Bachelor of Science in Computer Science & Engineering

Upscaling and Reconstructing Degraded Images Using Convolutional Neural Network

by

Fahmid Rahim Noor

ID: 1504028

Department of Computer Science & Engineering

Chittagong University of Engineering & Technology (CUET)

Chattogram-4349, Bangladesh.

May, 2021

Upscaling and Reconstructing Degraded Images Using Convolutional Neural Network

Submitted in partial fulfilment of the requirements for
Degree of Bachelor of Science
in Computer Science & Engineering

by

Fahmid Rahim Noor
ID: 1504028

Supervised by

Muhammad Kamal Hossen
Associate Professor
Department of Computer Science & Engineering

Chittagong University of Engineering & Technology (CUET)
Chattogram-4349, Bangladesh.

The thesis titled '**Upscaling and Reconstructing Degraded Images Using Convolutional Neural Network'** submitted by ID: 1504028, Session 2019-2020 has been accepted as satisfactory in fulfilment of the requirement for the degree of Bachelor of Science in Computer Science & Engineering to be awarded by the Chittagong University of Engineering & Technology (CUET).

Board of Examiners

_______________________________________ Chairman

Muhammad Kamal Hossen
Associate Professor
Department of Computer Science & Engineering
Chittagong University of Engineering & Technology (CUET)

_______________________________________ Member (Ex-Officio)

Dr. Md. Mokammel Haque
Professor & Head
Department of Computer Science & Engineering
Chittagong University of Engineering & Technology (CUET)

_______________________________________ Member (External)

Dr. Abu Hasant Mohammad Ashfak Habib
Professor
Department of Computer Science & Engineering
Chittagong University of Engineering & Technology (CUET)

Declaration of Originality

This is to certify that I am the sole author of this thesis and that neither any part of this thesis nor the whole of the thesis has been submitted for a degree to any other institution.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. I am also aware that if any infringement of anyone's copyright is found, whether intentional or otherwise, I may be subject to legal and disciplinary action determined by Dept. of CSE, CUET.

I hereby assign every rights in the copyright of this thesis work to Dept. of CSE, CUET, who shall be the owner of the copyright of this work and any reproduction or use in any form or by any means whatsoever is prohibited without the consent of Dept. of CSE, CUET.

Signature of the candidate

Date:

Acknowledgements

The student should acknowledge advice, service encouragement, library and information service support and source of financial support.

Abstract

We propose a deep learning method to produce reconstructed high quality image from low quality degraded image. We developed a convolution neural network that contains five layers. Features extraction and concatenation are performed in first three layer. The fourth layer maps these feature maps non linearly to high resolution patch representations. The last layer combines the predictions within a spatial neighbourhood to produce the final high quality image. The quality of the outputs of test images are measured by defining three image quality metrics PSNR, MSE and SSIM. Then the performance of our network is compared with other two networks.

Keywords: Image reconstruction, Convolution neural network, PSNR, MSE, SSIM

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

The aim of single image super resolution is to create a high-resolution image from a low-resolution image. This is a well-known problem of computer vision. When you only have a single low-resolution image, the challenge becomes much more challenging to reproduce its high-resolution output.

Super resolution is the measurement of a high resolution (hr) image from a single low resolution (lr) equivalent (sr). To put it another way, lr represents a single image input, hr represents the ground reality, and sr represents the projected high resolution output image. The lr images used in ML solutions are usually down sampled hr images with some blurring and noise applied.

Convolutional neural networks have been around for decades, but deep CNNs have recently exploded in popularity, thanks in part to their success in image classification. They've also been used in other computer vision fields including object detection, face recognition, and pedestrian detection with great success. Several factors are critical to this advancement: (i) efficient training on modern strong GPUs, (ii) the proposal of the rectified linear unit (ReLU), which speeds up convergence while maintaining high quality, and (iii) quick access to a large amount of data for training larger models. These advancements support our system in this project as well.

1.2 Image Reconstruction Neural Network Design

In the problem of reconstructing a super resolution image from a low resolution image, recent deep learning-based methods (especially with deeply and completely convolutional networks) have achieved high efficiency. This is because deep learning, through cascading CNNs and nonlinear layers, can gradually understand both local and global structures on a picture at the same time.

This project's neural network has five layers. From the input image, the first two convolution layers extract a collection of features. The low-level features are extracted by the first sheet, while the high-level features are extracted by the second. The third layer joins the first two layers together. The fourth layer converts these feature maps to high-resolution patch representations in a nonlinear fashion. To generate the final high-resolution reconstructed image, the final layer combines the predictions within a spatial neighborhood.

1.3 Difficulties

Reconstructing a high-resolution photorealistic image from a low-resolution equivalent has long been a difficult task in computer vision. There are numerous challenges in this operation, including:

1. Variations of images

2. Extracting features like edges, corners, interest points etc from different images in a generalized way to train the neural network

1.4 Applications

Single image super resolution was mainly used for specific fields like-

1. CCTV footage surveillance
2. Satellite image processing
3. Better medical image analysis
4. Better microscopic image analysis

However, as monitor resolutions continue to rise, it is now commonly used in television, video games, and websites.

1.5 Motivation

To keep up with the progress in technology in the field of computer vision, there is room to improve current approaches in order to achieve more accurate results. The following is a list of the key motivations for this thesis:

1. To produce better quality images for different analysis and research purposes
2. To overcome the limitations of the highly expensive sensors and optics manufacturing technology

1.6 Contribution of the thesis

Thesis or research work is done to meet a certain set of objectives, such as the quality of findings or the time it takes to process data. The thesis' areas of contribution can be summarized as follows:

1. Implement the neural network according to the design.
2. Train the network with a pretty large amount of images.
3. Set up the programs in such a way that we can directly upload degraded images to the network and get high-quality images in return.

1.7 Thesis Organization

The rest of this thesis report is organized as follows:

i. The second chapter provides a synopsis of previous studies in the field of image restoration from low-resolution degraded images.

ii. Chapter 3 describes the proposed methodology for image reconstruction using convolution neural network. The neural network in this project has five layers. The first two convolution layers are for feature extraction. The third layer concatenates the first two layers. The fourth layer is for second order mapping and the fifth layer is where the reconstruction happens.

iii. The working data set is defined in Chapter 4 along with an overview of the performance measure for the proposed system.

iv. The overall overview of this thesis work is presented in Chapter 5, along with some possible suggestions.

1.8 Conclusion

This chapter gives an overview of how to use a convolution neural network to upscale and recreate degraded photos. This chapter also covers the design of the neural network, as well as the difficulties. The inspiration for this work, as well as the contributions made, are also listed here. The history and current state of the problem will be discussed in the following chapter.

Chapter 2

Literature Review

2.1 Introduction

The focus of this thesis is to investigate the challenges that are to be faced while conducting the research on up scaling and reconstruction of degraded images and to present a comprehensive review of various image improving approaches. Through providing a brief summary of previous study, this chapter discusses various deep learning methods applied by different researches, and the performances of these researches on different data sets.

2.2 Related Literature Review

In [1] for single image super-resolution, the author proposed a deep learning system. Their method learns an end-to-end mapping between low-resolution and high-resolution images directly. The mapping is expressed by a deep convolutional neural network that takes the low-resolution image as input and produces the high-resolution image as output. The conventional sparse-coding-based SR methods can also be viewed as a deep convolutional network, according to the author. Unlike conventional approaches, which optimize each component separately, this approach optimizes all layers at the same time.

In [2] centered on sparse signal representation, the author suggested a new approach to single-image super resolution. Image patches can be well-represented as a sparse linear combination of elements from an appropriately chosen over-complete dictionary, according to image statistics research. They seek a sparse representation for each patch of the low-resolution input, and then use the coefficients of this representation to produce the high-resolution output, inspired by this observation.

In [3] the author suggested a perceptual loss feature that includes an adversarial loss as well as a material loss. Using a discriminator network that is equipped to distinguish between super-resolved images and original photo-realistic images, the adversarial loss moves the solution to the natural image manifold. Furthermore, the author used a content loss inspired by perceptual similarity rather than pixel space similarity.

In [4] the authors' method is based on the observation that patches in a natural image appear to recur several times within the image, both within the same scale and across scales. Classical super-resolution is achieved by recurrence of patches within the same image scale (at sub pixel misalignments), while example-based super-resolution is achieved by recurrence of patches across different scales of the same image. Their method tries to recover the best possible resolution increase for each pixel based on patch redundancy within and across scales.

In [5] to fix this issue, the author proposed a new probabilistic deep network architecture, a pixel recursive super resolution model, which is an extension of PixelCNNs. They show that this model can generate a wide range of plausible high-resolution images at high magnification factors. In addition, this paper discusses how previous approaches struggle to trick human observers in human assessment studies.

In [6] the author took a compressed sensing approach to solving this problem. The low-resolution image is presented as a down scaled version of a high-resolution image, with patches believed to have sparse representations in comparison to an over-complete dictionary of prototype signal-atoms. The compressed sensing theory guarantees that the sparse representation can be correctly retrieved from the down sampled signal under mild conditions. The authors also showed that sparsity can be used as a prior to regularize the otherwise ill-posed super-resolution problem.

2.3 Conclusion

This chapter contains a concise overview of the literature review. The proposed technique for up-scaling and reconstructing degraded images using neural networks is explained in detail in the following chapter.

Chapter 3

Methodology

3.1 Introduction

It's never easy to create a high-resolution restored image from a low-resolution degraded image. To complete this mission, we first create a neural network that learns a direct end-to-end mapping between low and high resolution images. After that, we take some low-resolution degraded images and feed them into our qualified neural network, which produces high-quality reconstructed images after some pre-processing.

3.2 Diagram/Overview of Framework

The working process of the neural network is given below:

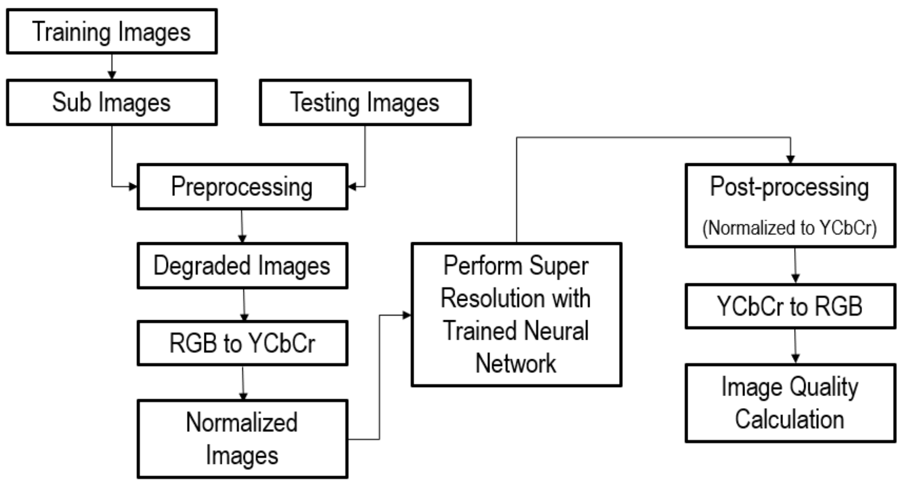

Figure 3.1: Steps of the proposed framework.

We take the original images of smaller amount and produce sub images (32X32) of larger amounts. The neural network is trained on sub images and the presumption of the whole image.

Low resolution degraded images are produced from these sub images of high quality. These images are in RGB format. Then they are converted to YCbCr color space which is a three channel image.

After that these images are normalized and fed into the neural network which contains five layers. These layers extracts features from the images, combines them, remap them and produce reconstructed output. Thus the neural network is trained with the sub images.

With the testing images we first make degraded images, convert RGB to YCbCr, normalize and feed them into the trained neural network. Neural network then produces reconstructed image gives us output which is normalized. So we post process the output, convert the normalized image to YCbCr color space and then convert it from YCbCr to RGB format.

After that we define three image quality metrics PSNR, MSE and SSIM. Then we compare the degraded image and the reconstructed image by using the ground truth original test image and see how good the result is.

3.3 Image Pre-processing

We take the original images of smaller amount and produce sub images (32X32) of larger amounts. The neural network is trained on sub images and the presumption of the whole image.

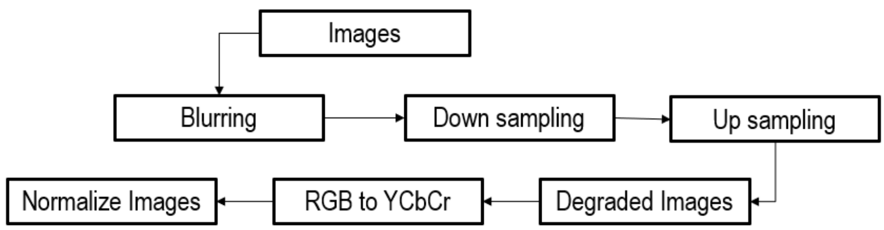

Figure 3.2: Overview of Pre-processing steps.

Low resolution degraded images are produced from these sub images of high quality. In order to do that, the sub images are first blurred using a Gaussian kernel. Then these images are down-sampled using the down scaling factor.After that they are again up-sampled using bicubic interpolation.

The neural network is trained on the luminance channel of the YCrCb color space. So we have to convert the image to YCrCb which is a three channel image ('Y' is the luminance, 'Cr' is the red difference and 'Cb' is the blue difference).

OpenCV is used in this project because it does a very good job of converting back and forth between this color space.

When OpenCV loads an image for the first time it is in BGR format, not RGB. That's why BGR to YCrCb is used.

These preprocessing steps are also applied on the test images.

3.4 Neural Network Training

The working process of the neural network is given below:

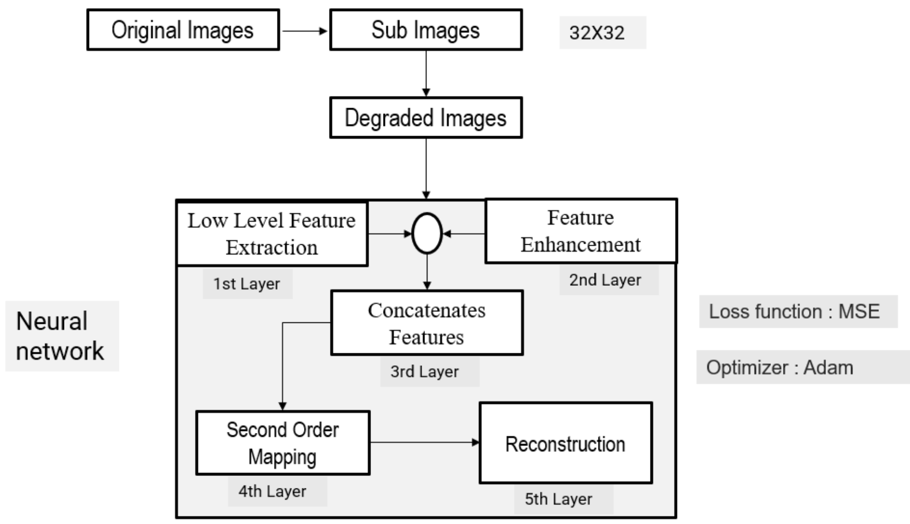

Figure 3.3: Working process of neural network.

The low resolution degraded images are then fed into the neural network which has five layers.

3.4.1 Low Level Feature Extraction

The first layer extracts different types of features from the input image. In this case, the first layer contains 64 feature maps or filters (n1) and filter size is 9X9. This layer then computes the low level features such as color, shape, texture etc.

$$F1(Y) = max(0, W1 * Y + B1) \tag{3.1}$$

W1 contains n1 filters of size c * f1 *f1 that produces n1 feature maps. c is number of channels of the image and f1 is the filter size. Here, c=1, f1=9, B1=n1=64.

3.4.2 Enhanced Feature Extraction

The second layer is a feature enhancement layer, which adds new features to the first layer's performance. This layer also contains 64 filters (n2) and filter size is 5X5.

$$F2(Y) = max(0, W2 * F1(Y) + B2) \tag{3.2}$$

W2 contains n2 filters of size c * f2 *f2 that produces n2 feature maps. Here, c=1, f2=5, B2=n2=64.

3.4.3 Concatenation

The features formed by the first two layers are concatenated in the third layer. To build a merged vector, this layer combines the low level and enhanced functionality. Since it allows for input of various sizes, concatenation is used as the merge method here. Different input sizes are not allowed in other merge operations such as summation, limit, averaging, or multiplication.

$$F3(Y) = merge(F1(Y), F2(Y)) \tag{3.3}$$

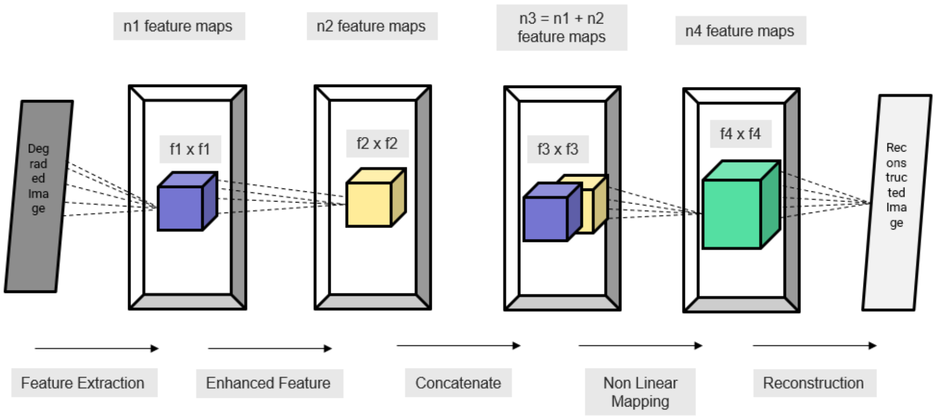

Figure 3.4: Neural network architecture.

3.4.4 Second Order Mapping

In the fourth layer, the second order mapping is performed. Each high-dimensional vector from the previous layer is mapped to another high-dimensional vector in the second hidden layer. ReLU activation function is applied in this layer with 128 feature maps (n4).

$$F4(Y) = max(0, W4 * F3(Y) + B4) \tag{3.4}$$

W4 contains n4 filters of size n2 * f4 *f4 that produces n4 feature maps. Here, n2=64, f4=1, B4=n4=128.

3.4.5 Reconstruction

The final fifth layer reconstructs the output high resolution image.

$$F5(Y) = W5 * F4(Y) + B5 \tag{3.5}$$

W5 contains c filters of size n4 * f5 *f5 that produces n4 feature maps. Here, n4=128, f5=5, B5=c=1.

3.5 Image Post-processing

The output is still normalized between 0 and 1 So we have to multiply every single pixel by 255 and convert the float values to integer. Thus the normalized images are back to YCbCr color space. Then we convert the image from YCbCr color space to RGB format.

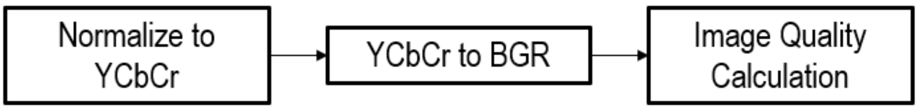

Figure 3.5: Post processing steps.

We define three image quality metrics: Peak signal to noise ratio (PSNR), Mean square error (MSE) and Structural similarity (SSIM). These image quality metrics compare the degrade images with the original images and compare the reconstructed images with the original images. By calculating these values we can see that how the neural network improved the quality of the degraded images by producing the reconstructed output images.

If the original image is not available, we work directly with the degraded image and generate the reconstructed output. We can't compare them mathematically to those image quality metrics in this case because we don't have the original ground truth image. To grasp the improvement, we can only compare them visually.

3.6 Conclusion

The methodology of up-scaling and reconstructing images are described in this chapter. The methodology basically have two parts, the neural network training and image pre-processing / post-processing. The next chapter is about the experimental result analysis of the proposed methodology.

Chapter 4

Results and Discussions

4.1 Introduction

In the previous chapter, a detailed description of the overall image reconstruction process and its various components were discussed. The performance of the proposed neural network on various data sets will be discussed in this chapter. Our main challenge is to observe how the proposed neural network improves the quality of the degraded images. The proposed method was implemented in Python language with Intel Core i5 processor and an 8GB RAM. Data images were collected from this website "https://source.unsplash.com/".

4.2 Dataset Description

To train the neural network we took 91 images from "https://source.unsplash.com/". These 91 images were then converted to sub images of 32*32 size. Total sub images were 24800. The network waas trained with those 24800 sub images.

To test the neural network we used two sets of images from the same website. Both set has 17 images each.

4.3 Evaluation of Framework

The neural network was trained for 200 epoches on 24800 images, with up scaling factor 2. Adam algorithm was used for optimization, with learning rate 0.0003 for all layers. OpenCV library was used to produce the training data and test data. For training the network, the loss function L is the average of mean square error (MSE) for the training samples (n), which is a kind of standard loss function.

Different learning rate was not set in different layer, but we found this network still works fine.

4.4 Evaluation of Performance

Three image quality metrics are defined to evaluate the performance of the framework. These three metrics are peak sinal to noise ratio (PSNR), mean square error (MSE) and structural similarity (SSIM).

If the PSNR value increases, the image quality gets better. If the MSE value decreases, the image quality gets better and as the SSIM value gets close to 1 (which is the highest for the original image), the image looks alike the original image.

Table 4.1: Results for 1st set of test images.

Degraded Img	**PSNR**	**MSE**	**SSIM**	**Reconst. Img**	**PSNR**	**MSE**	**SSIM**
baboon.bmp	22.28512	1152.62863	0.63272	baboon.bmp	36.25916	969.7034	0.70909
babyGT.bmp	34.25270	73.27084	0.93443	babyGT.bmp	23.03563	46.16211	0.95499
barbara.bmp	25.88885	502.70862	0.80869	barbara.bmp	26.57578	429.16439	0.85505
bird.jpg	33.05963	96.43593	0.93943	bird.jpg	35.06601	60.75754	0.95514
birdGT.bmp	32.97178	98.40651	0.95328	birdGT.bmp	36.54719	43.19992	0.96935
butterfly.bmp	24.75246	653.06341	0.87884	butterfly.bmp	30.38214	178.64371	0.95197
cguard.bmp	27.36883	357.53488	0.75278	cguard.bmp	28.67407	264.72503	0.82925
comic.bmp	23.63579	844.54299	0.82957	comic.bmp	25.88243	503.45273	0.89947
face.bmp	30.88222	159.21343	0.80036	face.bmp	31.64119	133.68469	0.82834
flowers.bmp	27.248692	367.56400	0.86906	flowers.bmp	29.66019	210.95109	0.89891
foreman.bmp	32.04483	121.81986	0.93839	foreman.bmp	36.14700	47.36976	0.96249
lenna.bmp	31.35857	142.67383	0.84594	lenna.bmp	33.10844	95.35802	0.86820
monarch.bmp	30.04416	193.10153	0.94329	monarch.bmp	35.13289	59.82916	0.96587
pepper.bmp	31.44722	139.79123	0.83767	pepper.bmp	32.78332	102.77070	0.84959
ppt3.bmp	24.65299	668.19450	0.92518	ppt3.bmp	27.14305	376.61372	0.94756
woman.bmp	29.38249	224.88078	0.93370	woman.bmp	33.62694	84.62668	0.96533
zebra.bmp	27.77654	325.49564	0.89229	zebra.bmp	30.40608	177.66182	0.93292

In the first set of test images, we took seventeen images randomly. From the table we see that all the reconstructed images have larger PSNR values, smaller MSE values and larger SSIM values than the degraded images.

In some particular images like (cguard.bmp), (woman.bmp), (comic.bmp), (foreman.bmp) and (butterfly.bmp), we observe a great improvement of image quality both visually and mathematically. The output results of the images are given below.

Table 4.2: Results for 2nd set of test images.

Degr. Img	PSNR	MSE	SSIM	Reconst. Img	PSNR	MSE	SSIM
athlete.jpg	29.30997	228.66746	0.89326	athlete.jpg	31.15325	149.58098	0.92725
baby2.jpg	26.31108	456.13491	0.92418	baby2.jpg	30.61073	169.48397	0.96108
breakfast.jpg	25.27533	578.98639	0.89104	breakfast.jpg	27.07853	382.2511	0.92901
car.jpg	30.49408	174.09813	0.94735	car.jpg	32.53689	108.77087	0.95993
couple.jpg	30.69027	166.40820	0.88471	couple.jpg	32.77497	102.96856	0.90382
dslr.jpg	34.25627	73.21079	0.97815	dslr.jpg	38.27571	29.01536	0.98705
flower.jpg	34.22856	73.67932	0.97649	flower.jpg	35.65770	53.01900	0.98075
free.jpg	33.81740	80.99559	0.95279	free.jpg	38.03362	30.67885	0.97149
hand.jpg	33.44234	88.30149	0.97279	hand.jpg	34.89942	63.13341	0.97907
idols.jpg	32.295229	114.99339	0.93264	idols.jpg	34.84551	63.92197	0.0.95804
pianist.jpg	28.40253	281.80516	0.93307	pianist.jpg	32.71181	104.47700	0.96537
plane.jpg	29.86930	201.03526	0.93011	plane.jpg	34.42019	70.49899	0.95644
random.jpg	31.43982	140.02968	0.89693	random.jpg	33.81783	80.98758	0.92270
redbull.jpg	31.35205	142.88839	0.94553	redbull.jpg	33.88027	80.83154	0.95657
sea.jpg	34.37775	71.19128	0.88498	sea.jpg	35.05932	60.85126	0.89188
socks.jpg	34.89241	63.23550	0.95610	socks.jpg	38.29137	28.91108	0.96899
stair.jpg	23.41202	889.20106	0.76867	stair.jpg	24.32803	720.10901	0.81035
town.jpg	21.51086	1377.57633	0.74057	town.jpg	21.85991	1271.19013	0.78255

We have randomly selected seventeen images for the second set of test images. All of the restored images have higher PSNR values, lower MSE values, and higher SSIM values than the degraded images, as seen in the table.

In some particular images like (baby2.jpg), (breakfast.jpg), (idols.jpg), (pianist.jpg), (plane.jpg) and (dslr.jpg) we observe a great improvement of image quality both visually and mathematically. The output results of the images are given below.

Figure 4.1: Experimental result of test image - 1.

Figure 4.2: Experimental result of test image - 2.

Figure 4.3: Experimental result of test image - 3.

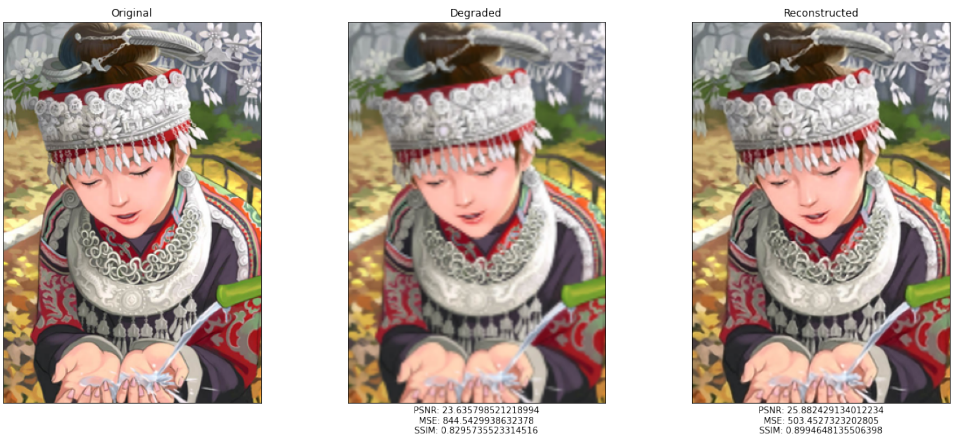

Figure 4.4: Experimental result of test image - 4.

Figure 4.5: Experimental result of test image - 5.

Figure 4.6: Experimental result of test image - 6.

Figure 4.7: Experimental result of test image - 7.

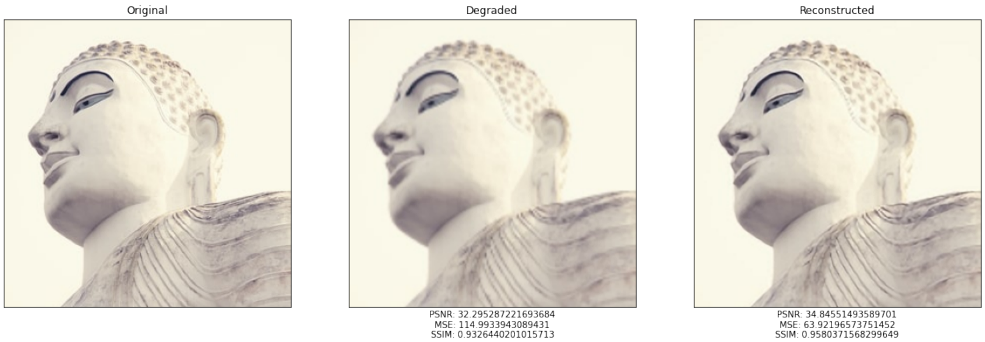

Figure 4.8: Experimental result of test image - 8.

Figure 4.9: Experimental result of test image - 9.

Figure 4.10: Experimental result of test image - 10.

We compared our network output for the first set of images with the output of the paper which introduced a network they named SRCNN [1] and another network DBSRCNN [7]:

Table 4.3: Performance comparison with other networks.

Neural Network	Scale	Avg. PSNR	Avg. SSIM
SRCNN	2	32.75	0.9090
Our Network	2	31.30	0.90256
DBSRCNN	2	28.38	0.8832

From the table, we can see our network performs better than DBSRCNN and the performance is almost equal to the SRCNN network.

Because it was trained with a large number of pictures (about 5 million) over a long period of time, the SRCNN network performs well. We don't have the same needs for training our network with such large amounts of images in our situation. As a result, the SRCNN network outperforms our network by a little margin. However, it also demonstrates that increasing the amount of training images does not result in significantly higher image quality.

We need to add more sustainable layers to the network to improve the image quality in order to improve the output. We extracted the features in two layers in our network and then concatenated them. Our network beat the DBSRCNN network when it came to modifying the feature extraction procedure.

In future work we will try to combine general adversarial network to our network to develop the guessing power of our network.

4.5 Conclusion

The chapter shows the result of upscaling and reconstruction of degraded images. We evaluate the performance by defining three most renowned image quality metrics and all the test image have showed improvement. Some images have shown improvement of a really satisfactory level and those output are shared above. In the next chapter, the conclusion is drawn on this thesis work.

Chapter 5

Conclusion

5.1 Conclusion

It's never easy to reconstruct a high-quality image from a low-quality degraded image. In this project, we created a five-layer neural network that successfully extracts low and high level features from a degraded input image, then combines these features and performs a second order mapping to produce a reconstructed output.

We perfectly implemented the neural network and trained it with a large number of images. The neural network was then put to the test with two sets of images, and it performed admirably on all of them. It enhanced the mathematical and visual quality of all the images. PSNR, MSE, and SSIM are three image quality metrics that we described. The proposed system will increase the PSNR of a degraded image, decrease the MSE value of a degraded image, and increase the SSIM of a degraded image for all images. Finally, we get a high-resolution picture that has been reconstructed.

5.2 Future Work

We will make an effort to do the following in the future:

Build a network with more layers.
Train the network with more pictures (at least a million).
Use GAN (Generative Adversarial Network) in the project.

References

[1] C. Dong, C. C. Loy, K. He and X. Tang, 'Image super-resolution using deep convolutional networks,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016. DOI: `10.1109/TPAMI.2015.2439281` (cit. on pp. 5, 20).

[2] J. Yang, J. Wright, T. S. Huang and Y. Ma, 'Image super-resolution via sparse representation,' *IEEE transactions on image processing*, vol. 19, no. 11, pp. 2861–2873, 2010 (cit. on p. 5).

[3] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, 'Photo-realistic single image super-resolution using a generative adversarial network,' in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690 (cit. on p. 6).

[4] H. Chang, D.-Y. Yeung and Y. Xiong, 'Super-resolution through neighbor embedding,' in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, IEEE, vol. 1, 2004, pp. I–I (cit. on p. 6).

[5] R. Dahl, M. Norouzi and J. Shlens, 'Pixel recursive super resolution,' in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 5439–5448 (cit. on p. 6).

[6] J. Yang, J. Wright, T. Huang and Y. Ma, 'Image super-resolution as sparse representation of raw image patches,' in *2008 IEEE conference on computer vision and pattern recognition*, IEEE, 2008, pp. 1–8 (cit. on p. 6).

[7] F. Albluwi, V. A. Krylov and R. Dahyot, 'Super-resolution on degraded low-resolution images using convolutional neural networks,' in *2019 27th European Signal Processing Conference (EUSIPCO)*, IEEE, 2019, pp. 1–5 (cit. on p. 20).