Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 27 (2024), No. 4, pp. 1259–1270
DOI : 10.47974/JDMSC-1980

A derived mathematical MAC-generative network for restoration and detection of images

Divya Sharma †
Shilpa Sharma *
Vaibhav Bhatnagar §
Department of Computer Applications
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

Harshal Patil ‡
Department of Computer Science and Engineering
Balaji Institute of Technology & Management (BITM)
Sri Balaji University
Pune
Maharashtra
India

Abstract

Detecting and recognizing the license plates in hazy or foggy weather conditions is a challenging task. The paper derived a mathematical MAC-Generative Network (MAC-GN) using Unpaired Multi-head Attention mechanism based on CycleGAN. Unlike the basic generator architecture of cycleGAN, the derived MAC-GN adds a multi-head attention mechanism to the encoder and the decoder for generating new sample images. CycleGAN uses a cycle-consistency loss to enforce the mapping between the noisy and clear license plate images. The derived MAC-GN used patch GAN as discriminator system with the ability to generate high-quality, high-resolution images with fine details and textures. The MAC-GN is evaluated using three image quality assessment metrics: Peak signal-to-Noise Ratio, Structural Similarity Index Measure and Perception based Image Quality Evaluator. The experimental result shows that the derived MAC-GN generates a restored, haze free

† *E-mail:* **divya.sharma2414@gmail.com**
* *E-mail:* **shilpa.sharma@jaipur.manipal.edu (Corresponding Author)**
§ *E-mail:* **vaibhav.bhatnagar15@gmail.com**
‡ *E-mail:* **pharshal2288@gmail.com**

license plate images with a natural image view. The results have shown an improvement of 13.1% on PSNR and 10.12% on SSIM.

Subject Classification: *Primary 68T07, Secondary 68M25.*

Keywords: *License plate recognition, Cycle GAN, Generator, Discriminator, Attention layer, U-Net, Patch GAN.*

1. Introduction

Leading-edge technologies and advancements are widely being used to transform developing nation into developed nation. With today's growing population and technical advancements, the task of Automatic License Plate detection and Recognition (ALPR) plays an important role in overseeing transportation safety and security [1]. ALPR helps in identifying high-speed vehicles, recognizing lost vehicles, and identifying vehicles involved in any crime activities [2]. The license plate images captured in the clear weather can be detected efficiently with the existing techniques. However, the license plate images captured in hazy or foggy weather result in a whitening and cloudy effect in the image. The atmospheric disturbance occurred due to the presence of complex weather conditions like haze, fog, rain, or snow, hides the important features and details of the affected images. Therefore, vehicle license plate recognition under such complex weather conditions becomes extremely important to deal with [3]. Some of the implemented image dehazing methods are multi-scale CNN for image dehazing. A dark channel prior algorithm calculates the least intensity values from the three color channels [4]; contrast assumptions via statistical findings and hypotheses [5]. Dehazing using a Generative Adversarial Network (GAN) generates fake clear images [6]. GAN proposed by Goodfellow et al. [7] in the year 2014. GAN mainly consists of two network architectures: the *generator* is the imitator who tries to generate fake data as real, while the *discriminator* acts as an investigator to classify whether the data is fake or real. The generator tries to reduce the losses and discriminator tries to increase the losses. The basic objective function for GAN is given in equation (1) [8].

$$\min_{gen}\max_{dis}(dis, gen) = E_{adata}[\log dis(a)] + E_{bdata}(1 - dis(gen(y)))] \tag{1}$$

Here, Gen(y), represents the generator outcome on noisy input images. Dis(a) denotes the probability of real data as 'a' .Dis(gen(y)) represents the probability of regenerated data gen(y) is real. E_{adata} and E_{bdata}

represents the likelihood all over real image and regenerated data. The main aim of this research work is to restore hazy and foggy vehicle license plates images. It is difficult to have a pair of clear and noisy image of the same vehicle at the same time and position. To eliminate the requirement of paired images, the derived network used the cycleGAN architecture. The mathematical MAC-GN is an enhanced cycleGAN architecture consisting of a generator-discriminator to generate clear, restored images. The network improves the traditional cycleGAN image restoration ability by combining the multi-head attention mechanism within the generator architecture. Then the restored image goes to the discriminator to analyze whether the fake regenerated image is similar to real or not. The paper is organized as follows: Section 2 presents the derived mathematical MAC-GN. The results are discussed in Section 3. The research paper concludes with future scope in Section 4.

2. Proposed Methodology

The mathematical derived MAC-GN follows the cycleGAN based generator-discriminator architecture. Figure 1 show the CycleGAN image restoration process. Here, GEN_{AB} is the generator generating dehazed images from input hazy images and GEN_{BA} regenerate hazy images. The discriminator D_B identifies real or fake images. The MAC-GN has modified the traditional basic U-Net Generator architecture. The derived networks generator has incorporated multi-head attention in the encoder and decoder regions in the U-Net architecture [9]. This improves the message passing in subsequent layers along with encoder-decoder skip connections. The modified generator architecture flow diagram is shown in Figure 2.

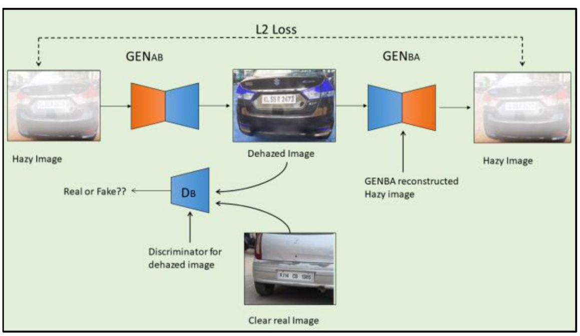

Figure 1

CycleGAN image restoration framework

This section discusses the derived MAC-GN generator architectural details followed by the discriminator.

2.1 *Mathematical MAC-GN Generator*

The MAC-GN generator architecture used Attention based U-Net architecture to generate new sample images. The encoder network of the generator contains 6 identical layers. Individual layers consist of self-attention and fully connected positioned feed-forward layers as sub layers. The sub layers are incorporated with residual layers and the batch normalization. The decoder also contains 6 layers with 3 sub-layers. The result of the convolution operation is a new sequence, which describes how the shape of one sequence is modified by the other. The subsequent stage involves the utilization of a max pooling layer, which partitions the input tensor into non-overlapping rectangular regions and chooses the maximum value from each region as the respective output value. It normalizes the activations of each neuron across a batch of data, which stabilizes and speeds up the training process.

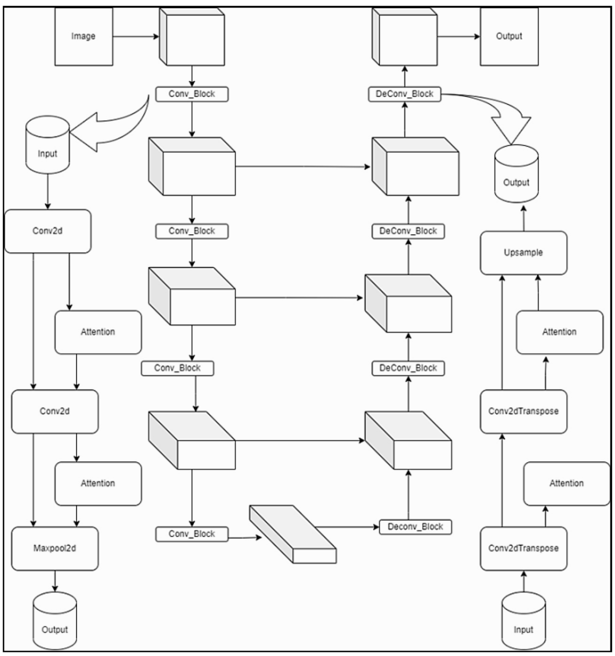

Figure 2

Derived MAC-GN Generator flow diagram

2.1.1 Mathematical Attention Mechanism

The derived MAC-GN generator architecture utilized the mathematical attention mechanism to perform mapping between the query data and a set of key-value pairs. The keys, values and queries are vectors. The generated output is computed as the values weighted addition. The weighted sum of value is defined by calculating the suitable function for the query with its associated key pair. It is constantly calculated as a set of queries and clubbed as a Q metrics [10]. The mathematical relations of attention mechanism to compute the output metrics for image is represented in equation (6).

$$F = \text{Softmax}\left(\frac{1}{\sqrt{d_k}} QK^T\right)V \tag{6}$$

The MAC-GN performs multi-head attention mechanism by projecting the query, key and value multiple times by defining diverse linear projection dimensions.

$$Attentation(A, B, C) = Concat(h1 \ldots, h_x)W^o \tag{7}$$

The concat() concatenates the outputs of each head and the W^O metrics is used to transform the concatenated output to the desired output dimension.

After encoding higher dimensional images to a lower dimension dense space the images are decoded through transpose convolution using equation (8).

$$Z_{i,j,k} = \sum_{u=0}^{K-1} \sum_{v=0}^{k-1} \sum_{l=1}^{L} w_{u,v,l,k} x(i+u-1), (j+v-1), l \tag{8}$$

Here, the paper provides a probabilistic approach to define the derived network. Let x ∈ R d is the input feature vector and v ∈ R b be the query vector [11]. The attention mechanism computes the attention weights a ∈ R n using mathematical equation (9)

$$w_i = \frac{\exp(score(v, qi))}{\sum_{j=1}^{n} \exp(ecore(v, qi))} \tag{9}$$

Here, score(v, qi) is a score function that computes the similarity between the query vector v and the input feature vector qi shown in equation (10). The derived algorithm for the MAC-GN is shown in algorithm 1.

$$Score(v, qi) = v^T qi \tag{10}$$

Algorithm 1: Derived Mathematical MAC-GN Generator Algorithm

1. Input: Input image X of size H ×W ×C, total classes K, number of filters F, dropout rate p, and number of attention blocks N
2. Output: Segmented images of size H × W × K

ENCODER

3. for i = 1 to N do
4. Convolution layer with F × 3 × 3 filters, stride 1, and padding 1
5. Batch normalization and ReLU activation
6. Convolution layer with F × 3 × 3 filters, stride 1, and padding 1
7. Batch normalization and ReLU activation
8. Max pooling with 2 × 2 kernel and stride 2
9. Save feature maps Ei for each attention block

end for

10. Convolution layer with F × 3 × 3 filters, stride 1, and padding 1
11. Batch normalization and ReLU activation
12. Convolution layer with F × 3 × 3 filters, stride 1, and padding 1
13. Batch normalization and ReLU activation
14. Dropout with rate p and Save feature maps EN+1

DECODER:

15. for i = N to 1 do
16. Up-sampling with factor 2 and nearest neighbor interpolation
17. Concatenate with corresponding feature maps from encoder Ei
18. Convolution layer with F × 3 × 3 filters, stride 1, and padding 1
19. Batch normalization and ReLU activation
20. Convolution layer with F × 3 × 3 filters, stride 1, and padding 1
21. Batch normalization and ReLU activation
22. Attention block using feature maps Ei and up-sampled feature maps

end for

23. Up-sampling with factor 2 and nearest neighbor interpolation
24. Concatenate with feature maps E1
25. Convolution layer with F × 3 × 3 filters, stride 1, and padding 1
26. Batch normalization and ReLU activation
27. Convolution layer with K × 1 × 1 filters, stride 1, and padding 0
28. Softmax activation
29. Return: Segmented image

3. MAC-GN Discriminator

PatchGAN comes under the category of Generative Adversarial Network (GAN) frequently used for tasks related to translating images, image segmentation, super-resolution, and style transfer. [12]. Unlike the traditional GAN model, PatchGAN operates on small patches of the image, allowing for more localized and fine-grained processing. The key advantage of using PatchGAN is its ability to generate high-quality, high resolution images with fine details and textures. Let x be an input image patch of size n × n × c, and let y be the output scalar. $x \to y$.

The output scalar y is given using equation (11)

$$y = \frac{1}{N} \sum_{i=1}^{N} D_i(xi) \tag{11}$$

Here, D_i is the output of the i-th discriminator block, which is a convolutional layer followed by a LeakyReLU activation function.

4. Loss Functions

Loss functions generate a comparison between the target images and the regenerated output images. Adversarial loss is mathematically defined to compare the generated restored images to the target clear haze free and fog free images. Cyclic consistency loss is used to prevent the contradicting mapping from domain A to B. The mathematical function for generator losses are provided by the following equation (12)

$$L_{Gen}(A,B,E_Y,E_X) = \lambda_{cycle} L_{cyc}(A,B) + \lambda_{adv} L_{adv}(A,E_y,Y,Z) + \lambda_{adv} L_{adv}(B,E_X,Z,Y) + \lambda_{iden} L_{iden}\ (A,B) \tag{12}$$

Here, L_{cyc} (A, B) is the cycle-consistency loss calculated as equation (13)[14].

$$L_{cyc}(A,B) = E_{x\sim pdata}(x)[\ ||B(A(x)) - x\,||\,1] + F_{y\sim pdata}(y)[||\,A(B(y)) - y\,||\,1] \tag{13}$$

$L_{adv}(A,C_y,Y,Z)$ is the adversarial loss for generator A and discriminator CY represented as (14)

$$L_{adv}(A,C_y,Y,Z) = F_{y\sim pdata}(y)[(C_Y(y)-1)^2] + F_{x\sim pdata}(x)[C_Y(A(x))^2] \tag{14}$$

$L_{iden}(A,B)$ is the identity loss represented as (15)

$$L_{iden}(A,B) = F_{y\sim pdata}(y)[||\,(A(y)) - y\,||\,1] + F_{x\sim pdata}(x)[||\,(B(x)) - x\,||\,1] \tag{15}$$

5. Results

To evaluate the accuracy and performance, the derived mathematical MAC-GN is compared with different state of art image restoration and image dehazing algorithms. The MAC-GN is trained using the manually collected hazy and foggy vehicle images. The MAC-GN is implemented and tested on three datasets: OHAZE [16], I-Haze [17] and Foggy-Hazy License Plate Images dataset [18]. This section discusses the qualitative and quantitative results of the derived network. The MAC-GN performs quantitative evaluation using three different image quality assessment metrics. First is the Peak-Signal-to Noise-Ratio i.e. PSNR; another is the Structural Similarity Index Measure (SSIM) and last is the PIQE that is perception-based image quality evaluation. Higher PSNR value and higher SSIM value indicate the better restored and generated image quality [19]. Lower PIQE indicates better image quality. Figure 3 showcases the randomly selected three sample images from the manually collected hazy and foggy vehicle images. The Ground truth images are the clear haze free images and the second column shows the dense hazy and foggy vehicle images. Column three gives the clear restored images generated by the MAC-GN and finally the detected and extracted license plates for character recognition are shown in column four. So, after restoring the whole image, the license plates from the generated images are detected and cropped using YOLO v5 model. The cropped license plates are clear enough to read the license plates characters. Table 1 shows the detailed quantitative comparison of the derived network with three different state-of-the-art dehazing algorithms.

Table 1

Qualitative results on foggy hazy license plate images

Methods	DCP	GAN	CycleGAN	MAC-GN
SSIM	0.6731	0.6166	0.7153	0.7915
PSNR	13.85	12.74	15.64	17.83
PIQE	40.76	38.27	35.49	30.32

While comparing the regenerated images with the ground truth images, the DCP algorithm generates more dark restored images. The average calculated results of PSNR, SSIM and PIQE of DCP algorithm states more color deviation between the ground truth and restored images. GAN performs the task of image restoration without prior knowledge of the dark channel, thus obtaining less PIQE value than the DCP algorithm.

The PSNR and SSIM value of GAN is less than DCP, which states that it is not very efficient in removing the hazy or foggy effect from the images. CycleGAN obtained better PSNR and SSIM in comparison to DCP and GAN and also the PIQE is less than the two. From the four discussed restoration algorithms the mathematical MAC-GN obtained the highest PSNR and SSIM and the least PIQE value. The highest value of PSNR and SSIM indicates that the derived mathematical MAC-GN performs a better image restoration task overall. Meanwhile, the least PIQE shows the fidelity concerning local color distortion.

Figure 3

Derived MAC-GN Results

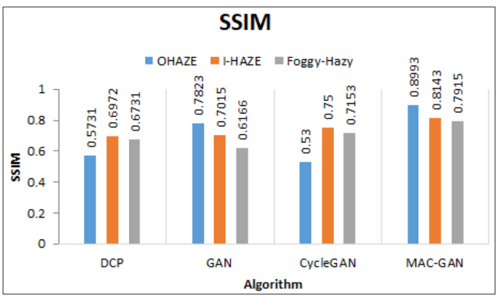

Figure 4

SSIM result from different dataset

Figure 4 shows the SSIM results of the four discussed restoration techniques on three different datasets. Figure 5 showcase the result of PSNR on three different datasets using four different techniques. Similarly Figure 6 represents the calculated PIQE values of four techniques implemented on three different datasets.

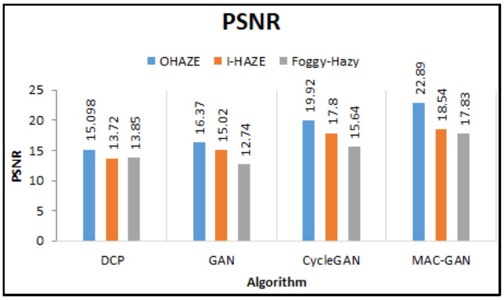

Figure 5

PSNR Result for different datasets

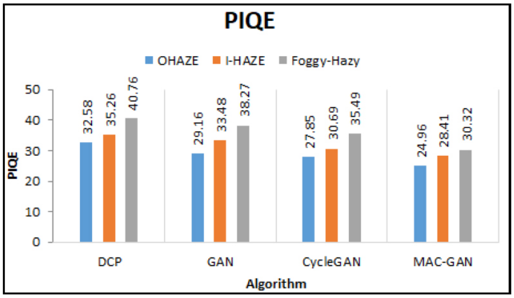

Figure 6

PIQE Result for different datasets

6. Conclusion

The derived mathematical MAC-GN for license plate image restoration and detection resolves the issue of image color loss or essential detail loss in the regenerated images. It introduced the concept of an intra-layer multi-head attention mechanism with CycleGAN.

The mathematics-based attention mechanism helps the network to achieve more robust and accurate feature extraction and image restoration results. The restored images give better texture details and natural view for the output images. The restored images are inputted to the YOLO v5 model for detecting the license plate for character recognition task. The MAC-GN is evaluated using three image quality assessment metrics that are PSNR, SSIM and PIQE. The derived network has obtained 17.83 PSNR, an improvement of 13.1% from other implemented techniques. SSIM achieved by the MAC-GN is 0.7915, an enhancement of 10.12%. The derived network achieves better results than the available state-of-the-art image restoration techniques. The future work will focus to reduce the image processing time the detection of license plate will be implemented first, followed by the image restoration for defogging and dehazing of the detected license plate.

References

[1] Agarwal, Parul, et al. “Implementing ALPR for detection of traffic violations: a step towards sustainability.” Procedia Computer Science 132 : 738-743 (2018).

[2] kumar, Ashutosh, et al. “Deep learning based highway vehicles detection and counting system using computer vision” *Journal of Information and Optimization Sciences*, 44:5, 997–1008

[3] Silva, Sergio Montazzolli, and Claudio Rosito Jung. ”License plate detection and recognition in unconstrained scenarios.” Proceedings of the European conference on computer vision (ECCV) (2018).

[4] He, Kaiming, Jian Sun, and Xiaoou Tang. “Single image haze removal using dark channel prior.” *IEEE transactions on pattern analysis and machine intelligence* 33.12 : 2341-2353 (2010).

[5] Tan, Robby T. “Visibility in bad weather from a single image.” 2008 IEEE conference on computer vision and pattern recognition. IEEE, (2008).

[6] Liu, Zheng, et al. “Single image dehazing with a generic model-agnostic convolutional neural network.” IEEE Signal Processing Letters 26.6 : 833-837 (2019).

[7] Goodfellow, Ian, et al. “Generative adversarial networks.” Communications of the ACM 63.11 : 139-144 (2020).

[8] Sachar, Silky, and Anuj Kumar. “DCGAN-based deep learning approach for medicinal leaf identification.”

[9] Ashish, Vaswani. “Attention is all you need.” arXiv preprint arXiv: 1706.03762 (2017).

[10] Liang, Meiyan, et al. “Multi-scale self-attention generative adversarial network for pathology image restoration.” The Visual Computer 39.9 : 4305-4321 (2023).

[11] Kim, Hwajoon. “The definition of convolution in deep learning by using matrix.” *Journal of Engineering and Applied Sciences* 14.7 : 2272-2275 (2019).

[12] Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." arXiv preprint arXiv:1409.0473 (2014).

[13] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition (2017).

[15] Pan, Yue, et al. "FDPPGAN: remote sensing image fusion based on deep perceptual patchGAN." Neural Computing and Applications 33 : 9589-9605 (2021).

[16] O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images [online] (accessed 1 October 2023). https://data.vision.ee.ethz.ch/cvl/ntire18//o-haze/

[17] I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images [online] (accessed 1 October 2023). https://data.vision.ee.ethz.ch/cvl/ntire18//i-haze/

[18] Sharma, Divya; Sharma, Shilpa; Bhatnagar, Vaibhav, "Foggy-Hazy License Plates Images", Mendeley Data, V1 (2023), doi: 10.17632/p3jr4555tf.1

[19] Saxena, Sameer, et al. "Comparative analysis between different edge detection techniques on mammogram images using PSNR and MSE." *Journal of Information and Optimization Sciences* 43.2 : 347-356 (2022).