Generative Adversarial Networks for Blind Image Deblurring

By

Ummama Aslam

(Registration No: 00000319239)

Thesis Supervisor: Dr. Arslan Shaukat

DEPARTMENT OF COMPUTER AND SOFTWARE ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD, PAKISTAN

(2023)

Annex A

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **NS Ummama Aslam** Registration No. **00000319239**, of College of E&ME has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the thesis.

Signature :

Name of Supervisor: **Dr Arslan Shaukat**

Date: 19-09-2023

Signature of HOD:
(Dr Usman Qamar)
Date: 19-09-2023

Signature of Dean:
(Brig Dr Nasir Rashid)
Date: 19 SEP 2023

Dedication

Dedicated to my exceptional parents, caring husband and adored siblings whose tremendous support and cooperation led me to this wonderful accomplishment

Acknowledgements

I am extremely thankful to Almighty Allah (the most merciful) who gave me guidance, patience, courage and the ability to carry out this work. Without his blessings I would not be able to complete this work. Indeed, he is the most glorified and worthy of praise.

I would like to express my sincere gratitude to my supervisor Dr. Arslan Shaukat for his guidance and motivation in my quest of knowledge. It's a blessing to have such great and co-operative person as mentor for my research.

I am also thankful to my entire thesis committee. Dr. Usman Akram and Dr. Ahsan Shehzad for their support and cooperation.

My acknowledgment would be incomplete without mentioning the biggest source of my strength, my family. I am grateful to my beloved parents, husband and siblings as their support has been indispensable in helping me to complete it.

Finally, I would like to express my gratitude to all the individuals who encouraged and supported me throughout the duration.

Abstract

In imaging applications, blurred images are of significant challenge and when the issue is addressed without the prior knowledge of blur kernel i-e blind image deblurring, it become more complex. A novel approach for blind image deblurring using Generative Adversarial Network (GANs) for the restoration of sharpness of images is presented in this thesis. A special type of attention module is added in the generator model of GANs which is helpful for finding co-relation among the pixels of image. A generator and discriminator network in GANs are trained in adversarial manner and the mapping between blurry and sharp image is learned in unsupervised manner in the proposed framework. Unlike, other traditional deblurring methods that are based on the estimation of blur kernel this method by-pass the need of blur kernel estimation and learns to generate visually pleasing images from the blurred input directly. Extensive experiments are carried out on GoPro benchmark dataset determines that GAN-based deblurring method out-performs the existing traditional state-of-the-art methods in terms of visual quality and objective metrics such as Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). The GAN-based approach shows superior results in complex scenarios when the blur patterns are complex. The ability to surpass the blur kernel makes this method more applicable in real world scenario as the accurate knowledge is challenging to achieve. The contribution in the field of blind image deblurring is made by the introduction of approach that addresses the traditional methods limitations and the results out-performs from the previous techniques. In the proposed frame work the potential of GANs is showcased in tackling the challenge of Blind Image Deblurring. Further avenues are opened for the advancement in the field of Image Processing.

Key Words: *Blind Image Deblurring, Generative Adversarial Networks, Image Restoration, Adversarial Training, GAN*

Table of Contents

List of Figures

List of Tables

List of Abbreviations

GAN	Generative Adversarial Network
CNN	Convolutional Neural Network
DCP	Dark Channel Prior
FPN	Feature Pyramid Network
GPU	Graphic Processing Unit
MAP	Maximum-a-Posterior
MSE	Mean Squared Error
PSF	Point Spread Function
PSNR	Peak Signal to Noise Ratio
ReLU	Rectified Linear Unit
SSIM	Structural Similarity Index Measurement

CHAPTER 1: INTRODUCTION

The blur in images degrades the quality of image. The undesirable drop in bandwidth results in reduction of quality of image and we cannot avoid it. This is mostly caused by atmospheric instability, shake in camera, motion of object or motion of capturing device. Image blur has significant value in many applications such as object detection, medical imaging, crime analysis, face recognition and many other visualization applications. The reason of blur is usually unknown making it more complex to retrieve the original image

This study aims to identify the model for the deblurring of images effectively. Earliest research that is carried out for this purpose has not produce significant results because the blur kernel is estimated here. It is then transferred to optimization problem. In optimization-based problem simplified models are used. It is not possible to achieve good results in real world problem.

With the enhancement of deep learning, many methods have been proposed to learn about the blur kernel which is estimated in non-Blind image deblurring. Deep learning methods are under rapid development because they achieve superior results with high computational system hardware.

We use GANs for Blind image deblurring, real looking images are generated. Two networks i-e Generator and Discriminator are trained. Generator tries to generate real looking images that are deblurred while the discriminator tries to differentiate between the images that are generated with the help of generator and real images.

By the use of GANs realistic images are generated that are deblurred without the prior knowledge of blurred kernel. The deblurring process is simplified with the help of this network and has many applications in real world problems. In the training of images using GANs, we use loss function such as MSE and adversarial loss. These losses help to generate the images that are more similar to original images.

1.1 Motivation

The motivation behind this research that is carried out in this thesis is to overcome the limitation of prior deblurring techniques which need information about the blur kernel to successively retrieve the image.

The blind image deblurring techniques are often challenging because we want to retrieve the original image from the degraded version of image. Motion blur is created in the image and it is difficult to avoid because sometimes we capture image in low light, when the object moves or when the camera is moving. We don't have any information about the blur kernel. In case when the blur present in image is complex it is difficult to estimate blur kernel and the results are time consuming and not so good.

We train a deep learning model known as Generative Adversarial Network (GANs) to retrieve real looking images. In blind image deblurring without the estimation of blur kernel the mapping between the generated image and original image is learned.

The advantage of using GANs for the deblurring purpose is that the resulted image has very good texture details and there is no halo and ringing effect present. We can handle various types of blurs by this technique with images having different features.

1.2 Problem Statement

Blind Image deblurring that uses GANs involves the removal of blur from the images that are caused due to movement of camera or atmospheric disorders. This type of blur is unknown as the source is not clear. Our purpose is to remove the blur from these images without the prior knowledge of blurred kernel.

For the achievement of realistic looking images two networks known as Generator Network and Discriminator Networks are trained against each other. The production of clear images is achieved with the help of Generator Model training and the difference between Generated images and Real images is done by Discriminator Training. A Generator model is able to generate realistic looking images that are deblurred in nature and resembles to original image.

1.3 Objectives and Contributions

Our objectives are:

- To develop a model based on GANs for the deblurring of images and to enhance texture details without the prior knowledge of blurred kernel.
- To explore the impact of different training strategies of GANs on the generated deblurred images.
- To estimate the blurred kernel by the image itself.
- To evaluate the performance of GANs. Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measurements (SSIM) quantitative metrics are used for this purpose.

- To compare the GANS based approach performance with other state of art methods for the deblurring of images.

We can produce the accurate images even if the blur kernel is complex with the help of GANs for blind image deblurring. We train Generator to generate real looking images and Discriminator for the discrimination of generated and real image. If we add some prior knowledge of blur kernel in this training process the results, we get are more stable and the accuracy of the results is increased.

As far as the contribution of Blind Image deblurring using GANs is concerned, it has many practical applications like medical Imaging, for generating underwater images, remote sensing images and many more. In these cases, images are blurred due to unknown reasons. Image based applications can also can improve their quality with the help of Blind Image deblurring using GANs.

1.4 Structure of Thesis

Thesis Report is structured as follow:

Chapter 2:

This chapter describes Blind Image Deblurring in detail. Different types of blurs are discussed along with the general introduction of deblurring techniques. Evaluation Parameters of Blind Image Deblurring are also discussed in this chapter along with the Applications.

Chapter 3:

This chapter covers the review of previous work done by researcher on the deblurring of images using traditional approaches. Different types of image restoration techniques are discussed in detail. Deep learning techniques for Image deblurring are also discussed. The general introduction of building an Adversarial Network in detail. Different frameworks are discussed here and literature review is presented here.

Chapter 4:

This chapter describes our proposed framework in detail. Also gives the overview of performance measures which is used for the evaluation of our model.

Chapter 5:

In this Chapter the details of dataset and experimentation with different parameters is described in detail. The results are also presented here.

Chapter 6:

Conclusion of research is given in this chapter and the scope of this research in future is discussed here.

CHAPTER 2: BLIND IMAGE DEBLURRING

Image deblurring is one of the essential tasks in the field of computer vision and image processing. The aim of this technique is the recovery of clear images from the degraded version of images. There can be several reasons of blur that is present in the image as various factors are involved while capturing e-g motion blur, shake of camera or noise can be present. The quality of the image has the significant impact due to these factors which makes the resultant image less informative and the analysis is difficult.

The process of deblurring of images estimates the blur and then it is removed. High-frequency information is lost making it a challenging problem for the researchers. The blur kernel is unknown in the process of blind image deblurring.

2.1 Understanding Blurry Images

The images that lack sharpness are often known as Blurry images. There are many factors that are responsible for the blur in the images e-g shake of camera, motion blur, noise or out-of-focus blur. For the understanding of blurry image these key aspects need to be considered:

2.1.1 Shake of Camera

Shake of camera occurs when the camera moves mistakenly while capturing the images. This leads to blurred image. This blur usually occurs due to the movement of hands or if the support of the camera is not stable. A random blurry effect is appeared in the image.

2.1.2 Motion Blur:

The motion blur usually appears when there is relative motion between the camera and the object. The reason of motion blur can be fast moving object or shake in camera. We need to understand the intensity of motion blur for the deblurring of images, successfully.

2.1.3 Noise:

A random variation in pixel value is usually known as noise. The image quality is degraded due to this variation. It appears usually in low-light condition. The image quality is degraded due to appearance of noise.

2.1.4 Out-of-focus Blur:

When the accurate focus of the object does not happen this type of blur is appeared. The degree of blur depends on the distance between the camera and the object.

Figure 2.1 shows different types of blurs.

(a) Camera shake blur **(b)** Out-of-focus blur

(c) Moving object blur **(d)** Mixed blur

Fig 2.1: Different types of blurs [1]

By the deep analysis of all the above factors and to understand the characteristics of blurred image an effective deblurring algorithm can be developed. With the advancements in the techniques of deblurring of images a valuable information is recovered from the blurred image and sharp image with appealing visual details is generated.

2.2 Techniques for Blind Image Deblurring

It is the challenging task and its purpose is the restoration of sharp image with good texture details when the blurred kernel is unknown. In the non-blind image deblurring the blur kernel is known but in this approach blur kernel and sharp image both needs to be estimated due to which blind image deblurring is complex problem and it has various applications.

Numerous algorithms and techniques have been developed by different researchers. These techniques includes both traditional methods and deep learning-based approaches recently.

The traditional techniques usually rely on the assumption in the process of blur for the recovery of sharp latent image. Optimization based algorithm are involved in these methods for the restoration of image details and for the estimation of blur kernel.

By the evaluation of deep learning-based approaches Blind image deblurring is possible in efficient manner. When the model is trained on large scale datasets, these deep learning model can capture the finer details effectively even if the blur kernel is unknown.

2.2.1 Traditional Image Deblurring Methods

Mathematical models and assumptions are employed in traditional image deblurring techniques for the restoration of sharp image from the blurred one. Techniques that are widely used for this purpose are regularization-based methods, inverse filtering and wiener deconvolution. For the restoration of image and suppression of noise blur kernel is necessary.

2.2.2 Deep Learning Approaches for Image Deblurring

Deep learning-based approaches are the powerful tool for the deblurring of images. In these methods a complex mapping has been learned between blurred image and sharp image. When the deep learning models are trained on large scale datasets the sharp details are captured even if the blur kernel is unknown. The deep learning-based methods handles complex types of blurs and impressive deblurring results are delivered.

2.3 Evaluation Metrics for Image Deblurring

The performance of the image deblurring algorithm is evaluated by some performance metrics which are utilized for the measurement of quality of deblurred image as compare to sharp image. The effectiveness of deblurring techniques is measured for the evaluation of quantitative measurements. The evaluation metrics for the deblurring of images are:

2.3.1 Peak Signal-to-Noise Ratio (PSNR)

It measures the quality of deblurred image as compared to latent sharp image and it is a widely used metrics. The ratio between the maximum pixel value and the mean squared error between the ground truth value and deblurred value is calculated. The formula for calculating the PSNR is as follow:

$$PSNR = 20 * log10(MAX) - 10 * log10(MSE) \tag{1}$$

Where MAX is the maximum possible pixel value e-g it is 255 for 8-bit image. MSE is mean squared error between original and deblurred image. In MSE average squared distance between the corresponding pixels is calculated.

The PSNR value is usually expressed in decibels (dB). If the value of PSNR is higher it indicates that the quality of image is better.

2.3.2 Structural Similarity Index (SSIM):

The perceptual metrics that measures the structural similarity between the ground truth image the deblurred image. Three parameters, luminous, contrast and structural similarity between two images are taken into account. The formula for the calculation of SSIM is as follow:

$$SSIM(x, y) = [l(x, y) * c(x, y) * s(x, y)]^\wedge\alpha \quad (2)$$

Where l(x,y), similarity of average pixel intensities, s(x,y), similarity of standard deviation of pixel intensities, c(x,y), covariance of pixel intensities and α is the parameter which controls the trade-off between these three components.

The values of luminous, contrast and structural similarity ranges from 0 to 1. The maximum value is 1 which indicates the perfect similarity. The more the value the better is the quality of deblurred image.

SSIM is the metrics that is used widely in the field of image processing. The perceptual quality of deblurred image is evaluated with the help of this metrics.

2.4 Applications of Image Deblurring

Image deblurring has various applications where the restoration of images is necessary. The below are some of the applications:

2.4.1 Photography and Image Enhancement:

The quality of captured image is improved which is degraded due to motion blur or shake in camera. These algorithm helps in the enhancement of sharp details which allows the photographers to deliver better quality images.

2.4.2 Medical Imaging:

In medical imaging, the quality of diagnostic images is improved which is affected by the artifact of motion or blurring which is caused when the image is acquired. With the help of these deblurring techniques the healthcare professionals can get the more detailed images which helps more accurate treatment.

2.4.3 Document Restoration:

The historic documents and photographs are often degraded and blurred. For the restoration of these valuable artifacts and to preserve these content image deblurring techniques are employed. With the help of these techniques' researchers are able to read historic data.

2.4.4 Aerial and Satellite Imaging

Satellite images are captured from the distance may suffer from blur. There are many factors that are involved e-g atmospheric conditions or the limitations of sensors. The enhancement is done in aerial imaging and satellite imagery with the help of these techniques.

2.4.5 Forensic Image Analysis

In the forensic investigation, image deblurring techniques plays an important role because in the images of the crime scenes some images need more enhancement. Deblurring techniques gives good details of facial features, vehicles number plates and other important evidences for the identification of suspects.

2.4.6 Video Processing

Image deblurring techniques are now extended to the sequence of videos. In the video sequence consecutive frames are deblurred for the reduction of blurring effect. The visual quality is improved in the sequence of videos which benefits video editing and video surveillance.

The applications of image deblurring is not limited to these applications only. It has broad implications in many fields. Improved image quality is obtained with the advancement in these techniques.

CHAPTER 3: LITERATURE REVIEW

In digital image processing and computer vision the concept of blurring is a fundamental problem as it has many applications in the field of medical imaging, radar imaging, astronomy and consumer photography. The quality of image is reduced and degradation may occur due to this artifact. The blur in images occur due to movement of camera, movement of object, missed focus or may be atmospheric disorder. The goal of deblurring of image is basically the removal of distortion from the image. Many techniques have been proposed over the years for the solution of this problems. The literature review section of this thesis document investigates the previous research on the topic of deblurring of images. A comprehensive overview is provided on the methodologies and challenges of this technique. By the analysis of this literature research gap is identified.

This section provides information on image restoration, denoising of images, dehazing of images and super resolution of images.

3.1 Image Restoration:

When the pictures are taken by some capturing device like mobile phone or camera, various type of degradation is experienced such as compression artifact, motion blurring, hazing or some noise. The technique that is used to improve the quality of degraded image is called image restoration. The basic Image Restoration process is shown below in Fig 3.1:

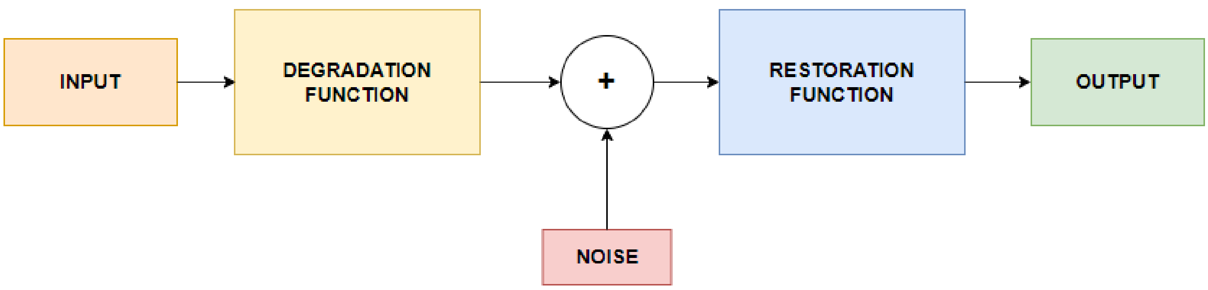


Figure 3.1: Basic Image Restoration

The unwanted distortion that is present in the image is removed and the image that resembles closely to the original image is restored. Image restoration can be divided into various categories depending on the different cases of kernels which includes denoising, dehazing and super resolution of images. We will discuss all of them in detail:

3.1.1 Denoising of Images

The underlying structure and details are preserved and noise is removed in this technique. An illustration of image denoising is given below in Fig 3.2 with 5 different methods:

Figure 3.2: An illustration of image denoising [152]: (a) Wiener filtering [153], (b) Bilateral filtering [154], (c) PCA method [155], (d) Wavelet transform domain method [156], (e) BM3D [157].

We take noisy image in input and clean image is retrieved without significantly affecting the sharpness and resolution of images. It has many real-world applications due to which it gains a lot of attention in the past few years. According to the analysis given in [32] the noising of images is distributed as blind image noising, hybrid noising and additive white noising images (AWNS). The last one gains a lot of attention.

Traditionally the techniques that are used for image denoising includes WNNM[33]. They have shown effective noise removing. In this technique singular value is assigned to different weights. MRF denoising [34] is a technique that models the image as a graph. Image model and optimization algorithm are combined in a single unit for the denoising of images.

Another method of denoising of images is described in [36] which explains trainable nonlinear reaction diffusion. In this technique all the framework including influence functions and filters are learned by training data. Learned simultaneous sparse coding is described in [35] we divide the input signal into cluster of similar features and then separate sparse representation is learned for each cluster.

A traditional technique which was used for image denoising and image restoration before deep learning was neural network. Many researchers carried out their research with neural network. In [38] a feed forward network is developed. Performance and denoising efficiency are balanced in this technique. They have used dynamic routing for this purpose. A network architecture is adjusted here according to noise present in the input signal. Signal integrity is maintained and the network works efficiently for the different range of tasks.

In [39] explained in which two optimization techniques known as maximum entropy and primal dual lagrangian multiplier are combined to improve neural network expressive ability. With the help of this combination complex relationships between input and outputs are learned resulting in improved performance than other neural networks.

Approach explained in [40] described two algorithm, greedy algorithm and asynchronous algorithm for the denoising of images. In the greedy algorithm optimal choices at each step is performed locally to find the global solution. A data is processed independently and in parallel in asynchronous algorithm. When both these algorithms are combined the denoising neural network removes noise effectively.

In the work explained in [41], multilayer perceptron and multilevel sigmoidal function are used for the improvement in the performance of neural network. Multi layers of nodes and activation function at each node allows the whole network to perform more efficiently and complex relationship is learned between the input and output variables.

After the arrival of deep neural network [44], learning based denoising methods are introduced rapidly. In [45] DnCNN method for the denoising of images is proposed. It consists of several layers which includes convolutional layers, batch normalization layer and rectified linear unit layer and also residual learning is adopted here [46]. These layers work together and remove noise from the image and to produce clear output.

Other important techniques for denoising network includes CBDNet[49] It consist of two networks performing different functions one is used for the estimation of noise in real image and other is responsible for the production of clear image. FFDNet[47] in this technique denoising speed of image is improved and blind image deblurring is processed by the use of different level of noise. They have the network for denoising of images and noisy image is the input of this network.

3.1.2 Dehazing of Images

The aim of Image dehazing is to remove the artifact of haze from the digital image. The amount of light that will reach the capturing device is reduced in case of haze that results in the decreased image quality. Haze that is present in the image is the serious issue as it changes the color of image and its processing is difficult. The issue of haze appears in different scenarios for example when we capture under water images. Image dehazing is mostly used in computer vision tasks where the accurate analysis of images is done. Image dehazing is complicated because of the atmospheric conditions that are difficult to avoid [50].

Image dehazing with 4 different methods is shown in Fig 3.3:

Figure 3.3: An illustration of image dehazing [75]

We will discuss single image dehazing in detail in this section. When a photo is captured, the lights that is reached to the sensor of camera is travelled through the medium known as atmosphere. The light is scattered due to the tiny particles that are present in the atmosphere that are known as aerosols. This effect is more visible in foggy scenes. Single image dehazing is a technique in which reduction of the haze from the image is done and clear latent image is obtained. This can be done if we remove the scattered light from the image that is caused by aerosols.

The dark channel prior (DCP) is a technique explained in [88], there are some areas in the image where the intensity of pixel is very low, they are known as dark channels. The amount of haze is estimated by these dark channels in this technique. High quality dehazed images are produced by this technique.

Another technique described in [48], this technique describes physics-based model that tells how uniform haze effects the quality of image when the physical properties of light are taken into account. The amount of haze is estimated and the original clear image is retrieved. This model performs well in complicated scenarios. A technique proposed in [52] is based on the observation that the clear image which is not hazy has the higher contrast value than hazy image. The local contrast is maximized to remove haze of the retrieved image.

After the rapid growth of deep learning techniques, many dehazing methods have been proposed by researchers. A direct method to create clear image from hazy image is explained in [53], light weight CNN which is a simple neural network is used for this technique. This process is based on mathematical model and explains how haze is created by the scattered light in the atmosphere. A technique proposed in [55] identifies the objects in images accurately by the use of method known as smoothed dilation. With the help of this technique small objects are detected accurately besides larger objects. A "gated sub-network'' which is the special type of neural network is used here. In this network the information is retrieved at different levels of network which results in the improvement of accuracy.

3.1.3 Super Resolution of Images

The process of image super resolution is actually the recovery of high-resolution image from low resolution image. Image super resolution is one of the kinds of image restoration. The images super resolution is attained with the help of down sampling with other kernels. The phenomenon of image super resolution is shown in Fig. 3.4

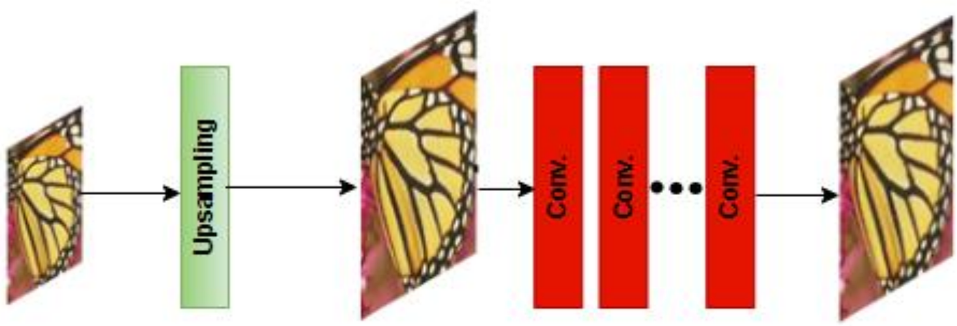


Figure 3.4: An Illustration of Super Resolution of Images.

SRCNN are proposed in [56], In this method with the help of deep neural network the images having less resolution looks more sharper because this network fills the missing details from the network by itself. A mathematical operation series is applied to blurry image and a sharp and clear image is produced at the output.

RDN are proposed in [57], In this technique the unnecessary modules are removed from the conventional residual network. The network resources are used more efficiently here. Feature representation is learned by residual learning. By the help of densely connected convolutional layers RDN achieved state of art performance.

3.2 Single Image deblurring

To remove blur from single image is known as single image deblurring. Blur in the image can be of any type e-g haze or noise is present in the image or the resolution of the image is very low. We need a generalized solution for this inverse problem so the main concern here in this thesis is image deblurring which can be on any type.

In the past years deblurring is performed with the help of mathematical model and is defined as:

$$b = k * s + n \tag{3}$$

Blur Kernel is convoluted with sharp image and the additive noise is added into it. The noise is usually white noise or gaussians noise.

There are some traditional ways to deblur the images and some learning-based methods are also used. Firstly, the analysis of some traditional approaches is done for the establishment of mathematical model.

The image can be blurred because in case of optical system when the light is passed through it is spread in different directions instead of converging into a single end point. Mathematically, this is represented by a point spread function (PSF). [7] The purpose of deblurring of image is to remove the blur from the image which is caused by PSF. There are two types of deblurring of image one is known as non-blind image deblurring and the other is blind image deblurring. We have the information of PSF in non-blind image deblurring and have no information about the PSF in blind image deblurring.

When the PSF is same throughout the image, this is known as uniform image deblurring. When the PSF is different in different patches of image, this is known as non-uniform image deblurring. Many traditional and deep learning methods are introduced in this research [7].

In [59,60], sharp image is estimated by statistical priors from the blurred image. A maximum-a-posteriori (MAP) estimation is used in this. The sharp image is estimated here that produced the observed unclear image. Richardson Lucy proposed an iterative algorithm. It always initiates with the initial guess. This guess is improved in this algorithm by the comparison of initial estimated image with the blurry image until the clear image is acquired. The desired accuracy is achieved by the repetition of this process. An improved version of Richardson Lucy was presented in [61] which is called Accelerated RL. An adaptive line searching technique is used in this improved version. A comprehensive survey presented in [62] gives more descriptive details on the development on this method is given.

In recent years more advanced digital cameras are introduced. The expectation of the quality of image is increased so more deblurring techniques are being introduced in [9, 63-65]. The aim of these techniques is to obtain clear images with these advanced cameras.

Many alternatives of sparse reconstruction are introduced to solve the problem of deblurring numerically. The problem is cut down into smaller iterative steps. The identification of dominant features is done and preserved for each iteration. Some features are enhanced and some are removed for the improvement in image quality. This was the most effective approach.

The parameters are estimated for the improvement of accuracy in Tikhonov Regularization explained in [7] where there is no required information is present. This technique adds bias in estimation in a very small amount but efficiency is increased.

Weiner deconvolution is one of the applications of this technique image spectrum is adjusted in Fourier domain for the reduction of blurring effect. A quadratic minimization technique along with fast Fourier transform are used to find the solution. The computational complexity is decreased with the help of this technique. Faster image restoration methods are developed explained in [66] under this context. Another technique known as tv-norm is explained in [67] blur kernel is controlled in this technique and the sharp edges are preserved. Stair casing artifact or visual blocking are present in the results. The edges in the final image are not appeared as smooth curves but the series of steps.

Different approaches are combined by researchers for the improvement in the quality of restoration of images [7] and multiple regularization prior are also used. In case of blind image restoration this single regularization framework is helpful.

The efficiency of the image restoration is improved by the common practice and the technique is known as split variable technique. The problem is broken into smaller parts which are independent to one another and are working in parallel. The processing time is reduced with the help of this technique. This technique of image restoration has many applications especially when multiple tasks are addressed simultaneously.[68]

3.3 Deep Learning for Image deblurring

Traditional deep learning techniques works well for simple deblurring tasks but it is difficult to handle heavy blurred image. A lot of time is needed in these optimization-based methods. After the rapid development of convolutional neural network which was explained in [16], image deblurring is learned by many learning-based methods explained in [17,18] single image deblurring performance is boosted with the help of these learning-based methods.

The author presented in research [19], presented a way to blur the images with the help of machine learning technique known as deep learning. A deep layered architecture is used for this purpose. This architecture is composed of many layers of neuron which better understand the features of images and try to make it clearer.

Another method is described in [20], used a similar approach but they have used convolutional neural network to address the same problem. The appearance of blurry image is learned by the help of neural network due to which it deblur the images more effectively.

These methods concludes that if we use deep learning with some specialized hardware the deblurring can be done more effectively.

The Author described a technique for the restoration of image in [21], the prediction is done about the quantity of blur in different patches of images. A mathematical model known as Markov random field is used here to find the distribution of blur across the image. The resulted information is used to find the way to undo the blur and to make the retrieved image clearer.

The Author in [22] predicted the specific kind of blur with the help of neural network in each part of image with the help of this it is easy to figure out how specific type of filter is applied to each part of image to make it clearer.

The methods described above accurately model the blur present in the image. The basic idea is same in both the research but the techniques are different. If we can figure out the blur it is easy to remove it.

The author in [23] described a machine learning approach known as 'fully connected neural network' for the estimation of sharp image from blurred version of image directly. A computer program is trained which take a blurred image at its input and sharp image is produced at its output. The detailed feature identification of image is done with the help of multiscale residual network. A better understanding for the removal of blur and to get the clear image is obtained.

Other researchers improve the performance of these models for getting better results.

The author in [27], described multiscale structures for the restoration of clear image. In this technique blurry image is taken and the size of blurry image is reduced to $1/4^{th}$. The size is increased gradually and the detailed information is added step by step.

This program works in steps. At every step the image is bit clearer and then when more information is added to it the image become clearer. The process is continued until the sharp image is achieved.

Many researchers carried out their research but the basic idea of scales is same in all these approaches.

Dark and bright channel prior are used in [70], for the restoration of clear image. These channel priors are the type of neural network.

DCP is a technique in which analysis in the areas of image that are either bright or dark is done and this evidence is used for the restoration of clear image from its blurred version. A special layer of dark and bright channel prior is embedded in neural network and the training is done to identify dark and bright channels in the image.

For the effective restoration of images and to improve the efficiency of neural network 'sparse regularization' is used.

When neural network is combined with DCP and sparse regularization is applied to it better results for the restoration are achieved.

The author in [71], for the deblurring of image used 'depth map' for the improvement in the restoration process of clear images. The distance between the two objects in the image is shown in depth map this additional information helps in the restoration process.

According to the research presented in [72], a special type of architecture known as MRPNet is used for the restoration purpose of clear images. The context features in the images are learned by using encoder-decoder approach. These features are combined with local details. The small and large context details are captured leads to get better results for restoration purpose.

In [73], a computer vision model known as DeFMO that can simulate the concept of capturing the motion blur by high-speed camera of fast-moving object. A sub-frame series is generated in which the position over time and appearance of object is shown. For the purpose of surveillance this can be a very useful application.

The approach presented in [74], does not rely on large external datasets but on single images. The model learns from a single image and train itself with the help of additional data that is generated from the original image. By the help of auxiliary network new data is generated for the training of main model. The scenarios where the large amount of labeled data is expensive the meta-auxiliary learning is used.

There are some other types of deblurring known as video deblurring or multi-image deblurring the task is to remove blur from video or to remove blur when multiple images are taken together. They are complex problem because of motion blur is different for every frame of image and the algorithm needs to find the relationship between each frame.

However, the purpose of our thesis is to remove the blur from the single image which is caused by motion.

Besides motion blur there are some other kinds of blur that can be present in the image which is out-of-focus blur. This blur is usually caused when the focus is on the wrong part or the adjustment of lens is not proper. This challenging task has many applications especially in medical imaging.

3.4 Building an Adversarial Network

3.4.1 Artificial Neural Network

A computer program that is design to assist the human beings. This network works like human brain. The purpose of this network is to solve complicated problems like the processing of signals or pattern recognition. Some of the examples in which neural network is used includes recognition of faces, conversion of spoken words into text and in the prediction of weather. Neural Network is one of the types of machines learning it improves its performance as the more training data is processed into it.

A neural network is actually the combination of computer processors that are small in size. These computer processors work together for a solution of the problem. The organization of the layers of processors are same like the organization of layers in the human brain. The basic architecture of artificial neural network is shown below in Fig 3.5:

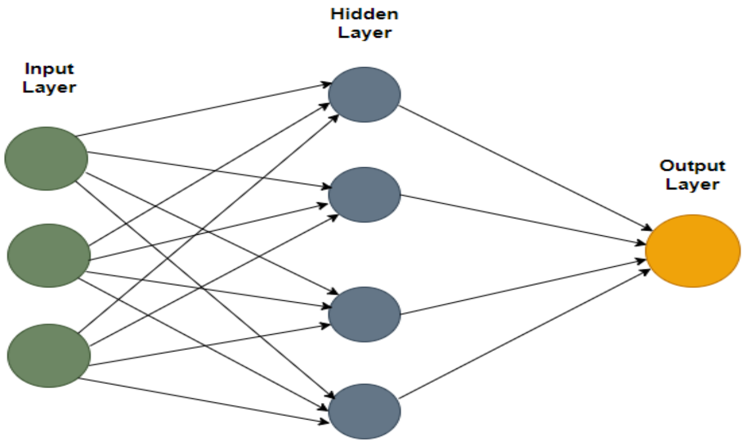


Figure 3.5: Artifacial Neural Network

The first layer of this network is responsible for receiving the initial information. The first layer performs the same function as our eyes receives initial information. Same like the communication of neurons in our brain each information of one layer is transferred to next and in the end final layer is responsible for producing the output result of the whole network.

This process is same as the working of our brain, something is seen and recognition is done.

Each small processor in the neural network is called a node and is responsible for specific task and it has its own area of expertise. Based on its rules of learning each successive node receives information from predecessor node, process it and to send it to the next node. All the nodes in the network are interconnected and the last node is responsible for the output of the system. The overall working of this system is like the team in which every individual is responsible for the specific task.

The neural network learns and improve themselves over each interval of time. The modification of network is done on the basis of the training data that is fed into it. This modified network has the improved performance.

Different input streams are assigned with weights in the network. Every processing unit in the network the evaluation is done about the input received from other processing units. The higher weights are assigned to the input which contribute more in getting the correct answers. The whole neural network pays more attention to the input which contribute more in solution of the problem for which it is designed. The better identification of important inputs that are responsible for accurate results are possible over time.

A large amount of data is needed for the creation and training of neural network. A dataset includes input-output pair, output is the correct answer for the input provided. For example, if the training of network is done for the identification of animals, dataset is consisting of pictures of animal as well as pictures of object or something else. Each picture in the dataset has some correct label. The internal weights are adjusted with the help of information provided; it determines how single input is helpful for providing correct output. The network learns the correct identification of animals by the adjustment of weights. The performance feedback and the amount of more data that is fed into the network the adjustment of weights takes place which improves the accuracy of the network.

3.4.2 Convolutional Neural Network

CNN [17] is one of the machine learning model. The working of this machine learning model is inspired by human brain as it processes the data the same way as human brain does. This model is designed for the effective and efficient analysis of images. It breaks the task into smaller parts for efficient processing. The recognition of patterns and features are not done manually but with the help of learning. A series of layers which are connected to one another, each layer performs different operation on initial data that is provided to the network. The specific features in the image are scanned by the filters provided in the network, a convolution operation is followed by this which is responsible for combining the features. After this pooling is performed for the reduction of the size of image. The basic architecture of convolutional neural network is shown below in Fig 3.6:

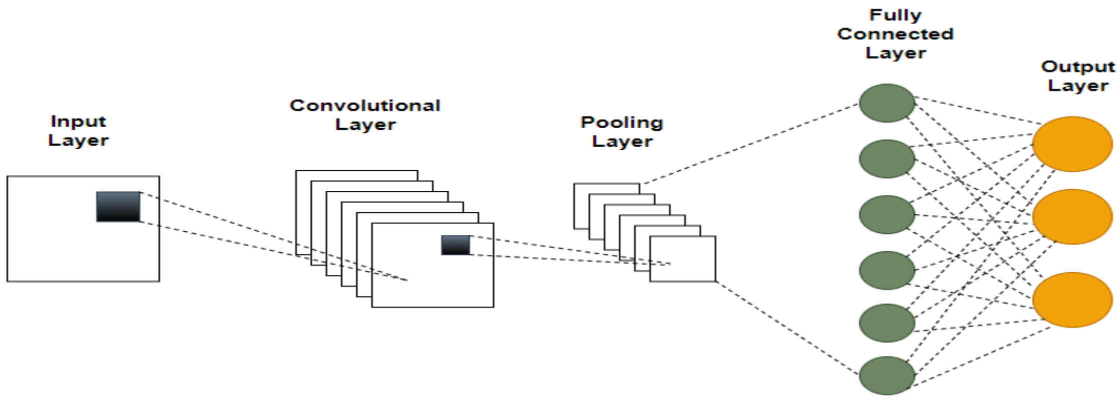


Figure 3.6: Convolutional Neural Network

The spatial structure of images is captured so they work successfully in the classification of images and in the detection of objects [11]. Millions of parameters can be processed with the help of convolutional neural network because of the advanced Graphic Processing Unit (GPU) hardware.

The main difference between the ordinary neural network and convolutional neural network is that the CNN are specially designed to deal with the data of images. Some assumptions are done on the initial input data like what type of array it is. With the help of these assumptions certain properties are encoded in the architecture and these properties are specific to the processing of images.

Convolutional layer, pooling layer and fully connected layer are the three main layers of this network. The first layer is responsible for applying filter to initial image and to extract useful features from the image.

The number of parameters is reduced with the help of pooling layer. Pooling can be done in two ways by taking average or by selecting the minimum number from small portion of image. The reduction in size of image is done to make the network more efficient. The last layer known as fully connected layer is responsible for the final classification.

A full architecture is formed by stacking these layers, different activation functions are applied to the network. Complex features on image dataset are learned by the combination of these activation functions and layers. For the processing of images CNN proves to be efficient because of feature selection and number of parameters they use. The literature table of the latest technique for image deblurring is as under in table 3.1:

Ref.	Author	Year	Technique	Dataset	PSNR	SSIM
[70]	Seo-Won Ji et al.*	2022	XY-Deblur	GoPro	30.97	0.95
[71]	Menghang Li et al.*	2021	L0 sparse representation	GoPro	29.89	-
[72]	Phong Tran et al.*	2021	SRN-Deblur	GoPro	30.2	-
[73]	Maitreva Suin et al.*	2020	Patch Hierarchical Network	GoPro	31.85	0.95
[74]	Jiawei Zhang et al.*	2018	Spatially variant recurrent neural network	GoPro	29.18	0.93
[66]	Orest Kupyn et al.*	2017	Deblur-GAN	GoPro	28.7	0.95
[75]	Dong Gong et al.*	2017	FCN	BDS-M	26.40	0.72
[76]	Seungjun Nah et al.*	2017	Multiscale-CNN	GoPro	29.08	0.91

[77]	Jian Sun et al.*	2015	CNN	Pascal VOC	24.64	-

Table 3.1: Image deblurring with CNN

3.4.3 Generative Adversarial Network

It is a deep learning model and has two neural networks known as Generator and Discriminator. These two networks are trained against one another in a way that generator generates image that are similar to real looking images and discriminator tries to identify if the image is generated from generator or real image. GANs generates realistic images when both the networks are trained. The basic architecture of GANs is shown below in Fig 3.7:

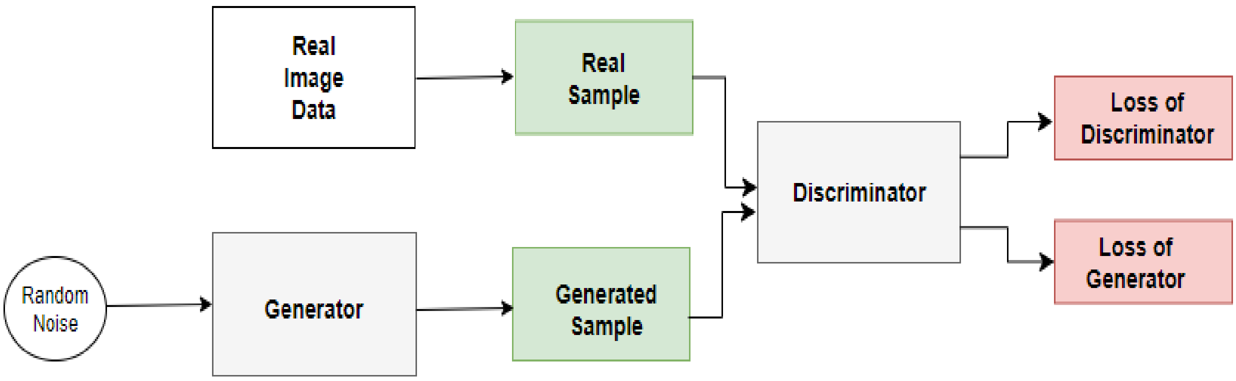


Figure 3.7: Generative Adversarial Network basic structure

The purpose of generator is to create fake images that resembles real images. Random noise is fed up in the generator with the help of which generator tries to generate images. The resolution of generated images is increased with the help of convolutional layer followed by up sampling layers that are present in this network. Loss function is being used that reduces the difference between the generated image and real image in the process of training.

The purpose of discriminator is to differentiate between real and generated images. The discriminator network takes an image –which may be real or fake- on the input and it has binary output that tells if the image is real or fake. The composition of discriminator network is same like the generator but here the convolution layer is followed by down sampling layer with the help of which resolution of input image that is fed to discriminator is decreased.

The GANs is trained on and off with generator and discriminator. On every single iteration of training the production of set of fake images is generated by generator and the role of discriminator is to classify these images on the basis of realism and evaluation is done. The process of adversarial training is continuing until the generator generates such images that are nor differentiable for the discriminator.

3.5 Summary of the Literature

Learning based methods brought a great change in the tasks of restoration of images. Deep learning approaches has several benefits for restoring the images which is degraded, successfully which includes:

- Learning based methods that use deep learning technique often shows superior results as compared to traditional approaches. In this approach complex pattern between clear and degraded image is learned which results in better restoration of images.
- This approach gives more realistic results because it fills the missing part of images which is usually not done in traditional methods. The results by this approach are usually natural looking and appealing.
- The efficient utilization of hardware is done in this approach as they take advantage of parallel processing GPUs due to which computational speed is increased. This approach is used where quick restoration of images is required e-g in real time applications.

3.6 Current Challenges and Gap in the Literature

- Deep learning-based methods are not able to restore the sharp details of images. To overcome this deficiency the improvement in training strategies and network architecture is needed for the reproduction of high-frequency components.
- The images that are severely blurred, particularly when there is the case in which the blur kernel is large deep learning methods are not able to accurately predict the blur kernel. More robust algorithm needs to be addressed for this deficiency that can handle complex blur types effectively.
- The complex motion blur is difficult to handle mathematically by traditional methods.
- The limited generalization ability is another deficiency. The deblurred network are trained on some specific blurs they are not be able to give good results on the blurred images that have different blurred characteristics.

- Deep learning-based methods can be computationally intensive, making them less suitable on devices with limited resources and there processing speed is too slow. For the practical applications researchers are working for making less computational model which performs good but computationally not that much expensive.
- The performance can be enhanced in deep leaning methods. The enhancement in the image restoration algorithm needs innovative approaches and improvement in training strategies and network architecture.
- In designing of deep leaning methods for the restoration of images, deep knowledge is required from other computer vision tasks.

3.7 Bridging the Gap with GANs for deblurring of Images

The aim of the thesis is to tackle these challenges in efficient way by the use of GANs for the deblurring of images. By the advent of GANs several improvements in the deblurring of images are made. It has two main components, the generator and discriminator that are trained against one another. The generator learns to generate clear image from the blurred image that is provided to the network as the input and the discriminator is responsible to distinguish the generated image and clear image.

The complex patterns of the blurred image are captured by the training of GANs and more accurate results are calculated.

Also, in terms of computations GANs are efficient, they have improved performance and visually more appealing results are produced.

3.8 GANs for reconstruction of blurred images

GANs are used for the reconstruction of blurred image without prior knowledge of blurred kernel. In this process sharp image is generated with the help of GANs and it looks same to the original image that is blurred.

We preprocess the image first by converting it to grayscale and the size is reduced so that it is less computational complex. The network is trained on the pair of blurred images and their corresponding clear images. We fed blurred image on input of generator and the result is sharp image and then discriminator discriminates whether it is generated image or real image. Both the networks are trained until the output of generator is not distinguishable for discriminator.

We can also use denoising or sharpening the images to get more improved results.

There are many evaluation parameters for evaluating the performance of Adversarial Networks but the mostly used are PSNR and SSIM.

The process is a bit complex but the results are encouraging making this model valuable tool in many image processing applications.

CHAPTER 4: PROPOSED METHODOLOGY

When the images are captured in low light condition or if the object is moving too fast which is being captured, in both these conditions deblurring of images may occur. The camera is not able to take clear picture. In the deblurring of images the edges of the object which is being captured are not clear. Our goal is to retrieve clear image from the blurred one. The PSF is estimated which is the cause of blur and the effect of this PSF is reversed for the retrieval of clear image from the blurred one.

A deep learning network known as GANs is designed to solve the problem of deblurring. It is based on the basic architecture of GANs. GANs consist of two components known as generator and discriminator which are trained against one another. Random noise is provided at the input and the data is generated with the help of generator which is real looking data, but it is fake. The job of discriminator is to discriminate between real and fake data. Both the networks are trained simultaneously, generator kept on producing fake data which is like real and discriminator tries to distinguish them well. Based on the previous output both the network improves their performance.

The problem formulation of using GANs for image deblurring involves GANs model training on the dataset which has clear as well as the corresponding blurred images. The blurry image is provided to generator as an input and it tries to deblur the image while discriminator tries to distinguish between clear image which is real and deblurred image which is fake and is generated by the generator. Our goal is to generate the deblurred images that are close to the real images.

The advantage of using GANs is that we do not need to estimate the blur kernel . The alignment of blurred and sharp image is done. For the extraction of high frequency features a residual block is used. To distinguish between different levels of image details a multi scale discriminator is used.

The difference between the generated image and clear image that is provided needs to be minimized. It can be done if the Mean Squared Error (MSE) is reduced between the images. The mean squared error estimates the average squared difference between the pixel values of two images. The generator and discriminator are trained with the help of GANS loss function. GANS loss function is the combination of generator and discriminator loss. The generator loss is actually the probability log of, when the discriminator classifies the fake image that is deblurred as real. The probability log, sum of all the outcomes when the discriminator classifies real image as real and generated image as generated is called discriminator loss. Some other loss functions are also used besides adversarial loss which includes pixel loss, perceptual loss and gradient loss. When the training is complete, and model is trained completely the images that are new and not present in the dataset are deblurred.

In the process of deblurring, a blurred image is fed to the input of generator It deblur it and then other artifacts are removed by denoising and sharpening. It is the most challenging problem and can be solved by GANs. The problem formulation includes the training of GANs on a dataset which has the pair of blurred and clear image.

4.1 Network Architecture

The task of deblurring of images are performed by the special type of GANs. The architecture of GANs model of proposed architecture for image deblurring is given below in Fig 4.1:

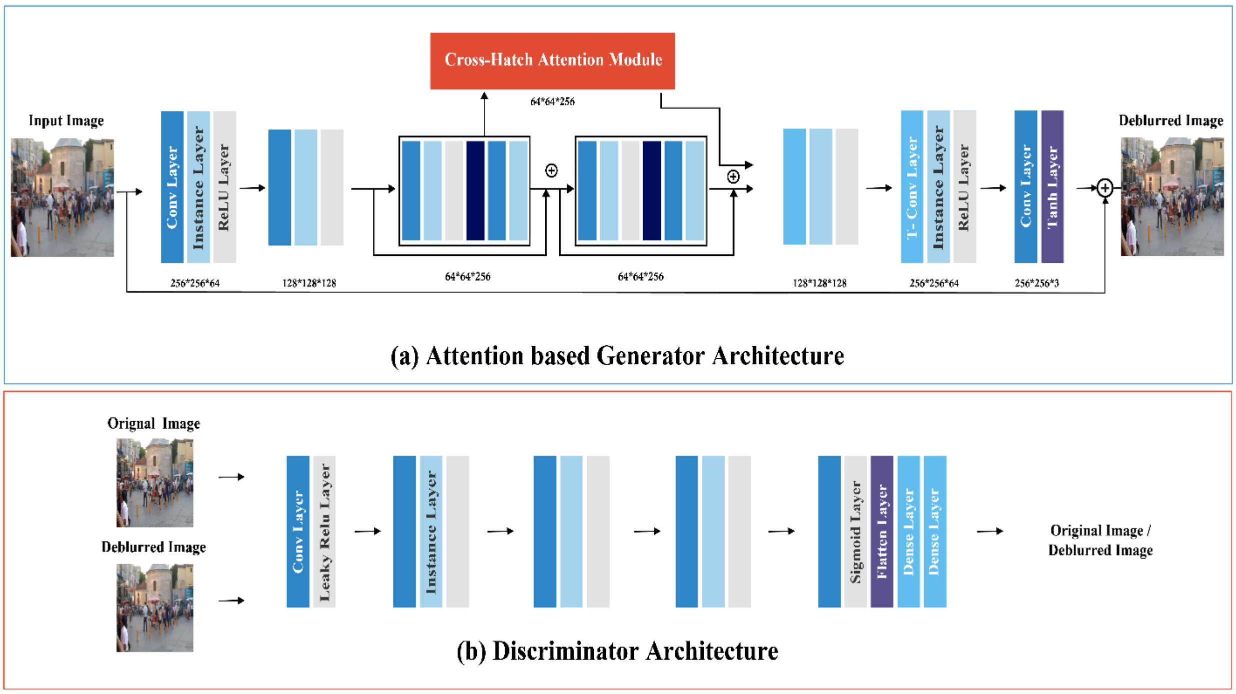


Figure 4.1: Architecture of GANs

In the architecture of GANs the two major components are Generator and Discriminator. The blurry image is provided at the input of Generator model and it aims to provide deblurred sharp image. The job of the discriminator is to distinguish between generated image and the sharp image that is provided in the dataset.

The Generator network of GANs follow the typical encoder-decoder architecture. The down sampling is done at the encoder part with the help of several convolutional layers and the decoder part is responsible for up sampling of encoded features for the generation of deblurred image. Skip connections are used for connecting encoder with their corresponding decoder layer. High quality details are preserved in these connections for the improvement of output (the deblurred generated image), The training of generated model is done to minimize the difference between generated image and sharp image. An additional attention module is added and described below in Figure 4.2:

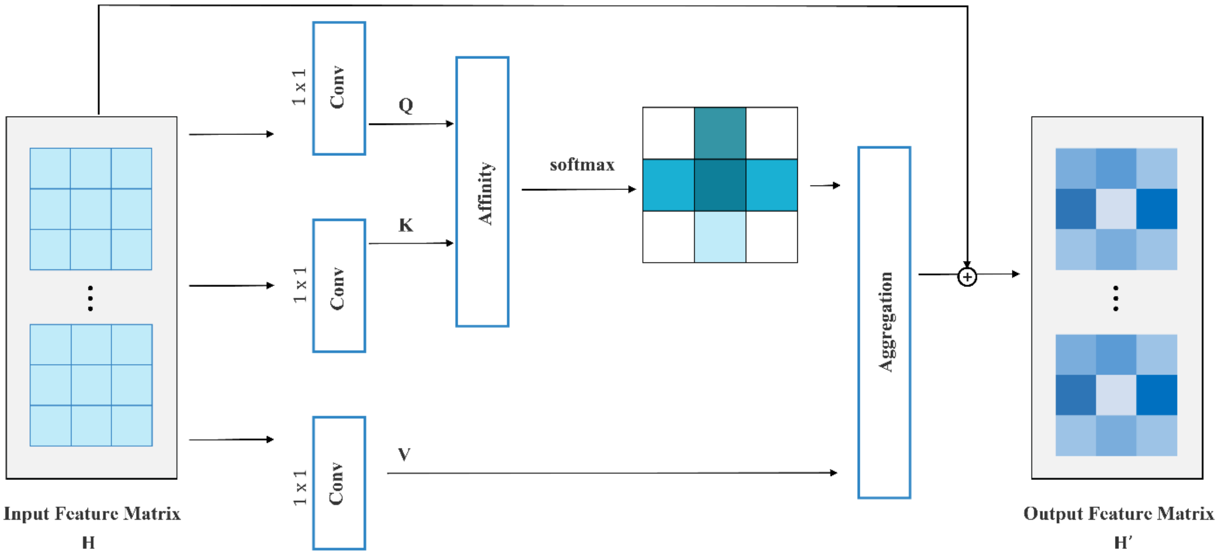


Figure 4.2: Attention Module[78]

In the attention module three types of transformations are query, key and value and they are obtained by applying 1*1 convolution. The horizontal relationship between the pixels is captured by transposing key information. The swapping between heights is width is done.

The vertical relationship between pixels is captured by original key features. After that attention map is calculated by the metrics multiplication of query with original key and query with transposed key. It tells the relationship of pixel is vertical and horizontal direction.

Attention map is combined using softmax so that they sums up to 1. Both the attention map is concatenated along specific axis. Concatenated attention map is divided into horizontal and vertical interaction. The interaction map which is obtained by attention is used to fuse the original value map. This process emphasize relevant features in the original value map which is based on pixel relationship which is learned.

The strength of fused interaction is controlled by gamma parameter. The fused pixel interaction map are added to original maps enhancing the representation of features. The complex relationship between pixels is obtained by incorporating both horizontal and vertical attention.

The Discriminator network of GANs follow the typical Convolutional Neural architecture. Multiple convolutional layers are followed by Fully connected layers in this network. Its output is usually a probability which indicated if the input is real or generated. The training of discriminator model is done so that its probability of classifying the real and fake sample is maximized. The main building blocks of both the architecture are:

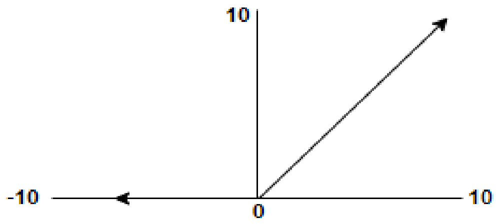
4.1.1 Convolutional Layer

The convolutional layer usually consists of several independent filters which are applied to the image. This filter slides over the whole image and dot product are calculated between the image pixel and filter. The initialization of the filter is random and these values are updated by subsequent learning. Lines and corner pattern are detected usually in the starting layers. As the continuation of training is done the network is actually taking dot product of the neurons of preceding layer with its corresponding weights.

4.1.2 ReLU (Rectified Linear Units) Layer

In convolutional neural network the convolutional layer is always followed by activation layer to introduce non-linearity in the linear system as only the linear operations are performed in convolution. The activation actually speeds up the training process and also the performance is not affected by this activation noticeably. The activation function we used in our architecture is ReLU which converts negative numbers into zero. The receptive fields of the convolution are not disturbed and the amount of non-linearity is increased as shown in eqn 4:

$$f(x) = \max(0, x) \tag{4}$$



4.1.3 Pooling Layer

The pooling layer is also known as down-sampling layer reduction of spatial resolution is done and the computation of the network are reduced. For the implementation of pooling non-linear functions are available. The L2-norm, average-pooling and max pooling are some types of pooling. The most commonly used among them is max pooling. The image is divided into non overlapping parts and the maximum value from each part is selected. The advantage of pooling is reduction in computation and overfitting. The average and maximum pooling is described in the below figure 4.3 and 4.4:

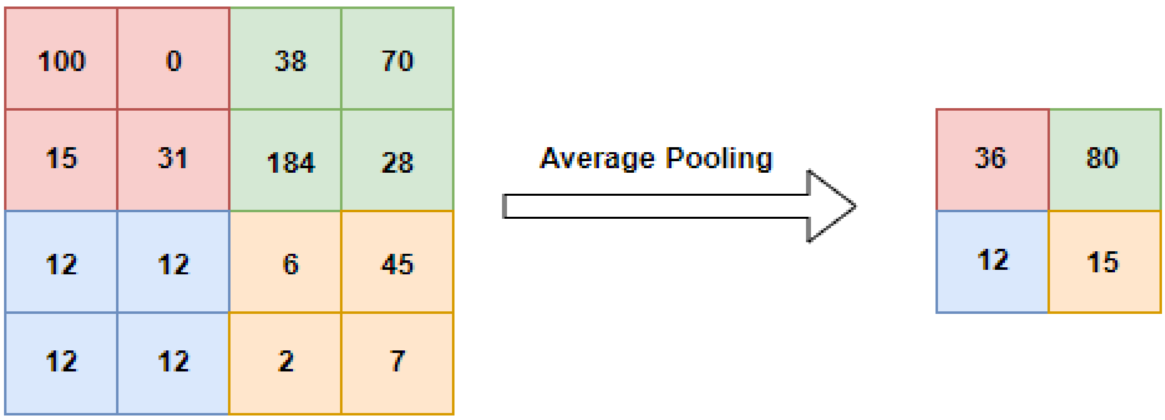


Figure 4.3: Average Pooling

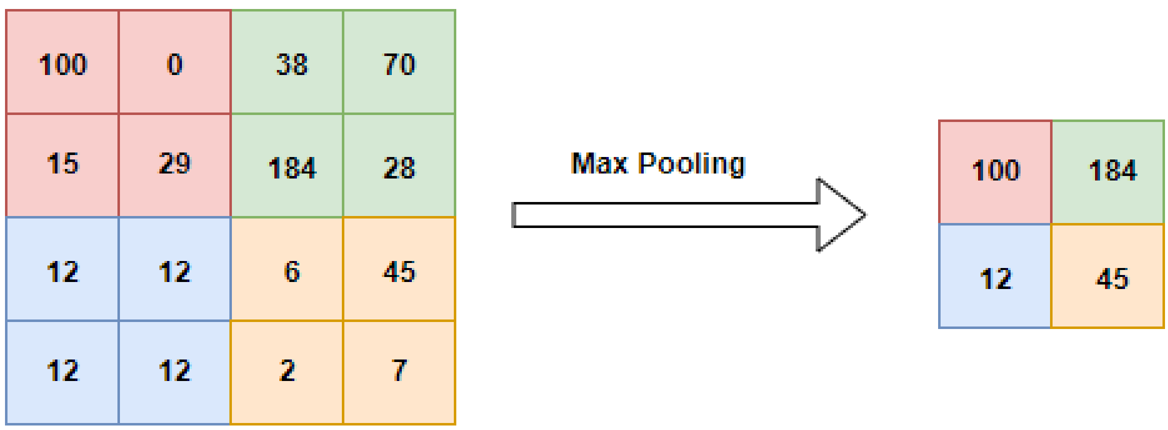


Figure 4.4: Max Pooling

4.1.4 Dropout Layer

It is a regularization technique mostly used in neural networks for the prevention of overfitting. During the training process some random fraction of neurons are dropped, improves generalization.

4.1.5 Batch Normalization

It is used for the stability and acceleration of training. The feature values are normalized in a mini batch by subtracting the mean of batch and dividing it with the standard deviation of batch. Better generalization and better convergence is achieved by this.

4.1.6 Tanh Layer

It is the special tool in neural network. It normalize the value between 1 and -1. It is capable of modeling positive and negative relationship. It is used in hidden layers for capturing of complex pattern.

4.2 Loss function

A loss function is the fundamental concept in machine learning and is used for optimization problems. It actually describes the performance of machine learning model and it measures the difference between the actual value and the predicted value that the model predicts. It quantifies the performance of machine learning model. Machine learning model aims to minimize the loss function.

The goal of machine learning model is the optimization of its parameters and the improvement in its predicted values. Different types of machine learning problems may require different types of loss function according to their objectives. Loss function plays an important role in the performance of machine learning model. The alignment between the loss function and problem is the important task.

The loss function has the significant role in machine learning tasks as it provides the feedback signal for updating the parameters, the optimization process guidance is performed with the help of this, the definition of learning objectives and it is also helpful in regularization purposes.

Loss function plays an important role in various type of machine learning models. We will see some in detail:

4.2.1 Linear Regression Loss

In linear regression, a linear model is used which fits on the set of data points. The main goal is to minimize the distance of actual and predicted value and to find the best fitting line. Using the mathematical techniques and formulas the intercept(a) and slope(b). The equation for the linear regression is expressed as:

$$y = \alpha + \beta x + \varepsilon \tag{5}$$

α is the intercept, β is the slope and ε is the error term between the actual and predicted value.

This equation minimizes the sum of square residual for finding the optimal value of α and β. The value of y is predicted on the basis of x.

The loss function used in linear regression is mean squared error. The average of the squared difference is estimated between the actual and predicted value in mean squared error. This loss is minimized to find the best fitting line for the data points. The slope and intercept are adjusted in a way that the predicted value become closer to actual value. The best approximation of linear relationship is achieved by the line with lowest MSE.

4.2.2 Logistic Regression Loss

The statistical model used for binary classification problem is known as Logistic Regression. Here the goal is the prediction of probability of event which belongs to any one of the two classes. In logistic regression binary classification problem is the major concern. Binary class entropy loss is usually used in these problems.

A linear combination of predicted value (x) and their corresponding (β) are represented as:

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{6}$$

The sigmoid function known as logistic function transforms the linear combination into value between 0 and 1 defined as:

$$p = \frac{1}{1 + e^{-z}} \tag{7}$$

This probability defines the chance of instance which belongs to class 1. The decision boundary is at p = 0.5 for the classification of instances. During the maximum likelihood estimation, the value of β is estimated and gradient descent minimize cost function for the training of model effectively.

In logistic regression, log loss is used. The dissimilarity between the predicted probabilities and actual binary labels is measured. By the minimization of log loss the actual probabilities of each class is learned. The accurate prediction is obtained by adjusting parameters.

4.2.3 Neural Networks Losses

Neural networks are able to perform different tasks such as object detection, natural language processing, and image classification. The selection of loss function is according to the nature of problem for example in regression problem of neural network the mean squared error and mean absolute error is usually. In the multi class classification problem categorical cross entropy is used. In the object detection problem L1 loss is used.

4.2.4 Reinforcement Learning Losses

In reinforcement learning the interaction is with the environment so different loss functions are used. The loss in reinforcement learning includes are Q-learning and Policy gradient.

The diverse role of loss functions is highlighted above. Now we will discuss the loss function for our specific application:

4.3 Loss Function for Deblurring of images using GANs

In GANs, the loss function is Adversarial loss which is actually the combination of Generator loss and Discriminator loss. The purpose of generator is to generate the real looking samples and the discriminator tries to differentiate the real and generated samples.

We use GANs for deblurring of images the commonly used loss functions are Perceptual loss and Adversarial loss.

4.3.1 Perceptual Loss

The difference between the features of the image that is generated and real image is calculated in this loss. This loss encourages the generator to produce more sharp images. The high-level features are extracted from both the generated and sharp images using CNN pre-trained model VGG-16.

$$Per\_Loss = \mu(true\ features - predicted\ features)^2 \tag{8}$$

The architecture of vgg-16 and vgg-19 is described below in Fig 4.5 and 4.6:

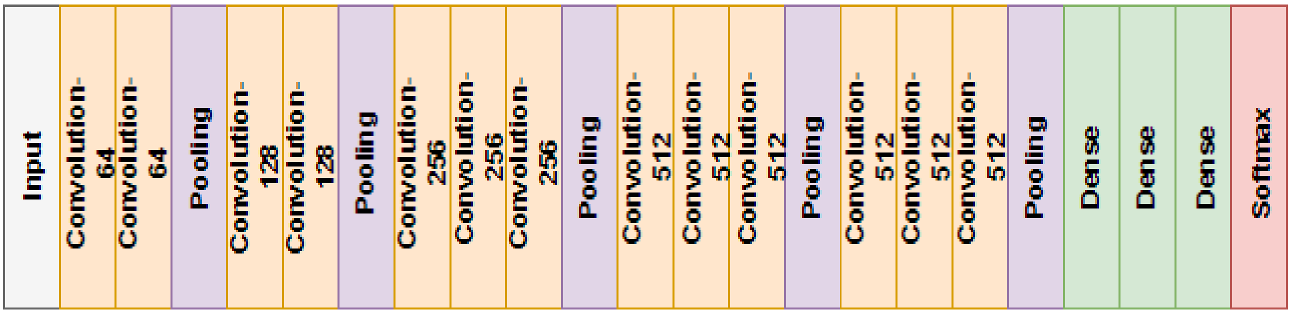


Figure 4.5: VGG-16 Architecture

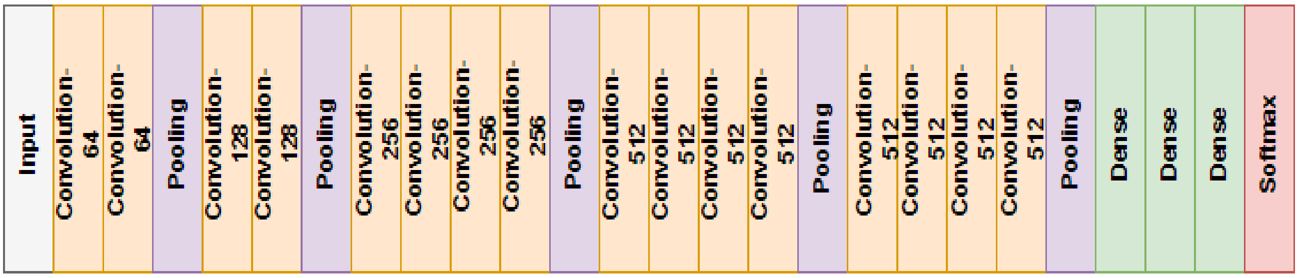


Figure 4.6: VGG-19 Architecture

The loss is defined as L1 distance between the feature representation of generated images and real images at multiple layers of pre trained model. The generator slowly learns to generate the higher-level texture details by considering feature level differences.

4.3.2 Adversarial Loss:

In Adversarial loss, the discriminator network is involved that differentiate generated and sharp images. The generator aims to produce deblurred images and discriminator job is to tell if the image is generated or real.

The Adversarial loss encourages the generator to produce real looking images so that the discriminator is not able to classify them. The higher loss is assigned in the feedback that discriminator gives to the generator to the images which are easily distinguishable by discriminator.

$$D\_Loss = \frac{1}{2} \cdot \mu[(real\ image\ - \mu(fake\ image) - 1)^2] + \frac{1}{2} \cdot \mu[(fake\ image\ - \mu(real\ image) + 1)^2] \qquad (9)$$

$$G\_Loss = \frac{1}{2} \cdot \mu[(fake\ image\ - \mu(real\ image) - 1)^2] \qquad (10)$$

The Loss function used in GANs is actually the combination of weighted sum of Perceptual and Adversarial Loss. The relative importance is determined by the weights that are adjusted after every iteration.

By the optimization of combination of Perceptual and Adversarial loss, the GANs learns to generate deblurred images with sharp details.

4.3.3 Pixel loss

Pixel loss is used to fix color and texture. It helps when images are different. It protects the important parts and ignores the unimportant part.

$$Pixel\ loss = 0.5 * (DB - S)^2 \tag{11}$$

If (DB-S) <1

(DB-S) – 0.5, otherwise

DB is deblurred image and S is sharp image.

4.3.4 Gradient difference loss

Image gradient difference is used to fix the gradient difference problem. The sobel operator is used for this purpose. The gradient image can be obtained using the derivative of x and y along x y direction

$$Gradient = |dx| + |dy| \tag{12}$$

The image difference gradient loss is calculated by:

$$Gradient_{loss} = \frac{1}{(h*w*c)} * |Gradient(S) + Gradient(DB)| \tag{13}$$

h,w and c are height, width and channel of gradient image This loss is used for making edges clear.

4.4 Implementation Details

The proposed architecture is implemented by TensorFlow[106] on NVIDEA GeForce GTX 1060 6GB with window 10 operating system. The version of python is 3.11.2 and the code is implemented in jupyter notebook 6.5.4. The original size of the images present in GOPRO dataset is 1280*720. The images are resized into 256*256 for processing. In the training of the GANs the batch size was set to 1 and runs for almost 150 epochs. The optimizer that is used in training process is Adam[109] and its variants and the values of learning rate is 0.0001.

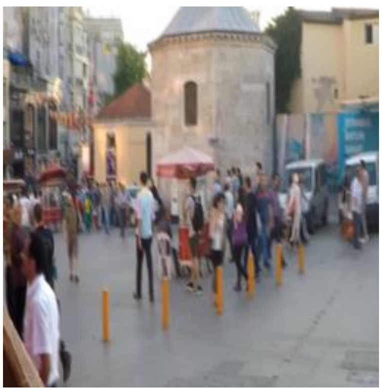
CHAPTER 5: EXPERIMENTS AND RESULTS

5.1 Dataset

The dataset usually used for the deblurring of images is GOPRO light dataset. It is popular for image deblurring algorithm. The focus of this dataset is mainly on the challenge of motion blur. Motion blur can be caused due to several reasons e-g shake in the camera or mat be low-light conditions. Some details of GoPro Datasets are:

- The images in the dataset are captured in lowlight condition. GoPro cameras are used for capturing images. Due to the longer exposure time motion blur is more prominent in the images of this dataset.
- The image pair, blurry image with its corresponding sharp image is available in this dataset. For training, testing and validation hundreds of images pairs are available.
- The blur type in GoPro light dataset is Motion blur which is caused due to the camera shake. This type of blur usually occurs when the unintentional movement of camera is present.
- The size of images in this dataset is 1280*720 and there are total of 3214 images that are present.
- Ground truth sharp images are present in the dataset that are used for evaluation.

This dataset is used mostly when deblurring is the major concern particularly when there is need to address motion blur in lowlight condition. This dataset helps the researchers to compare and evaluate the effectiveness of image deblurring models. Some Images from the dataset are:



Blurred Images

Clear Images

5.2 Performance Measures

5.2.1 Adversarial Loss

The adversarial loss is actually the measurement of how well the generator will be able to fool the discriminator. The goal of the training of GANs is to minimize this loss which leads to high-quality and visually pleasing generated images.

5.2.2 Perceptual Evaluation metrics

The quality of generated image is assessed by perceptual similarity. In this technique higher level features from the images are extracted and differences are measured rather than pixel wise difference. These metrics gives valuable results for the evaluation of visual similarity of generated image.5.2.3 Objective Evaluation metrics

The quality of generated image is assessed by quantifiable characteristics. PSNR and MSE (mean squared error) is usually used for evaluation purpose. Higher PSNR value and lowered MSE value indicates that the generated image is of better quality.

5.2.3 Visual Comparisons

In visual comparison, evaluation is done by human evaluators. The quality of generated image is measured against real images. Sometimes the objective metrics are not able to capture the small details in this type of cases visual comparison is of extreme importance.

5.3 Results

5.3.1 Graphical results on Training and validation dataset

In the Fig. 5.1, graphical representation of the training data is described. It has epochs at the x-axis and discriminator loss on y-axis. The loss first fluctuates between 0.6 and 1 and then keep on decreasing. We run our model for 150 epoch.

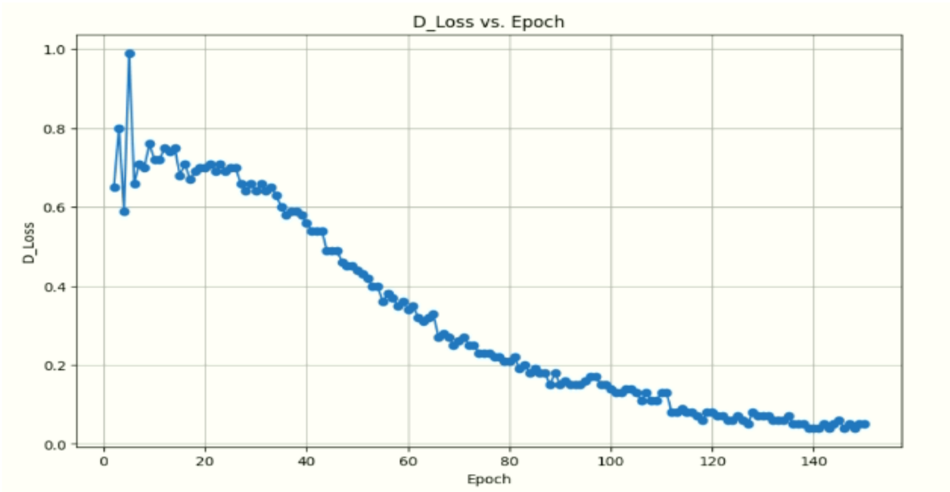


Figure 5.1: Graphical representation of training d_loss over epochs

In the Fig. 5.2, graphical representation of the training data is described. It has epochs at the x-axis and discriminator loss and generator loss (d on g loss) on y-axis. The loss starts decreasing from the start. The loss value is constant for some intervals and then again decreased. We run our model for 150 epoch.

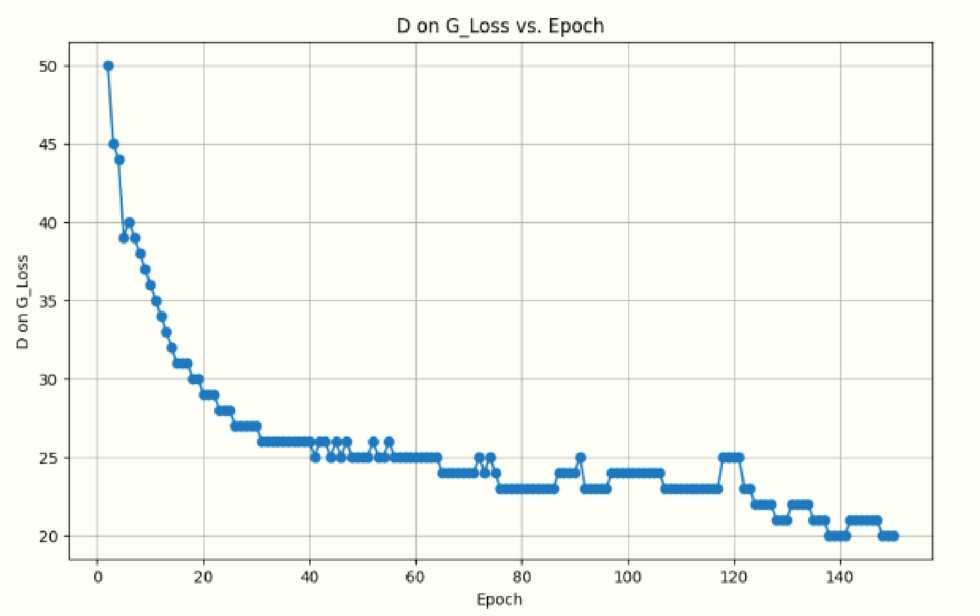


Figure 5.2: Graphical representation of training d on g_loss over epochs

In the Fig. 5.3, graphical representation of the validation data is described. It has epochs at the x-axis and discriminator loss on y-axis. The loss first remains constant and then keep on decreasing. After every three epoch we validates our results on validation dataset.

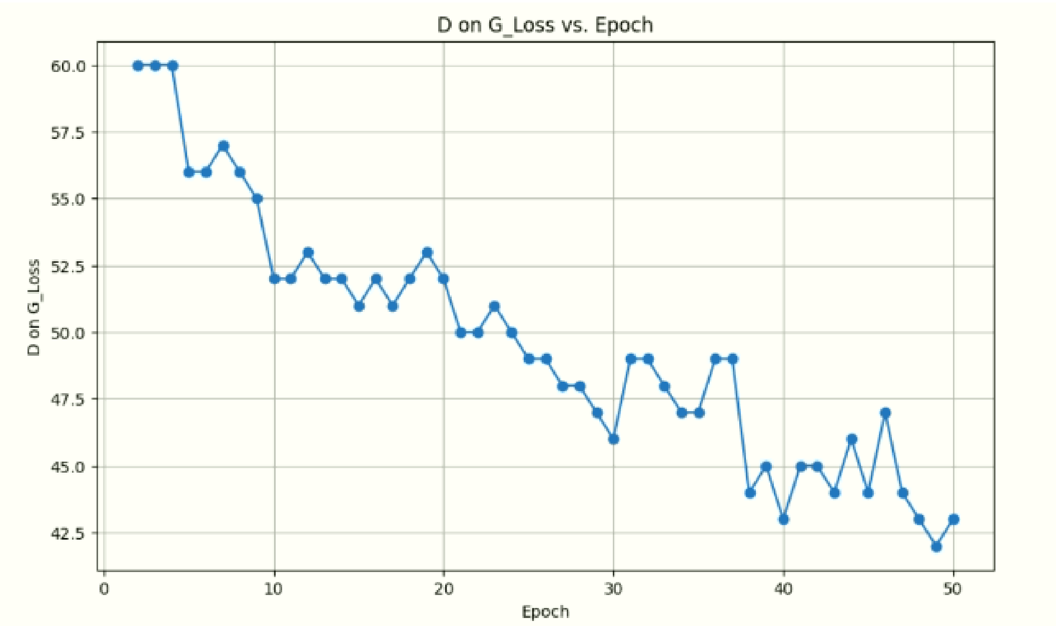


Figure 5.3: Graphical representation of validation d_loss over epochs

In the Fig. 5.4, graphical representation of the validation data is described. It has epochs at the x-axis and discriminator loss and generator loss (d on g loss) on y-axis. The loss fluctuates in the starts and then decreasing over progressive epochs. After every three epoch we validates our results on validation dataset.

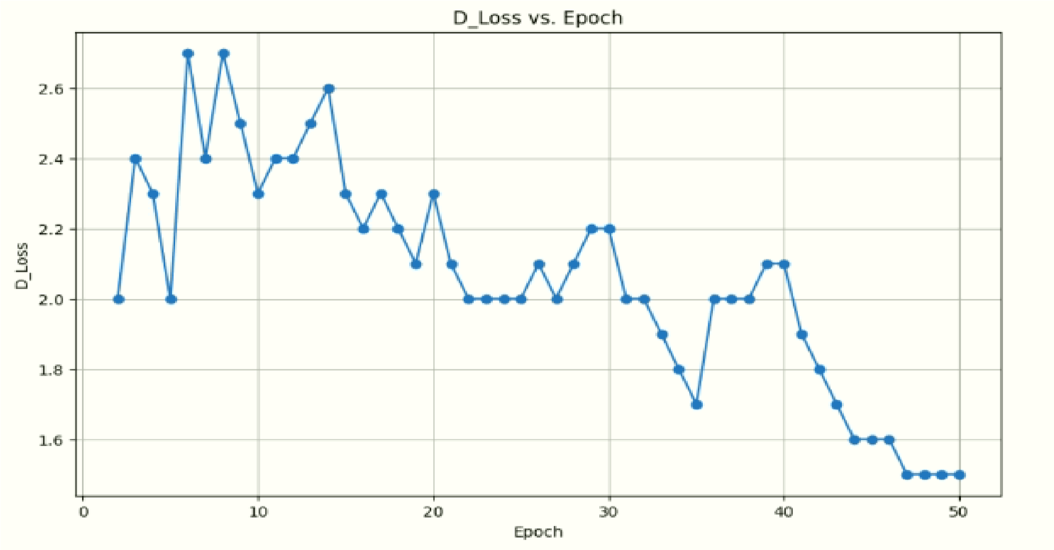
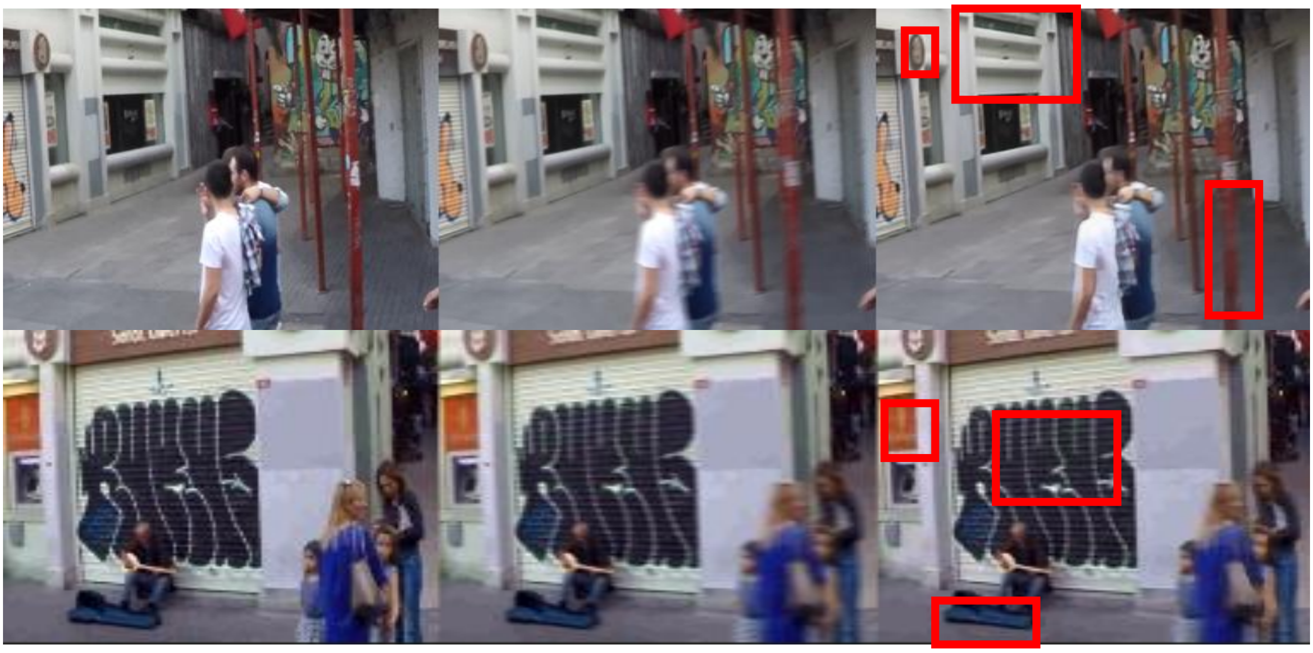


Figure 5.4: Graphical representation of validation d on g_loss over epochs

5.3.2 Visual results on Test dataset

These are the visual results on test dataset. First one is clear image second is blurred image and the third is generated. We can see the deblurring results in red box.

Original Blurred Generated



Original Blurred Generated

5.3.3 Objective metrics on Test dataset

The objective metrics that includes PSNR and SSIM is given below on table 5.1:

Dataset	Train/Test Data	Resolution	PSNR	SSIM
GoPro	**2103/1111**	**256*256*3**	**31.5**	**0.958**

Table 5.1: Results on GoPro dataset

5.4 Comparison with other State of Art methods

Our proposed attention based techniques shows remarkable results as compared to latest techniques.in the literature on GoPro benchmark dataset which is described in Table 5.2:

Author	Technique	PSNR/SSIM
Seo-Won Ji et al*[70]	XY- Deblur	30.97/0.95
Menhang Li et al*[71]	L0 Sparse Representation	29.89

Phong Tran et al*[72]	SRN Deblur	30.2
Suin M et al*[73]	Patch Hierarchical Network	31.85/0.95
Jaiwei Zhang et al*[74]	Spatially Variant RNN	29.18/0.93
Orest Kupyn et al*[66]	Deblur GAN	28.7/0.95
Dong Gong et al*[75]	FCN	26.40/0.72
Seungjun Nah et al*[76]	Multi Scale CNN	29.08/0.91
Jian Sun et al*[77]	CNN	24.64
Our Model	**Cross-Hatch Attention**	**31.5/0.95**

Table 5.2: Comparison with other State of art methods

CHAPTER 6: CONCLUSION & FUTURE WORK

6.1 Conclusion

In conclusion, the aim of this thesis is to address the challenging problem of image deblurring. Through extensive experimentation it is demonstrated successively that GAN-based approach outperforms in generating visually pleasing and high-quality images than the traditional approaches. The complex image texture details are captured by the utilization of GANs. The perceptual evaluation metrics such as PSNR and SSIM evaluates that the generated image has higher level of structural similarity with the real images. The proposed method outperformed several traditional techniques. Some limitations include computational requirements and sensitivity to different blur levels. This work contributes in the advanced image restoration methods and in many practical applications like medical imaging, satellite imagery and surveillance.

6.2 Contributions

- An in-depth investigation of different deblurring techniques.
- Development of GAN-based deblurring technique that addresses the challenging task of image deblurring which can be helpful in many practical applications such as medical imaging and surveillance.
- Comparison with different base line techniques.
- Use of hybrid loss function for capturing finer details.

6.3 Future Work

As a future work we can propose:

- The exploration of different methods which reduce the computational complexity of GANs for their practical applications.
- The investigation of different deblurring mechanism which are capable to handle varying blur kernel.
- The enhancement in the robustness of model is carried out for addressing the challenges that are encountered in deblurring process.
- The exploration of other regularization methods which improves the performance of deblurring further.
- The development of strategies that handle complicated blur types and motion blur which are common in real world.

References:

[1] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," arXiv:1406.2661 [stat.ML], Jun. 10, 2014

[2] A. Brock, J. Donahue, and K. Simonyan, “Large Scale GAN Training for High Fidelity Natural Image Synthesis,” arXiv preprint arXiv:1809.06222, 2018.

[3] M. Liu, F. Yuan, Y. Zhu, and E. Cheng, “Generating Underwater Images by GANs and Similarity Measurement,” in Proceedings of the OCEANS – MTS/IEEE Kobe Techno-Oceans (OCEANS), Publisher: IEEE, 2018]

[4] L. Courtrai, M.-T. Pham, C. Friguet, and S. Lefèvre, “Small Object Detection from Remote Sensing Images with the Help of Object-Focused Super-Resolution Using Wasserstein GANs,” in Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Year: 2020, Publisher: IEEE

[5] K. Zhang, W. Ren, W. Luo, W.-S. Lai, B. Stenger, M.-H. Yang, and H. Li, "Deep Image Deblurring: A Survey," International Journal of Computer Vision, vol. 130, pp. 1-28, Sep. 2022. doi: 10.1007/s11263-022-01633-5.

[6] S. Vasu, V. R. Maligireddy, and A. N. Rajagopalan, “Non-blind Deblurring: Handling Kernel Uncertainty with CNNs,” Indian Institute of Technology Madras,

[7] Fan L, Zhang F, Fan H, et al. “Brief review of image denoising techniques,” *Visual Computing for Industry*, Biomedicine, and Art, 2019, 2(1): 1-12.

[8] Baxes G A. “Digital image processing: principles and applications,” *John Wiley & Sons, Inc.*, 1994.

[9] Tomasi C, Manduchi R. “Bilateral filtering for gray and color images,” *International Conference on Computer Vision (ICCV)*. IEEE, 1998: 839-846.

[10] Muresan D D, Parks T W. “Adaptive principal components and image denoising,” *Proceedings International Conference on Image Processing (ICIP)*. IEEE, 2003, 1: I-101.

[11] Mallat S G. “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1989, 11(07): 674-693.

[12] Dabov K, Foi A, Katkovnik V, et al. “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Transactions on image processing*, 2007, 16(8): 2080-2095.

[13] Tian C, Fei L, Zheng W, et al. “Deep learning on image denoising: An overview,” *Neural Networks*, 2020.

[14] Gu S, Zhang L, Zuo W, et al. “Weighted nuclear norm minimization with application to image denoising,” *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014: 2862-2869.

[15] Schmidt U, Roth S. “Shrinkage fields for effective image restoration,” *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014: 2774-2781.

[16] Chen Y, Pock T. “Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration,” *IEEE transactions on pattern analysis and machine intelligence*, 2016, 39(6): 1256-1272.

[17] Mairal J, Bach F, Ponce J, et al. “Non-local sparse models for image restoration,” *2009 IEEE 12th international conference on computer vision*. IEEE, 2009: 2272-2279.

[18] Tamura S. “An analysis of a noise reduction neural network,” *International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 1989: 2001-2004.

[19] Bedini L, Tonazzini A. “Neural network use in maximum entropy image restoration,” *Image and Vision Computing*, 1990, 8(2): 108-114.

[20] Paik J K, Katsaggelos A K. “Image restoration using a modified Hopfield network,” *IEEE Transactions on image processing*, 1992, 1(1): 49-63.

[21] Sivakumar K, Desai U B. “Image restoration using a multilayer perceptron with a multilevel sigmoidal function,” *IEEE transactions on signal processing*, 1993, 41(5): 2018-2022.

[22] Krizhevsky A, Sutskever I, Hinton G E. “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, 2012, 25: 1097-1105.

[23] Zhang K, Zuo W, Chen Y, et al. “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE transactions on image processing*, 2017, 26(7): 3142-3155.

[24] He K, Zhang X, Ren S, et al. “Deep residual learning for image recognition,” *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770-778.

[25] Shwartz S, Schechner Y Y. “Blind haze separation,” *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. IEEE, 2006, 2: 1984-1991.

[26] Zhang K, Zuo W, Zhang L. “FFDNet: Toward a fast and flexible solution for CNN-based image denoising,” *IEEE Transactions on Image Processing*, 2018, 27(9): 4608-4622.

[27] Shaw S, Gupta R, Roy S. “A Review on Different Image De-hazing Methods,” *Emerging Technology in Modelling and Graphics*. Springer, Singapore, 2020: 533-540.

[28] Gu C, Lu X, He Y, et al. “Blur removal via blurred-noisy image pair,” *IEEE Transactions on Image Processing*, 2020, 30: 345-359.

[29] Mao X, Li Q, Xie H, et al. “Least squares generative adversarial networks,” *Proceedings of the IEEE international conference on computer vision*. 2017: 2794-2802.

[30] Narasimhan S G, Nayar S K. “Contrast restoration of weather degraded images,” *IEEE transactions on pattern analysis and machine intelligence*, 2003, 25(6): 713-724.

[31] Tan R T. “Visibility in bad weather from a single image,” *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008: 1-8.

[32] Li B, Peng X, Wang Z, et al. “An all-in-one network for dehazing and beyond,”*arXiv preprint*arXiv:1707.06543, 2017.

[33] Chen D, He M, Fan Q, et al. “Gated context aggregation network forimagedehazing and deraining,” *2019 IEEE winter conference on applications ofcomputer vision (WACV)*. IEEE, 2019: 1375-1383.

[34] Dong C, Loy C C, He K, et al. “Image super-resolution using deep convolutionalnetworks,”*IEEE transactions on pattern analysis and machine intelligence*, 2015, 38(2): 295-307.

[35] Lim B, Son S, Kim H, et al. “Enhanced deep residual networks for singleimagesuper-resolution,” *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017: 136-144.

[36] Hosseini M S, Plataniotis K N. “Convolutional deblurring for natural imaging,” *IEEE Transactions on Image Processing*, 2019, 29: 250-264.

[37] Richardson W H.“Bayesian-based iterative method of image restoration,”*JoSA*, 1972, 62(1): 55-59.

[38] Lucy L B. “An iterative techniquefor the rectification of observed distributions,”*The astronomical journal*, 1974, 79: 745.

[39] Biggs D S C, Andrews M. “Acceleration of iterative image restoration algorithms,”*Applied optics*, 1997, 36(8): 1766-1775.

[40] Kundur D, Hatzinakos D. “Blind image deconvolution,”*IEEE signal processing magazine*, 1996, 13(3): 43-64.

[41] Whyte O, Sivic J, Zisserman A, et al. “Non-uniform deblurring for shaken images,” *International journal of computer vision*, 2012, 98(2): 168-186.

[42] Amizic B, Molina R, KatsaggelosA K. “Sparse Bayesian blind image deconvolution with parameterestimation,”*EURASIP Journal on Image and VideoProcessing*, 2012, 2012(1): 1-15.

[43] Levin A, Weiss Y, Durand F, et al. “Efficient marginal likelihood optimization in blind deconvolution,” *CVPR 2011*. IEEE, 2011: 2657-2664.

[44] Levin A, Weiss Y, Durand F, et al. “Understanding blind deconvolution algorithms,” *IEEE transactions on pattern analysis and machine intelligence*, 2011, 33(12): 2354-2367

[45] Li J, Luisier F, Blu T. “PURE-LET image deconvolution,” *IEEE Transactions on Image Processing*, 2017, 27(1): 92-105.

[46] Afonso M V, Bioucas-Dias J M, Figueiredo M A T. “Fast image recovery using variable splitting and constrained optimization,” *IEEE transactions on image processing*, 2010, 19(9): 2345-2356.

[47] Liu G, Chang S, Ma Y. “Blind image deblurring using spectral properties of convolution operators,” *IEEE Transactions on image processing*, 2014, 23(12): 5047-5056.

[48] Krizhevsky A, Sutskever I, Hinton G E. “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, 2012, 25: 1097-1105.

[49] Sun J, Cao W, Xu Z, et al. “Learning a convolutional neural network for non-uniform motion blur removal,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015: 769-777.

[50] Wieschollek P, Hirsch M, Scholkopf B, et al. “Learning blind motion deblurring,” *Proceedings of the IEEE International Conference on Computer Vision*. 2017: 231-240.

[51] Schuler C J, Hirsch M, Harmeling S, et al. “Learning to deblur,” *IEEE transactions on pattern analysis and machine intelligence*, 2015, 38(7): 1439-1451.

[52] Li L, Pan J, Lai W S, et al. “Blind image deblurring via deep discriminative priors,” *International journal of computer vision*, 2019, 127(8): 1025-1043.

[53] Sun L, Cho S, Wang J, et al. “Edge-based blur kernel estimation using patch priors,” *IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2013: 1-8. [54]

Chakrabarti A. “A neural approach to blind motion deblurring,” *European conference on computer vision*. Springer, Cham, 2016: 221-235.

[55] Nah S, Hyun Kim T, Mu Lee K. ‘Deep multi-scale convolutional neural network for dynamic scene deblurring,” *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017: 3883-3891.

[56] Zhang H, Dai Y, Li H, et al. “Deep stacked hierarchical multi-patch network for image deblurring,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019: 5978-5986.

[57] Cai J, Zuo W, Zhang L. “Dark and bright channel prior embedded network for dynamic scene deblurring,” *IEEE Transactions on Image Processing*, 2020, 29: 6885-6897.

[58] Li L, Pan J, Lai W S, et al. ”Dynamic scene deblurring by depth guided model,” *IEEE Transactions on Image Processing*, 2020, 29: 5273-5288.

[59] Zamir S W, Arora A, Khan S, et al. “Multi-stage progressive image restoration,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 14821-14831.

[60] Rozumnyi D, Oswald M R, Ferrari V, et al. “DeFMO: Deblurring and Shape Recovery of Fast Moving Objects,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 3456-3465.

[61] Chi Z, Wang Y, Yu Y, et al. “Test-Time Fast Adaptation for Dynamic Scene Deblurring via Meta-Auxiliary Learning,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 9137-9146.

[62] https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/

[63] https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/

[64] M. M. Trivedi, C. X. Chen, and D. H. Cress, “Object detection by step-wise analysis of spectral, spatial, and topographic features,” Computer Vision, Graphics, and Image Processing, vol. 51, no. 3, pp. 235-255, September 1990.

[65] M. Mirza and S. Osindero, “Conditional Generative Adversarial Nets,” arXiv preprint arXiv:1411.1784, 2014.]

[66] DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks” Authors: Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, Jiří Matas Date: April 3, 2018 arXiv identifier: arXiv:1711.07064v4 [cs.CV]].

[67] S. Mascarenhas and M. Agarwal, “A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification,” in 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON), vol. 1, pp. 96-99, 2021, doi: 10.1109/CENTCON52345.2021.9687944.

[68] Abadi M, Agarwal A, Barham P, et al. “Tensorflow: Large-scale machine learning on heterogeneous distributed systems,” *arXiv preprint* arXiv:1603.04467, 2016.

[69]Kingma D P, Ba J. “Adam: A method for stochastic optimization,” *arXiv preprint* arXiv:1412.6980

[70]Ji S-W, Lee J, Kim S-W, Hong J-P, Baek S-J, Jung S-W, et al. XYDeblur:Divide and conquer for single image deblurring. In:Proceedings of theIEEE/CVF conference on computer vision and pattern recognition. 2022, p.17421–30.

[71] Li M, Gao S, Zhang C, Xu M, Zhang C. Blind motion deblurring via L0 sparse representation. Computer Graph 2021;97:248–57.

[72] Tran P, Tran AT, Phung Q, Hoai M. Explore image deblurring via encoded blur kernel space. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021, p. 11956–65.

[73] Suin M, Purohit K, Rajagopalan A. Spatially-attentive patch-hierarchical network for adaptive motion deblurring. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020, p. 3606–15.

[74] Zhang J, Pan J, Ren J, Song Y, Bao L, Lau RW, et al. Dynamic scene deblurring using spatially variant recurrent neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018, p.2521–9.

[75] Gong D, Yang J, Liu L, Zhang Y, Reid I, Shen C, et al. From motion blur to motion flow: A deep learning solution for removing heterogeneous motion blur. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 2319–28.

[76] Nah S, Hyun Kim T, Mu Lee K. Deep multi-scale convolutional neural network for dynamic scene deblurring. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 3883–91.

[77] Sun J, Cao W, Xu Z, Ponce J. Learning a convolutional neural network for non-uniform motion blur removal. In: Proceedings of the IEEE conference on computer vision and pattern

recognition. 2015, p. 769–77.
[78] Huang Z, Wang X, Huang L, Huang C, Wei Y, Liu W. CCNet: Criss-cross attention for semantic segmentation. In: Proceedings of the IEEE/CVF international conference on computer vision. 2019, p. 603–12.