Ancient Kawi Character Recognition Using Deep Learning

Abdulrazak Yahya Saleh[1], Sadia Afrin[2]

[1] Faculty of Cognitive Science & Human Development, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

[2] Faculty of Cognitive Science & Human Development, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

Abstract

In document analysis and recognition, recognizing characters is still an important problem. Document processing is gaining popularity in day by day in the field of pattern recognition. This study has been conducted to recognize Kawi character which is an ancient script of Malaysia by deep learning. As recently, deep learning has achieved more attention because of the ability to learn from raw data in character recognition. The Kawi character database has been developed for this research. Convolutional neural network models have recently proved an impressive classification performance towards recognizing characters. So, the Kawi character recognition system has been developed using CNN. The system has the capability to extract features and accuracy improvement. This research achieved 89% accuracy using CNN model.

Keywords: *Artificial Neural Network, Character Recognition, Convolutional Neural Network, Deep Learning, Kawi Script.*

1. Introduction

One of the most research applications in pattern recognition is character recognition [1]. Machine simulation of human reading considered as character recognition [2]. To identify characters by computer programs without the input given by human is the main aim of character recognition [3].

As mentioned earlier, character recognition is a popular and wide field of research and many ancient scripts have been recognized such as English script [4], Arabic script [5], Farsi script [6], Brahmi script [7], Bangla script [8], Malayalam script [9], Telegu script [10] etc. As individual has different handwriting styles, after sufficient research still open research problems remain [11]. However, also many old scripts exist which are not yet recognized. Kawi script is one of the scripts among them which had been used in Malaysia for long period. Due to this, many important documents were written in this script by Malaysians. Recognizing Kawi script will help to discover and understand many unrevealed facts about Malaysian history.

Brahmi script was introduced and spread into northern and southern during first millennium. After

Brahmi, Grantha script started to spread and became popular. Malayalam script, Tamil script and Pallava script derived from Grantha script. However, the parent of Kawi script was Pallava [12]. Fig. 1 shows the origin of the Kawi script whereas Fig. 2 shows the Vowels and consonants of Kawi script.
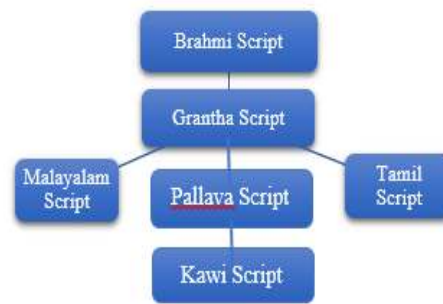

Fig. 1: Origin of Kawi script

Fig. 2: The Vowels and consonants of Kawi script

This is the only Austronesian language whose earlier forms are attested in written record in the shape of an extensive corpus of texts. This is considered as the great