

New Approach to Image Segmentation: U-Net Convolutional Network for Multiresolution CT Image Lung Segmentation

Sugiyarto Surono ^{1*}, Muhammad Rivaldi ¹, Deshinta Arrova Dewi ²,
Nursyiva Irsalinda ¹

¹ Department of Mathematics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia.

² Faculty of Data Science and Information Technology, INTI International University, Nilai, Malaysia.

Abstract

Image processing is the main topic of discussion in the field of computer vision technology. With the increase in the number of images used over time, the types of images with different resolution qualities are becoming more diverse. Low image resolution leads to uncertainty in the task of image processing. Therefore, a method with high performance is needed for image processing. In image processing, there is a Convolutional Neural Networks (CNN) architecture for semantic segmentation of pixels called U-Net. U-Net is formed by an encoder network and decoder network that will later produce segmented images. In this paper, researchers applied the U-Net architecture to the lung CT image dataset, which has different resolutions in each image, to segment the image that produces a segmented lung image. In this study, we conducted experiments for many training and testing data ratios while also comparing the model performances between the single resolution dataset and the multiresolution dataset. The results showed that the segmentation accuracy using a single resolution dataset is as follows: 5 to 5 ratio is 66.00%, 8 to 2 ratio is 88.96%, and 9 to 1 ratio is 94.47%. For the multiresolution dataset, the application is: 5 to 5 ratio is 82.42%, 8 to 2 ratio is 90.12%, and 9 to 1 ratio is 93.66%. And for the result, the training time using single resolution dataset are: 5 to 5 ratio is 59.94 seconds, 8 to 2 ratio is 87.16 seconds, and 9 to 1 ratio is 195.34 seconds, as for multiresolution data application are: 5 to 5 ratio is 49.60 seconds, 8 to 2 ratio is 102.08 seconds, and 9 to 1 ratio is 199.79 seconds. Based on those results, we obtained the best accuracy for single resolution at a 9:1 ratio and the best training time for multiresolution at a 5:5 ratio.

Keywords:

U-Net;
Image Segmentation;
Multiresolution Image;
Convolutional Neural Networks (CNN);
Resolution Image.

Article History:

Received:	25	August	2022
Revised:	02	December	2022
Accepted:	19	December	2022
Available online:	22	February	2023

1- Introduction

Computer vision is often involved when discussing image processing. In the process, the deep learning method is one that is often used in image processing [1]. The Deep Learning method has proven its success in its application in many fields [2]. The term "deep learning" or "deep neural network" refers to a multi-layer Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) one of the most popular deep neural network methods recently [3]. CNN has made tremendous progress, especially in image processing and vision-related tasks [4]. This will be very useful for technological developments in the field of artificial intelligence, which until now has become a trend, especially when discussing image processing. Image processing itself has many benefits that can be applied to needs, such as classification, clustering, prediction, and so on.

Semantic segmentation is capable of providing categorical information at the pixel level. Many real-world applications benefit from this task, such as self-driving vehicles, pedestrian detection, disability detection, therapy planning, and computer-assisted diagnosis. Semantic segmentation assigns a category label to each image pixel. Pixel-level semantic information helps intelligent systems understand spatial positions or make important judgments [5]. In the medical field, deep CNN is often used to find solutions to biomedical segmentation problems [6]. Recently, CNN

* **CONTACT:** sugiyarto@math.uad.ac.id

DOI: <http://dx.doi.org/10.28991/ESJ-2023-07-02-014>

© 2023 by the authors. Licensee ESJ, Italy. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<https://creativecommons.org/licenses/by/4.0/>).

has developed, and one of the developments is the segmentation model [7]. One of the CNN architectures used for image segmentation is the U-Net Convolutional Network [8]. This architecture was introduced in 2015 and has rapidly evolved over the years. U-Net is often used in the medical field for semantic segmentation purposes [9]. U-Net is based on the encoder-decoder architecture, whereby the encoder-decoder-based deep learning method is very effective in dealing with various problems in artificial intelligence applications [10]. U-Net incorporates a high-level semantic feature map from a low-level feature map decoder on the encoder by using a skip connection [7].

Research on image processing in the biomedical field is a rapidly emerging area that includes biomedical signal acquisition, image generation from signals, image processing, and image display for medical diagnosis. Clinical imaging devices combine hardware and software. The number of medical imaging sensors is of interest to researchers in the field of biomedical image processing. The number of medical imaging sensors is of interest to researchers in the field of biomedical image processing. Medical image classification is a sub-topic of image classification. Of the various types of cancer, lung cancer is the deadliest cause of death. This is because patients with non-small cell lung cancer (NSCLC) are diagnosed at an advanced stage. The death rate due to lung cancer even reached 18.4% of the total deaths caused by cancer worldwide, and this was calculated in 2019 [11]. Therefore, lung segmentation is important because it has a good impact on determining the lung area on CT scan images, which can be useful in treating medical problems related to the lungs.

2- Related Work and Dataset

2-1- Related Work

Identifying objects is an easy thing for humans. However, for computers, it is something complex [12]. The development of technology causes the amount of data available to increase. Nowadays, digital image processing is used in various application domains. Due to the inherent drawbacks associated with digital cameras and their image quality, there is great scope for developing techniques to improve image quality [13]. Thus, the number of images with more diverse types may increase [14]. As a result, a qualified image processing model will be needed so that it can be used for image processing with high performance. With the increasing number of image data taken by different optical sensors that produce images with various resolutions, a good multi-resolution image processing model will have a good impact on computer vision [15]. This will make the image processing model more flexible because it can be applied to various image resolution datasets.

Image segmentation is one of the important tasks when discussing computer vision or images [16]. Image segmentation can be applied for many tasks, such as medical image processing [17], scene understanding [18], autonomous driving cars [19], and augmented reality [20]. Image segmentation separates the semantic entities in the image by defining the boundaries between those semantic entities [21]. In this paper, we propose U-Net for image segmentation. We would like to apply the proposed method for the segmentation of multi-resolution images of lungs, which was inspired by Soomro et al. [22], to apply image segmentation techniques using multi-resolution image datasets. What will be discussed in this paper is the performance of the proposed method in performing image segmentation tasks using accuracy and computational time to examine how well the model performs lung segmentation. In the study by Soomro et al. (2019) that inspired us to do this research, we did segmentation using retinal images to segment retinal vessels [22]. The data used is an image that has a uniform resolution. Soomro et al. (2019) suggested applying the segmentation model to data that has different resolutions/multi-resolution [22]. Researchers want to do research related to image segmentation, whose dataset is a multi-resolution image. In this study, we want to segment the lungs on CT scan images. Due to the different types of data, namely [22] using the retinal dataset and our study using the lung CT scan dataset, what we did in this study was compare the performance of the model when applied to a multi-resolution and uniform-resolution image dataset.

2-2- Dataset

The dataset that we use to implement the proposed method is the NSCLC-Radiomics dataset from Zheng et al. (2020) [23]. The NSCLC-Radiomics dataset contains 422 CT scan images of Non-Small Cell Lung Cancer (NSCLC) patients (Figure 1), the mask images for the lung area, and its metadata. All images in the dataset are DICOM-format images with an image size of 512×512. For training, we used CT scan images and masking.

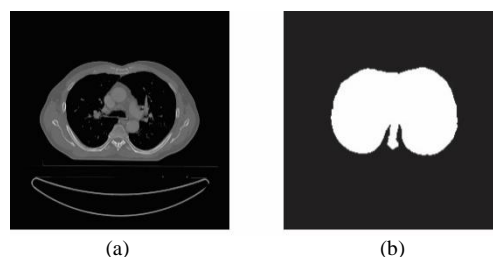


Figure 1. Preview CT scan and mask images sample: a) Sample of a CT scan image, b) The image mask of image

3- Method

3-1- Min Max Scaler

Min-max scaler is a normalization technique that adjusts all data numbers into probability numbers, i.e., 0 to 1, with min and max data [24]. The Min-Max scaler is denoted as Equation 1:

$$x_{scaled} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

3-2- Convolutional Neural Network

Convolutional Neural Network (CNN) performance has grown rapidly in the field of computer vision. CNN is formed by several layers of neural computing connections whose systematic processing is minimal [25]. Being a method that plays an important role, CNN is used to perform tasks related to image processing, such as image recognition, image segmentation, and object detection [26]. The first thing needed to understand CNN is convolution [27]. The step of classification for the U-Net convolutional network is described by the workflow in Figure 2. In this paper, we want to implement the U-Net Convolutional Network architecture on the various resolution image dataset to segment the lung area.

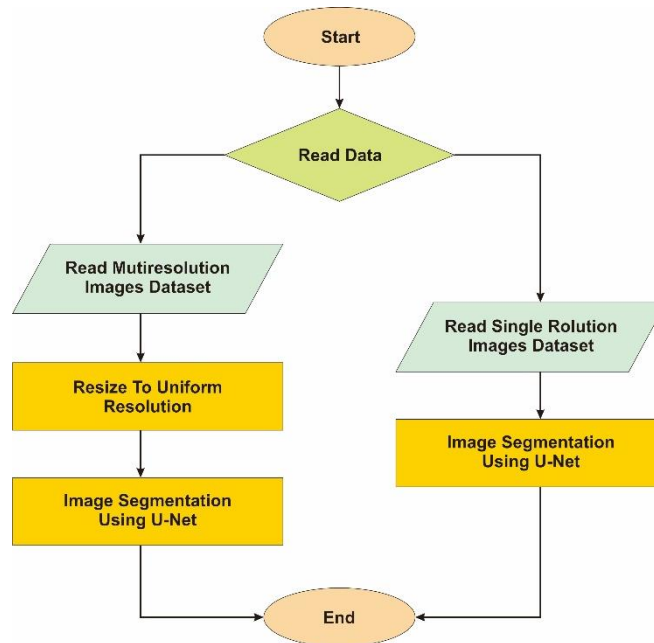


Figure 2. Purposed Research Flowchart

3-2-1- Convolutional Layer

Convolution is the process of applying a filter matrix or also known as a kernel to an image to reduce the size of the image, or it is also used to add several layers of padding to keep the size the same. Convolution is also used to perform feature extraction on the image [28]. Define V as a convolution function at positions x , and y in image I as follows:

$$V(x, y) = (I * K)(x, y) = \sum_{i=1}^n \sum_{j=1}^n I(x + i, y + j) K(i, j) \quad (2)$$

where m, n is the size of the kernel matrix K [29]. The kernel is one of the main components of CNN. The kernel is a square matrix of dimension $nk \times nk$, where nk is an integer and is usually a small number. The kernel is used to do sharpening, blurring, and so on [27].

3-2-2- ReLU Activation Function

The Rectified Linear Units (ReLU) activation function will convert a negative value to zero. However, it will make a value other than negative be fixed [30]. The ReLU activation function is defined as:

$$f(x) = \begin{cases} x & , \text{if } x \geq 0 \\ 0 & , \text{if } x < 0 \end{cases} \quad (3)$$

3-2-3- Dice Loss Function

The Dice coefficient is often used in Computer Vision work to calculate the similarity between two images which is then adapted into an error function called Dice loss [31]. Dice loss is denoted as:

$$D_{koeff} = \frac{2 \sum_{l=1}^n y_l t_l}{\sum_{l=1}^n y_l^2 + \sum_{l=1}^n t_l^2} \quad (4)$$

$$L_D = 1 - D_{koeff} \quad (5)$$

where $p1(x)$ is the probability that pixel x belongs to class l and $gl(x)$ is a vector of truth labels [32].

3-3- U-Net

Along with the development of deep neural networks, the performance of semantic segmentation also increases. Many researchers have made focused efforts to overcome the limitations of this field [33]. U-Net is one of the semantic segmentation architectures based on a fully convolutional network [34]. The architecture U-Net is shown in Figure 3 and the architecture is demonstrated in Tables 1 and 2.

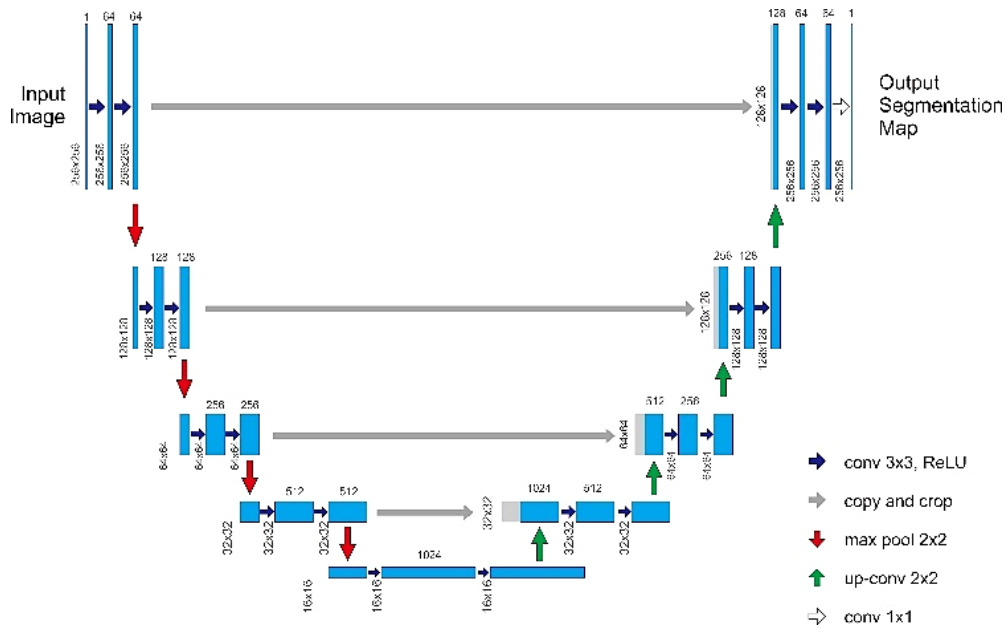


Figure 3. U-Net Architecture

Table 1. U-Net encoder layer

Layer	Ukuran Output	Layer U-Net
Convolution	$256 \times 256 \times 16$	conv2d
Dropout	$256 \times 256 \times 16$	dropout
Convolution	$256 \times 256 \times 16$	conv2d
Pooling	$128 \times 128 \times 16$	max pooling2d
Convolution	$128 \times 128 \times 32$	conv2d
Dropout	$128 \times 128 \times 32$	dropout
Convolution	$128 \times 128 \times 32$	conv2d
Pooling	$64 \times 64 \times 32$	max pooling2d
Convolution	$64 \times 64 \times 64$	conv2d
Dropout	$64 \times 64 \times 64$	dropout
Convolution	$64 \times 64 \times 64$	conv2d
Pooling	$32 \times 32 \times 64$	max pooling2d
Convolution	$32 \times 32 \times 128$	conv2d
Dropout	$32 \times 32 \times 128$	dropout
Convolution	$32 \times 32 \times 128$	conv2d
Pooling	$16 \times 16 \times 128$	max pooling2d
Convolution	$16 \times 16 \times 256$	conv2d
Dropout	$16 \times 16 \times 256$	dropout
Convolution	$16 \times 16 \times 256$	conv2d

Table 2. U-Net decoder layer

Layer	Ukuran Output	Layer U-Net
Deconvolution	$32 \times 32 \times 128$	conv2d transpose
Concatenate	$32 \times 32 \times 256$	concatenate
Convolution	$32 \times 32 \times 128$	conv2d
Dropout	$32 \times 32 \times 128$	dropout
Convolution	$32 \times 32 \times 128$	conv2d
Deconvolution	$64 \times 64 \times 64$	conv2d transpose
Concatenate	$64 \times 64 \times 128$	concatenate
Convolution	$64 \times 64 \times 64$	conv2d
Dropout	$64 \times 64 \times 64$	dropout
Convolution	$64 \times 64 \times 64$	conv2d
Deconvolution	$128 \times 128 \times 32$	conv2d transpose
Concatenate	$128 \times 128 \times 64$	concatenate
Convolution	$128 \times 128 \times 32$	conv2d
Dropout	$128 \times 128 \times 32$	dropout
Convolution	$128 \times 128 \times 32$	conv2d
Deconvolution	$256 \times 256 \times 16$	conv2d transpose
Concatenate	$256 \times 256 \times 16$	concatenate
Convolution	$256 \times 256 \times 32$	conv2d
Dropout	$256 \times 256 \times 16$	dropout
Convolution	$256 \times 256 \times 16$	conv2d
Convolution	$256 \times 256 \times 1$	conv2d

4- Experimental and Result

4-1 Experimental Setup

To implement the proposed method, we used Python Jupyter Notebook with 2-core Intel(R) Xeon(R) @ 2.20GHz CPU specifications, and 32GB RAM of the computer.

4-2- Training Process

Firstly, we convert all the image data we use into various pixel resolutions randomly, from 50×50 to 600×600 using the reshape tool. This process is to create conditions where the dataset that we use becomes the same as the dataset proposed in this study, which is the multiresolution dataset. We used CT images from 421 patients. For each patient, we used 3 CT images, so the total image data used in this study was 1263. We trained the data set to segment lung regions using the U-Net architecture for the encoder network using the convolutional network and the decoder using the deconvolution network. The output after implementing the U-Net architecture is an image with segmented regions. The color of the pixels in the lung region remains fixed and will be black for the rest of the pixels. Our experiment uses training and testing ratios, those are 5 to 5, 8 to 2, and 9 to 1 to compare the accuracy of two types of data sets, namely multi-resolution image data sets and single-resolution image data sets.

4-3- Result

Measurement is important to know the performance of one's model. We used the dice coefficient for the testing accuracy metric and also the training time. For the testing, performances are shown in Table 3, which shows the accuracy and computation time with three different training ratios. Figure 4 shows the change in model accuracy at each epoch of the training and validation process. As for the segmentation results, table 4 shows the predicted mask image and the cropped lung area.

Table 3. Proposed model performances

Dataset type	Ratio	Accuracy	Time
Single Resolution	5:5	66.00 %	59.94 seconds
	8:2	88.96 %	87.16 seconds
	9:1	94.47 %	195.34 seconds
Multiresolution	5:5	82.42 %	49.60 seconds
	8:2	90.12 %	102.08 seconds
	9:1	93.66 %	199.79 seconds

Table 3 is the result of the simulation to get the best model shown in Figure 4 where the ratio resolution training was already determined before.

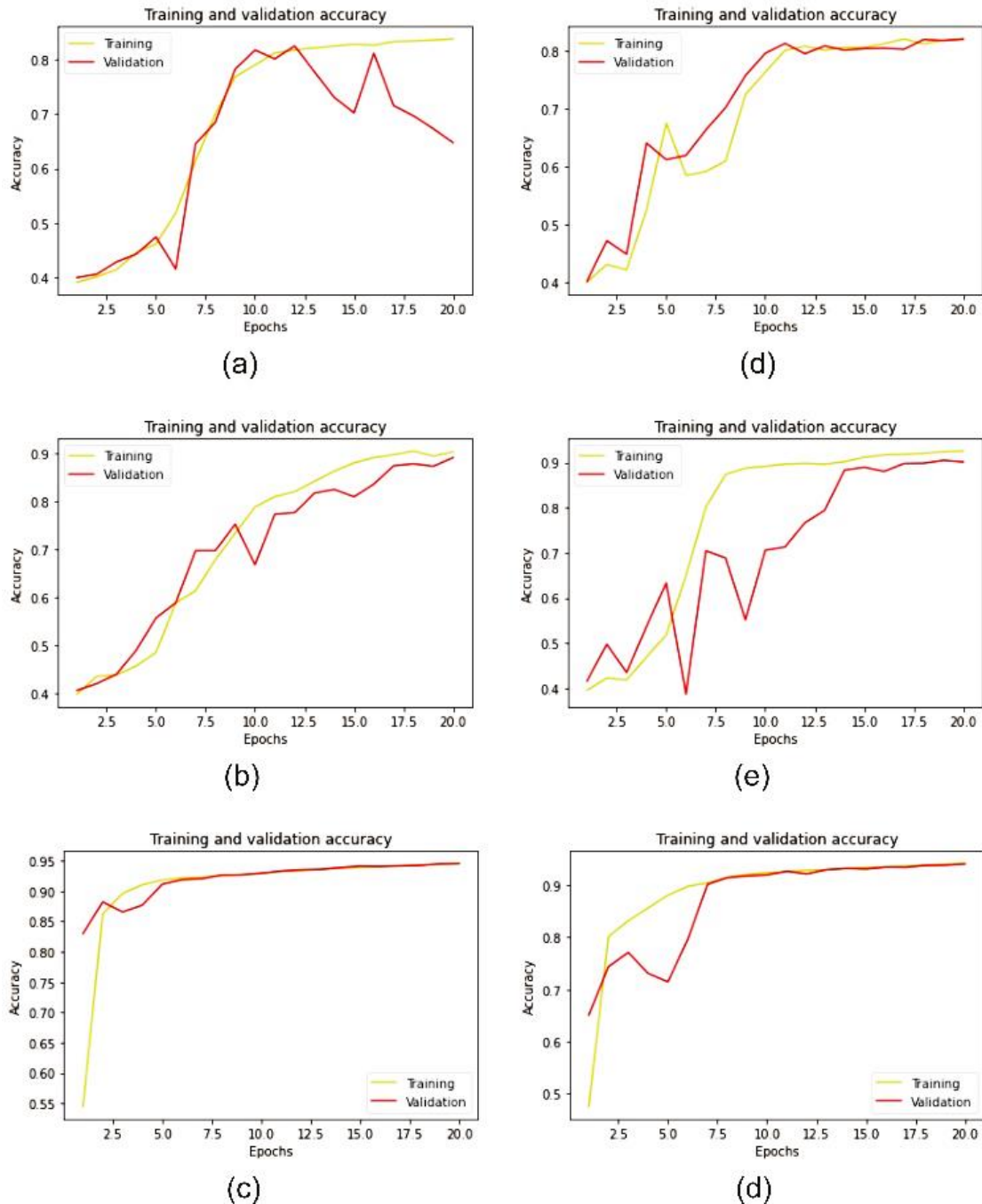



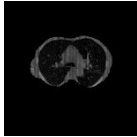
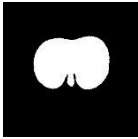



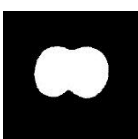
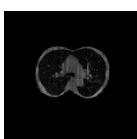
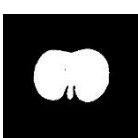
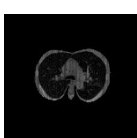


Figure 4. a) 5:5, b) 8:2, c) 9:1 ratio single resolution training performance, and d) 5:5, e) 8:2, f) 9:1 ratio multiresolution training performance

Table 4. Segmentation result

Dataset type	Ratio	Predicted mask	Segmented lung
Single Resolution	5:5		
	8:2		
	9:1		
Multiresolution	5:5		
	8:2		
	9:1		

5- Conclusion

In this study, we applied the U-Net network architecture to segment the lung areas on CT scan images of the lungs. From the experiments, we can see that the lung segmentation accuracy has the same pattern: the higher the training ratio, the higher the accuracy, but the longer the computation time. In the study of Soomro et al. [22], the resulting accuracy was above 90%. In this research, the highest accuracy is generated by a single-resolution dataset is using training: testing ratio of 9 to 1 with an accuracy of 94.47%, and the highest accuracy is generated by a single-resolution dataset is using training: testing ratio of 9 to 1 with an accuracy of 93.66%. And the fastest computation time is generated by a multi-resolution dataset with a computation time of 49.6 seconds. In terms of accuracy, our proposed model is able to do image segmentation quite well; the accuracy is above 90%. This indicates that our proposed model has successfully segmented well even when segmenting on a multi-resolution image dataset.

In the future, the authors want to conduct further experiments using other models to compare the performance of these other models. In addition, the authors also want to apply this segmentation model to segment other objects in multiresolution image datasets, or they can also use other loss functions such as IoU. It aims to see the performance of the model in segmenting with various architectures and other loss functions. In addition, it will also have a good influence in the field of medical image processing to be applied to medical treatment.

6- Declarations

6-1- Author Contributions

Conceptualization, S.S. and M.R.; methodology, S.S.; software, N.I.; validation, D.A.D., S.S., and N.I.; formal analysis, S.S.; investigation, N.I.; resources, S.S.; data curation, S.S., and M.R.; writing—original draft preparation, N.I.; writing—review and editing, S.S., and D.A.D.; visualization, M.R.; supervision, S.S.; project administration, N.I.; funding acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

6-2- Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6-3- Funding

This study was funded by the Ministry of Education and Culture for the fiscal year 2022 with the contract number: 157/E5/PG.02.00.PT/2022,1989.10/LL5-INT/PG.02.00/2022, 029/PB.PDUPT/BRIn.LPPM/VI/2022

6-4- Acknowledgements

The authors are grateful to the Ministry of Education and Culture for funding the PUDPT research for fiscal year 2022 with contract number 029/PB.PDUPT/BRIn.LPPM/VI/2022.

6-5- Institutional Review Board Statement

Not applicable

6-6- Informed Consent Statement

Informed consent is not applicable to this as the authors used a dataset from Zheng et al. (2020) [23].

6-7- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

- [1] Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, 7068349, 1-14. doi:10.1155/2018/7068349.
- [2] Soydaner, D. (2020). A Comparison of Optimization Algorithms for Deep Learning. *International Journal of Pattern Recognition and Artificial Intelligence*, 34(13), 1–26. doi:10.1142/S0218001420520138.
- [3] Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2018). Understanding of a convolutional neural network. *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017*, 1–6. doi:10.1109/ICEngTechnol.2017.8308186.
- [4] Khan, A., Sohail, A., Zahoor, U., & Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 53(8), 5455–5516. doi:10.1007/s10462-020-09825-6.
- [5] Hao, S., Zhou, Y., & Guo, Y. (2020). A Brief Survey on Semantic Segmentation with Deep Learning. *Neurocomputing*, 406, 302–321. doi:10.1016/j.neucom.2019.11.118.
- [6] Salehi, S.S.M., Erdogmus, D., Gholipour, A. (2017). Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks. *Machine Learning in Medical Imaging. MLMI 2017, Lecture Notes in Computer Science*, 10541, Springer, Cham, Switzerland. doi:10.1007/978-3-319-67389-9_44.
- [7] Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., & Wu, J. (2020). UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi:10.1109/icassp40776.2020.9053405.
- [8] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015, Lecture Notes in Computer Science*, 9351. Springer, Cham, Switzerland. doi:10.1007/978-3-319-24574-4_28.
- [9] Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., Wasserthal, J., Koehler, G., Norajitra, T., Wirkert, S., & Maier-Hein, K. H. (2019). Abstract: nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. *Bildverarbeitung Für Die Medizin 2019*, 22–22, Springer Vieweg, Wiesbaden, Germany. doi:10.1007/978-3-658-25326-4_7.
- [10] Yalçın, S., & Vural, H. (2022). Brain stroke classification and segmentation using encoder-decoder based deep convolutional neural networks. *Computers in Biology and Medicine*, 105941, 149. doi:10.1016/j.combiomed.2022.105941.
- [11] Nair, R. R., & Singh, T. (2019). Multi-sensor medical image fusion using pyramid-based DWT: A multi-resolution approach. *IET Image Processing*, 13(9), 1447–1459. doi:10.1049/iet-ipr.2018.6556.
- [12] Sultana, F., Sufian, A., & Dutta, P. (2020). Evolution of Image Segmentation using Deep Convolutional Neural Network: A Survey. *Knowledge-Based Systems*, 201–202. doi:10.1016/j.knosys.2020.106062.
- [13] Singh, S., Mittal, N., & Singh, H. (2022). Multifocus Image Fusion Based on Multiresolution Pyramid and Bilateral Filter. *IETE Journal of Research*, 68(4), 2476–2487. doi:10.1080/03772063.2019.1711205.
- [14] Peryanto, A., Yudhana, A., & Umar, R. (2020). Klasifikasi Citra Menggunakan Convolutional Neural Network DAN K Fold Cross Validation. *Journal of Applied Informatics and Computing*, 4(1), 45–51. doi:10.30871/jaic.v4i1.2017.

- [15] Duarte, D., Nex, F., Kerle, N., & Vosselman, G. (2018). Multi-resolution feature fusion for image classification of building damages with convolutional neural networks. *Remote Sensing*, 10(10). doi:10.3390/rs10101636.
- [16] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2022). Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3523–3542. doi:10.1109/TPAMI.2021.3059968.
- [17] Hesamian, M. H., Jia, W., He, X., & Kennedy, P. (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *Journal of Digital Imaging*, 32(4), 582–596. doi:10.1007/s10278-019-00227-x.
- [18] Lyu, H., Fu, H., Hu, X., & Liu, L. (2019). Esnet: Edge-Based Segmentation Network for Real-Time Semantic Segmentation in Traffic Scenes. *IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan. doi:10.1109/icip.2019.8803132.
- [19] Liu, Y., Chu, L., Chen, G., Wu, Z., Chen, Z., Lai, B., & Hao, Y. (2021). Paddleseg: A high-efficient development toolkit for image segmentation. *arXiv preprint, arXiv:2101.06175*. doi:10.48550/arXiv.2101.06175.
- [20] Ko, T. Y., & Lee, S. H. (2020). Novel method of semantic segmentation applicable to augmented reality. *Sensors (Switzerland)*, 20(6), 1737. doi:10.3390/s20061737.
- [21] Ghosh, S., Das, N., Das, I., & Maulik, U. (2019). Understanding deep learning techniques for image segmentation. *ACM Computing Surveys*, 52(4), 1-35. doi:10.1145/3329784.
- [22] Soomro, T. A., Afifi, A. J., Gao, J., Hellwich, O., Zheng, L., & Paul, M. (2019). Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation. *Expert Systems with Applications*, 134, 36–52. doi:10.1016/j.eswa.2019.05.029.
- [23] Zheng, M., Li, T., Zhu, R., Tang, Y., Tang, M., Lin, L., & Ma, Z. (2020). Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification. *Information Sciences*, 512, 1009–1023. doi:10.1016/j.ins.2019.10.014.
- [24] Phaladisailoed, T., & Numnonda, T. (2018). Machine Learning Models Comparison for Bitcoin Price Prediction. *2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE)*. doi:10.1109/icitied.2018.8534911.
- [25] Ting, F. F., Tan, Y. J., & Sim, K. S. (2019). Convolutional neural network improvement for breast cancer classification. *Expert Systems with Applications*, 120, 103–115. doi:10.1016/j.eswa.2018.11.008.
- [26] Yao, P., Wu, H., Gao, B., Tang, J., Zhang, Q., Zhang, W., Yang, J. J., & Qian, H. (2020). Fully hardware-implemented memristor convolutional neural network. *Nature*, 577(7792), 641–646. doi:10.1038/s41586-020-1942-4.
- [27] Michelucci, U. (2019). *Advanced applied deep learning: Convolutional neural networks and object detection*. Springer, Berkeley, United States. doi:10.1007/978-1-4842-4976-5.
- [28] Singh, H. (2019). *Practical Machine Learning and Image Processing*. Springer, Berkeley, United States. doi:10.1007/978-1-4842-4149-3.
- [29] Charniak, E. (2018). *Intro to Deep Learning*. MIT Press, Cambridge, Massachusetts, United States.
- [30] Wang, Y., Li, Y., Song, Y., & Rong, X. (2020). The influence of the activation function in a convolution neural network model of facial expression recognition. *Applied Sciences (Switzerland)*, 10(5), 1897. doi:10.3390/app10051897.
- [31] Jadon, S. (2020). A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. doi:10.1109/cibcb48159.2020.9277638.
- [32] Soomro, T. A., Afifi, A. J., Gao, J., Hellwich, O., Paul, M., & Zheng, L. (2018). Strided U-Net Model: Retinal Vessels Segmentation using Dice Loss. *IEEE Digital Image Computing: Techniques and Applications (DICTA)*, Canberra, ACT, Australia. doi:10.1109/dicta.2018.8615770.
- [33] Zheng, S., Lu, J., Zhao, H., Zhu, X., Luo, Z., Wang, Y., Fu, Y., Feng, J., Xiang, T., Torr, P. H. S., & Zhang, L. (2021). Rethinking Semantic Segmentation from a Sequence-to-Sequence Perspective with Transformers. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. doi:10.1109/cvpr46437.2021.00681.
- [34] Xiao, Z., Liu, B., Geng, L., Zhang, F., & Liu, Y. (2020). Segmentation of Lung Nodules Using Improved 3D-UNet Neural Network. *Symmetry*, 12(11), 1787. doi:10.3390/sym12111787.