International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Elevating Medical Image in Healthcare Through Deep Learning

Himali Patel[1], Prof. Appurva Kapil[2]

P.G. Student, Department of Computer Engineering, Silver Oak University, Ahmedabad, India
Assistant Professor, Department Of Information Technology, Silver Oak University, Ahmedabad, India

Abstract:
Medical imaging is essential to healthcare because it helps with disease diagnosis and treatment planning, and recent advances in deep learning (DL) have greatly increased the accuracy and efficiency of medical image analysis. This study examines how DL models, specifically convolutional neural networks (CNNs) and generative adversarial networks (GANs), can be used to improve the quality and classification of medical images. Using methods like data augmentation, transfer learning, and automated feature extraction, DL models achieve high accuracy in detecting diseases like cancer and heart failure. It also examines the effectiveness of various super-resolution techniques, such as SRCNN, SRGAN, and ESRGAN, for improving medical images. Experimental results show that ESRGAN performs better than other approaches, achieving the highest Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), indicating superior

Keywords: Deep Learning (DL), Medical Imaging, Convolutional Neural Networks (CNN), Generative Adversarial Networks (GANs), Data Augmentation, Super-Resolution, SRCNN, SRGAN, ESRGAN, Feature Extraction, Image Classification, Disease Diagnosis, Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM).

1. Introduction:

Medical imaging has revolutionized healthcare by enabling early disease detection, precise diagnosis, and effective treatment planning. However, analyzing medical images accurately remains a challenge due to variations in image quality, noise, and the complexity of anatomical structures. Traditional machine learning (ML) approaches have been used for image processing, but they often require extensive manual feature extraction and preprocessing.

Deep learning (DL), a subset of artificial intelligence (AI), has emerged as a powerful tool in medical image analysis by automating feature extraction and enhancing diagnostic accuracy. DL models, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable success in medical imaging tasks, such as detecting abnormalities in X-rays, MRIs, and CT scans. These models leverage large datasets and advanced architectures to learn complex patterns, significantly outperforming traditional methods.

In this study, we explore the impact of deep learning techniques in improving medical imaging. The research focuses on the application of CNNs, Generative Adversarial Networks (GANs), and Super-Resolution techniques for enhancing image quality and classification accuracy. Specifically, methods such as data augmentation, transfer learning, and feature extraction are employed to improve diagnostic

performance. Additionally, we evaluate the effectiveness of different super-resolution models—SRCNN, SRGAN, and ESRGAN—in generating high-quality medical images.
The results of this study indicate that deep learning models, particularly ESRGAN, achieve superior image resolution and classification accuracy. By leveraging DL-based techniques, healthcare professionals can benefit from more reliable and efficient medical image analysis, ultimately improving patient outcomes.

2. DEEP LEARNING:

DL models enable machines to achieve the accuracy by advancements in techniques to analyze medical images.The heart disease was diagnosed using the labelled chest X-Rays,cardiologist reviewed and relabelled all the data while discarding the data other than heart failure and normal images.

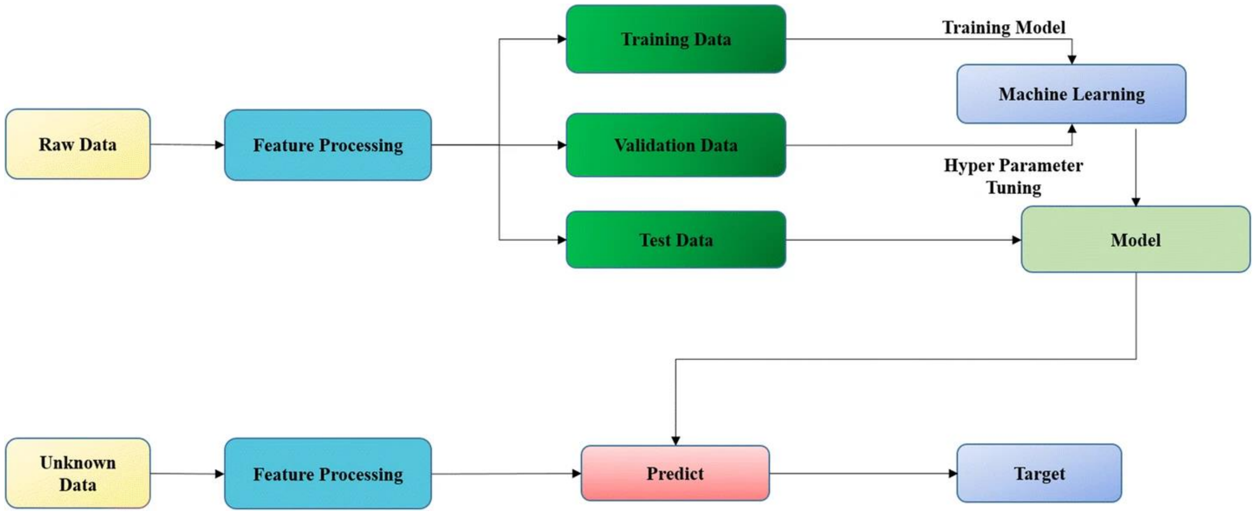


Fig. 1 ML process

To extract the exact features from the images, data augmentation and TL were used with 82% accuracy, 74% specificity and 95% sensitivity for heart failure. In an automatic feature selection, using histopathology images with the labelling of positive and negative cancer images, was developed with minimum manual work. Two networks named Deep Neural Network (DNN) 2-F and DNN1-F were used with PCA to reduce features in DNN whereas for unsupervised feature learning a single-layer network of K-means centroids was used. Later, the results of unsupervised (93.56%) and supervised (94.52%) learning were compared. The DL model automates the feature extraction procedure to handle data efficiently. Figure 2 depicts the process used by DL algorithms for the prediction and diagnosis of various diseases
To process the medical images for better prediction and accuracy, ML and DL techniques were used as shown in Figs. 1 and 2, respectively. As input, medical images from various modalities are taken into consideration, and then algorithms are applied to these images. Further, the input image is segmented based on various factors, these segments were used to extract the essential and maximum features using feature extraction techniques. After the extraction of the required features, they are further refined to obtain actual features used for the identification of diseases. Also, ML approaches were used to denoise the medical images for better prediction and accuracy in. Once the feature selection and noise removal from the data are achieved, the classification of the images according to the disease using classifiers like SVM, Decision Tree (DT), etc. was attained.

ML is the process where computers learn from data and use algorithms to carry out a task without being explicitly programmed. It uses pattern recognition to make predictions with new dataset. Alternatively, DL is modeled according to the human brain including a complex structure of algorithms enabling machines to process images, text and documents. It uses layered-structure algorithms such as Convolutional Neural Network (CNN), Artificial NeuralNetwork (ANN), etc., to analyze the data with logics. Comparatively, DL is more capable of processing huge amounts of data than ML models.

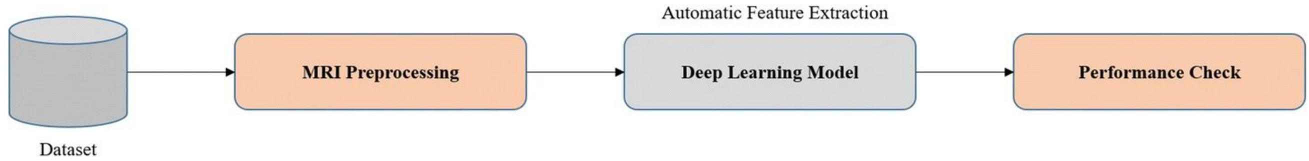


Fig. 2 DL process

3. Structure of CNN model:

In this study, we utilized the DenseNet-201 architecture, a convolutional neural network renowned for its image recognition capabilities. DenseNet-201 is characterized by its dense connectivity, where each layer receives input from all preceding layers in a block, enhancing feature propagation and reuse. This model consists of multiple dense blocks, transition layers, a global average pooling layer, and a fully connected layer for classification. Each dense block contains densely connected convolutional layers, facilitating the capture of both low-level and high-level features. Transition layers, equipped with batch normalization, pooling operations, and bottleneck convolutional structures, manage the dimensions and channels of feature maps. The global average pooling layer compacts these maps into a fixed-size feature vector, regardless of input size. Finally, the fully connected layer performs classification based on the number of target classes.

Additionally, DenseNet-201's design mitigates the vanishing gradient problem common in deep networks and maintains high accuracy with efficient parameter utilization. These attributes make DenseNet-201 particularly well-suited for complex image recognition tasks like detecting hidden pediatric elbow fractures in X-ray images.

Here's a breakdown of its key architectural components:

Dense Blocks: DenseNet-201 comprises multiple dense blocks, each containing a series of densely connected convolutional layers. In these blocks, each layer receives feature maps not just from the previous layer but also from all preceding layers within the same block. This dense connectivity promotes feature reuse, enabling the network to capture both low-level and high-level features effectively.

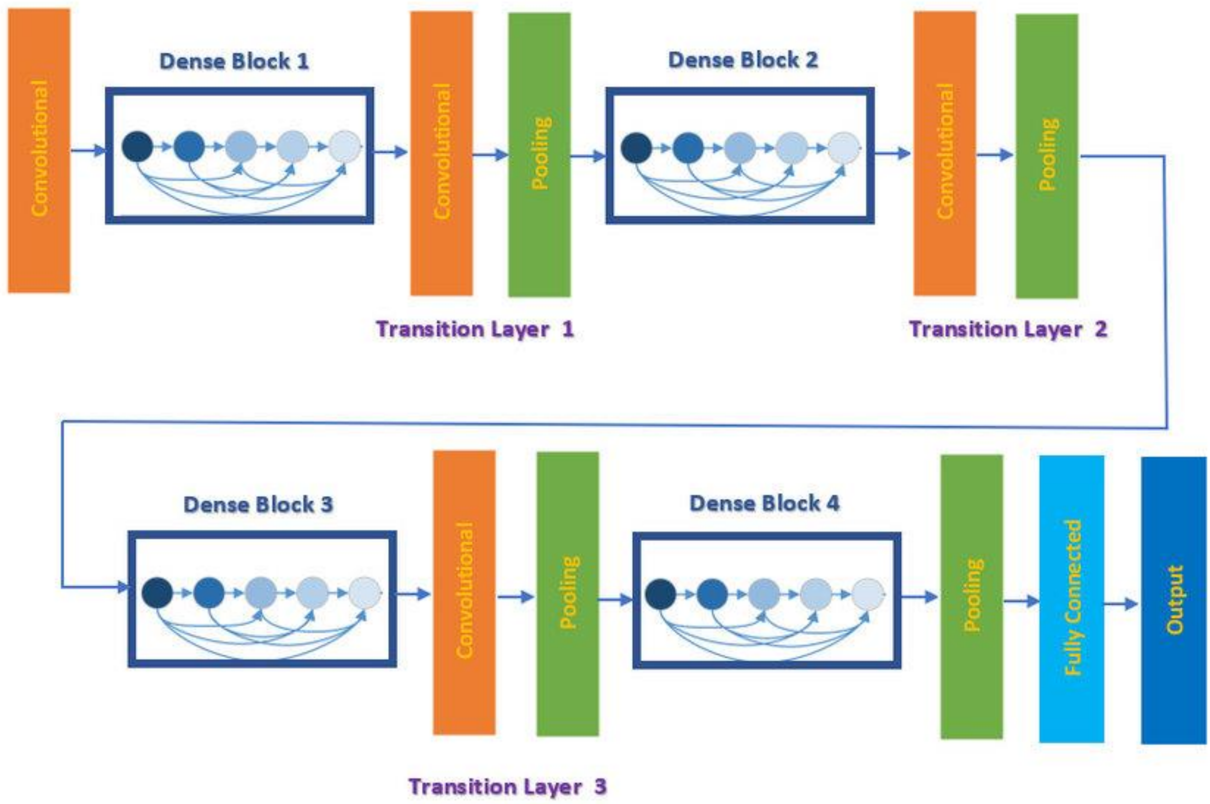


Fig-3 The network structure diagram of DenseNet-201.

Transition Layers: Between dense blocks, transition layers are inserted. These layers include batch normalization, a pooling operation (typically average pooling), and a convolutional layer with a bottleneck structure (1 × 1 convolution). Transition layers reduce the spatial dimensions of feature maps while increasing the number of channels,striking a balance between computational efficiency and expressive power.

Global Average Pooling (GAP): At the end of the network, a global average pooling layer is used to aggregate the feature maps spatially,resulting in a single vector for each feature map. This reduces the spatial dimension to 1 × 1, enabling the network to produce a fixed-size feature vector regardless of input size.

Fully Connected Layer: Following GAP, a fully connected layer performs the final classification. The number of neurons in this layer corresponds to the number of classes in the classification task.

Feature Reuse: DenseNet's dense connectivity allows for maximum feature reuse, which facilitates the learning of more compact and discriminative representations from the data

Mitigating Vanishing Gradient: The dense connections ensure the flow of gradients during training, mitigating the vanishing gradient problem often encountered in very deep networks

Efficient Parameter Utilization: DenseNet's parameter-efficient design enables it to maintain high accuracy while using fewer parameters compared to traditional architectures.

State-of-the-Art Performance: DenseNet-201 consistently achieves state-of-the-art performance in various image recognition challenges, outperforming many other architectures in terms of both accuracy and computational efficiency

In summary, DenseNet-201's unique architecture, characterized by dense connectivity, makes it a powerful tool for image recognition tasks.It efficiently reuses features, addresses gradient vanishing, and maintains competitive performance with fewer parameters compared to other architectures. The network structure diagram of DenseNet-201and detailed parameters can be seen in Fig. 3.

4. GANs- Generative Adversarial Networks:

Generative Adversarial Networks (GANs) are a powerful class of neural networks that are used for Generative Deep Learning Gan's were introduced by Ian Goodfellow et al. in 2014

Generative Adversarial Networks (GANs) can be broken down into three parts:

a) **Generative:** To learn a generative model, which describes how data is generated in terms of a probabilistic model

b) **Adversarial:** The training of a model is done in an adversarial setting.

c) **Networks:** Use deep neural networks as the artificial intelligence (AI) algorithms for training purpose

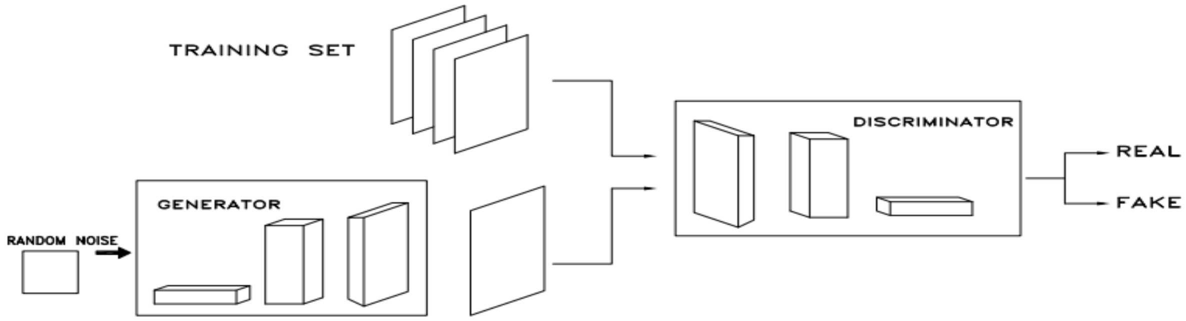


Fig 4:- Architecture Of GANs

Lower Resolution

Higher Resolution Image Generated By GANs

Fig 5:-Lower Resolution Vs Higher Resolution Using GANs

5. SR-CNN:

In Super Resolution Convolutional Neural Network, actually each association isn't profound. There are just 3 sections, patch extraction and representation, non-linear mapping, and reconstruction. The fundamental layer takes in feature maps from the lower resolution pictures. Resulting layer is considered as learning feature maps for the LR Super Resolution picture, including maps. Last layer uses these SR feature advisers for build up the genuine SR picture. The ReLU non-linearity is applied to the two beginning layers. The chief layer uses 64 channels of size 9x9, the resulting layer uses 32 channels of size 1x1, and the last layer uses a channel of size 5x5.

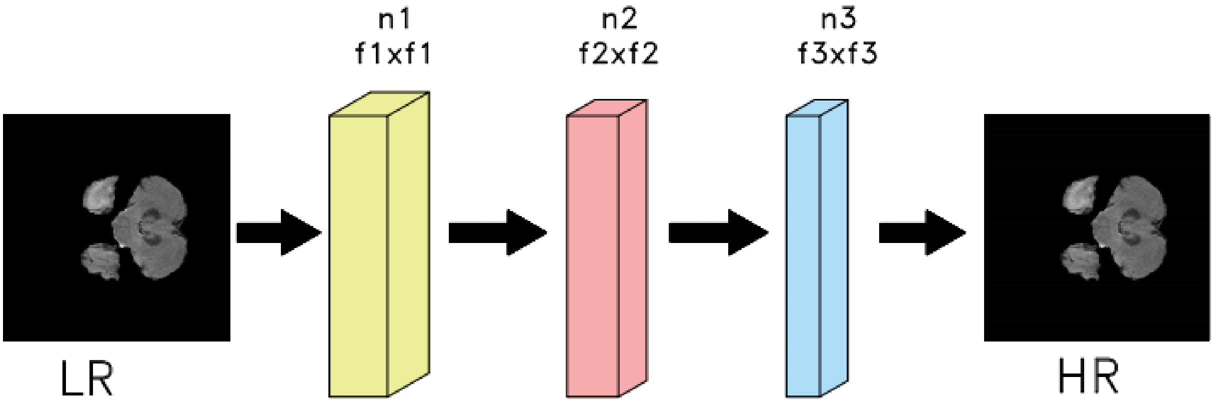


Figure 6:- Architecture of SR-CNN

Patch Extraction and Representation1: Realize that the LR input is first upscale to ideal size utilizing bicubic interpolation prior to contributing to the SRCNN network.
Non-Linear Mapping: It is used for mapping LR to HR vector.
Reconstruction: After mapping, we need to remake the picture. Henceforth, we do convolution once more.

6. SR-GAN:

SRGAN was put forward specialists at Twitter. The intention of this design is to recuperate better surfaces from the picture when we upscale it with the goal that its quality can't be undermined. There are different strategies, for example, Bilinear Interpolation that can be utilized to play out this assignment yet they experience the ill effects of picture data misfortune and smoothing. In this paper the creators proposed two designs the one without GAN (SRResNet) and one with GAN (SRGAN). It is inferred that SRGAN has better precision and produce picture more satisfying to eyes when contrasted with SRGAN.
SRGAN is the primary system equipped for surmising photograph reasonable normal pictures for 4× up-scaling factors. The SRGAN is the principal structure equipped for inducing photo-realistic common pictures for 4× up-scaling factors. Plan in Fig. 7 contains two segments: the generator and the discriminator associations.
The generative model is a significant extra association that recognizes a lower resolution picture and yields a super resolution pictures. The advantage of having a waiting organization with skip affiliations is that the generator swears off vanishing and exploding slants, which could arise as a result of the significance of the association. The generator is set up to fool the discriminator into tolerating that the yield super

resolution pictures are high resolution. The discriminator is set up to perceive super resolution pictures from the primary pictures.

The GAN approach uses an adversity work that is contained a perceptual incident, which urges super resolution amusements to move towards regions of the request space with a high probability of containing photo reasonable pictures and a substance hardship considering perceptual closeness using the verifiable level feature from a pre-arranged VGG19 network. The primary association works with Red Gray Blue pictures, while MR pictures are grayscale. We change in association, work with a one channel contribution, regardless, the VGG19 pre-arranged the association requires three channels and cannot be changed.

In oblige of objective, we reproduce the grayscale pictures to lay out three diverts just in the development that discovers the perceptual calamity. The generator has the irritating undertaking of making another picture, while the discriminator as it expected to deal with an immediate strategy issue.

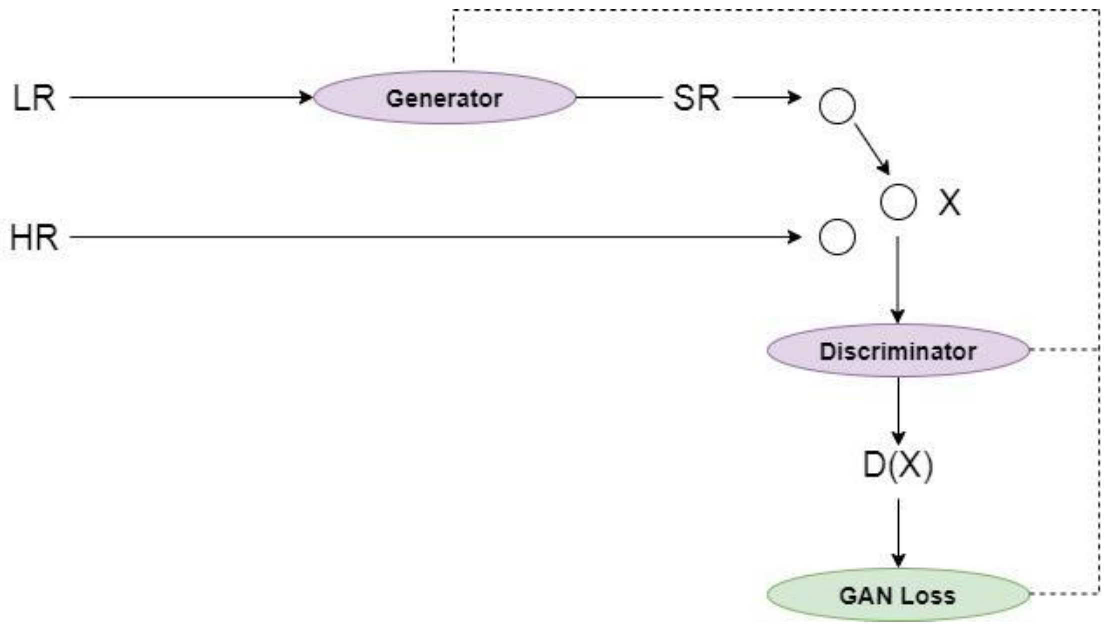


Figure:-7 SR-GAN Architecture

7. E-SR-GAN:

The SRGAN a fundamental work which fits for creating reasonable surfaces during a single picture super-goal. Be that as it may, the fantasized subtleties are frequently joined by disagreeable relics. To additional improve the visual quality, we altogether study three parts of Super Resolution Generative Adversarial Network – network plan, hostile adversity, and perceptual setback, and improve all of them to induce an ESRGAN. Specifically, we present the RRDB without bunch standardization from essential organization segment. In addition, we get the thought from RaGAN that allows discriminator to foresee relative realness rather than supreme worth.

The network structure of the Generator is improved by presenting the RRDB, which builds the limit of the organization and makes the preparation simpler as well.

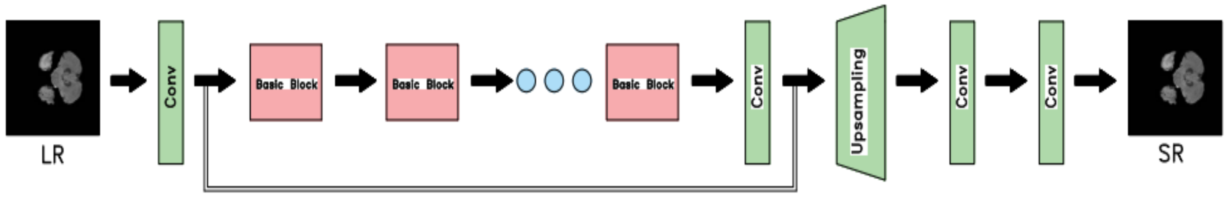


Figure :-8 Architecture of the ESRGAN-Generator (RRDB)

8. Result Analysis Parameters:

PSNR:-

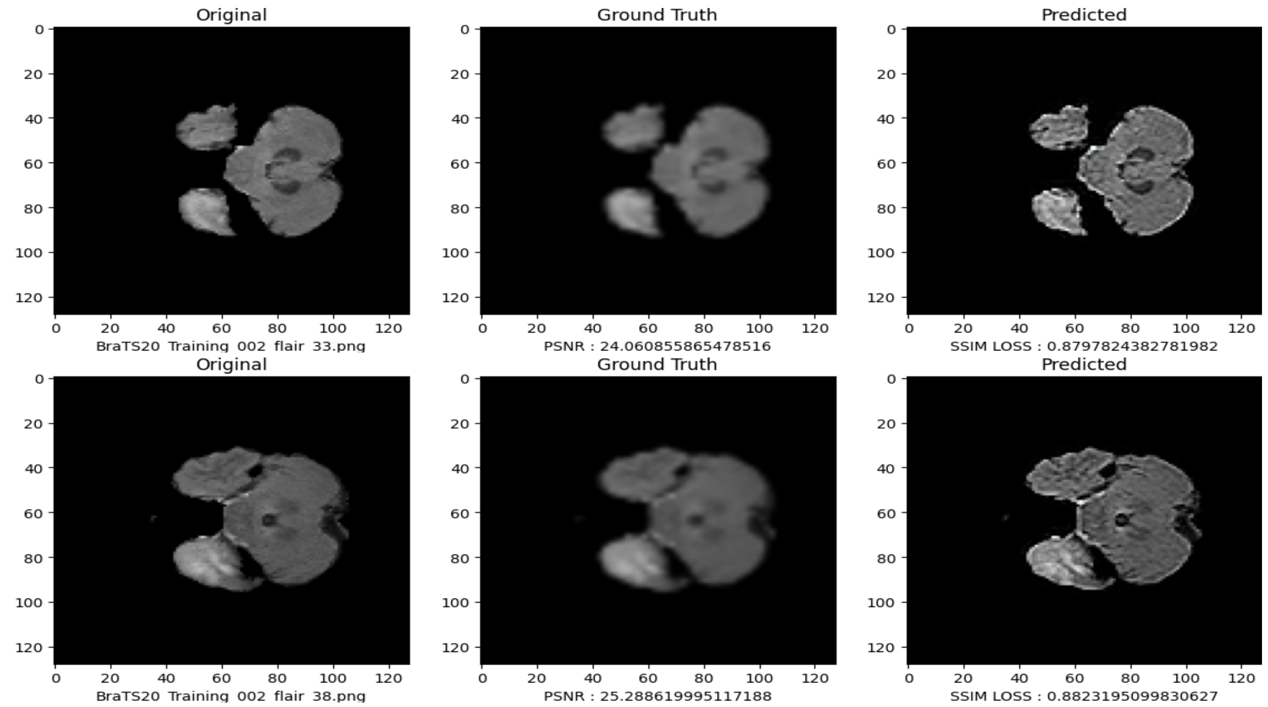
PSNR stands for Peak Signal Noise Ratio. It is most commonly used to measure the quality of reconstruction of an image. The more the PSNR, the better the image. PSNR is calculated in dB (decibel). The estimated range of PSNR lies between 30 – 50dB.

SSIM:-

SSIM stands for Structural Similarity Index and is a metric to measure visible structures in the images. Sometimes the SSIM is also calculated on how the image is visible to human eye. The SSIM values like 0.92, 0.97 – 0.99 are said to be good for the reconstructed images.

9. RESULT ANALYSIS:

SRCNN



Performance Evaluation-SRCNN:

SAMPLE	PSNR(dB)	SSIM
S1	25.28	0.88
S2	24.36	0.83

S3	**25.24**	**0.82**
S4	22.04	0.7
S5	21.70	0.76
S6	20.96	0.77

Our model has achieved the maximum PSNR of 25.28 dB, where as the minimum PSNR score was 20.96 dB also the maximum SSIM score was 0.88 and the minimum we got was 0.7.

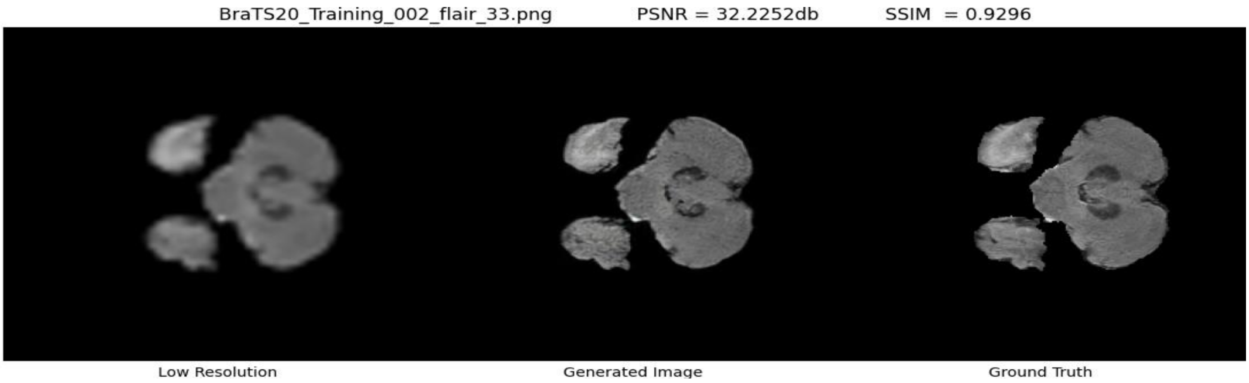
SRGANS

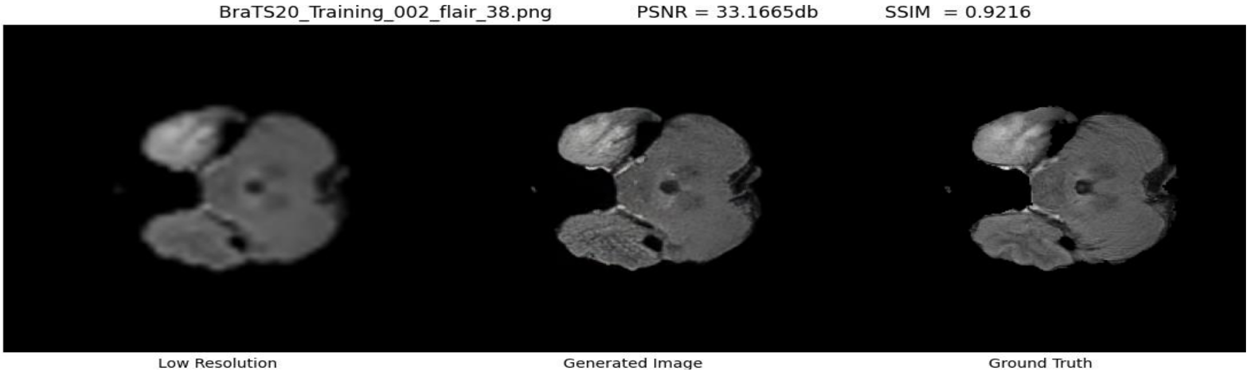
Performance Evaluation-SRGAN:

SAMPLE	PSNR(dB)	SSIM
S1	**23.43**	0.83
S2	18.30	0.75
S3	17.16	0.74
S4	14.50	0.63
S5	21.27	0.79
S6	12.57	0.62

Our model has achieved the maximum PSNR of 23.43 dB, where as the minimum PSNR score was 12.57 dB also the maximum SSIM score was 0.83 and the minimum we got was 0.62.

ESRGAN





Performance Evaluation-ESRGAN:

SAMPLE	PSNR(dB)	SSIM
S1	**33.16**	0.92
S2	30.56	0.85
S3	32.97	0.87
S4	28.81	0.82
S5	28.33	0.81
S6	28.36	0.82

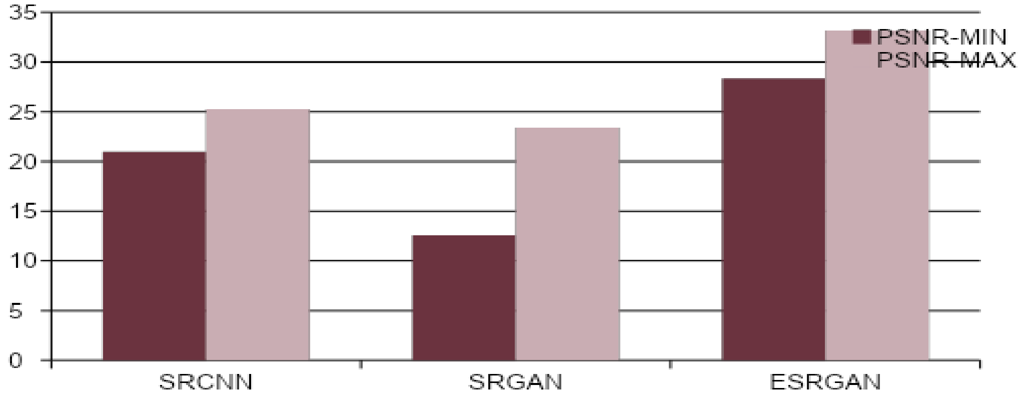
Our model has achieved the maximum PSNR of 33.16 dB, where as the minimum PSNR score was 28.33 dB also the maximum SSIM score was 0.92 and the minimum we got was 0.81.

Comparisons of Super-Resolution Variants - PSNR

Performance Evaluation - PSNR

FRAMEWORK	PERFORMANCE EVALUTION – PSNR (dB)
SR-CNN	20.96
	25.28
SRGAN	12.57

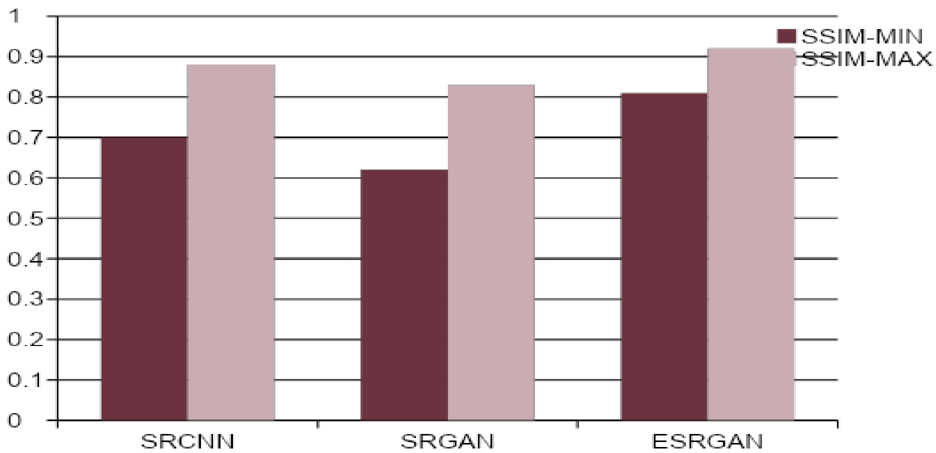
	23.43
ESRGAN	28.33
	33.16



Graphical representation – PSNR (min, max)

Comparisons of Super-Resolution Variants - SSIM
Performance Evaluation - SSIM

FRAMEWORK	PERFORMANCE EVALUTION - SSIM
SR-CNN	0.88
	0.7
SRGAN	0.83
	0.62
ESRGAN	0.92
	0.81



Graphical representation – SSIM (min, max)

10. Conclusion:

Deep learning has significantly transformed medical imaging by enhancing image quality, improving disease diagnosis, and reducing the need for manual intervention. This study explored various deep learning models, including Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), for medical image analysis and classification. Additionally, super-resolution techniques such as SRCNN, SRGAN, and ESRGAN were examined for their effectiveness in improving image clarity.

Experimental results demonstrated that ESRGAN outperformed other super-resolution methods, achieving the highest Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), which are critical metrics for image quality assessment. By leveraging deep learning techniques such as data augmentation, transfer learning, and automated feature extraction, medical image processing can be optimized for better accuracy and efficiency.

The findings of this study highlight the potential of deep learning in healthcare, enabling precise and automated medical diagnoses. Future research can focus on integrating deep learning models with real-time medical imaging systems, improving computational efficiency, and expanding datasets to further enhance model performance. The continued advancement of deep learning in medical imaging holds great promise for improving patient care and clinical decision-making.

References:

1. https://www.aidoc.com/blog/deep-learning-in-healthcare/
2. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5479722/
3. https://www.annualreviews.org/doi/abs/10.1146/annurev-bioeng-071516-044442
4. Chaudhari, A. S., Fang, Z., Kogan, F., Wood, J., Stevens, K. J., Gibbons, E. K., ... & Hargreaves, B. A. (2018). Super-resolution musculoskeletal MRI using deep learning. Magnetic resonance in medicine, 80(5), 2139-2154.
5. Gu, Y., Zeng, Z., Chen, H., Wei, J., Zhang, Y., Chen, B., ... & Lu, Y. (2020). MedSRGAN: medical images super-resolution using generative adversarial networks. MULTIMEDIA TOOLS AND APPLICATIONS.
6. Hamghalam, M., Wang, T., Qin, J., & Lei, B. (2020, April). Transforming Intensity Distribution of Brain Lesions Via Conditional Gans for Segmentation. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 1-4). IEEE.
7. Lyu, Q., Shan, H., Steber, C., Helis, C., Whitlow, C. T., Chan, M., & Wang, G. (2020). Multi-contrast super-resolution mri through a progressive network. IEEE Transactions on Medical Imaging.
8. Song, T. A., Chowdhury, S. R., Yang, F., & Dutta, J. (2020). PET image super-resolution using generative adversarial networks. Neural Networks, 125, 83-91.
9. Qiu, D., Zhang, S., Liu, Y., Zhu, J., & Zheng, L. (2020). Super-resolution reconstruction of knee magnetic resonance imaging based on deep learning. Computer methods and programs in biomedicine, 187, 105059.
10. Kora Venu, S. (2020). Evaluation of Deep Convolutional Generative Adversarial Networks for data augmentation of chest X-ray images. arXiv e-prints, arXiv-2009.
11. Cem Birbiri, U., Hamidinekoo, A., Grall, A., Malcolm, P., & Zwiggelaar, R. (2020). Investigating the Performance of Generative Adversarial Networks for Prostate Tissue Detection and Segmentation. Journal of Imaging, 6(9), 83.
12. Hamghalam, M., Wang, T., Qin, J., & Lei, B. (2020, April). Transforming Intensity Distribution of

Brain Lesions Via Conditional Gans for Segmentation. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 1-4). IEEE.

13. Gupta, R., Sharma, A., & Kumar, A. (2020). Super-Resolution using GANs for Medical Imaging. Procedia Computer Science, 173, 28-35.
14. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial networks. arXiv preprint arXiv:1406.2661.
15. Yi, X., Walia, E., Babyn, P., 2019. Generative adversarial network in medical imaging: A review. Medical image analysis 58, 101552.
16. Glasner, D., Bagon, S., & Irani, M. (2009, September). Super-resolution from a single image. In 2009 IEEE 12th international conference on computer vision (pp. 349-356). IEEE.
17. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4681-4690).
18. https://cognitiveresearchjournal.springeropen.com/articles/10.1186/s41235-019-0171-6/figures/1
19. Kazeminia, S., Baur, C., Kuijper, A., van Ginneken, B., Navab, N., Albarqouni, S., & Mukhopadhyay, A. (2020). GANs for medical image analysis. Artificial Intelligence in Medicine, 101938.
20. Hitawala, S. (2018). Comparative study on generative adversarial networks. arXiv preprint arXiv:1801.04271.
21. https://towardsdatascience.com/five-gans-for-better-image-processing-fabab88b370b
22. Sood, R., Topiwala, B., Choutagunta, K., Sood, R., & Rusu, M. (2018, December). An application of generative adversarial networks for super resolution medical imaging. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 326-331). IEEE.
23. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," CoRR, vol. abs/1609.04802, 2016
24. Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M.: Enhanced deep residual networks for single image super-resolution. In: CVPRW. (2017)
25. https://medium.com/syncedreview/microsoft-obj-gan-turns-words-into-complex-scenes5c6024f0f91d
26. Li, Y., Gan, Z., Shen, Y., Liu, J., Cheng, Y., Wu, Y., ... & Gao, J. (2019). Storygan: A sequential conditional gan for story visualization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6329-6338).
27. Yu, X. (2020). Emerging Applications of Generative Adversarial Networks. In IOP Conference Series: Materials Science and Engineering (Vol. 740, No. 1, p. 012132). IOP Publishing.
28. Stanton, G., & Irissappane, A. A. (2019). Gans for semi-supervised opinion spam detection. arXiv preprint arXiv:1903.08289.
29. Luc, P., Clark, A., Dieleman, S., Casas, D. D. L., Doron, Y., Cassirer, A., & Simonyan, K. (2020). Transformation-based adversarial video prediction on large-scale data. arXiv preprint arXiv:2003.04035.
30. Jiang, Y., Chang, S., & Wang, Z. (2021). Transgan: Two transformers can make one strong gan. arXiv preprint arXiv:2102.07074.
31. Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence, 38(2), 295-307.

32. Zhao, H., Gallo, O., Frosio, I., & Kautz, J. (2016). Loss functions for image restoration with neural networks. IEEE Transactions on computational imaging, 3(1), 47-57.
33. https://www.geeksforgeeks.org/super-resolution-gan-srgan/
34. Isaac, J. S., & Kulkarni, R. (2015, February). Super resolution techniques for medical image processing. In 2015 International Conference on Technologies for Sustainable Development (ICTSD) (pp. 1-6). IEEE.
35. Moran, M. B., Faria, M. D., Giraldi, G. A., Bastos, L. F., & Conci, A. (2021). Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs. Computers in biology and medicine, 129, 104139.
36. Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. (2019). Deep learning for single image super-resolution: A brief review. IEEE Transactions on Multimedia, 21(12), 3106-3121.
37. https://towardsai.net/p/machine-learning/reading-esrgan%E2%80%8A-%E2%80%8Aenhanced-super-resolution-generative-adversarial-networks-super-resolution
38. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., ... & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops (pp. 0-0).