

BDANET: BOOSTED DENSE ATTENTION HIERARCHICAL NETWORK FOR IMAGE DENOISING

Yun JIA, Lidan YU*, Jie ZHAO

*School of Information and Electronic Engineering
Shandong Technology and Business University
Yantai, 264003, China*

e-mail: jiayun0223@sdtbu.edu.cn, lidan_yu@163.com, jiezhao_99@163.com

Abstract. Deep convolutional networks have been widely applied in image denoising tasks with great success. However, many denoising models extract more feature information by increasing the network depth, which does not fully utilize the shallow features, but also makes it difficult to obtain accurate noise information. In this paper, we introduce a novel modified U-Net structure-based boosted dense attention neural network (BDANet) specifically designed for image denoising. The convolutional block within the encoding layer of BDANet incorporates dense connections and residuals, effectively circumventing the vanishing gradient issue through feature reuse and local residual learning. A boosting strategy is employed in the decoding layer to augment residual information in the noise map. To adeptly process edge details in images, BDANet deploys a polarized self-attentive mechanism to direct the densely connected blocks for depth feature extraction. The network is trained with Gaussian noise at random noise levels in the study to make it flexible to handle images with a wide range of noise levels. In experimental comparisons involving additive Gaussian noise, BDANet outperformed conventional denoising networks and attained competitive results relative to state-of-the-art image denoising networks, with an average improvement of approximately 1.03 dB in terms of PSNR values. Visualization results show that the image after denoising by BDANet network is sharper and richer in texture details than other methods.

Keywords: Image denoising, U-Net, dense connections, boosting strategy, attention mechanism

* Corresponding author

1 INTRODUCTION

Image denoising constitutes a pivotal and long-standing challenge in computer vision, primarily aimed at restoring high-quality images from their degraded, noise-laden counterparts. This process is commonly utilized in image restoration [1], image enhancement [2], and super-resolution [3], etc. As a critical preprocessing step in advanced computer vision tasks such as image recognition [4], target detection [5, 6], and image segmentation [7, 8], the quality of denoised images bears direct implications on subsequent tasks. To achieve superior denoising outcomes, a myriad of traditional algorithms have been developed in recent years. Conventional denoising methods employ image prior models based on the Bayesian perspective, including non-local means (NLM) [9], Markov random field (MRF) [10, 11], block-matching and 3D filtering (BM3D) [12], and weighted nuclear norm minimization (WNNM) [13]. While these approaches have demonstrated reasonable denoising efficacy, they often exhibit limitations due to time-consuming optimization algorithms and a reliance on manually-selected parameters, which increase model complexity.

As deep learning has advanced, researchers have explored the use of multilayer perceptrons [14] and convolutional neural networks (CNNs) [15, 16, 17] for image denoising. Learning-based denoising methods, with their enhanced modeling capabilities, network structure design, and parameter training, have surpassed prior-based models in terms of performance. For instance, DnCNN [16] implements image denoising with deeper CNN networks and surpasses previous traditional methods. Building upon this foundation, subsequent models [18, 19] have incorporated techniques such as residual learning and dense connections to increase network complexity and improve performance. Additionally, denoising models like MWCNN [20] and DHDN [21] have adopted the U-Net architecture [22] to facilitate multi-scale feature extraction. Boosting algorithm-based denoising models [23, 24, 25] offer a unique perspective on image denoising, iteratively enhancing image restoration by extracting residual signals or eliminating noise residue [26]. However, the denoising performance of boosting strategies alone remains inferior to that of deep learning-based methods. In [27], the boosting algorithm is integrated into a deep learning framework to achieve image denoising.

Despite the significant accomplishments of neural networks in image denoising, they still exhibit certain limitations. As network depth increases, the potential for gradient vanishing or exploding arises, which can lead to performance degradation and diminished information capture capabilities. Many denoising models have difficulty in obtaining accurate noise information when extracting features, and also ignore the edge information of images, resulting in over-smoothed images with lost texture details after denoising. Existing denoising networks need to be trained specifically for different noise levels and are not suitable for denoising a wide range of noise levels.

In this paper, we propose a boosted dense attention neural network based on U-Net hierarchical structure to solve the above problem. The BDANet structure design is based on the U-Net network, and the convolution of the encoding layer is

improved with densely connected noise extraction blocks to effectively avoid the gradient disappearance problem. BDANet employs a global residual learning approach to predict the residual noise of the noisy image, rather than predicting the denoised image directly. Furthermore, a boosting strategy is implemented in the decoder to extract residual noise information in a multi-scale fashion. To capture edge details more effectively, the polarized self-attention mechanism [28] is incorporated into the bottleneck layer for feature extraction, guiding the CNN in image denoising. Self-integration [29, 30] and model integration [31, 32] techniques are applied in many methods to improve the network performance, while the model integration technique requires averaging the outputs of more than two independent networks, which is too a tedious process, so we employ an efficient self-ensembling technique, generating eight output images through rotation and flipping of a single input image, with their average calculated to enhance output image quality.

The main contributions of the work in this paper are summarized as follows:

1. We propose a novel enhanced neural network for image denoising, building upon the improved U-Net model. The network is trained with random Gaussian noise, augmenting its adaptability to a wide range of noise levels. Modifications to the encoding layer contribute to superior denoising performance.
2. The boosting strategy is implemented within the decoding layer, aiming to iteratively refine the information contained in noisy images and enhance residual information extraction capabilities.
3. A polarized self-attention mechanism is incorporated in the proposed model to obtain the edge information of the image, which enhances the expressiveness of the denoising model. This leads to visually improved results compared to traditional denoising models that produce indistinct edges.
4. To prevent loss of spatial details due to consecutive downsampling during the sampling operation, strided convolution is used for downsampling instead of simple max pooling.

2 RELATED WORKS

2.1 CNN-Based Image Denoising

In recent years, neural network-based methods have been introduced into the field of image denoising and have demonstrated excellent denoising capabilities. DnCNN [16] is the first model that successfully applied CNNs to image denoising tasks. This model combines the ideas of residual learning [33] and normalization [34], but its convolution kernel size is singular, which limits the extracted image features. Zhang et al. [35] proposed the IRCNN model, which utilizes dilated convolutions to increase the receptive field and incorporates the half-quadratic splitting (HQS) algorithm. The common limitation of the aforementioned models is their inability to handle unknown noise using a single model. To tackle this challenge, the fast and

flexible denoising convolutional neural network (FFDNet) [36] is designed to train models using noise feature maps with different noise levels, enabling the handling of multi-level noise. However, the performance is not satisfactory when the noise level is unspecified. Inspired by the FFDNet network, in this study, different levels of noise are randomly added to the training dataset, enabling the trained single model to adapt to a wide range of noise.

Ronneberger et al. [22] proposed the U-Net architecture, which is composed of a top-down contraction path and a bottom-up expansion path connected by skip connections. This design effectively captures contextual information and has recently been applied to image denoising as well [37]. In DHDN [21], the convolution blocks in the contraction and expansion paths are replaced with dense residual blocks, doubling the number of feature maps at each layer of the network. Drawing on the multi-scale feature extraction ability of the U-Net network and its potential to decrease algorithm complexity, in this paper we improved the proposed denoising method by combining the advantages of the DHDN network.

2.2 Boosting Algorithm

The Boosting Algorithm is a robust approach that incrementally improves results by recursively using previous estimates as the input for the next step. It has been successfully employed in the field of image denoising, demonstrated in studies [26, 27, 38]. Various strategies have been explored to improve the denoising ability while preserving the original image information, including

1. the “twicing” technique [39], which iteratively filters the residuals to extract the remaining parts, Osher’s iterative regularization [40], and the Talebi-Milanfar strategy [25],
2. re-enhancement of the denoised image [24], and
3. iterative improvement of the signal-to-noise ratio, such as the Romano and Elad’s Strengthen-Operate-Subtract (SOS) strategy [27].

Recently, boosting strategies have been integrated with deep learning models to improve network performance. This combination is first applied to image classification, where IB-CNN [41] integrates boosting algorithms into CNNs by iteratively updating discriminative neurons starting from lower layers. Chen et al. [42] combined SOS and CNN for the denoising task, and the model used convolutional networks instead of enhancement units to form a deep enhancement framework to improve the denoising performance. DBDnet [38] extracts additional valuable information from residual images by iteratively updating the noise map via residual networks employing a boosting strategy. Xie et al. [43] propose a model-guided boosting framework, and enhance the interpretability of the framework using Regularization by Denoising (RED). Considering that the existing denoising networks rarely focus on information errors in the noise image when using residual learning to obtain the observed noise map, in this paper, the boosting algorithm is introduced

into the decoding layer to iteratively extract the residual noise map to improve the performance of the denoising network.

2.3 Attention Mechanism

For image denoising, extracting and selecting appropriate feature information is crucial. However, capturing valuable features within complex backgrounds can be challenging. Numerous recent studies have sought to enhance feature capture in denoising tasks by employing attention mechanisms [44, 45]. Tian et al. [46] proposed an attention-guided denoising convolutional neural network (ADNet) to extract noisy information hidden in complex backgrounds using a non-local self-similar attention mechanism. Li et al. [47] introduced an enhanced non-local cascading network with the attention mechanism (ENCAM) for noise removal in hyperspectral remote sensing images (HSIs).

The attention mechanism focuses on the impact of different branches within the same network, providing supplementary information to the previous stage network and guiding the feature learning of the subsequent stage. The Squeeze-and-Excitation Network (SE-Net) [48] leverages lightweight architectures to learn relationships among different channels in features, thereby enhancing the representational capacity of the network. In CBAM [49], Woo et al. devised a self-attention mechanism that combines channel and spatial attention, which yielded better results than self-attention that only considers channels. Liu et al. [28] proposed the polarized self-attention (PSA) module, which maintains high internal resolution in both the polarized channel and spatial attention branches, and showed promising results in semantic segmentation tasks. To restore the texture details of images, in this study, the PSA mechanism is incorporated into CNN to guide the model in learning pixel-level features, which resulted in denoised images with sharper edges.

3 THE PROPOSED METHOD

In this section, the architecture of BDANet is first introduced. The following subsections provide detailed descriptions of the shallow noise extraction block consisting of densely connected layers, the deep extraction guided by PSA, and the residual noise enhancement method based on boosting strategies. The loss function of this network is given at the end.

3.1 Network Architecture

The basic problem of image denoising is to recover a clean image x from a noisy image y , which can be formulated as:

$$x = y - v, \tag{1}$$

where v represents the additive noise map, typically modelled as white Gaussian noise with zero mean and standard deviation σ . Since the corrupted image contains a significant quantity of clean image structure, retaining this structure and estimating only the added noise is reasonable. Consequently, residual noise is predicted from noisy images by residual learning, which can be expressed as:

$$\hat{v} = F(y) = F(x + v), \quad (2)$$

where $F(\cdot)$ denotes the algorithm used to generate residual noise, and \hat{v} denotes the approximate value of noise v . The global residual learning method is adopted in BDANet network, assuming that the parameter mapping $R(y) \approx -v$. In addition, an approximate value of the clean image x can be calculated using the following formula:

$$\hat{x} = y + R(y), \quad (3)$$

where \hat{x} is the estimate of x .

Figure 1 shows the proposed BDANet architecture for learning $R(y)$, which improves the network structure based on the U-Net framework. Generally, a large convolution kernel is used to extract feature maps in images with high noise levels in order to capture more information, but this increases the amount of parameter computation. Therefore, our network begins by using two 3×3 convolutions to enhance the receptive field during image input.

Traditionally, U-Net adopts max pooling with a 2×2 kernel size and stride of 2 for downsampling. In this study, we use a convolution with a kernel size of 2×2 and a step size of 2 instead of pooling to prevent losing a large amount of feature information during downsampling. This is to prevent a significant loss of feature information during downsampling. To generate more feature mappings, we utilize transpose convolution instead of interpolation during upsampling. Moreover, we incorporate feature maps from the corresponding encoding layer at each upsampling step, merging shallow and in-depth information and performing 3×3 convolution to reduce the number of features while retaining critical information.

The contraction path of our proposed denoising network consists of three encoding layers, each containing two densely connected noise extraction blocks. The bottleneck layer between the contraction and expansion networks incorporates an attention mechanism to refine edge information and a decoding layer to improve the capture of residual information using the boosting algorithm. Further details of each component are described in the following subsections.

3.2 Noise Extraction Block

The proposed network replaces the regular convolution of the encoding layer in the U-Net architecture with a densely connected noise extraction block, as depicted in Figure 2. This approach is based on the DenseNet [50] model, which utilizes three

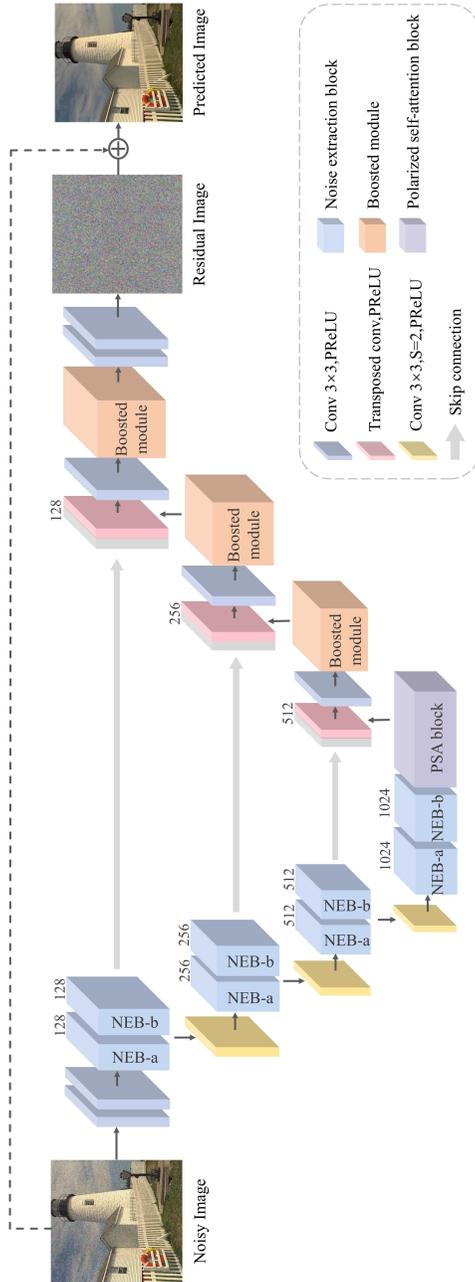


Figure 1. The architecture of the proposed BDANet network

3×3 convolutions and PreLU to map the features from the preceding layer to the subsequent layers. Each convolutional layer reduces the number of feature maps to half of the input features and utilizes the output of all preceding layers as input information. Finally, a 3×3 convolutional layer controls the output of the generated feature maps, combined with local residual learning to improve information flow and gradients. The spatial information of the noise is effectively extracted at the encoding layer with a densely connected noise extraction block, which achieves feature reuse and circumvents the gradient disappearance problem.

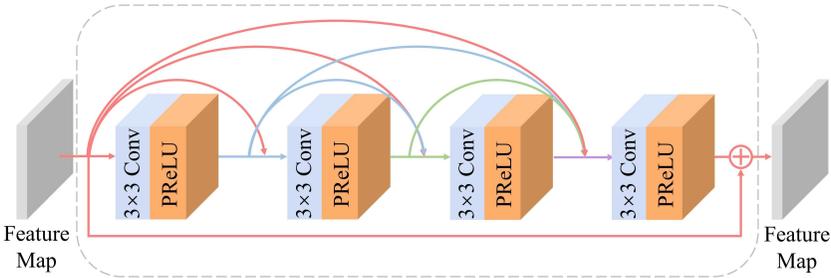


Figure 2. Architecture of noise extraction block

3.3 Boosting Perspective of Denoising

The denoising network proposed in this study performs denoising by predicting residual noise, and the quality of the predicted noise directly affects the final denoising effect. As the denoising network may mistakenly identify noise information as clean image information during the process of capturing information, the extracted noise map may contain some unrecovered noise information. Furthermore, high-frequency information such as edges and details in the clean image may be introduced into the estimated noise as noise during feature extraction. Based on the aforementioned possibilities, errors may occur in the predicted noise, leading to a discrepancy between the estimated noise and the added noise. This can be described by the following equation:

$$\hat{v} - v = x_r - v_r. \quad (4)$$

To extract the residual noise map from the noisy image, one can add unrecovered noise information to the estimated noise while removing high-frequency information, as described by the above equation, resulting in the true added noise map [27]. And using the boosting method with a recursive function, the noise map is iteratively enhanced. The process can be expressed as follows:

$$\hat{v}^n = \hat{v}^{n-1} + \hat{v}_r - \hat{x}_r, \quad (5)$$

where \hat{v}_r and \hat{x}_r are feature maps for iterations v_r and x_r , respectively, and \hat{v}^{n-1} is the noise map of the previous level. Repeated iterations make \hat{v}^n closer to the noise map v , gradually reducing the error in the predicted noise.

Drawing inspiration from the effective design of the boosting algorithm for noise progressive refinement in [38], we propose a noise enhancement module at the decoding layer in our network to gradually optimize the noise information extracted from the shallow layer, as depicted in Figure 3. To prevent mutual interference when extracting error information from the noise image, the enhancement module extracts the unrecovered information and high-frequency information separately. In the enhancement module, the unrecovered information is extracted through two 3×3 convolutions followed by PreLU activation, as the unrecovered information is a part of the noise added in the image and can be captured more easily from the image. In contrast, capturing high-frequency information that is mistakenly identified as noise in clean images requires obtaining details through dense connections at a deeper level. The structure is illustrated in the blue box on the right side of Figure 3, consisting of three densely connected 3×3 convolutions with PreLU, which enhances the capability of capturing information. Subsequently, the relationship between the noise map and the high-frequency information is measured by one 3×3 convolution and sigmoid convolution layer, and its output weight map is operated point by point with the feature values after extracting the unrecovered information, which enhances the complexity of module extraction to better capture the hidden high-frequency information from the noise map. After the experiments in [38], it is verified that its performance is best when the number of noise cancellation blocks is 6. In this paper, unlike the stepwise noise map noise cancellation blocks, the enhancement blocks are used in the decoding layer of the hierarchical denoising network to achieve multi-scale gradual denoising enhancement.

3.4 Polarized Self-Attention Block

As shown in the network architecture diagram in Figure 1, an attention block consisting of polarized self-attention mechanism (PSA) is added after the 1024th layer of the network. PSA has two advantages:

1. Filtering: it completely collapses the features in one direction while maintaining high resolution in its orthogonal direction;
2. High Dynamic Range: normalization is done by Softmax at the bottleneck tensor (the smallest feature tensor in the attention block), and then tone mapping is performed using the Sigmoid function to increase the dynamic range of attention [28].

This block connects the encoding layer and decoding layer simultaneously, enabling the network to effectively learn edge and complex texture information. Its module structure is shown in Figure 4. The blue dashed box above represents Channel-only Self-Attention, and the yellow dashed box below represents Spatial-only Self-

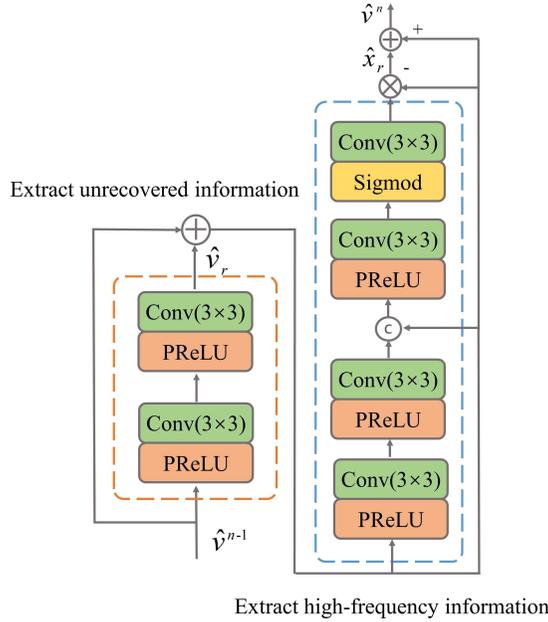


Figure 3. Architecture of boosted module

Attention, where $H \times W$ is the height and width of the image, and C is the number of channels. The model can be represented as follows:

Channel-only branch can be expressed as:

$$A^{ch}(X) = \text{Sigmoid} [LN (\sigma_1(W_\nu(X)) \times \text{Softmax}(\sigma_2(W_q(X))))]. \quad (6)$$

Spatial-only branch can be expressed as:

$$A^{sp}(X) = \text{Sigmoid} [\sigma_3 (\text{Softmax}(\sigma_1(O_{GP}(W_q(X)))) \times \sigma_2(W_\nu(X)))] , \quad (7)$$

where $A^{ch}(X) \in R^{C \times 1 \times 1}$. W_ν and W_q are 1×1 convolution, σ_1 , σ_2 and σ_3 are three tensor shaping operators, $O_{GP}(\cdot)$ is the global pooling operator, LN is the LayerNorm operator, and \times is the dot product of the matrix.

Incorporating LayerNorm in the attention mechanism can alter the data distribution, consequently affecting the subsequent convolution and hindering the adjustment of the decoding layer in the proposed network. Therefore, LayerNorm is not applied in the Channel-only branch, and the formula is as follows:

$$A^{ch}(X) = \text{Sigmoid} [\sigma_1(W_\nu(X)) \times \text{Softmax}(\sigma_2(W_q(X)))] . \quad (8)$$

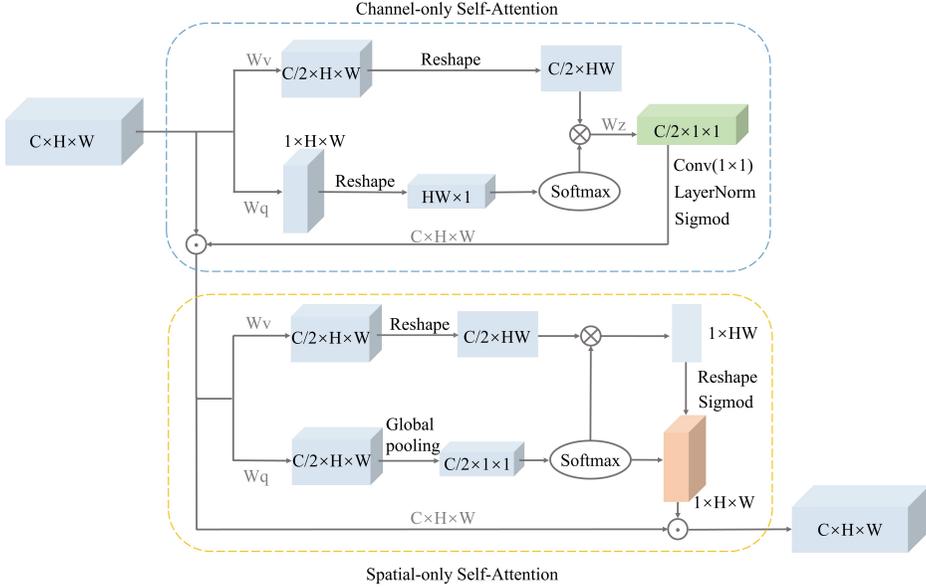


Figure 4. Architecture of Polarized self-attention mechanism

In a sequential arrangement, combine the results of the Channel-only branch and the Spatial-only branch discussed previously. The formula is as follows:

$$PSA(X) = A^{sp}(A^{ch}(X) \odot^{ch} X) \odot^{sp} A^{ch}(X) \odot^{ch} X, \tag{9}$$

where \odot^{ch} is the channel multiplication operator and \odot^{sp} is the space multiplication operator.

3.5 Loss Function

Let $f_{BDANet}(y; \theta)$ be the output of the network proposed in this study, where θ denotes the parameters of the network. The denoising model is trained using a paired dataset $\{(y_i, x_i)\}_{i=1}^N$, where y_i represents the i^{th} noisy image and x_i represents the i^{th} clear image. Through experimental research, it is found that the BDANet network obtains better results by using L1-norm training. The L1 loss function is shown as follows:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \| f_{BDANet}(y_i; \theta) - y_i - x_i \|_1, \tag{10}$$

where N is the number of noisy input images, and the Adam algorithm [51] is used to optimize the objective function stated above.

4 EXPERIMENTS

4.1 Experiment Datasets

4.1.1 Training Datasets

In recent years Agustsson and Timofte proposed the DIV2K [52] dataset, which contains 1000 high quality color images with 2 k resolution, of which 800 are used as the training dataset and 100 as the validation dataset. Initially designed for image super-resolution, the DIV2K dataset has been utilized by many image denoising algorithms [20, 21] for network training due to its sufficient number and high resolution. Therefore, we also utilize this dataset for network training and validation in our study. The training dataset is cropped to 64×64 size image blocks, which helps to improve the efficiency of the training model. To further enhance the denoising model’s robustness by increasing the diversity and complexity of the training samples, we expand the data by randomly flipping the original images horizontally and vertically and rotating the images 90 degrees counterclockwise. Additionally, we add white Gaussian noise to the training dataset with noise levels ranging from 5 to 50 to accommodate various noise levels.

4.1.2 Test Datasets

In this study, we aim to evaluate the denoising performance of the BDANet network by testing five datasets. The Set12 and BSD68 [53] datasets are utilized to validate the greyscale image denoising effect, where Set12 comprises 5 images with dimensions of 512×512 and 7 images with dimensions of 256×256 , and the BSD68 dataset includes 68 images measuring 321×481 pixels. For colour image denoising, three datasets are adopted: Kodak24, CBSD68, and McMaster [54]. The Kodak24 dataset comprises 24 images of size 768×512 , CBSD68 shares the same colour image dataset as the BSD68 scene, and the McMaster dataset comprises 18 images of size 500×500 .

4.2 Implementation Details

The initial learning rate of the network in this study is $1e-4$, which decays by half every three-period iterations over the whole dataset, stopping when it decays to $1e-6$. AdamW algorithm with parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$ is adopted to optimize the loss function. The BDANet denoising model is trained with a batch size of 16. The denoising model is trained and tested for denoising performance on PyTorch 1.10 and Python 3.8. Finally, all experiments are run on an Nvidia RTX 3090 GPU.

4.3 Performance Comparison

This section presents a comparison of the proposed model with other denoising methods, including model-based methods BM3D [12], WNNM [13], and TNRD [55], as well as CNN-based denoising methods DnCNN [16], IRCNN [35], FFDNet [36], DBDNet [38], ADNet [46], DHDN [21], SUNet [56], and MWDCNN [57]. All of the aforementioned methods are state-of-the-art denoising models. In this study, we compare the experimental results in terms of both greyscale and colour images. To further evaluate the network’s performance, we utilize PSNR [58] and SSIM [59] to quantify the denoising effect.

Methods	Set12		
	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
BM3D	32.37/0.8952	29.97/0.8504	26.72/0.7676
WNNM	32.70/0.8982	30.28/0.8557	27.05/0.7775
TNRD	32.50/0.8958	30.06/0.8512	26.81/0.7680
U-Net	32.91/0.9042	30.53/0.8649	27.39/0.7919
DnCNN	32.86/0.9031	30.44/0.8622	27.18/0.7829
IRCNN	32.77/0.9008	30.38/0.8601	27.14/0.7804
FFDNet	32.75/0.9027	30.43/0.8634	27.32/0.7804
DHDN	N/A	N/A	27.58/0.7984
ADNet	32.98/0.8916	30.58/0.8561	27.37/0.7908
DBDNet	33.03/–	30.65/–	27.46/–
MWDCNN	32.91/0.8972	30.55/0.8551	27.34/0.7882
BDANet	33.04/0.9064	30.70/0.8681	27.65/0.8022
BDANet+	33.08/0.9068	30.74/0.8688	27.70/0.8033

Table 1. Comparison of the average PSNR (dB)/SSIM performance of the different methods on the Set12 datasets for grayscale images with noise levels of $\sigma = 15, 25,$ and 50 . The best performance is indicated in red and the next best performance is highlighted in blue.

Regarding Gaussian denoising in state-of-the-art methods [16, 38, 46], three noise levels of 15, 25, and 50 determined by standard deviation σ are usually adopted to evaluate the method. In this study, the same noise level is also used for testing to ensure its fairness. As shown in Tables 1 and 2, we tested the grayscale images on the Set12 and BSD68 datasets, and BDANet achieved optimal and sub-optimal denoising performance at noise levels 15, 25, and 50, respectively, where red represents the ideal and blue represents the suboptimal. Table 1 shows that BDANet+ denotes the version with self-ensemble, and the experimental results indicate that the self-ensemble method leads to an improvement in denoising performance. In addition, we also consider an experimental comparison of U-Net networks to train denoising models for grey-scale images at different noise levels. Most of the results in the table are taken from the original literature or obtained from open code. The experimental results show that the BDANet in this paper improves 0.98 dB in PSNR

Methods	BSD68		
	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
BM3D	31.07/0.8722	28.57/0.8017	25.62/0.6869
WNNM	31.37/0.8766	28.83/0.8087	25.87/0.6982
TNRD	31.42/0.8769	28.92/0.8093	25.97/0.6994
U-Net	31.77/0.8769	29.29/0.8220	26.36/0.6994
DnCNN	31.73/0.8906	29.23/0.8278	26.23/0.7189
IRCNN	31.63/0.8881	29.15/0.8249	26.19/0.7171
FFDNet	31.63/0.8902	29.19/0.8289	26.29/0.7245
DHDN	N/A	N/A	26.44/0.7296
ADNet	31.74/0.8916	29.25/0.8294	26.29/0.7216
DBDNet	31.85/–	29.37/–	26.43/–
MWDCNN	31.77/0.8921	29.28/0.8299	26.29/0.7208
BDANet	31.81/0.8934	29.33/0.8301	26.41/0.7288
BDANet+	31.84/0.8939	29.36/0.8307	26.45/0.7300

Table 2. Comparison of the average PSNR (dB)/SSIM performance of the different methods on the BSD68 datasets for grayscale images with noise levels of $\sigma = 15, 25$, and 50. The best performance is indicated in red and the next best performance is highlighted in blue.

when the noise level is 50 compared with the conventional denoising model on the Set12 dataset. Our network improves 0.24 dB over DBDNet, which also uses the boosting strategy, and 0.34 dB compared to the latest technology, MWDCNN. The worst PSNR result of the difference is also less than 0.01 dB. BDANet+ exceeds the SSIM of each noise level in the competing networks and has better structural similarity.

The denoising performance of the BDANet model is evaluated on three widely used colour image datasets, namely CBSD68, Kodak24, and McMaster, and compared with several representative denoising models, as presented in Tables 3 and 4. The results demonstrate that BDANet outperforms other methods for different datasets and noise levels. When the noise level is 50, BDANet has better denoising effect than the U-Net based models (DHDN, SUNet) on the CBSD68 and Kodak24 datasets, with 0.21 dB and 0.3 dB improvement on the two data and compared with the latest denoising model MWDCNN, respectively. On the McMaster data set our model improves by 1.35 dB. Moreover, BDANet achieves a higher structural similarity than most competing networks when comparing the SSIM results on CBSD68 and Kodak24 datasets. Notably, BDANet+ exhibits the best performance on the McMaster dataset in terms of both peak signal-to-noise ratio and structural similarity, indicating its superior denoising capability.

To more directly reflect the denoising effect, we qualitatively analyse the denoised images on each dataset through subjective tests. Specifically, the visual graph of the enlarged area is obtained by locally zooming in on different locations of the test image, and the clearer the enlarged area is, the better the denoising

Methods	CBSD68		
	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
BM3D	33.52/0.9233	30.71/0.8719	27.38/0.7669
U-Net	33.80/0.9281	31.24/0.8820	27.95/0.7917
DnCNN	33.90/0.9291	30.44/0.8828	27.95/0.7882
IRCNN	33.86/0.9285	31.16/0.8824	27.86/0.7898
FFDNet	33.87/0.9295	31.21/0.8865	27.96/0.7920
DHDN	N/A	N/A	27.71/0.7874
ADNet	33.99/0.9330	31.31/0.8889	28.04/0.7974
BRDNet	34.10/ 0.9347	31.43/ 0.8917	28.16/0.8010
SUNet	N/A	N/A	27.85/0.7990
MWDCNN	34.18/0.9330	31.45/0.8867	28.13/0.7945
BDANet	34.19/0.9331	31.55/0.8896	28.30/0.8022
BDANet+	34.22/0.9334	31.58/0.8901	28.34/0.8034

Table 3. Comparison of the average PSNR (dB)/SSIM performance of the different methods on the CBSD68 datasets for color images with noise levels of $\sigma = 15, 25,$ and 50 . The best performance is indicated in red and the next best performance is highlighted in blue.

effect of the model. The graphs in Figures 5 and 6 show the denoising effect on grayscale images. Figure 5 demonstrates that the denoising model proposed in this study results in better clarity and contrast in image recovery. Meanwhile, the traditional model in Figure 6 causes deformation of the window contour due to excessive smoothing during denoising, while the proposed method produces smoother lines.

Methods	Kodak24		
	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
BM3D	34.28/0.9160	31.68/0.8684	28.46/0.7758
U-Net	34.51/0.9214	32.09/0.8783	29.04/0.7979
DnCNN	34.47/0.9204	32.14/0.8766	28.85/0.7915
IRCNN	34.55/0.9198	32.18/0.8741	28.91/0.7929
FFDNet	34.63/0.9211	32.11/0.8789	28.98/0.7938
DHDN	N/A	N/A	29.72/0.8170
ADNet	34.76/0.9247	32.26/0.8827	29.10/0.7994
BRDNet	34.88/0.9257	32.41/0.8862	29.22/0.8024
SUNet	N/A	N/A	29.54/0.8100
MWDCNN	34.91/0.9269	32.40/0.8862	29.26/0.8062
BDANet	35.10/0.9279	32.64/0.8889	29.50/0.8126
BDANet+	35.14/0.9284	32.69/0.8897	29.56/0.8141

Table 4. Comparison of the average PSNR (dB)/SSIM performance of the different methods on the Kodak24 datasets for color images with noise levels of $\sigma = 15, 25,$ and 50 . The best performance is indicated in red and the next best performance is highlighted in blue.

Methods	McMaster		
	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
BM3D	34.06/0.9111	31.66/0.8738	28.51/0.7945
DnCNN	33.44/0.9038	31.52/0.8691	28.62/0.7993
IRCNN	34.58/0.9196	32.18/0.8819	28.91/0.8067
FFDNet	34.66/0.9218	32.35/0.8894	29.18/0.8166
ADNet	34.93/0.9286	32.56/0.8942	29.36/0.8246
BRDNet	35.08/0.9297	32.75/0.8974	29.52/0.8280
BDANet	35.20/0.9296	32.91/0.8989	29.78/0.8369
BDANet+	34.26/0.9304	32.98/0.8998	29.86/0.8387

Table 5. Comparison of the average PSNR (dB)/SSIM performance of the different methods on the McMaster datasets for color images with noise levels of $\sigma = 15, 25,$ and 50 . The best performance is indicated in red and the next best performance is highlighted in blue.

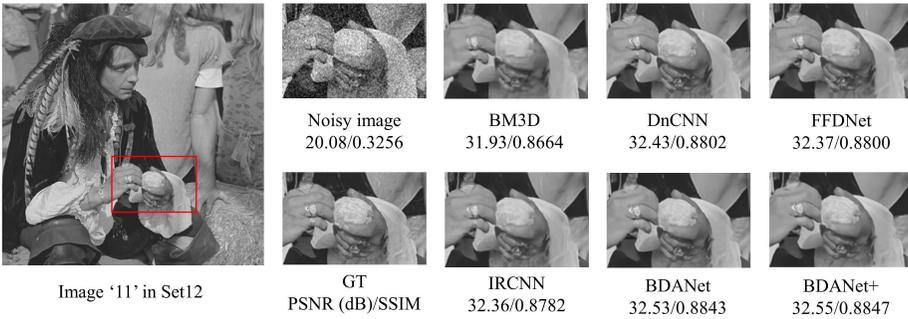


Figure 5. PSNR (dB)/SSIM values of grayscale image denoising for different methods at noise level $\sigma = 15$

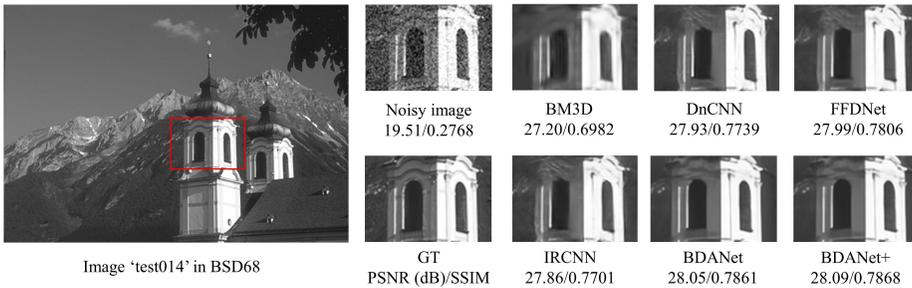


Figure 6. PSNR (dB)/SSIM values of grayscale image denoising for different methods at noise level $\sigma = 50$

Figures 7, 8 and 9 show the color image visualization results of the proposed method with the conventional model. In Figure 7, the traditional model produces a very blurry image after denoising, with the grid edge almost indistinguishable, while the proposed method better reconstructs the grid structure and preserves the edge information. Figure 8 shows that the suggested network preserves texture information while removing severe noise, whereas the existing model loses the hair texture on the squirrel’s tail after denoising. In Figure 9, the proposed network restores the texture on the wheel and recovers some background information from the grass, while the traditional denoising process smooths out the texture on the wheel and does not restore the background structure. The above qualitative analysis shows that the proposed denoising network outperforms the traditional denoising method and performs better.

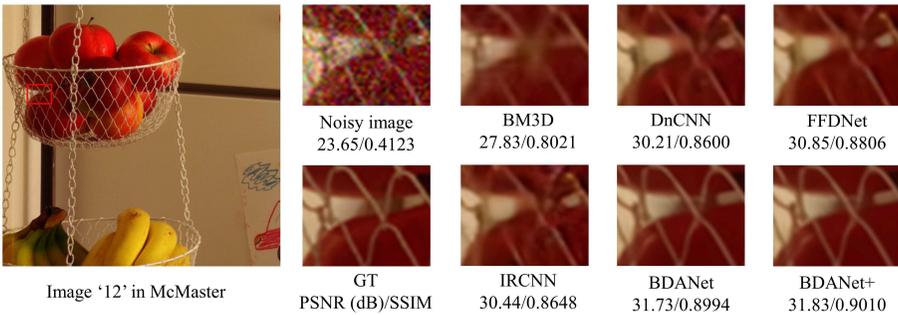


Figure 7. PSNR (dB)/SSIM values of color image denoising for different methods at noise level $\sigma = 50$

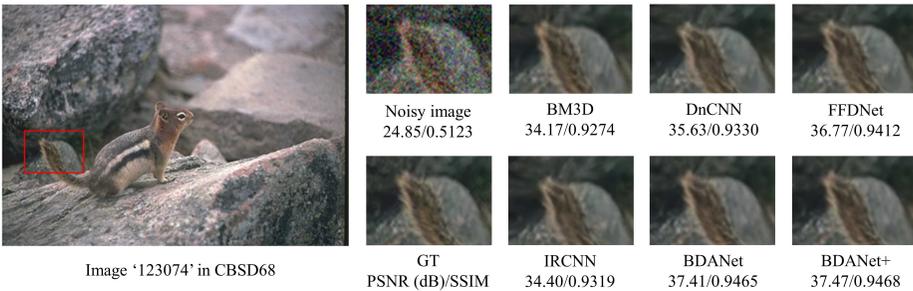


Figure 8. PSNR (dB)/SSIM values of color image denoising for different methods at noise level $\sigma = 15$

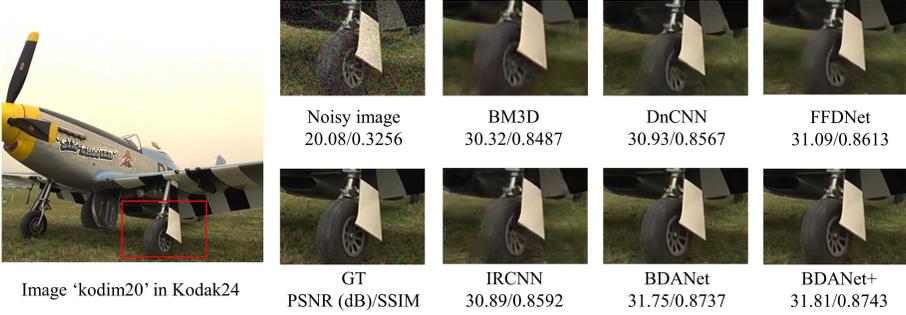


Figure 9. PSNR (dB)/SSIM values of color image denoising for different methods at noise level $\sigma = 50$

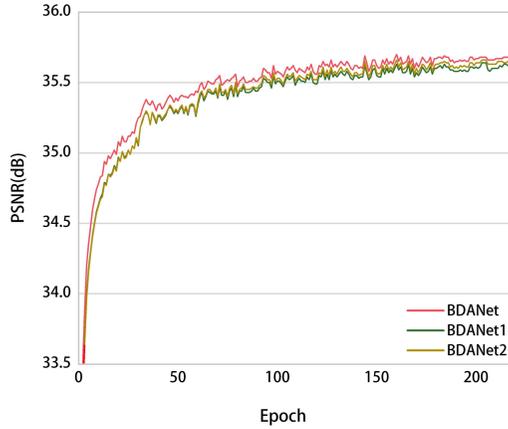
4.4 Ablation Experiments

Ablation experiments are designed in the study, as shown in Table 4, to evaluate the efficacy of the various modules in the BDANet network architecture. The table compares the effects of using the noise extraction block, the residual noise enhancement block based on boosting strategy, the polarized self-attention (PSA) mechanism, and the different downsampling methods, where BDANet $_i$ ($i = 1, 2$) represent versions of the various modules. And the basic network model U-Net and the DHDN model are evaluated together on the CBSD68 dataset for images with a noise level of 50.

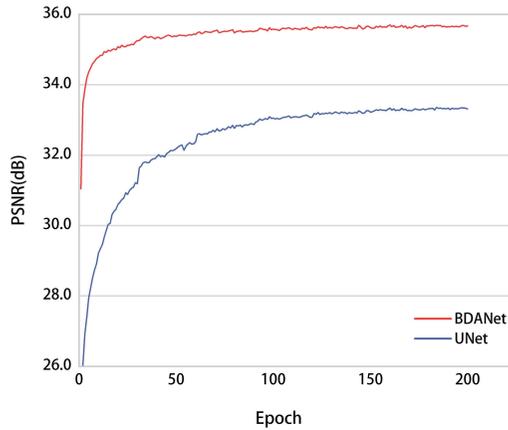
Methods	U-Net	DHDN	BDANet1	BDANet2	BDANet
MP $_{2 \times 2}$	✓	✓	✓		
Con $_{2 \times 2}$				✓	✓
Densely connected residual		✓			
Noise extraction			✓	✓	✓
Residual noise enhancement					✓
Polarized self-attention			✓	✓	✓
Parameters	29 M	168 M	160 M	163 M	152 M
MACs	15 G	64 G	46 G	48 G	51 G
PSNR (dB)	27.35	27.71	28.28	28.29	28.30

Table 6. The different blocks of the ablation study include max pooling downsampling MP, convolution downsampling, noise extraction block, residual noise cancellation block, and polarized self-attention. The comparison model averages PSNR (dB) values at noise level $\sigma = 50$ on the CBSD68 dataset.

The DHDN model uses a dense residual block in the layered network, while the BDANet1 uses a noise extraction and PSA mechanism, which improves the denoising effect and reduces the number of model parameters compared to the DHDN model by using a noise extraction block instead of a normal convolutional layer and combining the PSA mechanism. For downsampling, the max pooling downsampling in U-Net



a)



b)

Figure 10. Comparing the PSNR of various variant models as well as the base model U-Net at each training round: a) comparison of the training curves of the BDANet1 model with noise extraction block and polarized self-attention and BDANet2 model with convolutional downsampling and BDANet model with residual noise enhancement block, b) comparison of the training curves of BDANet model and U-Net model.

is replaced by stride convolution in BDANet2, although it slightly increases the number of parameters, considering that more feature information can be learned using stride convolution, while the pooling operation can only be simply mapped. The residual noise enhancement module is added to BDANet for the decoding layer to achieve multi-scale extraction of the noise map, and it can be seen in Table 4 that the addition of this module reduces the number of parameters and also significantly improves the network performance. Figure 10 a) displays the variation of the peak SNR of various variant models at each training round, where the red line represents the BDANet model, the green line BDANet1 and the yellow line BDANet2. The overall denoising effect of BDANet is better than the other two models. Figure 10 b) compares the base U-Net model and the BDANet model with the addition of noise extraction block, PSA mechanism, and noise enhancement block, demonstrating that the improved BDANet denoising performance is significantly better than that of U-Net.

5 CONCLUSION

In this paper, we propose a boosted dense attention neural network (BDANet) based on an improved U-Net structure for image denoising. To enhance the extraction of image noise features, the coding layer of the U-Net network is improved through the use of densely connected blocks, rather than conventional convolution methods. Additionally, local residual connections are employed to promote feature reuse and overcome the issue of gradient disappearance. Meanwhile iterative enhancement of the residual noise map at the decoding layer is combined with the theory of the boosting algorithm to improve the network's ability to capture information. A polarized self-attention mechanism is introduced at the bottleneck between the encoding and decoding layers, enabling the network to learn edge information and enhance the texture details of the image. The experimental results show that our method effectively improves image detail processing and denoising performance. In future work, we aim to extend BDANet to other types of images, such as near-infrared images.

REFERENCES

- [1] WANG, Z.—CUN, X.—BAO, J.—ZHOU, W.—LIU, J.—LI, H.: Uformer: A General U-Shaped Transformer for Image Restoration. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 17662–17672, doi: 10.1109/CVPR52688.2022.01716.
- [2] ZHANG, Z.—JIANG, S.—PAN, X.: CTransNet: Convolutional Neural Network Combined with Transformer for Medical Image Segmentation. Computing and Informatics, Vol. 42, 2023, No. 2, pp. 392–410, doi: 10.31577/cai.2023.2.392.
- [3] QIAO, C.—LI, D.—LIU, Y.—ZHANG, S.—LIU, K. et al.: Rationalized Deep Learning Super-Resolution Microscopy for Sustained Live Imaging of Rapid Sub-

- cellular Processes. *Nature Biotechnology*, Vol. 41, 2023, No. 3, pp. 367–377, doi: 10.1038/s41587-022-01471-3.
- [4] LI, Y.: Research and Application of Deep Learning in Image Recognition. 2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA), 2022, pp. 994–999, doi: 10.1109/ICPECA53709.2022.9718847.
- [5] ZENG, N.—WU, P.—WANG, Z.—LI, H.—LIU, W.—LIU, X.: A Small-Sized Object Detection Oriented Multi-Scale Feature Fusion Approach with Application to Defect Detection. *IEEE Transactions on Instrumentation and Measurement*, Vol. 71, 2022, pp. 1–14, doi: 10.1109/TIM.2022.3153997.
- [6] LI, S.—LIU, Y.—WU, S.—ZHANG, S. J.: MDM-YOLO: Research on Object Detection Algorithm Based on Improved YOLOv4 for Marine Organisms. *Computing and Informatics*, Vol. 42, 2023, No. 1, pp. 210–233, doi: 10.31577/cai.2023.1.210.
- [7] YU, T.—ZHU, M.: Image Enhancement Algorithm Based on Image Spatial Domain Segmentation. *Computing and Informatics*, Vol. 40, 2021, No. 6, pp. 1398–1421, doi: 10.31577/cai.2021.6.1398.
- [8] CHEN, Y.—WANG, M.—HEIDARI, A. A.—SHI, B.—HU, Z.—ZHANG, Q.—CHEN, H.—MAFARJA, M.—TURABIEH, H.: Multi-Threshold Image Segmentation Using a Multi-Strategy Shuffled Frog Leaping Algorithm. *Expert Systems with Applications*, Vol. 194, 2022, Art.No. 116511, doi: 10.1016/j.eswa.2022.116511.
- [9] BUADES, A.—COLL, B.—MOREL, J. M.: A Non-Local Algorithm for Image Denoising. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05), Vol. 2, 2005, pp. 60–65, doi: 10.1109/CVPR.2005.38.
- [10] LI, Y.—LI, C.—LI, X.—WANG, K.—RAHAMAN, M. M.—SUN, C.—CHEN, H.—WU, X.—ZHANG, H.—WANG, Q.: A Comprehensive Review of Markov Random Field and Conditional Random Field Approaches in Pathology Image Analysis. *Archives of Computational Methods in Engineering*, Vol. 29, 2022, No. 1, pp. 609–639, doi: 10.1007/s11831-021-09591-w.
- [11] ROTH, S.—BLACK, M. J.: Fields of Experts. *International Journal of Computer Vision*, Vol. 82, 2009, No. 2, pp. 205–229, doi: 10.1007/s11263-008-0197-6.
- [12] DABOV, K.—FOI, A.—KATKOVNIK, V.—EGIAZARIAN, K.: Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, Vol. 16, 2007, No. 8, pp. 2080–2095, doi: 10.1109/TIP.2007.901238.
- [13] GU, S.—ZHANG, L.—ZUO, W.—FENG, X.: Weighted Nuclear Norm Minimization with Application to Image Denoising. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2862–2869, doi: 10.1109/CVPR.2014.366.
- [14] BURGER, H. C.—SCHULER, C. J.—HARMELING, S.: Image Denoising: Can Plain Neural Networks Compete with BM3D? 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 2392–2399, doi: 10.1109/CVPR.2012.6247952.
- [15] TERAJ, S.—AHMED, S. S.—MESSALI, Z.: Comparative Study of Video/Image Denoising Algorithms Based on Convolutional Neural Network CNN. 2022 2nd International Conference on Advanced Electrical Engineering (ICAEE), 2022, pp. 1–6, doi: 10.1109/ICAEE53772.2022.9962112.
- [16] ZHANG, K.—ZUO, W.—CHEN, Y.—MENG, D.—ZHANG, L.: Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE*

- Transactions on Image Processing, Vol. 26, 2017, No. 7, pp. 3142–3155, doi: 10.1109/TIP.2017.2662206.
- [17] LIU, W.—ZHANG, C.—TAI, Y.: Research on CNN-Based Image Denoising Methods. In: Nakamatsu, K., Kountchev, R., Patnaik, S., Abe, J.M., Tyugashev, A. (Eds.): *Advanced Intelligent Technologies for Industry (ICAIT 2021)*. Springer, Singapore, Smart Innovation, Systems and Technologies, Vol. 285, 2022, pp. 475–481, doi: 10.1007/978-981-16-9735-7_46.
- [18] ZHANG, Y.—TIAN, Y.—KONG, Y.—ZHONG, B.—FU, Y.: Residual Dense Network for Image Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 43, 2021, No. 7, pp. 2480–2495, doi: 10.1109/TPAMI.2020.2968521.
- [19] ZHU, F.—WANG, Y.—TANG, H.: The Combination of GRU and Dense Block for Image Denoising Network. In: Li, K., Liu, Y., Wang, W. (Eds.): *Exploration of Novel Intelligent Optimization Algorithms (ISICA 2021)*. Springer, Singapore, Communications in Computer and Information Science, Vol. 1590, 2022, pp. 95–102, doi: 10.1007/978-981-19-4109-2_10.
- [20] LIU, P.—ZHANG, H.—ZHANG, K.—LIN, L.—ZUO, W.: Multi-Level Wavelet-CNN for Image Restoration. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018, pp. 886–88609, doi: 10.1109/CVPRW.2018.00121.
- [21] PARK, B.—YU, S.—JEONG, J.: Densely Connected Hierarchical Network for Image Denoising. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 2104–2113, doi: 10.1109/CVPRW.2019.00263.
- [22] RONNEBERGER, O.—FISCHER, P.—BROX, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.): *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer, Cham, Lecture Notes in Computer Science, Vol. 9351, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.
- [23] CHAREST, M. R.—ELAD, M.—MILANFAR, P.: A General Iterative Regularization Framework for Image Denoising. 2006 40th Annual Conference on Information Sciences and Systems, 2006, pp. 452–457, doi: 10.1109/CISS.2006.286510.
- [24] MILANFAR, P.: A Tour of Modern Image Filtering: New Insights and Methods, Both Practical and Theoretical. *IEEE Signal Processing Magazine*, Vol. 30, 2013, No. 1, pp. 106–128, doi: 10.1109/MSP.2011.2179329.
- [25] TALEBI, H.—ZHU, X.—MILANFAR, P.: How to SAIF-Ly Boost Denoising Performance. *IEEE Transactions on Image Processing*, Vol. 22, 2013, No. 4, pp. 1470–1485, doi: 10.1109/TIP.2012.2231691.
- [26] CHEN, C.—XIONG, Z.—TIAN, X.—ZHA, Z. J.—WU, F.: Real-World Image Denoising with Deep Boosting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42, 2019, No. 12, pp. 3071–3087, doi: 10.1109/TPAMI.2019.2921548.
- [27] ROMANO, Y.—ELAD, M.: Boosting of Image Denoising Algorithms. *SIAM Journal on Imaging Sciences*, Vol. 8, 2015, No. 2, pp. 1187–1219, doi: 10.1137/140990978.
- [28] LIU, H.—LIU, F.—FAN, X.—HUANG, D.: Polarized Self-Attention: Towards High-Quality Pixel-Wise Mapping. *Neurocomputing*, Vol. 506, 2022, pp. 158–167, doi: 10.1016/j.neucom.2022.07.054.

- [29] LIM, B.—SON, S.—KIM, H.—NAH, S.—MU LEE, K.: Enhanced Deep Residual Networks for Single Image Super-Resolution. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1132–1140, doi: 10.1109/CVPRW.2017.151.
- [30] TIMOFTE, R.—ROTHER, R.—VAN GOOL, L.: Seven Ways to Improve Example-Based Single Image Super Resolution. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1865–1873, doi: 10.1109/CVPR.2016.206.
- [31] GARISOV, T.—IZMAILOV, P.—PODOPRIKHIN, D.—VETROV, D. P.—WILSON, A. G.: Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs. In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R. (Eds.): Advances in Neural Information Processing Systems 31 (NeurIPS 2018). Curran Associates, Inc., 2018, pp. 8789–8798, doi: 10.48550/arXiv.1802.10026.
- [32] IZMAILOV, P.—PODOPRIKHIN, D.—GARISOV, T.—VETROV, D.—WILSON, A. G.: Averaging Weights Leads to Wider Optima and Better Generalization. CoRR, 2018, doi: 10.48550/arXiv.1803.05407.
- [33] HE, K.—ZHANG, X.—REN, S.—SUN, J.: Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [34] IOFFE, S.—SZEGEDY, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In: Bach, F., Blei, D. (Eds.): Proceedings of the 32nd International Conference on Machine Learning. Proceedings of Machine Learning Research (PMLR), Vol. 37, 2015, pp. 448–456, <http://proceedings.mlr.press/v37/ioffe15.html>.
- [35] ZHANG, K.—ZUO, W.—GU, S.—ZHANG, L.: Learning Deep CNN Denoiser Prior for Image Restoration. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2808–2817, doi: 10.1109/CVPR.2017.300.
- [36] ZHANG, K.—ZUO, W.—ZHANG, L.: FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising. IEEE Transactions on Image Processing, Vol. 27, 2018, No. 9, pp. 4608–4622, doi: 10.1109/TIP.2018.2839891.
- [37] ZHANG, J.—NIU, Y.—SHANGGUAN, Z.—GONG, W.—CHENG, Y.: A Novel Denoising Method for CT Images Based on U-Net and Multi-Attention. Computers in Biology and Medicine, Vol. 152, 2023, Art.No. 106387, doi: 10.1016/j.combiomed.2022.106387.
- [38] MA, J.—PENG, C.—TIAN, X.—JIANG, J.: DBDnet: A Deep Boosting Strategy for Image Denoising. IEEE Transactions on Multimedia, Vol. 24, 2021, pp. 3157–3168, doi: 10.1109/TMM.2021.3094058.
- [39] BÜHLMANN, P.—YU, B.: Boosting with the L2 Loss: Regression and Classification. Journal of the American Statistical Association, Vol. 98, 2003, No. 462, pp. 324–339, doi: 10.1198/016214503000125.
- [40] OSHER, S.—BURGER, M.—GOLDFARB, D.—XU, J.—YIN, W.: An Iterative Regularization Method for Total Variation-Based Image Restoration. Multiscale Modeling & Simulation, Vol. 4, 2005, No. 2, pp. 460–489, doi: 10.1137/040605412.
- [41] HAN, S.—MENG, Z.—KHAN, A. S.—TONG, Y.: Incremental Boosting Convolutional Neural Network for Facial Action Unit Recognition. 2016, pp. 109–117, doi:

- 10.48550/arXiv.1707.05395.
- [42] CHEN, C.—XIONG, Z.—TIAN, X.—WU, F.: Deep Boosting for Image Denoising. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.): *Computer Vision – ECCV 2018*. Springer, Cham, Lecture Notes in Computer Science, Vol. 11215, 2018, pp. 3–19, doi: 10.1007/978-3-030-01252-6_1.
- [43] XIE, Z.—LIU, L.—WANG, C.: Model-Guided Boosting for Image Denoising. *Signal Processing*, Vol. 201, 2022, Art.No. 108721, doi: 10.1016/j.sigpro.2022.108721.
- [44] ZHANG, J.—QU, M.—WANG, Y.—CAO, L.: A Multi-Head Convolutional Neural Network with Multi-Path Attention Improves Image Denoising. In: Khanna, S., Cao, J., Bai, Q., Xu, G. (Eds.): *PRICAI 2022: Trends in Artificial Intelligence*. Springer, Cham, Lecture Notes in Computer Science, Vol. 13631, 2022, pp. 338–351, doi: 10.1007/978-3-031-20868-3_25.
- [45] YU, J.—ZHANG, J.—GAO, Y.: MACFNet: Multi-Attention Complementary Fusion Network for Image Denoising. *Applied Intelligence*, Vol. 53, 2023, No. 13, pp. 16747–16761, doi: 10.1007/s10489-022-04313-6.
- [46] TIAN, C.—XU, Y.—LI, Z.—ZUO, W.—FEI, L.—LIU, H.: Attention-Guided CNN for Image Denoising. *Neural Networks*, Vol. 124, 2020, pp. 117–129, doi: 10.1016/j.neunet.2019.12.024.
- [47] LI, X.—YUAN, Y.—WANG, Q.: Hyperspectral and Multispectral Image Fusion via Nonlocal Low-Rank Tensor Approximation and Sparse Representation. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 59, 2021, No. 1, pp. 550–562, doi: 10.1109/TGRS.2020.2994968.
- [48] HU, J.—SHEN, L.—SUN, G.: Squeeze-and-Excitation Networks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.
- [49] WOO, S.—PARK, J.—LEE, J. Y.—KWEON, I. S.: CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.): *Computer Vision – ECCV 2018*. Springer, Cham, Lecture Notes in Computer Science, Vol. 11211, 2018, pp. 3–19, doi: 10.1007/978-3-030-01234-2_1.
- [50] HUANG, G.—LIU, Z.—VAN DER MAATEN, L.—WEINBERGER, K. Q.: Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
- [51] KINGMA, D. P.—BA, J.: Adam: A Method for Stochastic Optimization. *CoRR*, 2014, doi: 10.48550/arXiv.1412.6980.
- [52] AGUSTSSON, E.—TIMOFTE, R.: NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1122–1131, doi: 10.1109/CVPRW.2017.150.
- [53] ROTH, S.—BLACK, M. J.: Fields of Experts: A Framework for Learning Image Priors. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’ 05), Vol. 2, 2005, pp. 860–867, doi: 10.1109/CVPR.2005.160.
- [54] ZHANG, L.—WU, X.—BUADES, A.—LI, X.: Color Demosaicking by Local Directional Interpolation and Nonlocal Adaptive Thresholding. *Journal of Electronic Imaging*, Vol. 20, 2011, No. 2, Art.No. 023016, doi: 10.1117/1.3600632.

- [55] CHEN, Y.—POCK, T.: Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, 2017, No. 6, pp. 1256–1272, doi: 10.1109/TPAMI.2016.2596743.
- [56] FAN, C. M.—LIU, T. J.—LIU, K. H.: SUNet: Swin Transformer UNet for Image Denoising. *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2022, pp. 2333–2337, doi: 10.1109/ISCAS48785.2022.9937486.
- [57] TIAN, C.—ZHENG, M.—ZUO, W.—ZHANG, B.—ZHANG, Y.—ZHANG, D.: Multi-Stage Image Denoising with the Wavelet Transform. *Pattern Recognition*, Vol. 134, 2023, Art.No. 109050, doi: 10.1016/j.patcog.2022.109050.
- [58] HUYNH-THU, Q.—GHANBARI, M.: Scope of Validity of PSNR in Image/Video Quality Assessment. *Electronics Letters*, Vol. 44, 2008, No. 13, pp. 800–801, doi: 10.1049/el:20080522.
- [59] WANG, Z.—BOVIK, A. C.—SHEIKH, H. R.—SIMONCELLI, E. P.: Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing (TIP)*, Vol. 13, 2004, No. 4, pp. 600–612, doi: 10.1109/TIP.2003.819861.



Yun JIA received his B.Eng. degree from the Harbin Institute of Technology, Ph.D. in the Harbin Institute of Technology. He is currently Associate Professor at the School of Information and Electronic Engineering, Shandong Technology and Business University, Shandong, China. His research interests include deep learning and intelligent signal processing, network communication and image processing.



Lidan YU received her B.Eng. degree from the Qingdao University of Science and Technology. She is currently pursuing her M.Eng. degree at the School of Information and Electronic Engineering, Shandong Technology and Business University, Shandong, China. Her research interests include image processing and computer vision.



Jie ZHAO received his B.Eng. degree from the Shandong Youth University of Political Science. He is currently pursuing his M.Eng. degree at the School of Information and Electronic Engineering, Shandong Technology and Business University, Shandong, China. His research interests include saliency detection and segmentation, remote sensing image interpretation, and deep learning.