

# Squeeze-excitation half U-Net and synthetic minority oversampling technique oversampling for papilledema image classification

Wiharto, Angga Exca Pradipta Syaifuddin

Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia

## Article Info

### Article history:

Received Mar 18, 2024

Revised Oct 30, 2024

Accepted Nov 14, 2024

### Keywords:

Computational efficiency

Papilledema

Retinal

Squeeze-excitation half U-Net

SMOTE

## ABSTRACT

The emergence of various convolutional neural networks (CNN) architectures indicates progress in the computer vision field. However, most of the architectures have large parameters, which tends to increase the computational cost of the training process. Additionally, imbalanced data sources are often encountered, causing the model to overfit. The aim of this study is to evaluate a new method to classify retinal fundus images from imbalanced data into the corresponding classes by using fewer parameters than the previous method. To achieve this, squeeze-excitation half U-Net (SEHUNET) architecture, a modification of half U-Net with squeeze-excite process to provide attention mechanism on each feature maps channel of the model, in combination with synthetic minority oversampling technique (SMOTE) is proposed. The test accuracy of SEHUNET is 98.52% with area under the curve of receiver operation characteristic (AUROC) of 0.999. This result outperforms the previous study that used CNN with Bayesian optimization, achieving accuracy of 95.89% and AUROC of 0.992. SEHUNET is also able to compete with the transfer learning methods used in previous research such as InceptionV3 with 96.35% accuracy, visual geometry group (VGG) with 96.8%, and ResNet with 98.63%. This performance can be achieved by SEHUNET with only 0.268 million parameters compared to the architecture parameters used in previous research ranging from 11 million to 33 million.

This is an open access article under the [CC BY-SA](#) license.



## Corresponding Author:

Wiharto

Department of Informatics, Faculty of Information Technology & Data Science, Universitas Sebelas Maret  
St. Ir Sutami No. 36, Kentingan, Jebres, Surakarta, Central Java, Indonesia

Email: wiharto@staff.uns.ac.id

## 1. INTRODUCTION

Papilledema is a serious eye condition characterized by increased intracranial pressure that can cause permanent damage to the optic nerve if not detected and treated promptly [1]. The detection of papilledema requires a visual analysis of the fundus images taken from the patient's eye. The determining attributes for the analysis are mainly consisted of the clarity of the optical disc and the blood vessels around the center of the eye. However, this detection proves to be rather difficult because similar symptoms can be seen in other conditions, such as pseudopapilledema which really mimics the real disorder [2].

The fundamental difference between papilledema and pseudopapilledema lies in the causes and effects of both. Papilledema is caused by the presence of pressure from within the eye that causes the back of the eye to be pressed into the brain, causing the patient to possibly experience blurred and double vision accompanied by headaches or nausea. On the other hand, in pseudopapilledema, there is no pressure as in the

case of papilledema, but visually, the fundus image provided has some similarities. In the fundus image of a papilledema patient, the optic disc undergoes elevation and swelling, causing the edge of the disc to become blurred accompanied by dilation of the blood vessels around the disc, which is characterized by an effect such as being pressed or paler than other blood vessels. On the other hand, pseudopapilledema has characteristics that resemble the blurring of the disc edge, but tends to be without accompanied dilation of blood vessels. This can occur due to the factor of the appearance of drusen, dots that appear around the disc due to anatomical differentiation or aging [3].

As an effort to help doctors diagnose papilledema by observing the fundus images faster so that action for the patients can be done immediately, many artificial intelligence researches have been done to develop methods and tools that can support doctors in classifying papilledema. For example, research in [4], [5] where both of them use manual feature extraction that is then processed by support vector machine (SVM) [6]. The latest research with the same purpose also done Ahn *et al.* [7] that uses convolutional neural networks (CNN) with Bayesian optimization utilizing local average color subtraction for the preprocessing step.

In addition to the standard CNN architecture [8], the derivative architecture of CNN, especially U-Net [9], is also used to classify papilledema images. For instance, a study conducted by Milea *et al.* [10] employed the one versus rest (OVR) technique on the U-Net and DenseNet architecture. There is also a study conducted by Saba *et al.* [11] that uses a combination of U-Net and DenseNet accompanied by a Gabor filter. Furthermore, there are research conducted in [12], [13] who used the same dataset for papilledema classification, where the data used came from by Kim's Eye Hospital [7]. Kokulu and Göker [12] uses MobileNetV2 with histogram equalization and 3D box filtering techniques to preprocess the input images. Meanwhile, Al-Azzawi *et al.* [13] uses ResNet-50 in combination with segmentation optimization to prepare the inputs.

The use of the U-Net architecture is not limited to papilledema classification alone. Abedalla *et al.* [14] uses the Ens4B-UNet architecture that combines four transfer learning models as part of the encoder from the U-Net architecture to perform thorax image segmentation. This shows the potential offered by U-Net in the biomedical field. However, the main challenge of using CNN derivative architectures such as U-Net and transfer learning models is the need for large computational costs due to the large number of parameters. The standard U-Net has more than 30 million parameters [15], while transfer learning models such as VGG [16], ResNet [17], and InceptionV3 [18] have more than 24 million parameters [7]. The high numbers signal the need for parameter efficiency so that computational costs and model size can be minimized. In addition to the number of parameters, another issue that arises is the probability of imbalanced data, which makes the training process difficult. Data with unbalanced classes can easily cause the model to overfit for just one class, hence making it useless for detecting other classes. This can be addressed by synthetic minority oversampling technique (SMOTE) [19], which offers a way to synthesize data from imbalanced composition. The use of SMOTE makes the data balanced, thus providing better detection results.

The use of U-Net architecture is widely used for medical image segmentation but has high architectural complexity. In research conducted by Lu *et al.* [15] proposed the use of half U-Net for medical image segmentation which is a development of U-Net, namely by using ghost module as a substitute for convolutional blocks in U-Net. Test results conducted using mammography, lung nodule, and left ventricular magnetic resonance imaging (MRI) datasets show that, when using the standard Half U-Net architecture has 0.21 million parameters, while the standard U-Net has 31.04 million parameters. These results show that half U-Net is able to reduce the number of parameters while maintaining its performance compared to the standard U-Net. CNN are built on convolutional operations, which extract informative features by combining spatial and channel information together in a local receptive field. To improve the representational power of the network, several recent approaches have shown the benefits of improved spatial encoding. In the research conducted by Hu *et al.* [20] proposed a new architectural unit, namely squeeze-and-excitation (SE), which adaptively recalibrates channel feature responses explicitly modeling the interdependencies between channels. The performance of SE has been demonstrated in a competition by Abedalla *et al.* [14], where 4 of the top 5 rankings used SE blocks combined into the ResNeXt [21] model to improve encoder performance on standard U-Net. SE has also been shown to provide improvements to U-Net on USE-Net [22].

Referring to a number of studies that have been conducted, this study proposes a papilledema classification model using a combination of Squeeze-Excitation and Half U-Net, hereinafter referred to as squeeze-excitation half U-Net (SEHUNET). The combination of SE aims to provide weight calibration for each ghost module contained in the half U-Net, so that its output has an attention mechanism for certain feature map channels. To overcome data imbalance, the SMOTE method is applied. The system model is built using a dataset from the Kim Eye Hospital. Model performance is measured using accuracy parameters, area under the curve of receiver operation characteristic (AUROC), and the number of model parameters.

## 2. METHOD

Figure 1 shows the flow summary of the method used in this study. The aforementioned flow is consisted of dataset collection, dataset preprocessing, model building, model training, and performance evaluation. These steps are done sequentially, meaning that the flow is not able to be processed further if the previous step is not done.



Figure 1. Method flowchart

### 2.1. Collecting dataset

The dataset used in this study is fundus retinal dataset from Kim's Eye Hospital used in [7] which can be accessed on Kaggle. The available classes are normal with 779 images, papilledema with 295 images, and pseudopapilledema with 295 images. The dataset was collected from patients who underwent fundus photography as part of their eye examination.

### 2.2. Preprocessing dataset

This study adopts SMOTE as a crucial technique to tackle the imbalanced class challenge. It operates by flattening the input images into a one-dimensional numerical data, and then synthesizing artificial instances for the minority class by interpolating between its closest neighbors in the feature space [19]. Thus, using the SMOTE technique is able to equalize the class distribution and enrich the information for the model to learn from.

The preprocessing step is illustrated in Figure 2 where 20% of the data will be taken from the initial data to be the testing set. The rest then will go through the SMOTE oversampling process for synthesizing, creating new similar instances on every class based on the input images, hence forming a balanced composition between each class. When the synthesizing process is done, the synthetic data will enter the horizontal flip augmentation stage which will flip the image horizontally and add it to the training process every time the data generator is called. The result of the last process will be divided into 20% as validation data and 80% as the train data for the SEHUNET model.

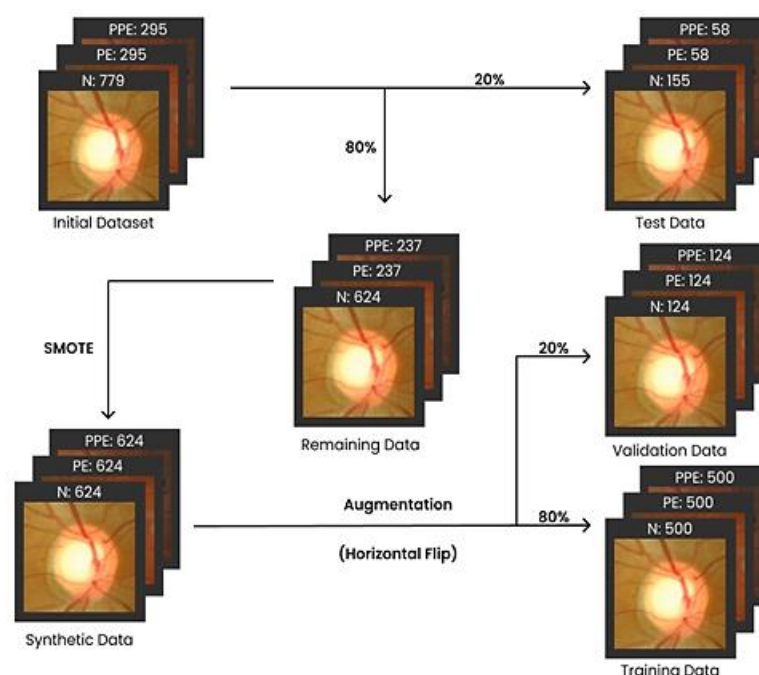


Figure 2. Dataset preprocessing (N: normal, PE: papilledema, PPE: pseudopapilledema)

### 2.3. Building model

The structure of SEHUNET architecture proposed in this study is illustrated in Figure 3. The SEHUNET architectural design is consisted of ghost module and SE block combination which forms the ghost-SE module, stacked to mimic the structure from half U-Net architecture. The ghost module [23] functions as generator to generate more feature maps from the input with lighter operations than the regular convolutional block. This makes the ghost module produces the same number of outputs compared to the regular convolutional block with much less computational cost and hence will contribute to reduce the final model size.

Half U-Net from Lu *et al.* [15] uses the ghost module to reduce computational complexity and avoid adding more parameters from the standard U-Net. In addition, half U-Net architecture standardizes the number of all channels to be combined in the decoder part of the architecture by performing full-scale feature fusion. This will further reduce the model size and the total computational cost of its training while still maintaining the performance.

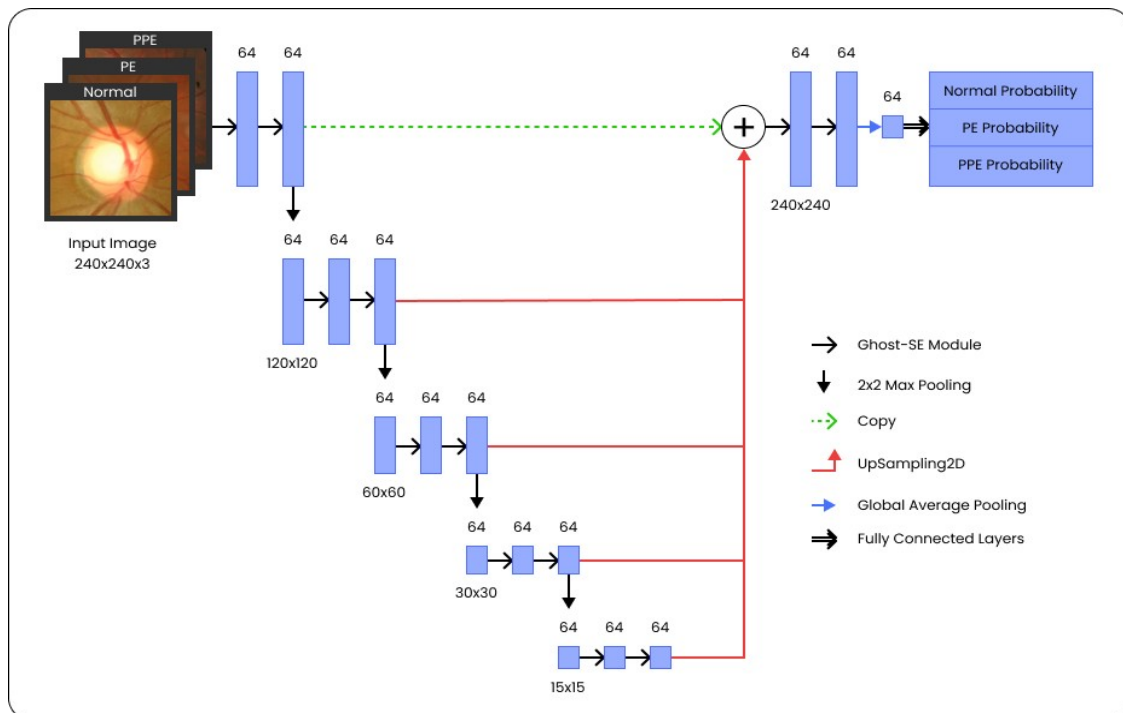


Figure 3. SEHUNET model architecture

The development carried out in this study involves the insertion of an SE block after each ghost module in half U-Net as ghost-SE module. The main purpose of the SE block addition is to provide an attention mechanism and the ability to recalibrate each feature maps by calculating their corresponding weights [20]. This enables the SE block to enhance the representational power of the feature maps by learning the interdependencies between them. The combination of ghost module and SE block allows SEHUNET to generate feature maps efficiently while having the ability to provide attention focus on each of these feature maps so that the model is able to achieve better performance than the regular half U-Net. The ghost-SE module structure can be seen in Figure 4.

### 2.4. Training model

In this study, a set of training experiments was conducted to determine which one is better between half U-Net and SEHUNET using the original data, then the best model from the previous trial will be trained using the SMOTE synthesized data. The training process was run on Google Colaboratory with the code fully written in Python programming language utilizing Tensorflow framework. It ran with GPU runtime type so that the training time can be minimized. The training used the Adam optimizer [24] with the starting learning rate value of  $10^{-3}$ . The batch size used for training was 32 with a preset of 300 epochs. The training system will reduce the learning rate to 1/10 of the previous value if the validation accuracy does not increase for 20

epochs with the minimum learning rate value being  $10^{-6}$ . In addition, the training will be stopped and the model will restore the best weights if there is no increase in validation accuracy for 50 epochs.

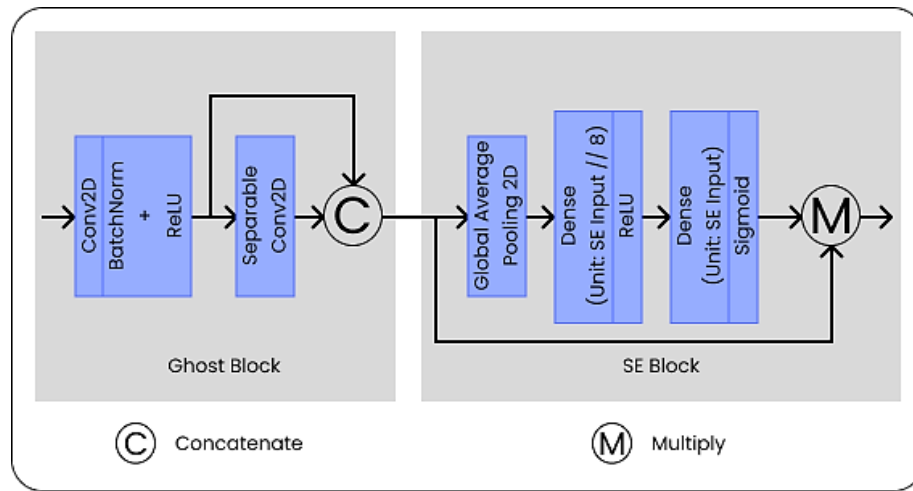


Figure 4. Ghost-SE module structure

## 2.5. Performance evaluation

The performance of the training experiments will be evaluated using testing set which consists of 20% from the original data. The main determining metric used is accuracy and followed by both categorical cross entropy (CCE) loss and AUROC [25] as the secondary metrics. The main factor of using accuracy as the primary metric involves the previous studies shown in [7], [12], [13] that collectively use accuracy as the main deterministic factor.

## 3. RESULTS AND DISCUSSION

The experiments conducted in this research will have three main steps. First, the experiment of training Half U-Net without SMOTE. Second, the experiment of training SEHUNET which also without SMOTE. Lastly, the experiment for the best architecture defined by comparison of the previous experiments, trained using SMOTE.

### 3.1. Experiments

The first experiment conducted was the training of Half U-Net using the original dataset, in the other word without SMOTE oversampling. The results of the tests conducted can be shown in Table 1. It performed well with the test accuracy score reaching 93.72%. This experiment lasted for 130 epochs, meaning that it reached its peak on epoch 80. The corresponding test loss graph is shown in Figure 5.

Table 1. Half U-Net performance without SMOTE

Metrics	Score (%)
Validation accuracy	96.78
Validation loss	0.15
Test accuracy	93.72
Test loss	0.16
AUROC	0.994

The second experiment conducted was the training of SEHUNET using original dataset without SMOTE oversampling. The result from this experiment will be compared to the result from the first experiment to determine which one performs the best on the same scenario. The results of the tests conducted can be shown in Table 2. It performed well with the test accuracy score reaching 95.94%. This experiment lasted for 159 epochs, meaning that it reached its peak on epoch 109. The corresponding test loss graph is shown in Figure 6.

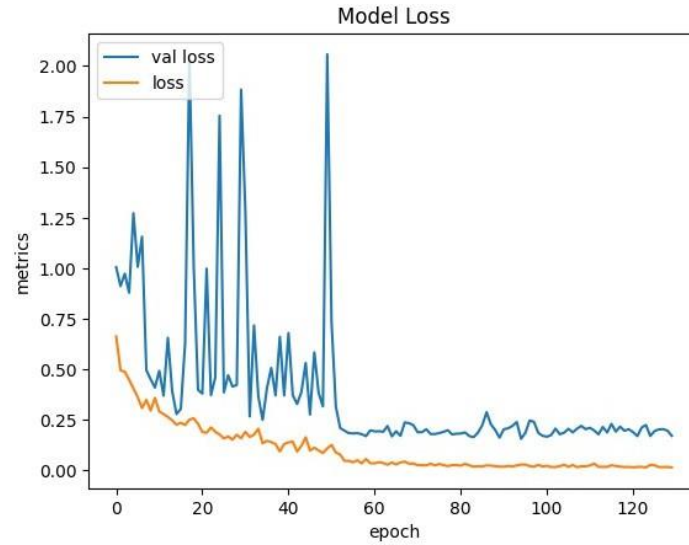


Figure 5. Half U-Net loss graph without SMOTE

Table 2. SEHUNET performance without SMOTE

Metrics	Score (%)
Validation accuracy	97.24
Validation loss	0.17
Test accuracy	95.94
Test loss	0.17
AUROC	0.993

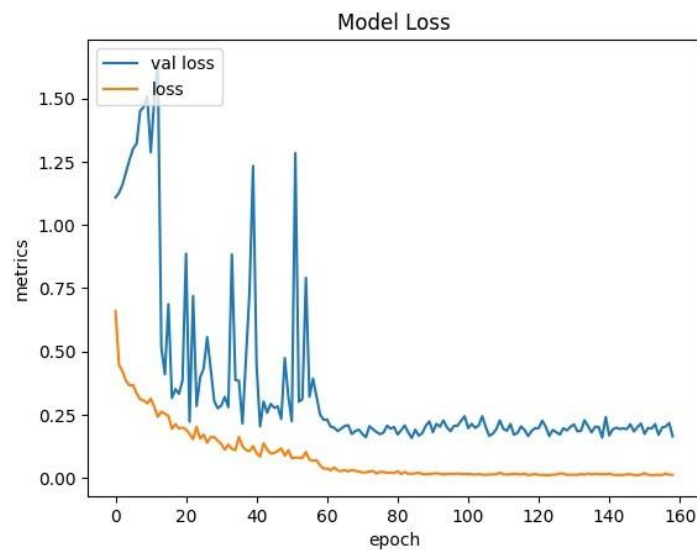


Figure 6. SEHUNET loss graph without SMOTE

From the results of two experiments above, SEHUNET model performed better than the original half U-Net model in the main determining metric collectively used in the previous studies, accuracy, for both validation and test data. Consequently, it come to the conclusion that SEHUNET has better performance than the standard half U-Net and will go through the next experiment using the SMOTE synthesized dataset. The SMOTE will create a synthetic image from the original image, in order to balance the data. The sample of the SMOTE results on a retinal image can be shown in Figure 7. If the original image is as shown in Figure 7(a), then the synthesized image is shown in Figure 7(b). The results of the SEHUNET experiment with SMOTE are shown in Table 3.

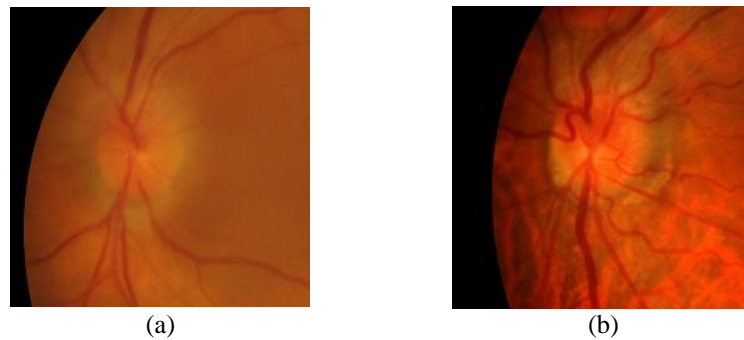


Figure 7. The sample of SMOTE result image: (a) original papilledema and (b) synthetic papilledema

Table 3. SEHUNET performance with SMOTE

Metrics	Score (%)
Validation accuracy	99.19
Validation loss	0.04
Test accuracy	98.52
Test loss	0.07
AUROC	0.997

This experiment lasted for 117 epochs, meaning that it reached its peak on epoch 67. The corresponding test loss graph is shown in Figure 8. SEHUNET with SMOTE gave the best result compared to the first two experiments, reaching the test accuracy score of 98.52%.

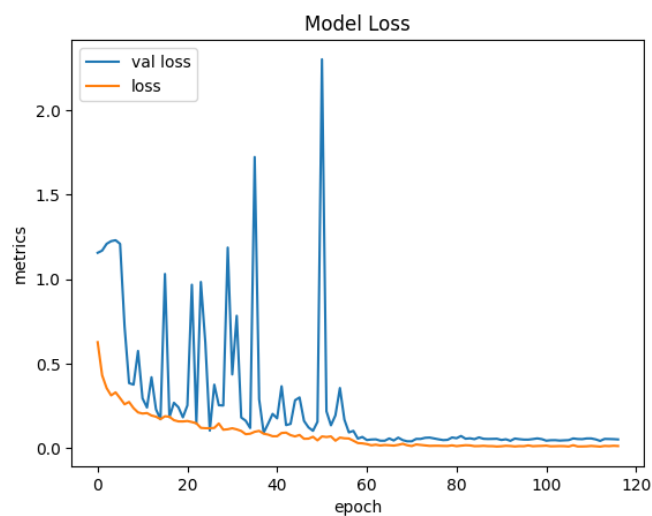


Figure 8. SEHUNET loss graph with SMOTE

### 3.2. Comparison with previous study

This study investigated the effects of SMOTE and an enhanced half U-Net architecture to image classification field in papilledema with a premise to achieve the tiniest model size and a considerable performance. While earlier studies have explored the impact of different transfer learning model on the classification performance based on Kim's Eye Hospital data, they have not addressed the parameter size minimization thoroughly. To gain some insight regarding that matter, the SEHUNET model in combination with SMOTE oversampling will be compared in terms of performance and the number of parameters with previous studies to prove that this architecture can achieve computational cost efficiency while still excelling in performance. The metrics used as evaluators are test accuracy, AUROC, and the number of parameters. The researches that will be compared to SEHUNET are research Ahn *et al.* [7] which uses the CNN Bayesian optimization, VGG, ResNet, and InceptionV3 architectures; research Kokulu and Göker [12] which uses MobileNetV2; and research Al-Azzawi *et al.* [13] which uses ResNet-50.



Table 4 shows the comparison between each study. CNN Bayesian optimization from Ahn *et al.* [7] and SEHUNET are the only architectures trained from scratch without utilizing any transfer learning model. In the study of Ahn *et al.* [7] achieved the best accuracy of 98.63% using the transfer learning ResNet model with 0.999 AUROC and 24.2 million parameters. In the study of Kokulu and Göker [12] achieved 96.83% accuracy using the transfer learning MobileNetV2 model. In the study of Al-Azzawi *et al.* [13] achieved 97.5% accuracy using the transfer learning ResNet-50 model. The proposed SEHUNET model achieved 98.52% with 0.997 AUROC and 0.268 million parameters.

Table 4. Performance comparison

Arch	Learning type	Preprocessing	Acc (%)	AUROC	Params
CNN Bayesian optim. [7]	Scratch	Local Avg color subtraction	95.89	0.992	11.6 million
VGG [7]	Transfer	Local Avg color subtraction	96.8	0.999	33.3 million
InceptionV3 [7]	Transfer	Local Avg color subtraction	96.35	0.997	24.8 million
ResNet [7]	Transfer	Local Avg color subtraction	98.63	0.999	24.2 million
MobileNetV2 [12]	Transfer	Histogram equalization 3D box filtering	96.83	-	-
ResNet-50 [13]	Transfer	Segmentation optim.	97.5	-	-
SEHUNET (proposed)	Scratch	SMOTE oversampling	98.52	0.997	0.268 million

Based on the comparison results above, it can be seen that SEHUNET becomes the second top performance architecture after almost matching the accuracy of ResNet transfer learning model from Ahn *et al.* [7] with much smaller parameters. Thus, SEHUNET successfully achieved computational cost efficiency with total percentage of parameters reduction from ResNet [7] by 98%. The performance of SEHUNET in the form of a confusion matrix [26] from the test set is shown in Figure 9.

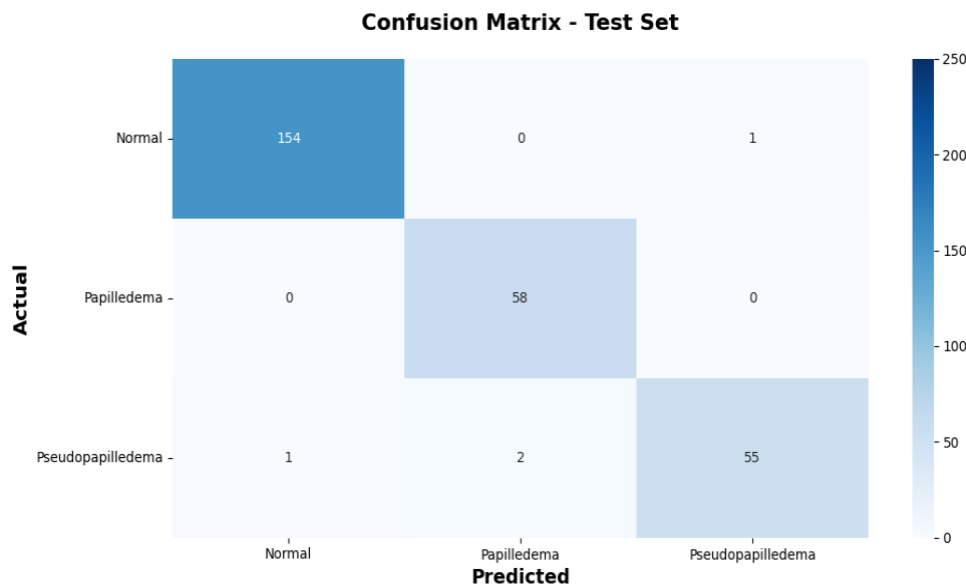


Figure 9. SEHUNET performance on confusion matrix

### 3.3. Model output

The following images contain the example of successful prediction results for each class collected in Figure 10. These images show that the SEHUNET model can accurately distinguish between normal, papilledema, and pseudopapilledema for the most cases. Figure 11 however, shows the example of failed prediction attempts by the model. The first correlated outputs are the prediction of normal image as pseudopapilledema and the opposite, which is suspected to be caused by the lack segmentation of drusen that make the disks appear to be blurry. The next prediction is a failed attempt to predict pseudopapilledema image as papilledema which is suspected to be caused by the high level of blurriness of the optic disc which very closely mimics the papilledema image from Figure 10, accompanied by the thinness of the blood vessels around the disc.



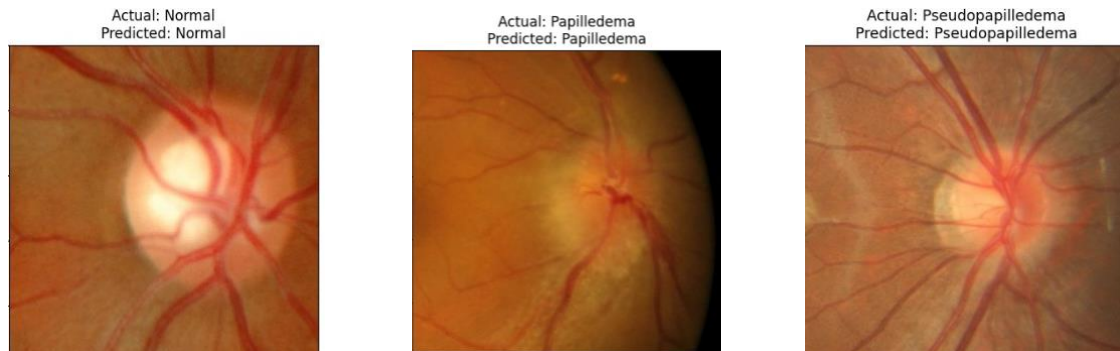


Figure 10. Correct predictions from SEHUNET

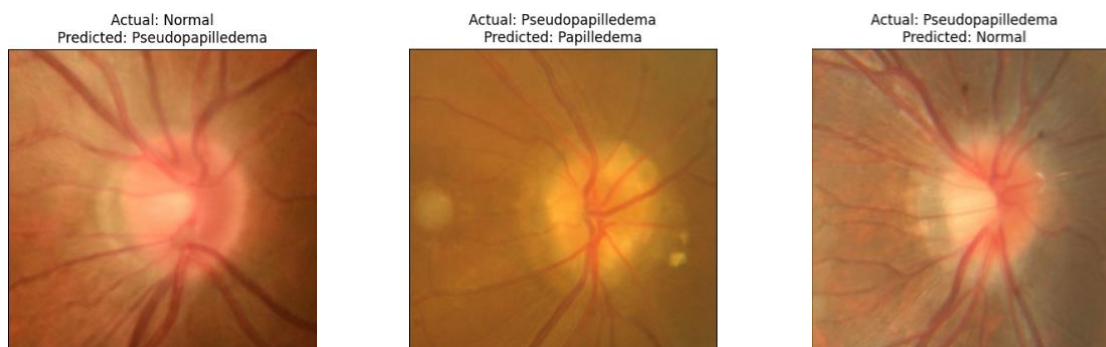


Figure 11. Failed predictions from SEHUNET

#### 4. CONCLUSION

Based on the results of this study, it can be concluded that SEHUNET is effective for papilledema image classification. The performance offered by SEHUNET matches and even exceeds the methods used previously in other studies with much lower computational cost in terms of the number of model parameters. This performance can be achieved using the combination of ghost module, half U-Net, and SE block as the main components of the proposed architecture. As for improvement suggestion in the future research, the images can be further subjected to additional augmentation such as color channel cropping and contrast limited adaptive histogram equalization (CLAHE). An additional segmentation process to help the model understand the disk better can also be applied, such as segmentation for the drusen, the blood vessels, or even the disk itself.

#### ACKNOWLEDGEMENTS

We would like to thank Universitas Sebelas Maret for providing research funding through the Hibah Riset Group scheme as stipulated in contract No. 194.2/UN27.22/PT.01.03/2024. We would also like to thank the Faculty of Information Technology and Data Science for facilitating the facilities and infrastructure that support the completion of this research.




#### REFERENCES

- [1] J. S. Xie, L. Donaldson, and E. Margolin, "Papilledema: A review of etiology, pathophysiology, diagnosis, and management," *Survey of Ophthalmology*, vol. 67, no. 4, pp. 1135–1159, 2022, doi: 10.1016/j.survophthal.2021.11.007.
- [2] J. J. Chen and M. T. Bhatti, "Papilledema," *International Ophthalmology Clinics*, vol. 59, no. 3, pp. 3–22, 2019, doi: 10.1097/IIO.0000000000000274.
- [3] G. L. Trick, S. S. Bhatt, D. Dahl, and B. Skarf, "Optic disc topography in pseudopapilledema: A comparison to pseudotumor cerebri," *Journal of Neuro-Ophthalmology*, vol. 21, no. 4, pp. 240–244, 2001, doi: 10.1097/00041327-200112000-00002.
- [4] S. Akbar, M. U. Akram, M. Sharif, A. Tariq, and U. U. Yasin, "Decision support system for detection of Papilledema through fundus retinal images," *Journal of Medical Systems*, vol. 41, no. 4, Apr. 2017, doi: 10.1007/s10916-017-0712-9.
- [5] K. N. Fatima, T. Hassan, M. U. Akram, M. Akhtar, and W. H. Butt, "Fully automated diagnosis of papilledema through robust extraction of vascular patterns and ocular pathology from fundus photographs," *Biomedical Optics Express*, vol. 8, no. 2, 2017, doi: 10.1364/boe.8.001005.
- [6] R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," *Analyst*, vol. 135, no. 2, pp. 230–267, 2010, doi: 10.1039/b918972f.




- [7] J. M. Ahn, S. Kim, K. S. Ahn, S. H. Cho, and U. S. Kim, "Accuracy of machine learning for differentiation between optic neuropathies and pseudopapilledema," *BMC Ophthalmology*, vol. 19, no. 1, 2019, doi: 10.1186/s12886-019-1184-0.
- [8] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017*, pp. 1–6, 2017, doi: 10.1109/ICEngTechnol.2017.8308186.
- [9] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *IEEE Access*, vol. 9, pp. 16591–16603, May 2015, doi: 10.1109/ACCESS.2021.3053408.
- [10] D. Milea *et al.*, "Artificial intelligence to detect Papilledema from ocular fundus photographs," *New England Journal of Medicine*, vol. 382, no. 18, pp. 1687–1695, 2020, doi: 10.1056/nejmoa1917130.
- [11] T. Saba, S. Akbar, H. Kolivand, and S. A. Bahaj, "Automatic detection of papilledema through fundus retinal images using deep learning," *Microscopy Research and Technique*, vol. 84, no. 12, pp. 3066–3077, 2021, doi: 10.1002/jemt.23865.
- [12] M. Kokulu and H. Göker, "Detection of Papilledema severity from color fundus images using transfer learning approaches," *Aksaray University Journal of Science and Engineering*, vol. 7, no. 2, pp. 53–61, 2023, doi: 10.29002/asujse.1280766.
- [13] A. Al-Azzawi, O. Bayat, S. Al-jumaili, S. Kurnaz, A. D. Duru, and O. N. Uçan, "Pseudopapilledema diagnosis using deep transfer learning approaches," *2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pp. 1–6, 2023.
- [14] A. Abedalla, M. Abdullah, M. Al-Ayyoub, and E. Benkhelifa, "Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures," *PeerJ Computer Science*, vol. 7, pp. 1–36, 2021, doi: 10.7717/peerj-cs.607.
- [15] H. Lu, Y. She, J. Tie, and S. Xu, "Half-UNet: a simplified U-Net architecture for medical image segmentation," *Frontiers in Neuroinformatics*, vol. 16, 2022, doi: 10.3389/fninf.2022.911679.
- [16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [18] C. Szegedy *et al.*, "Going deeper with convolutions," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.
- [19] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.
- [20] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, 2018, doi: 10.1109/CVPR.2018.00745.
- [21] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5987–5995, 2017, doi: 10.1109/CVPR.2017.634.
- [22] L. Rundo *et al.*, "USE-Net: incorporating squeeze-and-excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets," *Neurocomputing*, vol. 365, pp. 31–43, 2019, doi: 10.1016/j.neucom.2019.07.006.
- [23] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More features from cheap operations," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1577–1586, 2020, doi: 10.1109/CVPR42600.2020.00165.
- [24] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.
- [25] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, 1997, doi: 10.1016/S0031-3203(96)00142-2.
- [26] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, 2018, doi: 10.1016/j.aci.2018.08.003.

## BIOGRAPHIES OF AUTHORS



**Wiharto**    received obtained a bachelor's degree in electrical engineering from Universitas Telkom, Indonesia, in 1999. He obtained a master's degree in computer science from Universitas Gadjah Mada, Indonesia, in 2004 and a doctoral degree from the same University, in 2017. Currently he works as a lecturer in the Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia. His experience and areas of interest focus on artificial intelligence, computational intelligence, medical imaging, and expert systems. He can be contacted at email: wiharto@staff.uns.ac.id.



**Angga Exca Pradipta Syaifuddin**    holds Bachelor of Computer Science degree with 3.79 GPA. He received some recognition in the startup field such as fundings while still being an undergraduate. He holds two intellectual properties as the inventor of his works in the startup he founded. He can be contacted at email: excaangga@student.uns.ac.id.