UNIVERSITAS
GADJAH MADA
Analisis Perbandingan Kinerja EasyOCR dan PaddleOCR untuk Ekstraksi Data Drilling Report pada Dashboard Berbasis Web di PT Parama Data Unit
Erlangga Syifa Sutrisno, Dr.Eng. Ir. Ganjar Alfian, S.T., M.Eng.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

ANALISIS PERBANDINGAN KINERJA EASYOCR DAN PADDLEOCR UNTUK EKSTRAKSI DATA *DRILLING REPORT* PADA *DASHBOARD* BERBASIS *WEB* DI PT PARAMA DATA UNIT

Erlangga Syifa Sutrisno
22/493922/SV/20754

Pengelolaan arsip *Drilling Reports* di PT. Parama Data Unit saat ini menghadapi tantangan efisiensi akibat format dokumen digital yang bersifat tidak terstruktur (*unstructured data*). Ketergantungan pada proses rekapitulasi data secara manual dengan menyalin data satu per satu dari berkas PDF ke dalam format Excel atau sistem basis data tidak hanya memakan waktu, tetapi juga rentan terhadap kesalahan manusia (*human error*) yang dapat memengaruhi integritas data operasional. Penelitian ini bertujuan untuk mengembangkan sistem otomatisasi ekstraksi informasi berbasis web menggunakan teknologi *Optical Character Recognition* (OCR) dan *Deep Learning*. Penelitian diawali dengan analisis komparasi antara model EasyOCR dan PaddleOCR untuk menentukan performa terbaik dalam mengenali teks dan struktur tabel kompleks pada dokumen PDF digital *Drilling Reports*. Berdasarkan hasil pengujian, PaddleOCR menunjukkan performa yang unggul dengan tingkat akurasi rata-rata *Character Error Rate* (CER) sebesar 0,3607, *Word Error Rate* (WER) sebesar 0,3997, dan F1-Score sebesar 0,8589, serta kecepatan pemrosesan yang lebih efisien (rata-rata 15,00 detik per dokumen) dibandingkan EasyOCR. Sistem ini diimplementasikan menggunakan PaddleOCR sebagai mesin ekstraksi utama yang terintegrasi ke dalam dasbor visualisasi berbasis web, dan diharapkan dapat mempercepat proses digitalisasi arsip serta mempermudah pemantauan parameter pengeboran secara operasional.

Kata kunci : *Drilling Reports*, Ekstraksi Informasi, PaddleOCR, EasyOCR, Visualisasi Data.

UNIVERSITAS
GADJAH MADA
Analisis Perbandingan Kinerja EasyOCR dan PaddleOCR untuk Ekstraksi Data Drilling Report pada Dashboard Berbasis Web di PT Parama Data Unit
Erlangga Syifa Sutrisno, Dr.Eng. Ir. Ganjar Alfian, S.T., M.Eng.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

COMPARATIVE PERFORMANCE ANALYSIS OF EASYOCR AND PADDLEOCR FOR DRILLING REPORT DATA EXTRACTION ON A WEB-BASED DASHBOARD AT PT PARAMA DATA UNIT

Erlangga Syifa Sutrisno
22/493922/SV/20754

The management of Drilling Reports archives at PT. Parama Data Unit currently faces efficiency challenges due to the unstructured nature of digital documents. The reliance on manual data recapitulation processes, specifically copying data one by one from PDF files into Excel or database systems, is not only time-consuming but also prone to human error, which can compromise the integrity of operational data. This study aims to develop a web-based information extraction automation system utilizing Optical Character Recognition (OCR) and Deep Learning technologies. The research begins with a comparative analysis between EasyOCR and PaddleOCR models to determine the best performance in recognizing text and complex table structures in digital PDF Drilling Reports. Based on the test results, PaddleOCR demonstrated superior performance with an average Character Error Rate (CER) of 0.3607, Word Error Rate (WER) of 0.3997, and an F1-Score of 0.8589, alongside a more efficient processing speed averaging 15.00 seconds per document compared to EasyOCR. The system is implemented using PaddleOCR as the primary extraction engine, which is integrated into a web-based visualization dashboard. This system is expected to accelerate the archive digitization process and facilitate the operational monitoring of drilling parameters.

Keywords: *Drilling Reports, Information Extraction, PaddleOCR, EasyOCR, Data Visualization.*