REAL-TIME HUMAN EXPRESSION RECOGNITION USING DEEP LEARNING ON EMBEDDED SYSTEM

GOH YEN CHANG

UNIVERSITI TEKNOLOGI MALAYSIA

REAL-TIME HUMAN EXPRESSION RECOGNITION USING DEEP LEARNING ON EMBEDDED SYSTEM

GOH YEN CHANG

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Engineering (Computer and Microelectronic Systems)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JUNE 2018

This thesis is dedicated to my family and friends who encouraged and supported me throughout this research.

ACKNOWLEDGEMENT

Throughout the process in making this report, I received countless help and support from many people, lecturers and friends. Without their contribution, this report will never be completed. The very first person I wish to express my sincere gratitude is my supervisor, Assoc. Prof. Muhammad Mun'im Ahmad Zabidi. He gave me a lot of guidance, support and critics to ensure my project can be completed in timely manner. Without him, completing this report would have been difficult.

I am also indebted to Universiti Teknologi Malaysia (UTM) for all the teaching and knowledge throughout the 2 years courses of part time Master study in this university. I would also like to express my appreciation to Intel Microelectronic Sdn. Bhd. for sponsoring my part time Master study in UTM. Among my colleagues, I would like to thanks Teh Eng Keong and Ng Tze Han, as well as my manage Law Liang Ming, for constantly cover for my work in company so that I have sufficient time to perform my research and write this report.

Last but not least, I would like to express my sincere appreciation to my family members, especially my parents who gave me limitless support and encouragement.

ABSTRACT

Technology ease human life in every aspects. Some machine save human's effort, some machine save time and increase efficiency in work. Machine is designed to complete specific task or multiple tasks without any intelligence needed. The next level of machine is machine which has intelligent and capable to think like human being while doing jobs, moreover, learn by themselves. Recently, Machine Learning are becoming more and more popular in 21^{st}century. Machine learning can explores study and algorithms construction for making prediction. Data analytic by machine learning is a trend that used by Google, Facebook, Baidu and others big company nowadays. One of data analysis in machine learning which is Human Facial Expression Recognition is one of the hot topics now. Many researchers are proposed their techniques used in emotion recognition like PCA, LBP and etc. Goal in this project, is to analyze Inception v-3, the best performing high resolution image classifier based on Convolutional Neural Network, and also implement it in Raspberry Pi to see how it performs on detecting Facial Expressions.

ABSTRAK

Teknologi memudahkan kehidupan manusia dalam setiap aspek. Beberapa mesin menjimatkan usaha manusia, beberapa mesin menjimatkan masa dan meningkatkan kecekapan dalam kerja. Mesin direka untuk menyelesaikan tugas tertentu atau pelbagai tugas tanpa apa-apa kecerdasan yang diperlukan. Tahap seterusnya mesin adalah mesin yang pintar dan mampu berfikir seperti manusia ketika melakukan pekerjaan, apalagi belajar sendiri. Baru-baru ini, Pembelajaran Mesin menjadi semakin popular di abad ke-21. Pembelajaran mesin boleh meneroka pembinaan dan pembinaan algoritma untuk membuat ramalan. Analisis data oleh pembelajaran mesin adalah trend yang digunakan oleh Google, Facebook, Baidu dan syarikat besar lain sekarang. Salah satu analisis data dalam pembelajaran mesin yang merupakan Pengiktirafan Ungkapan Wajah Manusia adalah salah satu topik hangat sekarang. Ramai penyelidik mencadangkan teknik mereka yang digunakan dalam pengiktirafan emosi seperti PCA, LBP dan sebagainya. Matlamat dalam projek ini, adalah untuk menganalisis Inception v-3, pengeluar imej resolusi tinggi yang terbaik berdasarkan Rangkaian Neural Convolutional, dan juga melaksanakannya dalam Raspberry Pi untuk melihat bagaimana ia berfungsi dalam mengesan Ekspresi Wajah.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAM	-	Active Appearance Model
AU	-	Action Units
CSI	-	Camera Serial Interface
FACS	-	Facial Action Coding System
FER	-	Face Expression Recognition
GPU	-	Graphic Processing Unit
HD	-	High Definition
HOG	-	Histogram of Gradient Orientation
LBP	-	Local Binary Pattern
MP	-	Megapixel
PC	-	Personal Computer
PCA	-	Principle Component Analysis
ROI	-	Region of Interest
RP	-	Raspberry Pi
SDRAM	-	Synchronous Dynamic Random Access Memory

CHAPTER 1

INTRODUCTION

1.1 Problem Background

Emotion is a mental state which involves a lot of behaviors, actions, thoughts and feelings. In 1969, Charles Darwin wrote "The Expression of the Emotions in human and Animals" after recognizing the universality among emotions in different groups of people despite the cultural differences [7]. Ekman and Friesen classified six emotional expressions to be universal: happiness, sadness, disgust, surprise and fear [8]. Facial expressions can be considered as the most natural form of displaying human emotions and as a non-verbal communication technique . Implementation of efficient automatic facial expression recognition techniques may yield lot of improvements in the area of Human Computer Interaction.

The universality of facial expressions on the presence of emotions and micro-expressions can help people face face-to-face interactions in order to improve their skills and read other people's emotions in many different occupations. Reading emotional expressions, especially microcosmic expressions, can help develop relationships, beliefs and agreements; they can be used to assess credibility, assess truth and detect fraud; better information on emotional status is provided for better cooperation, advice or sales. Health professionals can build better relationships with patients, interact with patients with sympathy and sympathy, and make the right diagnosis by getting complete information. Teachers can read the emotions of the students to get a hint of progress on their lessons, so that they can make appropriate adjustments and communicate more effectively. School administrators who read teacher emotions can reduce fires, maintain and improve teacher effectiveness. Businessmen and consultants who can read the emotions of others can promote mutually beneficial cooperation. Researchers can improve qualitative data, and they give a true sense of what they feel when they can read the user's emotional assessment product, even though they have clues. Parents, spouses, friends, and everyone who is interested in building a strong and constructive relationship can benefit from improving their reading emotions.

Emotional expression may or may not be accompanied by self-awareness. Over the past 200 years, researchers have presented different, often competing, models to explain emotional and emotional expression and have been traced back to Charles Darwin. The different types of expressions namely joy, sadness, surprise, anger, disgust and fear are given below:

- Happy – Happy is an emotion caused by welfare, glory, luck, excitement or prospects.
- Sadness – Sadness is an emotional pain associated with or despair, helplessness, disappointment, sadness, related to defects, loss, and sadness. A person experiencing sadness may become quiet or drowsy and withdraw from others.
- Surprise – Surprise is defined as to cause of someone to feel in amazing feelings.
- Anger – Anger can occur when a person feels their personal boundaries are being or going to be violated.
- Disgust – Disgust is an aversion. Humans will feel disgusted from any sound, smell, taste, or difficulty.

1.2 Motivation

What makes embedded devices great?

- Small in size
- Low power consumption
- Large operating range
- Low per-unit cost
- Standalone system

Applications of facial expression recognition:

- Medicine sector
  - Rehabilitation
  - Companion
  - Counseling
  - Autism therapy
- E-learning
  - Tutoring system
- Monitoring
  - ATM machine
- Entertainment
  - Gaming
  - Music player
- Marketing
  - Impact of ads

1.3 Problem Statement

To lead the way in Human-Android communication. The purpose of this research is to improve the way of android understand human by only human expression. Many researchers already proposed their method in classify human expression by using machine learning. However, there is still has a lot of work needed to be done in order to achieved high performance and accuracy in human expression recognition. One of our non-verbal communication methods is to understand emotional as well as psychological states of people's facial expressions such as happiness, sadness, fear, disgust, surprise and anger [7, 9]. The exciting and challenging field of automatic facial recognition recognition (FER) [10] has become a branch of computer vision. Using the embedded device in FER is paving the way for the future. This is because embedded device is small in size and low power consumption, so it is widely applicable. The use of FER include mental identification [11], security [12], automated counseling systems, face synthesis expressions, lie detection, music mood detection [13], automated tutoring systems [14], operator fatigue detector [15], and others. Capturing spontaneous expressions on images and video is one of the biggest challenges.

1.4 Research Aim and Objectives

The aim of this research is to produce an embedded device that can preform FER in real-time. The objectives of research are:

- To investigate a method to run Real-time Face Expression Recognition (FER) based on deep learning in a low cost embedded system.
- To develop a method to train the deep neural network approach.
- To evaluate and compare the performance and accuracy between PC and embedded system.

1.5 Scope of Work

- Raspberry Pi 3 Model B is been using as the embedded device which has properties of low power consumption, small size, a wide range of overall work, and low unit cost.
- The Raspberry Pi is used only for the inference part of deep learning, the training is done on the PC
- Tensor-Flow will be used as tool in this project and the programming language is Python.
- One six basic facial expressions are recognized: happiness, sadness, surprise, fear, disgust, and anger.

1.6 Organization

The organization of this thesis consists of seven chapter.

Chapter one describes the project background, problem statement, objectives, scope of work and the outline of this thesis.

Chapter two describes the literature review on the existing human expression detection methods including face detection and pre-processing, and feature extraction and classification.

Chapter three discusses the research methodology of this project, as well as the methods are used for achieving objectives. The tools and dataset are used in this project also be discussed as well.

Chapter four illustrates the proposed design. The architecture of the proposed design will be discussed. The training details will be covered as well in this chapter.

Chapter five shows the result of this project. Discussion on the result will be carried on in this chapter.

Chapter six is the conclusion of this project. Some potential future work will be discussed on this chapter.

REFERENCES

1. Kanade, T., Cohn, J. F. and Tian, Y. Comprehensive database for facial expression analysis. *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*. IEEE. 2000. 46–53.
2. Dubey, M. and Singh, L. Automatic Emotion Recognition Using Facial Expression: A Review. *International Research Journal of Engineering and Technology (IRJET)*, 2016.
3. Hjelmås, E. and Low, B. K. Face detection: A survey. *Computer vision and image understanding*, 2001. 83(3): 236–274.
4. Ashraf, A. B., Lucey, S., Cohn, J. F., Chen, T., Ambadar, Z., Prkachin, K. M. and Solomon, P. E. The painful face–pain expression recognition using active appearance models. *Image and vision computing*, 2009. 27(12): 1788–1796.
5. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. *et al.* Going deeper with convolutions. Cvpr. 2015.
6. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016. 2818–2826.
7. Darwin, C. and Prodger, P. *The expression of the emotions in man and animals*. Oxford University Press, USA. 1998.
8. Ekman, P. and Friesen, W. V. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 1971. 17(2): 124.
9. Mehrabian, A. Communication without words. *Communication theory*, 2008: 193–200.
10. Fasel, B. and Luettin, J. Automatic facial expression analysis: a survey. *Pattern recognition*, 2003. 36(1): 259–275.
11. Mandal, M. K., Pandey, R. and Prasad, A. B. Facial expressions of emotions and schizophrenia: A review. *Schizophrenia bulletin*, 1998. 24(3): 399.
12. Butalia, M. A., Ingle, M. and Kulkarni, P. Facial expression recognition for

security. *International Journal of Modern Engineering Research (IJMER)*, 2012. 2(4): 1449–1453.

13. Dureha, A. An accurate algorithm for generating a music playlist based on facial expressions. *International Journal of Computer Applications*, 2014. 100(9): 33–39.

14. Wu, Y., Liu, W. and Wang, J. Application of emotional recognition in intelligent tutoring system. *Knowledge Discovery and Data Mining, 2008. WKDD 2008. First International Workshop on*. IEEE. 2008. 449–452.

15. Zhang, Z. and Zhang, J. A new real-time eye tracking for driver fatigue detection. *ITS Telecommunications Proceedings, 2006 6th International Conference on*. IEEE. 2006. 8–11.

16. Lyons, M. J., Budynek, J. and Akamatsu, S. Automatic classification of single facial images. *IEEE Transactions on pattern analysis and machine intelligence*, 1999. 21(12): 1357–1362.

17. Ojala, T., Pietikäinen, M. and Harwood, D. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 1996. 29(1): 51–59.

18. Turk, M. A. and Pentland, A. P. Face recognition using eigenfaces. *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*. IEEE. 1991. 586–591.

19. Bartlett, M. S., Movellan, J. R. and Sejnowski, T. J. Face recognition by independent component analysis. *IEEE Transactions on neural networks*, 2002. 13(6): 1450–1464.

20. Belhumeur, P. N., Hespanha, J. P. and Kriegman, D. J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, 1997. 19(7): 711–720.

21. Tong, Y., Chen, R. and Cheng, Y. Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle. *Optik-International Journal for Light and Electron Optics*, 2014. 125(16): 4186–4189.

22. Jabid, T., Kabir, M. H. and Chae, O. Facial expression recognition using local directional pattern (LDP). *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE. 2010. 1605–1608.

23. Hsu, C.-W., Chang, C.-C., Lin, C.-J. *et al.* A practical guide to support vector

classification. 2003.

24. Altman, N. S. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 1992. 46(3): 175–185.

25. Happy, S., George, A. and Routray, A. A real time facial expression classification system using local binary patterns. *Intelligent Human Computer Interaction (IHCI), 2012 4th International Conference on*. IEEE. 2012. 1–5.

26. Meher, S. S. and Maben, P. Face recognition and facial expression identification using PCA. *Advance Computing Conference (IACC), 2014 IEEE International*. IEEE. 2014. 1093–1098.

27. Friesen, E. and Ekman, P. Facial action coding system: a technique for the measurement of facial movement. *Palo Alto*, 1978.

28. Ekman, P. Methods for measuring facial action. *Handbook of methods in nonverbal behavior research*, 1982: 45–90.

29. He, K., Zhang, X., Ren, S. and Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision*. 2015. 1026–1034.

30. Schroff, F., Kalenichenko, D. and Philbin, J. Facenet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015. 815–823.

31. Chen, W., Wilson, J., Tyree, S., Weinberger, K. and Chen, Y. Compressing neural networks with the hashing trick. *International Conference on Machine Learning*. 2015. 2285–2294.

32. Lavin, A. and Gray, S. Fast algorithms for convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016. 4013–4021.

33. Ioffe, S. and Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv: 150203167. *arXiv preprint arXiv:1502.03167*, 2015.

34. Goodfellow, I., Bengio, Y. and Courville, A. Deep learning (adaptive computation and machine learning series). *Adaptive Computation and Machine Learning series*, 2016: 800.

35. Torrey, L. and Shavlik, J. Transfer learning. In: *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI Global. 242–264. 2010.

36. Ko, K.-E. and Sim, K.-B. Development of a Facial Emotion Recognition

Method based on combining AAM with DBN. *Cyberworlds (CW), 2010 International Conference on*. IEEE. 2010. 87–91.

37. Aswathy, M. R âA Literature review on Facial Expression Recognition Techniques. *IOSR Journal of Computer Engineering (IOSR-JCE)*, 2013. 11(1): 61–6.

38. Banu, S. M., Danciu, G. M., Boboc, R. G., Moga, H. and Bălany, C. A Novel approach for face expressions recognition. *Intelligent Systems and Informatics (SISY), 2012 IEEE 10th Jubilee International Symposium on*. IEEE. 2012. 537–541.

39. Zhen, W. and Zilu, Y. Facial expression recognition based on adaptive local binary pattern and sparse representation. *Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on*. IEEE. 2012, vol. 2. 440–444.

40. Kumar, S. and Gupta, A. Facial expression recognition: A review. *Proceedings of the National Conference on Cloud Computing and Big Data, Shanghai, China*. 2015. 4–6.

41. Yi, J., Mao, X., Chen, L., Xue, Y. and Compare, A. Facial expression recognition considering individual differences in facial structure and texture. *IET Computer Vision*, 2014. 8(5): 429–440.

42. Li, J., Zhao, C., Wang, H. and Ying, Z. Facial expression recognition based on completed local binary pattern and SRC. *Natural Computation (ICNC), 2013 Ninth International Conference on*. IEEE. 2013. 333–337.

43. Yi, J., Mao, X., Xue, Y. and Compare, A. Facial expression recognition based on t-SNE and adaboostM2. *Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing*. IEEE. 2013. 1744–1749.

44. Lee, J., Uddin, M. Z. and Kim, T.-S. Spatiotemporal human facial expression recognition using fisher independent component analysis and Hidden Markov Model. *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE. 2008. 2546–2549.

45. Zilu, Y. and Guoyi, Z. Facial expression recognition based on NMF and SVM. *Information Technology and Applications, 2009. IFITA'09. International Forum on*. IEEE. 2009, vol. 3. 612–615.

46. Dhavalikar, A. S. and Kulkarni, R. Face detection and facial expression recognition system. *Electronics and Communication Systems (ICECS), 2014*

International Conference on. IEEE. 2014. 1–7.