Dynamic Image Deblurring Based on Crosshatch Attention Adversarial Network and Hybrid Loss

Ummama Aslam, Arslan Shaukat and Basim Azam

January 2, 2024

Dynamic Image Deblurring based on Crosshatch Attention Adversarial Network and Hybrid Loss

Ummama Aslam
Dept. of Computer and Software Engineering
College of Electrical and Mechanical Engineering (CEME)
National University of Sciences and Technology (NUST)
H-12 Islamabad, Pakistan
umamma.malik@gmail.com

Arslan Shaukat
Dept. of Computer and Software Engineering
College of Electrical and Mechanical Engineering (CEME)
National University of Sciences and Technology (NUST)
H12 Islamabad, Pakistan
arslan.shaukat@ce.ceme.edu.pk

Basim Azam
Institute of Integrated and Intelligent Systems
Griffith University
Brisbane, Australia
basimazam0@gmail.com

***Abstract*—Image deblurring techniques that uses deep learning have shown great potential but due to low generalizability, noise immunity and the correlation among different pixels is not addressed in detail that results in unwanted artifact that appears in the deblurred image. To tackle this problem an end-to-end approach is proposed for the recovery of sharp image from blurred one without the estimation of blur kernel. A special type of attention module known as crosshatch attention is used after Residual Block of Generator model for removing noise and for the collection of correlation of different pixels in an image. Hybrid Loss function is defined which focus on different part of image and improve edges and texture details. The performance of the model for deblurring is measured on GoPro dataset. Our proposed model has slightly higher objective and subjective evaluation i-e PSNR, SSIM value and the visual results.**

Index Terms— adversarial Training, blind Image deblurring, image restoration, generative adversarial networks, GANs

I. INTRODUCTION

The deblurring of images is the critical task 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 an image as various factors are involved while capturing e-g motion blur, shake of camera or noise can be present. Blurred image can cause problem in facial recognition, [1] generating underwater images [2], medical image recognition [3] and video surveillance [4]. Blurred images are the result of moving object or due to the movement of camera. The prevention of image degradation and at the same time improvement in the quality of image are the major concerns. This paper solely focuses on dynamic image deblurring. The motion blur usually appears when there is dynamic movement between the camera and the captured object. The reason of motion blur can be fast moving object or shake in camera. The blurring in an image is the result of relative motion between the capturing device and object diminishes the precision of visual perception of human, but it can be challenging for further computer vision and image processing analysis. Additionally, to reduce the image degradation caused due to deblurring and the enhancement of quality of image simultaneously has been a major concern in the realm of image processing. Image deblurring is useful in the recovery of lost details and enhanced image analysis in many fields. It includes an extensive range of applications because it is a valuable asset in many fields. Generally, dynamic blur is defined as:

$$B = C.H + N \tag{1}$$

Where B is referred as the blurred image, C is the clear image, (.) defines convolution, H is the blurring filter and N is the additive noise that is present in the image. The image deblurring algorithm is categorized into two types: Blind image deblurring [5-10] and non-blind image deblurring [11-13]. Non-Blind image deblurring methods are not able to perform well due to estimation of blur kernel. 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 estimating the kernel that cause blur. 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. Xu et al. [13] introduced a technique for fixing the blur in an image using deep learning without knowing the details of blur. Later on, various deep learning approaches has been proposed for the estimation of blur kernel and the image quality of traditional methods is emerged in recent times [15-17]. Sun et al. [14] suggested a CNN technique for successfully estimating the blurred kernel but they assume that the image is blurred uniformly across its local regions Gong et al. [15] Employed a fully convolutional neural network for predicting directly the flow of motion within the blurred image. Chakrabarti et al. [16] developed a technique for blurred kernel estimation and image features are extracted by using different filters and are combined with neural networks. The blurred kernel is estimated of varying sizes, but it tends to blur edge feature in the resultant image. Ringing artifact is present in these techniques as the accuracy for estimating the kernel that cause blur is not sufficient results in non-uniform blur across the image. Nah et al. [5] suggested a multi-scale CNN approach for the enhancement of dynamic scene deblurring by the refining the image from coarse to fine level. However, this approach depends on the significant capacity neural network and the proposed network is built in a way that it does not pay attention of how the blurring in an image happens actually. Tao et al. [6] suggested a scaled recurrent network. The complex network structure are improved by the use of Recurrent Neural Networks (RNN) to share weight parameters at each scale. Kupyn et al. [7] suggested a novel technique for the deblurring of image. This method is built on Generative Adversarial Network (GAN) and called Deblur GAN. The core block for generator is Residual Network. In the objective evaluation, it yielded average performance. An updated version of Deblur GAN is presented called Deblur GAN-v2 [17]. It incorporates pyramid network in the frame work of GANs. Both the above network shows that the image details are preserved effectively through GAN architecture. Zhang et.al [8] created hybrid model which is the combination of RNN and three CNN to address the deblurring of images. This technique learns the importance of spatial variation weights and achieved comprehensive deblurring. An equilibrium between accuracy and runtime is achieved without multiscale architecture and adversarial training [12-14]. Tran et al. [20] suggested an enhanced technique operating in an encoded blur kernel space for the elimination of blurriness. However, the existing methods have intricate network structure, or they can be effective only in particular scenario. Furthermore, they are not suitable for handling complex blur. In contrast to conventional techniques of blind and non-blind image deblurring, the blind image techniques that are built on the methods of deep learning and can attain promising results without any assumptions. However, these methods still have some limitations. In dynamic scene deblurring, various types of blurs exist in different sections of images and the pixel connection is strong within each region. The random feature information processing at different location of images neglects the correlation among different pixels and the network is not able to learn effectively.

This paper primarily emphasizes on end-to-end GAN model and handles noise in context of image deblurring. The approaches based on deep learning for image deblurring offers strong generalizability. However, significant amount of data is required for this approach. A developed version of Deblur GAN is presented to address the issue of dynamic image deblurring. The Resblock is combined with the attention module. Different features are obtained from the images and correlation among them is calculated. A Hybrid loss function is designed to balance various aspects of image quality, pixel accuracy and structure preservation. This will enhance the performance of overall deblurring.

The main focus of this paper is as follow:

- A novel algorithm is developed specifically for the dynamic deblurring of image using GANs.
- Crosshatch attention is proposed after Residual block for capturing the correlation between the pixels of image, which enhances the deblurring capabilities of the architecture.
- A hybrid loss function is designed for the enhancement in the quality of overall image. The loss captures multiple aspects leading to better deblurring results.
- A comprehensive comparison is presented with other state-of-art techniques to showcase the impact of proposed architecture for the deblurring of images on GoPro benchmark dataset.

The subsequent sections of the paper are structured as: Section 2 provides the literature review of the related research. Section 3 outlines our suggested approach encompassing the architecture of network and the employed loss functions for our algorithm. Section 4 describes the comprehensive empirical results and verifies the proficiency of different techniques. Section 5 provides conclusive insights and findings from this paper.

II. RELATED WORK

In this chapter, the literature review of image deblurring algorithm is presented. The traditional image deblurring techniques usually focus on non-blind image deblurring. Blurry images are restored if we know about blur kernel. The blurred kernel estimation is based on assumptions. This assumption is used for the restoration of image. However, if the blur kernel is complex, it is not easy to estimate it completely. In simpler terms, the specific blur pattern for each pixel is a challenging problem in the real world. These traditional techniques often rely on making assumptions to create deblurred images. However, these techniques are complicated and unwanted artifacts are noticeable in the restored image.

In recent times, due to advancement in deep learning researchers have embraced techniques that are based on deep learning for the task related to deblurring of images. Different types of neural networks including Convolutional Neural Network (CNN) and various types of Generative Adversarial Network have gained popularity and extensively employed on the tasks related to deblurring of images.

In the previous years, the techniques that are based on CNN models are employed for the estimation of blurred kernel and this blurred kernel is used for the removal of blur that are non-uniform. For example, Gong et al. [15] has employed fully convolutional neural network for motion flow prediction in the entire image. The image is estimated by the prediction of motion flow. This approach uses blind image deblurring technique for the enhancement of image, but they stick to the non-blind image deblurring [31, 38, 39] idea i-e to perform particular operation on blurred image. The quality of the result of deblurring algorithm heavily depends on how accurately the blurred kernel is estimated. Sun et al. [14] introduced an approach for the elimination of non-uniform motion blur. The convolutional neural networks are employed for motion blur kernel estimation for different patches of image and applied Markov random field architecture to simulate motion blur that is uniform. The aim of this technique is to get improved clear image with changing motion blur pattern.

In recent times, there has been a growing trend for the use of end-to-end trainable deblurring techniques for dynamic deblurring. These methods bypass the need of estimation of blurred kernel making the process of deblurring more efficient and it does not rely on the determination of blur characteristics. For example, Tao et al. [6] introduced a scale recurrent network for single image deblurring. The number of parameters is decreased in this approach in multiscale framework because it utilize shared network weights across different scales. However, it's important to consider that this parameter sharing technique overlooks the characteristics of image feature that are scale specific and they are crucial for effective image restoration across different scales. Nah et al. [5] presented the idea for deblurring of images using multi scale CNN. Blurry image is taken as the input, and it produce deblurred image at the output. The training held in an end-to-end way, which means it learns to deblur the image from coarse to fine. The multiscale framework of the network allows to learn the features at different scales as it takes multi-resolution blurry image at the input which is important for deblurring of images that are blurred at various stages. Kupyn et al. [7] suggested DeblurGAN for the deblurring of images effectively. These GANs are based on Conditional Adversarial Networks [27]. They have used Generative Adversarial Networks with Gradient penalty. This technique follows an end-to-end technique for the process of deblurring i-e without the estimation of blurred kernel. Perceptual loss is added as optimization objective. In the perceptual loss calculation features are obtained from the images and their difference is measured. The empirical results shows that this method attains better visual effect than multi-scale convolutional networks. This method also has the simple network structure and faster running speed. However, it is still not capable to handle severely blurred images. Ji et al. [22] suggested a method for deblurring of images. U-Net structure with two decoders are used in this method. This method is called XY- Deblur. The encoder part extract features from blurry image and decoder part repairs the image. The two decoders learn to deblur the image in different ways. Output from both decoders is combined to reconstruct the blurred image. The two decoders improve the deblurring performance and model size is also not increased. The experiment results shows that this XY deblur shows better performance than the simple U-Net structure. Zhang et al. [8] suggested the technique that combines Recurrent Neural Network with three CNN for achieving end-to-end deblurring. This method does not require adversarial training and multi-scale architecture. A good balance between accuracy and speed is achieved but a complex structure is required which is not ideal for handling complex blurs.

The success of the mechanism of attention in the processing of natural language is an inspiration for the researchers to use them in the tasks related to computer vision. This mechanism can be used in classification of images, to focus on relevant parts of image. In object detection, they can also be used to detection and movement of object. In segmentation of images, they can be used for the identification of different regions in an image. In the restoration of image, they can be used for the removal of noise and blur from the image. Wang et al. [23, 24] proposed mechanism of attention into residual network for the improvement of performance of classification of images. They also proposed non-local neural networks. The dependencies in long-range are captured between the pixels of image. Hu et al. [25] used the excitation and squeeze block for the modeling of correlation between the channels of image. The expressiveness of model is improved by them which means more complex features are learned from the image. Fu et al. [26] presented a dual attention network. The integration of local features and global dependencies is done by modeling both inter-channel and spatial correlation. Feature expression ability of the network is improved for segmentation of scenes tasks. Zhang et al. [8] suggested a GAN- algorithm with self-attention which is often useful for the tasks related to the restoration of image. Suin et al. [40] suggested the attention mechanism for the tasks related to dynamic deblurring. The method is called Attention Guided Motion Deblurring Network. It is the combination of two networks attention network and deblurring network. Attention weights are computed for each pixel in blurred image in the first network and then these weights are used by the second network to focus on more relevant part of images. This network has shown effective results for motion deblurring but for high resolution image this network requires more GPU memory for high level calculations.

III. PROPOSED METHODOLOGY

In this chapter, the proposed method is defined. We first discuss the classical GANs architecture and then the improved version of GANs is described in detail.

GANs are proposed in 2014, initially in 2014 by Ian Good fellow et al. [27]. It represents a type of deep learning model. GANs have gained a great achievement in a variety of applications that includes image translation, conversion of one domain of image into some other domain [41], text-to-image synthesis, text information is converted into visual data [42] and Style transfer, artistic style of one image is applied to another [43]. This network also shown great success in image in-painting, restoration of missing parts of image [44, 45] Many GANs based variants have been proposed as the classic GANs has the issue of mode collapse and vanishing gradients.

The improved approach of GANs falls under the framework of c-GANs [21]. Unlike the classic GANs, the generator model takes blurred version of image at its input instead of random noise. The proposed framework comprises of four convolutional layer connected in series with nine Residual block and crosshatch attention block followed by two Deconvolutional layers. The configuration of layer one and four are identical. Layer one has 64 filters and layer 4 has 3 filters. Kernel size in stride in both layers are same i-e is 7 and 1 respectively. The two and three convolutional layers follows same structure. In the second layer, there are 256 filters and in the third layer they are 512. A kernel size and stride is same in both i-e 3 and 2 respectively. A Residual block comprises of a Batch normalization layer, ReLu layer and 3*3 Convolutional layer. 256 filters are present in this layer, the size of kernel is 3

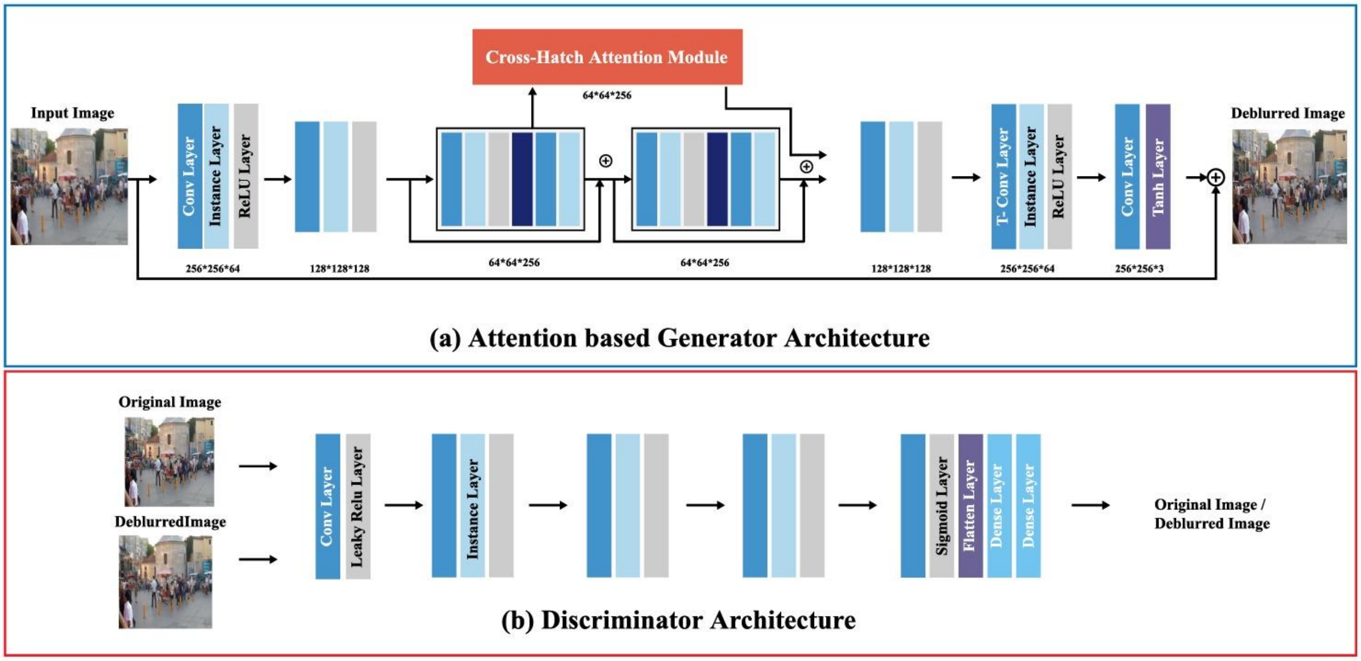


Fig. 1. Network overall Architecture.

and the stride is 1. The rate of dropout is 0.5 and identity connection are present followed by the crosshatch attention module inspired by [28].

3.1 Crosshatch Attention module

The suggested attention module incorporates the crosshatch attention technique inspired by [28] because it is efficient in capturing inter-pixel correlation of the whole image. The GPU memory usage and floating-point operations are reduced effectively. This attention module builds upon the concept of self-attention [29] this attention module gathers the information of spatial attention by the calculation of correlation between the concerned pixel and its other pixel within its row and column in a single feature map.

Initially, the given local feature map having the dimensions of $C * W * H$ is processed via three 1*1 convolutional layer for the production of three new information maps i-e Q, K and V. The new feature map has reduced channel dimensions than the original feature map. At every spatial position P in Q a vector Q_p is generated. Simultaneously, for every position in K, a vector K_p is generated which consists of Eigen vectors that corresponds to the other position in the same row and column. The correlation weights of the features are computed in this attention map using softmax function. The output is calculated by applying these attention scores to the feature map V. This is how we can calculate the horizontal attention. The vertical attention is calculated in the same way by taking the transpose of the feature map before passing it to softmax function. The output is again transposed to add it to the horizontal features. Horizontal and vertical outputs are added and scaled by the factor Gamma. The contextual details are accumulated by combining both the vertical and horizontal attention within the original feature map.

The blurred image is passed through one 7*7 filter and two 3*3 filters in the convolution operation. The features of image are extracted at pixel level. These features are than passed through 9 residual network and deep features are extracted. The features at the output of residual network are deconvolved again to keep the size of deblurred image and sharp image same. To avoid the chess board effect the deconvolution step is replaced up-sampling. The input of the first layer has the direct connection to the last layer with global skip connection. This is the generator network of the proposed architecture. It generates a deblurred image. This deblurred image is than combined with ground truth image and both these images are fed to discriminator network. The output of generator network acts as the input to the discriminator network i-e 256*256*3. The function of discriminator network is to determine if the input image is real or fake. It consists of series of convolutional layer along with batch Normalization and Activation. The activation function used here is Leaky Relu [42] and its parameter value is 0.2. In the end the output is passed through sigmoid, flatten and dense layer and the output at final stage is calculated within the range of 0 and 1 i-e if the image is clear or generated. The job of the discriminator is like the assistant of the generator which let the generator to produce deblurred image which is very near to its corresponding sharp image.

We design hybrid loss function in our proposed architecture. The selection of loss function is linked with the effectiveness of architecture. In the architecture of Deblur GAN, loss function is the combination of Adversarial loss and Content loss. We combined these two loss functions with two more loss functions i-e pixel loss and gradient loss to compute pixel wise difference and edge details for the extraction of more detailed overview of images. The objective of adding adversarial loss is restoration of texture details while the

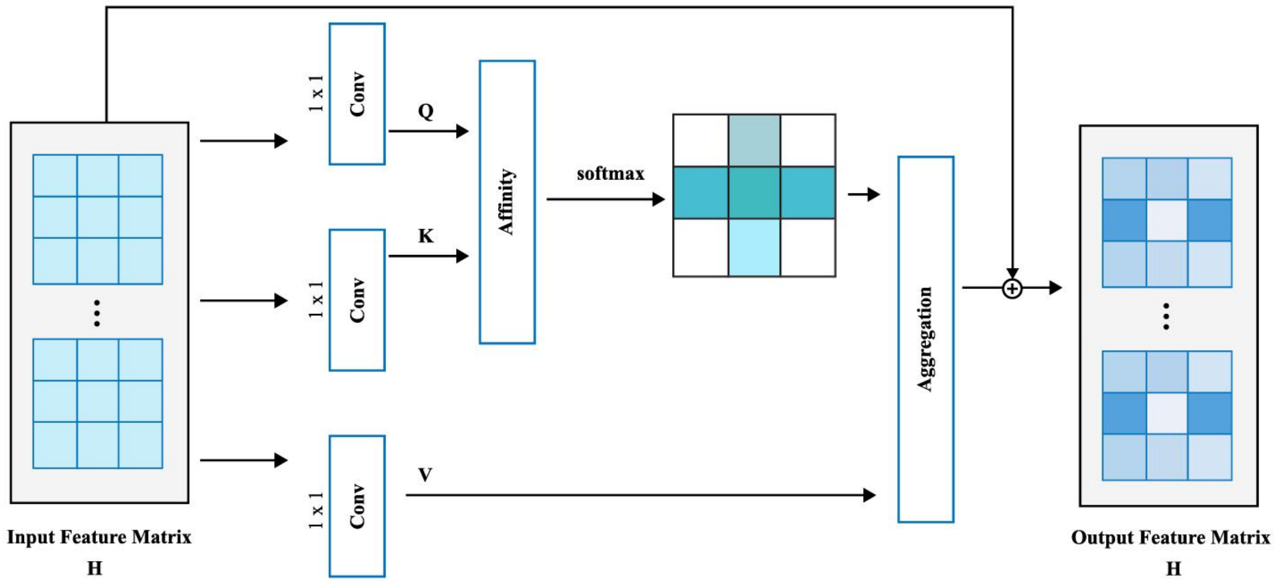


Fig. 2. Cross Hatch Attention Module

content loss is responsible for the restoration general content of image. With the help of loss function the discriminator and generator are trained in an adversarial manner to reach a state of Nash equilibrium. The details of loss functions are given below:

3.2 Hybrid Loss

The Hybrid loss in our proposed framework integrates different loss functions. Hybrid loss incorporates Adversarial loss, Content loss, Pixel loss and Gradient loss into one loss function whose weights can be adjusted for better results. Adversarial loss is responsible for training the model to generate deblurred image that looks convincing. Content loss is used for high-level feature mapping and to preserve texture details. Pixel wise similarity is ensured in pixel loss and smooth transitions are maintained by Gradient loss function. The Hybrid loss is defined as:

$$H_{Loss} = \alpha . adv_{loss} + \beta . Con_{loss} + \gamma . Pixel_{loss} + \delta . Grad_{loss} \tag{2}$$

Where H_{Loss} is the hybrid loss. The hyper parameters α, β, γ *and* δ are set empirically to 4, 0.6, 40, and 40. This collective loss function is responsible for guidance of model to generate realistic looking images that are rich in fine details and minimizing unnecessary artifacts. The details of these loss functions are given below:

3.2.1 Adversarial Loss

Adversarial loss is the combination of generator loss and discriminator loss. It is used for the reconstruction of details of texture of image. It makes the generated image looks more appealing. In the last few years, due to the rapid development of GANs numerous adversarial loss functions are emerged. We have used least mean squared loss because its effectiveness is superior to MSE. The adversarial loss for the discriminator and generator are expressed as:

$$D_{Loss} = \frac{1}{2} . \mu[(R(I) - \mu\big(F(I)\big) - 1)^2] + \frac{1}{2} . \mu[(F(I) - \mu(R(I) + 1)^2] \tag{3}$$

And

$$G_{Loss} = \frac{1}{2} . \mu[(F(I) - \mu(R(I)) - 1)^2] \tag{4}$$

Where R (I) is real image and F (I) is fake image.
This loss encourages the generator model to produce more visually appealing image by training the discriminator model to learn and differentiate between real and fake image.

3.2.2 Content Loss

The content loss is calculated to maintain content consistence of real and fake image. This loss is used commonly in the super-resolution of images for making the deblurred image close to human perception. This loss is calculated as:

$$Con\_Loss = \mu(true\ features - predicted\ features)^2 \tag{5}$$

The features are extracted with pre-trained VGG-19 model on image net. The feature difference is calculated in this. The focus is to preserve high level features by comparing the representation of real and fake image.

3.2.3 Pixel Loss

Pixel loss is used to confirm that the generated image resembles to the ground truth image closely in features details. This loss function is beneficial for the rectification of color and texture distortion. Smooth L1 function is used as a pixel loss function the pixel-wise difference between the target image and generated image is measured and this difference is minimized for the reconstruction of image.

$$Pixel\ Loss = SmoothL1[f(i), r(i)] \tag{6}$$

$$\{\frac{1}{2}\|f(i) - r(i)\|2, \text{ if } \|f(i) - r(i)\| < 1\|f(i) - r(i)\| - \frac{1}{2}, \text{ otherwise}$$

The L1 and L2 losses are combined together in smooth L1 loss function [34]. This loss function is promoting smoother convergence. It provides gradient explosion because it is less sensitive to outliers. It protects the relevant features and discard irrelevant features in motion image deblurring.

3.2.4 Gradient Loss

The edge features of sharp image are distinct and blurred image are chaotic. We have used image gradient difference loss function [30] for the enforcement of sharp edge difference. It enhances deblurred image and it also compensates the over-smoothness effect which is caused by pixel loss. The sobel operator is applied for the calculation of gradient.

$$Grad\_Loss = \mu|true\ gradient\ - predicted\ gradient| \tag{8}$$

This loss function is used for the preservation of edge information and for the reduction of unnecessary artifacts from the image.

IV. EXPERIMENTS AND RESULTS

The GoPro dataset [5] is known for its wide usage and high resolution, for the tasks of image deblurring research. It is the combination of 3214 image pair of blurred and sharp. The resolution of these images is 1280*720*3. The training data comprises of 2103 pair of images and test data has 1111 images. For faster training, the images that are present in the dataset are resized to 256*256*3. The framework of Keras deep learning is used for the implementation of development of model. A batch size is set to 1 for individual image restoration. ADAM [32] serves as an optimizer and its learning rate is set to 0.0001 and the other hyper parameters are set to default. The generator is trained once while the discriminator is trained for five iterations. The network model is trained until it reaches the stable convergence point. The system hardware configuration includes Intel core i7 CPU, 4.20 GHz and an NVIDIA GeForce GTX 1060 6GB GPU.

The effectiveness of this technique is evaluated by applying it for the restoration of GoPro dataset images. To access the quality of image restoration three evaluation parameters are taken into account which includes Peak Signal to Noise Ratio [33], Structural Similarity index [33] and visual comparisons. PSNR value measures the difference between the pixel value of deblurred image and ground truth image. More value of PSNR indicates that there is less distortion in the reconstructed image. SSIM access image similarity by taking the structural information, contrast and brightness taken under account. More the value is closed to one better is the quality of reconstructed image. Visual evaluation is the valuable step for accessing the performance of algorithm of deblurring because it complements quantitative metrics by offering insights into the extent to which the reconstructed image match human perception. The mathematical representation of PSNR and SSIM is given below:

$$PSNR = 10.log_{10}(\frac{MAX^2}{MSE}) \tag{9}$$

Methods	PSNR	SSIM
Sun et al.[14]	24.64	0.842
Xu et al.[13]	25.10	0.890
Gong et al.[15]	26.40	0.863
Kupyn et al.[7]	28.70	0.858
Nah et al.[5]	29.08	0.913
Zhang et al.[8]	29.19	0.930
Kupyn et al.[17]	29.55	0.934
Li et al.[9]	29.89	0.937
Tao et al.[6]	30.10	0.932
Tran et al.[20]	30.2	0.939
Ji et al.[22]	30.97	0.950
Ours	**31.50**	**0.958**

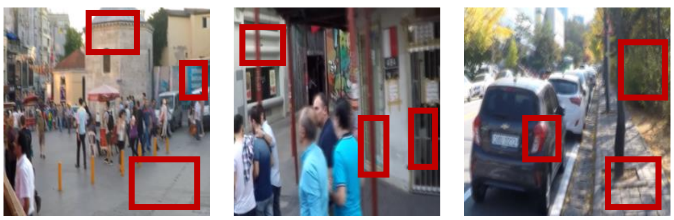
Table 1: Comparison of Average PSNR and SSIM on GoPro dataset

Where MAX is the maximum possible pixel value and MSE is the mean squared error between original and reconstructed image.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{10}$$

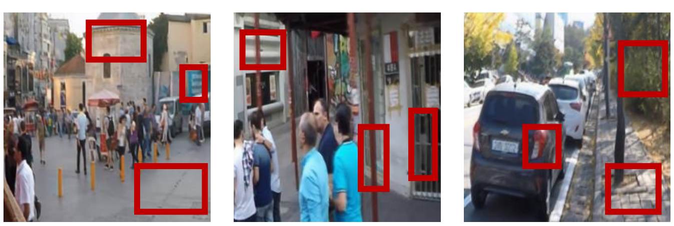
Where x and y are input images, μ_x and μ_y are means of x and y, σ_{xy} is the covariance of x and y, σ_x and σ_y are the standard deviations of x and y and C_1 and C_2 are the constant to stabilize division. The quantitative outcomes of PSNR and SSIM are given in the Table 1. The average Peak Signal to Noise Ratio (PSNR) [33] and average Structural Similarity Index Measure (SSIM) [33] is presented for the reconstruction of images by our proposed methodology. Various State-of-art techniques have used the same GoPro Dynamic Scene dataset. Our proposed approach out-performs from all the compared methods. Specifically, our proposed architecture achieved remarkable results with PSNR and SSIM values of 31.5 and 0.9582 respectively.

The Fig. 3 represents visual comparison of three examples which showcased that the results achieved by our proposed method using the GoPro dynamic scene dataset. It is evident by the results that the method, which aims to reconstruct blurry image by the estimation of motion blur kernel struggles to construct clear image and exhibit artifact in the results of deblurred image. Similarly, due to the limitation of receptive field methods for deblurring dynamic scenes which are based on CNN [35-37] encounters challenges with severe blur. In contrast our proposed method achieved finer texture details and the structure of images are clearer as compared to the several existing state of art techniques especially when the blur is severe.



(a) Blurred

(b) Clear



(c) Generated

Fig.3. Visual Comparison on GOPRO Dataset

V. Conclusion

In this research image deblurring technique is introduced that feature cross-grid attention module. This module is integrated with the Residual block for the better handling of complex blur present in an image. A discriminator is used to complement the grid attention module which enables the network to extract and utilize broader range of information of the image feature due to which the network learning capabilities are enhanced. Furthermore, a hybrid loss is designed for the guidance of whole GAN model towards the details of image and during the process of training the texture consistence is maintained between the deblurred image and clear image.

Extensive experiments are performed using the GoPro dynamic scene dataset the results exhibit better perceptual and quantitative performance as compared to the existing state of art methods. Notably, our proposed architecture has significantly faster runtime. In future more effective training strategies can be introduced for further enhancing the outcomes of deblurring.

References

.

[1] M. Nishiyama, A. Hadid, H. Takeshima, J. Shotton, T. Kozakaya, and O. Yamaguchi, “Facial deblur inference using subspace analysis for recognition of blurred faces,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 4, pp. 838–845, 2011.

[2] M. Liu, F. Yuan, Y. Zhu, and E. Cheng, "Generating Underwater Images by GANs and Similarity Measurement," Key Laboratory of Underwater Acoustic Communication and Marine Information Technology, Ministry of Education, Xiamen University, China, doi: 10.1109/ACCESS.2021.3086833.

[3] J. Ma, H. Xu, J. Jiang, X. Mei, and X. P. Zhang, “DDcGAN: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion,” IEEE Transactions on Image Processing, vol. 29, pp. 4980–4995, 2020.

[4] C. H¨ane, L. Heng, G. H. Lee et al., “3D visual perception for self-driving cars using a multi-camera system: calibration, mapping, localization, and obstacle detection,” Image and Vision Computing, vol. 68, pp. 14–27, 2017.

[5] 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.

[6] Tao X, Gao H, Shen X, Wang J, Jia J. Scale-recurrent network for deep image deblurring. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018, p. 8174–82.

[7] Kupyn O, Budzan V, Mykhailych M, Mishkin D, Matas J. Deblurgan: Blind motion deblurring using conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018,p. 8183–92.

[8] 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. 2521-2529 Yue T, Suo J, Dai Q. Non-uniform image deblurring using an optical computing system. Comput Graph 2013;37(8):1039–50.

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

[10] Fang Y, Zhang H, Wong HS, Zeng T. A robust non-blind deblurring method using deep denoiser prior. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022, p. 735–44.

[11] Nan Y, Ji H. Deep learning for handling kernel/model uncertainty in image deconvolution. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020, p. 2388–97.

[12] Vasu S, Maligireddy VR, Rajagopalan A. Non-blind deblurring: Handling kernel uncertainty with CNNs. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018, p. 3272–81.

[13] L. Xu, J. Ren, C. Liu, and J. Jia, “Deep convolutional neural network for image deconvolution,” Advances in Neural Information Processing Systems, vol. 27, no. 1, pp. 1790–1798, 2014.

[14] J. Sun, W. C. Wenfei, Z. X. Zongben, and J. Ponce, “Learning convolutional neural network for non-uniform motion blur removal,” in Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, pp. 769–777, Boston, MA, USA, June 2015

[15] D. Gong, J. Yang, L. Liu et al., “From motion blur to motion flow: a deep learning solution for removing heterogeneous motion blur,” in Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3806–3815, Honolulu, HI, USA, July 2017.

[16] A. Chakrabarti, “A Neural Approach to Blind Motion Deblurring,” in Proceedings of the European Conference on Computer Vision, pp. 221–235, Springer, Amsterdam, +e Netherlands, October 2016.

[17] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, “DeblurGANv2: deblurring (orders-of-magnitude) faster and better,” in *Proceedings of the 2019 IEEE International Conference on Computer Vision*, pp. 8877–8886, Seoul, Korea (South), October 2019.

[18] Hu X, Ren W, Yu K, Zhang K, Cao X, Liu W, et al. Pyramid architecture search for real-time image deblurring. In: Proceedings of the IEEE/CVF international conference on computer vision. 2021, p. 4298–307. Zhang T, Li J, Fan H. Progressive edge-sensing dynamic scene deblurring. Comput Vis Media 2022;8(3):495–508.

[19] Cho S-J, Ji S-W, Hong J-P, Jung S-W, Ko S-J. Rethinking coarse-to-fine approach in single image deblurring. In: Proceedings of the IEEE/CVF international conference on computer vision. 2021, p. 4641–50.

[20] 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.

[21] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” 2014.

[22] 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 the IEEE/CVF conference on computer vision and pattern recognition. 2022, p.17421–30.

[23] Wang F, Jiang M, Qian C, Yang S, Li C, Zhang H, et al. Residual attention network for image classification. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 3156–64.

[24] Wang X, Girshick R, Gupta A, He K. Non-local neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018, p. 7794–803.

[25] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018, p. 7132–41.

[26] Fu J, Liu J, Tian H, Li Y, Bao Y, Fang Z, et al. Dual attention network for scene segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019, p. 3146–54.

[27] I. J. Goodfellow, J. Pouget-Abadie, and M. Mirza, “Generative adversarial networks,” 2014.

[28] 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. 603–612.

[29] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Adv Neural Inf Process Syst 2017;30:5998–6008.

[30] Mathieu M, Couprie C, LeCun Y. Deep multi-scale video prediction beyond mean square error. 2015, arXiv preprint arXiv:1511.05440.

[31] Schmidt U, Rother C, Nowozin S, Jancsary J, Roth S. Discriminative nonblind deblurring. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2013, p. 604–11

[32] D. P. Kingma and J. L. Ba, "ADAM: A Method for Stochastic Optimization," in International Conference on Learning Representations (ICLR), 2015

[33] A. Horé and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in 2010 International Conference on Pattern Recognition.

[34] Girshick R. Fast R-CNN. In: Proceedings of the IEEE international conference on computer vision. 2015, p. 1440–8.

[35] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Visual Geometry Group, Department of Engineering Science, University of Oxford, Oxford, UK, 2015.

[36] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical Evaluation of Rectified Activations in Convolution Network," University of Alberta, Hong Kong University of Science and Technology, University of Washington, Carnegie Mellon University, 2015.

[37] Schmidt U, Roth S. Shrinkage fields for effective image restoration. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014, p. 2774–81.

[38] Schmidt U, Rother C, Nowozin S, Jancsary J, Roth S. Discriminative nonblind deblurring. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2013, p. 604–11.

[39] Zhang J, Pan J, Lai W-S, Lau RW, Yang M-H. Learning fully convolutional networks for iterative non-blind deconvolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 3817–25.

[40] 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.

[41] W. Xu, K. Shawn, and G. Wang, “Toward learning a unified many-to-many mapping for diverse image translation,” *Pattern Recognition*, vol. 93, pp. 570–580, 2019.

[42] X. Sun and J. He, “A novel approach to generate a large scale of supervised data for short text sentiment analysis,” *Multimedia Tools and Applications*, vol. 79, no. 9-10, pp. 5439–5459,2020.

[43] X. Chen, C. Xu, X. Yang, L. Song, and D. Tao, “Gated-GAN: adversarial gated networks for multi-collection style transfer,” *IEEE Transactions on Image Processing*, vol. 28, no. 2, pp. 546–560, 2019.

[44] Y. Chen, H. Zhang, L. Liu et al., “Research on image inpainting algorithm of improved GAN based on two-discriminations networks,” *Applied Intelligence*, vol. 51, no. 6, pp. 3460–3474, 2021.

[45] Y. Chen, L. Liu, J. Tao et al., “+e improved image inpainting algorithm via encoder and similarity constraint,” *Ee Visual Computer*, vol. 37, no. 7, pp. 1691–1705, 2020.