ImageGenie: A Magic Wand For Images Enhancement Using Deep Learning

Sue Chen Xiang

Faculty of Information and Communication Technology, Technical University of Malaysia Malacca, Jalan Hang Tuah Jaya, 76100 Durian Tunggal Malacca, Malaysia

B032010034@student.utem.edu.my

***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.**

Keywords – image enhancement, image super resolution, neural style transfer, deep convolutional neural network, peak signal to noise ratio

I. 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[16]. 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[14]. 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)[1], Efficient Sub-Pixel Convolutional Neural Network (ESPCN)[2] and deep Laplacian Pyramid Super-Resolution Network (LapSRN)[3] 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 [4].

Traditional image processing applications often lack the ability to perform advanced image enhancement techniques such as neural style transfer[5], 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)[6], 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[7].

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.

The remainder of this paper is organized into the following sections. Section 2 describes the methodology of the proposed work which is followed by Section 3 which discusses the experimental results. Lastly, Section 4 presents the overall findings of this work and conclusions.

II. METHODOLOGY

To gain insights into the current scenario of ImageGenie, an extensive investigation was conducted into existing methods and tools[8]. Image super resolution has gained increasing attention for decades[9][10][11]. Super resolution aims to recreate a high resolution image from a low resolution image. Recently, deep neural networks provide significantly improved performance in terms of PSNR in the super resolution problem[1]. EDSR, ESPCN and LapSRN can be used to increase the resolution of low quality images. The models used to super resolve the low quality image is from the deep Convolution Neural Network (CNN) family. They each have their own unique architecture that produces different results. The Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) metrics are calculated to evaluate the performance of the deep learning models to increase the quality of poor image. Based on the metrics, the models are compared, and the best model is chosen to be implemented in the system. Figure 1 illustrate the flowchart for the comparison of the performance of super resolution models EDSR, ESPCN and LapSRN.

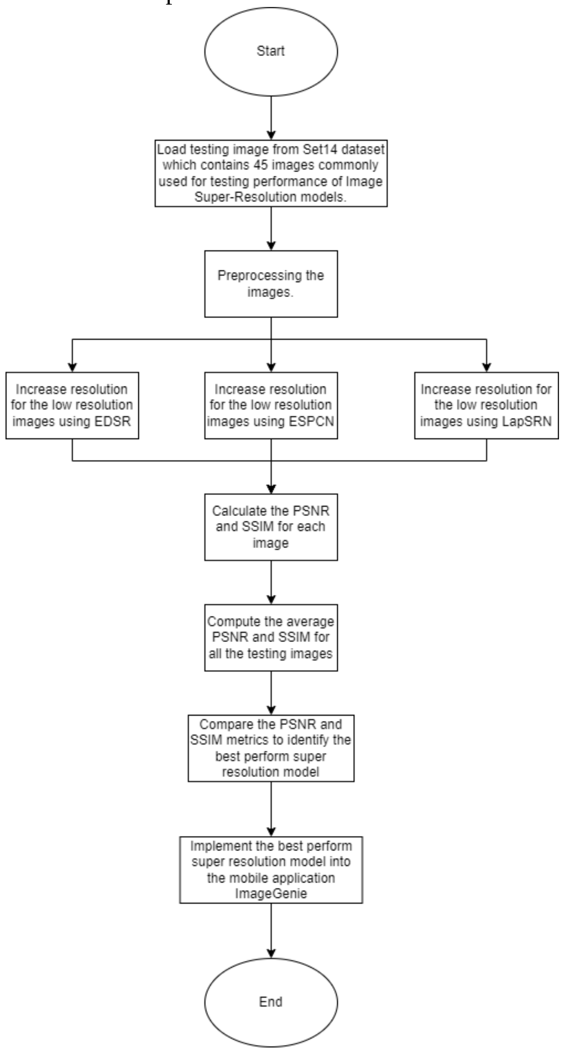

Figure 1: Flowchart for the Super Resolution Models Performance Comparison

For the implementation of neural style transfer on image, the neural representation of the content and style image need to be extracted clearly so that the reconstructed image can have the style of style image but do not lose their actual content. This can be accomplished by utilizing deep convolutional neural networks (CNN) that are trained on object recognition. Deep CNN models which are VGG19, MobileNet and ResNet are employed to capture the high level content and style features that are useful for our purpose. The ArtFID metric is used for assessing the quality of neural style transfer technique. The best deep CNN model was chosen to be implemented in the system. Figure 2 shows the flowchart for the comparison of the performance of deep CNN models in neural style transfer.

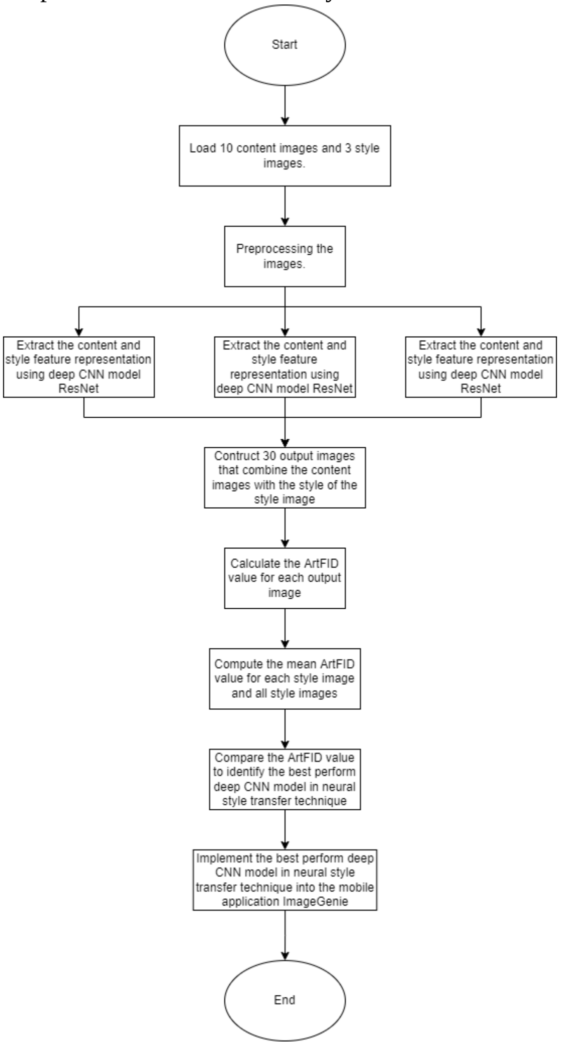

Figure 2: Flowchart for the deep CNN Models in Neural Style Transfer Performance Comparison

In this section, we identify the data requirements for this project. The system should have the capability to accept input images from users. These images can be uploaded through the application's user interface, providing a seamless experience for users to select images for enhancement. Additionally, the system needs to store various types of data in the server temporarily. This includes original images, output images and style images used for artistic effects.

The data used for testing the super resolution models are taken from publicly available benchmark dataset. The Set14 dataset is a dataset that consisting of 45 images commonly used for testing performance of image super resolution models [12]. It consists of low resolution image as well as the respective high resolution image version that are useful to evaluate the quality of the reconstructed image from image super resolution process.

Figure 3: Example 1 for Low Resolution Image Dataset

Figure 4: Example 2 for Low Resolution Image Dataset

On the other hand, the testing on the performance of neural style transfer CNN models ResNet, MobileNet and VGG19 use 10 custom content images with 3 style images. Each deep CNN models powered neural style transfer technique will experiment in processing 10 content images with 3 style images which will result in 30 output transfered image which combine the content images with the style of style images.

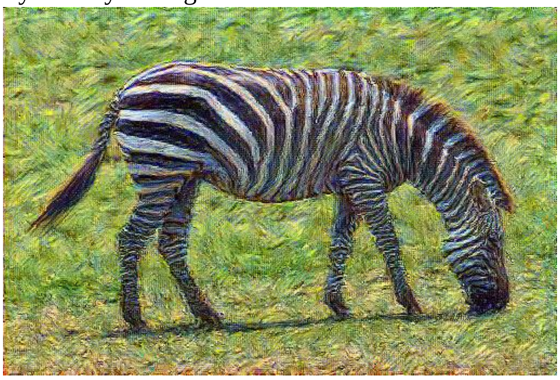

Figure 5: Output Image Example Using MobileNet

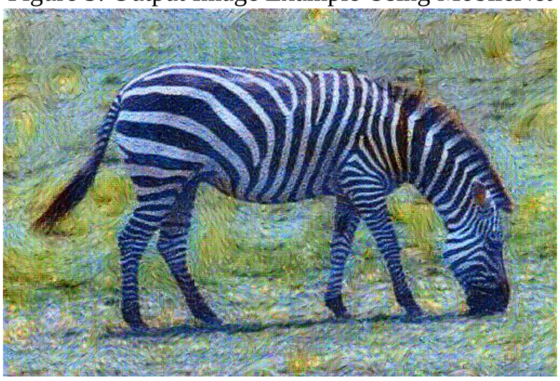

Figure 6: Output Image Example Using ResNet

Figure 7: Output Image Example Using VGG19

This project adopts the Agile methodology. Agile is an iterative development that emphasizes continuous incremental delivery of the application. At the same time, this allows action to incorporate any feedback to improve the application throughout the development process and adapts to any change in requirement. Figure 8 shows the Agile Methodology.

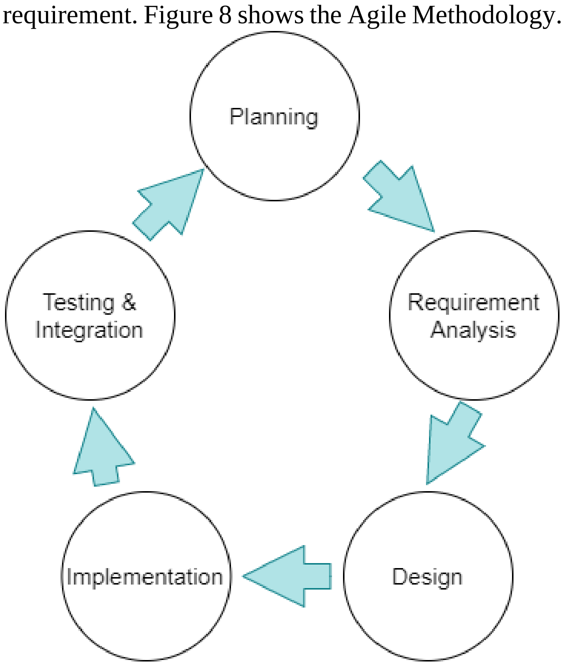

Figure 8: Agile Methodology

A. Planning Phase

The project planning phase includes choosing the project title, identify project objectives, scopes, milestones, and developing the Gantt Chart. The Gantt Chart serves as a visual representation of the project timeline, allowing for effective scheduling and task allocation. The functionality of the system is determined and the targeted operating system that the system will be running is also decided. The overall design architecture and the technology that are going to be used in the system is also identified.

B. Requirement Analysis

For requirement analysis, thorough research is done to analyze similar current applications and their advantages and limitations. Additionally, various methods applicable to the project's objectives are studied and compared to identify the most suitable approach. The respective research papers for all the deep learning models is studied and investigated in order to fully understanding the architecture and theory behind every deep learning models that enable it to perform respective image enhancement operations.

C. Design

This phase is the design of the project. For the backend part, Django is chosen as the framework that creates the Tensorflow deep learning api for processing the image. Django was chosen because of its modularity that makes it easy to maintenance when the code base is getting large. The api will be consumed by frontend or the mobile application using Flutter. Flutter was chosen because it is cross-platform, thus enables the code to be shared across multiple platforms (Android, IOS and window). Flutter is also easy to learn, and it provides material design applications which are attractive to the users. Moreover, Flutter can build applications that are as high-performance as native applications [13]. Figure 9 shows the architecture of the ImageGenie application.

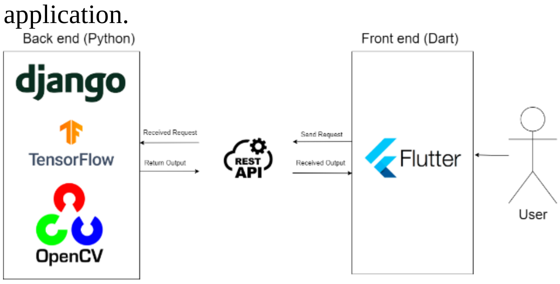

Figure 9: The design of the ImageGenie application

D. Implementation

For implementation, the mobile application as well as the Django backend is developed. The user can upload their images through the mobile application to the backend to carry out super resolution for the image and compose the image in the style of another image. The image received by the Django will pass to the Tensorflow and OpenCV deep learning api to carry out the task.

E. Testing and Integration

In this stage, it is crucial to ensure that the system runs smoothly and meets the expected requirements. The integration of the backend with the frontend is tested to prevent request time out problem. The result of each functionality provided by the api will be tested to ensure it is working and the processed image will be sent back to the mobile application successfully. Any issues or bugs identified during this phase are addressed, with the aim of ensuring a smooth and reliable user experience. In cases where the system encounters difficulties, the implementation phase is revisited to identify and rectify potential problems.

III. EXPERIMENT AND RESULT

A. User Interface Result

In this section, some interfaces of the proposed developed system will be shown and discussed. The result of the user interface for the developed mobile application is shown below. The mobile application is created using Flutter framework.

Figure 10: Splash Page

Figure 10 shows the splash page for ImageGenie which display a introduction page to user which consist of the icon of the application and animation of a circle running. After 2 seconds, the application will enter the home page.

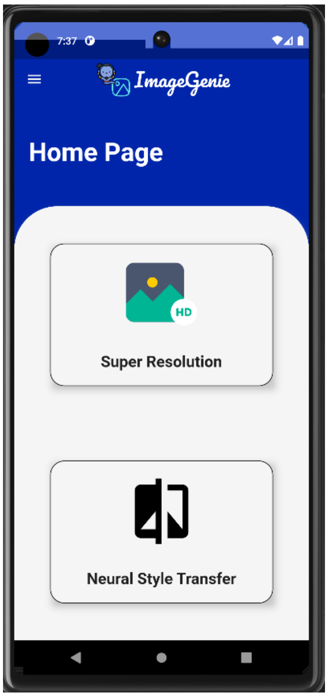

Figure 11: Home Page

Figure 11 shows the home page which consist of two card item which is clickable. The first card will navigate to the image super resolution page and the second card item will navigate to the neural style transfer page upon clicked.

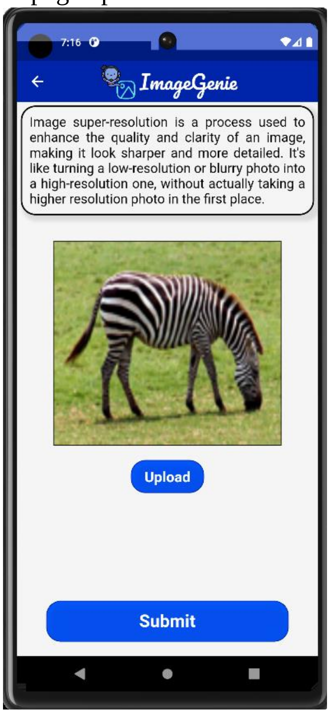

Figure 12: Image Super Resolution Page

Figure 12 illustrates the image super resolution page which include the description of what is image super resolution, a upload button which can select an image from the device file system to carry out the image super resolution task. At the bottom of the page, a submit button is used to send the image to the Django server to process the image.

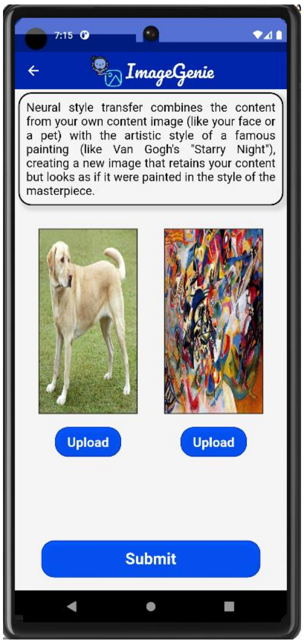

Figure 13: Neural Style Transfer Page

Figure 13 displays the neural style transfer page which consist of the description of what is neural style transfer, two upload button to upload content image and style image respectively. At the bottom of the page, a submit button is used to send the images to the Django server to process the image.

Figure 14: Loading Page After Submitted Input Image

Figure 14 shows the loading page for both the image super resolution and neural style transfer tasks. This page will indicate that the image is currently being processed at the server.

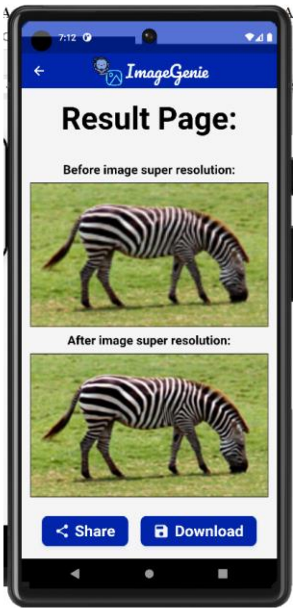

Figure 15: Result Page for Image Super Resolution

Figure 15 illustrates the result page for image super resolution task. The result shows the image before and after image super resolution. There is also share button which allow users to share the image to social media platform and a download button which can download the super resolved image to local device.

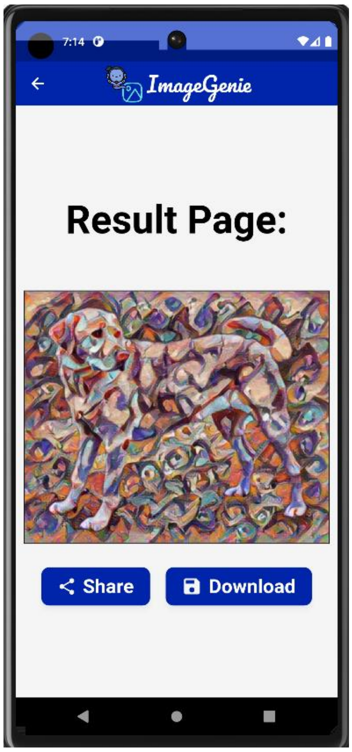

Figure 16: Result Page for Neural Style Transfer

Figure 16 displays the result page for neural style transfer task. There is also share button which allow users to share the image to social media platform and a download button which can download the super resolved image to local device.

B. Super Resolution Models Comparison Results

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. The PSNR is calculated using the mean squared error (MSE) between the original and reconstructed images and is typically expressed in decibels (dB). A higher PSNR value indicates a higher fidelity and lower distortion in the reconstructed image. The formula to calculate PSNR is shown below[4].

$$PSNR = 20log_{10}(\frac{\max_f}{\sqrt{MSE}})$$

Where the MSE (Mean Squared Error) is:

$$MSE = \frac{1}{mn}\sum_{0}^{m-1}\sum_{0}^{n-1}|f(i,j) - g(i,j)|^{2}$$

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. The final SSIM index value ranges between 0 and 1, with 1 indicating a perfect match and 0 indicating no similarity. Higher SSIM values correspond to images that are visually more similar. The formula to calculate SSIM is shown below[4].

$$SSIM(x,y) = \frac{2(\mu_x\mu_y + c1)(2\sigma_{xy} + c2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_X^2 + \sigma_y^2 + c2)}$$

Figure 17, Figure 18 and Figure 19 illustrate the examples of image before and after undergoing super resolution using different super resolution models which are EDSR, ESPCN and LapSRN with metrics PSNR and SSIM computed.

Figure 17: Example of Image Before and After Undergoing Super Resolution Using Different Super Resolution Models

In the example of Figure 17, we can see that the EDSR can obtain a fairly clear image espcially on the edges of the castle. The edges is clear and solid compare to that of ESPCN and LapSRN. The color of the super resolved image of EDSR is also more vivid. The leaves in the super resolved image of EDSR is also reverted to match the original image very well.

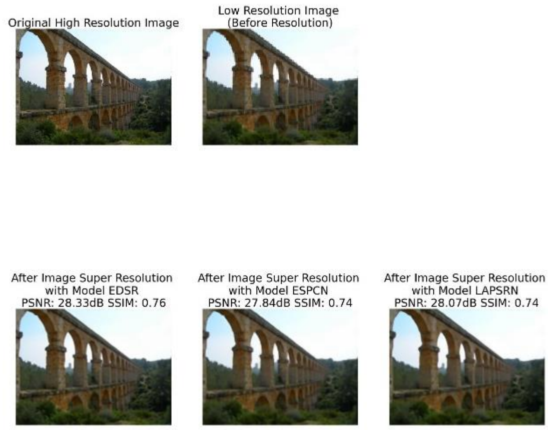

Figure 18: Example of Image Before and After Undergoing Super Resolution Using Different Super Resolution Models

In the example of Figure 18, we can see that the EDSR can obtain a fairly clear image espcially on the edges of the bridge. The edges is clear and solid compare to that of ESPCN and LapSRN. The color of the super resolved image of EDSR is also more vivid.

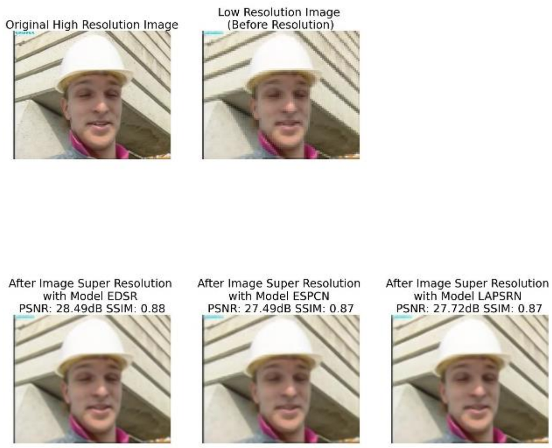

Figure 19: Example of Image Before and After Undergoing Super Resolution Using Different Super Resolution Models

In the example of Figure 19, we can see that the EDSR can obtain a fairly clear image espcially on the edges of the face. The edges is clear and solid compare to that of ESPCN and LapSRN. Table 1 below shows the comparison of EDSR, ESPCN and LapSRN on 45 testing images in term of PSNR metric.

Table 1: Comparison of EDSR, ESPCN and LapSRN on 45 testing images in term of PSNR

Testing Images	PSNR for EDSR	PSNR for ESPCN	PSNR for LapSRN
0	21.878399	21.449944	21.461944
1	23.315562	22.679062	22.735566
2	26.111846	25.313282	25.338782
3	28.331049	27.844156	28.073571
4	25.161231	24.436621	24.416166
5	24.352942	23.829916	23.892548
6	25.604176	25.391552	25.375444
7	25.746619	24.826793	24.889295
8	22.416045	21.423278	21.528038
9	24.121323	23.227436	23.210929
10	31.673301	31.020633	31.050307
11	29.853931	29.072989	29.123638
12	25.282401	24.149216	24.236310
13	35.376153	34.610614	34.602020
14	28.493322	27.494960	27.717854
15	23.246038	22.808271	22.829830
16	30.852457	29.913209	30.079588
17	28.971673	28.023187	28.053250
18	26.674827	26.089015	26.254745
19	24.294739	23.505603	23.593175
20	28.999248	27.205111	27.424549
21	25.095656	24.648887	24.682025
22	29.440361	28.488648	28.656179
23	27.699100	26.900454	26.888237
24	23.151439	21.905960	21.911420
25	23.581601	22.954031	22.955849
26	25.139445	24.023265	24.242495
27	30.320529	29.606569	29.604108
28	27.090382	26.480125	26.561322
29	23.193817	22.485449	22.519594
30	26.380273	25.700184	25.773560
31	23.599615	22.969184	23.033792
32	26.221630	25.722356	25.738186
33	19.742097	19.208276	19.265547
34	24.712881	24.069620	24.147578
35	22.491323	21.618717	21.774899
36	25.345032	24.810082	24.929146
37	23.374532	22.680546	22.778980
38	24.352436	23.756972	23.864585
39	29.037429	28.584109	28.596353
40	29.448629	28.780615	28.979260
41	24.223719	23.476029	23.534227
42	28.565813	28.091029	28.066656
43	19.889328	19.301676	19.382013
44	19.070460	18.629056	18.658960

Table 2 below shows the comparison of EDSR, ESPCN and LapSRN on 45 testing images in term of SSIM metric.

Table 2: Comparison of EDSR, ESPCN and LapSRN on 45 testing images in term of SSIM

Testing Images	PSNR for EDSR	PSNR for ESPCN	PSNR for LapSRN
0	0.472662	0.433278	0.431201
1	0.656359	0.620835	0.623640
2	0.758264	0.722497	0.722390
3	0.755064	0.738041	0.739238
4	0.644504	0.598876	0.596753
5	0.662501	0.637258	0.639125
6	0.572598	0.563236	0.560277
7	0.767148	0.737864	0.736743
8	0.632921	0.579156	0.584296
9	0.634278	0.602325	0.599813
10	0.782253	0.767527	0.766945
11	0.797340	0.782878	0.780250
12	0.745024	0.698185	0.699202
13	0.906031	0.896399	0.894557
14	0.882911	0.871294	0.872565
15	0.470871	0.439503	0.438019
16	0.842549	0.820028	0.821825
17	0.775000	0.749686	0.746631
18	0.750239	0.719723	0.721052
19	0.685578	0.653757	0.655652
20	0.925153	0.898761	0.902238
21	0.524979	0.498562	0.498662
22	0.867426	0.847479	0.847000
23	0.771575	0.746298	0.744603
24	0.887611	0.851408	0.851200
25	0.637554	0.608960	0.607616
26	0.735710	0.689339	0.684886
27	0.791200	0.768186	0.767234
28	0.776591	0.755821	0.756196
29	0.646438	0.600733	0.601888
30	0.682612	0.655961	0.656564
31	0.583133	0.543723	0.545462
32	0.617482	0.592250	0.589678
33	0.551272	0.506636	0.511104
34	0.688460	0.657782	0.660761
35	0.728327	0.691758	0.696020
36	0.634119	0.613759	0.615675
37	0.647141	0.609682	0.615518
38	0.652878	0.624623	0.626986
39	0.744774	0.726512	0.724134
40	0.822835	0.808772	0.810451
41	0.690474	0.655303	0.655718
42	0.733318	0.719072	0.715839
43	0.450788	0.402676	0.407999
44	0.430602	0.382027	0.386180

Figure 20 below shows the metric PSNR of EDSR, ESPCN and LapSRN to the testing images. Figure 21 below shows the metric SSIM of EDSR, ESPCN and LapSRN to the testing images. The results are tested on 45 low resolution images which each image will be super resolved using EDSR, ESPCN and LapSRN respectively and the results are compared. The average PSNR for EDSR is 25.82dB, 25.09dB for ESPCN and 25.16dB for LapSRN. The average SSIM for EDSR is 0.70, 0.66 for ESPCN and 0.66 for LapSRN. EDSR obtains the best results in both of the metrics with 25.82dB in PSNR and 0.70 in SSIM. Table 3 below shown the average PSNR and SSIM for each super resolution models tested on 45 images.

Table 3: Average PSNR and SSIM for Super Resolution Models

Super Resolution Models	**Average PSNR (dB)**	**Average SSIM**
EDSR	25.82	0.70
ESPCN	25.09	0.66
LapSRN	25.16	0.66

Based on Figure 20 and Figure 21, EDSR achieved the highest performance in terms of PSNR and SSIM. This means that EDSR can improve the quality of the image with the best results. Hence, we will utilize EDSR models since it provides the best result for image super resolution.

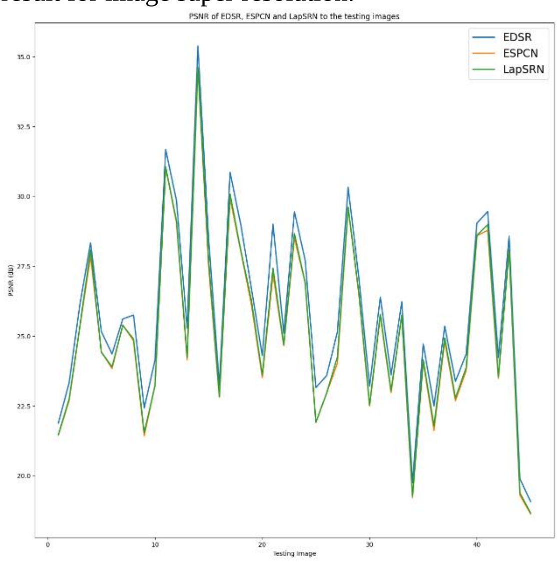

Figure 20: PSNR of EDSR, ESPCN and LapSRN to the testing images

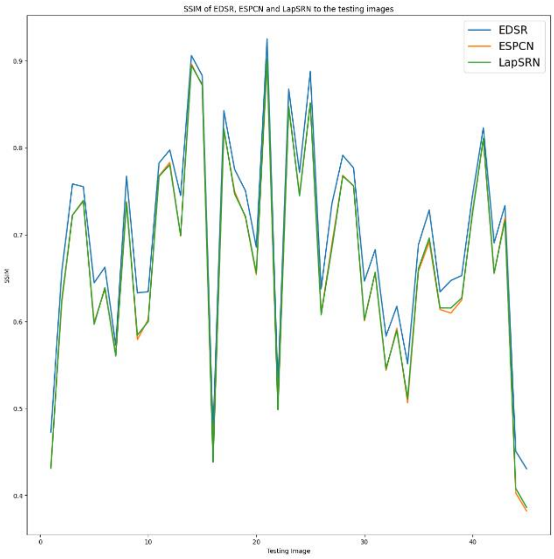

Figure 21: SSIM of EDSR, ESPCN and LapSRN to the testing images

C. Neural Style Transfer CNN Models Comparison Results

The test strategy in the neural style transfer is that different CNN models which are ResNet, MobileNet and VGG19 were employed and the performance or quality was measured using evaluation metric ArtFID. 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. The authors conduct extensive experiments using various neural style transfer algorithms and datasets that are labeled by artists to demonstrate the effectiveness of the ArtFID metric [7]. They compare ArtFID against other metrics, showcasing its ability to align better with human perception of style transfer quality. Equation below depicts the formula to calculate ArtFID [7].

$$ArtFID(X_g, X_c, X_s) = (1 + \frac{1}{N}\sum_{i=1}^{N} d\left(X_c^{(i)}, X_g^{(i)}\right)) \cdot (1 + FID(X_s, X_g))$$

Each CNN models powered neural style transfer technique will experiment in processing 10 content images with 3 style images which will result in 30 output image which combine the content images with the style of style images. The 30 images will then be evaluated using ArtFID metric with higher values indicating better performance.

A. Style Image 1

Figure 22: Style Image 1

Figure 22 above shows one of the style image used in evaluating the quality of the neural style transfer technique. Table 4 below displays the ArtFID value computer for all the CNN models for style image 1 with 10 content images.

Table 4: ArtFID Value of CNN Models for Style Image 1

Testing Image	ArtFID for ResNet	ArtFID for MobileNet	ArtFID for VGG19
1	48.61	47.00	51.80
2	54.74	42.43	54.94
3	59.67	52.33	62.55
4	61.90	54.27	67.69
5	76.65	61.49	74.59
6	76.88	62.30	62.35
7	58.84	64.41	60.92
8	57.67	52.44	58.81
9	62.33	53.95	70.65
10	65.96	57.15	71.48

Table 5: ArtFID Value Properties for Style Image 1

Models	Mean	Variance	Standard Deviation
ResNet	62.32	71.69	8.46
MobileNet	54.77	42.54	6.52
VGG19	63.57	49.63	7.04

Table 5 shows the ArtFID value properties which include mean, variance and standard deviation. MobileNet has the lowest mean ArtFID value which is 54.77 among the three models, indicating that it, on average, performs worse in preserving the style from style image 1 and content of the original images during the neural style transfer process. VGG19 has the highest mean ArtFID value which is 63.57, suggesting that it perform better than the other models in terms of style preservation. ResNet has the highest standard deviation among the models, indicating that its ArtFID values are more spread out, possibly resulting in a wider range of performance outcomes. MobileNet has the lowest standard deviation, suggesting that its performance is more consistent across the evaluated images.

Hence, VGG19 appears to be the most effective model among the three models, as it has the highest mean ArtFID value in the neural style transfer technique done toward 10 content images and style image 1.

B. Style Image 2

Figure 23: Style Image 2

Figure 23 above shows one of the style image used in evaluating the quality of the neural style transfer technique. Table 6 below displays the ArtFID value computer for all the CNN models for style image 2 with 10 content images.

Table 6: ArtFID Value of CNN Models for Style Image 2

Testing Image	ArtFID for ResNet	ArtFID for MobileNet	ArtFID for VGG19
11	29.74	27.27	37.83
12	30.74	37.66	39.22
13	26.70	37.20	40.55
14	34.48	37.59	44.94
15	34.85	49.55	49.33
16	29.73	45.25	34.33
17	33.68	35.02	48.95
18	35.50	38.44	44.76
19	43.69	50.53	50.67
20	33.85	37.69	49.34

Table 7: ArtFID Value Properties for Style Image 2

Models	Mean	Variance	Standard Deviation
ResNet	33.29	19.25	4.38
MobileNet	39.62	44.19	6.64
VGG19	43.99	29.37	5.41

Table 7 shows the ArtFID value properties which include mean, variance and standard deviation. ResNet has the lowest mean ArtFID value which is 33.29 among the three models, indicating that it, on average, performs worse in preserving the style from style image 2 and content of the original images during the neural style transfer process. VGG19 has the highest mean ArtFID value which is 43.99, suggesting that it perform better than the other models in terms of style preservation. MobileNet has the highest standard deviation among the models, indicating that its ArtFID values are more spread out, possibly resulting in a wider range of performance outcomes. ResNet has the lowest standard deviation, suggesting that its performance is more consistent across the evaluated images.

Hence, VGG19 appears to be the most effective model among the three models, as it has the highest mean ArtFID value in the neural style transfer technique done toward 10 content images and style image 2.

C. Style Image 3

Figure 24: Style Image 3

Figure 24 above shows one of the style image used in evaluating the quality of the neural style transfer technique. Table 8 below displays the ArtFID value computer for all the CNN models for style image 3 with 10 content images.

Table 8: ArtFID Value of CNN Models for Style Image 3

Testing Image	**ArtFID for ResNet**	**ArtFID for MobileNet**	**ArtFID for VGG19**
21	31.02	22.23	25.30
22	37.55	30.50	30.12
23	39.10	31.91	31.08
24	31.39	32.95	39.86
25	41.99	27.99	35.32
26	46.50	30.60	29.79
27	30.37	28.47	36.22
28	32.86	27.38	29.45
29	28.70	36.71	28.61
30	34.07	29.60	33.63

Table 9: ArtFID Value Properties for Style Image 3

Models	**Mean**	**Variance**	**Standard Deviation**
ResNet	35.35	29.86	5.46
MobileNet	29.83	13.14	3.62
VGG19	31.93	16.53	4.06

Table 9 shows the ArtFID value properties which include mean, variance and standard deviation. MobileNet has the lowest mean ArtFID value which is 29.83 among the three models, indicating that it, on average, performs worse in preserving the style from style image 3 and content of the original images during the neural style transfer process. ResNet has the highest mean ArtFID value which is 35.35, suggesting that it perform better than the other models in terms of style preservation. ResNet has the highest standard deviation among the models, indicating that its ArtFID values are more spread out, possibly resulting in a wider range of performance outcomes. MobileNet has the lowest standard deviation, suggesting that its performance is more consistent across the evaluated images.

Hence, ResNet appears to be the most effective model among the three models, as it has the highest mean ArtFID value in the neural style transfer technique done toward 10 content images and style image 3. In my opinion, ResNet stand out in style image 3 because the colour combination for style image 3 is more contrast and complex than that of style image 1 and 2.

D. All Datasets

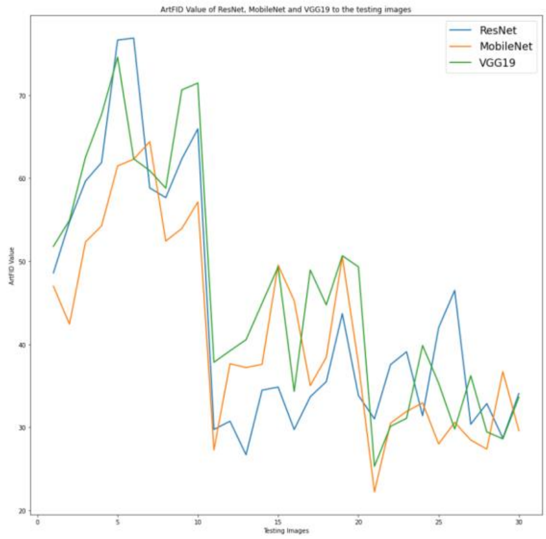

Figure 25: ArtFID Value for ResNet, MobileNet and VGG19 on Neural Style Transfer Technique Tested on 30 Images

Figure 25 above displays the graph of ArtFID value for ResNet, MobileNet and VGG19 on neural style transfer technique tested on 30 images.

Table 10: ArtFID Value Properties for all Dataset

Models	**Mean**	**Variance**	**Standard Deviation**

ResNet	42.10	40.26	6.10
MobileNet	39.96	33.29	5.59
VGG19	45.97	31.84	5.50

Table 10 shows the ArtFID value properties which include mean, variance and standard deviation for all three style images. MobileNet has the lowest mean ArtFID value which is 39.96 among the three models, indicating that it, on average, performs worse in preserving the style and content of the original images during the neural style transfer process. VGG19 has the highest mean ArtFID value which is 45.97, suggesting that it performs better than the other models in terms of style preservation. ResNet has the highest standard deviation among the models, indicating that its ArtFID values are more spread out, possibly resulting in a wider range of performance outcomes. VGG19 has the lowest standard deviation, suggesting that its performance is more consistent across the evaluated images.

In conclusion, VGG19 appears to be the most consistent and effective model among the three models, as it has the highest mean ArtFID value in the neural style transfer technique done toward 10 content images and three style images.

IV. CONCLUSION

Throughout the development of this project, we observed both strengths and weaknesses in the development of the system, ImageGenie, an mobile application designed to enhance image resolution and perform neural style transfer. First and foremost, we will discuss the project's strengths. The mobile application has a user friendly interface as it enables users to seamlessly engage with the image enhancement functions. Besides, this project provide a quick and easy way to enhance the quality of low-resolution image and perform neural style transfer. However, this project also have weaknesses, including huge computing power is required to perform the image enhancement operations. Without an excellent GPU, the inference time of the image operation will be longer. Other than that, while ImageGenie generally delivered reliable results, there was still variability in image enhancement outcomes. This variability, primarily observed in the performance of VGG19 in neural style transfers, suggests room for improvement.

In this project, several propositions for improvement can be considered. One of the improvement that can be done is continuously refining and incorporating state-of-the-art deep learning models for image resolution enhancement can elevate the application's performance. Another improvement that can also be done is to further enhance the quality and consistency of neural style transfer, exploring advanced neural style transfer models beyond VGG19 that could yield improved artistic effects. Last but not least, we can also expanding the application's capabilities to include a broader scope of image processing techniques, such as object recognition, segmentation, and image denoising which would increase its versatility and utility.

To sum up, this project can serve as a valuable reference for researchers and students interested in image processing and AI applications with the successful implementation and evaluation of various deep learning models for image enhancement, along with the development of a user-friendly application. ImageGenie make it a useful tool for individuals seeking to improve image quality and create captivating artistic images.

In conclusion, this project aimed to address the growing demand for image enhancement tools in the digital age. Through the integration of advanced deep learning models and a user-friendly interface, ImageGenie successfully achieved its objectives of enhancing image resolution and enabling neural style transfers. While the project demonstrated strengths in terms of effective image enhancement, consistency, and usability, there is room for improvement in reducing performance variability and expanding the application's feature set. Overall, ImageGenie represents a promising step toward providing users with accessible and powerful image enhancement capabilities. As technology continues to advance, there are exciting opportunities to further refine and expand this application to meet evolving user needs and expectations in the field of image processing and artificial intelligence.

REFERENCES

[1] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," Jul. 2017, [Online]. Available: http://arxiv.org/abs/1707.02921

[2] W. Shi *et al.*, "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network," Sep. 2016, [Online]. Available: http://arxiv.org/abs/1609.05158

[3] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Fast and Accurate Image Super-Resolution with Deep Laplacian Pyramid

Networks," Oct. 2017, [Online]. Available: http://arxiv.org/abs/1710.01992

[4] A. Horé and D. Ziou, "Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?," *IET Image Process*, vol. 7, no. 1, pp. 12–24, 2013, doi: 10.1049/iet-ipr.2012.0489.

[5] L. A. Gatys, A. S. Ecker, and M. Bethge, "A Neural Algorithm of Artistic Style," Aug. 2015, [Online]. Available: http://arxiv.org/abs/1508.06576

[6] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.1556

[7] M. Wright and B. Ommer, "ArtFID: Quantitative Evaluation of Neural Style Transfer," Jul. 2022, [Online]. Available: http://arxiv.org/abs/2207.12280

[8] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," Mar. 2016, [Online]. Available: http://arxiv.org/abs/1603.08155

[9] Ong Si Ci, "Image Super Resolution: A Comparison between Interpolation & Deep Learning-based Techniques to Improve Clarity of Low-Resolution Images." Accessed: Jun. 24, 2023. [Online]. Available: https://medium.com/htx-s-s-coe/image-super-resolution-a-comparison-between-interpolation-deep-learning-based-techniques-to-25e7531ab207

[10] S. Schulter, C. Leistner, and H. Bischof, "Fast and Accurate Image Upscaling with Super-Resolution Forests," 2015. Accessed: Jun. 24, 2023. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Schulter_Fast_and_Accurate_2015_CVPR_paper.pdf

[11] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs," Nov. 2017, [Online]. Available: http://arxiv.org/abs/1711.11585

[12] R. Zeyde, M. Elad, and M. Protter, "On Single Image Scale-Up Using Sparse-Representations," in *Curves and Surfaces*, J.-D. Boissonnat, P. Chenin, A. Cohen, C. Gout, T. Lyche, M.-L. Mazure, and L. Schumaker, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 711–730. Accessed: Sep. 19, 2023. [Online]. Available: https://doi.org/10.1007/978-3-642-27413-8_47

[13] A. Tashildar, N. Shah, R. Gala, T. Giri, and P. Chavhan, "APPLICATION DEVELOPMENT USING FLUTTER," 2020. Accessed: May 28, 2023. [Online]. Available: https://www.irjmets.com/uploadedfiles/paper/volume2/issue_8_august_2020/3180/1628083124.pdf

[14] Ipsos Malaysia. (2020). Digital Trends Survey. Available: https://www.ipsos.com/en-my

[15] Ipsos Malaysia. (2021). Digital 2021: Malaysia. Available: https://datareportal.com/reports/digital-2021-malaysia

[16] Malaysian Communications and Multimedia Commission (MCMC). (2021). Hand Phone Users Survey. Available: https://www.mcmc.gov.my/skmmgovmy/media/General/pdf2/FULL-REPORT-HPUS-2021.pdf