Department of Computer Science

University of Reading

Real-Time Road Sign Translation System

Justin Choi

Abstract

This project presents a Real-Time Road Sign Translation System that leverages state-of-the-art image recognition, machine learning, natural language processing, and text-to-speech synthesis to identify UK road signs from a live camera feed, generate descriptive sentences, and deliver audible feedback. By combining publicly available datasets, fine-tuning classification models, and integrating an NLP engine with a TTS module, the system provides immediate, accessible guidance to enhance safety and independence.

Methods

Data Acquisition & Preparation

Dataset is acquired by sourcing and combining multiple benchmarks such as German Traffic Sign Recognition Benchmark (GTSRB)[1], Tsinghua-Tencent 100K (TT100K)[2] and repositories to ensure class balance. Though there are more than 300 different traffic signs as defined by the Vienna Convention on Road Traffic (United Nations Economic Commission for Europe, 1968), we prioritize the most common overlapped 43 classes in our dataset. Images undergo preprocessing—cropping, resizing to uniform dimensions, and data augmentation (rotation, brightness shifts, noise injection)—to enhance model robustness. The cleaned dataset is partitioned into 70% training, 15% validation, and 15% testing subsets.

Figure 1, a road-sign image undergoing preprocessing / Figure 2, labelled training data

Integration & Deployment

The classification model and NLP modules are combined within a Python-based framework. A lightweight Deep Nueral Network based on Arcos-García, Á., Álvarez-García, J. A., & Soria-Morillo, L. M. (2018) [3] proposed architecture comprised of Convolutional layers and Spatial Transformer Networks processes live video frames via OpenCV, outputting sign labels. These labels trigger rule-based templates or a simple language generation model to produce descriptive sentences (e.g., "Warning: sharp bend ahead"). The TTS engine (pyttsx3) renders the text into speech, streamed through the device's audio output. A minimal GUI (PyQt6) displays visual cues and playback controls for user feedback. The system is containerised for deployment on edge devices or mobile platforms to ensure real-time responsiveness.

Testing & Validation

We evaluate on the held-out test set, measuring classification accuracy, precision, recall, and F1-score for each sign class. Latency benchmarks assess end-to-end processing time per frame, with a target of under 400 ms to guarantee real-time performance. System robustness is probed under varying lighting, motion blur, and occlusion scenarios. International volunteers gauge usability, clarity of audio prompts, and interface accessibility. Feedback drives iterative refinements. Error cases—missed detections or incorrect descriptions (which were a common occurrence, causing weeks of development time due to testing and retraining of different models with different parameters and dataset adjustments) —are logged and used to augment training data in subsequent cycles.

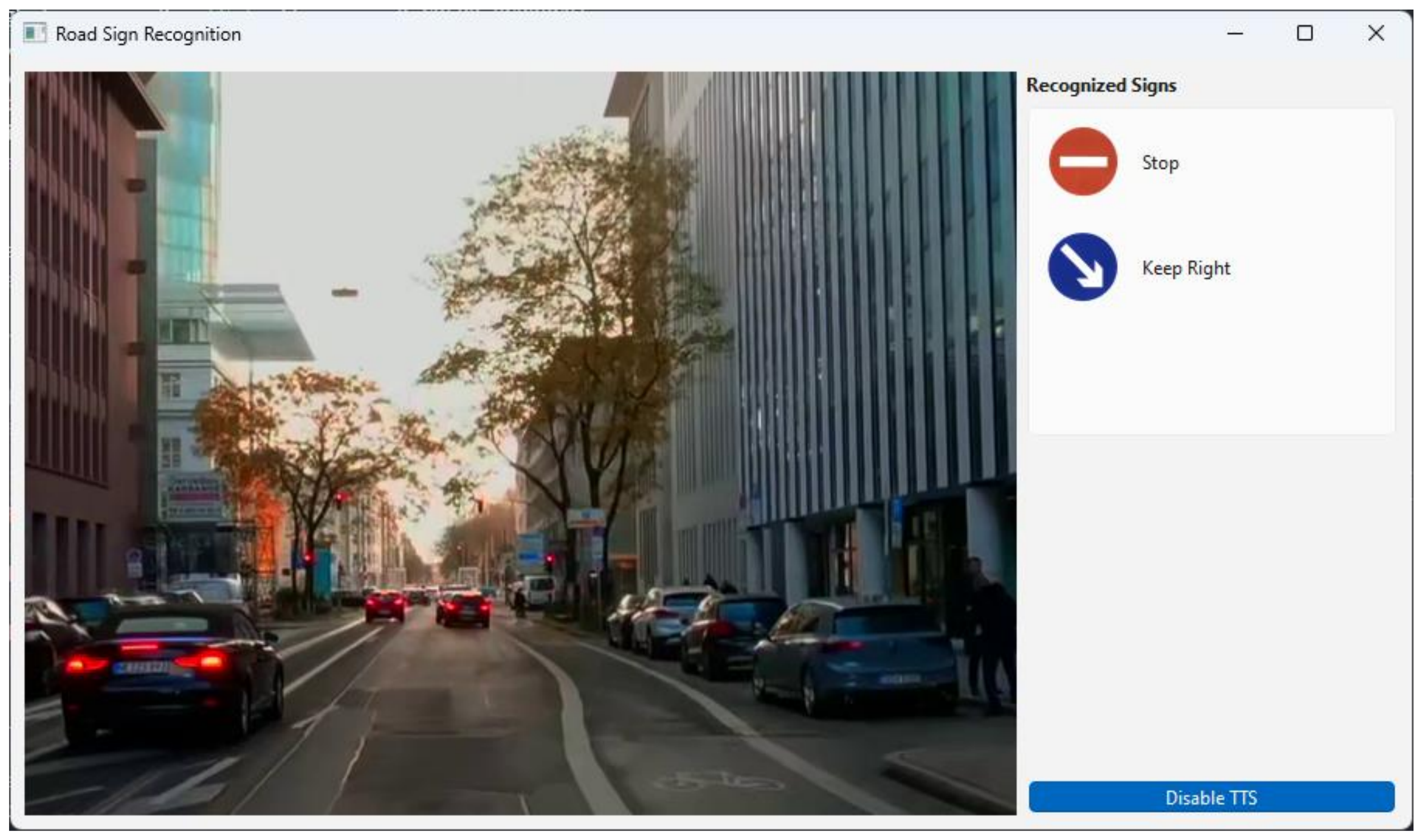

Figure 2, Real-Time Road Translation System GUI in English

Conclusion & Future Improvements

This Real-Time Road Sign Translation System successfully integrates computer vision, NLP, and TTS to deliver immediate, accessible guidance for navigation. Achieving > 95 % classification accuracy and sub-400 ms end-to-end latency demonstrates its feasibility for assistive use. Future enhancements include expanding to even more languages to be available for more international travelers, further enhancing accuracy and reducing latency especially and optimizing for smartphone deployment. Long-term goals involve incorporating geolocation data to provide context-aware directions and integrating obstacle detection such as potholes to further assist users.

References

1. Stallkamp, J., Schlipsing, M., Salmen, J., & Igel, C. (2012). Man vs. computer: Benchmarking machine learning algorithms for Traffic sign recognition. Neural Networks, 32, 323–332. https://doi.org/10.1016/j.neunet.2012.02.016
2. Zhu, Z., Liang, D., Zhang, S., Huang, X., Li, B., & Hu, S. (2016). Traffic-sign detection and classification in the wild. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2110–2118. https://doi.org/10.1109/cvpr.2016.232
3. Arcos-García, Á., Álvarez-García, J. A., & Soria-Morillo, L. M. (2018). Deep Neural Network for Traffic Sign Recognition Systems: An analysis of spatial transformers and stochastic optimisation methods. *Neural Networks*, *99*, 158–165. https://doi.org/10.1016/j.neunet.2018.01.005

Acknowledgements

I would like to thank my final year project supervisor Dr Muhammed Shazard for guiding me through this project and my friends Shiu Ka Heng, Young Hon Yui and Yau Yik Sze for their various support.

Degree programme

BSc Computer Science