IMPLEMENTASI ARSITEKTUR HIBRIDA YOLOv8 DAN VISION API UNTUK EKSTRAKSI DATA KEPENDUDUKAN

Yunia Ikawati[1)], Agus Wibowo[2)], dan Moch Rizaldhi Pratama[3)]
[1]Departemen Teknik Informatika, Politeknik Elektronika Negeri Surabaya
[2]Departemen Multimedia Broadcasting, Politeknik Elektronika Negeri Surabaya
[3]Departemen Teknik Informatika, Politeknik Elektronika Negeri Surabaya
E-mail: yunia@pens.ac.id

Abstract

Enhancing the efficiency of public services at the village level is a primary priority in the digital transformation era. However, population data management and administrative services in Rejosari Village are still conducted conventionally, leading to risks of input errors and slow service delivery. As a solution, a web-based population management information system was developed, integrating a hybrid Optical Character Recognition (OCR) pipeline. This system utilizes the YOLOv8 model to localize specific demographic text boundaries on Family Card (KK) images, followed by extraction via the Gemini Vision API. Additionally, the system provides a digital portal for residents to facilitate letter requests and monitor application status in real-time. Implementation results demonstrate that the hybrid OCR pipeline achieved a 96.9% average extraction score on successfully processed documents (77.8% successful processing rate across 18 test samples), significantly accelerating data entry and minimizing manual typographical errors. This development successfully optimizes public service quality and administrative data integrity at the village level.

Keywords: *Population Management, Hybrid OCR, YOLOv8, Generative AI, Village Administration.*

Abstrak

Peningkatan efisiensi pelayanan publik di tingkat desa merupakan salah satu agenda strategis dalam mendukung transformasi digital sektor pemerintahan. Namun, proses pengelolaan data kependudukan dan pelayanan administrasi di Desa Rejosari masih dilaksanakan secara konvensional, yang berpotensi menimbulkan kesalahan input data serta menghambat kecepatan pelayanan kepada masyarakat. Untuk mengatasi permasalahan tersebut, dikembangkan sebuah Sistem Informasi Manajemen Kependudukan berbasis web yang terintegrasi dengan hybrid Optical Character Recognition (OCR). Sistem ini memanfaatkan model YOLOv8 untuk mendeteksi dan melokalisasi area teks demografis pada citra Kartu Keluarga (KK), kemudian melakukan ekstraksi informasi menggunakan Gemini Vision API. Selain mendukung digitalisasi pengelolaan data kependudukan, sistem juga menyediakan portal layanan bagi masyarakat untuk mengajukan permohonan surat secara daring serta memantau status pengajuan secara real-time. Hasil implementasi menunjukkan bahwa hybrid OCR pipeline yang dikembangkan mampu mencapai tingkat akurasi ekstraksi rata-rata sebesar 96,9% pada dokumen yang berhasil diproses, dengan tingkat keberhasilan pemrosesan sebesar 77,8% dari 18 sampel pengujian. Temuan ini menunjukkan bahwa sistem yang dibangun mampu meningkatkan efisiensi proses entri data, mengurangi kesalahan pengetikan secara signifikan, serta mendukung peningkatan kualitas pelayanan publik dan integritas data administrasi di tingkat desa.

Kata Kunci: *Manajemen Kependudukan, OCR Hibrida, YOLOv8, Generative AI, Administrasi Desa.*

PENDAHULUAN

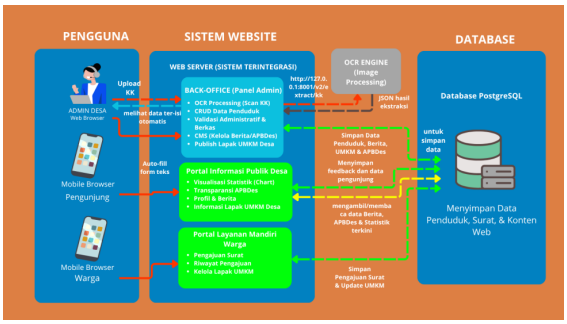
Peningkatan kualitas pelayanan publik di tingkat desa menjadi prioritas strategis, khususnya dalam mewujudkan e-government dan konsep Desa Pintar (Smart Village) (Aziiza et al., 2023; Manapa Sampetoding & Er, 2024). Penerapan e-government memerlukan evaluasi berkelanjutan terhadap kualitas sistem guna menjaga kepercayaan masyarakat (G. C. et al., 2024). Namun, pengelolaan data kependudukan di Desa Rejosari masih dilakukan secara manual, sehingga rentan terhadap kesalahan ketik saat entri 16 digit NIK, yang berdampak pada antrean layanan dan ketidakakuratan basis data kependudukan (Rabani et al., 2024).

Penelitian sejenis mengenai pengembangan sistem informasi berbasis web dengan evaluasi usability dan performa juga telah dilakukan, seperti sistem pakar rekomendasi diet dan olahraga berbasis Forward Chaining dengan framework Laravel yang mencapai skor System Usability Scale (SUS) sebesar 78,38 (Ikawati & Ramadhan, 2026), serta platform komunitas berbasis web dengan gamifikasi yang menunjukkan rata-rata waktu respons API di bawah 3 detik (Ikawati et al., 2025), menegaskan pentingnya evaluasi usability dan performa sistem secara sistematis sebagaimana diterapkan pada penelitian ini.

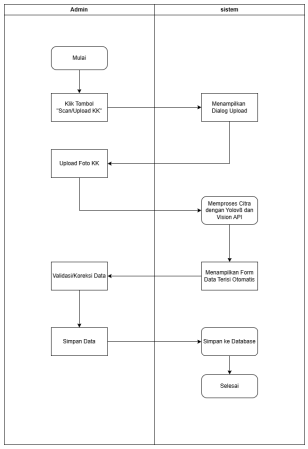
Oleh karena itu, penelitian ini mengusulkan pengembangan Sistem Informasi Manajemen Kependudukan berbasis web untuk Desa Rejosari yang terintegrasi dengan pipeline OCR hibrida. Sistem ini memanfaatkan YOLOv8 untuk melokalisasi area teks dokumen dan Vision-Language Model (Gemini API) untuk mengekstraksi karakter secara otomatis dari gambar Kartu Keluarga (KK) yang diunggah warga. Peralihan dari input manual ke ekstraksi otomatis ini bertujuan untuk meminimalisasi kesalahan ketik, mempercepat penerbitan surat pengantar, dan menyediakan portal layanan digital mandiri bagi penduduk guna memastikan efisiensi dan transparansi tata kelola pemerintahan desa.

METODE PENELITIAN

Pengembangan SIMK diawali dengan analisis kebutuhan menggunakan desain Use Case untuk dua aktor utama: Warga dan Admin Desa. Sistem dibangun dengan arsitektur Model-View-Controller (MVC) pada framework Laravel yang terbukti meningkatkan kecepatan dan transparansi layanan publik (Deri & Rahman, 2026; Ikawati & Ramadhan, 2026), dengan basis data PostgreSQL untuk menyimpan data kependudukan, status verifikasi, dan riwayat permohonan surat.



Gambar 1 Arsitektur Hibrida Sistem Ekstraksi Data Kependudukan



Gambar 2 Flowchart Proses OCR dan Ekstraksi Data

Model deteksi objek YOLOv8 dilatih secara khusus untuk melokalisasi 25 field data kependudukan pada citra Kartu Keluarga, meliputi data header (Nomor KK, Alamat, RT dan RW, Desa/Kelurahan, Kecamatan, Kabupaten/Kota, Provinsi, dan Kode Pos) serta data anggota keluarga (NIK, Nama, Jenis Kelamin, Tempat/Tanggal Lahir, Golongan Darah, Agama, Pendidikan, Jenis Pekerjaan, Kewarganegaraan, Status Hubungan, Status Kawin, Tanggal Kawin, dan Nama Orang Tua). Proses pelatihan menggunakan dataset internal berupa 14 citra KK asli yang dianotasi secara manual menggunakan platform Roboflow, kemudian diperluas melalui augmentasi data (rotasi ±9°, variasi brightness ±19%, dan blur ringan hingga 1,5 piksel) sehingga menghasilkan 32 citra efektif, dengan pembagian 27 citra data latih (training), 3 citra data validasi (validation), dan 2 citra data uji (testing). Model dilatih menggunakan arsitektur YOLOv8n selama 100 epoch pada platform Google Colaboratory dengan akselerasi GPU NVIDIA Tesla T4. Performa model dievaluasi menggunakan metrik mean Average Precision pada rentang Intersection over Union (IoU) 0,5 hingga 0,95 (mAP@0.5-0.95), dengan hasil sebesar 0,398 (39,8%), serta mAP@0.5 sebesar 0,506 (50,6%), precision 0,359, dan recall 0,581. Nilai ini mengindikasikan performa deteksi yang masih moderat, sejalan dengan keterbatasan jumlah data latih (27 citra untuk 25 kelas), di mana beberapa kelas dengan kemiripan

visual tinggi atau frekuensi kemunculan rendah pada data validasi (seperti Desa/Kelurahan, Kode Pos, dan RT/RW) memperoleh nilai Average Precision mendekati nol. Pada tahap produksi, sistem menetapkan confidence threshold sebesar 0,7 agar hanya deteksi bounding box dengan tingkat keyakinan tinggi yang diteruskan ke tahap ekstraksi karakter oleh Vision API.

Evaluasi sistem terdiri dari Black Box Testing (validitas input-output, keamanan, pencegahan duplikasi NIK), kalkulasi skor ekstraksi OCR dari 18 sampel, dan audit performa web menggunakan Lighthouse via Ngrok. Aspek usability menjadi indikator penting dalam menilai penerimaan sistem informasi oleh pengguna, sebagaimana diterapkan pada pengembangan sistem berbasis web serupa (Suria, 2024; Ikawati & Ramadhan, 2026; Ikawati et al., 2025).

HASIL DAN PEMBAHASAN

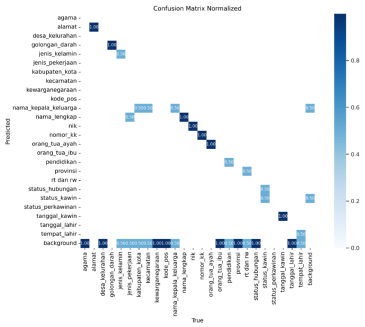
Pengujian kinerja ekstraksi menggunakan model YOLOv8 hasil pelatihan ulang (yolo_v1_kk_map398.pt) dilakukan pada 18 sampel citra KK yang beragam, dengan jumlah anggota keluarga bervariasi antara 2 hingga 6 orang per KK. Analisis tata letak dokumen sebelum OCR sangat menentukan keberhasilan ekstraksi, terutama pada dokumen berstruktur padat (Fleischhacker et al., 2025; Li et al., 2025). Seperti disajikan pada Tabel 1, dari 18 sampel yang diuji, 14 citra (77,8%) berhasil diproses hingga tuntas oleh sistem, sedangkan 4 citra (22,2%) gagal diproses akibat galat sistem (“Internal processing error”) yang belum sepenuhnya teridentifikasi penyebabnya, diduga terkait format atau kualitas berkas PDF tertentu. Dari 14 citra yang berhasil diproses, sistem mencapai skor ekstraksi rata-rata 96,9% (rentang 95-100%), dengan rata-rata waktu proses 30,0 detik per dokumen, membuktikan efektivitas pendekatan hibrida dalam mengurangi kesalahan entri manual pada kasus yang berhasil diproses.

Tabel 1
Evaluasi Ekstraksi Data KK dengan Model YOLOv8 Terbaru (18 Sampel)

Kategori Hasil	Jumlah Sampel	Persentase	Skor Rata-rata (%)
Berhasil diproses	14	77,8%	96,9
Gagal (galat sistem)	4	22,2%	–
Total	**18**	**100%**	**–**

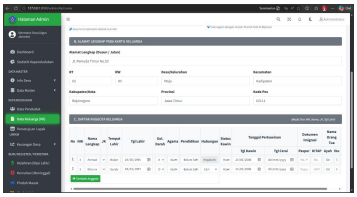
Selain evaluasi akurasi ekstraksi karakter, performa model deteksi YOLOv8 itu sendiri turut dievaluasi menggunakan metrik mAP@0.5-0.95 sebagaimana dijelaskan pada

bagian Metode Penelitian, dengan hasil sebesar 0,398 (39,8%) dan mAP@0.5 sebesar 0,506 (50,6%), tergolong moderat akibat keterbatasan data latih (27 citra untuk 25 kelas field). Meskipun demikian, skor ekstraksi akhir pada sampel yang berhasil diproses tetap tinggi (96,9%) karena pada mode operasi yolo_vlm, keluaran deteksi YOLOv8 (termasuk yang kurang presisi) tetap diteruskan ke Gemini Vision API dalam bentuk gambar asli beserta anotasi bounding box, bukan sebagai crop tunggal yang kaku, sehingga kemampuan pemahaman visual Gemini turut mengompensasi ketidaksempurnaan lokalisasi YOLOv8.

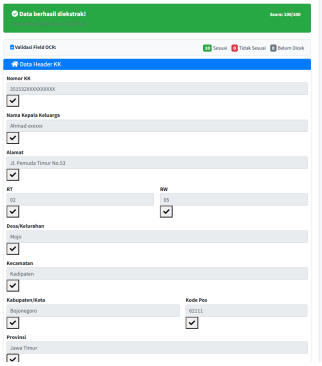


Gambar 3 Confusion Matrix Normalized Hasil Evaluasi Model YOLOv8

Seperti disajikan pada Gambar 3, sebagian besar kelas dengan jumlah instans terbatas pada data validasi tidak menghasilkan prediksi benar sama sekali, sementara kelas dengan pola visual lebih khas seperti nik, nomor_kk, alamat, dan tanggal_kawin mencapai akurasi klasifikasi 1,00. Pola ini konsisten dengan keterbatasan data latih yang telah dibahas, dan menegaskan perlunya penambahan volume serta variasi data latih pada penelitian lanjutan.

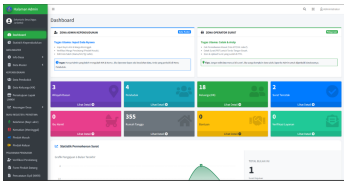


Gambar 4 Hasil Bounding Box YOLOv8



Gambar 5 Antarmuka Verifikasi Admin

Kinerja sistem dipengaruhi oleh kualitas citra dokumen. YOLOv8 dari Ultralytics unggul dalam ekstraksi fitur dan prediksi bounding box yang tangguh pada dokumen (Huang et al., 2024; Jocher et al., 2023; Wang, 2025). Pada citra jernih, YOLOv8 menghasilkan bounding box presisi dan Vision API membaca karakter secara konsisten (akurasi 100%). Sebaliknya, pada citra buram, noise dan distorsi visual menghambat lokalisasi optimal sehingga akurasi turun ke 80%.



Gambar 6 Antarmuka sistem untuk verifikasi permohonan surat warga

Tabel 2

Hasil Pengujian Black Box

Modul Fungsional	**Skenario Uji / Input**	**Hasil yang Diharapkan**	**Status**
Registrasi Warga	Mengunggah gambar Kartu Keluarga (KK) untuk OCR	Sistem secara akurat mengekstraksi kolom NIK, Nama, dan Alamat menggunakan YOLOv8.	Valid
Validasi NIK	Memasukkan NIK yang sudah ada di basis data	Sistem menolak registrasi dan menampilkan peringatan "Data sudah ada".	Valid
Verifikasi Admin	Admin meninjau pendaftaran warga dengan status "Pending"	Sistem memperbarui status akun menjadi "Active" dan memicu notifikasi.	Valid
Pemetaan Keluarga	Mengakses menu permohonan surat	Sistem menampilkan daftar dropdown seluruh anggota keluarga terverifikasi dari KK yang sama.	Valid
Penerbitan Surat	Mengajukan permohonan surat keterangan (misal: SKU)	Sistem membuat log permohonan dan mengirimkannya ke dasbor admin.	Valid
Penolakan Permohonan	Admin menolak permohonan surat	Sistem mewajibkan alasan penolakan spesifik dan menampilkannya di portal warga.	Valid
Direktori UMKM	Mengklik tombol "Pesan" pada sebuah produk	Sistem mengarahkan pengguna ke tautan WhatsApp penjual untuk komunikasi langsung.	Valid

Audit performa dilakukan menggunakan Google Chrome DevTools (Lighthouse) via tunneling Ngrok untuk mensimulasikan akses jaringan publik pada mode Desktop dan Mobile.

Tabel 3
Metrik Performa LightHouse

Metrik Evaluasi	Skor Desktop	Skor Mobile	Interpretasi
Performance	99	87	Waktu pemuatan cepat dan eksekusi sumber daya sangat efisien di berbagai perangkat.
Accessibility	100	100	Navigasi antarmuka yang sangat baik, teks yang dapat dibaca, dan dapat diakses penuh oleh semua pengguna.
Best Practices	100	100	Mematuhi standar pengembangan web modern dan protokol HTTPS yang aman.
SEO	100	100	Sangat optimal untuk pengindeksan mesin pencari dan kemudahan penemuan (discoverability).

Dibandingkan dengan pendekatan OCR tradisional tunggal seperti EasyOCR yang mengekstraksi seluruh teks tanpa mengetahui batas maupun jenis field secara struktural, pendekatan hibrida yang diusulkan terlebih dahulu melokalisasi area setiap field melalui YOLOv8 sebelum karakter diekstraksi oleh Vision API. Sebagai ilustrasi empiris, pengujian EasyOCR pada salah satu citra KK sampel menghasilkan 121 blok teks terdeteksi dengan rata-rata confidence hanya 0,264, di mana 83,5% di antaranya di bawah 0,5. Secara kualitatif, EasyOCR juga gagal membaca sejumlah label field secara akurat, misalnya “Desa/Kelurahan” terbaca sebagai “Desa kelirahar” dan “KALIMANTAN BARAT” terbaca sebagai “KalikaNiAN Barat”, akibat tidak adanya konteks struktural mengenai posisi field pada tabel yang padat. Hal ini mengindikasikan bahwa tahap lokalisasi berbasis object detection memberikan konteks struktural yang tidak dimiliki oleh OCR generik, sehingga meningkatkan ketepatan pemetaan hasil ekstraksi ke field basis data yang sesuai, meskipun pengujian pembanding ini masih bersifat indikatif karena baru dilakukan pada satu citra sampel dan perlu diperluas pada penelitian selanjutnya.

Perlu dicatat bahwa evaluasi ekstraksi pada penelitian ini masih menggunakan sampel terbatas, yaitu 18 citra KK dengan 14 di antaranya berhasil diproses, sehingga hasil yang diperoleh bersifat indikatif dan belum dapat digeneralisasi secara statistik untuk skala

layanan publik desa yang lebih luas. Tingkat kegagalan pemrosesan sebesar 22,2% pada pengujian ini juga menunjukkan perlunya perbaikan ketangguhan sistem terhadap variasi format dokumen. Pengujian pada jumlah sampel yang lebih besar dan bervariasi, disertai perbaikan penanganan galat, diperlukan pada tahap penelitian selanjutnya untuk memvalidasi keandalan sistem secara lebih menyeluruh.

SIMPULAN

Implementasi SIMK berbasis web berhasil mendigitalisasi alur kerja administrasi Desa Rejosari. Integrasi arsitektur hibrida YOLOv8 dan Vision API terbukti efektif mengotomatisasi ekstraksi data KK dengan skor rata-rata 96,9% pada sampel yang berhasil diproses (77,8% dari 18 sampel uji), meskipun performa YOLOv8 secara mandiri masih moderat (mAP@0.5-0.95 = 0,398) akibat keterbatasan data latih, serta skor Lighthouse yang sangat baik (Performance 99/100 desktop, 87/100 mobile). Sistem juga berhasil mencegah duplikasi NIK sehingga integritas basis data terjaga. Mengingat sampel yang masih terbatas dan tingkat kegagalan pemrosesan 22,2%, penelitian lanjutan perlu difokuskan pada penambahan volume dan variasi dataset, penguatan penanganan galat sistem, peningkatan ketangguhan OCR terhadap dokumen resolusi rendah, serta perluasan ekstraksi ke dokumen lain seperti KTP.

DAFTAR PUSTAKA

Aziiza, A. A., Sulistiyani, E., & Fitri, A. S. (2023). What is the Element of the Smart Village Model?: Domains, aspects and indicators. *INTENSIF: Jurnal Ilmiah Penelitian Dan Penerapan Teknologi Sistem Informasi,* *7*(1), 146–160. https://doi.org/10.29407/intensif.v7i1.18898

Deri, D. A., & Rahman, A. (2026). Rancang Bangun Sistem Informasi Manajemen Data Penduduk Desa Berbasis Web Menggunakan Laravel. *Jurnal Inovasi Komputer (INOKOM), 2*(1), 12–21. https://doi.org/10.71200/nktzmb34

Fleischhacker, D., Kern, R., & Göderle, W. (2025). Enhancing OCR in historical documents with complex layouts through machine learning. *International Journal on Digital Libraries, 26*(1). https://doi.org/10.1007/s00799-025-00413-z

G. C., S. B., Gurung, S. K., Rana, S. B. J. B., & Dhungana, B. R. (2024). e-Governance, citizen satisfaction and net benefits: the moderating effect of digital divide. *Cogent Business & Management, 11*(1). https://doi.org/10.1080/23311975.2024.2402512

Huang, X., Yang, S., Xiong, A., & Yang, Y. (2024). Enhanced YOLOv8 with VTR Integration: A Robust Solution for Automated Recognition of OCS Mast Number Plates. *IEEE Access, 12*, 179648–179663. https://doi.org/10.1109/ACCESS.2024.3505208

Ikawati, Y., Ferdiansyah, R., Mufid, M. R., Aditama, D., & Mawaddah, S. (2025). PaletteStream: A Promotional, and Community Web-Based Platform for Visual Artists with Gamification

Implementation. *Jurnal Teknologi Informasi Dan Terapan (J-TIT), 12*(2), 153–162. https://doi.org/10.25047/jtit.v12i2.422

Ikawati, Y., & Ramadhan, B. A. (2026). Interactive Web-Based Expert System for Personalized Diet and Exercise Recommendations Using Forward Chaining. *Journal of Advanced Vocational Information and Communication Technology, 1*(1), 1–14. https://doi.org/10.52158/javict.v1i1.1475

Jocher, G., Chaurasia, A., & Qiu, J. (2023). *Ultralytics YOLO.* https://github.com/ultralytics/ultralytics

Li, D. L., Lee, S. K., & Liu, Y. T. (2025). Printed document layout analysis and optical character recognition system based on deep learning. *Scientific Reports, 15*(1). https://doi.org/10.1038/s41598-025-07439-y

Manapa Sampetoding, E. A., & Er, M. (2024). Digital Transformation of Smart Village: A Systematic Literature Review. *Procedia Computer Science, 239*, 1336–1343. https://doi.org/10.1016/j.procs.2024.06.304

Rabani, H. F., Widiati, I. S., Agustin, T., & Turmudi, H. (2024). Perancangan Sistem Informasi Pengolahan Data Kependudukan Berbasis Website Desa Penggung Boyolali. *Jurnal Penelitian Sistem Informasi (JPSI), 2*(3), 190–198. https://doi.org/10.54066/jpsi.v2i3.2339

Suria, O. (2024). Assessing Financial Information System Usability Using System Usability Scale (SUS) and Usability Metric for User Experience-Lite (UMUX-Lite). *Jurnal Teknologi Sistem Informasi Dan Aplikasi, 7*(2), 538–547. https://doi.org/10.32493/jtsi.v7i2.38723

Wang, X. (2025). *A Review of Text Recognition in Complex Scene Images Based on Deep Learning* (pp. 776–783). https://doi.org/10.2991/978-94-6463-821-9_75