# CT scan Reconstruction Using Generative Adversarial Network

RAJLAXMI PATIL[1], TANVI KHARE[2], SANIKA INAMDAR[3], MADHURI MANE[4]
*[1, 2, 3] Computer Engineering Department, Pune Institute of Computer Engineering, Pune, India.*
*[4] Assistant Professor, Pune Institute of Computer Engineering, Pune, India.*

***Abstract- Computer Vision and Image Enhancement through Artificial Intelligence is being used for the improvement in the resolution of the images obtained from variety of sources such as surveillance camera, PAUS , CT scans, radiology results. Implemented method uses a machine learning algorithm, Gen-erative Adversarial Networks (GAN) for achieving the goal of enhancement and reconstruction of images. Super resolution of images allows us to obtain images with better resolution and less noise . Existing models tend to add artifacts to reconstructed im-ages or remove important details from enhanced scans. Cleaner images are obtained at cost of some information loss. Noise in scans should be removed while keeping important information intact and providing a better resolution.In this paper we explore research that has been done for reconstruction of CT scans using various methods.***

***Indexed Terms- Generative Adversarial Network, Image En-hancement,Computer Vision, noise reduction, Computed tomog-raphy, CNN***

## I. INTRODUCTION

CT scans have enabled imaging of different tissues and organs inside the human body in a fast, noninvasive and accu-rate manner. They are used for detecting lesions, tumors, and metastasis. It is a commonly used tool for detecting diseases and injuries within various regions of the body. Higher doses of radiation for CT scans introduce a risk of malignancies induced by radiation. Radiation dose in CT scans is lowered in an optimized way to reduce the ill effects of radiation, but this may lead to reduced image quality. In low-dose CT scans, radiation levels are reduced at the cost of additional LDCT noise and the introduction of artifacts resulting in loss of diagnostic information. CT scans can be improved by using two methods: 1. Hardware-oriented methods 2. Computational methods Hardware-oriented methods are generally expensive and decrease imaging speeds. Thus computational methods are preferable to hardware-oriented methods. Computational methods are classified into three categories: 1.Sinogram fil-tration techniques 2. Iterative reconstruction 3. Image post-processing-based technique In iterative reconstruction-based techniques, an objective function containing an accurate sys-tem model, a statistical noise model, and information about the image is optimized to produce high-quality images. They can successfully remove noise from low-dose scans but create blotchy images. Details are lost in iterative reconstruction, and artifacts are also not removed. With increasing strength of the iterative algorithm, these problems are worsened as images created have low contrast making the detection of lesions harder. Practical utilization of these algorithms is also not feasible due large computational power that is required for the process. Deep learning has been used for many image-to-image translation tasks recently. It is an image post-processing technique that has proved to be better than iterative reconstruction.DL systems do not require high computational power after the initial training and perform better.

## II. ARTIFICIAL INTELLIGENCE IN IMAGE RECONSTRUCTION

### A. Commercially used algorithms

There are multiple DL-based algorithms have been used for CT scan reconstruction. Two commercially used algorithms approved by FDA are Advanced Intelligent Clear-IQ engine (AiCE, Canon Medical Systems) and TrueFidelity(GE Health-care). AiCE uses a deep Convolutional neural network trained on pairs of low-dosage and high-dosage Ct scans. It learns to differentiate between noise ad high signal input from these image pairs. It takes a low-dosage scan as input and produces a corresponding scan with lesser noise as output. TrueFidelity uses a deep neural network to understand characteristics of high-dosage scans, such as noise, contrasts, and image texture, by comparing the high-dose image with multiple low-dose images. These techniques can efficiently reduce noise but require LR-HR image pairs for training. Creating these pairs requires subjecting patients to a high dose of radiation. There are other approaches to CT image reconstruction as well.

B. Previous Work

Artificial intelligence-based methods are being used for CT scan image reconstruction due to more data availability and improved computational capabilities for the training and testing of models. Algorithms proposed to improve SNR and resolution fall into broadly two categories: (a)Model-based reconstruction methods: These models are based on the image degradation process and produce optimal image quality (b) Learning-based method: In this method, models learn a non-linear mapping from datasets containing low-resolution and high-resolution image pairs. Machine learning systems can be trained for pattern recognition to suppress noise in low-dose CT scans. Multiple Deep learning approaches have been used for image reconstruction. Using deep learning models can extract important features containing diagnostic information and differentiate between signal data and noise present in images. These reconstruction techniques can work well with sparse data also. They extract information from labeled as well as unlabelled data. They can suppress noise and remove artifacts while avoiding the over-smoothening of images. A deep convolutional network model was proposed by [1], which tried to learn the mapping between a low-dose and normal-dose CT scan. It reconstructed the image patch by patch. This produced results faster and enhanced image quality significantly but led to the loss of subtleties in images. Reconstructed images were similar to normal dose scans but had differences in some structural details. These differences are marked by red arrows in image Fig.1 from Nie et al. [2] proposed a GAN model to construct a CT scan from a MRI. This model used MRI, a much safer option than CT scans, to create an accurate CT scan equivalent. The model uses a Full Convolutional Network to learn the non-linear mapping of radiation from MRI to CT scan using an adversarial training strategy. A CNN-based model proposed in [3] required 10s for testing per CT volume, suggesting that DL-based techniques perform not better just in terms of noise reduction but are a more practical approach. Noise reduction done by the model made the detection of phantom calcifications added to random images easier. Image-to-image networks may create inconsistent output in case of corrupted input. They are also susceptible to generating incorrect results if the training dataset is biased. Thus training dataset plays an important role in the working of these models. Due to positive results from these earlier works, many GAN models have emerged for CT scan reconstruction.

## III. GENERATIVE ADVERSARIAL NETWORK

GAN architecture was proposed by Ian Goodfellow in [4].GAN has two main components-generator and discrimi-nator. These two components are tightly coupled and trained together. The basic aim of a GAN model is to create data similar to training data. Working of Generator: The generator takes a vector as input and creates an output image from it. The generator tries to create a fake image as close as real as possible to fool the discriminator. If the discriminator is able to detect the image generated by the generator as fake, then

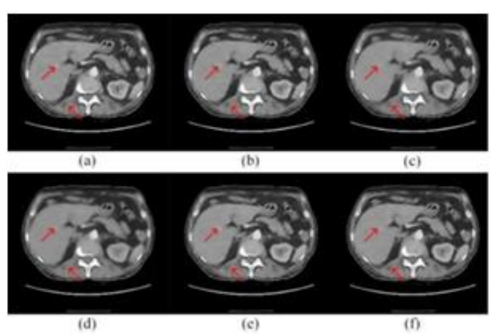


Fig. 1. Result of image processed by: (a)CNN200-1; (b)CNN200-4; (c)CNN200-1-DA; (d)CNN2000-1; (f)CNN2000-4

the loss of the generator increases. In this way, the generator keeps on learning iteratively about the training dataset so that it creates images good enough to fool the discriminator. After a number of iterations increase discriminator also continues to improve, so the images created by the generator also have to keep on improving to fool the discriminator. Working of Discriminator: Discriminator takes data which is an image in this case, as input and predicts whether it is a fake or real image. A fake image is an image generated by a generator, and a real image is an image taken from a training dataset. Loss for the discriminator is calculated using a cross-entropy function. Initially, fake images generated are easily able to fool the discriminator, but eventually, the discriminator learns to differentiate between real and fake images forcing the generator to learn to create better images. The generator and discriminator keep on getting better at their respective tasks till they reach an equilibrium point where images generated by the generator are quite good.

A. Minmax Loss

Minmax loss function was defined in [4] nu Ian Goodfellow as follows:

$$\min_G \max_D Ex[\log(D(x)] + Ex[\log(1 - D(G(z))]$$

The generator tries to minimize the loss while the discrimi-nator tries to maximize it. Adversarial loss causes the generator to produce images similar to the given dataset. But using only adversarial loss, we cannot teach the generator to detect and reduce noise. It also does not consider parameters such as PSNR or contrast. The generator may also remove details from images useful for lesion detection. Consistent results can't be produced using only the minmax function.

B. MSE Loss

To overcome the issue arising due to training pixel recon-struction loss is added. Loss is calculated using the mean squared error between image pixels. This reduces the noise in the image improving the PSNR(peak signal-to-noise ratio), but the visibility of the structural details in the image is reduced as denoising focuses only on trying to reduce the noise in the image. Using per-pixel MSE as a loss function leads to over-smoothed edges. Model tend to overlook features that can be understood and interpreted by a human. As it tries to reduce the loss per pixel, it does not understand structures present in the image may contain important diagnostic information. To reduce loss average values of pixels are calculated, and resultant images turn out a little blurry and may also contain artifacts.

C. Wasserstein distance

Q. Yang et al. [5] proposed a model based on GAN archi-tecture that uses Wasserstein distance and perceptual similarity to denoise an image. Wasserstein distance is a concept from optimal transport theory which calculates the effort required to convert one distribution into another. Thus GAN trained using Wasserstein distance focuses on migrating data noise from strong to weak statistically instead of trying to differentiate between noise and actual data. When a model uses per-pixel MSE as a measure to calculate the error between generated output and ground truth, subtleties in the image are ignored, and resultant images do not contain detailed structures. The proposed model in [5] was able to reduce image noise while keeping the details in the scan intact. As shown by the recent work [6], [7], this per-pixel MSE often results in over-smoothed edges and loss of details. As an algorithm tries to minimize per-pixel MSE, it overlooks subtle image textures/signatures containing critical diagnostic information.

D. Perceptual loss

As CT images are not uniformly distributed, using MSE leads to calculating only the Euclidean distance while com-paring images and not understanding the intrinsic similarity between structures present in images. For comparing the similarity of images in a way done by humans, geodesic distance should be calculated. Thus perceptual loss is used in the GAN model proposed by Q. Yang et al. [5] along with Wasserstein distance. While using only MSE only superficial differences between images are seen. While using the Percep-tual loss model learns features present in the image thus the output images contain better details. Using this method visual perception done by the feature extractor helped in maintaining image content while reducing the noise from LDCT. Pre-trained VGG network was used as feature extractor to calculate the perceptual loss. Wasserstein distance reduced the noise from the image while perceptual loss made the network keep important content in image, while the CNN-MSE model produced blurred images with waxy artifacts.WGAN-VGG images were more similar to NDCT images visually from LDCT images.

E. Style transfer loss

MEDGAN model proposed in [8] incorporated losses other than adversarial loss from image style transfer techniques to maintain the required amount of details in the image. Above mentioned losses do not take image properties such as image

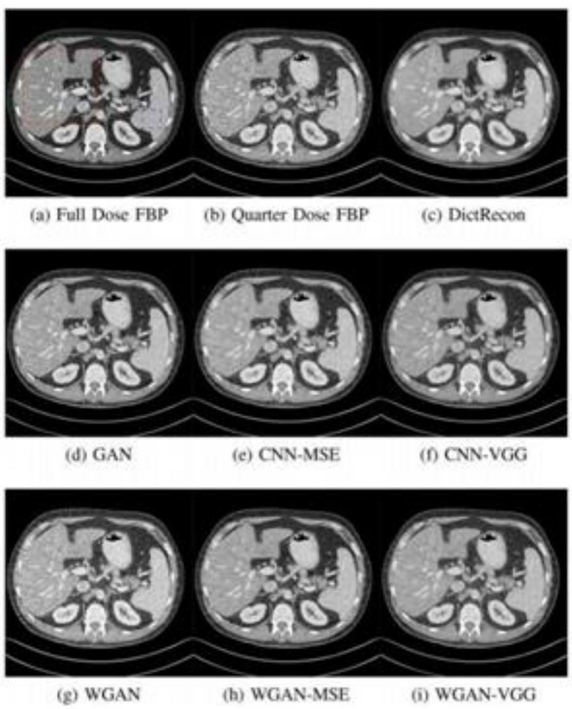


Fig. 2. Result of image processed by different models

texture into consideration. Image style-transfer losses train the model in such a way that the output image contains important details and maintains a image texture while reducing noise from the image. Similar to perceptual loss CNN is used for feature extraction. As the feature extractor is pretrained for the image classification, the resultant images contain rich features while preserving the global structure of the image. Feature ex-tractor network can deeper containing multiple convolutional blocks and be trained on more data to improve the translated images.

F. Cyclic learning ensemble

A novel CNN-based CycleGAN framework was presented in [9], which focused on preserving high-resolution anatomical details. Cycle-consistency constraint was utilized for strong domain consistency between input and output images. Multiple layers were cascaded to learn interpretable and disentangled hierarchical features. The network was designed to alleviate computational overheads. Wasserstein distance and L1 norm were used to refine deblurring such that noise in the image was reduced with minimal loss of anatomical information. To train the model to learn the mapping between LDCT and HDCT four loss function were used: adversarial loss, cycle-consistency loss, identity loss, joint sparsifying transform loss.

## IV. COMPARISION OF VARIOUS MODELS

Q. Yang et al. [5] compared CNN and GAN model along with different loss functions.CNN-MSE produced images that were blurred and contained minor streak artifacts as shown in fig(2) from [2].WGAN-MSE avoided over smooth image better than CNN-MSE but still contained minor streak artifacts which were prominent compared to images produced by models containing VGG loss. Images created by WGAN had stronger noise and white structures originating from streak artifact present in the LDCT image.As the VGG loss is computed in a feature space trained on a large dataset, images produced by CNN-VGG and WGAN-VGG were visually more similar to HDCT images in terms of noise levels and structural details in the image. Quantitative analysis was done by calculating peak signal-to-noise ratio and structural similarity(SSIM).CNN-MSE performed well in PSNR as it was trained using MSE as a loss function.MSE tries to reduce noise without considering structural details. Thus, CNN-MSE outperformed other models.WGAN ranks worst in PSNR and SSIM as it was trained on just Wasserstein distance thus neither reduces noise significantly nor does it detect arti-facts.VGG networks concentrate on structural details as well, so they performed better on SSIM than PSNR. This indicates that MSE loss-based networks are good at noise reduction while VGG-based networks are better at feature preservation. WGAN-VGG network was able to generate an image with less noise while keeping details important for lesion detection.

## V. DRAWBACKS OF GAN

GAN models have to be trained using an unbiased training dataset. The training dataset must be homogenized and should contain high-quality samples from different ages, races, body weight, and gender. The training dataset for most algorithms also requires LR-HR pairs for which a patient has to be subjected to a high dosage of radiation. Any imperfections in the training dataset will be reflected in the output im-ages generated.GANs are also sensitive to hyperparameter values. Thus they are vulnerable to perturbation and may give unstable results. Training deep networks takes a lot of time, and data and artifacts from imperfect data affect the PSNR value, which can be achieved using noise reduction. Data collection is expensive and risky for patients that are subjected to high radiation dosage. To reduce artifacts, images are smoothened, and subtle structures are removed, making diagnosis harder.The training data needs to be properly labeled to avoid any misinterpretation. For creating a reliable model a lot of high-quality images are required. Also, when using a DL-based method to improve reconstruction quality, high-quality raw data reduce the difficulty of the reconstruction task, facilitating images quality improvement

## CONCLUSION

Deep learning techniques are increasing in popularity due to their advantages such as greater image denoising and image texture improvement. Despite their drawbacks GAN models have performed better than other models and continue improving. Images generated by GAN have lesser noise and maintain the image details making lesion detection easier.

## REFERENCES


[1] hen H, Zhang Y, Zhang W, Liao P, Li K, Zhou J, Wang G. Low-dose CT via convolutional neural network. Biomed Opt Express. 2017 Jan 9;8(2):679-694. doi: 10.1364/BOE.8.000679. PMID: 28270976; PMCID: PMC5330597.

[2] Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q, Shen D. Med-ical Image Synthesis with Context-Aware Generative Adversarial Net-works. Med Image Comput Comput Assist Interv. 2017 Sep;10435:417-425. doi: 10.1007/978-3-319-66179-7-48. Epub 2017 Sep 4. PMID:30009283; PMCID: PMC6044459.

[3] Wolterink JM, Leiner T, Viergever MA, Isgum I. Generative Adversarial Networks for Noise Reduction in Low-Dose CT. IEEE Trans Med Imaging. 2017 Dec;36(12):2536-2545. doi: 10.1109/TMI.2017.2708987. Epub 2017 May 26. PMID: 28574346.

[4] Goodfellow, I., et al., 2014. Generative adversarial nets. Advances in Neural Information Processing Systems, pp. 2672–2680.

[5] Yang et al., ”Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss,” in IEEE Transactions on Medical Imaging, vol. 37, no. 6, pp. 1348-1357, June 2018, doi: 10.1109/TMI.2018.2827462.

[6] ohnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style trans-fer and super-resolution. 2016 [Online]. Available: arXiv:1603.08155.

[7] Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W. Photo-realistic single image super-resolution using a generative adversarial network. 2016 [Online]. Available: arXiv:1609.04802.

[8] Armanious, Karim and Yang, Chenming and Fischer, Marc and Kustner,¨ Thomas and Nikolaou, Konstantin and Gatidis, Sergios and Yang, Bin. (2018). MedGAN: Medical Image Translation using GANs.

[9] CT Super-resolution GAN Constrained by the Identical, Residual, and Cycle Learning Ensemble (GAN-CIRCLE) Chenyu You, Guang Li, Yi Zhang*, Senior Member, IEEE, Xiaoliu

Zhang, Hongming Shan, Mengzhou Li, Shenghong Ju, Zhen Zhao, Zhuiyang Zhang, Wenxiang Cong, Michael W. Vannier, Member, IEEE, Punam K. Saha, Senior Member, IEEE, Eric A. Hoffman, Member, IEEE, and Ge Wang*, Fellow, IEEE