

ABSTRAK

Penelitian ini bertujuan untuk mengukur performa sistem Optical Character Recognition (OCR) dalam mengekstraksi data dari Kartu Tanda Penduduk (KTP). Sistem yang dikembangkan dievaluasi berdasarkan tiga aspek utama, yaitu tingkat akurasi hasil ekstraksi, waktu pemrosesan, dan efisiensi dibandingkan metode input manual. Hasil pengujian menunjukkan bahwa sistem OCR mampu mencapai tingkat akurasi rata-rata sebesar 90%. Waktu rata-rata yang dibutuhkan untuk memproses satu gambar KTP adalah 1,68 detik, dengan waktu tercepat 1,26 detik dan waktu terlama 2,93 detik. Selain itu, penggunaan OCR dapat mereduksi waktu proses input data hingga 97% dibandingkan dengan metode manual. Namun, akurasi sistem bervariasi tergantung pada elemen data yang diekstrak, dengan elemen RT/RW memiliki tingkat keberhasilan terendah sebesar 70%. Berdasarkan hasil tersebut, disarankan untuk mengeksplorasi model OCR alternatif seperti EasyOCR atau TrOCR, serta meningkatkan dataset dan metode pra-proses untuk memperbaiki akurasi ekstraksi.

Kata Kunci: OCR, KTP, ekstraksi data, Tesseract, akurasi, efisiensi waktu

ABSTRACT

This study aims to evaluate the performance of an Optical Character Recognition (OCR) system in extracting data from Indonesian identity cards (KTP). The developed system was assessed based on three main aspects: extraction accuracy, processing time, and efficiency compared to manual data entry. The results show that the OCR system achieved an average accuracy rate of 90%. The average processing time for a single KTP image was 1.68 seconds, with the fastest time recorded at 1.26 seconds and the slowest at 2.93 seconds. Additionally, the implementation of OCR reduced data input time by up to 97% compared to manual methods. However, extraction accuracy varied across different data fields, with RT/RW achieving the lowest success rate at 70%. Based on these findings, it is recommended to explore alternative OCR models such as EasyOCR or TrOCR, as well as improve datasets and preprocessing methods to enhance extraction accuracy.

Keywords: OCR, KTP, data extraction, Tesseract, accuracy, time efficiency