GRAPH ATTENTION NETWORK FOR 3D OBJECT DETECTION IN AUTONOMOUS DRIVING

Paranamana Thilina Prabhath

218535U

Master of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa

Sri Lanka

July 2024

GRAPH ATTENTION NETWORK FOR 3D OBJECT DETECTION IN AUTONOMOUS DRIVING

Paranamana Thilina Prabhath

218535U

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree Master of Science in Artificial Intelligence

Department of Computational Mathematics
Faculty of Information Technology

University of Moratuwa
Sri Lanka

July 2024

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement 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 acknowledgement 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: *UOM Verified Signature* Date:09/07/2024

The above candidate has carried out research for the PhD/MPhil/Masters thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Dr. Thushari Silva

Signature of the Supervisor: Date:09/07/2024

DEDICATION

I dedicate my thesis to my parents for their boundless love, unwavering support and encouragement throughout my education journey. It is my sincere hope that this achievement fulfills the dreams they envisioned for me and brings them the pride and joy they deserve.

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere gratitude and appreciation to everyone who has supported me on this journey to its completion.

I would like to convey my special thanks to my supervisor Dr. Thushari Silva who has given me most valuable guidance and suggestions throughout the process of this research. The time and effort spent on the research will be highly appreciated.

I take this valuable opportunity to thank Prof. Asoka Karunananda, Dr. Subha Fernando, Dr. Sagara Sumathipala, Dr. Thilini Piyatilake, Dr. Romesh Thanuja and all academic staff of the Department of Computational Mathematics who gave knowledge and valuable guidance throughout this master program. Special thank goes to all my friends from my batch to had some valuable discussions about this research process and provided necessary documents from time to time and Non-academic staff who gave their hearties support to do the research.

Finally, my very special thanks and appreciation goes to my parents without their encouragement this would not have been possible.

ABSTRACT

Since 2007, the DARPA (Defense Advanced Research Projects Agency) Grand Challenge in Autonomous Driving has significantly surged the popularity of developing Autonomous vehicles (AVs) within the autonomous industry. Safely driving in complex and dynamic environments requires AVs to have accurate and precise localization of surrounding objects. The advancement of sensor technology, particularly LiDAR (Light Detection and Ranging) brings higher accuracy and avoids limitations associated with digital camera images. Recent Advancements in the Deep Learning (DL) models have shown good performance in LiDAR point cloud segmentation, classification, and object detection tasks. However, LiDAR data generates unstructured data as point clouds around 10^5 3D points per 360^o sweep and it's bringing major computation challenges for modern detectors to process this large amount of data in real-time. Most existing approaches use point-based, voxel-based, and view-based methods. However, point clouds are also unstructured and sparse around objects not like image pixels. Transforming to another representation causes to loss of important details related to objects.

This thesis presents our Graph Attention Network (GAT) based approach to 3D Object detection using point clouds in autonomous Driving applications. The main input to our model is raw LiDAR point clouds which contain x, y, z coordinates and intensity values. The major output from the developed model is bounding box of detected cars within a given frame. Our approach can be divided into two stages. First, we remove ground points from initial point clouds in order reduce number of points involved in proceeding steps to achieve real-time processing. The second stage, we utilize voxel downsampling further reduce points in 3D space and generate a graph using the nearest neighbor technique to apply the GAT DL model. Our model evaluates using the widely used KITTI benchmark dataset. The results indicate that our model achieves performance levels comparable to those of state-of-the-art LiDAR-based 3D detection methods.

Keywords: LiDAR, Point clouds, Graph Attention Network, Deep Learning

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
AI	Artificial Intelligence
AV	Autonomous Vehicle
DARPA	Defense Advanced Research Projects Agency
DL	Deep Learning
LiDAR	Light Detection and Ranging
GAT	Graph Attention Network
GNN	Graph Neural Network

LIST OF APPENDICES