

Deep Learning Image Classification Technical Report

1103200032 – Fajra Risqulla – TK-44-06

1.1 Abstrak

Tujuan dari laporan teknis ini adalah untuk mengevaluasi kemandirian jaringan saraf konvolusional (CNN) untuk tugas klasifikasi gambar menggunakan dataset MNIST dan FashionMNIST. Studi ini mengeksplorasi arsitektur model, strategi pelatihan, dan mengevaluasi kinerja mereka dalam mengklasifikasikan digit tulisan tangan dan item fashion, masing-masing. Metrik evaluasi dan teknik visualisasi digunakan untuk menilai akurasi model dan mengidentifikasi area untuk perbaikan.

1.2 Pendahuluan

Klasifikasi gambar memainkan peran penting dalam aplikasi visi komputer, yang mencakup tugas-tugas seperti pengenalan objek, klasifikasi digit, dan mengidentifikasi item mode. Himpunan data MNIST terdiri dari gambar skala abu-abu 28x28 dari digit tulisan tangan (0-9), berfungsi sebagai himpunan data tolok ukur untuk pengenalan digit. Dataset FashionMNIST terdiri dari gambar skala abu-abu dari item fashion, seperti T-shirt, celana panjang, gaun, dan sepatu, memberikan tugas klasifikasi yang lebih kompleks.

Laporan ini bertujuan untuk menilai efisiensi arsitektur CNN dalam mengklasifikasikan gambar secara akurat dari kedua dataset. Studi ini mengeksplorasi kinerja model, mengidentifikasi tantangan, dan menyarankan perbaikan untuk meningkatkan akurasi klasifikasi.

1.3 Metodologi

- **Model Architecture and Training**
Penelitian ini menggunakan arsitektur CNN yang terdiri dari lapisan konvolusional, fungsi aktivasi, lapisan pooling, dan lapisan klasifikasi. Untuk himpunan data MNIST, model dilatih menggunakan pengoptimal penurunan gradien stokastik (SGD) dengan fungsi kerugian entropi silang. Langkah-langkah preprocessing meliputi normalisasi dan augmentasi untuk meningkatkan generalisasi model.
- **Evaluation Metrics**
Kinerja model dinilai menggunakan metrik seperti akurasi dan matriks kebingungan. Teknik visualisasi digunakan untuk membandingkan prediksi model dengan label sebenarnya, memberikan wawasan tentang akurasi klasifikasi model untuk kelas yang berbeda.

1.4 Hasil

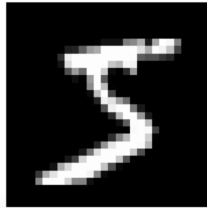
- **MNIST Dataset**
Model CNN terlatih mencapai akurasi 98% pada set tes MNIST. Visualisasi prediksi model menunjukkan akurasi tinggi dalam mengklasifikasikan sebagian besar digit, dengan kebingungan sesekali antara digit yang mirip secara visual seperti 4 dan 9. Akurasi tinggi yang dicapai pada dataset MNIST menunjukkan efektivitas CNN dalam pengenalan digit. Namun, kebingungan sesekali antara digit visual yang mirip menunjukkan area untuk perbaikan lebih lanjut, mungkin melalui peningkatan arsitektur atau teknik augmentasi data.

6. Visualize at least 5 different samples of the MNIST training dataset.

```
In [10]: import matplotlib.pyplot as plt
for i in range(5):
    img = train_data[i][0]
    print(img.shape)
    img_squeeze = img.squeeze()
    print(img_squeeze.shape)
    label = train_data[i][1]
    plt.figure(figsize=(3, 3))
    plt.imshow(img_squeeze, cmap="gray")
    plt.title(label)
    plt.axis(False)

torch.Size([1, 28, 28])
torch.Size([28, 28])
torch.Size([1, 28, 28])
torch.Size([28, 28])
torch.Size([1, 28, 28])
torch.Size([28, 28])
torch.Size([1, 28, 28])
torch.Size([28, 28])
torch.Size([1, 28, 28])
torch.Size([28, 28])
```

5



- FashionMNIST Dataset

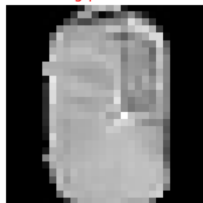
Evaluasi model CNN pada dataset FashionMNIST menghasilkan akurasi 86%. Namun, model menghadapi tantangan dalam membedakan antara item mode yang mirip secara visual, seperti T-shirt dan kemeja atau sandal dan sneakers. Tantangan dalam mengklasifikasikan item mode yang mirip secara visual mengungkapkan keterbatasan dalam pelabelan dataset dan arsitektur model. Ambiguitas dalam pelabelan, seperti membedakan antara T-shirt dan kemeja, menimbulkan kesulitan bagi model, menunjukkan perlunya label yang lebih berbeda atau peningkatan model.

```
In [42]: # Select random 9 wrong predictions and plot them
import random
random_selection = random.sample(list(wrong_pred_indexes), k=9)

plt.figure(figsize=(10, 10))
for i, idx in enumerate(random_selection):
    # Get true and pred labels
    true_label = fashion_mnist_class_names[fashion_mnist_test[idx][1]]
    pred_label = fashion_mnist_class_names[test_preds[idx]]

    # Plot the wrong prediction with its original label
    plt.subplot(3, 3, i+1)
    plt.imshow(fashion_mnist_test[idx][0].squeeze(), cmap="gray")
    plt.title(f'True: {true_label} | Pred: {pred_label}', c='r')
    plt.axis(False)
```

True: Bag | Pred: Pullover



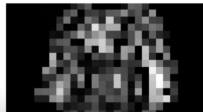
True: Coat | Pred: Shirt



True: Coat | Pred: Shirt



True: Shirt | Pred: Dress



True: T-shirt/top | Pred: Pullover



True: Shirt | Pred: T-shirt/top



1.5 Kesimpulan

Sebagai kesimpulan, penelitian ini menunjukkan kemandirian CNN dalam tugas klasifikasi gambar menggunakan dataset MNIST dan FashionMNIST. Sementara mencapai akurasi

tinggi pada pengenalan digit dalam dataset MNIST, tantangan tetap ada dalam membedakan item fashion yang mirip secara visual dalam dataset FashionMNIST. Pekerjaan di masa depan melibatkan penyempurnaan arsitektur model dan mengatasi ambiguitas dalam pelabelan dataset untuk meningkatkan akurasi klasifikasi.