As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Issue 02, pages 57-62.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue02.html
DOI: https://doi.org/10.5281/zenodo.18683096

Real-Time NIR-II Image Denoising via Spectrum-Shift Explainable GAN

Sri Lasya Hemadri | Sunitha Raavi | Aruna Dendukuri | Tejaswini Dondapati | Venkatesh Vijji | Venu Madhav Gosula

Department of Electronics and Communication Engineering, NRI Institute of Technology, Vijayawada, AP, India.

To Cite this Article
Sri Lasya Hemadri, Sunitha Raavi, Aruna Dendukuri, Tejaswini Dondapati, Venkatesh Vijji & Venu Madhav Gosula (2026). Real-Time NIR-II Image Denoising via Spectrum-Shift Explainable GAN. International Journal for Modern Trends in Science and Technology, 12(02), 57-62. https://doi.org/10.5281/zenodo.18683096

Article Info
Received: 15 January 2026; Revised: 10 February 2026; Accepted: 17 February 2026.

Copyright © The Authors ; This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

KEYWORDS

Biomedical image enhancement, cycle-consistent generative adversarial network, explainability, NIR-II imaging.

ABSTRACT

In Vivo Biomedical imaging, the Second Near Infrared Second Window (NIR-II) has real-time uses. As the wavelength of NIR-IIB (1500-1700 nm) imaging can produce higher resolution than NIR-IIA, the NIR-IIB window has been chosen for real-time clinical purposes. The process which is involved in the deduction of noise in the Low-Light images is inevitable. As in the earlier times, the usage of the fluorescent probes for the imaging of the low-light images of paired datasets, the technique of training a model using unpaired datasets of the low-light NIR-IIB images was to be used for the enhancement of the images, is the Explainable Spectrum Shift Generative Adversarial Network. This proposed method is used for tackling the problem aroused by a smaller number of unpaired datasets available in the real-time world. This model is trained such that, along with the enhancement of the images also generates the image metrics and their respective visual structural and tabular representations. The Python-based Deep Learning model can be utilized for the vivo biomedical imaging in real-time applications. The Results that are demonstrated by the model improve the reproducibility, flexibility, and analytical depth, also providing safe, accurate and scalable enhancement of the NIR-II images in the Clinical practice.

I. INTRODUCTION

In real-time and near real-time, the visualisation of the tissues inside a living organism needs to be captured in a window with a wavelength in nm. The modern option in order to promise the modality of the targeted tissues is the NIR-IIB window. Fluorescent Probes exhibit limited quality. This is because of the photon detection process in the NIR-IIA window, and also the scattering phenomena. Denoising of the NIR-IIB window images achieves maximum resolution in the tissues that are of millimetre depth. This also suppresses the scattering of the light.

The xAI-PSSGAN model developed earlier also used NIR-IIB window and our technique is much easier that the earlier approach as we have used Python-based Deep Learning Techniques for the training of the model to not only enhance and denoise the low-light unpaired images, but also generates the image metrics along with the visual and tabular representations for easy comparison of the image enhancements multiple levels of enhancement.

Traditional enhancement methods fail to recover fine textures, especially under heavy noise. A CycleGAN image enhancement framework that is designed to transform images from Domain A to Domain B without pairing the training data. This is followed by a quantitative assessment, and a report is generated automatically.

The dataset is divided into 2 domains: Domain A, which consists of original images or low-light images, and Domain B Target images. The model is trained in such a way that it is trained for taking a large dataset and dividing the dataset in such a way that the training data images be 70% and the testing and validation together be 30%.

The Generator and the Discriminator work like the process in which the Domain A dataset is converted to Domain B and then reconstructed back to Domain A, and compare the enhancement of the images. For the mismatched training dataset enhancement (unpaired data) the CycleGAN technique gives the best solution. Not only generating the enhanced image output, but also generating its theoretical representations can give the real-time users a keen knowledge and is easy for analyzing making this image increases, and also the size of the images remains the same. When the images are resized in the generators, some models may reduce the number of pixels. But this model ensures even during the process of enhancement in generators, the size of the image does not change.

The other 30% of the data is used for Validation and testing. These dataset images have been given to the CycleGAN Training model consisting of the ResNet-Generators and the PatchGAN Discriminators for the enhancement of the noisy images, generating the output enhanced images along with the metrics of the images in a detailed brief-out inference report.

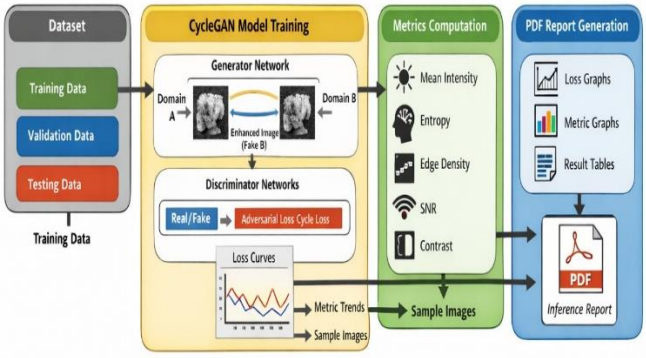


Figure 1 Block Diagram of the model working.

II. RELATED WORK

Near-infrared (NIR) imaging, particularly in the second near-infrared window (NIR-II), has gained significant attention in clinical oncology due to its superior imaging depth, high spatial resolution, and reduced scattering compared to conventional imaging modalities. Zhang et al. [1] highlighted the clinical opportunities and challenges of NIR-II imaging, emphasizing its potential for precise tumor visualization and surgical guidance. Similarly, Chang et al. [2] developed a phosphorescent probe capable of in vivo imaging in the NIR-II window, demonstrating improved sensitivity and imaging performance. Furthermore, Hu et al. [3] conducted the first human liver tumor surgery guided by multispectral fluorescence imaging, confirming the clinical feasibility and effectiveness of combined visible and NIR imaging techniques. Wang et al. [4] also emphasized the role of fluorescence image-guided surgery in enhancing tumor detection accuracy and improving surgical outcomes.

Recent advancements in artificial intelligence, particularly deep learning, have further enhanced medical imaging capabilities. Generative Adversarial Networks (GANs) have shown remarkable performance in medical image enhancement, reconstruction, and translation tasks. Yi et al. [5] provided a comprehensive review of GAN applications in medical imaging, highlighting their effectiveness in improving image quality and diagnostic accuracy. Zhu et al. [6] introduced the CycleGAN framework, which enables unpaired image-to-image translation and has been widely adopted for medical image enhancement and domain adaptation. In addition, Fu et al. [7] proposed a spectrum shift-based GAN model for perceptual quality enhancement in fluorescence imaging, demonstrating significant improvements in image clarity and detail preservation.

Explainable artificial intelligence (XAI) has also emerged as an important research area to improve the interpretability of deep learning models. Biran and Cotton [8] presented a comprehensive survey on explainability techniques, emphasizing their importance in building trust and transparency in machine learning systems. In the context of NIR imaging, Ma et al. [9] demonstrated the effectiveness of deep learning methods in improving in vivo NIR image quality and analysis. Moreover, visualization techniques such as saliency maps, introduced by Simonyan et al. [10], provide valuable insights into convolutional neural network (CNN) decision-making processes, enabling better understanding of feature importance in medical image analysis.

Although significant progress has been made in NIR imaging and deep learning-based enhancement techniques, challenges remain in improving image quality, interpretability, and real-time clinical applicability. Therefore, integrating advanced deep learning models with explainable frameworks offers a promising direction for enhancing fluorescence imaging performance and supporting clinical decision-making.

III. METHODS

A. ResNet-Generators

Residual Network Generators are mostly used for image enhancements of low-light to normal images, also used for image translations. The high-level architecture of the ResNet Generator uses convolutions, up-sampling and downsampling, along with the residual blocks, for the purpose of enhancing a low-light image. This is an encoder-transformer-decoder-based design. The first step in the Generator is the ReflectionPadding, which is used to avoid the edge artifacts.

The next step in the generator is the convolution. It's a 7X7 Convolution. This happens in the residual block with multiple steps. Invovles ReLU technique and the kernel used is 3X3, same number of channels.

$$y_{i,j,k} = \sum (x * W)_{i,j,k}$$

The convolution can also be written as follows:

$$y = x * W_{7X7}$$

The next step is the down-sampling. This is done for the purpose of spatial size reduction, feature depth improvement, and mainly the capturing of Semantic information. After N-Residual Blocks of Convolution(Say 6), each block preserves the dimensions and adds the skipped connections i.e, y=x+F(x). where x is the input feature map and F(x) is the learned residual mapping.

Upcoming will be the Upsampling. This is done for the restoration of the spatial resolutions and also helps in generating a high-quality output enhanced image. The output convolution image is then given to the Activation function. The activation function used here is the "tanh" function because the tanh function allows the output image to withstand the range of [-1,1] as that range is fixed to match the normalized input image. One of the major reasons to use the tanh function is that it produces stable GAN training.

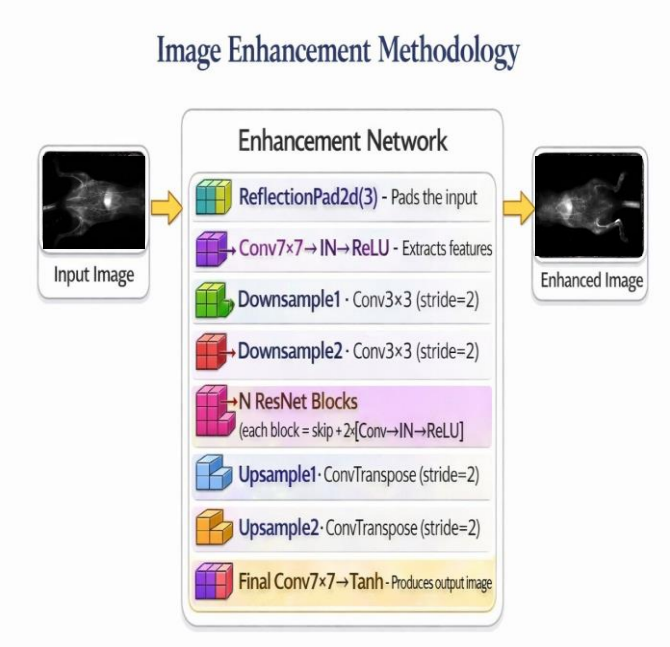


Figure 2(a) Res-Net Generator Architecture.

B. Patch-GAN Discriminator

The Discriminator used in the CycleGAN technique is the PatchGAN discriminator, which do not classify the entire image as fake or real but instead it divides the picture into small patches and identifies whether it's real or fake. Instead of considering single values, this considers probabilities of matrix values. This is what makes it different from the traditional Discriminators.

The PatchGAN Discriminator takes the input from the output of the Generator and undergoes multiple convolutions and compares the image with the fake image generated by the generator patch-wise. Unlike the generators, Discriminator used LeakyReLU to prevent dead neurons in the model and helps discriminator learn faster. The PatchGAN uses strided convolutions instead

of pooling. The layers are not fully connected. The ultimate reason is that it keeps spatial correspondence.

Later, the activation function used is either the Sigmoid function or no activation function is used. The convolution blocks used in discrimination are 4x4 kernels with Stride 2. Each value in the output is: $D(x)_{i,j} \in [0,1]$. The different types of patch sizes exist, but the 70X70 is used in CycleGAN, which provides the best balance.

The losses generated by the Generator and the Discriminator should be such that after multiple sample enhancements of the image, the generator loss must be reduced, and the discriminator losses should almost become zero.

The loss formulas that were used and trained for the model are:

GENERATOR LOSS:

$$\mathcal{L}_{GAN}(G) = \mathbb{E}[\log D(G(x))]$$

DISCRIMINATOR LOSS:

For each patch:

$$\mathcal{L}_D = \mathbb{E}[\log D(x)] + \mathbb{E}[\log(1 - D(G(z)))]$$

Total loss:

$$\mathcal{L}_D = \frac{1}{N} \sum_{i,j} \mathcal{L}_{D_{i,j}}$$

Where:

- i,j are patch locations
- Loss is averaged over all patches

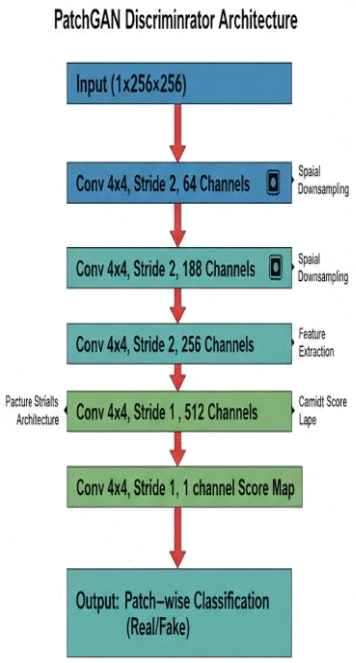


Figure 2(b) Patch-GAN Discriminator Architecture.

IV. RESULT

Firstly, the model has been trained for a large dataset of images, such that any number of images can be enhanced with the model. The images taken for the testing and validation compare with the images from the images given for the training, such that the unpaired image comparison makes the decisions based on the features.

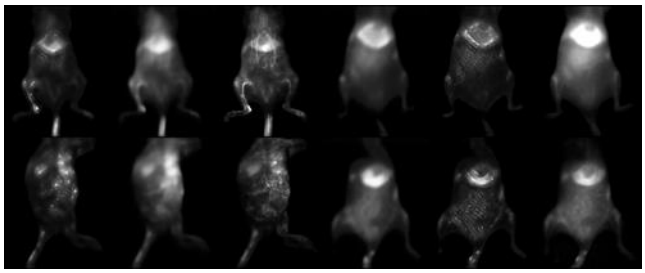


Figure 3(b) Enhanced Output of NIR-II low-light images.

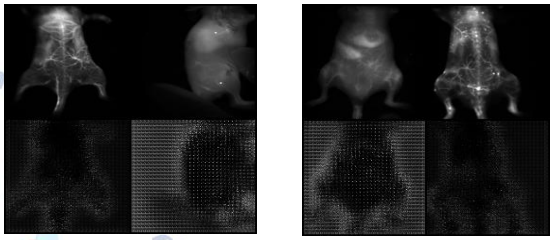


Figure 3(b) Step-wise enhancement.

The losses generated by the Generator and the Discriminator are generated, and for the easiest representation, the few samples are considered as a table of values, such that the graph for the appropriate values is to be plotted.

Sample	Generator-loss	DiscriminatorA-loss	DiscriminatorB-loss
0	12.0463	0.7877	0.7582
1	10.3233	0.7283	0.7129
2	9.6242	0.6625	0.6799
3	7.8868	0.6404	0.6436
4	7.6309	0.5784	0.6252
5	6.7991	0.5688	0.6221
6	5.6929	0.5749	0.6265
7	5.0870	0.6343	0.5962
8	4.3604	0.6092	0.5915
9	4.2952	0.6101	0.5841

Table 1 Losses in Generator and Discriminator

The above values are the theoretical values calculated in order to represent the losses in the generator and the discriminator. On observing the values in the table, we observe that the generator loss is gradually reducing from the samples as the number of iterations increases,

which means that the enhancement is increased for every iteration, and the discriminator loss is almost "0", which means that the image is not fake. The graphical representation of the values is as follows.

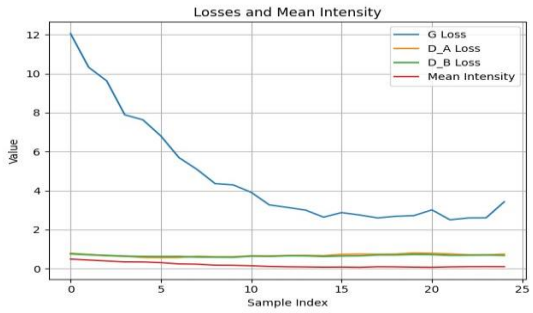


Figure 3(c) Graphical representation of the losses.

The mean intensity represented the average brightness of the image. NIR-II images generally have low mean intensity. Here, the mean intensity is almost 'Zero'.

V. EVALUATION METRICS

The performance of the image enhancement is assessed using the image metrics such as the signal-to-noise ratio (SNR), entropy, edge density, contrast, mean intensity…. The model is trained as the samples taken are not only enhanced but also calculate the image metrics generate the output values along with the graphical representation.

A. Signal-to-Noise Ratio (SNR): The evaluation of the strength of the signal to its background noise is detected and calculated using the Signal-to-Noise ratio. The low-light images usually have low SNR due to the sensor limitations and the photon noise. The higher SNR after enhancement means the higher, effective suppression of noise and preserves the integrity of the signal.

$$SNR = \frac{\mu s}{\sigma_n}$$

where μs is the mean of the signal

σ_n is the variance of the noise

B. Entropy: The amount of information present in an image or the randomness in an image is called as the entropy. The formula of entropy is the ratio of the number of pixels with a particular intensity to the total number of pixels. If the entropy is high, then the image is said to be more detailed. Hence symbolizing better information recovery from noise.

$$Entropy = \frac{number\ of\ pixels\ with\ intensity\ 'i'}{Total\ number\ of\ Pixels}$$

C. Edge Density: For defining the structural clarity, the proportion of the edge pixels are considerd. Mostly for low-light images like NIR-II images, the edges are mostly blurry. Inorder to suppress the smooth boundaries and attain a fine boundary, we use the edge density, which brings higher spatial resolution in NIR-II images. This is achieved by applying the edge pixel values to the edge detector.

$$Edge\ Density = \frac{Ne}{N}$$

where Ne is the number of edge pixels

N is normalised by the total number of pixels

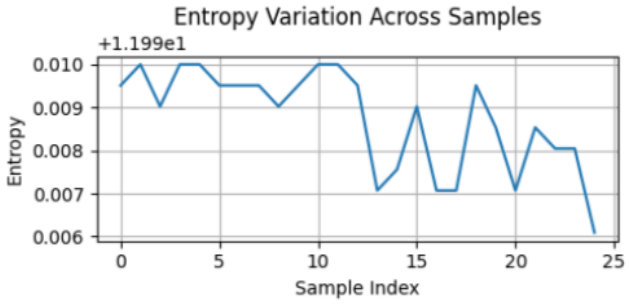
D. Contrast: Contrast is the variation in the intensity level. This measures the spread of intensities around the mean value. The main use of contrast is that it helps to separate the main image from the background more clearly by highlighting the darker regions.

$$C = \sqrt{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y) - \mu^2)}$$

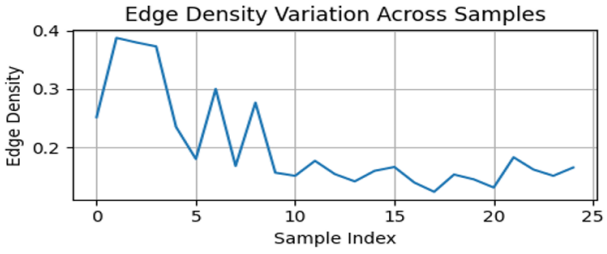
where I(x,y) is the pixel intensities

μ is the Mean intensity

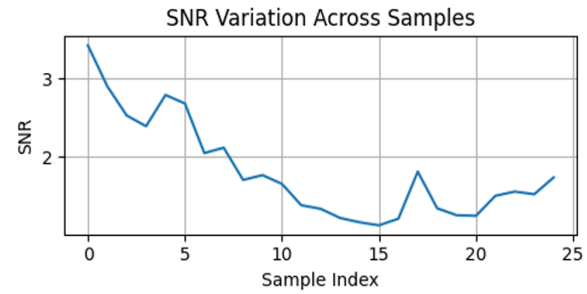
All these image metrics been calculated by the model along with a graphical representation for easier evaluation of the image enhancement.



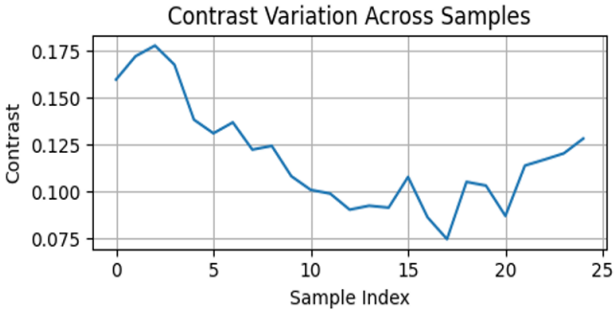
(a)



(b)



(c)



(d)

Figure 4 (a) Entropy, **(b)** Edge Density, **(c)** SNR, **(d)** Contrast.

For the above values of the image metrics, the graphical representation is shown.

VI. CONCLUSION

This work presented an Explainable Spectrum Shift Generative Adversarial Network (SSGAN) for enhancing low-light NIR-IIB in vivo biomedical images using unpaired datasets. The proposed model effectively addresses the limitations associated with the scarcity of paired training data while improving image quality, contrast, and structural visibility. In addition to image enhancement, the system provides explainable outputs in the form of quantitative image metrics and visual structural representations, thereby increasing the transparency and interpretability of the deep learning process. The experimental results demonstrate that the proposed Python-based deep learning framework improves reproducibility, flexibility, and analytical capability, making it suitable for real-time clinical imaging applications. By enabling safe, accurate, and scalable enhancement of NIR-II biomedical images, the model supports improved visualization and clinical decision-making. Therefore, the proposed approach offers a promising and reliable solution for advanced in vivo biomedical imaging, and it can be further extended to other medical imaging modalities to enhance diagnostic performance and clinical outcomes.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Z. Zhang, Y. Du, X. Shi, K. Wang, Q. Qu, and Q. Liang et al., "NIR-II light in clinical oncology: Opportunities and challenges," Nature Reviews Clinical Oncology, vol. 21, pp. 449–467, 2024.

[2] B. Chang, D. Li, Y. Ren, C. Qu, X. Shi, and R. Liu et al., "A phosphorescent probe for in vivo imaging in the second near-infrared window," Nature Biomedical Engineering, vol. 6, pp. 629–639, 2022.

[3] Z. Hu, C. Fang, B. Li, Z. Zhang, C. Cao, and M. Cai et al., "First-in-human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-infrared-I/II windows," Nature Biomedical Engineering, vol. 4, pp. 259–271, 2020.

[4] K. Wang, Y. Du, Z. Zhang, K. He, Z. Cheng, and L. Yin et al., "Fluorescence image-guided tumour surgery," Nature Reviews Bioengineering, vol. 1, pp. 161–179, 2023.

[5] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," Medical Image Analysis, vol. 58, pp. 101552, 2023.

[6] J. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232.

[7] L. Fu, B. Lu, J. Tian, and Z. Hu, "Pssgan: Towards spectrum shift based perceptual quality enhancement for fluorescence imaging," Computerized Medical Imaging and Graphics, vol. 107, pp. 102216, 2023.

[8] O. Biran and C. Cotton, "Explanation and justification in machine learning: A survey," in Proceedings of the International Joint Conferences on Artificial Intelligence workshop on explainable AI (XAI), 2017, pp. 8–13.

[9] Z. Ma, F. Wang, W. Wang, Y. Zhong, and H. Dai, "Deep learning for in vivo near-infrared imaging," Proceedings of the National Academy of Sciences, vol. 118, 2021.

[10] K. Simonyan, "Deep inside convolutional networks: Visualising imageclassification models and saliency maps," 2013.