

Abstraksi

Optical Character Recognition (OCR) adalah sebuah sistem komputer yang digunakan secara otomatis mengenali serangkaian karakter yang berasal dari mesin ketik, mesin cetak ataupun tulisan tangan. Dengan kata lain OCR adalah proses pengalihan dokumen teks menjadi file komputer tanpa harus pengeditan ulang, setiap karakter baik huruf, kata, kalimat dapat dikenali secara tepat dan dibaca oleh perangkat lunak yang lain, tanpa harus pengetikan ulang dan editing.

Pada tugas akhir ini dikembangkan suatu aplikasi untuk mengidentifikasi karakter pada suatu file gambar (bmp atau jpg) yang berisi karakter yang berasal dari pemindaian *hardcopy* atau dari sumber lainnya. Proses ekstraksi ciri menggunakan pendekatan vektor dan region. Pada proses tersebut akan ditentukan vektor penyusun garis karakter pada tiap area pengamatan (region), dimana tiap karakter dibagi menjadi 9 region yang sama besar dan simetris.

Untuk mengevaluasi performansi dari OCR dengan menggunakan metode tersebut, dilakukan pengujian terhadap beberapa sampel masukan baik yang berasal dari dokumen *hardcopy* maupun yang berasal dari sumber lainnya. Hasil analisis menunjukkan bahwa sistem OCR ini mempunyai tingkat akurasi sebesar 91,88% untuk font yang sudah dilatihkan, dan 52,26% untuk font yang belum dilatihkan.

Kata kunci : Pengenal huruf otomatis , ekstraksi ciri , vektor , region