Received Date:28th December, 2024
Revision Date: 16th January, 2025
Accepted Date: 5th March, 2025

Lumber Spine MRI Super-Resolution Using SRGAN's

Kritika Acharya[1*], Rashna K.C.[2], Suyog Acharya[3,] Yogesh Aryal[4], Suramya Sharma Dahal[5]
[1]Thapathali Campus, Institute of Engineering, E-mail: acharyakritika44@gmail.com
[2] Thapathali Campus, Institute of Engineering, E-mail: rashna.kc10@gmail.com
[3] Thapathali Campus, Institute of Engineering, E-mail: yogesharyal2525@gmail.com
[4] Thapathali Campus, Institute of Engineering, E-mail: asuyog042@gmail.com
[5]Associate Professor, Dept of Electronics, Communication & Information Engineering, Kathmandu Engineering College, E-mail: suramya.sharma@kecktm.edu.np

Abstract— Medical imaging is essential to modern healthcare because it provides accurate visual information for diagnosis and treatment planning. However, low-resolution images are often created due to inherent limitations of imaging technology, which impairs the ability of healthcare professionals to see small details. To overcome this difficulty, super-resolution - a method of improving image resolution - has become popular. This project investigates the use of Super-Resolution Generative Adversarial Networks (SRGAN) for medical image enhancement. In practice, adversarial training SRGANs, a method based on deep learning, produces high-resolution images of their low-resolution counterparts. In several fields, including medical imaging, the model's ability to learn complex mappings between low and high-resolution image spaces has produced impressive results. In this work, we specifically discuss how SRGANs can improve the resolution of medical images to help doctors diagnose patients with greater accuracy and detail. In this paper, we have discussed about our system that increases the resolution of low-resolution MRI imaging of Lumber Spine with high accuracy.

Keywords— AI, Deep Learning, GAN, SRGAN, Super-Resolution

Introduction

Medical imaging is essential to modern healthcare because it provides accurate visual information for diagnosis and treatment planning. However, low-resolution images are often created due to inherent limitations of imaging technology, which impairs the ability of healthcare professionals to see small details. To overcome this difficulty, super-resolution - a method of improving image resolution - has become popular. This project investigates Super-Resolution Generative Adversarial Networks (SRGAN) for medical images enhancement. In practice, adversarial training SRGANs, a method based on deep-learning, produces high resolution images of their low-resolution counterparts. In several fields, including medical imaging, the model's ability to learn complex relationship between low- and high-resolution images has produced impressive results. To improve the diagnostic precision and clinical usefulness of imaging data in healthcare settings, there is an urgent need to use Generative Adversarial Networks for super resolution of clinical report pictures. The capacity to identify minute characteristics in pictures is essential for precise interpretation and decision-making in medical diagnostics. However, medical pictures are frequently obtained at lesser resolutions because of things like hardware limits, financial concerns, and the shortcomings of outdated imaging equipment. This is a problem since reduced resolution might make it harder to see minute anomalies or anatomical features, which could affect the accuracy of diagnoses and the treatment that follows for the patient.

The medical images super resolution algorithm addresses the challenge of limited spatial resolution of medicinal images. Capturing high resolution image is time consuming sometimes not suitable for medical emergencies. Clinical obtained low-resolution images lack some detail that includes high frequency component, different patterns and tissues textures. Hence, Medical image super resolution with Generative Adversarial Networks not only preserves more texture detail but will also be able to generate more real patterns on generated super resolution images in limited time. In this work, we introduce a new system that uses SRGANs for enhancing the resolution of lumbar spine MRI images. Our system is designed in such a way that it will generate super-resolved high-quality images while preserving critical spinal structures, hence improving the visual clarity and possible diagnostic value of the obtained images. The proposed method attempts to fill the gap between low-resolution MRI scans and the standards for high-resolution imaging needed for such precise clinical analysis.

** Corresponding Author*

Materials and methods in sufficient detail

A. Dataset Overview

Our research relies heavily on the Lower Spine MRI Super-Resolution dataset, which was painstakingly assembled from the reliable academic portal Mendeley [9]. Because it draws from a wide range of academic sources, the dataset's quality and dependability are ensured by using the scholarly contributions that are accessible on Mendeley. There was a total of 48,345 MRI slice in the datasets out of which we have used 22532 images for the training purpose due to our limited computational resources. Manually removed images that were not clear in the form of high-resolution. We converted all the HR images into LR my using gaussian blur, and down-sampled the LR images by 4 times. We prepared HR-LR pair for all images before feeding them to the model. Out of the total dataset 10% dataset has been used for validation and 10% dataset has been used for testing. That is, 2253 images have been used for validation as well as testing.

Data Preprocessing

- Removal of duplicates: Finding and removing any duplicate rows from the dataset is essential throughout the data cleaning process to prevent biases or errors in the study. We eliminated duplicate rows from our sample to ensure that each user's information was utilized just once throughout the research. There are several reasons why there might be duplicate rows, including incorrect data entry and technological problems. The reliability and accuracy of our study may suffer if these duplicate rows are not found and eliminated as it might produce an overrepresentation of certain individuals or data items. We were able to make sure that every user's data was fairly represented in the study and that the absence of duplicate data did not skew or bias our findings by eliminating duplicate rows.
- Data Transformation: Finding and removing any duplicate rows from the dataset is essential throughout the data cleaning process to prevent biases or errors in the study. We eliminated duplicate rows from our sample to ensure that each user's information was utilized just once throughout the research. There are several reasons why there might be duplicate rows, including incorrect data entry and technological problems. The reliability and accuracy of our study may suffer if these duplicate rows are not found and eliminated as it might produce an overrepresentation of certain individuals or data items. We were able to make sure that every user's data was fairly represented in the study and that the absence of duplicate data did not skew or bias our findings by eliminating duplicate rows.

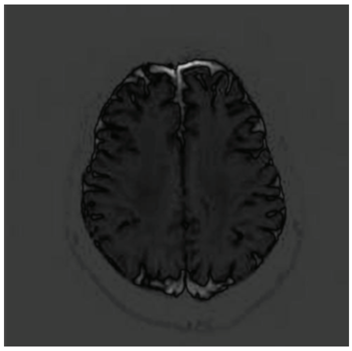

FIG. 1 Overly substantial black region dataset

- Relevant Information Retention: In the preprocessing phase, emphasis is placed on retaining critical information including spatial dimensions, anatomical structures, and intensity distributions within the lower spine region. Extraneous elements that do not contribute to the super-resolution task are removed, enhancing the dataset's focus and relevance.
- Conversion to .jpg format: The initial preprocessing step involved converting the MRI images from the. ima format to the more widely supported .jpg format. This conversion likely involved utilizing appropriate software or libraries capable of reading .ima files and saving them as .jpg files. The conversion ensures compatibility with common image processing tools and facilitates subsequent analysis.
- Loading and Resizing Images: After conversion, the MRI images were loaded into memory using libraries like PIL (Python Imaging Library). Depending on the original dimensions of the images, resizing may have been necessary to standardize the image sizes for consistency across the dataset. Resizing ensures that images are uniform in dimensions, which is crucial for subsequent analysis and model training.
- Enhancement Techniques: Preprocessing have included various enhancement techniques to improvize the qualities and clarity of image. This could involve applying filters (e.g., Gaussian blur) to reduce noise, enhance edges, or improve overall image quality. Such techniques help to enhance the interpretability of the images and aid in the identification of relevant features during analysis

Dataset Out contains single images taken from either sagittal or axial views of the three lowest vertebrae and the three lowest IVDs. The last three IVDs, including the one between the sacrum and the last vertebrae, are mostly used to obtain axial visual slices. The slices of the latter IVD are oriented to follow the curve of the spine, while the slices of earlier IVD are usually built in blocks or parallel to each other. Each IVD has four or five slices arranged from top to bottom. Many upper and lower slices penetrate the spine, leaving one to three slices that show only the image of the IVD and cleanly dissect the IVD.

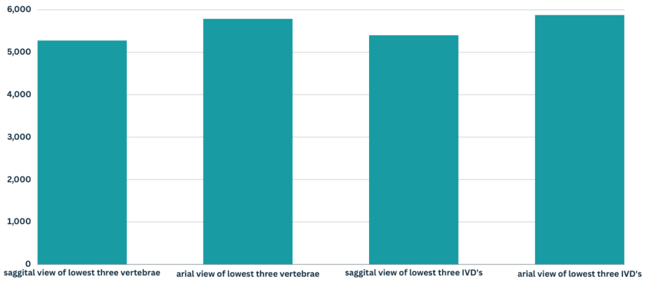

FIG. 2 Data distribution in dataset

B. Generator

Based on the RCAN design, the Residual Whole Map Attention Network (RWMAN) is modified generator for super-resolution images. To improve the attention mechanism, Residual Whole Map Attention Block (RWMAB) is used in place of the Residual Channel Attention Block (RCAB). RWMAB generates adaptive pixel weights across all channels by using a 1×1 conv layer and sigmoid activation, which enables model to concentrate on significant regions. The input conv, group of RWMAB with long, short residuals connection, an up sampled modules make up the generator. This change was motivated by the necessity to focus attention in medical imaging by selecting just regions that contain relevant information. The High-Resolution images generated by this neural network with parallel Low-Resolution image is name super resolution images in the given description:

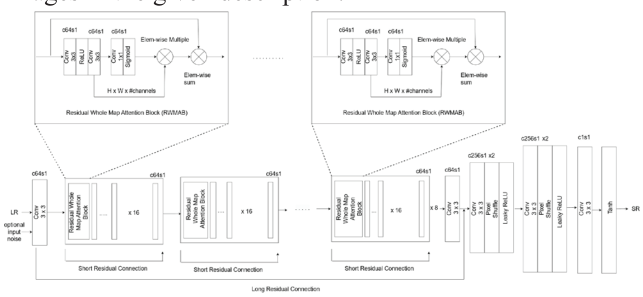

FIG. 3 Generator architecture [1]

Discriminator

In traditional GAN architectures, discriminator took single high-resolution images super-resolved images is an input. Then, when tried to differentiate if the images are true or generated SR image. In this SRGAN's architecture the discriminator uses pair of images consisting of low resolution (LR) images and its parallel HR or SR counterpart as an input. Discriminator architecture has been built to increase the use of paired data. Separate paths are applied to process the Low-Resolution images and matching High Resolution/SR images. Concatenated feature maps are obtained from the LR and HR routes. This indicates that the discriminator is picking up knowledge from both the LR and HR/SR images. The discriminator's goal is to catchup the pairwise information among Low Resolution and High Resolution/SR pictures with employing image pairings. The discriminator outputs the probable chances that provided image pair is a true (LR, HR) / (LR, SR). Parallel LR and HR images labelled as 1 (actual pair) with pairs of LR and SR images labelled as 0 (false pair) are used to train the discriminator while the training process. The goal to train the discriminator was to accurately identify whether or not the input LR image and its matching image (HR or SR) constitute a legitimate pair.

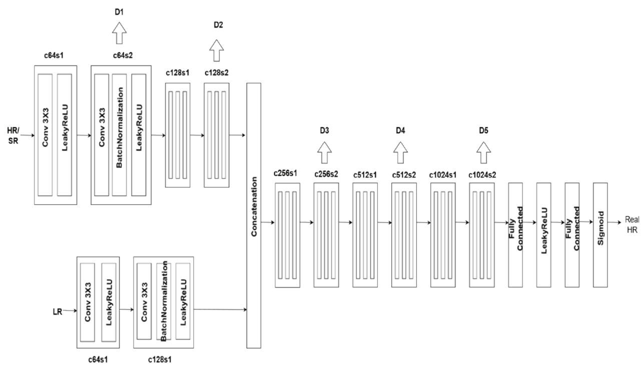

FIG. 4 Discriminator architecture [1]

C. Pretrained VGG-19

VGG-19 model is implemented as an extractor of feature to take high level quality from image. These features can be used to map perceptual resemblance among the supe resolved and truth images while guiding the optimization during training process. A specific layer of VGG-19 model has been employed to capture feature. The high resolution and super-resolved (SR) image are passed through layers which extracts feature maps for both HR and SR images.

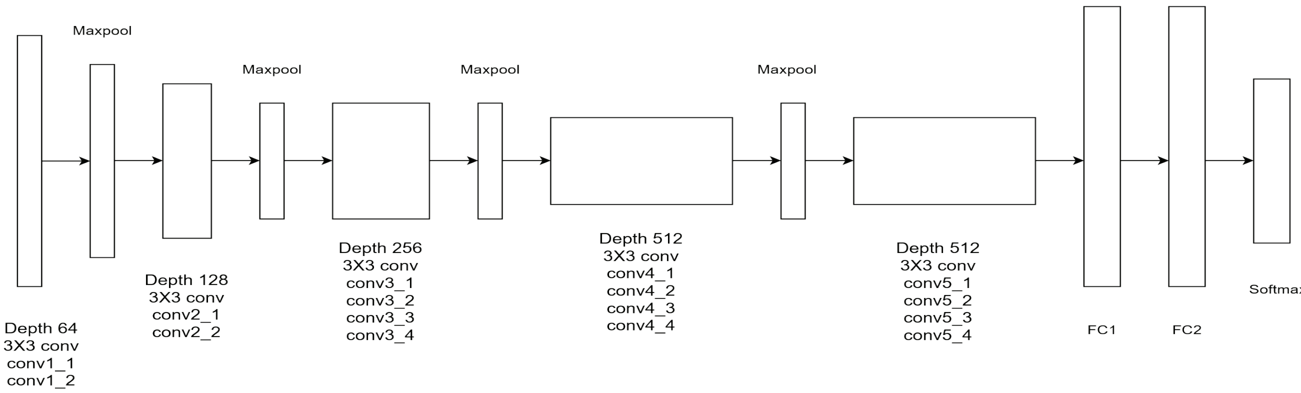

FIG. 5 Pre-trained VGG-19 architecture

Results

The table below presents a comparative analysis of various models' performance on an MRI image dataset, focusing on two key metrics: Structural-Similarity-Index and Peak Signal to Noise Ratio. Among the models evaluated, our proposed model achieved a SSIM score of 29.578 and a PSNR of 0.897.

Table i

Results

Method	Lumber spine MRI Test Set	
	PSNR	SSIM
Bicubic	25.126	0.984
RCAN	26.885	0.989
RWMAN	27.077	0.990
ESRGAN	26.153	0.987
MedSRGAN	26.184	0.987
Ours	29.578	0.897

In the pre-training phase, the model underwent training on a large dataset, specifically the lumber spine dataset, to acquire generic features. Over the 32 epochs, a decreasing trend in loss was observed as the model assimilated knowledge from the dataset. However, due to the disparity between the pre-training and task-specific datasets, the decline in loss might not have been as rapid as during fine-tuning. Nevertheless, a discernible overall enhancement in the model's performance was observed throughout the epochs.

Subsequent to pre-training, the fine-tuning process entailed the adjustment of parameters in the pre trained model using the specific task dataset. Across 32 epochs, a distinct reduction in loss was evident, signifying iterative improvement in the model's performance on the task. Additionally, it's likely that the model's accuracy or relevant performance metric displayed an ascending trend over the epochs, mirroring its learning trajectory.

During the pre-training phase, the model acquired generic features from the lumber spine dataset, serving as a robust foundation for subsequent fine-tuning. Although the loss decreased over epochs, the rate of decline might have been comparatively slower than in fine-tuning due to dataset differences. Nonetheless, the pre-trained model laid a crucial groundwork for the subsequent fine-tuning phase. Fine-tuning involved the adaptation of pre trained model's acquired features to the intricacies of the task-specific lumber spine dataset. The declining loss trend over epochs indicated the model's effective learning from the task-specific dataset. Throughout the fine-tuning process, a consistent improvement in the model's performance on the task was observed, facilitated by the refinement of its parameters. Further iterations are expected to lead to enhanced performance compared to training the model from scratch on the task dataset.

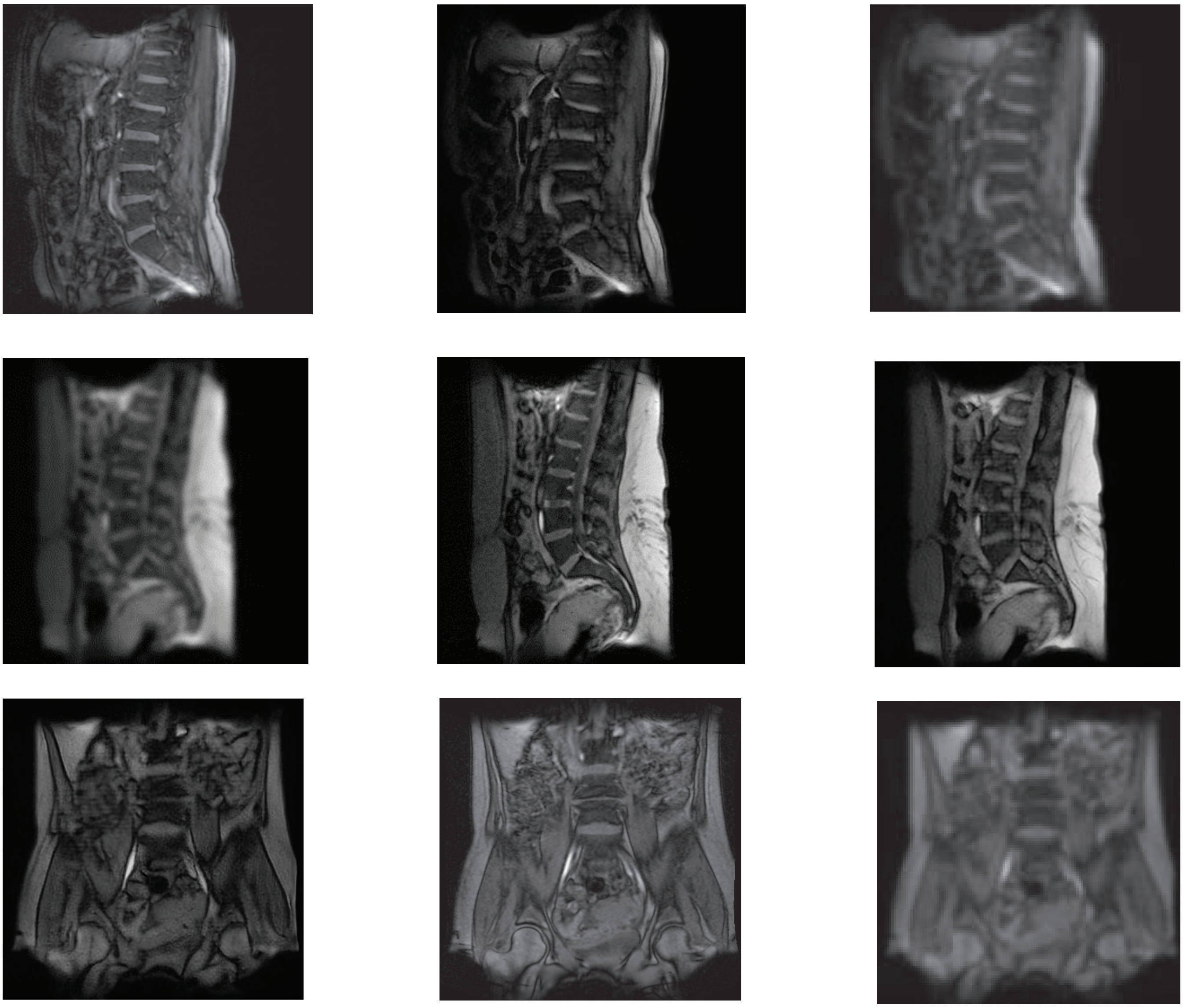

FIG. 6 LR images, S GT images, SR images

Our implementation of the Generative Adversarial Network for medical image super resolution, leveraging a pre trained model with extensive training on a large dataset, has yielded remarkable results with minimum number of epochs. Through rigorous evaluation, our model has demonstrated superior performance in improvising the resolution and qualities of low-resolution medical image, outperforming other conventional approaches.

MOS (Mean Opinion Score) is a subjective quality measurement used to assess the perceived visual quality of super-resolved images. It involves conducting experiments where human observers rate status of super-resolved images on numerical range. Purpose is to evaluate how well an image super-resolution algorithm performs from a human perceptual standpoint. Calculation of score in MOS test is done by calculating the mean of scores of every image used in the test. In our MOS test, we provided 3 set of images to the radiologist. The ratings were provided on the scale of 1 – 5.

Conclusions

Super-Resolution Generative Adversarial Networks (SRGANs) have been applied to lumbar spine MRI with promising results and implications in the quest to improve diagnostic imaging. The process of creating and assessing the “Lumbar Spine MRI Super Resolution using SRGANs” project has produced important discoveries and advancements in the area. We were capable to get psnr score of 29.578 and ssim score of 0.897. Upon conducting MOS test with radiologist working in Nepal we achieved a score of 2.88 . We were also able to create a user friendly website that is easy to use and is accessible to everyone. Our website has been deployed to azure cloud. This system with enhancements can be greatly used in medical field and help in future research.

Acknowledgement

We want to express our deep-seated gratitude to Institute of Engineering, Thapthali Campus for the infrastructure and facilities that made this research viable. We would like to express our deep-seated gratitude to our supervisor, Er. Suramya Sharma Dahal, for his continuous encouragement, constructive criticism, and expert guidance in the formulation of this research. His vast knowledge has been priceless in helping us shape the scope of this research.

We would like to acknowledge the Department of Electronics and Computer Engineering, Thapathali Campus, for the priceless technical support and resources that were so vital in implementing our super-resolution framework. First and foremost, we would like to express deep gratitude to our colleagues and research collaborators, for the enlightening discussions, invaluable criticism, and help throughout the experimentation and writing phases. Their expertise in deep learning or medical imaging areas of choice was of paramount importance for the purpose of fine-tuning these models so that the accuracy was just right.

We would like to thank, in addition, all the contributors to open-source libraries and tools used during this research. Especially, we note that the following pieces of software have greatly contributed to the development and execution of our super-resolution models such as TensorFlow, PyTorch, etc.

References

[1] Y. Gu, Z. Zeng, H. Chen, J. Wei, Y. Zhang, B. Chen, Y. Li, Y. Qin, Q. Xie, Z. Jiang and Y. Lu, "MedSRGAN: medical images super-resolution using generative adversarial networks," *Multimedia Tools and Applications,* 2020.

[2] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang and W. Shi, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 4681-4690, 2017.

[3] B.LIU and J.CHEN, "A super resolution based on attention mechanism and srgan network," *IEEE Acess,* 2021.

[4] Y. Ma, K. Liu, H. Xiong, P. Fang, X. LI, Y. Chen, Z. Yan, Z. Zhou and C. Liu, "Medical image super-resolution using a relativistic average generative adversarial network," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* 2021.

[5] X. Bing, W. Zhang, L. Zheng and Y. Zhang, "Medical Image Super Resolution Using Improved Generative Adversarial Networks," vol. 7, pp. 145030 - 145038, 2019.

[6] J. Hu, L. Shen and G. Sun, ""Squeeze-and-excitation networks," IEEE conference on computer vision and pattern recognition, 2018.," pp. 7132-7141, 2018.*FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[7] J. Kim, K. J. Lee and M. K. Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1646-1654, 2016.A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.

[8] B. Z. Demiray, M. Sit and . I. Demir, "D-SRGAN: DEM Super-Resolution with Generative Adversarial Networks," vol. 2, 2021.

[9] S. Sudirman, A. Al Kafri, F. Natalia, H. Meidia, N. Afriliana, W. Al-Rashdan, M. Bashtawi and M. Al-Jumaily, "Lumbar Spine MRI Dataset," Mendeley Data, 3 April 2019. [Online]. Available: https://data.mendeley.com/datasets/k57fr854j2/2.

[10] A. A. K. J. M. H. M. A. Z. K. a. N. M. D. Khaledyan, "Low-Cost Implementation of Bilinear and Bicubic Image Interpolation for Real-Time Image Super-Resolution," in *2020 IEEE Global Humanitarian Technology Conference (GHTC)*, Seattle, WA, USA, 2020.

[11] Z. Yulun, K. Li, K. Li, L. Wang, B. Zhong and Y. Fu, "Image Super-Resolution Using Very Deep Residual Channel Attention Networks," *Proceedings of the European Conference on Computer Vision (ECCV),* pp. 286-301, 2018.

[12] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao and C. C. Loy, "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," *Proceedings of the European Conference on Computer Vision (ECCV) Workshops,* pp. 0-0, 2018.