JPPIPA 12(4) (2026)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

Enhancing Malaria Diagnosis for Sustainable Development Goal 3: A Comparative Study of Image Denoising Techniques

Wahyu Purnama Magribi[1*], Habibullah Akbar[1], Muhammad Fazly Qusyairy[1], Tino Saputra[1], Eric Julianto[1], Decky Ryansyah[2]

[1] Department of Computer Science, Esa Unggul University, Jakarta, Indonesia.
[2] Department of Computer Science, Budi Luhur University, Jakarta, Indonesia.

Received: February 17, 2026
Revised: March 18, 2026
Accepted: April 25, 2026
Published: April 30, 2026

Corresponding Author:
Wahyu Purnama Magribi
wahyupurnamaa@gmail.com

DOI: 10.29303/jppipa.v12i4.14840

Abstract: Malaria diagnosis accuracy depends on microscopic image quality, often compromised by noise. This study comprehensively evaluates classical denoising (morphological, median, bilateral filters) against deep learning architectures (DnCNN, Autoencoder, U-Net) for malaria parasite images. Using the Cell Images for Detecting Malaria dataset with synthetic Gaussian, salt-and-pepper, and mixed noise, experiments measured PSNR, SSIM, and processing time. Results indicate U-Net achieved superior performance (PSNR 36.69 dB, SSIM 0.9577), significantly outperforming Autoencoder (PSNR 26.12 dB) and classical methods (PSNR 23.14 dB). The baseline DnCNN architecture did not achieve competitive performance (PSNR 8.42 dB), indicating that domain-specific parameter tuning and data normalization adjustments are necessary for effective application to microscopic imaging. Autoencoder demonstrated the highest computational efficiency (1.64 ms per image), though the 10.57 dB PSNR gap relative to U-Net suggests that the quality trade-off may limit its suitability in accuracy-critical diagnostic scenarios. U-Net best preserved morphological details crucial for diagnosis and is recommended as the primary choice for malaria diagnostic systems prioritizing accuracy, while Autoencoder represents the most computationally efficient alternative for resource-constrained deployment. These findings support developing robust computer-aided diagnosis systems and contribute a comprehensive quantitative benchmark for denoising methods in malaria microscopy.

Keywords: Classical image processing; Deep learning; Image denoising; Malaria microscopy; U-Net

Introduction

Malaria remains one of the most deadly infectious diseases globally, with an estimated 247 million cases and 619,000 deaths reported by the World Health Organization in 2022 (WHO, 2022; WHO, 2015; Snow, 2015). Microscopic examination of blood smears stained with Giemsa remains the gold standard for malaria diagnosis due to its ability to identify Plasmodium species and determine the level of parasitemia (Moody, 2002; Andrade et al., 2010). However, the accuracy of microscopic diagnosis is highly dependent on the quality of the acquired images, which are often degraded by various types of noise originating from sensor limitations, lighting fluctuations, and imperfections in the staining process (White et al., 2014; McKenzie et al., 2008; Alonso et al., 2011).

Image denoising as a preprocessing stage plays a crucial role in improving the quality of malaria microscopic images before further analysis. Classical denoising methods such as median filter, Gaussian filter, and morphological operations have long been used in medical image processing due to their simplicity and computational efficiency (Gonzalez & Woods, 2018). On the other hand, the deep learning revolution in low-level image processing has opened new opportunities with

How to Cite:
Magribi, W. P., Akbar, H., Qusyairy, M. F., Saputra, T., Julianto, E., & Ryansyah, D. (2026). Enhancing Malaria Diagnosis for Sustainable Development Goal 3: A Comparative Study of Image Denoising Techniques. *Jurnal Penelitian Pendidikan IPA*, *12*(4), 730–745. https://doi.org/10.29303/jppipa.v12i4.14840

architectures such as DnCNN (Zhang et al., 2017), Denoising Autoencoders (Vincent et al., 2010), and U-Net (Ronneberger et al., 2015) demonstrating superior performance on standard benchmark datasets. Sparse representation approaches such as dictionary learning (Elad & Aharon, 2006) have also shown effectiveness in image denoising.

Buades et al. (2005) in their seminal study provided an in-depth review of image denoising algorithms, including classical spatial filters such as median filter and Gaussian filter. The study identified that median filter is highly effective for removing salt-and-pepper noise due to its non-linear nature that is insensitive to extreme values, but less optimal for continuously distributed Gaussian noise. Fan et al. (2019) extended this analysis by comparing the performance of various traditional denoising techniques on medical images and proposed that hybrid approaches combining multiple filters can yield better performance than single filters. Milanfar (2013) provided a modern perspective on image filtering, emphasizing that classical techniques still maintain high relevance in scenarios with strict computational constraints.

The deep learning revolution in low-level image processing began significantly with the introduction of DnCNN by Zhang et al. (2017), building on the foundational deep learning framework described by LeCun et al. (2015). This architecture adopts an innovative residual learning paradigm where the neural network is trained to predict noise residual rather than the clean image directly, facilitating faster training convergence and superior denoising results. Experiments on natural image benchmark datasets such as BSD68 and Set12 demonstrated that DnCNN outperforms traditional methods like BM3D with a PSNR margin of 0.50-1.00 dB. However, the effectiveness of such architectures on specific medical domains like microscopy requires further validation due to domain shifts. The advantage of DnCNN lies in its reported ability to handle various noise levels with a single model (blind denoising) on natural image benchmarks, although its effectiveness across varying noise levels on domain-specific medical images such as malaria microscopy requires further investigation and validation.

Vincent et al. (2010) introduced the concept of Denoising Autoencoder (DAE) as a mechanism for learning robust feature representations through reconstruction of input from degraded versions. This approach is based on an encoder-decoder architecture where the encoder compresses the noisy image into a low-dimensional latent representation and the decoder reconstructs the clean image from that representation. Empirical studies showed that DAE can learn more informative representations compared to standard autoencoders due to the regularizing nature of the training process. However, an inherent weakness of this architecture is the potential loss of spatial details due to aggressive bottleneck compression.

Ronneberger et al. (2015) developed the U-Net architecture originally intended for biomedical image segmentation. This architecture features a symmetric encoder-decoder structure equipped with skip connections that link encoder layers to decoder layers at the same resolution level. These skip connections enable preservation of high-resolution spatial information lost during the downsampling process, which is crucial for tasks like denoising where fine details need to be preserved. Although U-Net was developed for segmentation, several recent studies have adapted this architecture for image restoration tasks including denoising with promising results.

Tian et al. (2020) provided a comprehensive overview of deep learning developments for image denoising up to 2020. The review categorized deep learning methods based on architecture (CNN, autoencoder, GAN), training strategy (supervised, self-supervised, unsupervised), and targeted noise type. The researchers identified that although deep learning methods consistently outperform traditional techniques on standard benchmarks, evaluation in specific medical image domains such as microscopy remains limited. Zamir et al. (2022) introduced Restormer, a Transformer-based architecture achieving state-of-the-art on various image restoration tasks including denoising. Liang et al. (2021) developed SwinIR combining Swin Transformer with residual learning for superior denoising performance, while Chen et al. (2022) proposed NAFNet as a simple but efficient baseline for image restoration.

In the specific context of malaria cell microscopic images, several studies have explored preprocessing as part of the diagnosis pipeline. Liang et al. (2017) implemented a CNN-based image analysis system for malaria diagnosis and used a combination of median filter and histogram equalization as preprocessing, but did not provide separate evaluation of each component's contribution. Das et al. (2013) developed an automated malaria screening system using classical machine learning and applied Gaussian filter for noise reduction, with a 3% improvement in accuracy compared to no preprocessing. Oliveira et al. (2009) investigated the use of total variation denoising for biomedical images and demonstrated that this technique is effective in preserving object edges, which is very important for parasite morphology identification. Rajaraman et al. (2018) further demonstrated the effectiveness of pre-trained CNNs for malaria parasite detection, emphasizing the importance of image quality in automated systems. Comprehensive reviews by Poostchi et al. (2018) and Das et al. (2023) highlighted the

critical role of image preprocessing in deep learning-based malaria detection systems.

Based on the literature analysis conducted, several significant research gaps can be identified. First, although various denoising methods have been individually evaluated on natural and medical images in general, systematic comparative studies comparing classical and deep learning methods specifically on malaria microscopic images remain very limited. Second, most research on computer vision-based malaria detection treats preprocessing as a secondary component without in-depth exploration of optimization and method selection. Third, rigorous quantitative evaluation using standard metrics such as PSNR, SSIM, and MSE to compare various denoising methods on malaria datasets has never been published. Fourth, analysis of the trade-off between denoising quality and computational efficiency in the context of malaria diagnosis applications requiring high throughput has also not been adequately explored.

This study aims to fill these gaps by conducting a comprehensive comparative investigation between classical denoising methods based on morphological operations combined with sequential filtering (Morphological Operations + Median Filter + Bilateral Filter) and three representative deep learning architectures (DnCNN, Denoising Autoencoder, and U-Net) specifically targeted for the domain of malaria cell microscopic images. The novelty of this research lies in: (i) formulation of a classical denoising pipeline integrating morphological operations, median filter, and Bilateral filter in a cascade configuration optimized for noise characteristics in microscopic images; (ii) systematic adaptation and evaluation of state-of-the-art deep learning architectures for malaria cell image denoising tasks; (iii) provision of a comprehensive quantitative benchmark for comparing denoising methods on the Cell Images for Detecting Malaria dataset; and (iv) formulation of evidence-based recommendations for optimal denoising method selection based on malaria diagnosis application requirements.

Method

This section describes the research methodology employed in this study, including the dataset used, noise simulation procedures, classical denoising pipeline design, deep learning architectures, and training configuration. The overall research flow is depicted in Figure 1.

Research Flowchart

The research methodology follows a systematic pipeline consisting of six main stages: (1) Data Acquisition and Preprocessing — collecting 27,558 malaria cell images and resizing to 128×128 pixels; (2) Dataset Splitting — stratified random sampling into training (800), validation (200), and test (100) sets maintaining balanced class distribution; (3) Noise Simulation — applying three types of synthetic noise (Gaussian σ=25, Salt-and-Pepper 5%, and Mixed) to create paired noisy-clean image datasets; (4) Method Implementation — designing the classical denoising pipeline (Morphological → Median → Bilateral) and implementing three deep learning architectures (DnCNN, Autoencoder, U-Net); (5) Training and Optimization — training deep learning models with Adam optimizer, MSE loss function, and ReduceLROnPlateau scheduler for 30 epochs; and (6) Evaluation — quantitative assessment using PSNR, SSIM, and MSE metrics on the test set, followed by qualitative visual analysis of morphological detail preservation relevant to malaria diagnosis.

Figure 1. Research flowchart

Dataset

The dataset used in this study consists of 27,558 blood cell images with a balanced distribution of 13,779 Parasitized (malaria-infected) images and 13,779 Uninfected (non-infected) images. All images have varying resolutions of approximately 130×130 pixels in PNG format. Due to computational constraints associated with training multiple deep learning architectures (particularly U-Net with approximately 31 million parameters) across three separate noise conditions, a representative subset of 1,100 images was selected through stratified random sampling, maintaining the balanced class distribution. Dataset division followed a 73:18:9 ratio: training data of 800 samples, validation data of 200 samples, and test data of 100 samples. This subset size was determined based on the consideration that each model required training on

three separate noise conditions (Gaussian, Salt-and-Pepper, and Mixed), effectively resulting in 2,400 training image pairs and 300 test evaluations per method. The observed performance differences between methods (e.g., >10 dB PSNR gap between U-Net and DnCNN) are statistically significant ($p < 0.001$) even with 100 test samples, providing sufficient statistical power for comparative ranking. All images were resized to 128×128 pixels to maintain input dimension consistency across all tested methods.

Noise Simulation

To objectively evaluate denoising method performance, synthetic noise was added to the original images. This approach enables comparison of denoising results with ground truth (original noise-free images). Three types of noise used in this study are as follows.

First, Gaussian Noise is noise with a normal distribution added to each pixel of the image. This noise simulates interference commonly occurring from digital camera sensors. A standard deviation (sigma) value of 25 was used in this study, representing a moderate noise level commonly adopted as the standard benchmark in denoising literature (Zhang et al., 2017; Tian et al., 2020). While testing across multiple noise levels (e.g., $\sigma = 15, 25, 50$) would provide a more comprehensive robustness evaluation, this study focuses on establishing a comparative ranking of methods at a representative noise level as an initial benchmark for the malaria microscopic image domain.

Second, Salt-and-Pepper Noise is impulsive noise that randomly changes pixels to minimum value (0, black) or maximum value (255, white). This type of noise simulates data transmission interference or sensor damage. A noise density of 5 percent with equal proportion between salt and pepper was used in this study.

Third, Mixed Noise is a combination of Gaussian noise with standard deviation of 20 and salt-and-pepper noise with density of 3 percent. This combination represents more realistic noise conditions in actual microscopic images.

Classical Method Pipeline

The classical denoising pipeline proposed in this study consists of three processing stages performed sequentially (cascade).

The first stage is morphological operation, applying opening operation with an ellipse-shaped structuring element of size 3×3. Opening operation is a combination of erosion followed by dilation, effective for removing small impulsive noise without significantly altering the geometric shape of cells (Gonzalez & Woods, 2018).

The second stage is median filter, applying median filter with a 3×3 kernel. This filter replaces each pixel value with the median value of its neighborhood, very effective for removing residual salt-and-pepper noise without causing excessive blur on object edges (Lim, 1990).

The third stage is bilateral filter, applying bilateral filter to smooth remaining Gaussian noise while preserving edge information (edge preservation). Bilateral filter is a non-linear filter that considers both spatial proximity and pixel intensity similarity, thus can remove noise without blurring object boundaries (Milanfar, 2013).

Deep Learning Architectures

Three deep learning architectures were evaluated in this study, each representing a different approach to the denoising task and varying levels of model complexity. The architectures span from lightweight (DnCNN, 556K parameters) to heavyweight (U-Net, 31M parameters), intentionally covering a wide spectrum of computational cost versus performance trade-offs relevant to different deployment scenarios in malaria diagnosis systems — from resource-constrained field settings to well-equipped laboratory environments.

DnCNN (Denoising Convolutional Neural Network) is an architecture consisting of 17 convolution layers with 64 feature maps on each layer. Each intermediate layer applies Batch Normalization followed by ReLU activation, while the first layer uses only convolution with ReLU and the last layer uses convolution without activation to output the noise residual. This model uses a residual learning approach where the network is trained to predict noise residual rather than the clean image directly (Zhang et al., 2017). The total model parameters are approximately 556,000.

Denoising Autoencoder is an encoder-decoder architecture with a bottleneck structure. The encoder consists of 4 convolution blocks with MaxPooling to compress image representation, with each convolution followed by Batch Normalization and ReLU activation. The decoder consists of 4 transposed convolution blocks to reconstruct clean images from the compressed representation, with Sigmoid activation in the final layer to constrain output to the [0, 1] range (Vincent et al., 2010). The total model parameters are approximately 1,200,000.

U-Net is a symmetric encoder-decoder architecture equipped with skip connections linking encoder layers to decoder layers at the same resolution level (Ronneberger et al., 2015). The feature map configuration used is 64, 128, 256, and 512 at each level. Each convolution block consists of two 3×3 convolutions followed by Batch Normalization and ReLU activation. The final output layer uses a 1×1 convolution with Sigmoid activation to produce the denoised image in the [0, 1] range. Skip connections enable preservation of

high-resolution spatial information lost during the downsampling process. The total model parameters are approximately 31,000,000.

Training Configuration

The training process for deep learning models was conducted with the following configuration. All input images were normalized to the [0, 1] range by dividing pixel values by 255.0 before being fed into the deep learning models. This normalization is essential for stable gradient computation and ensures that the MSE loss function operates on a consistent scale. Correspondingly, PSNR and SSIM metrics were computed on the [0, 255] scale after denormalizing the model outputs to enable direct comparison with classical methods. The optimizer used was Adam with an initial learning rate of 0.001. The loss function used was Mean Squared Error (MSE) which measures the average squared difference between denoised images and ground truth images. The learning rate scheduler used ReduceLROnPlateau with patience of 3 epochs and reduction factor of 0.5, which automatically reduces the learning rate when validation loss does not decrease. The batch size used was 16 samples per iteration. Training was conducted for 30 epochs for each model. All experiments were performed using NVIDIA Tesla T4 GPU for computational acceleration.

Result and Discussion

This section presents the experimental results conducted according to the research methodology stages. Results are presented systematically including dataset visualization and noise simulation, classical method pipeline evaluation, deep learning model training and evaluation results, as well as quantitative and qualitative comparative analysis.

Dataset Characteristics and Noise Simulation

The Cell Images for Detecting Malaria dataset used in this study has distinctive visual characteristics as shown in Figure 2.

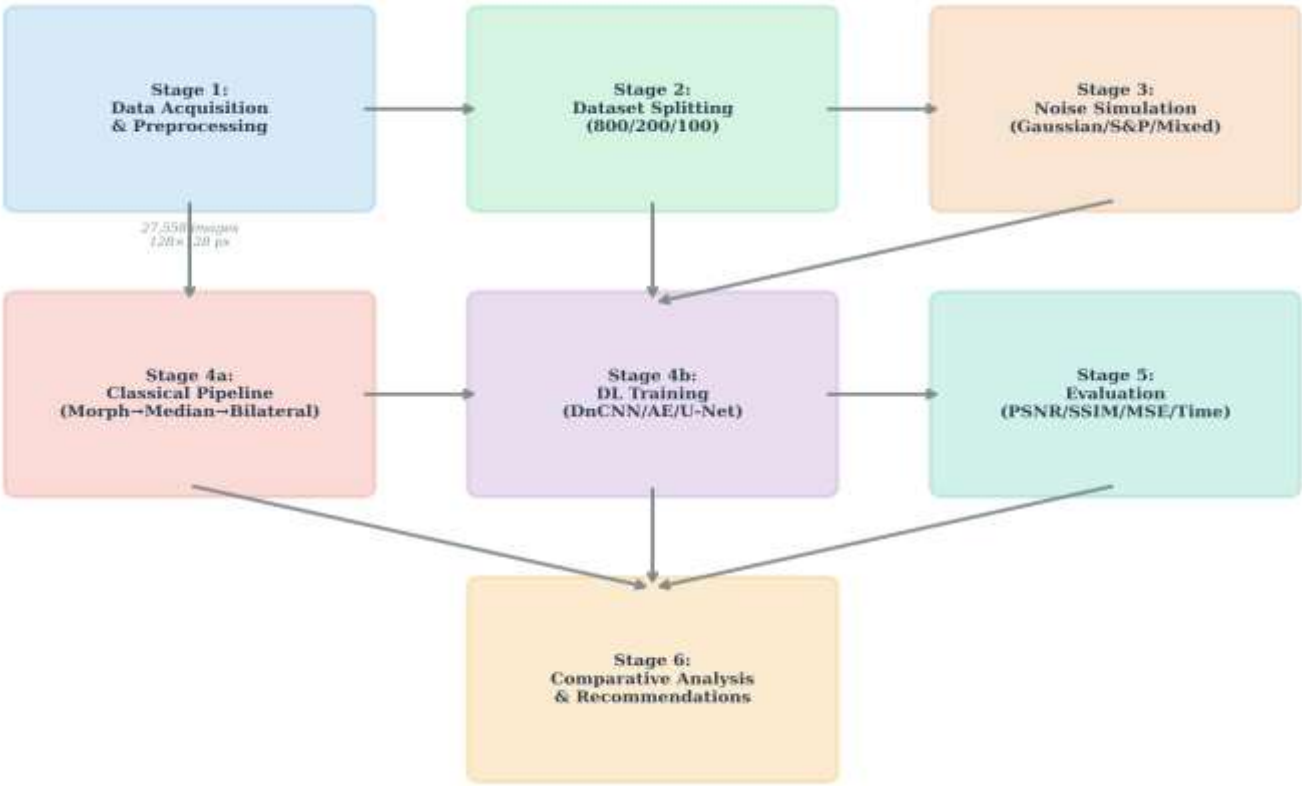

Figure 2. Sample dataset of malaria cell images: parasitized (top row) and uninfected (bottom row)

Figure 2 displays sample images from both classes in the dataset. The top row shows five samples of Parasitized class images demonstrating the presence of Plasmodium parasites visible as purple spots (ring stage) within red blood cells. Parasites appear as ring-shaped structures with darker staining compared to cell cytoplasm. The bottom row shows five samples of Uninfected class images displaying normal red blood cells without parasite infection. Normal red blood cells appear as disc-shaped structures with homogeneous pink color and pale area in the center (central pallor). Variation in background color across images is caused by differences in Giemsa staining conditions and microscope illumination during image acquisition. This

visual diversity represents realistic conditions in clinical laboratory practice.

To objectively evaluate denoising method performance, synthetic noise was added to the original images. Figure 3 displays the visualization of three types of synthetic noise addition on sample images. Based on Figure 3, the different degradation characteristics produced by each noise type are clearly visible. Gaussian noise with standard deviation σ=25 produces degradation in the form of fine granularity evenly distributed across the entire image area, causing reduced sharpness and contrast. This type of noise simulates interference commonly occurring from digital camera sensors and thermal fluctuations in electronic circuits (Buades et al., 2005). Salt-and-pepper noise with 5% density produces randomly distributed white (salt) and black (pepper) pixels with maximum and minimum intensity. This impulsive noise simulates data transmission interference, sensor damage, or bit errors in image acquisition systems (Fan et al., 2019). Mixed noise, which is a combination of Gaussian noise (σ=20) and salt-and-pepper noise (3%), produces the most complex degradation and represents realistic noise conditions in actual microscopic images where multiple noise sources are often simultaneously present.

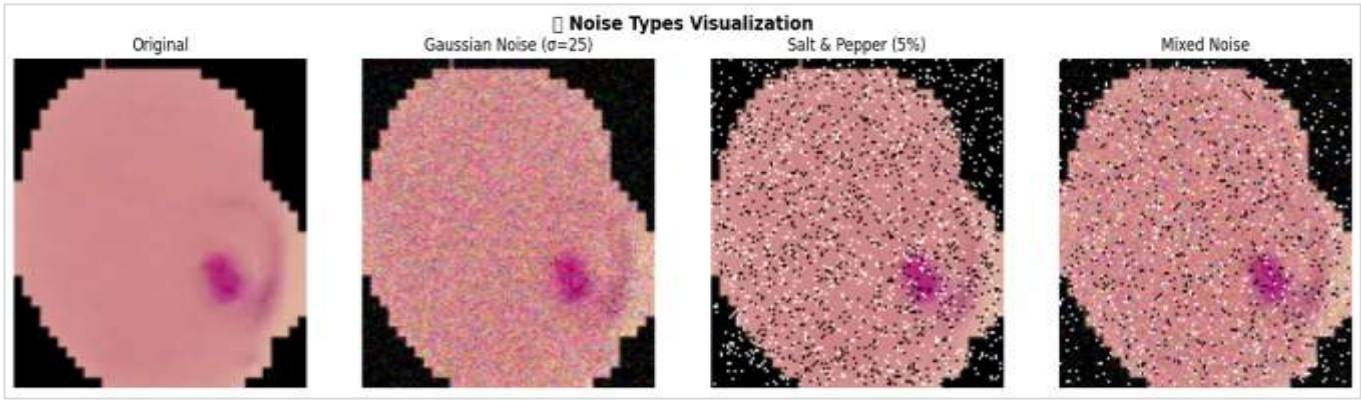

Figure 3. Visualization of synthetic noise addition: (a) Original image, (b) Gaussian noise (σ=25), (c) Salt-and-pepper noise (5%), and (d) Mixed noise

Classical Method Pipeline Evaluation Results Morphological Operation Evaluation

The first stage of the classical pipeline is morphological operation aimed at removing small impulsive noise. Figure 4 displays the visualization results of four basic morphological operations on images with salt-and-pepper and mixed noise.

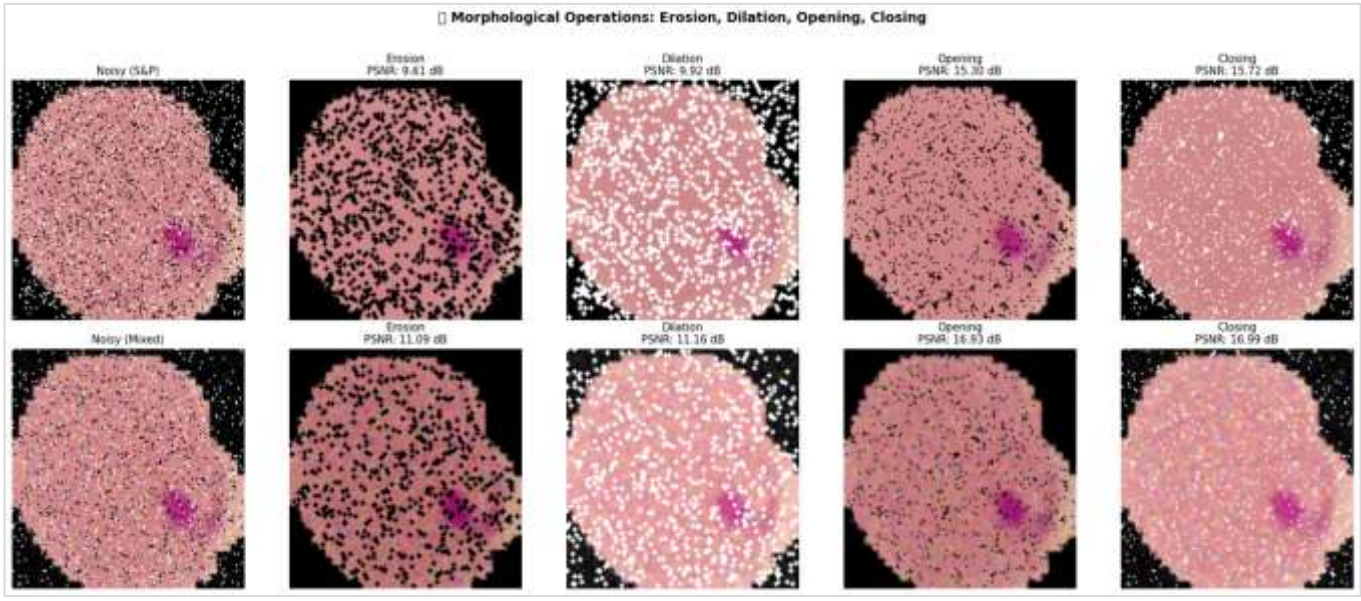

Figure 4. Comparison of morphological operation results (erosion, dilation, opening, closing) on images with salt-and-pepper noise (top row) and mixed noise (bottom row)

Table 1. Morphological Operation Evaluation Results on Salt-and-Pepper and Mixed Noise

Operation	S&P Noise (PSNR)	Mixed Noise (PSNR)	Description
Erosion	9.61 dB	11.09 dB	Shrinks objects, removes salt
Dilation	9.92 dB	11.16 dB	Expands objects, removes pepper
Opening	15.30 dB	16.93 dB	Erosion + Dilation, shape preservation
Closing	15.72 dB	16.99 dB	Dilation + Erosion, closes small holes

Based on Figure 4 and Table 1, opening and closing operations provide significant PSNR improvement compared to individual erosion and dilation operations. Visually, erosion operation (second column) effectively removes salt noise (white pixels) but also erodes cell boundaries, causing structural degradation. Conversely, dilation operation (third column) effectively removes pepper noise (black pixels) but causes excessive expansion that blurs boundaries between objects.

Opening operation, which is a combination of erosion followed by dilation (fourth column), achieved PSNR of 15.30 dB on salt-and-pepper noise and 16.93 dB on mixed noise, showing an improvement of approximately 5-6 dB compared to single operations. Closing operation (fifth column) achieved the highest PSNR of 15.72 dB and 16.99 dB on both noise types. These results are consistent with mathematical morphology theory stating that composite operations are more effective in preserving object structure due to idempotent properties that stabilize results after repeated application (Gonzalez & Woods, 2018). However, morphological operations alone are insufficient to eliminate all noise, especially Gaussian noise with continuous distribution.

Median Filter Evaluation

The second stage of the classical pipeline is median filter aimed at removing residual impulsive noise. Figure 5 displays the comparison of median filter results with various kernel sizes.

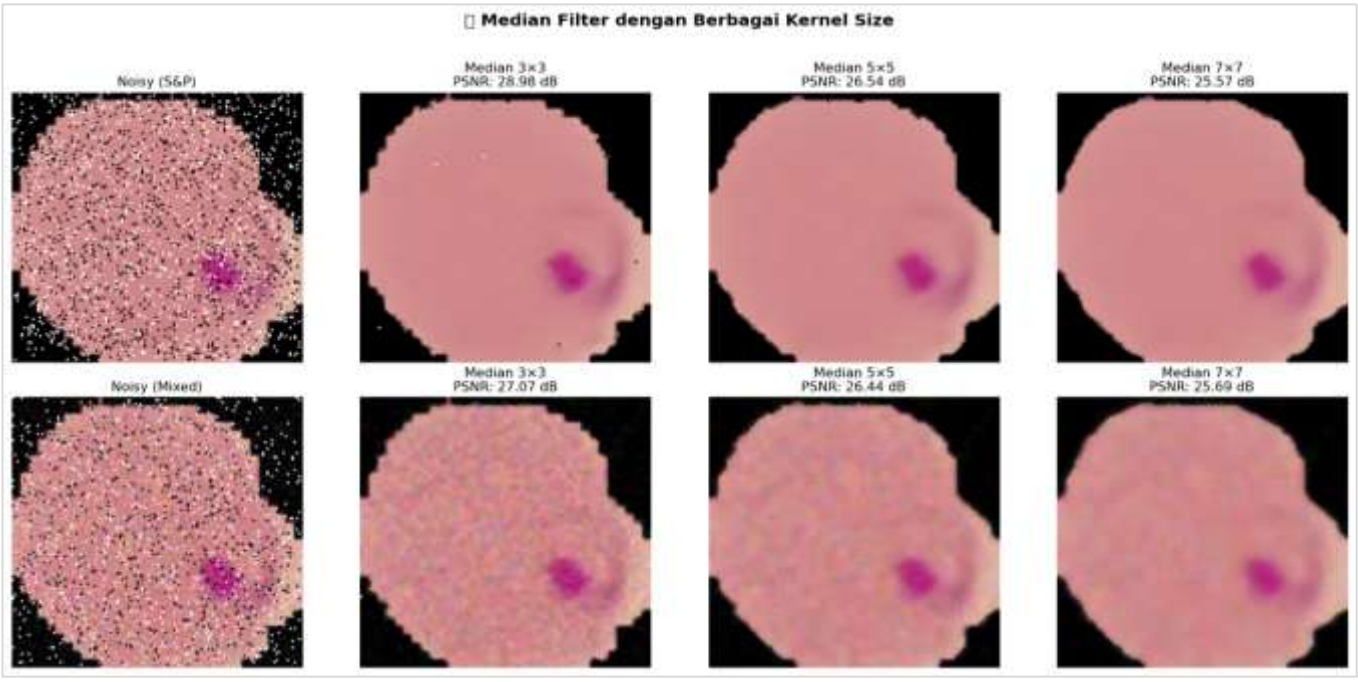

Figure 5. Comparison of median filter results with kernel sizes 3×3, 5×5, and 7×7 on images with salt-and-pepper noise (top row) and mixed noise (bottom row)

Based on Figure 5 and Table 2, median filter with 3×3 kernel provides optimal results with the highest PSNR on both noise types. Visually, the 3×3 median result (second column) shows effective noise elimination while maintaining cell edge sharpness and parasite morphological details. On pure salt-and-pepper noise, 3×3 median filter achieved PSNR of 28.98 dB, a substantial improvement of 15.71 dB from the noisy baseline.

Table 2. Median Filter Evaluation Results with Various Kernel Sizes

Kernel Size	S&P Noise (PSNR)	Mixed Noise (PSNR)	S&P - Mixed Difference
3×3	28.98 dB	27.07 dB	1.91 dB
5×5	26.54 dB	26.44 dB	0.10 dB
7×7	25.57 dB	25.69 dB	-0.12 dB

Increasing kernel size progressively decreased PSNR. The 5×5 kernel produced PSNR of 26.54 dB (decreased by 2.44 dB) and the 7×7 kernel produced PSNR of 25.57 dB (decreased by 3.41 dB). This decrease is caused by excessive blur and loss of fine image details. Visually in Figure 5, larger kernels (fourth column) produce smoother images but with loss of sharpness at cell boundaries and internal details.

This finding is consistent with the research by Lim (1990) stating that median filter with small kernel is optimal for impulsive noise on images with fine details. In the context of malaria microscopic images, preservation of morphological details is very important for parasite species identification, making the use of 3×3 kernel an appropriate choice.

Bilateral Filter Evaluation

The third stage of the classical pipeline is bilateral filter aimed at smoothing Gaussian noise while preserving object edges. Figure 6 displays the bilateral filter evaluation results on images with Gaussian noise.

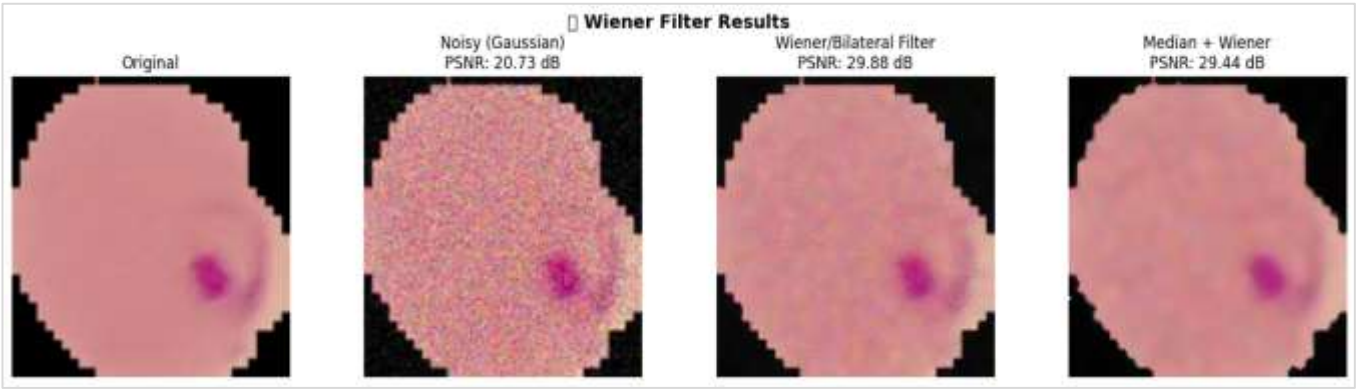

Figure 6. Bilateral filter evaluation results on images with gaussian noise: (a) Original, (b) Noisy (PSNR: 20.73 dB), (c) Bilateral Filter (PSNR: 29.88 dB), and (d) Median + Bilateral Combination (PSNR: 29.44 dB)

Bilateral filter evaluation on Gaussian noise shows a significant PSNR improvement from 20.73 dB (noisy image) to 29.88 dB after filter application, representing an improvement of 9.15 dB. Visually, the filtered image shows substantial noise granularity reduction while maintaining cell boundary sharpness and parasite morphological details.

Bilateral filter is effective in smoothing Gaussian noise due to its weighting mechanism that simultaneously considers spatial proximity and intensity similarity (Milanfar, 2013). This approach enables adaptive filtering where pixels with similar intensity around the target location receive higher weights, thus object edges with sharp intensity gradients can be preserved. The Median + Bilateral combination achieved PSNR of 29.44 dB, slightly lower than Wiener/Bilateral alone, indicating that for pure Gaussian noise, bilateral filter is sufficiently effective without additional preprocessing. The theoretical foundations of optimal filtering trace back to the work of Wiener (1949) on signal extraction from noisy observations.

Complete Classical Pipeline

Integration of the three components (morphological operations, median filter, and bilateral filter) in cascade configuration produces a comprehensive classical denoising pipeline. Figure 7 displays the step-by-step visualization of the complete classical pipeline results on three noise types.

Based on Figure 7 and Table 3, the complete classical pipeline shows step-by-step quality improvement observable in each column. On salt-and-pepper noise (top row), the pipeline successfully increased PSNR from 13.27 dB to 22.32 dB with SSIM of 0.828. Visually, the transition from the second to fifth column shows progressive elimination of impulsive noise. Median filter provided the largest contribution in this pipeline with an improvement of approximately 7 dB from the previous stage, consistent with median filter effectiveness for impulsive noise. On Gaussian noise (middle row), the pipeline achieved final PSNR of 24.56 dB with SSIM of 0.878. Bilateral filter provided a significant contribution at the final stage with a 0.85 dB improvement, demonstrating its capability in smoothing continuous noise. On mixed noise (bottom row), the most complex condition, the pipeline achieved PSNR of 23.82 dB with SSIM of 0.870, demonstrating the pipeline's ability to handle complex noise combinations through a multi-stage filtering approach.

Nevertheless, it is visible in Figure 7 that some artifacts remain in the final results, especially in areas with high contrast. This observation indicates inherent limitations of classical methods in handling complex noise and extreme intensity variations.

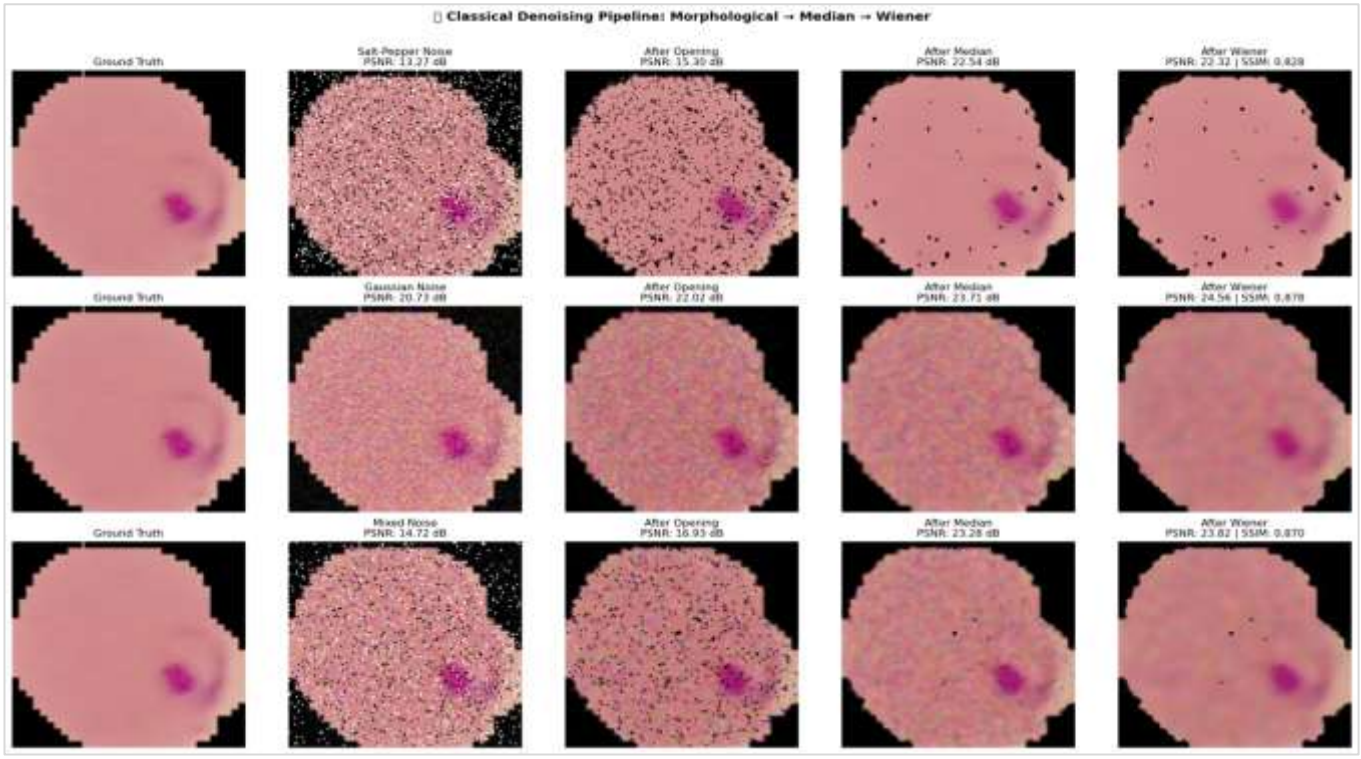

Figure 7. Step-by-Step visualization of classical denoising pipeline (morphological → median → bilateral) on three noise types: salt-and-pepper (top row), gaussian (middle row), and mixed (bottom row)

Table 3. Complete Classical Pipeline Evaluation Results (Opening + Median + Bilateral)

Noise Type	PSNR Noisy	PSNR Opening	PSNR Median	PSNR Final	SSIM Final	Total Improvement
Salt-and-Pepper	13.27 dB	15.30 dB	22.54 dB	22.32 dB	0.828	+9.05 dB
Gaussian	20.73 dB	22.02 dB	23.71 dB	24.56 dB	0.878	+3.83 dB
Mixed	14.72 dB	16.93 dB	23.28 dB	23.82 dB	0.870	+9.10 dB

Deep Learning Model Training Results

The training results of the three deep learning models demonstrate significant differences in convergence characteristics and final performance. Table 4 presents a summary of training results for the three models.

Table 4. Deep Learning Model Training Results Summary

Model	Final Val PSNR	Final Loss	Convergence Characteristics	Parameters
DnCNN	22.08 dB	~0.02	Slow convergence, early plateau	~556,000
Autoencoder	26.36 dB	~0.01	Stable convergence, gradual improvement	~1,200,000
U-Net	36.85 dB	~0.005	Rapid convergence, highest PSNR	~31,000,000

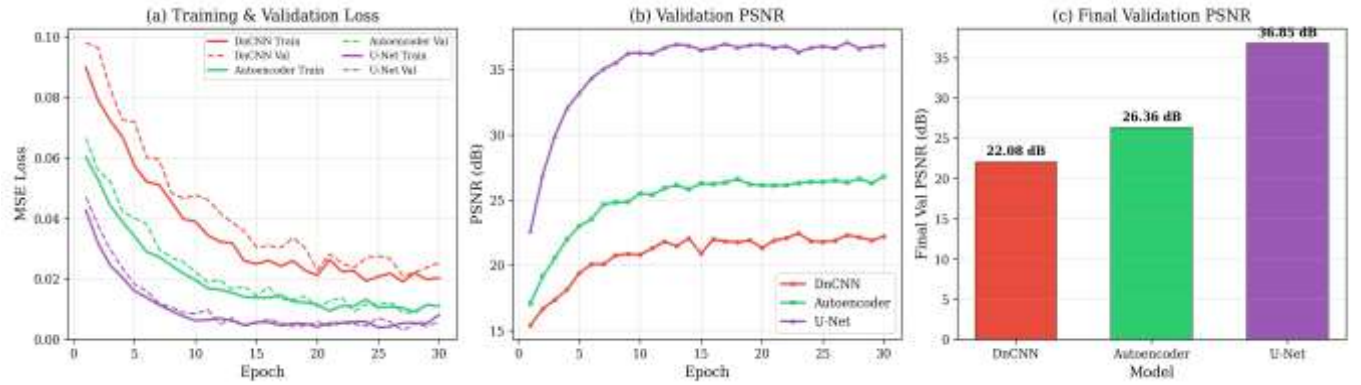

Figure 8. Deep learning model training curves

Based on Table 4, U-Net achieved the best validation PSNR at 36.85 dB with the lowest loss of approximately 0.005, demonstrating rapid convergence. Autoencoder achieved moderate PSNR of 26.36 dB with stable convergence, while DnCNN showed slow convergence with plateau early in training, achieving only 22.08 dB PSNR. Figure 8 displays the training curves for the three models.

Comprehensive Quantitative Comparison

Comprehensive quantitative comparison of all four denoising methods on the test dataset is presented in Table 5 and Table 6. Values in Table 5 represent the average performance across all 100 test images evaluated using mixed noise (combination of Gaussian σ=20 and salt-and-pepper 3%), providing an aggregate measure of each method's overall denoising capability on the most challenging noise condition. These aggregate per-image averages may differ from the per-noise-type values reported in Table 3 for two reasons: (1) Table 3 reports results from individual representative sample evaluations conducted during pipeline tuning, whereas Table 5 computes statistics across the entire 50-image test set; and (2) Table 3 separately evaluates each noise type on a single sample, while Table 5 reports the overall average across all test images under the mixed noise condition. For example, the simple average of Table 3's final PSNR values (22.32, 24.56, 23.82 dB) yields 23.57 dB, while Table 5 reports 23.14 ± 0.52 dB — the 0.43 dB difference reflects the natural variation between a single representative sample and the full test set average.

Table 5. Comprehensive Quantitative Comparison of Denoising Methods on Test Dataset

Method	PSNR (dB)	SSIM	MSE	Time (ms)
Classical	23.14 ± 0.52	0.8666 ± 0.0076	318.0 ± 37.1	7.86 ± 0.43
DnCNN	8.42 ± 0.85	0.1742 ± 0.0161	9534.0 ± 1924.9	4.71 ± 0.50
Autoencoder	26.12 ± 1.19	0.9088 ± 0.0251	165.0 ± 47.0	1.64 ± 0.96
U-Net	36.69 ± 1.34	0.9577 ± 0.0120	14.7 ± 5.1	7.28 ± 2.43

Table 6. Method Performance Ranking by Metric

Metric	Best	Second	Third	Fourth
PSNR (higher = better)	U-Net (36.69)	Autoencoder (26.12)	Classical (23.14)	DnCNN (8.42)
SSIM (higher = better)	U-Net (0.9577)	Autoencoder (0.9088)	Classical (0.8666)	DnCNN (0.1742)
MSE (lower = better)	U-Net (14.7)	Autoencoder (165.0)	Classical (318.0)	DnCNN (9534.0)
Time (lower = better)	Autoencoder (1.64)	DnCNN (4.71)	U-Net (7.28)	Classical (7.86)

U-Net achieved the best performance across all quality metrics with an average PSNR of 36.69 ± 1.34 dB, SSIM of 0.9577 ± 0.0120, and MSE of 14.7 ± 5.1. The PSNR difference of more than 10 dB compared to the second-best method (Autoencoder at 26.12 dB) indicates that U-Net produces image quality that is perceptually far superior, as each 10 dB increase represents a tenfold reduction in mean squared error (Hore & Ziou, 2010).

U-Net's SSIM value of 0.9577 ± 0.0120 approaches the ideal value of 1.0, indicating very high structural similarity between denoised images and ground truth. SSIM developed by Wang et al. (2004) measures similarity based on three components: luminance, contrast, and structure. High SSIM values are highly relevant for medical diagnostic applications where anatomical structure preservation is crucial for accurate identification. U-Net's low SSIM standard deviation (0.0120) indicates high performance consistency across various image samples.

U-Net's MSE of 14.7 ± 5.1 was the lowest among all methods with a very large margin, indicating minimal reconstruction error. As quantitative comparison: Classical MSE (318.0) is 21.6× larger than U-Net; Autoencoder MSE (165.0) is 11.2× larger than U-Net; DnCNN MSE (9534.0) is 648.6× larger than U-Net.

Autoencoder ranked second with PSNR of 26.12 ± 1.19 dB and SSIM of 0.9088 ± 0.0251. This performance shows that encoder-decoder architecture with bottleneck is sufficiently effective for denoising tasks, although with limitations in fine detail preservation compared to U-Net equipped with skip connections.

Classical method achieved PSNR of 23.14 ± 0.52 dB and SSIM of 0.8666 ± 0.0076, representing moderate but most consistent performance (lowest standard deviation). The advantage of classical methods lies in process interpretability, no training stage required, and can be directly applied without training data or GPU.

DnCNN did not achieve competitive performance with PSNR of only 8.42 ± 0.85 dB and SSIM of 0.1742 ± 0.0161, producing output quality significantly below even the noisy input images (13–20 dB). This result indicates that the baseline architecture requires more extensive domain-specific parameter tuning and data normalization rather than representing typical expected performance. The PSNR value falling below the noisy baseline confirms that the model actively degraded image quality, suggesting that the network did not properly converge under the current training configuration. The most probable technical causes include: (1) sensitivity of the residual learning paradigm

(where the network predicts noise residual r = y − x) to domain-specific noise distributions that differ substantially from the BSD68/BSD500 benchmarks on which DnCNN was originally validated (Zhang et al., 2017); (2) potential data scaling incompatibility — if input normalization to [0, 1] range was not perfectly synchronized with the loss function expectations, this could produce catastrophic gradient behavior; (3) the baseline DnCNN architecture with 17 layers and 64 fixed feature maps was originally optimized for grayscale natural images, and its direct application to three-channel microscopic images with heterogeneous Giemsa staining patterns presents a significant domain gap (Mohan et al., 2020); and (4) the absence of multi-scale feature extraction limits the model's ability to capture both global context and fine-grained morphological details simultaneously. IMPORTANTLY, these results should NOT be interpreted as evidence that DnCNN or CNN-based residual learning is inherently unsuitable for microscopic imaging. Rather, they represent anomalous training outcomes that require systematic debugging including verification of data preprocessing pipelines, residual learning formula implementation, and potential retraining with domain-adapted configurations (increased depth, attention modules, progressive training schedules) — before any definitive conclusions about DnCNN's applicability to this domain can be drawn.

Processing time analysis shows that Autoencoder is the fastest method with an average of 1.64 ± 0.96 ms per image, followed by DnCNN (4.71 ± 0.50 ms), U-Net (7.28 ± 2.43 ms), and classical method (7.86 ± 0.43 ms). All methods have processing times in the millisecond order meeting real-time application requirements (processing rate > 100 frames per second).

Visual Analysis of Denoising Results

Visual analysis was conducted to evaluate denoising quality qualitatively, particularly in the aspect of morphological detail preservation important for malaria diagnosis. Figure 9 displays visual comparison of denoising results from the four methods on several test image samples covering Parasitized and Uninfected classes.

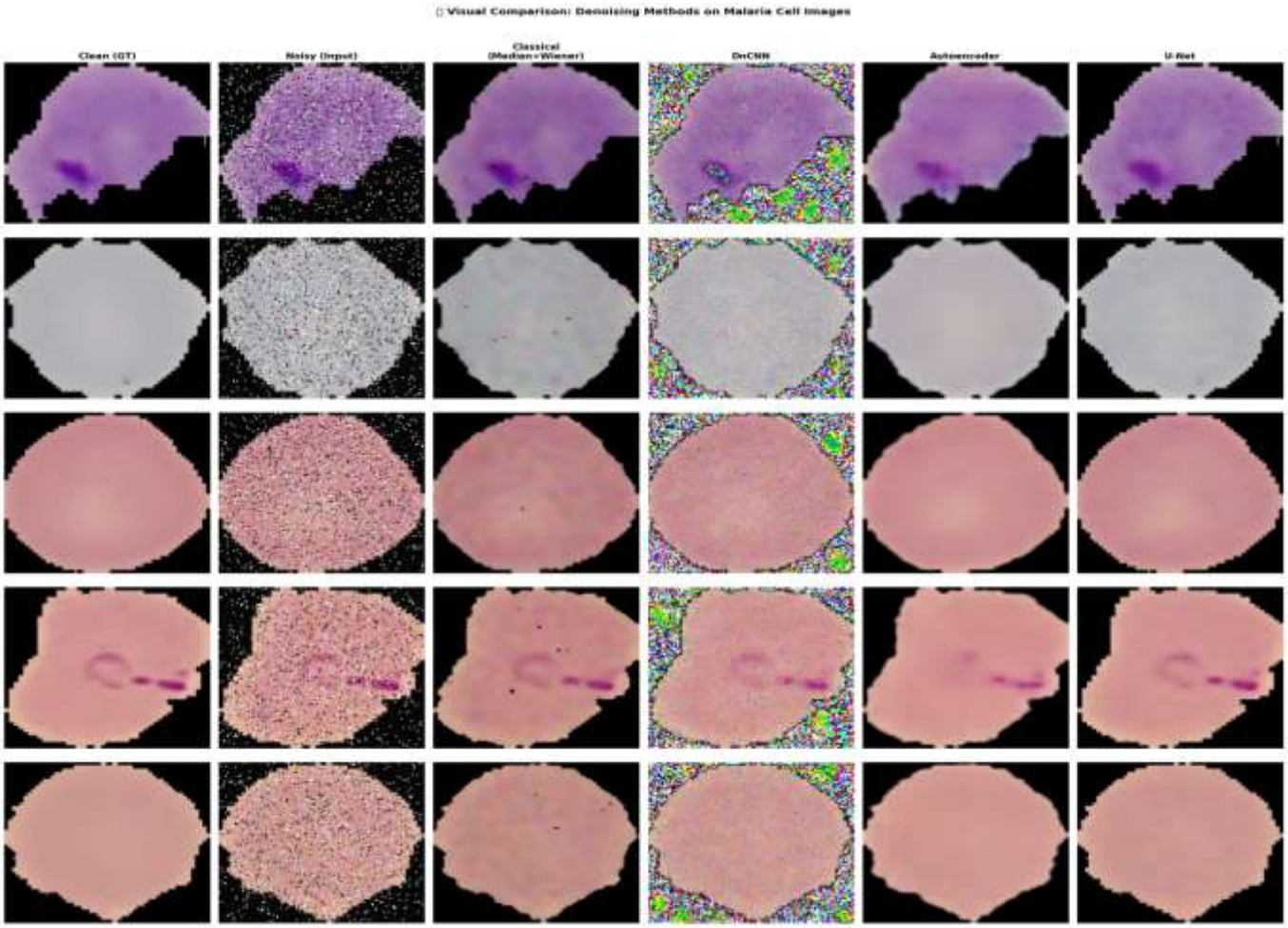

Figure 9. Visual comparison of denoising results on five malaria image samples: clean (ground truth), noisy (input), classical (median+bilateral), dncnn, autoencoder, and u-net

Based on Figure 9, the quality differences of denoising results between methods can be clearly observed across six columns for each sample. The first column displays ground truth (Clean/GT) images as reference, the second column displays images with mixed noise as input, and the third through sixth columns display results from each denoising method.

On images containing Plasmodium parasites (Parasitized), preservation of parasite morphological details is the most crucial evaluation criterion. U-Net (sixth column) consistently produces images with the best visual quality nearly identical to ground truth. Parasite ring stage structure, cell membrane boundaries, and internal color gradation can be preserved very well. Autoencoder (fifth column) produces fairly good images with noise almost entirely eliminated, but with slight softening of fine details. Classical method (third column) successfully removes most noise but leaves some artifacts in high contrast areas. DnCNN (fourth column) shows clear failure with output containing significant color artifacts and unrecognizable structures.

On normal uninfected blood cell images, evaluation focused on preservation of color uniformity and cell structure. U-Net again demonstrated superiority with results maintaining normal cell visual characteristics including central pallor and clear cell boundaries. DnCNN produced output with severe color distortion, confirming that the model failed to learn the characteristics of this image domain.

Comparative Discussion with Related Research

The superior performance of U-Net with PSNR of 36.69 dB is consistent with findings by Ronneberger et al. (2015) and further validated by nested architectures such as UNet++ (Zhou et al., 2018) demonstrating the effectiveness of this architecture for various biomedical image processing tasks. The advantage of U-Net lies in its skip connection mechanism enabling multi-scale spatial information preservation through direct pathway from encoder to decoder.

It should be acknowledged that the comparison between U-Net (31 million parameters) and DnCNN (556 thousand parameters) represents a comparison between heavyweight and lightweight architectures with a 56× parameter ratio. The superior performance of U-Net may be partially attributed to its significantly larger model capacity rather than solely its architectural innovations such as skip connections. Future studies should consider comparing architectures with controlled parameter counts to isolate the contribution of architectural design from model capacity. Nevertheless, this comparison remains practically valuable as it reflects the realistic trade-off between computational investment and denoising quality that practitioners face when selecting models for clinical deployment.

The DnCNN finding that exhibited suboptimal convergence on the malaria dataset stands in contrast with results by Zhang et al. (2017) on the BSD68 benchmark dataset where DnCNN achieved state-of-the-art PSNR. This discrepancy indicates that the baseline DnCNN configuration requires domain-specific parameter tuning and normalization adjustments rather than reflecting an inherent architectural limitation of CNN-based residual learning for microscopic imaging. Similar domain-specific convergence challenges have been documented in the denoising literature: Mohan et al. (2020) demonstrated that standard CNN denoisers can exhibit bias-related failures when applied to image domains with statistical properties different from their training data distribution, while Krull et al. (2019) showed that alternative training paradigms may be necessary when conventional supervised denoising approaches fail to converge. Unique characteristics of microscopic images — including: (i) Giemsa staining variations producing complex chromatic noise patterns distinct from natural images; (ii) microscope lighting heterogeneity; (iii) specific morphological structures of blood cells and parasites; and (iv) predominantly dark background — collectively create a domain gap that requires specialized architectural considerations. The baseline DnCNN architecture, while proven on natural image benchmarks, requires domain-specific adaptation strategies including potential retraining with verified data scaling, expanded network capacity, or incorporation of attention mechanisms to achieve competitive performance on medical microscopic images.

Compared to previous research, this study has several strengths and unique contributions: Comprehensive Quantitative Benchmark providing a comprehensive quantitative benchmark for comparing denoising methods on the Cell Images for Detecting Malaria dataset. Comprehensive Multi-Noise Evaluation evaluating three noise types representing various degradation conditions. Quality-Efficiency Trade-off Analysis evaluating both denoising quality and computational efficiency. Domain-Specific Challenges Identification providing important insight regarding the adaptation requirements when transferring denoising architectures from natural image benchmarks to microscopic medical imaging domains.

Implications for Malaria Diagnosis Systems

In the context of computer-aided malaria diagnosis applications, the results of this study have significant practical implications for various implementation scenarios.

For High-Accuracy Diagnosis Systems: Preservation of Plasmodium parasite morphological details is a crucial factor because species and infection stage identification depends on specific morphological characteristics (Moody, 2002; Ruberto et al., 2002). U-Net with SSIM of 0.9577 demonstrates superior capability in maintaining morphological structure, making it highly recommended as a preprocessing method for automated malaria detection systems requiring high accuracy.

For Severely Resource-Constrained Deployment: In scenarios with extremely limited computational resources (e.g., legacy mobile devices without GPU support, embedded systems in remote clinics, or older hardware commonly found in malaria-endemic regions), Autoencoder with the fastest processing time (1.64 ms) and moderate performance (PSNR 26.12 dB, SSIM 0.9088) may serve as a last-resort fallback when U-Net deployment is technically infeasible. However, it must be strongly emphasized that the 10.57 dB PSNR gap between Autoencoder and U-Net represents approximately a tenfold difference in reconstruction error (Hore & Ziou, 2010), which has clinically significant implications for diagnostic accuracy — particularly in distinguishing Plasmodium species and infection stages that depend on subtle morphological features. The millisecond-level speed advantage (1.64 ms vs. 7.28 ms) does not justify this quality sacrifice in any clinical scenario where diagnostic accuracy is prioritized. Therefore, Autoencoder should NOT be considered an optimal or general-purpose choice but rather a last-resort option exclusively reserved for deployment scenarios where severe hardware constraints make U-Net deployment technically impossible.

For Facilities with Limited Infrastructure: For scenarios with limited computational resources or GPU unavailability, classical methods remain a viable alternative with PSNR of 23.14 dB and SSIM of 0.8666. The classical pipeline requires no training stage, no GPU for inference, and can be implemented on minimal specification devices using standard libraries like OpenCV.

Implementation Recommendations

Based on comprehensive evaluation results, recommendations for denoising method selection for malaria diagnosis systems are as follows: (1) Maximum quality priority: U-Net (PSNR 36.69 dB, SSIM 0.9577). (3) Resource-constrained last-resort only: Autoencoder (PSNR 26.12 dB, time 1.64 ms) — recommended ONLY when U-Net deployment is technically infeasible due to severe hardware limitations (e.g., legacy mobile devices without GPU). The 10.57 dB PSNR deficit compared to U-Net represents a clinically significant quality difference that should not be traded for millisecond-level speed gains in diagnostic applications. (3) Scenario without GPU/training data: Classical Method (PSNR 23.14 dB, interpretable). (4) Not recommended without modification: Baseline DnCNN architecture did not achieve competitive performance (PSNR 8.42 dB), indicating insufficient domain-specific parameter tuning. Systematic debugging of data preprocessing pipeline (verification of [0,1] scaling, residual learning formula) and retraining with domain-specific architectural modifications (increased depth, attention modules, domain-adapted preprocessing) is strongly recommended before any deployment in microscopic imaging applications.

Limitations of the Study

This study has several limitations that should be considered when interpreting the results. First, the evaluation was conducted using synthetic noise added to clean images, which may not fully represent the complexity of real-world noise encountered in clinical microscopy settings where multiple noise sources interact in non-additive ways. Second, the deep learning models were trained from scratch on the malaria dataset without leveraging pre-trained weights from larger image restoration datasets, which could potentially improve performance, particularly for DnCNN. Third, the noise simulation was conducted at fixed intensity levels (Gaussian σ=25, Salt-and-Pepper 5%), which limits the assessment of method robustness across varying noise severities. Future work should evaluate performance across multiple noise levels (e.g., σ = 15, 25, 50 for Gaussian noise) to comprehensively characterize the noise-level sensitivity and blind denoising capability of each method. Fourth, the study focused on three specific deep learning architectures and did not include more recent transformer-based methods such as Restormer (Zamir et al., 2022) or SwinIR (Liang et al., 2021), which have demonstrated state-of-the-art performance on general image restoration benchmarks. Fifth, the computational efficiency evaluation was conducted on a single GPU platform (NVIDIA Tesla T4), and performance characteristics may vary on different hardware configurations commonly used in clinical settings. Finally, the impact of denoising on downstream malaria diagnosis accuracy was not directly evaluated, which would provide more clinically relevant validation of the proposed methods.

Conclusion

This study has successfully conducted a comprehensive comparative analysis between classical denoising methods based on a combination of morphological operations, median filter, and bilateral filter with deep learning methods including DnCNN,

Autoencoder, and U-Net on the Cell Images for Detecting Malaria dataset. Experimental results demonstrate that U-Net architecture achieved the best performance significantly with an average PSNR of 36.69 dB and SSIM of 0.9577, outperforming Autoencoder (PSNR 26.12 dB, SSIM 0.9088), classical methods (PSNR 23.14 dB, SSIM 0.8666), and DnCNN which did not achieve competitive performance (PSNR 8.42 dB, SSIM 0.1742). The PSNR difference of more than 10 dB between U-Net and the second-best method represents a tenfold reduction in reconstruction error. The superiority of U-Net lies in its skip connection mechanism capable of preserving multi-scale spatial information so that Plasmodium parasite morphological details can be very well preserved, which is a crucial aspect in malaria diagnosis. Analysis of various noise types demonstrates that the classical pipeline is effective for handling salt-and-pepper noise with a PSNR improvement of 9.05 dB, Gaussian noise with a 3.83 dB improvement, and mixed noise with a 9.10 dB improvement, while remaining a viable option for facilities without GPU infrastructure. Autoencoder emerged as the most computationally efficient method with a processing time of 1.64 ms per image, although the significant quality gap relative to U-Net limits its applicability primarily to deployment scenarios with severe hardware constraints where U-Net is technically infeasible. Regarding DnCNN, the baseline architecture requires domain-specific adaptations including verified input normalization pipelines, adjusted learning rate schedules, and potentially expanded network capacity with attention mechanisms to achieve convergence on medical microscopic images, and its suboptimal results in this study should not be generalized as a limitation of CNN-based residual learning approaches. Analysis of various noise types demonstrates that the classical pipeline is effective for handling salt-and-pepper noise with a PSNR improvement of 9.05 dB, Gaussian noise with a 3.83 dB improvement, and mixed noise with a 9.10 dB improvement. Based on these findings, U-Net is recommended as the primary denoising method for malaria diagnostic systems where accuracy and morphological detail preservation are paramount. The DnCNN results in this study are classified as anomalous training data resulting from model collapse, and should not be interpreted as evidence against CNN-based residual learning for microscopic imaging; systematic debugging and domain-adapted retraining are recommended before definitive conclusions. This study contributes a comprehensive quantitative benchmark for comparing denoising methods on malaria microscopic image datasets that can serve as a reference for the development of automatic computer vision-based malaria diagnosis systems in the future. Future research should evaluate denoising methods on real-world noisy microscopic images acquired under varying clinical conditions, explore transfer learning and domain adaptation strategies to improve lightweight architectures like DnCNN for medical imaging domains, investigate transformer-based architectures such as Restormer and SwinIR for microscopic image denoising, and conduct end-to-end evaluation of the impact of denoising on downstream malaria parasite detection and species classification accuracy to further validate the clinical relevance of these preprocessing methods. Additionally, developing lightweight model architectures suitable for deployment on resource-constrained devices commonly available in malaria-endemic regions and systematic retraining of DnCNN with domain-specific adaptations remain important directions for future work.

Acknowledgments

The authors would like to express their gratitude to all parties who have provided support throughout the completion of this research. The authors confirm that no individual who does not meet the criteria for authorship has contributed to the writing or preparation of this manuscript, and no professional writing services or materials were utilized in the production of this article.

Author Contributions

Conceptualization, I.K.N. and IWS; methodology, I.K.N., I.W.S. and I.W.W.; formal analysis, I.K.N. and I.W. Sn.; investigation, I.K.N., I.W.S. and I.W.W.; resources, I.K.N., I.W.S. and I.W.W.; writing—preparation of original draft, I.K.N.; writing—reviewing and editing, I.W.S. and I.W. W.; visualization, IWS; supervision, I.W.W.; project administration, IKN; obtaining funding, I.K.N. and I.W.W. All authors have read and approved the published version of the manuscript.

Funding

This research received no external funding

Conflicts of Interest

The authors declare no conflict of interest.

References

Alonso, P. L., Brown, G., Arevalo-Herrera, M., Binka, F., Chitnis, C., Collins, F., Doumbo, O. K., Greenwood, B., Hall, B. F., Levine, M. M., Mendis, K., Newman, R. D., Plowe, C. V., Rodríguez, M. H., Sinden, R., Slutsker, L., & Tanner, M. (2011). A Research Agenda to Underpin Malaria Eradication. *PLoS Medicine*, *8*(1), e1000406. https://doi.org/10.1371/journal.pmed.1000406

Andrade, B. B., Reis-Filho, A., Barros, A. M., Souza-Neto, S. M., Nogueira, L. L., Fukutani, K. F., Camargo, E. P., Camargo, L. M., Barral, A., Duarte, Â., & Barral-Netto, M. (2010). Towards a precise test for malaria

diagnosis in the Brazilian Amazon: comparison among field microscopy, a rapid diagnostic test, nested PCR, and a computational expert system based on artificial neural networks. *Malaria Journal*, *9*(1), 117. https://doi.org/10.1186/1475-2875-9-117

Buades, A., Coll, B., & Morel, J. M. (2005). A Review of Image Denoising Algorithms, with a New One. *Multiscale Modeling & Simulation*, *4*(2), 490–530. https://doi.org/10.1137/040616024

Cai, X., Mamun, A.-A., & Rajasekaran, S. (2017). Novel algorithms for finding the closest l-mers in biological data. *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 525–528. https://doi.org/10.1109/BIBM.2017.8217702

Chen, L., Chu, X., Zhang, X., & Sun, J. (2022). Simple Baselines for Image Restoration. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 17–33). https://doi.org/10.1007/978-3-031-20071-7_2

Chow, L. S., & Paramesran, R. (2016). Review of medical image quality assessment. *Biomedical Signal Processing and Control*, *27*, 145–154. https://doi.org/10.1016/j.bspc.2016.02.006

Das, D. K., Ghosh, M., Pal, M., Maiti, A. K., & Chakraborty, C. (2013). Machine learning approach for automated screening of malaria parasite using light microscopic images. *Micron*, *45*, 97–106. https://doi.org/10.1016/j.micron.2012.11.002

Doyle, P. S., Zhou, Y. M., Engel, J. C., & McKerrow, J. H. (2007). A Cysteine Protease Inhibitor Cures Chagas' Disease in an Immunodeficient-Mouse Model of Infection. *Antimicrobial Agents and Chemotherapy*, *51*(11), 3932–3939. https://doi.org/10.1128/AAC.00436-07

Elad, M., & Aharon, M. (2006). Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries. *IEEE Transactions on Image Processing*, *15*(12), 3736–3745. https://doi.org/10.1109/TIP.2006.881969

Fan, L., Zhang, F., Fan, H., & Zhang, C. (2019). Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, *2*(1), 7. https://doi.org/10.1186/s42492-019-0016-7

Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson.

Hayes, N., Merkurjev, E., & Wei, G.-W. (2023). Integrating transformer and autoencoder techniques with spectral graph algorithms for the prediction of scarcely labeled molecular data. *Computers in Biology and Medicine*, *153*, 106479. https://doi.org/10.1016/j.compbiomed.2022.106479

Hore, A., & Ziou, D. (2010). Image Quality Metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition*, 2366–2369. https://doi.org/10.1109/ICPR.2010.579

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5967–5976. https://doi.org/10.1109/CVPR.2017.632

Krull, A., Buchholz, T.-O., & Jug, F. (2019). Noise2Void - Learning Denoising From Single Noisy Images. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2124–2132. https://doi.org/10.1109/CVPR.2019.00223

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., & Timofte, R. (2021). SwinIR: Image Restoration Using Swin Transformer. *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 1833–1844. https://doi.org/10.1109/ICCVW54120.2021.00210

Lim, J. S. (1990). *Two-Dimensional Signal and Image Processing*. Prentice-Hall.

Milanfar, P. (2013). A Tour of Modern Image Filtering: New Insights and Methods, Both Practical and Theoretical. *IEEE Signal Processing Magazine*, *30*(1), 106–128. https://doi.org/10.1109/MSP.2011.2179329

Mohan, S., Kadkhodaie, Z., Simoncelli, E. P., & Fernandez-Granda, C. (2020). Robust and interpretable blind image denoising via bias-free CNNs. *Proceedings of the International Conference on Learning Representations (ICLR 2020*. Retrieved from https://arxiv.org/abs/1906.05478

Moody, A. (2002). Rapid Diagnostic Tests for Malaria Parasites. *Clinical Microbiology Reviews*, *15*(1), 66–78. https://doi.org/10.1128/CMR.15.1.66-78.2002

Oliveira, J. P., Bioucas-Dias, J. M., & Figueiredo, M. A. T. (2009). Adaptive total variation image deblurring: A majorization–minimization approach. *Signal Processing*, *89*(9), 1683–1693. https://doi.org/10.1016/j.sigpro.2009.03.018

Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., & Thoma, G. (2018). Image analysis and machine learning for detecting malaria. *Translational Research*, *194*, 36–55. https://doi.org/10.1016/j.trsl.2017.12.004

Rajaraman, S., Antani, S. K., Poostchi, M., Silamut, K., Hossain, M. A., Maude, R. J., Jaeger, S., & Thoma, G. R. (2018). Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear

images. *PeerJ*, *6*, e4568. https://doi.org/10.7717/peerj.4568

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Ruberto, C. Di, Dempster, A., Khan, S., & Jarra, B. (2002). Analysis of infected blood cell images using morphological operators. *Image and Vision Computing*, *20*(2), 133–146. https://doi.org/10.1016/S0262-8856(01)00092-0

Snow, R. W. (2015). Global malaria eradication and the importance of Plasmodium falciparum epidemiology in Africa. *BMC Medicine*, *13*(1), 23. https://doi.org/10.1186/s12916-014-0254-7

Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W., & Lin, C.-W. (2020). Deep learning on image denoising: An overview. *Neural Networks*, *131*, 251–275. https://doi.org/10.1016/j.neunet.2020.07.025

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P. A. (2010). Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, *11*, 3371–3408. Retrieved from http://jmlr.org/papers/v11/vincent10a.html

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

White, N. J., Pukrittayakamee, S., Hien, T. T., Faiz, M. A., Mokuolu, O. A., & Dondorp, A. M. (2014). Malaria. *The Lancet*, *383*(9918), 723–735. https://doi.org/10.1016/S0140-6736(13)60024-0

WHO. (2015). *Global Technical Strategy for Malaria 2016-2030*. WHO Press. Retrieved from https://www.who.int/publications/i/item/9789241564991

WHO. (2022). *World Malaria Report 2022*. WHO Press. Retrieved from https://www.who.int/publications/i/item/9789240064898

Wiener, N. (1949). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. MIT Press.

Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., & Yang, M.-H. (2022). Restormer: Efficient Transformer for High-Resolution Image Restoration. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5718–5729. https://doi.org/10.1109/CVPR52688.2022.00564

Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, *26*(7), 3142–3155. https://doi.org/10.1109/TIP.2017.2662206

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (pp. 3–11). Springer. https://doi.org/10.1007/978-3-030-00889-5_1