BUILDERFORMER: A PANOPTIC SEGMENTATION APPROACH TO AUTOMATIC FLOOR PLAN ANALYSIS USING VISION TRANSFORMERS

Mestiyage Don Pathum Pramuditha Goonathilake

218525N

Master of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa

Sri Lanka

July 2024

BUILDERFORMER: A PANOPTIC SEGMENTATION APPROACH TO AUTOMATIC FLOOR PLAN ANALYSIS USING VISION TRANSFORMERS

Mestiyage Don Pathum Pramuditha Goonathilake

218525N

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.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 01/07/2024

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

Name of Supervisor: Dr. ALARR Thanuja

Signature of the Supervisor: Date: 01/07/2024

DEDICATION

I would like to express my heartfelt dedication of this thesis to my parents, who have served as the pioneering figures and unwavering pillars that have provided support and inspiration throughout my educational journey. Furthermore, I am deeply grateful to extend this dedication to all the lecturers of the Department of Computational Mathematics, Faculty of Information Technology, University of Moratuwa.

ACKNOWLEDGEMENT

I convey my gratitude to everyone who has contributed to completing this research project. Their constant support, guidance and encouragement have greatly benefited me.

Firstly, I want to express my heartfelt gratitude to my supervisor, Dr. ALARR Thanuja, Senior Lecturer in the Department of Computational Mathematics at the Faculty of Information Technology, University of Moratuwa. I thank him for his expert advice, exceptional mentoring, and continuous support. He was instrumental in shaping the direction and quality of this research through his deep knowledge, insightful feedback, and constant commitment to academic excellence.

In addition, I would like to extend my deepest gratitude to my family and friends for their unwavering support. Their continuous encouragement, understanding, and support are determined as a consistent source of motivation and inspiration throughout the research journey. They have been instrumental in my belief that I can do it, and their constant support at times of failure has been indispensable. I am deeply grateful for the loyalty they have shown me.

Finally, I would also like to thank BuilderBid founder Mr. Bo Dillon for accepting and developing my idea to integrate this research into their production-grade software, which will help them streamline the takeoff process of floor plans. I have been able to focus on the correct direction thanks to his immense feedback throughout the research project.

Throughout this research, I referred to numerous research publications. I sincerely appreciate the authors of these works' notable contributions to the field, which greatly enhanced the scope and quality of this project.

ABSTRACT

Automatic floor plan analysis, rooted in computer vision and pattern recognition, seeks to extract meaningful insights from architectural and interior design drawings. Its significance spans across various sectors, including architecture, interior design, real estate, and construction. However, conventional methods, whether rule-based or learning-based, encounter constraints such as manual annotation requirements and the challenge of achieving precise segmentation. This research project endeavors to evaluate the efficacy of vision transformers in tasks crucial to floor plan analysis, including identifying unique objects like doors and windows, extracting line segments to represent room layouts, and segmenting object areas such as room spaces. A novel panoptic segmentation approach is proposed to achieve a holistic understanding of scenes by integrating object detection with image segmentation. This approach involves identifying and delineating individual object instances within regions of interest and subsequently segmenting images into meaningful regions. Key to this methodology is the utilization of the Segment Anything Model for segmentation, guided by input prompts generated through object detection models. The proposed approach introduces a mask classification mechanism within the segmentation stage, diverging from conventional methods of assigning semantic labels to individual pixels. The effectiveness of this approach is demonstrated through promising results obtained on existing datasets. Real-time object detection models like YOLOv8 are employed alongside end-to-end object detection models like RT-DETR to provide visual prompts for the Segment Anything Model. This innovative approach is prepared to revolutionize automatic floor plan analysis in real-world applications by bridging the gap between object detection and image segmentation. As the final step of this research project, the proposed methodology integrates into a web application, enabling users to visualize results for various floor plan types, thus providing a practical tool for professionals in architecture, interior design, real estate, and construction to analyze floor plans with enhanced accuracy and efficiency.

Keywords: Automatic Floor Plan Analysis, Panoptic Segmentation, Real-Time Object Detection, Visual Prompt Engineering, Vision Transformers

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
API	Application Programming Interface
CAD	Computer-Aided Design
CNN	Convolutional Neural Network
DL	Deep Learning
GAN	Generative Adversarial Network
GNN	Graph Neural Network
GPU	Graphics Processing Unit
ML	Machine Learning
RAG	Room Adjacency Graph
RT-DETR	Real-Time Detection Transformer
SAM	Segment Anything Model
SG	Shape Grammar
SOTA	State Of The Art
VIT	Vision Transformer
YOLO	You Only Look Once

LIST OF APPENDICES