INTERSECTION WAYPOINT IDENTIFICATION FOR VISION BASED INDOOR NAVIGATION OF NANO-SCALE DRONES

H.K.A.Y.D. Kodithuwakku

248111U

Master of Science(Major Component of Research)

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

February 2026

INTERSECTION WAYPOINT IDENTIFICATION FOR VISION BASED INDOOR NAVIGATION OF NANO-SCALE DRONES

H.K.A.Y.D. Kodithuwakku

248111U

Dissertation submitted in partial fulfillment 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

February 2026

DECLARATION

I declare that this is my own work and this 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: Date: 18.02.2026

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

Name of Supervisor: Dr. Chathuranga Hettiarachchi

Signature of the Supervisor: Date: 18/02/2026

Name of Supervisor: Dr. Sulochana Sooriyaarchchi

Signature of the Supervisor: Date: 18/02/2026

ACKNOWLEDGEMENT

This thesis would not have been feasible without the help, direction, and inspiration of numerous people and organizations, for which I am extremely grateful.

First of all, I would like to express my sincere gratitude to my supervisors, Dr. Chathuranga Hettiarachchi and Dr. Sulochana Sooriyaarachchi. Throughout this work, their visionary insights and meticulous dedication to detail improved the work significantly. I am especially grateful for the valuable time they dedicated to guiding me, despite their demanding schedules. Their unwavering faith in my abilities, helpful criticism, and patient coaching inspired me to overcome obstacles and hone my concepts into a solid, unified system.

Additionally, I owe a debt of gratitude to my examiners, Professor Chathura de Silva and Dr. Sunimal Rathnayake. I strengthened my arguments and clarified my contributions as a result of their thorough examination of my study during progress reviews and their insightful inquiries.

I would like to express my sincere gratitude to Senior Professor N.D. Gunawardena, Vice Chancellor of University of Moratuwa and Professor J.M.A. Manatunge, Dean of the Faculty of Engineering for their continued support towards administrative matters. I also gratefully acknowledge Dr. Uthayasanker Thayasivam, Head of the Department of Computer Science and Engineering, University of Moratuwa for his support. I would like to express my appreciation to all of the faculty and administrative staff for their effective management of academic and administrative issues, which enabled me to devote all of my attention to my project. I sincerely appreciate the generous support from the University of Moratuwa Senate Research Committee grant SRC/LT/2024/02.

I gratefully acknowledge Mr. Ignatious Peiris for the technical and moral support. I would also like to thank Mr. Kenul Perera for his efforts in creating video to simulation tool. I would also like to thank Mr. Vihanga Wickramage, Mr. Sandun Dushyantha for the synthetic environment building. I am grateful to my colleagues at the IntelliSense Lab for their technical support, early-stage discussions, and the camaraderie that made the research process both productive and enjoyable.

Personally, I am extremely thankful to my parents and my brother. My biggest source of strength has been their faith in my abilities. I want to express my sincere gratitude to my friends for their support and encouraging remarks.

Lastly, I would like to thank everyone who helped with this research, whether directly or indirectly: the creators of open-source datasets and tools, the peer reviewers whose work influenced my literature review, and the larger academic community whose findings served as the foundation for this thesis.

ABSTRACT

Nano-scale drones represent a promising platform for indoor inspection and search operations in GPS-denied, constrained environments, as their compact form factor enables access to narrow spaces. However, they require lightweight, resource-efficient autonomous navigation algorithms. Vision-based navigation offers an infrastructure-free substitute to GPS and external positioning systems. Simple vision-based navigational algorithms such as ring attractor, optical flow-based approaches, suffer from drift and cumulative error over time. Introducing semantic waypoints provides a mechanism to mitigate such drift.

This work proposes an egocentric, vision-based waypoint navigation framework that formulates waypoint detection as an intersection understanding problem, cast as a concept learning task. We first formulate the intersection identification problem as a multi-class classification. To ensure a model can be trained with limited data and still preserve the conceptual understanding, each intersection is decomposed into its fundamental directional components: left, right, and forward. This novel multi-label classification is advantageous because it can extend to zero-shot inference and complex junction types. Further, we augment the limited real data with synthetic visual data collected from a safer digital twin environment which replicates the real-world scenes. We design five experimental setups in which model weights are initialized from Kaiming He, ImageNet or synthetic pre-trained weights obtained using the Digital Twin dataset.

We show that, (i) Intersections can be used as waypoints for the navigation of nano-drones. (ii) Multi-label classification outperforms multi-class formulation in intersection identification. (iii) Model trained using digital twin data alone can yield an F1 score of 0.6 in real-world predictions, compared to 0.5 from a random classifier. (iv) Introducing limited real-world training data can bring the model up to near-perfect accuracy with faster convergence. (v)The multi-label classification of junctions provides insightful visual explanations ensuring reliability and generalizability of the conceptual framework. Finally, the proposed framework is effective in environments such as libraries and supermarkets where structured aisle naturally exists and GPS is unavailable. This method offers an infrastructure and drift-resistant solution for the navigation of nano-scale drones by leveraging intersections as semantic waypoints.

Keywords: ego-centric vision, explainability, concept learning, nano-scale drones

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	**Description**
AVL	Audio-Visual-Language
CfLib	Crazyflie Python Library
CNN	Convolutional Neural Network
COTS	Commercial-off-the-shelf
CRTP	Crazyflie Real Time Protocol
DORY	Deployment ORiented MemorY
EVC	Early Visual Cortex
FLA	Fast Lightweight Autonomy
fMRI	Functional Magnetic Resonance Imaging
GPS	Global Positioning System
IIR	Infinite Impulse Response
IMU	Inertial Measurement Units
INDI	Incremental Nonlinear Dynamic Inversion
IoT	Internet of Things
MCU	Microcontroller Units
NEMO	NEural Minimizer for pytOrch
NUC	Next Unit of Computing
OPA	Occipital Place Area
PCB	Printed Circuit Board
PID	Proportional-Integral-Derivative
PPA	Parahippocampal Place Area
PULP	Parallel Ultra Low Power
RSC	Retrosplenial Complex
SaR	Search and Rescue
SLAM	Simultaneous Localization and Mapping
ToF	Time-of-Flight
TSM	Temporal Shift Module
UAV	Unmanned Aerial Vehicles
VLA	Vision-Language-Action Navigation
VQA	Visual Question Answering
WPN	Waypoint Prediction Networks
YOLO	You Only Look Once

LIST OF APPENDICES