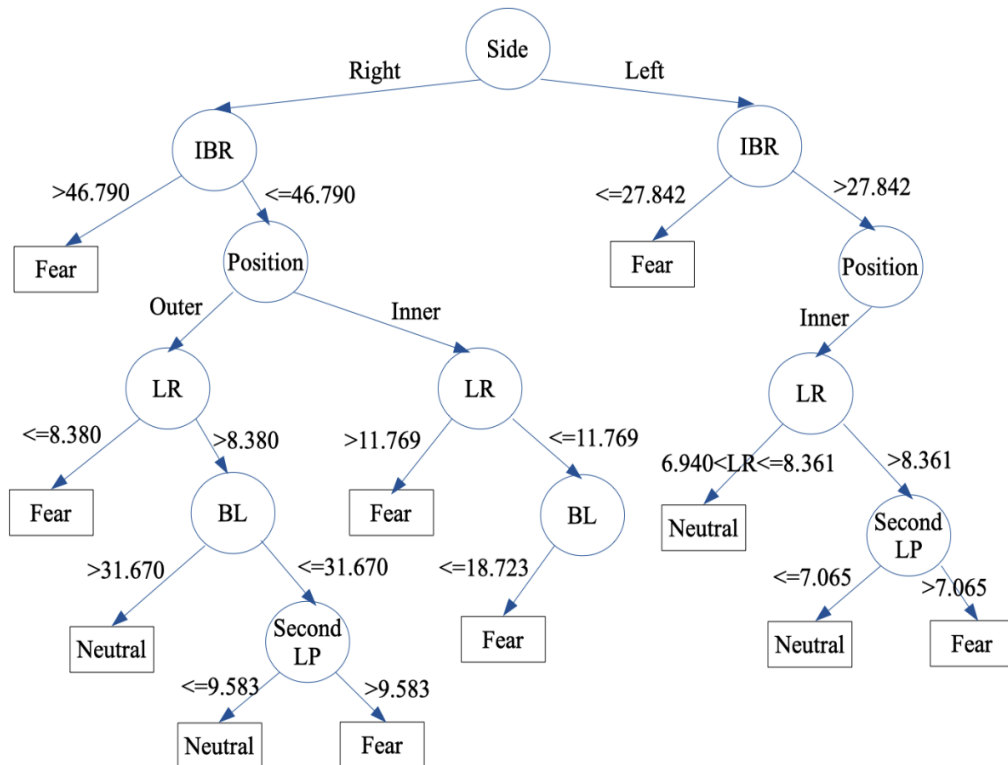


Figure 36 Decision Tree Algorithm for Emotion Classification

4.2.3.2. The emotion simulation based on decision tree classification model

The training led to learn for the emotion classification model while the testing dataset was used to measure the performance of the emotional classification model and displayed in the confusion matrix as shown in Table 27.

Table 27 Confusion Matrix of the Emotion Classification Decision Tree

Predict/Actual	Neutral	Fear	Class Precision
Neutral	15	3	83.33 %
Fear	3	15	83.33 %
Class Recall	83.33 %	83.33 %	

From the experimental results in Table 27, the number of True Positive (*TP*) that refers to the right prediction of the emotion from the model when comparing with the true value is 15. The number of False Positive (*FP*) which means the prediction is not correct is 3. The number of False Negative (*FN*) is always equal to *FP* is 3. The number of True Negative (*TN*) is always equal to *TP* is 15. Consequently, the precision value is 83.33%, the recall is 83.33%, and the accuracy of this model is 83.33%. Moreover, the comparison of the performance and classification error from the same datasets with the decision tree model and random forest model is displayed in Table 28.

Table 28 Comparison of the results between Decision Tree and Random Forest Model

Model	Performance	Classification Error
Decision Tree	83.33 %	16.67 %
Random Forest	75.00 %	25.00 %

Figure 37 shows the receiver operating characteristic (ROC) curves, where the performance of the decision tree model, the red line, and the random forest model, the blue line, has been compared. So, it can be interpreted that the decision tree has better performance than the random forest.

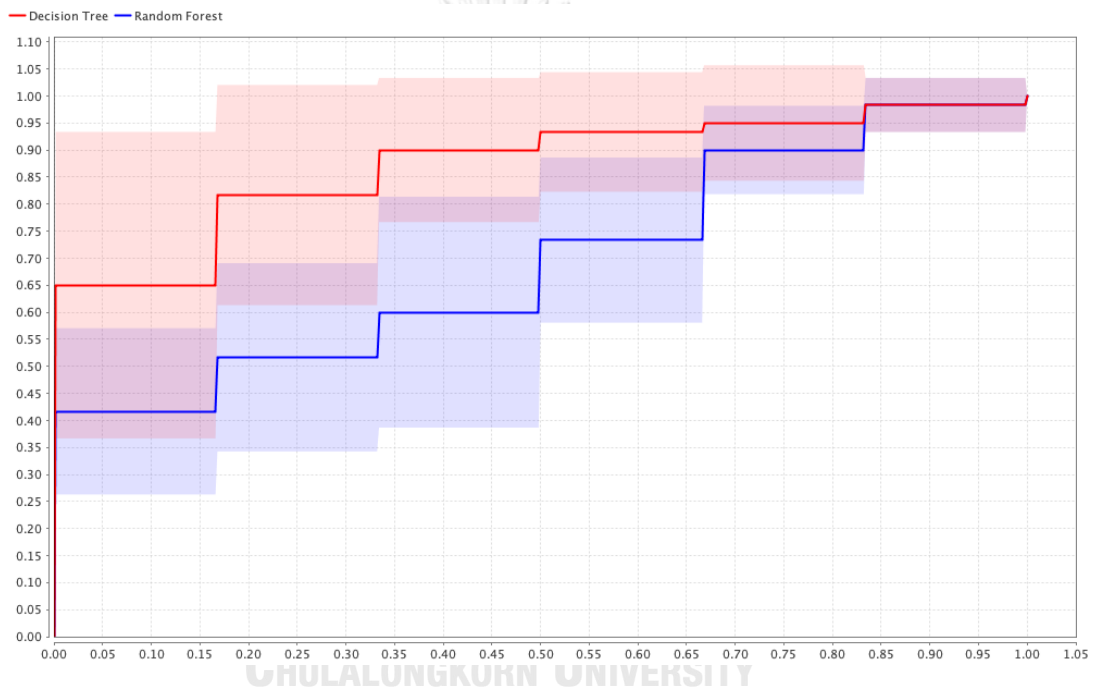


Figure 37 Comparison of Receiver Operating Characteristics (ROC) between Decision Tree and Random Forest

4.3 Parameter Selection: Third Phase

This is the significant phase since the real data from patients who are selected by the medical staff in Chulalongkorn Comprehensive Stroke Center of Excellent, King Chulalongkorn Memorial Hospital, is tested under the obtained model mentioned above. Since all parameters from the first two phases can distinguish between normal and fear emotions, thus, these parameters will be collected from the stroke patients for the same purpose. Therefore, parameters to be collected are the Interpalpebral Fissure (IPF), Palpebral Fissure Length (PFL), Palpebral Fissure

Region (PFR), Inner Brow (IBR), Brow Lower (BL), Lid Raiser (LR), and Lip Part (LP), including gender and age.

Then, the data of these patients are separated into the training set and the testing set, similar to the previous processes. In the training set, 54 images, 9 are 44.44% male and 55.56% female. Additionally, the highest age range is more than 76 years old, 50% of the entire samples. In the statistical testing, the non-parametric Wilcoxon signed ranks test, is deployed to compare the differences between the normal status and the abnormal status. The Wilcoxon signed ranks test is performed employing a 95% confidence level as detailed below.

4.3.1 Statistical Results

This part contains the results of the determination on the difference of average and standard deviation values of the IPF, PFL, PFR, IBR, BL, LR, and LP.

4.3.1.1 Test of the IPF means and standard deviations differences

Considering the distribution of the data, the distribution of the interpalpebral fissure is not normal as displayed in Figure 38. Thus, the non-parametric analysis is applied to analyze the differences between normal and abnormal status. As a result, using the Wilcoxon signed ranks test with a 95% confidence level, the statistical values indicated that, for both males and females, there are no significant difference of the means between the IPFs of the right eye and the left eye under the normal situation and abnormal situation, with $p\text{-value} = 0.362 > 0.05$ and $p\text{-value} = 0.148 > 0.05$, respectively.

Considering the differences of means of the IPFs under the normal status and abnormal status, the non-parametric Wilcoxon signed ranks test, it can determine that there is a significant difference of the means of the males' IPFs when the state has been changed from normal to abnormal, with $p\text{-value} = 0.00 < 0.05$, like females. Moreover, there is a significant difference between the means of the right IPFs of males when their temperament was altered from normal to the abnormal status, with $p\text{-value} = 0.00 < 0.05$, which is the same as the left IPFs. Like males, when status was changed from normal to the critical state, there is also a significant difference between the means of the right IPFs of females, which is the same as the left IPFs, with $p\text{-value} = 0.00 < 0.05$. The IPF means with standard deviations of both eyes are displayed in Table 29.

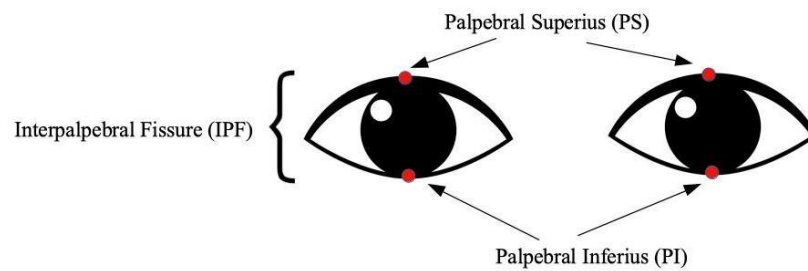


Figure 38 Interpalpebral Fissure

Table 29 The IPF means with standard deviations (Millimeters)

Status	IPF	Gender		Total Mean
		Male	Female	
Normal	Right	8.13 ± 1.82	8.20 ± 0.71	8.17 ± 1.27
	Left	7.90 ± 1.50	8.17 ± 0.95	8.04 ± 1.23
Abnormal	Right	10.62 ± 1.86	11.23 ± 0.72	10.93 ± 1.29
	Left	10.69 ± 2.07	11.43 ± 0.62	11.06 ± 1.35

4.3.1.2 Test of the PFL means and standard deviations differences

Considering the PFLs on Figure 39, which is the distance between the exocanthion (EX) to the endocanthion (EN), the means of the PFLs in both gender under two statuses were calculated and compared for differences, as shown in Table 30. With the same reason in the previous section, running the Wilcoxon signed ranks test indicates that there is no significant difference of the means between the PFLs of the right eye and the left eye under the normal situation and abnormal situation in males, with p -value = 0.542 > 0.05. Nevertheless, the means of the PFLs for both sides in female, under abnormal status, have at least one significant difference, with p -value = 0.021 < 0.05. Furthermore, in both males and females, the mean values of the PFLs are significantly differences under normal mode and abnormal mode, with p -value = 0.000 < 0.05.

Using the non-parametric Wilcoxon signed ranks test to prove the differences of PFLs means under the normal situation and abnormal situation, it can determine that there is at least one significant difference between of the right PFLs of both males and females, with p -value = 0.000 < 0.05. Moreover, there is also a significant difference between the means of the left PFLs of both males and females, with p -value = 0.000 < 0.05.

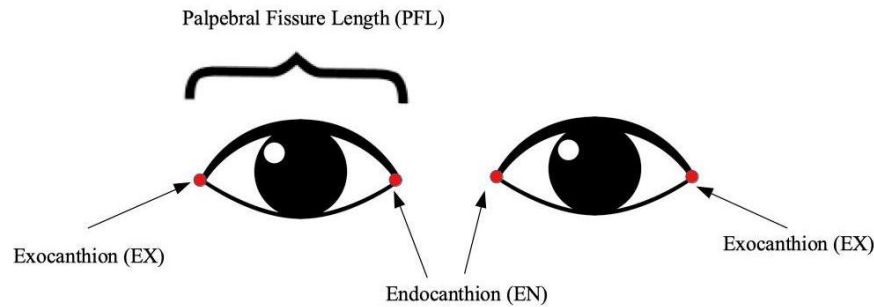


Figure 39 Palpebral Fissure Length (PFL)

Table 30 The PFL means with standard deviations (Millimeters)

Status	PFL	Gender		Total Mean
		Male	Female	
Normal	Right	31.60 ± 7.64	31.05 ± 5.94	31.32 ± 6.79
	Left	32.90 ± 6.58	30.87 ± 6.73	31.89 ± 6.66
Abnormal	Right	38.21 ± 8.02	40.13 ± 7.90	39.17 ± 7.96
	Left	42.49 ± 7.32	40.69 ± 7.98	41.59 ± 7.65

4.3.1.3 Test of the PFR means and standard deviations differences

This section focuses on the palpebral fissure region that is calculated from the IPF and the PFL; the formula is presented in Eq. (4.3). Since the distribution of the PFRs of both eyes from all samples is not normal, the Wilcoxon signed ranks test with a 95% confidence level, which is a non-parametric method, is applied. As a result, the statistical values indicated that, for both males and females, there are no significant difference of the means between the PFRs under the normal situation, with p -value = 0.802 > 0.05. However, the mean of the PFRs for both sides under the abnormal situation shows a significant difference in both males and females, with p -value = 0.016 < 0.05.

$$PFR = (\pi/4) * IPF * PFL \quad \text{-----} \quad (4.3)$$

Considering the differences of the PFR means under the normal mode and the abnormal mode, there is a significant difference between the PFR mean of both sides for both males and females when the situation has been altered from normal state to abnormal state, with p -value = 0.000 < 0.05. Moreover, there is a significant difference between the means of the right PFRs of both males and females, with p -value = 0.000 < 0.05, similar to the left PFRs.

Table 31 The PFR means with standard deviations (Millimeters)

Status	PFR	Gender		Total Mean
		Male	Female	
Normal	Right	209.23 ± 97.21	199.18 ± 36.53	204.21 ± 66.87
	Left	207.74 ± 73.34	199.57 ± 55.02	203.65 ± 64.18
Abnormal	Right	317.03 ± 81.83	353.28 ± 68.74	335.16 ± 75.29
	Left	354.62 ± 79.71	365.94 ± 76.50	360.28 ± 78.11

4.3.1.4 Test of the IBR means and standard deviations differences

Considering the distribution of the data, the distribution of the inner brow raisers (IBRs) is not normal. Thus, the non-parametric analysis is applied to analyze the differences between normal mode and abnormal mode. As a result, using the Wilcoxon signed ranks test with a 95% confidence level, the statistical values indicated that, for both males and females, there are no significant difference of the means between the IBRs under the normal situation and abnormal situation, with p -value = 0.546 > 0.05 and p -value = 0.125 > 0.05, respectively.

Considering the differences of the means IBRs under the normal status and abnormal status, the non-parametric Wilcoxon signed ranks test, for both sides, there is a significant difference between the means of the IBRs of males and females, with p -value = 0.000 < 0.05.

Table 32 The IBR means with standard deviations (Millimeters)

Status	IBR	Gender		Total Mean
		Male	Female	
Normal	Right	40.80 ± 8.54	40.59 ± 7.68	40.70 ± 8.11
	Left	41.90 ± 7.35	39.57 ± 7.46	40.73 ± 7.41
Abnormal	Right	47.92 ± 7.73	49.95 ± 7.56	48.94 ± 7.65
	Left	52.22 ± 7.58	49.47 ± 8.06	50.85 ± 7.82

4.3.1.5 Test of the BL means and standard deviations differences

Considering the distribution of the blow lower, which is the distance between the right end of the left eyebrow to the left end of another eyebrow. With the same reason in the previous section, running the Wilcoxon signed ranks test indicates that there is a significant difference between the mean values of the BLs of males and females under the normal mode and the abnormal mode, with p -value = 0.000 < 0.05. Moreover, there is a significant difference between the means of the BLs of both males and females under the normal situation and the abnormal situation, with p -value = 0.015 < 0.05 and p -value = 0.010 < 0.05, respectively. The distance average and standard deviation of the BLs on all states of both males and females are displayed in Table 33.

Table 33 The average with standard deviation of the BL (millimeters)

Status	Gender		Total Mean
	Male	Female	
Normal	26.54 ± 6.70	27.10 ± 4.73	26.82 ± 5.71
Abnormal	30.60 ± 7.06	31.32 ± 4.96	30.96 ± 6.01

4.3.1.6 Test of the LR means and standard deviations differences

Based on Figure 32, this section focuses on the positions of the lid raisers (LRs) that can be classified as the inner lid and the outer lid for each eye. Thus, the outer LRs of the right eye are the number pairs (37, 41) and the outer LRs of the left eye are the number pairs (44, 46). The inner LRs of the right eye are the number pairs (38, 40) and the inner LRs of the left eye are the number pairs (43, 47).

Since the distribution of the LRs of both eyes is not normal, the Wilcoxon signed ranks test, which is a non-parametric method, is applied. The result shows that there is no significant difference of the means values of the LRs under the normal status from both eyes of both genders in the inner position, with p -value = 0.256 > 0.05. There is no significant difference between the means values of the LRs under the abnormal status from both eyes of both genders, with p -value = 0.600 > 0.05. Moreover, in the outer position, there is also no significant difference of the means values of the LRs under the normal situation and the abnormal situation from both eyes of both genders, with p -value = 0.581 > 0.05 and p -value = 0.629, respectively.

However, under the normal state and the abnormal state, there is a significant difference between the means of the right LRs and the left LRs of males and females, with p -value = 0.000 < 0.05. The average and standard deviation of the inner lid raiser and the average and standard deviation of the outer lid raiser are shown in Table 34 and Table 35, respectively.

Table 34 The average with standard deviation of the Inner Lid Raiser (LR) (millimeters)

Status	Inner LR	Gender		Total Mean
		Male	Female	
Normal	Right	8.43 ± 1.86	8.50 ± 0.77	8.47 ± 1.31
	Left	8.21 ± 1.56	8.39 ± 1.30	8.30 ± 1.43
Abnormal	Right	10.91 ± 1.98	11.54 ± 1.06	11.23 ± 1.52
	Left	10.97 ± 2.21	11.77 ± 1.18	11.37 ± 1.70

Table 35 The average with standard deviation of the Outer Lid Raiser (LR) (millimeters)

Status	Outer LR	Gender		Total Mean
		Male	Female	
Normal	Right	7.87 ± 1.99	7.94 ± 0.95	7.90 ± 1.47
	Left	7.62 ± 1.64	7.99 ± 0.82	7.80 ± 1.23
Abnormal	Right	10.37 ± 1.91	10.94 ± 0.85	10.66 ± 1.38
	Left	10.44 ± 2.02	11.14 ± 0.81	10.79 ± 1.42

4.3.1.7 Test of the LP means and standard deviations differences

This section focuses on the six positions of the lip parts (LPs), as displayed in Figure 33. There are three order pairs to study the change of the lip shape: (61, 65), (62, 66), (63, 67). Considering the distribution of the LPs, all links are not normal, the comparisons among links' means must reply on the non-parametric method. Using Wilcoxon signed ranks test with 95% confident level, it can determine that the means of all lengths for order pairs (61, 65), (62, 66), and (63, 67) of both males and females under the normal state are significantly different form the abnormal state, with p -value = 0.000 < 0.05. The average and standard deviation f all LPs' links are displayed in Table 36.

Table 36 The Distance Link's Lip Part (LP) average with standard deviation (millimeters)

The First Link			
Status	Gender		Total Mean
	Male	Female	
Normal	20.51 ± 8.94	16.16 ± 4.25	18.34 ± 6.60
Abnormal	31.04 ± 12.79	25.18 ± 5.66	28.11 ± 9.23
The Second Link			
Status	Gender		Total Mean
	Male	Female	
Normal	8.21 ± 7.23	5.19 ± 1.63	6.70 ± 4.43
Abnormal	18.09 ± 12.91	11.96 ± 4.44	15.03 ± 8.67
The Third Link			
Status	Gender		Total Mean
	Male	Female	
Normal	20.30 ± 8.33	16.17 ± 3.90	18.24 ± 6.11
Abnormal	28.94 ± 12.36	24.04 ± 5.36	26.49 ± 8.86

4.3.2 Choosing Parameters

Based on the statistical results, there are eight parameters for classification detection of situation: gender, the IPF, the PFL, the PFR, the IBR, the BL, the LR, and the LP, which is similar to the defined parameters those selected from the movies. Nevertheless, there are small details for each parameter that have to be concerned, such as defining the values of the PFR are depended on calculation from the values of the IPF and the values of the PFR, or the values of the IBR are on the right or the left side on the face, or the values of the LR is associated with the outer lid or the inner lid, the including of age, and the links over the LP. Therefore, the parameters those are applied for classification of facial expression as gender, age, side, position, the IPF, the PFL, the PFR, the IBR, the BL, the LR, and the LP. Table 37 displays all values of all parameters. Moreover, Figure 40 shows relationships between independent variables and dependent variable; the red dot line represents the effect of factors towards the status. Furthermore, the relationship between gender and status, and the relationship between age and status can be determined by calculating the correlation.

Table 37 The Values of All Parameters

Independent Variable	Values
Gender	Male, Female
Age-Range	40-57, 58-67, 68-79, 80-100
Side	Right, Left
Position	Outer, Inner
Link	First, Second, Third
Interpalpebral Fissure (IPF)	[7.90, 11.73]
Palpebral Fissure Length (PFL)	[30.87, 42.49]
Palpebral Fissure Region (PFR)	[199.18, 365.94]
Inner Brow Raiser (IBR)	[39.57, 52.22]
Brow Lower (BL)	[25.54, 31.32]
Lid Raiser (LR)	[7.62, 11.54]
Lip Part (LP)	[5.19, 31.04]
Dependent Variable	Values
Status	Normal, Abnormal

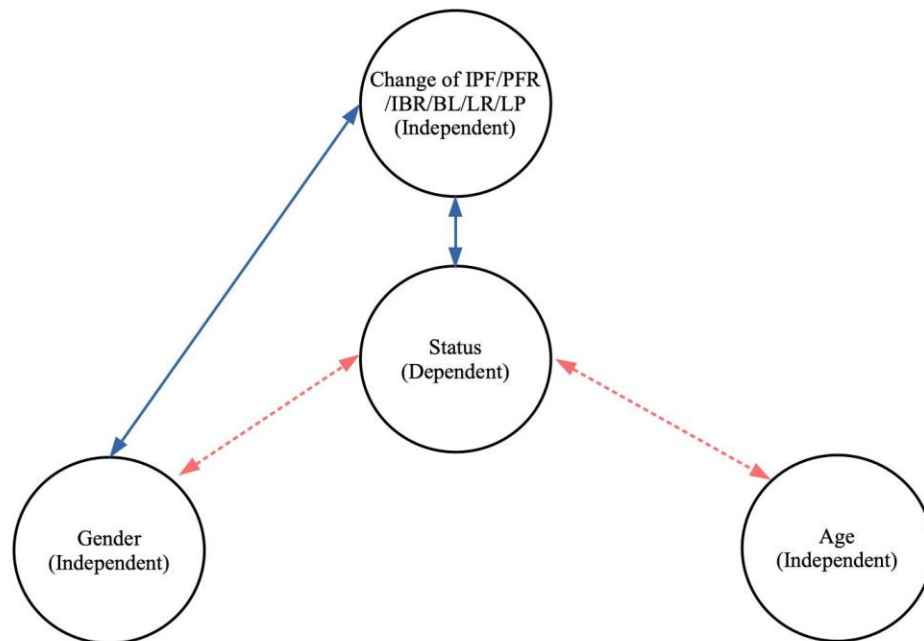


Figure 40 Relationships among independent and dependent variables

Since the third phase classified IPF, PFL, PFR, IBR, BL, LR, and LP as the independent variables, the selected dependent variable stores either the value of “normal” or “abnormal”. Thus, the statistical testing must be performed to determine the significant difference of IPF, PFL, PFR, IBR, BL, LR, and LP both two situations using all these independent variables. Although there are various factors that are related to the dependent variable, the created classification model must be optimized and implemented by the smallest numbers of factors to maintain simplicity of the classification modeling. Thus, statistical analysis to identify factors and their impacts towards the values of IPF, PFL, PFR, IBR, BL, LR, or LP was performed. However, the distributions of IPF, PFL, PFR, IBR, BL, LR, and LP were determined as non-normal distribution. Therefore, the non-parametric analysis, namely Kruskal-Willis test, is applied using 95% to be the confident level.

4.3.2.1 IPF Consideration

This process is to analyze the IPF value using non-parametric analysis, Kruskal-Wallis test, with 95% confident level. The basic testing assumptions are listed as follow.

- At least one levels of age-range has different effect to the IPF mean values.
- Different genders, male and female, have different effects to the IPF mean values.

- Different statuses, normal and abnormal, have different effects to the IPF mean values.
- Different eye-sides, right eye and left eye, have different effects to the IPF mean values.

The results of these basic testing found that are two variables that do not have affect to the mean of IPF: the variables of genders and sides, with p -value $> 0.05 = \alpha$. On the other hand, other variables that can categorize the differences of the IPF mean values are age-range, statuses, and eye-sides, with p -value $< 0.05 = \alpha$. Therefore, 14 assumptions can be derived to determine the interactions among factors as listed below.

- At least one combination between ages and genders has significant different from others towards the mean values of the IPF.
- At least one combination between ages and statuses has significant different from others towards the mean values of the IPF.
- At least one combination between ages and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination between genders and statuses has significant different from others towards the mean values of the IPF.
- At least one combination between genders and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination between statuses and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination among ages, genders, and statuses has significant different from others towards the mean values of the IPF.
- At least one combination among ages, genders, and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination among ages, statuses, and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination among ages, genders, and statuses has significant different from others towards the mean values of the IPF.
- At least one combination among genders, statuses, and eye-sides has significant different from others towards the mean values of the IPF.
- At least one combination among ages, genders, statuses, and eye-sides has significant different from others towards the mean values of the IPF.

From the assumption above, the testing results discover that only the combination between genders and eye-sides has no significant difference in the mean values of the IPF, with p -value = $0.617 > 0.05 = \alpha$. However, the combination between ages and eye-sides have at least one significant difference between the IPF's mean values, with p -value = $0.015 < 0.05 = \alpha$. Furthermore, there is at least one significant difference among the IPF's mean values under the combination of ages, genders, statuses, and eye-sides with p -value = $0.000 < 0.05 = \alpha$.

4.3.2.2 PFL Consideration

This process is to analyze the value of the PFL using the non-parametric analysis, Kruskal-Wallis test, with 95% confident level. The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effect to the PFL mean values.
- Different genders, male and female, have different effects to the PFL mean values.
- Different statuses, normal and abnormal, have different effects to the PFL mean values.
- Different eye-sides, right eye and left eye, have different effects to the PFL mean values.

The results of these fundamental testing found that there are two variables that do not have effect to the PFL: genders and eye-sides, with $p\text{-value} > 0.05 = \alpha$. On the contrary, another variable that can categorize the differences of the PFL mean values is age-range and statuses, with $p\text{-value} < 0.05 = \alpha$. Thus, age-range and statuses are combining with others where the interactions among factors are listed below.

- At least one combination between ages and genders has significant different from others towards the PFL mean values.
- At least one combination between ages and eye-sides has significant different from others towards the PFL mean values.
- At least one combination between ages and statuses has significant different from others towards the PFL mean values.
- At least one combination between genders and statuses has significant different from others towards the PFL mean values.
- At least one combination between eye-sides and statuses has significant different from others towards the PFL mean values.
- At least one combination between ages, genders, and eye-sides has significant different from others towards the PFL mean values.
- At least one combination between ages, genders, and statuses has significant different from others towards the PFL mean values.
- At least one combination between ages, eye-sides, and statuses has significant different from others towards the PFL mean values.
- At least one combination between genders, eye-sides, and statuses has significant different from others towards the PFL mean values.
- At least one combination between ages, genders, eye-sides, and statuses has significant different from others towards the PFL mean values.

According to all assumptions above, the statistical results indicated that there is at least one significant difference among the mean values of PFL under combination of ages and genders, with $p\text{-value} = 0.012 < 0.05 = \alpha$; and there is at least one significant difference among the mean values of PFL under combination of ages and eye-sides, with $p\text{-value} = 0.027 < 0.05 = \alpha$. Furthermore, there is at least one significant difference among the mean values of PFL under ages, genders, statuses, and eye-sides, with $p\text{-value} = 0.000 < 0.05 = \alpha$.

4.3.2.3 PFR Consideration

The third process is to analyze the PFR value using non-parametric analysis, Kruskal-Wallis test, with 95% confident level. The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effect to the PFR mean values.
- Different genders, male and female, have different effects to the PFR mean values.
- Different statuses, normal and abnormal, have different effects to the PFR mean values.
- Different eye-sides, right eye and left eye, have different effects to the PFR mean values.

The results of these fundamental testing found that there are three variables that do not have effect to the PFR: age-range, genders, and eye-sides, with $p\text{-value} > 0.05 = \alpha$. On the other hand, another variable that can categorize the differences of the PFR mean values is statuses, with $p\text{-value} < 0.05 = \alpha$. Therefore, statuses are used to combine with other where the interactions among factors are listed below.

- At least one combination between ages and statuses has significant different from others towards the PFR mean values.
- At least one combination between genders and statuses has significant different from others towards the PFR mean values.
- At least one combination between eye-sides and statuses has significant different from others towards the PFR mean values.
- At least one combination between ages, genders and statuses has significant different from others towards the PFR's mean values.
- At least one combination between ages, eye-sides and statuses has significant different from others towards the PFR mean values.
- At least one combination between genders, eye-sides and statuses has significant different from others towards the PFR mean values.
- At least one combination between ages, genders, eye-sides and statuses has significant different from others towards the PFR mean values.

The results of the interaction testing discover that all assumptions above have significant different with $p\text{-value} = 0.00 < 0.05 = \alpha$.

4.3.2.4 IBR Consideration

The fourth process is to analyze the IBR value using non-parametric analysis, Kruskal-Wallis test, with 95% confident level. The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effect to the IBR mean values.
- Different genders, male and female, have different effects to the IBR mean values.

- Different statuses, normal and abnormal, have different effects to the IBR mean values.
- Different brow-sides, right brow and left brow, have different effects to the IBR mean values.

The results of these testing found that all combinations between ages, genders, statuses, and brow-sides have at least one significant difference in the mean values of the IBR, with $p\text{-value} = 0.00 < 0.05 = \alpha$; or, the combination between ages, genders, statuses, and brow-sides have effect to the mean values of the IBR. Thus, all assumptions can be derived as listed below.

- At least one combination between ages and genders has significant different from others towards the IBR mean values.
- At least one combination between ages and brow-sides has significant different from others towards the IBR mean values.
- At least one combination between ages and statuses has significant different from others towards the IBR mean values.
- At least one combination between genders and brow-sides has significant different from others towards the IBR mean values.
- At least one combination between genders and statuses has significant different from others towards the IBR mean values.
- At least one combination between brow-sides and statuses has significant different from others towards the IBR mean values.
- At least one combination between ages, genders and brow-sides has significant different from others towards the IBR mean values.
- At least one combination between ages, genders and statuses has significant different from others towards the IBR mean values.
- At least one combination between ages, brow-sides and statuses has significant different from others towards the IBR mean values.
- At least one combination between genders, brow-sides and statuses has significant different from others towards the IBR mean values.
- At least one combination between ages, genders, brow-sides and statuses has significant different from others towards the IBR mean values.

Based on the assumptions above, the testing results discover that the combination between ages and genders has at least one significant difference in the IBR mean values, with $p\text{-value} = 0.001 < 0.05 = \alpha$. Furthermore, all combinations among ages, genders, brow-sides, and statuses have at least one significant difference in the IBR mean values, with $p\text{-value} = 0.000 < 0.05 = \alpha$.

4.3.2.5 BL Consideration

This process aims to analyze the BL value using non-parametric analysis, Kruskal-Wallis test, with 95% confident level. The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effects to the BL mean values.

- Different genders, male and female, have different effects to the BL mean values.
- Different statuses, normal and abnormal, have different effects to the BL mean values.

The result of these basic testing found that there are two variables that do not have effect to the BL mean values, with $p\text{-value} > 0.05 = \alpha$; these variables are genders and statuses. On the other hand, the variables of ages and statuses can categorize the differences of the BL mean values, with $p\text{-value} < 0.05 = \alpha$. Thus, four assumptions can be derived as listed below.

- At least one combination between ages and genders has significant different from others towards the BL mean values.
- At least one combination between ages and statuses has significant different from others towards the BL mean values.
- At least one combination between genders and statuses has significant different from others towards the BL mean values.
- At least one combination between ages, genders, and statuses has significant different from others towards the BL mean values.

Based on the assumptions above, the testing results discover that only the combination between genders and statuses has no significant difference in the BL mean values, with $p\text{-value} = 0.057 > 0.05 = \alpha$. However, all combinations between ages and statuses have at least one significant difference between the BL mean values, with $p\text{-value} = 0.028 < 0.05 = \alpha$; and there is at least one significant difference among the BL mean values under the combination of ages and genders, with $p\text{-value} = 0.001 < 0.05 = \alpha$. Moreover, there is at least one significant difference between ages, genders, and statuses in the BL mean values, with $p\text{-value} = 0.001 < 0.05 = \alpha$.

4.3.2.6 LR Consideration

This process likes the IBR analysis where the non-parametric namely Kruskal-Wallis test with 95% confident level is applied. The basic testing assumptions are listed as follows.

- At least one levels of age-range has different effect to the LR mean values.
- Different genders, male and female, have different effects to the LR mean values.
- Different statuses, normal and abnormal, have different effects to the LR mean values.
- Different eye-sides, right and left eyes, have different effects to the LR mean values.
- Different positions, in and out positions, have different effects to the LR mean values.

The results of these fundamental testing found that there are three variables: eye-sides, statuses, and positions, that do not have effect towards the differences in LR mean values, with $p\text{-value} > 0.05 = \alpha$. On the other hand, other variables that can categorize the differences of the LR values are genders and ages, with $p\text{-value} < 0.05$

= α . Thus, genders and ages are used to combine with others where the interactions among factors are listed below.

- At least one combination between ages and genders has significant different from others towards the LR mean values.
- At least one combination between ages and statuses has significant different from others towards the LR mean values.
- At least one combination between ages and sides has significant different from others towards the LR mean values.
- At least one combination between ages and positions has significant different from others towards the LR mean values.
- At least one combination between genders and statuses has significant different from others towards the LR mean values.
- At least one combination between genders and sides has significant different from others towards the LR mean values.
- At least one combination between genders and positions has significant different from others towards the LR mean values.
- At least one combination between ages, genders and statuses has significant different from others towards the LR mean values.
- At least one combination between ages, genders and sides has significant different from others towards the LR mean values.
- At least one combination between ages, genders and positions has significant different from others towards the LR mean values.
- At least one combination between ages, genders, statuses and sides has significant different from others towards the LR mean values.
- At least one combination between ages, genders, statuses and positions has significant different from others towards the LR mean values.
- At least one combination between ages, genders, statuses, sides and positions has significant different from others towards the LR mean values.

The results of these interactions testing indicate that all combinations between genders, ages, statuses, sides, and positions have at least one significant difference between the LR mean values, with p -value = $0.000 < 0.05 = \alpha$. Therefore, all variables: ages, genders, statuses, sides, and positions, are applied to create the classification model.

4.3.2.7 LP Consideration

The last process is to analyze the LP value using non-parametric analysis, Kruskal-Wallis test, with 95% confident level. In addition, this process also includes links; those are the connection between the lower-left mouth to the upper-right mouth, the connection between the lower-middle mouth and the upper-middle mouth, and the connection between upper-left mouth to the lower-right mouth. Therefore, basic testing assumptions are listed as follow.

- At least one level of age-range has different effect to the LP mean values.

- Different genders, male and female, have different effects to the LP mean values.
- Different statuses, normal and abnormal, have different effects to the LP mean values.
- Different links, the first link, the second link, and the third link, have different effects to the LP mean values.

The results of these elementary testing found that all variables have effect to the LP mean values with $p\text{-value} < 0.005 = \alpha$. Thus, assumptions can be derived as listed below.

- At least one combination between ages and genders has significant different from others towards the LP mean values.
- At least one combination between ages and statuses has significant different from others towards the LP mean values.
- At least one combination between ages and links has significant different from others towards the LP mean values.
- At least one combination between genders and statuses has significant different from others towards the LP mean values.
- At least one combination between genders and links has significant different from others towards the LP mean values.
- At least one combination between statuses and links has significant different from others towards the LP mean values.
- At least one combination between ages, genders and statuses has significant different from others towards the LP mean values.
- At least one combination between ages, genders and links has significant different from others towards the LP mean values.
- At least one combination between genders, statuses and links has significant different from others towards the LP mean values.
- At least one combination between ages, genders, statuses, and links has significant different from others towards the LP mean values.

The results of these interactions testing indicated that all combinations between ages and genders have at least one significant difference between the LP mean values, with $p\text{-value} = 0.046 < 0.05 = \alpha$. Moreover, there is at least one significant difference among the LP mean values under the combination of ages and statuses, with $p\text{-value} = 0.000 < 0.05 = \alpha$. Furthermore, there is at least one significant difference among the LP mean values under the combination of genders and statuses, and genders and links, with $p\text{-value} = 0.000 < 0.05 = \alpha$. Therefore, all variables: ages, genders, statuses, and links, are applied to create the classification model.

Referring to all statistical results mention previously, there are relationships among parameters in Table 37. So, the relationship diagram that illustrates relationships between the dependent variable and the independent variables is drawn in Figure 41. From this figure, the direct effect to the change of Status is the Change of IPF/PFL/PFR/IBR/BL/LR/LP, while the indirect effects occur from Age, Side, Position, Link, and Gender.

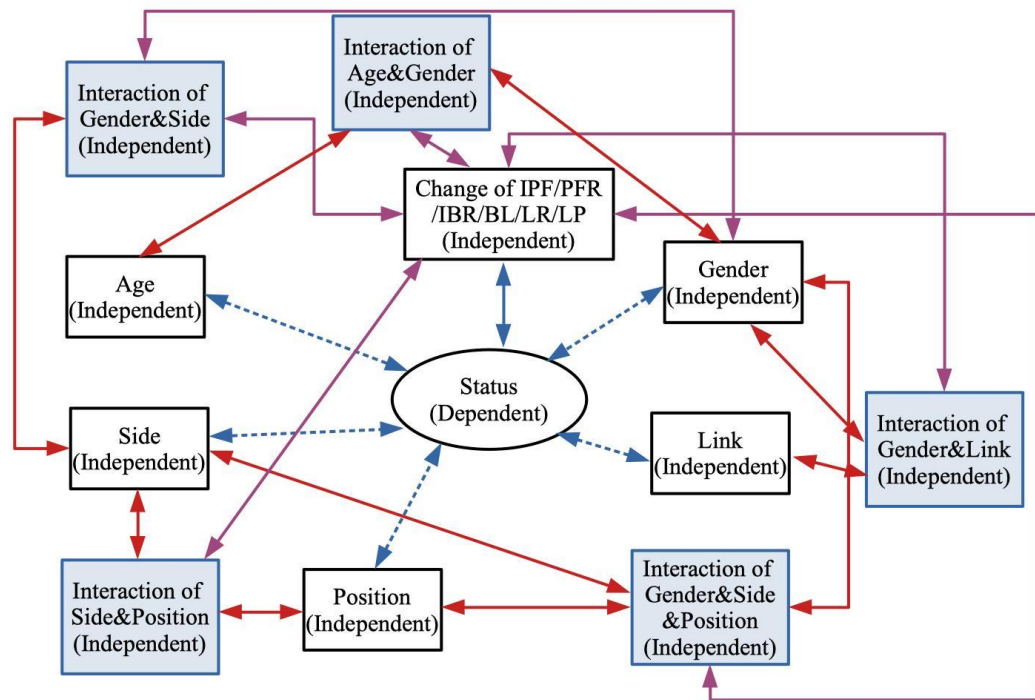


Figure 41 Relationships among variables

4.3.3 Machine Learning Results

4.3.3.1 Creating Model

After defining all parameters, the next step is to implement a model based on the discovered parameters. To accomplish the objective, this phase deploys the Scikit-Learn, a machine learning library to interpret all parameters via the PyCharm Community Edition version 2018.2 with Python language. The pattern is modified to a general classification model, as presented in Eq (4.4),

$$Status = M(\text{gender}, \text{age}, \text{side}, \text{position}, \text{link}, \text{IPF}, \text{PFL}, \text{PFR}, \text{IBR}, \text{BL}, \text{LR}, \text{LP}) \text{ -- (4.4)}$$

From Eq (4.4), *Status* is obtained from the model *M* which *M* refers to the model with parameters *gender*, *age*, *side*, *position*, *link*, *IPF*, *PFL*, *PFR*, *IBR*, *BL*, *LR*, and *LP*; the possible values of all parameters are defined in Table 4.22. The output of the *Status* classification model is either normal or abnormal, using a decision tree algorithm. Then, the derivation for patients' facial classification model starts.

Once the classification model was determined, 78 images from patients were imported to PyCharm Community Edition with Scikit-Learn.

As a result of grid search optimization, the dataset was separated into the training dataset and the testing dataset, with a ratio of 75:25. So, 20 images of the

entire dataset are the training dataset that contains the cluster attributes, for which the expected result after the learning process, namely a label, is obtained. Furthermore, the decision tree algorithm received as another outcome of the learning process is displayed in Figure 42.

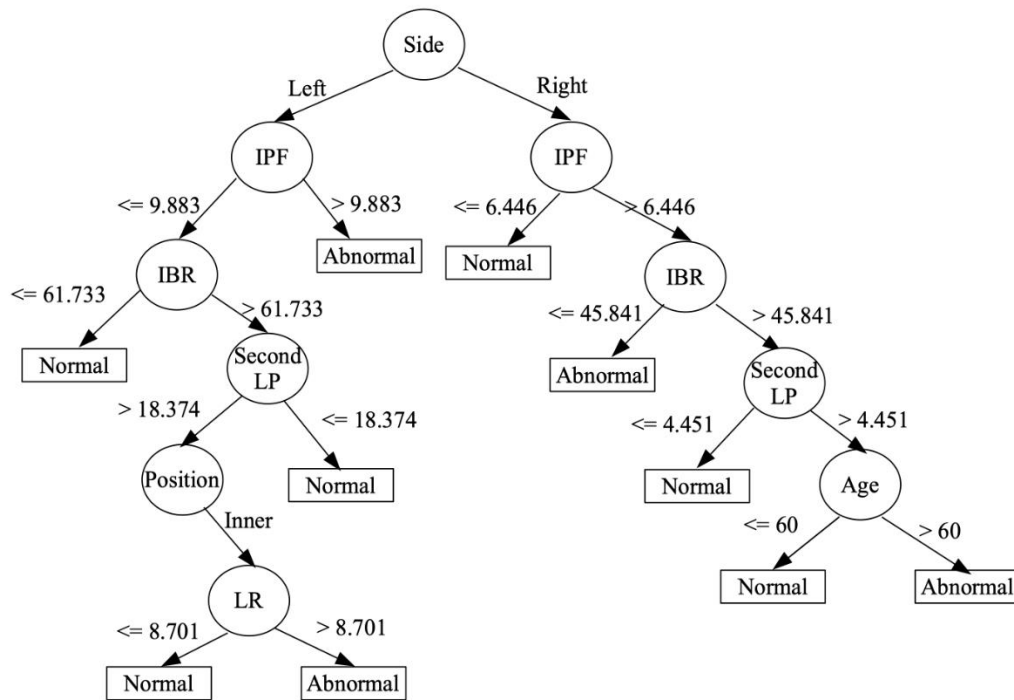


Figure 42 Decision Tree Algorithm for Status Classification

4.3.3.2 The *Status* simulation based on decision tree classification model

The training led to learn for the *Status* classification model, while the testing dataset was used to measure the performance of the *Status* classification model and displayed in the confusion matrix as shown in Table 38.

Table 38 Confusion Matrix of the *Status* Classification Decision Tree Model

Predict/Actual	Normal	Abnormal	Class Precision
Normal	10	1	91.00 %
Abnormal	0	9	91.00 %
Class Recall	100.00 %	100.00 %	

From the experimental results in Table 38, the number of True Positive (*TP*) that refers to the right prediction of the status from the model when comparing with the true value is 10. The number of False Positive (*FP*) which means the prediction is not correct is 1. The number of False Negative (*FN*) is always equal to *FP* is 0. The number of True Negative (*TN*) is always equal to *TP* is 9. Consequently, the precision value is 91.00%, the recall is 100.00%, and the accuracy of this model is 95.00%. Moreover, the comparison of the performance and classification error from the same

datasets with the decision tree model and random forest model is displayed in Table 39.

Table 39 Comparison of the results between Decision Tree and Random Forest Model

Model	Performance	Classification Error
Decision Tree	95.00 %	5.00 %
Random Forest	94.00 %	6.00 %

Figure 43 shows the receiver operating characteristic (ROC) curves, where the performance of the decision tree model, the red line, and the random forest model, the blue line, has been compared. So, it can be interpreted that the decision tree has better performance than the random forest.

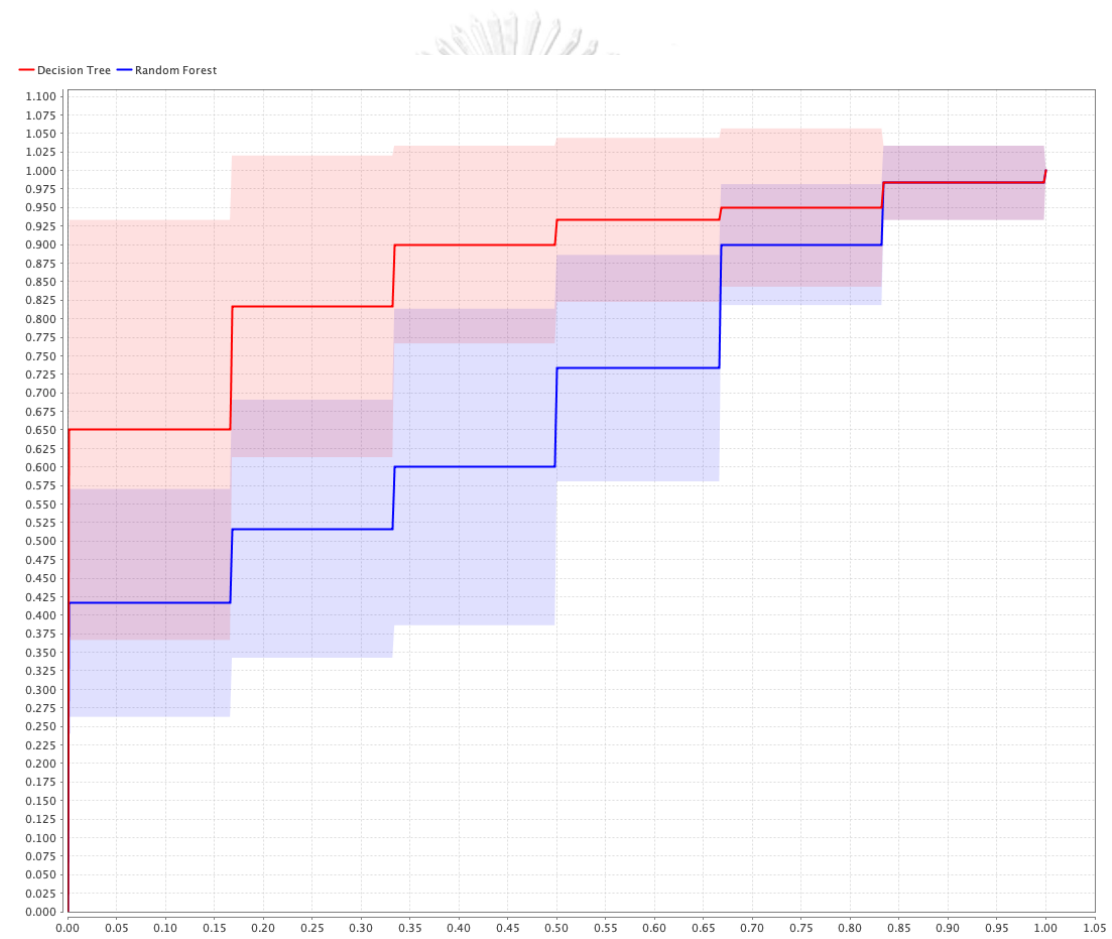


Figure 43 Comparison of Receiver Operating Characteristics (ROC) between Decision Tree and Random Forest

CHAPTER V

DISCUSSIONS AND CONCLUSIONS

This chapter composes of the discussions and conclusions in the dissertation; details of such topics are described in Section 5.1 and 5.2 respectively.

5.1 Discussions

Most stroke patients cannot move their bodies and speechless. Consequently, they cannot cry for help or push any button when they need, but they can express emotion through their faces. Thus, the aim of this research is to investigate facial features for creating a classification model that can differentiate between normal situations and abnormal situations. Therefore, the first phase of the preliminary study is to focus on gender, age-range, eye-side, interpalpebral fissure (IPF), palpebral fissure length (PFL), and palpebral fissure region (PFR). The results indicated that the parameters are customized to a general model as an emotional classification model between neutral emotion and fearful emotion. Then, the second phase of the preliminary study is to focus on gender, age-range, side, position, link, inner brow raisers (IBR), brow lower (BL), inner and outer lid raiser (LR), and lip part (LP) for the emotional classification model. The results showed that these parameters could classify emotion between normal emotion and fearful emotion. Then, the output of the classification model in two phases is either neutral or fearful, using a decision tree algorithm. These two phases use the dataset from horror-thriller movies on the Internet Movie Database (IMDb) website. Therefore, the experiment of the third phase applied all parameters from the first two stages mentioned previously. The results indicated that using all parameters could classify status between normal status and abnormal status with stroke patients in Chulalongkorn Comprehensive Stroke Center of Excellent, King Chulalongkorn Memorial Hospital. Comparison of all phases are presented in Table 40.

From Table 40, Phase III is divided into three different criteria based on the facial features to be applied. For the first criteria, there are 3 parameters to be involved which are IPF, PFL, and PFR. In the second criteria, the parameters are IBR, BL, LR, and LP, while the third phase many parameters, which are IPF, PFL, PFR, IBR, BL, LR, and LP. The accuracy of the first part is 94.82%. The second part is to apply the facial features as IBR, BL, LR, and LP from Phase II. The accuracy of the second part is 95%. The third part is to apply the facial features from Phase I and Phase II and it can obtain accuracy of 95.0%. Although the accuracy of using data in phase I is as high as 95% that is equivalence to the accuracy of Phase III, the image in Phase III uses only real stroke patients while Phase 1 uses only the figures from the films. Therefore, the most suitable set of variables for creating a classification model should contain all variables as same as Phase III.

Table 40 Facial features and accuracy of each phase

Phase	Facial Features	Accuracy
I	Interpalpebral Fissure (IPF), Palpebral Fissure Length (PFL), Palpebral Fissure Region (PFR)	95.00 %
II	Inner Brow Raiser (IBR), Brow Lower (BL), Lid Raiser (LR), Lip Part (LP)	83.33 %
III (Experiment)	Interpalpebral Fissure (IPF), Palpebral Fissure Length (PFL), Palpebral Fissure Region (PFR)	94.82 %
	Inner Brow Raiser (IBR), Brow Lower (BL), Lid Raiser (LR), Lip Part (LP)	95.00 %
	Interpalpebral Fissure (IPF), Palpebral Fissure Length (PFL), Palpebral Fissure Region (PFR), Inner Brow Raiser (IBR), Brow Lower (BL), Lid Raiser (LR), Lip Part (LP)	95.00 %

Therefore, the proposed algorithm can detect the facial paralysis of patients from facial features for normal situations and abnormal situations using decision tree. Moreover, the tool sets for facial detection have comfortable and safety for stroke patients, including it is not disturbing the treatment. After facial detection, the facial images are analyzed with facial features and classified two situations from decision tree algorithm as situation classification model. From the result of the experiment that decision tree is appropriate for classification of two states more than other algorithms. The accuracy of this model is 95%, it is better than the algorithms in related works. Furthermore, the precision of this research is 91%, and the recall is 100%.

5.2 Conclusions

The paralysis refers to the inability to move some part of the body's muscle when needed. For example, a stroke patient may be unable to move their arms and legs or speak. Unfortunately, the number of stroke patients is continuously increasing (Avan et al., 2019 and Areechokchai et al., 2017); in contrast, the number of caregivers is not commensurate with the increasing rate of patients, including the mental effects and burden on caregivers (Meretoja et al., 2017 and Sugii et al., 2019). Thus, these issues cause a negative impact on patients' safety. Therefore, a Smart Monitoring System (SMS) must be provided as soon as their lives are in danger; so, the quality of life of a patient is increase. Consequently, the idea of implementing small cameras to monitor a patient's expressions using the altering of all external elements of the eyes and the

sub-structures of the human face. The external elements of the eyes can be classified as the interpalpebral fissure (IPF), the palpebral fissure length (PFL), and the palpebral fissure region (PFR). The sub-structures comprise of the inner brow raiser (IBR), brow lower (BL), lid raiser (LR), and lip part (LP). These components are applied to detect and identify the situation either normal or abnormal.

The preliminary study of this thesis separates two phases; the first phase is to study the changes of the external elements of the eyes under neutral emotion and fear emotion. The second phase is to investigate the transformation of the sub-structures of the face under normal status and fear status. These two phases use the dataset from horror-thriller movies on the Internet Movie Database (IMDb) website. Thus, these parameters are used to classify the emotion differences between normal emotion and fearful emotion using a decision tree algorithm. The accuracy of the first phase was 95%, and the second phase was 83.33%.

Then, the third phase is the experiment with the external element and the sub-structures from a stroke patient, medical staff in Chulalongkorn Comprehensive Stroke Center of Excellent, King Chulalongkorn Memorial Hospital. As same two phases, the IPF, the PFR, the IBR, the BL, the LR, and the LP are applied to analyze the status differences between normal status and abnormal status using a decision tree algorithm. The accuracy of the third phase was 95% for classifying normal state and abnormal state. Furthermore, this research is published by the Department of Intellectual Property, Thailand (Bhattarakosol et al., 2019).

REFERENCES

Uncategorized References

Akash, A. A., et al. (2016). "Improvement of Haar Feature Based Face Detection in OpenCV Incorporating Human Skin Color Characteristic." J. Comput. Sci. Appl. Inf. Technol **1**(1): 1-8.

Amato, G., et al. (2018). A comparison of face verification with facial landmarks and deep features. Proceedings of the 10th International Conference on Advances in Multimedia (MMEDIA 2018).

Ancillao, A., et al. (2016). "Quantitative evaluation of facial movements in adult patients with hemiplegia after stroke." Int. J. Signal Image Process(2016): 1.

Areechokchai, D., et al. (2017). "Population attributable fraction of stroke risk factors in Thailand: Utilization of non-communicable disease surveillance systems." OSIR Journal **10**(1): 1-6.

Avan, A., et al. (2019). "Socioeconomic status and stroke incidence, prevalence, mortality, and worldwide burden: an ecological analysis from the Global Burden of Disease Study 2017." BMC medicine **17**(1): 191.

Avery, P., et al. (2011). Computational intelligence and tower defence games. 2011 IEEE Congress of Evolutionary Computation (CEC), IEEE.

Azzakhnini, S., et al. (2018). "Combining Facial Parts For Learning Gender, Ethnicity, and Emotional State Based on RGB-D Information." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) **14**(1s): 1-14.

Barbosa, J., et al. (2019). "paraFaceTest: an ensemble of regression tree-based facial features extraction for efficient facial paralysis classification." BMC medical imaging **19**(1): 1-14.

Benitez-Quiroz, C. F., et al. (2018). "Discriminant functional learning of color features for the recognition of facial action units and their intensities." IEEE transactions on pattern analysis and machine intelligence **41**(12): 2835-2845.

Chang, C.-Y., et al. (2018). Application of Machine Learning for Facial Stroke Detection. 2018 IEEE 23rd International Conference on Digital Signal Processing

(DSP), IEEE.

Cooksley, T., et al. (2018). "A systematic approach to the unconscious patient." Clinical Medicine **18**(1): 88.

Dixit, A., et al. (2017). "Emotion detection using decision tree." Development **4**(2): 145-149.

Ekman, P. and W. V. Friesen (1976). "Measuring facial movement." Environmental psychology and nonverbal behavior **1**(1): 56-75.

Ekman, R. (1997). What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS), Oxford University Press, USA.

Elmer, P., et al. (2015). Exploring compression impact on face detection using haar-like features. Scandinavian Conference on Image Analysis, Springer.

Goyal, K., et al. (2017). Face detection and tracking: Using OpenCV. 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), IEEE.

Guarin, D. L., et al. (2020). "Toward an Automatic System for Computer-Aided Assessment in Facial Palsy." Facial Plastic Surgery & Aesthetic Medicine **22**(1): 42-49.

Handelman, G. S., et al. (2019). "Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods." American Journal of Roentgenology **212**(1): 38-43.

Hansen, D. W. and Q. Ji (2009). "In the eye of the beholder: A survey of models for eyes and gaze." IEEE transactions on pattern analysis and machine intelligence **32**(3): 478-500.

Hao, Y., et al. (2017). "Predictors for symptomatic intracranial hemorrhage after endovascular treatment of acute ischemic stroke." Stroke **48**(5): 1203-1209.

Hsu, G.-S. J., et al. (2018). Hierarchical network for facial palsy detection. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE.

Hsu, G.-S. J., et al. (2018). "Deep hierarchical network with line segment learning for

quantitative analysis of facial palsy." IEEE access **7**: 4833-4842.

Hutter, F., et al. (2019). Automated machine learning: methods, systems, challenges, Springer Nature.

Kadir, K., et al. (2014). A comparative study between LBP and Haar-like features for Face Detection using OpenCV. 2014 4th International Conference on Engineering Technology and Technopreneuship (ICE2T), IEEE.

Kadri, F., et al. (2016). "Seasonal ARMA-based SPC charts for anomaly detection: Application to emergency department systems." Neurocomputing **173**: 2102-2114.

Kundu, T. and C. Saravanan (2017). Advancements and recent trends in emotion recognition using facial image analysis and machine learning models. 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICECCOT), IEEE.

Liu, D., et al. (2003). "Automatic mood detection from acoustic music data."

Lu, H., et al. (2018). "Brain intelligence: go beyond artificial intelligence." Mobile Networks and Applications **23**(2): 368-375.

Lucey, P., et al. (2011). Painful data: The UNBC-McMaster shoulder pain expression archive database. Face and Gesture 2011, IEEE.

Manfredonia, J., et al. (2019). "Automatic recognition of posed facial expression of emotion in individuals with autism spectrum disorder." Journal of autism and developmental disorders **49**(1): 279-293.

Melzer, J. E. and K. Moffit (2001). "Head-mounted displays." Digital Avionics Handbook.

Meretoja, A., et al. (2017). "Stroke doctors: who are we? A World Stroke Organization survey." International Journal of Stroke **12**(8): 858-868.

Michailidis, L., et al. (2019). Combining Personality and Physiology to Investigate the Flow Experience in Virtual Reality Games. International Conference on Human-Computer Interaction, Springer.

Moieni, A., et al. (2017). "Regression Facial Attribute Classification via simultaneous

dictionary learning." Pattern Recognition **62**: 99-113.

Morales-Vargas, E., et al. (2019). "On the use of Action Units and fuzzy explanatory models for facial expression recognition." Plos one **14**(10): e0223563.

Mozaffarian, D., et al. (2016). "Executive summary: heart disease and stroke statistics—2016 update: a report from the American Heart Association." Circulation **133**(4): 447-454.

Nasser, I. M. and S. S. Abu-Naser (2019). "Predicting Tumor Category Using Artificial Neural Networks."

Nasser, I. M., et al. (2019). "Artificial Neural Network for Diagnose Autism Spectrum Disorder."

Park, S.-W., et al. (2012). "Augmentative and alternative communication training using eye blink switch for locked-in syndrome patient." Annals of rehabilitation medicine **36**(2): 268.

Puklavage, C., et al. (2010). Imitating personalized expressions in an avatar through machine learning. Twenty-Third International FLAIRS Conference.

Rudnicki, W. R., et al. (2015). All relevant feature selection methods and applications. Feature Selection for Data and Pattern Recognition, Springer: 11-28.

Sajid, M., et al. (2018). "Automatic grading of palsy using asymmetrical facial features: A study complemented by new solutions." Symmetry **10**(7): 242.

Sajjad, M., et al. (2019). "Raspberry Pi assisted facial expression recognition framework for smart security in law-enforcement services." Information Sciences **479**: 416-431.

Salmam, F. Z., et al. (2016). Facial expression recognition using decision trees. 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV), IEEE.

Shim, H., et al. (2018). "Multi-tasking deep convolutional network architecture design for extracting nonverbal communicative information from a face." Cognitive Systems Research **52**: 658-667.

Shreve, M. A., et al. (2017). Human identity verification via automated analysis of

facial action coding system features, Google Patents.

Sondhi, P. (2009). "Feature construction methods: a survey." sifaka. cs. uiuc. edu **69**: 70-71.

Spitzer, M. (1998). "The history of neural network research in psychopathology." Neural networks and psychopathology: connectionist models in practice and research: 14.

Storey, G., et al. (2019). "3DPalsyNet: a facial palsy grading and motion recognition framework using fully 3D convolutional neural networks." IEEE access **7**: 121655-121664.

Sugii, N., et al. (2019). "Stroke mimics and accuracy of referrals made by emergency department doctors in Japan for patients with suspected stroke." Journal of Stroke and Cerebrovascular Diseases **28**(4): 1078-1084.

Summers, D., et al. (2009). "Comprehensive overview of nursing and interdisciplinary care of the acute ischemic stroke patient: a scientific statement from the American Heart Association." Stroke **40**(8): 2911-2944.

Sun, X., et al. (2018). "Face detection using deep learning: An improved faster RCNN approach." Neurocomputing **299**: 42-50.

Sutton, R. S. and A. G. Barto (2018). Reinforcement learning: An introduction, MIT press.

Takalkar, M., et al. (2018). "A survey: facial micro-expression recognition." Multimedia Tools and Applications **77**(15): 19301-19325.

Umirzakova, S., et al. (2019). Fully Automatic Stroke Symptom Detection Method Based on Facial Features and Moving Hand Differences. 2019 International Symposium on Multimedia and Communication Technology (ISMAT), IEEE.

Umirzakova, S. and T. K. Whangbo (2018). STUDY ON DETECT STROKE SYMPTOMS USING FACE FEATURES. 2018 International Conference on Information and Communication Technology Convergence (ICTC), IEEE.

Vinay, A., et al. (2019). Face Recognition using Invariant Feature Vectors and Ensemble of Classifiers. Soft Computing and Signal Processing, Springer: 769-779.

Wang, N., et al. (2018). "Facial feature point detection: A comprehensive survey." Neurocomputing **275**: 50-65.

Wintermark, M. and T. Rizvi (2018). Principles of Clinical Diagnosis of Hemorrhagic Stroke. Stroke Revisited: Hemorrhagic Stroke, Springer: 109-132.

Xu, F., et al. (2017). "Microexpression identification and categorization using a facial dynamics map." IEEE Transactions on Affective Computing **8**(2): 254-267.

Xu, X. and I. A. Kakadiaris (2017). Joint head pose estimation and face alignment framework using global and local CNN features. 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), IEEE.

Zafeiriou, S., et al. (2015). "A survey on face detection in the wild: past, present and future." Computer Vision and Image Understanding **138**: 1-24.

Zhang, Y., et al. (2018). Classifier learning with prior probabilities for facial action unit recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Zöllner, J. P., et al. (2020). "National Institutes of Health Stroke Scale (NIHSS) on admission predicts acute symptomatic seizure risk in ischemic stroke: a population-based study involving 135,117 cases." Scientific Reports **10**(1): 1-7.



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

VITA

NAME Rawinan Praditsangthong

DATE OF BIRTH 13 August 1980

PLACE OF BIRTH Bangkok

INSTITUTIONS ATTENDED She received a Bachelor's degree in Computer Science from Rangsit University and Master of Information Technology from Rangsit University. Now she is a Doctor of Philosophy in Computer Science and Information Technology, Faculty of Science, Chulalongkorn University.

HOME ADDRESS 24 Soi. Rangsit-Pathum 12 soi 5. Moo.2. Rangsit-Pathum Road. Prachathipat Thanyaburi Pathum-thani 12130

PUBLICATION M-Learning System: A mobile learning for asthma,
A fear detection method based on palpebral fissure