2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE) | 979-8-3503-2095-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCoSITE57641.2023.10127803

A Comparison Between Interpolation Method and Neural Network Approach in 3D Digital Imaging and Communications in Medicine

Muhammad Ibadurrahman Arrasyid Supriyanto
Institut Teknologi Sepuluh November
Surabaya
ibad.arrasyid@gmail.com

Riyanarto Sarno
Institut Teknologi Sepuluh November
Surabaya
riyanarto@if.its.ac.id

Chastine Fatichah
Institut Teknologi Sepuluh November
Surabaya
chastine@if.its.ac.id

Aziz Fajar
Institut Teknologi Sepuluh November
Surabaya
azizfajar519@gmail.com

***Abstract*— Higher image reconstruction with excellent structural detail allows experts to perform accurate analysis, especially on the smallest organ details. The interpolation method that approaches the problem of medical image reconstruction, especially 3D, still causes serious problems. The medical image produced by the interpolation method produces blurred or smooth lines on some parts of the organ. This can cause errors in the medical analysis that will be carried out if the reconstruction results are problematic. For this reason, a method is needed that can reconstruct images well without producing blur but does not require very large computer resources. This study aims to evaluate and compare the quality of 3D magnetic resonance imaging medical images reconstructed using interpolation methods and artificial neural network architectures in the DICOM data format. This study evaluates and compares the quality of 3D magnetic resonance imaging medical images reconstructed using interpolation methods and artificial neural network architectures. The test scenario was performed using images from the ADNI dataset and comparing the output results using a variational autoencoder and a multi-level densely connected super-resolution network on 3D data with existing interpolation methods. The evaluation was done using two metrics, i.e., SSIM and PSNR. The results showed that the variational autoencoder method has the highest SSIM and PSNR values, indicating it has the highest image quality among the three methods, while the mDCSRN method has the lowest SSIM and PSNR values, meaning it has the lowest image quality.**

Keywords—DICOM, Reconstruction, MRI Brain, Computer Vision

I. INTRODUCTION

Higher image reconstruction with excellent structural detail allows experts to perform accurate analysis, especially on the smallest organ details. However, to produce high-resolution images, experts usually use devices with superconducting magnets that are very large, which is very burdensome in terms of infrastructure, size, and cost. Not all health facilities can meet these needs. Advances in technology can overcome these needs using reconstructing using the interpolation method. The interpolation method is possible because it does not require a large device to run. The interpolation method can also be performed with data with 2D and 3D dimensions, so this method is very flexible.

However, the interpolation method [10] that approaches the problem of medical image reconstruction, especially 3D, still causes serious problems. The medical image produced by the interpolation method produces blurred or smooth lines on some parts of the organ as seen in Figure 1. In Figure 1, which is shown by the direction of the small arrow, the parts should intersect each other but what happens instead is that they merge, causing the parts to become one organ. This can cause errors in the medical analysis that will be carried out if the reconstruction results are problematic. For this reason, a method is needed that can reconstruct images well without producing blur but does not require very large computer resources.

Some research has been done on medical image processing such as super resolution. Super-resolution methods have potential benefits, thus many focus on optimizing them to improve accuracy and speed using deep learning convolutional networks. however, this study only focuses on 2D brain MRI data. One approach to enhance accuracy is to minimize the value Mean Squared Error (MSE). Research [12], which was influenced by the Resnet model, resulted in the SRResNet model which optimizes MSE values. This study resulted in a high Peak Signal-to-Noise Ratio (PSNR) value of 23.53db, surpassing the conventional interpolation method which yields only 21.59db. Another study [17] tried to reconstruct the image using Generative Adversarial Network (GAN) with Perceptual Loss, resulting SRGAN (Super Resolution Generative Adversarial Network) model. While the SRGAN model was able to reconstruct the image but PSNR values were lower than conventional SRResnet and Interpolation, which is 21.15db.

Digital Imaging and Communications in Medicine (DICOM) is a widely adopted data communication standard used by medical devices, particularly for Brain MRIs. While many researchers tend to use the NIFTI data format in their studies, especially [3]. In this study, the researcher will compare the reconstruction results between the interpolation, study [3], and the VAE method [25] using 3D shaped DICOM format dataset. The evaluation will be performed using Structural Similarity Index Measurement (SSIM) and PSNR metrics.

This research aims to evaluate and compare the quality of 3D magnetic resonance imaging medical images reconstructed using interpolation methods and artificial neural network architectures in the DICOM format. The test scenario is carried out by taking images per slice from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset and then comparing the output results of high-resolution medical images using several neural networks used in this case

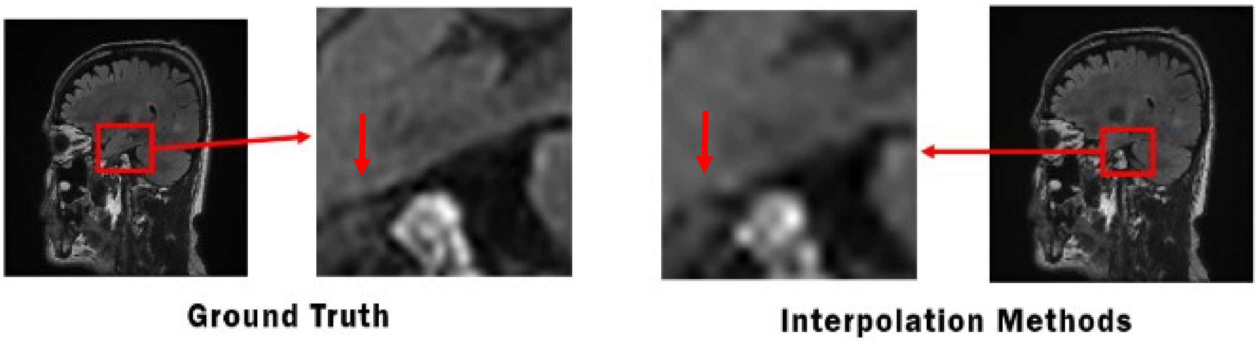

Figure 1. Comparison of images resulting from the reconstruction of interpolation and ground truth

variational autoencoder (VAE) and multi-level densely connected super-resolution network (mDCSRN) on 3D data with existing interpolation methods. Image evaluation metrics to be used such as SSIM and PSNR. If the evaluation result of the SSIM metric is close to 1.0, the image is similar to the original image, while for PSNR the higher the decibel value of PSNR, the better the image.

II. LITERATURE REVIEW

Recent advancements in machine learning have contributed to the development of successful image reconstruction techniques using low-resolution inputs to produce high-resolution images. [5] proposed a method that differs from previous approaches, [14] in which patch space modeling is done using knowledge obtained from dictionaries and manifolds. Instead, the proposed method obtains knowledge during the learning process via hidden layer. Optimization is done by extracting patches with a convolution technique called Super-Resolution Convolutional Neural Network (SRCNN). The SRCNN method is particularly advantageous as it has a simple neural network structure with only three layers, high accuracy compared to state of the art, fewer filters and layers. It is also fast in computation, even when only using a CPU, due to its full feed-forward design. However, the method does have limitations, including the fact that deeper layers in the method may not improve image quality, resulting in decreased contrast even though it has a high Structure Similarity Index Measure value, and is limited to 2D images. In another study, researchers proposed a new neural network model for 3D data, namely the multi-level densely connected super-resolution network (mDCSRN) model with training performed by the generative adversarial network (GAN) model. The resulting mDCSRN model produces fast image reconstruction and realistic image output that is almost indistinguishable from the original image. However, this study used the nifti dataset and used a 64 x 64 x 64 patch.

III. METHODOLOGY

A. *Super Resolution*

According to [3] Single Image Super Resolution (SISR) is a process of estimating in image to convert a low-resolution image Y into an actual X image with high resolution, the relationship between and can be seen in equation (1) as follows:

$$Y = f(X) \tag{1}$$

A previous study [3] utilized 2D models on 3D medical images by processing them slice by the slice or by combining results from different views. However, this approach fails to fully utilize the structural information present in the 3D form of medical images. 3D models, on the other hand, allow for a more thorough understanding of the data, for example by enabling the detection of small blood vessels bending in adjacent parts of the image, which might be difficult to discern using a 2D SR model.

B. *Variational Autoencoder*

In this section, the researcher will provide details of the VAE method [3]. VAE is consisted of three sections: the encoder, decoder, and Latent Space. The encoder is used to extract features from low-resolution (LR) and ground-truth images. The encoder's feature extraction results in a latent space, which is randomly sampled into mean and variance. Then reparameterization is applied, so that backpropagation can be used to estimate the features and reconstruct the initial image dimensions on the decoder layer.

$$q_{\emptyset}(z|Y) = \mathcal{N}\,(z; \mu_{\emptyset}(Y), \sigma_{\emptyset}(Y), I) \tag{1}$$

For the generative model pθ(x,z), we assume p(z) is assigned to the multivariate Gaussian units, ie p(z) = N(0,I). Assuming that these two nonlinear functions μθ and σθ map from z to the mean and standard deviation vector respectively, we get that

$$p_{\theta}(\hat{Y}|z) = \mathcal{N}\,(\hat{Y}; \mu_{\theta}(z), \sigma_{\theta}(z), I) \tag{2}$$

Looking at the network architecture of this model, we can see why this model is called an autoencoder.

$$Y \xrightarrow{q_{\emptyset}(z|Y)} z \xrightarrow{p_{\theta}(Y|z)} \hat{Y} \tag{3}$$

Input x is probabilistically mapped to code z by the encoder $q_{\emptyset}$ which is ultimately mapped probabilistically back to the input space by the decoder $p_{\emptyset}$. We take samples $z^{i}, l = 1 \dots . l$ from $q_{\emptyset}(z|Y)$ to reduce the variance of our gradient estimates, we apply a reparameterization trick to the multivariate Gaussian distribution $q_{\theta}(z|Y)$ using the noise distribution of the random epsilon the sample is taken from the Gaussian multivariate unit, i.e., $p(\epsilon) \sim N(0,1)$. we can write the reparameterization for qφ as follows :

$$z^{(l)} = \mu_{\emptyset}(Y) + \varepsilon \odot \sigma_{\emptyset}(Y) \tag{4}$$

Where:

$$\epsilon^{(l)} \sim N(0,1) \tag{5}$$

C. SSIM

Based on research [1], the SSIM is a method for evaluating the quality of reconstructed images. This metric was invented by [29] to measure the degradation in the quality of reconstructed images. SSIM can be calculated using formula (6) as follows:

$$SSI = \frac{(2\mu_S\mu_{SR} + C_1)(2\sigma_{SSR} + C_2)}{(\mu_S^2 + \mu_{SR}^2 + C_1) + (\sigma_S^2 + \sigma_{SR}^2 + C_2)} \quad (6)$$

The SSIM formula uses standard deviation (σ) values and the mean (μ) values of the reference and target images, along with constants C1 and C2, to stabilize the calculation. A SSIM value of 1 suggests that the reconstructed SR MRI image is identical to the HR MRI image.

D. PSNR

Research [1] indicates that the PSNR is used to calculate the decibel value of the maximum signal to noise ratio between the ground truth and the reconstructed images. A higher PSNR value indicates better image quality. The calculation of PSNR can be done using formula (7):

$$PSNR = -10log_{10}\frac{e_{mse}}{S^2} \quad (7)$$

E. Dataset

This study uses a dataset of public brain MRI data from ADNI, an organization that focuses on biomedical research, particularly related to the brain. This data is in 3D format with DICOM format, the brain MRI data includes x, y, and z values from each point in a 3D volume, also known as a voxel, a total of 92 datasets were used in this study and divided into several groups based on AD (Alzheimer's Disease) diagnosis, such as 6 men with age ranging from 55 to 57 years old, CN (Cognitively Normal) with 43 people, which includes 13 men and 30 women with age ranging from 55 to 88 years old, EMCI (Early Mild Cognitive Impairment) which consist of 54 men and 1 woman with age ranging from 73 to 89 years old, LMCI (Late Mild Cognitive Impairment) has 6 people consist of 6 men and 6 women with age ranging from 72-83 years, MCI (Mild Cognitive Impairment) has 27 people, 6 men, and 21 women, with age ranging from 55 to 88 years old, SMC (Significant Memory Concern) has 5 men, 13 battery people and 2 women with age ranging from 73 to 80 years old. Visualization of sample datasets are presented in Figure 2, 3 and 4.

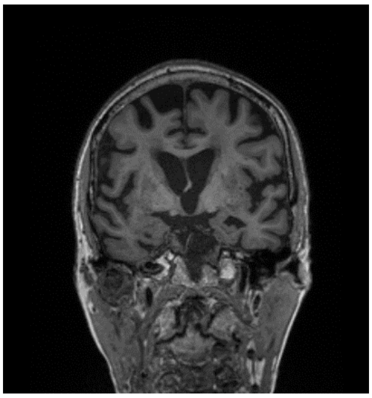

Figure 2. Axial Dataset ADNI

Figure 3. Coronal Dataset ADNI

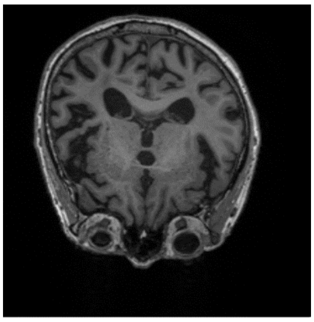

Figure 4. Coronal Dataset ADNI

IV. EXPERIMENTAL

This section will explain the methodology used in the research which consists of the stages of the research and the schedule of the research being conducted.

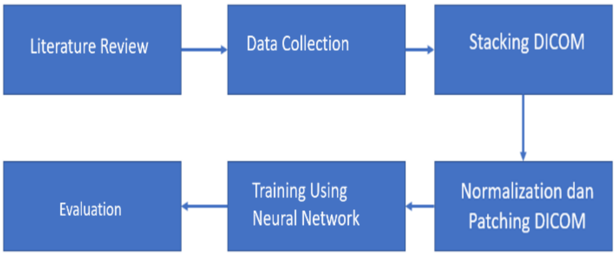

Figure 5. Research Flow

This section will explain the methodology used for relation to the research which consists of the stages of the research and the schedule of the research being conducted. Figure 5 is an overview of the stages of the research to be carried out. At the literature review stage, the researcher conducts a literature review process where the literature is used as a reference from journals, conferences, and related books. At the data collection stage, researchers explored public datasets that can be accessed, one of which is the public dataset originating from ADNI, an organization that focuses on biomedical development, especially the brain. Furthermore, the researcher carried out the data preprocessing stage, namely the process of preparing data before conducting research. After preprocessing the data, the training process is carried out using the neural network architecture and the metric evaluation calculation process is carried out using SSIM and PSNR

A. *Preprocessing*

The data preprocessing used in this study uses DICOM stacking, normalization, and patching. The DICOM stacking carried out in this study refers to previous research [10] where stacking must be in accordance with the slice sequence obtained in the DICOM metadata. The medical MRI image obtained is a slice-shaped image with a DICOM file extension. The data must be stacked to obtain a 3D image. Normalization in this study uses min-max normalization. Due to the limited memory used, only 16 GB RAM, the researcher performs the patching process on each 3D dataset that will be trained as done in previous research [26].

B. Training and Parameter

The experimental process will be carried out using Google Collab with a Tesla K80 GPU, with 16 GB of RAM. The training process will involve using low-resolution data from a patching dataset with dimensions of 256 x 256 x 256 which will be split into smaller patches of 32 x 32 x 32, resulting in 8 patches per dimension (x,y,z) with a total of 512 patches. A total of 93.184 datasets with dimensions of 32 x 32 x 32 will be obtained. The training set will be split into a 70:30 ratio, with 65.229 datasets for training and 27.955 datasets for testing. The down sampling and up sampling process will use the 3D Trilinear Interpolation method. The sigma value using is one. The scaling factor used in this experiment is 2 and the ground truth used is the original image from the dataset. This study uses 128 latent dimensions. Loss calculations in this research will be carried out using MSE and Reconstruction Loss where the loss value will be calculated when each image has been generated after batch size, then backpropagation will update each weight. For the optimizer at super-resolution which will be trained using the Adam optimizer using initial learning rate $\alpha_0 = 1e-3$ the total number of epochs (100 in this case).

C. Implementation and Evaluation

At this stage, the researcher will explain the results and tests carried out by taking images per slice from the dataset and then comparing them with the original images using several state-of-art methods resulting from the reconstruction, in this case, the interpolation method, mDCSRN, and VAE.

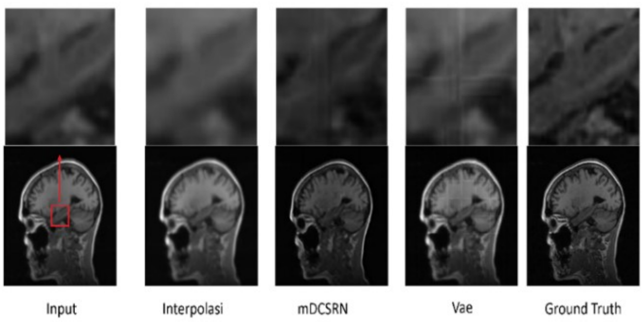

Figure 6. Comparison of Image Reconstruction Results

At this stage, the researcher will present the results and tests conducted by selecting images per slice from the dataset. The researcher compared the results of image reconstruction as shown in Figure 6, which depicts the output of SR images from various methods applied in this study. The inputs in this research are low-resolution images, as seen in the image. Low-resolution images are obtained by applying a sigma value to each patch, resulting in a blur effect. The neural networks then perform the process of reconstructing high-resolution images to produce a clear image. The results of the reconstruction reveal that the sigma value affects the pixel intensity value. It can be seen that the image reconstruction results from the VAE neural network are better than interpolation and mDCSRN methods. However, the reconstructed images still have some level of blur. Additionally, the VAE neural network generates images that have artifacts or shadow lines, unlike the other models.

The VAE and mDCSRN models both Improve image sharpness by reducing blur and allowing for clear visualization of brain details, as shown by the high PSNR values between the two models. However, the mDCSRN mode''s reconstruction results show a significant decline, as indicated by the low SSIM value, compared to the results of reconstructing local data. This decline results in the reconstructed image appearing darker.

TABLE I. TABLE OF IMAGE RECONSTRUCTION COMPARISON RESULTS BETWEEN METHODS USING SSIM AND PSNR EVALUATION METRICS

No	Methods	Sigma	Dataset ADNI	
			SSIM	PSNR
1	Interpolation	1	0.87	22.89
2	multi-level densely connected super-resolution network	1	0.75	19.20
3	Variational Autoencoder	1	0.88	23.19

Table 1 compares the outcomes of the interpolation method, mDCSRN , and VAE using the evaluation metrics of SSIM and PSNR. A high SSIM value of close to 1.0 indicates that the image is similar to the identical or original image, and a high PSNR value suggests a better-quality image.

Based on the table 1 the VAE method has the highest SSIM and PSNR values, indicating it has the highest image quality among the three methods. On the other hand, the MDCSRN method has the lowest SSIM and PSNR values, meaning it has the lowest image quality

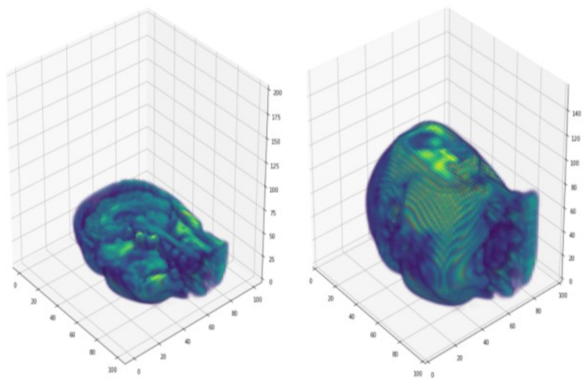

Figure 7. 3D Medical Image Visualization

Figure 7 (right) is a 3D visualization of the medical image when a 3D plot is performed using the python programming language with the matplotlib library. Figure 7 (left) is a visualization of the image of the inner organ of the brain. It can be seen that the inside of the brain has 3 main parts, namely the cerebrum, cerebellum, and brainstem. In addition to the inside of the brain, an image of the skull or skull bone is also formed. The skull is a set of bones that form the structure of the face and head while protecting the brain from impact.

V. CONCLUSION

From the research that has been carried out, it can be concluded that to reconstruct high-resolution images using neural network VAE models, mDCSRN and interpolation methods can be performed on the ADNI 3D dataset. Testing of reconstruction results is done using evaluation metrics. Evaluation metrics in this study were carried out using SSIM and PSNR. Then the results show that the VAE method has the highest SSIM and PSNR values, indicating it has the highest image quality among the three methods. On the other hand, the mDCSRN has the lowest SSIM and PSNR values, meaning it has the lowest image quality.

ACKNOWLEDGMENT

This study was financially supported by *Riset Inovatif-Produktif* (RISPRO) program under the Indonesia Endowment Funds for Education (LPDP); *Penelitian Terapan Unggulan Perguruan Tinggi* (PTUPT) program under the Indonesian Ministry of Education and Culture; Writing and IPR Incentive Program (PPHKI) under *Institut Teknologi Sepuluh November* (ITS).

REFERENCES

[1] Andrew, J., Mhatesh, T. S. R., Sebastin, R. D., Sagayam, K. M., Eunice, J., Pomplun, M., & Dang, H. (2021). Super-resolution reconstruction of brain magnetic resonance images via lightweight autoencoder. *Informatics in Medicine Unlocked*, *26*, 100713. https://doi.org/10.1016/j.imu.2021.100713

[2] Bahrami, K., Shi, F., Rekik, I., & Shen, D. (2016). Convolutional Neural Network for Reconstruction of 7T-like Images from 3T MRI Using Appearance and Anatomical Features (pp. 39–47). https://doi.org/10.1007/978-3-319-46976-8_5

[3] Chen, Y., Xie, Y., Zhou, Z., Shi, F., Christodoulou, A. G., & Li, D. (2018). *Brain MRI Super Resolution Using 3D Deep Densely Connected Neural Networks*. https://doi.org/10.1109/ISBI.2018.8363679

[4] Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

[5] Dong, C., Loy, C. C., He, K., & Tang, X. (2014a). *Learning a Deep Convolutional Network for Image Super-Resolution* (pp. 184–199). https://doi.org/10.1007/978-3-319-10593-2_13

[6] Dong, C., Loy, C. C., He, K., & Tang, X. (2014b). Image Super-Resolution Using Deep Convolutional Networks.

[7] Dong, C., Loy, C. C., & Tang, X. (2016a). Accelerating the Super-Resolution Convolutional Neural Network.

[8] Dong, C., Loy, C. C., & Tang, X. (2016b). Accelerating the Super-Resolution Convolutional Neural Network.

[9] Dosovitskiy, A., & Brox, T. (2016). Generating Images with Perceptual Similarity Metrics based on Deep Networks.

[10] Fajar, A., Sarno, R., Fatichah, C., & Fahmi, A. (2020). Reconstructing and resizing 3D images from DICOM files. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2020.12.004

[11] Goshtasby, A., Turner, D. A., & Ackerman, L. V. (1992). Matching of tomographic slices for interpolation. *IEEE Transactions on Medical Imaging*, *11*(4), 507–516. https://doi.org/10.1109/42.192686

[12] He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*.

[13] Jianchao Yang, Wright, J., Huang, T. S., & Yi Ma. (2010). Image Super-Resolution Via Sparse Representation. *IEEE Transactions on Image Processing*, *19*(11), 2861–2873. https://doi.org/10.1109/TIP.2010.2050625

[14] Jianchao Yang, Wright, J., Huang, T., & Yi Ma. (2008). Image super-resolution as sparse representation of raw image patches. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8. https://doi.org/10.1109/CVPR.2008.4587647

[15] Katsigiannis, S., Scovell, J., Ramzan, N., Janowski, L., Corriveau, P., Saad, M. A., & van Wallendael, G. (2018). Interpreting MOS scores, when can users see a difference? Understanding user experience differences for photo quality. *Quality and User Experience*, *3*(1), 6. https://doi.org/10.1007/s41233-018-0019-8

[16] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, *1*(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

[17] Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2016). *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*.

[18] Lim, B., Son, S., Kim, H., Nah, S., & Lee, K. M. (2017). Enhanced Deep Residual Networks for Single Image Super-Resolution. *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1132–1140. https://doi.org/10.1109/CVPRW.2017.151

[19] Lyu, Q., Shan, H., & Wang, G. (2020). MRI Super-Resolution With Ensemble Learning and Complementary Priors. *IEEE Transactions on Computational Imaging*, *6*, 615–624. https://doi.org/10.1109/TCI.2020.2964201

[20] Lyu, Q., You, C., Shan, H., & Wang, G. (2018). *Super-resolution MRI through Deep Learning*.

[21] Mane, V., Jadhav, S., & Lal, P. (2020). Image Super-Resolution for MRI Images using 3D Faster Super-Resolution Convolutional Neural Network architecture. *ITM Web of Conferences*, *32*, 03044. https://doi.org/10.1051/itmconf/20203203044

[22] Myronenko, A. (2018). 3D MRI brain tumor segmentation using autoencoder regularization.

[23] Park, S., Gach, H. M., Kim, S., Lee, S. J., & Motai, Y. (2021). Autoencoder-Inspired Convolutional Network-Based Super-Resolution Method in MRI. *IEEE Journal of Translational Engineering in Health and Medicine*, *9*, 1–13. https://doi.org/10.1109/JTEHM.2021.3076152

[24] Pihlgren, G. G., Sandin, F., & Liwicki, M. (2020). Improving Image Autoencoder Embeddings with Perceptual Loss.

[25] Pu, Y., Gan, Z., Henao, R., Yuan, X., Li, C., Stevens, A., & Carin, L. (2016). *Variational Autoencoder for Deep Learning of Images, Labels and Captions*.

[26] Sanchez, I., & Vilaplana, V. (2018). Brain MRI super-resolution using 3D generative adversarial networks.

[27] Shao, Z., Wang, L., Wang, Z., & Deng, J. (2019). Remote Sensing Image Super-Resolution Using Sparse Representation and Coupled Sparse Autoencoder. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(8), 2663–2674. https://doi.org/10.1109/JSTARS.2019.2925456

[28] Supriyanto, M. I. A., Fajar, A., Sarno, R., Fatichah, C., Fahmi, A., Utomo, S. A., & Notopuro, F. (2021). Slice Reconstruction on 3D Medical Image using Optical Flow Approach. *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, 242–246. https://doi.org/10.1109/APWiMob51111.2021.9435238

[29] 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. https://doi.org/10.1109/TIP.2003.819861

[30] Wu, Y., & He, K. (2018). *Group Normalization*.

[31] Xue, X., Wang, Y., Li, J., Jiao, Z., Ren, Z., & Gao, X. (2020). Progressive Sub-Band Residual-Learning Network for MR Image Super Resolution. *IEEE Journal of Biomedical and Health Informatics*, *24*(2), 377–386. https://doi.org/10.1109/JBHI.2019.2945373