Chest X-Ray Image Segmentation Using 2D V-Net Algorithm to Improve Diagnosis of Lung Disease

Rima Tri Wahyuningrum*
Department of Informatics Engineering (Universitas Trunojoyo Madura)
Bangkalan, Indonesia
rimatriwahyuningrum@trunojoyo.ac.id

Firman Maulana
Department of Informatics Engineering (Universitas Trunojoyo Madura)
Bangkalan, Indonesia
fm28496@gmail.com

Ari Kusumaningsih
Department of Informatics Engineering Universitas Trunojoyo Madura
Bangkalan, Indonesia
ari.kusumaningsih@trunojoyo.ac.id

Budi Dwi Satoto
Department of Informatics Engineering (Universitas Trunojoyo Madura)
Bangkalan, Indonesia
budids@trunojoyo.ac.id

Amillia Kartika Sari
Radiologic Imaging Technology, Departement of Health, Faculty of Vocational Studies (Universitas Airlangga)
Surabaya, Indonesia
amillia.kartika.sari@vokasi.unair.ac.id

Anggraini Dwi Sensusiati
Dept of Radiology, Medical Faculty, Universitas Airlangga Hospital
Surabaya, Indonesia
anggraini-d-s@fk.unair.ac.id

2024 International Seminar on Intelligent Technology and Its Applications (ISITIA) | 979-8-3503-7857-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISITIA63062.2024.10667916

***Abstract*— Various kinds of lung diseases can occur such as asthma, pneumonia, bronchitis, tuberculosis (TB), and many others. This disease is typically characterised by symptoms including wheezing, chest pain, shortness of breath, and chronic cough. The world has recently been impacted by the deadly COVID-19 outbreak. WHO stated that pandemic status was increasing with the number of cases reaching 118,000 infections and more than 4000 deaths in 114 countries. Therefore, detection tools are needed to find out which citizens are infected with this virus, so that they can suppress and reduce the growth rate of daily cases by immediately providing assistance. In this research, the method used to obtain the system algorithm with the best performance is a Convolutional Neural Network (CNN) using the V-Net algorithm as a segmentation method which was tested on the chest x-ray dataset. This method can be correlated with the Reverse Transcription-Polymerase Chain Reaction examination or known as RT-PCR. In this research we have obtained the best model that can produce accurate, fast and efficient segmentation using epoch 15, 2D V-Net obtained evaluation results of Dice Coefficient and IoU metric values of 0.95769 and 0.91882, with a testing time of 1 second on the Qatar University (QU) Kaggle dataset.**

Keywords—Lung Disease, Deep Learning, X-Ray, V-Net

I. INTRODUCTION

Lung diseases, such as tuberculosis, bronchitis, pneumonia, and asthma, arise from infection, smoking, and genetics. Recently, the COVID-19 outbreak, caused by the SARS-CoV-2 virus, has impacted the world, leading to severe respiratory issues [1].

Coronavirus, is a group of viruses that cause infections in the respiratory tract of humans and animals. This virus is classified as a zoonotic group and can be transmitted through direct contact from infected animals to humans [2]. This disease has become a public concern since it first appeared in China in 2019, with thousands of victims dying from the virus which has become the focus in several countries, including Indonesia. This virus includes respiratory diseases such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) because it is infected by SARS-COV-2 (the coronavirus that causes the COVID-19 respiratory infection) [3]. In 2003, SARS-CoV-1 caused over 8,000 cases in 27 countries with a 10% mortality rate and has not been detected in humans since [4]. On January 7, 2020, Chinese authorities identified a new coronavirus, later named SARS-CoV-2, after analyzing respiratory samples [5] [6].

During the COVID-19 pandemic, countries followed WHO guidelines and used Reverse Transcription-Polymerase Chain Reaction (RT-PCR) tests for initial screening to detect infected citizens. [7]. According to D. Wang et al., RT-PCR is the widely used COVID-19 detection tool, examining respiratory tract or blood samples [8]. Overall RT-PCR itself has advantages and disadvantages. The advantage is that it can detect antigens even in low concentrations. Meanwhile, the disadvantages are: it requires a lot of time to process samples, laboratory personnel must have special skills, in terms of sample processing requirements special laboratory equipment and space are required, testing costs are high, and the risk of exposure to stakeholders is high.

In 2021, research on deep learning carried out by Bonechi, et. al [9] shows that the automatic segmentation process is also very useful, especially in the medical world, making it possible to make a time-consuming diagnosis process faster. This research uses a Convolutional Neural Network (CNN) architectural model and uses two other different architectures, namely U-Net and LinkNet with different encoders (ResNet 34 and Inception ResNetV2). Jalali et al. [10] highlighted the significance of Computerized Tomography (CT) scan image segmentation for applications like lung cancer detection. Their study introduced Res BCDU-Net, a modified U-Net with a pre-trained ResNet-34 encoder and Bidirectional Convolutional Long Short-Term Memory (BConvLSTM), achieving a dice coefficient of 97.31%.

In other research, a method to detect COVID-19 is through chest X-rays, which can be analyzed to identify infected patients [11]. Meanwhile, Wahyuningrum et. al used the same dataset in this study, but with the Fully Convolutional Attention Network (FCANet) method to segment Chest X-Ray (CXR) images, obtaining a Dice coefficient accuracy of 97.24% and an IoU of 94.66% on the test dataset [12].

In recent years, Deep Learning networks have contributed to the development of newer image segmentation models with improved performance. Deep neural networks have achieved high levels of accuracy on different popular datasets. Image segmentation techniques can be broadly classified as semantic segmentation and instant segmentation. Semantic segmentation can be considered as a pixel classification

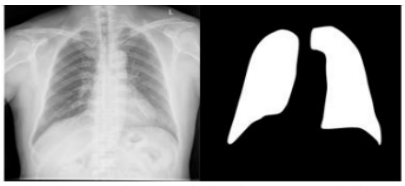

Fig. 1. Segmentation image dataset (left: chest x-ray, right: masking area)

problem. In this segmentation technique, each pixel in the image is labeled with a certain class. Meanwhile, instant segmentation detects and describes every object of interest in the input image [13].

II. METHOD

The dataset comes from the Kaggle Qatar University (QU) Database repository [14] which contains chest x-ray images along with area masks, the amount of original data from the source is 33,920 image data x-ray. The following is an example of x-ray image data and masking area in Figure 1.

. Detailed data distribution and class characteristics are in Table I. This table provides a detailed description that is an important basis for the development of chest x-ray image segmentation models.

TABLE I. SEGMENTATION DATASETS

No	Class	Amount of research data	Training (90%)	Testing (10%)
1	COVID-19 chest X-ray image	500	450	50
2	Chest X-ray Image of Non-COVID-19 (Bacterial Pneumonia)	500	450	50
3	Normal chest X-ray image	500	450	50
	Total	1500	1350 (*random*)	150

This research divides 1500 images into training (1350) and testing (150) data in a 9:1 ratio. The test set includes 50 images each of COVID-19, Non-COVID-19, and Normal, selected randomly from the randomized dataset.

TABLE II. DISTRIBUTION OF TRAINING DATA, VALIDATION, TESTING

K-fold cross validation process			
K-fold (k=5)	**Training**	**Validation**	**Testing**
Fold 1-5	1080	270	150
Total		1500	

The 1350 training images were randomized for cross-validation. K-Fold Cross Validation (k=5) was used to detect overfitting, splitting the data into 1080 training and 270 validation images. These, along with 150 test images, were used to train and test the 2D V-Net model. Following are the details of the distribution of Training, Validation, and Testing data in Table II.

This research was also tested on a chest X-ray dataset obtained from Airlangga University Hospital (RSUA) in Surabaya, Indonesia. The total data amounts to 299 images, consisting of three classes: 32 non-COVID images, 53 non-COVID pneumonia images, and 214 COVID images. Just like the Kaggle dataset, the ratio of training and testing data is 9:1.

By increasing the number of channels, the model can learn to recognize finer and more specific features. At the first level of convolution, the model learns to recognize simple local features, such as edges and corners. At the second level of convolution, the model learns to recognize more complex features, such as shapes and patterns. At the third level of convolution, the model learns to recognize the most complex features, such as texture and structure.

Encoder Process [16]

At the first convolution level, the encoder identifies simple features like edges and corners using 16 channels, which allows for quick and efficient learning. A pooling layer then reduces the image size by half, decreasing the amount of information the model needs to process.

At the second convolution level, the model recognizes more complex features like shapes and patterns, with channels increasing from 16 to 32 for greater detail. A pooling layer then reduces the image size to a quarter of its original size.

At the third convolution level, the model identifies complex features like texture and structure, with channels increasing from 32 to 64 for greater accuracy.

A pooling layer then reduces the image size to one-eighth of its original size.

At the fourth convolution level, channels increase from 64 to 128, enabling the model to capture global features from the input image. Subsequently, a pooling layer reduces the image size to one-tenth of its original size.

Process on Bridge

After that on the bridge, there are 256 channels, which serve to connect the encoder and decoder. The encoder extracts features from the input image, and the decoder reconstructs the output image. The bridge integrates features from both the encoder and decoder parts.

At the first connection level, the model combines features from the third convolution level on the encoder with features from the first deconvolution level on the decoder. At the second connection level, the model combines features from the second convolution level in the encoder with features from the second deconvolution level in the decoder.

Decoder Process

In the decoder, the model uses an upsampling process to rebuild the output image from the features extracted by the encoder. Upsampling is performed using a transpose convolution layer. A transpose convolution layer is a convolution layer that uses small filters to increase the size of the input image. In the 2D V-Net architecture image, upsampling is carried out using a deconvolution layer. Deconvolution layers are transpose convolution layers specifically designed for use in 2D V-Net architectures.

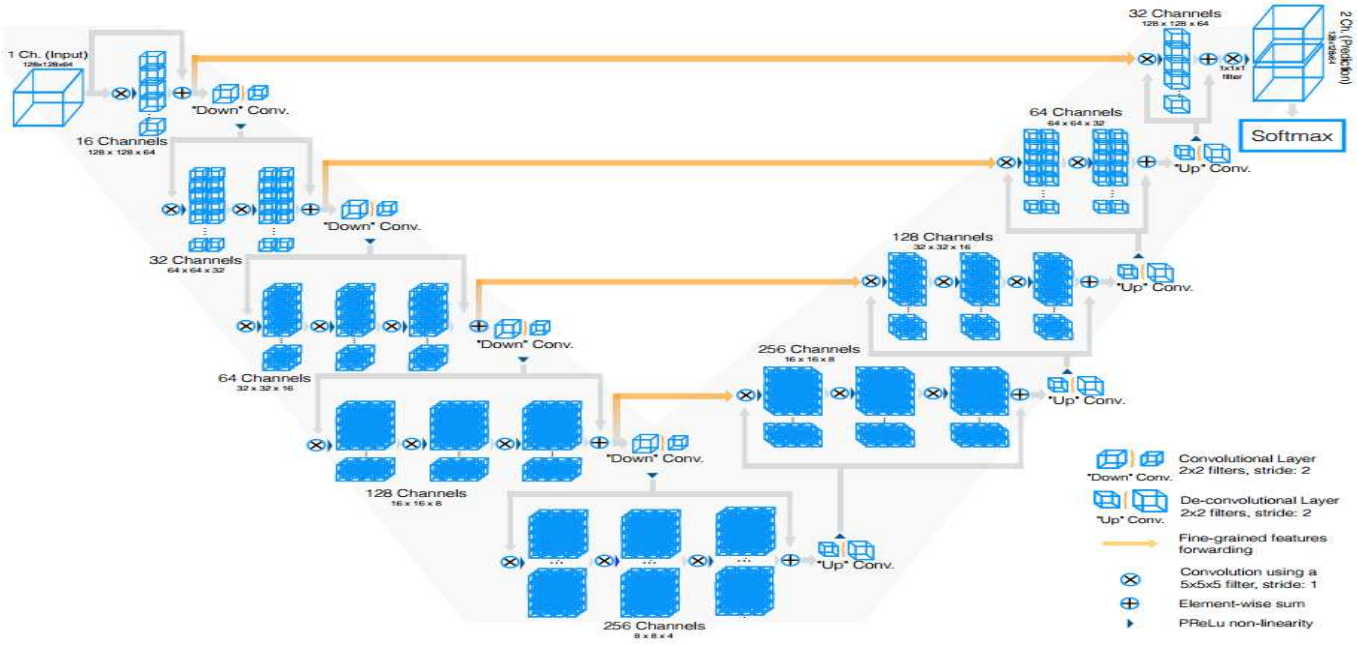

Fig. 2. 2D V-Net Architecture [15]

At the first deconvolution level, the number of channels is reduced from 256 to 128 to streamline information processing. Following this, a deconvolution layer enlarges the input image to twice its original size.

At the second level of deconvolution, the number of channels is reduced from 128 to 64. After that, the model uses a deconvolution layer to increase the size of the input image to four times the original size.

At the third level of deconvolution, the number of channels is reduced from 64 to 32. After that, the model uses a deconvolution layer to increase the size of the input image to eight times the original size.

At the last deconvolution level, the number of channels is reduced from 32 to 1. This last deconvolution layer produces an output image with the same size as the input image. After that, the model uses a softmax layer to produce an output that has a value between 0 and 1. A value of 0 indicates that the pixel is not included in the target, while a value of 1 indicates that the pixel is included in the target.

III. TEST SCENARIO

The test scenario aims to evaluate the accuracy and computation time of using the 2D V-Net architecture for segmenting lung CXR image data. Figure 3 is a flowchart of the segmentation that will be tested.

From the flowchart diagram Figure 3 includes the process:

a. Performs dataset and label input. For the input dataset, data pre-processing has been carried out, we will resize the image size to 256 x 256 and normalize the data.
b. The preprocessed dataset is divided into three sets: training set, validation set, and test set.
c. The 2D V-Net model architecture is defined to determine the number and configuration of layers, filters, and other architectural options.
d. The 2D V-Net model is trained on the training set using an appropriate optimization algorithm (e.g., stochastic gradient descent) and a specified loss function (e.g., binary cross-entropy).
e. The results of the preprocessing dataset stage will be carried out in the training and testing stages, which will later train the V-Net network with chest x-ray images that have been preprocessed.
f. In the model, Hyperparameters (Validation Set), such as learning rate, dropout rate, or batch size, are set using the validation set. Different hyperparameter combinations are tested, and model performance is evaluated on the validation set to select the best hyperparameter configuration.
g. Once the model is trained and tuned, it is evaluated on the test set to assess its performance on unseen data. Model predictions are compared with ground truth labels to measure segmentation accuracy.
h. Compute Evaluation Matrices
i. Evaluation matrices and visual inspection of the segmented output are used to analyze and interpret the results.

From the trials above, there were 2 scenarios carried out:

a. Scenario 1
   The training process uses the Adam optimizer with a batch number of 16, learning rate of 0.0001 and an epoch of 10 epochs
b. Scenario 2
   The training process uses the Adam optimizer with several batches of 16, learning rate of 0.0001 and epochs of 15 epochs. For more details, see Figure 4.

IV. RESULTS AND DISCUSSION

Tables III – VII are the results of experiments on the Kaggle dataset. Table III is presented as the model training results from 5-fold cross validation in scenario 1. This table includes matrix values such as Dice, IoU, Val Dice, Val IoU and records the computational time required during the training process. Meanwhile, Table IV shows the results of the testing process in scenario 1.

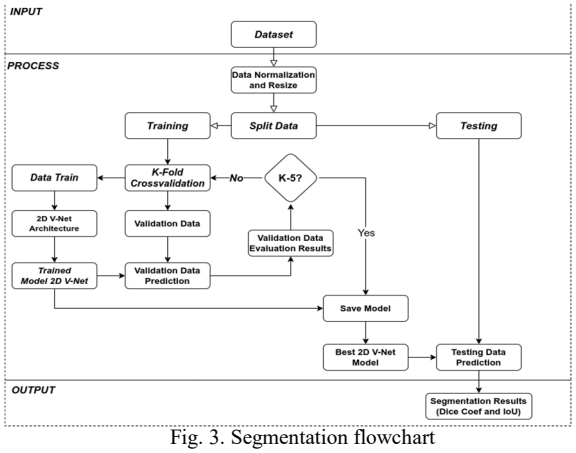

Fig. 3. Segmentation flowchart

TABLE III. RESULTS OF THE TRAINING PROCESS IN SCENARIO 1 FOR KAGGLE DATASET

Fold	Dice	IoU	Val Dice	Val IoU	Time (seconds)
1	0.96081	0.92462	**0.95601**	**0.91585**	24.38617
2	0.96104	0.92505	0.94809	0.90151	23.77582
3	0.95622	0.91616	0.93369	0.87591	23.71611
4	0.95622	0.91621	0.54373	0.37737	23.77297
5	0.96234	0.92744	0.94318	0.89273	24.55466

TABLE IV. RESULTS OF THE TESTING PROCESS IN SCENARIO 1 FOR KAGGLE DATASET

Fold	*Dice*	*IoU*	*Time (seconds)*
1	**0.95292**	**0.91007**	6
2	0.95209	0.90856	1
3	0.92230	0.85581	2
4	0.56089	0.38975	1
5	0.94916	0.90324	2

The best model has the highest Val Dice and Val IoU matrix results. In the training process for scenario 1, the best model was obtained, namely at fold 1 with the results of the Val Dice and Val IoU matrix being 0.95601 and 0.91585 respectively. The best model resulting from training will be applied to carry out testing on new data, which is referred to as previously prepared test data. Meanwhile, in the test results for scenario 1 in Table IV of the five folds that have been tested, the best Dice and IoU results are found in fold 1 of 0.95292 and 0.91007 with a testing time of 6 seconds. This is because the model used is the best model from the training results. The higher Dice and IoU values, the better the predicted image results in approaching ground truth in segmentation. Furthermore, Tables V and VI are the results of the training and testing process in scenario 2, respectively.

TABLE V. RESULTS OF THE TRAINING PROCESS IN SCENARIO 2 FOR KAGGLE DATASET

Fold	*Dice*	*IoU*	*Val Dice*	*Val IoU*	*Time (seconds)*
1	0.97116	0.94396	0.96022	0.92366	23.43004
2	0.96690	0.93595	0.94772	0.90075	23.40821
3	0.97441	0.95012	0.91990	0.85244	24.19717
4	0.97403	0.94939	**0.96221**	**0.92729**	24.27112
5	0.96828	0.93855	0.92877	0.86769	23.56606

TABLE VI. RESULTS OF THE TESTING PROCESS IN SCENARIO 2 FOR KAGGLE DATASET

Fold	*Dice*	*IoU*	*Time (seconds)*
1	0.95359	0.91130	12
2	0.95317	0.91053	1
3	0.90431	0.82533	1
4	**0.95769**	**0.91882**	1
5	0.93695	0.88138	2

In Table V, the best model was obtained, namely in fold 4 with the results of the Val Dice and Val IoU metrics being 0.96221 and 0.92729 respectively. Meanwhile, in the test results for scenario 2 in Table VI of the five folds that have been tested, the best Dice and IoU results are found in fold 4 of 0.95769 and 0.91882 with a testing time of 1 second.

Next, based on the highest results in Tables V and VI, they can be seen that there is a positive change in increasing metric values (Val Dice and Val IoU) as training progresses, while computing time tends to decrease. This phenomenon shows that the model consistently improves the quality of its segmentation, and there is an increase in agreement between Val Dice and Val IoU values at each training epoch. Therefore, it can be concluded that the overall quality of the training data improved in the second trial.

TABLE VII. COMPARISON OF TEST RESULTS FOR SCENARIOS 1 AND 2

Scenario	Dice	IoU	Time (seconds)
1	0.95292	0.91007	6
2	0.95769	0.91882	1

Based on Tables VII show that the test results in scenario 2 obtained higher Dice and IoU results compared to scenario 1. The greater the number of epochs, the more complex the model.

In scenario 1, 2D V-Net achieves an IoU metric value of 0.91882, with a test time of 1 second. It can be seen that the IoU metric results increase in the same direction, but the computing time decreases. Therefore, the test data was better in the second trial. This can be seen from the increase in the IoU value from 0.91007 to 0.91882 with testing time from 6 seconds to 1 second.

In scenario 2, 2D V-Net achieves a DSC metric value of 0.95769, with a testing time of 1 second. It can be seen that the DSC metric results values increase in the same direction as each other, but the computing time decreases. Therefore, the test data was better in the second trial. This can be seen from the increase in the DSC value from 0.95292 to 0.95769 with the testing time from 6 seconds to 1 second.

The segmentation results from the first and second scenarios can be seen in Figure 5 and Figure 6, respectively.. Meanwhile, experiments using the RSUA dataset are presented in Tables VIII – XI and Fig.7 - 8. Table VIII is presented as the model training results from 5-fold cross validation in scenario 1. Meantime, Table IX shows the results of the testing process in scenario 1.

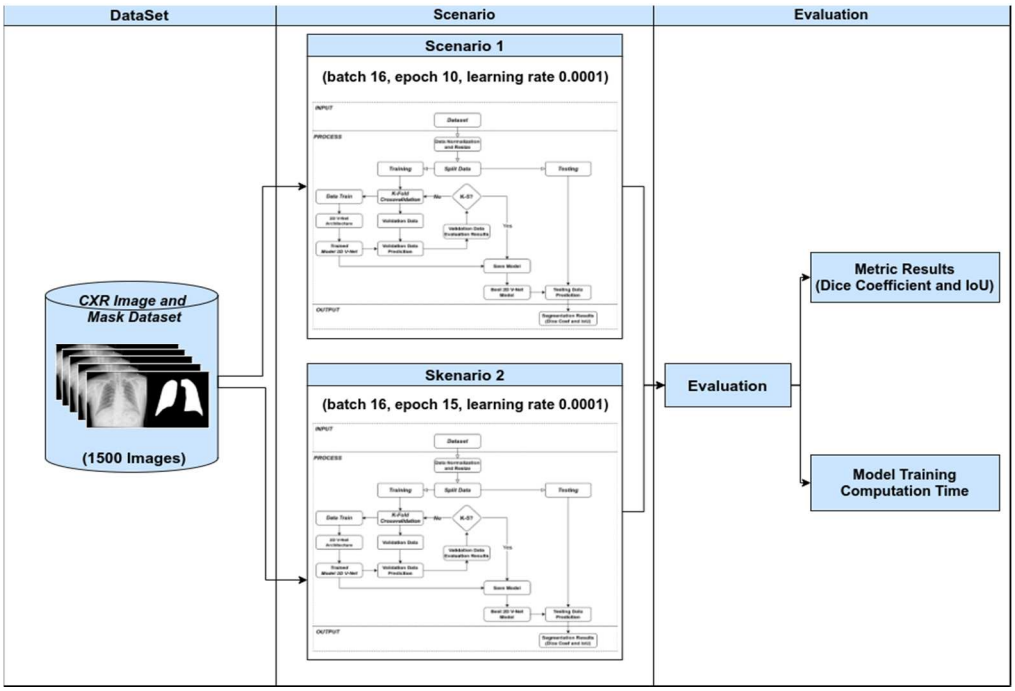

Fig. 4. Test scenario

Based on the test results, it shows that the Dice and IoU RSUA dataset results are smaller than the Kaggle dataset.

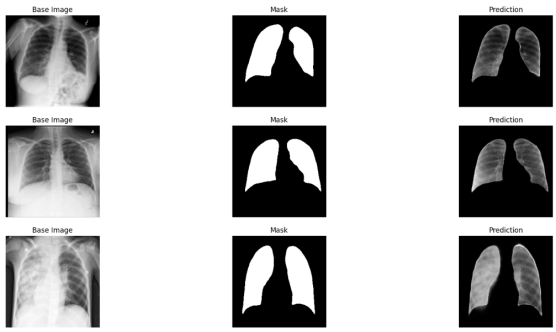

Fig. 5. Segmentation results for scenario 1 for Kaggle dataset

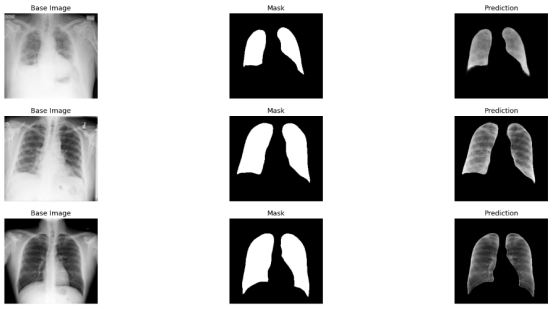

Fig. 6. Segmentation results for scenario 2 for Kaggle dataset

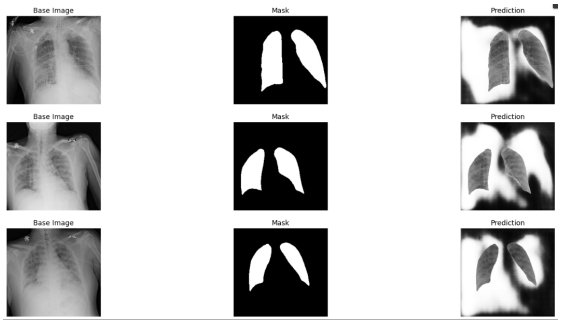

Fig. 7. Segmentation results for scenario 1 for RSUA dataset

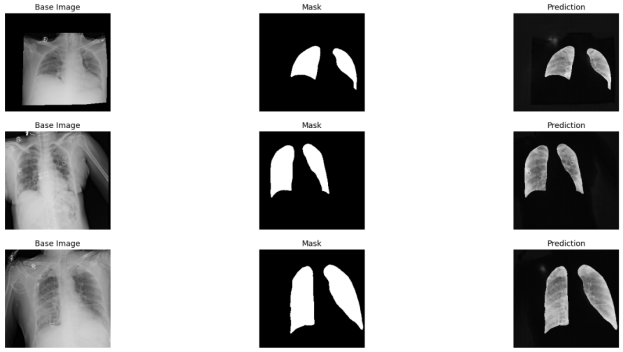

Fig. 8. Segmentation results for scenario 2 for RSUA

TABLE VIII. RESULTS OF THE TRAINING PROCESS IN SCENARIO 1 FOR RSUA DATASET

Fold	*Dice*	*IoU*	*Val Dice*	*Val IoU*	*Loss*
1	0.1659	0.0905	0.16594	0.09048	0.4232
2	0.2186	0.1227	0.21862	0.12272	0.4530
3	0.0657	0.0340	0.06568	0.03395	0.5897
4	0.1561	0.0846	0.15608	0.08465	0.4999
5	0.2492	0.2492	**0.39901**	**0.24923**	1.8278

TABLE IX. RESULTS OF THE TESTING PROCESS IN SCENARIO 1 FOR RSUA DATASET

Fold	*Dice*	*IoU*	*Time (seconds)*
1	0.16594	0.09048	1
2	0.21862	0.12272	1
3	0.06568	0.03395	1
4	0.15608	0.08465	1
5	**0.39901**	**0.24923**	1

TABLE X. RESULTS OF THE TRAINING PROCESS IN SCENARIO 2 FOR RSUA DATASET

Fold	*Dice*	*IoU*	*Val Dice*	*Val IoU*	*Loss*
1	0.7552	0.6066	0.16594	0.09048	0.1566
2	0.7186	0.5614	0.21862	0.12272	0.2491
3	0.0657	0.2472	0.06568	0.03395	1.3624
4	0.1561	0.0899	0.15608	0.08465	0.4255
5	0.2492	0.0891	**0.39901**	**0.24923**	0.5015

TABLE XI. RESULTS OF THE TESTING PROCESS IN SCENARIO 2 FOR RSUA DATASET

Fold	*Dice*	*IoU*	*Time (seconds)*
1	**0.75517**	**0.60664**	4
2	0.71911	0.56142	3
3	0.39641	0.24720	3
4	0.16495	0.08989	3
5	0.16365	0.08912	3

V. CONCLUSIONS

The best test evaluation results were obtained using the Adam optimizer using batch 16, epoch 15 and a learning rate of 0.0001 in scenario 2. The test uses 150 data.

In this test there was an increase in the time value, this value shows that the model has learned better and can produce more accurate segmentation. The metric results increase as the number of epochs increases. The greater the model complexity.

The 2D V-Net model can be trained using limited data and produces high metric values. The more test data there is, the more reliable the consideration of metric results is for automatically segmenting new data.

ACKNOWLEDGMENT

This work was supported by Directorate of Research, Technology and Community Service - Directorate General of Higher Education in the National Competition Applied Research scheme [grant numbers: 101/E5/PG.02.00.PL/2024 and 037/UN46.4.1/PT.01.03/2024].

REFERENCES

[1] P. Zhou, X. Lou Yang, X. Guang Wang, B. Hu, L. Zhang, W. Zhang, H. Rui Si, Y. Zhu, B. Li, C. Lin Huang, H. Dong Chen, J. Chen, Y. Luo, H. a Guo, R. Di Jiang, M. Qin Liu, Y. Chen, X. Rui Shen, X. Wang, X. Shuang Zheng, K. Zhao, Q. Jiao Chen, F. Deng, L. Lin Liu, B. Yan, F. Xian Zhan, Y. Yi Wang, G. Fu Xiao & Zheng-Li Sh, "A pneumonia outbreak associated with a new coronavirus of probable bat origin," *Nature*, vol. 579, no. 7798, pp. 270–273, 2020, doi: 10.1038/s41586-020-2012-7.

[2] C. Martinez-Perez, C. Alvarez-Peregrina, C. Villa-Collar, and S. anchez-Tena, "Citation Network Analysis of the Novel Coronavirus Disease 2019 (COVID-19)," International Journal of Environmental Research and Public Health, vol. 17, no. 20, p. 7690, 0 2020.

[3] T. P. Velavan and C. G. Meyer, "The COVID19 epidemic," Tropical Medicine & International Health, vol. 25, no. 3, pp. 278–280, 0 2020

[4] S. Ochoa and S. Cueto, "Enfermedad por coronavirus 2019 (COVID-19)," Medicina UPB, vol. 40, no. 2, pp. 41–49, 2021.

[5] F. Wu, S. Zhao, B. Yu, Y. Mei Chen, W. Wang, Z. Gang Song, Y. Hu, Z. Wu Tao, J. Hua Tian, Y. Yuan Pei, M. Li Yuan, Y. Ling Zhang, F. Hui Dai, Y. Liu, Q. Min Wang, J. Jiao Zheng, L. Xu, Edward C. Holmes & Y. Zhen Zhang, "A new coronavirus associated with human respiratory disease in China," Nature, vol. 579, no. 7798, pp. 265–269, 0 2020.

[6] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," The Lancet, vol. 395, no. 10223, pp. 497–506, Feb. 2020.

[7] W. G. Note, "Laboratory testing for coronavirus disease (COVID-19) in suspected human cases: interim guidance, September 11, 2020. URL://who.int/publications/i/item/10665-331501, accessed 14 August 2023.

[8] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, BinbinWang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, XinghuanWang, Z. Peng,, "Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China," JAMA, vol. 323, no. 11, p. 1061, Mar. 2020, doi: 10.1001/jama.2020.1585.

[9] S. Bonechi, P. Andreini, A. Mecocci, N. Giannelli, F. Scarselli, E. Neri, M. Bianchini and G. Maria Dimitri, "Segmentation of Aorta 3D CT Images Based on 2D Convolutional Neural Networks," Electronics, vol. 10, no. 20, p. 2559, Oct. 2021, doi: 10.3390/electronics10202559.

[10] Y. Jalali, M. Fateh, M. Rezvani, V. Abolghasemi, and M. H. Anisi, "ResBCDU-Net: A Deep Learning Framework for Lung CT Image Segmentation," Sensors, vol. 21, no. 1, p. 268, Jan. 2021, doi: 10.3390/s21010268.

[11] D. Cozzi, M. Albanesi, E. Cavigli, C. Moroni, A. Bindi, S. Luvarà, S. Lucarini, S. Busoni, L. Nicola Mazzoni, V. Miele, "Chest X-ray in new Coronavirus Disease 2019 (COVID-19) infection: findings and correlation with clinical outcome," La radiologia medica, vol. 125, no. 8, pp. 730–737, Jun. 2020, doi: 10.1007/s11547-020-01232-9.

[12] R. T. Wahyuningrum, I. Yunita, I. A. Siradjuddin, B. D. Satoto, A. K. Sari, and A. D. Sensusiati, "Improvement of chest X-ray image segmentation accuracy based on FCA-Net," Cogent Engineering, vol. 10, no. 1, Jun. 2023, doi: 10.1080/23311916.2023.2229571.

[13] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G. Zhong Yang, "Deep Learning for Health Informatics," IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 1, pp. 4–21, Jan. 2017, doi: 10.1109/jbhi.2016.2636665.

[14] A. M. Tahir, "COVID-QU-Ex Dataset." [Online]. 2022. URL://www.kaggle.com/datasets/anasmohammedtahir/covidqu, accessed May 9, 2023

[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.

[16] T. Panboonyuen, "Semantic segmentation on remotely sensed images using deep convolutional encoder-decoder neural network." Office of Academic Resources, Chulalongkorn University.