LB/TH/30/2026
TH6280

VISION-BASED ACCIDENT ANTICIPATION WITH DEEP NEURAL NETWORKS

Vipulananthan Vipooshan

248048G

Master of Science(Major Component of Research)

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

June 2026

VISION-BASED ACCIDENT ANTICIPATION WITH DEEP NEURAL NETWORKS

Vipulananthan Vipooshan

248048G

Thesis submitted in partial fulfilment of the requirements for the degree
Master of Science(Major Component of Research)

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

June 2026

DECLARATION

I declare that this is my own work and this Thesis does not incorporate without acknowledgment any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 10/06/2026

The supervisor should certify the Thesis with the following declaration.

The above candidate has carried out research for the Master of Science(Major Component of Research) Thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Dr. Charith Chitraranjan

Signature of the Supervisor: Date: 10/06/2026

ACKNOWLEDGEMENT

This thesis would not have been possible without the support, guidance, and encouragement of many individuals and institutions, to whom I am deeply grateful.

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Charith Chitraranjan. His visionary insights and meticulous attention to detail provided the ideal balance of breadth and depth throughout this work. His patient mentorship, constructive feedback, and unwavering confidence in my abilities continuously motivated me to overcome challenges and develop my ideas into a cohesive and robust system.

I am also thankful to my examiners, Dr. Ranga Rodrigo and Dr. Thanuja D. Ambegoda, for their careful evaluation of my research during progress reviews. Their insightful questions and valuable suggestions encouraged me to strengthen my arguments and clearly articulate my contributions, significantly enhancing the overall quality of this work.

I am sincerely grateful to the academic staff of the Department of Computer Science & Engineering, Faculty of Engineering, University of Moratuwa, for providing a strong academic foundation and a supportive research environment. Their guidance throughout my MSc by Research program has been instrumental in the successful completion of this thesis.

I would also like to extend my heartfelt thanks to my family for their unwavering love, patience, and moral support. Their constant encouragement and understanding made this journey possible.

Finally, I would like to acknowledge all those who contributed, directly or indirectly, to this research, including the contributors and maintainers of open-source datasets, the broader research community whose prior work informed and inspired this study, and all individuals whose efforts laid the foundation for this thesis. I am honoured to build upon their contributions.

ABSTRACT

Traffic accidents remain a leading cause of fatalities and injuries worldwide, making the ability to anticipate hazardous situations in advance critical for improving road safety. Accident anticipation systems enable timely preventive actions, such as driver alerts and collision avoidance maneuvers, and play a key role in advanced driver assistance systems (ADAS) and autonomous driving. While many existing approaches rely on multiple sensors such as LiDAR, radar, and GPS, using dash-cam videos offers a cost-effective and easily deployable alternative, albeit with significant challenges due to the complexity of real-world driving environments and real-time constraints.

In this thesis, we investigate efficient and accurate accident anticipation from dash-cam videos through two complementary approaches. First, we propose an enhanced spatio-temporal graph-based framework (STAGNet), which improves accident prediction by incorporating richer spatio-temporal features and aggregating them using recurrent neural networks. Second, we introduce VAGNet, a novel architecture that leverages global scene representations instead of computationally expensive object-level features, combining transformer and graph modules with the vision foundation model VideoMAE-V2 for efficient feature extraction.

Extensive experiments conducted on multiple benchmark datasets, including DAD, DoTA, DADA, and Nexar, demonstrate that the proposed methods achieve superior performance in terms of average precision and mean time-to-accident compared to existing approaches. Notably, the global feature-based framework significantly improves computational efficiency while maintaining high predictive accuracy. Together, these contributions advance the development of scalable, real-time accident anticipation systems using only dash-cam video, providing a practical pathway toward safer intelligent transportation systems.

Keywords: Advanced driver assistance systems, Computer vision, Transformers, Graph neural networks, Foundation models.

TABLE OF CONTENTS

Declaration of the Candidate & Supervisor i

Acknowledgement ii

Abstract iii

Table of Contents iv

1 Introduction 1

2 Literature Review 3

2.1 Methods based on convolutional neural networks 3

2.2 Convolutional and recurrent neural networks 3

2.3 Incorporating attention mechanisms 4

2.4 Graph neural networks 6

2.5 Reinforcement learning-based methods 7

2.6 Transformer-based and Multi-modal Methods 8

3 Methodology 9

3.1 STAGNet: A Spatio-Temporal Graph and LSTM Framework for Accident Anticipation 9

3.1.1 Spatio-Temporal Object Graph Learning 9

3.1.2 Spatio-Temporal Frame-level Feature Learning 10

3.1.3 Frame Graph Learning 10

3.1.4 Model Pruning 12

3.2 VAGNet: Vision-based Accident Anticipation with Global Features 12

4 Experimental Evaluation 16

4.1 Experimental Settings 16

4.1.1 Datasets 16

4.1.2 Networks and Baselines 17

4.1.3 Evaluation Metrics 19

4.1.4 Results and Discussion 20

5 Conclusion and Future Work 31

5.1 Limitations and Future work 31

5.2 Conclusion 31

References 33