This is the cover page

劉宇舜 Yu-Shun Liu

Student ID: 413551030

Department of Computer Science and Information Engineering
National Yang Ming Chiao Tung University

Course: This is the course name

April 3, 2025

Contents

1 Introduction

This is the introduction section. MaskGIT [1] is a novel approach to image generation using transformers. The transformer architecture [3] has revolutionized many areas of deep learning. Recent advances in vision-language models [2] have shown promising results in various tasks.

2 Implementation Details

This is the implementation details of the project.

2.1 Dataset

This is the dataset used in the project.

2.2 Model

This is the model used in the project.

3 Experiment Results

This is the results section.

3.1 Experiment 1

This is a subsubsection.

3.2 Experiment 2

This is another subsubsection.

4 Discussion

This is a discussion section.

4.1 Experiment 1

This is a subsubsection.

4.2 Experiment 2

This is another subsubsection.

References

[1] Huiwen Chang et al. "MaskGIT: Masked Generative Image Transformer". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022, pp. 11315–11325.

[2] Alec Radford et al. "Learning transferable visual models from natural language supervision". In: *arXiv preprint arXiv:2103.00020* (2021).

[3] Ashish Vaswani et al. "Attention is all you need". In: *Advances in neural information processing systems* 30 (2017).