International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 04 | Apr 2025 www.irjet.net p-ISSN: 2395-0072

Real-ESRGAN Based Super-Resolution for Low-Quality Chest X-Ray

Ayush Patel[1], Prof. (Dr.) Tejas Shah[2]

[1]Ayush Patel, Research Scholar, Instrumentation & Control Dept., L.D College of Engineering, Ahmedabad, Gujarat, India

[2]Prof. (Dr.) Tejas Shah, Associate Professor, Instrumentation & Control Dept., L.D College of Engineering, Ahmedabad, Gujarat, India

---***---

Abstract - *This work introduces a super-resolution method for improving low-quality chest X-ray images with a fine-tuned Real-ESRGAN model. High-resolution chest X-rays from a public database were artificially degraded to produce low-resolution images, which were restored using the trained super-resolution model. The quality of the restored images was assessed by comparison with ground truth using PSNR, SSIM, and MSE metrics. Also, a pre-trained CNN-based pneumonia classifier was evaluated with both low-resolution and super-resolved images. Outcomes show that the Real-ESRGAN model effectively enhances image quality without sacrificing diagnostic characteristics.*

Key Words: **Real-ESRGAN, Medical Image Enhancement, Chest X-Ray, Convolutional Neural Networks (CNN), Medical Imaging**

1.INTRODUCTION

Chest X-ray imaging is arguably the most applied diagnostic tool used to detect diverse lung conditions, thanks to its speed, affordability, and access. Nevertheless, in most realistic situations i.e., in mobile health setups, remote settings, or telemedicine the resultant images can exhibit low resolution, which can result in impeded accurate interpretation as well as computerized analysis. It is paramount to improve such low-resolution image quality to help ensure diagnostic credibility in such limiting environments.

Super-resolution (SR) methods, especially deep learning-based SR methods, have been highly promising in recovering visual details from degraded medical images. Here, we investigate the application of Real-ESRGAN, a strong real-world image restoration generative adversarial network, to improve low-resolution chest X-rays.

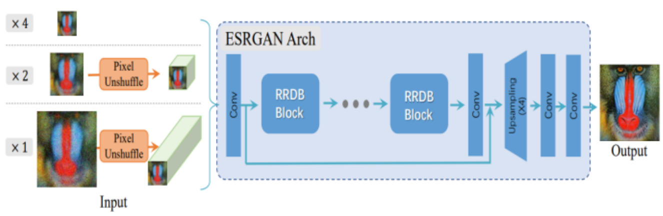


Fig -1: Architecture of the Real-ESRGAN generator based on RRDB blocks and pixel unshuffle input handling.

This generator structure processes multi-scale low-resolution inputs via pixel unshuffle, extracts deep features using Residual-in-Residual Dense Blocks (RRDB), and upsamples to produce high-quality super-resolved images.

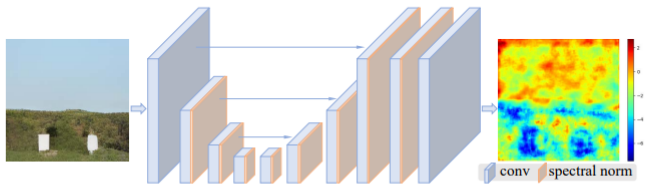


Fig -2: U-Net-based discriminator architecture with spectral normalization used in Real-ESRGAN.

The discriminator adopts a U-Net style layout to capture both local and global image artifacts, with spectral normalization stabilizing training and improving adversarial feedback quality.

We used the pre trained CNN model on a pneumonia X-ray image dataset and measure the quality of reconstructed images using conventional metrics including PSNR, SSIM, and MSE. CNN consists of Convolutional layers, pooling layers & ANN consists of hidden layers and output layers.

To develop the model author have used 5 Conv2D layers which are followed by maxpooling layers for every Conv2D layer [1] Then for classification Artificial Neural Network with 2 hidden layers and one output layer which has a single neuron is used.

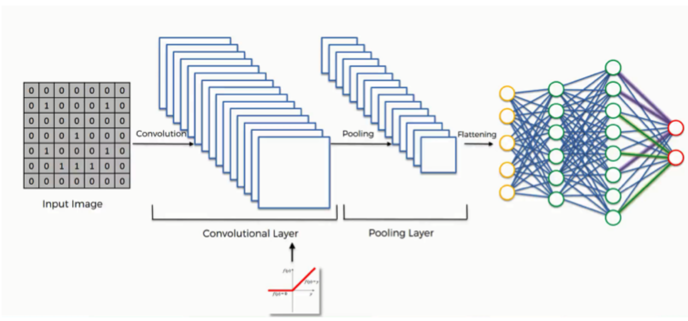


Fig -3: Convolutional Neural Network (Reproduced from [1])

We also measure the on-classification accuracy using a pre-trained CNN model we get by reconstructed images and to confirm that diagnostic features are not affected during upscaling.

2. METHODOLOGY

2.1 Workflow Overview

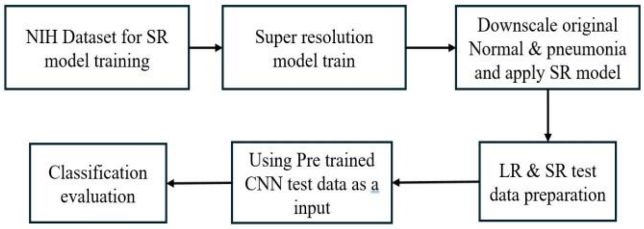


Fig -4: Flow process of Methodology

The overall workflow consists of training a super-resolution model using the NIH chest X-ray dataset. After training, the model is applied to synthetically downscaled images from a pneumonia dataset to generate enhanced versions. Both low-resolution (LR) and super-resolved (SR) images are then used as input to a pre-trained CNN classifier.

Classification performance is evaluated on both LR and SR versions to assess the effectiveness of the enhancement process.

2.2 Dataset Preparation

A random sample from the NIH Chest X-ray Dataset was used to fine-tune the Real-ESRGAN model for super-resolution tasks. Total 8100 images were used to fine tune Real-ESRGAN model. This dataset provided a variety of chest X-ray images as multiscale of original & multiscale cropped images were generated using data augmentation which is suitable for training the model on general radiographic patterns.

For classification evaluation. A balanced Chest X-ray Dataset was used. Subset of 1000 normal and 1000 pneumonia images was selected. High-resolution images were downscaled using bicubic interpolation to generate low-resolution inputs.

The trained SR model was then applied to restore these images, and both LR and SR versions were prepared for testing with a pre-trained CNN classifier

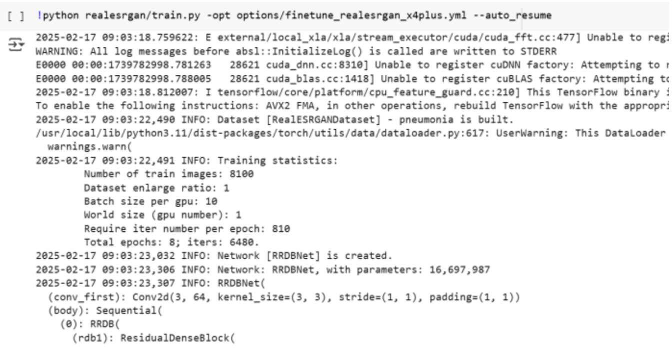
2.3 Real-ESRGAN Fine-Tuning

The chosen generator in this setup is RRDBNet, using the pre-trained RealESRGAN_x4plus.pth. It is set with 3 input and output channels, 64 feature maps, 23 residual-in-residual dense blocks, and 32 growth channels. sThe discriminator employed is UNetDiscriminatorSN, which is loaded from RealESRGAN_x4plus_netD.pth, also having 3 input channels and 64 feature maps, and supports skip connections for better learning stability and performance.

The datadataset with 8100 high-resolution images which consist original 3600 ,3600 multiscale images and 900 multiscale cropped images, whose paths are enumerated in a metadata file. Simple data augmentation is applied, namely horizontal flipping, which is turned on to increase diversity at training time. Rotation is not used. These augmentations are small but useful for making the model more robust on the medical imaging data.

A sophisticated two-stage degradation pipeline is employed to synthesize low-resolution inputs from high-resolution images. During the first stage, there is a 20% probability of upscaling, a 70% probability of downscaling, and a 10% probability of retaining the resolution.

Gaussian noise is added 50% with a noise range from 1–30; Poisson noise is randomly scaled between 0.05–3; grayscale noise is used 40% of the time; and JPEG compression is between quality 30 and 95. During the second stage, blur is used 80% of the time with different kernel types, resizing is done with probabilities 30% up, 40% down, and 30% keep, and the same noise and compression methods are employed with slightly different parameters.



```
[ ] !python realesrgan/train.py -opt options/finetune_realesrgan_x4plus.yml --auto_resume
2025-02-17 09:03:18.759622: E external/local_xla/xla/stream_executor/cuda/cuda_fft.cc:477] Unable to regi
WARNING: All log messages before absl::InitializeLog() is called are written to STDERR
E0000 00:00:1739782998.781263   28621 cuda_dnn.cc:8310] Unable to register cuDNN factory: Attempting to r
E0000 00:00:1739782998.788005   28621 cuda_blas.cc:1418] Unable to register cuBLAS factory: Attempting to
2025-02-17 09:03:18.812007: I tensorflow/core/platform/cpu_feature_guard.cc:210] This TensorFlow binary i
To enable the following instructions: AVX2 FMA, in other operations, rebuild TensorFlow with the appropri
2025-02-17 09:03:22,490 INFO: Dataset [RealESRGANDataset] - pneumonia is built.
/usr/local/lib/python3.11/dist-packages/torch/utils/data/dataloader.py:617: UserWarning: This DataLoader
  warnings.warn(
2025-02-17 09:03:22,491 INFO: Training statistics:
        Number of train images: 8100
        Dataset enlarge ratio: 1
        Batch size per gpu: 10
        World size (gpu number): 1
        Require iter number per epoch: 810
        Total epochs: 8; iters: 6480.
2025-02-17 09:03:23,032 INFO: Network [RRDBNet] is created.
2025-02-17 09:03:23,306 INFO: Network: RRDBNet, with parameters: 16,697,987
2025-02-17 09:03:23,307 INFO: RRDBNet(
  (conv_first): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (body): Sequential(
    (0): RRDB(
      (rdb1): ResidualDenseBlock(
```

Fig -5: Fine tuning hyperparameter settings

The training dataset contains 8100 images. Using a batch size of 10 on a single GPU, the training runs for 8 epochs, totaling 6480 iterations with 810 iterations per epoch.

2.4 CNN Classification

For CNN-based pneumonia classification, a pre-trained model from the publicly available GitHub repository [1] was utilized. The model was used directly to test the super-resolved images generated from our pipeline, allowing evaluation of classification performance on enhanced inputs.

The author used a chest X-ray dataset of 5216 training images (3815 pneumonia, 1341 normal) and 624 test images (390 pneumonia, 234 normal) to train a Convolutional Neural Network (CNN) model for binary classification.

The input images to the model were preprocessed using the ImageDataGenerator function with rescaling, shear transformation, zoom augmentation, and horizontal flipping. These enhancements were intended to enhance the model's robustness and generalizability during training. The training data was resized to 300×300 pixels with a batch size of 128 and binary class labels for normal (0) and pneumonia (1) conditions.

The CNN model used a sequential architecture with five convolutional layers with growing filter depths (16, 32, 64, 128, 128), each followed by max pooling layers. These convolutional layers were preceded by a flattening layer and three dense layers, and the last output layer gave a binary prediction.

The number of trainable parameters in the model totaled 1,983,009. Training was performed for 50 epochs and validation done using an independent validation generator. At the end of training, the model recorded a training accuracy of 98.16% and a flawless validation accuracy of 100%.

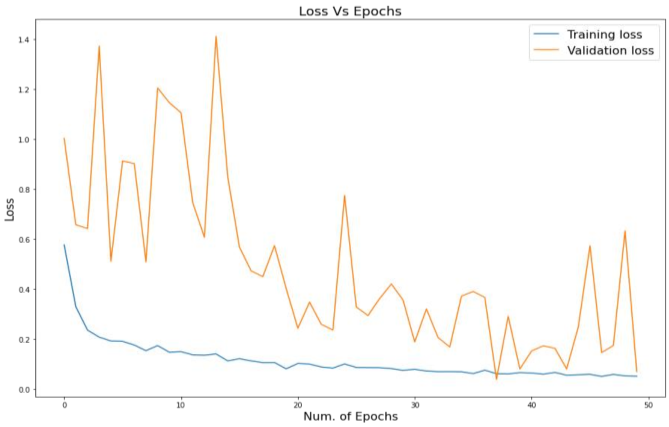


Fig -6: Training and validation loss over epochs for CNN model (Reproduced from [1])

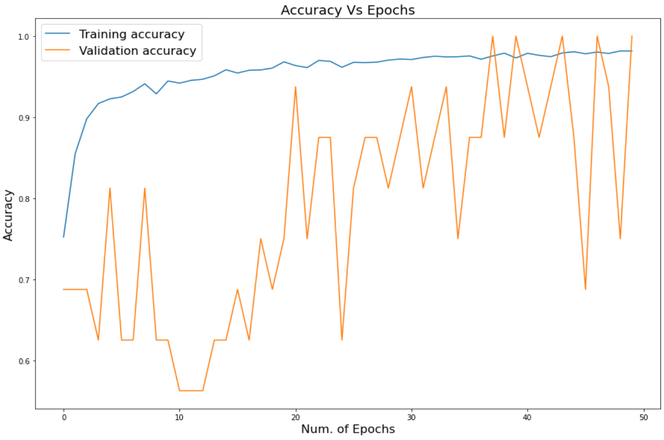


Fig -7: Training and validation accuracy over epochs for CNN model (Reproduced from [1])

3. RESULTS AND DISCUSSION

3.1 Qualitative Results of SR model

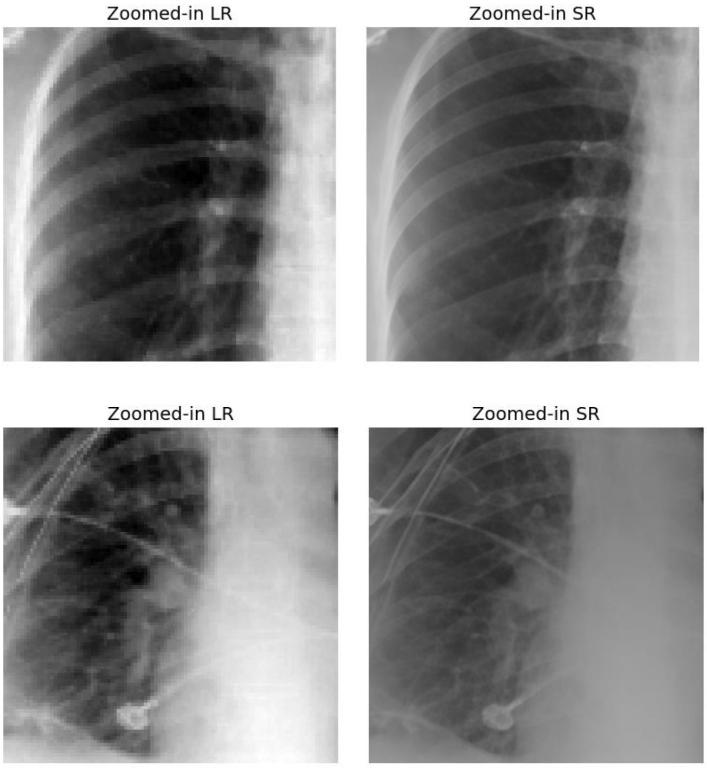


Fig -8: Zoomed in comparison of LR & SR images

Zoomed-in regions from LR (left) and corresponding SR (right) images are shown in figure 8. The SR outputs are generated using our fine-tuned super-resolution model, highlighting improved detail and clarity.

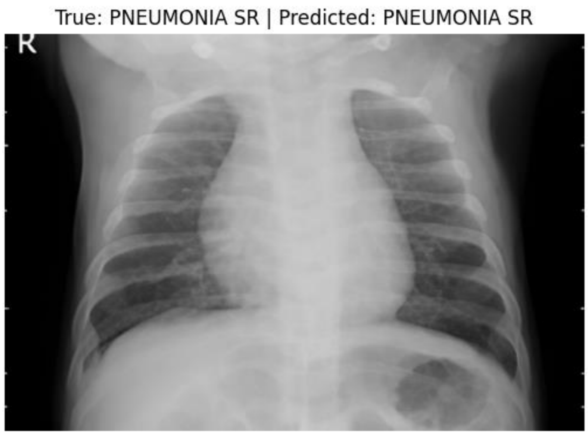


Fig -9: Test SR image prediction using pre trained CNN

3.2 Quantitative Evaluation of SR model

Evaluation of SR fine-tuned model shown in Fig 10.

Metric	Dataset	Between Original & LR	Between SR & LR	Improvement (%)
PSNR	Same dataset	34.463706	34.2113	-0.73%
SSIM		0.940185	0.9494	**0.98%**
MSE		28.383241	26.4873	**-6.68%**
Metric	**Dataset**	**Between Original & LR**	**Between SR & LR**	**Improvement (%)**
PSNR	Unseen	34.947938	35.0723	**0.36%**
SSIM		0.873238	0.9495	**8.73%**
MSE		22.056937	20.5166	**-6.98%**

Fig -10: Test SR image prediction using pre trained CNN

As per the evaluation we have tested the PSNR, SSIM, MSE on within the same type of training dataset and completely unseen dataset also. With which we can able to check the robustness of the model

3.3 Classification Performance on CNN

Original input	Loss	0.236
	Accuracy	0.921
SR input	Loss	0.279
	Accuracy	0.923

Fig -11: Test SR image prediction using pre trained CNN

The original input images achieved an accuracy of 92.1% with a loss of 0.236, while the super-resolved (SR) images achieved an accuracy of 92.3%, with a marginally higher loss of 0.279. This indicates that applying the SR model preserved same clinical details as original images with enhancing the perceptual quality of an image.

4. CONCLUSION

This research assessed the effect of using a fine-tuned Super-Resolution (SR) model on low-resolution medical images before classification. Although the classification accuracy on SR images (0.923) was almost the same as that of the original inputs (0.921).

Notably, essential clinical details were maintained in the SR outputs, which ensured diagnostic reliability. These findings indicate that SR can improve visual sharpness and perceptual quality without degrading classification accuracy

REFERENCES

[1] Y. Yashwanth, "Pneumonia detection using CNN," GitHub repository, Available: https://github.com/Yashwanth-23/Pneumonia-detection-using-CNN

[2] Ayush Patel, Prof. (Dr.) Tejas Shah. *"Chest X-ray Image Super-Resolution Using Artificial Intelligence"*, Volume 13, Issue III, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 401-411, ISSN : 2321-9653, www.ijraset.com

[3] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in Proc. 13th Eur. Conf. Comput. Vis.,Zurich, Switzerland. Springer, 2014, pp. 184–199.

[4] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jul. 2017, pp. 136–144.

[5] T. Iqbal and H. Ali, "Generative adversarial network for medical images (MI-GAN)," J. Med. Syst., vol. 42, no. 11, pp. 1–11, Nov. 2018.

[6] J. Zhu, G. Yang, and P. Lio, "How can we make GAN perform better in single medical image super-resolution? A lesion focused multi-scale approach," in Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI), Apr. 2019, pp. 1669–1673.

[7] X. Bing, W. Zhang, L. Zheng, and Y. Zhang, "Medical image super resolution using improved generative adversarial networks," IEEE Access, vol. 7, pp. 145030–145038, 2019.

[8] Xintao Wang, Liangbin Xie, Chao Dong, Ying Shan. Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. arXiv preprint arXiv:2107.10833, 2021.

[9] Random sample of NIH Chest X-ray Dataset. (2017, November23).Kaggle.https://www.kaggle.com/datasets/nih-chest-xrays/sample

[10] Chest X-Ray images (Pneumonia). (2018, March24).Kaggle.https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

[11] X. Wang, Y. Yu, K. Gu, S. Wu, J. Gu, C. Dong, Y. Qiao, and C. C. Loy, "ESRGAN: Enhanced super-resolution generative adversarial networks," in Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018. Available: https://arxiv.org/abs/1809.00219