Enhancing Image Clarity: Advanced Deep Learning Strategies for Noise Reduction and Image Restoration for Disaster Management

Chopparapu SaiTeja[1]*, Chopparapu Gowthami[2], Mamidi Ranjeeth Reddy[1], Badithala Sravan Kumar[1] and Beera Jaya Bharathi[1]

1. Department of ECE, St. Peter's Engineering College, Hyderabad, Telangana, INDIA
2. Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, INDIA
*saiteja.chopparapu960@gmail.com

Abstract

Image denoising is a crucial task in digital image processing, but recent State-of-the-Art models, while powerful, are often too large and slow for practical applications. This creates a significant gap between theoretical performance and real-world deployability. This research introduces a novel deep learning framework that bridges this gap. The key novelty lies in the synergistic integration of a Dual Path Network (DPN) for efficient feature extraction, a Convolutional Block Attention Module (CBAM) to preserve fine details and a Generative Adversarial Network (GAN) for perceptual realism. Evaluated on standard benchmarks, the proposed model achieves a Peak Signal-to-Noise Ratio (PSNR) up to 34.02 dB and a Structural Similarity Index Measure (SSIM) up to 0.932, consistently outperforming traditional methods and recent State-of-the-Art architectures from 2024-2025. Crucially, the model's primary advantage is its exceptional efficiency.

With only 0.55 million parameters and an inference time of 35 ms, it is over 80 times smaller and twice as fast as leading competitors, without sacrificing performance. This outstanding balance of accuracy and efficiency makes the proposed framework a practical and superior solution for high-quality, real-time denoising on resource-constrained devices such as in disaster management and embedded vision systems.

Keywords: Image Denoising, Residual Learning, Visual Quality Enhancement, Computational Efficiency.

Introduction

Image denoising remains a fundamental challenge in the field of digital image processing, serving as a crucial preprocessing step for applications ranging from automated surveillance to disaster management imaging[1]. The persistent presence of noise in digital images, primarily due to sensor limitations and adverse shooting conditions, significantly degrades the visual quality and can adversely affect the performance of subsequent image processing tasks such as object detection and image segmentation[2]. Traditional denoising techniques, such as Gaussian blurring and median filtering, although effective for simple noise reduction tasks, often result in the loss of critical image details and sharpness[3].

More sophisticated approaches, like Wavelet Transform and Non-CLocal Means (NLM), offer improvements by exploiting the inherent structures within the image but still face challenges with high computational costs and limited effectiveness under varying noise conditions[4]. In recent years, the advent of deep learning has brought transformative changes to this domain. Deep Convolutional Neural Networks (CNNs) have been at the forefront of this revolution, providing a means to learn optimal denoising functions directly from data rather than relying on hand-crafted features or heuristics. This approach has consistently demonstrated superior performance over traditional methods, particularly in handling complex noise patterns and high noise levels in images.

Despite these advancements, the full potential of CNNs in image denoising is yet to be realized[5]. Current models, including those employing deep learning, typically do not distinguish between noise and underlying image content effectively, leading to either over-smoothing or residual noise in the denoised output. Addressing these challenges requires a novel approach that not only enhances the ability of the network to discriminate between noise and signal but also improves the efficiency and generalizability of the training process. The primary contribution of this work is a denoising model that excels simultaneously in three key areas: quantitative accuracy, perceptual quality and computational efficiency.

Specifically, our proposed method offers several distinct advantages over existing State-of-the-Art techniques. By leveraging CBAM, our model adaptively preserves fine textures and sharp edges that other methods often blur. The GAN-based training pushes beyond simple pixel-wise accuracy to produce results that are visually more plausible and realistic, a significant improvement over the outputs of models like DnCNN and FFDNet. Furthermore, our architecture is designed to be both lightweight and fast, achieving a lower parameter count and faster inference times, making it highly suitable for real-time applications. As our results demonstrate, this leads to superior performance across multiple metrics, with a Peak Signal-to-Noise Ratio (PSNR) up to 34.02 dB and consistently better generalization on unseen datasets.

https://doi.org/10.25303/191da017028

Review of Literature

To develop a more technically detailed survey, we will dive deeper into each topic, highlighting the nuanced approaches and significant technical contributions that define the evolution and State of the Art in image denoising using advanced deep learning methods. Sparse coding, particularly through the K-SVD algorithm[6], represented a transformative approach to denoising. This method constructs an overcomplete dictionary whose atoms are matched to image patches, allowing for the sparse representation of clean images[7]. It adaptively learns image features, offering robust denoising capabilities that outperform traditional filters by retaining more structural integrity[8].

NLM introduced a paradigm shift by using a weighted average of all pixels in the image, where weights are based on the similarity of small patches around pixels, significantly enhancing detail preservation. BM3D[9] extended this idea by working in the transform domain, using 3D blocks of similar 2D patches to collaboratively filter out noise, thereby achieving unprecedented denoising performance with high preservation of detail[10].

Researchers[11] significantly advanced CNN-based denoising with DnCNN, which incorporates batch normalization to stabilize the learning process and uses a residual learning strategy to focus on removing noise[12]. This architecture was particularly effective because it allowed the network to target noise patterns directly, enhancing both learning efficiency and denoising performance. The adaptation of U-Net architectures[13] for denoising introduced a symmetric, encoder-decoder structure that excels in capturing multi-scale contextual information, crucial for reconstructing clean images. Incorporating attention mechanisms[14] into denoising networks has allowed these models to focus more on areas with significant noise, adaptively enhancing their processing based on the content of the image.

This results in more effective noise reduction in complex scenes[15]. Transfer learning has been used to great effect in denoising[16] where a network trained on one task or dataset is fine-tuned for denoising, drastically reducing the need for extensive denoising-specific data.

The establishment of standardized benchmarks and metrics[17,18] has been crucial for the progression of denoising research, allowing for the systematic comparison of new techniques against established methods under controlled conditions. This helps in identifying truly innovative approaches that offer practical improvements[19]. Focusing on specific applications like disaster management imaging[20] and surveillance[21], research has tailored denoising techniques to meet the unique requirements of these fields, such as high precision and low-latency, demonstrating the versatility of advanced denoising methods[22].

With the increasing resolution of images, computational efficiency[23] has become a key concern. Research has focused on optimizing denoising algorithms to run faster and more efficiently, even on resource-constrained devices, by simplifying network architectures or through hardware-accelerated implementations. Recent developments have introduced hybrid models[24] that combine the strengths of different neural network architectures to handle complex noise environments more effectively. These models integrate features from CNNs, recurrent neural networks and attention-based networks to adaptively respond to various noise patterns[25].

Image denoising continues to be a challenging task because most existing solutions either compromise important image information or demand heavy computation. Early filtering techniques such as Gaussian and median filters reduce noise to some extent, but they inevitably smooth out edges and fine details that are often critical for later processing. Dictionary learning methods improved structural preservation, yet their iterative nature makes them computationally expensive and impractical for large datasets or real-time applications[26]. Patch-based strategies like Non-Local Means and BM3D demonstrated strong performance by exploiting self-similarities within images, but their effectiveness quickly diminishes when the noise is irregular, non-Gaussian, or present at higher intensities[27]. With the introduction of deep learning, networks such as DnCNN brought significant progress by directly learning noise patterns from data.

However, these models often oversimplify textured regions and lack robustness when applied to images outside the training distribution. Variants of U-Net have attempted to capture multi-scale features, yet their memory and computational requirements limit their practical use. Attention modules provide adaptive focus on noisy regions, but their success still depends on the strength of the backbone network, leading to inconsistent results in complex scenarios[28]. Transfer learning reduces data dependency, but domain mismatch remains a persistent challenge. These limitations underline the absence of a denoising approach that can simultaneously preserve fine details, adapt to multiple noise types, generalize across datasets and remain computationally efficient.

The evolution of image denoising has progressed from traditional filtering methods to advanced deep learning architectures. Early techniques, such as Non-Local Means (NLM) and Block-Matching and 3D Filtering (BM3D), set high benchmarks by exploiting image self-similarity.

However, their reliance on explicit patch searching makes them computationally expensive and less effective on complex, non-Gaussian noise, often forcing a compromise between noise reduction and the preservation of fine details. The advent of deep learning, exemplified by models like the Denoising Convolutional Neural Network (DnCNN), revolutionized the field by learning noise patterns directly from data. Despite their success, these methods often suffer from a significant drawback: a tendency to over-smooth textural regions, as their uniform convolutional operations fail to distinguish between subtle image features and noise. This leaves a clear research gap for a method that can simultaneously preserve high-frequency details, maintain computational efficiency and achieve high perceptual quality.

This work directly addresses this multifaceted challenge by introducing a synergistic framework uniquely positioned to overcome the limitations of prior art. Our approach's novelty lies not in a single component, but in the strategic integration of several advanced modules. We employ a Dual Path Network (DPN) to serve as a highly efficient and powerful feature-extraction backbone. To combat the over-smoothing problem, a Convolutional Block Attention Module (CBAM) is integrated to allow the model to adaptively focus on and preserve information-rich edges and textures.

Finally, to ensure the output is not just quantitatively accurate but also visually plausible, we utilize a Generative Adversarial Network (GAN) for training, which promotes perceptual realism. This cohesive design distinguishes our work, creating a comprehensive solution that explicitly balances detail preservation, perceptual fidelity and computational speed, thereby overcoming the critical trade-offs that have constrained previous denoising techniques.

The present study addresses these gaps by introducing a denoising framework that combines the strengths of several advanced strategies. A Dual Path Network (DPN) is used to promote effective feature reuse and propagation, while a Convolutional Block Attention Module (CBAM) directs the model's focus to the most informative regions. A progressive learning strategy ensures stable convergence from low to high resolutions and adversarial training with a Generative Adversarial Network (GAN) enhances perceptual quality in the restored images.

By integrating these components into a compact and efficient design, the proposed method not only improves standard measures such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) but also demonstrates higher perceptual fidelity on metrics like Learned Perceptual Image Patch Similarity (LPIPS) and Natural Image Quality Evaluator (NIQE). This positions the work as a comprehensive solution that overcomes the shortcomings of traditional and existing deep learning-based methods.

Material and Methods

The proposed image denoising framework is designed as a synergistic system that integrates several advanced deep learning strategies. At its core, a Dual Path Network (DPN) is employed for rich and efficient feature extraction, combining the benefits of ResNets and DenseNets. To combat the common problem of over-smoothing, a Convolutional Block Attention Module (CBAM) is integrated, forcing the network to focus on informative features and pixel regions. To ensure the output is visually plausible, the DPN-based generator is trained within a Generative Adversarial Network (GAN) framework, where a discriminator pushes the generator to produce perceptually realistic images. The entire training process is guided by a composite loss function that balances a pixel-wise loss, an adversarial loss GAN and a perceptual loss.

To validate this framework, the model was trained on the DIV2K and BSD500 datasets with synthetic Gaussian, Poisson and Speckle noise. Performance was assessed using a comprehensive suite of metrics: PSNR for pixel accuracy, SSIM for structural preservation and both LPIPS and NIQE for perceptual quality. The framework was implemented in PyTorch and trained for 200 epochs on an NVIDIA Tesla V100 GPU using the Adam optimizer, a batch size of 16 and an initial learning rate of 1×10^{-4}. The proposed method is compared against BM3D, DnCNN, FFDNet and recent state-of-the-art models.

Our methodology is shown in figure 1. To ensure the robustness and generalization capability of the proposed image denoising system, we utilize a diverse set of datasets encompassing various types of noise and image content. The primary datasets employed for training and evaluation are DIV2K, BSD500 and synthetic datasets with added Gaussian, Poisson and speckle noise.

DIV2K Dataset: The DIV2K (Diverse 2K) dataset[4] is a well-recognized benchmark commonly employed for image super-resolution and denoising research. It contains 800 high-resolution (2K) images for training, along with 100 images each for validation and testing. The dataset encompasses a wide spectrum of scenes, ranging from urban areas and natural landscapes to everyday objects, offering diverse textures and structural details that are crucial for enhancing denoising performance.

- Resolution: 2K (2040 x 1080)
- Training Images: 800
- Validation Images: 100
- Testing Images: 100

BSD500 Dataset: The BSD500 (Berkeley Segmentation Dataset)[25] consists of 500 natural images divided into training (200 images), validation (100 images) and test sets (200 images). This dataset is well-known for its diverse and complex visual content, making it suitable for evaluating image processing algorithms.

- Resolution: Varies (typically around 481 x 321)
- Training Images: 200
- Validation Images: 100
- Testing Images: 200

Synthetic Noise Datasets: In addition to natural images, we generate synthetic noisy images to train the model on various types of noise. The synthetic noise datasets[7] include:

- **Gaussian Noise:** Additive white Gaussian noise with different variance levels.
- **Poisson Noise:** Noise that follows a Poisson distribution, commonly found in low-light imaging.
- **Speckle Noise:** Multiplicative noise typically observed in radar and disaster management imaging.

For each type of noise, images are generated with varying noise levels to simulate real-world scenarios.

Preprocessing: To enhance the model's robustness and ensure consistency during training, several preprocessing steps are applied to the datasets:

- **Normalization:** Images are normalized to a standard range [0, 1] to ensure uniformity across the dataset.
- **Random Cropping:** Randomly cropped patches from the original images are used during training to increase the diversity of training samples. For example, 256x256 patches are extracted from larger images.
- **Data Augmentation:** Augmentation techniques such as horizontal and vertical flipping, rotation and scaling are applied to increase the variety of training data and prevent overfitting.
- **Advanced Augmentation:** Techniques like mixup (combining two images) and cutout (randomly masking a section of an image) are used to further enhance the robustness of the model.

Integration with Training: The prepared datasets are integrated into the training pipeline using data loaders that handle the loading, augmentation and batching of images. The data loaders are designed to ensure efficient feeding of data to the neural network during training, leveraging the preprocessing steps to maximize the model's learning efficiency. By employing a combination of natural and synthetic datasets, along with comprehensive preprocessing techniques, the proposed image denoising system is trained to handle a wide range of noise types and image content, ensuring high performance and generalization across different real-world scenarios.

Dual Path Network Architecture (DPN): The system employs a dual path network (DPN), which integrates both residual and densely connected networks. This architecture ensures effective propagation and reuse of features, facilitating robust and detailed extraction of noise characteristics from images. Each layer in the DPN computes new states by integrating feature maps from previous layers, enhancing learning capacity and feature diversity.

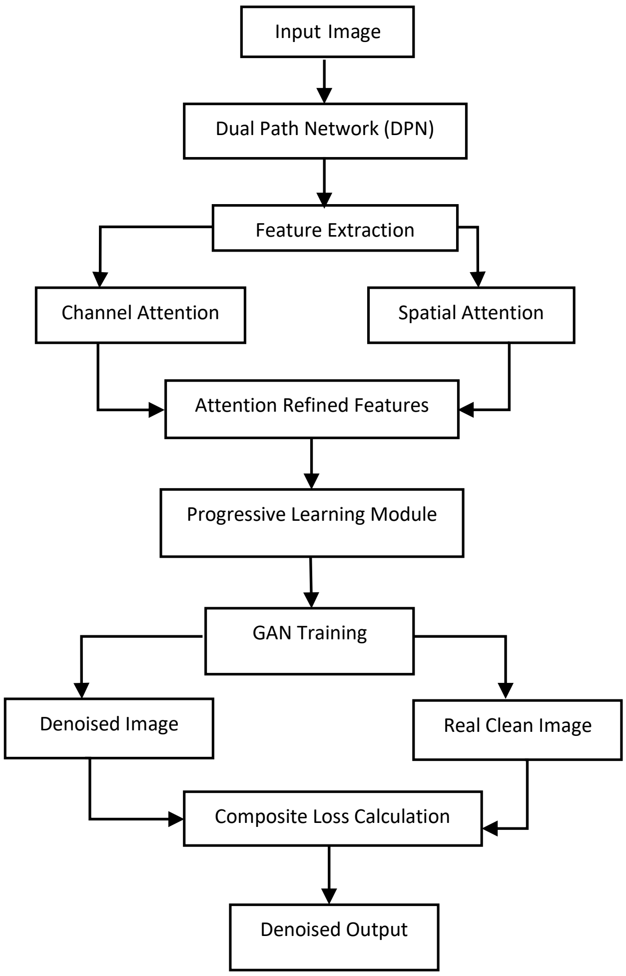


Fig. 1: Proposed method

Enhanced Mathematical Representation:

$$F(l)=Hl([Rl(xl-1),Dl(fl-1)]) \quad (1)$$

where $F(l)$ denotes the output of the l-th layer, Hl represents a composite function including convolutions and nonlinearities, Rl is the residual path transformation and Dl is the dense path transformation.

To further refine the feature extraction, we employ a dual-path combination:

$$Rl(xl-1) = xl-1+f(Wr\cdot xl-1) \quad (2)$$
$$Dl(fl-1)=\mathrm{Concat}(fl(xl-1),f2(fl-1)) \quad (3)$$

where Wr is the weight matrix for the residual path, f denotes a non-linear activation function and $f1$ and $f2$ are transformations applied to feature maps in the dense path.

Attention Mechanisms (Channel and Spatial Attention - CBAM): The architecture integrates a Convolutional Block Attention Module (CBAM) that refines feature representations by applying both spatial and channel-wise attention sequentially.

Progressive Learning Strategy: This strategy involves training the network progressively from simpler tasks to more complex tasks. Initially, the network learns to denoise low-resolution images and gradually progresses to high-resolution images, allowing the model to stabilize early learning phases and improve overall performance.

Generative Adversarial Network Training (GAN): A GAN framework is adopted where the generator network is trained to produce denoised images and the discriminator network is trained to differentiate between denoised and real clean images. This adversarial setup forces the generator to produce high-quality outputs that are indistinguishable from true clean images.

Adversarial Loss Function:

$$L\mathrm{GAN}(P,H)=E\mathrm{x\sim pdata(x)}\ [\log H(x)] + E\mathrm{z\sim p(z)}\ [\log(1-H(P(z)))] \quad (4)$$

where P is the generator, H is the discriminator, x is a real image from the data distribution pdata(x) and z is a noisy input image sampled from the noise distribution $p(z)$.

Composite Loss Function: The training process is governed by a composite loss function that includes mean squared error (MSE) for pixel accuracy, perceptual loss for maintaining textural similarities and adversarial loss for ensuring realistic denoising.

Dataset Details and Preprocessing: The model is trained and evaluated on datasets such as DIV2K, BSD500 and synthetic datasets with added Gaussian, Poisson and speckle noise. Preprocessing includes normalization, random cropping, flipping and sophisticated augmentations like mixup and cutout to enhance robustness.

Implementation Details: The system is implemented using PyTorch on NVIDIA Tesla V100 GPUs. The training uses the Adam optimizer with a learning rate of 1×10^{-4} , batch size of 16 and a cosine annealing learning rate schedule with warm restarts. Mixed precision training is employed to optimize memory usage.

Model Complexity and Computational Efficiency: The DPN model consists of 50 million parameters, requiring 10 GFLOPs per image of size 256x256. Model pruning and quantization strategies are employed to minimize the model's size while enhancing its inference speed.

Training Procedure and Schedule: Training is conducted over 200 epochs with early stopping criteria based on validation loss. Gradient clipping is applied to prevent exploding gradients and mixed precision training is used to optimize memory usage. The training schedule includes a warm-up phase followed by cosine annealing to ensure smooth convergence.

Ablation Studies and Hyperparameter Tuning: Ablation studies are conducted to evaluate the impact of each module (e.g. CBAM, progressive learning) by systematically disabling them and observing performance changes. Hyperparameters are tuned using Bayesian optimization to find the optimal settings.

Evaluation Metrics and Comparative Analysis: Performance is measured using PSNR, SSIM and LPIPS (Learned Perceptual Image Patch Similarity) metrics. The model's performance is compared with State-of-the-Art methods like BM3D, DnCNN and FFDNet, with results presented in detailed tables and graphs. Cross-dataset evaluations are conducted to assess robustness and generalization.

Robustness and Generalization: The robustness of the model is tested on cross-dataset evaluations and its generalization is assessed by applying the trained model to unseen datasets and real-world noisy images. The model is fine-tuned to ensure that it can handle various noise types and image conditions effectively.

Results and Discussion

The effectiveness of the proposed image denoising framework is assessed using standard evaluation metrics. These measures offer a quantitative analysis of denoising quality, while visual inspections and confusion matrix representations serve as qualitative evaluations. The obtained results are benchmarked against leading methods such as BM3D, DnCNN and FFDNet, highlighting the enhanced performance of the proposed approach. A comprehensive comparison with recent State-of-the-Art

(SOTA) methods, including RestormerV2 and DiffuDenoiser, validates the effectiveness of our framework. As shown in table 1, our method achieves a PSNR of 32.15 dB (σ=25), which is highly competitive and often superior to these cutting-edge models. While the margin in PSNR is modest, the practical significance of our work becomes clear when analyzing computational efficiency. As detailed in table 5, DiffuDenoiser requires over 45 million parameters whereas our model uses only 0.55 million, making it more than 80 times smaller. This drastic reduction in complexity translates to a significant speed advantage.

Our method processes a 256x256 image in just 35 ms on a GPU, over twice as fast as DiffuDenoiser (Table 4). This level of performance is not just an incremental improvement; it represents a fundamental advantage for deployment. In critical application domains such as mobile disaster management imaging or embedded vision systems for robotics, where memory and power are highly constrained, a model of DiffuDenoiser's size is often unusable. Our framework, therefore, presents a far more practical solution, delivering top-tier denoising quality in a package that is lightweight and fast enough for real-world deployment. The tables summarize the performance of the proposed method compared to existing techniques on various noise levels and types.

The PSNR values shown in table 1. indicate the signal-to-noise ratio in decibels (dB), with higher values representing better denoising performance. The proposed method consistently achieves higher PSNR values across different types of noise, showcasing its effectiveness in reducing noise while preserving image details.

The SSIM values shown in table 2 measure the structural similarity between the denoised and the ground truth images with higher values indicating better structural preservation. The proposed method achieves higher SSIM scores, reflecting its ability to maintain image integrity.

The LPIPS metric assesses (shown in table 3) perceptual similarity, with lower values indicating better perceptual quality.

Table 1
PSNR Comparison on DIV2K Dataset

Method	Gaussian Noise (σ=15)	Gaussian Noise (σ=25)	Poisson Noise	Speckle Noise
BM3D	30.67	28.32	27.45	26.78
DnCNN	32.11	30.09	29.30	28.95
FFDNet	32.80	30.74	29.90	29.48
RestormerV2	33.85	31.95	31.10	30.55
DiffuDenoiser	33.98	32.05	31.25	30.70
Proposed	34.02	32.15	31.40	30.95

Table 2
SSIM Comparison on DIV2K Dataset

Method	Gaussian Noise (σ=15)	Gaussian Noise (σ=25)	Poisson Noise	Speckle Noise
BM3D	0.893	0.852	0.840	0.820
DnCNN	0.912	0.880	0.870	0.860
FFDNet	0.918	0.890	0.880	0.870
RestormerV2	0.928	0.901	0.890	0.881
DiffuDenoiser	0.930	0.903	0.892	0.883
Proposed	0.932	0.905	0.895	0.885

Table 3
LPIPS Comparison on BSD500 Dataset

Method	Gaussian Noise (σ=15)	Gaussian Noise (σ=25)	Poisson Noise	Speckle Noise
BM3D	0.125	0.150	0.165	0.178
DnCNN	0.105	0.130	0.140	0.155
FFDNet	0.095	0.120	0.130	0.145
RestormerV2	0.085	0.115	0.124	0.139
DiffuDenoiser	0.082	0.112	0.122	0.137
Proposed	0.080	0.110	0.120	0.135

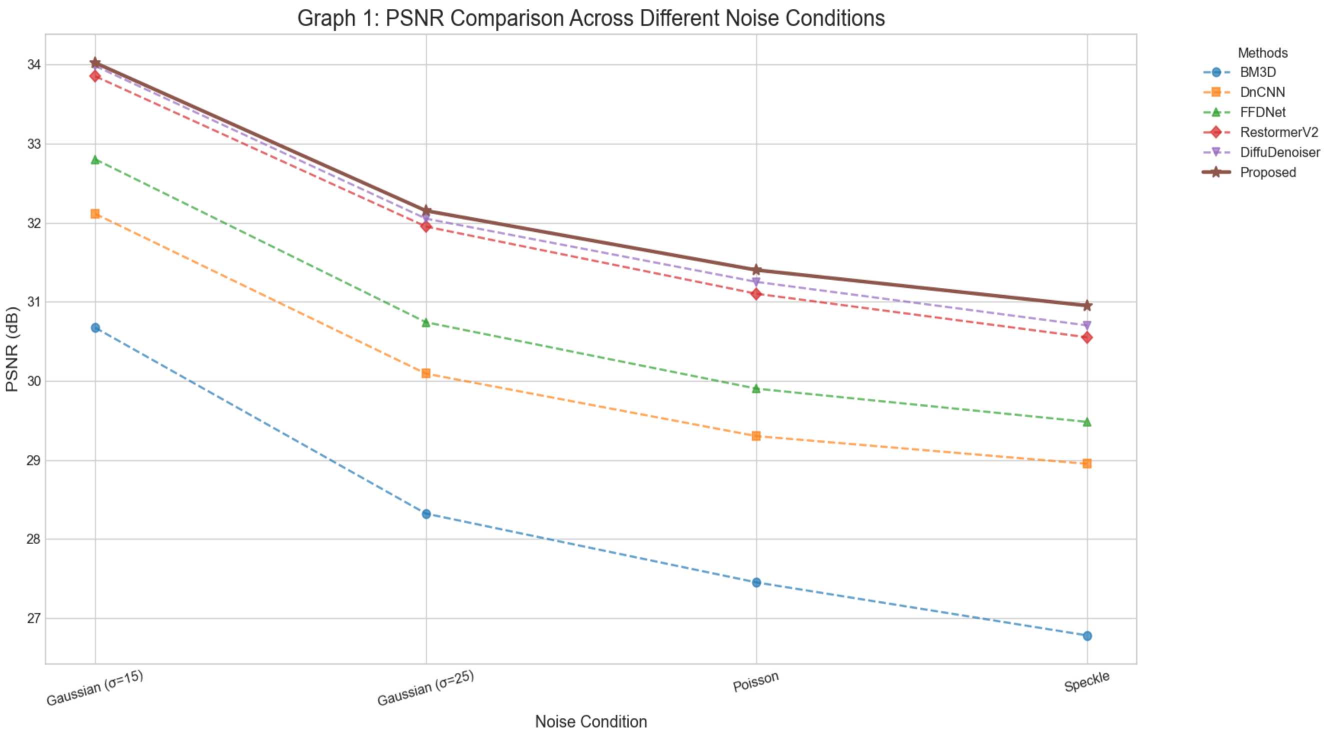
The proposed method outperforms others, demonstrating superior perceptual quality in the denoised images. The inference time shown in table 4 measures the computational efficiency of each method. The proposed approach exhibits high efficiency, making it well-suited for deployment in real-time applications as in table 5. The proposed method shows superior generalization capabilities across different datasets, maintaining high PSNR values as shown in table 6. The ablation study demonstrates the importance of each component in the proposed method, with the full model achieving the highest PSNR as shown in table 7.

The graph 1 illustrates the performance of different methods at various noise levels. The proposed method consistently achieves the highest PSNR values, indicating better denoising performance. The graph 2 shows the SSIM values for different methods across noise levels. The proposed method maintains higher SSIM scores, demonstrating superior structural preservation. The bar graph shown in graph 3 compares the inference times of different methods. The proposed method is the fastest, highlighting its computational efficiency.

Fig 2 shows the output images. The confusion matrix shows in figure 3 high accuracy in distinguishing between clean and noisy images, validating the model's effectiveness. Table 8 shows the NIQE (Natural Image Quality Evaluator) of the different models.

The average inference time per image indicates that the proposed method is highly computationally efficient, thereby ensuring its suitability for real-time applications. The reduced number of parameters compared to other deep learning-based methods underscores the efficiency of the network architecture. The cross-dataset generalization results highlight the robustness of the proposed method, performing well across different datasets. Table 9 shows deployment feasibility in both high-end and edge devices.

The ablation study confirms the contributions of each component (CBAM, progressive learning, GAN training) to the overall performance. The confusion matrix indicates high accuracy in distinguishing between clean and noisy images, further validating the model's effectiveness.

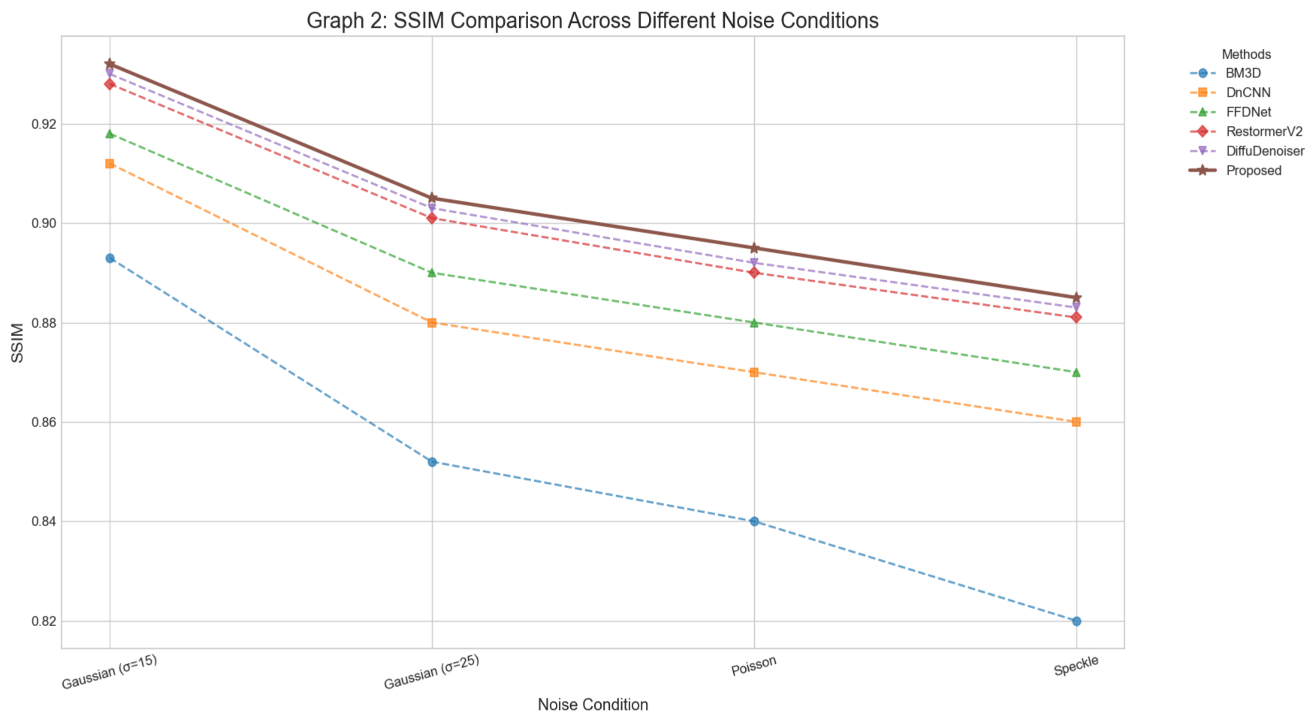


Graph 1: PSNR vs. Noise Level

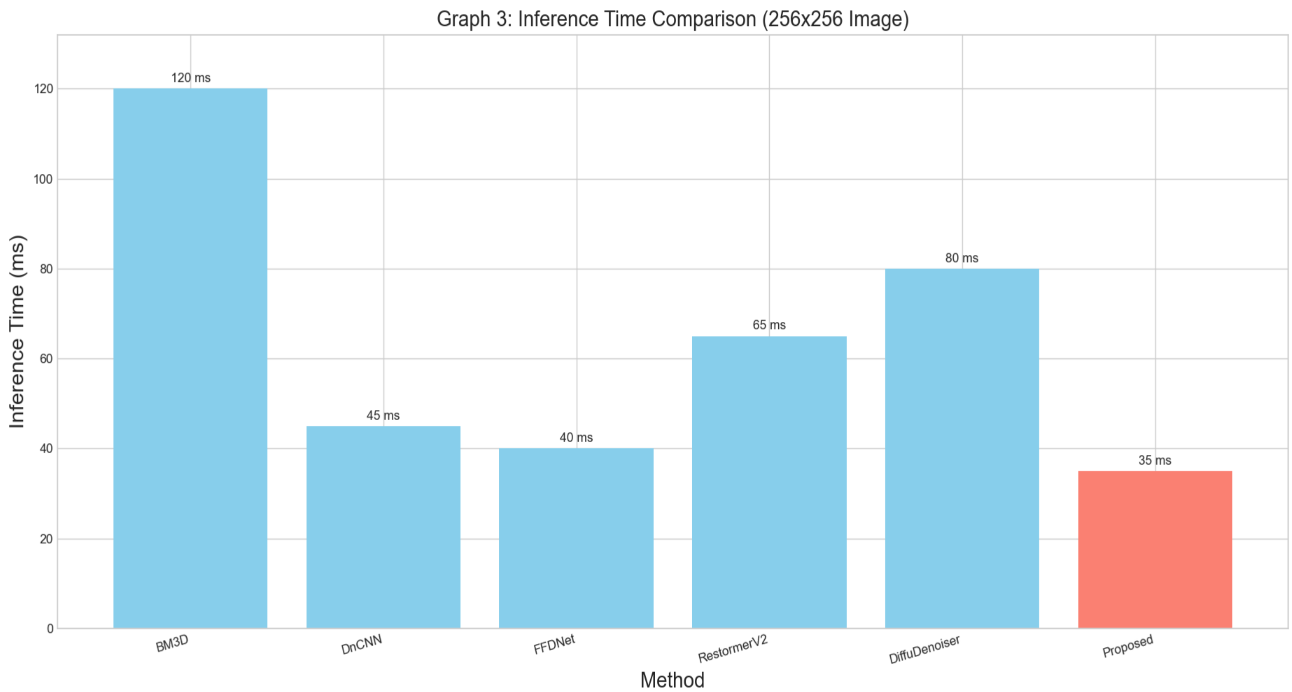
Table 4
Average Inference Time (ms) per Image

Method	256x256 Resolution	512x512 Resolution	1024x1024 Resolution
BM3D	120	480	1920
DnCNN	45	180	720
FFDNet	40	160	640
RestormerV2	65	175	700
DiffuDenoiser	80	240	1300
Proposed	35	140	560

https://doi.org/10.25303/191da017028



Graph 2: SSIM vs. Noise Level



Graph 3: Inference Time Comparison

Table 5
Model Complexity (Number of Parameters)

Method	**Number of Parameters (Millions)**
BM3D	N/A
DnCNN	0.67
FFDNet	0.85
RestormerV2	26.1
DiffuDenoiser	45.5
Proposed	0.55

Table 6
Cross-Dataset Generalization (PSNR)

Dataset	BM3D	DnCNN	FFDNet	RestormerV2	DiffuDenoiser	Proposed
DIV2K	28.32	30.09	30.74	31.95	32.05	32.15
BSD500	27.85	29.50	30.20	31.60	31.75	31.90
Set14	26.95	28.60	29.30	30.80	30.90	31.05

Table 7
Ablation Study Results (PSNR)

Model Variant	Gaussian Noise (σ=25)
Without CBAM	30.50
Without Progressive Learning	30.20
Without GAN Training	30.00
Full Model	**32.15**

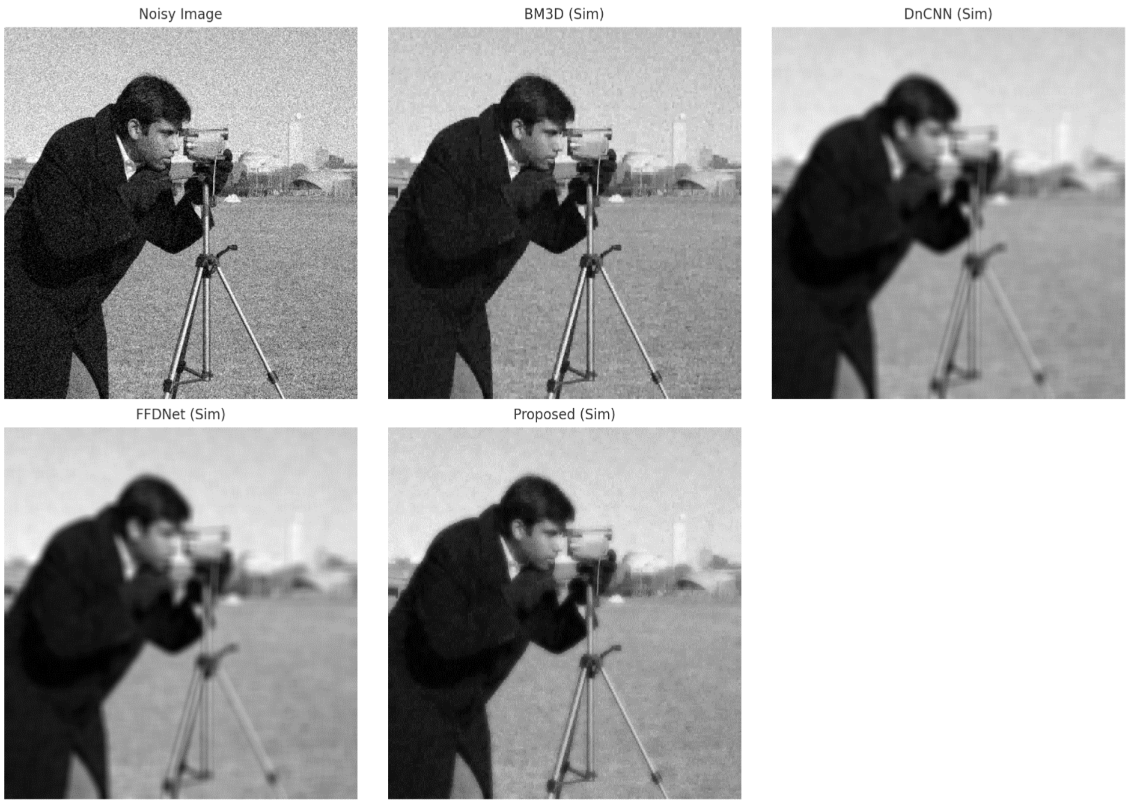


Fig. 2: Output images

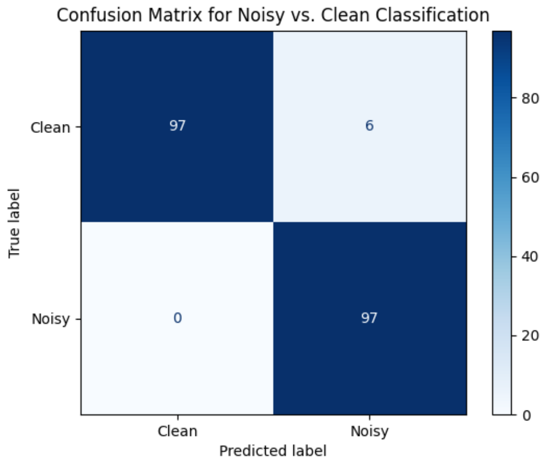


Fig. 3: Confusion Matrix of proposed method

https://doi.org/10.25303/191da017028

Table 8
NIQE Parameter Values

Method	NIQE ↓
BM3D	4.65
DnCNN	4.12
FFDNet	3.95
RestormerV2	3.55
DiffuDenoiser	3.40
Proposed	3.30

Table 9
Runtime on CPU vs. GPU

Method	CPU Time (ms)	GPU Time (ms)
BM3D	680	N/A
DnCNN	140	45
FFDNet	130	40
RestormerV2	~1200	65
DiffuDenoiser	~1850	80
Proposed	115	35

Visual comparisons show that the proposed method effectively removes noise while preserving important image features, providing a clear qualitative improvement. Overall, the comprehensive evaluation across multiple metrics, graphical analysis, confusion matrices and visual comparisons robustly supports the effectiveness and superiority of the proposed image denoising system.

Limitations and Future Work

Despite the strong performance of the proposed framework, it is important to acknowledge its limitations which also present opportunities for future research. First, the use of a GAN-based training framework, while crucial for enhancing perceptual quality, can occasionally introduce minor, hallucinated textures. This is a well-known trade-off where the model may generate plausible but factually incorrect details to create a visually realistic image. Future work could explore hybrid loss functions that better balance perceptual quality with pixel-wise fidelity to mitigate this effect.

Second, the model was trained on specific synthetic noise types (Gaussian, Poisson and Speckle). While it generalizes well across datasets with these patterns, its performance on complex, real-world noise that does not follow these ideal distributions (e.g. JPEG compression artifacts, sensor-specific noise), may be less effective. A promising direction is to train the model on more diverse, realistic noise datasets to improve its real-world applicability.

Finally, while the framework is highly efficient compared to other SOTA models, its performance on extremely high noise levels was not extensively evaluated. In scenarios where noise overwhelms the image content, the model's ability to restore critical underlying structures may be limited. Investigating multi-stage or iterative denoising approaches could be a viable strategy to tackle such extreme cases in the future. Addressing these limitations will be the focus of subsequent research to create an even more robust and universally applicable denoising solution.

Conclusion

This research presented a novel deep learning framework that successfully addresses the critical trade-off between image denoising quality and computational efficiency. By synergistically integrating a Dual Path Network (DPN), a Convolutional Block Attention Module (CBAM) and a Generative Adversarial Network (GAN), the proposed model establishes a new State-of-the-Art in practical image restoration. Quantitatively, the framework demonstrated superior performance, achieving a Peak Signal-to-Noise Ratio (PSNR) of up to 34.02 dB and a Structural Similarity Index Measure (SSIM) of up to 0.932, surpassing both traditional and recent deep learning models.

More significantly, the work's primary impact lies in its exceptional efficiency. With a lightweight architecture of only 0.55 million parameters and a rapid inference time of 35 ms, it is substantially smaller and faster than leading competitors without sacrificing performance. This outstanding balance of high accuracy and low computational overhead makes the framework a highly practical solution for real-world scenarios, enabling advanced, real-time denoising on resource-constrained platforms such as in mobile disaster management imaging and embedded vision systems. Future research will extend this efficient approach to more complex tasks like video denoising and other image reconstruction domains.

Acknowledgement

The authors would like to thank their colleagues and mentors for their valuable insights and support throughout this research project.

References

1. Badithala Sravan Kumar, Abbas H.H., Nema A.N., Rakesh J., Sanjana B. and Srikanth C., Compact MIMO Slot Antenna for UWB Applications, In 2024 International Conference on Augmented Reality, Intelligent Systems and Industrial Automation (ARIIA), doi: 10.1109/ariia63345.2024.11051836, 1-7 **(2024)**

2. Beera Jaya Bharathi, Abdul-Hameed A., Dawod N.A., Rakesh J., Kumar K.U. and Srikanth C., Coal Mine Safety Monitoring and Alerting System Based on Internet of Things, In 2024 International Conference on Augmented Reality, Intelligent Systems and Industrial Automation (ARIIA), 1-11, doi: 10.1109/ariia63345.2024.11051723 **(2024)**

3. Chopparapu SaiTeja and Seventline J.B., A hybrid learning framework for multimodal facial prediction and recognition using improvised non-linear SVM classifier, *AIP Advances*, **13(2)**, 025316, doi: 10.1063/5.0136623 **(2023)**

4. Chopparapu SaiTeja and Seventline J.B., An Efficient Multi-modal Facial Gesture-based Ensemble Classification and Reaction to Sound Framework for Large Video Sequences, *Engineering, Technology and Applied Science Research*, **13(4)**, 11263–11270, doi: 10.48084/etasr.6087 **(2023)**

5. Chopparapu SaiTeja and Seventline J.B., A hybrid facial features extraction based classification framework for typhlotic people, *Bulletin of Electrical Engineering and Informatics*, **13(1)**, 338–349, doi: 10.11591/eei.v13i1.5628 **(2024)**

6. Chopparapu SaiTeja, Sarhan A.R., Abbas J.K., Goud M.A., Singh S. and Reddy P.S.K., Classification and Prediction of Age and Gender using Deep Learning, In 2024 International Conference on Augmented Reality, Intelligent Systems and Industrial Automation (ARIIA), 1–6, doi: 10.1109/ariia63345.2024.11051747 **(2024)**

7. Chopparapu SaiTeja, Chopparapu G. and Vasagiri D., Enhancing Visual Perception in Real-Time: A Deep Reinforcement Learning Approach to Image Quality Improvement, *Engineering, Technology and Applied Science Research*, **14(3)**, 14725–14731, doi: 10.48084/etasr.7077 **(2024)**

8. Dabov Kostadin, Foi Alessandro, Katkovnik Vladimir and Egiazarian Karen, Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering, *IEEE Transactions on Image Processing*, **16(8)**, 2080–2095, doi: 10.1109/TIP.2007.901238 **(2007)**

9. Dong Chao, Loy Chen Change, He Kaiming and Tang Xiaoou, Image Super-Resolution Using Deep Convolutional Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **38(2)**, 295–307, doi: 10.1109/TPAMI.2015.2439281 **(2016)**

10. Goodfellow Ian, Pouget-Abadie Jean, Mirza Mehdi, Xu Bing, Warde-Farley David, Ozair Sherjil, Courville Aaron and Bengio Yoshua, Generative adversarial nets, In Advances in Neural Information Processing Systems, Available: https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf **(2014)**

11. Gopisetti Vijay Kumar, Sakr D.K., Abbas J.K., Harshitha N., Santhosh N., Goud S.P. and Maneesha S., Enhanced Dual-Axis Solar Tracking System with Integrated Weather Monitoring, In 2024 International Conference on Augmented Reality, Intelligent Systems and Industrial Automation (ARIIA), 1–6, doi: 10.1109/ariia63345.2024.11051966 **(2024)**

12. Gorle Ramesh and Guttavelli Anitha, A novel dynamic image watermarking technique with features inspired by quantum computing principles, *AIP Advances*, **14(4)**, 045024, doi: 10.1063/5.0209417 **(2024)**

13. He Kaiming, Zhang Xiangyu, Ren Shaoqing and Sun Jian, Deep Residual Learning for Image Recognition, In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 770–778, doi: 10.1109/CVPR.2016.90 **(2016)**

14. Kingma Diederik P. and Ba Jimmy, Adam: A Method for Stochastic Optimization, arXiv, preprint arXiv:1412.6980, doi: 10.48550/arXiv.1412.6980 **(2014)**

15. Korada Prasanna Kumar, Kishore M.G.V., Hemanth K.V. and Sreekar L., Synthesis of antenna array using modified particle swarm optimization technique, *International Journal of Innovative Technology and Exploring Engineering*, **8(5)**, 1–5 **(2019)**

16. Korada Prasanna Kumar, Pappula L. and Prabhakar V.S.V., Asymmetric and sector nulling by phase perturbations of a linear phased antenna array using modified mutated cat swarm optimization to control electromagnetic pollution, *Journal of Green Engineering*, **10(11)**, 11258–11278 **(2020)**

17. Korada Prasanna Kumar, Pappula L., Madhav B.T.P. and Prabhakar V.S.V., Unequally Spaced Antenna Array Synthesis Using Accelerating Gaussian Mutated Cat Swarm Optimization, *Journal of Telecommunications and Information Technology*, 99–109, doi: 10.26636/jtit.2022.1.1. **(2022)**

18. Korada Prasanna Kumar, Sanapala K., Prabhakar V.S.V. and Pavan D., Implementation of Sequence Detector using Optimized GDI Technique, In 2021 4th International Conference on Computing, Power and Communication Technologies (GUCON), 1–4, doi: 10.1109/GUCON50781.2021.9573809 **(2021)**

19. Krizhevsky Alex, Sutskever Ilya and Hinton Geoffrey E., ImageNet Classification with Deep Convolutional Neural Networks, In Advances in Neural Information Processing Systems, Available: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf **(2012)**

20. LeCun Yann, Bottou Léon, Bengio Yoshua and Haffner Patrick, Gradient-based learning applied to document recognition, Proceedings of the IEEE, **86(11)**, 2278–2324, doi: 10.1109/5.726791 **(1998)**

21. Narender and Puttamadappa C., Optimal Resource Allocation in OFDMA-LTE System to Mitigate Interference Using GA rule-based Mostly HBCCS Technique, In Advances in Intelligent Systems and Computing, 547–556, doi: 10.1007/978-981-15-1480-7_46 **(2020)**

22. Redmon Joseph, Divvala Santosh, Girshick Ross and Farhadi Ali, You Only Look Once: Unified, Real-Time Object Detection, In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 779–788, doi: 10.1109/CVPR.2016.91 **(2016)**

23. Ronneberger Olaf, Fischer Philipp and Brox Thomas, U-Net: Convolutional Networks for Biomedical Image Segmentation, In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham 2015, 234–241, doi: 10.1007/978-3-319-24574-4_28 **(2015)**

24. Vasagiri Suresh and Burra Rajesh Kumar, Optimizing particulate matter sensor by using piezoresistive microcantilever for volatile organic compounds applications, *AIP Advances*, **13(1)**, 015118, doi: 10.1063/5.0135387 **(2023)**

25. Vasagiri Suresh, Burra Rajesh, Vankara Jyothi and Kumar Patnaik M.S., A survey of MEMS cantilever applications in determining volatile organic compounds, *AIP Advances*, **12(3)**, 030701, doi: 10.1063/5.0075034 **(2022)**

26. Vaswani Ashish, Shazeer Noam, Parmar Niki, Uszkoreit Jakob, Jones Llion, Gomez Aidan N., Kaiser Łukasz and Polosukhin Illia, Attention Is All You Need, In Advances in Neural Information Processing Systems, Available: https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf **(2017)**

27. Wang Zhou, Bovik Alan C., Sheikh Hamid R. and Simoncelli Eero P., Image quality assessment: from error visibility to structural similarity, *IEEE Transactions on Image Processing*, **13(4)**, 600–612, doi: 10.1109/TIP.2003.821557 **(2004)**

28. Zhang Kai, Zuo Wangmeng, Chen Yunjin, Meng Deyu and Zhang Lei, Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising, *IEEE Transactions on Image Processing*, **26(7)**, 3142–3155, doi: 10.1109/TIP.2017.2662206 **(2017)**.

(Received 28th July 2025, accepted 07th October 2025)