

**SUSTAINED CONTEXT-AWARE IMAGE  
GENERATION FOR THE DUNGEONS &  
DRAGONS DOMAIN**

WAG Weerasundara

229403G

Master of Science in Computer Science & Engineering

Department of Computer Science & Engineering  
Faculty of Engineering

University of Moratuwa  
Sri Lanka

April 2024

**SUSTAINED CONTEXT-AWARE IMAGE  
GENERATION FOR THE DUNGEONS &  
DRAGONS DOMAIN**

WAG Weerasundara

229403G

Thesis submitted in partial fulfillment of the requirements for the degree  
Master of Science in Computer Science & Engineering

Department of Computer Science & Engineering  
Faculty of Engineering

University of Moratuwa  
Sri Lanka

April 2024

## DECLARATION

I declare that this is my own work and this Thesis 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: 06/30/2024

The supervisor should certify the Thesis with the following declaration.

The above candidate has carried out research for the Master of Science in Computer Science & Engineering Thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Nisansa de Silva

Signature of the Supervisor:

 Digitally signed by  
Nisansa de Silva  
Date: 2024.06.30  
22:38:48 +05'30'

Date: 30/06/2024

## **ACKNOWLEDGEMENT**

I would like to express my heartfelt gratitude to my supervisor, Nisansa de Silva, for his invaluable contribution to the development of this project. His in-depth knowledge and expertise in both Dungeons and Dragons and Machine learning have been instrumental in guiding me through the entire project development process. Without his support and guidance, this would not have been possible.

## ABSTRACT

*Dungeons & Dragons* (D&D) is a fantasy tabletop role-playing game which has become a subculture phenomenon due to its immense popularity. In a D&D adventure images play a major role in guiding the players' emotions and providing additional information regarding the setting. The proposed research attempts to use text extracted from pre-made adventure to generate cohesive and contextually accurate images according to the given setting. The core of the proposed methodology involves two strategic components. First, the project seeks to develop an *Natural Language Processing* (NLP) model capable of deep textual analysis to identify and understand the underlying context and key elements within the D&D text. This model will focus on extracting salient features and narrative cues from the text, which will then be used to generate precise prompts. These prompts are designed to encapsulate the essential elements needed to guide the image generation process, ensuring that the resulting visuals are not only relevant but also enrich the storytelling by aligning closely with the D&D lore. The second component is a comprehensive pipeline that generates consistent images in a zero-shot manner. Additionally, this study proposes the creation of an end-to-end pipeline that not only generates images but also automates the creation of complete D&D adventures. This pipeline will integrate multi-agent workflows, an image generation framework, and NLP techniques to produce a comprehensive suite of adventure materials—including story narratives, gameplay guidelines, gameplay-related tables, contextual images, and maps. This holistic approach is designed to streamline the preparation process for Dungeon Masters, enabling them to deliver richly detailed sessions with less preparation time.

**Keywords:** NLP, Image generation, Diffusion, D&D

## TABLE OF CONTENTS

Declaration of the Candidate & Supervisor	i
Acknowledgement	ii
Abstract	iii
Table of Contents	iv
List of Appendices	viii
1 Introduction	1
1.1 Tabletop role-playing games	1
1.2 Dungeons and Dragons	1
1.2.1 Setting	2
1.2.2 Importance of Lore	2
1.2.3 Role of Images in DnD	3
2 Research Problem	5
3 Research Objectives	6
4 Literature Survey	7
4.1 Key Phrase extraction	7
4.1.1 Text Pre-processing	7
4.1.2 Text Summarization	8
4.1.3 Keyphrase extraction using machine learning	8
4.1.4 Named Entity Recognition	9
4.2 Prompt generation	9
4.2.1 Prompt Engineering and Design Principles	10
4.2.2 Control over Text Generation with Templates and Constraints	10
4.2.3 Generative Pretrained Transformer (GPT) Models	10
4.2.4 Evaluation Metrics for Text Generation	11
4.3 Image generation	11
4.3.1 Image Synthesis with Generative Adversarial Networks	11
4.3.2 Convolutional Neural Networks (CNNs) for Image Generation	12

4.3.3	Variational Autoencoders (VAEs) for Image Generation	12
4.3.4	Diffusion models for Image Generation	13
4.3.5	Image Style Transfer	17
4.3.6	Image-to-Image Translation	19
4.4	Image matching	20
4.4.1	Evaluation Metrics for Image Generation	20
4.4.2	Image Similarity Metrics	20
4.5	Sustaining context across sessions	21
4.5.1	Document-Level Sentiment Analysis	21
4.5.2	Cross-Document Event Co-reference	22
4.5.3	Document-Level Relation Extraction	22
4.6	Consistency of generated images	23
4.6.1	Image Temporal Consistency	23
4.7	Conclusion	23
5	Dataset Preparation	25
5.1	Introduction	25
5.2	Datasets	26
5.3	Limitations	27
5.4	Conclusion	28
6	Testing LoRA & ControlNet	30
6.1	Introduction	30
6.1.1	LoRA	30
6.1.2	ControlNet	31
6.2	Testing LoRA	32
6.3	Testing ControlNet	33
6.4	Limitations	34
6.5	Conclusion	34
7	Dungeon Generation	41
7.1	Force Directed Graph Drawing	41
7.1.1	Graph Aesthetics and Force-Directed Principles	41
7.1.2	Algorithmic Approaches	41

7.1.3	Applications in Dungeon Generation	42
7.1.4	Enhancing Dungeon Layouts with Planarity	43
7.2	Cellular Automata and Perlin Noise	44
7.3	Limitations	48
7.4	Conclusion	48
8	Consistent Image Generation	49
8.1	Introduction to Consistent Image Generation	49
8.2	Content Generation for D&D Campaigns	49
8.2.1	Automated Content Creation Process	50
8.3	Application of Named Entity Recognition (NER)	52
8.3.1	Comparative Analysis of Named Entity Recognition in the Dungeons and Dragons Domain	53
8.4	Image Generation Pipeline	62
8.4.1	Training a General Art Style	63
8.4.2	Training a Specific Character Generation	63
8.4.3	Experimental Setup	74
8.4.4	Results	76
8.5	Limitations	78
8.6	Conclusion	79
9	Research Outcome	80
9.1	Limitations	87
9.2	Conclusion	88
10	Ethical Considerations	89
11	Overall Limitations	90
11.1	Lack of High-Quality Dataset for LoRA	90
11.2	Generation of Redundant LoRA	90
11.3	Time and Resources Consumed for Training	90
11.4	Variability in Generation Time	91
12	Conclusion	92
	References	93



## LIST OF APPENDICES

Appendix	Description	Page
Appendix -A	Generated Adventure Modules	106