SHAHIN SULAIMAN M

07.04.2025

CNN-Based Road Signs Classification

[Document subtitle]

# Objective

The main objective of this project is to design and implement a Convolutional Neural Network (CNN) model capable of classifying road signs accurately into 30 distinct categories. With the growing demand for autonomous driving systems, a reliable road sign recognition mechanism is essential for ensuring safety and compliance with traffic regulations. By leveraging the power of deep learning and image processing, the model aims to identify road signs such as speed limits, stop signs, and pedestrian crossings, among others. This will not only aid in the development of self-driving cars but also assist advanced driver-assistance systems (ADAS) in interpreting traffic scenarios. The use of CNNs, known for their ability to capture spatial hierarchies in images, ensures that the model can generalize well to real-world scenarios. The model is trained and evaluated using a structured dataset of labeled road sign images. The goal is to achieve high performance in terms of accuracy, precision, recall, and F1-score. The successful deployment of this project could contribute significantly to the field of intelligent transportation systems, making road travel safer and more efficient by minimizing human error and improving response times through automation.

# Dataset Description

The dataset used in this project contains labeled images of 30 classes of road signs, including commonly seen types like speed limit indicators, stop signs, and pedestrian warnings. The dataset has been curated to include a diverse set of road sign images captured in varying lighting conditions, orientations, and levels of clarity to simulate real-world conditions. It is split into three main subsets to facilitate training, validation, and testing of the CNN model: 70% of the dataset is used for training, 15% for validation, and the remaining 15% for testing. This split ensures that the model is exposed to enough data during training while being evaluated on unseen data during validation and testing to measure generalization capability. Each image in the dataset is labeled with its corresponding class, allowing supervised learning algorithms to map visual features to specific categories. The diversity in the dataset helps in training a robust model that can handle variations in real-life scenarios, such as changes in angle, lighting, or partial obstructions. The comprehensive nature of the dataset is critical in building a CNN model that performs consistently across different environments, contributing to its reliability and potential deployment in autonomous systems.

# Data Preprocessing

Data preprocessing plays a vital role in ensuring the quality and effectiveness of the input data fed into the Convolutional Neural Network (CNN). For this project, the preprocessing pipeline begins with resizing all images to a uniform dimension of 64x64 pixels. This step standardizes the input size and reduces computational load while maintaining sufficient detail for classification. Next, the pixel values are normalized to a range between 0 and 1 by dividing by 255. This normalization speeds up the training process and helps the model converge more efficiently. Labels associated with each image are encoded using one-hot encoding, which transforms categorical labels into binary vectors, allowing them to be interpreted correctly by the neural network. To further improve the model’s generalization, data augmentation techniques such as rotation, horizontal and vertical flipping, zooming, and shifting are applied. These techniques synthetically expand the training dataset and make the model robust to various distortions and transformations that might occur in real-world conditions. Overall, preprocessing not only ensures that the input data is clean and consistent but also enhances the diversity and quality of the training data, ultimately leading to a more accurate and resilient classification model.

# CNN Architecture

The architecture of the Convolutional Neural Network (CNN) designed for road sign classification is structured to effectively learn spatial hierarchies from input images. The input layer accepts RGB images of size 64x64x3, representing height, width, and color channels. The network includes a series of convolutional layers that apply various filters to detect edges, textures, and other spatial features. These layers are typically followed by max-pooling layers, which reduce the spatial dimensions and retain the most significant features, thereby improving computational efficiency and reducing overfitting. After several convolutional and pooling operations, the feature maps are flattened into a one-dimensional vector. This flattened vector is then passed through fully connected (dense) layers that combine the learned features to make classification decisions. To prevent overfitting, dropout layers are introduced, randomly deactivating a fraction of the neurons during training. The final layer is a softmax output layer with 30 neurons, corresponding to the 30 road sign categories. It outputs a probability distribution over the classes, enabling the model to make predictions. The overall architecture is optimized to balance complexity and performance, ensuring that the model learns effectively from the data while maintaining the ability to generalize to unseen inputs.

# Evaluation & Results

To measure the performance of the CNN model, various evaluation metrics are used, including accuracy, precision, recall, and F1-score. Accuracy reflects the overall correctness of the model, while precision and recall provide insights into how well the model distinguishes between different classes. The F1-score, which balances precision and recall, is especially useful when dealing with class imbalances. During training, the model’s learning progress is monitored using line plots that show training and validation accuracy and loss over epochs. These plots help in identifying issues such as overfitting or underfitting, guiding decisions about model improvements. A confusion matrix is also generated to visualize how well the model performs across all 30 classes, highlighting any misclassifications. The final trained model achieves high accuracy and exhibits consistent performance across the validation and testing datasets. This indicates that the model not only memorized the training data but also learned generalizable features useful for recognizing new, unseen images. The detailed evaluation ensures that the CNN is both reliable and practical for deployment in real-world scenarios, such as autonomous vehicles and traffic monitoring systems.

# Testing with New Images & Optimization Techniques

After training the CNN model, it was tested using new, previously unseen images of road signs to assess its real-world applicability. The model performed well, correctly identifying most of the road signs and demonstrating its capability to generalize beyond the training dataset. To further improve performance and robustness, various optimization techniques were employed. Hyperparameter tuning was conducted using both grid search and random search approaches. Parameters such as the number of filters, kernel size, dropout rate, and learning rate were adjusted to identify the most effective configuration. Additionally, dropout was used as a regularization method to prevent overfitting by randomly disabling a fraction of neurons during training. Experiments with different layer depths and dense units were carried out to evaluate their impact on accuracy and computational efficiency. These optimization efforts resulted in a model that is not only accurate but also stable and efficient. By thoroughly testing and refining the architecture, the model becomes more reliable for practical deployment. These steps are essential for ensuring the model’s effectiveness in varied real-world scenarios, especially in dynamic environments encountered by autonomous vehicles.

# Conclusion

In conclusion, the CNN-based road sign classification model developed in this project has proven to be both accurate and efficient. By leveraging deep learning techniques and a well-structured dataset, the model was able to classify road signs into 30 categories with high reliability. The preprocessing techniques and data augmentation strategies played a critical role in improving the model’s performance. Additionally, the carefully designed CNN architecture and the application of optimization techniques ensured that the model was not only accurate but also capable of generalizing well to new data. The successful evaluation and testing of the model demonstrate its potential for real-world applications, particularly in the domain of autonomous driving and intelligent transportation systems. The project lays a strong foundation for future work, including deploying the model as a web application using Streamlit. This would allow for broader accessibility and practical use, especially in vehicle systems and traffic management solutions. Overall, the project contributes meaningfully to the field of computer vision and presents a scalable solution for automated road sign recognition that enhances road safety and navigation efficiency.