

## Abstrak

Optical Character Recognition (OCR) adalah sistem komputer yang mampu membaca suatu karakter, baik yang berasal dari sebuah mesin pencetak (scanner atau komputer) dan juga yang berasal dari tulisan tangan. OCR membantu meringankan pekerjaan manusia karena proses pengetikan ulang data yang telah ditulis pada selembar kertas menjadi lebih efektif dilihat dari segi waktu dan tenaga.

Terdapat beberapa tahapan dalam OCR yaitu preprocessing didalamnya terdapat proses segmentasi dan normalisasi, ekstraksi ciri, dan klasifikasi atau pengenalan karakter. Segmentasi yaitu membagi citra kedalam segmen-segmen, proses ini akan mempengaruhi akurasi pada saat pengenalan karakter, sedangkan pada tahap normalisasi terdapat teknik dilasi/penebalan dan centering pada tiap karakter tunggal. Ekstraksi ciri yaitu proses mengekstrak karakter serta mengambil ciri khusus dari karakter tunggal yang akan dikenali, metode yang digunakan pada tahap ini adalah zoning. Zoning merupakan metode ekstraksi ciri dengan membagi karakter kedalam  $N \times M$  wilayah. Tahap pengenalan karakter adalah tahap pengklasifikasikan karakter. Metode yang digunakan adalah KNN(K-Nearest Neighbour). Metode ini mencocokkan karakter yang akan dikenali dengan karakter yang sudah tersimpan didalam database.

Berdasarkan implementasi yang telah dilakukan, dengan menerapkan metode zoning dan klasifikasi KNN pada penginputan nilai matakuliah mahasiswa di dalam OCR, hasil pengujian dapat dilihat dari tingkat akurasi per-angka dan akurasi per-nilai. Dari pengujian yang telah dilakukan, ukuran dan bentuk karakter menjadi faktor yang mempengaruhi tingkat akurasi pada saat pengenalan karakter.

**Kata Kunci:** OCR , *preprocessing*, segmentasi, normalisasi, *zoning*, KNN.