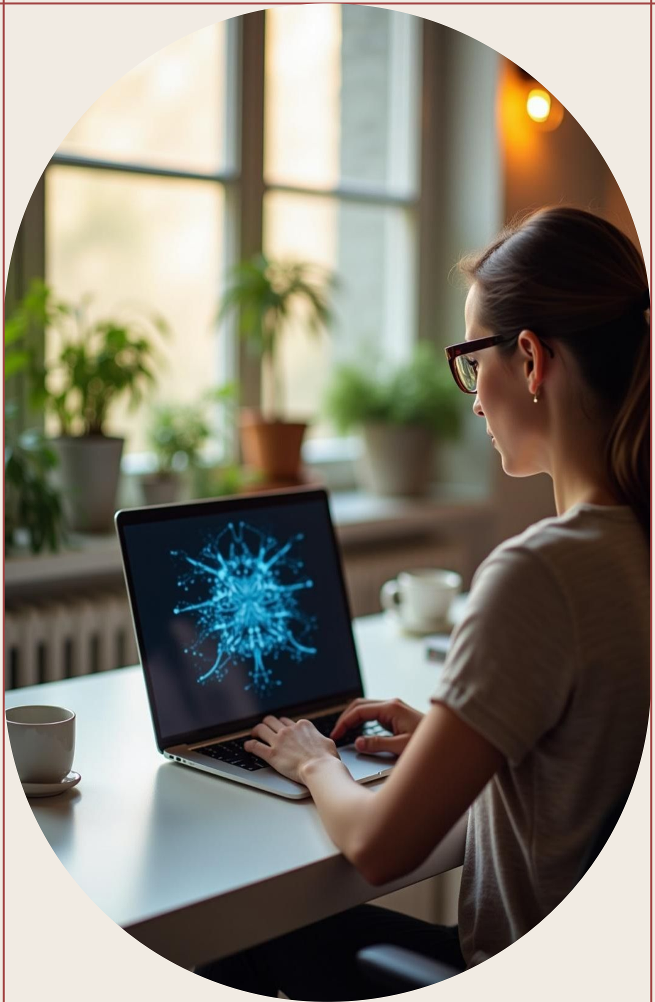
Traffic Sign Recognition with Deep Learning

Muhammad Mustafa Qaiser (2220058)
Deep Learning WS 2025

Introduction

- Traffic sign recognition is essential for autonomous driving and advanced systems.
- This project implements a compact CNN model to classify traffic signs with high accuracy.
- Our aim is to demonstrate practical deep learning applications for traffic sign understanding in real-time scenarios.



Dataset Description and Preprocessing


```
traffic_sign_small/
    labels.csv
    traffic_Data/
        DATA/
            0/
            1/
            2/
            ...
```

- The dataset consists of over 50 traffic sign classes, each represented by 32×32 RGB images.
- Data variation includes lighting, blur, and perspective challenges.
- Images are normalized and split into 80% training and 20% validation sets to ensure reliable performance evaluation.

Overview of Convolutional Neural Networks (CNNs)

- **Definition:** A class of deep neural networks specifically designed for processing grid-like data such as images.
- **The "Flashlight" Analogy:** A kernel scans an image like a flashlight in a dark room, noting specific features at each spot to create a "feature map".
- **Spatial Patterns:** CNNs exploit the spatial structure of images to detect local patterns like edges, textures, and shapes.
- **Feature Hierarchy:** Early layers detect simple lines/edges, while deeper layers combine these into complex objects.

Mathematics of CNN

Discrete Convolution: The primary operation involves a learnable kernel W sliding over an input tensor X

Forward Pass Equation:

$$Y_{i,j} = \sigma\left(\sum_{m=0}^{H-1}\sum_{n=0}^{W-1} X_{i+m,j+n} \cdot W_{m,n} + b\right)$$

Activation Function: The Rectified Linear Unit (ReLU), introduces sparsity for deep representation learning.

$$\sigma(x) = \max(0, x),$$

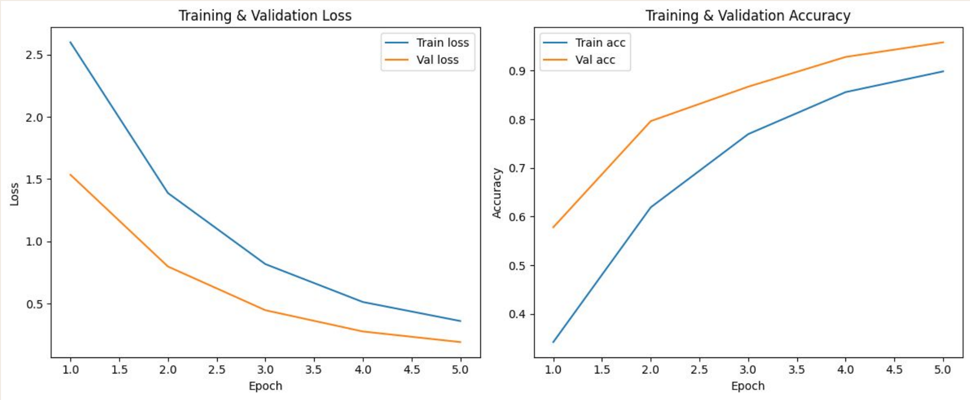
Dimensionality: Output size O is determined by input size N, kernel size K, padding P, and stride S:

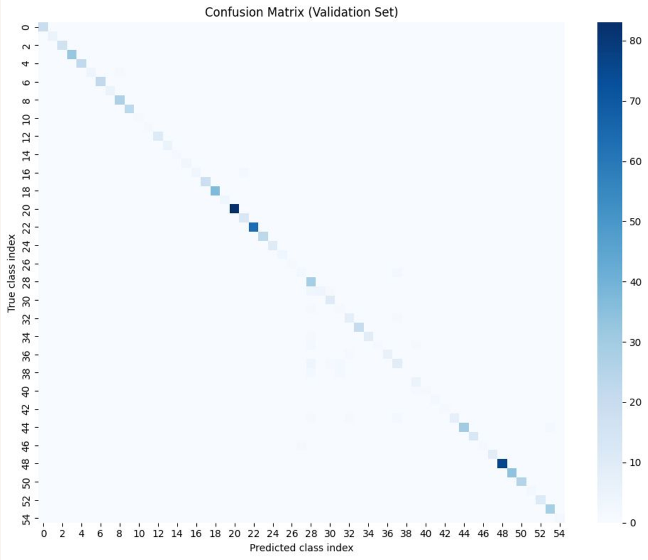
$$O = \lfloor \frac{N-K+2P}{S} \rfloor + 1$$

Implementation Model Architecture

Convolutional Layer: uses small learnable kernels to detect low-level features through a multiplication and summation process
Activation Layer (ReLU): introduces the non-linearity and sparsity necessary for deep representation learning.
Batch Normalization Layer: used to stabilize the training process
Pooling Layer (Max Pooling): reduce computational complexity & introduce translation invariance
Flatten Layer: transforms the feature maps into a single vector to prepare the data for the classification stages.
Fully Connected (Linear) Layer: act as the final classifier, mapping the extracted features to the specific class labels
Dropout Layer: randomly "drops" neurons during training to prevent overfitting and improve generalization.

Results





- The model achieves high accuracy with tightly correlated training and validation metrics
- The confusion matrix reveals accurate classification with only a few confusions for visually similar signs.
- Inference takes approximately 4.51 milliseconds per

Predicted Images

Conclusions

- The compact CNN model effectively classifies traffic sign images
- Limitations include sensitivity to data diversity and lighting changes

THANK YOU

Do you have any questions?

Citations

I. Čordašić, M. Vranješ, R. Grbić and M. Herceg, "Algorithm for Extracting Signposts from the Scene and Understanding Text on Them," 2024 Zooming Innovation in Consumer Technologies Conference (ZINC), Novi Sad, Serbia, 2024, pp. 60-65, doi: 10.1109/ZINC61849.2024.10579407.

A. Biswas, A. R. Chowdhury and R. Selvanambi, "Lane detection and traffic sign detection for autonomous driving systems," 2nd International Conference on Computer Vision and Internet of Things (ICCVIoT 2024), Coimbatore, India, 2024, pp. 206-211, doi: 10.1049/icp.2024.4425.

S. Khalid, J. H. Shah, M. Sharif, F. Dahan, R. Saleem and A. Masood, "A Robust Intelligent System for Text-Based Traffic Signs Detection and Recognition in Challenging Weather Conditions," in IEEE Access, vol. 12, pp. 78261-78274, 2024, doi: 10.1109/ACCESS.2024.3401044.

Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.

G. Cui, "Research on Recognition and Classification Technology Based on Deep Convolutional Neural Network," 2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 2021, pp. 353-357, doi: 10.1109/ECICE52819.2021.9645706.

W. Wang, G. Wen and Z. Zheng, "Design of Deep Learning Mixed Language Short Text Sentiment Classification System Based on CNN Algorithm," 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkur, Karnataka, India, 2022, pp. 1-5, doi: 10.1109/ICMNWC56175.2022.10031786.

***All Images are from google stock images or created by implementation or screenshots from my paper**