EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

Enhanced visual insights and diagnosis of interstitial lung diseases via Mufinet-DCGAN framework

Jayalakshmi Ramachandran Nair*[1)], Sumathy Pichai Pillai[1)] and Rajkumar Narayanan[2)]

[1)]Department of Computer Science & Applications, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India
[2)]Department of Sciences, St. Claret College, Bangalore, Karnataka, India

Received 13 September 2024
Revised 28 February 2025
Accepted 1 April 2025

Abstract

Accurate identification and classification of medical images are pivotal in recent medical diagnostics. Despite considerable advancements in deep learning, current methodologies face challenges in capturing nuanced details, particularly from the perspective of interstitial lung diseases (ILDs). Moreover, there is a prominent gap in the investigation of integrating sophisticated image enhancement techniques, such as contrast-limited adaptive histogram equalization (CLAHE), and classification strategies leveraging convolutional neural networks (CNNs). This study proposes a novel methodology that synergistically combines the MufiNet-DCGAN approach to enhance image resolution and refine ILD classification. Through rigorous experimentation, our proposed method achieves commendable accuracy (98.75%), precision (98.01%), recall (98.63%), and F1 score (97.99%). These results underscore the potential of the proposed approach to advance medical diagnostics by furnishing robust tools for precise disease detection.

Keywords: Contrast limited adaptive histogram equalization (CLAHE), Interstitial lung diseases (ILDs), Convolutional neural network (CNN), Deep Learning, MufiNet-DCGAN

1. Introduction

In modern clinical practice, acquiring high-quality and high-resolution medical images is essential. However, challenges such as recording time and equipment limitations often impact image spatial resolution. To overcome these challenges, super-resolution (SR) techniques have emerged as post-processing solutions to enhance medical image quality without incurring additional scanning expenses [1]. These techniques employ super-resolving algorithms to generate additional pixels based on low-resolution (LR) images and prior knowledge, thereby improving the resolution of LR images, which are considered degraded versions of high-resolution (HR) images. SR techniques are commonly classified as Single-Image Super Resolution (SISR) or Multi-Image Super Resolution (MISR), depending on the number of input and output images involved.

In contrast to LR images, HR images exhibit intricate informational structures that are typically captured using sophisticated imaging equipment, leading to prolonged acquisition times and reduced signal-to-noise ratios [2]. This study aims to reconstruct HR images from LR counterparts using SR methods, thus circumventing the need for costly imaging equipment. Image SR represents a method to recover high-frequency information in LR images, thereby simplifying the image retrieval process by relying solely on LR images. However, achieving super-resolution remains challenging, particularly in medical imaging, where subtle anatomical details can provide valuable diagnostic insights [3]. For instance, in brain MRI scans, the microscopic features surrounding a tumour can facilitate the diagnosis of its growth rate and origin [4]. Similarly, in retinal imaging, precise identification of fine vessel structures aids in diagnosing vessel swelling, a hallmark of hypertensive retinopathy. Therefore, mitigating artefacts introduced by SR approaches is crucial to ensuring accurate medical image interpretation.

Existing approaches to interstitial lung disease (ILD) classification often struggle to capture subtle features and patterns crucial for accurate diagnosis. Moreover, the scarcity of high-resolution medical images impedes precise ILD analysis and interpretation. This study addresses these challenges by introducing an integrated framework by merging advanced feature extraction and image generation techniques, thus enhancing both ILD classification and image quality. Through the synergistic combination of multiscale fusion residual network (MufiNet) and deep convolutional generative adversarial network (DCGAN) architectures, the study overcomes the limitations of conventional methods, offering a comprehensive solution for ILD analysis and super-resolution. The primary contributions to the proposed work include:

- Introducing a novel hybrid approach that harnesses the strengths of the MufiNet architecture for feature extraction and the DCGAN framework for super-resolution images.
- Training the generator network on LR images and MufiNet-derived attributes to generate high-quality, realistic HR images.
- Employing a comprehensive set of performance metrics, including precision, accuracy, F1-score, recall, structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and mean-square error (MSE), to evaluate both the ILD classification model and the quality of SR images.

*Corresponding author.
Email address: jayabinoy2021@gmail.com
doi: 10.14456/easr.2025.31

2. Background of the study

2.1 Super-Resolution

Super-resolution (SR) techniques aim to enhance the resolution of low-resolution (LR) images to produce high-resolution (HR) versions [5]. While LR images can be derived from downsampling HR images [6], restoring LR images to HR quality poses challenges, particularly in preserving fine details and sharp edges. Convolutional neural networks (CNNs) with deep learning have emerged as a promising approach for SR due to their capacity to model complex transformations [7]. Recent advancements in SR predominantly rely on CNNs for image and video enhancement [8]. However, current CNN-based SR methods often face issues related to perceptual quality and feature loss during training. Many exercise pixel-wise loss functions such as l2 to curtail mean squared error (MSE) and improve peak signal-to-noise ratio (PSNR) [9]. Nonetheless, these metrics may not adequately consider aesthetic quality, leading to visually inferior and blurred images. Generative adversarial networks (GANs), inspired by CNNs, have indicated remarkable performance in various computer vision tasks [10].

2.2 Generative Adversarial Networks (GANs)

GANs play a pivotal role in the SR of medical images, leveraging LR-HR image pairs for training [11]. A typical GAN comprises a generator (G) and a discriminator (D). The generator learns to produce realistic images to deceive the discriminator and to distinguish between real and fake images. As training progresses, the generator improves output to fool the discriminator, resulting in realistic image generation [12]. This adversarial process leads to the generator producing images closely resembling real data, while the discriminator struggles to differentiate between real and generated samples. Figure 1 illustrates the architecture of a GAN.

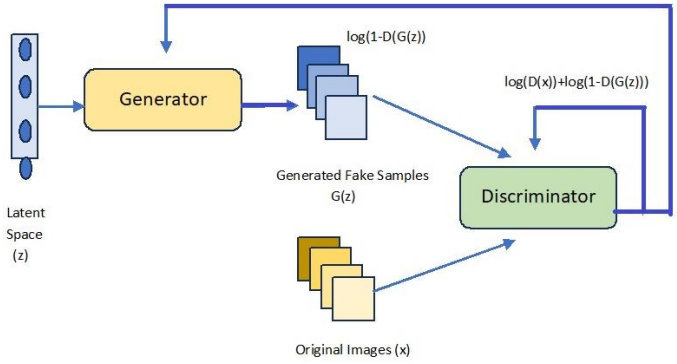

Figure 1 Architecture of GAN [13]

Formula (1) shows the expression for the training of GAN

$$\text{minimax (D, G)} = \mathbb{E}_{x\sim \text{pdata(x)}}\ [\text{logD(x)}] + \mathbb{E}_{z-p_{z(z)}}\ [\log\ (1 - \text{D(G(z))})] \quad (1)$$

where x is the initial image, z is a vector of random values in d dimensions, and $pdata(x)$and *pz(z)* are the probability distributions of x and z, respectively. The likelihood that the input is an image formed from *pdata(x)* is given by D(x) and the likelihood that it is generated from pz(z) is given by *(1- D(G(z)).* G is trained to fool D by decreasing log *(1 – D(G(z))*, while D is trained to raise the correct response.

2.3 Deep Convolutional Generative Adversarial Network (DCGAN)

In 2015, Radford et al. introduced DCGAN, an extension of GAN that utilizes CNNs for unsupervised training [14]. DCGAN enhances the training of the generator network by leveraging CNN's feature extraction capabilities. The generator network comprises several layers, including 2D transposed convolutional layers (fractionally-strided convolutions) to up-sample the noise vector into a recognizable image. Batch normalization is applied to each layer for training stability, except for the output layer. The discriminator, exposed to both real and fake images during training, learns to differentiate between them, driving the generator to produce more realistic outputs over time. Figure 2 showcases the architecture of the DCGAN generator, and Figure 3 depicts the architecture of the DCGAN discriminator.

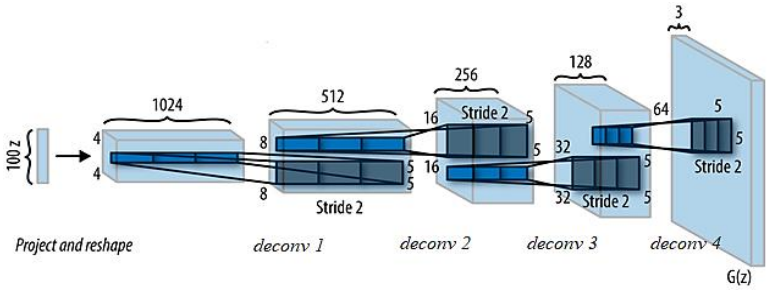

Figure 2 Architecture of DCGAN Generator [14, 15]

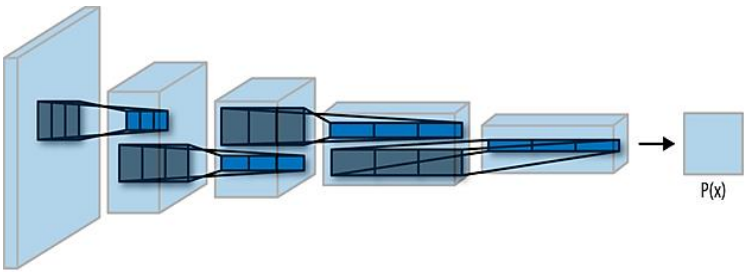

Figure 3 DCGAN Discriminator Architecture [14]

2.4 Super-Resolution Convolutional Neural Network (SRCNN)

SRCNN [16] was the first deep learning approach introduced for single-image super-resolution (SR). It employs a straightforward three-layer convolutional neural network to establish a direct mapping between low-resolution (LR) and high-resolution (HR) images. The input image undergoes bicubic interpolation before entering the network to match the HR image size. The first layer extracts small patches and image features, the second layer applies a non-linear transformation to these features, and the final layer reconstructs the high-resolution output. However, increasing the number of layers does not necessarily enhance performance due to the gradient vanishing issue.

2.5 Fast Super-Resolution Convolutional Neural Network (FSRCNN)

FSRCNN [17] builds upon SRCNN by eliminating the bicubic interpolation step, thereby reducing computational cost and improving processing speed. The network is structured into five main components: feature extraction, shrinking, mapping, expanding, and deconvolution. The shrinking layer before the mapping step compresses the feature dimensions to simplify processing, while the expanding layer afterward refines details for improved HR image generation. A key improvement in FSRCNN is the deconvolution layer, which allows direct learning of the transformation from LR to HR without resizing the input first. This results in both faster processing and enhanced image quality compared to SRCNN.

2.6 Very Deep Super-Resolution Convolutional Networks (VDSR)

VDSR [18] takes a different approach by significantly increasing network depth, using a 20-layer structure with small convolutional filters to expand the receptive field. Training deep networks can be challenging due to slow convergence, so VDSR incorporates residual learning, where the network focuses on predicting the difference between LR and HR images. This approach allows for faster training with higher learning rates. The input images require bicubic interpolation, and zero-padding ensures that feature maps remain the same size, leading to better performance, particularly at image edges. Experiments indicate that using 12 filters per layer is adequate for reconstructing images of space objects.

2.7 Deeply-Recursive Convolutional Networks (DRCN)

DRCN [19] introduces the concept of deep recursive layers to improve SR performance while maintaining a relatively low number of parameters. Unlike traditional deep networks, where each layer has its own parameters, DRCN shares parameters across multiple recursive layers, making the model more efficient. The final output is derived from a weighted combination of multiple recursive layers. Similar to VDSR, bicubic interpolation is applied before inputting the image into the network. While deeper recursion can lead to better results, experiments suggest that using more than five recursive layers provides little additional improvement when reconstructing space object images.

3. Literature review

The study conducted by Hwang et al. [20] aimed to improve the accuracy and consistency of ILD evaluation through computed tomography (CT) image conversion using a routable generative adversarial network (RouteGAN). They assessed the impact of CT image conversion on ILD quantification across various scan settings and manufacturers. The study included ILD patients who underwent thin-section CT scans, and a RouteGAN was utilized to standardize CT images from different acquisition conditions. The results showed that CT conversion using RouteGAN enhanced accuracy and consistency in quantifying ILD, particularly for total abnormalities and specific features like fibrosis score, honeycombing, and reticulation.

Shamrat et al. [21] proposed a novel approach for diagnosing various lung diseases using a neural network algorithm. They focused on classifying 14 lung conditions from the ChestX-ray14 dataset and employed a fine-tuned MobileLungNetV2 framework. Pre-processing techniques such as CLAHE for contrast enhancement and Gaussian Filter for denoising were applied. The study achieved an exceptional classification accuracy of 96.97% with MobileLungNetV2, demonstrating high precision, recall, and specificity scores. Pawar and Talbar [22] introduced a unique two-stage hybrid method for ILD classification using high-resolution computed tomography (HRCT) images. Their approach integrated a conditional generative adversarial network (c-GAN) with multiscale features for accurate lung segmentation and a pre-trained ResNet50 for feature extraction. This led to significant enhancements in accuracy without manual region of interest (ROI) extraction. The study achieved impressive results, including a peak accuracy of 94.65% for healthy cases and an accuracy of 84.12% for the consolidation class.

Teramoto et al. [23] presented an innovative deep-learning approach for classifying various idiopathic interstitial pneumonia (IIP) types, emphasizing the distinction between idiopathic pulmonary fibrosis (IPF) and other pneumonia types [19]. They utilized a two-step training strategy with a progressive growth generative adversarial network (PGGAN) to generate synthetic IIP images alongside real ones, achieving promising results in IPF detection with different CNN models. David et al. [24] proposed a hybridized method

combining local binary pattern (LBP) and histogram of oriented gradients (HOG) feature extraction with a CNN for accurate classification of medical images, particularly for diseases like COVID-19. They achieved high accuracy, precision, and recall rates, highlighting the potential of GANs in generating high-resolution medical images for accurate diagnosis.

Chao et al. [25] addressed the need for high-resolution medical imaging through a novel approach of fusing two orthogonal CT planes and employing SR modelling on a third plane. Their method consistently outperformed other resolution enhancement methods, demonstrating potential clinical and research applications for improving image quality for CT scans. Ma et al. [26] emphasized the significance of early detection of lung diseases through medical imaging using artificial intelligence and deep learning techniques. They discussed the application of deep learning approaches for diagnosing various lung infections, highlighting the potential of artificial intelligence (AI) in enhancing accuracy in detecting and classifying lung infections. Wang et al. [27] proposed a three-phase automated system for pneumonia diagnosis using chest X-rays, demonstrating improved accuracy and potential for further enhancements. Gu et al. [28] introduced a deep learning-based approach, named Medical Images SR using GANs Adversarial Networks (MedSRGAN), for medical image SR, showcasing improved texture preservation and realism in SR images. Kim et al. [29] explored the potential of DCGANs for generating high-resolution medical images from low-resolution ones, demonstrating improved image quality and lesion complexity in breast MRI. Deepak and Ameer [30] utilized deep learning techniques for brain tumour characterization in a computer-aided diagnosis (CAD) system, demonstrating improved performance of classifiers to balance accuracy scores. Baur et al. [31] successfully generated high-resolution, realistic images of skin lesions using GANs, addressing the class imbalance in skin lesion classification through synthesized high-resolution melanoma samples.

Han et al. [32] demonstrated the capability of GANs to generate realistic medical images, particularly multi-sequence brain magnetic resonance (MR) images, showcasing the potential clinical applications of GAN-based image generation in medical imaging. Chen and Tong [33] investigated the application of Wasserstein distance as a training objective for GANs, particularly focusing on Single Image Super-Resolution (SISR), and highlighted the stability and effectiveness of Wasserstein GAN in training progress assessment.

The work by Al-Mekhlafi and Liu [34] in single image super-resolution (SISR) has been driven by deep learning, particularly CNNs, which effectively reconstruct high-resolution images. The team also analysed the role of Generative Adversarial Networks (GANs) and attention mechanisms in improving image quality by capturing finer details. Other major findings of the work include the significance of knowledge distillation techniques in helping to optimize SR models, making them more efficient without losing accuracy, the important applications of SISR in medical imaging, satellite imagery, and security monitoring, where high-resolution images are essential. However, challenges like preserving high-frequency details and reducing computational complexity remain.

4. Materials and methods

The proposed methodology, outlined in Figure 4, delineates a systematic workflow for accurate diagnosis of ILDs leveraging a fusion of computational techniques. Initiating with low-resolution CT images, the process traverses through an integrated architecture combining MufiNet and DCGAN. Within this architectural framework, the initial low-resolution images undergo a transformative journey culminating in the generation of high-resolution CT counterparts. The sophisticated MufiNet-DCGAN architecture facilitates this pivotal enhancement step. Subsequently, the high-resolution CT images are subjected to feature extraction leveraging the prowess of ResNet-50, a pre-trained convolutional neural network (CNN) esteemed for its feature extraction capabilities.

The output of ResNet-50 manifests as a feature vector, encapsulating intricate attributes extracted from the high-resolution images. This feature vector is a bridge, facilitating seamless integration between the enhanced image data and the subsequent ILD classification task. The classification is executed precisely through a support vector machine (SVM) classifier compared with the other classifier, thereby enabling accurate diagnosis of ILD conditions. In the SVM classifier, kernel functions help transform the input data into a higher-dimensional space, making it easier to separate classes using a straight line or hyperplane. To analyse complex patterns, the common kernel functions considered in this study, are the polynomial kernel and radial basis function (RBF) kernel functions. The other parameters of SVM that are considered in this work are regularization parameter, kernel coefficient (gamma) for polynomial & RBF kernels, degree and tolerance.

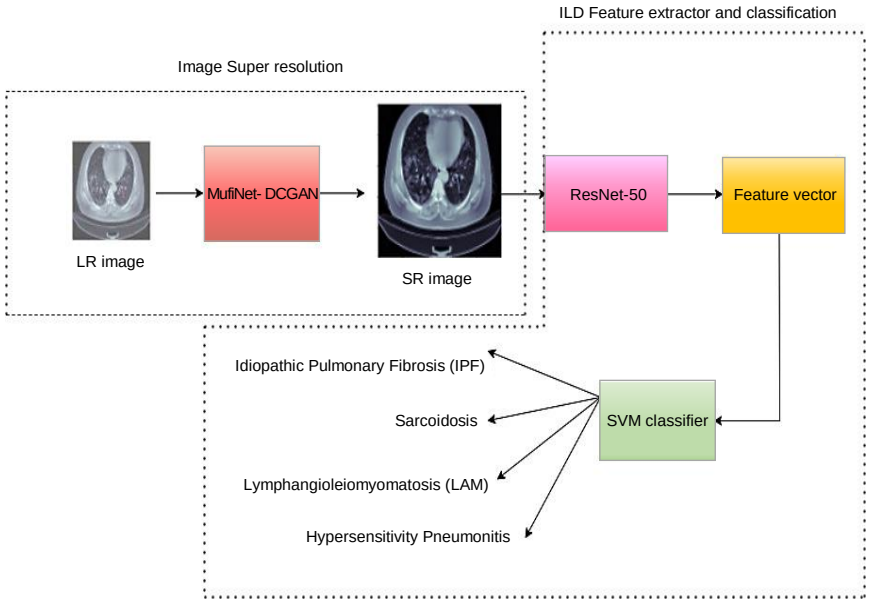

Figure 4 Architecture of the proposed approach

4.1 Dataset overview

The database contains lung CT scan images, collected from KIMS Hospital in Thiruvananthapuram, Kerala, India. It focuses mainly on treating interstitial lung diseases, including idiopathic pulmonary fibrosis, sarcoidosis, lymphangioleiomyomatosis, and hypersensitivity pneumonitis. The data were collected from July 2023 to December 2023 and included the scanning of 500 patients between the ages of 18 and 85 years; there was an approximately equal representation of about 60% male and 40% female among the patient population, typical of the region's diverse ethnic representation. Additionally, to ensure robust model evaluation, a standard 9:1 split is applied to divide the dataset into training and testing subsets.

The CT scans were acquired on a Siemens SOMATOM Definition AS scanner, using a tube voltage of 120 kVp, tube current of 200 mA, and slice thickness of 1 mm. Images were reconstructed using an iterative reconstruction algorithm, with a field of view (FOV) set at 350 mm. All relevant scans are in digital imaging and communications in medicine (DICOM) format and stored in the picture archiving and communication system (PACS). The annotations were carried out by the team of radiologists and pulmonologists from KIMS Hospital diligently, using home-developed software compatible with DICOM.

This clinical research maximally respects ethical considerations. Informed written consents were obtained from the patients and, wherever applicable, had patient-identifiable particulars and images expunged to protect their privacy. It was first approved by the hospital's Institutional Review Board: KIMS Hospital.

4.2 Preprocessing and data augmentation

Preprocessing of the collected CT images precedes their utilization in the model. This preprocessing involves the application of CLAHE to enhance contrast and detail within the images. Additionally, cropping techniques are employed to isolate and focus on the relevant lung regions affected by ILD. These preprocessing steps aim to enhance the discriminatory features within the images, thus facilitating more effective training of the subsequent classification model. Furthermore, to augment the diversity and robustness of the dataset, various augmentation techniques are applied. These techniques encompass operations such as flipping, scaling, and rotating the images. However, to maintain the integrity of ILD characteristics, transformations that may distort the shape of the images, such as shearing, are deliberately excluded. Figure 5 and Figure 6 provide the visualisation of proposed preprocessed and augmented images respectively. By amalgamating these preprocessing and augmentation techniques, the dataset is meticulously prepared to ensure optimal training efficacy and robustness of the subsequent ILD classification model.

4.3 Proposed methodology

The integration of MufiNet and DCGAN in this study introduces innovative advancements in image processing, particularly in enhancing the quality of medical images for the diagnosis of ILDs. The overall framework of the proposed MufiNet-DCGAN network is illustrated in Figure 7.

4.3.1 Architecture of the MufiNet DCGAN

MufiNet serves as a crucial component in the fusion of LR and HR CT images [35]. LR CT images provide information about the overall lung structure, while HR CT images offer details about fine lung structures. The fusion process in MufiNet integrates these complementary sources of information to generate super-resolved CT images that preserve both overall structure and fine details.

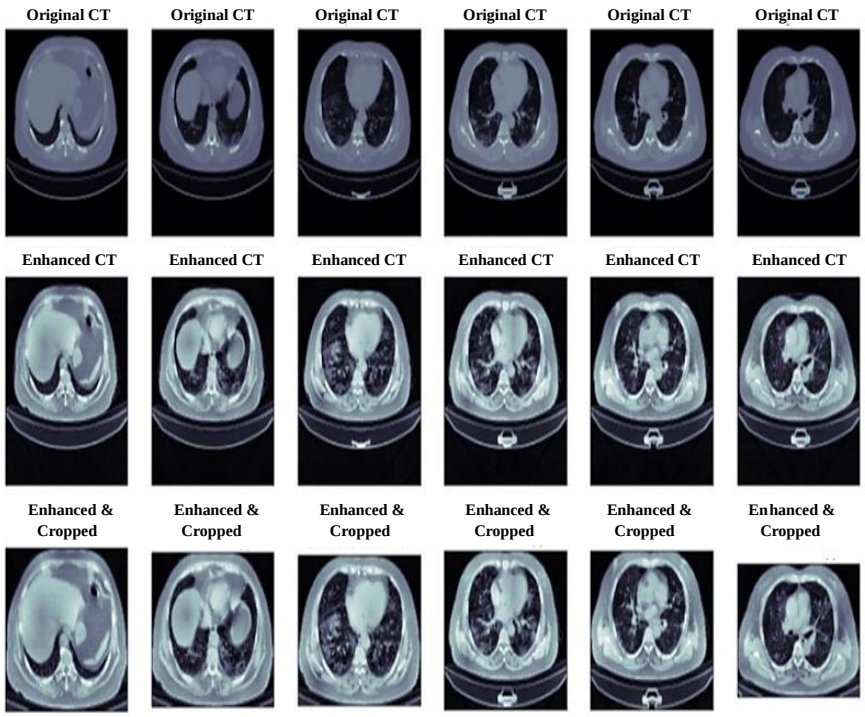

Figure 5 Preprocessed images

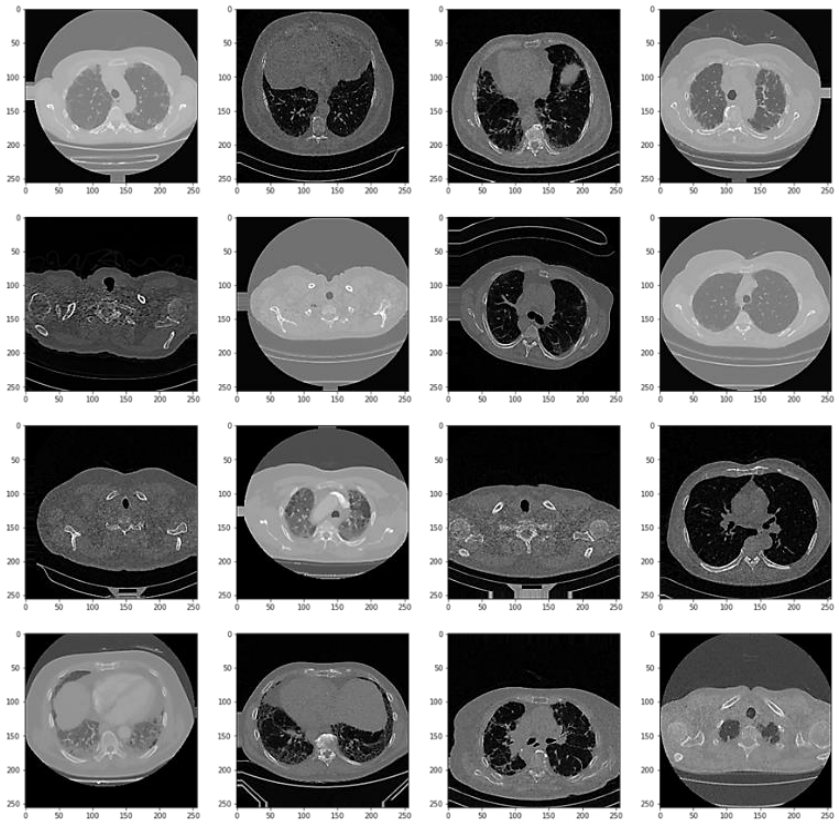

Figure 6 Augmented images

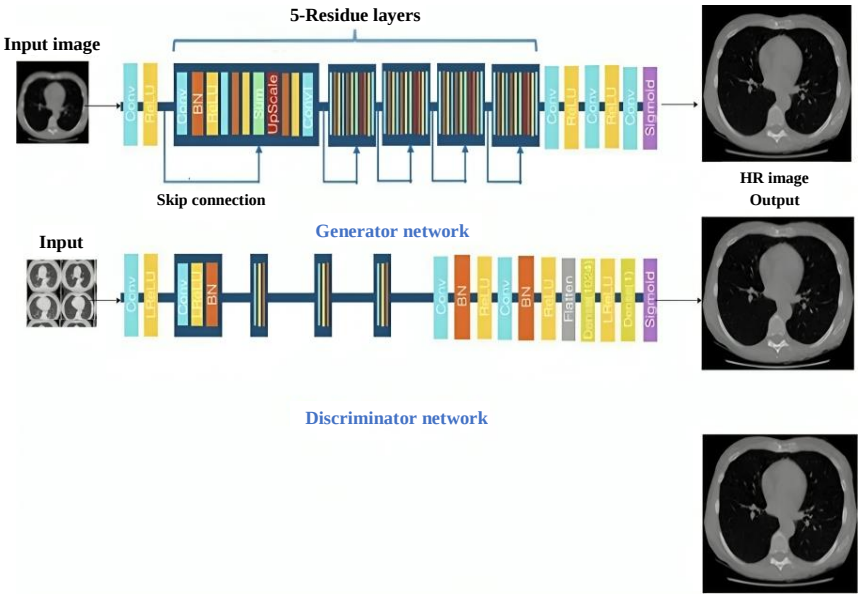

Figure 7 Architecture of the MufiNet – DCGAN

MufiNet comprises two main modules: the feature fusion module and the residual learning module. The feature fusion module extracts feature from both LR and HR images and merges them through concatenation and fully connected layers. This merged feature set retains crucial details from both sources, enhancing the diagnostic potential of ILD images. Meanwhile, the residual learning module, presented in Figure 8, enhances the quality of reconstructed images by learning the difference between LR and HR images. Through a series of convolutional layers, this module removes noise and artefacts, contributing to the generation of high-quality reconstructed images.

For the classification of ILD, SR CT lung images serve as input, and the output is the corresponding ILD class label [36]. The process involves enhancing lung regions within HR CT images and extracting features for classification using a memory-efficient classifier. ResNet50, a pre-trained CNN, effectively extracts deep features from SR images, aiding in classification tasks.

Based on the "fit class error-correcting output codes (fitcecoc)" function, the classifier utilizes deep features extracted by ResNet50 for multi-class error-correcting output. This approach leverages diverse deep features to accurately classify different ILDs, contributing to a robust and precise diagnosis. Overall, the integrated MufiNet-DCGAN architecture, coupled with ResNet50 and classifiers, offers a comprehensive and efficient framework for enhancing medical images and accurately diagnosing ILDs.

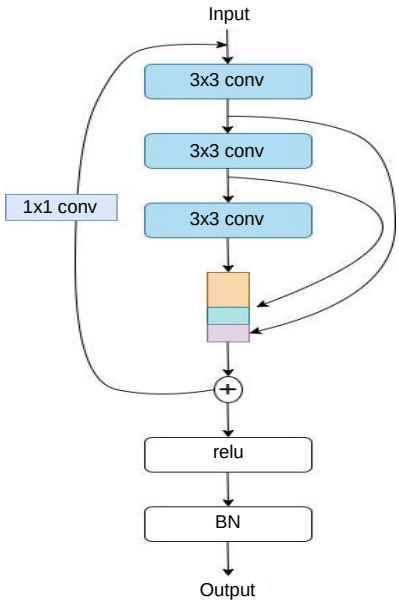

Figure 8 Multi-scale fusion residual block

5. Results and discussion

5.1 System specifications

The experiments were conducted on a machine equipped with an Intel Core i7 processor, 16GB RAM, and an NVIDIA GeForce GTX 1080Ti GPU. Google Colab, a cloud-based platform, facilitated efficient deep-learning model training by providing access to high-performance GPUs. The implementation of the MufiNet-DCGAN architecture was carried out using Python, leveraging popular deep-learning libraries such as TensorFlow and Keras.

Hyperparameters, the essential settings predefined before training a machine learning model, played a pivotal role in guiding the learning process. Their values were fixed throughout training and were predefined as shown in Table 1.

Table 1 List of hyperparameters.

Hyperparameters	Values
Batch size	32
Learning rate	0.001
Dropout	0.2
Optimizer	Adam
Epochs	280
Activation	ReLU
Loss function	Categorical cross-entropy

5.2 Data Specifications

The complete acquired dataset consisted of a total of 4100 images, and we have divided it into training and testing sets in a 9:1 ratio. Thus, there would be 3690 training samples and 410 test samples in the dataset. The distribution of each class among the four classes is provided in Table 2.

Table 2 Distribution of samples among train and test sets.

Class Name	Train Samples	Test Samples
IPF	909	101
Sarcoidosis	927	103
LAM	925	105
Hypersensitivity Pneumonitis	909	101

5.3 Evaluation protocol

We have utilized different evaluation methods to ensure the effectiveness of our proposed system. We have used performance metrics such as accuracy, precision, recall and f-measure to assess the robustness of the proposed classification methodology. Additionally, in the case of the SR technique, image comparison techniques such as PSNR, SSIM along MSE have been examined. The equations of all metrics used are given in the equations (2)-(8).

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Samples} \quad (2)$$

$$Precision = \frac{True\ Positives}{True\ Positives+False\ Positives} \quad (3)$$

$$Recall = \frac{True\ Positives}{True\ Positives+False\ Negatives} \quad (4)$$

$$F - Measure = \frac{2}{(1/Precision + 1/Recall)} \quad (5)$$

Where true positives are the number of correctly predicted positive samples, true negatives are the number of correctly predicted negative samples, false positives are the samples incorrectly predicted as positive and false negatives are the samples incorrectly predicted as negative class.

$$MSE = \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n} \quad (6)$$

$$PSNR = 10.log_{10}(\max(x)^2/MSE) \quad (7)$$

$$SSIM = \frac{(2\mu_x\mu_y+c_1)+(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \quad (8)$$

where x and y are pixel values in their corresponding images, μ refers to the mean value, σ refers to the covariance and c are variables.

5.4 Super-Resolution of images

To deceive the discriminator initially, a matrix of uniform plane values was assigned as presented in Figure 9, establishing a baseline for the generator to produce images. As training progressed, the generator refined its ability to create images that successfully fooled the discriminator, resulting in more realistic and accurate images. This iterative process enabled the generator to learn and capture complex features and patterns specific to ILDs.

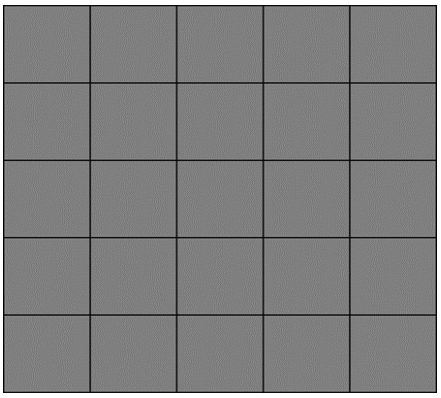

Figure 9 Initial grid of plane values for discriminator deception

Subsequently, the model underwent a training process spanning 280 epochs, with careful monitoring of its progression. At epoch 1, the generator's loss was typically high, reflecting its struggle to produce coherent and realistic images. The resolution of input at epoch 1 is shown in Figure 10. Conversely, the discriminator's loss was lower, indicating its capability to differentiate between real and generated images. As training progressed, both losses evolved, ideally converging to equilibrium, where the generator effectively produced images that deceived the discriminator with increased proficiency. Figures 11 and 12 present insightful visualizations of the evolution of both the generator and discriminator losses at two distinct epochs.

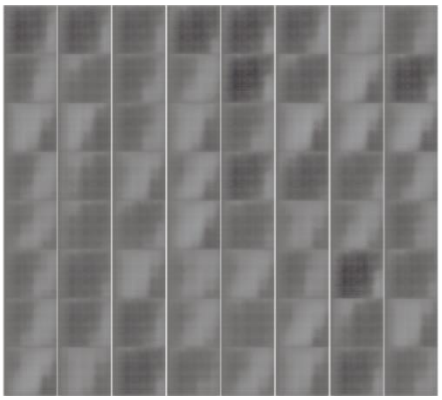

Figure 10 Resolution of the image at epoch 1

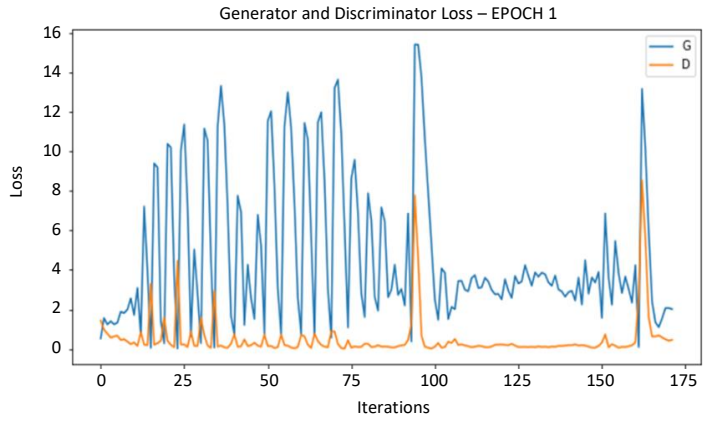

Figure 11 Loss of the framework at epoch 1

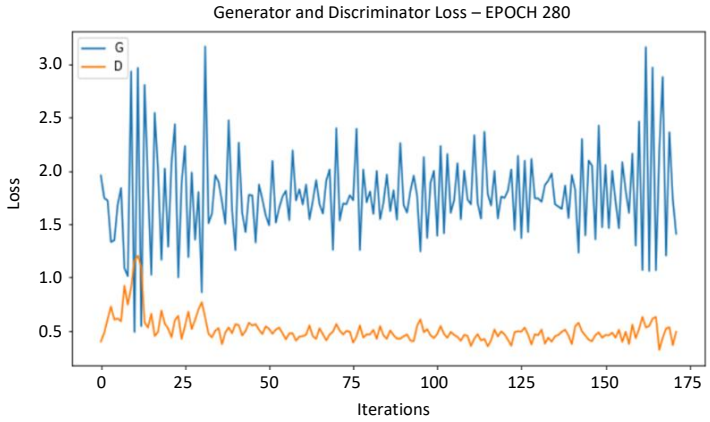

Figure 12 Loss of the framework at epoch 280

Figure 13 presents a comparative evaluation of image quality metrics for different images. The results underscored the positive impact of the proposed super-resolution technique on image quality. The consistent increase in PSNR and SSIM values, coupled with the decrease in MSE values, highlighted the successful enhancement of images, imperative for accurate diagnosis and interpretation of ILDs. Table 3 presents a comparison of our methodology against existing models.

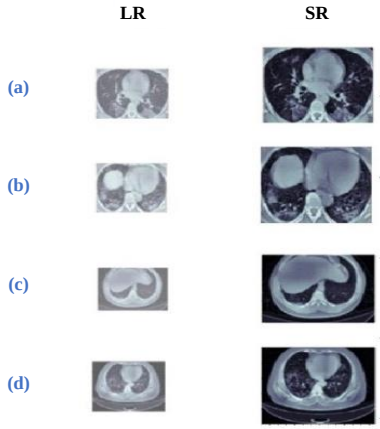

Image	Original PSNR	SR PSNR	Original SSIM	SR SSIM	Original MSE	SR MSE
(a)	30.14	34.67	0.85	0.91	0.0082	0.0035
(b)	28.76	33.42	0.82	0.89	0.0105	0.0041
(c)	29.89	35.21	0.84	0.92	0.0091	0.0032
(d)	27.65	32.08	0.80	0.87	0.0123	0.0049

Figure 13 Performance evaluation of super-resolution images

Table 3 Performance comparison

Method	PSNR
MedSRGAN+	31.70
SRCNN	31.88
FSRCNN	32.62
VSDR	32.74
DRCN	32.91
DCGAN LR and HR Pre and Post	33.13
Proposed	34.08

5.5 ILD Classification

In the preceding sections, the deployment of deep learning techniques for the super-resolution enhancement of low-resolution CT images of ILDs was discussed, enhancing the visual quality and diagnostic potential of the images. Building upon this foundational enhancement, the focus shifted to the subsequent step in the study: the classification of ILDs using enriched and enhanced images. The loss and accuracy is identified using the training and testing data set as shown in Figure 14.

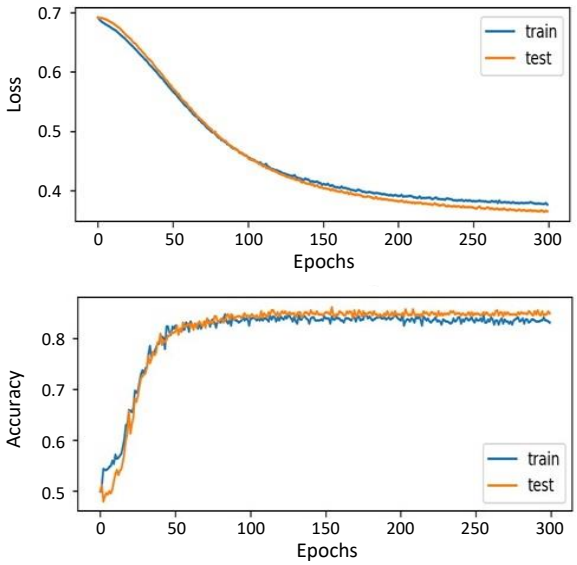

Figure 14 Accuracy and loss plots

Additionally, the proposed approach provides the confusion matrix in Figure 15 along with a comparison graph of the obtained performance metrics in Figure 16.

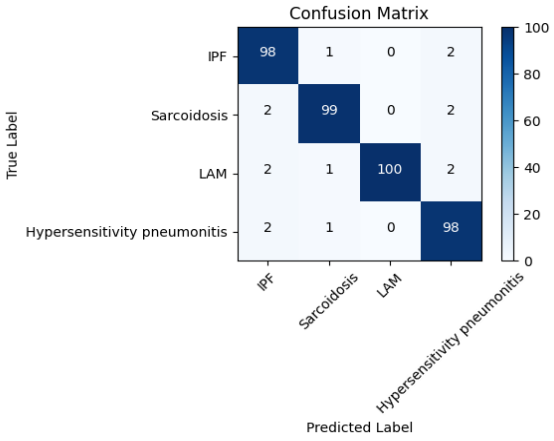

Figure 15 Confusion matrix obtained

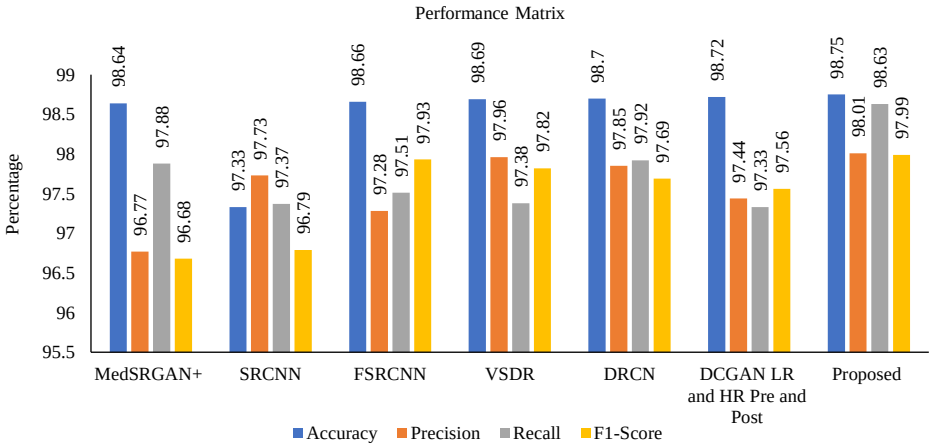

Figure 16 Comparison of different performance metrics.

6. Conclusion

The study presents a comprehensive exploration of deep learning methodologies tailored to enhance medical image quality and facilitate precise disease diagnosis, specifically focusing on ILDs. By integrating advanced techniques such as super-resolution using the MufiNet-DCGAN architecture, the research endeavours to bridge the gap in capturing nuanced details essential for accurate medical diagnostics. Through rigorous experimentation and evaluation, the proposed framework demonstrates commendable performance metrics, including high accuracy, precision, recall, and F1 score, underscoring its potential to advance medical diagnostics. Overall, this research represents a significant stride towards enhancing medical imaging techniques and refining disease classification methodologies, thereby furnishing robust tools for precise disease detection and improved patient care.

7. Novelty and Future Enhancement

The novelty of this study lies in proposing the use of SR image generation specifically for ILD classification, which, to our knowledge, has not been extensively applied to this particular medical imaging problem. While SR techniques have been utilized in other medical image classification problems, their application to ILDs presents a unique challenge and opportunity. Additionally, the use of a different GAN and CNN architecture, specifically the MufiNet-DCGAN, distinguishes this work from current methods applied for ILD image classification. This innovative approach allows for enhanced image resolution and better feature extraction, which are crucial for accurate diagnosis.

While the current study achieved promising results in super-resolution enhancement and ILD classification, there are opportunities for future research and enhancements. Future research can improve the limitations of the proposed methodology, such as high computational cost and the risk of overfitting for better outcome and complex diagnosis of ILDs. Further investigation into alternative architectures, dataset augmentation techniques, and optimization strategies could potentially enhance the robustness and generalization capability of the proposed framework.

8. Acknowledgements

The research work was supported by the Department of Computer Science and Applications at Bharathidasan University and KIMS Hospital.

9. References

[1] Li Q, Qu H, Liu Z, Zhou N, Sun W, Sigg S. AF-DCGAN: Amplitude feature deep convolutional GAN for fingerprint construction in indoor localization systems. IEEE Trans Emerg Top Comput Intell. 2021;5(3):468-80.
[2] Plenge E, Poot DHJ, Bernsen MR, Kotek G, Houston G, Wielopolski P, et al. Super-resolution methods in MRI: Can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time?. Magn Reson Med. 2012;68(6):1983-93.
[3] Guo K, He Y, Kui X, Sehdev P, Chi T, Zhang R, et al. LLTO: Towards efficient lesion localization based on template occlusion strategy in intelligent diagnosis. Pattern Recognit Lett. 2018;116:225-32.
[4] Tanno R, Worrall DE, Ghosh A, Kaden E, Sotiropoulos SN, Criminisi A, et al. Bayesian image quality transfer with CNNs: exploring uncertainty in dMRI super-resolution. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins D, Duchesne S, editors. Medical Image Computing and Computer Assisted Intervention − MICCAI 2017. Lecture Notes in Computer Science. Cham: Springer; 2017. p. 611-9.
[5] Xiang Q, Peng L, Pang X. Image DAEs based on residual entropy maximum. IET Image Process. 2020;14(6):1164-9.
[6] Kaji S, Kida S. Overview of image-to-image translation by use of deep neural networks: denoising, super-resolution, modality conversion, and reconstruction in medical imaging. Radiol Phys Technol. 2019;12(3):235-48.
[7] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [Internet]. arXiv [Preprint]. 2015 [cited 2024 Sep 3]. Available from: https://arxiv.org/abs/1502.03167.
[8] Li Y, Sixou B, Peyrin F. A review of the deep learning methods for medical images super resolution problems. IRBM. 2021;42(2):120-33.
[9] Zareapoor M, Emre Celebi M, Yang J. Diverse adversarial network for image super-resolution. Signal Process Image Commun. 2019;74:191-200.
[10] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Networks [Internet]. arXiv [Preprint]. 2014 [cited 2024 Sep 3]. Available from: https://arxiv.org/abs/1406.2661.
[11] Xu K, Cao J, Xia K, Yang H, Zhu J, Wu C. Multichannel residual conditional GAN-leveraged abdominal pseudo-CT generation via Dixon MR images. IEEE Access. 2019;7:163823-30.
[12] Kim DW, Chung JR, Jung SW. GRDN: Grouped residual dense network for real image denoising and GAN-based real-world noise modeling [Internet]. arXiv [Preprint]. 2019 [cited 2024 Sep 3]. Available from: https://arxiv.org/abs/1905.11172.
[13] Abe K, Iwana BK, Holmér VG, Uchida S. Font creation using class discriminative deep convolutional generative adversarial networks. Proceedings of the 4th Asian Conference on Pattern Recognition (ACPR); 2017 Nov 26-29; Nanjing, China. Piscataway: IEEE; 2017. p. 238-43.
[14] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [Internet]. arXiv [Preprint]. 2015 [cited 2024 Sep 3]. Available from: http://arxiv.org/abs/1511.06434.
[15] Cao Y, Jia L, Chen YX, Lin N, Yang C, Zhang B. Recent advances of generative adversarial networks in computer vision. IEEE Access. 2019;7:14985-5006.
[16] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. IEEE Trans Pattern Anal Mach Intell. 2016;38(2):295-307.
[17] Dong C, Loy CC, Tang X. Accelerating the super-resolution convolutional neural network. In: Leibe B, Matas J, Sebe N, Welling M, editors. Computer Vision–ECCV 2016. Lecture Notes in Computer Science. Cham: Springer; 2016. p. 391-407.
[18] Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27-30; Las Vegas, USA. Piscataway: IEEE; 2016. p. 1646-54.

[19] Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27-30; Las Vegas, USA. Piscataway: IEEE; 2016. p. 1637-45.
[20] Hwang HJ, Kim HJ, Seo JB, Ye JC, Oh G, Lee SM, et al. Generative adversarial network-based image conversion among different computed tomography protocols and vendors: effects on accuracy and variability in quantifying regional disease patterns of interstitial lung disease. Korean J Radiol. 2023;24(8):807-20.
[21] Shamrat FMJM, Azam S, Karim A, Ahmed K, Bui FM, De Boer F. High-precision multiclass classification of lung disease through customized MobileNetV2 from chest X-ray images. Comput Biol Med. 2023;155:106646.
[22] Pawar SP, Talbar SN. Two-stage hybrid approach of deep learning networks for interstitial lung disease classification. Biomed Res Int. 2022;2022:7340902.
[23] Teramoto A, Tsukamoto T, Michiba A, Kiriyama Y, Sakurai E, Imaizumi K, et al. Automated classification of idiopathic pulmonary fibrosis in pathological images using convolutional neural network and generative adversarial networks. Diagnostics (Basel). 2022;12(12):3195.
[24] David SI, Bashir SA, Mohammed AD. A concept-based review on generative adversarial network for generating super-resolution medical image using SWOT analysis. The 5th Information Technology for Education and Development (ITED); 2022 Nov 1-3; Abuja, Nigeria. Piscataway: IEEE; 2022. p. 1-6.
[25] Chao HL, Wu Y, Siana L, Chen YM. Generating high-resolution CT slices from two image series using deep-learning-based resolution enhancement methods. Diagnostics. 2022;12(11):2725.
[26] Ma J, Song Y, Tian X, Hua Y, Zhang R, Wu J. Survey on deep learning for pulmonary medical imaging. Front Med. 2019;14(4):450-69.
[27] Wang Z, Hall J, Haddad RJ. Improving pneumonia diagnosis accuracy via systematic convolutional neural network-based image enhancement. SoutheastCon 2021; 2021 Mar 10-13; Atlanta, USA. Piscataway: IEEE; 2021. p. 1-6.
[28] Gu Y, Zeng Z, Chen H, Wei J, Zhang Y, Chen B, et al. MedSRGAN: medical images super-resolution using generative adversarial networks. Multimed Tools Appl. 2020;79(29-30):21815-40.
[29] Kim S, Qu L, Lian C, Pan Y, Hu D, Xia B, et al. High-resolution breast MRI reconstruction using a deep convolutional generative adversarial network. J Magn Reson Imaging. 2020;52(6):1852-8.
[30] Deepak S, Ameer PM. MSG-GAN based synthesis of brain MRI with meningioma for data augmentation. 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT); 2020 Jul 2-4; Bangalore, India. Piscataway: IEEE; 2020. p. 1-6.
[31] Baur C, Albarqouni S, Navab N. MelanoGANs: high-resolution skin lesion synthesis with GANs [Internet]. arXiv [Preprint]. 2018 [cited 2024 Sep 3]. Available from: https://arxiv.org/abs/1804.04338.
[32] Han C, Hayashi H, Rundo L, Araki R, Shimoda W, Muramatsu S. GAN-based synthetic brain MR image generation. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); 2018 Apr 4-7; Washington, USA. Piscataway: IEEE; 2018. p. 734-8.
[33] Chen Z, Tong Y. Face super-resolution through Wasserstein GANS [Internet]. arXiv [Preprint]. 2017. [cited 2024 Sep 3]. Available from: https://arxiv.org/abs/1705.02438.
[34] Al-Mekhlafi H, Liu S. Single image super-resolution: a comprehensive review and recent insight. Front Comput Sci. 2023;18(1):181702.
[35] Wang C, Wang Z, Xi W, Yang Z, Bai G, Wang R. MufiNet: multiscale fusion residual networks for medical image segmentation. International Joint Conference on Neural Networks (IJCNN); 2020 Jul 19-24; Glasgow, UK. Piscataway: IEEE; 2020. p. 1-7.
[36] Exarchos KP, Gkrepi G, Kostikas K, Gogali A. Recent advances of artificial intelligence applications in interstitial lung diseases. Diagnostics (Basel). 2023;13(13):2303.