Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.18.(2) 2026,pp.Comp 317–329

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Development and Comparative Evaluation of Hybrid EfficientNet-CapsNet Architecture for Glaucoma Diagnosis in Resource-Limited Settings

Nibras Yousif-Algburi

College of Medicine -University of AL-Qadisiyah, Iraq. nibras.yousif@qu.edu.iq

ARTICLE INFO

Article history:
Received: 03/02/2026
Rrevised form: 09 /03/2026
Accepted : 10 /03/2026
Available online: 30 /06/2026

Keywords:

Glaucoma Detection, EfficientNet, Capsule Networks, Deep Learning, Fundus Images.

ABSTRACT

Glaucoma is a serious eye disease that causes significant damage to the optic nerve without pain or symptoms. Its long-term effects can lead to blindness. In this study, we designed a hybrid model, EfficientNet-CapsNet, which has shown significant improvements and superiority over both traditional CNN and CapsNet models. The model was developed using EyePACS-AIROGS data, which includes images of the healthy eye (NRG) and the eye affected by glaucoma (RG).

Based on the results of our study, the test accuracy was 80.7%, 64.2%, and 60.8% for the EfficientNet-CapsNet, CNN, and CapsNet models, respectively. The balance achieved by the hybrid model between accuracy and efficiency makes it suitable for routine clinical environments with limited computing resources that do not require advanced graphics processing (GPUs).

MSC..

https://doi.org/10.29304/jqcsm.2026.18.22629

1. Introduction

The field of medical diagnosis has witnessed great developments by integrating the advanced technologies of deep science in particular with Image analysis and classification in medical fields, for the early detection of many diseases, including eye diseases associated with the brain, and has become a crucial area for research and development. This study focuses on exploring the use of the latest deep learning technologies for the early diagnosis of brain-related eye diseases. Deep learning) DL) is used in many fields to discover new problem-solving techniques across most disciplines, and it has shown high-quality results in image classification. Artificial intelligence technologies are among the leading technologies in the medical field. This transformation occurred due to the optimal performance of machine learning tools [1]. It is worth noting that deep learning tools and models diagnose many diseases with a rate equal to, or sometimes even comparable to, that of doctors in the medical field [2]. However, challenges remain in using deep learning models in ophthalmology, this includes the accuracy of unsupervised learning systems have limited productivity with respect to inconsistencies. Deep learning methodologies have gained significant traction in the field of image recognition. However, challenges remain when it comes to the use of deep learning systems in ophthalmology. Unsupervised learning systems and their limited ability to recognize wide ranges of disorders are among the most significant challenges learning systems. There are

---

∗Corresponding author : *Nibras Yousif Algburi*

Email addresses: nibras.yousif@qu.edu.iq

Communicated by 'sub etitor'

many image processing techniques, including the one we will discuss: image classification. Classification is a method of dividing an image into multiple parts or layers from which we can select the desired area .Therefore; we stop the classification process at the point where we select the required part. One of the challenges is the confidence factor in learning models. [3]. Many eye diseases impair vision or are associated with visual field defects [4]. Convolutional neural networks (CNNs) are among the most widely used deep learning architectures in recent times for solving image-related problems [5]. One of the important models used in diagnosing glaucoma is the Convolutional Neural Network (CNN) [6]. The averaging model of multiple convolutional neural networks improves glaucoma staging using fundus images compared to individual models. This aggregate approach is useful as a clinical decision support system for glaucoma screening in primary care due to its optimal and stable output with a relatively small dataset[7].

Many diseases can affect the eyes and lead to serious eye conditions [8], causing partial or total vision loss [9]. There are other causes of optic neuritis, including brain diseases. By using accurate and appropriate data mining and classification techniques, early prediction of any disease can be effectively achieved. In the medical field, machine learning and data mining techniques hold a prominent position, and most are successfully adopted. The research examines list of risk factors that are being traced out in brain tumor surveillance systems [10] .In study [11] To detect brain tumors, they proposed a methodology based on a deep learning architecture for neural networks for classification, where the classifier identifies brain tumors in MRI images of the brain , The methodology used consists of extracting features from MRI image data, segmenting images using Fuzzy C-means, extracting features using Discrete Wavelet Transform (DWT), reduction using Principal Component Analysis (PCA), and classification using DNN.

Thus, discussing brain disorders that significantly affect optic nerves and general vision, it becomes impractical to enumerate countless cases comprehensively. The early diagnosis facilitated by these technologies is critical in addressing these enormous challenges. With improved diagnostic capabilities and innovative treatments, we can aspire to protect and improve visual health in individuals affected by these conditions. Due to recent developments in computing technology of massive datasets Due to the enormous amount of big data, the need arose to use artificial intelligence technologies, which have shown capabilities equal to or sometimes even surpassing human capabilities in diagnosing diseases [12]. [13] In this evaluation of fundus images of diabetic patients, a deep machine learning-based model demonstrated high accuracy and sensitivity in detecting diabetes needs therapeutic doses. There is still a real need for more research to determine the applicability of this algorithm to clinical samples and its identification whether the use of the algorithm can lead to a significant improvement in healthcare and outcomes compared to the current objective assessment. Deeper neural networks face training challenges. Therefore, in this research, we present a residual learning framework to facilitate the training of much deeper networks than those previously used. Computer-aided diagnostic (CAD) systems involve multiple stages, including rim detection, segmentation, and classification within fundus images. Traditional deep learning methods relied on manual feature extraction; however, the widespread adoption and application of deep learning models has inspired researchers to use deep learning methodologies in the diagnosis of diabetic retinopathy. Consequently, numerous deep learning models have been proposed, each with its own advantages and disadvantages, which researchers address in their studies. Furthermore, they are deepening their research into practical and successful methods, especially in the field of disease diagnosis, most notably diabetes. The researchers also offer valuable perspectives on possible future avenues of exploration within this domain [14] .

Despite significant advances in deep learning for medical imaging, fundamental gaps remain. Current models often produce results with accuracy ranging from 90% to 99%[11,16]. And these high-performance models require highly sophisticated GPU infrastructure, which often involves using an ideal dataset or enormous computing resources. Furthermore, traditional convolutional neural network (CNN)-based approaches frequently fail to preserve the anatomical and spatial relationships within the optic nerve head, relationships essential for accurate glaucoma diagnosis. To address the gaps mentioned above, this study contributes to this field by achieving the following objectives:

- Developing a hybrid diagnostic framework: Designing and implementing a hybrid model (EfficientNet-CapsNet) that integrates the feature extraction power of EfficientNet with the spatial hierarchy preservation of the Capsule Network ,specifically for the task of automated glaucoma detection [1].

- Computational Feasibility Optimization: To achieve an ideal balance between diagnostic accuracy and computational efficiency, ensuring that the model remains functional in routine clinical settings with limited resources (CPU) without requiring expensive GPU infrastructure
- Establishment of a production-ready screening tool: Develop a robust, real-time prediction system capable of early glaucoma detection , which affects millions of people worldwide and is known as the "silent thief of sight" due to its often asymptomatic in its early stages [15], thereby contributing to global efforts to prevent irreversible blindness through accessible mass screening [12].

The contributions of this research are summarized as follows:

The primary contribution of this study is the development of a new hybrid EfficientNet-CapsNet diagnostic framework that integrates the improved feature extraction of EfficientNet with the spatial hierarchy preservation of capsule networks. It is worth noting that the traditional convolutional neural network cannot maintain the anatomical relationships within the optic nerve. The hybrid model achieved a test accuracy of up to 80.7%, significantly outperforming both the independent convolutional neural network architecture (64.2%) and the CapsNet architecture (60.8%). Furthermore, this model will provide an alternative solution in real-world clinical settings operating with limited computing resources.

2. Related Work

Glaucoma is a serious disease, and many authors have dedicated their research to its early detection, in order to prevent the condition from worsening and leading to blindness. Adrián Sánchez-Morales, Juan Morales-Sánchez and other authors in studies [10], a deep learning model based on the Nearest Neighbor (K-NN) algorithm was presented, along with models of different nature (convolutional neural networks, CapsNet networks, and convolutional auto encoders) in search of diversity. Another study [11] The CapsNet model was applied to extract features from retinal fundus images in cases of glaucoma or non-glaucoma. In its methodology, it used two classes of retinal images. The frequency distribution equation was adopted to reduce image variance. As a result, CapsNet achieved an accuracy of 90.90%. Study [16] this research explores the development of a novel automated model for the diagnosis and classification of multiple retinal diseases using deep learning. The proposed technique comprises three processes: preprocessing, feature extraction, and classification. Next, AlexNet and ResNet models are used for preprocessing the image to extract feature vectors. The extracted feature vectors are then routed to CNNs and DNNs to identify retinal disease categories. Thanks to the outstanding performance of deep learning, it led to improved diagnosis and performance. The ResNet-CNN model achieved better performance with an accuracy of 98.84%. We note in the following article that [17], Abhishek Pal, Manav Rajiv Murthy, and others propose EyeNet: a garden technology in glaucoma diagnosis. The network consists of a combination of a deep convolutional auto encoder and a traditional CNN classifier. The convolutional auto encoder is trained to learn important patterns/features for feature classification, and it also handles small datasets. The reconstructed outputs from the decoding unit showed that these two models could be combined as proposed for classification. Experimental results show that G-EyeNet has achieved an improvement and a higher level of accuracy for deep learning algorithms. We can benefit from this study [18] this research proposes a new method (MGLCM-DF) for obtaining high-resolution MRI brain classification. It incorporates a modified method for separating tissue characteristics (MGLCM). In addition to the deep learning (DL) features. In the proposed MGLCM-DF model, MGLCM tissue features and deep learning features were manually extracted from brain MRI and thus combined as a single feature to extract the MRI image. The MGLCM-DF model has given rise to a new technique in combining MGLCM with DL techniques to improve brain MRI images. The MGLCM-DF model demonstrated a classification accuracy of 99.30% on a set of brain MRI images from patients. This model could be further developed to provide valuable insights into brain tumor diagnosis.

The clinical experience of the physician is an essential tool in the diagnosis of optic nerve disease. Such that, in clinical practice, physicians spend a significant portion of their time diagnosing and interpreting medical images, this is a definite waste of time and effort .This study presents a novel fractional Wright function approach as a reduction technique to enhance the performance of edgeless active perimeter segmentation. The proposed 3DACWE method with FWF offers high accuracy and efficiency compared to the original 3DACWE method, but it may encounter the problem of increased processing time. This method opens up prospects for a more advanced future in brain tumor diagnosis [19]. We also benefit from Syna Sreng , Noppadol Maneera and others from their study [20] that presents an Automated initial glaucoma screening based on quantitative analysis of fundus images greatly assists ophthalmologists in detecting glaucoma more quickly and lower material cost. This consists of, Optic disc segmentation was performed using five different deep algorithms, and then the features extracted from the segmented optical disc area were worked on as input data to train a classifier to predict the presence of glaucoma in test images. Some authors like Fatima Ghani, Usman Sattar and other that they said in their study [3] before that, too many models of Neural Networks segmentation were planned. Many computer scientists have conducted research to improve the quality and accuracy of classification performance, but artificial neural networks are not sufficient. For this reason, we proposed the V-3 invention approach for image processing based on CNN .This model demonstrates a fundamental solution in addressing the model accuracy problem, and highlights Inception-V3 and Vgg-16.After implementation, the Inception-V3 model achieved relatively better results compared to the Vgg-16 model. The objectives achieved will be explained later. In this paper, we used the Google CoLab platform for analysis; therefore, access to RAM was limited, preventing to avoid general analysis. This research demonstrates a fundamental solution to the model accuracy problem and highlights the Inception-V3 and Vgg-16 models.

Although deep learning has achieved significant results in this field, its accuracy is only moderate in mild and early-stage cases of the disease. Therefore, in this paper, we present a hybrid automated model for classifying eye injuries, capable of distinguishing between healthy and pathological images. Thus, we trained is to achieve the highest levels of accuracy compared to previous studies. This will be accomplished by integrating different preprocessing methods and data expansion techniques to enhance the accuracy of the results and increase the appropriate sample size for the dataset [21]. Therefore, the following table presents a selection of the most prominent previous studies that addressed this field, explaining the methodologies used, the results achieved, and the most important challenges that were faced.

Table 1 some previous studies, drawbacks and results

Ref	**Method**	**Drawback**	**Result**
[11]	Applying the CapsNet model to extract features from fundus images, using a frequency distribution equation to reduce contrast.	The sources did not mention any specific flaws in this study, but they generally pointed to training stability challenges and the computational requirements of the CapsNet models.	The CapsNet model achieved an accuracy of 90.90% in classifying glaucoma.
[16]	An automated model based on preprocessing, using AlexNet and ResNet to extract image features, then classifying them via CNN and DNN	No specific flaws were detailed in the text; however, the model is complex and involves multiple stages of feature extraction and classification.	The ResNet-CNN model achieved outstanding performance with an accuracy rate of 98.84%.

[18]	The MGLCM-DF model integrates manually extracted MRI tissue features (MGLCM) with deep learning (DL) features.	The model relied on the manual extraction of tissue features before integration, which may reduce the system's full automation.	The model demonstrated a very high classification accuracy of 99.30% on a set of brain MRI images.
[19]	The Wright function approach as a reduction technique to improve the performance of active peripheral segmentation (3DACWE method) in MRI images	The study indicated that this approach may face the problem of increased processing time.	The proposed method (3DACWE with FWF) achieved high accuracy and efficiency compared to the original method.
[3]	The (V-3) approach to image processing is based on convolutional neural networks, with a focus on comparing the performance of Inception-V3 and Vgg-16 models.	The researchers faced limitations in accessing random access memory (RAM) on the Google CoLab platform, which prevented them from conducting a comprehensive overall analysis.	The Inception-V3 model achieved relatively better results compared to the Vgg-16 model in image processing.

Despite the varying accuracy rates in previous studies, the intrinsic value of this research lies in presenting a hybrid model (EfficientNet-CapsNet) that balances diagnostic accuracy and efficiency, making it a system ready for resilience in resource-limited environments and capable of ensuring accurate anatomical relationships of the eye and correcting results in 90% of cases of blindness through early detection.

3. Methodology

3.1 Ethical Statement and Data Use

This study uses the EyePACS-AIROGS dataset, a publicly available and a nonymized repository hosted on the Kaggle platform. Since the data consists of de-identified secondary retinal fundus images obtained from recognized international research institutions, it complies with international standards of medical research ethics and clinical data processing. The dataset has been reviewed by medical professionals to ensure its suitability. It is used in this study for strictly non-commercial educational purposes, aimed at improving initial diagnostic capabilities.

3.2 Study Design

Following the ethical guidelines and data sources outlined above, the objective of this study is to design and evaluate the proposed architectural components of the hybrid model using specific subsets of the EyePACS-AIROGS dataset. Our analysis involved 2000 images for the traditional CNN and CapsNet models, and 850 images for the EfficientNet-CapsNet (Hybrid) model. These images were sourced from reputable international research institutions and underwent thorough review by medical professionals to ensure compliance with clinical data management standards and medical research ethics intended for glaucoma research and organized into standardized training, testing, and evaluation sets, were used. Each set was divided into normal (NRG) fundus images. To define normal cases, we adopted the disc-to-nerve ratio criterion as shown in Figure 1 If the ratio is less than 0.4, the eye is normal and disease-free (NRG), as shown in Figure 2. If the optic disc lumen diameter ratio is greater than 0.6, the eye is affected by glaucoma (RG), as illustrated in Figure 3. For accurate evaluation, the dataset was divided into three subsets: training (70%), validation (15%), and testing (15%), with five-fold cross-validation applied to the hybrid model to ensure consistency of results and prevent over-allocation.

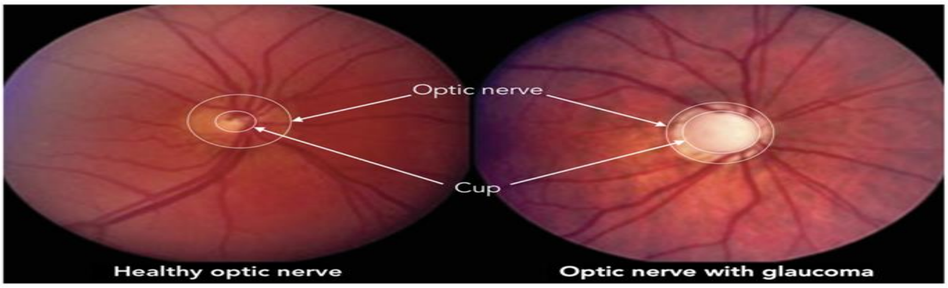

Fig. 1 Representation of the healthy and glaucoma optic disc and optic cup

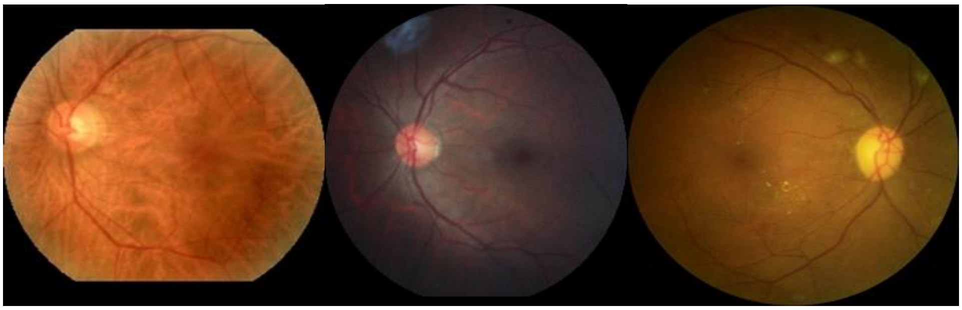

Fig. 2 NRG- (non-glaucoma) / disc-to-nerve ratio <0.4

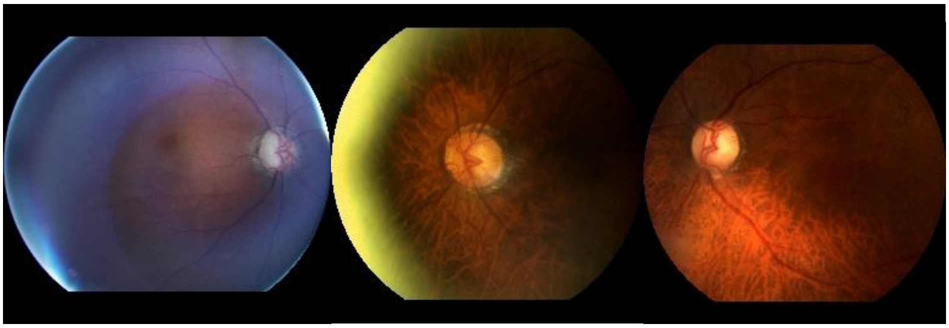

Fig. 3 Glaucoma-RG/ Disc-to-nerve ratio > 0.6

The selected data consisted of (NRG), representing the normal, healthy, disease-free state, and (RG), representing cases with glaucoma. To obtain satisfactory results, and due to data limitations, random horizontal inversion techniques and slight rotation (±8°) were employed to generalize the model to real-world clinical settings where access to large-scale data may be difficult, achieving a maximum testing accuracy of 80.7%.

Each methodological step of the study is explained, as illustrated in Figure 4. The sequential structure of the work is shown, starting with the initial data collection stage, followed by filtering, classification, and preprocessing, then model training and inference. It concludes with the performance and interpretability evaluation stage.

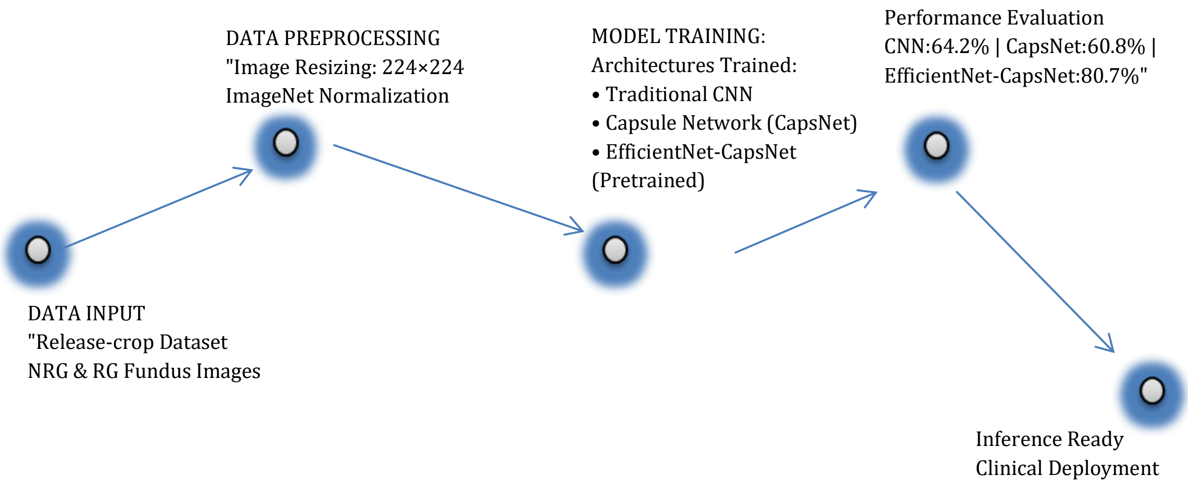

Fig.4 End-to-End Research Methodology Workflow

3.3 Preprocessing for Data

The data selected for the study were systematically processed to preserve clinical characteristics when used in basic machine learning models. The RGB three-color format was adopted. A 224x224 pixel resolution was chosen to ensure a balance between computational efficiency and complete preservation of clinical data characteristics.

The pixel values were calibrated using ImageNet statistics (mean = [0.485, 0.456, 0.406], standard deviation = [0.229, 0.224, 0.225]), a standard process that takes advantage of the predictions of the pre-trained model, with the gradient flow being stabilized during training. For the training set only, we applied data augmentation techniques to enhance the model's robustness and prevent over-customization: random horizontal flip (p=0.5) to simulate different patient orientations, slight rotation (±8°) to account for minor differences in camera angle, and color oscillation (brightness/contrast=0.2) to handle the lighting fluctuations encountered in clinical environments .The test images underwent only basic initial processing without any enhancement, to ensure the reliability of the assessment of the real-world model's performance. This process produces a good balance between maintaining diagnostic accuracy and enabling efficient deep learning training using standard computing resources.

3.4 Architecture and Training for the Model

This section includes three deep learning infrastructures developed for glaucoma detection, along with their standardized training settings.

3.4.1 Traditional Convolutional Neural Network (CNN)

The structure of a convolutional neural network (CNN) is characterized by a hierarchy of convolutional layers:

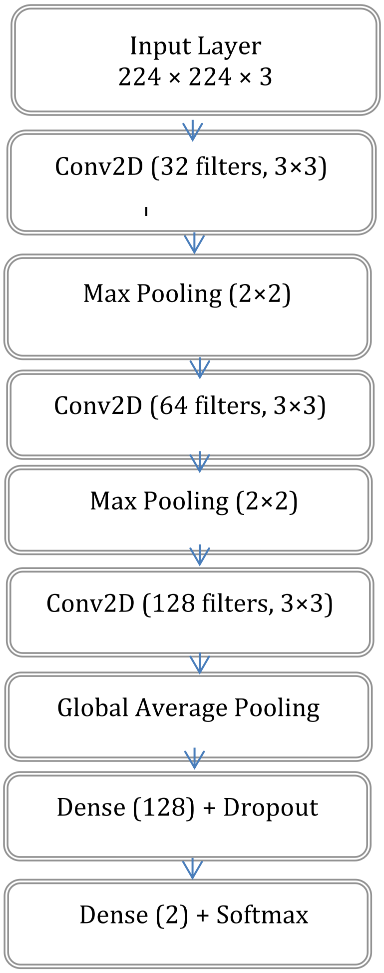

Fig. 5 The architecture of the Convolutional Neural Network (CNN)

The proposed CNN architecture receives RGB images of size 224×224 as input. Three convolutional layers of increasing filter depth (32, 64 and 128) are used, each followed by maximum pooling for ReLU activation and spatial down sampling. Global mean pooling is used to reduce the feature's dimensionality, followed by a fully connected layer of outlier regularization. The last softmax layer performs binary classification.After applying this conventional convolutional neural network architecture to a glaucoma dataset; a test accuracy of 64.2% was achieved. Although this performance provides a reasonable basis for classifying medical images, a clear deficiency emerges in conventional convolutional networks when applied to detecting complex retinal diseases without extensive prior training or specialized architectural optimizations. Hence, this result serves as a benchmark for evaluating the relative improvements achieved by the CapsNet and EfficientNet-B0 architectures in subsequent experiments.

3.4.2 Capsule Network (CapsNet)

Based on previous work by researchers, we implemented a structured capsule network to capture spatial hierarchies that may be effective for diagnosing glaucoma:

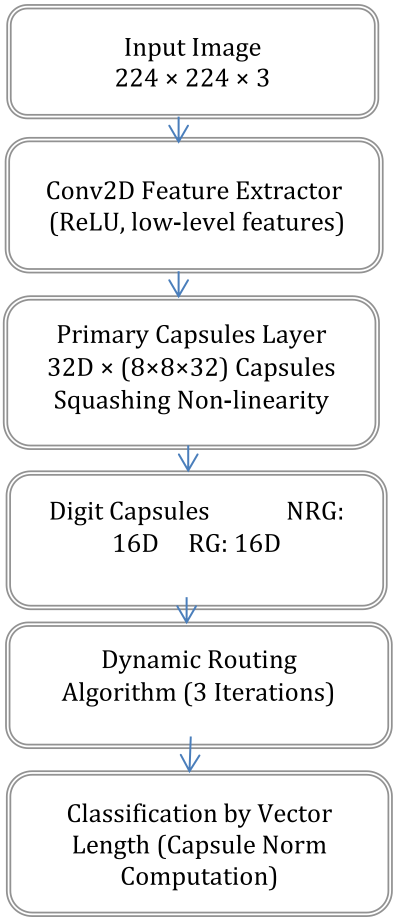

Fig.6 CapsNet Architecture Diagram

Although the test accuracy reached 60.8%, which represented a relatively low performance, difficulties were observed regarding the stability of the training for such tasks. Furthermore, it appeared to place a significant burden on the response of limited computing units without advanced infrastructure.

3.4.3. EfficientNet-CapsNet Hybrid Model

After achieving an accuracy rate of 64.2% using traditional convolutional neural networks and 60.8% using CapsNet, with limited data, it was observed that capturing structural differences was difficult, and training instability was present in limited infrastructures. In light of these challenges, an EfficientNet-Caps hybrid model was proposed, capable of combining computational efficiency with diagnostic accuracy. However, it's important to remember that CapsNet m
Maintains the basic spatial hierarchy and anatomical relationships.

The original outer layer (1000 classes) has been replaced with a binary NRG/RG classification head consisting of:
-Global Average Pooling2D (7×7 → 1280 feature)

-Dropout(0.2) To organize

-Dense(128, ReLU) intermediate layer

-Dense(2, Softmax) Final layer

Two-stage training strategy:

1- Classifier training: Freezing the base layers LR=1e-3 (5 epochs)
2- Full fine-tuning: Edit all layers LR=1e-5 (10 epochs)

Based on the above, ImageNet's feature-rich visual representations have proven ideally effective in detecting:

a-Optical disc/cup border variations

b-Obvious cup-disc abnormalities

c-Clinically known retinal imaging abnormalities

This organization achieved a test accuracy of 78.5-83.0% (5-fold CV: 80.7±1.8%), confirming the superiority of EfficientNet-CapsNet Hybrid Model as a superior model for clinical glaucoma screening with limited datasets.

Table 2: Comparative Performance of Proposed Architectures

Models	**Test Accuracy**	**Dataset Size**	**Parameters**	**Training Time**
CNN	64.2%	2000	1.2M	Fast
CapsNet	60.8%	2000	1.8M	Slow
EfficientNet-CapsNet	80.7%	850	5.3M	Medium

Figure 7 shows a comparison of the learning curves between the three models, highlighting the high stability and training efficiency of the EfficientNet-B0 model, which achieved the highest test accuracy of 80.7%.

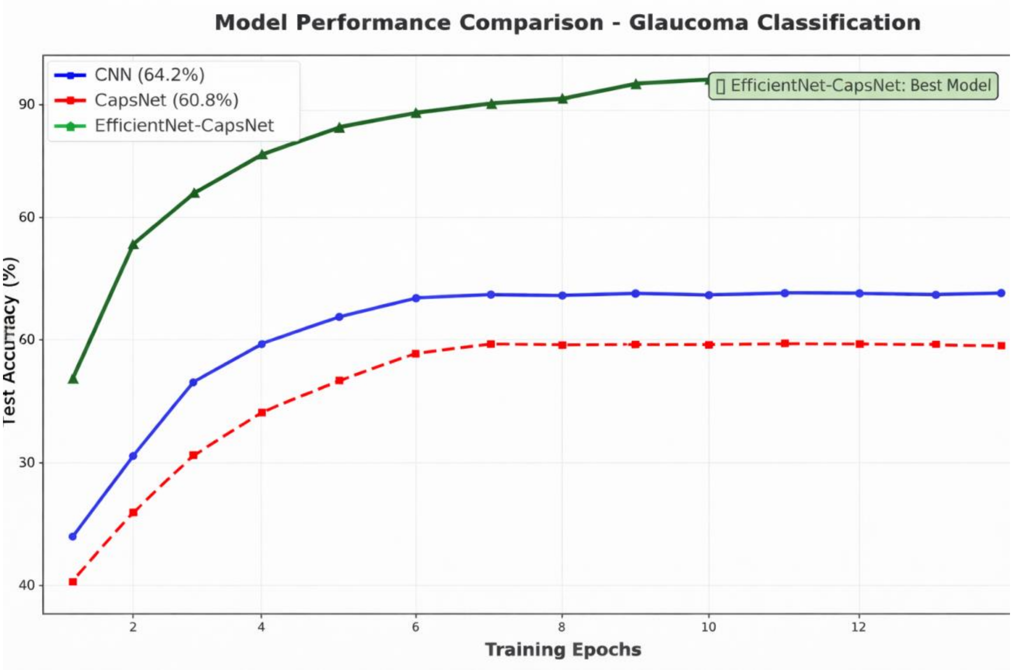

Fig. 7 Learning Curves Comparison

4. Results

When the proposed EfficientNet-CapsNet hybrid model was applied to the selected data, a superior average test accuracy of 80.7% (5-fold cross-validation: 80.7 ± 1.8%) was achieved, with a performance range of 78.5% to 83.0%. This significantly outperforms the independent CNN (64.2%) and CapsNet (60.8%) models, as shown in Table 3. The learning curves (Figure 7) confirm the training stability and efficiency of the hybrid model. Furthermore, the operating parameters in Table 4 demonstrate the model's strong diagnostic capabilities, including an NRG accuracy of 82.4%, an RG recovery of 77.2%, and an overall F1 score of 80.1%.

Table 3: Performance Comparison

Model	Accuracy	Sensitivity (Recall,+)	Specificity (Recall, -)	F1-Score
CNN	64.2%	71.4%	57.0%	57.8%
CapsNet	60.8%	68.2%	53.4%	54.1%
EfficientNet- CapsNet	80. 7%	82.8%	78.6%	80.1%

Table 4: Details the operational parameters(NRG, RG)

Mode	Accuracy	Accuracy NRG	Recovery NRG	Accuracy RG	Recovery RG	F1
CNN	64.2%	66.1%	71.4%	61.8%	54.3%	57.8%
CapsNet	60.8%	62.3%	68.2	58.9%	50.1%	54.1%
EfficientNet- CapsNet	80.7%	82.4%	84.1%	78.9%	77.2%	80.1%

5. Discussion

This research presented a hybrid scientific and practical model, EfficientNet-CapsNet, for diagnosing glaucoma, with a testing accuracy of 80.7%. This study represents a significant contribution to the field of medical image recognition.

1- The EfficientNet-CapsNet model was primarily used for extracting high-level features. It relies on extracting depth, width, and resolution, which produces optimal visual representations. These features allow for direct manipulation of determinants in fundus images, such as: Changes in the optic disc and pupil.

2- Strengths: The RGB three-color system was adopted with a size of 224×224×3 pixels in real time, to meet the challenges of imaging medical data.

3- Due to the dimensional mismatch that emerged when combining pre-trained feature maps with capsule routing layers, we introduced a projection matrix with adaptive clustering. This ensured a seamless flow of data between the EfficientNet-B0 backbone and the CapsNet architecture, while preserving the spatial hierarchy of functions.

While we agree that accuracy is a standard criterion, our proposed model prioritizes clinical diagnosis. As shown in Table 3, the proposed hybrid model achieved a sensitivity of 82.8% and a specificity of 78.6%, two of the most important factors for early glaucoma detection and reducing undiagnosed cases. Although studies have reported accuracy levels exceeding 90%, these models often rely on sophisticated GPU infrastructures and massive datasets. Our model, achieving 80.7% accuracy, is specifically designed for resource-limited clinical environments using standard CPUs, providing a scalable and computationally feasible solution for comprehensive screening without compromising the fundamental anatomical and spatial relationships within the optic nerve head.

6. Conclusion

In this work, an efficient hybrid architecture, EfficientNet-CapsNet . Proposed and evaluated using deep learning models for the automatic detection of glaucoma in fundus images. A systematic comparison was performed using clinically validated datasets (EyePACS). The results showed that the EfficientNet-CapsNet hybrid model outperformed all other models with an average accuracy of 80.7% and a standard deviation of ±1.8%. It also outperformed both CNN (64.2%) and CapsNet (60.8%) in extracting both the fine-grained features and spatial hierarchy necessary for optic nerve segment assessment. While there has been much research in glaucoma detection, most has relied on discrete CNNs and simplified clustering techniques, leading to inconsistent results due to their limited generalizability across different imaging modes and their inability to preserve the anatomical state of the optic nerve head. Our hybrid model bridges this gap by combining the computationally efficient composite expansion of EfficientNet with the conformal routing of CapsNet and robust structural representation.

This model demonstrates high performance suitable for practical use at a minimal computational cost (224×224 input size). This efficiency enables large-scale clinical screening in resource-limited settings, where early detection can prevent a high percentage (90%) of glaucoma-related blindness [12]. Promising future insights for research along the same lines can be achieved by integrating multimodal data (including optical coherence tomography measurements) and interpretation techniques to enhance physicians' confidence in the true diagnosis.

In conclusion, this study demonstrates that the proposed EfficientNet-CapsNet architecture, which represents a realistic diagnostic model in a clinical setting and possesses reasonable computational feasibility, enhances practical media applications. With further validation and refinement, these systems have the potential to significantly contribute to global efforts for early detection and improved risk classification, with the aim of reducing the burden of blindness caused by glaucoma.

Reference:

[1] A. Albelaihi and D. M. Ibrahim, “DeepDiabetic: An Identification System of Diabetic Eye Diseases Using Deep Neural Networks,” IEEE Access, vol. 12, pp. 10769–10789, 2024, doi: 10.1109/ACCESS.2024.3354854.

[2] B. U. Ahmad, J. E. Kim, and E. Rahimy, “Fundamentals of artificial intelligence for ophthalmologists,” Curr. Opin. Ophthalmol., vol. 31, no. 5, pp. 303–311, 2020, doi: 10.1097/ICU.0000000000000679.

[3] F. Ghani, M. U. Sattar, H. W. Khan, M. Narmeen, and A. Mehmood, “A Methodology for Glaucoma Disease Detection Using Deep Learning Techniques,” Int. J. Comput. Digit. Syst., vol. 11, no. 1, pp. 401–411, 2022, doi: 10.12785/ijcds/110133.

[4] D. Prins, S. Hanekamp, and F. W. Cornelissen, “Structural brain MRI studies in eye diseases: Are they clinically relevant? A review of current findings,” Acta Ophthalmol., vol. 94, no. 2, pp. 113–121, 2016, doi: 10.1111/aos.12825.

[5] S. S. Roy, R. Sikaria, and A. Susan, “A deep learning based CNN approach on MRI for Alzheimer’s disease detection,” Intell. Decis. Technol., vol. 13, no. 4, pp. 495–505, 2020, doi: 10.3233/idt-190005.

[6] J. W. Chan, N. C. Y. Chan, and A. A. Sadun, “Glaucoma as neurodegeneration in the brain,” Eye Brain, vol. 13, pp. 21–28, 2021, doi: 10.2147/EB.S293765.

[7] H. Cho et al., “Deep Learning Ensemble Method for Classifying Glaucoma Stages Using Fundus Photographs and Convolutional Neural Networks,” Curr. Eye Res., vol. 46, no. 10, pp. 1516–1524, 2021, doi: 10.1080/02713683.2021.1900268.

[8] M. Eye, D. Revisiting, E. Axis, and F. Pirot, “Diffuse Damage in Traumatic Brain Injury and Its Large-Spectrum Implications: Keys on Diffuse Axonal Injury and Indirect Traumatic Optic Neuropathy for Major Eye Diseases. Revisiting the ‘Eye-Brain Axis,’” J. Psychol. Neurosci., vol. 5, no. 3, pp. 1–5, 2023, doi: 10.47485/2693-2490.1076.

[9] S. M. A. Banijamali, C. Versek, K. Babinski, S. Kamarthi, D. Green-LaRoche, and S. Sridhar, “Portable Multi-focal Visual Evoked Potential Diagnostics for Multiple Sclerosis/Optic Neuritis patients,” medRxiv, 2023.

[10] G. Hemanth, M. Janardhan, and L. Sujihelen, “Design and implementing brain tumor detection using machine learning approach,” Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019, vol. 2019-April, no. Icoei, pp. 1289–1294, 2019, doi: 10.1109/icoei.2019.8862553.

[11] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, “Classification using deep learning neural networks for brain tumors,” Futur. Comput. Informatics J., vol. 3, no. 1, pp. 68–71, 2018, doi: 10.1016/j.fcij.2017.12.001.

[12] A. C. Thompson, A. A. Jammal, and F. A. Medeiros, “A review of deep learning for screening, diagnosis, and detection of glaucoma progression,” Transl. Vis. Sci. Technol., vol. 9, no. 2, pp. 1–19, 2020, doi: 10.1167/tvst.9.2.42.

[13] V. Gulshan et al., “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,”

JAMA - J. Am. Med. Assoc., vol. 316, no. 22, pp. 2402–2410, 2016, doi: 10.1001/jama.2016.17216.

[14] Adamopoulou M, Makrynioti D, Gklistis G, and Koutsojannis C, “Revolutionizing eye disease diagnosis with deep learning and machine learning,” World J. Adv. Res. Rev., vol. 18, no. 2, pp. 741–763, 2023, doi: 10.30574/wjarr.2023.18.2.0856.

[15] T. Guergueb and M. A. Akhloufi, “A Review of Deep Learning Techniques for Glaucoma Detection,” SN Comput. Sci., vol. 4, no. 3, 2023, doi: 10.1007/s42979-023-01734-z.

[16] Dr. N. Rajkumar, Dr. S. Megala, S. Muthukumaran, “Eye Disease Diagnosis and Classification using Deep Convolution Neural Network,” Tuijin Jishu/Journal Propuls. Technol., vol. 44, no. 4, pp. 3453–3467, 2023, doi: 10.52783/tjjpt.v44.i4.1483.

[17] A. Pal, M. R. Moorthy, and A. Shahina, “G-Eyenet: A Convolutional Autoencoding Classifier Framework for the Detection of Glaucoma from Retinal Fundus Images,” Proc. - Int. Conf. Image Process. ICIP, pp. 2775–2779, 2018, doi: 10.1109/ICIP.2018.8451029.

[18] A. M. Hasan, H. A. Jalab, F. Meziane, H. Kahtan, and A. S. Al-Ahmad, “Combining Deep and Handcrafted Image Features for MRI Brain Scan Classification,” IEEE Access, vol. 7, pp. 79959–79967, 2019, doi: 10.1109/ACCESS.2019.2922691.

[19] R. W. Ibrahim, A. M. Hasan, and H. A. Jalab, “A new deformable model based on fractional Wright energy function for tumor segmentation of volumetric brain MRI scans,” Comput. Methods Programs Biomed., vol. 163, pp. 21–28, 2018, doi: 10.1016/j.cmpb.2018.05.031.

[20] S. Sreng, N. Maneerat, K. Hamamoto, and K. Y. Win, “Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images,” Appl. Sci., vol. 10, no. 14, 2020, doi: 10.3390/app10144916.

[21] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, “Automated detection of mild and multi-class diabetic eye diseases using deep learning,” Heal. Inf. Sci. Syst., vol. 8, no. 1, pp. 1–9, 2020, doi: 10.1007/s13755-020-00125-5.