

“LAPORAN PROYEK CNN: KLASIFIKASI ANGKA TULISAN TANGAN”

Kelompok 6:

Ghulam Mushthofa

Muhammad Rafi Aditya

Hasbi Abdullah

Rafii Abdurahman

Muslih hadi

Universitas Darussalam Gontor

May 20, 2025

PENDAHULUAN

Angka tulisan tangan merupakan bentuk input visual yang umum digunakan dalam pengujian algoritma pengenalan pola dan computer vision. Dataset MNIST adalah benchmark populer dalam machine learning yang terdiri dari gambar grayscale 28x28 piksel, masing-masing mewakili angka 0 sampai 9. Dalam proyek ini, kami mengembangkan model Convolutional Neural Network (CNN) untuk mengklasifikasikan angka-angka tersebut dengan akurasi tinggi menggunakan PyTorch.

LANDASAN TEORI

1. Convolutional Neural Network (CNN)

CNN adalah arsitektur deep learning yang dirancang khusus untuk pengolahan data grid seperti gambar. CNN menggunakan lapisan konvolusi (convolutional layers) yang mengekstrak fitur spasial dari gambar, lapisan pooling untuk mengurangi dimensi, dan lapisan fully connected untuk klasifikasi. CNN sangat efektif dalam menangani gambar karena kemampuannya mengenali pola spasial.

2. MNIST Dataset

Dataset MNIST berisi 60.000 data latih dan 10.000 data uji berupa gambar angka tulisan tangan. Dataset ini telah diproses sebelumnya dan menjadi standar benchmark untuk algoritma klasifikasi.

3. Fungsi Aktivasi dan Loss Function

- Fungsi aktivasi ReLU digunakan untuk menambah non-linearitas.
- Fungsi loss yang digunakan adalah CrossEntropyLoss, cocok untuk tugas klasifikasi multikelas.

METODOLOGI

1. Pengumpulan & Persiapan Data

- Dataset MNIST dimuat dari file binary standar (train-images-idx3-ubyte, train-labels-idx1-ubyte, dst.).
- Gambar dikonversi ke tensor dan dinormalisasi.
- Data dibagi ke dalam batch menggunakan DataLoader.

2. Arsitektur Model CNN

Model didefinisikan dalam model.py dengan struktur sebagai berikut:

Layer	Tipe	Ukuran	Aktivasi
Input	Conv2D	$1 \times 28 \times 28 \rightarrow 16 \times 26 \times 26$	ReLU
Hidden	Conv2D + MaxPool	$16 \times 26 \times 26 \rightarrow 32 \times 12 \times 12$	ReLU
Fully Connected Linear		$32 \times 12 \times 12 \rightarrow 128 \rightarrow 10$	ReLU + Softmax (implicit)

- Optimizer: Adam (learning rate 0.001)
- Loss Function: CrossEntropyLoss
- Epoch: 10 (dapat disesuaikan)
- Implementasi pelatihan dilakukan dalam train.py.

3. Training

- Model dilatih menggunakan DataLoader dalam loop epoch.
- Loss dicatat per batch.
- Parameter disimpan setelah pelatihan (model_cnn.pth).

4. Prediksi

- Gambar baru dapat diuji menggunakan prediksi_gambar.py.
- Input berupa path gambar, hasil dikonversi dan diprediksi oleh model.

HASIL DAN ANALISIS

Visualisasi Loss

- Selama pelatihan, nilai loss mengalami penurunan signifikan hingga stabil di bawah ambang batas tertentu, menunjukkan konvergensi yang baik.
- Grafik pelatihan dapat ditambahkan untuk memvisualisasi tren ini (tidak tersedia di kode saat ini, tapi bisa ditambahkan dengan matplotlib).

Akurasi Model

- Berdasarkan hasil pengujian (print output train.py), akurasi model CNN mencapai **98%+** pada dataset MNIST.
- Model cukup robust terhadap data uji dan hanya keliru pada beberapa prediksi angka yang bentuknya ambigu.

Analisis

- CNN unggul dalam menangani data gambar dibanding metode klasik karena kemampuan ekstraksi fitur otomatis.
- MNIST menjadi dataset yang ideal untuk eksplorasi CNN awal karena kompleksitasnya sedang dan hasilnya representatif.
- Kemampuan generalisasi model bisa ditingkatkan dengan data augmentasi atau tuning hyperparameter seperti jumlah filter, ukuran kernel, dan jumlah epoch.