

INTISARI

REKOGNISI KARAKTER AKSARA SUNDA TULISAN TANGAN MENGGUNAKAN EKSTRAKSI CIRI HOG DAN METODE KLASIFIKASI SVM

M Taufiqurahman

14/366919/PA/16276

Aksara Sunda merupakan sebuah jenis tulisan yang digunakan oleh suku Sunda. Pada era modern ini, bukan hal yang asing untuk menerapkan *Optical Character Recognition* (OCR) dalam rangka melestarikan suatu jenis tulisan, baik itu cetak maupun tulisan tangan. Pengenalan tulisan tangan masih cukup sulit dilakukan karena setiap orang memiliki gaya penulisannya masing-masing.

Tahap prapengolahan yang dilakukan yaitu reduksi derau, pengubahan citra grayscale menjadi citra biner, segmentasi, dan normalisasi ukuran. Metode ekstraksi ciri yang digunakan yaitu metode *Histogram of Oriented Gradient* (HOG). Metode klasifikasi ciri yang digunakan yaitu metode *Support Vector Machine* (SVM). Tahap pascapengolahan yang dilakukan yaitu konversi label, pembentukan suku kata, dan pembentukan kata. Sistem dilatih dengan berbagai variasi keadaan ukuran *cell* dan jenis kernel untuk mengetahui kondisi optimalnya kemudian diuji dengan citra huruf dan citra kata.

Hasil pengujian sistem dengan berbagai variasi keadaan ukuran *cell* dan jenis kernel menunjukkan akurasi terbaik yang didapat yaitu 91,98% dengan pengaturan sistem berupa ukuran *cell* bernilai 12 dan penggunaan kernel linear. Hasil pengujian sistem dengan citra karakter dan citra kata menunjukkan akurasi sebesar 90,83% dan 49,5%. Sistem OCR yang dibangun masih memiliki permasalahan rekognisi untuk beberapa karakter yang memiliki komponen karakter yang mirip dengan karakter lainnya.

Kata kunci— OCR, Aksara Sunda, Tulisan Tangan , HOG, SVM

ABSTRACT

HANDWRITTEN SUNDANESE SCRIPT CHARACTER RECOGNITION USING HOG FEATURE AND SVM CLASSIFIER

M Taufiqurahman

14/366919/PA/16276

Sundanese Script is one of script that has been used by Sunda tribe. In this modern era, it is a common thing to using Optical Character Recognition (OCR) for preserving a script, either it is a printed scripts or handwritten scripts. Handwritten recognition is difficult to implement because every people have their own writing style.

Preprocessing steps are noise reduction, converting grayscale image to binary image, segmentation, and size normalization. Feature extraction is Histogram of Oriented Gradient (HOG). Classification method is Support Vector Machine (SVM). Postprocessing steps are label conversion, syllable construction, and word construction. The system is trained by varying its cell size and kernel type to identified its optimal condition then system is tested by using character images and word images.

The result of testing system by varying its cell size and kernel type shows that the best accuracy it can get is 91,98% by using cell size is 12 and linear kernel. The result of testing system by using character images and word images shows that the accuracy it can get is 90,83% and 49,5%. OCR system that being build still has a problem to recognize some characters that have character components that seems like the other characters has.

Keywords— OCR, Sundanese Script, Handwritten , HOG, SVM