

INTISARI

SISTEM INFORMASI PERSURATAN DIGITAL BERBASIS WEB DENGAN TEKNOLOGI OCR (*OPTICAL CHARACTER RECOGNITION*) UNTUK PENGARSIPAN DOKUMEN

Luthfia Nisa Azzahra

21/474456/SV/18930

Sistem penyimpanan dan pencarian dokumen digital yang efisien hingga saat ini masih menjadi permasalahan serius, khususnya dalam mengelola volume surat yang terus bertambah. Proses pengajuan surat yang kurang terstruktur juga berpotensi menyebabkan kesalahan dan inkonsistensi. Salah satu instansi yang mengalami kendala tersebut adalah Bagian Akademik Departemen Teknik Elektro dan Informatika (DTEDI) UGM. Oleh sebab itu, dikembangkan sebuah sistem bernama SuratTEDI, yang memiliki fitur penyimpanan, pembuatan, dan pengajuan surat. Teknologi OCR (*Optical Character Recognition*) digunakan untuk mengatasi kendala pencarian surat dan memperkuat fitur penyimpanan. Penelitian ini mengembangkan sistem manajemen surat digital yang mengintegrasikan *framework* Laravel dengan Tesseract OCR melalui *wrapper* PyTesseract berbasis Python. Integrasi ini memungkinkan ekstraksi otomatis bagian-bagian surat, memfasilitasi pencarian, dan manajemen surat digital dengan lebih mudah. Pengujian akurasi OCR pada 20 dokumen (Tugas, Keterangan, Permohonan, Tidak Terklasifikasi) menunjukkan *Character Error Rate* (CER) sebesar 5,71% dan *Word Error Rate* (WER) sebesar 9,59%. Persentase tersebut membuktikan sistem mampu mengonversi dokumen berkualitas baik. Dengan *User Acceptance Testing* (UAT) mencapai 94,29%, SuratTEDI terbukti diterima dan layak digunakan oleh staf Bidang Akademik maupun mahasiswa DTEDI UGM. Diharapkan sistem ini dapat menjadi solusi efektif yang meningkatkan efisiensi administrasi DTEDI.

Kata kunci : Persuratan, OCR, Laravel, Python, PyTesseract

ABSTRACT

WEB-BASED DIGITAL CORRESPONDENCE INFORMATION SYSTEM USING OCR (OPTICAL CHARACTER RECOGNITION) TECHNOLOGY FOR DOCUMENT ARCHIVING

Luthfia Nisa Azzahra

21/474456/SV/18930

Efficient digital document storage and retrieval systems remain a significant challenge, especially with the continuously increasing volume of correspondence. Unstructured letter submission processes also carry the risk of errors and inconsistencies. The Academic Section of the Department of Electrical and Information Engineering (DTEDI) at UGM is one institution experiencing these obstacles. To address these issues, a system named SuratTEDI was developed, featuring letter storage, creation, and submission functionalities. Optical Character Recognition (OCR) technology is employed to overcome letter retrieval difficulties and enhance the storage feature. This research involved developing a digital letter management system that integrates the Laravel framework with Tesseract OCR via a Python-based PyTesseract wrapper. This integration enables the automatic extraction of letter components, facilitating easier digital letter search and management. Accuracy testing of the OCR on 20 documents (Tasks, Statements, Applications, Unclassified) revealed a Character Error Rate (CER) of 5.71% and a Word Error Rate (WER) of 9.59%. These percentages demonstrate the system's capability to accurately convert good-quality documents. With a User Acceptance Testing (UAT) score of 94.29%, SuratTEDI is proven to be accepted and suitable for use by both the Academic Staff and students of DTEDI UGM. It is anticipated that this system will serve as an effective solution, enhancing the administrative efficiency of DTEDI.

Keywords: Letters, OCR, Laravel, Python, PyTesseract