AI-Powered Traffic Signal Optimization: A Three-Stage Pipeline for Urban Intersection Delay Reduction

Mohammad Usaid[1], Asmith Reddy[1], Archak Mittal[1]*
[1]*Department of Civil Engineering, Indian Institute of Technology, Bombay, Mumbai, 400076, India*
usaid3655@gmail.com, 22b0663@iitb.ac.in, archak@iitb.ac.in

Abstract Body:

1. Introduction: Urban traffic congestion is a critical challenge in India, leading to significant economic losses, increased travel time, and environmental pollution. Delays at signalized intersections are a primary contributor to this inefficiency. This paper addresses this problem by proposing an intelligent, multi-stage traffic signal optimization system designed to enhance urban mobility in Indian cities.

2. Objectives: The primary objective of this research is to develop, implement, and validate a novel three-stage AI-powered pipeline that minimizes vehicle delays and improves traffic flow at urban intersections. The system is specifically engineered to handle the complex and heterogeneous traffic conditions found in India, using real-world data from Jaipur as a testbed.

3. Methodology: Our proposed system integrates computer vision, spatiotemporal forecasting, and reinforcement learning into a cohesive pipeline. Ground truth data is generated by manually annotating the centre point of each vehicle using the VGG Image Annotator (VIA) software. A ground-truth density map is then produced by convolving these annotation points with a Gaussian kernel, which preserves object count information even in dense scenes.

- **Stage 1: Vehicle Density Estimation:** A custom ResNeXt-based deep learning architecture is employed for robust, real-time vehicle density estimation. This model enhances the standard Counting CNN by incorporating aggregated residual transformations (grouped convolutions), which improves feature representation without significantly increasing model complexity. Figure 1 shows a sample image from our dataset and its corresponding ground-truth density map, while Figure 2 details the model architecture.

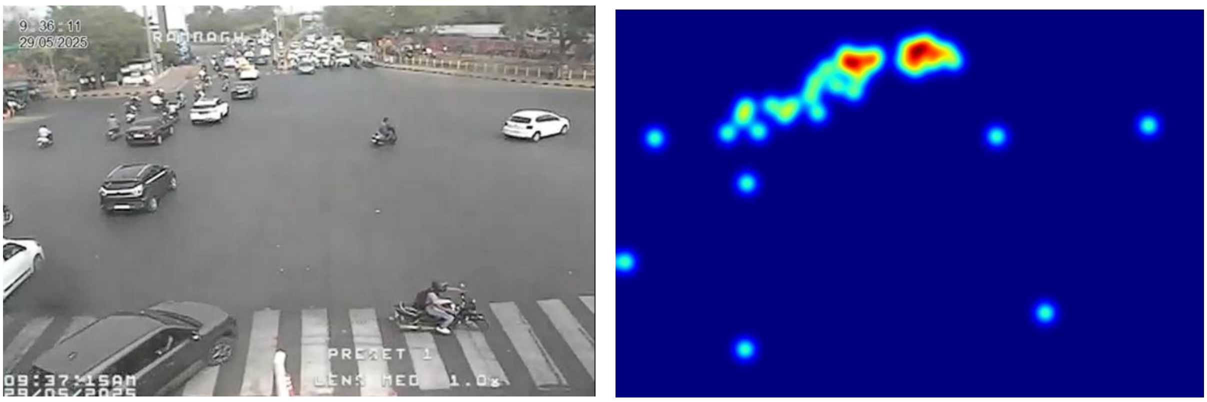
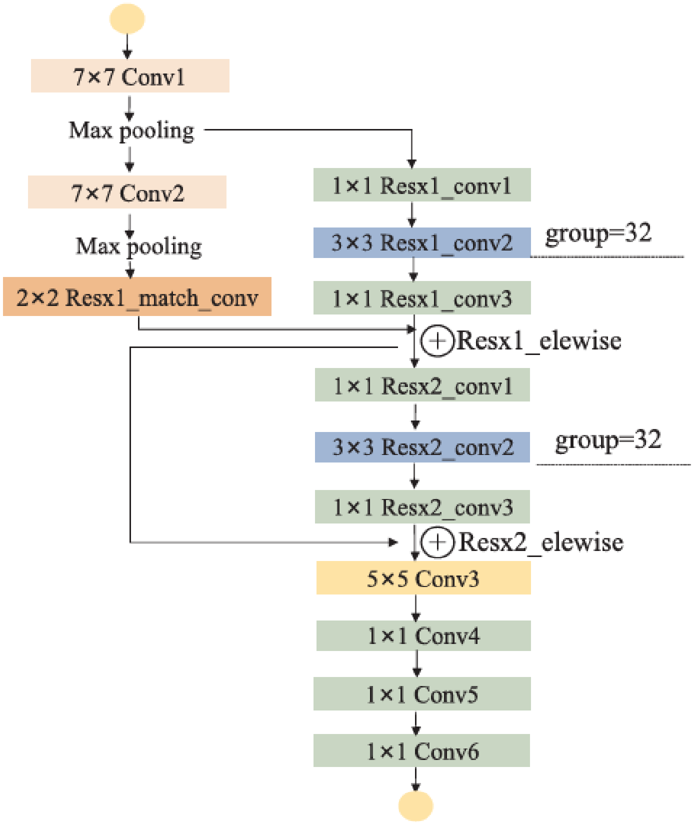


Figure 1(a): Sample traffic image from the Jaipur dataset.
Figure 1(b): Corresponding ground truth vehicle density map.



Layer Name/Stride	Output Size
Input	3@72×72
7×7 Conv1/1	32@72×72
Max pooling/2	32@36×36
7×7 Conv2/1	32@36×36
Max pooling/1	32@36×36
1×1 Resx1_conv1/1	32@36×36
3×3 Resx1_conv2/1	32@36×36
1×1 Resx1_conv3/2	32@18×18
2×2 Resx1_match_conv/2	32@18×18
Resx1_elewise	32@18×18
1×1 Resx2_conv1/1	32@18×18
3×3 Resx2_conv2/1	32@18×18
1×1 Resx2_conv3/1	32@18×18
Resx2_elewise	32@18×18
5×5 Conv3/1	64@18×18
1×1 Conv4/1	1000@18×18
1×1 Conv5/1	400@18×18
1×1 Conv6/1	1@18×18

Figure 2: The ResNeXt-based architecture for Stage 1 vehicle density estimation.

- **Stage 2: Traffic Flow Forecasting:** A pre-trained Spatial-Temporal-Decoupled Masked Autoencoder (STD-MAE) is utilized to predict near-future traffic flow (Figure 3). This framework uses two decoupled autoencoders with self-attention mechanisms to learn representations separately along the spatial and temporal dimensions. This allows the model to capture long-range spatiotemporal dependencies and heterogeneity effectively, enabling proactive rather than reactive traffic management.

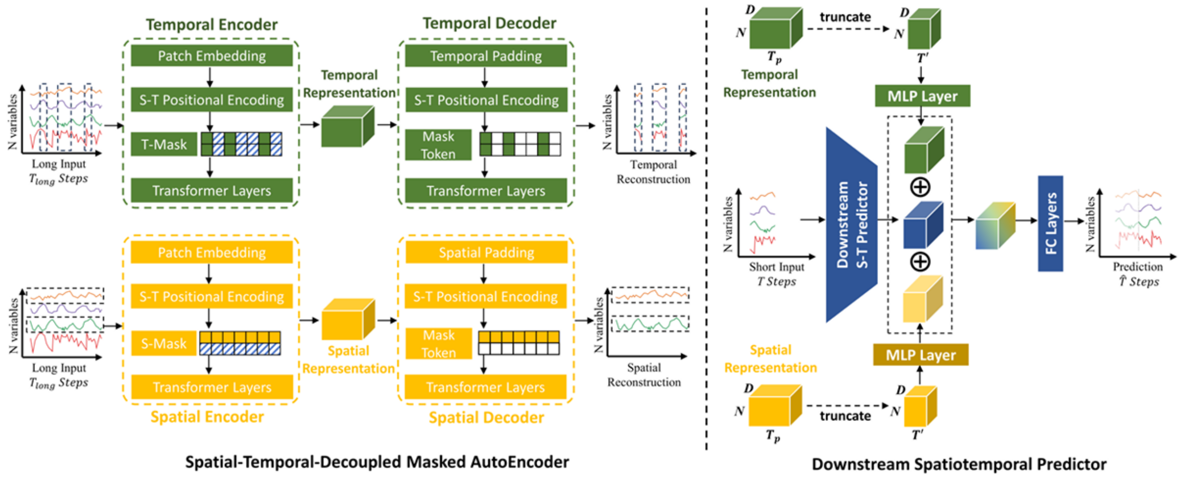


Figure 3: The STD-MAE architecture for Stage 2 traffic flow forecasting.

- **Stage 3: Dynamic Signal Optimization:** A sophisticated reinforcement learning (RL) controller serves as the decision-making core. It synthesizes the real-time density estimates and traffic flow predictions to dynamically optimize signal timings. The RL agent continuously learns and adapts its control policy to minimize network-wide delays in response to fluctuating traffic conditions.

4. Results: Preliminary evaluations indicate that the proposed system has the potential to significantly reduce average intersection wait times and queue lengths compared to traditional fixed-time and actuated signal controllers. The modular design of the pipeline ensures scalability and facilitates seamless integration with existing traffic management infrastructure.

5. Conclusions: This research presents a technologically advanced, data-driven solution for intelligent traffic signal control. By leveraging state-of-the-art AI techniques, the system offers a cost-effective and scalable pathway for transforming urban mobility. It stands as a significant contribution to India's smart cities mission, paving the way for more efficient, sustainable, and resilient urban transportation networks.

Keywords: Traffic Signal Control, Reinforcement Learning, Spatiotemporal Forecasting, Urban Mobility, Smart Cities

References:

1. Abhishek Dutta and Andrew Zisserman. 2019. The VIA Annotation Software for Images, Audio and Video. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19), October 21–25, 2019, Nice, France. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3343031.3350535.
2. Gao, H., Jiang, R., Dong, Z., Deng, J., Ma, Y., and Song, X., "Spatial-Temporal-Decoupled Masked Pre-training for Spatiotemporal Forecasting", <i>arXiv e-prints</i>, Art. no. arXiv:2312.00516, 2023. doi:10.48550/arXiv.2312.00516.
3. Tian, M., Guo, H., Chen, H., Wang, Q., Long, C., Ma, Y., 2019. Automated pig counting using deep learning. Comput. Electron. Agric. 163, 104840. https://doi.org/10.1016/j.compag.2019.05.049.
4. Guerrero-Gómez-Olmedo, Ricardo & Torre-Jiménez, Beatriz & López-Sastre, Roberto & Maldonado-Bascón, Saturnino & Oñoro, Daniel. (2015). Extremely Overlapping Vehicle Counting. 10.1007/978-3-319-19390-8_48.
5. S. M. V M, K. S P and P. Mohandas, "Real-Time Traffic Signal Prediction and Control using Deep Q-Network," *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)*, Srinagar Garhwal, India, 2023, pp. 1-6, doi: 10.1109/IC2E357697.2023.10262818.
6. Hayashi, K. and Sugimoto, M. (1999). Signal control system (moderato) in japan. In Proceed ings 199 IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems (Cat. No.99TH8383), pages 988–992.
7. Eamthanakul, B., Ketcham, M., and Chumuang, N. (2017). The traffic congestion investigating system by image processing from cctv camera. In 2017 International Conference on Digital Arts, Media and Technology (ICDAMT), pages 240–245.