Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

BLINDNESS DETECTION USING MACHINE LEARNING APPROACHES

MUHAMMAD JAVAID IQBAL
Department of Computer Science and Information Technology, The Superior University Lahore, Pakistan.
Email: javaid.iqbal@superior.edu.pk

MUHAMMAD USMAN NASIR*
Department of Computer Science, COMSATS University, Islamabad, Pakistan.
Email: usmannasir@cuisahiwal.edu.pk

MUHAMMAD UMER
Department of Computer Science, COMSATS University, Islamabad, Pakistan.
Email: mumar@cuisahiwal.edu.pk

MUHAMMAD WASEEM IQBAL
Ph.D., Associate Professor Department of Software Engineering, Superior University, Lahor,e Pakistan.
Email: Waseem.iqbal@superior.edu.pk

ARFAN JAFFAR
Ph.D. Department of Computer Science, Superior University, Lahore, Pakistan.
Email: ARFAN.JAFFAR@ superior.edu.pk

ALI ASIF
Department of Computer Science, Superior University, Lahore 54000, Pakistan.
Email: a.asif83@gmail.com

Abstract:

Vision impairment and blindness are chronic diseases where blindness is a complete or partial loss of vision. Blindness occurs suddenly or over a period. The primary reasons for blindness occurrence are diabetes and secondly eye diseases. Older people in developing countries are more affected than other age people. The major problem with blindness is no proper guidelines or precautions for the people. According to the World Health Organization (WHO), 2.2 Billion people are suffering from near or distance impairment. Due to age, the leading causes are uncorrected refractive errors and cataracts. Diabetes patients mostly face vision problems due to diabetic retinopathy. The high blood glucose level in the eye blood vessels increases the chances of vision problems. We proposed the model using deep learning algorithms to detect blindness in the early stages. We apply pre-processing approaches to manage the dataset. Then apply ResNet, DenseNet, Xception, and InceptionResNet models to train the model. The trained model was used for the testing and evaluate the proposed model using accuracy, precision, recall, and f1-score. The proposed model outperformed using the ResNet model compared to the other models. This model can be utilized for clinical purposes after testing on different datasets. The proposed model evaluated for accuracy, precision, recall, and f-measure are 0.93, 0.94, 0.98, and 0.94 respectively. The results show proposed model outperforms blindness detection.

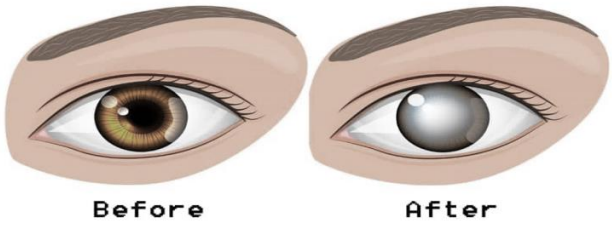
Keywords: Blindness Detection, Machine Learning, ResNet, Proliferative Diabetic Retinopathy, Transfer Learning.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

1. INTRODUCTION

Blindness is the complete or partial loss of vision in the human eye. Blindness has many causes, such that it may occur due to traffic accidents, diabetes, and also due to common eye diseases. Blindness occurs suddenly or over a period. The primary reasons for blindness are due to diabetes and eye diseases. Typically, blindness has no pain when it happens in the eyes. When blindness suddenly occurs, it needs emergency treatment to diagnose the actual reason for this infection. Sometimes eye infections occur due to the worst weather or any other accidents. Blindness has three major types: color blindness, night blindness, and snow blindness. It may cure using different approaches, but few of them are curable. Blindness impacts a significant effect on human life. Due to this disease, he cannot read, cook, or work properly. Sometimes people use different glasses, which may also impact the human eyes and lead to blindness [1]. Figure 1 shows the human eye before the blindness and when blindness occurred. We can see the eye retina is covered with white color material, which impacts the eyesight. Due to this white color layer, the human eye becomes blind, and humans can't see entirely or partially because it depends upon the intense impact on the eye.

Figure 1: Eye before blindness and after [2]



Blindness has more causes, such as age-developed countries. Due to age, the leading causes are uncorrected refractive errors and cataracts. Both exist due to the age factor in humans, and most aged persons face blindness problems. Diabetes patients mostly face vision problems due to diabetic retinopathy. The high blood glucose level in the eye blood vessels increases the chances of vision problems. If someone feels blurry vision or cannot see things correctly, he should contact the optician. Blindness affects almost all the world in different ratios, but according to the World Health Organization (WHO), 2.2 Billion people suffer from near or distance impairment [3]. Blindness and vision impairment is caused to all age people but most people above 50 years suffer a lot. Blindness and vision impairment are classified into two groups. The first is distance vision impairment which has four types of blindness, such as mild visual acuity worse than 6/12 to 6/18, and the second type of distance impairment is moderate, which has visual acuity worse than 6/18 to 6/60. The third type is if the distance impairment is severe, which has visual acuity worse than 6/60 to 3/60, and the last is blindness which has visual acuity worse than 3/60. Near vision, impairment has only one type, and their near visual acuity is worse than N6 or M.08 at 40cm. Some factors impact every person to their impairment. For example, a person experiences a problem with an accident, building, transport, and information [4].

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

Globally, significant causes of vision impairment or blindness are uncorrected refractive errors, contracted age, diabetic retinopathy, corneal capacity, and trachoma. Vision impairment impact individual, community, and the economy. All these effects are due to vision impairment. Some strategies are implemented to avoid vision impairment and blindness problems. Each eye disease needs early prevention so maximum people can save themselves such that disease. If a person is affected, then the best way is to diagnose early and suggest proper treatment [5].

1.2 Research Objectives

The research objectives are to give the optimum solution for blindness detection. A few main objectives were identified, which are discussed below.

- To provide the solution for early detection of vision impairment and blindness.
- To analyze the pre-processing techniques like removing noisy data.
- To extract the necessary features for a blindness types classification.
- To develop an intelligent algorithm to classify blindness types.
- To explore the feasibility of the solution using deep learning approaches.

1.3 Problem Statement

Vision impairment and blindness are the most spreading disease all over the world. We need to detect blindness on an early basis. People don't get treatment in developing countries and suffer from this type of disease. Using basic precautions, we can save people from blindness, and there are need to give the solution for early detection and early precautions to save themselves [11]. This is a real-world problem where people face such vision problems with their minor negligence or lack of resources.

1.4 Main Contribution

We observe that vision impairment and blindness are significant problems in developing countries, specifically in people aged. We analyze some communities' scientific research productivity and international collaborations because we do not find any comprehensive study conducted to measure and identify using only local hospital data. Open-source data under machine learning techniques to promote research and international collaboration among people. The primary purpose is to measure this viral and rapid spread of vision impairment and blindness at an early stage to control and overcome this disease. Every second million of people suffer from this disease. Data is available in open source, so we need to give the solution for early detection of blindness. We have to use this data positively and classify it at an early stage before it grows to the final stages. To calculate this, we use machine learning and artificial intelligence methods.

The remaining portion of this report is divided into the literature review, data design, results, and discussion. Finally, we discuss the conclusion of the dissertation.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

2. RELATED WORK

This chapter deeply elaborates on the existing approaches to blindness prediction. Vision impairment and blindness with their types of identification need time, specifically for people who don't have many resources. Some already applied approaches are discussed in this chapter in detail. Approaches are based on machine learning algorithms, deep learning, and image processing. We explore these approaches which help people in blindness detection and the early precautions to avoid this alarming situation.

Diabetic Retinopathy (DR) is a chronic disease that impacts vision loss and blindness. This research study [6] highlights the problem of early detection of blindness loss, so the treatment can be carried out to save people's vision. They suggest that the early detection of blindness plays a vital role in adequately treating patients. They explored that various machine learning approaches were applied to give the solution timely, but there is still a gap for the accurate valued system. Where this research model suggests the deep learning model using the open-source dataset UCI machine learning repository for blindness detection. In this approach, they used pre-processing technique. Principal Component Analysis (PCA) approach is used to extract the most valuable features to train on the open-source dataset. The given feature vector pass to the Deep Neural Networks for training and then tests the model using an unseen dataset. The proposed model was evaluated using different measures such as precision, recall, and f1. After evaluation, the system's performance was good, which can be used in the real-time solution for early detection. This approach used the traditional deep learning approach, which should be updated according to the new features in the dataset so the system can give a more generalized solution.

Thippa et al. [7] discussed that diabetic retinopathy is the primary cause of blindness and vision loss. Older people are affecting more than the child or youngsters worldwide. Mostly already approaches using fungal retinal imaging for blindness detection. This study discussed the principal component analysis-based deep neural network learning approaches to detect blindness in images. Grey Wolf Optimal (GWO) algorithm is used for the detection of blindness and classification purposes. The DNN model was updated to optimize its parameters and use in the proposed approach to applying the detection process for blindness. This paper used a standardized optimization approach to reduce the feature using principal component analysis and then choose optimal hyperparameters in the GWO algorithm. DNN algorithm was used to train the model and this trained model was used for testing purposes. Using the open-source dataset and the evaluation measures, the proposed devaluated calculate their values. This model was compared with traditional machine learning algorithms such as SVM, Decision Tree, and Naïve Bayes. The statistical results show the performance of the proposed system, which was reasonable compared to the traditional machine learning algorithms.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

Diabetic retinopathy is the primary cause of the spreading of blindness and vision globally. A lot of people are affected due to not being aware of diabetes, accidents, and transport. There is a lack of early prevention and detection for blindness. This paper [8] elaborates on these problems in detail and suggests a solution for detecting diabetic retinopathy in CT-scan images. The deep learning algorithm VGG-16 was used for the detection of blindness in the CT images of the eye. They proposed fully automated optical coherence tomography (OCT), which can detect blindness in CT images. A large volume dataset was used for training the model and the trained model was evaluated using the unseen dataset. The evaluation measures calculate the values that prove the proposed model's performance. The accuracy value was 98% which gives the maximum performance on this system.

Balyen et al. [9] proposed an approach to deal with blindness detection. Vision impairment and blindness are the most spreading disease all over the world. There was a need to detect blindness on an early basis. They suggest that the early detection of blindness plays a vital role in adequately treating patients. They explored that various machine learning approaches were applied to give the solution timely, but there is still a gap in the accurate valued system [12]. Deep learning and machine learning approaches are suggested in this research work. Where a large volume dataset was used to train the model and the procedure for this implementation was discussed in detail. Optical Coherence Tomography (OCT) was used with the eye's CT-Scan images and the deep learning algorithms to train the models. The proposed system was evaluated using different evaluation measures. The proposed system outperformed the training and testing data. In the above table, we have summed up the condition of the artistry overview of papers. Everyone has his reasoning, information, and examination measure [13]. In particular papers, we couldn't discover all data we referenced above in various sections. So we need to go through hours to gather this information. Some have worked so great we may discover everything in overseen and coordinated way; few have missing segments in papers. We have referenced all online media, author names, procedures, and highlights used to investigate information distribution year meeting, diary, and assessment measures utilized.

3. DATASET

This section discusses the dataset for the proposed model. The first step is to choose a more accurate and authentic data source. As we all know that some fake datasets people are uploaded for different purposes online. There is a vast quantity of data available on UCI machine learning. As each dataset has its format and structure [9]. Our research found that progressively exact and bonafide information APTOS 2019 Blindness Detection dataset [10]. This dataset is open source and available on the Kaggle website.

This dataset is the large volume of retina images taken through fundus photography. And these images were taken under imaging conditions. After collecting the required images, the clinician rated all the images manually [14-16]. This was based on diabetic

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

retinopathy with scale values of 0 to 4. The scale values were 0 for the No Diabetic Retinopathy, 1 for Mild, 2 for Moderate, 3 for Severe, and the last 4 was Proliferative DR. available dataset has different files such as the train images folder and test images folder [17-19].

So we pass this dataset to the preprocessor to make it understandable and remove noisy data. Data pre-processing is self-explanatory terminology, a data analysis technique to transform raw data into an understandable or readable format. Real-world data is inconsistent primarily, incomplete, containing noise and unnecessary information. It is a well-known technique to resolve such kinds of issues. It prepares such kinds of data for additional processing [20-21].

4. PROPOSED METHODOLOGY

Various approaches are discussed in the literature review section, where people used machine learning and deep learning approaches on the CT-scan images. We proposed the system by using deep learning algorithms for fake blindness detection. The detailed approach we used for blindness detection is discussed below in the following points.

- Exploratory Data Analysis (EDA)
- Loading the open-source dataset for processing
- Data visualization and EDA process
- Analysis of the train dataset labels
- Image Data visualization of the different classes
- Greyscale image analysis
- Image Cropping
- Data Pre-Processing
- Image Data Generator
- Deep Learning Models
- Models Training and Evaluation

5. Experiments and Results

After applying the proposed methodology for blindness detection, we did experiments to get the optimum solution for the proposed system. We apply all the mentioned points in the above-proposed methodology and the results obtained after each point are discussed in detail. As we discussed, our first point is exploratory data analysis (EDA). We analyzed the whole dataset to check its images and the distribution against the label images in training. Figure 1 shows below bar chart, which clearly shows that the data set is quite imbalanced. And even it’s expected in the medical domain. Both training and

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

testing datasets are not too large. The training dataset is about three times greater than the testing dataset.

Figure 1: Data visualization and EDA

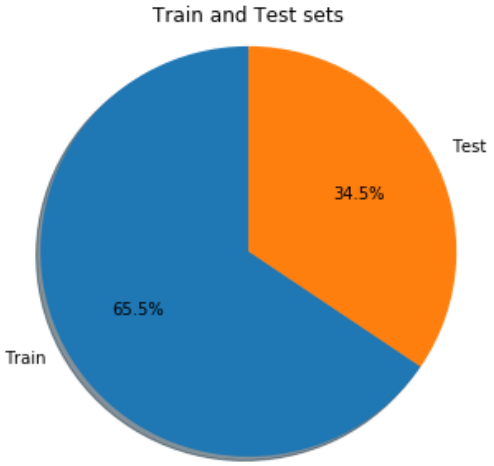


Figure 2 shows the diabetes stages percentages: NO DR, Mild, Moderate, Severe, and Proliferative DR. Diabetic retinopathy labels with their percentage of values of the images available in the dataset are shown here. No DR has greater values than others, and this is an unbalanced dataset.

Figure 2: Diabetic retinopathy condition labels

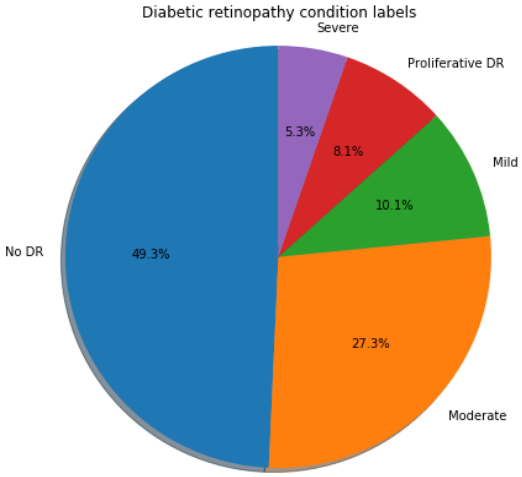
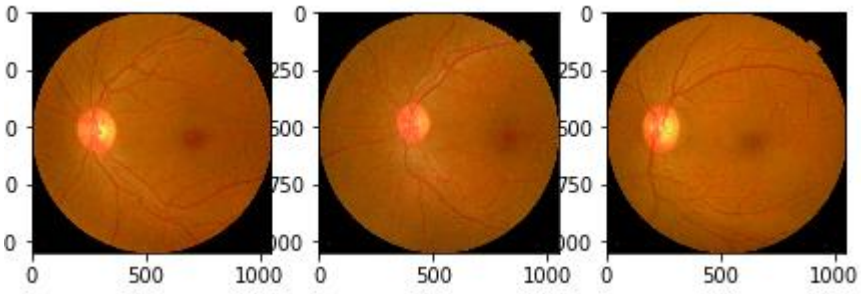


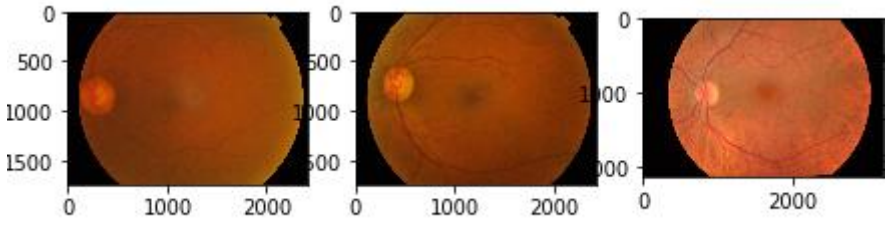
Figure 3 shows the visuals for some images from the training labeled dataset. These images are from different classes: NO DR, Mild, Moderate, Severe, and Proliferative DR.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

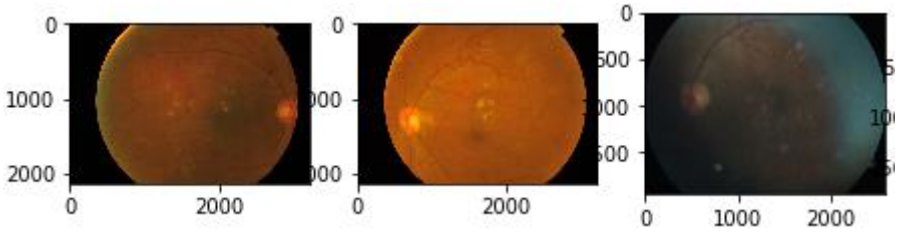
Figure 3: Visual for some images for different classes (a) No DR, (b) Mild, (c) Moderate, (d) Severe, (e) Proliferative DR



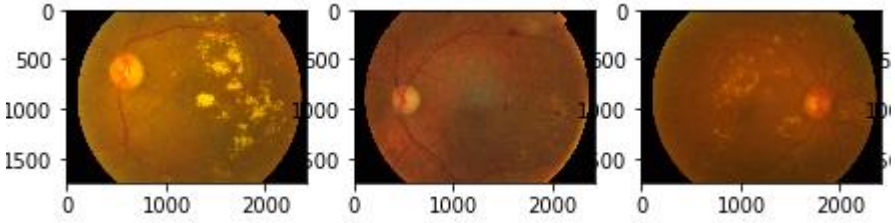
(a) No DR



(b) Mild

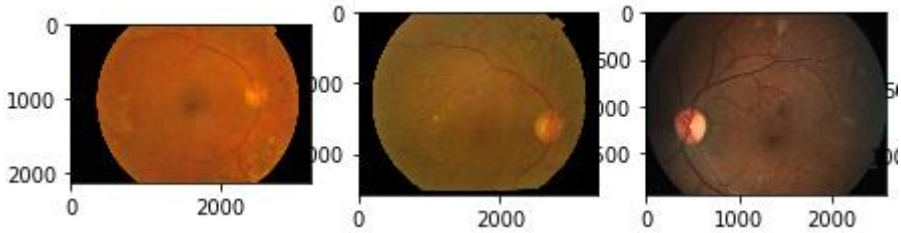


(c) Moderate



(d) Severe

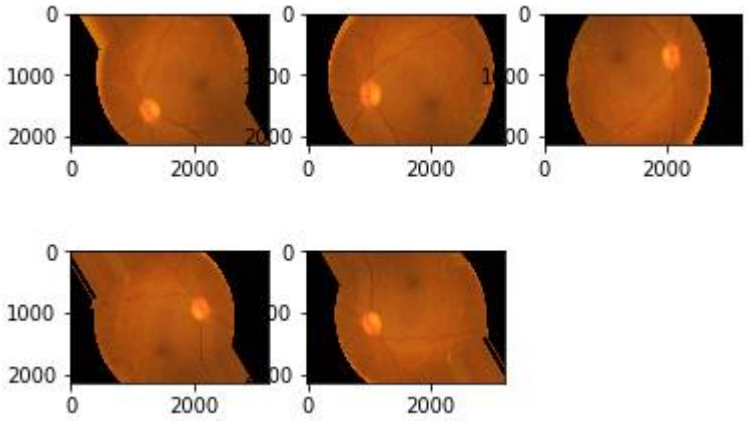
Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU



(e) Proliferative DR

Here we use the Keras ImageDataGenerator class to generate data for the Keras model. It is used for data generation, increasing the data size. With the help of ImageDataGenerator, we did image “augment” via several random transformations so that our model would never see twice the same picture. Training Deep Learning model can perform better with more data, and augmentation technique can create data variations that can increase the ability to fit the model to the gene. The results of the image augmentation are shown in figure 4.

Figure 4: Results of the augmented images



We applied deep learning models ResNet, DenseNet, Xception, and InceptionResNet. Table 1 shows the obtained results for each algorithm separately. First, we train each model separately to pass the same dataset, and the trained model is used to test the unseen images. Evaluation measures are used to calculate the statistical values for each model. Accuracy, precision, recall, and f1-score for each model were calculated [22]. All the model’s performances were measured after complete execution, where different results obtained are shown below.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

Table 1: Deep learning algorithms evaluation results

No	Algorithm	Accuracy	Precision	Recall	F1-Score
1	ResNet	0.97	0.91	0.90	0.90
2	DenseNet	0.95	0.94	0.93	0.89
3	Xception	0.94	0.90	0.92	0.87
4	InceptionResNet	**0.93**	**0.94**	**0.98**	**0.96**

After evaluating the proposed system using ResNet, DenseNet, Xception, and InceptionResNet models, we observe the values against each deep algorithm. We calculate each algorithm's accuracy, precision, recall, and f-score. So there is no significant difference in the results for each algorithm, but overall, the ResNet model performance was outstanding. The overall results for the accuracy measure with all values are shown in table 1.

6. CONCLUSION

It is concluded that blindness detection was necessary at this time due to its significant impact on humans. Because blindness was a chronicle disease, where blindness complete or partial loss of vision. Blindness occurs suddenly or over a period. Primary reasons for blindness occurrence diabetes and secondly eye diseases. Older ages people in developing countries are more affected than other age people. The major problem with blindness was no proper guidelines or precautions for the people. Globally, significant causes of vision impairment or blindness were uncorrected refractive errors, contracted age, diabetic retinopathy, corneal capacity, and trachoma. Vision impairment impact individual, community, and the economy. All these effects were due to vision impairment. Some strategies were implemented to avoid such as vision impairment and blindness problems. Due to age, the leading causes are uncorrected refractive errors and cataracts.

Diabetes patients mostly face vision problems due to diabetic retinopathy. The high blood glucose level in the eye blood vessels increases the chances of vision problems. We proposed the model using deep learning algorithms to detect blindness in the early stages. We apply pre-processing approaches to manage the dataset. Then apply ResNet, DenseNet, Xception, and InceptionResNet models to train the model. The trained model was used for the testing and evaluate the proposed model using accuracy, precision, recall, and f1-score. The proposed model outperformed using the ResNet model compared to the other models. This model can be utilized for clinical purposes after testing on different datasets.

In the future, if we get a more clinical real-time dataset to test these models for clinical purposes will be beneficial to save the human vision. There will be best to use deep learning and other approaches to make a more generalized model.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

References

1. Atwany, M. Z., Sahyoun, A. H., & Yaqub, M. (2022). Deep Learning Techniques For Diabetic Retinopathy Classification: A Survey. IEEE Access.
2. Trial Runners [Online Available] “https://trialrunners.com/blindness-seeing-believing/”
3. World Health Organization (WHO) [Online Available] “https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment”
4. Lakshminarayanan, V., Kheradfallah, H., Sarkar, A., & Jothi Balaji, J. (2021). Automated Detection and Diagnosis of Diabetic Retinopathy: A Comprehensive Survey. Journal of Imaging, 7(9), 165.
5. Hasan, D. A., Zeebaree, S. R., Sadeeq, M. A., Shukur, H. M., Zebari, R. R., & Alkhayyat, A. H. (2021, April). Machine Learning-based Diabetic Retinopathy Early Detection and Classification Systems-A Survey. In 2021 1st Babylon International Conference on Information Technology and Science (BICITS) (pp. 16-21). IEEE.
6. Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., Ra, I. H., & Alazab, M. (2020). Early detection of diabetic retinopathy using PCA-firefly based deep learning model. Electronics, 9(2), 274.
7. Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., & Srivastava, G. (2020). Deep neural networks to predict diabetic retinopathy. Journal of Ambient Intelligence and Humanized Computing, 1-14.
8. Hamid, K. Iqbal, M. W. Nazir, Z. Muhammad, H. A. B. Fuzail, Z. (2022). Usability empowered by user’s adaptive features in smart phones: the rsm approach. Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/Journal of Tianjin University Science and Technology. 55. 285-304. 10.17605/OSF.IO/6RUZ5.
9. Li, F., Chen, H., Liu, Z., Zhang, X., & Wu, Z. (2019). Fully automated detection of retinal disorders by image-based deep learning. Graefe’s Archive for Clinical and Experimental Ophthalmology, 257(3), 495-505.
10. Balyen, L., & Peto, T. (2019). Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. *The Asia-Pacific Journal of Ophthalmology*, *8*(3), 264-272.
11. Dataset [Online Available] “https://www.kaggle.com/competitions/aptos2019-blindness-detection/data”
12. Mustafa, Saleem & Jaffar, Arfan & Iqbal, Muhammad waseem & Abubakar, Asma & Alshahrani, Abdullah & Alghamdi, Ahmed. (2022). Hybrid Color Texture Features Classification Through ANN for Melanoma. 10.32604/iasc.2023.029549.
13. Moshfeghi, D. M., & Trese, M. T. (2021). Reducing blindness resulting from retinopathy of prematurity using deep learning. Ophthalmology, 128(7), 1077-1078.
14. Shen, B., Huang, C., Xu, W., Yang, T., & Cui, S. (2021). Blind channel codes recognition via deep learning. IEEE Journal on Selected Areas in Communications, 39(8), 2421-2433.
15. Gangwar, A. K., & Ravi, V. (2021). Diabetic retinopathy detection using transfer learning and deep learning. In Evolution in Computational Intelligence (pp. 679-689). Springer, Singapore.
16. Iqbal, Muhammad waseem & Shahzad, Syed Khuram & Ahmad, Nadeem & Amelio, Alessia & Brodic, Darko. (2018). Adaptive Interface for Color-blind People in Mobilephones. 10.1109/ICACS.2018.8333488.

ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 12-2022
DOI 10.17605/OSF.IO/T3XDU

17. Hamid, K., Iqbal, M. waseem, Muhammad, H., Fuzail, Z., Ahmad, Z.: Anova Based Usability Evaluation Of Kid's Mobile Apps Empowered Learning Process. Qingdao Daxue XuebaoGongcheng JishubanJournal Qingdao Univ. Eng. Technol. Ed. 41, 142–169 (2022).
18. Rahman, M. W., Tashfia, S. S., Islam, R., Hasan, M. M., Sultan, S. I., Mia, S., & Rahman, M. M. (2021). The architectural design of smart blind assistant using IoT with deep learning paradigm. Internet of Things, 13, 100344.
19. Tymchenko, B., Marchenko, P., & Spodarets, D. (2020). Deep learning approach to diabetic retinopathy detection. arXiv preprint arXiv:2003.02261.
20. K. Hamid, M. W. Iqbal, H. A. B. Muhammad, Z. Fuzail, Z. T. Ghafoor et al., "Usability evaluation of mobile banking applications in digital business as emerging economy," International Journal of Computer Science & Network Security, vol. 22, no. 1, pp. 250–260, 2022.
21. Naqvi, M. R., Iqbal, M. W., and et al. (2020). Emotion Detection of Contextual Text using Deep learning. 10.1109/ISMSIT50672.2020.9255279.
22. Iqbal, M. J., Bajwa, U. I., Gilanie, G., Iftikhar, M. A., & Anwar, M. W. (2022). Automatic brain tumor segmentation from magnetic resonance images using superpixel-based approach. Multimedia Tools and Applications, 1-19.