Semantic Segmentation of Venous on Deep Vein Thrombosis (DVT) Case using UNet-ResNet

by Arta Kusuma Hernanda

Semantic Segmentation of Venous on Deep Vein Thrombosis (DVT) Case using UNet-ResNet

1st Arta Kusuma Hernanda
Department of Electrical Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
6022211010@student.its.ac.id

2nd I Ketut Eddy Purnama
Department of Electrical Engineering
Department of Computer Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
ketut@te.its.ac.id

3rd Eko Mulyanto Yuniarno
Department of Electrical Engineering
Department of Computer Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
ekomulyanto@ee.its.ac.id

4th Johanes Nugroho
Departement of Cardiology
Airlangga University
Surabaya, Indonesia
j.nugroho.eko@fk.unair.c.id

***Abstract*—Deep Vein Thrombosis (DVT) is caused by an abnormal condition of blood clots in the network of blood vessels. No accurate profile data has been found on the number of common DVT cases in Indonesia. Several studies were conducted in several hospitals but with small sample sizes. In common cases, the diagnosis of DVT is made using Doppler Ultrasonography to monitor the condition of blood flow through the veins. This study uses the UNet-ResNet Deep Learning model to semantically segment the venous area on a 2D ultrasound image. The segmentation model is built from the pre-trained model UNet with the encoder ResNet-34. The dataset is taken from phantoms, a human body parts simulation tool. Ultrasound image acquisition on the Phantom will use Ultrasound Telemed SmartUs EXT-1M, which is directly connected to a PC. The segmentation model from the training process was evaluated with the Intersection-over-Union score (IoU) and Dice Loss. The result of the IoU evaluation on the standard UNet model resulted in an IoU score of 81.22% and an assessment of the dice loss of 0.1341. The UNet segmentation model assessment results with the ResNet-34 encoder using the IoU score of 84.50% and the dice loss matrix evaluation of 0.0857. The ResNet-34 model as an encoder in the UNet architecture can improve segmentation accuracy.**

***Index Terms*—Semantic Segmentation, Ultrasound Image, UNet-ResNet**

I. Introduction

Deep Vein Thrombosis (DVT) is a disease caused by offorming of a blood clot (thrombus) in a deep vein. These blood clots can block blood vessels in the lungs, potentially causing serious conditions, such as Pulmonary Embolism (PE) [1]. There is no official confirmation regarding the total number of DVT cases in Indonesia. Still, in 2020 a research at RSUP Dr. Kariadi Semarang reported an increased risk of DVT in patients with glioblastoma and COVID-19 infection due to hypercoagulability and coagulopathy due to tumor cells and the SARS-CoV-2 virus [2].

Medical imaging of DVT cases can use MRI, CT Scan, and Ultrasound (USG) to diagnose DVT areas. The ultrasound modality is the doctor's initial choice in diagnosing a group of patients who are most likely to be positive for DVT [3]. The choice of ultrasound as a diagnostic modality for DVT does not cause pain to the patient, does not pose a risk of spreading radiation exposure to the doctor or patient, and the cost of imaging using ultrasound is much cheaper than CT Scan and MRI. Doppler ultrasound is a type of ultrasound used in diagnosing blood clots based on the results of the evaluation of blood flow in the veins.

A study that clustered DVT on ultrasound was conducted by Berthomier et al. [4]. However, they did not show the clustering results, which proved that the segmentation was successful. Cluster-based segmentation on ultrasound images does not give good results. Studies on blood vessel segmentation were carried out using an elliptical shape approach by Sunarya et al. [5] and Guerrero et al. [6]. In patients with DVT, the use of an ellipse to determine the contours of blood vessels on ultrasound images is difficult because blood clots block it. The weakness of the elliptical approach method is that it is only an estimated outline of the detection area and not the actual area. Research conducted by M. Ikhsan et al. [7] showed the results of identification and classification of blood vessels could be done using the cascading classifier method. Research by Jianfeng et al. [8] finger-vein segmentation based on KFCM and active contour model succeeded in combining the traditional active contour method with a fuzzy kernel to produce an optimal segmentation process. The technique used in the two studies that have been mentioned is still using the semi-automatic method.

Semantic segmentation is the task of producing pixel-level labelling of different object categories. In medical image segmentation, this translates to separating the dominant background class from the smaller Region Of Interest (ROI) class, to produce a binary segmentation map. Studies on automated semantic segmentation using the Deep Learning model Xin Yang et al. [9] and Nabila et al. [10]. The research shows the success of the UNet model training for detecting and

978-1-6654-8165-6/22/$31.00 ©2022 IEEE

segmenting outliners in the specified area. According to Yang Ji et al., the accuracy of Dual-Path UNet was 4% - 5% greater than that of Seg-Net and UNet utilizing an intravascular dataset from Ultrasound. Intravascular modalities can only be achieved since ultrasound must be put into the body and into the vessels [11]. Lang Yuan et al. [12] discovered that when compared to the standard UNet architecture, the modified UNet design with ResNet can recognize many objects with a greater degree of accuracy. The use of deep learning models with multi-object detection capabilities is very important in the case of medical imaging DVT which will detect two object outlines, the vein area, and the blood clot area.

Specifically, we employ the UNet as our segmentation framework and perform some experiments on UNet architecture's encoder. To improve the segmentation accuracy performance we provide use ResNet-32 to construct the contraction part of U-Net on Figure 3. Therefore, the residual structure guarantees the network a better description capability when contracting images with deeper layers. This study aims to produce a comparative analysis of UNet and UNet-ResNet in the segmentation of ultrasound images of veins. The best-performing model will then be used in the development stage of a clot determination system (thrombus) in the DVT case.

II. Materials and Methods

A. Data Acquisition

The ultrasound Telemed SmartUs EXT-1M modality and Echowave II software collected venous ultrasound image data. Researchers collected ultrasound data on a lab-scale and handled it directly. Data is taken from phantom body parts. The total data used is 536 ultrasound images, where 445 data are used for training, 81 data for validation of training, and 10 data for testing. The probe used for the image data collection on Phantom is a linear probe-type L15-7L40H-5. This probe has a frequency range of 7.5 MHz - 15 MHz. Figure 1 shows phantom, modality to acquired data ultrasound image, and probe.

B. Preprocessing

Ultrasound image data acquired using Ultrasound Telemed SmartUs EXT-1M is still raw data, so the acquired image needs to be refined in the preprocessing stage. Preprocessing has several stages. The first is removing unnecessary parts from the image by cropping. This method can eliminate unnecessary parts and focus the image on the object of the vein to be detected. Next is the stage of equalizing the size of the image by rescaling, the cropped image has a size that is not the same between one image and another. The ultrasound image size is set to 224x224 pixels.

Ultrasound imaging exhibits considerable difficulties for medical visual inspection and for the development of automatic analysis methods due to speckle, which negatively affects the perception of tissue boundaries and the performance of automatic segmentation methods. Anisotropic diffusion filter was chosen to reduce image noise without removing significant parts of the image content. Anisotropic diffusion resembles the process that creates a scale space, where an image generates a parameterized family of successively more and more blurred images based on a diffusion process. The method needs the estimate of a parameter q0(t) related to the coefficient of noise variation. The estimation of this parameter is precisely the weakest point of the method [13]. Aisotropic diffusion filter in image filtering is shown in 1.

$$\frac{\partial u(\mathbf{x},t)}{\partial t} = \nabla \cdot (c(q)\nabla u(\mathbf{x},t)),$$
$$u(\mathbf{x},t=0) = u_0(\mathbf{x}), \qquad \left.\frac{\partial u(\mathbf{x},t)}{\partial \hat{n}}\right|_{\partial\Omega} = 0, \tag{1}$$

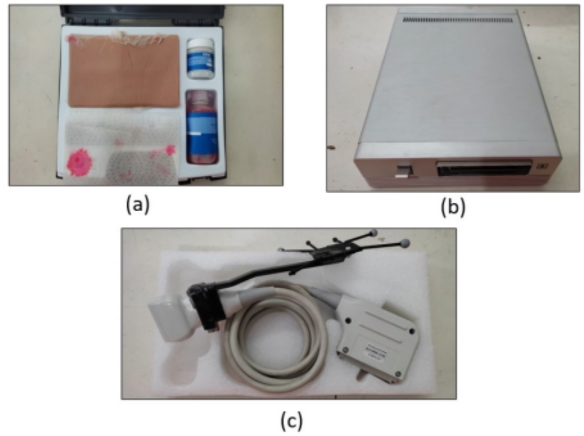

Fig. 1. (a) Phantom human body part standart, (b) USG modality Telemed SmartUs EXT-1M, and (c) Linear Probe type L15-7L40H-5.

The last step in preprocessing ultrasound images is creating a masking image (labelling) on ROI of venous. The labelling process using the freehand method. Figure 2 show an overlay ultrasound image of venous ROI.

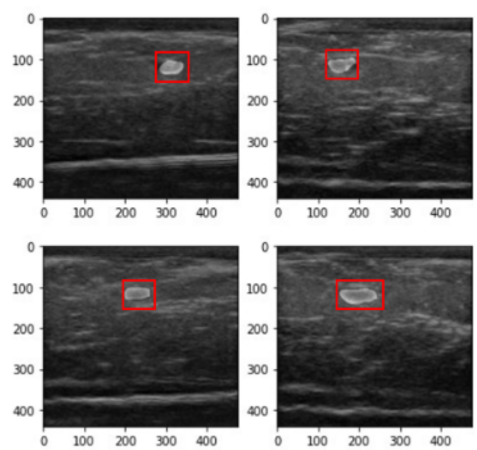

Fig. 2. Venous ROI on Overlay Ultrasound Image.

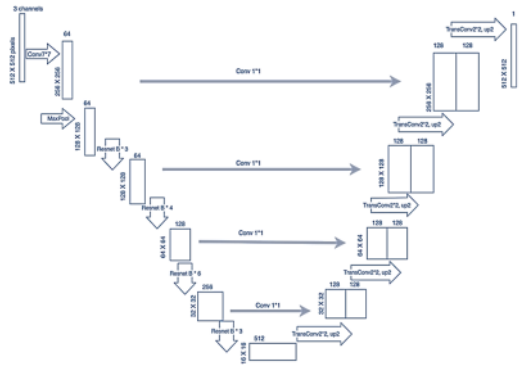

Fig. 3. UNet Architecture with ResNet Encoder

C. Model

UNet is a fully convolutional neural network for semantic image segmentation. Convolutional neural networks are commonly used for classification tasks, where the output to an image is a single class label. However, in many visual jobs, particularly in biomedical image processing, the intended result should contain localization, and each pixel should be allocated a class label. Furthermore, thousands of training images in biomedical tasks are generally out of reach. Trained a network in a sliding-window setup to predict the class label of each pixel by providing a local region (patch) around that pixel [14]. Comprises encoder and decoder elements linked via skip connections. The encoder extracts characteristics of varying spatial resolution (skip connections), which are then employed by the decoder to produce an appropriate segmentation mask. Concatenation is used to join decoder blocks with skip connections.

In theory, an increase in the number of layers in a neural network should increase the quality of the model as long as over-fitting is taken care of. However many architectures have the problem of ”vanishing gradients”. The ResNet architecture is designed to solve this problem. It is based on the idea of shortcut connections. They make it possible to guarantee that with an increase in the number of layers of a neural network it will not need to learn the identical transformation to remain no worse than its counterpart with fewer layers. This is because immediately add a direct connection between the output of each layer with the input of the next to it layer. Thus the neural network remains to learn only residuals. In this regard, ResNet has a great advantage: it is relatively easy to optimize and thereby increases accuracy by adding more layers [15]. Figure 3 shows UNet architecture with ResNet encoder backbone.

D. Segmentation Evaluation Matrics

To evaluate the model trained performance, use Intersection over Union (IoU) and Dice Loss. The quality measure used is IoU, also known as the Jaccard index; it is one of the most used metrics in semantic segmentation, shown in equation 2. IoU is a stan extremely powerful and very simple metric. IoU equals the overlap between the predicted segmentation and the terrain truth, divided by the joining area between the predicted segmentation and the terrain truth [16].

$$IoU = \frac{B_1 \cap B_2}{B_1 \cup B_2} \tag{2}$$

Dice equation 3 as a reproductive measure [17] as a form of validation of manual annotation statistics with segmenters repeatedly annotating with the same ultrasound image, then the pair-wise overlap of segmentation performed repeated calculations using Dice.

$$DICE = \frac{2(S_a|S_b)}{|S_a| + |S_b|} \tag{3}$$

III. Result

This research was conducted on a computer-based on Anaconda with the processor specification of AMD Ryzen 5 4600H and NVIDIA GeForce GTX 1650 Ti. In training the UNet-ResNet model, several input parameters will be defined. Batch size will be defined as 16, epoch 40, initial learning rate of 0.001, the optimizer will use adam, and activation layer in the model architecture will use sigmoid. The first step is to divide the dataset from the data acquisition results with the proportion of 80% (445 images) for training data, 15% (81 images) for validation data and 5% (10 images) for testing data. The training process will use two different segmentation models, first using the standard UNet architectural model and then using the UNet model with the ResNet-34 encoder. The results of the IoU evaluation on the standard UNet model resulted in IoU score of 81.22% shown in Figure 4. The evaluation of the dice loss is 0.1341 shown in in Figure 5.

Training the UNet model with the ResNet-34 encoder using an ultrasound image dataset of venous shows an increase in the accuracy of segmentation of the venous area in each training epoch. The graph of the increase in IoU score in each training epoch can be seen in Figure 6. The results of the model performance evaluation using the IoU score show that the UNet-ResNet model can achieve IoU score of 84.50% and the dice loss value is at 0.0857 shown in Figure 7.

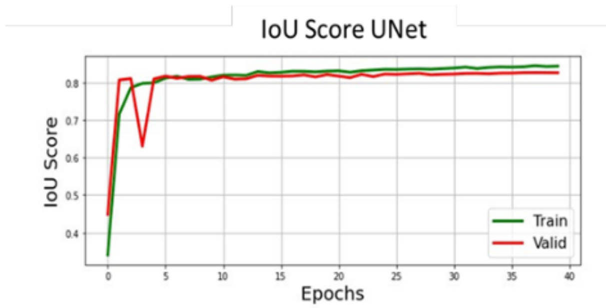

Fig. 4. IOu Score UNet Architecture

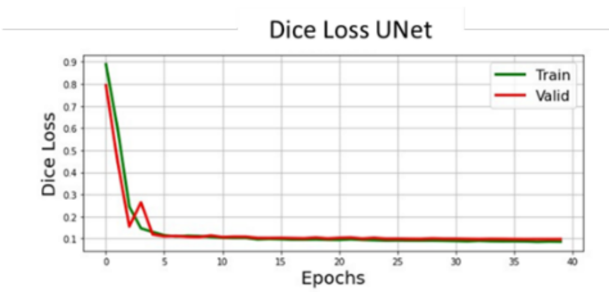

Fig. 5. Dice Loss UNet Architecture

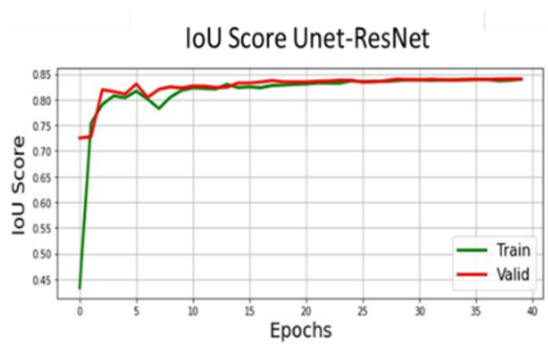

Fig. 6. IOu Score UNet-ResNet34 Architecture

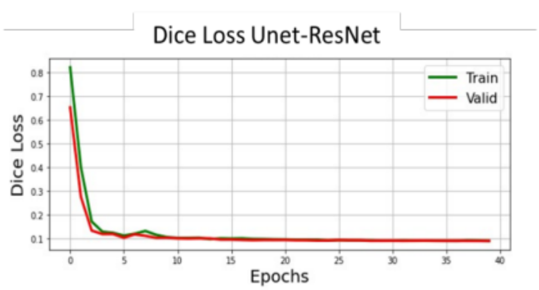

Fig. 7. Dice Loss UNet-ResNet34 Architecture

Different results were obtained from the training results of the two segmentation models, UNet and UNet-ResNet. Figure 8 shows the comparison results of segmentation visualization between UNet and UNet-ResNet model, on the left is an ultrasound image of veins as input, column 2 is a masking image (ground truth), column 3 is a predictive image of segmentation results using the UNet model, and column 4 is a predictive image of segmentation results using the UNet-ResNet model. After the model has gone through the training phase, the model will then be tested using 10 testing images. At this testing stage, the model shows the predicted results of the segmentation of the vein area which is shown in the yellow area.

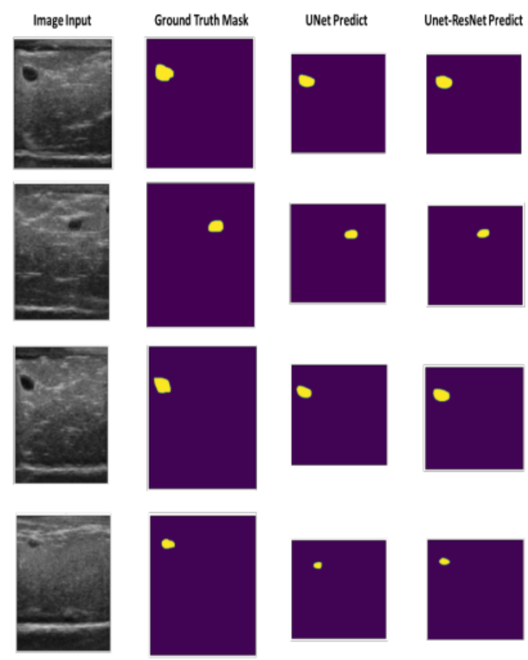

Fig. 8. Comparison Results of Segmentation Visualization between UNet and UNet-ResNet Model

Based on the graphical analysis (Figure 4) of the UNet training process, there is a vanishing gradient problem. This problem occurs because of the use of the sigmoid activation function. Because the value of the sigmoid function ranges from 0 to 1, the value of w will get smaller over time. The longer the layer, the faster the weight (w) value will decrease, especially if repeated epoch processes. Thus, over time the value of the gradient Δw will approach zero, and the adjustment of the w value will no longer be significant. If the w value adjustment is not significant, the decrease in the epsilon value (error) will stagnate, even though the solution has not yet converged. The vanishing gradient problem can be solved by using ResNet architecture as UNet encoder, where a neuron skips several neurons in front of it. The results of the UNet-ResNet model training process in Figure 6 do not indicate the occurrence of the vanishing gradient problem. Training results comparison between UNet and UNet-ResNet can be seen in Table I.

TABLE I
TRAINING RESULTS COMPARISON UNET AND UNET-RESNET

Evaluation	Model	
	UNet	UNet-ResNet
IoU Score	81.22%	84.50%
Dice Loss	0.1341	0.0857

IV. CONCLUSION

Based on the comparison of the IoU scores for the segmentation of the venous area using the UNet-ResNet model with the standard UNet architecture, it shows that the segmentation accuracy performance of the UNet-ResNet model is better than the standard UNet model. The result of the IoU evaluation on the standard U-Net model resulted in IoU score of 81.22% and an assessment of the dice loss is 0.1341. The evaluation UNet segmentation model with the ResNet-34 encoder using the IoU score is 84.50% and evaluation of the dice loss matrix is 0.0857. ResNet-34 architecture for the UNet encoder can solve vanishing gradient problems during the training process. The segmentation of the UNet-ResNet34 model can provide better performance when compared to the UNet architecture. ResNet-34 model as an encoder in the UNet architecture can improve segmentation accuracy.

REFERENCES

[1] L. Mazzolai, V. Aboyans, W. Ageno, G. Agnelli, A. Alatri, R. Bauersachs, M. P. A. Brekelmans, H. R. B¨uller, A. Elias, D. Farge, S. Konstantinides, G. Palareti, P. Prandoni, M. Righini, A. Torbicki, C. Vlachopoulos, and M. Brodmann, "Diagnosis and management of acute deep vein thrombosis: a joint consensus document from the European Society of Cardiology working groups of aorta and peripheral vascular diseases and pulmonary circulation and right ventricular function," European Heart Journal, 02 2017.

[2] D. Tugasworo, A. Kurnianto, R. Retnaningsih, Y. Andhitara, R. Ardhini, D. Priambada, and D. Daynuri, "Glioblastoma dengan deep vein thrombosis pada pasien covid-19: Sebuah laporan kasus," Medica Hospitalia : Journal of Clinical Medicine, vol. 7, no. 1A, p. 181–188, 2020.

[3] P. Wells and D. Anderson, "The diagnosis and treatment of venous thromboembolism," Hematology, vol. 2013, no. 1, pp. 457–463, 12 2013.

[4] T. Berthomier et al., "Unsupervised clustering of DVT Ultrasound Images using High Order Statistics," 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 2018, pp. 2495- 2501.

[5] I. Made Gede Sunarya, Eko Mulyanto Yuniarno, Tri Arief Sardjono, Ismoyo Sunu, P. M. A. (Peter) van Ooijen I. Ketut Eddy Purnama (2020) 3D reconstruction of carotid artery in B-mode ultrasound image using modified template matching based on ellipse feature, Computer Methods in Biomechanics and Biomedical Engineering: Imaging Visualization, 8:3, 301-312.

[6] J. Guerrero, S. E. Salcudean, J. A. McEwen, B. A. Masri and S. Nicolaou, "Real-Time Vessel Segmentation and Tracking for Ultrasound Imaging Applications," in IEEE Transactions on Medical Imaging, vol. 26, no. 8, pp. 1079-1090, Aug. 2007.

[7] M. Ikhsan, K. K. Tan, A. S. Putra, T. H. Sophia Chew and C. F. Kong, "Automatic identification of blood vessel cross-section for central venous catheter placement using a cascading classifier," 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2017.

[8] J. Zhang, Z. Lu and M. Li, "Finger-vein image segmentation based on KFCM and active contour model," 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2019.

[9] X. Yang et al., "Towards Automated Semantic Segmentation in Prenatal Volumetric Ultrasound," in IEEE Transactions on Medical Imaging, vol. 38, no. 1, pp. 180-193, Jan. 2019.

[10] N. Abraham, K. Illanko, N. Khan and D. Androutsos, "Deep Learning for Semantic Segmentation of Brachial Plexus Nervesin Ultrasound Images Using U-Net and M-Net," 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), 2019.

[11] Yang J, Faraji M, Basu A. Robust segmentation of arterial walls in intravascular ultrasound images using Dual-Path U-Net. Ultrasonis. 2019 Jul.

[12] L. Yuan et al., "Multi-Objects Change Detection Based on Res-Unet," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021.

[13] G. Ramos-Llordén, G. Vegas-Sánchez-Ferrero, M. Martin-Fernandez, C. Alberola-López and S. Aja-Fernández, "Anisotropic Diffusion Filter With Memory Based on Speckle Statistics for Ultrasound Images," in IEEE Transactions on Image Processing, vol. 24, no. 1, pp. 345-358, Jan. 2015

[14] O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", MICCAI 2015, 2015.

[15] V. Atliha and D. Šešok, "Comparison of VGG and ResNet used as Encoders for Image Captioning," 2020 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), 2020.

[16] C. A. R. Goyzueta, J. E. C. De la Cruz and W. A. M. Machaca, "Integration of U-Net, ResU-Net and DeepLab Architectures with Intersection Over Union metric for Cells Nuclei Image Segmentation," 2021 IEEE Engineering International Research Conference (EIRCON), 2021.

[17] Li X, Hong Y, Kong D, Zhang X. Automatic segmentation of levator hiatus from ultrasound images using U-net with dense connections. Phys Med Biol. 2019.

Semantic Segmentation of Venous on Deep Vein Thrombosis (DVT) Case using UNet-ResNet

ORIGINALITY REPORT

16%	11%	10%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to National Institute of Technology Karnataka Surathkal Student Paper	2%
2	Umi Laili Yuhana, Laszlo T. Koczy, Tri Arief Sardjono, I Ketut Eddy Purnama, Eko Mulyanto Yuniarno, Mauridhi Hery Purnomo. "Classifying the Complexity of Competency in Elementary School based on Supervised Learners", 2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM), 2018 Publication	1%
3	lmi.med.harvard.edu Internet Source	1%
4	Zhengquan Chu, Tian Tian, Ruyi Feng, Lizhe Wang. "Sea-Land Segmentation With Res-UNet And Fully Connected CRF", IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 2019 Publication	1%

5	serialsjournals.com Internet Source	1 %
6	www.arxiv-vanity.com Internet Source	1 %
7	Dewinda Julianensi Rumala, Eko Mulyanto Yuniarno, Reza Fuad Rachmadi, Supeno Mardi Susiki Nugroho et al. "Bilinear MobileNets for Multi-class Brain Disease Classification Based on Magnetic Resonance Images", 2021 IEEE Region 10 Symposium (TENSYMP), 2021 Publication	1 %
8	medicahospitalia.rskariadi.co.id Internet Source	1 %
9	Dewinda Julianensi Rumala, Eko Mulyanto Yuniarno, Reza Fuad Rachmadi, Supeno Mardi Susiki Nugroho et al. "Activation Functions Evaluation to Improve Performance of Convolutional Neural Network in Brain Disease Classification Based on Magnetic Resonance Images", 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), 2020 Publication	1 %
10	Submitted to Sheffield Hallam University Student Paper	1 %

11	web.archive.org Internet Source	1 %
12	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2020", Springer Science and Business Media LLC, 2020 Publication	1 %
13	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2021", Springer Science and Business Media LLC, 2021 Publication	1 %
14	Submitted to Columbia University Student Paper	<1 %
15	tel.archives-ouvertes.fr Internet Source	<1 %
16	Xi Li, Zhangyong Li, Dewei Yang, Lisha Zhong, Lian Huang, Jinzhao Lin. "Research on Finger Vein Image Segmentation and Blood Sampling Point Location in Automatic Blood Collection", Sensors, 2020 Publication	<1 %
17	iopscience.iop.org Internet Source	<1 %
18	Sumaira Hussain, Xiaoming Xi, Inam Ullah, Syeda Wajiha Naim, Kashif Shaheed, Cuihuan	<1 %

Tian, Yilong Yin. "Difficulty-aware prior-guided hierarchical network for adaptive segmentation of breast tumors", Science China Information Sciences, 2023
Publication

19	ensorisations.firesidegrillandbar.com Internet Source	<1 %
20	www.iis.sinica.edu.tw Internet Source	<1 %
21	www.notebookcheck.net Internet Source	<1 %
22	"Simplifying Medical Ultrasound", Springer Science and Business Media LLC, 2021 Publication	<1 %
23	hal.archives-ouvertes.fr Internet Source	<1 %
24	ijcseonline.org Internet Source	<1 %
25	riset.its.ac.id Internet Source	<1 %

Exclude quotes Off
Exclude bibliography On
Exclude matches Off

Semantic Segmentation of Venous on Deep Vein Thrombosis (DVT) Case using UNet-ResNet

GRADEMARK REPORT

FINAL GRADE

/100

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5