

Evaluating the impact of downsampling on 3D MRI images segmentation results based on similarity metrics

Aziz Fajar¹, Riyanarto Sarno², Chastine Fatichah²

¹Data Science Technology, Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Surabaya, Indonesia

²Department of Informatics, Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Article Info

Article history:

Received Dec 29, 2023

Revised Apr 16, 2024

Accepted May 6, 2024

Keywords:

3D image segmentation

3D MRI

Deep learning

Image downsampling

Similarity measurement

ABSTRACT

Medical imaging plays a crucial role in diagnosing patient conditions, with magnetic resonance imaging (MRI) standing as a significant modality for numerous years. However, leveraging convolutional neural network (CNN) architectures like U-Net and its variations for anatomical segmentation demands considerable memory, particularly when working with full 3D image sets. Therefore, downsampling 3D MRIs proves advantageous in reducing memory consumption. Nevertheless, downsampling leads to a reduction in voxel count, potentially impacting the performance of commonly used segmentation metrics. The jaccard similarity index (JSI), dice similarity coefficient (DSC), and structural similarity index (SSIM) are extensively employed in image segmentation contexts. Hence, this study employs all three metrics to assess downsampled images and evaluate the robustness of the metrics when used to evaluate the downsampled 3D MRI images. The results show that JSI and DSC are more robust than SSIM when handling the downsampled data.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Riyanarto Sarno

Department of Informatics, Faculty of Intelligent Electrical and Informatics Technology

Institut Teknologi Sepuluh Nopember

Sukolilo, 60111 Surabaya, Indonesia

Email: riyanarto@if.its.ac.id

1. INTRODUCTION

Advancements in medical imaging have revolutionized diagnostic procedures, with 3D magnetic resonance imaging (MRI) becoming an indispensable tool for precise anatomical visualization. However, the sheer volume and resolution of MRI data present computational challenges, often necessitating downsampling for efficient storage and processing. Downsampling involves reducing the spatial resolution of images, yet its impact on the accuracy of subsequent analyses remains a critical concern. In this study, we delve into the repercussions of downsampling on the assessment of 3D MRI images, specifically focusing on the evaluation metrics of jaccard similarity index (JSI) [1]-[3], Sorensen dice similarity coefficient (DSC) [4]-[8], and structural similarity index (SSIM) [9], [10]. These metrics, widely employed in quantifying similarity between images, serve as the foundation for assessing the effects of downsampling on image analysis. A study performed by previous study [11] shows that JSI and DSC can be used as a metric to measure blood vessel segmentation. Another study performed [12] also proved that JSI and DSC to gauge machine learning segmentation performance. Other studies only uses DSC as the metric to measure the segmentation results of generalized medical image [13]. While SSIM are used in other studies to measure the segmentation results [9], [14].

To predict brain MRI image segments, this research employs convolutional neural network (CNN) architectures, namely U-Net, Res U-Net, and Dense U-Net. These architectures and its modifications are chosen for their proficiency in semantic segmentation tasks, aiming to replicate the delineation established by medical experts [15]-[18]. All the architecture have been used to segment biomedical images. U-Net have been used in several studies both as a base model [19] or on its own [20]-[22]. While, Dense U-Net [23] and Res U-Net [24] as modified architecture of U-Net are also used in various segmentation cases. The CNN-based predictions are then compared with ground truth segments curated by experienced doctors. These comparisons are performed at three distinct downsampling levels: 100%, 75%, and 50% of the original image resolution. It is important to note that the datasets utilized in this study are sourced from medical institutions, namely Dr. Soetomo and National Hospital in Surabaya, Indonesia. These datasets comprise a set of MRI images in the form of slices, capturing a spectrum of brain anatomical variations. Since the dataset we got is in the form of 2D images, we need to reconstruct each MRI image into 3D MRI image using our previous study [25]. The primary objective of this research are as follows: i). Testing the robustness of evaluation metrics on different downsampling level. ii). Testing the evaluation metrics on different segmentation results. This study investigated the effects of downsampling on similarity metrics. While earlier studies have explored the impact of the use of similarity metrics on segmentation results, they have not explicitly addressed its influence on the downsampled image. In the subsequent sections, we outline the methodology employed, including the utilization of U-Net, Res U-Net, and Dense U-Net architectures, the datasets obtained from Dr. Soetomo and National Hospital in Surabaya, Indonesia, and the specific experimental setups designed to do a comparison between CNN-predicted segments and ground truth across varying downsampling levels. This study are organized as follows: Introduction, Method, Result and Discussion, and Conclusion.

2. METHOD

In this section we will explain our approach to segment MRI images acquired from the hospitals in Surabaya, Indonesia. Then, the trained models are used to predict image segments. The segmentation results are then used to test similarity metrics robustness to downsampled images. Figure 1 shows how this study performed.

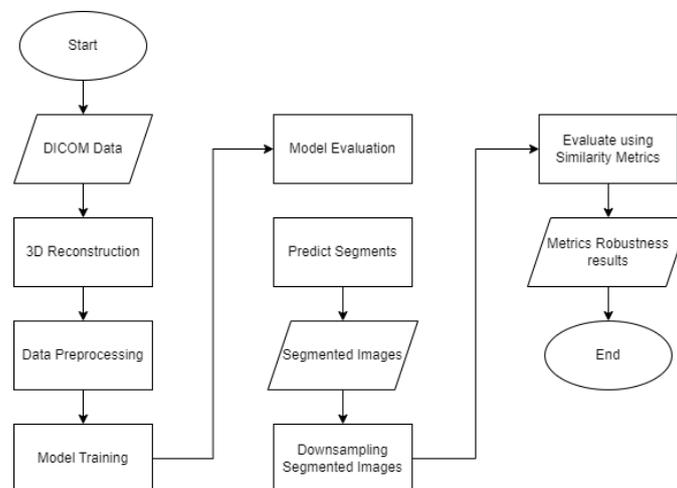


Figure 1. System diagram

Based on Figure 1, this study aims to test the robustness of the similarity metrics, namely, JSI, Sørensen DSC, and SSIM. The metrics are tested on the downsampled ground truth and segmentation results of CNN models. This is the contribution of this study compared to those of previous studies.

2.1. Data annotation

To create the ground truth, we created a python program to annotate the 3D image. The doctors from both hospitals in Surabaya annotate each slice of the 3D image until all images are annotated, then the doctor will annotate the next image until all 50 images are annotated. Figure 2 shows the annotation result.

Figure 2(a) shows a slice of the reconstructed 3D image of an MRI sequence, while Figure 2(b) shows the annotation done by doctors of the National Hospital Surabaya. The segments annotated are Cerebrum, Cerebellum, Brain stem, Corpus Callosum, and Ventricle. We augment the ground truth along the original data. Thus, we got the same orientation with the original image. To increase the number of data, we used data augmentation technique to the original 3D data and the ground truth. We used random flip and random rotate to increase the numbers of data. The total data we got after the data augmentation process are 200 data.



Figure 2. The annotation result (a) original image and (b) annotation result

2.2. Data preprocessing

In performing brain anatomy segmentation on 3D images, we utilized 3D images reconstructed from DICOM data obtained from hospitals in Surabaya using method used in previous study [25]. The reconstruction process comprises of histogram equalization, 3D reconstruction, trilinear interpolation, and 3D image resizing. The reconstructed data has a resolution of 240 x 240 x 240. The total data we got from both hospitals are 50 MRI images. These 3D images will be processed using the Python and TensorFlow library. For the hardware we used, we employed an NVIDIA QUADRO 8000 GPU with GPU memory of 48 GB DDR6. The MRI images are augmented to increase the amount of data in the dataset using random flip and random rotate. The final dataset got 200 3D MRI images. Then, each data is divided into 6 images based on the annotated anatomy along with the background. This is done to make each image a binary image. Figure 3 shows the binary images produced.

Based on Figure 3, the images are divided into 6 binary images. Figure 3(a) shows the original image. Figures 3(b) background, (c) cerebrum, (d) cerebellum, (e) corpus callosum, (f) brain stem, and (g) ventricle. This is needed since we used softmax activation function on the final layer of each of the CNN architectures.

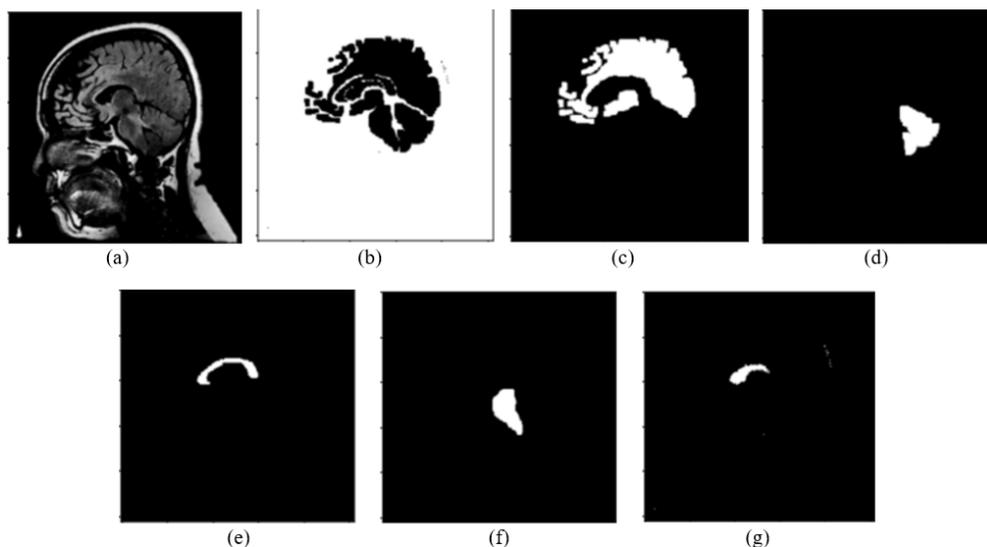


Figure 3. U-Net architecture; (a) shows the original image, (b) background, (c) cerebrum, (d) cerebellum, (e) corpus callosum, (f) brain stem, and (g) ventricle

2.3. CNN architectures

The architectures we used in this study are U-Net, Dense U-Net, and Res U-Net. We used DSC as the loss function since previous studies generally used DSC compared to other metrics. For the optimizer, this study used Adam with a learning rate of 0.001. All layers used in this study are 3D convolution layers and 3D Maxpooling layers, since the data used in this study are 3D MRI images. The GPU used in this study is NVIDIA QUADRO 8000 GPU with GPU memory of 48 GB. For the loss function of the CNN architectures, this study used 1-DSC to get the loss value.

2.3.1. U-Net

The U-Net is one of the deep learning architectures commonly used for image segmentation [20]-[22]. In this research, the U-Net architecture will be employed for segmenting brain images using 3D image data created using the proposed methodology. Since the data used is 3D images, the layers used in U-Net architecture are 3D convolutional and 3D pooling layers. The U-Net architecture consists of downsampling blocks, a bridge, and upsampling blocks. Figure 4 shows the U-Net architecture.

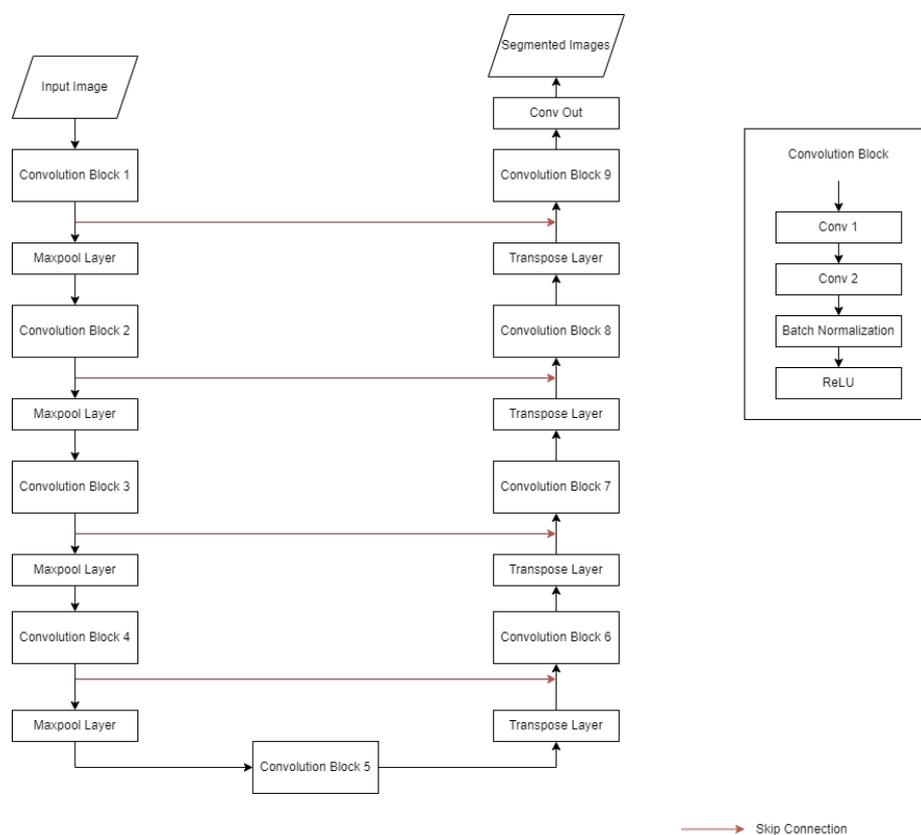


Figure 4. U-Net architecture

As shown in Figure 4, U-Net architecture comprised of 3 components. Downsampling blocks made of double convolution layer and maxpooling layer, upsampling blocks comprised of a transpose layer and 2 convolution blocks, and bridge comprised of 2 convolution layers. Each convolution layer in the upsampling block will be concatenated with the result of the downsampling layer with the same resolution. Thus, doubling the number of channels before acting as the input of the next convolution layer of the upsampling block.

2.3.2. Res U-Net

The residual U-Net has a similar architecture to the U-Net [24]. However, Res U-Net uses addition after its second convolutional layer of each convolutional block, creating a residual block. Figure 5 illustrates the residual block.

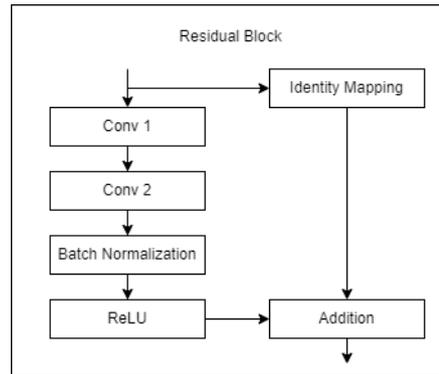


Figure 5. Residual block

Figure 5 shows the residual block used by the Res U-net. By adding an addition layer, the input needs to be identity mapped. this process ensures the number of channel of the input image is the same with the number of channel produced by the activation layer of the residual block.

2.3.3. Dense U-Net

Dense U-Net uses dense block which is a convolution block that utilizes the concatenation of the input image of each of its convolution layers [23]. Figure 6 shows the dense block used by the Dense U-Net architecture. The Dense U-Net architecture used dense block shown in Figure 6. this block demands significantly more computational power than the blocks used by previous 2 architectures. This is due to the concatenation done after each convolutional layer. However, this concatenation helps in retaining the input image characteristics, theoretically.

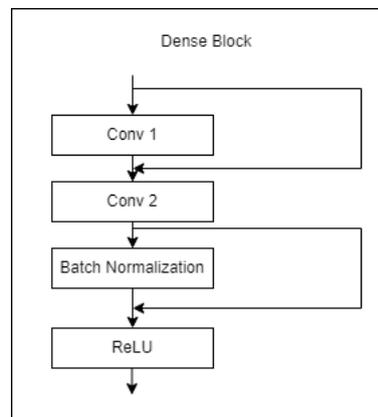


Figure 6. Dense block

2.4. 3D MRI downsampling

The downsampling process typically involves mapping each voxel in the output image to a corresponding voxel in the input image. For each voxel in the output image, I_{out} , find the nearest neighbor in the input image I_{in} using Euclidean distance. The nearest neighbor calculation can be represented as shown in (1).

$$I_{out}[i, j, k] = \underset{(x,y,z) \in I_{in}}{\operatorname{argmin}} \sqrt{(i-x)^2 + (j-y)^2 + (k-z)^2} \quad (1)$$

Where, $I_{out}[i, j, k]$ = represents the voxel at position (i, j, k) in the output image $x, y =$ voxels in I_{in} argmin = voxel in the input image that minimizes the Euclidean distance from the current voxel in the output image.

In (1) describes the process of finding the nearest neighbor in the input image for each voxel in the output image using Euclidean distance. This method essentially maps each voxel in the output image to the closest voxel in the input image based on Euclidean distance.

2.5. Segmentation result evaluation

2.5.1. JSI

JSI, also called intersection over union is one of the most used similarity indexes as shown in previous study [1]-[3]. In (2) shows the calculation of JSI.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (2)$$

Where, J = Jaccard similarity index value A, B = Images being measured. Based on (2), JSI considers the total membership of voxels of both sets compared to the denominator.

2.5.2. Sørensen DSC

DSCs behave in similar ways to JSI. DSC compares the ratio of intersection to the total of both sets. In (3) shows the calculation of DSC.

$$DSC(A, B) = \frac{2 \times |A \cap B|}{|A| + |B|} \quad (3)$$

As shows in (3), the difference between JSI and DSC is that the denominator of the DSC is the sum of cardinalities, not the union of the sets as those of JSI.

2.5.3. SSIM

The SSIM is a widely used metric in image processing for quantifying the similarity between two images. SSIM evaluates three key components: luminance, contrast, and structure. Luminance represents the brightness of the image, while contrast measures the difference in intensity. Structure reflects the spatial arrangement of image features. SSIM compares corresponding image patches and computes a similarity score for each component, yielding an overall index that ranges from -1 to 1, with 1 indicating identical images. In (4) shows how to calculate SSIM.

$$SSIM(x, y) = (2\mu_x\mu_y + c^1) (2\sigma_{xy} + c^2) / ((\mu_x^2 + \mu_y^2 + c^1) (\sigma_x^2 + \sigma_y^2 + c^2)) \quad (4)$$

Where, x, y = the compared images. μ_x, μ_y = the means of x and y , respectively. σ_x^2, σ_y^2 = the variances of x and y , respectively. σ_{xy} = is the covariance between x and y . c^1, c^2 = constants added to stabilize the division when the denominator is close to zero. In (4) yields an SSIM value between -1 and 1, where 1 denotes identical images and values closer to 1 indicate higher similarity between the images. Adjusting the constants fine-tunes the formula.

3. RESULTS AND DISCUSSION

This section focuses on the results of the study. This section shows the results of training process and the downsampling results. Then, the results are presented and discussed to provide insights.

3.1. Data preprocessing

To produce the segments of the brain anatomy, this study trained the CNN models with the architectures mentioned. Since the image obtained from the hospitals is in the form of 2D images, the images need to be reconstructed to 3D images using our previous study and resize the image into 240 x 240 x 240 for the resolution.

3.2. Training results

The model training processes produces loss graphs. Each of the architecture produce a loss graph. The loss graphs of all of the architecture are shown in Figure 7. Figure 7 shows the loss graphs of all the architectures with different training length to achieve convergence. the result of U-Net as shown in Figure 7(a) shows that the architecture is not the fastest to converge. Figure 7(b) shows that Res U-Net is the fastest to reach convergence, while Dense U-Net as shown in Figure 7(c) needs much longer to achieve it. This is might due to the size of the architecture that employs dense blocks. Thus, harder to train and much slower. Res U-Net being the fastest might be due to the efficient use of concatenation in the architecture. The training time of each of the architecture are shown in Table 1.

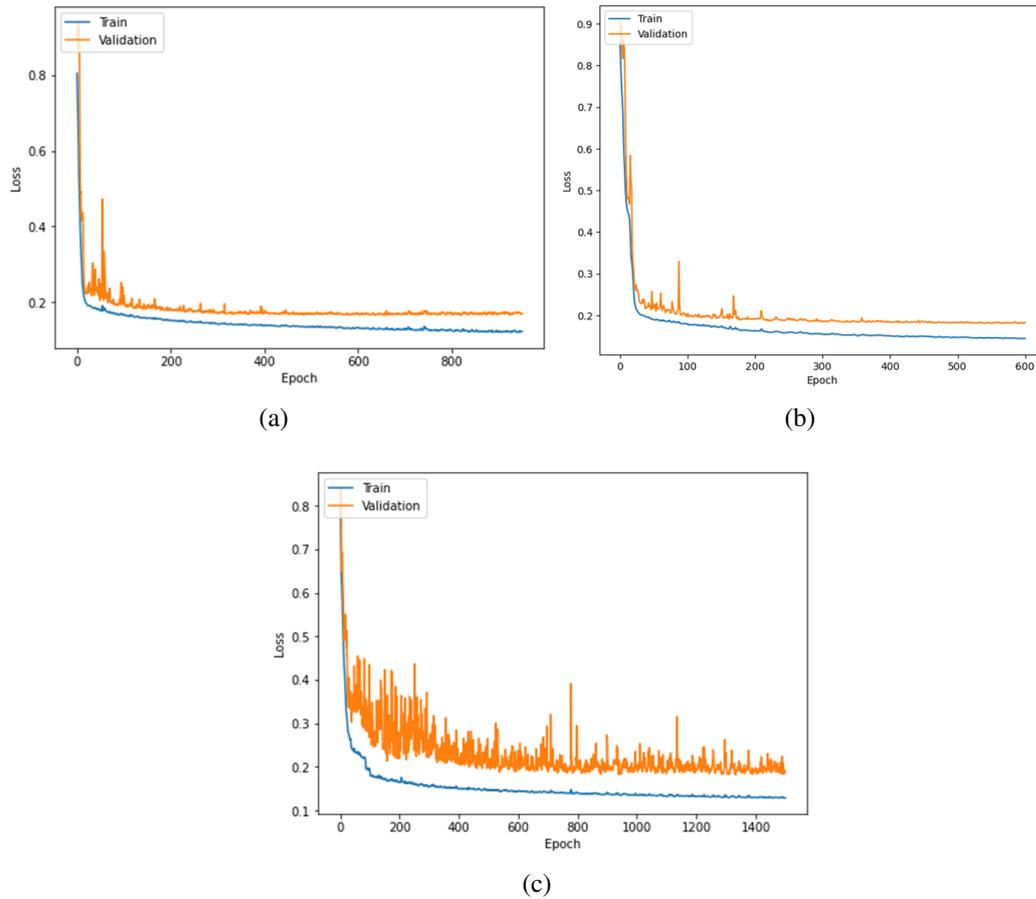


Figure 7. Loss graph of; (a) U-Net, (b) Res U-Net, and (c) Dense U-net

As shown in Table 1, Res U-Net is significantly faster compared to the other architecture. Res U-Net almost twice as fast as U-Net and almost 4 times faster than Dense U-Net. The loss value produced by each architecture during the training process are shown in Table 2.

Based on the results of the training process as shown in Table 2, we can conclude that U-Net produce the lowest loss value both in the training and validation set. Then, followed by Dense U-Net and Res U-Net. This might indicate that this dataset does not need to retain the information from the input image through concatenation. This is due to the results of the architectures utilizing concatenation results in higher loss value.

Table 1. Length of model training of each architecture

Architecture	Epoch to convergence
U-Net	836
Res U-Net	485
Dense U-Net	1721

Table 2. Length of model training of each architecture

Architecture	Loss value	
	Training	Validation
U-Net	0.1242	0.1661
Res U-Net	0.1337	0.1819
Dense U-Net	0.1264	0.1813

3.3. Downsampling results

Since we already trained the model, we could predict the segments using the trained model. Since we trained all the architectures to segment each of the anatomy using softmax on the final layer, the model produces 6 images, namely, Cerebrum, Cerebellum, Brain Stem, Corpus Callosum, and Ventricle. Figures 8-10 shows some of the predicted MRI slices using different architectures.

As shown in Figure 9 and Figure 10, the downsampled images loses its resolution and made the images more pixelated compared to the full resolution shown in Figure 8. This is due to the reduced number of voxels. The ground truth shown in (a) Figure 8, in are compared to the segmentation result of (b) U-Net, (c) Res U-Net, and (d) Dense U-Net in different voxel counts as described in Figures 8(a)-(d), Figures 9(a)-(d), and Figures 10(a)-(d). The original image has $240 \times 240 \times 240$ resolution. However, for the images with the number of voxels reduced to 75% of the original images, the resolution is $180 \times 180 \times 180$. while, images which are 50% of the original images have resolution of $120 \times 120 \times 120$.

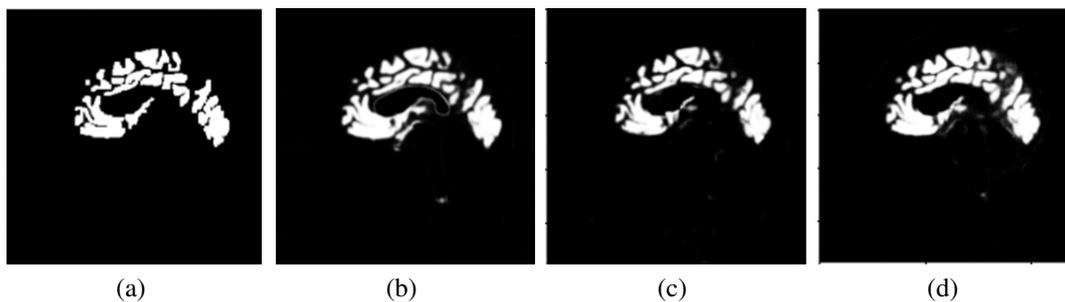


Figure 8. Full resolution of; (a) ground truth, (b) U-Net, and (c) Res U-net, and (d) Dense U-net

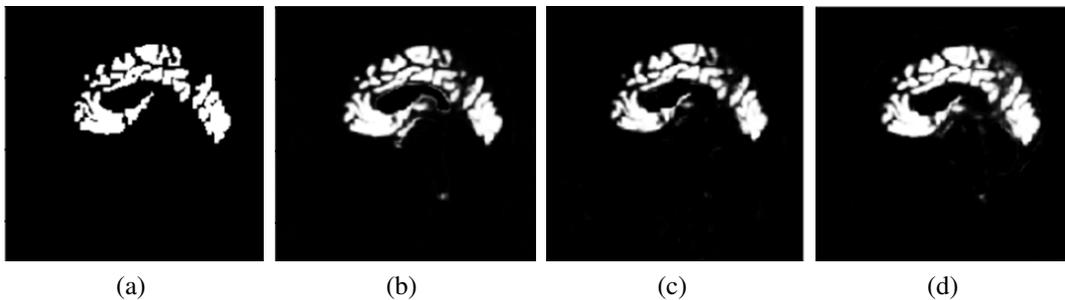


Figure 9. Downsampled; (a) ground truth, (b) U-Net, and (c) Res U-net, and (d) Dense U-net to 75% of original resolution

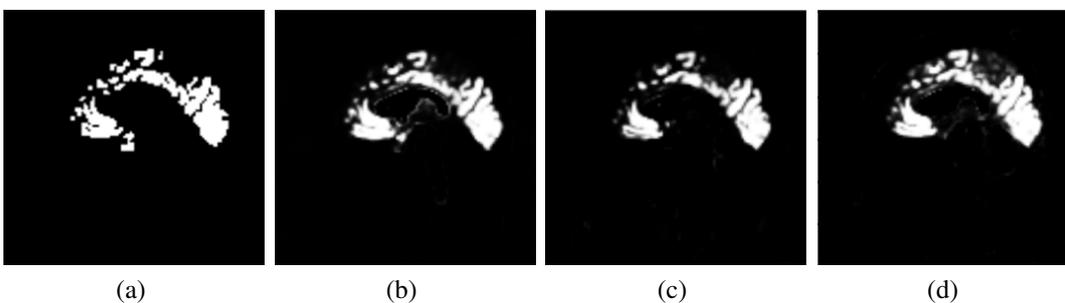


Figure 10. Downsampled; (a) ground truth, (b) U-Net, and (c) Res U-net, and (d) Dense U-net to 50% of original resolution

3.4. Metrics comparison

The full image and the downsampled images can be calculated for its metrics by comparing the predicted images to the ground truth made by doctors. Since in this study we used JSI, DSC, and SSIM as the similarity metrics, the predicted images with both full resolution and the downsampled images will be compared to those of the ground truth with similar resolutions. Table 3 shows how the metric perform on full and downsampled images

Table 3. Results of evaluation of similarity metrics on full and downsampled images

Architecture	Resolution	JSI	DSC	SSIM
U-Net	100%	0.497	0.854	0.887
	75%	0.497	0.855	0.881
	50%	0.498	0.856	0.887
Res U-Net	100%	0.498	0.842	0.886
	75%	0.498	0.842	0.879
	50%	0.499	0.843	0.860
Dense U-Net	100%	0.498	0.846	0.887
	75%	0.498	0.846	0.880
	50%	0.498	0.846	0.860

Our experimentation involved downsampling images to 75% and 50% of their original resolution. Based on Table 3, the downsampling process exhibited minimal alteration in the computed similarity metrics, indicating robustness to downsampling. The JSI values remained largely consistent, showcasing high image fidelity despite resolution reduction. Similarly, the DSC, used notably in image segmentation, showcased stability even at reduced resolutions. Comparatively, the SSIM, even if not by much, exhibits higher inconsistency in its measurements when contrasted with the DSC and JSI in this study. These findings imply that downsampling images to 75% and 50% of their original resolution does not significantly impact the computed similarity metrics. Therefore, while achieving computational efficiency and reduced memory footprint, downsampling maintains fidelity in similarity assessments, underscoring its viability in various image processing applications.

4. CONCLUSION

In this study we used several CNN architectures, namely, U-Net, Res U-Net, and Dense U-Net to produce predicted 3D MRI images. Since this study focuses on testing the robustness of the metrics such as JSI, DSC, and SSIM, the results shows that downsampling the images and the ground truth to 75% and 50% for images produced by the same architecture does not affect the similarity metrics results by much. However, when compared to JSI and DSC, SSIM exhibits higher inconsistency. This is due to SSIM which use not only spatial similarity like the other metrics, since SSIM also consider luminance and contrast of the compared images. Thus, using JSI and DSC on downsampled image might be a better options. Future studies may explore the effects of other similarity metrics or modify available metric to be more resilient on downsampled images.

ACKNOWLEDGEMENTS

This work was funded by Institut Teknologi Sepuluh Nopember (ITS) under Penelitian Dana Departemen Program and Penelitian Keilmuan Program.

REFERENCES

- [1] İ. Ataş, "Performance evaluation of jaccard-dice coefficient on building segmentation from high resolution satellite images," *Balkan Journal of Electrical and Computer Engineering*, vol. 11, no. 1, pp. 100–106, Jan. 2023, doi: 10.17694/bajece.1212563.
- [2] Y. Xu et al., "Boundary restored network for subpleural pulmonary lesion segmentation on ultrasound images at local and global scales," *Journal of Digital Imaging*, vol. 33, no. 5, pp. 1155–1166, Oct. 2020, doi: 10.1007/s10278-020-00356-8.
- [3] M. George and A. H. Bhimasena, "Entropy based segmentation model for Kidney Stone and Cyst on ultrasound image," *International Journal of Computing*, pp. 450–455, Dec. 2022, doi: 10.47839/ijc.21.4.2780.
- [4] A. Fajar, R. Sarno, C. Fatchah, R. I. Susilo, and G. Pangestu, "Cyclical learning rate optimization on deep learning model for brain tumor segmentation," *IEEE Access*, vol. 11, pp. 119802–119810, 2023, doi: 10.1109/ACCESS.2023.3326475.
- [5] L. Atikah, N. A. Hasanah, R. Sarno, A. Fajar, and D. Rahmawati, "Brain segmentation using adaptive thresholding, k-means clustering and mathematical morphology in MRI Data," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Sep. 2020, pp. 161–167, doi: 10.1109/iSemantic50169.2020.9234303.

- [6] C. Chen, Y. Huang, P. Fang, C. Liang, and R. Chang, "A computer-aided diagnosis system for differentiation and delineation of malignant regions on whole-slide prostate histopathology image using spatial statistics and multidimensional DenseNet," *Medical Physics*, vol. 47, no. 3, pp. 1021–1033, Mar. 2020, doi: 10.1002/mp.13964.
- [7] H.-Y. Hsu *et al.*, "Automatic segmentation of retinal fluid and photoreceptor layer from optical coherence tomography images of diabetic macular edema patients using deep learning and associations with visual acuity," *Biomedicine*, vol. 10, no. 6, p. 1269, May 2022, doi: 10.3390/biomedicine10061269.
- [8] S. Klein, U. A. van der Heide, I. M. Lips, M. van Vulpen, M. Staring, and J. P. W. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Medical Physics*, vol. 35, no. 4, pp. 1407–1417, Apr. 2008, doi: 10.1118/1.2842076.
- [9] M. I. A. Supriyanto *et al.*, "Slice reconstruction on 3D medical image using optical flow approach," in *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Apr. 2021, pp. 242–246, doi: 10.1109/APWiMob51111.2021.9435238.
- [10] P. Kancharla and S. S. Channappayya, "Improving the visual quality of generative adversarial network (GAN)-generated images using the multi-scale structural similarity index," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Oct. 2018, pp. 3908–3912, doi: 10.1109/ICIP.2018.8451296.
- [11] D. N. H. Thanh, D. Sergey, V. B. S. Prasath, and N. H. Hai, "Blood vessels segmentation method for retinal fundus images based on adaptive principal curvature and image derivative operators," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2/W12, pp. 211–218, May 2019, doi: 10.5194/isprs-archives-XLII-2-W12-211-2019.
- [12] T. S. R. Priya and A. Manickavasagan, "Evaluation of segmentation methods for RGB colour image-based detection of Fusarium infection in corn grains using support vector machine (SVM) and pre-trained convolution neural network (CNN)," *Canadian Biosystems Engineering*, vol. 64, no. 1, pp. 7.09-7.20, Dec. 2022, doi: 10.7451/CBE.2022.64.7.9.
- [13] S. Ahmed and M. K. Hasan, "Twin-SegNet: dynamically coupled complementary segmentation networks for generalized medical image segmentation," *Computer Vision and Image Understanding*, vol. 240, p. 103910, Mar. 2024, doi: 10.1016/j.cviu.2023.103910.
- [14] M. Xu, L. Cao, D. Lu, Z. Hu, and Y. Yue, "Application of swarm intelligence optimization algorithms in image processing: a comprehensive review of analysis, synthesis, and optimization," *Biomimetics*, vol. 8, no. 2, p. 235, Jun. 2023, doi: 10.3390/biomimetics8020235.
- [15] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, Oct. 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.
- [16] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-Net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749–753, May 2018, doi: 10.1109/LGRS.2018.2802944.
- [17] M. Kolařík, R. Burget, V. Uher, K. Říha, and M. Dutta, "Optimized high resolution 3D Dense-U-Net network for brain and spine segmentation," *Applied Sciences*, vol. 9, no. 3, p. 404, Jan. 2019, doi: 10.3390/app9030404.
- [18] B. Wu, Y. Fang, and X. Lai, "Left ventricle automatic segmentation in cardiac MRI using a combined CNN and U-net approach (DOUBLE)," *Computerized Medical Imaging and Graphics*, vol. 82, p. 101719, Jun. 2020, doi: 10.1016/j.compmedimag.2020.101719.
- [19] Y. Huo *et al.*, "3D whole brain segmentation using spatially localized atlas network tiles," *NeuroImage*, vol. 194, pp. 105–119, Jul. 2019, doi: 10.1016/j.neuroimage.2019.03.041.
- [20] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9901 LNCS, Springer, 2016, pp. 424–432.
- [21] H.-J. Bae *et al.*, "Fully automated 3D segmentation and separation of multiple cervical vertebrae in CT images using a 2D convolutional neural network," *Computer Methods and Programs in Biomedicine*, vol. 184, p. 105119, Feb. 2020, doi: 10.1016/j.cmpb.2019.105119.
- [22] N. Reamaroon *et al.*, "Robust segmentation of lung in chest x-ray: applications in analysis of acute respiratory distress syndrome," *BMC Medical Imaging*, vol. 20, no. 1, p. 116, Dec. 2020, doi: 10.1186/s12880-020-00514-y.
- [23] N. Alalwan, A. Abozeid, A. A. ElHabshy, and A. Alzahrani, "Efficient 3D deep learning model for medical image semantic segmentation," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 1231–1239, Feb. 2021, doi: 10.1016/j.aej.2020.10.046.
- [24] T. D. Bui, J. Shin, and T. Moon, "Skip-connected 3D DenseNet for volumetric infant brain MRI segmentation," *Biomedical Signal Processing and Control*, vol. 54, p. 101613, Sep. 2019, doi: 10.1016/j.bspc.2019.101613.
- [25] A. Fajar, R. Sarno, C. Faticah, and A. Fahmi, "Reconstructing and resizing 3D images from DICOM files," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3517–3526, Jun. 2022, doi: 10.1016/j.jksuci.2020.12.004.

BIOGRAPHIES OF AUTHORS



Aziz Fajar    was born in Malang, Jawa Timur, Indonesia. He is a doctorate student in Institut Teknologi Sepuluh Nopember Surabaya focusing on computer vision, especially medical image processing. He is working on projects with neurosurgeon to process MRI, CT scan, and MRA images to help neurosurgeons treat Parkinson diseases and brain tumor since 2019. Up to now, the author has several publications on the subjects in the form of journals and conferences using both machine learning and CNN methods. He is working in Department of Data Science Technology of Airlangga University as a lecturer since 2023. He can be contacted at email: azizfajar519@gmail.com.



Riyanarto Sarno    received his Ph.D. degree in 1992. He is currently a Professor in the Informatics Department of Institut Teknologi Sepuluh Nopember (ITS). His research projects involve machine learning, deep learning, generative AI, the Internet of Things, and knowledge engineering, which have been implemented in electronic noses, medical equipment, and health care systems in recent years. He wrote more than eight books and more than 573 scientific articles, which led him to the top 2% world ranking scientist in 2020-2023, by Stanford University. He can be contacted at email: riyanarto@if.its.ac.id.



Chastine Fatichah    was born in Pasuruan, December 1975. in year 2000, She received the bachelor's degree in informatics. Then, she continue her study for master's degree in computer science and graduated in 2008. Then, she got her Ph.D. degree in computational intelligence in 2012. currently, she is a Professor in the Department of Informatics of Institut Teknologi Sepuluh Nopember (ITS). She also got some awards in her name in 2014, such as Young Researcher Award from Advanced Computational Intelligence and Intelligent Informatics (JACIII) by Fuji Press in Tokyo, Japan. She can be contacted at email: chastine@if.its.ac.id.