

Penerapan Algoritma Optical Character Recognition (OCR) Untuk Pelaporan Surat Pengantar Pengiriman BBM Di PT. Mandiri Melangkah Sejahtera

Muhammad Al Fajar¹, Yuvi Darmayunata², Roki Hardianto³

^{1,2,3}Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Lancang Kuning

^{1,2,3}Jl. Yos Sudarso KM. 8 Rumbai, Pekanbaru, Riau, telp. 0811 753 2015

e-mail: ¹muhammadalfajar218@gmail.com, ²yuvidarmayunata@unilak.ac.id,

³roki@unilak.ac.id

Abstrak

Pengelolaan dokumen di PT. Mandiri Melangkah Sejahtera, sebuah perusahaan transportir BBM, masih mengandalkan proses manual untuk input data Surat Pengantar Pengiriman BBM, yang mengakibatkan inefisiensi dan tingginya potensi kesalahan. Penelitian ini bertujuan untuk merancang dan membangun sebuah aplikasi mobile berbasis Optical Character Recognition (OCR) guna mengotomatisasi dan mempercepat proses pelaporan. Metode penelitian yang digunakan adalah metode prototype dengan mengembangkan aplikasi Android native menggunakan Kotlin. Pipeline OCR diimplementasikan dengan mengkombinasikan library OpenCV untuk tahap pra-pemrosesan gambar (seperti deskewing dan adaptive thresholding) dan Tesseract OCR untuk pengenalan teks. Pengujian sistem dilakukan melalui metode Black Box, White Box, serta pengujian kuantitatif terhadap akurasi OCR. Hasil pengujian menunjukkan bahwa implementasi pra-pemrosesan dengan OpenCV sangat krusial, terbukti berhasil meningkatkan Rata-rata Akurasi Kata dari 34.17% menjadi 76.9% dan menurunkan Word Error Rate (WER) secara signifikan dari 65.27% menjadi 24.7%. Kesimpulannya, aplikasi OCR berbasis mobile ini terbukti menjadi solusi yang layak dan efektif untuk meningkatkan efisiensi dan akurasi pelaporan dokumen di perusahaan.

Kata Kunci: Optical Character Recognition, Aplikasi Mobile, OpenCV, Tesseract, Efisiensi Operasional.

Abstract

Document management at PT. Mandiri Melangkah Sejahtera, a fuel transportation company, still relies on manual processes for data entry of Fuel Delivery Notes, resulting in inefficiencies and a high potential for errors. This research aims to design and build a mobile application based on Optical Character Recognition (OCR) to automate and accelerate the reporting process. The research method used is the prototype method by developing a native Android application using Kotlin. The OCR pipeline was implemented by combining the OpenCV library for the image preprocessing stage (such as deskewing and adaptive thresholding) and Tesseract OCR for text recognition. System testing was conducted using Black Box, White Box, and quantitative testing of OCR accuracy. The test results show that the implementation of preprocessing with OpenCV is crucial, proven to increase the average Word Accuracy from 34.17% to 76.9% and significantly reduce the Word Error Rate (WER) from 65.27% to 24.7%. In conclusion, this mobile-based OCR application proves to be a viable and effective solution for improving the efficiency and accuracy of document reporting in the company.

Keywords: Optical Character Recognition, Mobile Application, OpenCV, Tesseract, Operational Efficiency.

1. PENDAHULUAN

Pengelolaan dan pencatatan dokumen merupakan aspek krusial untuk memastikan kelancaran operasional serta kepatuhan terhadap regulasi di suatu perusahaan, terutama pada sektor yang diawasi ketat seperti minyak dan gas [1]. Dokumen tidak hanya berfungsi sebagai sumber informasi, tetapi juga menjadi bukti resmi dalam proses audit internal maupun eksternal, di mana kelengkapan dan keakuratannya menjadi penjamin keandalan audit keuangan [2]. Meskipun demikian, banyak perusahaan masih menghadapi tantangan efisiensi akibat proses pencatatan yang masih manual, yang diperparah oleh ketersediaan teknologi informasi perkantoran yang belum optimal sehingga data tidak tersimpan secara ideal [3].

Masalah ini secara spesifik terjadi di PT. Mandiri Melangkah Sejahtera, sebuah perusahaan transportir BBM, di mana proses input data Surat Pengantar Pengiriman BBM dilakukan secara manual. Proses ini terbukti lambat dan sangat rentan terhadap kesalahan input data. Untuk mengatasi hal tersebut, diperlukan sebuah inovasi berbasis teknologi Optical Character Recognition (OCR), yang mampu mengubah teks pada dokumen cetak menjadi format digital secara otomatis dan efisien [4]. Penerapan OCR pada sistem otomatis terbukti dapat mengurangi kesalahan dalam penginputan data, seperti yang telah ditunjukkan pada proses impor dan ekspor [5].

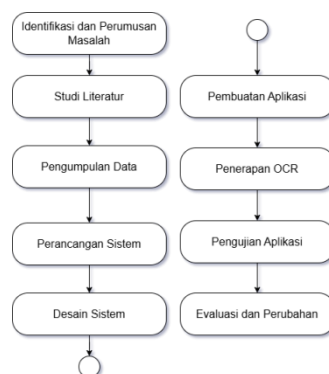
Penelitian ini berfokus pada penerapan teknologi OCR berbasis mobile untuk memberikan solusi yang fleksibel bagi staf lapangan (supir) di PT. Mandiri Melangkah Sejahtera. Kebaruan dari penelitian ini terletak pada implementasi pipeline OCR yang mengkombinasikan library OpenCV untuk tahap pra-pemrosesan gambar guna meningkatkan akurasi, dengan Tesseract OCR sebagai mesin pengenalan teksnya [6]. Selain itu, aplikasi ini memanfaatkan Firebase sebagai platform penyimpanan berbasis cloud yang mendukung sinkronisasi data secara real-time [7]. Dengan pendekatan ini, diharapkan dapat tercipta sebuah sistem pelaporan yang tidak hanya cepat dan akurat, tetapi juga mendukung kebutuhan transformasi digital di sektor industri.

2. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan yang terstruktur menggunakan metode pengembangan prototype.

2.1. Tahapan Penelitian

Proses penelitian dimulai dari identifikasi masalah di PT. Mandiri Melangkah Sejahtera, dilanjutkan dengan studi literatur, analisis kebutuhan sistem, perancangan, implementasi, hingga pengujian dan evaluasi akhir.



Gambar 1. Tahapan Penelitian

Berikut ini adalah 9 tahapan penelitian yaitu:

1. Identifikasi Masalah
Mengidentifikasi masalah proses input data yang manual, lambat, dan rentan kesalahan di PT. Mandiri Melangkah Sejahtera melalui observasi dan wawancara.
2. Studi Literatur
Melakukan kajian pustaka mengenai teknologi OCR, pengembangan aplikasi mobile, dan pemanfaatan cloud sebagai landasan teori.
3. Pengumpulan Data
Mengumpulkan data kebutuhan pengguna serta sampel dokumen Surat Pengantar Pengiriman BBM untuk dijadikan dataset pengujian.
4. Perancangan Sistem
Merancang arsitektur dan alur kerja sistem secara konseptual menggunakan metode prototype dan pemodelan UML.
5. Desain Sistem
Merancang antarmuka pengguna (UI) dan pengalaman pengguna (UX) yang intuitif dan efisien untuk aplikasi.
6. Pembuatan Aplikasi
Mengembangkan aplikasi Android secara native menggunakan Kotlin, Firebase, dan library terkait.
7. Penerapan OCR
Mengimplementasikan pipeline OCR yang mengkombinasikan OpenCV untuk pra-pemrosesan gambar dan Tesseract untuk pengenalan teks.
8. Pengujian Aplikasi
Melakukan pengujian fungsional (Black Box & White Box) dan pengujian akurasi OCR untuk memvalidasi sistem.
9. Evaluasi dan Perbaikan
Mengevaluasi seluruh hasil pengujian untuk melakukan perbaikan minor dan finalisasi aplikasi.

2.2. Metode Pengujian OCR

Pengujian ini bertujuan untuk menilai secara kuantitatif efektivitas pipeline pra-pemrosesan gambar. Pengujian dilakukan dengan membandingkan akurasi ekstraksi data dari dua skenario:

1. Skenario A: Ekstraksi teks menggunakan Tesseract OCR tanpa pra-pemrosesan gambar.
2. Skenario B: Ekstraksi teks menggunakan Tesseract OCR dengan pra-pemrosesan gambar penuh menggunakan OpenCV.

2.2.1. Metodologi Pengujian Akurasi

Akurasi dievaluasi menggunakan tiga metrik standar: Character Error Rate (CER), Word Error Rate (WER), dan Akurasi OCR Keseluruhan.

1. Character Error Rate (CER)

Mengukur tingkat kesalahan pada tingkat karakter. Metrik ini sangat peka terhadap kesalahan sekecil apa pun. Semakin rendah nilai CER, semakin tinggi tingkat akurasi OCR [8].

Rumus :

$$CER\% = \frac{S_{char} + D_{char} + I_{char}}{N_{char}} \quad (1)$$

Dimana :

Schar = Jumlah Substitusi (karakter yang diganti).

Dchar = Jumlah Delesi (karakter yang hilang).

Ichar = Jumlah Inseri (karakter tambahan).

Nchar = Jumlah total karakter pada teks referensi (Ground Truth).

2. Word Error Rate (WER)

Mengukur tingkat kesalahan pada tingkat kata. Metrik ini lebih mencerminkan pemahaman kontekstual. Sebuah kata dianggap salah jika terdapat satu karakter yang salah, hilang, atau ditambahkan[9].

Rumus :

$$WER\% = \frac{S+D+I}{N} \quad (2)$$

Dimana:

S = Jumlah Substitusi Kata.

D = Jumlah Delesi Kata.

I = Jumlah Inseri Kata.

N = Jumlah total kata pada teks referensi (Ground Truth).

3. Akursasi OCR secara Keseluruhan

Metrik ini menghitung tingkat akurasi dengan membandingkan jumlah kata yang teridentifikasi dengan benar dalam hasil OCR (output) terhadap total kata yang terdapat dalam teks referensi (ground truth). Tujuan dari pengukuran ini adalah untuk menilai seberapa efektif sistem OCR dalam mengenali kata-kata secara keseluruhan [10].

Rumus :

$$Akurasi \% = \frac{Kata\ Benar}{Total\ Kata}$$

2.2.2. Data Uji dan Ground Truth

Sebagai dasar perbandingan, digunakan satu dokumen sampel dengan data yang benar (*Ground Truth*) [11] yang telah ditetapkan sebagai berikut:

Tabel 1. Data Uji (Ground Truth)

Field	Ground Truth	Total N _{char}	Total N
Nomor Polisi	BM39880U	8	1
Shipment No	37323254 Ref:7903932	20	2
Nama Pengemudi	APMS_3863 DONI EFENDI	22	3
	907893 PT. NAJISA YUSNA KARYA ABADI	90	15
Tujuan	SPBU 15283047		
	DS. KIJANG JAYA, KEC. TAPUNG HILIR, KAMPAR		
Produk	PERTALITE	9	1
Nomor SO/SA	4031814711	10	1
Nomor LO	8116414143	10	1
Jml Pemesanan (KL)	8.0	3	1
Dens/Temp (OBS)	0.75/29.0	9	1
Tanggal Pengiriman	2024-11-26	10	1
Jam Keluar	10:24:37	8	1
Nomor Segel	A-2699428,A-2699429,A-2699430,	30	3

3. HASIL DAN PEMBAHASAN

3.1. Implementasi Pipeline OCR Proses OCR

Dalam penelitian ini merupakan alur kerja terstruktur yang melibatkan persiapan library, pra-pemrosesan, ekstraksi teks, hingga pengolahan data.

1. Enkapsulasi Logika dengan TessOCRHelper: Seluruh logika terkait Tesseract OCR dikemas dalam sebuah kelas bantuan. Kelas ini menangani inisialisasi cerdas (memastikan file data terlatih ind.traineddata siap digunakan), menyediakan konfigurasi lanjutan (seperti setPageSegMode dan setCharWhitelist), serta mengelola proses ekstraksi teks dan pelepasan sumber daya untuk mencegah kebocoran memori.
2. Alur Proses Ekstraksi:
 - a. Pra-pemrosesan Gambar dengan OpenCV: Sebelum diserahkan ke Tesseract, gambar bitmap menjalani tahap pra-pemrosesan yang krusial. Tahapan ini mencakup:
 - 1) Scaling: Menyesuaikan ukuran gambar ke resolusi optimal.
 - 2) Konversi Grayscale: Mengubah gambar menjadi skala abu-abu.
 - 3) Deskewing: Memperbaiki kemiringan gambar untuk meluruskan teks.
 - 4) Adaptive Thresholding: Mengubah gambar menjadi format biner (hitam-putih) yang bersih, sangat efektif untuk kondisi pencahayaan tidak merata.
 - b. Ekstraksi dan Parsing Data: Setelah gambar bersih, TessOCRHelper mengekstrak blok teks mentah. Teks ini kemudian diolah menggunakan Regular Expressions (Regex) untuk menemukan kata kunci (seperti "NO.POLISI", "NAMA PENGEMUDI", dll.) dan mengekstrak nilai terkait, mengubah data tidak terstruktur menjadi informasi terstruktur.

3.2. Analisis Kinerja Skenario A (Tanpa OpenCV)



Gambar 2. Log Output OCR Tanpa Pre-processing

Berikut adalah contoh output OCR mentah yang mewakili log sebelum penerapan OpenCV. Teks tersebut menunjukkan tingkat noise yang tinggi, kurang terstruktur dengan baik, dan beberapa kata hilang.

Tabel 2. Hasil Perhitungan CER (%) pada Skenario A

Field	Hasil Ekstraksi	S _{char}	D _{char}	I _{char}	CER (%)
Nomor Polisi	null	0	8	0	100.0%
Shipment No	37323254	0	10	0	44.4%
Nama Pengemudi	null	0	19	0	100.0%
Tujuan	907893 PT NAJISA ...ma : SPBU 15283047 V : ! DS, KIJANG JAYA, KE C TW : HILIR,KAMPAR — ?	3	29	21	58.8%
Produk	PERTAUTE	1	1	0	
Nomor SO/SA	4031814711	0	0	0	22.2%
Nomor LO	8116414143	0	0	0	0.0%
Jml Pemesanan (KL)	8 0	1	0	0	0.0%
Dens/Temp (OBS)	(1751291)	5	0	0	33.3%
Tanggal Pengiriman	20244 1—26	3	0	0	55.5%
Jam Keluar	10:24:37	0	0	0	30.0%

Nomor Segel	A-2699428,A—2699429,A— 2699430	2	1	0	0.0%
Rata - Rata					70.83%

Tabel 3. Hasil Perhitungan WER (%) pada Skenario A

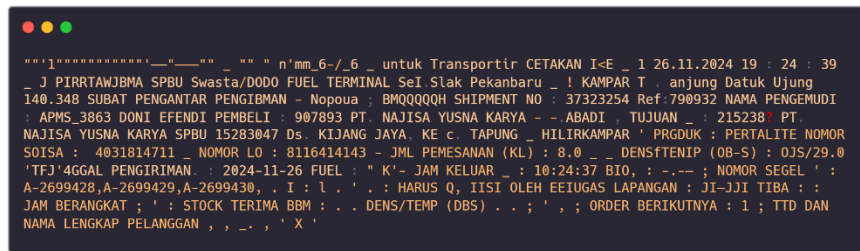
Field	Hasil Ekstraksi	S	D	I	WER (%)
Nomor Polisi	null	0	8	0	100.0%
Shipment No	37323254	0	10	0	50.0%
Nama Pengemudi	null	0	19	0	100.0%
Tujuan	907893 PT NAJISA ...ma : SPBU 15283047 V : ! DS, KIJANG JAYA, KE C TW : HIL!R,KAMPAR — ?	3	29	21	133.3%
Produk	PERTAUTE	1	1	0	100.0%
Nomor SO/SA	4031814711	0	0	0	0.0%
Nomor LO	8116414143	0	0	0	0.0%
Jml Pemesanan (KL)	8 0	1	0	0	100.0%
Dens/Temp (OBS)	(1751291)	5	0	0	100.0%
Tanggal Pengiriman	20244 1—26	3	0	0	100.0%
Jam Keluar	10:24:37	0	0	0	0.0%
Nomor Segel	A-2699428,A—2699429,A— 2699430	2	1	0	66.6%
Rata - Rata					70.83%

Tabel 4. Hasil Perhitungan Akurasi OCR pada Skenario A

Field	Hasil Ekstraksi	Jumlah Kata Benar	Akurasi OCR (%)
Nomor Polisi	null	0	0.0%
Shipment No	37323254	1	50.0%
Nama Pengemudi	null	0	0.0%
Tujuan	907893 PT NAJISA ...ma : SPBU 15283047 V : ! DS, KIJANG JAYA, KE C TW : HIL!R,KAMPAR — ?	4	26.7%
Produk	PERTAUTE	0	0.0%
Nomor SO/SA	4031814711	1	100.0%
Nomor LO	8116414143	1	100.0%
Jml Pemesanan (KL)	8 0	0	0.0%
Dens/Temp (OBS)	(1751291)	0	0.0%
Tanggal Pengiriman	20244 1—26	0	0.0%
Jam Keluar	10:24:37	1	100%
Nomor Segel	A-2699428,A—2699429,A— 2699430	1	33.3%
Rata - Rata			34.17%

Berdasarkan analisis yang dilakukan pada Skenario A, dapat disimpulkan bahwa penerapan OCR tanpa melalui tahap pre-processing gambar menghasilkan tingkat akurasi yang sangat rendah, sehingga tidak dapat diandalkan untuk penggunaan praktis.

3.3. Analisis Kinerja Skenario B (Dengan OpenCV)



Gambar 3. Log Output OCR Tanpa Pre-processing

Berikut adalah contoh output OCR yang representatif setelah gambar diproses melalui pipeline pre-processing OpenCV. Teks yang dihasilkan tampak lebih rapi dan terstruktur.

Tabel 5. Hasil Perhitungan CER (%) pada Skenario B

Field	Hasil Ekstraksi	S _{char}	D _{char}	I _{char}	CER (%)
Nomor Polisi	BMQQQQQH	8	0	0	100.0%
Shipment No	37323254 Ref:790932	1	0	0	5.0%
Nama Pengemudi	APMS_3863 DONI EFENDI	0	0	0	0.0%
Tujuan	215238? PT. NAJISA YUSNA KARYA SPBU 15283047 Ds. KIJANG JAYA, KE c. TAPUNG _ HILIRKAMPAR	9	7	1	18.5%
Produk	PERTALITE	0	0	0	0.0%
Nomor SO/SA	4031814711	0	0	0	0.0%
Nomor LO	8116414143	0	0	0	0.0%
Jml Pemesanan (KL)	8.0	0	0	0	0.0%
Dens/Temp (OBS)	OJS/29.0	3	0	0	33.3%
Tanggal Pengiriman	2024-11-26	0	0	0	0.0%
Jam Keluar	10:24:37	0	0	0	0.0%
Nomor Segel	A-2699428,A-2699429,A-2699430,	0	0	0	0.0%
Rata - Rata					13.1%

Tabel 6. Hasil Perhitungan WER (%) pada Skenario B

Field	Hasil Ekstraksi	S	D	I	WER (%)
Nomor Polisi	BMQQQQQH	1	0	0	100%
Shipment No	37323254 Ref:790932	1	0	0	50.0%
Nama Pengemudi	APMS_3863 DONI EFENDI	0	0	0	0.0%
Tujuan	215238? PT. NAJISA YUSNA KARYA SPBU 15283047 Ds. KIJANG JAYA, KE c. TAPUNG _ HILIRKAMPAR	2	2	3	46.7%
Produk	PERTALITE	0	0	0	0.0%
Nomor SO/SA	4031814711	0	0	0	0.0%
Nomor LO	8116414143	0	0	0	0.0%
Jml Pemesanan (KL)	8.0	0	0	0	0.0%
Dens/Temp (OBS)	OJS/29.0	1	0	0	100%
Tanggal Pengiriman	2024-11-26	0	0	0	0.0%
Jam Keluar	10:24:37	0	0	0	0.0%

Nomor Segel	A-2699428,A-2699429,A-2699430,	0	0	0	0.0%
Rata - Rata					24.7%

Tabel 7. Hasil Perhitungan Akurasi OCR pada Skenario B

Field	Hasil Ekstraksi	Jumlah Kata Benar	Akurasi OCR (%)
Nomor Polisi	BMQQQQQH	0	0.0%
Shipment No	37323254 Ref:790932	1	50.0%
Nama Pengemudi	APMS_3863 DONI EFENDI	3	100.0%
Tujuan	215238? PT. NAJISA YUSNA KARYA SPBU 15283047 Ds. KIJANG JAYA, KE c. TAPUNG _ HILIRKAMPAR	11	73.3%
Produk	PERTALITE	1	100.0%
Nomor SO/SA	4031814711	1	100.0%
Nomor LO	8116414143	1	100.0%
Jml Pemesanan (KL)	8.0	1	100.0%
Dens/Temp (OBS)	OJS/29.0	0	0.0%
Tanggal Pengiriman	2024-11-26	1	100.0%
Jam Keluar	10:24:37	1	100.0%
Nomor Segel	A-2699428,A-2699429,A-2699430,	3	100.0%
Rata - Rata			76.9%

Dengan penerapan pre-processing menggunakan OpenCV, kualitas hasil OCR mengalami peningkatan yang signifikan. Meskipun belum mencapai kesempurnaan, sebagian besar data kini dapat diekstraksi dengan akurasi yang baik atau hanya dengan kesalahan yang minimal.

3.4. Analisis Kinerja Skenario B (Dengan OpenCV)

Perbandingan langsung antara kedua skenario menunjukkan dampak signifikan dari penerapan OpenCV.

TABEL 8. Hasil Perhitungan Akurasi OCR pada Skenario B

Metrik	Sebelum OpenCV	Sesudah OpenCV	Peningkatan Kinerja
Rata-rata Akurasi Kata	34.17%	76.9%	+42.73 poin
Rata-rata CER	37.85%	13.1%	Penurunan 24.75 poin
Rata-rata WER	70.83%	24.7%	Penurunan 46.13 poin

Berdasarkan analisis terhadap Skenario A, dapat disimpulkan bahwa penerapan OCR tanpa melalui tahap pre-processing gambar menghasilkan tingkat akurasi yang sangat rendah dan tidak dapat diandalkan untuk aplikasi praktis. Hal ini terlihat dari Rata-rata Akurasi Kata yang hanya mencapai 34,17%, serta tingginya tingkat kesalahan rata-rata pada CER (37,85%) dan WER (70.83%). Angka kesalahan yang tinggi ini menunjukkan bahwa tanpa pre-processing, Tesseract OCR mengalami kesulitan dalam menangani noise dan variasi tata letak dokumen, yang mengakibatkan banyak informasi penting tidak berhasil diekstraksi. Temuan ini menegaskan bahwa pre-processing gambar merupakan langkah yang sangat penting untuk memastikan sistem OCR dapat beroperasi secara efektif.

4. KESIMPULAN

Berdasarkan analisis kuantitatif yang telah dilakukan, dapat ditarik beberapa kesimpulan utama. Pertama, implementasi *Optical Character Recognition* (OCR) menggunakan Tesseract tanpa adanya tahap pra-pemrosesan menghasilkan tingkat akurasi yang sangat rendah (34.17%), sehingga tidak efektif untuk penggunaan praktis di lingkungan industri. Kedua, penerapan *pipeline* pra-pemrosesan gambar menggunakan OpenCV terbukti secara signifikan meningkatkan kinerja OCR. Hal ini ditunjukkan oleh peningkatan Rata-rata Akurasi Kata menjadi 76.9% dan penurunan drastis pada Word Error Rate (WER) sebesar 46.13 poin. Temuan ini menegaskan bahwa tahap persiapan gambar, seperti *deskewing* dan *adaptive thresholding*, merupakan langkah krusial yang menentukan keberhasilan sistem ekstraksi data otomatis.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Bapak Yuvi Darmayunata, ST, M.Kom, selaku dosen pembimbing yang telah memberikan bimbingan, arahan, dan dukungan yang sangat berharga selama proses penelitian dan penyusunan artikel ini. Ucapan terima kasih juga ditujukan kepada pihak PT. Mandiri Melangkah Sejahtera yang telah memberikan izin dan kesempatan untuk melakukan penelitian, serta seluruh pihak yang telah membantu hingga artikel ini dapat terselesaikan.

DAFTAR PUSTAKA

- [1] F. Tjiptabudi and R. Ndaumanu, "Implementasi Aplikasi Arsip Dokumen Untuk Mendukung Proses Bisnis Pada Pt. XXX Implementation of Document Archive Application to Support Business Processes at PT. XXX," Pontianak, May 2024.
- [2] L. Sava and E. Petreanu, "Audit Documentation – An Important Factor In Financial Audit Mission," *JOURNAL OF SOCIAL SCIENCES*, vol. 7, no. 2, pp. 15–24, Jul. 2024, doi: 10.52326/jss.utm.2024.7(2).02.
- [3] M. F. L. Ambarwati and D. Adianti, "Efisiensi dalam Pengelolaan Dokumen Berbasis Digital," *Jurnal Administrasi dan Kesekretarisan*, vol. 7, no. 1, pp. 95–109, Mar. 2022, doi: 10.36914/jak.v7i1.767.
- [4] S. Sharmin, T. Mim, and M. Rahman, "Bangla Optical Character Recognition for Mobile Platforms: A Comprehensive Cross-Platform Approach," *American Journal of Electrical and Computer Engineering*, vol. 8, no. 2, pp. 31–42, Sep. 2024, doi: 10.11648/j.ajece.20240802.12.
- [5] H. Cho, J. Jeong, C. Lee, S. Yun, K. Bang, and J. Byun, "Design and Implementation for BIC Code Recognition System of Containers using OCR and CRAFT in Smart Logistics," *WSEAS Transactions on Computer Research*, vol. 11, pp. 62–72, 2023, doi: 10.37394/232018.2023.11.6.
- [6] Dr. S. Mubeen, J. Brahmani, D. P. Kalyan, A. Jagirdar, and A. P. Kumar, "Optical Character Recognition Using Tesseract," *Int J Res Appl Sci Eng Technol*, vol. 10, no. 11, pp. 672–675, Nov. 2022, doi: 10.22214/ijraset.2022.47414.
- [7] A. Hudawi, M. Kamalio, and M. Syafi'i, "Sistem Informasi Pengarsipan Dokumen Berbasis Framework Codeigniter di Tata Usaha Fakultas Teknik Universitas Nurul Jadid Probolinggo," *TRILOGI: Jurnal Ilmu Teknologi, Kesehatan, dan Humaniora*, vol. 3, no. 3, Dec. 2022, doi: 10.33650/trilogi.v3i3.3893.
- [8] M. N. Darpito, K. Firdausy, and A. Fadlil, "Perbandingan Unjuk Kerja Library Optical Character Recognition (OCR) dalam Pengenalan Teks pada Dokumen Digital," *JIP (Jurnal Informatika Polinema)*, vol. 11, no. Vol. 11 No. 3 (2025): Vol. 11 No. 3 (2025), pp. 273–81, May 2025, doi: 10.33795/jip.v11i3.7025.
- [9] A. F. I. Setyadi and Y. A. Susetyo, "Implementasi Algoritma LSTM pada Aplikasi Optical Character Recognition Berbasis Website Menggunakan Tesseract OCR," *Jurnal*

- Teknologi Sistem Informasi dan Aplikasi*, vol. 6, no. 2, pp. 63–71, Apr. 2023, doi: 10.32493/jtsi.v6i2.29235.
- [10] H. Holila, A. R. Pratama, S. A. P. Lestari, and J. Indra, “INTRODUCTION NATIONAL IDENTIFICATION NUMBER AND NAME ON ID CARD USING OCR (OPTICAL CHARACTER RECOGNITION) METHOD,” *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 4, pp. 1191–1196, Jul. 2024, doi: 10.52436/1.jutif.2024.5.4.2242.
- [11] Y. Maurer, P. Schneider, and R. Marschall, “Nautilus – An End-To-End METS/ALTO OCR Enhancement Pipeline,” *LIBER Quarterly*, vol. 33, no. 1, Feb. 2023, doi: 10.53377/lq.13330.



Prosiding- SEMASTER: Seminar Nasional Teknologi Informasi & Ilmu Komputer is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)