

Abstrak

Optical Character Recognition merupakan suatu teknik untuk merubah data tulisan tangan, mesin ketik maupun mesin cetak (tulisan dalam bentuk *non-digital*) menjadi bentuk tulisan digital yang dapat dirubah menggunakan komputer. Pada OCR terdapat beberapa tahap yang harus dilakukan antara lain *Image Acquisition*, *Image Preprocessing*, *Feature Extraction* dan *Classification*.

Dalam tugas akhir ini akan dikembangkan suatu sistem yang dapat membaca tulisan tangan huruf Jepang (KANJI) dengan menerapkan metode *Independent Component Analysis* (ICA) sebagai *image transformer* pengganti tahap *feature extraction* dan *Learning Vector Quantization* (LVQ) sebagai classifier. Hasil akurasi terbaik yang dihasilkan oleh sistem ini ialah 69.5652% untuk data latih hiragana, 61.087% untuk data latih katakana, 43.913% untuk data uji hiragana dan 33.4783% untuk data uji katakana. Hasil akurasi ini didapatkan dengan menggunakan 5 IC (*independent component*) dari 8 IC yang dibangkitkan untuk hiragana dan 2 IC dari 5 IC yang dibangkitkan untuk katakana. Kemudian pada tahap pelatihan besar *learning rate* terbaik yang digunakan ialah 0.0015 baik pada hiragana maupun katakana dan jumlah hidden neuron yang terbaik untuk hiragana ialah 184 hidden neuron sedangkan untuk katakana ialah 230 hidden neuron. Hasil akurasi yang kurang baik ini disebabkan tingkat kemiripan yang cukup tinggi antar huruf sehingga sangat sulit bagi *classifier* untuk mengenali huruf-huruf Jepang tersebut.

Kata kunci : optical character recognition, hiragana, katakana, independent component, learning vector quantization.