

Abstrak

Optical Character Recognition (OCR) adalah konversi mekanik atau elektronik dari gambar scan tulisan tangan, ketik, atau cetak ke dalam *encoded text* untuk komputer. Kajian tentang OCR akhir-akhir ini memberikan metode dengan akurasi yang baik dan kecepatan komputasi yang semakin ringan. Komputasi yang semakin ringan memungkinkan mesin dengan spesifikasi rendah seperti smart phone sanggup mengerjakan OCR.

Pada Tugas Akhir ini diterapkan metode OCR, dengan menggunakan *Contour Analysis* dengan representasi contour berupa bilangan kompleks dan metode *feature extraction* *Intercorelation* dan *Autocorrelation function* dan dengan *feature preselection* menggunakan algoritma *longest common subsequence* (LCS). Untuk penghalusan contour digunakan algoritma Rames Douglas Peucker (RDP).

Hasil pengujian menunjukkan penghalusan contour dengan menggunakan RDP dengan konfigurasi epsilon optimal sebesar 2,0 dan titik contour sebanyak 32 akurasi yang dihasilkan lebih besar 4% dari pada tidak menggunakan RDP. Penggunaan *feature preselection* menggunakan LCS lebih efektif menyeleksi daripada *Euclidean distance*. LCS dengan konfigurasi toleransi optimal sebesar 0,1 menghasilkan 68 *contour* hasil preseleksi dari 676 *contour*. Euclidean distance dengan konfigurasi *threshold* jarak 0,3 menghasilkan 158 *contour* hasil preseleksi dari 676 *contour*.

Kata kunci: OCR, *feature preselection*, algoritma Rames-Douglas-Peucker, *longest common subsequence*, *contour analisis*.

Commented [MDS1]: Do this one!