CONTOURLET TRANSFORM-BASED GENERATIVE ADVERSARIAL NETWORK FOR ENHANCED FACIAL IMAGE DENOISING

Adamu, Z.M.[1*], Abdulrazaq, A.A.[1], Sani, B.H.[2], Ismail, A.A.[1], Suleiman, A.A.[3] and Olawoye, T.O.[4]

[1]Department of Computer Engineering, University of Maiduguri, Borno, Nigeria

[2]Department of Computer Engineering, Ahmadu Bello University, Zaria

[3]Department of Electrical Electronics Engineering, Edo State University, Iyahmo, Edo State

[4]Department of Electrical Electronics Engineering, Rufus Giwa Polytechnic, Owo, Ondo State

*Corresponding email: zainabmbarkindo@unimaid.edu.ng

ABSTRACT

Facial image denoising is a critical preprocessing step in computer vision applications such as face recognition, surveillance and biometric authentication, where noise can significantly degrade recognition accuracy and visual quality. While adoption of deep learning techniques particularly the Generative Adversarial Network (GAN)-based approach has significantly improved the image denoising process, integration of GAN with other approaches still enhances it performance especially in preserving fine facial structures such as edges, textures and contours. This study proposes a Contourlet Transform–Based Generative Adversarial Network (CT-GAN) for enhanced facial image denoising by integrating the contourlet transform into a GAN architecture to exploit multi-resolution and multi-directional feature representations for improved noise suppression and structural preservation. The model was evaluated using the FER2013 facial image dataset, where synthetic Gaussian, salt-and-pepper and Poisson noise were introduced at varying intensities to simulate real-world degradation. Experimental results demonstrate that the proposed CT-GAN achieves superior performance compared to conventional CNN-based denoisers and standard GAN models, with improvements in Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Mean Squared Error (MSE). The proposed CT-GAN method achieved a PSNR of 32.1 dB and SSIM of 0.94 alongside the lowest MSE value of 0.005. The integration of contourlet-domain features enhances edge preservation and texture fidelity, resulting in more visually coherent reconstructed facial images.

Keywords: Facial Image Denoising, Generative Adversarial Network, Contourlet Transform, Deep Learning, Image Restoration

1 INTRODUCTION

Facial image denoising has become an essential preprocessing step in modern computer vision systems, particularly in applications such as facial recognition, surveillance, human-computer interaction and biometric authentication[1, 2]. In real-world scenarios, facial images are often degraded by various types of noise introduced during image acquisition, transmission, compression, or environmental conditions such as low illumination and sensor limitations. These distortions can significantly reduce image quality and negatively affect the performance of downstream tasks by obscuring important facial features. Traditional image denoising techniques, including spatial filtering, wavelet-based methods and statistical approaches, have been widely studied. While these methods are computationally efficient, they often struggle to balance noise removal and detail preservation [3]. In many cases, excessive smoothing leads to the loss of fine facial structures such as edges, textures and subtle geometric patterns that are crucial for identity preservation. This limitation has motivated the development of more advanced learning-based approaches capable of modeling complex image distributions [4]. With the rapid advancement of deep learning, Convolutional Neural Networks (CNNs) have demonstrated strong performance in image restoration tasks by learning hierarchical feature representations directly from data [5, 6, 7]. However, CNN-based denoisers may still fail to effectively preserve directional and structural information, especially in highly corrupted images [8, 9]. More recently, Generative Adversarial Networks (GANs) have emerged as powerful frameworks for image generation and restoration, offering improved perceptual quality through adversarial learning between a generator and a discriminator [5, 10, 11]. Despite their success, standard GAN-based denoising models may still produce artifacts or lose fine structural details due to limited feature representation in the spatial domain [12, 13]. To address these challenges, transform-domain techniques have been introduced to enhance feature representation in image restoration tasks [2, 4, 14]. Among them, the Contourlet Transform provides a powerful multi-resolution and multi-directional decomposition that effectively captures edges, contours and geometric structures in images[15, 16, 18]. Unlike traditional wavelet transforms, which are limited in directional selectivity, the contourlet transform offers improved capability in representing smooth contours and directional textures, making it particularly suitable for facial image analysis [1]. In this study, we propose a Contourlet Transform–Based Generative Adversarial Network (CT-GAN) for enhanced facial image denoising. The proposed framework integrates contourlet-domain feature representation into a GAN architecture to improve the reconstruction of noisy facial images. The combination of adversarial learning strengths with directional multi-scale analysis, the model aims to suppress noise while preserving critical facial structures. The main contributions of this work are summarized as follows: (i) Development of a contourlet-enhanced GAN framework for facial image denoising. (ii) Integration of multi-resolution and directional feature extraction into the generator network. (iii) Demonstration of improved structural preservation and image quality over existing denoising methods. The remainder of this paper is organized as follows: Section 2 presents the related work in image denoising and GAN-based restoration methods. Section 3 presents the proposed methodology, including data preprocessing, network architecture and contourlet transformation. Section 4 discusses experimental results and performance evaluation. Finally, Section 5 concludes the study.

2 LITERATURE REVIEW

Literatures have demonstrated that contourlet-based denoising improves edge preservation and directional feature representation, making it suitable for facial image restoration tasks where subtle geometric patterns are critical [19, 20]. Subsequent works extended contourlet frameworks with adaptive thresholding and hybrid filtering strategies to enhance noise suppression performance while maintaining structural fidelity [21]. Nevertheless, transform-based methods remain limited by handcrafted assumptions and lack strong data-driven learning capability. The emergence of deep learning significantly advanced image denoising performance. Convolutional Neural Networks (CNNs) such as DnCNN, RED-Net and residual encoder–decoder architectures learn nonlinear mappings between noisy and clean images directly from data. CNN-based models demonstrated superior performance compared to traditional algorithms by automatically extracting hierarchical features [22]. More recent works introduced attention mechanisms and multi-scale feature fusion to better preserve facial textures under varying noise conditions [15, 23]. Despite their effectiveness, purely CNN-based approaches sometimes generate overly smooth outputs because optimization is dominated by pixel-wise reconstruction losses such as Mean Squared Error (MSE). Generative Adversarial Networks (GANs) introduced a paradigm shift by incorporating adversarial learning to improve perceptual realism. Since the seminal work of [24], GANs have been widely adopted for image restoration tasks including super-resolution, inpainting and denoising. GAN-based denoising models employ a generator–discriminator framework in which the generator produces clean images while the discriminator enforces realism constraints. Approaches such as adversarial autoencoders, conditional GANs and perceptual-loss-guided GAN architectures have demonstrated improved preservation of facial identity and high-frequency details compared to CNN regression models [6]. Recent studies further integrated attention modules, residual dense blocks and perceptual feature losses derived from pretrained networks to enhance visual fidelity in facial restoration tasks [25]. Although GANs improve perceptual quality, training instability and artifact generation remain persistent challenges. Consequently, hybrid frameworks combining transform-domain priors with deep generative models have recently attracted research interest. Transform-assisted deep networks embed handcrafted representations into learning pipelines to guide feature extraction. For example, wavelet-guided GANs and frequency-aware restoration networks demonstrated improved convergence and enhanced reconstruction of fine textures [21, 26]. These studies indicate that incorporating directional and multiscale representations into adversarial learning can effectively balance noise removal and structural preservation. Within facial image denoising specifically, maintaining identity information is particularly important for downstream applications such as face recognition and biometric verification. Recent research emphasizes combining geometric priors with deep generative models to preserve semantic facial structures under severe noise conditions [5, 27]. However, limited attention has been given to exploiting contourlet-based directional representations within GAN frameworks for facial denoising. Existing works either rely solely on deep networks or employ frequency transforms without fully leveraging adversarial perceptual optimization. Motivated by these gaps, the proposed CT-GAN integrates multiscale directional contourlet features with adversarial learning to enhance facial image denoising by combining the structural representation strength of contourlet decomposition with the perceptual reconstruction capability of GANs. This approach aims to achieve superior noise suppression while preserving identity-relevant facial details.

3 METHODOLOGY

The development of an improved contourlet transform generative adversarial network (CT-GAN) involves the integration of the contourlet transform module to the generator (G) network of the GAN. Figure 1 depicts the architecture of the developed CT-GAN.

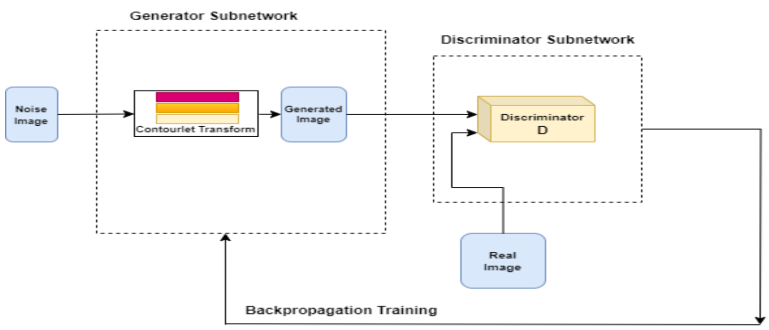

Figure 1. Architecture of the Developed CT-GAN for Image Denoising

The architecture of the developed CT-GAN illustrated in Figure 1 contain two (3) module components which are the generator (G) network block and the discriminator (D) network block. The steps in achieving the denoising output is presented in Algorithm 1.

Table 1. The steps in achieving the denoising output

Algorithm 1: CT-GAN Training Algorithm	
Input:	The training set $\mathcal{L} = (B_I, P_i)_{i=1}^{L}$ L: number of total iterations; weight parameter.
Output	The Generator able to denoise an input sample
Set	m to the total number of training data points
for	i =1: L do
	Noise Addition and Preprocessing:
	Add Poisson noise, Gaussian noise and salt-and-pepper noise to the real ground truth images y, generating noisy images x.
	Contourlet Transform: Decompose the noisy images x into 3-multi-level frequency subbands using the Contourlet Transform to enhances the representation of edges, textures and curve details across multiple scales. Generator Update: Draw m samples $(z^{(1)}, z^{(2)}, \ldots, z^{(m)})$ from a noise prior p(z). Fix the Discriminator D and perform gradient ascent on the parameters of D to maximize the adversarial loss V(G, D). Loss Computation: Compute the adversarial loss using Wasserstein loss to improve the realism of the generated images.
End for	
Return	The denoised images G(x).

3.1 Data Acquisition and Preprocessing

The FER2013 dataset was adopted in this study due to its grayscale image format to ensures uniformity in the input representation. To evaluate the robustness of the model under different degradation conditions, synthetic noise was deliberately introduced into the clean facial images. Three noisy variants of the dataset were generated by adding Gaussian noise, salt-and-pepper noise and Poisson noise at multiple intensity levels, thereby simulating common real-world image acquisition and transmission distortions. The prepared dataset was partitioned into training, validation and testing subsets following a 70:20:10 ratio. This data division enabled effective model learning, hyperparameter tuning and unbiased performance assessment using previously unseen samples. Prior to training, all images underwent preprocessing in which pixel values were normalized to the range [0,1], improving numerical stability and facilitating faster convergence of the Generative Adversarial Network (GAN). Since the original FER2013 images have a spatial resolution of 48×48, they were resized to 128×128 pixels to enhance feature representation while maintaining manageable computational cost. The increased resolution allowed the model to capture finer facial structures essential for accurate noise removal without significantly increasing training complexity. Following preprocessing, the images were stored as NumPy arrays to enable efficient data handling, minimize disk access overhead and support rapid batch loading during the training process.

3.2 Design of the Generator Network Architecture

The Generator network architecture is as depicted in Figure 2. The generator forms the central component of the proposed denoising framework and was specifically redesigned to incorporate contourlet-domain information into the Generative Adversarial Network learning process. It receives a noisy facial image as input and outputs a corresponding denoised reconstruction through a sequence of processing stages that include initial convolutional layers for low-level feature extraction, a contourlet-domain embedding module for capturing multiscale directional information, deep residual convolutional blocks for effective feature learning, feature fusion layers for integrating spatial and transform-domain representations and final reconstruction layers that generate the restored image. Instead of directly reconstructing the clean image, the generator adopts a residual learning strategy in which the network estimates the noise residual, leading to faster convergence and improved reconstruction accuracy. The incorporation of contourlet features further enhances structural preservation by maintaining edge continuity and directional texture information, which are critical for high-quality facial image restoration.

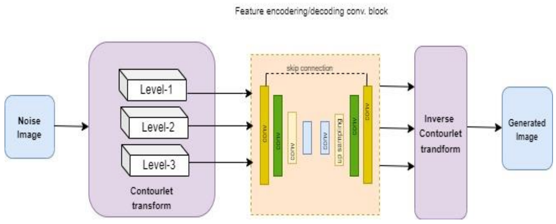

Figure 2. Generator Network Architecture

3.3 Contourlet Transformation

A three-level contourlet decomposition was applied to transform input images into a multi-resolution and directional representation space. The contourlet transform performs Laplacian Pyramid decomposition for multi-scale analysis and Directional Filter Bank decomposition for directional feature extraction. This process separates the image into frequency subbands representing edges, contours and textures at different orientations. Within this transformed domain, the high-frequency components emphasize edges and fine details, the directional information preserves curves and structural patterns while the feature weighting enhances significant textures while suppressing noise components. During reconstruction, inverse contourlet transformation fuses these features, allowing the generator to synthesize denoised images with improved clarity and minimal information loss.

3.4 Discriminator Network

The discriminator distinguishes between real clean images and generator-produced denoised images. It consists of stacked convolutional layers followed by nonlinear activation and down-sampling operations. The discriminator guides the generator through adversarial learning, encouraging perceptually realistic reconstructions. The adversarial loss function was employed to maintain training stability and reserve image salient feature

3.5 Training Procedure

The proposed CT-GAN was trained using a supervised learning strategy based on paired noisy and clean images. During training, the generator learns to reconstruct noise-free images from corrupted inputs, while the discriminator simultaneously learns to distinguish between real clean images and generated denoised outputs through adversarial optimization. Prior to training, the preprocessed dataset was shuffled and divided into mini-batches to ensure stable gradient updates and improved generalization. Model optimization was performed using the Adam optimizer, selected for its adaptive learning capability and effectiveness in GAN training. The initial learning rate was set to 2×10^{-4}, with momentum parameters configured to promote stable convergence of both generator and discriminator networks. Training was conducted using mini-batch gradient descent, where each iteration involved alternating updates between the generator and discriminator. The generator parameters were updated by minimizing a composite loss function comprising adversarial loss and reconstruction loss, while the discriminator was optimized to maximize classification accuracy between real and generated samples. An early stopping strategy based on validation loss monitoring was employed to prevent overfitting and unnecessary computational cost. Training was terminated when no significant improvement in validation performance was observed over consecutive epochs. The optimization process continued until convergence was achieved, indicated by stabilization of adversarial loss and reconstruction error, as well as consistent improvement in denoising performance on validation data.

4 RESULTS AND DISCUSSION

To evaluate the denoising performance of the CT-GAN model, images from the FER2013 dataset were employed. Controlled amount of noise was deliberately introduced to simulate different types of noises for image degradation. Specifically, salt-and-pepper noise, Gaussian noise and Poisson noise were applied as primary noise types, each added at a noise factor of 0.05 to corrupt the selected images. The CT-GAN model was then used to regenerate denoised images and the restored quality images were assessed using two quantitative metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Metric (SSIM). Figure 3 present sampled of corrupted and denoised images at 0.05 noise factor for the three-noise type.

Noise type	Corrupted image	Denoised image
Gaussian		
Poisson		
Salt and pepper		

Figure 3. Sampled of corrupted and denoised images at 0.05 noise factor

The comparison of the PSNR and SSIM values obtained for the three noise types is also presented in table 1.

Table 1. PSNR and SSIM at 0.05 Noise Factors

Noise type	PSNR		SSIM	
	Before denoised	After denoised	Before denoise	After denoised
Gaussian	32.0318	33.87578	0.8895	0.99948
Poisson	35.9951	37.27578	0.8989	0.99996
Salt and pepper	37.0414	39.32478	0.8995	0.99984

The Performance of the CT-GAN was also compared with other existing methods and evaluated using, Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Mean Squared Error (MSE) (Table 2).

Table 2. present the detail comparison analysis

Model	PSNR (dB)	SSIM	MSE
CNN Denoiser	27.8	0.86	0.012
Standard GAN	29.6	0.89	0.009
Wavelet-GAN	30.4	0.91	0.007
Proposed CT-GAN	32.1	0.94	0.005

Table 1 summarizes the denoising performance of the proposed CT-GAN at a noise factor of 0.05 using PSNR and SSIM metrics. The results show consistent improvement in image quality across all noise types after applying the CT-GAN model. For Gaussian noise, PSNR increases from 32.03 dB to 33.88 dB, while Poisson noise improves from 35.99dB to 37.28dB. The highest enhancement is observed for salt-and-pepper noise, where PSNR rises from 37.04 dB to 39.32 dB, indicating the model's strong capability in removing impulse noise. Similarly, SSIM values improve significantly from approximately 0.89 before denoising to nearly perfect structural similarity (0.99948–0.99996) after reconstruction. These results confirm that incorporating contourlet-domain features enables the CT-GAN to effectively suppress noise while preserving facial structures, edges and texture details, demonstrating robust and consistent denoising performance across varying noise conditions. Similarly, Table 2 presents a comparative performance analysis between the CT-GAN and existing denoising approaches using PSNR, SSIM and MSE metrics. The conventional CNN denoiser records the lowest performance with a PSNR of 27.8 dB and SSIM of 0.86, indicating limited capability in preserving fine facial structures during noise removal. The Standard GAN improves reconstruction quality, achieving 29.6 dB PSNR and 0.89 SSIM, demonstrating the advantage of adversarial learning in generating more realistic image details. Further improvement is observed with the Wavelet-GAN model, which leverages multi-resolution frequency information to reach 30.4 dB PSNR and 0.91 SSIM while reducing reconstruction error. The proposed CT-GAN achieves the best overall performance, recording the highest PSNR of 32.1 dB and SSIM of 0.94 alongside the lowest MSE value of 0.005. These results highlight the effectiveness of integrating contourlet-domain representations within the GAN framework. Unlike conventional CNN or wavelet-based approaches, the contourlet transform captures multidirectional and anisotropic image features, enabling superior preservation of facial contours, edges and texture information. The reduced MSE further confirms that CT-GAN produces reconstructions that are closer to the ground-truth images. In general, the tables demonstrate that the proposed CT-GAN provides more accurate, perceptually consistent and structurally realistic denoising performance compared to existing baseline methods.

5 CONCLUSIONS

This study introduced a CT-GAN model to enhance facial image denoising performance. The proposed framework integrates contourlet-domain feature representation with adversarial learning to effectively model multi-scale and multidirectional characteristics of facial images. The incorporation contourlet decomposition into the GAN architecture preserves critical facial structures, edges and texture information while removing various forms of image noise, thereby maintaining identity-relevant features during restoration. Experimental results obtained using the FER2013 dataset under multiple noise conditions demonstrate that the CT-GAN consistently surpasses conventional CNN-based denoisers, standard GAN models and wavelet-based approaches. The method achieves superior PSNR and SSIM values alongside reduced reconstruction error. therefore, the proposed CT-GAN framework offers an effective and robust solution for denoising degraded facial images, particularly in challenging acquisition environments where noise significantly affects visual quality and biometric reliability.

REFERENCES

1. Kimbrough, T., Tian, P., Liao, W. and Yu, W. (2023). Performance of GAN-Based denoising and restoration techniques for adversarial face images. *2023 IEEE/ACIS 21st International Conference on Software Engineering Research, Management and Applications*, pp. 335-340.

2. Biswas, S. and Sil, J. (2020). An efficient face recognition method using contourlet and curvelet transform. *Journal of King Saud University-Computer and Information Sciences*, vol. 32, no. 6, pp. 718-729.

3. Asmare, M.H., Asirvadam, V.S. and A.F.M. (2015). Hani, Image enhancement based on contourlet transform. *Signal, Image and Video Processing*, vol. 9, no. 7, pp. 1679-1690.

4. Vetrivelan, P. and Kandaswamy, A.K.A. (2019). Image Denoising Using Wavelet Based Contourlet Transform. *Journal of Electrical Engineering*, vol. 19, no. 2, pp. 1-5.

5. Zhu, M.L., Zhao, L.L. and Xiao, L. (2022). Image denoising based on GAN with optimization algorithm. *Electronics*, vol. 11, no. 15, pp. 24-45.

6. Zhang, K., Zuo, W. and Zhang, L. (2018). FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608-4622.

7. Liu, Z., Yan, W.Q. and Yang, M.L. (2018). Image denoising based on a CNN model. *2018 4th International conference on control, automation and robotics, IEEE*.

8. Tripathi, M. (2021). Facial image denoising using AutoEncoder and UNET. *Heritage and Sustainable Development*, vol. 3, no. 2, pp. 1-89.

9. Ilesanmi, A.E. and Ilesanmi, T.O. (2021). Methods for image denoising using convolutional neural network: a review. *Complex & Intelligent Systems*, vol. 7, no. 5, pp. 2179-2198.

10. Gong, P., Liu, J. and Lv, S. (2020). Image denoising with GAN based model. *Journal of Information Hiding and Privacy Protection*, vol. 2, no. 4, pp. 155.

11. Tran, L.D., Nguyen, S.M. and Arai. M. (2020). GAN-based noise model for denoising real images. *Proceedings of the Asian Conference on Computer Vision*, 2020.

12. Din, N.U., Javed, K., Bae, S. and Yi, J. (2020). Effective removal of user-selected foreground object from facial images using a novel GAN-based network. *IEEE Access,* vol. 8, 109648-61.

13. Banerjee, S. and Das, S. (2020). SD-GAN: Structural and Denoising GAN reveals facial parts under occlusion. *arXiv preprint arXiv:2002.08448*, 2020.

14. Xu, C. and Zou, Y. (2024). Contour wavelet diffusion–a fast and high-quality facial expression generation model. *Connection Science*, vol. 36, no. 1, pp. 2316023.

15. Chen, G. and Xie. W. (2020). Noise Robust Illumination Invariant Face Recognition via Contourlet Transform in Logarithm Domain. *International Conference on Intelligent Computing, Springer*.

16. Biswas, S. and Sil. J. (2015). An efficient face recognition method using contourlet transform. *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks, IEEE*.

17. Anila, S., Sivaraju, S. and Devarajan, N. (2017). A new contourlet based multiresolution approximation for MRI image noise removal. *National Academy Science Letters*, vol. 40, no. 1, pp. 39-41.

18. AlAsadi, A.H.H. (2015). Contourlet transform based method for medical image denoising. *Journal of AL-Qadisiyah for computer science and mathematics*, vol. 7, no. 1, pp. 146-159.

19. Do, M.N. and Vetterli, M. (2006). The contourlet transform: an efficient directional multiresolution image representation. *Official Journal-European Union Information and Notices C*, vol. 49, no. 27A, pp. 2091.

20. Da Cunha, A.L., Zhou, J. and Do, M.N. (2006). The nonsubsampled contourlet transform: theory, design, and applications. *IEEE transactions on image processing*, vol. 15, no. 10, pp. 3089-3101.

21. Zhang, M., Yang, C., Yuan, Y., Guan, Y., Wang, S. and Liu, Q. (2021). Multi-wavelet guided deep mean-shift prior for image restoration. *Signal Processing: Image Communication*, vol. 99, pp. 116449.

22. Zhang, K., Zuo, W., Chen, Y., Meng, D. and Zhang, L. (2017). Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142-3155.

23. Li, Y., Zhou, T., He, K., Zhou, Y. and Shen, D. (2023). Multi-scale transformer network with edge-aware pre-training for cross-modality MR image synthesis. *IEEE transactions on medical imaging*, vol. 42, no. 11, pp. 3395-407.

24. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets, advances in neural information processing systems, vol. 3, no. 6.

25. Liu, J., Shang, J., Liu, R. and Fan, X. (2022). Attention-guided global-local adversarial learning for detail-preserving multi-exposure image fusion. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 8, pp. 5026-5040.

26. Wang, J., Li, P., Xiang, L., Wang, R., Zhang, Z., Tian, Q. and He, Z. (2025). FDNet: A Frequency-Aware Decomposition Network for Robust Face Super-Resolution Against Adversarial Attacks. *IEEE Transactions on Information Forensics and Security*, vol. 20, pp. 6175-6189.

27. Yue, Z., Yong, H., Zhao, Q., Zhang, L., Meng, D. and Wong, KY. (2024). Deep variational network toward blind image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 11, pp. 7011-7026.