ABSTRACT

Indonesia is a country that has a cultural diversity. One of them is the diversity of regional languages. Java language is one of the regional languages which is used in Indonesia. To read the Java Characters are not easy. Because we have to memorize the shapes of java characters and its pronunciation. This because the shape of Java characters are very complex.

In this final project, the writer has made a tool that can help people to read Java characters easier. This tool identifies the Java characters and then translated them into Latin characters using Optical Character Recognition (OCR). OCR is a computer system that can recognize handwritten characters or typed manuscript which is produced by scanner , into an editable text using a computer application. A series of Java characters which will be identified in this final project comes from the Javanese character texts from the print-out or block letters. Those characters will be directly translated by the system after capturing it by webcam. The result of this process is in the editable text form. The process does a few stages. In the preprocessing stage, the image which is produced by the webcam, is changed its color from RGB to grayscale, and then changed it to black and white form. Black and white image is segmented into one syllable, in order to be detected. At this feature extraction stage, the traits of image which is produced in preprocessing stage is taken of each character. And at this character recognition stage using Self Organizing Map neural network, the determination of vector compiler character line is done by connecting the neurons dots in a region character.

The test which is done to the system involves taking a picture of the fonts that have been trained. And from tests above, the best accuracy is 98% for aksara Murda and Swara, 96% for Basic Aksara, 55% for Aksara with Sandhangan, 53% for Aksara with punctuations.

Keywords: Java character, OCR, webcam, feature extraction, SOM artificial neural network.