Urban Traffic Sign Recognition and Analysis System

MSA 8395 – Computer Vision
Instructor: Dr. Peter Molnar
Student: Harshal Prashant Kamble
Submission Date: November 12, 2025

Abstract / Executive Summary

This project develops a comprehensive computer vision pipeline for urban traffic sign recognition using the Mapillary Traffic Sign Dataset. The system integrates classical vision techniques (color-based region proposals, HOG features, and SVM/RF classifiers) and deep learning models (CNN and ResNet-18 transfer learning). Through preprocessing, feature extraction, and comparative analysis, the project demonstrates how classical methods outperform deep models on small datasets, while deep models scale better with more data and augmentation.

1. Introduction & Dataset Description

The Mapillary Traffic Sign Dataset contains over 100,000 real-world images with 320,000+ annotated traffic signs across 300+ categories. A subset of 5 frequent classes (100 images each) was sampled, balanced, and split into train/val/test (70/15/15). The dataset presents diverse challenges such as occlusions, lighting variations, and multiple sign instances per image.

2. Methodology

Phase 1: Dataset preparation, preprocessing (CLAHE, bilateral filter), and region proposals using HSV-based color segmentation and shape detection (Hough transforms). IoU and recall metrics were computed for validation. Phase 2: HOG feature extraction and classical ML training (SVM, Random Forest) with cross-validation. Phase 3: Deep Learning CNN trained on standardized 128×128 crops with ReLU, BatchNorm, and Dropout. Bonus Phase: ResNet-18 fine-tuned on subset data for transfer learning comparison.

3. Results & Discussion

Model	Validation Accuracy	Test Accuracy
SVM (HOG)	0.89	0.88
Random Forest (HOG)	0.87	0.86
CNN (Custom)	0.50	0.30
ResNet-18 (Transfer Learning)	0.55	0.33

The SVM and Random Forest models achieved >85% accuracy, outperforming the CNN due to limited data. ResNet-18 transfer learning improved validation accuracy to ~55%, confirming the effectiveness of pretrained features even with small datasets.

4. Failure Analysis

Misclassifications primarily occurred within the “other-sign” class due to its visual heterogeneity. Limited training samples and imbalance reduced CNN generalization. The confusion matrix highlighted overlap between regulatory and informational signs under poor lighting or occlusion conditions.

5. Insights & Recommendations

• Classical models remain strong baselines for small datasets with engineered features like HOG. • Deep CNNs require larger datasets or augmentation for optimal learning. • Transfer learning with ResNet-18 shows promising scalability and robustness. • Future work: incorporate data augmentation, unfreeze deeper ResNet layers, explore hybrid HOG+CNN ensembles.

6. Conclusion

An end-to-end vision system for traffic sign detection and recognition was implemented. The project demonstrates the evolution from classical computer vision to modern deep learning and highlights the trade-offs in performance under dataset constraints. The final codebase, results, and report are available in the organized GitLab repository.

7. References

1. Mapillary Traffic Sign Dataset (https://www.mapillary.com/dataset/trafficsign) 2. OpenCV Documentation (https://docs.opencv.org) 3. Scikit-learn Documentation (https://scikit-learn.org) 4. PyTorch Documentation (https://pytorch.org) 5. scikit-image Documentation (https://scikit-image.org)