

## Abstract

Arabic Alphabet is one of the most used writing system in the source documents and reference manuscripts around the world besides the Latin Alphabet. In the pattern recognition field, Arabic Alphabet (Abjad) has been received numerous concern, but some approach that has been done until today only covers Arabic Alphabet pattern recognition that not using *harakat*. *Harakat* is a diacritic in Arabic Alphabet that tells how to pronounce certain letters to minimize misinterpretation. The example of most cited and definitely the most influential documents that uses Arabic Alphabet with *harakat* is The Holy Qur'an, which was one of the first major works of Arabic literature.

In this study, pattern recognition has been performed on *Hijaiyah* Letters (isolated/non-cursive Arabic letter) with *harakat* by using Modified Direction Feature Extraction (M-DFE) for feature extraction method, and Hidden Markov Models (HMM) for classification method. Morphologically, each of *Hijaiyah* Letters have a fixed stroke direction. Meanwhile M-DFE itself extract directional and transitional feature from whole structure of a letter. Thus by using M-DFE, a good feature could be extracted from each *Hijaiyah* Letters. HMM, in other hand has been showing good result when implemented on several studies to recognize Arabic Alphabet without *harakat*, whether it's using online image acquisition, offline image acquisition, or using *multilevel classifier* in the classification process.

More over, handling the *harakat* in *Hijaiyah* Letters is becoming another challenge in this study. *Harakat* expanding the class enormously which is affecting the system accuracy and processing time as well. To solve this problem identifying the *harakat* and main letter from the image to be processed separately is done. By performing segmentation technique combined with rule-based clustering, the number of class can be reduced to minimum so that the processing time could be shorter, while still maintaining to obtain good accuracy. After testing several scenarios, the best result was obtained with 4 *maximum transitions* and 70 *states* HMM parameter on 69,4% accuracy rate.

**Keywords:** *Hijaiyah Letters, Harakat, Pattern Recognition, Modified Direction Feature Extraction, Hidden Markov Models.*