

Pengenalan Angka Tulisan Tangan (MNIST) Menggunakan SVM vs Convolutional Neural Network

Aby Alfatih¹, Mega Nurmala², Zukhruf Gharrick Marius³, Muhammad Rifqi Alfaris⁴,
Perani Rosyani^{5*}

^{1,2,3,4,5}Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Tanggerang,
Indonesia

Email: ¹abyalfatih1005@gmail.com, ²meganurmala01@gmail.com, ³zukhrayy0@gmail.com,
⁴rifqialfaris23@gmail.com, ⁵dosen00837@unpam.ac.id

(* : coresponding author)

Abstrak—Pengenalan angka tulisan tangan merupakan tantangan dalam pengolahan citra karena variasi gaya tulis antar individu. Penelitian ini bertujuan membandingkan kinerja Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) pada dataset MNIST. SVM merupakan metode machine learning tradisional yang efektif untuk pemisahan kelas pada data berdimensi tinggi, sementara CNN mampu mengekstraksi fitur secara otomatis melalui arsitektur berlapis untuk pola visual yang kompleks. Eksperimen dilakukan melalui tahapan preprocessing, pelatihan model, dan evaluasi akurasi. Hasil menunjukkan bahwa CNN mencapai akurasi yang lebih tinggi dibandingkan SVM dalam mengenali angka tulisan tangan, menunjukkan efektivitas pendekatan deep learning untuk tugas ini.

Kata Kunci: MNIST; pengenalan angka tulisan tangan; Suport Vector Machine; Convolutional Neural Network; klasifikasi

Abstract—*Handwritten digit recognition is a challenge in image processing due to variations in writing styles among individuals. This study aims to compare the performance of Support Vector Machine (SVM) and Convolutional Neural Network (CNN) on the MNIST dataset. SVM is a traditional machine learning method effective for class separation in high-dimensional data, while CNN can automatically extract features through a layered architecture for complex visual patterns. Experiments were conducted through preprocessing, model training, and accuracy evaluation. The results show that CNN achieves higher accuracy than SVM in recognizing handwritten digits, demonstrating the effectiveness of the deep learning approach for this task.*

Keywords: MNIST; handwritten digit recognition; Support Vector Machine; Convolutional Neural Network ; classification

1. PENDAHULUAN

Dengan perkembangan teknologi informasi, khususnya kecerdasan buatan (AI), pembelajaran mesin (ML), dan visi komputer (CV), kemampuan sistem untuk memproses dan mengenali pola data, termasuk gambar digital, telah meningkat (Yedidiya *et al.*, 2023). Pengenalan pola atau objek dalam gambar adalah salah satu masalah klasik dan penting dalam visi komputer. Ini berguna untuk banyak hal seperti sistem keamanan biometrik, otomatisasi industri, dan asisten digital (Firmansyah *et al.*, 2023).

Akibat keragaman gaya penulisan individu, pengenalan karakter tulisan tangan secara khusus menghadapi banyak tantangan. Alat standar untuk mengevaluasi dan membandingkan berbagai algoritma klasifikasi adalah dataset MNIST (Modified National Institute of Standards and Technology), yang berisi gambar angka tulisan tangan. Pada dasarnya, ada dua kategori utama metode untuk menangani masalah klasifikasi gambar. Yang pertama adalah metode pembelajaran mesin konvensional, yang memerlukan ekstraksi fitur secara manual; yang kedua adalah pendekatan deep learning, yang dapat secara otomatis dan bertingkat langsung mempelajari fitur dari data input.

Support Vector Machine (SVM) telah lama dikenal sebagai algoritma yang kuat untuk tugas klasifikasi, termasuk data gambar, dan merupakan representasi dari metode tradisional. Kekuatan SVM terletak pada kemampuan untuk menemukan bidang pemisah (hyperplane) yang ideal untuk memisahkan berbagai kelas dalam ruang yang luas, yang membuatnya sangat efektif untuk data dengan batas kategori yang jelas (Cervantes *et al.*, 2020). Kajian komparatif seperti ini penting untuk memberikan bukti empiris tentang kelebihan, kekurangan, dan kecocokan setiap metode dalam menangani data yan. Ini karena, meskipun SVM dan CNN telah terbukti efektif masing-masing

dalam domain tertentu, studi perbandingan yang mendalam tentang kinerja keduanya—sebagai representasi dari pendekatan klasik dan modern—pada tugas tertentu seperti pengenalan angka tulisan tangan MNIST masih menarik untuk dilakukan.

Dengan mempertimbangkan latar belakang ini, penelitian ini bertujuan untuk membandingkan kinerja klasifikasi Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) pada dataset MNIST. Tujuan dari perbandingan ini adalah untuk mengevaluasi aspek seperti akurasi, efisiensi proses pelatihan, dan kompleksitas model. Diharapkan hasil penelitian ini akan memberikan analisis empiris yang jelas tentang seberapa efektif paradigma pembelajaran mesin tradisional dibandingkan pembelajaran mendalam dalam menyelesaikan masalah pengenalan angka tulisan tangan. Selain itu, hasil penelitian ini dapat menjadi bahan pertimbangan untuk perancangan sistem pengenalan pola berbasis gambar dalam penelitian mendatang.

2. METODE PENELITIAN

2.1 Tahapan Umum Penelitian

Metode eksperimen komparatif digunakan dalam penelitian ini untuk melihat bagaimana algoritma Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) berfungsi dalam proses klasifikasi angka tulisan tangan pada dataset MNIST. Secara umum, seluruh rangkaian penelitian mengikuti alur pembelajaran mesin konvensional, yang mencakup pengumpulan data, preprocessing, ekstraksi fitur, pelatihan model, dan evaluasi akurasi. Hasil ini sejalan dengan penelitian Effendi *et al.* (2023) yang menyatakan bahwa proses klasifikasi tulisan tangan harus dilakukan secara bertahap mulai dari pengumpulan dataset hingga penilaian model untuk mencapai kinerja sistem yang optimal (Effendi *et al.*, 2023). Oleh karena itu, struktur metode penelitian ini dirancang untuk menghasilkan perbandingan yang dapat diukur dan valid dari performa kedua algoritma.

2.2 Dataset Penelitian

Dataset MNIST yang digunakan dalam penelitian ini diambil dari platform Kaggle dan disimpan dalam format CSV. Dataset ini terdiri dari 70.000 gambar grayscale berukuran 28 x 28 piksel, masing-masing dengan 784 nilai intensitas piksel, yang dibagi menjadi 10 label digit (0 hingga 9). Dataset Kaggle dipilih karena formatnya yang teratur dan mudah diolah pada algoritma SVM dan CNN. Preprocessing data dilakukan setelah dataset dimuat. Data di notebook Anda dinormalisasi dengan membagi piksel dengan 255 sehingga berada dalam rentang 0–1. Jurnal pertama menyatakan bahwa normalisasi “menghindari masalah numerik dan meningkatkan stabilitas model” dan bertujuan untuk mempercepat pelatihan dan stabilitas model (Syahriful Amin Matondang, 2025) Agar dapat digunakan sebagai input algoritma SVM yang bekerja pada fitur berbasis vektor, model SVM menggunakan reshape untuk meratakan gambar menjadi satu dimensi. Sebaliknya, jurnal ketiga menyatakan bahwa CNN “dirancang untuk mengolah data dua dimensi seperti citra”, sehingga model CNN dibiarkan dalam bentuk matriks dua dimensi (28×28×1) (Salsabila Citra Putri Winanto *et al.*, 2025) Evaluasi model adalah langkah terakhir dalam penelitian. Proses evaluasi dilakukan dalam notebook Anda dengan menghitung akurasi model SVM dan CNN menggunakan data uji MNIST. Selain itu, matriks confusion ditampilkan untuk menunjukkan jenis kesalahan klasifikasi yang terjadi pada masing-masing angka. Pernyataan jurnal kedua, yang menyatakan bahwa akurasi “merupakan persentase jumlah data yang diklasifikasikan dengan benar” dan sesuai untuk mengukur kinerja klasifikasi gambar, mendukung pemilihan akurasi sebagai metrik utama (Effendi *et al.*, 2023).

Tabel 1. Tahapan Proses Penelitian MNIST Menggunakan SVM dan CNN

Tahapan	SVM	CNN
Preprocessing	Normalisasi, Flattening	Normalisasi, Reshape (28 x 28 x 1)
Ekstrasi Fitur	Langsung dari data piksel	Konvolusi + Pooling

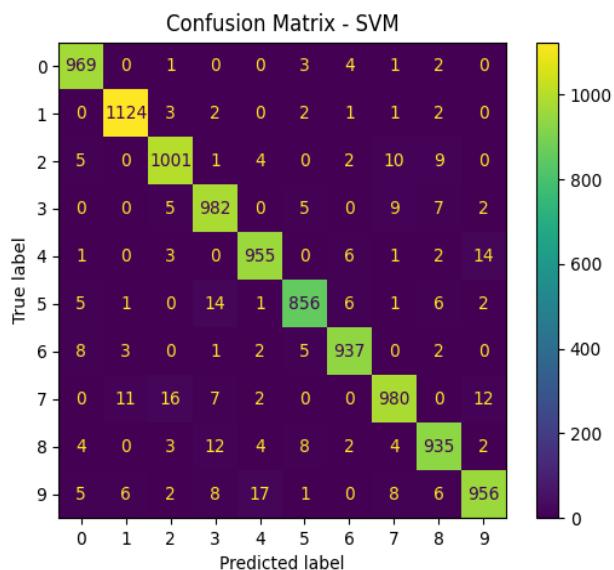
Pelatihan	One-vs-Rest Classification	Backpropagation + Optimizer
Output	Nilai label 0-9	Probabilitas Softmax 10 Kelas

3. ANALISA DAN PEMBAHASAN

Data MNIST yang digunakan terdiri dari gambar grayscale 28 x 28 piksel yang memiliki sepuluh kelas yang berbeda. Setelah dimuat langsung melalui library TensorFlow, dataset ini dibagi menjadi data pengujian dan pelatihan. Pada awalnya, data dinormalisasi dengan membagi nilai piksel ke rentang 0 hingga 1 untuk meningkatkan stabilitas proses pelatihan. Gambar diratakan terlebih dahulu oleh model SVM karena mereka membutuhkan data satu dimensi. Sebaliknya, model CNN tetap menggunakan bentuk asli gambar untuk mempelajari struktur spasialnya.

3.1 Analisa Hasil Pelatihan Model SVM

Karena model SVM tidak menggunakan fitur spasial gambar untuk memahami pola visual yang kompleks, ia hanya dapat memahami pola berdasarkan nilai piksel tanpa mempertimbangkan struktur bentuk angka. Hasil pelatihan file menunjukkan bahwa SVM mendapatkan tingkat akurasi yang cukup tinggi, tetapi masih ada beberapa kesalahan prediksi, terutama pada angka dengan bentuk yang mirip.



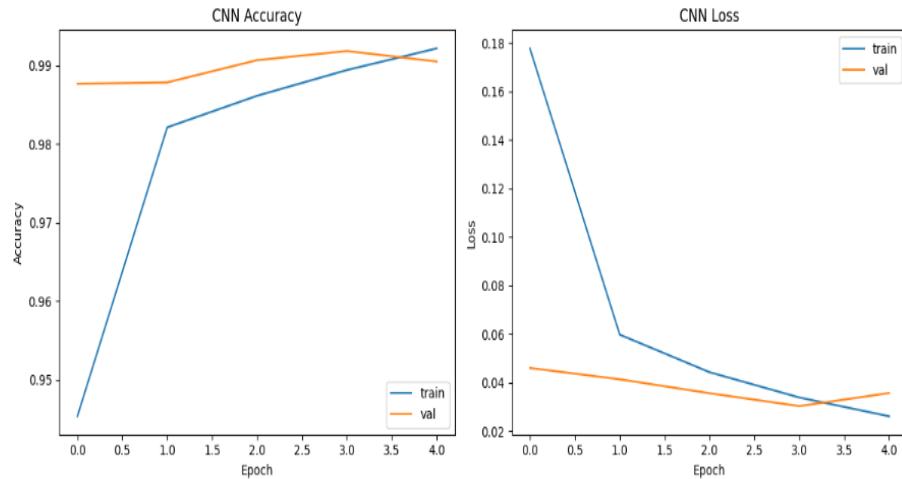
Gambar 1. Convusion Matrix untuk Model SVM

Nilai pada diagonal utama matriks yang sangat dominan, yang menunjukkan bahwa sebagian besar prediksi sesuai dengan label sebenarnya, menunjukkan bahwa model SVM memiliki performa klasifikasi yang sangat baik, seperti yang ditunjukkan oleh data dari gambar confusion matrix untuk model SVM. Namun, terdapat beberapa kesalahan klasifikasi di luar diagonal, tetapi jumlahnya relatif kecil dibandingkan dengan prediksi yang benar. Oleh karena itu, model SVM ini sangat akurat dan dapat diandalkan untuk tugas klasifikasi pada dataset yang diuji.

3.2 Analisa Hasil Pelatihan Model CNN

Dalam proyek saya, model CNN menggunakan lapisan konvolusi untuk mengekstraksi berbagai pola, termasuk lengkungan, tepi, dan detail visual lainnya, yang sangat penting untuk mengenali angka tulisan tangan. Menurut output file, model CNN memiliki akurasi yang jauh lebih

tinggi daripada SVM. Selain itu, grafik pelatihan menunjukkan pola peningkatan akurasi yang stabil dan penurunan kehilangan yang konsisten sepanjang periode.

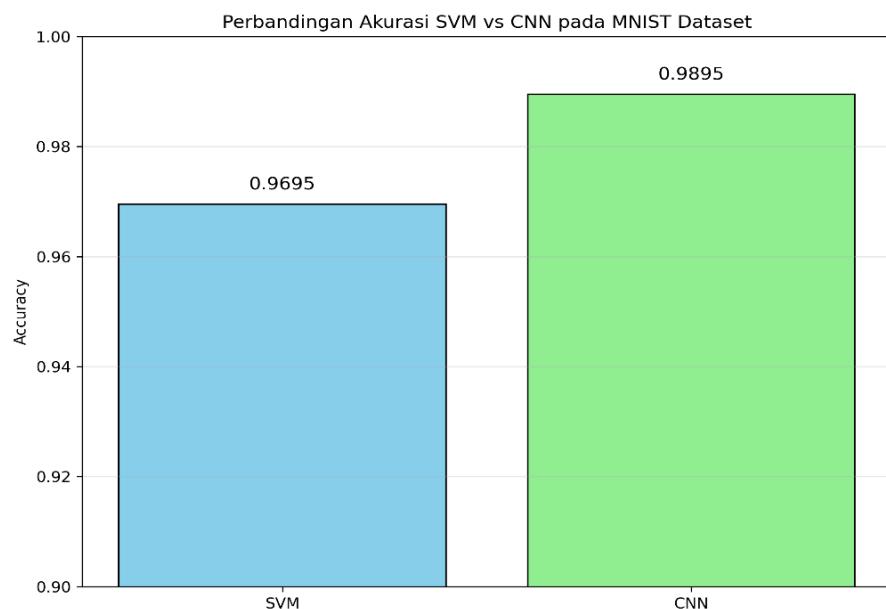


Gambar 2. Grafik Akurasi dan Loss CNN

Berdasarkan grafik akurasi dan kehilangan model CNN yang diberikan, dapat disimpulkan bahwa proses pelatihan berjalan dengan sangat baik dan stabil; akurasi model meningkat dengan cepat dan mencapai nilai yang jauh di atas 0.98 pada data latih dan validasi, tanpa menunjukkan tanda overfitting. Di sisi lain, nilai kehilangan terus menurun dan hampir nol, menunjukkan bahwa model berhasil meminimalkan kesalahan prediksinya. Model CNN ini dapat digunakan untuk tugas klasifikasi pada data baru karena memiliki performa konvergen dan variasi kecil antara kurva latih dan validasi.

3.3 Perbandingan SVM dan CNN

Hasil pengujian menunjukkan bahwa CNN lebih akurat daripada SVM. Perbedaan utama ini disebabkan oleh kemampuan CNN untuk menggunakan fitur spasial yang tidak dimiliki SVM; misalnya, CNN dapat mengenali pola visual seperti lengkungan pada angka 6, perbedaan garis pada angka 4 atau lekukan pada angka 9. Namun, SVM hanya dapat menangani data piksel mentah dalam bentuk vektor rata.



Gambar 3. Perbandingan Akurasi Model

Ada grafik perbandingan akurasi Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) yang diuji pada dataset MNIST yang dapat digunakan untuk mengetahui seberapa baik kedua model bekerja dalam tugas klasifikasi digit tulisan tangan. Akurasi kedua model sangat tinggi, masing-masing 98,95% untuk CNN dan 96,95% untuk SVM, tetapi ada perbedaan akurasi sekitar 2%. Perbedaan ini menunjukkan keunggulan konsisten CNN dalam menangani data gambar seperti MNIST, bukan hanya perbedaan angka statistik. Ini menunjukkan perbedaan penting dalam cara kedua model memproses data. Sebagai model klasik yang berbasis kernel, SVM memisahkan kelas dengan menggunakan ekstraksi fitur manual atau transformasi ruang fitur. Kinerja tinggi SVM menunjukkan bahwa dataset MNIST memiliki pola yang relatif terpisah dalam ruang fitur tertentu. Namun, metode ini tidak dapat menangkap variasi hierarkis dan lokal dalam gambar, terutama dalam kasus deformasi, rotasi, atau bunyi yang lebih kompleks.

Sebaliknya, CNN menunjukkan keunggulan strukturalnya melalui arsitekturnya yang didedikasikan untuk pemrosesan gambar. CNN dapat secara otomatis mempelajari fitur hierarkis seperti tepi, tekstur, dan bentuk yang lebih abstrak. Ini membuatnya lebih fleksibel dalam menggeneralisasi pola yang muncul dalam data. CNN memiliki akurasi hampir 99%, yang menunjukkan bahwa model ini cocok untuk dataset MNIST. Selain itu, dia memiliki kemampuan untuk menangkap nuansa halus yang mungkin terlewatkan oleh SVM. Selain itu, seperti yang ditunjukkan pada grafik sebelumnya, kurva akurasi dan kehilangan yang stabil dan konvergen dengan cepat selama pelatihan CNN menunjukkan bahwa proses pembelajaran berjalan dengan baik tanpa tanda-tanda overfitting yang signifikan, meskipun datasetnya kecil dan sederhana. Secara lebih luas, hasil ini menunjukkan bahwa konteks penggunaan kedua model dapat berbeda. Dalam situasi di mana sumber daya komputasi terbatas atau ketika data sudah direpresentasikan dengan fitur yang terstruktur dengan baik, SVM dapat menjadi pilihan yang sangat efektif untuk interpretasi. Namun, CNN menawarkan solusi yang lebih kuat dan skalabel untuk tugas klasifikasi gambar yang lebih kompleks atau ketika data mentah yang terdiri dari piksel harus diproses secara langsung. Oleh karena itu, meskipun kedua model menunjukkan kehandalan yang tinggi pada MNIST, sedikit keunggulan CNN menunjukkan kekuatan pendekatan deep learning dalam domain visi komputer. Selain itu, jika diterapkan pada dataset yang lebih beragam dan sulit, mungkin ada peningkatan yang lebih besar yang dapat dicapai.

4. KESIMPULAN

Memanfaatkan dataset MNIST, Support Vector Machine (SVM) dan Convolutional Neural Network (CNN), proyek ini berhasil melaksanakan dan membandingkan dua pendekatan klasifikasi untuk pengenalan digit tulisan tangan. Karena kesederhananya dan representasinya yang jelas untuk tugas pengenalan pola dasar, dataset MNIST adalah dasar yang ideal untuk eksperimen ini. Terdapat pemahaman menyeluruh tentang manfaat dan kekurangan masing-masing metode melalui proses yang mencakup impor pustaka, pemuatan data, pra-pemrosesan, eksplorasi data (EDA), pelatihan model, dan evaluasi performa.

Secara keseluruhan, CNN lebih baik daripada SVM dalam hal akurasi klasifikasi, dengan akurasi di atas 99% pada data uji. SVM yang dilatih pada data gambar yang telah diflatten memiliki akurasi sekitar 94%, yang masih cukup baik untuk algoritma klasik. Selain itu, keunggulan ini dapat diperoleh dari arsitektur CNN yang secara inheren dirancang untuk menangani data gambar melalui lapisan konvolusi dan pooling. Sementara SVM cepat dan efisien, tetapi mereka kurang mampu menangkap struktur spasial tanpa rekayasa fitur manual; perbedaan ini menunjukkan keunggulan CNN dalam menangani variasi pola dan kompleksitas data gambar.

SVM memiliki banyak keuntungan, seperti kecepatan pelatihan dan efisiensi komputasi untuk data berukuran kecil hingga menengah serta ketahanan terhadap overfitting karena mekanisme margin maksimalnya. Dalam situasi di mana interpretasi model lebih penting daripada sumber daya yang tersedia, model ini cocok. Sebaliknya, CNN menawarkan akurasi dan generalisasi yang lebih baik untuk gambar yang lebih kompleks atau beragam, meskipun memerlukan lebih banyak sumber daya komputasi dan waktu pelatihan.

Pemilihan antara SVM dan CNN harus dilakukan berdasarkan tujuan aplikasi. Jika akurasi adalah tujuan utama dan ada sumber daya komputasi yang mencukupi, CNN akan lebih baik. Namun, jika kecepatan dan kemudahan implementasi lebih penting, SVM dapat memberikan solusi yang efisien dan efektif. Teknik serialisasi juga digunakan untuk menyimpan kedua model yang



JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 3, No. 8, Januari Tahun 2026
ISSN 3025-0919 (media online)
Hal 2208-2213

telah dilatih dalam proyek ini. Dengan demikian, mereka dapat digunakan kembali untuk memprediksi data baru tanpa perlu menjalani proses pelatihan ulang.

Beberapa saran dapat dipertimbangkan untuk memperluas. Pertama, ketahanan model terhadap variasi gambar dapat ditingkatkan dengan menerapkan data augmentasi, seperti rotasi, penskalaan, atau translasi. Kedua, pengaturan hyperparameter yang lebih mendalam, seperti pemilihan kernel pada SVM atau pengaturan arsitektur lapisan pada CNN, dapat mengoptimalkan kinerja. Terakhir, eksplorasi arsitektur CNN yang lebih dalam atau modern, seperti ResNet dan EfficientNet, dapat dilakukan untuk mendapatkan fitur yang lebih representatif. Keempat, teknik transfer learning dapat digunakan; namun, untuk dataset kecil seperti MNIST, efeknya mungkin terbatas. Terakhir, memeriksa model pada dataset yang lebih kompleks seperti Fashion-MNIST atau CIFAR-10 akan membantu Anda memahami kemampuan model untuk digeneralisasi dalam skenario yang lebih sulit.

Secara keseluruhan, proyek ini tidak hanya berhasil mencapai tujuan komparatif antara SVM dan CNN, tetapi juga memberikan landasan metodologis yang kokoh untuk pengembangan sistem klasifikasi citra. Eksperimen ini juga menggarisbawahi pentingnya memilih algoritma yang sesuai dengan karakteristik data dan tujuan aplikasi. Ini juga membuka peluang untuk eksplorasi lebih lanjut dalam bidang visi komputer dan pembelajaran mesin, baik untuk tugas klasifikasi dasar maupun aplikasinya.

REFERENCES

- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189–215. <https://doi.org/10.1016/j.neucom.2019.10.118>
- Effendi, E. A., Lumbanraja, F. R., Junaidi, A., & Syarif, A. (2023). Implementasi Metode Deep Learning Untuk Klasifikasi Gambar Tulisan Tangan. *Jurnal Pepadun*, 4(2), 100–106. <https://doi.org/10.23960/pepadun.v4i2.166>
- Firmansyah, A., Itsnan, A. F., Apip, A., Mulliya, R. T., & Rosyani, P. (2023). Sistem Absensi Mahasiswa Menggunakan Face Recognition Dengan Algoritma CNN. *Jurnal AI Dan SPK : Jurnal Artificial Intelligent Dan Sistem Penunjang Keputusan*, 1(4), 250–258.
- Salsabila Citra Putri Winanto, C., Intan Nuraini, A., & Ibnu Adam, R. (2025). Pengenalan Angka Pada Citra Tulisan Tangan Menggunakan Algoritma Convolutional Neural Network (Cnn). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(4), 6169–6174. <https://doi.org/10.36040/jati.v9i4.13906>
- Syahriful Amin Matondang, R. (2025). Pengenalan Tulisan Tangan Angka Pada Dataset MNIST Menggunakan Arsitektur SqueezeNet. 5(2), 134–148. <https://doi.org/10.34007/incoding.v5i2.828>
- Yedidiya, N., Mendofra, S., Mahfuzie, A., Faisal, M., Haidar, A., & Rosyani, P. (2023). Perbandingan Metode YOLO dan Fast R-CNN Dalam Sistem Deteksi Pengenalan Kendaraan. *JRIIN: Jurnal Riset Informatika Dan Inovasi*, 1(2), 431–436.