

Konversi Tulisan Tangan Huruf Kapital Menjadi Teks Menggunakan Metode Deep Learning Berbasis YOLOv8 dan CTC

Makmur Jaya Nur¹ | Rizki Yusliana Bakti^{*2} | Fahrin Irhamna Rachman²

¹ Mahasiswa Program Studi Informatika,
Fakultas Teknik, Universitas
Muhammadiyah Makassar, Indonesia.
Email:
105841104521@student.unismuh.ac.id

² Program Studi Arsitektur, Fakultas Teknik,
Universitas Muhammadiyah Makassar,
Indonesia.
Email:
fachrim141020@unismuh.ac.id,
rizkiyusliana@unismuh.ac.id

Korespondensi:
Rizki Yusliana Bakti
rizkiyusliana@unismuh.ac.id

ABSTRAK

Penelitian ini mengkaji pengembangan sistem konversi tulisan tangan ke teks digital menggunakan metode deep learning dengan mengombinasikan arsitektur Convolutional Neural Network (CNN), YOLOv8, dan Connectionist Temporal Classification (CTC). Dataset yang digunakan terdiri dari 700 citra tulisan tangan huruf kapital (A–Z) yang diperoleh dari dokumen resmi Dinas Kependudukan dan Pencatatan Sipil Kabupaten Barru. Tahapan penelitian meliputi prapemrosesan citra berupa grayscale, normalisasi, perataan teks, serta augmentasi data, dilanjutkan dengan anotasi bounding box menggunakan Roboflow. Dataset kemudian dibagi menjadi data pelatihan, validasi, dan pengujian. Model YOLOv8 dilatih untuk mendeteksi karakter dan hasilnya diproses menggunakan CTC untuk menghasilkan teks akhir. Evaluasi menunjukkan performa yang baik dengan precision 98,38%, recall 87,25%, F1-score 92,44%, serta mAP@0.5 sebesar 87,19%. Hasil ini menunjukkan metode yang diusulkan efektif untuk mendukung digitalisasi dokumen administrasi publik.

Kata Kunci:

YOLOv8, Konversi Tulisan Tangan, Deep Learning, Citra Digital, Administrasi Publik, Roboflow, CNN, CTC

ABSTRACT

This study investigates the development of a handwritten text-to-digital text conversion system using deep learning by combining Convolutional Neural Network (CNN), YOLOv8, and Connectionist Temporal Classification (CTC) architectures. The dataset consists of 700 images of uppercase handwritten letters (A–Z) obtained from official documents of the Department of Population and Civil Registration of Barru Regency. The research stages include image preprocessing such as grayscale conversion, normalization, text alignment, and data augmentation, followed by bounding box annotation using Roboflow. The dataset is then divided into training, validation, and testing sets. The YOLOv8 model is trained to detect characters, and the outputs are processed using CTC to generate the final text. Evaluation results demonstrate strong performance, achieving a precision of 98.38%, recall of 87.25%, an F1-score of 92.44%, and an mAP@0.5 of 87.19%. These findings indicate that the proposed method is effective in supporting the digitalization of public administrative documents.

Keywords:

YOLOv8, Handwriting Conversion, Deep Learning, Digital Image, Public Administration, Roboflow, CNN, CTC

1 | PENDAHULUAN

Menurut Kamus Besar Bahasa Indonesia tulisan tangan adalah tulisan yang ditulis oleh tangan bukan diketik. Tulisan tangan merupakan hal yang unik bagi setiap orang. Tulisan tangan manusia memiliki perbedaan antara satu dengan yang lainnya dan sering terjadi tulisan tangan sulit terbaca atau sulit dikenali dan hal ini dapat menghambat aktivitas sehari-hari, misalnya aktivitas transaksi yang memerlukan tulisan tangan (Effendi et al., 2023).

Berdasarkan pendekatan ini banyak digunakan di bidang ini tetapi seringkali mengalami kesulitan dalam hal akurasi karena variasi gaya tulisan tangan, distorsi, dan noise pada gambar yang dipindai (Kumar & Dharshan, 2025). Tantangan utama dalam pengenalan karakter huruf pada gambar tulisan tangan adalah variasi yang tinggi dalam gaya tulisan tangan, ukuran karakter, orientasi, dan kejelasan gambar. (Swasono et al., 2024).

Teknologi saat ini dapat dikatakan sebagai bagian dari kehidupan sehari-hari. Teknologi terus dikembangkan seiring tingginya penggunaan operasi digital. Operasi digital dianggap memberikan kemudahan bagi penggunanya seperti pemesanan online, pertukaran informasi, pengenalan objek, dan sebagainya (Rahmadani et al., 2016). Aplikasi komputer untuk pengenalan tulisan tangan (Handwriting recognition) adalah kemampuan komputer untuk menerima dan menafsirkan input tulisan tangan yang dapat dimengerti dari sumber seperti teks tertulis, foto, layar sentuh dan perangkat lainnya (Mahmudi et al., 2017).

Perkembangan teknologi pengolahan citra dan kecerdasan buatan pada masa kini semakin mendorong penerapan sistem otomatis untuk mengidentifikasi tulisan tangan digital (Tanjung et al., 2025). Tradisi menulis di atas kertas menggunakan alat tulis mulai ditinggalkan karena sudah adanya aplikasi untuk mengetik dokumen dan menyimpannya di tempat yang lebih aman, dibandingkan sebuah kertas atau buku yang dapat rusak jika dibandingkan dengan teks digital (Giamiko & Tjiong, 2024).

Dalam hal karakter tulisan saat ini menjadi banyak diadakannya sebuah penelitian dalam dunia kecerdasan buatan dengan menggunakan berbagai metode penelitian (Andono & Rachmawanto, 2022). Berbagai algoritma telah digunakan untuk menangani klasifikasi angka tulisan tangan, antara lain Convolutional Neural Network (CNN), K-Nearest Neighbor (KNN), dan Support Vector Machine (SVM) (Rosyada et al., 2025).

Penelitian pengenalan tulisan tangan sudah dilakukan puluhan tahun dengan fokus pembahasan yang beragam. Ada banyak survey artikel yang dipublikasikan tentang penelitian ini. Paradigma merupakan perspektif atau cara pandang melihat sesuatu. Konsep paradigma pada aktivitas ilmiah dibahas oleh Thomas Kuhn (Teori et al., 2023). Mendeteksi individu berdasarkan tulisan tangan adalah jenis identifikasi biometrik perilaku yang diakui dengan baik oleh psikolog, ahli saraf, ahli paleografi, analisis forensik, analisis dokumen, dan peneliti ilmu komputer (Alaci, 2023). Para peneliti "Pengenalan dan Konversi Teks Tulisan Tangan Menggunakan Model Deep Learning Berbasis Convolutional Neural Network (CNN)" menggunakan teks tulisan tangan sebagai gambar masukan, kemudian mengonversinya menjadi teks digital (Odeh et al., 2022). Pengolahan citra dapat berupa manipulasi citra dalam bidang visi komputer. Dengan perkembangan teknologi, terdapat banyak teknik untuk memanipulasi foto (Manchala, 2020).

Model invariance memanfaatkan sparsity alami pada variabel keadaan dengan membangun abstraksi keadaan yang lebih kasar pada tingkat per variabel, memungkinkan kemampuan generalisasi baru atas komposisi baru dari nilai variabel keadaan (Zhang et al., n.d.).

Klasifikasi karakter tunggal dan angka telah mencapai hasil yang tinggi menggunakan model CNN sebagai pengklasifikasi karakter. Misalnya, Ahlawat dkk. menggunakan model CNN untuk mencapai akurasi 99,87% pada basis data angka yang sangat populer, yaitu dataset MNIST (Chen, 2023).

Penggunaan deep learning pada zaman ini sangatlah meningkat. Salah satu arsitektur deep learning yaitu Convolutional Neural Network (CNN) dapat digunakan untuk pengenalan pola seperti karakter, angka, dan penyelesaian kasus-kasus computer vision. CNN dapat digunakan sebagai basis dalam membangun sistem pengenalan karakter yang ditulis oleh manusia (Briliantio et al., 2020). Beberapa peneliti menyebutkan bahwa pengenalan tulisan tangan menggunakan teknik deep learning telah menghasilkan akurasi yang lebih tinggi dari pada menggunakan teknik Conventional Machine Learning (Abdiansah et al., 2025).

Berbagai inovasi teknologi kini mendukung kemajuan di bidang pendidikan. Misalnya, sistem e-learning memungkinkan siswa mengakses materi pembelajaran kapan saja dan di mana saja, sementara platform video konferensi memfasilitasi pembelajaran jarak jauh (Apriani & Pratiwi, 2025). Seiring meningkatnya digitalisasi, sistem pengenalan objek semakin banyak digunakan oleh masyarakat umum. Teknologi ini penting dalam berbagai aplikasi, seperti digitalisasi dokumen, sistem pembayaran otomatis, dan analisis data numerik (Salsabila et al., 2025). Dari permasalahan di atas dibutuhkan sebuah sistem yang dapat mengenali tulisan tangan (Pradika et al., 2020).

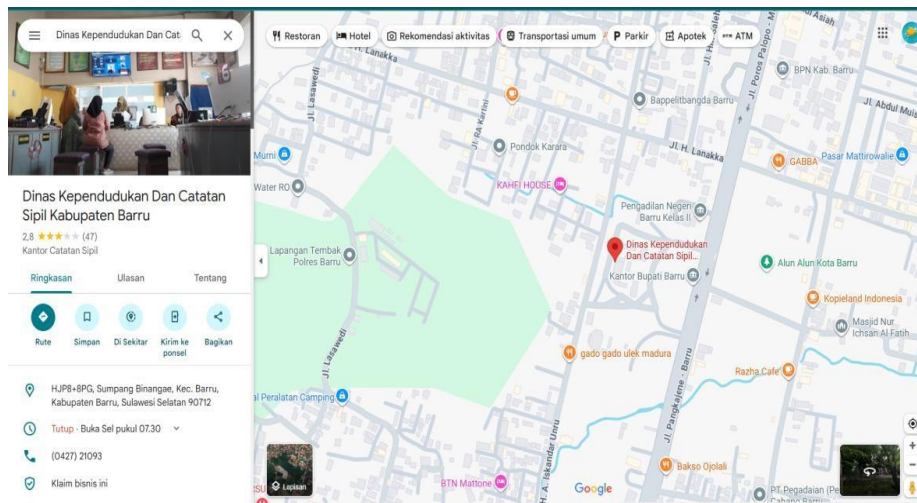
2 | METODE

2.1 | Lokasi Penelitian

Penelitian terkait konversi tulisan tangan menjadi teks menggunakan metode Deep Learning dapat dilakukan di berbagai instansi yang masih menggunakan dokumen tulisan tangan dalam proses pencatatan data, seperti kantor pemerintahan atau lembaga administrasi publik. Dataset yang digunakan terdiri dari 700 gambar tulisan tangan yang telah melalui proses pelabelan. Waktu pelaksanaan penelitian dapat bervariasi tergantung pada kompleksitas sistem, ketersediaan data, dan tahapan pengujian model.

Penentuan lokasi merupakan aspek penting dalam proses penelitian karena akan memudahkan peneliti dalam memperoleh data yang relevan dan sesuai dengan kebutuhan penelitian. Lokasi pengambilan data pada penelitian ini dilakukan di Dinas

Kependudukan dan Pencatatan Sipil (Dukcapil) Kabupaten Barru, yang beralamat di Jalan Sultan Hasanuddin No. 93, Kabupaten Barru, Sulawesi Selatan, dan masih memiliki banyak arsip dokumen dalam bentuk tulisan tangan.



GAMBAR 1. Lokasi Penelitian

2.2 | Teknik Pengumpulan Data

Pertama, data tulisan tangan asli dikumpulkan secara langsung di Dinas Kependudukan dan Catatan Sipil (Dukcapil) Kabupaten Barru. Data tersebut diambil dari dokumen resmi dan formulir yang diisi dengan tangan oleh petugas serta masyarakat. Proses pengumpulan ini bertujuan untuk memperoleh contoh tulisan tangan yang otentik dan mencerminkan kondisi nyata di lingkungan Dukcapil Kabupaten Barru, agar model pembelajaran mendalam yang dibuat dapat memahami dan mengenali karakter tulisan tangan yang digunakan dalam konteks administrasi.

2.3 | Analisis Desain

Pertama, data tulisan tangan asli dikumpulkan secara langsung di Dinas Kependudukan dan Catatan Sipil (Dukcapil) Kabupaten Barru. Data tersebut diambil dari dokumen resmi dan formulir yang diisi dengan tangan oleh petugas serta masyarakat. Proses pengumpulan ini bertujuan untuk memperoleh contoh tulisan tangan yang otentik dan mencerminkan kondisi nyata di lingkungan Dukcapil Kabupaten Barru, agar model pembelajaran mendalam yang dibuat dapat memahami dan mengenali karakter tulisan tangan yang digunakan dalam konteks administrasi.

Penelitian ini menggunakan data tulisan tangan yang diperoleh dari dokumen resmi di Dinas Kependudukan dan Pencatatan Sipil (Dukcapil) Kabupaten Barru, yang berlokasi di Jalan Sultan Hasanuddin No. 93, Kabupaten Barru, Sulawesi Selatan. Data tersebut berasal dari berbagai arsip dan formulir yang diisi secara manual oleh petugas maupun masyarakat. Bentuk data yang dikumpulkan berupa citra hasil pemindaian atau foto dokumen, dengan variasi gaya tulisan, ukuran huruf, serta kualitas kertas yang berbeda-beda.

Meski data yang terkumpul cukup beragam, kondisi awalnya masih mentah dan belum siap untuk langsung dipakai dalam pelatihan model. Beberapa kendala yang ditemui antara lain adanya noise pada gambar, orientasi tulisan yang tidak sesuai, perbedaan resolusi gambar, serta jarak antar huruf yang tidak konsisten. Oleh karena itu, sebelum data digunakan dalam proses pelatihan, dilakukan tahapan pra- pemrosesan yang meliputi konversi format, normalisasi ukuran, penyelarasan orientasi teks, dan penghilangan noise.

3 | HASIL PENELITIAN

3.1 | Pengumpulan Dataset

Penelitian ini menggunakan data tulisan tangan yang diperoleh dari dokumen resmi di Dinas Kependudukan dan Pencatatan Sipil (Dukcapil) Kabupaten Barru, yang berlokasi di Jalan Sultan Hasanuddin No. 93, Kabupaten Barru, Sulawesi Selatan. Data tersebut berasal dari berbagai arsip dan formulir yang diisi secara manual oleh petugas maupun masyarakat. Bentuk data yang dikumpulkan berupa citra hasil pemindaian atau foto dokumen, dengan variasi gaya tulisan, ukuran huruf, serta kualitas kertas yang berbeda-beda.

Proses pengumpulan data ini menjadi langkah penting dalam penelitian karena kualitas dan variasi data sangat berpengaruh terhadap performa model deep learning yang dikembangkan untuk mengenali dan mengubah tulisan tangan menjadi teks digital dengan akurasi tinggi.

Tabel 1. Dataset yang berhasil terkumpul

1.	MALUSETASI	100
2.	PUJANANTING	100
3.	SOPPENG RIAJA	100
4.	TANETE RIAJA	100
5.	TANETE RILAU	100
6.	BALUSU	100
7.	BARRU	100

3.2 | Analisis Kebutuhan Ruang

Upload Dataset

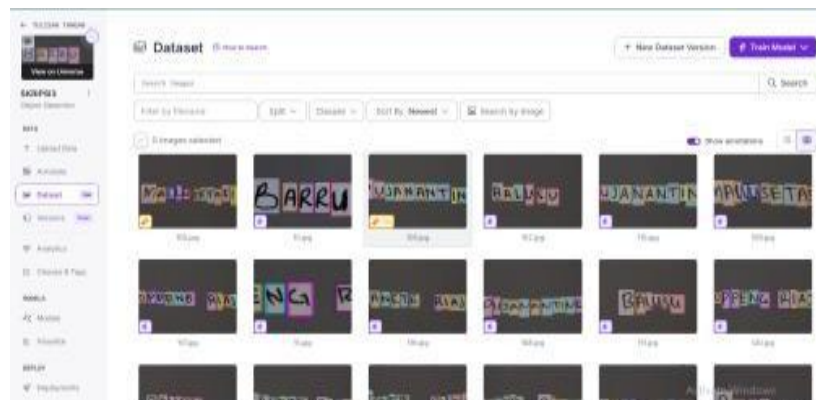
Data gambar yang telah dikumpulkan menggunakan kamera smartphone diunggah di Roboflow sebelum proses dilakukan



Gambar 2. Upload Dataset

Proses Anotasi Gambar

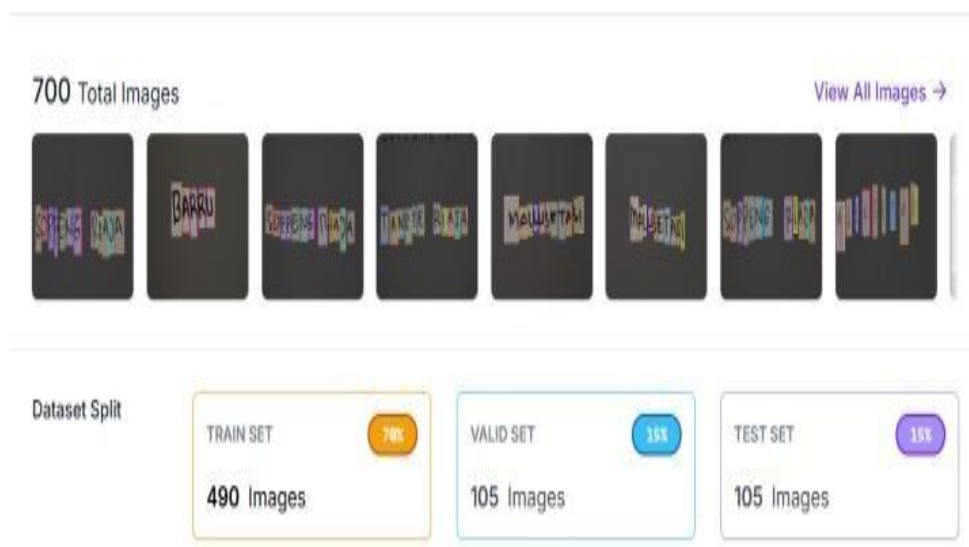
Pada Gambar 6 di bawah ini, terlihat bahwa setiap gambar dalam dataset telah dilengkapi dengan anotasi. Setiap huruf dibatasi oleh kotak berwarna yang menunjukkan area objek yang dideteksi. Warna kotak tersebut disesuaikan dengan kelas huruf yang diwakili, sehingga memudahkan proses identifikasi.



Gambar 3. Pembuatan Kelas

3.3 | Pembagian Dataset

Dataset yang digunakan terdiri dari 700 gambar tulisan tangan yang telah melalui proses pelabelan. Dataset ini kemudian dibagi menjadi tiga bagian sesuai dengan standar pembagian untuk pelatihan model deep learning.

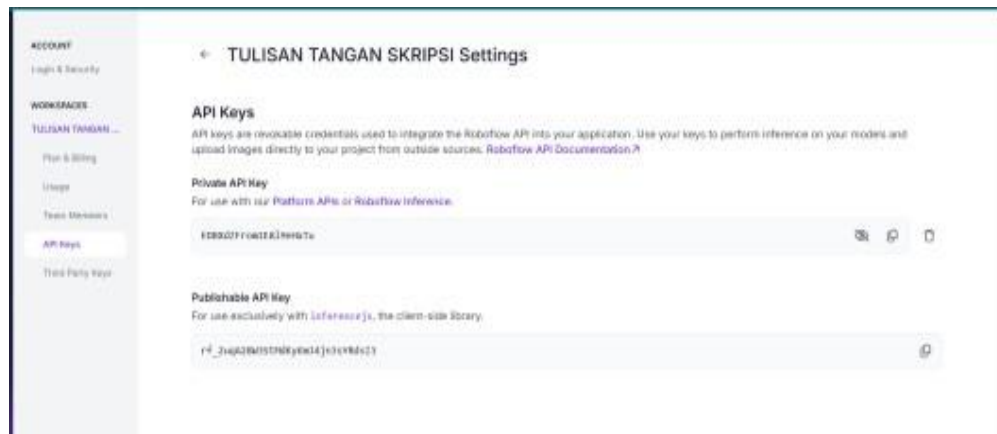


Gambar 4. Pembagian Dataset

3.4 | Pelatihan Model

Persiapan Data Untuk Pelatihan

Dataset ini diambil secara otomatis melalui API Roboflow yang terhubung dengan Google Colab. Dengan menggunakan API tersebut, dataset diambil dan diekstrak langsung ke folder kerja Google Colab, sehingga mempermudah proses pelatihan tanpa memerlukan pengunduhan manual.



Gambar 5. API key

Konfigurasi Model

Model yang digunakan dalam penelitian ini adalah YOLOv8, sebuah model deteksi objek terbaru yang lebih efisien dibandingkan versi sebelumnya. YOLOv8 dapat mengidentifikasi objek dalam gambar secara langsung dengan akurasi tinggi berkat peningkatan arsitektur backbone dan kepala deteksi yang lebih responsive

```
from ultralytics import YOLO

model = YOLO("yolov8n.pt") # kamu bisa ganti dengan yolov8s.pt dll

model.train(
    data="/content/SKRIPSI3-1/data.yaml",
    epochs=100,
    imgsz=640,
    batch=16,
    name="tulisan_tangan_model"
)
```

Gambar 6. Kode Pelatihan Model YOLOv8

Proses Pelatihan

Pelatihan dilakukan dengan parameter berikut: epoch sebanyak 100, ukuran batch 16, dan resolusi gambar 640×640 piksel. Seluruh proses diadakan di Google Colab menggunakan GPU NVIDIA Tesla T4.

```

Epoch      GPU_mem  loss      cls_loss  dfl_loss  Instances  Size
99/100      2.79G    1.257    0.4897    1.346     189        640: 100% | 11/31 [00:03:00:00, 3.131it/s]
Class      Images  Instances  BoxP      mAP50  mAP50-95: 100% | 16/16 [00:03:00:00, 4.171it/s]

Epoch      GPU_mem  loss      cls_loss  dfl_loss  Instances  Size
100/100     2.79G    1.248    0.484     1.337     92         640: 100% | 12/31 [00:03:00:00, 3.431it/s]
Class      Images  Instances  BoxP      mAP50  mAP50-95: 100% | 16/16 [00:03:00:00, 4.161it/s]

100 epochs completed in 0.381 hours.
Optimizer stripped from runs/detect/tulisan_tangan_model/weights/last.pt, 6.390
Optimizer stripped from runs/detect/tulisan_tangan_model/weights/best.pt, 6.390

Validating runs/detect/tulisan_tangan_model/weights/best.pt...
Ultralytics 8.3.176 Python 3.11.11 torch 2.6.0dev20 CUDA0 (Tesla T4, 15919MiB)
Model summary (fused): 72 layers, 1,908,863 parameters, 0 gradients, 8.1 GByte

Class      Images  Instances  BoxP      mAP50  mAP50-95: 100% | 16/16 [00:03:00:00, 1.931it/s]
all         400      4591     0.871    0.877    0.875    0.586
A           400      895     0.992    0.991    0.991    0.678
B           148      148     0.994     1     0.995    0.668
E           273      401     0.996    0.998    0.991    0.647
O           134      134     0.979     1     0.975    0.651
I           336      339     0.979    0.978    0.985    0.567
J           195      197     0.97     0.981    0.981    0.605
L           210      254     0.975    0.996    0.988    0.668
U           1         1         0         0         0         0
H           74       75     0.966    0.987    0.991    0.691
B           263      403     0.991    0.991    0.992    0.698
O           67       67     1     0.995    0.995    0.653
P           137      204     0.998     1     0.995    0.687
B           272      368     0.996    0.994    0.992    0.681
S           215      292     0.992    0.997    0.991    0.646
T           272      402     0.991    0.996    0.991    0.656
U           354      428     0.985    0.998    0.991    0.671
w           1         1         0         0         0         0

Speed: 0.5ms preprocess, 2.7ms inference, 0.8ms loss, 4.1ms postprocess per image

```

Gambar 7. Proses Pelatihan

3.5 | Hasil Pelatihan Model

Proses dimulai dengan memuat gambar 16.jpg ke dalam sistem, lalu model melaksanakan tahap prapemrosesan untuk menyesuaikan gambar dengan format input yang diinginkan, yang dalam kasus ini memakan waktu sekitar 3,5 milidetik. Setelah itu, model melakukan inferensi atau proses prediksi untuk mengidentifikasi posisi dan jenis huruf dalam gambar, yang memerlukan waktu sekitar 14,9 milidetik.

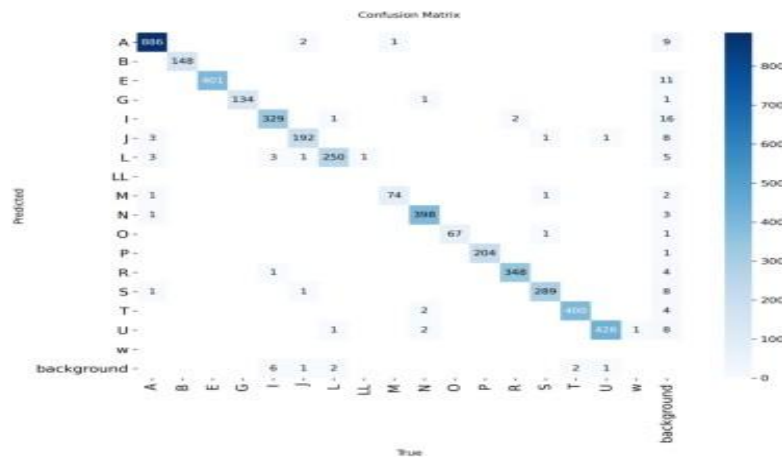


GAMBAR 8. Hasil Prediksi

Pada gambar hasil deteksi tampak bahwa setiap huruf dari kata "MALUSETASI" telah teridentifikasi dengan baik oleh model. Setiap huruf dikelilingi oleh kotak pembatas berwarna berbeda untuk memisahkannya secara visual, dan di atas setiap kotak terdapat label huruf beserta nilai confidence score yang menunjukkan seberapa yakin model terhadap prediksi tersebut.

Visualisasi Hasil Prediksi Menggunakan Confusion Matrix

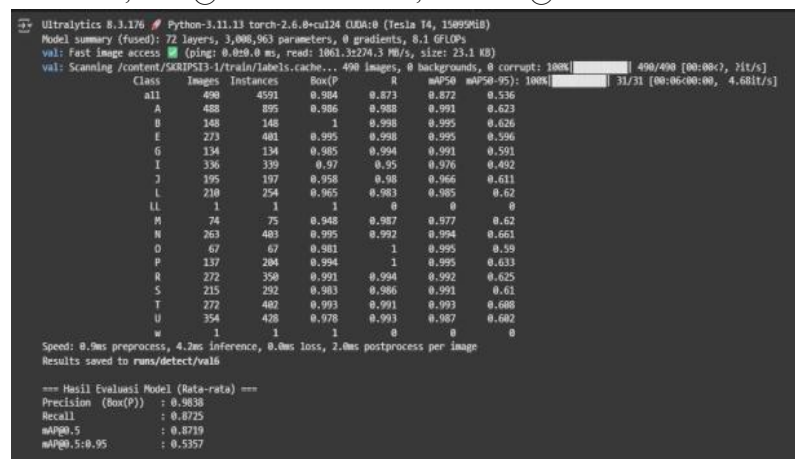
Gambar di bawah memperlihatkan Matriks Kebingungan dari evaluasi model deteksi huruf kapital tulisan tangan. Dalam matriks itu, sumbu horizontal (True) menggambarkan label asli dari dataset, sedangkan sumbu vertikal (Predicted) merefleksikan label yang diprediksi oleh model. Setiap sel pada diagonal utama (dari kiri atas ke kanan bawah) menggambarkan total prediksi yang akurat untuk setiap kelas huruf, sementara sel di luar diagonal menunjukkan jumlah prediksi yang keliru atau huruf yang tertukar dengan kelas yang berbeda.



Gambar 9. Confusion matrix

Hasil Akurasi

Gambar dibawah ini menunjukkan hasil evaluasi dari model YOLOv8 setelah dilatih, yang mencakup informasi spesifikasi model seperti jumlah layer (72), total parameter (3.008.963), dan penggunaan GPU Tesla T4. Evaluasi dilakukan pada 490 gambar uji dengan metrik utama Precision, Recall, mAP@0.5, dan mAP@0.5:0.95 untuk setiap kelas huruf. Nilai Precision menggambarkan seberapa akurat prediksi model, sedangkan Recall menggambarkan kemampuan model dalam mengidentifikasi semua objek yang ada. mAP@0.5 mengukur ketepatan deteksi pada ambang IoU 50%, sedangkan mAP@0.5:0.95 mengukur ketepatan pada ambang IoU yang lebih ketat. Berdasarkan rata-rata total, model mendapatkan Precision sebesar 0.9838, Recall 0.8725, mAP@0.5 sebesar 0.8719, dan mAP@0.5:0.95 sebesar 0.5357



GAMBAR 2 Hasil evaluasi model

4 | PEMBAHASAN

Tulisan tangan huruf kapital menjadi teks digital dengan memanfaatkan pendekatan deep learning melalui kombinasi arsitektur Convolutional Neural Network (CNN), YOLOv8, dan Connectionist Temporal Classification (CTC). Pendekatan ini dipilih untuk mengatasi tantangan utama dalam pengenalan tulisan tangan, yaitu variasi bentuk karakter, perbedaan gaya penulisan antarindividu, serta ketidakteraturan jarak dan ukuran huruf yang umum ditemukan pada dokumen tulisan tangan, khususnya dokumen administrasi publik. Penilaian manual terhadap aspek ini rentan terhadap subjektivitas dan kurang efisien. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja tiga versi model You Only Look Once (YOLO), yaitu YOLOv5, YOLOv8, dan YOLO 11 dalam mendeteksi jarak antar baris tulisan tangan(et al., 2025)

Tahap pra-pemrosesan citra memiliki peran penting dalam meningkatkan kualitas data sebelum proses pelatihan model. Proses konversi citra ke grayscale bertujuan untuk menyederhanakan informasi visual dengan menghilangkan komponen warna yang tidak relevan, sehingga model dapat lebih fokus pada pola bentuk karakter. Normalisasi dilakukan untuk menyamakan skala intensitas piksel, sementara perataan teks membantu mengurangi kemiringan tulisan yang dapat memengaruhi akurasi deteksi karakter. Selain itu, augmentasi data diterapkan untuk memperkaya variasi dataset, mengingat

jumlah data yang relatif terbatas, sehingga model dapat belajar dari berbagai kemungkinan bentuk tulisan tangan dan meningkatkan kemampuan generalisasi.

Penggunaan YOLOv8 sebagai metode deteksi karakter terbukti efektif dalam mengidentifikasi posisi dan batas setiap huruf pada citra tulisan tangan. YOLOv8 memiliki keunggulan dalam mendeteksi objek secara cepat dan akurat karena mampu melakukan proses deteksi dalam satu tahap (single-stage detector). Hal ini sangat sesuai untuk kebutuhan pengenalan karakter tulisan tangan yang memerlukan kecepatan serta ketelitian dalam menentukan bounding box setiap karakter. Hasil evaluasi menunjukkan nilai precision yang sangat tinggi, yang menandakan bahwa sebagian besar karakter yang terdeteksi oleh model merupakan karakter yang benar dan relevan.

Tahap pengenalan urutan karakter menggunakan Connectionist Temporal Classification (CTC) berperan sebagai komponen penting dalam menyusun hasil deteksi karakter menjadi teks yang utuh dan bermakna. Keunggulan CTC terletak pada kemampuannya mengenali urutan karakter tanpa memerlukan proses segmentasi eksplisit antarhuruf. Dalam konteks tulisan tangan, di mana jarak antarhuruf sering kali tidak konsisten, CTC mampu memberikan solusi yang efektif dan fleksibel. Hasil pengujian menunjukkan bahwa kombinasi YOLOv8 dan CTC mampu menghasilkan teks akhir yang cukup akurat dan sesuai dengan tulisan tangan asli. Pengenalan karakter tulisan tangan dibedakan menjadi dua jenis, yaitu pengenalan karakter tulisan tangan offline dan online (Rahmadani et al., 2016)

Penggunaan dataset yang berasal dari dokumen resmi Dinas Kependudukan dan Pencatatan Sipil Kabupaten Barru memberikan nilai praktis yang tinggi terhadap penelitian ini. Data yang digunakan mencerminkan kondisi nyata dokumen administrasi publik, sehingga hasil penelitian lebih relevan untuk diterapkan dalam lingkungan pemerintahan. Sistem yang dikembangkan berpotensi membantu proses digitalisasi arsip, mempercepat pengolahan data, serta mengurangi kesalahan akibat input data manual.

Secara keseluruhan, hasil penelitian menunjukkan bahwa kombinasi CNN–YOLOv8–CTC merupakan pendekatan yang efektif dan menjanjikan untuk sistem konversi tulisan tangan ke teks digital. Meskipun demikian, masih terdapat beberapa keterbatasan, seperti jumlah dataset yang terbatas pada huruf kapital serta belum mencakup angka, huruf kecil, maupun kata dan kalimat lengkap. Oleh karena itu, pengembangan lebih lanjut dengan dataset yang lebih besar dan beragam diharapkan dapat meningkatkan kinerja sistem secara keseluruhan.

5 | KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa pengembangan sistem konversi tulisan tangan menjadi teks digital menggunakan pendekatan deep learning dengan mengombinasikan arsitektur Convolutional Neural Network (CNN), YOLOv8, dan Connectionist Temporal Classification (CTC) mampu memberikan kinerja yang sangat baik dalam mengenali dan mengonversi tulisan tangan huruf kapital. Penelitian ini berhasil membuktikan bahwa integrasi antara metode deteksi objek berbasis CNN melalui YOLOv8 dan metode pengenalan urutan karakter menggunakan CTC merupakan pendekatan yang efektif untuk menangani permasalahan pengenalan tulisan tangan yang memiliki variasi bentuk, ukuran, dan gaya penulisan yang berbeda-beda.

Hasil evaluasi menunjukkan bahwa sistem yang diusulkan mencapai nilai precision rata-rata sebesar 98,38%, recall sebesar 87,25%, dan F1-score sebesar 92,44%, yang mengindikasikan tingkat akurasi dan keandalan sistem yang tinggi dalam mendeteksi dan mengenali karakter tulisan tangan. Selain itu, nilai mAP@0.5 sebesar 87,19% menunjukkan kemampuan model yang baik dalam mendeteksi karakter pada tingkat kepercayaan yang moderat, sementara nilai mAP@0.5:0.95 sebesar 53,57% mencerminkan konsistensi kinerja model pada tingkat ketelitian yang lebih ketat. Secara keseluruhan, hasil ini menunjukkan bahwa metode yang diusulkan mampu bekerja secara stabil dan efisien dalam berbagai skenario pengujian.

Dengan mempertimbangkan seluruh hasil yang diperoleh, dapat disimpulkan bahwa kombinasi CNN–YOLOv8–CTC merupakan pendekatan yang efektif dan potensial untuk diterapkan dalam proses digitalisasi dokumen administrasi publik, khususnya dokumen yang masih menggunakan tulisan tangan huruf kapital. Sistem ini diharapkan dapat membantu meningkatkan efisiensi pengelolaan arsip, mempercepat proses input data, serta mengurangi kesalahan manusia dalam proses penyalinan data secara manual. Ke depan, penelitian ini masih dapat dikembangkan lebih lanjut dengan memperluas jumlah dan variasi dataset, mencakup huruf kecil, angka, serta kata atau kalimat lengkap, serta mengintegrasikan sistem ke dalam aplikasi berbasis web atau mobile agar dapat digunakan secara lebih luas dalam lingkungan pemerintahan maupun institusi lainnya.

Penelitian selanjutnya disarankan untuk memperluas dataset mencakup huruf kecil, angka, serta kata atau kalimat utuh, dan mengintegrasikan sistem ke dalam aplikasi berbasis web atau mobile.

Daftar Pustaka

- Abdiansah, L., Eviyanti, A., & Azizah, N. L. (2025). Implementation of Convolutional Neural Networks Algorithm for Javanese Handwriting Recognition Penerapan Algoritma Convolutional Neural Networks untuk Pengenalan Tulisan Tangan Aksara Jawa. 5(April), 496–504.
- Alaei, F. (2023). Review of age and gender detection methods based on handwriting analysis. *Neural Computing and Applications*, 35(33), 23909–23925. <https://doi.org/10.1007/s00521-023-08996-x>
- Andono, P. N., & Rachmawanto, E. H. (2022). DETEKSI KARAKTER HIRAGANA MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI* | 184. 11, 183–192.
- Apriani, E., & Pratiwi, N. (2025). Perancangan Sistem Pengenalan Tulisan Tangan pada Jawaban Esai Menggunakan Metode CNN-LSTM Berbasis Android. 385–396. <https://doi.org/10.47002/metik.v9i2.1094>
- Briliantio, J., Santosa, N., Ardian, G., Hakim, L., Studi, P., Informatika, T., & Mulia, U. B. (2020). Penerapan Convolutional Neural Network untuk Handwriting Recognition pada Aplikasi Belajar Aritmatika Dasar Berbasis Web. 05, 137–146.
- Chen, A. (2023). Handwritten Text Classification Based on Convolutional Neural Network. *Journal of Artificial Intelligence Research*, 34, 39–44.
- Effendi, A., Lumbanraja, F. R., Junaidi, A., & Syarif, A. (2023). Implementasi metode deep learning untuk klasifikasi gambar tulisan tangan. *Jurnal Pepadun*, 4(2), 100–106.
- Giamiko, E. S., & Tjong, E. L. (2024). Pengembangan Aplikasi Pengenalan Tulisan Tangan Abjad dan Angka Berbasis Convolutional Neural Network. 11(02), 22–30.
- Kumar, J. P., & Dharshan, H. D. (2025). Handwritten text recognition using deep learning: A CNN–LSTM approach. *Intelligent Systems and Applications*, 13(1), 14–22.
- Mahmudi, A., Maghfiroh, R. El, & Sasmito, A. P. (2017). Aplikasi MATLAB untuk Mengenali Karakter Tulisan Tangan. 9(1), 18–22.
- Manchala, S. Y. (2020). Handwritten Text Recognition using Deep Learning with TensorFlow. 9(05), 594–600.
- Odeh, A., Odeh, M., Odeh, H., & Odeh, N. (2022). Handwritten text recognition methods: A review study. *Revue d'Intelligence Artificielle*, 36(2), 333–339.
- Pradika, S. I., Nugroho, B., & Puspaningrum, E. Y. (2020). Pengenalan Tulisan Tangan Huruf Hijaiyah Menggunakan Convolutional Neural Network Dengan Augmentasi Data. 1, 129–136.
- Rahmadani, O., Rozikin, C., & Karawang, U. S. (2016). KINERJA TESSERACT-OCR PENGENALAN DALAM. 13(3).
- Rosyada, M., Putri, I. Z., Virda, S. A., Aditya, D., & Saputra, A. (2025). Pengenalan Angka Tulisan Tangan Menggunakan Metode Machine Learning Random Forest. 9675, 74–80.
- Salsabila, C., Putri, C., Nuraini, A. I., & Adam, R. I. (2025). PENGENALAN ANGKA PADA CITRA TULISAN TANGAN MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN). 9(4), 6169–6174.
- Swasono, N. E., Himamunanto, A. R., Budiati, H., & Immanuel, U. K. (2024). Recognition of Letter Characters in Handwritten Images Using Convolutional Neural Network and K-Means Clustering Algorithm Pengenalan Karakter Huruf pada Gambar Tulisan Tangan Menggunakan Algoritma Convolutional Neural Network dan K-Means Clustering. 4(October), 1646–1656.
- Tanjung, D. A., Lestari, T. A., & Harahap, L. S. (2025). Implementasi Algoritma Convolutional Neural Network (CNN) Untuk Pengenalan Angka Tulisan Tangan. 5(2), 93–100.
- Teori, B., Thomas, P., Budiawan, R., Ichwani, A., Munir, R., & Mahayana, D. (2023). Pergeseran Paradigma pada Penelitian Pengenalan Tulisan Tangan. 6(2), 170–179.
- Zhang, A., Park, M., Taylor, M. E., Pineau, J., Jun, L. G., Calandra, R., & Park, M. (n.d.). Model-Invariant State Abstractions for Model-Based Reinforcement Learning. 1–20.