Journal of Advanced Zoology

ISSN: 0253-7214

*Volume **44** Issue **S-2 Year 2023** Page **3857:3866***

A STUDY COMPARING THE FUNDUS IMAGE-GENERATING MODEL'S PERFORMANCE

YongSuk Kim[1]

[1]Konyang University, 158 Gwanjeodong-ro, Seo-gu, Daejeon, South Korea

*yongsuk@konyang.ac.kr.

Article History

Received: 12 July 2023
Revised: 10 September 2023
Accepted:30 October 2023

ABSTRACT

Artificial intelligence technologies have been used much more often in recent years for processing images in medical research, including the analysis of medical data. However, the existence of personal and medical information presents several challenges for deep learning studies in the medical industry. Research projects suffer significant delays and financial losses as a result. Consequently, in order to stimulate research on deep learning in the medical domain, a medical image creation study was carried out in this study to identify the features of lesions using aberrant medical pictures inside medical data. A total of 356 fundus photos were used in this investigation, and three lesions, comprising images that were normal, were identified utilizing the 'Ocular Disease Intelligent Recognition' open dataset. The Res U-Net deep learning model, which generates images like the real fundus image by adding Residual Blocks to the U-Net structure which produces current fundus images, was employed in this work to generate fundus images. Furthermore, in order to compare the effectiveness of the current Res U-Net model and U-Net model in this investigation, three picture similarity indicators and ophthalmologist verification were used to objectively analyze and assess the fundus images produced by each model. The comparative assessment findings demonstrated that Res U-Net outperformed traditional models in all picture similarity measures, with Fréchet inception distance (FID) demonstrating an 8-fold improvement in performance. The average area under the curve, also known as or AUC, for all four lesions after the fundus picture produced in this study was validated by an ophthalmologist is Res U-Net 0.7415, U-Net 0.7705, indicating that the image that was generated was more similar to the original image. Subsequent studies will look into lesion insertion and deletion as well as patient data generation models with fundus images.

Keywords: *Fundus Image, Fréchet inception distance (FID), Res U-Net Ophthalmologist, Deep Learning, Image Generation*

1. INTRODUCTION

Numerous research in a variety of sectors have benefited from the development of deep learning-based image processing technology (Dhamo et al., 2020). several research projects are also underway in the field of medicine. Of particular note are the several published examples of deep learning research employing medical imaging (Shen et al., 2017). Nevertheless, acquiring the medical picture data needed for deep learning models to be created in the realm of medicine has several challenges (Kim et al., 2020). A number of legal processes, namely the Institutional Review Board, or IRB, approval process, must be completed before the medical imaging study may move forward. It has a lengthy and intricate process that must be completed in order to move on with the research (Dyrbye et al., 2007). De-identification via anonymous and data verification for healthcare staff members—including physicians—must also be carried out. Research projects suffer significant delays and financial losses as a result of this. Furthermore, the ratio of aberrant (i.e., lesioned) data to normal data is unbalanced, and data security is a complete challenge. The performance if the algorithm used for deep learning is determined by the quantity and quality of the data (Sun et al., 2017). Artificial intelligence research in the healthcare domain should address data acquisition challenges, and effective study need to be carried out by gathering superior medical data. Thus, a model for creating fundus pictures from medical photos is investigated in this study by expanding on the current research on medical image generation. Three image resemblance indicators and ophthalmologists are used to compare this model with the current generation model.

2. REVIEW OF LITERATURE

2.1 U-Net

Deep learning network U-Net was created for quick and precise segmentation of pictures in the biomedical industry. Cell structural segmentation of electron microscopy was its initial application, based on picture extraction of features and segmentation performance (Ronneberger et al., 2015). The U-Net gets its name from the fact that its structure is a 'U' formed network with a left-to-right symmetrical layout. Two processes make up the U-Net: the contraction process and the expansion process. Down sampling is used in the contraction procedure to get information between adjacent pixels and to extract features from the input picture. Convolution layers make up each layer in the contraction process, and as learning proceeds, the number of channels doubles and the dimension of the map of features is half. The contraction procedure ends with a connection with the extension path called Skip Connection. While connecting between the contraction phase to a prolongation process, the skip connection helps to preserve image information. This portion of the expansion process combines feature extraction data from an image that was acquired by up sampling during the contraction phase. In contrast with the contraction procedure, the extension path likewise uses a mixture of convolutional layers within each layer, but as learning advances, the dimension of the map of features double while the number of channel half. U-Net is also employed in picture creation investigations, such as image segmentation studies, because to its outstanding efficacy for image feature extraction. In order to create pictures, a study was done in 2018 to extract characteristics from photographs using U-Net (Esser et al., 2018).

2.2 Residual Block

There are issues with deep learning networks because when the network's steps rise and go deeper, the weight value of the model learning changes unnaturally. Gradient Exploding, whereby the value that is weighted gives off and the gradient steadily grows, is similar to Gradient Vanishing, during which that gradient gradually lowers. These issues lead to a degradation of deep learning networks. Residual Block is one approach of solving this issue. The accumulated convolution procedure is used successively in the procedure of learning of the network of deep learning. However, Residual Block transfers parameters by setting the current convolutional layer block, adding values that were previously received (He et al., 2016). It is feasible to enhance performance by fixing issues that arise as the network used for deep learning gets deeper through the addition of blocks with straight forward computations.

3. Materials and Methods

This study uses two types of materials and methodologies. The procedure of analyzing and classifying research data is explained in the initial data analysis. The study's Res U-Net is described in the second the fundus image generating model.

3.1 Data Analysis

'Oculative Disease Intelligent Recognition', a fundus image dataset gathered by Shangong Medical Technology Co., Ltd., a subsidiary of Kaggle, is used in this research (Larxel et al., 2020). China's hospitals and healthcare facilities provided the data for this collection. It is taken at different picture resolutions using different fundus cameras, including Zeiss, Kowa, and Cannon. Fundus pictures of both eyes, together with the patient's gender and years of age, were gathered in 5,000 patients based on their age. Additionally, there is an ophthalmic database with the number of patient's as well as diagnosis from the treating physician. The 6,392 fundus pictures in the collection are categorized into 8 lesions in total. Normal, diabetic retinopathy, diabetic glaucoma, cataract, macular degeneration, excessive hypertension, pathologic myopia, and other disorders are among the identified lesions. This study utilizes four sorts of these conditions: macular degeneration, glaucoma, diabetic retinopathy, and normal.

The Korean Ophthalmological Society has identified three lesions that, when absented from normal circumstances, are indicative of blinding disorders. The average incidence of each of these lesions was 19.6% for retinopathy caused by diabetes, 3.4% for glaucoma, and 13.4% for macular degeneration, according to the 2018 National Health as well as Nutrition Survey, which was carried out by the Korea Disease Control and Prevention Agency and the Korean Ophthalmological Society (Korea Disease Control and Prevention Agency, 2019). Consequently, the decision was made to aggressively investigate the fundus picture of the lesion in this research, and four lesions, including the normal fundus image, were selected as study subjects

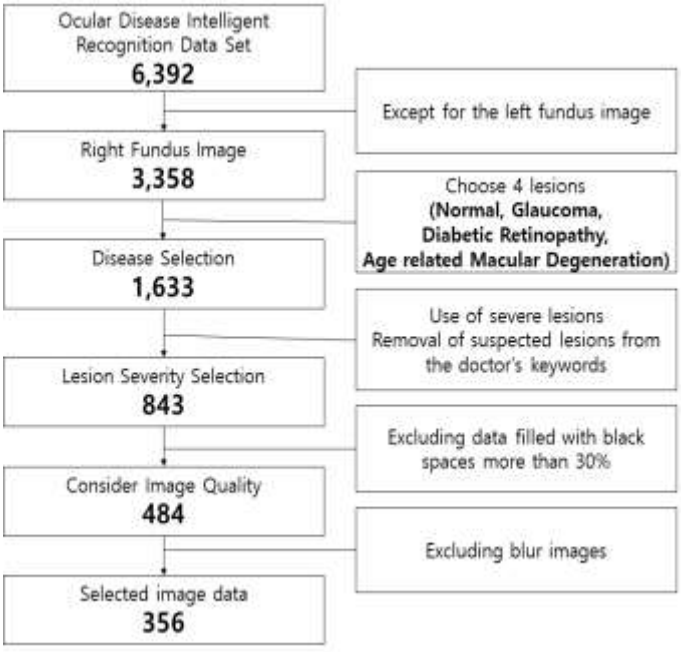

Figure 1: 1: Data Classification Process

The dataset utilized in this paper's classification procedure is shown in Fig. 1. We build a classification technique of 6,392 fundus pictures to enhance the quality of produced images and the efficacy of deep learning models. Only the appropriate fundus picture is chosen in the first stage. This took into account the learning data's consistency. The lesion that will be investigated is chosen in the second stage. Four lesions in all, including the normal the fundus, was chosen. Lesions of a certain severity or more were taken into account while choosing doctor keywords in the third phase.

Excluding data with at least 30% of space (black area) in the fundus picture is the fourth stage. In the sixth phase, fundus photos with blurry edges were removed based on image quality.

356 of the 6,392 data were chosen for this paper's research after categorization. There are 97 macular degeneration, 78 diabetic retinopathy, 67 glaucoma, and 114 normal fundus. Out of around 2,000 photos, the normal lesion was chosen at random since its number was greater compared to the number of other lesions. Due to low picture quality in glaucoma and the little quantity of fundamental data in the dataset for diabetic retinopathy, relatively little data was employed for data categorization. Afterwards, a sharpening filter was used to preprocess the picture and reduce its size from 512x512 to 256x256 in order to address the issue of the model not recognizing the features of small blood vessels or lesions within the low-weight fundus image during learning.

3.2 Fundus Image Generation Model

In 2018, Zhang et al. developed Res U-Net to extract roads from high-resolution aerial pictures. It has since been applied in a variety of sectors, including medicine (Xiao et al., 2018). Residual Block and U-Net are combined to create Res U-Net. The model's general architecture uses U-Net to extract picture characteristics effectively from a little quantity of input. We use Residual Block to address Gradients Vanishing & Exploding in the U-Net contraction and expansion procedures. Additionally, it lessens the degradation of low-weight data that happens as the structure of the machine learning algorithm gets deeper. When utilizing current U-Net, a more advanced deep learning network is possible.

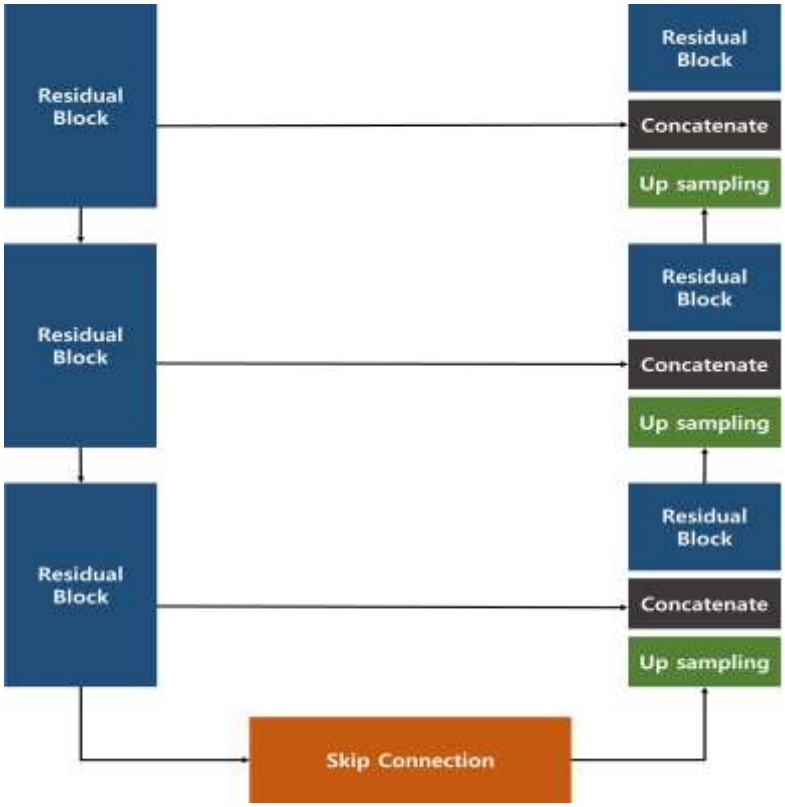

Figure 2: Res U-Net Model Configuration

Image of the fundus The Res U-Net architecture for creating fundus pictures is depicted in Fig. 2. Three Residual Block1 layers make up the contraction phase on the left. It helps to extract features from the input picture during the contraction process. After then, a skip connection is made to avoid losing picture data. The structure of the skip connection and the residual block are identical. The extension procedure is shown on the right portion of Figure 2. It is composed of three layers, just as the contraction process. There are three steps in every stratum. The Up Sampling procedure is the initial phase. The total amount of channels was half and the feature map's size are increased.

Secondly, by employing the distinctive information from the contraction stage in the expansion process, the 'Concatenate' operation reduces the amount of information lost. Third, Residual Block is utilized similarly to the shrinking procedure. The 1x1 convolution layer in the final output layer produces an image with the same dimensions and channels as the input picture.

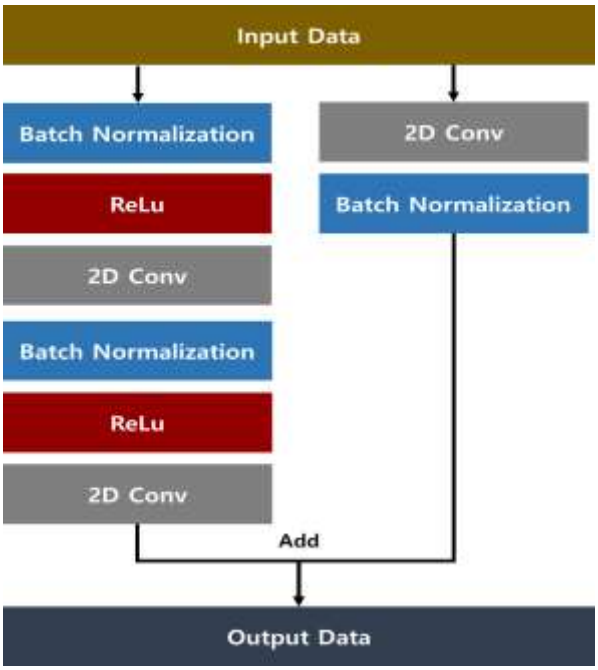

Figure 3: Residual Block Configuration

Res U-Net's Residual Block structure is seen in Fig. 3. Once data is received, it is sent through the Shortcut Connection to the right and the layer of convolution on the left by the Residual Block. Batch normalization is initially carried out via the convolution layer on the left side of image. By stabilizing the model's learning process, batches normalization is a technique for increasing learning speed. Utilizing the process of activating functional ReLu (Rectified Linear Unit), the operation of convolution is carried out in the second stage. Learning speed can be increased by using the ReLu function. The two-dimensional convolution operation is performed in the third phase. Convolution techniques are used to identify the image's properties. Subsequently, repeat the preceding procedure. Diagram 3 To communicate the importance of the incoming data to the left convolution layer's computational output, there is a Shortcuts Connection on the right. Batch Normalization and 2D convolution make up the setup. The convolution layer's total computation results plus the input data that was transmitted from the appropriate shortcut connection make up the output data.

The focus on the outside border section of the fundus picture was caused by the sharpening filter's impact during the image preparation process, as seen by the significant difference in pixel values among the black blank portion and the fundus image. Blurring is done by post-processing the picture to address this issue. When the filter known as the Gaussian is applied, the element's value is 1 and the kernel's size is 3. By dividing every component by 1/9, the mean filtering was applied, and the sum was set to 1.

4. RESULTS AND DISCUSSION

This research compares the fundus image generating model's performance using three different approaches. In the first, the u-net and res u-net produced pictures are visually compared with the original image. Three image similarity evaluations are used to compare the second in a quantitative manner. An ophthalmologist compares the third by verifying it.

4.1 Compare the Outcomes of Fundus Image Generation

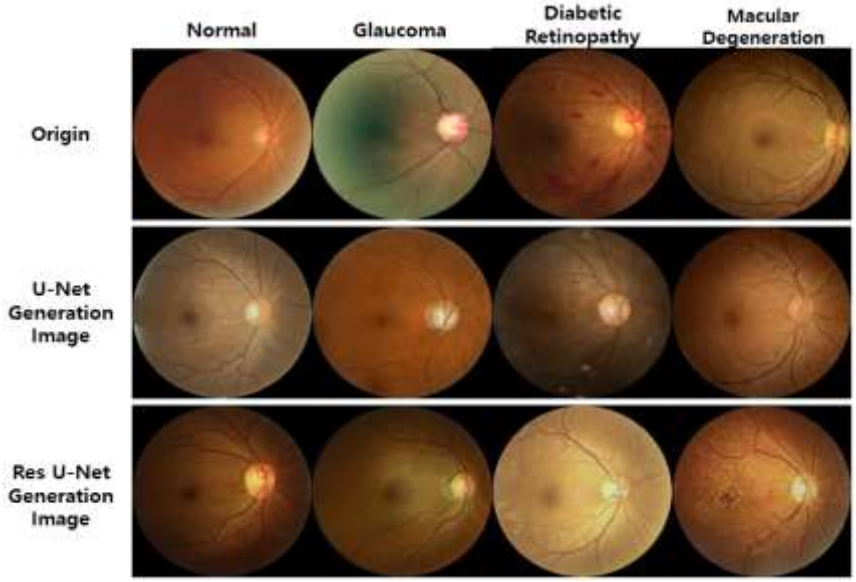

Figure 4: Compare generated fundus images

Figure 4 displays an analysis of the original images vs those generated through the High-Res U-Net as well as U-Net models. The U-Net-generated picture in the prior study had an issue where the fundus's general color was similarly changed (Kim et al., 2022). The retina and blood vessels in the original image display a low color when compared to the U-Net produced image. This issue arises from the loss of the original weight value of the input image's retina and blood vessels when the U-Net model learns further deeply within the deep learning network. When microvascular vessels have low weights compared to the backdrop, they may become invisible. It gets harder to tell what color the backdrop is when the retina's brightest region's pixel value drops. The contrast on the blood vessels and retinal in the produced image, however, is identical to the original image since Res U-Net employs Residual Block to keep initial weighted values. Furthermore, there is a noticeable difference in the image quality compared to the U-Net produced image. When compared with the original image, the Res U-Net produced image does not appear visibly altered.

4.2. Comparing the Findings of the Fundus Image Similarities Evaluation

The difference in picture similarity between the produced and original images is expressed quantitatively in this research. Among methods for evaluating image similarity, the Root Mean Square Error (RMSE), Structural Similarity Index Map (SSIM), and Fréchet inception distances (FID) are used to confirm variations in similarity across pictures. As the produced picture and the original image get closer in comparison, the RMSE and FID values decrease. Because the created picture and the original image are comparable, the SSIM returns a result that is near to 1.

Table 1: Evaluation of image similarity

		RMSE	SSIM	FID
Res U-Net	Normal	15.38	0.94	6.37
	Glaucoma	27.94	0.75	63.48
	Diabetic Retinopathy	20.09	0.90	20.66
	Macular	30.4	0.84	38.09

	Degeneration			
U-Net	Normal	38.58	0.65	254.32
	Glaucoma	39.93	0.65	310.90
	Diabetic Retinopathy	36.76	0.61	284.28
	Macular Degeneration	37.78	0.64	253.49

The created picture's image similarity assessment index findings are compared in Table 1. The RMSE test was the lowest normal lesion while the prevalence of macular degeneration was over two times as high, using the Res U-Net indications. Glaucoma had the lowest value in SSIM at 0.75, whereas the normal lesion value was the one that was closest to 1. Glaucoma had the greatest score and FID closest the normal lesions value to 0. On the picture similarity index, there was a disparity between the RMSE with FID scores and the original image's similarity. Still, every signal performs better when seen through the lens of U-Net's picture similarity evaluation index. Comparing Res U-Net's RMSE 23.45, SSIM 0.85, and FID 32.15 to the average for the parameters by lesion, and U-Net's RMSE 38.25, SSIM 0.63, & FID 275.74. More than 15 RMSE, 0.22 SSIM, and more than 240 FID differences are observed. There is a noticeable discrepancy between the three picture similarity metrics. This indicates that the picture produced from the Res U-Net model is more effective at generating images than the one from the prior study, and that it is quantitatively equivalent to the original image.

4.3 Analysis of Ophthalmologist Verification Outcomes Comparing

Three ophthalmologists in this work sought data verification in order to compare and assess the therapeutic effectiveness of fundus pictures produced by two deep learning models. Finding the produced picture amongst randomly mixed fundus photos is how the needed data is found. It comprises of 45 actual fundus images and 5 created fundus images per lesion. The produced fundus picture's patient information was not disclosed during verification by an ophthalmologist, and the original fundus image was created in the same manner. By doing this, the impact of ophthalmologists' omission of patient data on fundus images verification will be reduced. The average of the verification ratings from three ophthalmologists was used to express each figure. This is due to the fact that understanding the average reading proficiency of all ophthalmologists is more essential than assessing the proficiency of individual ophthalmologists in reading.

The Area Under the Curve (AUC) and the Receiver Operating Characteristic (ROC) Curves display the findings of the verification conducted by ophthalmologists. According to Fawcett et al. (2006), the ROC Curve is a graph that shows how well artificial intelligence models perform in terms of categorization across all classification thresholds. The categorization index point is the area of the chart where the line breaks. The ROC Curve's proximity to the upper left corner of the graph indicates how well the model performs in terms of classification. According to Huang et al. (2005), the AUC computes the area beneath the ROC Curve and offers a thorough performance indicator for each and every classification criterion. When the AUC value approaches 1, it indicates that the classification model is doing better; when it falls below 0.5, it indicates no classification performance at all.

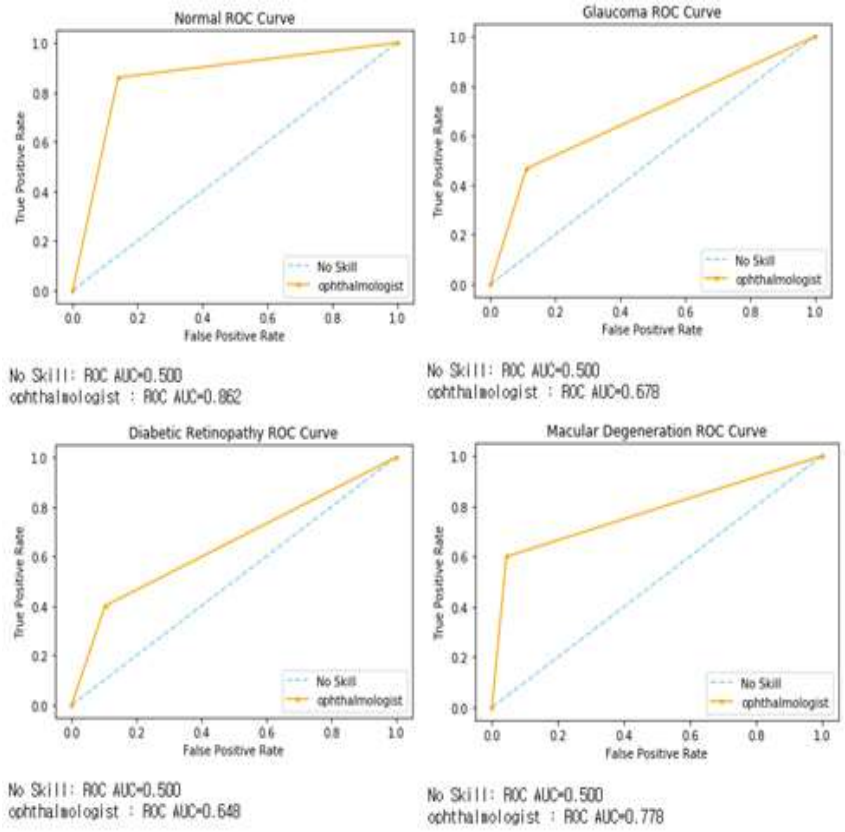

Figure 5: Res U-Net Image Verification Outcomes

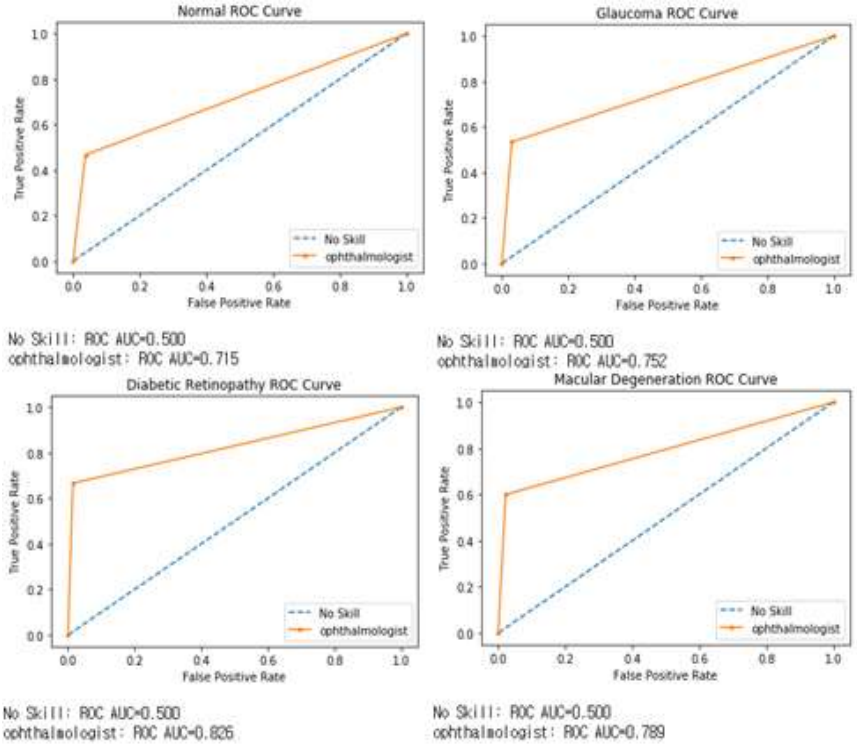

Figure 6: Results of U-Net Image Verification

Ophthalmologist verification findings of pictures produced using Res U-Net and U-Net models are displayed in Figs. 5 and 6. All lesions, with the exception of normally produced pictures, have high AUC values upon verification by an ophthalmologist. Under typical circumstances, Res U-Net 0.862 and U-Net 0.715 were noted. Res U-Net 0.678 and U-Net 0.752 for glaucoma were noted. Res U-Net 0.648 and U-Net 0.826 exhibit the most difference in the context of diabetic retinopathy. Furthermore, the AUC values for macular degeneration are identical for Res U-Net 0.778 and U-

Net 0.789. As the average AUC score was used to compare the entire statistics of ophthalmology data verification, the Res U-Net generated picture scored 0.7415 while the U-Net generated image scored 0.7705. This indicates that compared to the U-Net generated picture, the Res U-Net produced image is around 3% harder to identify. One may argue that there is little to no distinction between the two models. When used in actual clinical settings, however, it indicates that 2,585 produced fundus photos among 10,000 fundus images—a relatively high number—were missed by the ophthalmologist as a result, it is evident that the visual difference between the produced picture by the Res U-Net model in the present research and the actual fundus picture is so tiny that the ophthalmologist only finds around 30% of the image.

5. Conclusion

This work addresses the challenge of acquiring medical imaging data needed for deep learning model research in the medical domain. Fundus photos were used to create data for deep learning models, and fundus images produced by models from previous generations were contrasted. Preprocessing of the images was done in addition to the data categorization procedure. A Res U-Net-based generating model was then used to create a fundus picture. When image similarity evaluation was used to compare the created model of this study to the present model, all three of the picture similarity indicators showed better values. The verification findings for the ophthalmologist revealed Res U-Net generated pictures with a score of 0.7415 and U-Net generated image with a score of 0.7705. Thus, it may be concluded that there is no visual, quantitative, or clinical difference between the fundus picture produced by this paper's generation model and the real fundus image.

In a subsequent investigation, a model based on a fundus picture will be examined for the purposes of adding and eliminating lesions and producing patient data. Additionally, it intends to carry out research on the creation of medical information in a variety of contexts, such as but not limited to the fundus pictures and images from chest radiation therapy.

6. References

1. Dhamo, H., Farshad, A., Laina, I., Navab, N., Hager, G.D., Tombari, F., & Rupprecht, C. (2020). Semantic image manipulation using scene graphs. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5213-5222.
2. Shen, D., Wu, G., & Suk, H.I. (2017). Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering, 19(1), 221–248.
3. Kim, Y.J., & Kim, K.K. (2020). Development of an optimized deep learning model for medical images. Journal of the Korean Society of Radiology, 81(6), 1274-1289.
4. Dyrbye, L.N., Thomas, M.R., Mechaber, A.J., Eacker, A., Harper, W., Massie, F.S., David, V., & Shanafelt, T.D. (2007). Medical education research and IRB review: an analysis and comparison of the IRB review process at six institutions. Academic Medicine, 82(7), 654-660.
5. Sun, C., Shrivastava, A., Singh, S., & Gupta, A. (2017). Revisiting unreasonable effectiveness of data in deep learning era. In Proceedings of the IEEE international conference on computer vision, 843-852.
6. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, 234-241.
7. Esser, P., Sutter, E., & Ommer, B. (2018). A variational u-net for conditional appearance and shape generation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 8857-8866.
8. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.

9. Larxel. (2020). Ocular Disease Recognition: Right and left eye fundus photographs of 5000 patients. WEB: https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k/
10. Korea Disease Control and Prevention Agency. (2019). resource book of the results of the 2018 National Health and Nutrition Survey and Youth Health Behavior Survey. WEB: https://knhanes.kdca.go.kr/knhanes/sub04/sub04_04_03.do/
11. Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual u-net. IEEE Geoscience and Remote Sensing Letters, 15(5), 749-753.
12. Xiao, X., Lian, S., Luo, Z., & Li, S. (2018). Weighted res-unet for high-quality retina vessel segmentation. In 2018 9th international conference on information technology in medicine and education (ITME), 327-331.
13. Kim, Y.S., Song, H.J., & Han, J.H. (2022). A Deepfake-Based Deep Learning Algorithm for Medical Data Manipulation Detection. Journal of System and Management Sciences, 12(1), 13-26
14. Fawcett, T. (2006). An introduction to ROC analysis. Pattern recognition letters, 27(8), 861-874.
15. Huang, J., & Ling, C. X. (2005). Using AUC and accuracy in evaluating learning algorithms. IEEE Transactions on knowledge and Data Engineering, 17(3), 299-310.