IMAGEGENIE: A MAGIC WAND FOR IMAGES ENHANCEMENT USING DEEP LEARNING

SUE CHEN XIANG

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS LAPORAN

JUDUL: IMAGEGENIE: A MAGIC WAND FOR IMAGES USING DEEP LEARNING

SESI PENGAJIAN: 2022 / 2023

Saya: ____SUE CHEN XIANG___________________

mengaku membenarkan tesis Projek Sarjana Muda ini disimpan di Perpustakaan Universiti Teknikal Malaysia Melaka dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan unituk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. * Sila tandakan (✓)

______	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
______	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi / badan di mana penyelidikan dijalankan)
✓	TIDAK TERHAD	

(TANDATANGAN PELAJAR)

Alamat tetap: 157, Jln Love Lane, Kg Baru Sg Buloh, 47000 Selangor

Tarikh: _19/9/2023__________________

(TANDATANGAN PENYELIA)

FAUZIAH KASMIN

Nama Penyelia

Tarikh: 22/9/2023

CATATAN: * Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

IMAGEGENIE: A MAGIC WAND FOR IMAGES USING DEEP LEARNING

SUE CHEN XIANG

This report is submitted in partial fulfillment of the requirements for the Bachelor of Computer Science (Artificial Intelligence) with Honours.

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2023

DECLARATION

I hereby declare that this project report entitled

IMAGEGENIE: A MAGIC WAND FOR IMAGES USING DEEP LEARNING

is written by me and is my own effort and that no part has been plagiarized

without citations.

STUDENT : ________________________________ Date : _19/9/2023___

(SUE CHEN XIANG)

I hereby declare that I have read this project report and found

this project report is sufficient in term of the scope and quality for the award of

Bachelor of Computer Science (Artificial Intelligence) with Honours.

SUPERVISOR : ________________________________ Date : _22/9/2023

(DR. FAUZIAH BINTI KASMIN)

DEDICATION

This final year project is dedicated to my supervisor, Dr. Fauziah binti Kasmin, for her guidance and support throughout this project. She has been the source of my strength and on Her wings only have I soared. I would also dedicate this to my parents who encouraged me all the way. Finally, I would like to dedicate this project to my friends for their encouragement and assistance during my university studies.

ACKNOWLEDGEMENTS

First and foremost, I would like to take this opportunity to express my gratitude to the people who helped and guided me throughout this project. I would like to show my greatest appreciation to Dr. Fauziah binti Kasmin. I can't say thank you enough for her support and guidance. Without her encouragement and guidance this project would not have materialized.

I would like to thank my friends for their assistance, support and encouragement in completing this project. I am grateful for their constant support and help.

I would also like to thank my wonderful parents for their encouragement throughout this project. Without their motivation, I cannot fulfill this project before the deadline.

ABSTRACT

The goal of this project is to develop an application that can increase the resolution of low quality images and to perform neural style transfer that composes one image in the style of another image. In today's digital age, images play a crucial role in communication and self-expression. However, low-quality images are a common problem that many people encounter, particularly when taking pictures with their mobile phones. The lack of image quality can result in images that are blurry, pixelated, or have poor resolution. These low-quality images can be frustrating for users who want to share their images on social media platforms or use them for personal or professional purposes. In this project, several pretrained deep convolutional neural networks model that are the Enhanced Deep Residual Network (EDSR), Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and deep Laplacian Pyramid Super-Resolution Network (LapSRN) are used for increasing the resolution of image and the result of each model is analyzed and compared. The EDSR was chosen as the best model as it achieved the highest average PSNR and SSIM in testing 45 images which are 25.82 dB and 0.70 respectively. Traditional image processing applications often lack the ability to perform advanced image enhancement techniques such as neural style transfer, which can be used to create artistic effects on images. The neural style transfer functionality is achieved by extracting the style of the style image using the VGG19 network architecture which is a pretrained image classification network and apply to the content image to create artistic effects on images. VGG19 was employed because it obtained the highest average ArtFID in 30 testing images which is 45.97 when compare to MobileNet and ResNet. Finally, an application is built using Flutter that combines all the functions above.

ABSTRAK

Tujuan projek ini adalah untuk membangunkan satu aplikasi yang mampu meningkatkan resolusi imej berkualiti rendah dan melakukan pemindahan gaya neural yang menggabungkan satu imej dalam gaya imej yang lain. Dalam era digital ini, imej memainkan peranan yang penting dalam komunikasi dan ekspresi diri. Walau bagaimanapun, imej berkualiti rendah merupakan masalah yang biasa dihadapi oleh ramai orang, terutamanya apabila mengambil gambar menggunakan telefon bimbit mereka. Kekurangan kualiti imej boleh menghasilkan imej yang kabur, berpiksel, atau mempunyai resolusi yang rendah. Imej berkualiti rendah ini boleh menyebabkan frustrasi kepada pengguna yang ingin berkongsi imej mereka di platform media sosial atau menggunakannya untuk tujuan peribadi atau profesional. Dalam projek ini, beberapa model rangkaian neural konvolusi mendalam pra-latih iaitu *Enhanced Deep Residual Network (EDSR)*, *Efficient Sub-Pixel Convolutional Neural Network (ESPCN)*, dan *deep Laplacian Pyramid Super-Resolution Network (LapSRN)* digunakan untuk meningkatkan resolusi imej, dan hasil setiap model telah dianalisis dan dibandingkan. *EDSR* dipilih sebagai model terbaik kerana ia telah mencapai nilai purata *PSNR* dan *SSIM* tertinggi dalam pengujian 45 imej iaitu 25.82 dB dan 0.70 masing-masing. Aplikasi pemprosesan imej tradisional sering kali tidak mempunyai keupayaan untuk melaksanakan teknik penambahbaikan imej yang canggih seperti pemindahan gaya neural, yang digunakan untuk mencipta kesan seni pada imej. Fungsi pemindahan gaya neural dicapai dengan mengekstrak gaya imej gaya menggunakan seni bina rangkaian *VGG19* yang pra-latih dan menggunakannya pada imej kandungan untuk mencipta kesan seni pada imej. *VGG19* digunakan kerana ia mencapai *ArtFID* purata tertinggi dalam 30 imej ujian, iaitu 45.97 berbandingkan kepada *MobileNet* dan *ResNet*. Akhirnya, satu aplikasi dibangunkan menggunakan *Flutter* yang menggabungkan semua fungsi di atas.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

PAGE

LIST OF FIGURES

PAGE

LIST OF ABBREVIATIONS

FYP	**-**	**Final Year Project**
CNN	**-**	**Convolution Neural Network**
PSNR	**-**	**Peak Signal to Noise Ratio**
SSIM	**-**	**Structural Similarity Index Measure**
MSE	**-**	**Mean Squared Error**
EDSR	**-**	**Enhanced Deep Residual Networks**
ESPCN	**-**	**Efficient Sub-Pixel Convolutional Neural Network**
LapSRN	**-**	**Deep Laplacian Pyramid Networks**
MCMC	**-**	**Malaysian Communications and Multimedia Commission**
API	**-**	**Application Programming Interface**
VGG19	**-**	**Very Deep Convolutional Networks for Large-Scale Image Recognition**
ResNet	**-**	**Residual Neural Network**
ArtFID	**-**	**Art Fréchet Inception Distance**

CHAPTER 1: INTRODUCTION

1.1 Introduction

In today's digital age, images play a crucial role in communication and self-expression. People take and share hundreds of photos every day on various social media platforms, and the demand for image enhancement tools has never been higher. According to a report by the Malaysian Communications and Multimedia Commission (MCMC) in 2021, approximately 94.8% of Malaysians own a smartphone, indicating a widespread adoption of mobile devices. Among the activities of smartphone users, 74.8% use smartphone to take photos or videos (MCMC Hand Phone Users Survey, 2021). This high penetration rate of smartphones has contributed to a significant increase in the number of photos taken and shared by Malaysians on various social media platforms. Furthermore, a study conducted by Ipsos Malaysia in 2020 revealed that 78% of Malaysians consider the visual quality of images to be essential when sharing them online (Ipsos Malaysia Digital Trends Survey, 2020). This statistic emphasizes the importance of image enhancement tools that can enhance the visual appeal and quality of photos, enabling users to create captivating and engaging content.

In this project, several pretrained deep convolutional neural networks model that are the Enhanced Deep Residual Network (EDSR), Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and deep Laplacian Pyramid Super-Resolution Network (LapSRN) are employed for increasing the resolution of image and the result of each model is analyzed and compared. The performance for each model is measured using metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). PSNR is a widely used metric to evaluate the quality of a reconstructed image. It measures the ratio between the maximum possible power of a signal (usually the original, unaltered image) and the power of the noise or distortion introduced by reconstruction. On the other hand, SSIM is a metric used to assess the similarity between two images. It is designed to capture both structural information and perceived changes in luminance, contrast, and structure (Horé & Ziou, 2013).

Traditional image processing applications often lack the ability to perform advanced image enhancement techniques such as neural style transfer, which can be used to create artistic effects on images. These techniques require a deep understanding of image processing and advanced technical knowledge, which is beyond the scope of most users. The neural style transfer functionality is achieved by extracting the style of the style image using the pretrained image classification network and applying it to the content image to create artistic effects on images. Several pretrained image classification networks such as Very Deep Convolutional Networks for Large-Scale Image Recognition (VGG19), MobileNet and Residual Neural Network (ResNet) network architecture are implemented to compare the quality of the output image between these deep CNN models. The performance for each deep CNN models on neural style transfer technique is evaluated using Art Fréchet Inception Distance (ArtFID) metric. The ArtFID metric is used for assessing the quality of neural style transfer technique and is inspired by the Fréchet Inception Distance (FID) that is used to evaluate the quality of generated images. ArtFID measures the perceptual similarity between the stylized image and a reference image, capturing both the content and style aspects (Wright & Ommer, 2022).

Therefore, there is a need for an application that combines multiple image enhancement tasks in a single, user-friendly platform. This is where ImageGenie comes in. ImageGenie aims to provide users with a unified solution for enhancing image resolution and applying artistic styles to their images.

1.2 Problem Statement

While image processing applications have become more accessible and user-friendly, they still fall short in meeting the needs of users who want to enhance their images without having advanced technical knowledge. Furthermore, low-quality images are a common problem that many people encounter, particularly when taking pictures with their mobile phones. The lack of image quality can result in images that are blurry, pixelated, or have poor resolution. These low-quality images can be frustrating for users who want to share their images on social media platforms or use them for personal or professional purposes. According to a survey conducted by Ipsos in Malaysia (2021), approximately 65% of social media users express dissatisfaction with the quality of images they encounter online. This statistic highlights the need for a solution that addresses image enhancement and empowers users to enhance their low quality images.

With the increasing popularity of social media platforms, people will share images that they think are attractive or interesting to social media. Moreover, many individuals appreciate the beauty of visually captivating artwork, artistic images have the power to evoke emotions, spark creativity, and inspire others. However, for individuals that are lacking the advanced drawing skills, creating such artistic artwork can be quite challenging.

1.3 Objective

This project embarks on the following objectives:

1. To increase the resolution of low quality images using EDSR.

2. To allow users to perform neural style transfer that can compose one super resolved image in the style of another image using VGG19.

1.4 Scope

The scope of the project includes development of a deep learning-based image processing application that can perform a range of image enhancement tasks, such as increasing the resolution of an image and applying neural style transfer.

1.5 Project Significance

By providing a comprehensive solution for image enhancement, ImageGenie will allow people to take their digital images to the next level, making it easier than ever to create and share high-quality, visually appealing content.

1.6 Expected Output

The expected output of the project is a powerful, all-in-one tool for image enhancement that will revolutionize the way people work with their digital images.

1.7 Conclusion

In conclusion, this project will use deep learning models to increase the resolution of low quality images and allow users to perform neural style transfer that can compose one image in the style of another image.

CHAPTER 2: LITERATURE REVIEW AND PROJECT METHODOLOGY

2.1 Introduction

In today's digital age, images have become a vital means of communication and self-expression. People capture and share hundreds of photos daily on various social media platforms, leading to a growing demand for image enhancement tools. Low-quality images are a common problem, particularly when capturing pictures with mobile phones. The lack of image quality can result in blurry, pixelated, or poorly resolved images, causing frustration for users who aim to share their visuals on social media platforms or use them for personal and professional purposes. This project aims to address these challenges by developing an application that employs pretrained deep convolutional neural network models to enhance image resolution and allow users to perform neural style transfer that can compose one image in the style of another image.

This chapter provides a comprehensive review of the literature relevant to this project. The method behind every image enhancement function which is super resolution and neural style transfer is studied. The research paper for every method for the purpose above is read and review. In order to achieve a better result, a thorough review of the references is also conducted.

All of the articles and journals that are related to this project are discussed and the method proposed in these articles is also analyzed in this chapter.

2.2 Facts and Findings

In this section, the detail of this project are concentrated in order to gain a better understanding of the project concept.

2.2.1 Domain

This project focuses on the domain of image enhancement, image processing, and neural networks. It explores various concepts and advancements in these domains to lay the foundation for the project's objectives. Key areas of interest include image resolution enhancement, neural style transfer and deep learning techniques. Understanding these domains is crucial for effectively implementing the project's objectives and achieving high-quality results.

2.2.2 Existing System

Other related projects that are related to this project will also be taken into consideration.

2.2.2.1 Image Super Resolution

Image super resolution not only can be done with deep learning technique, there are also other methods such as interpolation or super resolution forests method. The deep learning method of super resolution is employed in this project because the methods stated above cannot achieve good results of obtaining high quality of super resolved image compare to deep learning method.

(a) Interpolation Upsampling Methods (Ong Si Ci, 2023)

1. Nearest Neighbour Interpolation

This method is straightforward and efficient as it requires minimal calculations. It involves adding pixels based on the intensity values of neighboring pixels. While this approach increases image resolution, it may result in blocky and unnatural-looking images. Figure 2.1 below shows an example of a nearest neighbour interpolation with an upsampling factor of 2 applied to a 2x2 array.