Received: 19 November 2023 | Revised: 7 March 2024 | Accepted: 23 April 2024 | Published online: 26 April 2024

RESEARCH ARTICLE

Artificial Intelligence and Applications
2024, Vol. 0(0) 1–9
DOI: 10.47852/bonviewAIA42022106

BON VIEW PUBLISHING

Advancements in Medical Image Analysis: A Comprehensive Method of AI-Based Classification and Segmentation Technique

Ankur Biswas[1,*] and Rita Banik[2]

[1]*Computer Science & Engineering Department, Tripura Institute of Technology, India*

[2]*Electrical Engineering Department, ICFAI University, India*

Abstract: Medical imaging is a vital tool in the existing healthcare system for precise analysis and treatment of various medical conditions. The use of artificial intelligence particularly machine learning and deep learning approaches has changed the interpretation of medical images and resulted in substantial advances in the area. With the increasing occurrence of eye diseases and the imperative need for early diagnosis, artificial intelligence presents potential solutions for the automatic, precise, and early detection of various eye problems. New advances in machine learning and deep learning allow different ocular disorders to be automatically analyzed, classified, and segmented which leads to early and precise diagnosis. This study attempts to give a comprehensive overview of the rapidly evolving field by reviewing the most recent approaches, challenges, and possible uses of artificial intelligence in medical imaging for ocular illnesses with a focus on segmentation and classification. The research also presents advanced methodologies such as transfer learning with MobileNet and U-Net for automating the diagnosis of various ocular problems using classification and segmentation. The study highlights the importance of early and accurate diagnoses for better patient outcomes.

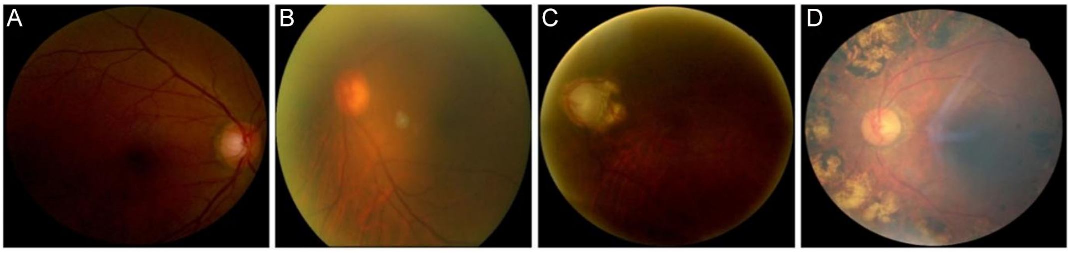
Keywords: ocular disease, artificial intelligence, deep learning, classification, segmentation

1. Introduction

Medical images are crucial in modern healthcare for accurate diagnosis and treatment of a wide range of medical disorders. The application of artificial intelligence (AI), particularly deep learning (DL), techniques has resulted in breakthroughs in the field of medical image interpretation [1, 2]. AI-powered medical image analysis has demonstrated tremendous promise in automating the identification and characterization of anomalies in medical images such as X-rays, magnetic resonance imaging (MRI) scans, computed tomography (CT) scans, and histopathology slides. Techniques such as convolutional neural networks (CNNs) and recurrent neural networks have outperformed humans in tasks including cancer detection, organ segmentation, and illness categorization, creating new standards in the field [3–5]. Deep fusion clustering (scDFC) is an innovative method introduced by [6] intended for the analysis of single-cell RNA sequencing data. This protocol powers DL approaches to improve the accuracy of single-cell transcriptomic data clustering by integrating data from numerous sources. The authors offer insightful information on the processing of single-cell RNA-seq data by signifying how scDFC advances clustering results when compared to conventional methods. These advances have significantly improved the accuracy and efficiency of medical image interpretation resulting in better patient care and results. Recent advances in AI for medical image analysis have resulted in the creation of sophisticated algorithms capable of detecting minute patterns and abnormalities that the human eye may miss [7, 8]. As a result of the advancements in this area, medical image analysis has become much more accurate and effective, improving patient care and results [9–11]. Even with the advancements, there are still issues with the current strategies. Obstacles including interpretability, scalability, and resilience in many therapeutic settings continue to exist. Understanding these limitations is essential to create solutions that not only deal with the present issues but also foresee future medical image analysis needs. Thus, in this rapidly evolving setting, it becomes imperative to show how new solutions contribute to ongoing advancements in addition to addressing the modern problems that current methods are facing. The purpose of this study is to examine the rapidly evolving field by reviewing the most recent approaches, challenges, and possible uses of AI in medical imaging for ocular illnesses with a focus on segmentation and classification, and advance the area by filling in important gaps in the existing approaches by enhancing accuracy in challenging imaging conditions or improving interpretability for clinical practitioners through segmentation and classification. The study, hence examines the critical issue of early and accurate diagnosis of eye illnesses by concentrating on the application of AI, particularly DL techniques, in automating the segmentation

*__Corresponding author:__ Ankur Biswas, Department of Computer Science & Engineering, Tripura Institute of Technology, India. Email: abiswas.tit@gmail.com

Figure 1
Different eye diseases: (A) Normal, (B) Cataract, (C) Glaucoma, and (D) Diabetic retinopathy



and classification of ocular disorders. By exploiting medical imaging techniques such as fundus photography, the research examines the application of CNNs trained on huge datasets of annotated healthy and diseased eye pictures. Moreover, the research investigates the effectiveness of transfer learning and fine-tuning pre-existing models like MobileNet, explicitly for eye disease datasets to achieve high accuracy in classification. This approach contributes to the development of AI-driven solutions for automatically classifying eye disorders aiming to improve the timely detection and prevention of visual impairment globally. Additionally, this study focuses on the use of DL models, like the U-Net architecture, for the segmentation of eye illnesses. The goal of the research is to better understand the complexities involved in accurately identifying and isolating afflicted areas within ocular images by exploring the nuances of segmentation approaches, especially when using cutting-edge models such as U-Net. The research explores the utilization of U-Net and similar models with a focus on their efficiency for future use in medical image segmentation.

The rest of the paper is organized as follows: Section 2 presents a detailed review of related literature. Section 3 furnishes details of the methodologies available and employed in classification and segmentation in the context of ocular diseases. In Section 4, the results of the study based on the selected model are presented. Finally, Section 5 concludes the paper by providing a summary of the study and exploring the future scope in this field.

2. Literature Review

The domain of medical imaging has seen enormous growth in the recent decades. It has become a crucial component of modern healthcare for illness diagnosis, treatment planning, and monitoring. Various imaging modalities such as X-rays, CT, MRI, and ultrasound offer doctors with all-inclusive visual representations of interior anatomical structures and clinical conditions. Presently, eye disorders constitute a major global health problem that impacts millions of persons and imposes a significant cost on healthcare systems worldwide. Timely and precise identification of ocular illnesses is critical for optimal management and the prevention of visual impairment. Traditional techniques of diagnosing eye disorders generally rely on manual interpretation of fundus images, which can be time-consuming and vulnerable to inter-observer variability. In recent years, there have been substantial improvements in the application of AI [12], notably DL approaches [13, 14], for the automated segmentation and classification of eye disorders [15, 16]. CNNs, a kind of DL model, have exhibited excellent performance in the analysis and categorization of medical pictures, notably those linked to eye disease [17–19]. These CNNs are trained on huge datasets including annotated images of healthy and affected eyes, enabling them to recognize subtle patterns indicative of various ocular disorders with a high degree of accuracy. Moreover, transfer learning is a technique that embraces fine-tuning pre-trained models on particular eye ailment datasets and has been proven to be important for attaining high classification accuracy even with minimal training data [20–22]. Fundus images have given excellent sources of image data for training and evaluating AI models in the context of eye illness diagnosis. Given the potential of AI to automate the segmentation and classification of eye disorders, the intent is to examine and explicate the use of AI-driven solutions in the context of medical imaging for ocular pathology. By using the capabilities of DL and medical image analysis, the work hopes to contribute to the evolution of automated diagnostic tools for eye illnesses, ultimately aiming to enhance the prompt detection and prevention of eye diseases on a global basis. By incorporating both segmentation and classification into the ongoing breakthroughs, the present challenges are not only solved but also actively contribute to the ongoing advancements in medical image analysis. The method is intended to set new norms, demonstrating its potential to redefine accuracy, efficiency, and clinical applicability. When provided with an eye image as input, the system accurately classifies the image into one of the predefined categories as shown in Figure 1.

The classes are as follows: normal, cataract, glaucoma, and diabetic retinopathy (DR). This classification is based on the training of the model, where the model learns to differentiate between the various characteristics and features associated with each class through the analysis of a labeled dataset.

2.1. AI-based segmentation

Medical image segmentation is the process of dividing a medical image into several segments in order to detect and outline certain structures or regions of interest [23–25]. Treatment planning, surgical guiding, and disease monitoring entirely rely on segmentation. DL approaches such as U-Net and SegNet, have shown great promise in precise and fast medical image segmentation [26, 27]. However, the need for robust models that can handle variations in image quality, noise, and anatomical variation are the most challenging aspects of segmentation. The integration of real-time segmentation into clinical operations remains a barrier that must be addressed. Figure 2 illustrates different retinal structures like hemorrhages, microaneurysms, hard exudates, soft exudates, and optic discs. The ability of the system to properly segment these features is highlighted as essential for its functionality.

2.2. AI-based classification

The AI-based classification in medical image analysis requires categorizing the images into specific classes or spotting irregular

Figure 2
Different retinal structure: (A) Hemorrhages, (B) Microaneurysms, (C) Hard exudates, (D) Soft exudates, and (E) Optic disc

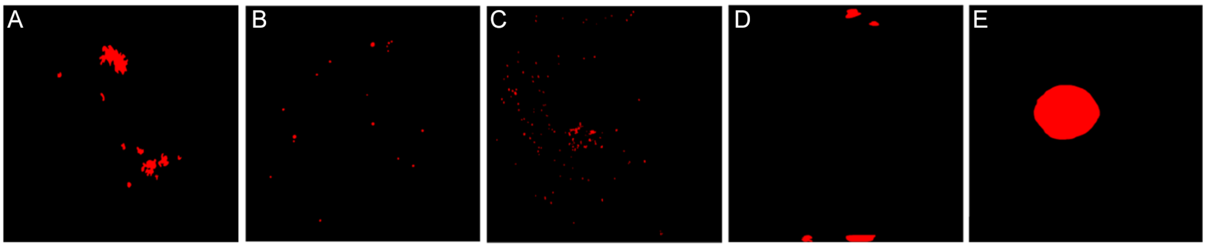
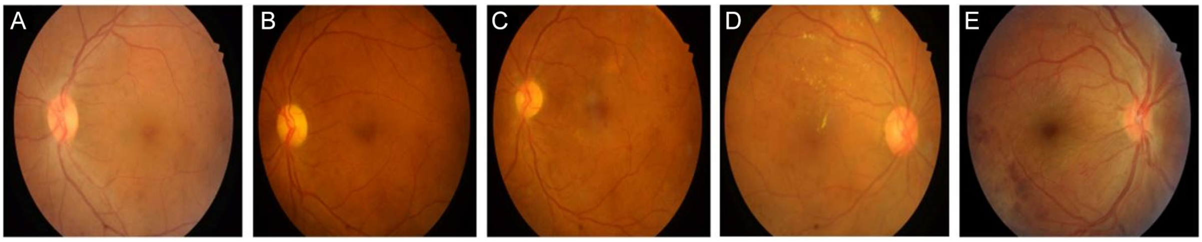


Figure 3
Different stages DR: (A) Normal, (B) Mild, (C) Moderate, (D) Severe, and (E) Proliferative



patterns. CNNs have emerged as a powerful technique for image classification due to their ability to automatically learn hierarchical features. CNNs have been used to categorize medical images such as X-rays, MRIs, CT scans, and histopathology slides in many investigations. Despite the advancement of AI-based categorization, it confronts challenges. The need for large annotated datasets, model conclusion interpretability, and the risk of biased predictions are a few of them. Overcoming these obstacles is critical for the effective implementation of AI in clinical practice. Figures 3 and 4 showcase different stages of DR and macular edema respectively. It indicates the need for the system to accurately classify and differentiate between these critical conditions based on distinctive features extracted from retinal images.

3. Methodology

The research utilizes a pre-trained CNN model applied to medical images to automatically segment and classify eye disorders. A variety of annotated eye images were gathered from fundus photography. These images covered a range of ocular pathologies such as glaucoma, DR, and cataracts. The foundation for the proposed method is formed through key findings and insights into the existing gaps and opportunities for innovation. Hence, the proposed method introduces a comprehensive AI-based technique for eye disease detection that goes beyond existing standards. The approach is designed to tackle the identified challenges and provide an innovative solution in medical image analysis utilizing both segmentation and classification.

3.1. Datasets for classification

Researchers frequently use datasets like the Kaggle DR Detection dataset EyePACs [28]. It contains high-resolution retina images labeled with DR severity for the categorization of DR. The APTOS 2019 Blindness Detection from Kaggle is another popular color image dataset that provides different degrees of DR [29]. Other well-liked datasets for DR are the Messidor dataset and RIGA dataset [30]. Furthermore, the DIARETDB1 and DIARETDB0 databases include a variety of images with DR annotations. On the other hand, the IDRiD dataset concentrates on DR in Indian patients [31]. The terms of use for the dataset should be followed by researchers, and they even need to preprocess images to make them compatible with particular ML models.

Figure 4
Different stages of Macular edema: (I) Stage 0, (II) Stage I, and (III) Stage II

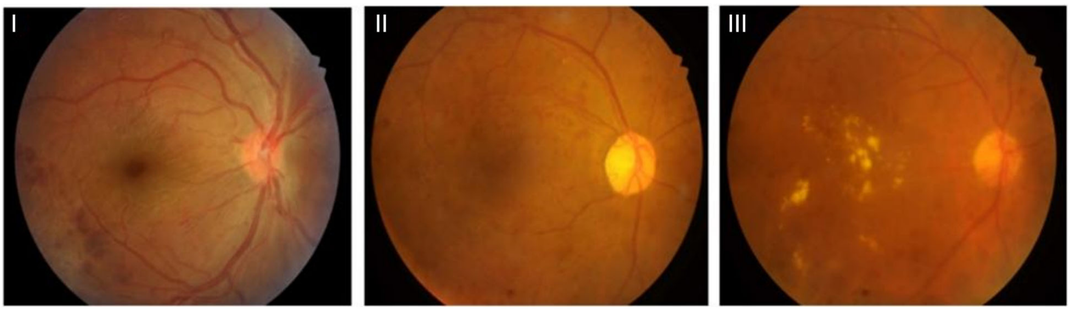


Table 1
Comparison of pre-trained models for classification

Model	Architecture	Complexity	Accuracy	Parameters	Pros	Cons
AlexNet	Eight layers (5 conv, 3 fully connected)	High	Moderate	~62 million	Effective feature extraction	Limited depth compared to newer models
VGGNet	CNN – 19 layers with 3 × 3 convolutional filters	High	High	~138 million (VGG16)	Simplicity, easy to understand.	Prone to overfitting due to a large number of parameters.
Inception	CNN – Inception module with filters of various sizes	Moderate	High	~23 million (InceptionV3)	Efficient parameter usage, reduced overfitting.	Complex architecture may be challenging to train.
Xception	CNN with depth-wise separable convolutions	Moderate	High	~22 million (Xception)	Parameter-efficient architecture in terms of both speed and memory.	Computationally demanding, during training as separable convolutions' added depth and complexity.
ResNet	CNN – Residual connections to address vanishing gradient	Moderate	High	~25 million (ResNet50)	Very deep architectures, mitigates degradation	Increased complexity, may be computationally expensive.
DenseNet	CNN – Dense connectivity, each layer connected to every other	Moderate	Moderate	~6 million (DenseNet121)	Improved parameter efficiency, reduced vanishing gradient.	Increased memory consumption due to dense connectivity
MobileNet	CNN – Depth-wise separable convolutions	Low	Moderate	~4 million (MobileNetV2)	Lightweight, suitable for mobile and edge devices	Sacrifices some accuracy compared to larger models
EfficientNet	CNN – Compound Scaling (depth, width, and resolution)	Low/ Moderate	High	Varies based on the version	State-of-the-art accuracy with fewer parameters	Training may require more resources compared to smaller models

3.2. Datasets for segmentation

Several datasets are used in the field of DR segmentation to train and assess segmentation algorithms. For segmentation, the same EyePACs can be used as was previously used in classification. Two well-known datasets for retinal vascular segmentation which is a critical component of DR segmentation are DRIVE [32] and DRIONS-DB [33]. Another useful resource that offers images for the segmentation of optic discs and retinal vessels is the STARE dataset [34]. These datasets are frequently used by researchers to create and evaluate segmentation algorithms. They are intended to detect and demarcate areas of interest (lesions or blood arteries within retinal pictures) accurately. However, it is critical to follow dataset usage standards. The data must be preprocessed to make it compatible with the model. Once the images are preprocessed to ensure consistent format and quality, an appropriate pre-trained CNN model is chosen in the next step. The model is then fine-tuned to focus on identifying patterns associated with different eye diseases. To enhance the model's resilience and ability to generalize the new data, data augmentation techniques can be applied to training data. The performance of the model is then assessed using critical metrics like accuracy, precision, recall, and $F1$ score. The discriminatory power can also be evaluated by calculating the area under the curve (AUC) and creating receiver operating characteristic (ROC) curves. The effectiveness of the pre-trained model can be confirmed through comprehensive comparisons with other AI models and traditional diagnostic methods.

The objective of this study is to enhance the diagnostic tools for eye health by utilizing the potential of pre-trained CNN models through transfer learning to provide an AI-driven solution for the automated categorization of ocular illnesses. The steps involved in AI-driven classification are as follows.

a. Data collection and pre-processing

A dataset of annotated fundus images covering multiple ocular disorders, including glaucoma, DR, and cataracts, has been gathered. To ensure consistency in quality, resolution, and format, a comprehensive pre-processing method is needed. This encompasses optimizing the images by removing artifacts and inconsistencies. It also involves multiple techniques for noise reduction and contrast enhancement.

b. Selection of pre-trained models

The selection of the appropriate pre-trained model involves picking a model that has been trained on large datasets such as ImageNet. The chosen model is evaluated based on its ability to capture related features and meet the criteria for classifying eye diseases. A thorough evaluation is conducted considering the functionality and design of the models to determine the best pre-trained model for transfer learning.

c. Transfer learning with fine-tuning

The selected model is initialized and its architecture is tuned to suit the objective of classifying eye diseases. To adapt the model for the classification of eye diseases, the modifications are incorporated in the fully connected layer and output layer. Then, transfer learning is applied based on the characteristics learned from the model. The annotated image collection is then used to refine the model's parameters to enable it to become more proficient at identifying patterns indicative of different ocular illnesses.

d. Training with validation

A balanced partition of various eye conditions is ensured in the training, validation, and testing data of the dataset. These subsets ensure a fair and unbiased evaluation of the model. The modified model integrating data augmentation is trained on the designated training data. The addition of data augmentation is aimed to improve robustness and enable it to successfully generalize unseen data. On the validation subset, the model's performance is thoroughly assessed. The iterative process of hyperparameter tuning is accomplished to maximize classification accuracy.

e. Evaluation and performance metrics

The pre-trained model's performance was assessed on the test set using essential metrics including accuracy, precision, recall, and $F1$ score for every category of ocular condition [10, 35, 36]. To evaluate the discriminating power of the model, ROC curves were created and the AUC was computed. The pre-trained model's advantage in automating the classification of ocular disorders was validated by comparative comparisons with other AI models and traditional diagnostic techniques [37]. The comparison of the pre-trained DL models for image classification is presented in Table 1.

Pre-trained CNN architectures such as AlexNet, VGG, Inception, Xception, ResNet, DenseNet, MobileNet, and EfficientNet are briefed in the table. AlexNet and VGG having millions of parameters and moderate to high processing needs are known for their depth and simplicity respectively. Inception and Xception use complex architectures to make effective use of parameters. ResNet uses residual connections to reduce the effects of vanishing gradients. DenseNet increases memory usage while improving parameter efficiency because of its extensive connection. EfficientNet provides advanced accuracy with fewer parameters, but training could need more resources. MobileNet trades accuracy for a lightweight architecture suitable for mobile devices. Vision transformer is a recently introduced architectures that show promise in image classification problems. It uses a self-attention mechanism akin to transformers in natural language processing. Furthermore, models like Swin Transformer and MLP-Mixer are attracting attention due to their creative architectural layouts and ability to process both token-wise and spatial data which might boost computer vision research further [38].

CNN architectures also provide new avenues for improving segmentation performance in the field of image segmentation. These models can increase segmentation performance and accuracy by combining creative architectural layouts with self-attention techniques. The steps for AI-based segmentation are as follows:

- Data Collection and Preprocessing: A substantial collection of data must be gathered to train the model. This collection must have a diverse array of images that illustrate different situations and anomalies. The images should be preprocessed to ensure consistent and ready for input into the model. This covers tasks like scaling, normalization, noise reduction, etc.,
- Annotation: The pictures must be annotated to generate ground truth data for the training of the segmentation model. This means defining the regions of interest in the images that the model will be taught to distinguish and identify.
- Model Selection: This involves a selection of an appropriate model for segmentation like U-Net, Fully Convolutional Network, or DeepLab [39]. Wide-ranging models with varied architecture and complexity are available for this task.
- Training with validation: The selected model is trained on the annotated dataset. The images are fed into the model and its parameters are adjusted during training to accurately segment the regions of interest. The validation set is extracted from the dataset to ensure that the trained model performs well on new and untested data. Tuning the hyperparameters is essential for improving the performance of the model.
- Implementation and Assessment: The model is used to segment new images after it has been trained and verified. This involves integrating the model into a software application or a more comprehensive medical imaging system. Lastly, new data is continuously added to assess the performance of the model and modifications are made as necessary. The comparison of the available pre-trained DL models for image classification is presented in Table 2.

The table lists several segmentation models with distinct architecture and sets of trade-offs. Both U-Net and SegNet use encoder-decoder architecture. SegNet is strong in capturing spatial information but struggles with complex object boundaries. U-Net excels in segmentation but lacks context for large objects. DeepLabV3 uses spatial pyramid pooling and atrous convolution to produce high-resolution outputs, although it needs large processing power. Mask R-CNN combines object recognition with instance segmentation for accurate results with its slower inference speeds. FCN with end-to-end convolutional architecture preserves spatial information but may result in coarse segmentations due to downsampling.

4. Results and Discussion

AI has substantially improved medical image analysis, especially classification and segmentation tasks. This discipline has the potential to fundamentally change diagnostic and treatment approaches. This section discusses the results of employing AI for medical image segmentation and classification.

Table 2
Comparison of deep learning models for segmentation

Model	Architecture	Pros	Cons
U-Net	Encoder-decoder with skip connections	Effective for medical image segmentation	Limited context information, may struggle with large objects
SegNet	Encoder-decoder with max-pooling indices	Captures spatial information, good for road scenes	Limited performance on complex object boundaries
DeepLabV3	Atrous convolution, spatial pyramid pooling	High-resolution output, good for fine segmentation	Computationally expensive, may require substantial resources
Mask R-CNN [40]	Combines object detection with instance segmentation	Precise instance segmentation, versatile	Slower inference speed compared to simpler models
FCN (Fully Convolutional Network) [41]	End-to-end convolutional architecture	Retains spatial information, adaptable to input sizes	May suffer from coarse segmentation due to downsampling

Figure 5
Accuracy and loss graph during training

Table 3
Classification summary of the pre-trained model

Eye diseases	Precision	Recall	*f*1-score
normal	0.91	0.85	0.88
cataract	0.97	0.89	0.93
diabetic_retinopathy	0.98	0.99	0.98
glaucoma	0.83	0.95	0.88

- Classification: The study found that AI-based classification using pre-trained DL architectures such as MobileNet correctly categorized fundus images into multiple categories. The accuracy and loss per epoch graph throughout the training process evaluates the training procedure's efficacy and convergence, as shown in Figure 5. The graph demonstrates the model's learning by displaying an increase in accuracy and decreases in loss across successive epochs.

The accuracy curve indicates the percentage of correctly categorized occurrences, whereas the loss curve reveals the model's inaccuracy. Analyzing these curves allows us to examine the convergence and stability of the training process. Table 3 shows the full results of classifying eye diseases using a pre-trained network.

The model performed admirably in the classification of eye diseases and obtained the highest accuracy of 98% in detecting "DR" and 83% in detecting "glaucoma" which is the least. The effectiveness of AI in automating the classification of medical pictures is demonstrated by the high recall and precision rates seen in our classification findings. Furthermore, by identifying distinct patterns of misclassification across various classes, linking these measures with a confusion matrix provides deeper insights into the model's performance and helps identify areas of success and errors. The confusion matrix for the classification is presented in Figure 6.

- Segmentation: This research explores the use of AI-driven segmentation techniques and shows promise in precisely identifying and dividing areas of interest within medical images.

Figure 6
Confusion matrix for the classification of eye diseases

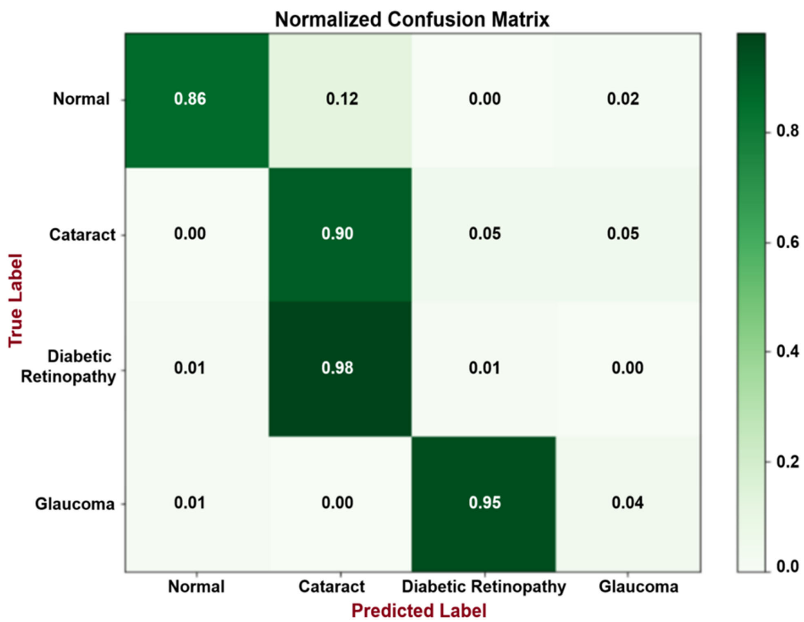
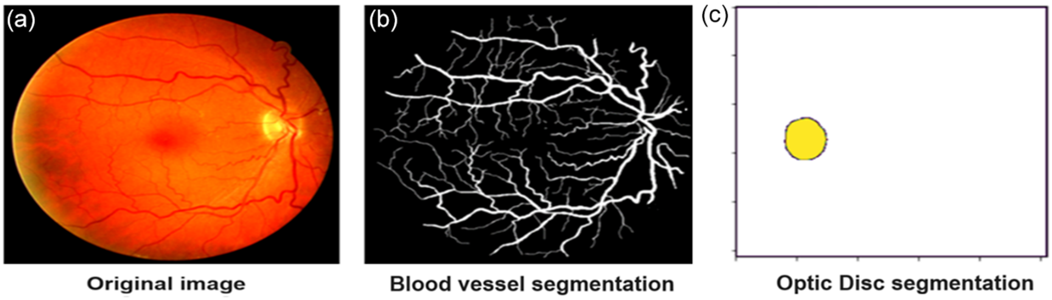


Figure 7
Segmentation results of eye diseases



The trained segmentation models outline anatomical structures, abnormalities, and lesions with remarkable precision, suggesting that they could help medical professionals locate and analyze volumes of interest with greater accuracy. This could result in better surgical techniques, treatment plans, and continuous disease surveillance. All of these might lead to better patient outcomes. The results of segmentation using U-net architecture are shown in Figure 7.

The findings highlight how AI is revolutionizing medical image processing. The effective use of AI in segmentation and classification tasks has the potential to enhance clinical judgment, maximize resource allocation, and enhance patient care. AI has several drawbacks, including interpretability, validation, and moral questions concerning patient privacy and data security. These are issues that must be resolved before implementing AI in the healthcare industry.

5. Conclusion and Future Scope

The application of AI in fundus image classification and segmentation is an advancement in the field of medical imaging and diagnostics. This study demonstrates the potential of AI to automate the interpretation and analysis of medical images. It aids in the detection of various ocular conditions such as cataracts, DR, and glaucoma. An effective integration of AI into medical imaging improves the effectiveness and precision of identifying and treating eye illnesses. It can automate the process, resulting in improved patient care and outcomes. In this study, AI-based segmentation and classification algorithms collaborate to provide a comprehensive framework for extracting useful information from medical photos. The framework provides doctors with critical information for diagnosis and predictions.

The widespread use of AI in medical imaging needs thoughtful consideration of ethical, legal, and interpretability issues. Concerns around computational openness, data privacy, and the appropriate application of AI technology in clinical practice must be addressed. Further study and collaboration across interdisciplinary teams are also necessary to improve AI models. It is also critical to prove their efficacy across a variety of patient populations. Continuous evaluation and development are required to increase their efficacy and usefulness in clinical practice.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in public repository as cited and also available on request from the corresponding author upon reasonable request.

Author Contribution Statement

Ankur Biswas: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization, Supervision, Project administration, Funding acquisition. **Rita Banik:** Methodology, Investigation, Visualization, Supervision, Project administration, Funding acquisition.

References

[1] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., . . . , & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, *35*, 18–31. https://doi.org/10.1016/j.media.2016.05.004

[2] Anaya-Isaza, A., Mera-Jiménez, L., & Zequera-Diaz, M. (2021). An overview of deep learning in medical imaging. *Informatics in Medicine Unlocked*, *26*, 100723. https://doi.org/10.1016/j.imu.2021.100723

[3] Sanghavi, J., & Kurhekar, M. (2024). Ocular disease detection systems based on fundus images: A survey. *Multimedia Tools and Applications*, *83*(7), 21471–21496. https://doi.org/10.1007/s11042-023-16366-x

[4] Sengar, N., Joshi, R. C., Dutta, M. K., & Burget, R. (2023). EyeDeep-Net: A multi-class diagnosis of retinal diseases using deep neural network. *Neural Computing & Applications*, *35*(14), 10551–10571. https://doi.org/10.1007/s00521-023-08249-x

[5] Egger, J., Gsaxner, C., Pepe, A., Pomykala, K. L., Jonske, F., Kurz, M., . . . , & Kleesiek, J. (2022). Medical deep learning—A systematic meta-review. *Computer Methods and Programs in Biomedicine*, *221*, 106874. https://doi.org/10.1016/j.cmpb.2022.106874

[6] Hu, D., Liang, K., Zhou, S., Tu, W., Liu, M., & Liu, X. (2023). scDFC: A deep fusion clustering method for single-cell RNA-seq data. *Briefings in Bioinformatics*, *24*(4), bbad216. https://doi.org/10.1093/bib/bbad216

[7] Kumar, S. V. M., & Gunasundari, R. (2023). Computational intelligence in eye disease diagnosis: A comparative study.

Medical & Biological Engineering & Computing, *61*(3), 593–615. https://doi.org/10.1007/s11517-022-02737-3

[8] Muchuchuti, S., & Viriri, S. (2023). Retinal disease detection using deep learning techniques: A comprehensive review. *Journal of Imaging*, *9*(4), 84. https://doi.org/10.3390/jimaging9040084

[9] Lee, Y., Ryu, J., Kang, M. W., Seo, K. H., Kim, J., Suh, J., . . . , & Han, S. S. (2021). Machine learning-based prediction of acute kidney injury after nephrectomy in patients with renal cell carcinoma. *Scientific Reports*, *11*(1), 15704. https://doi.org/10.1038/s41598-021-95019-1

[10] Dai, L., Wu, L., Li, H., Cai, C., Wu, Q., Kong, H., . . . , & Jia, W. (2021). A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature Communications*, *12*(1), 3242. https://doi.org/10.1038/s41467-021-23458-5

[11] Bhandari, S., Pathak, S., & Jain, S. A. (2023). A literature review of early-stage diabetic retinopathy detection using deep learning and evolutionary computing techniques. *Archives of Computational Methods in Engineering*, *30*(2), 799–810. https://doi.org/10.1007/s11831-022-09816-6

[12] Saminu, S., Xu, G., Zhang, S., Kader, I. A. E., Aliyu, H. A., Jabire, A. H., . . . , & Adamu, M. J. (2023). Applications of artificial intelligence in automatic detection of epileptic seizures using EEG signals: A review. *Artificial Intelligence and Applications*, *1*(1), 11–25. https://doi.org/10.47852/bonviewAIA2202297

[13] Simon, K., Vicent, M., Addah, K., Bamutura, D., Atwiine, B., Nanjebe, D., & Mukama, A. O. (2023). Comparison of deep learning techniques in detection of sickle cell disease. *Artificial Intelligence and Applications*, *1*(4), 252–259. https://doi.org/10.47852/bonviewAIA3202853

[14] Bali, A., & Mansotra, V. (2024). Analysis of deep learning techniques for prediction of eye diseases: A systematic review. *Archives of Computational Methods in Engineering*, *31*(1), 487–520. https://doi.org/10.1007/s11831-023-09989-8

[15] Tajudin, N. M. A., Kipli, K., Mahmood, M. H., Lim, L. T., Awang Mat, D. A., Sapawi, R., . . . , & Hoque, M. E. (2022). Deep learning in the grading of diabetic retinopathy: A review. *IET Computer Vision*, *16*(8), 667–682. https://doi.org/10.1049/cvi2.12116

[16] Jabbar, M. K., Yan, J., Xu, H., Rehman, Z. U., & Jabbar, A. (2022). Transfer learning-based model for diabetic retinopathy diagnosis using retinal images. *Brain Sciences*, *12*(5), 535. https://doi.org/10.3390/brainsci12050535

[17] Li, M., Jiang, Y., Zhang, Y., & Zhu, H. (2023). Medical image analysis using deep learning algorithms. *Frontiers in Public Health*, *11*, 1273253. https://doi.org/10.3389/fpubh.2023.1273253

[18] Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., . . . , & Mellit, A. (2023). A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. *Sustainability*, *15*(7), 5930. https://doi.org/10.3390/su15075930

[19] Iqbal, S., Qureshi, A. N., Li, J., & Mahmood, T. (2023). On the analyses of medical images using traditional machine learning techniques and convolutional neural networks. *Archives of Computational Methods in Engineering*, *30*(5), 3173–3233. https://doi.org/10.1007/s11831-023-09899-9

[20] Iman, M., Arabnia, H. R., & Rasheed, K. (2023). A review of deep transfer learning and recent advancements. *Technologies*, *11*(2), 40. https://doi.org/10.3390/technologies11020040

[21] Theodoris, C. V., Xiao, L., Chopra, A., Chaffin, M. D., Al Sayed, Z. R., Hill, M. C., . . . , & Ellinor, P. T. (2023). Transfer learning enables predictions in network biology. *Nature*, *618*(7965), 616–624. https://doi.org/10.1038/s41586-023-06139-9

[22] Biswas, A., & Banik, R. (2023). Estimating risk levels and epidemiology of diabetic retinopathy using transfer learning. In *2023 Third International Conference on Secure Cyber Computing and Communication*, 287–292. https://doi.org/10.1109/ICSCCC58608.2023.10176908

[23] Aljabri, M., & AlGhamdi, M. (2022). A review on the use of deep learning for medical image segmentation. *Neurocomputing*, *506*, 311–335. https://doi.org/10.1016/j.neucom.2022.07.070

[24] Nagamani G. M., & Rayachoti, E. (2024). Deep learning network (DL-Net) based classification and segmentation of multi-class retinal diseases using OCT scans. *Biomedical Signal Processing and Control*, *88,* 105619. https://doi.org/10.1016/j.bspc.2023.105619

[25] Liu, X., Song, L., Liu, S., & Zhang, Y. (2021). A review of deep-learning-based medical image segmentation methods. *Sustainability*, *13*(3), 1224. https://doi.org/10.3390/su13031224

[26] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention: 18th International Conference,* 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

[27] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

[28] Dugas, E., Jorge, J., & Cukierski, W. (2015). *Diabetic retinopathy detection*. Retrieved from: https://www.kaggle.com/c/diabetic-retinopathy-detection

[29] Karthik, M., & Dane, S. (2019). *APTOS 2019 blindness detection*. Retrieved from: https://www.kaggle.com/competitions/aptos2019-blindness-detection

[30] Almazroa, A., Alodhayb, S., Osman, E., Ramadan, E., Hummadi, M., Dlaim, M., . . . , & Lakshminarayanan, V. (2018). Retinal fundus images for glaucoma analysis: The RIGA dataset. *Proccedings of Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, *10579*, 105790B. https://doi.org/10.1117/12.2293584

[31] Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., & Meriaudeau, F. (2018). Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research. *Data*, *3*(3), 25. https://doi.org/10.3390/data3030025

[32] Staal, J., Abràmoff, M. D., Niemeijer, M., Viergever, M. A., & van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, *23*(4), 501–509. https://doi.org/10.1109/TMI.2004.825627

[33] Carmona, E. J., Rincón, M., García-Feijoó, J., & Martínez-de-la-Casa, J. M. (2008). Identification of the optic nerve head with genetic algorithms. *Artificial Intelligence in Medicine*, *43*(3), 243–259. https://doi.org/10.1016/j.artmed.2008.04.005

[34] Hoover, A. D., Kouznetsova, V., & Goldbaum, M. (2000). Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, *19*(3), 203–210. https://doi.org/10.1109/42.845178

[35] Qayyum, W., Ehtisham, R., Bahrami, A., Camp, C., Mir, J., & Ahmad, A. (2023). Assessment of convolutional neural network pre-trained models for detection and orientation of cracks. *Materials*, *16*(2), 826. https://doi.org/10.3390/ma16020826

[36] Park, D. J., Park, M. W., Lee, H., Kim, Y. J., Kim, Y., & Park, Y. H. (2021). Development of machine learning model for diagnostic disease prediction based on laboratory tests. *Scientific Reports*, *11*(1), 7567. https://doi.org/10.1038/s41598-021-87171-5

[37] Stančić, A., Vyroubal, V., & Slijepčević, V. (2022). Classification efficiency of pre-trained deep CNN models on camera trap images. *Journal of Imaging*, *8*(2), 20. https://doi.org/10.3390/jimaging8020020
[38] Ishihara, K., & Matsumoto, K. (2023). Comparing the robustness of ResNet, Swin-Transformer, and MLP-Mixer under unique distribution shifts in fundus images. *Bioengineering*, *10*(12), 1383. https://doi.org/10.3390/bioengineering10121383
[39] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 834–848. https://doi.org/10.1109/TPAMI.2017.2699184
[40] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. *arXiv Preprint*: *1703.06870.*
[41] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *arXiv Preprint: 1411.4038.*

How to Cite: Biswas, A., & Banik, R. (2024). Advancements in Medical Image Analysis: A Comprehensive Method of AI-Based Classification and Segmentation Technique. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA42022106