IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 June 2022, accepted 2 July 2022, date of publication 6 July 2022, date of current version 12 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3188860

RESEARCH ARTICLE

FIBS-Unet: Feature Integration and Block Smoothing Network for Single Image Dehazing

MING-HWA SHEU[1], (Member, IEEE), S. M. SALAHUDDIN MORSALIN[1], (Graduate Student Member, IEEE), SZU-HONG WANG[1], (Member, IEEE), YU-TENG SHEN[1], SHIH-CHANG HSIA[1], (Member, IEEE), AND CHUAN-YU CHANG[2], (Senior Member, IEEE)

[1]Department of Electronic Engineering, National Yunlin University of Science and Technology, Douliu, Yunlin 64002, Taiwan
[2]Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, Douliu, Yunlin 64002, Taiwan

Corresponding author: S. M. Salahuddin Morsalin (s.morsalin10@gmail.com)

This work was supported in part by the Ministry of Science and Technology, Taiwan, under Project 110-2221-E224-052; and in part by the National Yunlin University of Science and Technology, Taiwan.

ABSTRACT The dehazing algorithms are based on the hazy simulation equation to remove haze and restore the input image feature maps by estimating the intensity coefficient of the atmospheric light source and the scattering coefficient of the atmosphere. However, the coefficient prediction isn't good, resulting in artifact noise in the dehazed output image. The increasing expansion of deep learning algorithms in computer vision applications to combat noise and interference in the hazy picture is growing. This paper proposed an efficient framework for Feature Integration and Block Smoothing (FIBS-Unet) Unet architecture using encoder-decoder processing with intensity attention block. We modified the Res2Net residual block with customized convolution and added instance normalization to improve the encoder feature extraction efficiency. Besides, we designed the Intensity Attention Block (IAB) using Sub-Pixel Layer and convolution (1×1) to amplify input feature and fusion feature maps. We developed an efficient decoder employing sub-pixel convolutions, concatenations, contrive convolutions, and multipliers to recover smooth and high-quality feature maps at the framework. The proposed FIBS-Unet has minimized the Mean Absolute Error (MAE) at perceptual loss function with the RESIDE dataset. We calculated the Peak Signal-to-Noise Ratio (PSNR), the Similarity Index Measure (SSIM), and a subjective visual color difference to evaluate the model's effectiveness. The proposed FIBS-Unet achieved better quality dehazing image results of PSNR:34.122 and SSIM:0.9890 in the outdoor scenarios at dense haze and backlight image for the Synthetic Objective Testing Set (SOTS). Our extensive experimental results specify that proposed FIBS-Unet is extendable to real-time applications.

INDEX TERMS Image dehazing, feature extraction, feature fusion, intensity attention block, instance normalization, residual connection, sub-pixel convolution.

I. INTRODUCTION

Rapid industrial expansion, as well as population growth, have a significant environmental impact. As a result, traffic exhaust and air pollution have become serious issues worldwide in recent years. The haze, fog, smoke, and mist all are atmospheric particles that degrade the image's visual quality [1]. Lower eyesight, hazy, and foggy weather conditions may have various effects on our everyday life. For example, driving a car needs wide-field views, regular weather change has a significant impact on the sustainable city, navigation control, and the haze and foggy weather can hamper military security. However, Artificial Intelligence (AI) increases computer vision visibility. The

The associate editor coordinating the review of this manuscript and approving it for publication was Badri Narayan Subudhi.

deep learning approach has significant contributions to computer vision applications, including object detection and class counting [2], and semantic segmentation [3] that we uncovered our previous works. For the outdoor scenarios, the foggy and hazy occurred as a severe problem for improving visibility and usability. Outside images easily suffer from atmospheric scattering, which reduces faded colors due to the presence of the illumination media in the atmosphere. Color balance and clean edges work together to increase visibility in computer vision applications [4], resulting in dehazed pictures that appear realistic. Image dehazing process enhances and improves the visual appearance for human scenes with measuring the images features and atmospheric current structures. The proposed FIBS-Unet framework effectively performs to remove the atmospheric moisture effects and boost the visibility of outdoor image. Besides, the proposed FIBS-Unet continues the sustainable development through low-complexity, improves degraded visibility, and high-performance on image dehazing techniques in advances autonomous systems and platforms. In addition, the proposed FIBS-Unet image dehazing is an increasingly widespread approach to address the degradation of images of the environment effects, low-visibility, dust, and other phenomena. The proposed approach is extendable to traffic controlling system, military surveillance, security managements, object recognition, medical image processing, remote sensing applications and so on. Therefore, image processing and computer vision applications require haze elimination from an image with low complexity and high-performance dehazing algorithms to increase visibility.

Dehazing has several real-world applications, including designing unique vehicles, sustainable cities, surveillance, navigations, control system, and many more research fields which have been investing in image dehazing [5]. There are two types of dehazing methods: prior-based and learning-based approaches. The prior-based technique [6] expressed a physical ground for effective dark channel prior, which is acceptable for the scattering model. The learning-based solution uses supervised learning to estimate the essential brightness from the foggy input image. The physical scattering model is often used to depict picture generation in most dehazing methods. The coefficients of atmospheric light and transmission map generate the dehazed image quality. The coefficients optimization process influences the performance of the dehaze image techniques. Details are discussed on the following section.

In this paper, we define the image dehazing task as encoder-decoder feature maps for calculating the coefficients of atmospheric light and transmission task. We proposed a lightweight network, FIBS-Unet framework to enhance the image dehaze application. The proposed FIBS-Unet provides haze-free realistic pictures in terms of state-of-the-art performance. The main contributions of this work are summarized as follows:

1. To enhance the dehazing technique, we proposed the FIBS-Unet architecture by updating the encoder with M_Res2net, Instance Normalization, convolution function, and decoder modified by sub-pixel convolutions, intensity attention blocks, and customized convolution.
2. We have designed an Intensity Attention Block to amplify the input feature maps and fuse them in different stages at the decoder to calculate the atmosphere scattering coefficient and atmospheric light source coefficient. Besides, we feed the sub-pixel layers color space to robust haze removal performance.
3. The proposed FIBS-Unet architecture is a firster and well-tailored to dense haze images and backlight environmental images with high-speed performance on the GPU platform for real-time applications.

The remainder of this paper is organized as follows. In Section II, we review the related work, which provides background knowledge to understand the design of FIBS-Unet. In Section III, we present details of the proposed FIBS-Unet, and discuss how it relates to existing methods. Data analysis and Experimental results have been presented in Section IV. The conclusion has been drawn in Section V.

II. RELATAD WORK

Many solutions have been developed to address the image dehaze concerns during the last few decades. Traditional and machine learning (ML)-based approaches may be divided into two groups. Histogram-based [7], contrast-maximizing [8], [9], and saturation-based [10] structures are examples of classic approaches. Due to the use of stronger assumptions and priors, single-picture haze reduction has made tremendous progress recently. The conventional category includes items such as [6], [11], and [10]. They address the underdetermined problem by using previous information if the haze-free image has a substantially greater local contrast than the foggy one. Despite advances, state-of-the-art approaches are confined to haze-relevant priors or heuristic signals, making them ineffective for outdoor photos. On the other hand, the machine learning approach using convolutional neural networks (CNNs) for high-level vision tasks such as object identification, tracking, classification, and segmentation can swiftly recognize individual items from natural scenery without taking a long time. Machine learning-based techniques have also been used in certain related research such as [12], [13], [13], and [15]. To solve the picture dehaze problem, they used breakthroughs in traditional and deep learning methods.

The characteristics of color features are not the same for each channel due to the object's shape, background color, and substance relationship of the RGB input channel feature maps in a frame. The dust, smoke, fog, or dry particles impair the clarity of a scene in the haze, an atmospheric phenomenon. The dehaze images are visually more attractive and pleasant, whereas hazy effects decree overall color contrast and occluding photo quality in photographs [16]. The K. He *et al.* developed the Dark Channel Prior [6] technique to get a haze-free picture by predicting the depth information in image restoration technology at atmospheric

deterioration models. With the development of artificial intelligence, the RefineDNet [17] has haze noise removal capability, a unified variational model [18] obtained the transmission map and recovered the scene, the PDR-Net [19] demonstrated comparable performance on the original and synthetic images. An Artificial Image Fusion Method [20] eliminated the visual degradation caused by haze, the Task-Oriented Network [21] achieved good performance. The wavelet U-net [22] two-stage end-to-end network discrete wavelet transforms, and inverse wavelet transforms for replating down-sampling and up-sampling features.

The semi-supervised learning algorithm [23] for single image dehazing works with synthetic training datasets and real-world images. The Single image dehazing [24] for nonlinear bounding function interprets transmission and estimation techniques. An artificial multi-exposure image fusion method [25] constructed the pixel-wise weight maps for global and local exposedness fusion. In addition, an optimization algorithm [26] merged radiance and reflectance to recognize dehaze images. Besides, the pixel-wise autoencoder [27] produced haze-free images. The retinex dehazing network [28] estimates residual illumination maps and haze-free pictures. The MSRCR algorithm [29] extended the retinex algorithm balancing dynamic range, edge, color constancy of different images. The denoising of low-frequency and high-frequency pictures was the focus of the multi-scale wavelet [30] for single image dehazing. The fusion heterogeneous adversarial networks (GAN) [31], [32] produced haze-clear and textural details images. The color-transfer convolutional neural network [33] removed haze obscuration and unnecessary information regarding the coefficients. The closed shared source residual attention fusion network [34]–[36] presented to fuse high-frequency information of different level features.

The AOD-Net [37] is an end-to-end lightweight architecture that explored the picture dehazing effect through CNN. The DehazeNet [13] end-to-end architecture used a nonlinear activation function called BReLU, which combines two ReLUs that enable numerical linear activation at a bounded maximum and minimum value. The GCANet [38] architecture added separable convolutional operations. This approach enhances the correlation between pixels, effectively avoids artifact noise, and designs a multi-layered network architecture with significant improvements. The feature fusion attention network (FFA-net) [39] considered different weigh values between channels and the uneven haze division pixels. The FFA-net extracted the relationship between the weight values of each input channel and the output pixels values.

It is very difficult for human eyes to distinguish the element of haze effects due to atmospheric light scattering particles, color distortion, contrast imbalance, and noise interference in various aspects of the pictures. Considering those issues, may researchers have developed AI algorithm for image dehazing. Many of the earlier image dehazing frameworks require large memory, more computation, and complex architecture. The estimation of the global atmospheric light intensity, improvement of medium transmission map, have good performance to remove the haze effects but the limitation is real-time proceeding application. However, the predicting approaches' on dehazing performance may not always predicts enough image characteristics. To alleviate those issues for single image dehazing, we have focused on global atmospheric light intensity coefficient, medium transmission, and image feature maps restoration process. We designed FIBS-Unet, estimating medium transmission framework trainable CNN-based Feature Integration and Block Smoothing network. The FIBS-Unet takes a hazy image and generates a medium transmission map, which is utilized to recover the haze-free image using the simple pixel-wise feature integration and block smoothing operation. We have adopted the U-net [40] architecture and modified the encoder and decoder for feature integration and block smoothing.

III. THE PROPOSED NETWORK

Indoor sceneries often contain a single type of light source and fixed color variations, and the types of items that occur in the scenes are regular, with more monotonous textures and color shifts. In contrast, the sun is generally the light source for outdoor images, and the angle of lighting illumination depends on the geographical location, season, and time.

Equation (1) defines the physical model of the light source, haze, object, lens, and distance. $I(x)$ stands for the hazy input picture, $\mathcal{J}(x)$ represents the output image after feature maps restoration, $t(x)$ is the medium of transmission, α is the global atmospheric light intensity coefficient, and the x signifies the index pixel of amplifying image.

$$I(x) = \mathcal{J}(x)\, t(x) + \alpha\,(1 - t(x)) \tag{1}$$

The hazy image $I(x)$ consists of two variables: the medium transmittance $t(x)$ and the global atmospheric light intensity coefficient α. Hence, in the equation (2), $t(x)$ has a scattering coefficient β and depth-of-field $d(x)$ particles. The depth estimation method [41] was followed to obtain the $d(x)$ depth field of the training images. In addition, the haze image feature map adjusts different concentrations by calculating the α and β parameters. The $t(x)$ is simulated by calculating the β and depth value to finalize the dehazing effect on equation (4). The β for atmospheric scattering coefficient, and $d(x)$ for the distance between camera and object.

$$t(x) = e^{-\beta d(x)} \tag{2}$$

The haze removal parameters require accurate prediction of α and $t(x)$calculation for effective dehazing using a neural network. The offset constant 'b' is added in equation (4) during the determination of the parameters. Equation (5) shows the integration of a unique parameter $K(x)$ conversion and process; the dehazing framework only needs to forecast one parameter to restore the hazy image. The dehazing framework can complete the dehazing effect using end-to-end

convolutional computation.

$$\mathcal{J}(x) = \frac{1}{t(x)}(I(x) - \alpha) + \alpha \tag{3}$$

The offset constant 'b' is added during the calculation.

$$\mathcal{J}(x) = \frac{1}{t(x)}(I(x) - \alpha) + (\alpha - b) + b \tag{4}$$

$$\mathcal{J}(x) = [I(x) - 1]\left[\frac{\frac{1}{t(x)}(I(x) - \alpha) + (\alpha - b)}{I(x) - 1}\right] + b$$
$$\mathcal{J}(x) = K(x)(I(x) - 1) + b \tag{5}$$

$$where, K(x) = \frac{\frac{1}{t(x)}(I(x) - \alpha) + (\alpha - b)}{I(x) - 1}$$

In this work, we have adopted the Unet model and modified their encoder and decoder framework shown in figure 1. The autoencoder separates the down-sampling operation into five different stages to extract significant feature maps of hazy images. The encoder uses M_Res2net residual block, instance normalization, and convolutional (3 × 3) operations to extract important input features. The encoder effectively shrinks convolutional operation for the down-sampling of the input feature maps. At the end of encoder, we placed at intensity attention block to amplify the input feature maps. The up-sampling procedure is split into five phases by the decoder operation, which enlarges the length, width, and channel of the image feature maps. The decoder employs several functional operations for successful dehazing image feature restoration. The decoder includes intensity attention blocks, sub-pixel convolutions, concatenations, customizable convolutional (1 × 1) operations, multipliers, and instance normalization.

Table 1 shows the autoencoder convolutional operation characteristics analysis for the down-sampling procedure details of proposed FIBS-Unet. The auto-encoder operation splits into five individual stages with the M_Res2net residual block, the instance normalization, and the stride convolutional operation. We have employed the M_Res2net, the IN, and the stride convolutional for better feature extraction. The convolution calculation reduces the hazy image size while down-sampling operation. The down-sampling converts a high-resolution hazy image to a low-resolution image feature. Each step has different input & output feature maps and channels separately. Since the letters H, W, and C stand for height, width, and channel number, respectively. However, the height and width of the hazy image gradually reduce during the down-sampling operation that helps to focus on a larger receptive field. Besides, the number of channels gradually increases, which helps to extract more complex features from the image. At the end of the autoencoder down-sampling operation, we have placed an intensity attention block as an encoder feature code amplifier.

Table 2 shows the feature investigation of the proposed FIBS-Unet auto-encoder down-sampling operations.

TABLE 1. Proposed FIBS-Unet framework auto-encoder convolutional analysis.

Auto-encoder operation		
Stage	Block / Layer	Output Feature Size (H×W×C)
Input	Input Layer	256×256×3
1st Stage	M_Res2net	256×256×64
	Instance Normalization (IN)	256×256×64
	Stride Conv.	128×128×64
2nd Stage	M_Res2net, IN, Stride Conv.	64×64×128
3rd Stage	M_Res2net, IN, Stride Conv.	32×32×256
4th Stage	M_Res2net, IN, Stride Conv.	16×16×512
5th Stage	M_Res2net, IN, Stride Conv.	8×8×1024

The encoder distinguishes the foreground and background features in detail at the building texture, sky, and sun halo features. Each stage of the encoder discredits those different features separately. The foreground characteristics in the (a) section imply building texture information, but their specification is unclear due to fog obstructions in front of the building structure. Section (b) and (c) focus on the image's background due to the sun's halo gradation, the sky, and the textural details of the building not being clear. The view hasn't effectively differentiated each stage features of the photos. The sun produces an extreme light source at a specific location, which makes it unable to distinguish at this stage. However, the foreground and background features of the 4th and 5th are smaller than other states that can't show in detail in this table.

Table 3 demonstrates auto decoder convolutional operation characteristics analysis for the up-sampling procedure details. The auto decoder consists of the up-sampling feature maps and concatenation followed by sub-pixel convolutions, customized convolution, intensity attention block, multiplier, and instance normalization. The decoder performs the up-sampling operation that restores discriminative features for the target resolution feature. There are many ways to recover the up-sampling feature maps for dehaze images. We used the sub-pixel convolutional calculation method to reduce the computational complexity. In addition, the performance of up-sampling used the same size feature, customized convolution (1 × 1) by adding them together through the multiplier to reduce the loss function.

The following table 4 shows the auto-encoder feature integration analysis. The auto-encoder mainly eliminates hazy effects and fuse the feature maps to enhance the dehazing technique. The features of the 7th and 8th stages are smaller than other states that can't show in detail in this table.

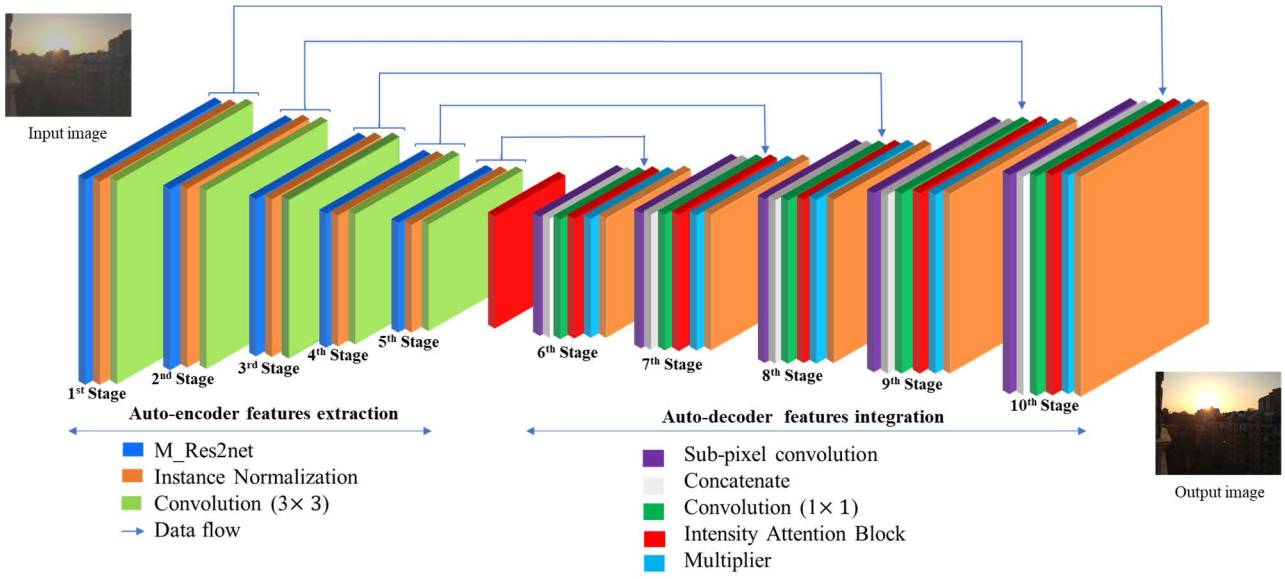


FIGURE 1. The proposed FIBS-Unet architecture with customized auto-encoder and auto-decoder operation.

The auto-decoder distinguishes the foreground and background features in the building, sky, and halo features. The IAB block amplifies the decode features and fuses by sub-pixel convolution. The concatenation, customized convolution, and multiplier enrich the dehazing process in every stage. We can see the features integration visual effects at the 8th and 9th stages. After the haze elimination at the 10^{th}stage, the foreground features expose the building texture information properly in section (d). Correspondingly, sections (e) and (f) vividly depict the sky and the sun's halo background features. This study demonstrates that this framework can deal with light pollution induced by extreme light sources. The proposed auto-encoder significantly restores foreground and background features while also completing the haze removal effect in different environments.

A. M_RES2NET RESIDUAL BLOCK

The Res2Net [42] follows the bottleneck residual block model architecture. The Res2Net residual block shows multiple scales of feature maps through this function. We have modified the Res2Net residual block using the special convolutional (1 × 1) layer, which adds the original data into the output data to avoid gradient problems. The modified M_Res2Net residual block does customize convolutional (1 × 1) operation and data concatenation instead of data addition. The granular information expresses multi-scale properties and broadens the range of receptive fields available to each layer of this network. Besides, it improves the residual block integration in the original Res2Net and makes the layer design more consistent with the overall architecture. Figure 2 illustrates the M_Res2Net residual block function. The modified M_Res2Net residual block simultaneously considers verification loss and identification loss while learning the discriminative embedding with a similar measurement at the same time.

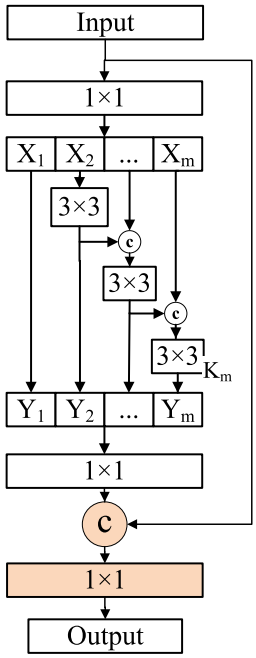


FIGURE 2. M_Res2Net block diagram details.

B. DOWN SAMPLE (STRIDE CONVOLUTION 2 × 2)

The convolution operation is computed by functional matrix calculation, which extracts information from the input pictures. An image kernel controls the convolutional operation. The Max Pooling function saves the maximum values in the feature blocks, which miss the average and

TABLE 2. Proposed FIBS-Unet framework auto-encoder feature investigation.

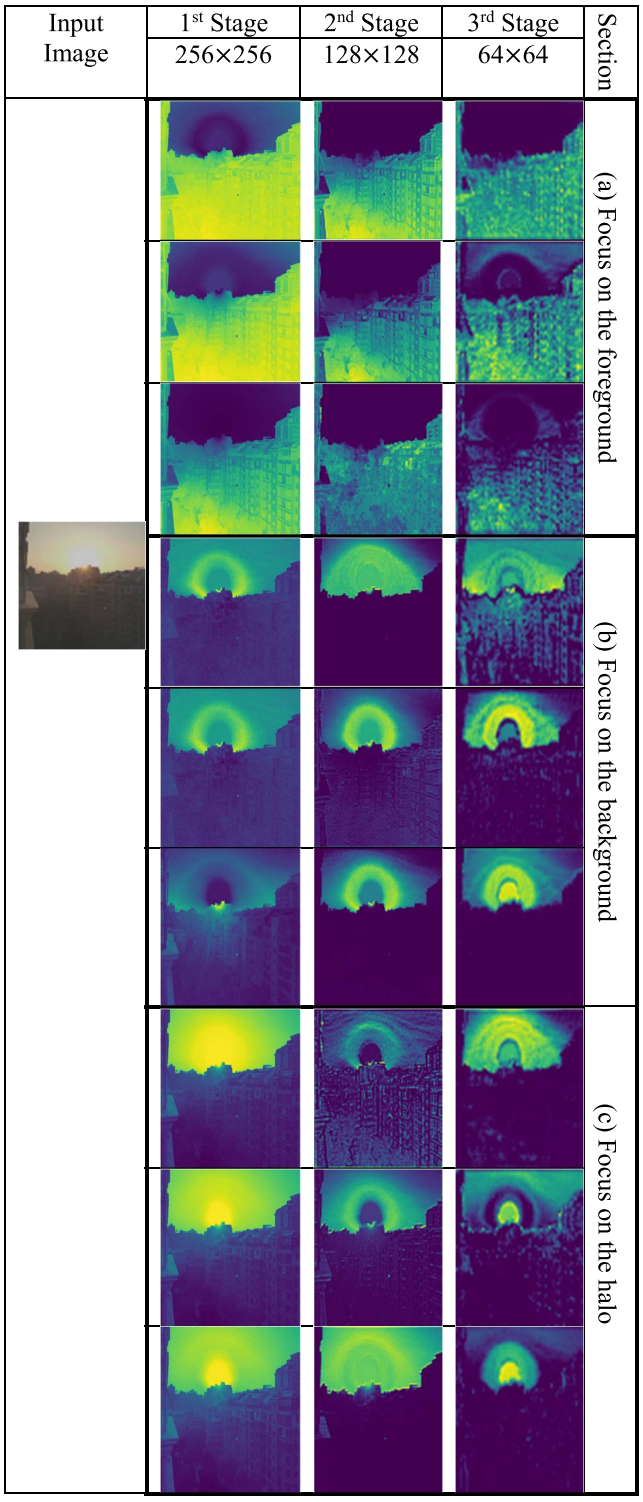


TABLE 3. Proposed FIBS-Unet framework auto-decoder convolutional analysis.

Auto-decoder operation		
Stage	Block / Layer	Output Feature Size (H×W×C)
6th Stage	Sub-Pixel Conv	16×16×256
	Concatenate	16×16×1280
	Conv (1×1)	16×16×1024
	IAB	16×16×1024
	Multiply	16×16×1024
	IN	16×16×1024
7th Stage	Sub-Pixel Conv, Concatenate, Conv (1×1), IAB, Multiply, IN	32×32×512
8th Stage	Sub-Pixel Conv, Concatenate, Conv (1×1), IAB, Multiply, IN	64×64×256
9th Stage	Sub-Pixel Conv, Concatenate, Conv (1×1), IAB, Multiply, IN	128×128×128
10th Stage	Sub-Pixel Conv, Concatenate, Conv (1×1), IAB, Multiply, IN, Conv (3×3)	256×256×3
Output	Final Image	256×256×3

lower feature maps to preserve all the spatial information. The stride convolution has several advantages over max-pooling [43] to restore accurate feature data. The stride convolution layer optimizes weights and biases to the accelerator for real-time application. For minimizing model size and increasing precision, the proposed approach intends to replace max-pooling layers with stride convolution layers utilizing the same filter size and stride operational function. Stride convolution necessitates a fewer computation process, intended to minimize feature loss during down-sampling.

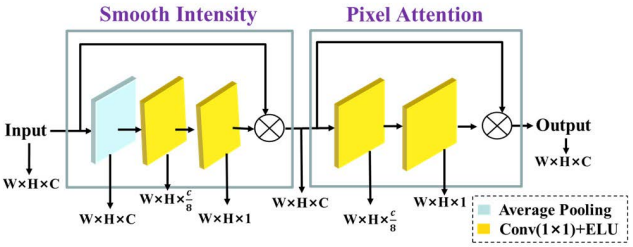
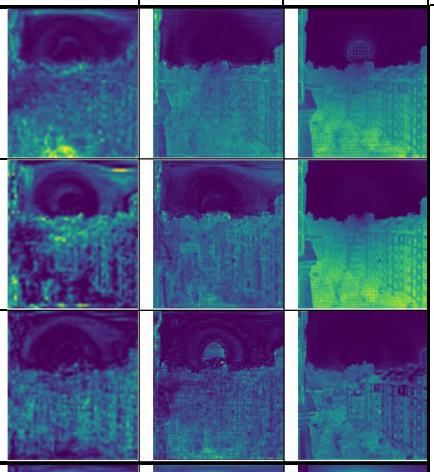
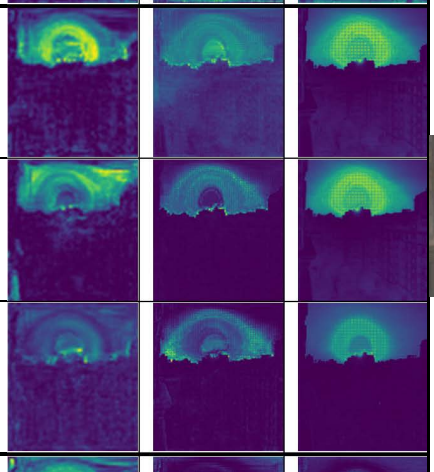
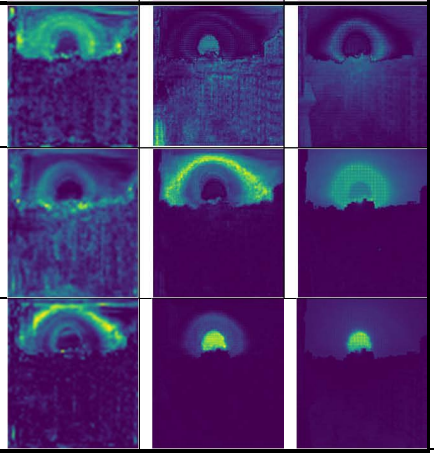


FIGURE 3. Two phase of Intensity attention Block diagram.

C. INTENSITY ATTENTION BLOCK

Figure 3 demonstrates the design of intensity attention block details. The Smooth Intensity (SI) features and Pixel Attention (PA) are two aspects of the steady operation of the intensity attention block. The average pooling function

TABLE 4. Proposed FIBS-Unet framework auto-decoder feature integration.

Sectio	Stage8	Stage9	Stage10	Output Image
	64×64	128×128	256×256	256×256
(d) Focus on the foreground				
(e) Focus on the background				
(f) Focus to the halo				

extracts initial feature maps from the average variation of the hazy images. Which progressively reduces the spatial size of the representation to reduce the network complexity and computational cost. In addition, the multiple 1 × 1 convolutional operations and Exponential Linear Unit (ELU) activation functions extract smooth intensity feature maps. The ELUs contain negative values, which drive mean unit activations closer to zero, like batch normalization with less computing cost. The pixel attention sub-block extracts input data by residual computation of two more 1 × 1 convolutional operations and ELU activation function. The residual analysis restores additional information in terms of relevant color features. We have considered the standard α value of 0.2 for equation (6).

$$\mathrm{ELU(x)} = \begin{cases} x, & x > 0 \\ \alpha(e^x - 1), & x < 0 \end{cases} \tag{6}$$

D. UP SAMPLE LAYER (SUB-PIXEL CONVOLUTION)

We designed an efficient sub-pixel convolution layer that learns an array of low-resolution feature maps into the high-resolution upscaling filters. The framework employs Sub-Pixel Convolution [44], a method of upscaling pictures linked to feature maps, to prevent the initialization value from impacting the training process. The sub-pixel convolution magnifies W and H values by r times of each, the data of W × H × (C × r^2) rearrange into the data of r × W × r × H × C. To calculate the up-sampling output pixel feature regulation, use the following equation (7).

Input: $\boldsymbol{I}^{\boldsymbol{LR}}_{\boldsymbol{W}\times\boldsymbol{H}\times(\boldsymbol{C}\times\boldsymbol{r^2})}$

Output: $\boldsymbol{I}^{\boldsymbol{LR}}_{\boldsymbol{r}\times\boldsymbol{W}\times\boldsymbol{r}\times\boldsymbol{H}\times\boldsymbol{C}}$

$$I^{SR}_{w,h,\hat{c}} = I^{LR}_{p} \tag{7}$$

whereas $p = |{}^{w}/_{r}|, |{}^{w}/_{r}|,\ \mathrm{C.mod(h,r)} + \mathrm{C.mod(w.r)} + \hat{c}$

The input data I^{LR} represents low-resolution images features, whereas the variables W and H stand for width and height of the $\mathrm{I^{LR}}$ feature map. The output data I^{SR} signifies the high-resolution images, and C is the channel numbers of I^{SR}. The variable r is the needed magnification to amplify I^{LR} to I^{SR}, while variables w, h, ĉ are the width position, height position, and channel number of current I^{SR} feature maps. All values are required to fill the I^{SR} amplification area by the I^{LR} channel supplementation. Fig. 4 shows the I^{LR} input data size of 2 × 2×4 using channel depth enhanced into 4 × 4×1 I^{SR} output data.

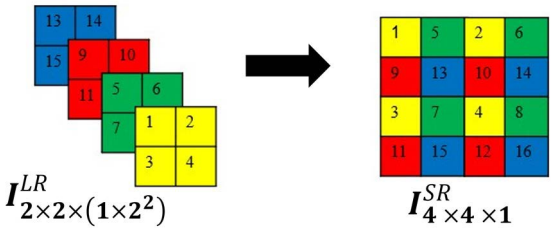


FIGURE 4. The up-sampling enhancement using sub-pixel convolution.

E. INSTANCE NORMALIZATION LAYER

The normalization of activations requires shifting and scaling the activation function via mean and standard deviation.

The Batch Normalization [45] calculates the mean and variance value across all samples and both dimensions for the learning process and significantly decreases the amount of training data to train networks. However, the Instance Normalization [46] calculates the mean and variance values for each channel for each sample and both spatial dimensions. The instance normalization feature maps are more suitable for the proposed FIBS-Unet architecture to accelerate the speed and improve the performance accuracy.

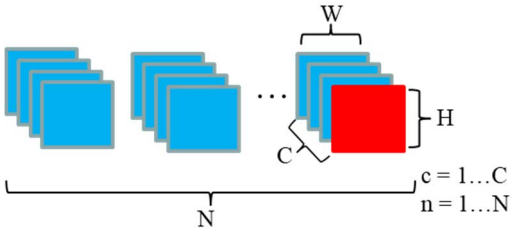


FIGURE 5. Instance normalization operational diagram.

Figure 5 shows the schematic instance normalization diagram, where $x \in \mathrm{R}^{N\times C\times W\times H}$ is the input training data. Let x be that tensor consisting of a batch of N images, for each color block represents one data, the variables c and n represent the current channel position and batch number. The read block signifies each time normalized data, the unit of instance normalization feature data is different from batch normalization. The instance normalization can accelerate training and make the network converge faster.

IV. DATA ANALISYS AND EXPERIMENTAL RESULTS

This section presents the performance and effectiveness of the proposed FIBS-Unet model, providing experimental datasets, training details, evaluation metrics, and experimental results. Several networks considered testing the dehazing outcomes and perceived color impacts of outdoor photos. The experimental verification using synthetic and backlight picture dehazing confirmed the usefulness of the proposed method's efficacy.

A. THE DATASET

For these experiments, we considered a large-scale benchmark dataset consisting of both synthetic and real-world hazy images named "*REalistic Single Image DEhazing*" (RESIDE) [47] is the haze simulation dataset with indoor and outdoor scenes. The outdoor training set of RESIDE-β consists of 2061 real-world outdoor images and 72135 haze images generated from the corresponding outdoor images with different atmospheric lights A. The global atmospheric illumination ranges from 0.8 to 1.0, and the scattering parameter β ranges from 0.04 to 0.2. The RESIDE outdoor testing dataset (500) has composed of the Synthetic Objective Testing Set (SOTS) for test performance.

B. TRAINING DETAILS

The proposed FIBS-Unet model architecture experimental results and comparison test are programed in Tensor-Flow 2.2.0 environment on Python 3.8.3 frameworks. The AMD Ryzen 7 3700x 8-Core processor with Nvidia Titan RTX 24 GB graphics card. The training procedure employs CUDA version 10.0 and cuDNN version 7.6.5. The overall number of parameters of this architecture is 18.84M, the computation volume is 21.6 G flops, and the memory usage is 6323MB. The Adaptive moment estimation (Adam optimizer) updates the weight data. The Mean Absolute Error (MAE) has been estimated for the training loss calculation. The initial learning rate set to 1e-4 values for training later gradually decreased with the number of the training period. The training period runs until the loss function value is minimized and completes the convergence.

C. EVALUATION METRICS

Many real-world applications require firster inference speed into the production environment; hence network latency time correctly is one of the most significant aspects of installing a deep network. The Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM) highest values represent the best performance of the model. The PSNR is the proportion of an image's maximal achievable value to the power of reducing noise that degrades its quality of representation. The better image feature reconstruction calculates by the following equation (8). Whereas MSE is the mean square error of input and output images, MAX_I is the peak value of image size. For example, the 8 bits image means MAX_I is 255.

$$\boldsymbol{PSNR} = \mathbf{20}\log_{10}\left(\frac{\boldsymbol{MAX_I}}{\sqrt{\boldsymbol{MSE}}}\right) \tag{8}$$

The SSIM is a perceptual metric that measures picture quality loss due to processing such as data compression or data transmission losses. The mathematical term quantifies the difference between the sample image and the reference image pixels values measuring Mean Squared Error. The metric equation mimics the visual perception system's ability to recognize structural information in a scene, allowing it to perform better on tasks that require a comparison of sample and reference picture pixels. The SSIM value is calculated by the following equation (9). Where x, y is the input and output signals, and α, β, γ are the essential parameters for adjusting the stability. The $l(x, y)$ represents the luminance comparison, $c(x, y)$ is the contrast comparison, and $s(x, y)$ is the structural comparison.

$$\boldsymbol{SSIM}\,(\boldsymbol{x},\boldsymbol{y}) = [\boldsymbol{l}\,(\boldsymbol{x},\boldsymbol{y})]^{\boldsymbol{\alpha}}\,[\boldsymbol{c}\,(\boldsymbol{x},\boldsymbol{y})]^{\boldsymbol{\beta}}\,[\boldsymbol{s}\,(\boldsymbol{x},\boldsymbol{y})]^{\boldsymbol{\gamma}} \tag{9}$$

$$(\boldsymbol{\alpha} > \mathbf{0}, \boldsymbol{\beta} > \mathbf{0}, \boldsymbol{\gamma} > \mathbf{0})$$

$$\boldsymbol{l}\,(\boldsymbol{x},\boldsymbol{y}) = \frac{\mathbf{2}\boldsymbol{\mu_x}\boldsymbol{\mu_y}+\boldsymbol{C_1}}{\boldsymbol{\mu_x^2}+\boldsymbol{\mu_y^2}+\boldsymbol{C_1}} \tag{10}$$

FIGURE 6. The dehazing performance of the proposed FIBS-Unet model at the dense-haze and backlight image.

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{11}$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \tag{12}$$

The μ_x and μ_y are the mean values, σ_x and σ_y are the standard deviations, and σ_{xy} is the covariance. Whereas the C_1, C_2, and C_3 are the stability adjustment parameter for the equations (10), (11), and (12).

D. EXPERIMENTAL RESULTS

The proposed FIBS-Unet furnishes haze-free realistic pictures in terms of state-of-the-art performance. Figure 6 demonstrates the execution results of the proposed FIBS-Unet model for the dense haze images and backlight outdoor images. Hence, column (a) represents the hazy input images, the middle column (b) displays the original scenario without the haze effects that defined as ground truth (GT) images, and lastly the right-side column (c) images demonstrate the output images which obtained by the proposed FIBS-Unet framework. We accepted the peak signal to reduces the noise ratio and the structural similarities index for experimental results evaluation.

Table 5 illustrates the performance of the proposed FIBS-Unet model and recent work on the RESIDE (SOTS) image datasets. We have compared the proposed FIBS-Unet model performance with the DCP [6], AOD-Net [37], DehazeNet [13], GCANet [38] and FFA-Net [39] state-of-the-art image dehazing methods. The DCP, AOD-Net, and DehazeNet models have achieved lower PSNR and SSIM values that are below PSNR: 30 and SSIM: 0.9000. Hence, the GCANet shows good performance on SSIM value exceeds SSIM:0.9000, but the PSNR is an average value for outdoor scenarios. However, the FFA-Net exceeds PSNR: 30 and SSIM:0.9000 values but is still lower than the proposed FIBS-Unet. In terms of PSNR and SSIM, our proposed FIBS-Unet model surpasses all five models and outperforms the state-of-the-art approaches by a wide margin. The peer approached data have been obtained from reference papers.

TABLE 5. The state-of-the-art performance at the RESIDE-Standard (SOTS) image datasets.

Model	PSNR	SSIM
DCP [6]	19.13	0.8148
AOD-Net [37]	20.29	0.8765
DehazeNet [13]	22.46	0.8514
GCANet [38]	30.23	0.9800
FFA-Net [39]	33.38	0.9840
Proposed FIBS-Unet	33.62	0.9860

Table 6 displays the dehazing quantitative performance of outdoor images for the RESIDE (SOTS) datasets. We compared the proposed FIBS-Unet against several state-of-the-art contemporary image dehazing approaches. We evaluate these algorithms using two measurable assessments: the Peak Signal to Noise Ratio (PSNR) and the Structure Similarity Index Measure (SSIM). The PSNR and SSIM score illustrates

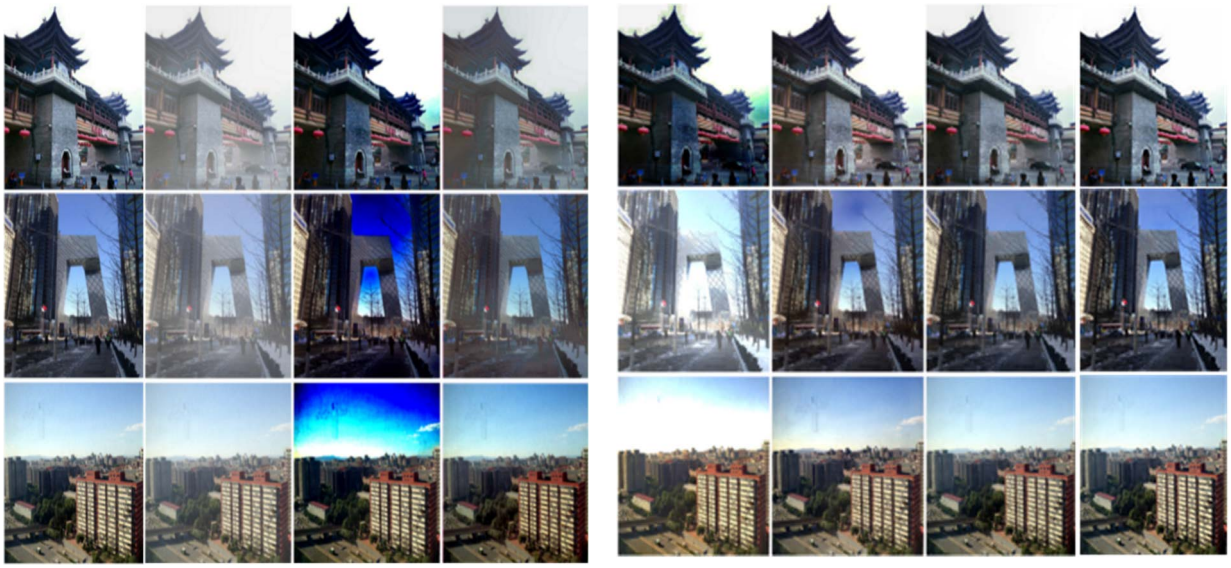


FIGURE 7. The visual color effects of outdoor image test results.

TABLE 6. Dehazing quantitative evaluations result in terms of average PSNR (dB) and SSIM.

Methods	Outdoor Image in SOTS	
	PSNR	SSIM
MSCNN [50]	19.56	0.8630
GFN [48]	21.55	0.8440
Deep Dehazing Net [32]	25.73	0.9496
Fusion Network [34]	25.35	0.7798
GMAN [51]	28.58	0.9090
pWAE [27]	28.67	0.9670
DRHNet [36]	30.23	0.9730
GridDehazeNet [49]	30.86	0.9819
Proposed FIBS-Unet	34.13	0.9890

the outdoor image testing performance between the proposed FIBS-Unet and other methods. Hence, the MSCNN [50] has the lowest PSNR: 19.56, and Fusion Network got below-average SSIM: 0.7798 values. However, the GFN [48], Deep Dehazing Net, Fusion Network, GMAN [51], and pWAE models achieved below PSNR: 30, whereas the DRHNet and GridDehazeNet [49] only acquire PSNR: 30. In terms of SSIM values, the Deep Dehazing Net, GMAN, pWAE, DRHNet, and GridDehazeNet achieved SSIM:0.90 value, but all are still lower than the proposed method. The proposed FIBS-Unet outperforms the state-of-the-art dehazing algorithms with a large margin using the RESIDE (SOTS) dataset's 500 synthetic outdoor photos.

Figure 7 demonstrates the color effects quantitative analysis for the outdoor scenes of RESIDE (SOTS) image datasets. We can observe that the DCP model endures from significant color distortion due to previous fundamental assumptions resulting in loss of the depth information of the image. The AOD-Net model is inefficient to remove haze clearly from input images in low-illumination environments. On the other hand, the DehazeNet model restores image features with higher brightness concerning reality which is not pleasant to vision. The GCANet model and FFA-Net model reinstate the high-frequency detail information. The processing capabilities are insufficient for reliably processing detailed features like textures and edges. The proposed FIBS-Unet framework has a significant improvement for the image dehazing technique. For an instant, the dense haze and backlight outdoor images. The testing effect of FIBS-Unet is increasingly better than the previous framework nearly 500 scenes in the test dataset. Finally, the overall quality of the dehazing image has improved in dense haze and backlight scenarios.

The outdoor image dehazing evaluations [50-55] considered and compared with proposed FIBS-Unet model.

TABLE 7. Average running time (Seconds) on GPU platform for SOTS image datasets.

Model	Running Time
DehazeNet [13]	0.30
FFA-Net [39]	1.10
GFN [48]	0.37
GridNet [49]	0.57
MSCNN [50]	0.26
Proposed FIBS-Unet	0.13

Table 7 compares runtimes on the SOTS outdoor image dataset. We have divided the dataset into two groups: the

training images and the validation image for test. The qualitative and quantitative evaluation for single image dehazing, it is also important to evaluate running time to show the architecture's computational complexity. The execution environment, implementation tools, and average running time for each technique to dehaze a single hazy image are described in the table above. It is observed that our method achieves comparable efficiency to other deep learning-based methods. Whereas the proposed FIBS-Unet acquired 62.5 frame per second (FPS), which is real-time proceeding speed.

V. CONCLUSION

In this paper, we have designed a novel image dehazing framework by encoder-decoder of Unet architecture modification utilizing M_Res2Net residual block, customized convolution, Instance Normalization, Intensity Attention Block, Sub-pixel Convolution, Concatenation, and multiplier. The encoder enhances input image features extraction and makes it more effective through nonlinear normalization, residual computation, and depth extraction, which has an excellent dehazing effect. The Intensity Attention Block amplifies the extracted feature maps, the Sub-pixel addition, and (1×1) convolution fuse the front-end output feature maps expanded number of channel channels. Besides, the Concatenation and multiplier helped the decoder restore the high-frequency feature maps. The Instance Normalization enhances the effect of input feature normalization in the framework and eliminates interference during the training process. The proposed FIBS-Unet achieved state-of-the-art performance for the Synthetic Objective Testing Set (SOTS) test dataset at the outdoor scene. The proposed approach has the best PSNR: 34.122 and SSIM: 0.983 values among the existing methods for dense haze and backlight pictures.

Despite the existing several dehazing algorithms, image dehazing still faces numerous obstacles. For example, it is hard to solve the color temperature at a low illuminance problem. In the future, we will focus on better performance for quality image feature restoration through different networks, reducing the number of parameters and computation, which makes it more practical for image dehazing applications in embedded systems. Besides, we are interested in the variety of image feature restoration on the extension for video dehazing tasks in a real-time dehazing framework for engineering applications.

REFERENCES

[1] Y. Wang and C. Fan, "Single image defogging by multiscale depth fusion," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4826–4837, Nov. 2014, doi: 10.1109/TIP.2014.2358076.
[2] M.-H. Sheu, S. M. S. Morsalin, J.-X. Zheng, S.-C. Hsia, C.-J. Lin, and C.-Y. Chang, "FGSC: Fuzzy guided scale choice SSD model for edge AI design on real-time vehicle detection and class counting," *Sensors*, vol. 21, no. 21, p. 7399, Nov. 2021, doi: 10.3390/s21217399.
[3] M.-H. Sheu, S. M. S. Morsalin, S.-H. Wang, L.-K. Wei, S.-C. Hsia, and C.-Y. Chang, "FHI-UNet: Faster heterogeneous images semantic segmentation design and edge AI implementation for visible and thermal images processing," *IEEE Access*, vol. 10, pp. 18596–18607, 2022, doi: 10.1109/ACCESS.2022.3151375.
[4] A. Mehra, M. Mandal, P. Narang, and V. Chamola, "ReViewNet: A fast and resource optimized network for enabling safe autonomous driving in hazy weather conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4256–4266, Jul. 2021, doi: 10.1109/TITS.2020.3013099.
[5] L. Jiao, C. Hu, L. Huo, and P. Tang, "Guided-Pix2Pix: End-to-end inference and refinement network for image dehazing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3052–3069, 2021, doi: 10.1109/JSTARS.2021.3061460.
[6] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011, doi: 10.1109/TPAMI.2010.168.
[7] T. K. Kim, J. K. Paik, and B. S. Kang, "Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering," *IEEE Trans. Consum. Electron.*, vol. 44, no. 1, pp. 82–87, Feb. 1998.
[8] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 889–896, May 2000.
[9] R. Eschbach and B. W. Kolpatzik, "Image-dependent color saturation correction in a natural scene pictorial image," U.S. Patent 5 450 217, Sep. 12, 1995.
[10] R. T. Tan, "Visibility in bad weather from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.
[11] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3522–3533, Nov. 2015.
[12] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2995–3000.
[13] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016.
[14] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *Eur. Conf. Comput. vision. Springer*, vol. 2016, pp. 154–169.
[15] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "An all-in-one network for dehazing and beyond," 2017, *arXiv:1707.06543*.
[16] D. Yang and J. Sun, "A model-driven deep dehazing approach by learning deep priors," *IEEE Access*, vol. 9, pp. 108542–108556, 2021, doi: 10.1109/ACCESS.2021.3101319.
[17] S. Zhao, L. Zhang, Y. Shen, and Y. Zhou, "RefineDNet: A weakly supervised refinement framework for single image dehazing," *IEEE Trans. Image Process.*, vol. 30, pp. 3391–3404, 2021, doi: 10.1109/TIP.2021.3060873.
[18] Y. Liu, J. Shang, L. Pan, A. Wang, and M. Wang, "A unified variational model for single image dehazing," *IEEE Access*, vol. 7, pp. 15722–15736, 2019, doi: 10.1109/ACCESS.2019.2894525.
[19] C. Li, C. Guo, J. Guo, P. Han, H. Fu, and R. Cong, "PDR-Net: Perception-inspired single image dehazing network with refinement," *IEEE Trans. Multimedia*, vol. 22, no. 3, pp. 704–716, Mar. 2020, doi: 10.1109/TMM.2019.2933334.
[20] M. Zheng, G. Qi, Z. Zhu, Y. Li, H. Wei, and Y. Liu, "Image dehazing by an artificial image fusion method based on adaptive structure decomposition," *IEEE Sensors J.*, vol. 20, no. 14, pp. 8062–8072, Jul. 2020, doi: 10.1109/JSEN.2020.2981719.
[21] R. Li, J. Pan, M. He, Z. Li, and J. Tang, "Task-oriented network for image dehazing," *IEEE Trans. Image Process.*, vol. 29, pp. 6523–6534, 2020, doi: 10.1109/TIP.2020.2991509.
[22] H.-H. Yang and Y. Fu, "Wavelet U-Net and the chromatic adaptation transform for single image dehazing," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 2736–2740, doi: 10.1109/ICIP.2019.8803391.
[23] L. Li, Y. Dong, W. Ren, J. Pan, C. Gao, N. Sang, and M.-H. Yang, "Semi-supervised image dehazing," *IEEE Trans. Image Process.*, vol. 29, pp. 2766–2779, 2020, doi: 10.1109/TIP.2019.2952690.
[24] S. C. Raikwar and S. Tapaswi, "Lower bound on transmission using non-linear bounding function in single image dehazing," *IEEE Trans. Image Process.*, vol. 29, pp. 4832–4847, 2020, doi: 10.1109/TIP.2020.2975909.
[25] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, and N. Mazur, "A novel fast single image dehazing algorithm based on artificial multiexposure image fusion," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–23, 2021, doi: 10.1109/TIM.2020.3024335.
[26] J. Shin, M. Kim, J. Paik, and S. Lee, "Radiance–reflectance combined optimization and structure-guided ℓ_0-norm for single image dehazing," *IEEE Trans. Multimedia*, vol. 22, no. 1, pp. 30–44, Jan. 2020, doi: 10.1109/TMM.2019.2922127.

[27] G. Kim, S. W. Park, and J. Kwon, "Pixel-wise Wasserstein autoencoder for highly generative dehazing," *IEEE Trans. Image Process.*, vol. 30, pp. 5452–5462, 2021, doi: 10.1109/TIP.2021.3084743.
[28] P. Li, J. Tian, Y. Tang, G. Wang, and C. Wu, "Deep retinex network for single image dehazing," *IEEE Trans. Image Process.*, vol. 30, pp. 1100–1115, 2021, doi: 10.1109/TIP.2020.3040075.
[29] J. Wang, K. Lu, J. Xue, N. He, and L. Shao, "Single image dehazing based on the physical model and MSRCR algorithm," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 9, pp. 2190–2199, Sep. 2018, doi: 10.1109/TCSVT.2017.2728822.
[30] W.-Y. Hsu and Y.-S. Chen, "Single image dehazing using wavelet-based haze-lines and denoising," *IEEE Access*, vol. 9, pp. 104547–104559, 2021, doi: 10.1109/ACCESS.2021.3099224.
[31] J. Park, D. K. Han, and H. Ko, "Fusion of heterogeneous adversarial networks for single image dehazing," *IEEE Trans. Image Process.*, vol. 29, pp. 4721–4732, 2020, doi: 10.1109/TIP.2020.2975986.
[32] Y. Li, Y. Liu, Q. Yan, and K. Zhang, "Deep dehazing network with latent ensembling architecture and adversarial learning," *IEEE Trans. Image Process.*, vol. 30, pp. 1354–1368, 2021, doi: 10.1109/TIP.2020.3044208.
[33] J.-L. Yin, Y.-C. Huang, B.-H. Chen, and S.-Z. Ye, "Color transferred convolutional neural networks for image dehazing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 11, pp. 3957–3967, Nov. 2020, doi: 10.1109/TCSVT.2019.2917315.
[34] Z. Yang, D. Pan, and P. Shi, "Joint image dehazing and super-resolution: Closed shared source residual attention fusion network," *IEEE Access*, vol. 9, pp. 105477–105492, 2021, doi: 10.1109/ACCESS.2021.3100328.
[35] M. Tan, T. Fang, Y. Fan, W. Wu, Q. She, and H. Gan, "Image-dehazing method based on the fusion coding of contours and colors," *IEEE Access*, vol. 7, pp. 147857–147871, 2019, doi: 10.1109/ACCESS.2019.2945108.
[36] C. Wang, Z. Li, J. Wu, H. Fan, G. Xiao, and H. Zhang, "Deep residual haze network for image dehazing and deraining," *IEEE Access*, vol. 8, pp. 9488–9500, 2020, doi: 10.1109/ACCESS.2020.2964271.
[37] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-Net: All-in-one dehazing network," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4780–4788.
[38] D. Chen, M. He, Q. Fan, J. Liao, L. Zhang, D. Hou, L. Yuan, and G. Hua, "Gated context aggregation network for image dehazing and deraining," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 1375–1383.
[39] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "FFA-Net: Feature fusion attention network for single image dehazing," in *Proc. AAAI*, 2020, pp. 11908–11915.
[40] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.
[41] F. Liu, C. Shen, G. Lin, and I. Reid, "Learning depth from single monocular images using deep convolutional neural fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 10, pp. 2024–2039, Oct. 2016, doi: 10.1109/TPAMI.2015.2505283.
[42] S. H. Gao, M. M. Cheng, and K. Zhao, "Res2Net: A new multi-scale backbone architecture," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 2, pp. 652–662, Feb. 2021, doi: 10.1109/TPAMI.2019.2938758.
[43] R. Ayachi, M. Afif, Y. Said, and M. Atri, "Strided convolution instead of max pooling for memory efficiency of convolutional neural networks," in *Proc. Int. Conf. Sci. Electron., Technol. Inf. Telecommun. Smart Innov., Syst. Technol.*, vol. 146. Berlin, Germany: Springer, 2020, pp. 234–243.
[44] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1874–1883, doi: 10.1109/CVPR.2016.207.
[45] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)* vol. 1, 2015, pp. 448–456.
[46] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," 2016, *arXiv:1607.08022*.
[47] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019, doi: 10.1109/TIP.2018.2867951.
[48] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang, "Gated fusion network for single image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3253–3261, doi: 10.1109/CVPR.2018.00343.
[49] X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-based multi-scale network for image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7313–7322, doi: 10.1109/ICCV.2019.00741.
[50] W. Ren, J. Pan, H. Zhang, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks with holistic edges," *Int. J. Comput. Vis.*, vol. 128, no. 1, pp. 240–259, Jan. 2020.
[51] Z. Liu, B. Xiao, M. Alrabeiah, K. Wang, and J. Chen, "Single image dehazing with a generic model-agnostic convolutional neural network," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 833–837, Jun. 2019.

MING-HWA SHEU (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the National Cheng Kung University, Taiwan, in 1989 and 1993, respectively. From 2008 to 2011, he was the Chairperson of the Department of Electronic Engineering, National Yunlin University of Science and Technology, Taiwan. From 2015 to 2018, he was a Supervisor at the Taiwan IC Design Association. He is currently a Full Professor at the Department of Electronic Engineering, National Yunlin University of Science and Technology. His research interests include CAD/VLSI, digital signal process, algorithm analysis, edge AI, and embedded systems. He has served as a Committee Chair of E.E. Course Planning for Technical High School, Ministry of Education, Taiwan. He has also served as a Review Committee of the Department of Engineering, Ministry of Science and Technology (MOST).

S. M. SALAHUDDIN MORSALIN (Graduate Student Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from Daffodil International University, Bangladesh, in 2015, and the M.Sc. degree in green technology for sustainability (major in electronics) from Nanhua University, Taiwan, in 2020. He is currently pursuing the Ph.D. degree with the National Yunlin University of Science and Technology, Taiwan. He is also working as a Research Assistant with the Department of Electronic Engineering, National Yunlin University of Science and Technology, and an Adjunct Lecturer at the Department of Computer Science and Information Engineering, Nanhua University. His research interests include image and video processing, data analytics, deep learning, artificial intelligence, and edge AI systems design.

SZU-HONG WANG (Member, IEEE) received the M.S. degree from the Department of Computer and Communication Engineering, National Kaohsiung University of Science and Technology, Taiwan, in 2005, and the Ph.D. degree from the Institute of Engineering Science and Technology, National Kaohsiung First University of Science and Technology, in 2010. He is currently an Associate Professor with Bachelor Program in Interdisciplinary Studies at the National Yunlin University of Science and Technology. His research interests include image processing, DSP/VLSI architecture design, and embedded systems.

YU-TENG SHEN received the B.Sc. degree from the Department of Electronic Engineering, National United University, Taiwan, in 2018, and the M.Sc. degree from the Department of Electronic Engineering, National Yunlin University of Science and Technology, Taiwan, in 2022. His research interests include digital signal process, image processing, object detection, semantic segmentation, deep learning, embedded systems, and their applications.

SHIH-CHANG HSIA (Member, IEEE) received the Ph.D. degree from the Department of Electrical Engineering, National Cheng Kung University, Taiwan, in 1996. From 1986 to 1989, he was an Engineer at the Research and Development Department, Microtek International Inc. From 1991 to 1998, he was an Instructor and an Associate Professor at the Department of Electronic Engineering, Chung Chou Institute of Technology. From 1998 to 2010, he was a Professor at the Department of Computer and Communication Engineering and the Department of Electronics Engineering, National Kaohsiung First University of Science and Technology Kaohsiung. He was elected as the Chairman at the Department of Electronics Engineering, in 2007. He is currently a Professor with the Department of Electronics Engineering, National Yunlin University of Science and Technology. His research interests include VLSI/SOC designs, video/image processing, HDTV/Stereo TV systems, LED lighting systems, and electrical sensors.

CHUAN-YU CHANG (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the National Cheng Kung University, Taiwan, in 2000. From 2009 to 2011, he was the Chair of the Department of Computer Science and Information Engineering. From 2011 to 2019, he was the Dean of Research and Development and the Director of the Incubation Center for Academia-Industry Collaboration and Intellectual Property. He is currently the Deputy General Director of the Service Systems Technology Center, Industrial Technology Research Institute, Taiwan. He is also a Distinguished Professor at the Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, Taiwan. His current research interests include computational intelligence and their applications to medical image processing, automated optical inspection, emotion recognition, and pattern recognition. In the above areas, he has more than 200 publications in journals and conference proceedings. He is an IET Fellow, a Life Member of IPPR, and TAAI. He served as the Program Co-Chair for TAAI 2007, CVGIP 2009, from 2010 to 2019 International Workshop on Intelligent Sensors and Smart Environments, and the third International Conference on Robot, Vision and Signal Processing. He also served as the General Co-Chair for the 2012 International Conference on Information Security and Intelligent Control, from 2011 to 2013 Workshop on Digital Life Technologies, CVGIP2017, WIC2018, ICS2018, and WIC2019. From 2015 to 2017, he was the Chair of the IEEE Signal Processing Society Tainan Chapter and the Representative for Region 10 of IEEE SPS Chapters Committee. He is currently the President of Taiwan Association for Web Intelligence Consortium.