HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 3 (2024)

Machine Learning-Based Biomarker for Enhanced Autism Diagnosis Using Fundus Images

Ellie Sul[1] and Ashley Um[#]

[1]Yongsan International School of Seoul, Republic of Korea
[#]Advisor

ABSTRACT

Numerous children are suffering from neurodevelopmental disorders. According to the CDC, about 1 in 36 children were diagnosed as having ASD. The seriousness of this problem is also seen through the rapid increase in the prevalence of ASD with the combined prevalence per 1,000 children increasing from 6.7 in 2000 to 27.6 in 2020. The complete cure of neurodevelopmental disorders such as ASD is the key to early diagnosis of the disorder. One of the diagnostic methods using fundus images is an effective, non-invasive method; however, they need more research on how to capture key parts of the fundus images such as the optic nerve and the blood vessels. To address this problem, this study proposes a novel advanced image classification system with the use of an adversarial attack where the adversarial attack will be used to improve the function of the system and determine which parts of the fundus images are most crucial to getting an accurate result. To carry this out, the adversarial attack was applied to three distinct parts of the fundus image–the blood vessel isolated image, optic cup isolated image, and optic disc isolated image. Through extensive experiments, it is shown that the proposed method can be utilized as a biomarker for diagnosing ASD.

Introduction

Numerous children are suffering from neurodevelopmental disorders. In fact, about one in six children in the U.S. have one or more developmental disabilities or other developmental delays (CDC 2022). The early diagnosis of neurodevelopmental disorders is very important as it is the key to increasing the percentage of being cured of the disorders. There are many methods of diagnosing children with neurodevelopmental disorders, but the method using fundus image is very effective as it is a non-invasive method that can be used to diagnose children with neurodevelopmental disorders such as Autistic Spectrum Disorder(ASD). And because this is a newly discovered method, there is a need for improvement with more research.

There were a number of previous studies conducted on the diagnosis of various neurodevelopmental disorders using fundus images. For example, the research proposed by Diaz et al. found that fundus images can be used to diagnose Parkinson's disease (Diaz et al. 2020). Another example is the diagnosis of ASD using fundus images (Kim et al. 2023) that was discovered in the research proposed by Kim et al.. Although these researches brought very important findings, they could be improved through more research on how to better capture the important parts of a fundus image such as the optic nerve and blood vessels.

To address this problem, I proposed a more advanced image classification system using the adversarial attack method. This system, specifically the adversarial attack method will be used to improve the system's function and determine which parts of the fundus images, out of three distinct parts (the blood vessel isolated image, optic disc isolated image, and the entire image), plays the most important role in giving accurate results.

HIGH SCHOOL EDITION
Journal of Student Research

Related Work

Retinal Image

A fundus image is a medical photograph that captures a detailed view of the back of the eye, including the retina, optic nerve, and blood vessels. This image is often used in ophthalmology for diagnostic and monitoring purposes. Fundus images are filmed using a small, convenient device, and it could even be filmed using a smartphone. Fundus images are non-invasive ways to capture the blood vessels closest to the brain. Observing the previous research about fundus images, they are found to be useful in checking the presence of numerous diseases or disorders.

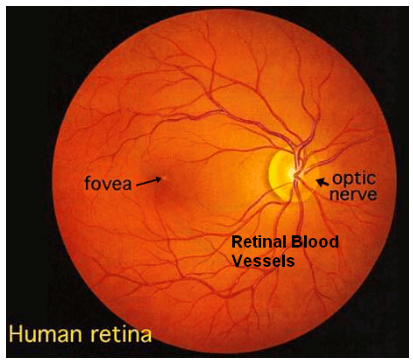


Figure 1. Fundus image (Eyecarepro 2021)

As for these advantages of fundus images, numerous research studies have been conducted to diagnose a variety of diseases including cardiovascular disease, Parkinson's, and even neurodevelopmental disorders such as ASD.

Diaz et al. proposed research where they studied the diagnosis of Parkinson's disease using a machine learning-based system that takes fundus images as the input and the presence of the disease as the output (Diaz et al. 2020). They compared the shapes of the retinal blood vessels of people with and without Parkinson's disease to create a system to diagnose Parkinson's disease. Kim et al. studied the diagnosis of ASD using a fundus image to diagnose the presence of ASD (Kim et al. 2023). The researchers studied if the shapes of the retinal blood vessels could diagnose the presence of ASD and created a system that could diagnose ASD using fundus images.

Motivated by these studies, I propose a novel machine learning-based neurodevelopmental disorder diagnosis system in this research. A comprehensive description of the proposed system is provided in Chapter 3.

Image Classification

An image classification is the process of the machine learning systems reading the pixels of an image, processing those numbers into a score, multiplying it to weight value and adding bias value, and using cross-entropy loss function to check if the system categorized the image correctly. The inputs for an image classification system are digital images, and the outputs are the images getting categorized into a fixed category.

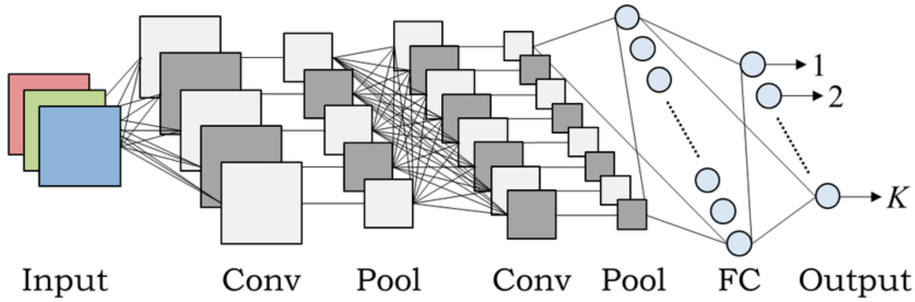


Figure 2. Architecture of convolutional neural network (Hidaka et al. 2017)

Generally, an image classifier using Convolutional Neural Network(CNN) is used for image classification processes. CNNs have proven highly effective in many computer vision tasks such as image classification. The CNN extracts the feature maps from the input images using convolutional operations. Then, the CNN subsamples those image maps for a certain number of times and flattens them to get a linear layer output.

In this study, the CNN image classifier will be used to classify the presence of a neurodevelopmental disorder. The proposed system will take a fundus image as its input to have an output of those images getting categorized into the categories of whether or not the disorder is present.

Proposed Method

Overall Architecture

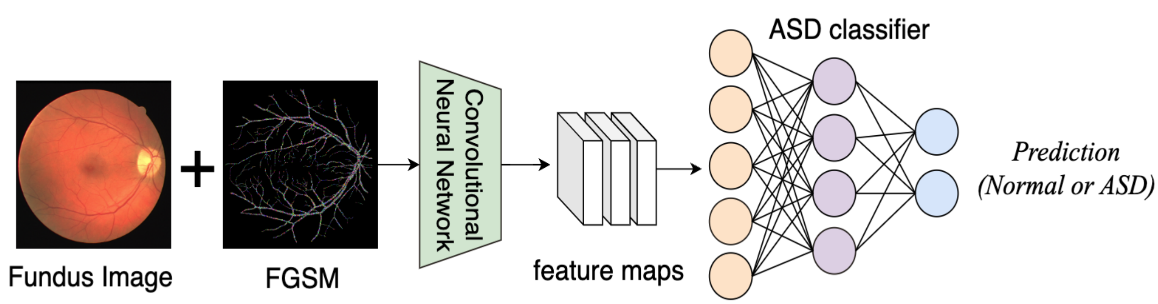


Figure 3. The architecture of the proposed ASD diagnosis network.

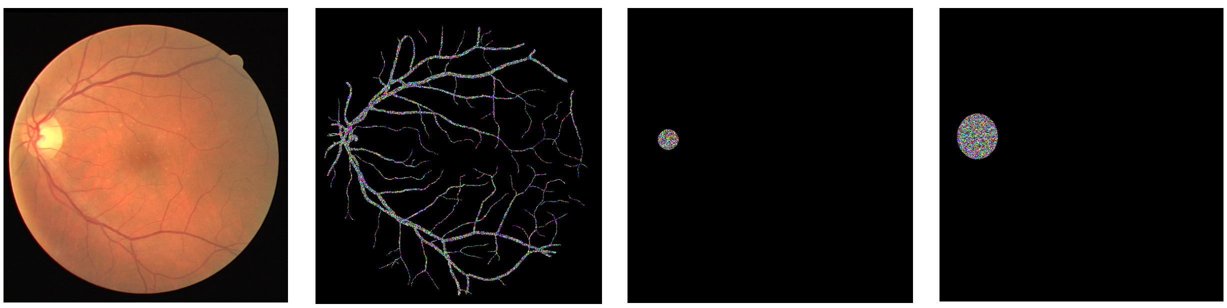
Figure 3 illustrates the architecture of the proposed ASD diagnosis network. The proposed system comprises three main components: adversarial noise generation, convolutional neural network, and ASD classifier.

The adversarial noise generation component creates perturbations within input images, strategically manipulating them to disrupt accurate predictions. These perturbations are designed to induce misclassification or produce incorrect outputs from the machine learning model. The effectiveness of the proposed adversarial noise approach will be detailed in Chapter 3.2. In this study, the Fast Gradient Sign Method (FGSM) algorithm (Goodfellow et al. 2014) was employed to generate adversarial noise which is a widely used technique in the adversarial attack task.

Adversarial noise is added to the input image before it is passed through the convolutional neural network to generate feature maps. These feature maps contain significant visual characteristics relevant to ASD. Finally, the ASD classifier analyzes these feature maps to predict whether the input fundus image belongs to an individual with ASD or to a normal case.

Adversarial Noise

Adversarial noise is the factor that decreases the performance of the machine learning system as noise is added to the image. In this study, this adversarial attack will be used to advance the performance of the machine learning by training it with images that have noise added to it that will increase its loss function, creating an image that is much harder for a machine learning to classify. This can be expressed with the function: I(normal fundus image) + N(noise) = I^N(fundus image with noise added). The noise will be added to different parts of the fundus image to observe which parts of the fundus image have the most correlation with the effective classification of ASD. This noise will each be added to the blood vessel isolated image(I_v^N), optic disc isolated image(I_d^N), and optic cup isolated image(I_c^N). As mentioned, there will be four types of noises(N): the one added to the blood vessel isolated image(N_v), the one added to the optic disc isolated image(N_d), and the one added to the optic cup isolated image(N_c). The noise will then be added by each of these images, getting a noise-added value for the parts that are being observed, and getting a value of 0 for the background.



(a) (b) (c) (d)

Figure 4. Different types of FGSM (Goodfellow et al. 2014) noise map
(a): Fundus image, (b): retinal vessel, (c): optic cup area, and (d): optic disc area

Figure 4 displays three types of adversarial noise used in this research. The experimental outcomes of applying each type of noise will be elaborated upon in Chapter 4.

Convolutional Neural Network and ASD Classifier

For the convolutional neural network, I employed state-of-the-art architectures known for their high performance in image classification tasks, including VGG-19 (Simonyan et al., 2014), MobileNetV2 (Sandler et al., 2018), ResNeXt-50 (Xie et al., 2017), SE-ResNet-50 (Hu et al., 2018), ResNet-50 (He et al., 2016), and DenseNet-121 (Huang et al., 2017). Among these architectures, ResNet-50 demonstrated the most accurate results. Further elaboration on this selection will be provided in Chapter 4.

To implement the ASD classifier, I constructed a two-layered neural network with a hidden size of 512 and utilized the rectified linear unit activation function. While experimenting with deeper neural network architectures, such as a three-layered neural network, no improvement in accuracy was observed.

Experimental Results

Dataset

To train and evaluate the proposed system, I looked for a fundus dataset from a government-supported large-scale dataset platform called AI Hub (AI Hub 2024) with publicly available datasets. This platform has a dataset of the fundus images of children with ages ranging from under 7 to 21. The proportion of the various age groups shows that there are 27.7% of children under 7, 43.57% of children from 7 to under 13, and 28.73% of children from 13 to under 21. This dataset contains 57,195 samples, giving a large dataset to work with to understand the fundus images with labels of normal and ASD. In the dataset, there are 64.94% normal children and 35.06% with ASD.

Figure 5. Sample fundus image (AI Hub 2024)

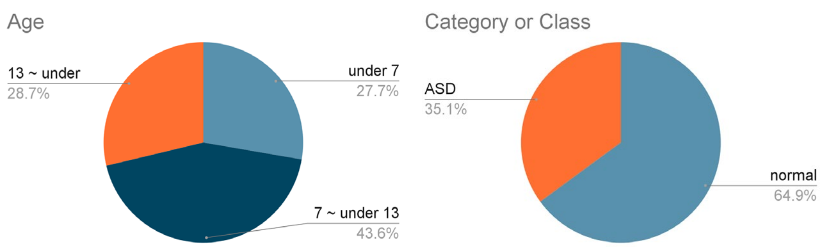


Figure 6. Age and category distribution of the dataset

Inference Metrics

In classification, accuracy, recall, precision, and F1-Score are metrics that are used to evaluate how effective the trained model is. Accuracy measures the proportion of correct predictions made by the model by finding the ratio of the number of correct predictions to the total number of predictions made by the model. Recall measures the proportion of correct predictions of how many true positive samples are found upon the total actual positive samples. Precision measures the proportion of actual positive samples from the predicted positive samples determined by the model. Lastly, the F1-Score measures how stable the model is by finding the harmonic mean of recall and precision.

Comparison with State-Of-The-Art Methods

Table 1. Evaluation comparison with various convolutional neural network architectures.

	Accuracy	**Recall**	**Precision**	**F1-Score**
VGG-19 (Simonyan et al. 2014)	0.9104	0.8416	0.6702	0.7462
MobileNetV2 (Sandler et al. 2018)	0.9176	0.8386	0.6727	0.7465

ResNeXt-50 (Xie et al. 2017)	0.9491	0.8608	0.6891	0.7654
SE-ResNet-50 (Hu et al. 2018)	0.9501	0.8617	0.6948	0.7693
ResNet-50 (He et al. 2016)	0.9523	0.8615	0.6926	0.7679
DenseNet-121 (Huang et al. 2017)	0.9615	0.8722	0.7232	0.7907
Proposed Method	**0.9824**	**0.9018**	**0.7473**	**0.8173**

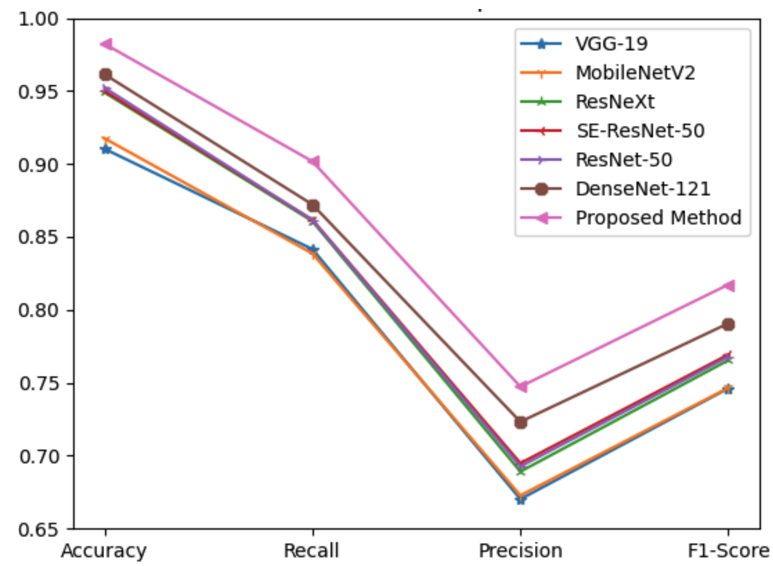


Figure 7. Evaluation comparison with various convolutional neural network architectures.

In Table 1, the accuracy, recall, precision, and F1-Score for 6 different convolutional neural networks and the proposed method. Looking at the figure, one can easily notice that the performance of the proposed method is much better compared to the other networks as it has significantly higher values of accuracy, recall, precision, and F1-Score. Because the accuracy is much higher than 91%, which is the cut-off for whether or not the network could be used in real-life diagnosis process, it could be used as an aid for psychiatrists in the diagnostic process of ASD although it is important to consider that the precision value is quite low, signifying that it may predict normal patients with ASD.

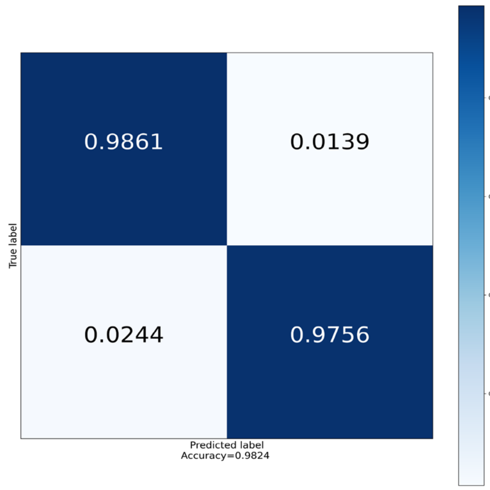


Figure 8. Confusion matrix of the proposed approach.

In Figure 8, the confusion matrix shows the ratios of true positive, people who have ASD and are classified as ASD, to false positive, people who don't have ASD but are classified as ASD, and of true negative, people who don't have ASD and are classified as normal, to false negative, people who have ASD but are classified as normal. In the confusion matrix of the proposed approach, the values of true positive and negative are very close to 1.0, signifying how effective the proposed approach is.

Table 2. Result of an ablation study

	Baseline (Accuracy)	**Retina Vessel (Accuracy)**	**Optic Cup (Accuracy)**	**Optic Disc (Accuracy)**
ResNeXt-50	0.9491	0.9553	0.9557	**0.9692**
SE-ResNet-50	0.9501	0.9558	0.9569	**0.9731**
ResNet-50	0.9523	0.9574	0.9751	**0.9824**
DenseNet-121	0.9615	0.9662	0.9670	**0.9866**

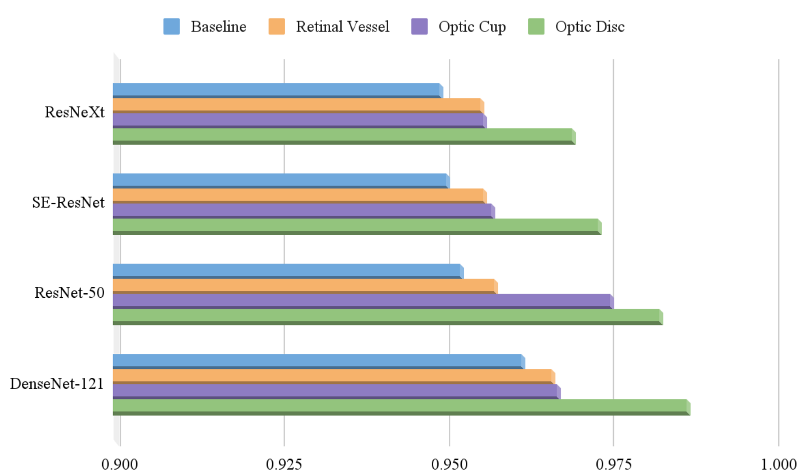


Figure 9. Result of an ablation study.

In Table 2 and Figure 9, the results of the ablation study are shown. Comparing the accuracy of each convolutional neural network to the baseline results, all of the networks trained with images that contain noise (Retina Vessel, Optic Cup, and Optic Disc) showed higher accuracy than the baseline network without any training for images that contain noise. This clearly shows that the accuracy has shown an increasing trend in the presence of the adversarial attack technique. Also, it is shown that the network trained with noise in the optic disc of the fundus images showed the highest accuracy among networks trained with noise in the retina vessel, optic cup, and optic disc. This indicates that the optic disc plays the most significant role in diagnosing ASD using fundus images.

Conclusion

In this study, I investigated ways to more effectively diagnose neurodevelopmental disorders such as ASD using fundus images. Therefore, I proposed an image classification system using CNN that utilizes the adversarial attack method to create a system that could accurately categorize the fundus images even in the presence of noise in the image. The evaluation of the proposed method compared to six different CNN systems and the confusion matrix clearly indicated that the proposed method was able to classify the fundus images more effectively. In addition, through the evaluation of the different kinds of noise added to the fundus images, it was found that the CNN systems trained with noise added to the optic disc isolated image had the highest accuracy in its classification. These results suggest that health professionals could use the proposed method for a more accurate diagnosis of neurodevelopmental disorders using fundus images. However, if it will be used in the diagnostic process, it is important for health professionals to carefully consider the presence of a disease since the precision value is quite low, signifying that it may predict normal patients with ASD. In conclusion, this non-invasive, effective method of diagnosing neurodevelopmental disorders such as ASD should be further developed and used for the early diagnosis of young children to put a stop to the rapid increase in ASD prevalence rate among children.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

CDC. (2022, Apr). "*Developmental disabilities*":Centers for Disease Control and Prevention Preventio https://www.cdc.gov/ncbddd/developmentaldisabilities/index.html#:~:text=Developmental%20disabilities%20are%20a%20group,disabilities%20or%20other%20developmental%20delays

Diaz, M., Tian, J., & Fang, R. (2020, December). Machine Learning for Parkinson's Disease Diagnosis Using Fundus Eye Images. In Annual Meeting of Radiology Society of North America (RSNA).

Eyecarepro (2021, Oct). "*retina-photography*": Eyecarepro https://www.eyeqdr.com/eye-care-services/management-of-ocular-diseases/retina-photography/

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572.

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 (pp. 770-778).
https://doi.org/10.48550/arXiv.1512.03385

Hidaka, A., & Kurita, T. (2017, May). Consecutive dimensionality reduction by canonical correlation analysis for visualization of convolutional neural networks. In Proceedings of the ISCIE international symposium on stochastic systems theory and its applications (Vol. 2017, pp. 160-167). The ISCIE Symposium on Stochastic Systems Theory and Its Applications.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708). https://doi.org/10.48550/arXiv.1608.06993

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7132-7141). https://doi.org/10.48550/arXiv.1709.01507

Kim, J. H., Hong, J., Choi, H., Kang, H. G., Yoon, S., Hwang, J. Y., ... & Cheon, K. A. (2023). Development of Deep Ensembles to Screen for Autism and Symptom Severity Using Retinal Photographs. JAMA Network Open, 6(12), e2347692-e2347692.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4510-4520). https://doi.org/10.48550/arXiv.1801.04381

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1492-1500).
https://doi.org/10.48550/arXiv.1611.05431