

Automated Data Extraction From Student Id Cards (Ktm) Using Optical Character Recognition (Ocr) And Regular Expressions Based On Tkinter Gui

Andes Putra Gunawan ¹⁾; Yulia Darmi ²⁾; Ardi Wijaya ³⁾; Yuza Reswan ⁴⁾

^{1,2,3,4)}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Bengkulu

Email: ¹⁾ andesputragunawan96@gmail.com

How to Cite :

Gunawan, A., P., Darmi, Y., Wijaya, Ardi., Reswan, Y. (2026). Automated Data Extraction from Student ID Cards (KTM) Using Optical Character Recognition (OCR) and Regular Expressions Based on Tkinter GUI. Jurnal Media Computer Science, 5(4).

ARTICLE HISTORY

Received [31 Agustus 2026]

Revised [25 Oktober 2026]

Accepted [28 Oktober 2026]

KEYWORDS

Data Extraction, Student ID Card, Optical Character Recognition (OCR), Regular Expression, Tesseract, Confusion Matrix, Tkinter.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



ABSTRAK

Penelitian ini bertujuan merancang dan mengimplementasikan sistem ekstraksi data otomatis dari citra KTM menggunakan kombinasi *Optical Character Recognition* (OCR) berbasis Tesseract dan *Regular Expression*, dengan antarmuka pengguna grafis berbasis Tkinter. Sistem yang dibangun terdiri atas tahap pra-pemrosesan citra (pemangkasan zona teks, koreksi kemiringan dengan penjagaan terhadap latar belakang bertekstur, dan binerisasi Otsu), ekstraksi teks OCR menggunakan strategi multi-sumber dan multi-mode segmentasi halaman, serta ekstraksi *field* data menggunakan pola *Regular Expression* bertingkat yang dilengkapi kamus pemetaan kata kunci untuk *field* Program Studi dan Fakultas. Pengujian dilakukan terhadap 100 citra KTM Universitas Muhammadiyah Bengkulu menggunakan metrik evaluasi *confusion matrix* berupa *Precision*, *Recall*, *F1-Score*, dan *Accuracy* untuk setiap *field* data. Hasil pengujian menunjukkan sistem mencapai akurasi keseluruhan sebesar 93,8%, dengan *Precision* 94,86%, *Recall* 98,72%, dan *F1-Score* 96,75%. Akurasi per-*field* menunjukkan NPM mencapai 100%, Fakultas 98%, Program Studi 95%, TTL 94%, dan Nama 82%. Penelitian ini juga menemukan dan mengatasi permasalahan lintas platform pada pustaka OCR yang menyebabkan hilangnya spasi dan tanda baca pada hasil ekstraksi ketika dijalankan pada sistem operasi Windows. Hasil penelitian menunjukkan bahwa kombinasi OCR dan *Regular Expression* yang dirancang khusus sesuai karakteristik visual KTM mampu menghasilkan sistem ekstraksi data yang akurat dan layak digunakan untuk mendukung proses administrasi akademik.

ABSTRACT

This research aims to design and implement an automatic data extraction system from student ID card (KTM) images using a combination of Tesseract-based Optical Character Recognition (OCR), and Regular Expression, with a graphical user interface built using Tkinter. The developed system consists of an image preprocessing stage (text-zone cropping, skew correction with protection against textured backgrounds, and Otsu binarization), OCR text extraction using a multi-source and multi-page-segmentation-mode strategy, and data field extraction using tiered Regular Expression patterns supported by a keyword-mapping dictionary for the Study Program and Faculty fields. Testing was conducted on 100 student ID card images from Universitas Muhammadiyah Bengkulu using confusion matrix evaluation metrics, namely Precision, Recall, F1-Score, and Accuracy for each data field. The test results show that the system achieved an overall accuracy of 93.8%, with a Precision of 94.86%, Recall of 98.72%, and F1-Score of 96.75%. Per-field accuracy results show that the Student ID Number (NPM) field reached 100%, Faculty 98%, Study Program 95%, Place/Date of Birth 94%, and Name 82%. This research also identified and resolved a cross-platform issue in the OCR library that caused the loss of spaces and punctuation marks in the extraction results when run on the Windows operating system. The findings indicate that the combination of OCR and Regular Expression, specifically designed to match the visual characteristics of student ID cards, is able to produce an accurate and reliable data extraction system suitable for supporting academic administration processes.

PENDAHULUAN

Di era transformasi digital yang terus berkembang pesat, institusi pendidikan tinggi dituntut untuk mampu mengelola data mahasiswa secara lebih efisien dan akurat. Salah satu dokumen identitas penting yang diterbitkan oleh setiap perguruan tinggi adalah Kartu Tanda Mahasiswa (KTM), yang memuat informasi vital seperti nama lengkap, Nomor Induk Mahasiswa (NIM), program studi, dan fakultas (Reswan *et al.*, 2024). Proses pencatatan data dari KTM masih banyak dilakukan secara manual oleh petugas administrasi, yakni dengan menyalin informasi yang tercetak di kartu tersebut ke dalam sistem komputer satu per satu. Cara kerja seperti ini tentu memakan banyak waktu dan berpotensi menimbulkan kesalahan entri data yang tidak disadari, terutama ketika volume data yang harus diproses cukup besar. Proses pembuatan dan pengelolaan KTM yang masih dilakukan secara manual mengharuskan data mahasiswa diinput dan diverifikasi satu per satu, sehingga rawan terjadi kesalahan dan memerlukan waktu yang cukup lama. Kondisi ini menjadi salah satu penghambat utama dalam upaya meningkatkan kualitas layanan administrasi akademik di berbagai kampus, khususnya yang masih mengandalkan proses konvensional.

Permasalahan inti dari penelitian ini adalah bagaimana merancang sebuah sistem yang mampu mengotomatiskan proses pembacaan dan ekstraksi data dari gambar KTM secara akurat, cepat, dan mudah digunakan. Secara khusus, persoalan teknis yang muncul meliputi bagaimana memastikan mesin dapat mengenali karakter pada citra KTM dengan presisi tinggi, bagaimana memisahkan setiap field data secara tepat dari teks mentah hasil OCR, serta bagaimana menyajikan antarmuka pengguna yang intuitif agar dapat dioperasikan oleh staf non-teknis sekalipun.

Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem ekstraksi data otomatis dari citra KTM menggunakan kombinasi teknologi OCR dan *Regular Expression* dengan desain antarmuka grafis berbasis Tkinter. Secara lebih rinci, tujuan yang ingin dicapai yaitu membangun alur pemrosesan citra yang mampu mempersiapkan gambar KTM sebelum dikenali oleh mesin OCR, mengimplementasikan Tesseract OCR untuk mengubah citra menjadi teks, merancang pola *Regular Expression* yang tepat untuk mengekstrak field data spesifik seperti nama, NIM, dan program studi, serta menyajikan seluruh fungsionalitas tersebut dalam aplikasi desktop berbasis GUI Tkinter yang dapat dioperasikan dengan mudah (Julian Mulyadi and Nuri David Maria Veronika, 2025).

Pendekatan yang digunakan dalam penelitian ini memadukan beberapa teknologi secara berurutan. Kebaruan utama penelitian ini terletak pada integrasi menyeluruh antara tiga komponen teknis pra-pemrosesan citra, OCR berbasis Tesseract, dan *parsing Regular Expression* ke dalam satu aplikasi desktop yang dapat langsung digunakan tanpa keahlian pemrograman (Setiawan, Guntara and Purwaamijaya, 2025). Penggunaan pola *Regular Expression* yang dirancang khusus untuk variasi format teks KTM (Nanda Aprillia *et al.*, 2025). Sementara implementasi Tesseract OCR untuk pengenalan karakter pada objek dokumen terstruktur telah menunjukkan tingkat akurasi rata-rata yang tinggi, penelitian ini mengembangkan lebih lanjut dengan menambahkan lapisan parsing berbasis pola dan antarmuka GUI yang belum dikombinasikan secara terpadu sebelumnya (Al Fajar, Darmayunata and Hardianto, 2025).

Hasil penelitian ini diharapkan membawa manfaat yang konkret dari dua sisi. Dari sisi praktis, aplikasi yang dihasilkan dapat langsung digunakan oleh petugas administrasi kampus untuk mempercepat proses pencatatan data mahasiswa, mengurangi risiko kesalahan manusiawi, dan menghemat waktu kerja secara signifikan. Dari sisi akademis, penelitian ini berkontribusi pada pengembangan literatur lokal mengenai penerapan OCR dan *Regular Expression* untuk dokumen identitas berbahasa Indonesia, serta menjadi rujukan bagi peneliti lain yang ingin mengembangkan sistem serupa dengan cakupan yang lebih luas, misalnya untuk jenis dokumen lain atau dengan mekanisme ekspor data ke berbagai format file.

Kartu Tanda Mahasiswa (KTM)

Kartu Tanda Mahasiswa (KTM) merupakan dokumen identitas resmi yang diterbitkan oleh perguruan tinggi sebagai bukti status kemahasiswaan seseorang. KTM memuat informasi vital mahasiswa seperti Nomor Induk Mahasiswa (NIM), nama lengkap, program studi, fakultas, dan tempat tanggal lahir (Reswan *et al.*, 2024). Informasi tersebut tercetak dengan format yang bervariasi antar institusi, namun secara umum mengikuti pola penulisan label diikuti tanda titik dua dan nilai data.

Optical Character Recognition (OCR)

Optical Character Recognition (OCR) adalah teknologi yang memungkinkan komputer untuk mengenali dan mengubah teks yang terdapat pada gambar atau dokumen hasil pemindaian menjadi teks yang dapat diedit dan diproses secara digital. OCR bekerja dengan menganalisis pola visual dari karakter pada citra dan mencocokkannya dengan basis data karakter yang dikenal (Rahmadani, Rozikin and Karawang, 2021).

Tesseract OCR Engine

Tesseract adalah mesin OCR open-source yang dikembangkan oleh Hewlett-Packard pada tahun 1985 dan kemudian dilanjutkan pengembangannya oleh Google sejak tahun 2006. Tesseract mendukung lebih dari 100 bahasa termasuk bahasa Indonesia, dan menyediakan dua mode mesin utama melalui parameter `--oem`, yaitu LSTM (Long Short-Term Memory) dan mode warisan (legacy engine) (Kusnantoro, Kusumaningrum, Sulistya and Rohana, 2022).

Prapemrosesan Citra Digital

Prapemrosesan citra merupakan serangkaian operasi yang diterapkan pada citra digital sebelum diproses lebih lanjut oleh sistem OCR. Tujuannya adalah meningkatkan kualitas citra sehingga karakter pada dokumen lebih mudah dikenali oleh mesin OCR.

OpenCV

OpenCV (Open Source Computer Vision Library) adalah pustaka open-source untuk pengolahan citra dan visi komputer yang dikembangkan oleh Intel. OpenCV menyediakan lebih dari 2500 algoritma yang dioptimalkan untuk berbagai tugas pengolahan citra termasuk filtering, transformasi geometris, segmentasi, dan deteksi fitur.

Regular Expression (Regex)

Regular Expression atau Regex adalah urutan karakter yang mendefinisikan pola pencarian dalam teks. Regex merupakan alat yang sangat kuat untuk pencarian, pencocokan, dan ekstraksi pola teks tertentu dari string yang lebih panjang (Yusuf *et al.*, 2025). Dalam konteks penelitian ini, Regex digunakan untuk mengekstrak field-field data terstruktur dari teks mentah hasil OCR.

Bahasa Pemrograman Python

Python adalah bahasa pemrograman tingkat tinggi yang bersifat *interpreted*, dinamis, dan mendukung berbagai paradigma pemrograman termasuk pemrograman berorientasi objek, prosedural, dan fungsional (Adawiyah Ritonga and Yahfizham Yahfizham, 2023). Python dipilih sebagai bahasa implementasi dalam penelitian ini karena ekosistem pustakanya yang luas untuk pengolahan citra (OpenCV), OCR (pytesseract), ekspresi reguler (modul *re* bawaan), pembuatan antarmuka grafis (Tkinter), serta pengolahan berkas Excel (openpyxl). Sifat Python yang bersifat lintas platform juga memungkinkan aplikasi yang dibangun dapat dijalankan pada berbagai sistem operasi tanpa perubahan kode yang signifikan.

Tkinter

Tkinter merupakan pustaka antarmuka pengguna grafis (*Graphical User Interface/GUI*) standar bawaan Python yang menyediakan lapisan berorientasi objek di atas *toolkit* Tk. Karena statusnya sebagai pustaka bawaan, Tkinter tidak memerlukan instalasi tambahan dan menjadikannya pilihan yang praktis untuk membangun aplikasi desktop tanpa dependensi eksternal. Tkinter menyediakan berbagai *widget* siap pakai seperti label, tombol, kotak masukan teks, area teks, serta *notebook* untuk antarmuka bertab, yang seluruhnya dimanfaatkan dalam penelitian ini untuk membangun tiga tab fungsional yaitu Pengujian Satu Citra, Pengujian Seluruh Data, dan Confusion Matrix.

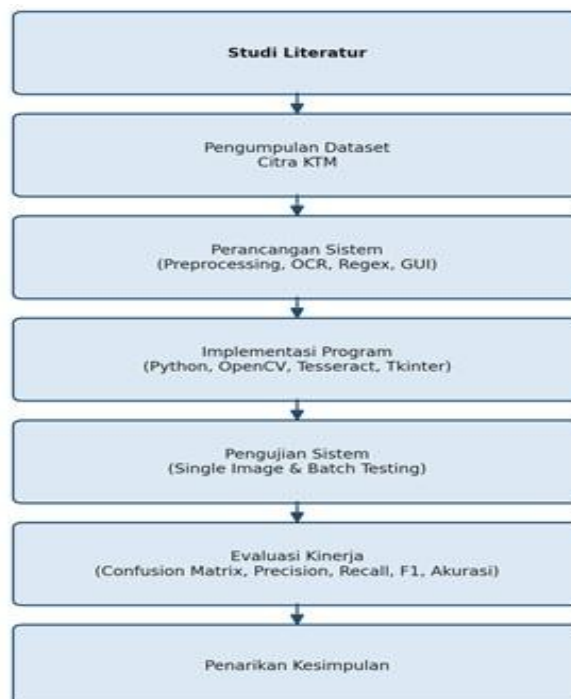
Evaluasi Kinerja Sistem (*Confusion Matrix*)

Confusion matrix adalah tabel yang digunakan untuk mengukur kinerja suatu sistem klasifikasi atau ekstraksi dengan membandingkan hasil prediksi sistem terhadap nilai sebenarnya (*ground truth*). Dalam konteks penelitian ini, *confusion matrix* diterapkan pada level *per-field* data (NPM, Nama, Program Studi, Fakultas, dan TTL) untuk mengetahui seberapa akurat hasil ekstraksi OCR dan Regex dibandingkan data referensi yang telah diverifikasi secara manual.

METODE PENELITIAN

Metode Pelaksanaan

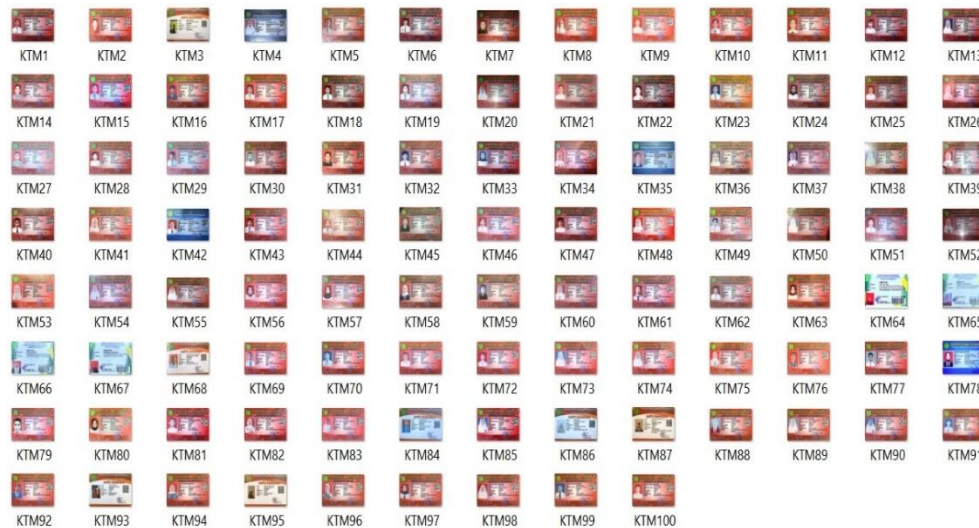
Metode pelaksanaan dalam penelitian ini menggunakan metode eksperimental yang berbasis pengembangan perangkat lunak (Hidayati *et al.*, 2023). Metode ini dipilih karena penelitian ini tidak hanya menganalisis suatu fenomena, tetapi juga merancang, membangun, menguji, dan memperbaiki sebuah sistem perangkat lunak secara iteratif berdasarkan hasil pengujian terhadap data nyata.



Gambar 1 Proses Tahap Pelaksanaan Eksperimen

Studi Literatur, mengkaji penelitian terdahulu terkait OCR, Regex, dan pengolahan citra digital. Studi literatur tersebut dari berbagai sumber jurnal penelitian lima tahun terakhir.

Pengumpulan Data, Penelitian ini menggunakan data penelitian dengan sumber data primer sebanyak 100 citra KTM mahasiswa Universitas Muhammadiyah Bengkulu yang diambil langsung menggunakan kamera handphone. Data citra KTM tersebut masing masing dari berbagai Fakultas yang ada pada UM Bengkulu seperti KTM mahasiswa Fakultas Teknik, Fakultas Manajemen, Fakultas FKIP, Fakultas Ilmu Sosial dan Politik. Data citra tersebut kemudian dilakukan proses cropping untuk mendapatkan gambar yang presisi sesuai ukuran KTM, dan selanjutnya dilakukan normalisasi dengan resolusi 400*400 secara manual.



Gambar 2 Data Penelitian 100 Citra KTM Mahasiswa

Perancangan Sistem, merancang arsitektur, alur data, dan antarmuka pengguna sebagaimana diuraikan pada subbab 3.3.

Implementasi, membangun *pipeline* prapemrosesan citra, integrasi OCR Tesseract, pola ekstraksi Regex, serta antarmuka GUI Tkinter.

Pengujian dan Evaluasi, menguji sistem terhadap 100 citra KTM dan mengevaluasi hasilnya menggunakan *confusion matrix* dibandingkan data ground truth.

Analisis dan Perbaikan, menganalisis kesalahan yang ditemukan pada tahap pengujian (sebagaimana diuraikan pada subbab 3.1), memperbaiki bagian sistem yang bersangkutan, kemudian mengulangi tahap pengujian hingga diperoleh hasil yang memuaskan.

HASIL DAN PEMBAHASAN

Hasil

Hasil pengujian menunjukkan sistem mencapai akurasi keseluruhan sebesar 93,8%, dengan *Precision* 94,86%, *Recall* 98,72%, dan *F1-Score* 96,75%. Akurasi per-*field* menunjukkan NPM mencapai 100%, Fakultas 98%, Program Studi 95%, TTL 94%, dan Nama 82%. Penelitian ini juga menemukan dan mengatasi permasalahan lintas platform pada pustaka OCR yang menyebabkan hilangnya spasi dan tanda baca pada hasil ekstraksi ketika dijalankan pada sistem operasi Windows. Hasil penelitian menunjukkan bahwa kombinasi OCR dan *Regular Expression* yang dirancang khusus sesuai karakteristik visual KTM mampu menghasilkan sistem ekstraksi data yang akurat dan layak digunakan untuk mendukung proses administrasi akademik.

Uji Coba

Uji coba antarmuka pengguna diimplementasikan menggunakan Tkinter dengan struktur *notebook* tiga tab. Modul *main_window.py* bertindak sebagai jendela utama yang memuat ketiga tab

(*tab_single.py*, *tab_batch.py*, dan *tab_confusion.py*), sedangkan modul *styles.py* menyediakan konfigurasi warna dan tipografi yang konsisten di seluruh antarmuka. Setiap tab berkomunikasi dengan lapisan logika melalui pemanggilan fungsi *process_image* pada modul *ocr_engine.py*, sehingga logika pemrosesan citra tidak perlu diduplikasi pada setiap tab

Uji Coba Citra KTM

Tab Pengujian Satu Citra merupakan antarmuka yang dirancang untuk menguji proses ekstraksi data pada satu berkas citra KTM secara individual. Pengguna memulai proses dengan menekan tombol "Pilih Gambar KTM" untuk memuat berkas citra yang akan diuji, kemudian menekan tombol "Proses OCR" untuk menjalankan seluruh rangkaian tahapan preprocessing hingga ekstraksi data. .

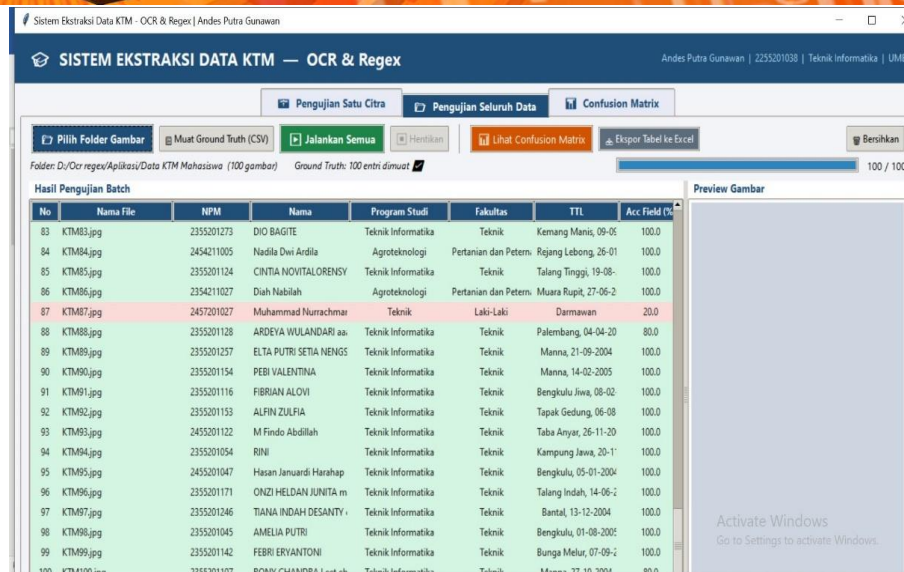


Gambar 3 Uji Coba Tunggal

Gambar di atas memperlihatkan tampilan tab Pengujian Satu Citra yang digunakan untuk menguji sistem ekstraksi data KTM berbasis OCR dan Regex terhadap satu sampel citra bernama KTM1.jpg. Pada bagian kiri jendela aplikasi ditampilkan empat tahap preprocessing citra secara berurutan, yaitu citra asli, citra grayscale, citra hasil Gaussian blur, dan citra hasil binarisasi Otsu, yang masing-masing merepresentasikan proses transformasi citra sebelum dilakukan pembacaan teks. Pada bagian kanan jendela, sistem menampilkan hasil teks mentah dari proses OCR beserta hasil akhir ekstraksi data menggunakan regex, seperti NPM, nama, dan program studi mahasiswa. Pengujian ini berhasil menunjukkan bahwa sistem mampu mengekstraksi data KTM dengan cukup cepat meskipun teks pada bagian header citra masih mengandung noise akibat proses OCR.

Proses Uji Keseluruhan

Tab Pengujian Seluruh Data berfungsi untuk melakukan pengujian secara batch (kelompok) terhadap banyak citra KTM sekaligus, berbeda dengan tab sebelumnya yang hanya menguji satu citra. Pada bagian atas antarmuka terdapat sederet tombol kontrol utama, yaitu "Pilih Folder Gambar" untuk memuat seluruh berkas citra KTM dari suatu direktori, "Muat Ground Truth (CSV)" untuk mengimpor data acuan (data sebenarnya) yang akan dibandingkan dengan hasil ekstraksi sistem, "Jalankan Semua" untuk memulai proses OCR dan ekstraksi terhadap seluruh citra dalam folder secara otomatis, serta "Hentikan" untuk menghentikan proses yang sedang berjalan. Di sisi kanan tersedia tombol "Lihat Confusion Matrix" untuk menampilkan evaluasi performa sistem, "Ekspor Tabel ke Excel" untuk menyimpan hasil pengujian ke dalam berkas Excel, dan "Bersihkan" untuk mengosongkan hasil pengujian yang sedang ditampilkan.



Gambar 4 Proses Uji Keseluruhan

Gambar di atas menunjukkan tampilan tab Pengujian Seluruh Data yang digunakan untuk menjalankan pengujian secara batch terhadap kumpulan citra KTM yang tersimpan dalam satu folder. Melalui tab ini, pengguna dapat memuat folder gambar beserta data ground truth dalam format CSV, kemudian menjalankan proses OCR dan ekstraksi regex secara otomatis untuk seluruh citra sekaligus. Hasil pengujian ditampilkan dalam bentuk tabel yang memuat nama file, NPM, nama, program studi, fakultas, tempat tanggal lahir, dan persentase akurasi setiap field, sementara panel di sisi kanan menampilkan pratinjau citra yang sedang dipilih. Fitur ini memudahkan proses evaluasi performa sistem secara menyeluruh, karena hasilnya dapat langsung dianalisis melalui confusion matrix maupun diekspor ke dalam berkas Excel untuk keperluan dokumentasi dan analisis lebih lanjut pada laporan skripsi.

Setelah seluruh 100 data citra KTM selesai diproses, sistem secara otomatis menampilkan ringkasan statistik pada panel log yang merangkum hasil evaluasi keseluruhan dataset. Ringkasan tersebut mencakup seluruh metrik evaluasi yang dihitung dari keseluruhan 100 citra sebagaimana ditampilkan pada Tabel berikut:

Tabel 1 Tabel Hasil Pengujian

Nama File	NPM	Nama	Program Studi	Fakultas	TTL	Acc Field (%)	Acc Character (%)	Waktu (s)
KTM1.jpg	2354231016	DEKY KAMSIRDIN	Peternakan	Pertanian dan Peternakan	Manna, 18-03-2004	100.0	100.0	1.47
KTM2.jpg	2255201121	GENTADEDIKUMARA isi	Teknik Informatika	Teknik	Taba Padang, 13-06-2004	80.0	93.7	1.45
KTM3.jpg	2454201006	Randi Kurniawan	Agribisnis	Pertanian dan Peternakan	Padang Tambak, 18-01-2005	100.0	98.2	1.27
KTM4.jpg	2188201036	SEPTI JUWITA TGL.	NURMA TENPAT	Keguruan dan Ilmu Pendidikan	Gunung Kembang, 30-16-2002	60.0	68.1	1.33

KTM5.jpg	225520107 1	M SHIDIQ ALFAJAR	Teknik Informatika	Teknik	Padang Jawi, 19-09-2003	100.0	100.0	1.33
KTM6.jpg	226120114 7	MUHAMMAD IHSAN IQBAL	Ekonomi net	Ekonomi dan Bisnis	Kota Agung, 11-11-2003	80.0	91.8	1.40
KTM7.jpg	226120103 0	MUHAMAD FITER GANDI	Manajemen	Ekonomi dan Bisnis	Padang Tambak, 13-10-2004	80.0	90.6	1.70
KTM8.jpg	225520115 2	MARYA OVI SESA	Teknik Informatika	Teknik	Gunung Tiga, 23-03-2004	100.0	100.0	1.32
KTM9.jpg	295520124 4	RIZKY RAHMAT SAPUTRA	Teknik Informatika	Teknik	Tanjung Betung, 23-04-2004	100.0	98.0	1.53
KTM10.jpg	225520115 4	USANDI MUSLIM	Teknik Informatika	Teknik	Karang Dapo, 10-03-2004	100.0	99.1	1.49
KTM11.jpg	225520106 1	ARIPUN	Teknik Informatika	Teknik	Kayu Kunyit, 09-03-2002	100.0	100.0	1.43
KTM12.jpg	235423100 2	MUHAMMAD FADLI AZK	Peternakan	Pertanian dan Peternakan	Lubuk Rumbai, 27-07-2005	80.0	83.6	1.58
KTM13.jpg	235423101 4	FITRI PERMATA SARI	Peternakan	Pertanian dan Peternakan	Bengkulu, 01-04-2006	100.0	98.0	1.60
KTM14.jpg	225520101 2	AGENG PAMORA	Teknik Informatika	Teknik	Ulak Agung, 23-03-2004	100.0	98.2	1.28
KTM15.jpg	225720101 3	MELZAN SABRI	Sistem Informasi	Teknik	Gedung Agung, 03-02-2004	100.0	99.1	1.42
KTM16.jpg	225520120 6	SANDI PERDANA PUTRA	Teknik Informatika	Teknik	Babatan Ulu, 15-04-2004	100.0	100.0	1.33
KTM17.jpg	235423103 1	HERU FERDIANYAH	Peternakan	Pertanian dan Peternakan	Gunung Tiga, 26-12-2005	100.0	98.8	1.50
KTM18.jpg	225520107 2	DIMAS ANDREAN SAPUTRA	Teknik Informatika	Teknik	Gunung Agung, 29-09-2003	100.0	100.0	1.50
KTM19.jpg	225520107 3	ARIAS JAKA PRATAMA	Teknik Informatika	Teknik	Talang Benuang, 23-09-2004	100.0	100.0	1.45
KTM20.jpg	225520115 1		Teknik Informatika	Teknik	Air Long, 26-02-2004	80.0	80.0	1.52
KTM21.jpg	225520104 6	ARDEKAPRANSISK A Pog J end	Teknik Informatika	Teknik	Lawang Agung, 03-11-2004	80.0	91.2	1.66
KTM22.jpg	225520111 3	DENDY PRANATA	Teknik Informatika	Teknik	Sungai Naik, 08-05-2004	100.0	98.5	1.45
KTM23.jpg	225520103 8	ANDES PUTRA GUNAWAN	Teknik Informatika	Teknik	Batu Kuning, 14-12-2003	100.0	99.1	1.33
KTM24.jpg	225520412 5	MELATI PUSPITA SARI capt	Teknik Informatika	Teknik	Manna, 15-04-2004	80.0	91.3	1.76

KTM25.jp g	225520122 5	MALFARISI i.	Teknik Informatika	Teknik	Manna, 26- 07-2003	80.0	94.5	1.30
KTM26.jp g	236220100 3	PUTRI REZA UTAMI	Akuntansi	Ekonomi dan Bisnis	Mentiring, 23-11-2005	100. 0	100. 0	1.25
KTM27.jp g	225520112 9	MUHAMMAD HAIKAL ZAQI	Teknik Informatika	Teknik	Suka Bumi, 11-07-2003	80.0	93.0	1.50
KTM28.jp g	227020100 7	PEKI JULIANTO	Ilmu Komunikasi	imu Sosial dan Politik	Tanjung Eran, 06-07- 2004	80.0	95.6	1.32
KTM29.jp g	225520115 3	AKBAR FEBRIANSYAH PUTRA	Teknik Informatika	Teknik	Gunung Tiga, 01-02- 2003	80.0	94.8	1.33
KTM30.jp g	225520122 4	MUHAMMAD REYPI	Teknik Informatika	Teknik	Manna, 06- 12-2002	80.0	90.0	1.50
KTM31.jp g	225520116 5	ANDIKAAKBAR	Teknik Informatika	Teknik	Tanjung Alam, 30-03- 2003	100. 0	98.3	1.65
KTM32.jp g	225520121 6	ADITHYO RAMADHAN	Teknik Informatika	Teknik	Serambi, 25- 10-2004	100. 0	100. 0	1.35
KTM33.jp g	235520108 4	ELVINA UTAMI	Teknik Informatika	Teknik	Suka Rami, 19-10-2004	100. 0	100. 0	1.54
KTM34.jp g	238420500 3	WINDA ROSALINDA	Pendidikan Biologi ake	Keguruan dan Ilmu Pendidik an	Tanjung Bulan, 15- 04-2005	80.0	95.5	1.59
KTM35.jp g	216120118 3	ADEFRI AL AHMAD NUR	Ekonomi Manajemen	Ekonomi dan Bisnis	Gedung Agung, 29- 11-2003	100. 0	99.1	1.50
KTM36.jp g	221320102 8	YOWANDA ENGGAL ARROZQA	Kesehatan Masyarakat	Ilmu Kesehata n	Bengkulu, 19-04-2004	100. 0	100. 0	1.60
KTM37.jp g	225520116 4					20.0	20.0	0.99
KTM38.jp g	221320103 1	DINDA METRIZA	Kesehatan Masyarakat	Ilmu Kesehata n	Manna, 17- 05-2004	100. 0	100. 0	1.47
KTM39.jp g	725520109 7	GENI APRIYANI	Teknik Informatika	Teknik	Anggut, 16- 04-2004	100. 0	98.0	1.36
KTM40.jp g	225520114 8	ZOILUDIN RAYHAN Ihe	Teknik Informatika	Teknik	Marga Sakti, 09-11-2003	80.0	95.8	1.50
KTM41.jp g	221320102 3	MELANI MAWADDAH	Kesehatan Masyarakat	Ilmu Kesehata n	Pasir Laweh, 23-03-2003	100. 0	100. 0	1.43
KTM42.jp g	215420102 6	NOPRI SAPUTRA.	ADE Agribisnis	Pertanian dan Petrnak an	Gunung Tiga, 17-11- 2002	100. 0	98.2	1.53
KTM43.jp g	238420501 3	REGEL RAHMATULLAH	Pendidikan Bioogi ae F	Keguruan dan Ilmu Pendidik an	Bengkulu, 03-03-2006	80.0	94.8	1.55
KTM44.jp g	228720503 9	KIKI APRIANTO	Pendidikan Pancasila dan	Keguruan dan Ilmu Pendidik	Tambangan, 10-11-2003	100. 0	97.0	1.57

			Kewargane an s						
KTM45.jp g	228820101 3	ANDRE WIHNATA	Pendidikan Bahasa dan Sastra	Keguruan dan Ilmu Pendidik an	Pasar Pino, 29-03-2004	80.0	94.7	1.49	
KTM46.jp g	225520120 3	NYKO CONDRO DINOTO M.N	Teknik Informatika	Teknik	Penogo, 23- 12-2004	100. 0	100. 0	1.45	
KTM47.jp g	225520124 2	OJLHERWANDAPU TRA	Teknik Informatika	Teknik	Palak Bengkerung, 23-03-2004	100. 0	96.7	1.55	
KTM48.jp g	225520107 7	MESY SUCI YATI bor	Teknik Informatika	Teknik	Pelangki, 23- 11-2004	80.0	95.6	3.70	
KTM49.jp g	225520105 6	SELTA JAYA PUTRA	Teknik Informatika	Teknik	Palembang, 03-01-2004	100. 0	100. 0	1.48	
KTM50.jp g	225720101 4	HELEN ANGGRAINI	Sistem Informasi	Teknik	Talang Lais, 01-03-2004	100. 0	97.8	1.34	
KTM51.jp g	225520116 9	DICKOVEANDREY	Teknik Informatika	Teknik	Masmamba ng, 06-08- 2080	100. 0	96.7	1.62	
KTM52.jp g	227020107 6	Ilmu Komunikasi	Ilmu Komunikasi	fimu Sosial dan Politik	Air Putih, 27-07-2004	60.0	77.3	1.60	
KTM53.jp g	235520106 1	AISYAH RAMADHANI	Teknik Informatika	Teknik	Fajar Baru, 10-30-2005	100. 0	98.1	1.43	
KTM54.jp g	272882040 51	Pendidikan Bahasa	Pendidikan Bahasa dan Sastra Indonesia	Keguruan dan Ilmu Pendidik an	Jeranglah Tinggi, 28- 01-2003	80.0	82.2	1.37	
KTM55.jp g	225520116 6	NABELLA HERMY NATHANIA	Teknik Informatika	Teknik	Rigangan l, 14-03-2004	100. 0	99.1	1.45	
KTM56.jp g	228820103 0	LARAS SHANTY ere	Pendidikan Bahasa dan Sastra inde	Keguruan dan Ilmu Pendidik an	Argamakmu r, 27-05- 2003	80.0	92.4	1.77	
KTM57.jp g	228820104 3	NANDA PUTRI SALSABILLAH	Pendidikan Bahasa dan Sastra Ind	Keguruan dan Ilmu Pendidik an	Bengkulu, 17-06-2004	100. 0	96.8	1.67	
KTM58.jp g	228820103 8	DELYA APRILINA	Pendidikan Bahasa dan Sastra indones	Keguruan dan Ilmu Pendidik an	Pariaman, 14-04-2004	100. 0	99.0	1.41	
KTM59.jp g	225520120 8	NAIM MUKARROM NAWWAR	Teknik Informatika	Teknik	Bengkulu, 26-02-2004	100. 0	99.0	1.60	
KTM60.jp g	235520114 1	PEPTI KUMALA DEWI	Teknik Informatika	Teknik	Padang Niur, 20-02- 2005	100. 0	97.1	1.65	
KTM61.jp g	225520108 6	FREDY PRAYOGA	Teknik Informatika	Teknik	Pelabuhan Talang Leak, 12-11-2004	100. 0	100. 0	1.57	
KTM62.jp g	228620801 2	DIGO PRAMANA eed	Pendidikan Agamalsla m (PAI	Agama Islam	Tanjung Aur, 31-12-2002	60.0	90.2	1.53	
KTM63.jp	225423101	RANI OKTAVIANTI	Peternakan	Pertanian	Bengkulu,	100.	97.0	1.48	

g	5	RAHMA PUTRI.			dan Petrnak an	08-10-2004	0		
KTM64.jp g	258820103 2	Laura Tri Ayu					66.7	66.7	1.17
KTM65.jp g	258820103 7	Chelsea Nurfadhilla					66.7	66.7	1.20
KTM66.jp g	258820105 6	MIE NIP.					33.3	38.1	1.39
KTM67.jp g	258820105 2						33.3	33.3	1.16
KTM68.jp g	245720101 9	Khosyifah Rahmadani Berliawa fale	Sistem Informasi	Teknik	Kota Bengkulu, 04-10-2005		80.0	92.7	1.45
KTM69.jp g	235520110 6	DHEZAN SAPUTRA	Teknik Informatika	Teknik	Bengkulu, 16-01-2006		100. 0	98.2	1.67
KTM70.jp g	235520111 4	AGUS TRI WAHYUDI	Teknik Informatika	Teknik	Bengkulu, 19-08-2004		100. 0	100. 0	1.39
KTM71.jp g	235520103 0	M. ARYA SAPUTRA	Teknik Informatika	Teknik	Bekasi, 01- 08-2005		100. 0	100. 0	1.55
KTM72.jp g	235520127 0	GUNAWAN SAPUTRA	Teknik Informatika	Teknik	Ketenong I, 29-03-2005		100. 0	99.0	1.49
KTM73.jp g	235520102 2	OKTI FELINA ANGREINI	Teknik Informatika	Teknik	Bengkulu, 07-10-2004		100. 0	99.0	1.56
KTM74.jp g	235520102 5	DIAKOMALA el tod	Teknik Informatika	Teknik	Tanjung Beringin, 17- 08-2005		80.0	90.0	1.56
KTM75.jp g	235520121 7	NADYA FEFBRIANTI	Teknik Informatika	Teknik	Batam, 10- 08-2004		100. 0	100. 0	1.68
KTM76.jp g	235520119 4	BENI ALPAJRI	Teknik Informatika	Teknik	Kampai, 08- 10-2005		100. 0	100. 0	1.48
KTM77.jp g	225720100 6	CALVIN LEORENZA	Sistem Informasi	Teknik	Talang Alai, 30-03-2003		100. 0	100. 0	1.72
KTM78.jp g	218420201 8	SEPTIA AGUSTINA	Pendidikan vegan sone	Keguruan dan Ilmu Pendidik an	Suka Makmur, 24- 09-2003		80.0	89.1	2.08
KTM79.jp g	225520110 5	REDHO PUTRA RAMADHAN	Teknik Informatika	Teknik	Bengkulu, 27-10-2003		100. 0	100. 0	1.77
KTM80.jp g	228720301 0	AZIZAH HERLIANA	Pendidikan Ekonomi	Keguruan dan Ilmu Pendidik an	Muara Enim, 23-10-2003		100. 0	100. 0	1.69
KTM81.jp g	235520124 3	MUHAMMAD DAFA SADIKIN	Teknik Informatika	Teknik	Tes, 29-09- 2004		100. 0	100. 0	1.55
KTM82.jp g	235520113 8	FEBBY IBTYAHDITHA LESTARI sare	Teknik Informatika	Teknik	Bengkulu, 07-05-2005		100. 0	96.7	1.65
KTM83.jp g	235520127 3	DIO BAGITE	Teknik Informatika	Teknik	Kemang Manis, 09- 09-2003		100. 0	100. 0	1.64
KTM84.jp g	245421100 5	Nadila Dwi Ardila	Agroteknol ogi	Pertanian dan Petrnak an	Rejang Lebong, 26- 01-2007		100. 0	100. 0	1.71

KTM85.jpg	2355201124	CINTIA NOVITALORENSYA	Teknik Informatika	Teknik	Talang Tinggi, 19-08-2005	100.0	99.1	1.55
KTM86.jpg	2354211027	Diah Nabilah	Agroteknologi	Pertanian dan Peternakan	Muara Rupit, 27-06-2005	100.0	98.3	1.55
KTM87.jpg	2457201027	Muhammad Nurrachman Hesu	Teknik	Laki-Laki	Darmawan	20.0	40.7	1.47
KTM88.jpg	2355201128	ARDEYA WULANDARI age	Teknik Informatika	Teknik	Palembang, 04-04-2005	80.0	93.3	1.45
KTM89.jpg	2355201257	ELTA PUTRI SETIA NENGSI	Teknik Informatika	Teknik	Manna, 21-09-2004	100.0	100.0	1.47
KTM90.jpg	2355201154	PEBI VALENTINA	Teknik Informatika	Teknik	Manna, 14-02-2005	100.0	100.0	1.52
KTM91.jpg	2355201116	FIBRIAN ALOVI	Teknik Informatika	Teknik	Bengkulu Jiwa, 08-02-2005	100.0	100.0	1.43
KTM92.jpg	2355201153	ALFIN ZULFIA	Teknik Informatika	Teknik	Tapak Gedung, 06-08-2005	100.0	98.3	1.47
KTM93.jpg	2455201122	M Findo Abdillah	Teknik Informatika	Teknik	Taba Anyar, 26-11-2006	100.0	97.8	1.56
KTM94.jpg	2355201054	RINI	Teknik Informatika	Teknik	Kampung Jawa, 20-11-2005	100.0	100.0	1.42
KTM95.jpg	2455201047	Hasan Januardi Harahap	Teknik Informatika	Teknik	Bengkulu, 05-01-2004	100.0	100.0	1.43
KTM96.jpg	2355201171	ONZI HELDAN JUNITA med	Teknik Informatika	Teknik	Talang Indah, 14-06-2004	100.0	96.4	1.48
KTM97.jpg	2355201246	TIANA INDIAH DESANTY cae	Teknik Informatika	Teknik	Bantal, 13-12-2004	100.0	96.5	1.56
KTM98.jpg	2355201045	AMELIA PUTRI	Teknik Informatika	Teknik	Bengkulu, 01-08-2005	100.0	100.0	1.39
KTM99.jpg	2355201142	FEBRI ERYANTONI	Teknik Informatika	Teknik	Bunga Melur, 07-09-2004	100.0	100.0	1.45
KTM100.jpg	2355201107	RONY CHANDRA Leet eben	Teknik Informatika	Teknik	Manna, 27-10-2004	80.0	90.9	1.41

Pengujian batch dilakukan terhadap 100 citra KTM mahasiswa menggunakan sistem ekstraksi data berbasis OCR dan Regular Expression. Hasil pengujian menunjukkan bahwa sistem memperoleh rata-rata akurasi per-field (Acc Field) sebesar 90,40% dan rata-rata akurasi per-karakter (Acc Char) sebesar 93,32%, dengan waktu pemrosesan rata-rata 1,51 detik per citra. Dari 100 sampel yang diuji, sebanyak 67 citra (67%) berhasil diekstraksi dengan akurasi sempurna (100%), sedangkan 24 citra memperoleh akurasi 80% akibat kesalahan pada satu field tertentu, umumnya disebabkan oleh noise OCR yang menambahkan karakter asing pada field Nama atau Program Studi (misalnya "GENTADEDIKUMARA isi" atau "Pendidikan Biologi ake").

Hanya terdapat 4 citra dengan akurasi di bawah 40%, yaitu KTM37.jpg, KTM66.jpg, KTM67.jpg, dan KTM87.jpg, yang seluruhnya mengalami kegagalan signifikan pada pembacaan field Nama maupun kesalahan pemetaan field akibat kualitas citra atau tata letak KTM yang berbeda dari format standar. Secara keseluruhan, hasil ini mengindikasikan bahwa sistem ekstraksi data KTM yang dibangun memiliki tingkat akurasi dan kecepatan pemrosesan yang baik untuk digunakan pada skenario pengujian dalam skala besar.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil perancangan, implementasi, dan pengujian sistem ekstraksi data otomatis dari citra Kartu Tanda Mahasiswa (KTM) menggunakan Optical Character Recognition (OCR) dan Regular Expression berbasis GUI Tkinter yang telah dilakukan, dapat ditarik beberapa kesimpulan sebagai berikut.

1. Sistem yang dibangun berhasil mengotomatiskan proses ekstraksi lima field data (NPM, Nama, Program Studi, Fakultas, dan TTL) dari citra KTM melalui *pipeline* yang terdiri atas prapemrosesan citra (*resize*, *deskew* dengan penjagaan, *crop* zona teks, dan binarisasi Otsu), pengenalan teks menggunakan Tesseract OCR dengan strategi multi-sumber dan multi-PSM disertai optimasi *early-exit*, serta ekstraksi field terstruktur menggunakan pola Regex bertingkat dengan mekanisme *fallback*.
2. Penyederhanaan *pipeline* prapemrosesan citra dengan menghilangkan CLAHE dan *sharpening kernel* yang pada rancangan awal justru merusak keterbacaan teks pada citra berlatar belakang bertekstur, serta menggantinya dengan binarisasi Otsu global — terbukti secara signifikan meningkatkan keterbacaan hasil OCR dibandingkan pendekatan prapemrosesan yang lebih agresif.
3. Hasil pengujian terhadap 100 citra KTM menunjukkan sistem mencapai akurasi Overall sebesar 93,8%, dengan rincian akurasi NPM 100%, Fakultas 98%, Program Studi 95%, TTL 94%, dan Nama 82%, yang membuktikan bahwa pendekatan OCR dan Regex mampu mengekstraksi data KTM secara akurat pada mayoritas kasus tanpa memerlukan model *machine learning* yang kompleks.

Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan, berikut beberapa saran untuk pengembangan lebih lanjut.

1. Field Nama masih memiliki akurasi terendah di antara kelima field yang diekstraksi (82%), sehingga penelitian selanjutnya dapat mengeksplorasi mekanisme pembersihan token yang lebih adaptif atau pemanfaatan basis data nama untuk memvalidasi hasil ekstraksi, tanpa mengorbankan risiko memotong nama pendek yang sah.
2. Tahap prapemrosesan citra pada penelitian ini menggunakan binarisasi Otsu global yang terbukti efektif meningkatkan keterbacaan OCR dibandingkan pendekatan CLAHE dan *sharpening* yang lebih agresif, namun efektivitas ini baru diuji pada kondisi pencahayaan dan latar belakang templat KTM UMB. Penelitian selanjutnya disarankan menguji efektivitas kombinasi teknik prapemrosesan lain, misalnya adaptive thresholding atau image enhancement berbasis deep learning, khususnya untuk menangani citra dengan pencahayaan tidak merata, bayangan, atau kualitas kamera yang lebih rendah.
3. Sistem saat ini bergantung sepenuhnya pada mesin OCR Tesseract berbasis aturan dan pola. Penelitian lanjutan dapat mengeksplorasi penggunaan model OCR berbasis *deep learning* atau model *Vision-Language* untuk menangani citra dengan kualitas rendah atau kemiringan ekstrem yang sulit ditangani oleh pendekatan berbasis Regex.
4. Waktu pemrosesan batch, meskipun telah dipercepat melalui strategi *early-exit*, masih dapat dioptimalkan lebih lanjut misalnya melalui pemrosesan paralel (*multiprocessing*) pada pengujian data dalam jumlah besar

DAFTAR PUSTAKA

- Adawiyah Ritonga and Yahfizham Yahfizham (2023) "Studi Literatur Perbandingan Bahasa Pemrograman C++ dan Bahasa Pemrograman Python pada Algoritma Pemrograman," *Jurnal Teknik Informatika dan Teknologi Informasi*, 3(3), pp. 56–63. Available at: <https://doi.org/10.55606/jutiti.v3i3.2863>
- Bachiar Rizki, H., Pratomo, S. and Wiwit Agus, T. (2026) "Sistemasi: Jurnal Sistem Informasi Implementasi Image Processing dalam Pemindaian Data KTP menggunakan Optical Character Recognition (OCR) Implementation of Image Processing in Scanning KTP Data using Optical Character Recognition (OCR)," *Sistemasi: Jurnal Sistem Informasi*, 15(1), pp. 334–343. Available at: <https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/5856/1140>.
- Dwi, M. *et al.* (2026) "Penerapan OCR (Optical Character Recognition) pada Transformasi Digital," *Jurnal Sains Dan Teknologi (JSIT)*, 6(1), pp. 197–201. Available at: <https://doi.org/https://doi.org/10.47233/jsit.v6i1.4375>.
- Al Fajar, M., Darmayunata, Y. and Hardianto, R. (2025) "Penerapan Algoritma Optical Character Recognition (OCR) Untuk Pelaporan Surat Pengantar Pengiriman BBM Di PT. Mandiri Melangkah Sejahtera," *Prosiding-Seminar Nasional Teknologi Informasi & Ilmu Komputer (SEMASTER)*, 4(1), pp. 312–321.
- Firin, A.S. and Sriani, S. (2025) "Analisis Algoritma Canny Edge Detection dengan Tesseract OCR untuk Mendeteksi Pelat Nomor Kendaraan Bermotor," *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, 13(3), pp. 381–393. Available at: <https://doi.org/10.37905/euler.v13i3.34868>.
- Hidayati, N. *et al.* (2023) "Ekstraksi Fitur dengan Color Histogram dan Classifier Random Forest pada Citra Kupu-Kupu diusulkan mengungguli pendekatan segmentasi gambar berbasis deep learning yang citra kupu-kupu dengan menggunakan teknik CNN dan mengevaluasi performansi," *Jami: Jurnal Ahli Muda Indonesia*, 4(2), pp. 148–157. Available at: <https://doi.org/https://doi.org/10.46510/jami.v4i2.172>.
- Inal Achyar, M., Yasir, A. and Satria, W. (2023) "Pengolahan Citra Berwarna Dalam Mendeteksi Asap Menggunakan Metode Segmentasi Foreground Dan Color Models," *SISKOMTI: Jurnal Sistem Informasi Komputer dan Teknologi Informasi*, 5(2), pp. 1–10. Available at: <https://doi.org/10.54342/fmy91476>.
- Julian Mulyadi and Nuri David Maria Veronika (2025) "Penerapan Optical Character Recognition (OCR) Dalam Pengambilan Teks Judul Buku Untuk Digitalisasi Arsip Perpustakaan," *JSai (Journal Scientific and Applied Informatics)*, 8(2), pp. 447–454. Available at: <https://doi.org/10.36085/jsai.v8i2.8457>.
- Kurnia, D. (2024) "Perancangan aplikasi berbasis GUI dalam pengolahan citra gambar kulit buah kelapa sawit dengan metode konvolusi," *Jurnal Vokasi Teknik*, 5(1), pp. 12–19. Available at: <https://mentech.id/jurnal/index.php/juvotek/article/view/44>.
- Kurnia Meiliyani *et al.* (2025) "Penggunaan Ekspresi Regular dalam Meningkatkan Efisiensi Pencarian Produk pada Aplikasi E-Commerce," *Neptunus: Jurnal Ilmu Komputer Dan Teknologi Informasi*, 3(1), pp. 169–177. Available at: <https://doi.org/10.61132/neptunus.v3i1.698>.

- Kusnantoro, Kusumaningrum, Sulistya, D. and Rohana, T. (2022) "Implementasi Metode Tesseract OCR(Optical Character Recognition) untuk Deteksi Plat Nomor Kendaraan Pada Sistem Parkir," *Scientific Student Journal for Information, Technology and Science*, III, pp. 59-67.
- Latifa Khoirani, Rino Ariansyah and Supiyandi Supiyandi (2024) "Implementasi Algoritma OCR (Optical Character Recognition) untuk Ekstraksi Informasi dari Citra Struk Transaksi Keuangan," *Mars: Jurnal Teknik Mesin, Industri, Elektro Dan Ilmu Komputer*, 2(6), pp. 151-160. Available at: <https://doi.org/10.61132/mars.v2i6.537>.
- Nanda Aprillia, M. et al. (2025) "Implementasi Teknologi QR Code pada Sistem Absensi Karyawan Berbasis Website," *Jurnal Sistem Informasi Galuh*, 3(1), pp. 39-50. Available at: <https://doi.org/10.25157/jsig.v3i1.4445>.
- Nashiruddin, M. et al. (2025) "Implementasi OCR berbasis Tesseract untuk Ekstraksi Data Kartu Mahasiswa UMKLA," *Jurnal Keilmuan Teknologi Informasi (JKTI)*, 01(1), pp. 11-14.
- Nasution, M.U. et al. (2025) "TINJAUAN METODE PENGOLAHAN CITRA DIGITAL," *Jurnal Informatika dan Teknik Elektro Terapan*, 13(3), pp. 1020-1028.
- Novendra Sulu, V. et al. (2024) "Implementasi Komunikasi Peer-to-Peer dengan Mengembangkan Aplikasi Menggambar Kolaboratif Implementation of Peer-to-Peer Communication by Developing Collaborative Drawing Application," *Jurnal Ilmiah Informatika dan Komputer*, 1(2), pp. 247-254.
- Octaviani, T., Setiawan, H. and Kelana, O.H. (2023) "Pytesseract dan Template Matching Untuk Otomatisasi Input Data KTP 147 dan Teknologi," *Jurnal Buana Informatika*, 14(2), pp. 147-156. Available at: <https://doi.org/https://doi.org/10.24002/jbi.v14i02.7612>.
- Pramudiya, R. et al. (2024) "Analisi Gambar Menggunakan Metode Grayscale dan HSV," *Jurnal Sistem Informasi, Teknologi Informatika dan Komputer*, 14(3), pp. 174-180.
- Qonita, V.A. et al. (2026) "PERBANDINGAN GAUSSIAN BLUR , MEDIAN , DAN BILATERAL FILTER UNTUK REDUKSI NOISE CITRA DIGITAL," *Jurnal PROSISKO*, 13(1).
- Rahmadani, O., Rozikin, C. and Karawang, U.S. (2021) "Kinerja Tesseract-Ocr Dalam Pengenalan Teks Tulisan Tangan Menggunakan Dataset Kustom," *Jurnal Informatika dan Teknik Elektro Terapan*, 13(3), pp. 1097-1104.
- Reswan, Y. et al. (2024) "Penerapan Algoritma Ocr Untuk Ekstraksi Informasi Dari Citra Kartu Tanda Mahasiswa (Ktm)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(5), pp. 11004-11009. Available at: <https://doi.org/10.36040/jati.v8i5.11006>.
- Ristanto, A.F. et al. (2023) "Pemantauan Posisi Multi-Node Menggunakan Aplikasi GUI pada LoRaWAN," *Jurnal Teknologi Elektro*, 14(2), p. 67. Available at: <https://doi.org/10.22441/jte.2023.v14i2.002>.
- Sebatubun, M. and Haryawan, C. (2024) "Implementasi Metode Otsu dan Momen Hu pada Citra Keris," *Jurnal Informatika Polinema*, 10(2), pp. 285-290. Available at: <https://doi.org/10.33795/jip.v10i2.4978>.

- Setiawan, A.A., Guntara, R.G. and Purwaamijaya, B.M. (2025) "Ekstraksi Informasi Struk Belanja Melalui Pemanfaatan Tesseract dan Regular Expressions," *RIGGS: Journal of Artificial Intelligence and Digital Business*, 4(2), pp. 6586–6594. Available at: <https://doi.org/10.31004/riggs.v4i2.1731>.
- Yusuf, A. et al. (2025) "Ekstraksi Teks Metadata Artikel dengan Metode Regex," *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, 8(5), pp. 2420–2434. Available at: <https://doi.org/10.32672/jnkti.v8i5.9569>.