

UNIVERSITI TEKNOLOGI MARA

**MALAY WORDS AND DIALECT
IDENTIFICATION USING LONG
SHORT-TERM MEMORY AND
CONVOLUTIONAL NEURAL
NETWORKS ON TRAINED MEL
FREQUENCY CEPSTRAL
COEFFICIENT**

MOHD AZMAN HANIF BIN SULAIMAN

Thesis submitted in fulfillment
of the requirements for the degree of
Master of Science
(Electrical Engineering)

College of Engineering

November 2022

ABSTRACT

As Malaysia moves towards to the Industrial Revolution (IR 4.0), and as machines become more intelligent and autonomous, man and machine interaction are becoming inevitable. In general, the machine robustness towards dialect identification will be the main one of the many practical methods for interaction is using spoken language. However, there are many limitations of this type of interaction, particularly for native speakers other than English among them is dialect criteria for the system development. The complexity of dialects requires a new paradigm/generation of Artificial Intelligence (AI) – based classifiers methods capable of adapting to the linguistic of the language. These proposed techniques are recent and relatively unexplored in the field of dialect identification. This research explores two types of methods for dialect identification, namely Convolution Neural Network (CNN) for Malay dialect identification and MFCC feature extraction technique will be used to extract the features. Next, these features will be transferred to the CNN to be trained. For Long short-term Memory (LSTM), the inputs are fed directly from the recorded dataset for training. This research is to design and implement the CNN and LSTM network for Malay language dialect classification. In support of this, several objectives need to be achieved is to perform features extraction using MFCC on Malaysian Dialect, then to classify the Malaysian Dialect using CNN and LSTM neural networks and compare the performance of CNN and LSTM neural networks on Malay dialect identification.

As there are currently no publicly available Malay language dialect datasets, the data was collected from volunteer subject fluent with the dialect. To ensure optimal recording conditions, recording was performed in a controlled environment inside a silent room without any disturbance. Each subject (six male and six female native dialect speakers) was required to utter the each of same word ten times. Each recording was saved into the raw wav format, resulting in 2,400 samples divided into standard, eastern and northern dialects. Additionally, the age of the subjects ranged from 20 to 35 years old in order to obtain clear and fluent pronunciation of their native dialect. The results suggest that both methods were able to accurately classify the different words spoken in different Malay dialects with above 85% accuracy, with LSTM performing the best with 98.2% accuracy.

ACKNOWLEDGEMENT

In the name of Allah, most beneficent, most merciful, all the praised and thanks to Allah most almighty, lord of the universe and peace be upon to His messenger Prophet Muhammad PBUH, the last of the prophets. Foremost, I am indebted and extremely grateful to my supervisor, Ir. Ts. PM. Dr. Ahmad Ihsan Bin Mohd Yassin, Ts. Dr. Megat Syahirul Amin B. Megat Ali, Dr. Azlee B. Zabidi and also Dr. Farzad Eskandari who has gave a lot of his time supervising the research upon this completion. Not forget their understanding, encouraging and personal guidance has provided a good basis for the present thesis. I could not have imagined having a better supervisor and mentor for my research.

I would also like to extend my appreciation to the speakers from Selangor, Kelantan and Kedah for their willingness to become my subject in gathering speech data for the purpose of this research.

I thank to my wife, Nurhakimah Abd Aziz and children Naurah, Adam and Maryam who sacrifice their right to my time and attention throughout the difficult period of completing my study.

Finally, this thesis is dedicated to my mother and the loving memory of my very dear late father for the vision and determination to educate me. This piece of victory is dedicated to both of you. Alhamdulillah.

Last but not least, I would like to express my sincerest appreciation to the people who have directly and indirectly contributed to the successful completion of this thesis.

TABLE OF CONTENTS

	Page
CONFIRMATION BY PANEL OF EXAMINERS	ii
AUTHOR'S DECLARATION	iii
ABSTRACT	iv
ACKNOWLEDGEMENT	v
TABLE OF CONTENTS	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF SYMBOLS	x
LIST OF ABBREVIATIONS	xi
CHAPTER ONE INTRODUCTION	1
1.1 Research Background	1
1.2 Problem Statement	3
1.3 Objective	4
1.4 Contribution of the Thesis	5
1.5 Research Scope	5
1.6 Organization of the Thesis	6
CHAPTER TWO LITERATURE REVIEW	7
2.1 The Malay Language & Dialects in Malaysia	7
2.2 Mel Frequency Cepstral Coefficients (MFCC)	8
2.2.1 MFCC Feature Extraction	9
2.3 Dynamic Time Warping	9
2.4 Deep Learning	11
2.4.1 CNN	13
2.4.2 LSTM	15
2.5 Relevant Research	18
2.6 Previous Research on Dialects	22
2.7 Chapter Summary	23

CHAPTER THREE RESEARCH METHODOLOGY	24
3.1 Introduction	24
3.2 Data Collection	25
3.3 Hardware Specification	28
3.4 Features Extraction using MFCC	28
3.5 Classification	31
3.5.1 CNN	32
3.5.2 LSTM	34
3.6 Performance comparison	37
CHAPTER FOUR RESULTS AND DISCUSSION	40
4.1 Introduction	40
4.2 Data collection	40
4.3 Result for Dynamic Time Warping	46
4.4 MFCC Feature Extraction Results	53
4.4.1 CNN	53
4.5 Classification results	57
4.5.1 CNN network	57
4.5.2 LSTM network	66
4.6 Performance analysis	76
4.6.1 Network performance analysis	76
4.6.2 Comparison with Similar Recent Works	82
CHAPTER FIVE CONCLUSION & FUTURE WORK	84
5.1 Conclusions	84
5.2 Future Work	85
REFERENCES	86
APPENDICES	95
AUTHOR'S PROFILE	104