



UNIVERSITI TEKNIKAL MALAYSIA MELAKA

**EMOTIONAL RECOGNITION VIA SUBTLE HUMAN
INTERACTION (FACIAL EXPRESSION) USING
CONVOLUTIONAL NEURAL NETWORK**

This report is submitted in accordance with the requirement of the Universiti Teknikal Malaysia Melaka (UTeM) for the Bachelor of Electrical Engineering Technology (Industrial Automation & Robotics) with Honours

by

KOAY XIAN HONG

B071510264

950706-07-5485

FACULTY OF ELECTRICAL AND ELECTRONIC ENGINEERING
TECHNOLOGY

2019

BORANG PENGESAHAN STATUS LAPORAN PROJEK SARJANA MUDA

Tajuk: EMOTIONAL RECOGNITION VIA SUBTLE HUMAN INTERACTION
(FACIAL EXPRESSION) USING CONVOLUTIONAL NEURAL NETWORK

Sesi Pengajian: 2019

Saya **KOAY XIAN HONG** mengaku membenarkan Laporan PSM ini disimpan di Perpustakaan Universiti Teknikal Malaysia Melaka (UTeM) dengan syarat-syarat kegunaan seperti berikut:

1. Laporan PSM adalah hak milik Universiti Teknikal Malaysia Melaka dan penulis.
2. Perpustakaan Universiti Teknikal Malaysia Melaka dibenarkan membuat salinan untuk tujuan pengajian sahaja dengan izin penulis.
3. Perpustakaan dibenarkan membuat salinan laporan PSM ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. **Sila tandakan (X)

SULIT*

Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia sebagaimana yang termaktub dalam AKTA RAHSIA RASMI 1972.

TERHAD*

Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan.

TIDAK TERHAD

Yang benar,

Disahkan oleh penyelia:

.....
KOAY XIAN HONG

.....
IRIANTO

Alamat Tetap:
2-12, Jalan Bukit Belah Satu,
Krystal County Homes,
11900 Bayan Lepas, Pulau Pinang
Tarikh:

Cop Rasmi Penyelia

Tarikh:

*Jika Laporan PSM ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan PSM ini

DECLARATION

I hereby, declared this report entitled EMOTIONAL RECOGNITION VIA SUBTLE HUMAN INTERACTION (FACIAL EXPRESSION) USING CONVOLUTIONAL NEURAL NETWORK is the results of my own research except as cited in references.

Signature:

Author : KOAY XIAN HONG

Date:

APPROVAL

This report is submitted to the Faculty of Mechanical and Manufacturing Engineering Technology of Universiti Teknikal Malaysia Melaka (UTeM) as a partial fulfilment of the requirements for the degree of Bachelor of Electrical Engineering Technology (Industrial Automation & Robotics) with Honours. The member of the supervisory is as follow:

Signature:

Supervisor : IRIANTO

ABSTRAK

Pembelajaran mesin telah berkembang pesat baru-baru ini. Terutama dengan kewujudan GPU dan CPU yang lebih baik (yang membolehkan kuasa pengiraan yang lebih baik), pembelajaran mendalam, teknologi pembelajaran mesin yang memerlukan kuasa pengkomputeran yang besar, kini boleh memproses dengan berkesan sejumlah besar data yang kita ada sekarang. Ini membolehkan pelbagai aplikasi, salah satunya boleh digunakan untuk penglihatan mesin. Menggunakan teknologi pembelajaran mesin, projek ini ditubuhkan untuk mewujudkan satu program yang belajar untuk mengenali emosi melalui interaksi manusia yang halus menggunakan kerumitan ekspresi wajah. Satu akan dapat melaksanakan program tersebut dalam kehidupan seharian menggunakan penilaian masa nyata, membolehkannya mengesan tahap awal kemurungan, menyediakan persahabatan kepada orang tua dan bantuan dalam industri perkhidmatan. Menggunakan Rangkaian neural Convolutional dan algoritma pembelajaran mendalam, projek ini berjaya mewujudkan satu program yang mengiktiraf ekspresi manusia dan mengenal pasti emosi dalam masa nyata. Hasil projek ini menunjukkan sejumlah janji, projek mencapai ketepatan pengesanan 40% hingga 50%. Walau bagaimanapun, penilaian masa sebenar menggunakan model terlatih adalah agak tepat pada masanya, ketepatannya jatuh secara drastik untuk manusia dengan cermin. Dengan lebih banyak masa dan sumber, adalah mungkin untuk memperbaiki projek ini lebih lanjut daripada apa yang sedang berlaku sekarang.

ABSTRACT

Machine learning have been booming recently. Especially with the existence of better GPU and CPU (which allow better computational power), deep learning, a machine learning technology that need the large computational power, can now effectively process the large amount of data that we currently have. This allow wide array of applications, one of which can be apply to machine vision. Using the technology of machine learning, this project set out to create a program that learn to recognize emotion via subtle human interaction using the intricacy of facial expression. One would be able to implement said program in daily life using real time evaluation, allowing it to detect early stage of depression, provide companionship to the elderly and aid in service industry. Using Convolutional Neural Network and deep learning algorithm, this project managed to create a program that recognize human expression and identify the emotion in real time. The result of this project shows certain amount of promises, the project achieved a validation accuracy of 40% to 50%. Despite that, the real time evaluation using the trained model is quite accurate most of the time, the accuracy drops drastically for human with spectacles, however. With more time and resources, it is possible to improve this project further than what currently is.

DEDICATION

To my beloved parents, my mother, Ooi Lay Hoon and my father, Koay Ting Hoo.
Thank you for unconditional love that you all have shower me with. Thank you for the patient upbringing. Thank you for the education you provided me with. Thank you for creating what am I today.

To my siblings, thank you for supporting and helping me throughout the project. Thank you for aiding me on things outside this project which allows me to concentrate on the project.

ACKNOWLEDGEMENTS

I like to thanks to all those who supported, encouraged, challenged, and inspired me. A special thanks to my supervisor, Sir Irianto for his guidance and encouragement. Thanks to my senior and my mentor for supporting me and aiding me in the problem that I couldn't solve on my own. Thanks also to my family, which provided me with the resources I need to continue with the project.

TABLE OF CONTENTS

	PAGE
TABLE OF CONTENTS	ix
LIST OF TABLES	xiv
LIST OF FIGURES	xvi
LIST OF APPENDICES	xviii
LIST OF SYMBOLS AND ABBREVIATIONS	xix
CHAPTER 1 INTRODUCTION	1
1.1 Background	1
1.2 Problem Statement	3
1.3 Objectives	3
1.4 Project scope	4
CHAPTER 2 LITREATURE REVIEW	5
2.1 Deep Learning	5
2.2 Artificial Neural Network	6
2.3 Learning paradigms	8
2.3.1 Supervised learning	8
2.3.2 Unsupervised learning	9

2.3.3	Reinforcement learning	10
2.4	Backpropagation Algorithm	11
2.5	Activation function	17
2.5.1	Sigmoid function	17
2.5.2	Hyperbolic Tangent (Tanh)	18
2.5.3	Rectified Linear Unit (ReLU)	20
2.5.4	Leaky ReLU	22
2.6	Dropout	23
2.7	Softmax cross entropy	23
2.8	Convolutional Neural Network (CNN)	24
2.8.1	Convolutional Layer	25
2.8.2	Pooling layer	26
2.8.3	Fully Connected Layer	28
CHAPTER 3 METHODOLOGY		29
3.1	Project Methodology	29
3.2	Project Software	30
3.2.1	Python programming language	30
3.2.2	TensorFlow	32
3.2.3	OpenCV	33
3.3	Project Flow	34
3.3.1	Data acquisition and processing	35

3.3.2	Designing a Convolutional Neural Network (CNN)	37
3.3.3	Training the CNN	38
3.3.4	Validation Test	39
3.3.5	Live Video Feed Design	39
CHAPTER 4	RESULT & DISCUSSION	41
4.1	Summary	41
4.1.1	Result, analysis and discussion	41
4.2	First attempt	43
4.2.1	Train and Validation dataset	43
4.2.2	CNN design	44
4.2.3	Training the CNN and Result	45
4.2.4	Loss Graph	45
4.2.5	Observation	46
4.2.6	Analysis and discussion	46
4.3	Second attempt	47
4.3.1	Train and Validation dataset	48
4.3.2	CNN design	49
4.3.3	Training the CNN and Result	50
4.3.4	Accuracy throughout the training	50
4.3.5	Final Accuracy Table	51
4.3.6	Observation	51

4.3.7	Analysis and discussion	52
4.4	Third, Fourth and Fifth Attempt	53
4.4.1	Train and Validation dataset	54
4.4.2	CNN design	56
4.4.3	Training the CNN and Result	57
4.4.4	Accuracy throughout the training	58
4.4.5	Final Accuracy Table	59
4.4.6	Observation	61
4.4.7	Analysis and discussion	62
4.5	Sixth attempt	64
4.5.1	Train and Validation dataset	65
4.5.2	CNN design	66
4.5.3	Training the CNN and Result	67
4.5.4	Accuracy throughout the training	67
4.5.5	Final Accuracy Table	68
4.5.6	Observation	68
4.5.7	Analysis and discussion	69
4.6	Seventh attempt	70
4.6.1	Train and Validation dataset	71
4.6.2	CNN design	72
4.6.3	Training the CNN and Result	73
4.6.4	Accuracy throughout the training	73
4.6.5	Final Accuracy Table	74

4.6.6	Observation	74
4.6.7	Analysis and discussion	75
4.7	Eighth attempt	76
4.7.1	Train and Validation dataset	76
4.7.2	CNN design	77
4.7.3	Training the CNN and Result	78
4.7.4	Accuracy throughout the training	78
4.7.5	Final Accuracy Table	79
4.7.6	Observation	79
4.7.7	Analysis and discussion	80
4.8	Video feed and real-time evaluation architecture	81
4.7.1	Video feed test and the result	81
CHAPTER 5 CONCLUSION		83
5.1	Summary of Result and Research	83
5.2	Possible improvement for the future	86
REFERENCES		88
APPENDIX		92

LIST OF TABLES

TABLE	TITLE	PAGE
Table 4.1.1:	Goal of each attempt	42
Table 4.2.1:	Attempt 1 dataset	43
Table 4.2.3:	Training information and Result for attempt 1	45
Table 4.3.1:	Attempt 2 dataset	48
Table 4.3.3:	Training information and Result for attempt 2	50
Table 4.3.5:	Accuracy table for attempt 2	51
Table 4.4.1.1:	Attempt 3 dataset	54
Table 4.4.1.2:	Attempt 4 dataset	54
Table 4.4.1.3:	Attempt 5 dataset	55
Table 4.4.3.1:	Training information and Result for attempt 3	57
Table 4.4.3.2:	Training information and Result for attempt 4	57
Table 4.4.3.3:	Training information and Result for attempt 5	57
Table 4.4.5.1:	Accuracy table for attempt 3	59
Table 4.4.5.2:	Accuracy table for attempt 4	60
Table 4.4.5.3:	Accuracy table for attempt 5	60
Table 4.5.1:	Attempt 6 dataset	65
Table 4.5.3:	Training information and Result for attempt 6	67
Table 4.5.5:	Accuracy table for attempt 6	68
Table 4.6.1:	Attempt 7 dataset	71

Table 4.6.3:	Training information and Result for attempt 7	73
Table 4.6.5:	Accuracy table for attempt 7	74
Table 4.7.1:	Attempt 8 dataset	76
Table 4.7.3:	Training information and Result for attempt 8	78
Table 4.7.5:	Accuracy table for attempt 8	79
Table 4.8.1:	Video feed Test on models	81

LIST OF FIGURES

FIGURE	TITLE	PAGE
Figure 2.2.1:	Artificial Neural Network	6
Figure 2.4.1:	Effect of learning rate	13
Figure 2.4.2:	Gradient descent algorithm with 2 weights	14
Figure 2.4.3:	Gradient descent algorithm as seen from top	14
Figure 2.4.4:	Gradient descent algorithm based on size of batch	16
Figure 2.5.1:	Sigmoid function graph	17
Figure 2.5.2.1:	Tanh function graph	19
Figure 2.5.2.2:	Tanh function graph as compared to sigmoid function graph	19
Figure 2.5.3:	ReLU function graph	20
Figure 2.5.4:	Leaky ReLU function graph	22
Figure 2.6.1:	Before and after applying dropout in neural network	23
Figure 2.8.1:	Convolutional Neural Network (CNN)	25
Figure 2.8.2:	Convolution	26
Figure 2.8.3:	Max pooling	27
Figure 2.8.4:	Average pooling	28
Figure 3.3.0:	Project Flowchart	34
Figure 3.3.2:	CNN architecture	38

Figure 3.3.5:	Video Feed Architecture	40
Figure 4.1.2:	CNN design for attempt 1	44
Figure 4.1.4:	Loss Function graph for attempt 1	45
Figure 4.2.2:	CNN design for attempt 2	49
Figure 4.2.4:	Accuracy graph for attempt 2	50
Figure 4.3.2:	CNN design for attempt 3, 4 and 5	56
Figure 4.3.4.1:	Accuracy graph for attempt 3	58
Figure 4.3.4.2:	Accuracy graph for attempt 4	58
Figure 4.3.4.3:	Accuracy graph for attempt 5	59
Figure 4.4.2:	CNN Design for attempt 6	66
Figure 4.4.4:	Accuracy graph for attempt 6	67
Figure 4.5.2:	CNN Design for attempt 7	72
Figure 4.5.4:	Accuracy graph for attempt 7	73
Figure 4.6.2:	CNN Design for attempt 8	77
Figure 4.6.4:	Accuracy graph for attempt 8	78

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
Appendix 1:	AffectNet Database User Agreement	92
Appendix 2:	CK and CK+ Database User Agreement	93
Appendix 3:	pic_processor.py	94
Appendix 4:	convert_to_hdf.py	96
Appendix 5:	Convolutional Neural Network architecture	99
Appendix 6:	Deep learning architecture	101
Appendix 7:	Model training architecture	103
Appendix 8:	Train data manual sorting and feed into model	105
Appendix 9:	Video feed architecture	106
Appendix 10:	matplotlib labelled image display in real time	108

LIST OF SYMBOLS AND ABBREVIATIONS

DNN	Deep Neural Network
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
Tanh	Hyperbolic tangent
ReLU	Rectified Linear Unit
NN	Neural Network
API	Application Programming Interface
w	weight
b	bias
z	logits
a	Activation function
η	Learning rate
σ	sigmoid

CHAPTER 1

INTRODUCTION

1.1 Background

Machine learning have been booming of late. This is due to the wide application of what machine learning can do. Especially with the existence of better GPU and CPU (which allow better computational power), deep learning can now be apply effectively by processing the large amount of data we currently have. This allow wide array of application such as: Automatic speech recognition, Image recognition, Visual art processing, Natural language processing, Drug discovery and toxicology, Customer relationship management, Recommendation systems, Bioinformatics, Mobile advertising, and Image restoration. This project will mainly concentrate on the ability of deep learning for image recognition.

While Deep learning was introduced to the machine learning community by Rina Dechter in 1986 and Artificial Neural Network was introduced by Igor Aizenberg and his colleagues in 2000, Deep learning truly take off in 2012 whereas a team led by Dahl won the "Merck Molecular Activity Challenge" using multi-task deep neural networks to predict the bio-molecular target of a drug. And just 2 years later, Hochreiter's group won the "Tox21 Data Challenge" of the National Institutes of Health (NIH), the Food and Drug Administration (FDA) and

the National Center for Advancing Translational Sciences (NCATS). The group is able to detect off-target and toxic effects of environmental chemicals in nutrients, household products and drugs by using deep learning architecture. We currently live in era of deep learning revolution where deep learning is receiving more breakthrough constantly. For example, DeepMind and WaveNet by Google and OpenAI founded by Elon Musk.

As for Convolutional Neural Network (CNN), it can be trace back to 1980, which at that time is call neocognitron. Neocognitron does not have multiple layer. One of the important breakthrough for CNN is the LeNet-5, this CNN was pioneered by Yann LeCun in 1998. LeNet-5 have the ability to process high resolution images, however it requires larger and more convolutional layers. This cause LeNet-5 to be relatively unpopular at the time due to the large computational power it required as the network gets bigger. Today, LeNet-5 is commonly used in CNN. In 2005, scientist is able to use GPU instead of CPU to train CNN and it has shown to be very efficient. By 2011, GPU implemented CNN was really well-refined, producing extremely satisfying result. Later one year, AlexNet is introduced. To run with GPU support, AlexNet was written with CUDA, which was created by Nvidia. The SuperVision group designed the AlexNet, the group consist of Alex Krizhevsky, Geoffrey Hinton, and Ilya Sutskever. The SuperVision group used the AlexNet to compete in the ImageNet Large Scale Visual Recognition Challenge 2012, it manage to achieve a top-5 error of 15.3%, which is about 10.8 % ahead of the runner up.

1.2 Problem Statement

Extracting information of emotion from human has proven to be extremely challenging problem in many years. Despite the success of facial recognition and being widely implemented. Emotional recognition via facial expression was proven to be very hard due to the subtle features of human faces. And with the evolution of deep learning in computer vision, emotion recognition has become a widely-tackled research problem. However, none have been truly been implemented.

1.3 Objectives

The objectives of this project is:

- 1) To create a convolutional neural network model that is able to learn facial expression with good probability of success.
- 2) Using deep learning algorithm to recognize emotion of human being via facial expression.
- 3) To have video feed into model and output result.

1.4 Project scope

For this project, in order for the model to work, several factor must be taken into account. This include:

- the size of the datasets
- number of pixel of training data
- the size of model
- the number of convolution layer
- the arrangement of layers in the neural network
- the type of programming language used

CHAPTER 2

LITERATURE REVIEW

2.1 Deep Learning

Deep Learning is one of the field in machine learning that undergo learning using algorithm. It using multiple levels of layers as to create complicated relationships between data, sometime also known as model. Lower-level features and concepts can define higher-level features and concepts can defined, this pyramid of features is called a deep structure or architecture.

Various high-level definition of deep learning have shown to have two common key aspects:

- Models that consist of multiple layers
- Can undergo unsupervised learning or supervised learning of feature representation, which usually becomes more abstract each continuous layer.

Deep learning can discover patterns in large datasets by using the backpropagation algorithm. The backpropagation algorithm is the core algorithm of deep learning, this algorithm is what dictates the changes to the internal parameter of the model. This internal parameter is what allow the deep learning model to find patterns from given datasets.