MULTIMODAL PERSON RECOGNITION SYSTEM USING CONVOLUTIONAL NEURAL NETWORK

Jagathis A/L Karunakaran

Bachelor of Engineering

Electrical and Electronics Engineering with Honours

2022

UNIVERSITI MALAYSIA SARAWAK

Grade: _____

Please tick (√)

Final Year Project Report ☑

Masters ☐

PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the ……29……..day of……July……..2022.

Student's Declaration:

I Jagathis A/L Karunakaran, 66213, Faculty of Engineering

(PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled Multimodal Person Recognition System Using Convolutional Neural Network is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

29/07/2022
Date submitted

Jagathis A/L Karunakaran (66213)
Name of the student (Matric No.)

Supervisor's Declaration:

I Dr Annie anak Joseph (SUPERVISOR'S NAME) hereby certifies that the work entitled Multimodal Person Recognition System Using Convolutional Neural Network (TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of Bachelor of Engineering in Electrical and Electronics Engineering with Honours (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.

Received for examination by: Dr Annie anak Joseph
(Name of the supervisor)

Date: 29/07/2022

I declare that Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*****
☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*****
☑ **OPEN ACCESS**

Validation of Project/Thesis

I therefore duly affirm with free consent and willingly declare that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abiding interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies for the purpose of academic and research only and not for other purpose.
- The Centre for Academic Information Services has the lawful right to digitalise the content for the Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic exchange between Higher Learning Institute.
- No dispute or any claim shall arise from the student itself neither third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student except with UNIMAS permission.

Student signature _____ (Date) 29/07/2022

Supervisor signature: _____ (Date)

Current Address:
Faculty of Engineering, University of Malaysia Sarawak, Jalan Datuk Mohammad Musa, 94300 Kota Samarahan, Sarawak.

Notes: ***** If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the period and reasons of confidentiality and restriction.

[The instrument is duly prepared by The Centre for Academic Information Services]

MULTIMODAL PERSON RECOGNITION SYSTEM USING CONVOLUTIONAL NEURAL NETWORK

JAGATHIS A/L KARUNAKARAN

A dissertation submitted in partial fulfilment
of the requirement for the degree of
Bachelor of Engineering
Electrical and Electronics Engineering with Honours

Faculty of Engineering
Universiti Malaysia Sarawak

2022

ACKNOWLEDGEMENT

I would like to convey my profound appreciation to everyone who has assisted me throughout this research project. A special gratitude I give to Dr Annie Joseph, the project supervisor for her advice and inputs. The supervisor has provided me with helpful information and insights to this research project based on her expertise as a lecturer and engineering. Thanks to Final Year Project course coordinator, Dr Dayang Nur Salmi Dharmiza binti Awang Salleh, I was able to correctly organise and finish my project responsibilities on time despite my hectic schedule. Throughout I was taught on how to be an excellent engineering student with what it takes to be a future engineer and possess a strong professional ethics. I am also grateful to my friends for their aid in developing the research report from a third-person perspective by giving their thoughts and comments. They were willing to take the time to listen to my ideas and provide input based on their knowledge. Finally, I would like to acknowledge and express my gratitude to my family for providing me resources and support when I needed it most. I was able to devote my attention and effort towards the research project because they provided me the required space and encouragement.

ABSTRACT

Multimodal person recognition system in the current age is an alternative to unimodal biometric system. Recognition systems routinely used for person identification in various segments and commonly resort to daily life use. In this project, two pretrained Convolutional Neural Networks were trained, tested and evaluated using MATLAB R2021b. The networks are AlexNet with 25 layers and VGG16 with 16 layers. Then an optimal network was opted, AlexNet and used to train face and fingerprint dataset separately with different variation of hyperparameter. The face and fingerprint dataset used to train and test the networks are self-created face dataset and NIST Special Database 302 fingerprint dataset. The accuracy of the AlexNet for face and fingerprint were 95.00% and 98.67% respectively. The AlexNet model was evaluated with an accuracy of 98.67% and 100% in the 5-fold Validation test. The accuracy of the confusion matrix was the same for the AlexNet networks of face and fingerprint. The face and fingerprint networks were later fused in decision-level fusion to produce an overall multimodal recognition network. AlexNet has an average high accuracy of 96.84% setting a high standard for future work in multimodal person recognition suggesting AlexNet as an effective pretrained network for classification.

ABSTRAK

Sistem pengecaman orang multimodal pada zaman sekarang adalah alternatif kepada sistem biometrik unimodal. Sistem pengecaman yang digunakan secara rutin untuk mengenal pasti orang dalam pelbagai segmen dan biasanya menggunakan penggunaan kehidupan harian. Dalam projek ini, dua Rangkaian Neural Konvolusi yang telah dilatih telah dilatih, diuji dan dinilai menggunakan MATLAB R2021b. Rangkaian tersebut ialah AlexNet dengan 25 lapisan dan VGG16 dengan 16 lapisan. Kemudian rangkaian optimum telah dipilih, AlexNet dan digunakan untuk melatih set data muka dan cap jari secara berasingan dengan variasi hiperparameter yang berbeza. Set data muka dan cap jari yang digunakan untuk melatih dan menguji rangkaian ialah set data muka ciptaan sendiri dan set data cap jari Pangkalan Data Khas NIST 302. Ketepatan AlexNet untuk muka dan cap jari masing-masing adalah 95.00% dan 98.67%. Model AlexNet dinilai dengan ketepatan 98.67% dan 100% dalam ujian Pengesahan 5 kali ganda. Ketepatan matriks kekeliruan adalah sama untuk rangkaian muka dan cap jari AlexNet. Rangkaian muka dan cap jari kemudiannya digabungkan dalam gabungan peringkat keputusan untuk menghasilkan rangkaian pengecaman multimodal keseluruhan. AlexNet mempunyai purata ketepatan tinggi sebanyak 96.84% menetapkan standard yang tinggi untuk kerja masa hadapan dalam pengiktirafan orang berbilang mod yang mencadangkan AlexNet sebagai rangkaian pralatihan yang berkesan untuk pengelasan.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

2D	–	Two-Dimensional
3D	–	Three-Dimensional
ANN	–	Artificial Neural Network
CNN	–	Convolutional Neural Network
FC	–	Fully Connected
FYP	–	Final Year Project
GPU	–	Graphics Processing Unit
GUI	–	Graphical User Interface
IEEE	–	Institute of Electrical and Electronics Engineers
LBP	–	Local Binary Pattern
MATLAB	–	Matrix Laboratory
NN	–	Neural Network
PIN	–	Personal Identification Number
PNN	–	Probabilistic Neural Network
PCA	–	Principal Component Analysis
ReLU	–	Rectified Linear Unit
ResNet	–	Residual Network
SVM	–	Support Vector Machine
VGG16	–	Visual Geometry Group 16
VGGNet	–	Visual Geometry Group Network

Chapter 1

INTRODUCTION

1.1 Background

Since the dawn of modern computer application, the concepts of digital authenticators and person recognition system have been a fundamental part of digital security evolution. Secure access has been enabled in many applications, such as smartphone, banks, and government related procedures due to security related policies and reasons to preserve information and details of a personal as per law implementation. In conjunction with this, person recognition system was introduced during industrial revolution and further elevated due to its prominence for users and by users. In simple words, person recognition system is an application used to recognise and identify a person using certain authentication techniques.

In the older norms, passwords, PIN, and patterns were used as recognition technique in unlocking digital devices. However, existing person recognition system was upgraded with the use human physical traits such as fingerprints of human that is also known as biometric authentication. The biometric were later alternated with physical traits of human such as face, iris, and voice. The existing person recognition systems were labelled as unimodal recognition system due to its singular use of human biometric. The existing unimodal systems are widely utilised in commercial applications, government administrations, and forensic applications such as biometric authentication on mobile phones, national ID cards, and identifying criminals respectively [1].

Unimodal systems were received greatly by many parties; however, the technology evolution has made the system weak, the unimodal biometric system fails due to several flaws and limitations [1]. The biometric(s) such as fingerprint and voice were easily imitated by offenders. Even though the computer and software can decipher the biometric, but it was easily exploited with third party tool for identity bypass related purposes. The implementation of unimodal system could not withstand duplicates due to only single

biometric use, and security breaches and scams rise. This is when multimodal person recognition system was introduced.

A multimodal recognition system integrates two or more biometric features with the raw data of the multiple biometrics traits [1]. The multimodal person recognition works similarly to unimodal system, but it has two or more input data and the algorithm developed has more precision. Biometric identification can be either one-dimensional or multi-dimensional. Thus, combining various sources of biometric data creates a multibiometric recognition system that could overcome certain weakness exists in unimodal systems [2]. Multimodal systems are greater due to high performance, greater robustness, and higher resistance to spoofing [3]. In current generation, the multimodal person recognition systems are vastly used. Thus, a multimodal system that uses face and fingerprint could play a major role in recognising patterns and image processing. Facial and fingerprint recognition technology are the techniques used to identify an individual based on their facial characteristics and fingerprints respectively. In identity verification, facial recognition is the most inherent and accepted method [3], whereas fingerprint has great accuracy and user friendly. A general multimodal person recognition system has the following architecture as depicted in Figure 1.1.

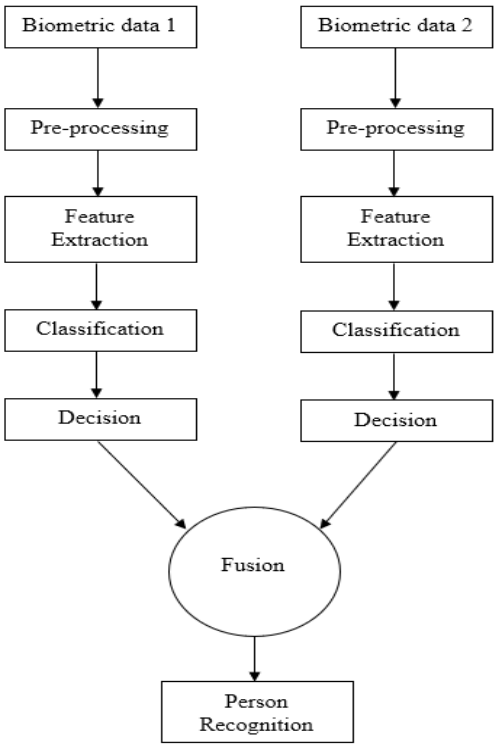

Figure 1.1: The multimodal person recognition system's overall structure [1]

In general, the basic unimodal recognition system using neural network has three stages known as pre-processing, feature extraction and classification. However, based on Figure 1.1, it can be seen a person recognition system with two biometrics can be identified through two separate detection and feature extraction for face and fingerprint respectively, moving onto classification and then fused together in through implying suitable fusion technique. The features are classified from picture sequences or static images, then the person is recognised as an output. In last few years' time, deep learning approaches have been implemented to summarise such complicated and technical procedures to form an algorithm.

Lately, convolutional neural network (CNN) which is the subset of deep learning method is used to construct multimodal recognition system [2]. In contrast to ordinary neural networks, image-based neural networks use pixels of an input image [3]. In this paper, the aim to implement a multimodal person recognition system by applying deep learning approach that could analyse pixel input for image processing and recognition.

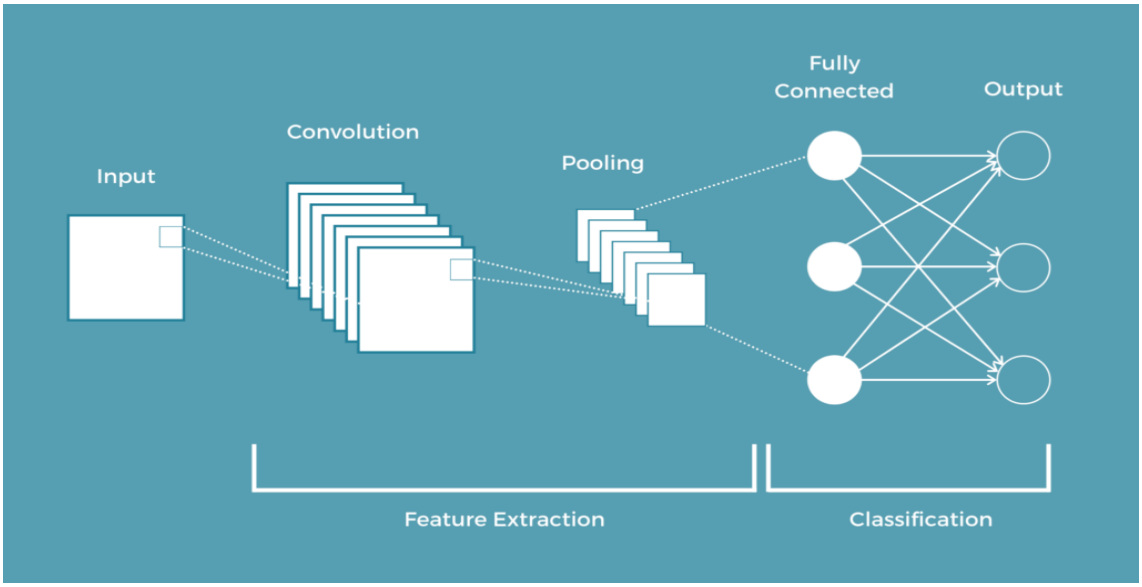

Figure 1.2: Convolutional Neural Network

In this research, a multimodal person recognition system based on computer and sophisticated technique of CNN, will evaluate faces and fingerprints in photos through mathematical algorithms. The artificial neural network is invented with numerous levels or layers, each of which can include one or more neurons [3], and uses convolution to extract characteristics from pictures while keeping spatial correlation [4].

1.2 Problem Statement

Recognising people for authentication purpose are important and needs accuracy for true recognition. Multiple types of recognition system exist in current electronic driven century involving traditional authenticator, digital authenticator, and systems through human-computer interaction (HCI). However, there is always a different accuracy between digital authenticator based on physiological or behavioural characteristics and conventional ways of passwords, signature, and fingerprint.

People are still using common authenticator such as password, pin number and pattern recognition. For example, computer logins are still highly dependent on password and pin number. The credentials are easily guessed, weak and does not provide sufficient protection against vulnerabilities. The vast growth of technology introduced the biometric recognition system that involves fingerprint, voice, and iris. The government and any legal procedure that requires identity verification are still highly using fingerprint as recognition system.

Problem arises when common authenticators and biometrics such as fingerprint and voice that computers and software can decipher are being exploited for use of identity bypass related purposes. What are some of the ways that developing a person recognition system might benefit a variety of sectors? Firstly, a recognition system built with convolutional neural network enhances the security measures through automated recognition with increased efficiency and accuracy due to complex nature of the network making it achieve sophistication through network training. Even with fast and easy recognition systems available, the necessity for well-secured identity and personal recognition technology is increasingly obvious as the number of security breaches and scams rises.

Next, the algorithm developed avoids imitations of a personal's features such as fingerprint, signature, and voice that are easily imitated through implementation of unimodal biometric systems. Unimodal biometric system only has single metric of accuracy delivered where it can vary from low to high, making the system less dependable. Thus, through multimodal recognition, the individuality of a person makes it difficult to discover another person with similar traits that would match your digital print and have second metric of accuracy that could compensate if the first biometric could not produce high accuracy. It is also important to create a multimodal system without

complicated structures through integration of various techniques. This is due to available multimodal person recognition that is not completely based on deep learning are built through various integration of separately introduced third party techniques. The architecture varies widely which has an occurrence of relatively different recognition rates and inconsistent accuracy that could be enhanced through deep learning.

1.3 Objectives

The main goal of this study is to create an algorithm that can recognise person using multimodal biometrics. The following research aims will help to achieve this goal:

- To investigate the existing works related to person recognition system
- To develop an algorithm of multimodal person recognition system using convolutional neural network
- To validate the algorithm and compare to an existing person recognition system

1.4 Project Motivation

The motivation of the project is to develop a multimodal person recognition system using deep learning method by eliminating the problems faced in existing algorithms. Besides, the project motivates to solve recognition issues faced by users in a daily life use of available recognition systems widely. Throughout the project, it is important to understand a user's perspective while implementing an algorithm that is better than previous works in order future works are based on this project's evolution.

1.5 Project Scope

The overall project is about a multimodal person recognition system that can be used by personals associated for secure access requiring personal identification in various segments and sectors. The ability of the system is to recall information or a person's characteristics in order to later identify the individual, who has saved their information in the database. This multimodal person recognition system uses two physical biometric traits of a human, thus face and fingerprint are intended to be included in this system expansion. The algorithm is developed using convolutional neural network, a deep learning approach with the use of MATLAB and C++ programming language.

The project is expected to provide various outcomes, one such is the extensive literature study. Literature study is important to any research and this context all the related works from 2017-2021 is thoroughly studied and reviewed to provide a bigger picture of all aspects involved in this project. Through algorithm design and implementation, an algorithm for the multimodal person recognition system is expected. As mentioned earlier, the algorithm is on the grounds of deep learning approach with complex computational techniques. The algorithm must be able to extract features from inserted image accurately and matched with the system's database. Moreover, the efficiency is prioritised too, where it should work better than existing available algorithms.

Next, there are certain acceptance criteria to be ticked before the project is declared successful. Firstly, the algorithm could detect the biometric inputs, such as face and fingerprint then extract their features. This is very crucial as without achieving the basic function of the algorithm, the project will not progress to next category of algorithm development. The algorithm produced must has higher recognition rate in term of accuracy and duration. Accuracy is highly anticipated for this recognition system as it acts as the core of the system, then moving on duration, the recognition process should only take seconds. This recognition duration includes time of input, system processing time and the time to display recognition result. The multimodal person recognition system should be user-friendly, so that users are convenient to use anytime and anywhere if required without much of hassle.

A project will have many exclusions to avoid inefficient methods and mistakes highlighted through literature review. In this case, the development of algorithm did not implement traditional authenticators. The authenticators such as pin and password are outdated for current industrial revolution, instead human physical traits are integrated for much secure and reliable authentication. Next, the existence of neural network is widely exploited, and initial recognition systems use complicated neural network structure. There were combinations of multiple techniques of pre-processing, feature extraction, and classification. However, this issue is not expected to recur in this project due to advanced neural networks that are simplified and open-ended.

Finally, as the project's design and implementation are expected to take off, there are few constraints to be faced and requires a fix. The algorithm requires face and fingerprint dataset for network training and test, however there are limited resources in the online