

AN ARCHITECTURAL MODEL FOR THE SKETCH IMAGE COLORIZATION WITH DEEP LEARNING

W A Rashmi Shehana

229389P

M.Sc. in Computer Science

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

June 2024

AN ARCHITECTURAL MODEL FOR THE SKETCH IMAGE COLORIZATION WITH DEEP LEARNING

W A Rashmi Shehana

229389P

Thesis submitted in partial fulfillment of the requirements for the degree

M.Sc in Computer Science

Department of Computer Science and Engineering

Faculty of Engineering

University of Moratuwa

Sri Lanka

June 2024

DECLARATION

I declare that this is my own work and this thesis does not incorporate without acknowledgment 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 acknowledgment 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).

W A Rashmi Shehana

Signature:

Date: 16/June/2024

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

Name of Supervisor: Prof Indika Perera

Signature of the Supervisor:

UOM Verified Signature

Date:

ACKNOWLEDGEMENT

I would like to give my special thanks to,

Professor Indika Perera, Senior Lecturer at the University of Moratuwa.

My loving father and mother, for always being my strength and courage.

ABSTRACT

Image colorization can be considered a vital task in several industries such as the comic, animation, and technical illustration industries. Because of that, it is a field of research with a considerable market demand. There are several fields of image coloring available such as grayscale image coloring and line sketch image coloring. Unlike grayscale image coloring, sketch image colorization is a challenging task because sketch images do not have any color-related detail such as the texture and color gradient. This research proposes the software architecture of a novel technique to automate sketch image colorization with deep learning concepts. The research aims to automate the sketch colorization process while maintaining a higher accuracy compared to existing image coloring methods. This also tries to reduce the identified color errors in the majority of the currently available sketch colorization methods including water-color blurring, color distortion and bleeding. The method which is proposed for sketch image coloring contains two steps. The first step is called the coloring step which is responsible for initially coloring the sketch image. In this stage, the model assumes the color ranges and sprays a rich selection of colors on the sketch. The output of this stage is almost completely free from color faults such as color bleeding and watercolor blurring. But it may include mistakes such as unmatching differences in luminance or the unfitting colors that make an object distinguishable from other objects. To correct that issue, the output of the colorization stage is sent to the refinement stage. In that stage, the image is enhanced with a suitable contrast value and outputs the image with good quality and higher accuracy. To develop this model, two types of deep neural architectural models were adopted. For colorization model, the architecture of conditional Generative Adversarial Network (cGAN) has been used. This is a kind of Generative Adversarial Network architecture which additionally makes use of the input labels. For the refinement model, Convolutional Neural Network (CNN) architecture is used in this research. In this step, the model learns to predict the most suitable contrast value for the given input image. With the new contrast value, it mostly fix the luminance issues and unfitting color issues that can be remaining in the output of the coloring step. The components in the two steps mentioned, are developed in a way that they can be used independently as well. For verifying the accuracy of the results, two types of evaluation methods have been used. They are peak-signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) methods which are widely used in many of the image colorization researches. With these evaluation methods the ground truth is compared with the output image and obtained a value to measure how similar these images are. Also in this research, outputs of a state-of-the-art methods over the several years are compared with the outputs of the proposed method in this research.

Keywords: Sketch Image Colorization, Convolutional Neural Network, Conditional Generative Adversarial Network

TABLE OF CONTENTS

Declaration	1
Acknowledgement	2
Abstract	3
Table of Contents	4
List of Figures	6
List of Tables	8
List of Abbreviations	8
List of Appendices	8
Chapter 1	1
Introduction	1
1.1 Research Problem	4
1.2 Significance of the Project	4
1.2.1 For animation and comic industries	4
1.2.1 For Computer Science	4
1.3 Goals and Objectives	5
1.4 Justification of the research	5
1.5 Methodology	5
1.5.1 Data Source	5
1.5.2 Approach Selection	6
1.5.3 Approach	6
1.5.4 Training Phase	7
1.6 Delimitations of the scope	9
1.6.1 Justification for Outscope	9
1.6 Outline of the thesis	10
Chapter 2	11
Literature Review	11
2.1 Introduction	11
2.2 Related Work	11
2.3 Summary of the literature review	18
Chapter 3	18
Research Design	18
3.1 Introduction	18
3.2 Colorizing stage	19
3.2.1 Generator	20
3.2.1 Discriminator	21
3.3 Refinement stage	23
3.4 Summary	24

Chapter 4	25
Implementation	25
4.1 Introduction	25
4.2 Language and Tools	25
4.3 Data Processing	25
4.3 Implementation of the Colorization stage	27
4.3.1 Generator	27
4.3.2 Discriminator	27
4.3.3 The loss function of the cGAN model	28
4.4 Implementation of the Refinement stage	29
Chapter 5	31
Results and Evaluation	31
5.1 Introduction	31
5.2 Results of the methods proposed	31
5.2.1 Summary	35
5.3 Evaluation	35
5.3.1 PSNR - Peak signal-to-noise ratio	35
5.3.2 SSIM - Structure similarity	35
5.3.1 Evaluation Results	36
5.3.1.1 Colorizing stage	36
5.3.1.2 Refinement stage	37
5.4 Quantitative comparison with the other similar research work	38
5.3.1.2 Further evaluation results	39
5.4 Results Obtained in the Training Process	42
Chapter 6	44
Conclusion	44
6.1 Introduction	44
6.2 Conclusion about the research question	44
6.3 Conclusion about the model architecture used in research	44
6.4 Limitations	45
6.5 Implications for further research	46
References	47
Appendix A - Code Listing	51
1. Convolutional Block	51
2. Deconvolutional Block	51
3. Generator Submodel	52
4. Discriminator SubModel	53
5. Contrast Finding for refinement model	54
6. Refinement Model	55

LIST OF FIGURES

Figure	Description
Fig 1.1	Results of Active colorization
Fig 1.2	Results of grayscale image colorization
Fig 1.3	Result of language based sketch colorization system
Fig 1.4	High Level overview of the Final system
Fig 1.5	Formation of the colored image
Fig 1.6	High level architecture of the training model
Fig 1.7	Project scope
Fig 3.1	Architecture of the colorization model
Fig 3.2	Overview of the generator model
Fig 3.3	High level architecture of generator model
Fig 3.4	Overview of the discriminator model
Fig 3.5	High level architecture of discriminator model
Fig 3.6	High level overview of the refinement stage
Fig 3.7	Architecture of the CNN model
Fig 4.1	Example of a Ktaebum anime sketch colorization dataset
Fig 4.2	Folder structure of data after preprocessing
Fig 4.3	Generator loss
Fig 4.4	Discriminator loss
Fig 4.5	Epoch loss
Fig 5.1 (a)	Sketch Image
Fig 5.1 (b)	Output of Colorization stage
Fig 5.1 (c)	Output of Refinement stage (Final output)
Fig 5.1 (d)	Ground truth
Fig 5.2 (a)	Sketch Image
Fig 5.2 (b)	Output of Colorization stage
Fig 5.2 (c)	Output of Refinement stage (Final output)
Fig 5.2 (d)	Ground truth
Fig 5.3 (a)	Sketch Image
Fig 5.3 (b)	Output of Colorization stage
Fig 5.3 (c)	Output of Refinement stage (Final output)
Fig 5.3 (d)	Ground truth

Fig 5.4 (a)	Output of the colorizing stage
Fig 5.4 (b)	Ground truth image
Fig 5.5	SSIM and PSNR values of the output of colorizing stage
Fig 5.6 (a)	Output of the refinement stage
Fig 5.6 (b)	Ground truth image
Fig 5.7	SSIM and PSNR values of the output of the refinement stage
Fig 5.8	Screenshots of training step Epoch 0
Fig 5.9	Screenshots of training step Epoch 2
Fig 5.10	Screenshots of training step Epoch 30
Fig 5.11	Screenshots of training step Epoch 45

LIST OF TABLES

Table	Description
Table 5.1 Refinement steps	SSIM and PSNR values for Colorization and
Table 5.2 values	Comparison of average SSIM and average PSNR
Table 5.3 (a) refinement stages - 1	Accuracy values of outputs of colorizing,
Table 5.3 (b) refinement stages - 2	Accuracy values of outputs of colorizing,
Table 5.3 (c) refinement stages - 3	Accuracy values of outputs of colorizing,

LIST OF ABBREVIATIONS

Abbreviation	Description
GAN	Generative Adversarial Network
WGAN	Wasserstein Generative Adversarial Networks
cGAN	Conditional Generative Adversarial Networks
CNN	Convolutional Neural Network
AC-GAN	Auxiliary Classifier GAN
WGAN-GP	Wasserstein GAN + Gradient Penalty

LIST OF APPENDICES

Appendix	Description
Appendix - A	Code Listing