

**IMPLEMENTASI *OPTICAL CHARACTER RECOGNITION* BERBASIS
CONVOLUTIONAL NEURAL NETWORK UNTUK EKSTRAKSI DATA
KTP-EL PADA MODUL PENGISIAN DATA DIRI E-MAGANG
KEMENTERIAN LUAR NEGERI**

Nuansa Syafrie Rahardian
H1D021083

ABSTRAK

Proses pengisian data manual pada platform e-magang Kementerian Luar Negeri yang tidak efisien dan rentan kesalahan diatasi melalui penelitian ini dengan mengembangkan dan mengintegrasikan sistem ekstraksi data otomatis dari KTP elektronik (KTP-el). Sistem ini menggunakan *metode Optical Character Recognition* (OCR) berbasis *Convolutional Neural Network* (CNN) yang didahului oleh tahap pra-pemrosesan dan segmentasi citra yang krusial untuk memisahkan baris, kata, dan karakter secara individual. Arsitektur sistem mencakup tiga model CNN yaitu, validasi citra, OCR NIK, dan OCR karakter umum yang diimplementasikan melalui API Flask terhubung dengan *backend* Laravel dan *frontend* React. Hasil pengujian menunjukkan performa model yang sangat tinggi, dengan akurasi model validasi KTP 100%, OCR NIK 100%, dan OCR karakter umum 99,95%. Namun, kinerja *end-to-end* sistem yang mencapai akurasi pembacaan karakter 93,31% membuktikan bahwa kualitas proses segmentasi menjadi faktor penentu utama keberhasilan ekstraksi akhir. Kesimpulannya, sistem ini berhasil meningkatkan efisiensi secara signifikan dengan waktu ekstraksi rata-rata 14,48 detik (lebih dari 50% lebih cepat dari input manual) dan terbukti akurat untuk otomatisasi verifikasi identitas, dengan catatan performa sangat bergantung pada keberhasilan segmentasi citra *input*.

Kata kunci: segmentasi karakter, CNN, e-magang, flask API, KTP-el, OCR

**IMPLEMENTATION OF CONVOLUTIONAL NEURAL NETWORK-BASED
OPTICAL CHARACTER RECOGNITION FOR KTP-EL DATA
EXTRACTION IN THE PERSONAL DATA ENTRY MODULE OF THE
MINISTRY OF FOREIGN AFFAIRS E-MAGANG SYSTEM**

Nuansa Syafrie Rahardian
H1D021083

ABSTRACT

The inefficient and error-prone manual data entry process on the Ministry of Foreign Affairs' e-magang platform is addressed in this research by developing and integrating an automatic data extraction system from electronic ID cards (KTP-el). This system uses the Optical Character Recognition (OCR) method based on a Convolutional Neural Network (CNN), preceded by a crucial preprocessing and image segmentation stage to separate rows, words, and characters individually. The system architecture includes three CNN models: image validation, NIK OCR, and general character OCR, which are implemented via a Flask API connected to a Laravel backend and a React frontend. Test results show very high model performance, with the KTP validation model achieving 100% accuracy, the NIK OCR model 100%, and the general character OCR model 99.95%. However, the system's end-to-end performance, which achieved a character reading accuracy of 93.31%, proves that the quality of the segmentation process is the main determining factor for final extraction success. In conclusion, the system successfully improves efficiency significantly with an average extraction time of 14.48 seconds (over 50% faster than manual input) and is proven accurate for identity verification automation, with the note that performance is highly dependent on the success of input image segmentation.

Kata kunci: *character segmentation, CNN, e-magang, flask API, KTP-el, OCR*