

Multi-Class of Retinal Diseases Classification via Deep Learning Techniques Based on Fundus Images

Ahmed Tuama Khalaf* and Salwa Khalid Abdulateef

Computer Science Department Computer Science and Mathematics College, Tikrit University, Iraq



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<https://doi.org/10.54153/sjpas.2024.v6i3.881>

Article Information

Received: 04/04/2024

Revised: 30/04/2024

Accepted: 25/05/2024

Published: 30/09/2024

Keywords:

Retinal disease, deep learning, fundus image, and YOLOv8.

Corresponding Author

E-mail:

ahmed.t.khalaf@st.tu.edu.iq

Mobile:

Abstract

Early detection of retinal disorders is highly important in making the decision at the right moment because of the importance of treatment and prevention of diseases that are a significant cause of blindness. Experts in diagnosing these diseases frequently encounter a number of obstacles, including time, as well as variances across observers because of the occurrence of a comparable set of symptoms among groups, and some may share Categories in similar instances to address these challenges. This work proposes a novel model of deep learning approaches concerning object detection and classification. By utilizing YOLOv8, YOLOv8m-cls was used, which is distinguished by excellent efficiency, accuracy, and speed and outperforms earlier investigations. The model was trained using publicly available data from the Kaggle database, which contains 4217 fundus images, including glaucoma, cataracts, diabetic retinopathy, and the final health condition. The accuracy was 94, the precision was 95%, the recall was 97%, and the F1-score was 96 after 0.5 hours of training. This study demonstrates, based on the data obtained, particularly with the new model YOLOv8m-clsIn, revolutionizing the diagnosis of eye diseases, which leads to better patient outcomes and health care provision.

Introduction:

According to a World Health Organization (WHO) report, almost 2.2 billion people worldwide have vision impairment.[1] Diabetes is a primary risk factor for diabetic retinopathy, which causes damage to the retina's blood vessels. High blood sugar levels associated with diabetes can weaken and destroy the small blood vessels in the retina, causing visual issues and probable blindness if left untreated.[2] According to one report, India has approximately 12,000 ophthalmologists for 60 million diabetics with eye disorders. The main reason for the growing number of patients is that most people are unaware that they have this condition. They also exhibit insensitivity and a careless attitude toward this sickness. Approximately 18% of people with diabetes suffer from Diabetic Retinopathy (DR), and the probability of acquiring DR in a diabetic person is 25 times greater than that of a healthy [3]. Glaucoma is a lesion inside the optic nerve that can lead to impaired vision or blindness. Early identification and therapy help prevent disease progression, while there is currently no cure for this lesion.[4] Glaucoma, nicknamed "the silent thief of sight," is the world's second leading

cause of blindness. This illness affects around 60 million people worldwide; by 2020, that amount is expected to rise to 79.6 million.[5]. A cataract is a lenticular opacity that obscures the clear lens in the eyes. Typically, the lens focuses light onto the retina. Cataracts stop light from reaching the lens, decreasing visual acuity [6]. Cataracts affect around 90% of persons by the age of 65, and half of those aged 75-85 experience vision loss owing to them [7]. Deep Learning (DL) has demonstrated notable benefits in many real-world applications lately [8]. Various attempts were made to use retinal fundus images and DL approaches to identify retinal eye disorders. Classifying multi-class retinal eye diseases is still difficult, despite the fact that DL approaches in machine learning (ML) have produced excellent results in binary classification tasks that distinguish between healthy and diseased retinal fundus images [9,10]. In previous years, most previous studies used old models to classify retinal diseases. These algorithms require a large amount of storage space because they contain a huge number of parameters. The main reason is that accuracy is one of our most important priorities in this field. In our paper, we worked hard to overcome these challenges. We used a revolutionary model. It is lightweight and is characterized by achieving a balance between accuracy and speed. It is known as YOLOv8, and especially one of the five YOLOv8 models that was used is YOLOv8m-cls, as it contains a medium storage space, in addition to containing 15,777,460. This novel approach outperforms its predecessors in terms of both performance and accuracy, representing a significant advancement in the sector. Through thorough experimentation and research, we have proved YOLOv8's exceptional capabilities, showing its potential to change existing approaches and establish norms in the subject of our study.

LITERATURE SURVEY

Researchers have attempted to classify eye illnesses using advanced machine learning and image-processing techniques. Researchers have lately begun to use deep learning approaches to identify appropriate answers to this problem and achieve acceptable performance in the early diagnosis of irreversible eye illnesses. This section will go over many classification techniques for eye illnesses. A. Singh et al (2019) [11] In this study, researchers proposed early diagnosis of glaucoma using the pre-trained model Inception-v3 on RIM-ONE, a freely available dataset, including 624 monocular and 159 stereoscopic retinal fundus images. To approximate natural deformations, Fundus images underwent data augmentation, CLAHE, and normalization. Accuracy was 92.51%. M. Smaida et al (2020) [12] This study uses deep learning approaches to diagnose numerous types of eye illnesses, including diabetic retinopathy, glaucoma, myopia, and health status, in comparison to three neural network structures: CNN, VGG16, and InceptionV3. They are compared using a packing set. InceptionV3 was the most accurate of the three approaches. Rating reached 87%. N. Chea et al (2021) [13] In this study, they used different methods based on convolutional neural networks, including the Resnet50 algorithm. Image pre-processing techniques include region of interest reduction, adaptive histogram equation with limited contrast at ISO brightness level, and data augmentation. Resnet50 obtained a peak and average accuracy of 91.16% and 85.79% in retinopathy (DR), glaucoma (GLC), and age-related macular degeneration (AMD), respectively. R. Sarki et al. (2021) [14] this study introduced a new CNN model that can classify Diabetic Eye Disease (DED) into various groups. 1,748 data points, the Messidor, DRISHTI-GS, Messidor-2, and Kaggle cataract databases, and Retinal fundus images were enhanced with contrast by using mathematical morphology. By using the right model parameters and enhancing the quality of the training images, the proposed model in the

RMSprop optimizer yielded the maximum sensitivity and specificity. The suggested CNN model achieved 81% accuracy in the test dataset. T. Babaqi et al (2023) [15] the study aimed to distinguish between normal eyes and those with diabetic retinopathy, cataracts, or glaucoma. The following resources were used to acquire information about using CNN and transfer learning for multi-class classification. The dataset contains approximately 4200 colored photos of normal eyes, cataracts, diabetic retinopathy, and glaucoma. The images were resized from 512×512 to 224×224. The suggested transfer learning approach is based on a pre-trained Efficient Net CNN architecture model, which obtained 84% accuracy.

Research methodology:

Our research methodology has five phases. The first phase involves data collection. In the second phase, we discuss preprocessing the data and then dividing it. The third phase involves utilizing a model, while the final phase is evaluating the model's performance. Fig. 1 shows a detailed explanation of the research methodology used.

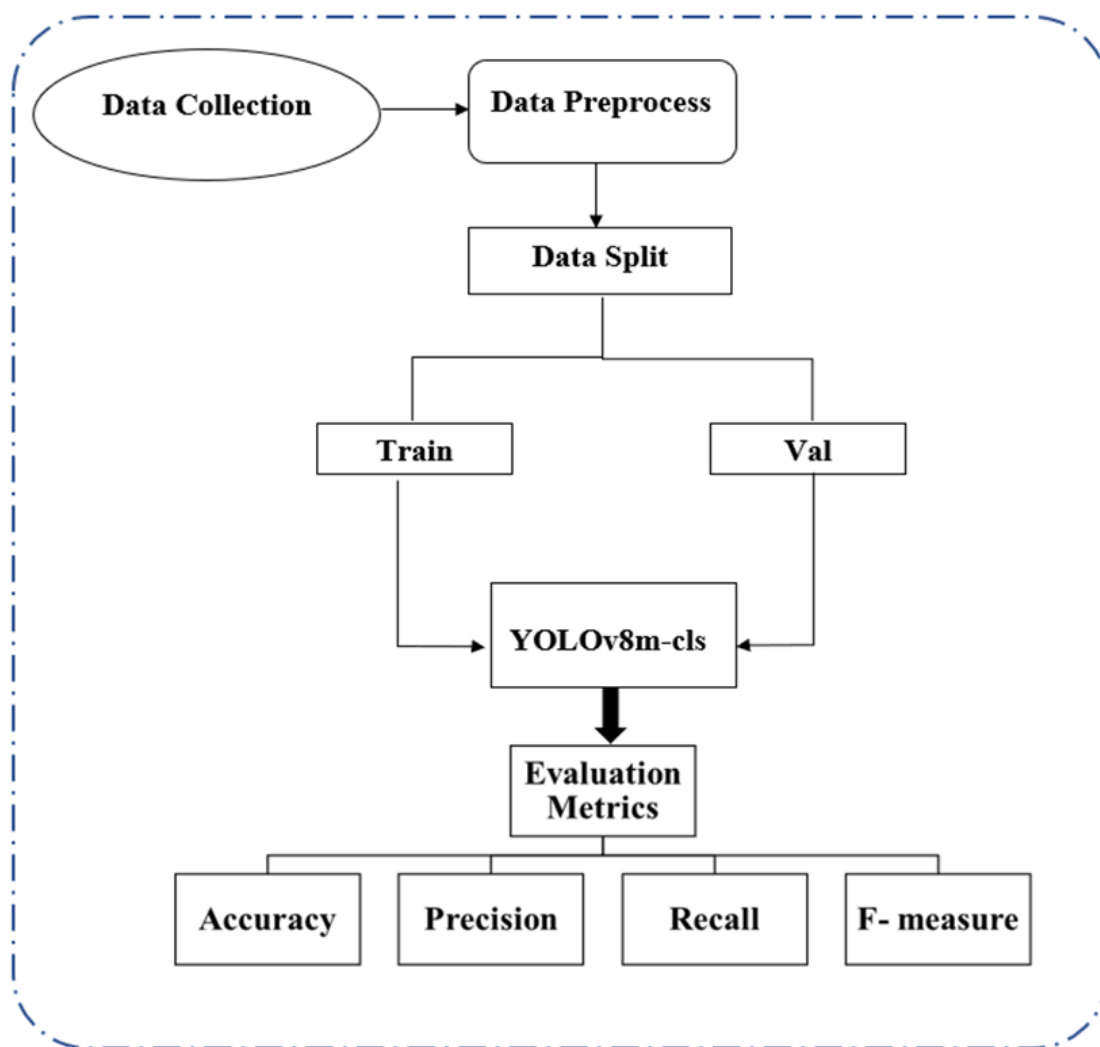


Fig. 1 Research methodology

Data Collection:

Our framework utilizes a dataset sourced from Kaggle [16], which includes a total of 4,217 fundus images categorized into four distinct classes, which is considered the most dangerous and the most famous disease. The first category is for glaucoma, comprising 1,007 fundus images stored in JPG and PNG format, with each image containing eight bits for every color channel. All other categories had eight bits for each color, and the image format was JPG; cataracts are represented by 1038 images, and the third category, diabetic retinopathy, includes 1098 images. The fourth and last category, health conditions, has 1074 images. Fig. 2 shows a group of fundus images for different diseases.

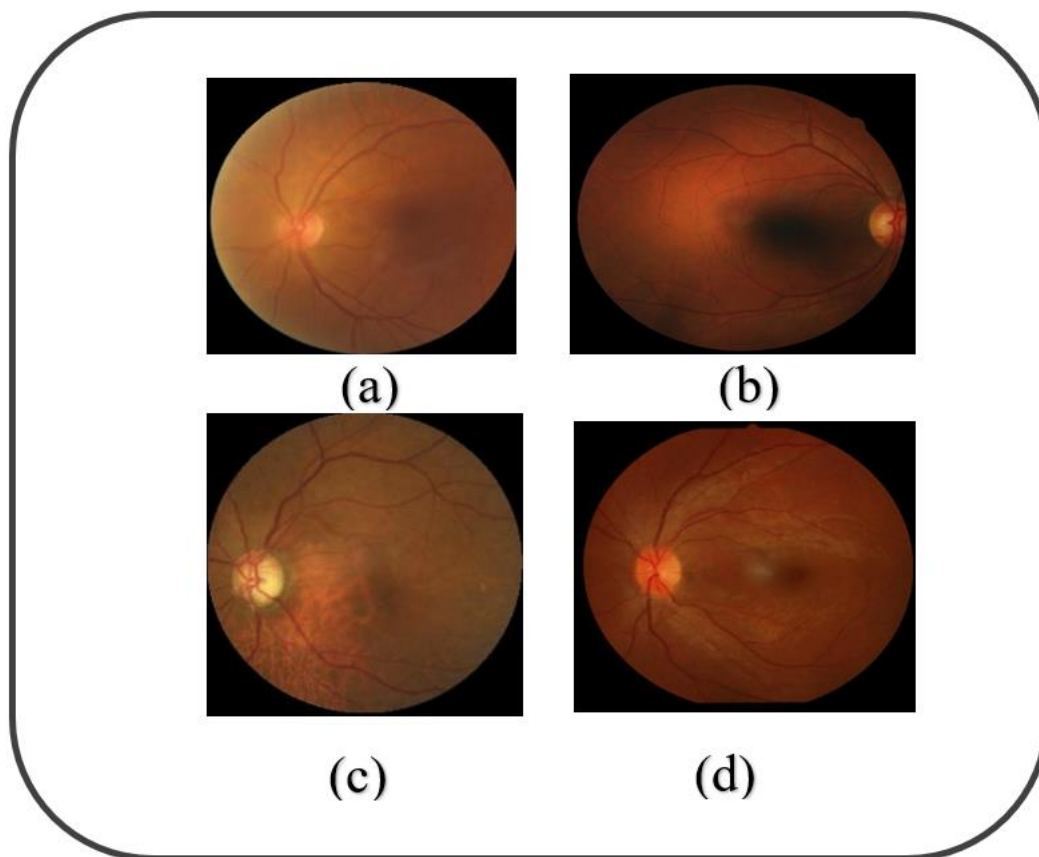


Fig. 2 Fundus images for different diseases (a) cataract;(b) diabetic retinopathy;(c) glaucoma;(d) normal

Data Pre-processing

Data pre-processing is the most essential and crucial factor in the generalization performance of a Supervised Machine Learning Algorithm [17]. Below are the processing steps used in our research methodology.

Resize image and Enhance Contrast.

Before training, the size and dimensions of all fundus pictures were changed by unifying them, with some having 256 categories and others having 512, whose dimensions were standardized to 224. To keep the quality and resolution of the fundus images after reducing the size, the degree of detail (contrast) is raised by (0.6,1.4) from the original image. That is, the brightness of the images will increase by a factor of no less than 0.6. The intention is that the dark color will be darker and the light color will be lighter but not excessively, while the factor will improve the contrast by 1.4. Fig. 3 demonstrates Image processing steps.

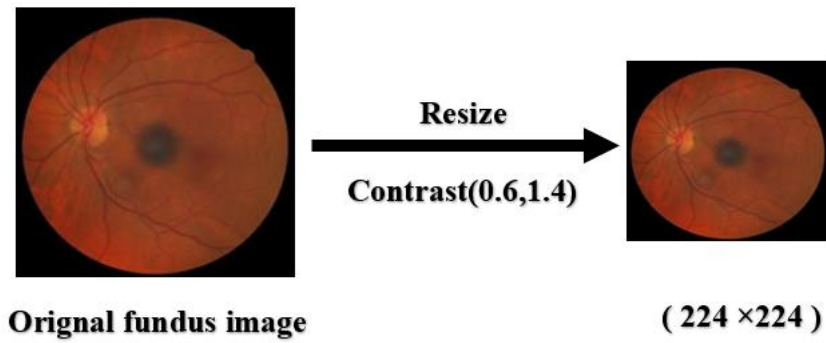


Fig. 3 image processing steps

Data Split

In the data division stage, we divided the dataset into two independent subsets, allocating 80% for training and the remaining 20% for validation. This split method was chosen based on well-established principles in machine learning and deep learning workflows, with the goal of striking a compromise between supplying enough data for model training and ensuring robust evaluation of model performance. By allocating a significant amount of the dataset to training, our model was able to successfully understand the underlying patterns and relationships contained in the data, enhancing its capacity to generalize to previously unseen samples.

Model

The model used in our research is from the YOLOv8 family, specifically the YOLOv8M variation. This model variant is defined by its architecture, which includes 141 layers and a remarkable 15,777,460 parameters. Notably, YOLOv8m-cls has a computational efficiency of around 41.9 GFLOPS. We used publicly available datasets for its training as described earlier and used optimizer Adam. The training process was fine-tuned by setting the momentum parameter to 0.9 and the learning rate 0.000714, ensuring excellent convergence throughout training. A batch size of 16 was used to enable efficient processing of data batches during training cycles. The training environment made use of the Python programming language in the Colab environment, taking advantage of the computational capabilities of the T4(GPU) infrastructure. This complete approach to model selection, training, and environment setup highlights our dedication to systematically creating and assessing deep learning solutions to ensure robust and accurate performance in real-world applications. Table 1 shows Adjust the hyperparameters of the YOLOv8m-cls

Table 1: Adjust the hyperparameters of the model.

Parameter	Value	Parameter	Value
epoch	25	momentum	0.9
weight(decay)	0.0005	Patch size	16
Adam learning rate	0.000714	Image size	224

Model Evaluation

In machine learning or deep learning, a performance evaluation metric gauges the model's effectiveness in comparison to a predetermined dataset. A performance evaluation metric measures the capability of a trained model to resolve a specific problem. The confusion

matrix is an efficient metric for assessing the performance of a deep learning or machine learning model [18]. The efficiency of the classification assignment is computed using the confusion matrix, which is established through the tally of correct and incorrect categorizations for each variant [19,20]. Illustrated are the classifiers utilized for the two-category confusion matrix in Fig. 4. We limited ourselves to using four of the most popular performance measures, namely Accuracy, Precision, Recall, and F-Measure

		Predictive	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

Fig. 4 two-category confusion matrix

Accuracy

Accuracy indicates how frequently the classifier properly predicts. as the ratio of the number of correct forecasts to the total number of predictions [21]. The formula for calculating accuracy is given below.

$$Accuracy = \left(\frac{TP+TN}{TP+TN+FP+FN} \right) \quad (1)$$

Precision

Precision is a metric utilized in machine learning to evaluate the efficacy of a classification model. The model computes the proportion of accurate positive predictions relative to the total number of positive predictions [22]. The description of the precision model is as follows:

$$Precision = \left(\frac{TP}{TP+FP} \right) \quad (2)$$

Recall

Recall, also known as sensitivity, measures a classification system's ability to predict the number of positive cases[18] accurately. The formula for calculating Recall is given below.

$$Recall = \left(\frac{TP}{TP+FN} \right) \quad (3)$$

F-Measure

It provides a combined view of Precision and Recall measures. It reaches its greatest when Precision Equals Recall [23]. The formula for calculating F1-score is given below.

$$F1\text{-score} = \left(2 * \frac{Precision * Recall}{Precision + Recal} \right) \quad (4)$$

Result and discussion

Our findings show that the suggested approach is highly effective at classifying retinal disorders. We achieved a 94% accuracy by training the model on the data it reported at a pace of 0.1 across 25 epochs. The Precision value was 95%, the recall value was 97%, and the F-score was 96. It is important to note that the training period is merely 0.5 hours. Table 2 displays the most important outcomes. Meanwhile, Fig. 5 shows the accuracy of the validation

in conjunction with the epoch number. And the loss values for the training and validation process and Fig. 6 Image prediction using YOLOv8m-cls.

Table 2: Evaluation performance for the YOLOv8m-cls and time.

Model	Dataset	Accuracy	Precision	Recall	F1-Score	Time
YOLOv8-cls	Eye Disease Retinal Images	94%	95%	97%	96%	0.5(H)

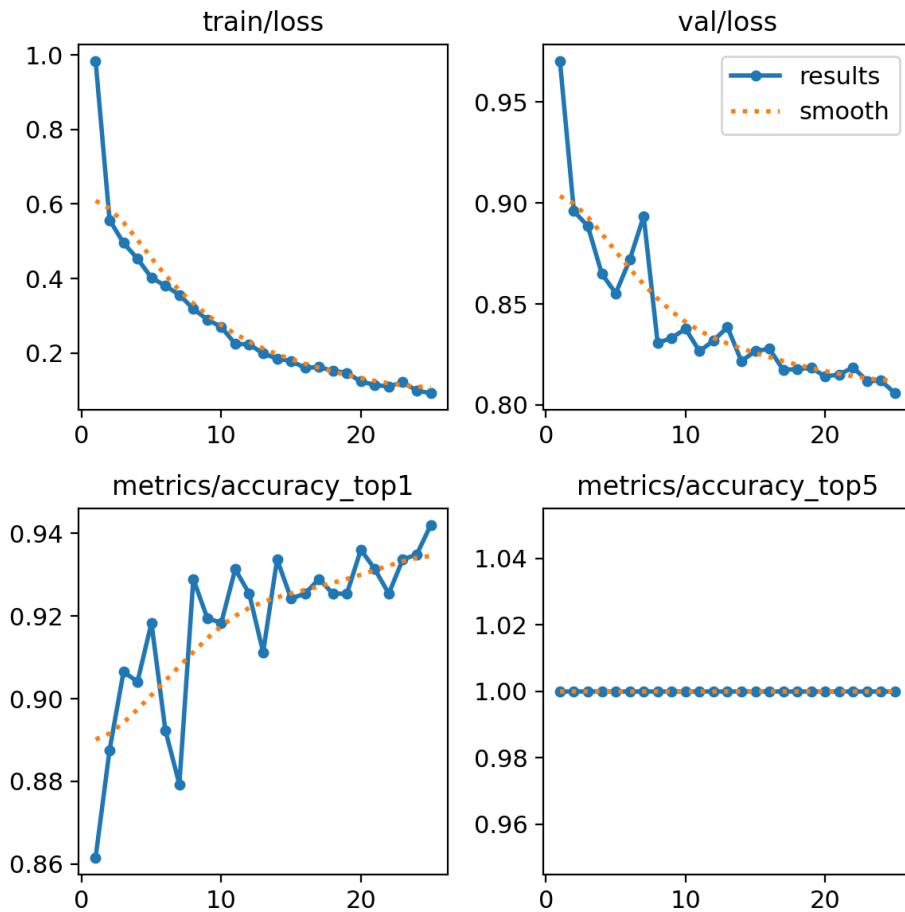


Fig. 5 Validation accuracy, train loss, and val loss against each epoch

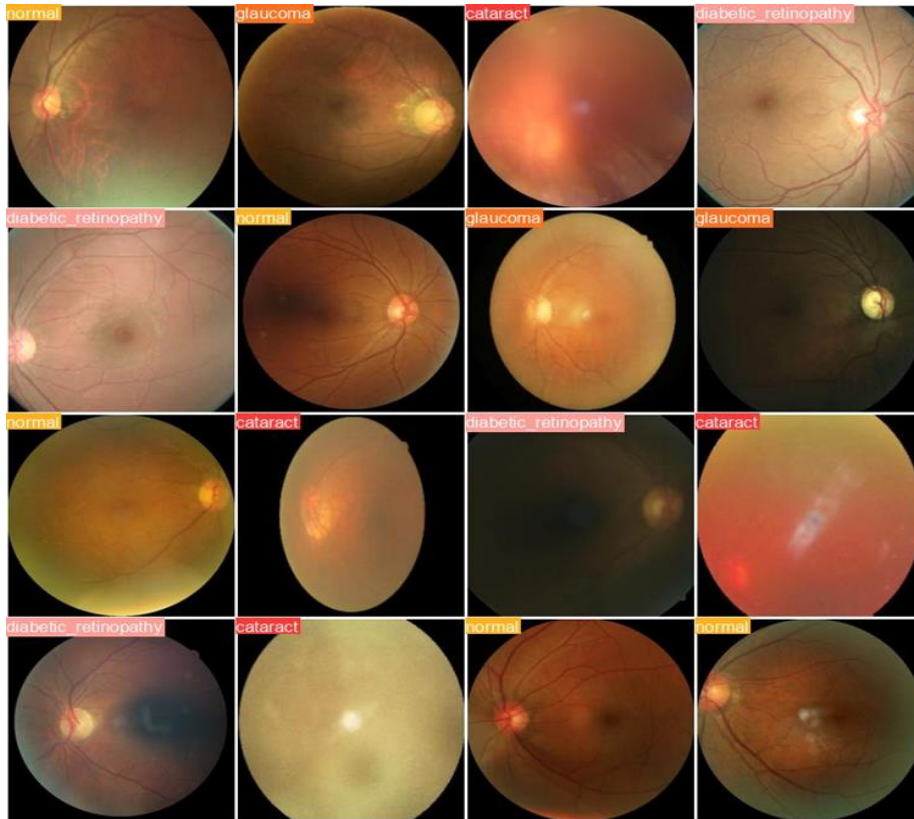


Fig. 6 Predicting various diseases using the YOLOv8m-cls model

The confusion matrix, which details the performance of our classification model across many categories, is shown in the picture. Notably, diabetic retinopathy had the highest accuracy rate, with a faultless score of 100%. This outstanding success is owed to our model's rigorous classification of all 220 test photos in this category, demonstrating its robustness in reliably diagnosing diabetic retinopathy cases. Glaucoma had the lowest accuracy rate, although it was still commendable at 88%. Despite this, there were some cases of misclassification within the glaucoma group, resulting in somewhat reduced accuracy compared to other categories. In Fig. 7, we show the confusion matrix for the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) for the YOLOv8m-cls predictions.

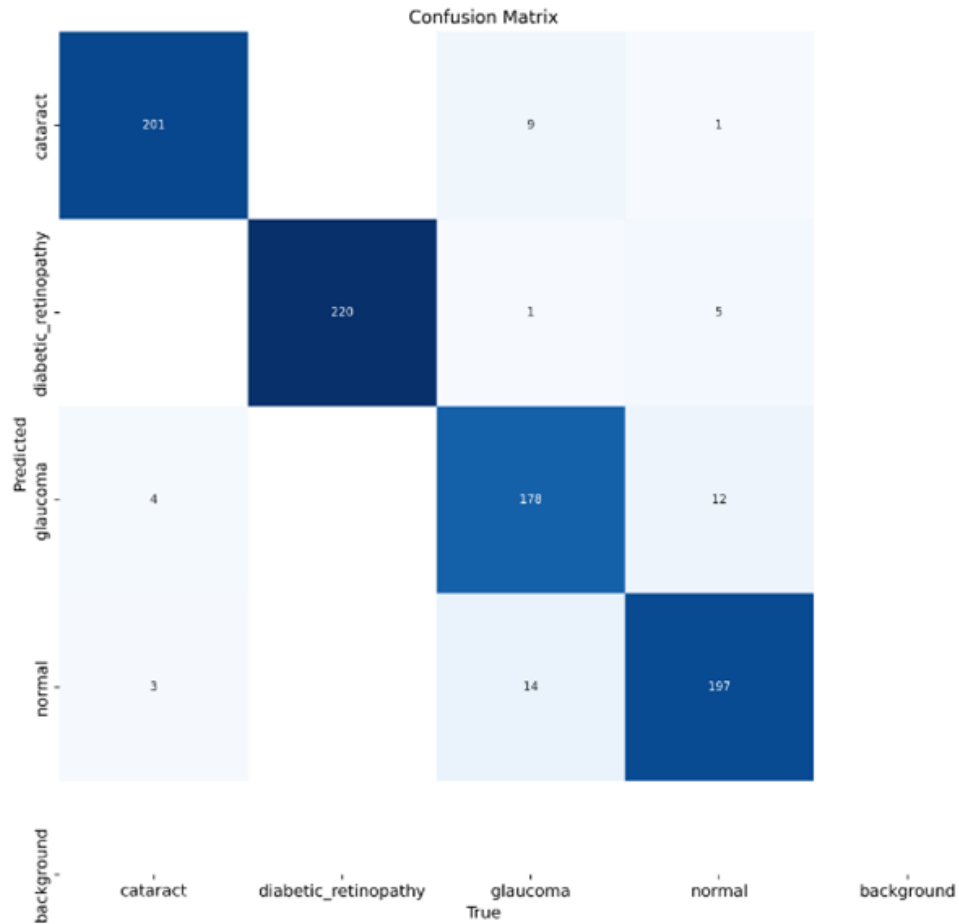


Fig. 7 Confusion matrix for the YOLOv8m-cl classification.

Comparison with previous works

The provided model was compared to some earlier CNNs that use traditional structures and are similar in classes. It exceeds the model in terms of accuracy and efficiency. Furthermore, the model requires less storage space than Resnet50 and Inceptionv3. The lightweight model has 141 layers, but the 15,777,460 million parameters are enough for training. Through the following table, study number three [13] used Resnet50 for four categories, and the total number of images was 4789. The model obtained an accuracy of 91%. As for the first study [11], the model was limited to diagnosing glaucoma with health status and binary classification using an available data set, the RIM-ONE dataset, with a number of 783. The image of the model was used in Inceptionv3, where the model obtained an accuracy of 92.51%. As for study number four [14], he used only 1748 images to classify five categories, and he obtained an accuracy of 81%. As in the fifth study [15], they used four categories for classification using the Alexnet algorithm. The number of images was 4200. The model obtained an accuracy of 84. Table 3 shows the results and number of categories for classification with some Previous works.

Table 3: Comparison with previous works.

S. N	Paper	Model	Dataset	Number of images	Number of classes	Accuracy
1	[11]	Inceptionv3	RIM-ONE dataset	783	2	92.51%
2	[12]	Inceptionv3	Kaggle	2781	4	87.2%
3	[13]	Resnet 50	(NOISE STRESS)	3,789	4	91%
4	[14]	CNN	Messidor, Messidor-2, DRISHTI-GS, and Kaggle	1748	5	81%
5	[15]	Alexnet	public	4200	4	84%
6	Proposed Model	YOLOv8-cls	Eye Disease Retinal Images	4217	4	94%

Conclusion

Retinal illnesses are regarded as a major cause of irreversible blindness, particularly if they are not diagnosed early on or if visual diagnosis is hampered by identical symptoms. The suggested model, based on convolutional neural networks, allows for the early diagnosis of disorders like glaucoma, cataracts, and diabetic retinopathy. Our findings show that the YOLOv8m-cls model can effectively diagnose, take action early in the condition, and ensure that the patient receives the appropriate medication. These classifications were determined using 4217 fundus photos from the Kaggle database. The model demonstrates its advantage over previous traditional methods, as it was obtained with an accuracy of 94%. Future work will focus on increasing accuracy and testing for other diseases and can be developed into an application that can be used for early diagnosis and assistance to human doctors.

References

- [1] <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>
- Y. Elloumi, M. Akil, and H. Boudegga, "Ocular diseases diagnosis in fundus images using a deep learning: approaches, tools and performance evaluation," in *Real-Time Image Processing and Deep Learning 2019*, 2019, vol. 10996: SPIE, pp. 221-228.

3. G. Mushtaq and F. Siddiqui, "Detection of diabetic retinopathy using deep learning methodology," in *IOP conference series: materials science and engineering*, 2021, vol. 1070, no. 1: IOP Publishing, p. 012049.
4. M. Juneja, S. Singh, N. Agarwal, S. Bali, S. Gupta, N. Thakur, and P. Jindal, "Automated detection of Glaucoma using deep learning convolution network (G-net)," *Multimedia Tools and Applications*, vol. 79, pp. 15531-15553, 2020.
5. R. Kashyap, R. Nair, S. M. P. Gangadharan, M. Botto-Tobar, S. Farooq, and A. Rizwan, "Glaucoma detection and classification using improved U-Net Deep Learning Model," in *Healthcare*, 2022, vol. 10, no. 12: MDPI, p. 2497.
6. M. S. Junayed, M. B. Islam, A. Sadeghzadeh, and S. Rahman, "CataractNet: An automated cataract detection system using deep learning for fundus images," *IEEE access*, vol. 9, pp. 128799-128808, 2021.
7. T. Tasin and M. A. Habib, "Computer-Aided Cataract Detection Using Random Forest Classifier," in *Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM 2021*, 2022: Springer, pp. 27-38.
8. S. Sengupta *et al.*, "A review of deep learning with special emphasis on architectures, applications and recent trends," *Knowledge-Based Systems*, vol. 194, p. 105596, 2020.
9. M. Elkholy and M. A. Marzouk, "Deep learning-based classification of eye diseases using Convolutional Neural Network for OCT images," *Frontiers in Computer Science*, vol. 5, p. 1252295, 2024.
10. A. Ali *et al.*, "Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image," *Entropy*, vol. 22, no. 5, p. 567, 2020.
11. A. Singh, S. Sengupta, and V. Lakshminarayanan, "Glaucoma diagnosis using transfer learning methods," in *Applications of machine learning*, 2019, vol. 11139: SPIE, pp. 223-232.
12. M. Smaida and S. Yaroshchak, "Bagging of Convolutional Neural Networks for Diagnostic of Eye Diseases," in *COLINS*, 2020, pp. 715-729.
13. N. Chea and Y. Nam, "Classification of Fundus Images Based on Deep Learning for Detecting Eye Diseases," *Computers, Materials & Continua*, vol. 67, no. 1, 2021.
14. R. Sarki, K. Ahmed, H. Wang, Y. Zhang, and K. Wang, "Convolutional neural network for multi-class classification of diabetic eye disease," *EAI Endorsed Transactions on Scalable Information Systems*, vol. 9, no. 4, 2021.
15. T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer, and D. Won, "Eye disease classification using deep learning techniques," *arXiv preprint arXiv:2307.10501*, 2023.
16. <https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification>.

17. K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91-99, 2022.
18. D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.
19. M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, p. 1, 2015.
20. S. K. Abdulateef, A. N. Ismael, M. D. Salman, " Feature weighting for parkinson's identification using single hidden layer neural network," *International Journal of Computing*, vol.22, no.2, pp.225-230, 2022. doi:10.47839/ijc.22.2.3092
21. S. Visa, B. Ramsay, A. L. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection," *Maics*, vol. 710, no. 1, pp. 120-127, 2011.
22. Ž. Vujović, "Classification model evaluation metrics," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 599-606, 2021.
23. D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC genomics*, vol. 21, pp. 1-13, 2020.

تصنيف متعدد الفئات لأمراض الشبكية عبر تقنيات التعلم العميق بناءً على صور قاع العين

أحمد طعمه خلف*، سلوى خالد عبداللطيف

قسم علوم الحاسوب كلية علوم الحاسوب والرياضيات جامعة تكريت، العراق

الخلاصة:

يعد الكشف المبكر عن اضطرابات الشبكية أمراً في غاية الأهمية من أجل اتخاذ القرار في اللحظة المناسبة، وذلك لأهمية العلاج والوقاية من الأمراض التي تعد سبباً مهماً في الإصابة بالعمى. كثيراً ما يواجه الخبراء في تشخيص هذه الأمراض عدداً من العقبات، بما في ذلك الوقت، فضلاً عن الفروق بين المراقبين بسبب حدوث مجموعة مماثلة من الأعراض بين المجموعات، وقد يتشارك البعض في الفئات في حالات مماثلة لمواجهة هذه التحديات. يقترح هذا العمل نموذجاً جديداً لأساليب التعلم العميق المتعلقة باكتشاف الأشياء وتصنيفها. وباستخدام YOLOv8، تم استخدام YOLOv8m-cls، الذي يتميز بالكفاءة والدقة والسرعة الممتازة، ويتفوق على التحقيقات السابقة. تم تدريب النموذج باستخدام البيانات المتاحة للجمهور من قاعدة بيانات Kaggle، التي تحتوي على 4217 صورة لقاع العين، بما في ذلك الجلوكوما، وإعتام عدسة العين، واعتلال الشبكية السكري، والحالة الصحية النهائية. كانت الدقة 94، والدقة 95%، والاستدعاء 97%، وكانت درجة F1 بعد 0.5 ساعة من التدريب. توضح هذه الدراسة، بناءً على البيانات التي تم الحصول عليها، ولا سيما مع النموذج الجديد YOLOv8m-cls، إحداث ثورة في تشخيص أمراض العين، مما يؤدي إلى نتائج أفضل للمرضى وتوفير الرعاية الصحية.

معلومات البحث:

تاريخ الاستلام: 2024/04/04

تاريخ التعديل: 2024/04/30

تاريخ القبول: 2024/05/25

تاريخ النشر: 2024/09/30

الكلمات المفتاحية:

مرض الشبكية، التعلم العميق وصورة

قاع العين و YOLOv8

معلومات المؤلف

الايمل:

ahmed.t.khalaf@st.tu.edu.iq

الموبايل: