ISSN: 2508-7894
KJAI website: http://accesson.kr/kjai
doi: http://dx.doi.org/10.24225/kjai.2024.12.1.1

Feasibility Study of CNN-based Super-Resolution Algorithm Applied to Low-Resolution CT Images

Doo Bin KIM[1], Mi Jo LEE[2], Joo Wan HONG[3]

Received: January 08, 2024. Revised: February 07, 2024. Accepted: February 14, 2024.

Abstract

Recently, various techniques are being applied through the development of medical AI, and research has been conducted on the application of super-resolution AI models. In this study, evaluate the results of the application of the super-resolution AI model to brain CT as the basic data for future research. Acquiring CT images of the brain, algorithm for brain and bone windowing setting, and the resolution was downscaled to 5 types resolution image based on the original resolution image, and then upscaled to resolution to create an LR image and used for network input with the original imaging. The SRCNN model was applied to each of these images and analyzed using PSNR, SSIM, Loss. As a result of quantitative index analysis, the results were the best at 256×256, the brain and bone window setting PSNR were the same at 33.72, 35.2, and SSIM at 0.98 respectively, and the loss was 0.0004 and 0.0003, respectively, showing relatively excellent performance in the bone window setting CT image. The possibility of future studies aimed image quality and exposure dose is confirmed, and additional studies that need to be verified are also presented, which can be used as basic data for the above studies.

Keywords: AI, CNN, Super-Resolution, Windowing, CT

Major Classifications: Artificial Intelligence, Convolution Neural Network, Super-Resolution

1. Introduction

Computed tomography (CT) is a fundamental and important examination for diagnosing diseases. It is used effectively in the medical field. CT can depict anatomical structures of bones, soft tissues, and blood vessels inside the human body in three dimensions. Moreover, conditions such as the window width, window level, and kernel can be adjusted on the raw data acquired from CT scans to obtain images that have high diagnostic value and are adapted to the purpose while reconstructing the data. Among the various types of CT, brain CT provides soft tissue setting images and bone setting images by examining patients with head injuries caused by traffic accidents and similar incidents. Soft tissue setting images are utilized to observe intracerebral hemorrhage and cerebral infarction (Brenner et al., 2007). Meanwhile, bone setting images are used to identify fractures in the bones that form the skull (Hsieh et

* This paper was supported by Eulji University in 2023(EJRG-23-15)
1 First Author. Radiographer, Department of Radiology, Uijeongbu Eulji Medical Center, Korea, Email: kkddbb2011@eulji.ac.kr
2 Second Author. Professor, M.D,Department of Radiation Oncology, Catholic Kwandong University International ST.MARY's Hospital, Korea, Email: minsk7144@hanmail.net
3 Corresponding Author. Professor, Department of Radiological Science, Eulji University, Korea, Email: jwhong@eulji.ac.kr

al., 2013). Sufficient studies have been conducted to demonstrate the clinical necessity and effectiveness of brain CT. Furthermore, according to statistical data published by the Health Insurance Review and Assessment Service (HIRA), the number of brain CT scans conducted for diagnosis, treatment, and health checkups has been increasing (Health Insurance Review & Assessment Service, 2022). However, CT equipment uses X-rays, and exposure to radiation occurs. Furthermore, the radiation dose is higher than that in general X-ray examinations. In addition, the risk of adverse health effects of radiation on the head and neck areas examined in brain CT is high. The International Commission on Radiological Protection (ICRP), which specifies radiation dose limits, has stipulated new tissue weighting factors (ICRP Publication 103, 2007). In recent years, Fourth Industrial Revolution technologies have been applied to various fields, and the utility of these technologies has been demonstrated. In the medical imaging field, studies incorporating various technologies are actively underway. In particular, a convolutional neural network (CNN; one of the major deep learning algorithms) is an artificial neural network that extracts features from images. When applied to medical images, CNN have demonstrated remarkable performance (Shen et al., 2017). A super-resolution technique has also been implemented based on CNN (Dong et al., 2015). This technique improves the resolution of low-resolution (LR) images and outputs these as high-resolution (HR) images. In general, when the amount of X-rays decreases in CT, the resolution of CT images reduces correspondingly. This produces LR images (Whiting et at., 2002). If the CNN-based super-resolution technique is applied to such CT images, it can potentially produce images that have similar image quality as conventional CT images. That is, it may be feasible to generate HR images from LR images obtained using a lower amount of X-rays, thus reducing the radiation dose. Therefore, there is a need to conduct research on the radiation dose and image quality by applying the CNN-based super-resolution technique to brain CT, which has the drawback of high radiation dose but has been demonstrated to be clinically effective and is being used increasingly for examination. To provide foundation data for such research, this study aims to preliminarily assess the effectiveness of applying the super-resolution technique.

2. Related Work

2.1. Research Subjects

The images used in the experiment were collected retrospectively using the anonymization function of the picture archiving and communication system (PACS) from patients who underwent brain CT scans without contrast media from January to February 2021.

2.2. Data Composition and Preprocessing

The dataset was divided into 615 images of the brain and 615 images of bones. Each image was constructed using multi-planar reformation (MPR) with a matrix size of 512×512. Moreover, it was aligned parallelly to the skull base, encompassing the area from the skull base to the vertex. To generate LR images, the ground truth (GT) of the dataset was down-sampled to matrix sizes of 16×16, 32×32, 64×64, 128×128, and 256×256 using the resize function in openCV. Then, the image was upscaled back to the size of the GT (512×512). Hence, the data were designed such that training could be performed with five conditions for each dataset. The training process is shown in Figure 1.

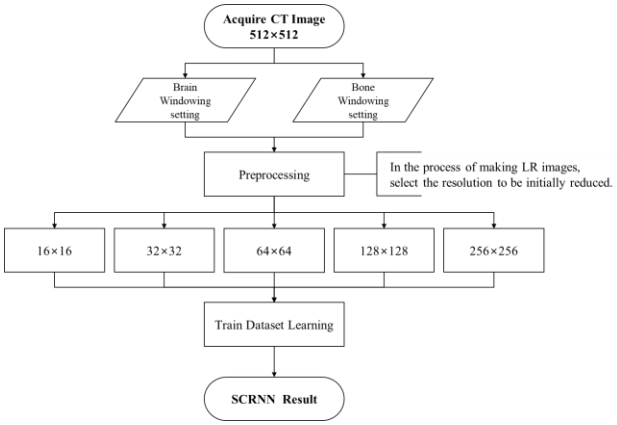


Figure 1: Low-Resolution CT images generation and SRCNN model train dataset

2.3. Modeling and Implementation Environment

The model used in the experiment is a super-resolution convolutional neural network (SRCNN). It is the first model that applied the super-resolution technique to deep learning (Dong et al., 2015). SRCNN has a simple structure and consists of 3 processes patch extraction, non-linear mapping, and reconstruction. This is shown in Figure 2.

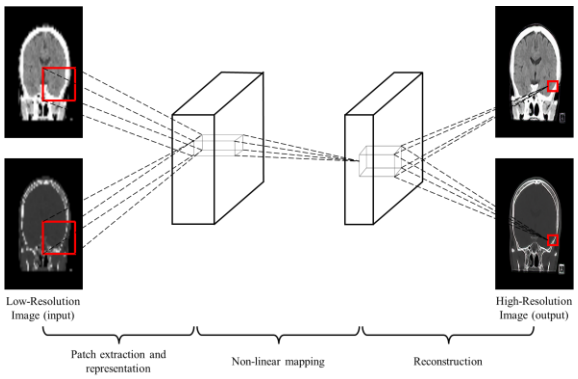


Figure 2: SRCNN network architecture

To implement this model, the system utilized Ubuntu (ver. 20.04.6 LTS), a CPU (Intel Core X i9-9940X), two GPUs (GeForce RTX 3090, D6X 24G), and four 32G RAMs (DDR4 32G PC-425600). In addition, Python (ver. 3.6) and PyCharm Integrated Development Environment (IDE, ver. 2021.2) were used. Moreover, TensorFlow (ver. 2.1.0), Keras (ver. 2.3.1), Matplotlib (ver. 3.2.2), Seaborn (ver. 0.11.2), NumPy (ver. 1.19.2), and OpenCV (ver. 3.3.1) libraries were used for analysis. Rectified linear unit (ReLu) and mean squared error (MSE) were used as the functions (He et al., 2015, Gudivada et al., 2017). Adaptive moment estimation (Adam) was used as the optimizer (Kingma et al., 2014). The equation for each function is shown in Equation (1) – (3).

$$Relu = \begin{Bmatrix} x, x > 0 \\ 0, x \leq 0 \end{Bmatrix} = \max(0, x) \quad (1)$$

Where x is input value, if input value is less than 0, it is 0 and if input value was greater than 0, the input value is sent as is.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (2)$$

Where y_i is correct label and $\hat{y}_i$ is value estimated by the neural network.

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1-\beta_1)\nabla_\theta J(\theta) \\ v_t &= \beta_2 v_{t-1} + (1-\beta_2)\left(\nabla_\theta J(\theta)\right)^2 \\ \theta &= \theta - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}}\hat{m}_t \end{aligned} \quad (3)$$

Where m_t is moment v is exponential averaged, v is exponential average of the slope squared and is G value of RMSProp. And generally, β_1 is 0.9, β_2 is 0.999, ϵ is 10^{-8}.

2.3. Evaluation Metrics

Quantitative evaluation metrics were used to evaluate the results. The evaluation was performed by calculating the peak signal-to-noise ratio (PSNR), structural similarity index map (SSIM), and loss measured with MSE for each result (Johnson et al., 2006; Wang et al., 2004; Mathieu et al., 2015). The equation for each metric is shown in Equation (4) and (5).

$$PSNR = 10\log_{10}\left(\frac{R^2}{MSE}\right) \quad (4)$$

Where R is maximum value of the pixel.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C1)(2\sigma_{xy} + C2)}{(\mu_x^2 + \mu_y^2 + C1)(\sigma_x^2 + \sigma_y^2 + C2)} \quad (5)$$

Where μ_x, μ_y is local means, σ_x, σ_y is standard deviations, σ_{xy} is cross-covariance for images x, y. And $C1$ is $(K_1 L)^2$, $C2$ is $(K_2 L)^2$, K_1 is 0.01, K_2 is 0.03, L is specified dynamic range value.

3. Performance Evaluation and Results

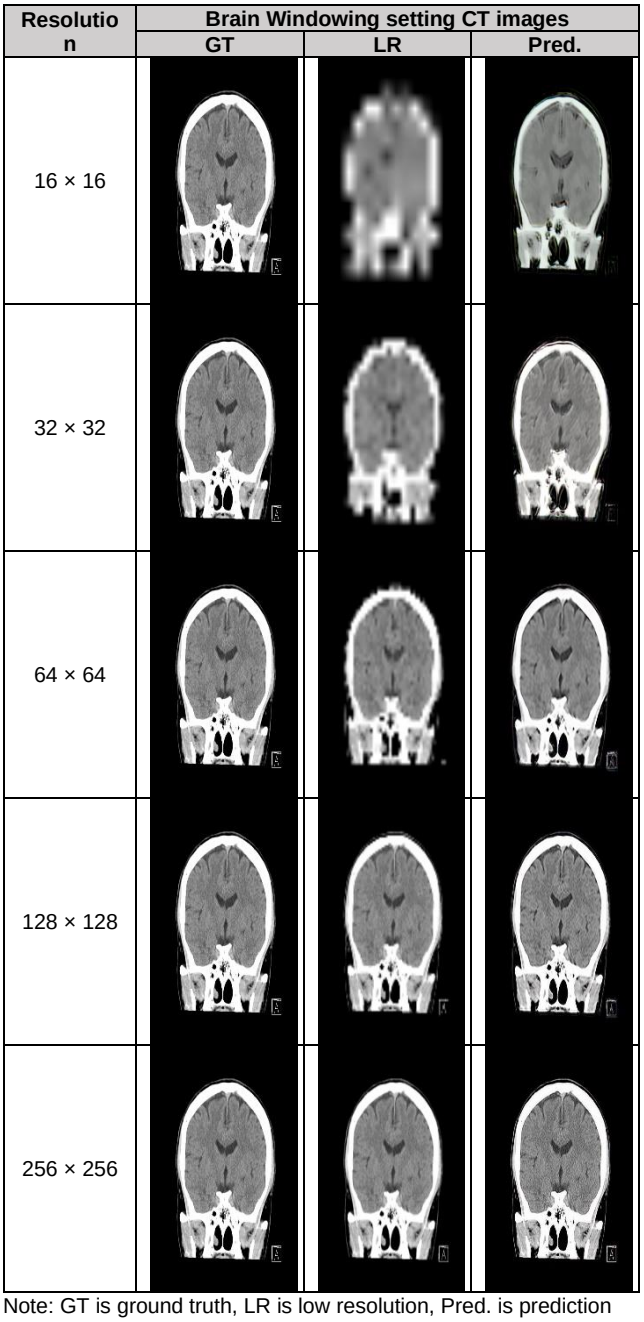
3.1. Brain Windowing Setting CT Images

Brain windowing setting CT image quality evaluation metric of PSNR and SSIM of CT images increased as the resolution increased, and the loss decreased, so that LR images were generated into HR images through the application of SRCNN model, and the image quality of the generated images was improved at the same time. Additionally, the percentage differences between PSNR, SSIM and Loss minimum and maximum values increased by 219.76%, 142.83%, and decreased by 1.45%, respectively. Table 1 shows the results obtained using the quantitative evaluation metrics. In addition, Figure 3 shows the images for each training result obtained using the brain windowing setting CT dataset.

Table 1: Evaluation metrics for results of brain windowing setting CT dataset

Resolution	PSNR(dB)	SSIM	Loss
16 × 16	15.34	0.68	0.0299
32 × 32	18.36	0.82	0.0152
64 × 64	22.16	0.92	0.0063
128 × 128	27.44	0.96	0.0019
256 × 256	33.72	0.98	0.0004

Note: PSNR is peak signal-to-noise ratio, SSIM is structural similarity index map



Note: GT is ground truth, LR is low resolution, Pred. is prediction

Figure 3: Results images of SRCNN according to CT brain windowing setting

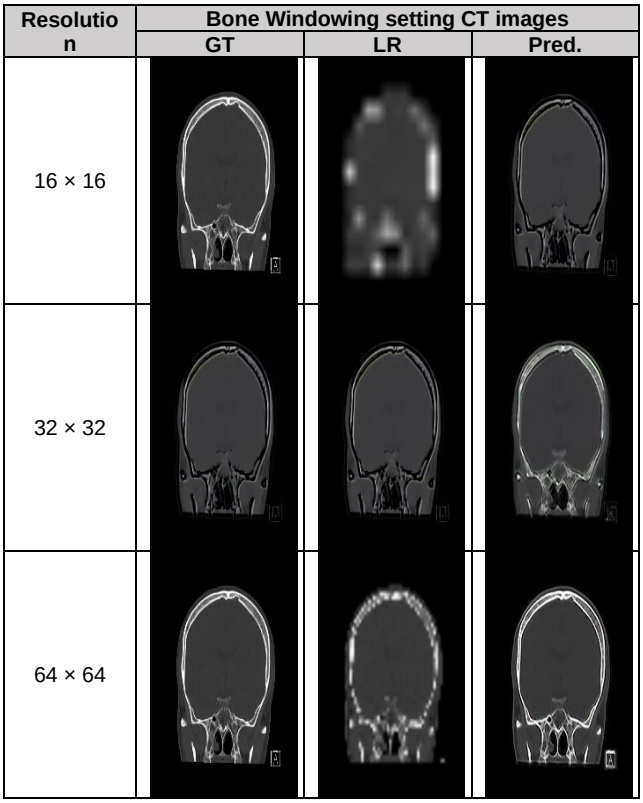
3.2. Bone Windowing Setting CT Images

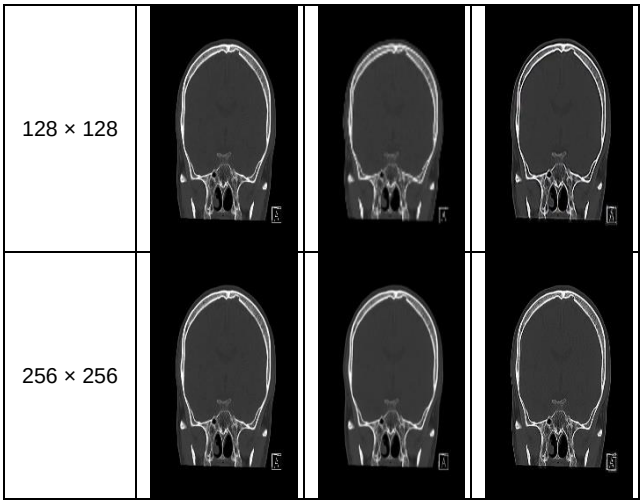
Bone windowing setting CT image quality evaluation metric of PSNR and SSIM of CT images increased as the resolution increased, and the loss decreased, so that LR images were generated into HR images through the application of SRCNN model, and the image quality of the generated images was improved at the same time. Additionally, the percentage differences between PSNR, SSIM and Loss minimum and maximum values increased by 193.16%, 155.91%, and decreased by 2.01%, respectively. Table 2 shows the results obtained using the quantitative evaluation metrics. In addition, Figure 4 shows the images for each training result obtained using the bone windowing setting CT dataset.

Table 2: Evaluation metrics for results of bone windowing setting CT dataset

Resolution	PSNR(dB)	SSIM	Loss
16 × 16	18.22	0.63	0.0153
32 × 32	20.35	0.84	0.0094
64 × 64	23.89	0.93	0.0042
128 × 128	29.04	0.96	0.0013
256 × 256	35.20	0.98	0.0003

Note: PSNR is peak signal-to-noise ratio, SSIM is structural similarity index map





Note: GT is ground truth, LR is low resolution, Pred. is prediction

Figure 4: Results images of SRCNN according to CT bone windowing setting

4. Discussion

This study was conducted to apply the CNN-based super-resolution model to brain CT images to obtain visible images, assess the effectiveness of applying the super-resolution model by analyzing the results using the evaluation metrics, and provide foundational data for future research related to radiation dose and image quality. For both the datasets, remarkable PSNR, SSIM, and loss values were achieved by the training of LR images down-sampled using a higher matrix size than by the training of LR images down-sampled using a lower matrix size. For the bone dataset, the LR images down-sampled with a matrix size of 256×256 achieved the best performance (PSNR = 35.20183, SSIM = 0.976, and loss = 0.000307). Moreover, the results for the bone dataset were generally higher than those for the brain dataset. This outcome can be attributed to the fact that the portion of the brain image occupied by emphasized pixels is smaller than that of the bone image, which is advantageous for computations required for the training and facilitates the extraction of features in the convolution layer (Brownlee et al., 2016). These results were compared with the most remarkable results of similar previous studies. Georgescu et al. used a brain CT dataset and obtained the following super-resolution results: PSNR = 36.39 and SSIM = 0.9291(Georgescu et al., 2020). Wang et al. applied the super-resolution technique to brain MRI images. They attained a PSNR of 37.38 and an SSIM of 0.9623(Wang et al., 2020). The PSNR value of this study was lower than those of the previous studies, whereas its SSIM value was higher. The parameters, datasets, training methods, and structure of the model used differed across the studies.

Therefore, this is not an effective comparison. However, a simple comparison could be performed using the same evaluation metrics, and it can be determined that the results of this study are significant based on the comparison result.

A notable aspect of the method used in this study is that the previous studies trained their models using images focused on soft tissues inside the brain, whereas the model in this study was trained on soft tissue images as well as bone images. This approach is significant because it is based on real-world clinical cases. Moreover, it supplements the deficiency when assuming the application of this method in clinical practice. Furthermore, by utilizing the LR images obtained using the GT and by varying the matrix size from 16×16 to 256×256 to train the model, this approach implemented images with varying resolution according to a decrease in radiation. As a result, it provided foundational data and reference data for comparing the results when applying the super-resolution technique to images obtained according to different radiation doses. The limitations of this study are as follows. In general, better results can be obtained by using a large amount of data to train the model. Many cases support this assertion, and studies have been conducted that utilized data augmentation to perform training for better results (Cubuk et al., 2018, Zhang et al., 2018). However, the experiment in this study was not conducted with a relatively large amount of data. It was difficult to identify cases that quantitatively indicated the amount of data appropriate for the training method used in this study. Furthermore, because this study was an experimental approach to research, few reference papers were considered. In addition, similar previous studies used different amounts of data. Therefore, the amount of data was determined by considering the performance of the workstation used for training. It is likely that results with better metrics than those of the results of this experiment can be obtained if the amount of data increases dramatically.

However, it is considered that the results of this study were sufficiently significant in that these aligned with the objectives of this study. An additional limitation of this study is that the evaluation metrics used in this experiment (PSNR, SSIM, and loss) are generally used to evaluate super-resolution. However, there are challenges with fully evaluating super-resolution quantitatively (Wang et al., 2004). CT is a very important imaging examination in diagnosing diseases, but radiation dose reduction is necessary considering the side effects caused by radiation. However, a decrease in radiation dose results in a decrease in the signal required for image acquisition, which results in increased noise and reduced image quality, resulting in low-quality images that are unsuitable for diagnosis, making it

difficult to distinguish between anatomical structures and lesions. This study is because super-resolution is a technique that predicts and presents solutions to unsolved problems.

Consequently, the results can be interpreted differently depending on the purpose or the individual receiving the results, which is a subjective assessment. Hence, there is a limitation in performing the evaluation simply with quantitative metrics. Therefore, it is considered that additional analysis, in conjunction with the existing evaluation metrics, is required to perform important evaluations for the required purpose.

References

Brenner, D. J., & Hall, E. J. (2007). Computed tomography—an increasing source of radiation exposure. *New England journal of medicine*, *357*(22), 2277-2284.

Brownlee, J. (2016). Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras. *Machine Learning Mastery*.

Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., & Le, Q. V. (2018). Autoaugment: Learning augmentation policies from data. *arXiv preprint arXiv:1805.09501*.

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.

Georgescu, M. I., Ionescu, R. T., & Verga, N. (2020). Convolutional neural networks with intermediate loss for 3D super-resolution of CT and MRI scans. *IEEE Access*, *8*, 49112-49124.

Gudivada, V., Apon, A., & Ding, J. (2017). Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations. *International Journal on Advances in Software*, *10*(1), 1-20.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision* (pp. 1026-1034).

Hsieh, J., Nett, B., Yu, Z., Sauer, K., Thibault, J. B., & Bouman, C. A. (2013). Recent advances in CT image reconstruction. *Current Radiology Reports*, *1*, 39-51.

Jhnson, D. H. (2006). Signal-to-noise ratio. *Scholarpedia*, *1*(12), 2088.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Mathieu, M., Couprie, C., & LeCun, Y. (2015). Deep multi-scale video prediction beyond mean square error. *arXiv preprint arXiv:1511.05440*.

National Health Insurance Service. (2022). 2022 Major Health Insurance Statistics. Retrived from: https://www.hira.or.kr/bbsDummy.do?pgmid=HIRAJ030000007001&brdScnBltNo=4&brdBltNo=7&pageIndex=1&pageIndex2=1#none

Protection, R. (2007). ICRP publication 103. *Ann ICRP*, *37*(2.4), 2.

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual review of biomedical engineering*, *19*, 221-248.

Wang, J., Chen, Y., Wu, Y., Shi, J., & Gee, J. (2020). Enhanced generative adversarial network for 3D brain MRI super-resolution. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 3627-3636).

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, *13*(4), 600-612.

Whiting, B. R. (2002, May). Signal statistics in x-ray computed tomography. In *Medical Imaging 2002: Physics of Medical Imaging* (Vol. 4682, pp. 53-60). SPIE.

Zhang, Y., & An, M. (2017). Deep learning-and transfer learning-based super resolution reconstruction from single medical image. *Journal of healthcare engineering*, *2017*.