



Pemeriksaan KTP Menggunakan Optical Character Recognition (OCR) dan Pengenalan Background serta Komponen KTP

INFO PENULIS	INFO ARTIKEL
Khairun Nisha Universitas Muhammadiyah Makassar khairunnisha1522@gmail.com	ISSN: 3026-3603 Vol. 2, No. 2 Oktober 2024 http://jurnal.ardenjaya.com/index.php/ajst
Titin Wahyuni Universitas Muhammadiyah Makassar	
Muhyiddin A M Hayat Universitas Muhammadiyah Makassar	

© 2024 Arden Jaya Publisher All rights reserved

Saran Penulisan Referensi:

Nisha, K., Wahyuni, T., & Hayat, M. A M. (2024). Pemeriksaan KTP Menggunakan Optical Character Recognition (OCR) dan Pengenalan Background serta Komponen KTP. *Arus Jurnal Sains dan Teknologi*, 2 (2), 490-495.

Abstrak

Penelitian ini bertujuan untuk mengembangkan sistem verifikasi identitas berbasis Kartu Tanda Penduduk (KTP) yang lebih aman dan akurat dengan memanfaatkan teknologi Optical Character Recognition (OCR), khususnya Tesseract OCR. Sistem ini dikembangkan untuk menangani peningkatan kebutuhan akan perlindungan data pribadi di era digital, di mana risiko penyalahgunaan data KTP semakin meningkat. Penelitian ini dilakukan di Laboratorium Informatika Universitas Muhammadiyah Makassar selama bulan Juli hingga Agustus 2024, dengan menerapkan teknik pra-pemrosesan gambar seperti grayscale, Gaussian blur, dan thresholding untuk memperbaiki hasil ekstraksi teks. Hasil pengujian menunjukkan bahwa sistem ini mampu mengekstraksi elemen-elemen penting pada KTP, seperti NIK, nama, dan jenis kelamin, dengan tingkat akurasi yang tinggi. Dari 10 KTP yang diuji, 9 di antaranya berhasil diekstraksi dengan benar, memberikan akurasi keseluruhan sebesar 90%. Namun, beberapa elemen seperti alamat dan RT/RW masih menimbulkan tantangan dalam proses ekstraksi, terutama pada gambar dengan kualitas rendah atau format penulisan yang tidak konsisten. Kesimpulan dari penelitian ini menegaskan bahwa teknologi OCR efektif dalam mendeteksi dan menganalisis informasi utama pada KTP, namun memiliki keterbatasan saat berhadapan dengan variasi kualitas gambar. Oleh karena itu, disarankan untuk melakukan pengujian tambahan di lingkungan operasional nyata untuk mengidentifikasi dan mengatasi tantangan yang mungkin muncul dalam penerapan aplikasi ini di dunia nyata.

Kata kunci : Optical Character Recognition (OCR), Tesseract OCR, KTP, Pra-pemrosesan Gambar.

Abstract

This study aims to develop a safer and more accurate identity verification system based on Resident Identity Cards (KTP) by utilizing Optical Character Recognition (OCR) technology, especially Tesseract OCR. This system was developed to address the increasing need for personal data protection in the digital era, where the risk of misuse of KTP data is increasing. This research was conducted at the Informatics Laboratory of the University of Muhammadiyah Makassar from July to August 2024, by applying image pre-processing techniques such as grayscale, Gaussian blur, and thresholding to improve text extraction results. The test results show that this system is able to extract important elements on KTPs, such as NIK, name, and gender, with a high level of accuracy. Of the 10 KTPs tested, 9 of them were successfully extracted correctly, giving an overall accuracy of 90%. However, some elements such as addresses and RT/RW still pose challenges in the extraction process, especially in images with low quality or inconsistent writing formats. The conclusion of this study confirms that OCR technology is effective in detecting and analyzing key information on KTP, but has limitations when dealing with variations in image quality. Therefore, it is recommended to conduct additional testing in a real operational environment to identify and overcome challenges that may arise in the implementation of this application in the real world.

Keywords: Optical Character Recognition (OCR), Tesseract OCR, KTP, Image Pre-processing.

A. Pendahuluan

Teknologi selalu memengaruhi cara orang berpikir dan menjalani gaya hidup mereka, kemajuan dalam ilmu pengetahuan dan teknologi telah membawa masyarakat ke arah kehidupan modern. Teknologi pada dasarnya diciptakan untuk meningkatkan kualitas hidup manusia dan mempermudah aktivitas manusia menjadi lebih efisien dan efektif. Namun, tidak dapat dipungkiri bahwa teknologi memiliki sisi positif dan negatif. (Badruzaman, 2019).

Oleh karena itu Kemajuan teknologi komunikasi dan informasi telah mempersulit pengamanan privasi seseorang dan mempermudah kebocoran informasi pribadi. (Sulistianingsih et al., 2023), Kejahatan pemanfaatan informasi pribadi orang lain tanpa izin semakin marak seiring berjalannya waktu. Perilaku ini menunjukkan bahwa seseorang melakukan penipuan demi keuntungan dirinya sendiri dengan mengorbankan orang lain. (Yuliany Siahaan, 2023)

Kartu identitas elektronik adalah selebaran kertas sensitif yang dapat disalahgunakan. Terkait penyelagunaan KTP elektronik, banyak sekali masyarakat yang tertipu, baik dengan membuat sendiri maupun memanfaatkan data orang lain. Banyak kejadian yang melibatkan penyalahgunaan data pribadi terjadi, antara lain jual beli data pribadi, pengungkapan informasi rekening nasabah bank, meluasnya penggunaan data pribadi untuk tujuan penipuan, dan peretasan rekening pribadi melalui jaringan online. (Permadi & Rokhman, 2023)

OCR (Optical Character Recognition) digunakan untuk menerjemahkan karakter dari suatu gambar ke dalam bentuk karakter aslinya, atau teks. Teks ini biasanya berbentuk Word, Notepad, atau program sejenis lainnya. Selain itu, OCR dapat dimanfaatkan untuk mengidentifikasi karakter, huruf, dan bahkan simbol pada gambar yang dikonversi. Ada beberapa langkah atau prosedur OCR yang terlibat di sini, termasuk pra-pemrosesan, segmentasi, normalisasi, dan pengenalan teks di akhir. (Kebutuhan et al., 2020)

Sedangkan penelitian ini memfokuskan dalam OCR yang membandingkan gambar masukan dengan salah satu gambar preset untuk melakukan pengenalan otomatis. Tesseract OCR dapat mengidentifikasi wajah serta sidik jari, foto, dan tanda tangan. (Kusnantoro et al., 2022). Adapun penelitian sebelumnya oleh Hendradito Dwi Aprillian & Hindriyanto Dwi Purnomo, S.T., MIT, Ph.D yaitu memastikan OPTICAL CHARACTER RECOGNITION (OCR) dalam memeriksa kecepatan dan keakuratan Kartu Tanda Penduduk (KTP), untuk mendaftarkan kartu kredit secara lengkap pada sistem registrasi, tergantung kualitas foto (warna, hitam putih, kontras, dan ukuran piksel) dan efisiensi waktu. (Purnomo, 2019)

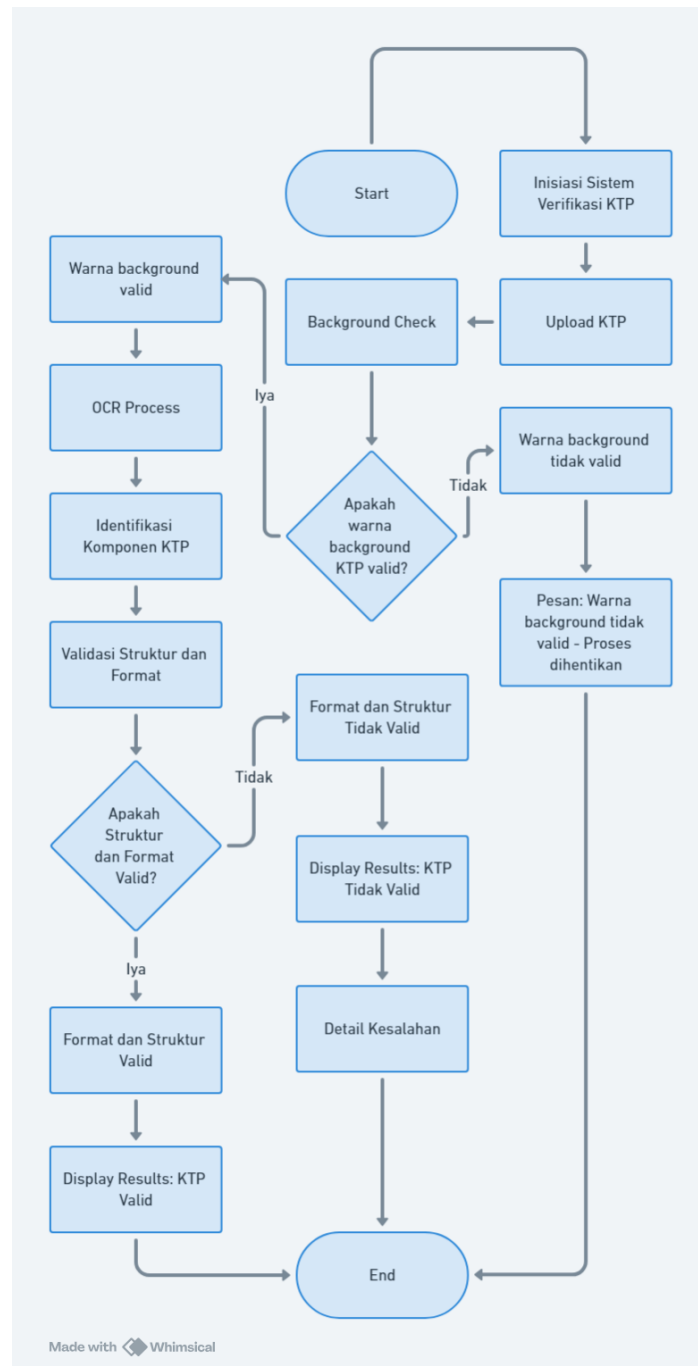
Dengan demikian, penelitian ini diharapkan dapat memberikan solusi yang efektif dan efisien dalam meningkatkan keamanan dan akurasi verifikasi identitas menggunakan teknologi OCR, serta mengurangi risiko penyalahgunaan data pribadi di era digital. Penelitian ini juga akan

membahas bagaimana teknologi OCR dapat diintegrasikan dalam sistem verifikasi KTP untuk meningkatkan perlindungan data pribadi dan mencegah penipuan. Selain itu, penelitian ini akan mengeksplorasi pendeteksian Background dan komponen-komponen spesifik pada KTP, seperti logo, tanda tangan, elemen mikroprinting, nama, jenis kelamin, dan elemen identitas lainnya yang akan diuji secara menyeluruh. Hal ini diharapkan dapat menambah lapisan keamanan dalam proses verifikasi, sehingga lebih efektif dalam mencegah penyalahgunaan KTP.

B. Metodologi

a. Perancangan Sistem

Berikut merupakan perancangan sistem.



Gambar 1. Perancangan Sistem

Flowchart Sistem Verifikasi KTP Menggunakan OCR

1. Start

Proses dimulai dengan inisiasi sistem verifikasi KTP.

2. Upload KTP

Pengguna mengunggah gambar KTP yang akan diperiksa. Gambar KTP dapat diunggah dari perangkat lokal atau URL.

3. *Background Check*

- a. *Warna Background* : Sistem pertama-tama memeriksa apakah warna *Background* KTP sesuai dengan warna yang telah ditentukan sebagai standar (misalnya, biru atau merah).
- b. Validasi: Jika warna *Background* sesuai, proses dilanjutkan. Jika tidak, sistem memberikan pesan bahwa warna *Background* KTP tidak valid dan proses dihentikan.

4. OCR Process

Setelah *Background* KTP divalidasi, gambar KTP diproses menggunakan OCR untuk membaca dan mengidentifikasi komponen utama KTP.

- a. OCR: Digunakan untuk membaca teks yang terdapat pada KTP, seperti NIK, nama, tempat/tanggal lahir, jenis kelamin, alamat, agama, status perkawinan, pekerjaan, dan kewarganegaraan.
- b. Validasi Struktur: Sistem memverifikasi bahwa format dan struktur elemen-elemen tersebut sesuai dengan standar yang telah ditetapkan.

5. *Display Results*

Sistem menampilkan hasil pemeriksaan kepada pengguna.

- a. Valid: Jika elemen-elemen KTP terdeteksi dengan format yang sesuai, sistem akan memutuskan bahwa KTP tersebut valid.
- b. Tidak Valid: Jika elemen-elemen KTP tidak sesuai dengan format yang diharapkan, sistem akan memutuskan bahwa KTP tersebut tidak valid.

Hasil dari proses pemeriksaan KTP akan ditampilkan kepada pengguna, yang dapat berupa:

- a. Pesan yang menyatakan bahwa KTP tersebut valid.
- b. Pesan yang menyatakan bahwa KTP tersebut tidak valid.
- c. Informasi detail tentang kesalahan format atau struktur pada KTP yang dideteksi.

6. *End*

Proses pemeriksaan selesai dan sistem kembali ke status siap untuk menerima input berikutnya.

b. Teknik Pengujian Sistem

Pengujian sistem ini sekarang mencakup tahapan tambahan untuk memvalidasi warna *Background* KTP sebelum proses OCR dilakukan. Berikut adalah teknik pengujian yang diperbarui:

1. Persiapan Dataset

- a. Kumpulkan Dataset: Mengumpulkan dataset yang terdiri dari gambar KTP asli dengan variasi warna *Background* yang sesuai dan tidak sesuai.
- b. *Labeling*: Memberi label pada setiap gambar dalam dataset untuk menentukan apakah gambar tersebut memiliki warna *Background* yang valid atau tidak.

2. *Background Check*

- a. *Warna Background* : Memeriksa apakah warna *Background* KTP sesuai dengan standar yang telah ditentukan.
- b. Validasi *Background* : Jika warna *Background* sesuai, gambar diproses lebih lanjut. Jika tidak sesuai, gambar ditandai sebagai tidak valid dan tidak diproses lebih lanjut.

3. Pemrosesan Gambar

- a. *Resizing*: Mengubah ukuran gambar agar sesuai dengan input yang diterima oleh *Tesseract OCR*.
- b. *Grayscale Conversion*: Mengonversi gambar ke skala abu-abu untuk meningkatkan akurasi OCR, setelah validasi *Background* .
- c. *Noise Reduction*: Menerapkan teknik reduksi *Noise* untuk meningkatkan kejelasan gambar.

C. Hasil dan Pembahasan

A. Hasil Mengunggah Gambar KTP



Gambar 2. Hasil OCR

Proses hasil validasi dari proses pengunggahan dan pengecekan KTP. Dalam tampilan ini, aplikasi menunjukkan status validasi dengan status "VALID", menandakan bahwa KTP yang diunggah telah lolos pemeriksaan dan dianggap sah oleh sistem. Detail hasil validasi juga ditampilkan, yang berisi teks yang berhasil diekstraksi dari gambar KTP menggunakan OCR (Optical Character Recognition). Meskipun demikian, terlihat bahwa beberapa bagian dari teks hasil ekstraksi mungkin mengalami kesalahan atau distorsi, yang merupakan tantangan umum dalam proses OCR, terutama jika gambar tidak memiliki kualitas yang optimal.

Tombol "Unggah gambar lain" juga disediakan untuk memungkinkan pengguna mengulangi proses dengan gambar KTP yang berbeda, jika diperlukan. Desain ini memberikan kemudahan bagi pengguna untuk langsung melanjutkan ke tahap berikutnya setelah proses validasi, atau mengulanginya jika mereka merasa perlu. Hal ini mencerminkan siklus penggunaan yang intuitif dan efisien dalam aplikasi verifikasi KTP.

D. Kesimpulan

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan pada aplikasi validasi KTP menggunakan teknologi OCR (*Optical Character Recognition*), beberapa kesimpulan utama dapat diambil:

1. Keberhasilan Identifikasi KTP

Aplikasi ini berhasil mengidentifikasi elemen-elemen kunci seperti NIK, nama, dan jenis kelamin dengan tingkat akurasi yang cukup tinggi. Dari 10 KTP yang diuji, 9 di antaranya berhasil diekstraksi dengan benar, menunjukkan kemampuan OCR dalam mendeteksi dan menganalisis informasi utama pada KTP. Namun, tantangan masih ada dalam hal mengidentifikasi komponen pada elemen seperti alamat dan RT/RW, terutama pada gambar dengan kualitas rendah atau format penulisan yang tidak

konsisten. Ini menegaskan bahwa teknologi OCR efektif dalam kondisi ideal, namun memiliki keterbatasan ketika berhadapan dengan variasi format dan kualitas gambar.

2. Tingkat Akurasi Identifikasi KTP dengan OCR

Tingkat akurasi dalam proses identifikasi dan verifikasi elemen-elemen utama pada KTP menunjukkan hasil yang positif. Teknik pra-pemrosesan gambar seperti *thresholding* dan *blurring* secara signifikan meningkatkan akurasi OCR, khususnya dalam menangani gambar dengan pencahayaan yang rendah. Namun, masih terdapat ruang untuk peningkatan, terutama dalam menghadapi gambar dengan resolusi rendah atau kualitas buruk. Dari 10 KTP yang diuji, tingkat keberhasilan identifikasi dan validasi mencapai 90%, menandakan bahwa OCR dapat menjadi alat yang efektif, namun tetap perlu dioptimalkan lebih lanjut untuk menghadapi berbagai kondisi.

Saran

Penelitian ini hanya menggunakan dataset gambar KTP yang telah disiapkan dengan variasi kualitas gambar, disarankan untuk melakukan pengujian tambahan di lingkungan operasional nyata. Hal ini akan membantu mengidentifikasi tantangan yang mungkin muncul dalam penggunaan aplikasi di dunia nyata, seperti variasi kualitas kamera, pencahayaan, dan kondisi fisik KTP.

E. Referensi

- Alfarizi, M. R. S., Nugraha, M. A., Putra, R., Basri, A., Jannuartha, R. A. R., Pratama, M. R., Syagara, G. W., Pratama, A. P., & Kalam, M. J. (2023). Menggali Bahasa Pemrograman Populer: Karakteristik Utama dan Penggunaan yang Luas. *Karimah Tauhid*, 2(4), 1191-1197. <https://ojs.unida.ac.id/karimahtauhid/article/view/8816>
- Badruzaman, D. (2019). Kajian Hukum Tentang Internet Mobile dalam Upaya Pencegahan Dampak Negatif Teknologi Informasi dan Komunikasi di Indonesia. *Ajudikasi : Jurnal Ilmu Hukum*, 3(2), 135. <https://doi.org/10.30656/ajudikasi.v3i2.1657>
- Kebutuhan, D., Untuk, N., Pupuk, J., Tanaman, P., Daun, B. W., Support, M., Machine, V., & Darajat, G. F. (2020). Program studi teknik informatika fakultas teknik dan ilmu komputer universitas komputer indonesia 2020. *Researchgate.Net*, 10115277. file:///C:/Users/andik/Downloads/UNIKOM_GUMILAR FAJAR DARAJAT_JURNAL DALAM BAHASA INGGRIS.pdf
- Kusnantoro, Rohana, T., & Kusumaningrum, D. S. (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, 59-67.
- Permadi, I. B., & Rokhman, A. (2023). Implementasi Identitas Kependudukan Digital Dalam Upaya Pengamanan Data pribadi. *JOPPAS: Journal of Public Policy and Administration Silampari*, 4(2), 80-88. <https://doi.org/10.31539/joppas.v4i2.6199>
- Purnomo, H. D. (2019). *Pemanfaatan Teknologi OPTICAL CHARACTER RECOGNITION pada Pembacaan Kartu Tanda Penduduk Artikel Ilmiah Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Satya Wacana Salatiga Februari 2019 Pemanfaatan Teknologi Optical Char. 672015123.*
- Sulistianingsih, D., Ihwan, M., Setiawan, A., & Prabowo, M. S. (2023). Tata Kelola Perlindungan Data Pribadi Di Era Metaverse (Telaah Yuridis Undang-Undang Perlindungan Data Pribadi). *Masalah-Masalah Hukum*, 52(1), 97-106. <https://doi.org/10.14710/mmh.52.1.2023.97-106>
- Sulistiyo, A., & Nugraha Saian, P. O. (2019). *Perancangan Sistem Text Extraction menggunakan library Tesseract OCR untuk mengambil Nomor Induk Kependudukan pada foto Kartu Tanda Penduduk (Studi Kasus : PT. Bank ABC, Tbk). April.*
- Teguh Ainul Darajat, Agus Pratondo, & Fery Prasetyanto. (2023). Penerapan Fitur Face Recognition dan Ekstraksi Data dari Citra KTP pada Sistem Informasi Penerimaan Tamu di PT. Cilegon Fabricators. *Jurnal Sains Dan Teknologi*, 9(3), 1410-1413. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=QK3N1kcAAAAJ&citation_for_view=QK3N1kcAAAAJ:W70EmFMy1HYC
- Yuliany Siahaan, A. (2023). Sanksi Pidana Terhadap Pelaku Perbuatan Penyalahgunaan KTP Orang Lain Untuk Pinjaman Online. *Nanggroe : Jurnal Pengabdian Cendikia*, 84(4), 2986-7002. <https://doi.org/10.5281/zenodo.8133281>