Acta Scientific COMPUTER SCIENCES

Volume 4 Issue 6 June 2022

Research Article

Efficient Super-resolution For Chest X-rays

Karthik Sivarama Krishnan* and Koushik Sivarama Krishnan

Gen Nine Inc, United States

***Corresponding Author:** Karthik Sivarama Krishnan, Gen Nine Inc, United States.

Received: April 30, 2022

Published: May 18, 2022

Abstract

High-resolution images are really helpful in various applications like medical diagnosis and hence the need for super-resolution has also increased significantly. Increasing the image resolution on various medical images like a chest X-ray or cell images can improve the accuracy of diagnosis by revealing previously unseen details. Using Image super-resolution also reduces the number of X-ray radiations required to render ultra high-quality imaging. Hence we applied super-resolution on X-rays using fine-tuned Swift-SRGAN architecture, which significantly improved the details on the chest X-rays. This helps in rendering super-resolution images from low-resolution images with less computational requirements. The proposed approach achieves a Structural Similarity Index Measure(SSIM) of 0.893 and a Peak Signal-to-Noise Ratio (PSNR) of 32.10.

Keywords: X-rays; Super-resolution; Swift-SRGAN; Generative Adversarial Network (GAN); Chest X-ray; Medical Imaging; Mobile Computing; Peak Signal-to-Noise Ratio (PSNR); Structural Similarity Index Measure (SSIM)

Abbreviations

SSIM: Structural Similarity Index Measure; PSNR: Peak Signal-to-Noise Ratio; Swift-SRGAN: Swift-Super-Resolution Generative Adversarial Network; GAN: Generative Adversarial Network; CLAHE: Contrast Limited Adaptive Histogram Equalization

Introduction

Since the invention of Convolutional Neural Networks for image processing applications, image super-resolution has attracted the interest of numerous experts. Many research communities have focused on the reconstruction of a single high-resolution image from a low-quality low-resolution image [17]. The process of recovering and reconstructing the resolution of a noisy low-quality image into a very high-quality high resolution image is known as image super-resolution. For example, Reconstructing and recovering a low-quality 256x256 pixels image into a higher quality super-resolution 1024x1024 pixel image is an example of an Image super-resolution task. Image super-resolution has been an area of keen research for some time now through various advancements in the field of deep learning and growing computational power.

With the evolution of Deep Learning techniques, the requirement for high-end computing power to execute calculations of such sophisticated neural network topologies is increasing. With the growing costs of higher-end computing devices, there's a huge demand to make the applications run in real-time and are mobile-friendly. When it comes to the domains of wearable health tech [16], robotics [6], and augmented reality, there is a growing need for real-time applications with minimal latency and very low footprint requirements.

This strategy can also be used in a variety of situations. When it comes to medical imaging, issues like scanning duration, spatial convergence, and signal-to-noise ratio complicate capturing a very high-resolution Magnetic Resonance Imaging (MRI) [1]. As a result,

adopting image super-resolution principles solves these issues by upscaling low-resolution photos into high-resolution images. X-rays and CT scans are in the same boat. Several researches have shown that adding image super-resolution models to chest X-ray (CXR) pictures can enhance pulmonary illness detection greatly [2,3].

Existing imaging super resolution techniques are computationally intensive, time-consuming processes, and may need the optimization of a large number of training parameters. In addition, existing approaches for imaging super-resolution and visual task analysis involve independently trained models.

This paper focuses on improving the performance in upscaling medical X-rays. The architecture of the Swift-SRGAN is fine-tuned to render higher resolution X-ray images from low resolution in near real-time [5]. By using Depth-wise separable convolutions in place of standard convolution layers in the architecture, the required trainable parameters are significantly reduced. This helps in improving the latency of the architecture and thus boosts the performance in real-time.

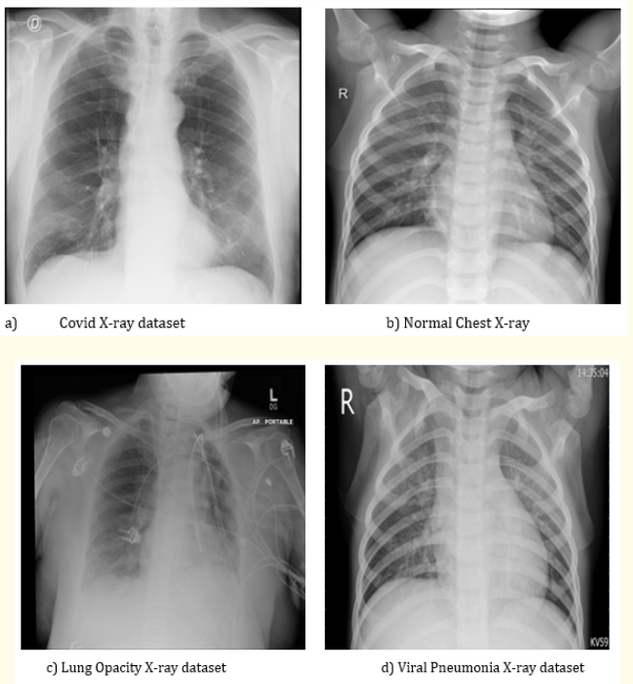

Figure 1: Combined X-ray datasets consisting of a) Covid X-ray data, b) Normal Chest X-ray, c) Lung opacity X-ray dataset and d) Viral Pneumonia X-ray dataset.

Materials and Methods

Dataset

We train this system on the COVID-19 radiography database dataset from Kaggle [4,10,11]. This dataset was a collective effort of various researchers in cooperation with medical practitioners of Malaysia. This dataset has 3616 covid chest X-rays, 6012 lung opacity X-rays, 10,200 Normal chest X-rays, and 1345 Viral Pneumonia chest X-rays. We combine them to get a total of 21,173 chest X-ray images. We use 20,000 chest X-ray images for training the system, 586 chest X-ray images for validation, and 587 chest X-ray images for testing the system.

Data preprocessing

We used several data augmentation techniques with the help of the albumentations library [7]. At first, we randomly crop a 256x256 image from the original image. We randomly rotated the images with a maximum limit of 90 degrees and a minimum limit of 35 degrees. We also created fake RGB channels by repeating the same channel thrice since the pre-trained Swift-SRGAN was trained on RGB images. We also applied CLAHE (Contrast Limited Adaptive Histogram Equalization) as an enhancement technique to remove excessive contrast from the images [8]. Finally, we randomly flipped the images horizontally 50% of the time and randomly rotated the image vertically 20% of the time. All these image augmentation techniques combined yielded a better-generalized model.

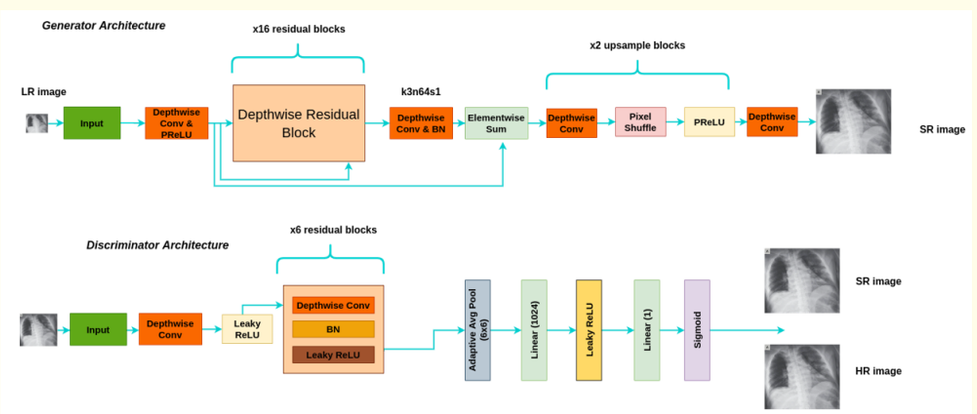

Figure 2: Architecture diagram of Swift-SRGAN's generator and discriminator.

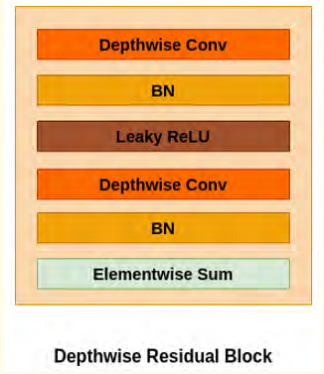

Figure 3: The above figure shows the depthwise residual block used on the generator part of the Swift-SRGAN.

Fine tuned swift-SRGAN

The proposed system is based on the fast and efficient Swift-SRGAN architecture that can perform real-time super-resolution even on low-end computing devices [5]. The Swift-SRGAN architecture consists of generator and discriminator architectures. The generator architecture takes in the input low-resolution chest X-ray image and passes it through a series of Depth-wise Separable Convolutions, Batch Normalization, and PReLU as the activation function. Figure 2 explains the generator and discriminator architecture of the fine tuned Swift-SRGAN network. There are a total of sixteen residual blocks in the generator architecture. This residual block is then passed through a series of Depth-wise Separable convolutions, batch Normalization, and element-wise summation operations to add the outputs of the previous and the current block. Figure 3 shows the residual block contents. The image is then up-sampled twice by running it through the up-sample Block. The depth-wise separable convolutions are followed by two Pixel Shuffle layers with PReLU as the activation function in the up-sample block. The final Depth-wise Separable Convolution layer is then applied, with a kernel size of Nine, three output channels, and a stride of one.

Eight Depth-wise Separable Convolution blocks make up the discriminator architecture. A Depth-wise Separable Convolution [9] is used in all blocks, followed by batch normalization and the activation function LeakyReLU (negative slope = 0.2). There is no batch normalization layer in the first block. The output of the final Convolution block is sent to the adaptive average pooling layer, which produces a 6x6 output. The adaptive average pooling layer's output is flattened and fed into a 1024-neuron linear layer.

Experimental setup

The experimental setup consists of an NVIDIA RTX 2060 Graphics Processing Unit (GPU) which is used to train the fine-tuned Swift-SRGAN architecture. The architecture takes in the low-resolution images of dimensions 256x256 and up-samples it to the high resolution of dimensions 1024x1024. AdamW optimizer [15] is used here along with the ReduceLROnPlateau learning rate scheduler in-order to generalize the model. A batch size of 32 images is considered. The perceptual loss function is used here to calculate the generator loss. Mixed-precision support of PyTorch is used to support the training process of the model.

The low-resolution image is passed to the generator and the perceptual loss is calculated. Then the high resolution and super-resolution images are fed to the discriminator block which helps in calculating the discriminator loss. The goal of the generator model is to create as realistic as possible super-resolution images and fool the discriminator into classifying super-resolution images as high-resolution original images. And the goal of the discriminator is to better distinguish between high-resolution and super-resolution images. We use the Swift-SRGAN generator and the discriminator that is pre-trained on DIV2K [13,14] and FLICKR2K [12] datasets and fine-tune them on the COVID-19 radiography database dataset [10,11].

Performance metrics

We benchmark the fine-tuned model on the standard performance metrics Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) [5]. These two metrics together define the super-resolution quality of our proposed system by comparing it with the high-resolution image. Typically, the higher the PSNR and SSIM, the better the super-resolution results. While PSNR uses the signal-to-noise ratio to compare the images, SSIM uses the contrast, luminance, and structure of the images to compare the images.

Results and Discussion

The proposed system uses the pretrained Swift-SRGAN to perform super-resolution on medical images. We achieve a PSNR

of 32.10 and SSIM of 0.893 for medical image super-resolution. The use of depthwise separable convolution on the Swift-SRGAN allows us to perform real-time super-resolution of medical images on low-end computing devices.

The PSNR is formulated as,

$$PSNR = 10 \cdot \log 10(MAX_1{}^2 \div MSE)$$
$$= 20 \cdot \log 10(MAX_1) - 10 \cdot \log 10(MSE)$$

And SSIM is formulated as,

Luminance comparison function,

$$l(x,y) = (2\,\mu_x\,\mu_y + c1) / (\mu^2{}_x, \mu^2{}_y + c1)$$

Contrast comparison function,

$$l(x,y) = (2\,\sigma_x\,\sigma_y + c2) / (\sigma^2{}_x, \sigma^2{}_y + c2)$$

Structure comparison function,

$$l(x,y) = (2\,\sigma_{xy} + c3) / (\sigma_x\sigma_y + c3)$$

Metric	Performance
PSNR	32.10
SSIM	0.893

Table 1: Evaluation metrics for our proposed approach.

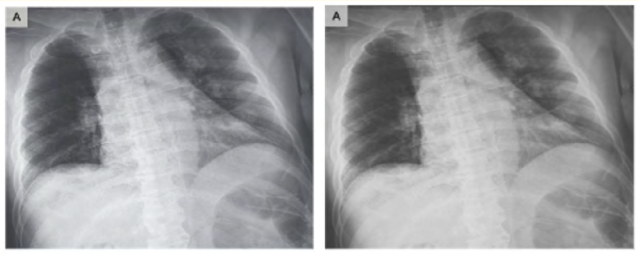

Figure 4: The image on the left is the original high-resolution image and the image on the right is the super-resolution image from the proposed system.

The figure 4 clearly demonstrates the performance of rendering a super resolution image from a single low resolution image.

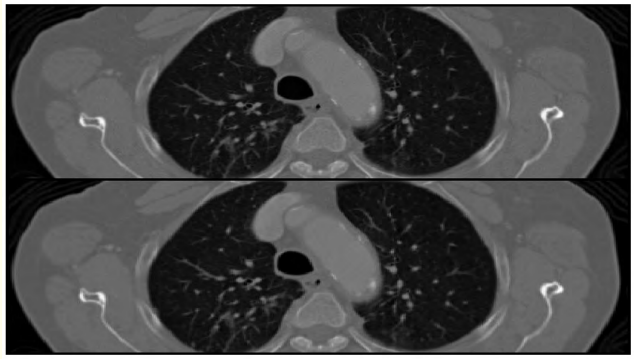

Figure 5: The image compares the super-resolution CT scan (top) and the original CT scan (bottom).

In figure 5, we can see that the super-resolution CT scan has much more details and, hence helps in better diagnosis. The proposed system works really well even on chest CT scans even though it was trained only on chest X-ray images. Thus the proposed system generalizes to all medical images and performs equally well on all kinds of medical images.

Conclusion

Medical super-resolution has proven to be a great help for several medical practitioners to efficiently analyze medical images and perform diagnoses. Using super resolution also reduced the amount of X-ray radiation required to achieve ultra high quality images. The use of Swift-SRGAN allows us to perform real-time and efficient super-resolution even on low-edge computing devices making it perfectly suitable for the task of medical image super-resolution. Combining this with some data augmentation techniques and efficient training strategies like mixed-precision, we achieve a highly efficient system. From table 1, we can see that our proposed lightweight real-time system has achieved a PSNR of 32.10 dB and an SSIM of 0.893. The resultant images from the system are highly detailed and hence help medical practitioners to detect even the slightest anomaly on the X-rays and better diagnose the patient.

Although the system achieves a significant performance while maintaining speed and efficiency, there is still scope for improving the performance by tweaking the Swift-SRGAN architecture. We are planning to extend our work by improving the proposed system's

performance by adding visual attention mechanisms into the Swift-SRGAN architecture and reducing the number of convolutions used without affecting the performance of the architecture.

Conflict of Interest

There is no conflict of interest at this time.

Bibliography

1. Haque Md Inzamam Ul., *et al*. "The effect of image resolution on automated classification of chest x-rays". *medRxiv* (2021).

2. Zamzmi Ghada., *et al*. "Accelerating super-resolution and visual task analysis in medical images". *Applied Sciences* 10.12 (2020): 4282.

3. Zhao Chao-Yue., *et al*. "Chest X-ray images super-resolution reconstruction via recursive neural network". *Multimedia Tools and Applications* 80.1 (2021): 263-277.

4. Krishnan Koushik Sivarama and Karthik Sivarama Krishnan. "Vision Transformer based COVID-19 Detection using Chest X-rays". 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC). IEEE, (2021).

5. K S Krishnan and K S Krishnan. "SwiftSRGAN - Rethinking Super-Resolution for Efficient and Real-time Inference". 2021 International Conference on Intelligent Cybernetics Technology and Applications (ICICyTA) (2021): 46-51.

6. Krishnan Karthik Sivarama and Ferat Sahin. "ORBDeepOdometry-A feature-based deep learning approach to monocular visual odometry". 2019 14th Annual Conference System of Systems Engineering (SoSE). IEEE (2019).

7. Buslaev Alexander., *et al*. "Albumentations: fast and flexible image augmentations". *Information* 11.2 (2020): 125.

8. Reza Ali M. "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement". *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology* 38.1 (2004): 35-44.

9. Krishnan Karthik Sivarama and Koushik Sivarama Krishnan. "Benchmarking Conventional Vision Models on Neuromorphic Fall Detection and Action Recognition Dataset". 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC). IEEE (2022).

10. MEH Chowdhury., *et al*. "Can AI help in screening Viral and COVID-19 pneumonia?" *IEEE Access* 8 (2020): 132665-132676.

11. Rahman T., *et al*. "Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images" (2021).

12. Bee Lim., *et al*. "Enhanced deep residual networks for single image super-resolution". In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops (2017): 136-144.

13. Eirikur Agustsson and Radu Timofte. "Ntire 2017 challenge on single image super-resolution: Dataset and study". In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops (2017): 126-135.

14. Radu Timofte., *et al*. "Ntire 2017 challenge on single image super-resolution: Methods and results". In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops (2017): 114-125.

15. Ilya Loshchilov and Frank Hutter. "Fixing Weight Decay Regularization in Adam". In: CoRR abs/1711.05101 (2017). arXiv: 1711.05101.

16. Karthik Sivarama Krishnan., *et al*. "Recognition of human arm gestures using Myo armband for the game of hand cricket". In: 2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS) (2017): 389-394.

17. Van Ouwerkerk JD. "Image super-resolution survey". *Image and vision Computing* 24.10 (2006): 1039-1052.