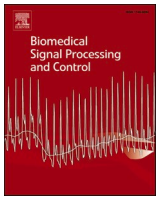
Biomedical Signal Processing and Control 92 (2024) 106120

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



An evaluation of AI-based methods for papilledema detection in retinal fundus images

Ahmed M. Salaheldin [a,b,*], Manal Abdel Wahed [a], Manar Talaat [c], Neven Saleh [b,d]

[a] Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt
[b] Systems and Biomedical Engineering Department, Higher Institute of Engineering, EL Shorouk Academy, Cairo, Egypt
[c] Department of Ophthalmology, Faculty of Medicine, South Valley University, Qena, Egypt
[d] Electrical Communication and Electronic Systems Engineering Department, Engineering Faculty, October University for Modern Sciences and Arts, Giza, Egypt

ARTICLE INFO

Keywords:
Papilledema detection
Retinal fundus images
Deep learning models
Multi-paths CNN
LSTM
Occlusion sensitivity

ABSTRACT

The complexities inherent in diagnosing papilledema, particularly within the realm of neuro-ophthalmology, emphasize the pressing need for sophisticated diagnostic methodologies. This study highlights the application of novel models tailored explicitly for papilledema detection, distinguishing it from pseudo-papilledema and normal cases, through the strategic utilization of deep learning frameworks, specifically convolutional neural networks (CNNs) and recurrent neural networks (RNNs). Leveraging hierarchical feature extraction from retinal images, the multi-paths CNN model successfully identifies crucial indicators of papilledema, while the cascaded model, integrating ResNet-50 and long short-term memory (LSTM), effectively captures sequential features. Leveraging a meticulously curated dataset comprising 18,258 fundus images, the presented models exhibit exceptional performance, with the multi-paths CNN achieving an accuracy of 99.97 %, and the LSTM model demonstrating an accuracy of 99.81 %. Comparative analysis showcases the unparalleled efficacy of the models, underscoring their potential in clinical diagnostics. Notably, they demonstrate robustness in occlusion sensitivity tests, highlighting their resilience in scenarios involving obscured image components. This pioneering study represents a significant milestone in papilledema detection, with the promise of advancing patient outcomes and streamlining healthcare practices. The proposed deep learning models not only offer precise diagnoses but also hold the potential to automate elements of the diagnostic workflow, alleviating the workload of healthcare professionals and enhancing overall patient care outcomes.

1. Introduction

The realm of medical diagnostics, particularly within the field of neuro-ophthalmology, has witnessed a growing need for precise and efficient methods to detect and assess a range of ocular conditions [1,2]. Among these, the detection of papilledema, a condition characterized by optic disc swelling often associated with underlying neurological disorders, stands out as a critical diagnostic challenge [3]. Timely and accurate identification of papilledema is pivotal in providing appropriate medical care and preventing potential vision loss or further neurological complications [4]. However, the diagnosis of papilledema is far from straightforward. It requires a comprehensive evaluation of fundus images to discern the subtle but critical signs of optic disc swelling [5]. This intricate diagnostic task often necessitates the expertise of skilled ophthalmologists and, increasingly, the assistance of advanced technologies [6,7].

The intricacies of discerning papilledema extend beyond traditional diagnostic methods, prompting the exploration of innovative technologies and methodologies [8]. In recent years, as the importance of early and accurate diagnosis has become increasingly apparent, researchers have fervently addressed this diagnostic challenge. They have embarked on a journey to develop and refine novel techniques and methodologies that harness the power of deep learning models, particularly CNNs [9,10] and RNNs [11] in a synergistic manner, alongside various imaging modalities.

The utilization of CNNs allows for the automatic extraction of hierarchical features from retinal fundus images. These networks are adept at identifying intricate patterns, edges, and structural nuances within the images, enabling the model to discern subtle but critical signs of papilledema [12]. Concurrently, the incorporation of RNNs introduces a

* Corresponding author.
E-mail address: ahmed.201920235@eng-st.cu.edu.eg (A.M. Salaheldin).

https://doi.org/10.1016/j.bspc.2024.106120
Received 5 November 2023; Received in revised form 30 January 2024; Accepted 14 February 2024
Available online 20 February 2024

temporal aspect to the analysis. RNNs are well-suited for capturing sequential dependencies within the data, which is valuable in tracking changes in the feature sequence for papilledema cases [13].

It is crucial to recognize the prominence of transformer models in the realm of sequence analysis. Transformer models, renowned for their self-attention mechanism, have demonstrated prowess in extracting powerful sequence features. LSTM's strength lies in its adeptness at capturing temporal dependencies and subtle changes within the image sequence, making it particularly well-suited for certain aspects of fundus image classification. Conversely, transformers leverage self-attention mechanisms to achieve a global understanding of context, potentially excelling in scenarios where long-range dependencies and contextual information are paramount [14,15]. We delve into considerations such as computational efficiency, model interpretability, and the specific challenges each architecture may encounter in the nuanced task of papilledema detection.

Introducing temporal features into fundus images can significantly enhance diagnostic results from a medical perspective, particularly in conditions like papilledema. Fundus images, which provide a detailed view of the back of the eye, capture dynamic information about the retina, optic nerve, and surrounding structures. Temporal features, related to changes over time, become crucial for understanding the progression and subtle alterations in pathological conditions. In medical practice, the introduction of temporal features allows the diagnostic model to capture these sequential changes, providing a more nuanced and accurate representation of the disease progression [16].

In the context of fundus images, LSTM can learn and retain information about sequential variations in optic disc appearance, edema progression, or other dynamic features indicative of papilledema. This temporal awareness enables the model to better differentiate between normal variations and pathological changes, ultimately improving diagnostic accuracy. Moreover, fundus images are influenced by factors such as lighting conditions, image quality, and patient-specific variations. Temporal features allow the model to adapt to these variations, learning patterns and trends that might be obscured in individual static images. The incorporation of temporal information, therefore, aligns with the clinical reality that certain diagnostic indicators may not be apparent in a single snapshot but become discernible over a series of images taken over time [17].

Multi-paths CNNs are versatile architectures with diverse applications across various domains [18]. One notable use of multi-paths CNNs is in computer vision tasks, particularly image classification and object detection. The parallel pathways in these networks allow for the extraction of different hierarchical features from input data, enabling a more comprehensive understanding of complex visual patterns [19]. In medical imaging, multi-paths CNNs find applications in the detection and diagnosis of various diseases, leveraging their ability to capture intricate details in medical images. Additionally, in natural language processing, multi-path CNNs can be employed for text classification and sentiment analysis by processing text data through different pathways to capture diverse linguistic features [20]. Overall, the uses of multi-path CNNs are diverse, ranging from image analysis in healthcare to text processing in natural language understanding, making them valuable tools in the realm of artificial intelligence and machine learning.

This combination of CNNs and RNNs offers a comprehensive approach that capitalizes on both spatial and temporal information present in the image data, ultimately enhancing diagnostic precision. These advancements aim to not only enhance the diagnostic precision of papilledema but also streamline the process, making it more accessible to medical practitioners and patients alike [21]. The integration of deep learning models with medical imaging not only holds the promise of more accurate and timely diagnoses but also offers the potential for automating aspects of the diagnostic workflow, thereby alleviating the burden on healthcare professionals and improving patient care outcomes [22].

This study aims to deploy and evaluate the most recent deep learning approaches for the identification of papilledema in retinal fundus images. The study makes significant contributions: (1) by presenting a newly curated fundus image dataset, specifically assembled for this study and associated with diverse medical assessments to ensure accurate diagnosis; (2) introducing, for the first time to the best of our knowledge, a multi-paths CNN model for papilledema detection; (3) design a novel multi-paths CNN architecture that was designed especially for the detection of papilledema from fundus images as an assistive tool for neuro ophthalmic diseases detection; (4) As far as we know, this is the first time to develop a cascaded model for papilledema detection, integrating transfer learning concepts and RNN using LSTM network; (5) obviously the proposed models can rapidly classify an image according to its class in few milliseconds; (6) demonstrating the superiority of the proposed models over existing methods in detecting papilledema, pseudo papilledema, and normal cases according to the comparison results; (7) external validation of the proposed models with public available data yielded robustness of the models.

The structure of the study is as follows: Section 2 offers a comprehensive review of related works that have addressed similar challenges. Section 3 delves into the materials and methods employed in the study. Section 4 presents the experimental setup and configuration. The results obtained from the proposed models are presesnted in Section 5. A thorough discussion of these results and a comparative analysis with related works can be found in Section 6. Finally, Section 7 serves as the conclusion of the study, shedding light on potential avenues for future research.

2. Relevant works

In light of the significant implications posed by this challenge, several recent studies have endeavored to tackle it from various angles. Notably, Sun et al. [23] introduced a comprehensive end-to-end deep learning system for the detection of different ophthalmic diseases including papilledema, leveraging the capabilities of three pre-trained CNN models, namely, Efficient B7, DenseNet, and ResNet 101. The implementation of these models was carried out utilizing ultra-widefield fundus images, marking a noteworthy departure from prior methodologies. The system yielded a classification accuracy of 93 % when utilizing the Efficient B7 model as the most proficient classifier among the ensembles. Biousse et al. [24] conducted a study evaluating the performance of the BONSAI deep learning system in discerning papilledema from non-mydriatic fundus photographs. The study utilized a dataset comprising 1608 images from 828 patients. The BONSAI system demonstrated notable accuracy in distinguishing between normal and abnormal optic discs, effectively identifying cases of papilledema. These findings underscore the BONSAI system's potential as a valuable diagnostic tool in settings beyond ophthalmology clinics, particularly in non-ophthalmic healthcare facilities and emergency departments equipped with non-mydriatic fundus cameras.

Kokulu et al. [25] proposed transfer learning based model for the detection of papilledema from fundus images. The study employs image preprocessing techniques such as histogram equalization and 3D box filtering. Transfer learning approaches, including EfficientNet-B0, GoogleNet, MobileNetV2, NASNetMobile, and ResNet-101, were compared, with MobileNetV2 exhibiting the highest performance, achieving 0.96 overall accuracy and 0.94 Cohen's Kappa.

An additional work is credited to Badawi et al. [26], who introduced a cascaded deep learning (DL) framework. This innovative model operates in a two-stage process: initially, it undertakes the task of segmenting retinal blood vessels, distinguishing between arteries and veins. Subsequently, it proceeds to quantify the width of the segmented vascular network within the region of interest encompassing the optic disc. Remarkably, this approach yielded a remarkable accuracy rate of 96.82 %, showcasing the model's proficiency in this multifaceted task.

In their study, Vasseneix et al. [27] introduced a tailored CNN meticulously crafted for the precise assessment of papilledema severity.

The proposed model is structured around two fundamental phases. Initially, it focuses on the segmentation of the region of interest within the fundus images. Subsequently, the model is trained to classify the papilledema severity into two distinct grades, achieving commendable results with an accuracy of 87.9 %, sensitivity of 91.8 %, and specificity of 86.2 %. Furthermore, Saba et al. [28] introduced a multi-stage model designed to address the detection and grading of papilledema. Their approach involves initial segmentation of the area of interest using the U-Net architecture, followed by the utilization of DenseNet for the subsequent classification process. Their model achieved outstanding classification performance, boasting an accuracy rate of 99.17 %, sensitivity of 98.63 %, and specificity of 97.83 % with total training time of 6 h and 37 min.

Cen et al. [29] harnessed the power of CNN techniques for the detection of a diverse spectrum of 39 distinct retinal diseases from fundus photographs. Following this, the authors employed a heat map to pinpoint the regions contributing to the diagnostic process. The performance of their system was meticulously assessed through various evaluation criteria, including an impressive area under the curve (AUC) score of 0.9984, a sensitivity rate of 97.8 %, and an exceptional specificity rate of 99.6 %. In another study, Farazdaghi et al. [30] introduced a strategy for distinguishing between papilledema and pseudo papilledema. Their study incorporated both B-scan ultrasound (US) images and Optical Coherence Tomography (OCT) images. Through the extraction of Retinal Nerve Fiber Layer (RNFL) thickness from the OCT images, medical experts were able to make disease diagnoses without the need for a Computer-Aided Detection (CAD) model. In evaluating the B-scan ultrasound images, the study achieved an 86 % sensitivity and an 88 % specificity. In the context of RNFL thickness measurement, the reported sensitivity was 83 %, accompanied by a specificity of 76 %.

In examining research employing the LSTM model for fundus image classification, a limited number of studies have specifically addressed the retinal disorders classification. DL- based model has been introduced by Demir et al. [31] for classification of fundus images into diabetic retinopathy, glaucoma, age-related macular degeneration (AMD), myopia, and hypertension from fundus image through the integration of R-CNN and LSTM to attain classification accuracy of 89.54 %.

Li et al. [32] delved into a multifaceted classification challenge, exploring the detection of numerous retinal disorders, including the intricate realm of neuro-ophthalmic conditions like papilledema. Within their study, they introduced a tailored CNN model to tackle this intricate classification task. Their model achieved notable results, boasting an AUC score of 0.95, coupled with a sensitivity rate of 85.05 % and specificity rate of 92.15 %. A hybrid model combining CNN and RNN concepts for the detection and grading of cataract from fundus images is proposed by Imran et al. [33]. The authors utilized four pretrained CNN models for the feature extraction phase to generate multilevel feature vector and training bi-directional LSTM model for the classification phase. The system achieved a classification accuracy of 97.39 %.

Yang et al. [34] addressed the task of detecting optic disc pallor, embarking on a comprehensive exploration. Within their study, the authors introduced a model comprising two pivotal phases. Initially, the model engages in the segmentation of the optic disc, followed by the subsequent utilization of a classical machine learning (CML) classifier for the detection process. The proposed model achieved noteworthy results, boasting an accuracy rate of 96.1 %, a sensitivity rate of 95.3 %, and a commendable specificity rate of 96.7 %. Ahn et al. [35] proposed multiple DL models for the classification of fundus images into neuropathies, pseudo papilledema, and normal cases. These models utilized various pre-trained models as well as customized CNN architecture. The proposed models achieved impressive classification accuracies, ranging from 95.89 % to 98.63 %. Most of the studies that were covered in the literature review used local collected datasets and are not publicly available. A noticeable research gap exists, as no studies have yet delved into the application of LSTM model for the detection of neuro-ophthalmic diseases, particularly papilledema. For more investigation of the literature, comparison between the related works is presented in Table 1.

3. Materials and methods

The proposed research endeavors to introduce an AI-based model aimed at the precise detection of papilledema within retinal fundus images. This endeavor encompasses the deployment of two distinct approaches, meticulously constructed upon the foundational principles of deep learning. By harnessing the sophisticated intricacies of deep learning, we aim to add a contribution in the area of neuro-ophthalmic diagnostics, characterized by elevated accuracy and enhanced diagnostic efficacy in the domain of papilledema detection, setting new benchmarks in the field of medical image analysis.

3.1. Dataset

Over a span of seven months, a meticulously curated dataset of retinal fundus images was procured exclusively for this research initiative. The source of this data was the Department of Ophthalmology within the Faculty of Medicine at South Valley University. All procedures performed in the study were in accordance with ethical standards and the data is collected based on patient consent and stored securely to ensure patient privacy. Throughout the data collection process, a thorough examination of the patients was conducted, encompassing comprehensive medical assessments. These evaluations encompassed the measurement of various parameters, including best-corrected visual acuity (BCVA), and the identification of symptoms such as headache, blurring, and vomiting. Pupil reactions were systematically tested and categorized as normal, sluggish, or indicative of an afferent pupillary defect (APD). Additionally, measurements were recorded for parameters like RNFL thickness and total peripapillary retinal thickness. Visual field tests were systematically administered to determine results, classifying them as either normal, displaying an enlarged blind spot, scattered scotomas, incomplete arcuate scotoma, or complete arcuate scotoma. This dataset was thoughtfully stratified into three distinct classes: normal, papilledema, and pseudo papilledema, encompassing the diverse spectrum of retinal conditions. Fig. 1 presents a sample of the data for the three classes.

The initial dataset comprised a total of 1167 images distributed across these specified classes. In preparation for the subsequent deep learning model implementation, essential data preprocessing was carried out. One pivotal step involved resizing the images to dimensions of $224 \times 224 \times 3$, aligning them with the input layer requirements of the deep learning architecture.

Furthermore, to enhance the model's capacity to generalize across a multitude of clinical scenarios, data augmentation was diligently executed. This augmentation procedure encompassed two primary approaches. Firstly, image rotation was systematically applied within the range of -15° to 15°, in discrete 5° increments. In conjunction with these rotations, a suite of image processing techniques, including histogram adjustments, sharpening, and blurring, were judiciously employed. Secondly, a parallel augmentation approach was initiated by performing horizontal mirroring on the images. Subsequently, the same rotational transformations were systematically applied, mirroring the criteria of the initial approach. This augmentation scheme was complemented by the concurrent application of the aforementioned image processing techniques. This process led to extending of the dataset to be 18,258 fundus images over the three classes. Fig. 2 provides a comprehensive visual overview of the augmentation phase, which succinctly encapsulates the entire augmentation process.

3.2. Classification models

In this study, we present two distinct approaches tailored to the intricate classification task of fundus images, aimed at the precise

Table 1
Comparison between the related works.

Study	Objective	Scheme	Dataset	Contributions	Limitations
Sun et al., 2023	Detection of several common fundus diseases	EfficientNet B7, DenseNet, and ResNet 101	4,574 ultra-widefield images from local database	The proposed models outperform the human expert in the data classification. The originality of the used dataset, as it was collected specially for the study	The study's training and test sets consist of a limited number of images. The study's accuracy falls short, particularly in correctly identifying certain conditions, like optic nerve disease and vitreous opacity
Biousse et al., 2023	Detection of papilledema on nonmydriatic ocular fundus photographs	U-Net architecture with a ResNet-34 encoder	17,368 multi-ethnic ocular color fundus photographs from local dataset	The study covers dataset from different centers which provide generalization of the cases	The study was a secondary retrospective analysis requiring reclassification of the photographs into the 3 categories used by the BONSAI-DLS
Kokulu et al., 2023	Detection of papilledema severity	Transfer learning-based model using 4 pretrained models	The image dataset includes 295 papilledema images, 295 pseudo-papilledema images, and 779 control images.	An image processing-based solution was proposed for the detection of papilledema severity from color fundus images using transfer learning, histogram equalization, and 3-D box filtering	Unbalancing data distribution among the targeted classes which may bias the system performance to specific class
Badawi et al., 2022	Detection and grading hypertensive retinopathy	Retinal vasculature segmentation and vessel fragment extraction using morphometric analysis	504 retinal vessel morphometry images specially collected for the study	Combining both medical image processing techniques with the calculation of the clinical measurement, which contribute directly to the diagnosis	In the segmentation phase, the authors used traditional methods instead of deep learning, which is the state of the art for this task
Vasseneix et al., 2021	Papilledema severity detection	Customized CNN model	Training dataset: 2,103 images collected from 16 sites and 11 countries Testing dataset: 214 images collected from 4 sites	Applying two phases by segmenting the area of interest and then classifying the images to their corresponding classes.	Retrospective data collection. Lack of objective data like OCT retinal nerve fiber thickness
Saba et al., 2021	Detection and grading of papilledema	A cascaded system to localize the optic disc using U-Net and classify the abnormalities using DenseNet	60 papilledema and 40 normal fundus images taken from STARE dataset	The proposed method provides two stages of work, including first extracting the region of interest, then detecting the papilledema and grading it, which adds potential success in the recognition phase	The small amount of data through Fundus images
Cen et al., 2021	Detection of 39 fundus diseases and conditions in retinal photographs	A deep learning-based model using CNN techniques is presented to classify the data	275,543 fundus images collected from multiple sources cover 39 retinal diseases from local database	The proposed model has the ability to differentiate between 39 classes including their variability.	There is no universal diagnostic standard or consensus for reference of most fundus diseases if based only on fundus images.
Farazdaghi et al., 2021	Differentiating between papilledema and pseudo papilledema	Clinical based study	B scan ultrasonography and OCT images specially collected for the study	Using OCT imaging modalities, which provide quantitative data for the differentiation between papilledema and pseudo papilledema	The study was designed based on the use of OCT features only, such as retinal nerve fiber layer thickness, without implementing any computer-based system for the automated diagnosis
Demir et al., 2021	Detection of different retinal disorders	Cascaded model consists of R- CNN + LSTM	The utilized dataset consists of 6426 fundus images covering 8 classes from local database	The residual layer information of the R-CNN model and the LSTM model's ability to keep important data in memory was utilized.	The proposed R-CNN + LSTM model contains too many learnable parameters. Therefore, powerful hardware is required for fast prediction results
Li et at al., 2020	Evaluation of an appropriate deep learning system (DLS) for the detection of 12 major fundus diseases using color fundus images	Deep learning-based model was designed specially for this task using CNN techniques	The DLS was developed with 56 738 images and tested with 8176 images from one internal test set and two external test sets.	Multiple classification problem is presented by the authors using enough images, up to 56,738 images.	When compared with human doctors, the DLS obtained a higher diagnostic sensitivity but a lower diagnostic specificity.
Imran et al., 2020	Detection and grading of cataract	Hybrid model consists of CNN phase and RNN phase	8030 fundus images covering all classes	Combined different DL approaches	The result is biased to the classes contributing with higher number of images
Ahn et al., 2019	Differentiation between optic neuropathies, pseudo papilledema, and normal	Customized deep learning model and multiple transfer learning CNN models	295 images of optic neuropathies, 295 images of pseudo papilledema, and 779 normal images	The authors proposed their own CNN model and compared it with different pre trained models	The small number of used data using Fundus images
Yang et al., 2018	Detection of optic disk pallor	The model consists of two phases. First, optic disc segmentation and image analysis. Second, an artificial intelligence model using logistic regression	A total of 230 photographs with variable degree of optic disc pallor, and 123 normal	First, the CAD system detected optic disc pallor through a fully automated process. Second, the combination of corrected green channel images and blue channel images provided a high contrast of the optic disc.	The reproducibility of the CAD system could not be determined by multiple photographs taken from one person, as this was a retrospective study

a b c

Fig. 1. Sample of the data (a) normal, (b) papilledema, (c) pseudo papilledema.

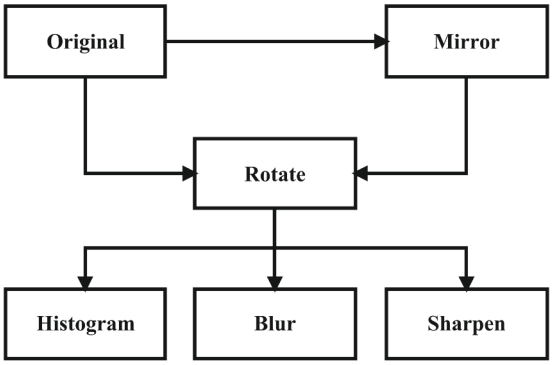


Fig. 2. Schema for the augmentation process of selected fundus images.

detection of papilledema. Our innovative methodology encompasses the incorporation of cutting-edge deep learning paradigms, notably CNN and RNN specially with LSTM applications, strategically harnessed for the dual purpose of feature extraction and classification. These advanced techniques promise to significantly elevate the accuracy and reliability of the proposed papilledema detection system.

A) Multi-paths CNN Model

Specially designed multi-paths CNN model was built to perform end-to-end deep learning model. The proposed multi-paths CNN architecture was designed to enhance the discriminative capabilities of image classification tasks through the parallel extraction and fusion of multiple levels of features. Comprising three distinct paths, each path progressively captures distinct hierarchical representations of features in images. Fig. 3 presents the process starting from the raw data to the classification process. Based on local, intermediate, and global feature extraction techniques, the three models were distinguished from one another. The local features path prioritizes the capture of detailed, spatially proximate information. In contrast, the intermediate features path focuses on broader yet intricate characteristics. The global features path is designed to capture comprehensive, context-rich information. This intricate multi-path architecture empowers the model to discern fine details, intermediate characteristics, and the overall contextual information inherent in fundus images.

The proposed model comprises three primary paths, each with a unique approach to feature extraction. The first path focuses on capturing local features: it starts with a 5×5 convolutional layer with 32 filters, followed by batch normalization and ReLU activation. Subsequently, a max-pooling layer with a 2×2 window size down samples the feature maps, and another 5×5 convolutional layer with 64 filters is applied, followed by batch normalization and ReLU activation, followed by another max-pooling layer.

The second path, dedicated to intermediate feature extraction, begins with a 3×3 convolutional layer with 16 filters, followed by batch normalization and ReLU activation. Max-pooling is used to down sample the feature maps, and two additional 3×3 convolutional layers with 32, and 64 filters follow consecutively, each accompanied by batch normalization and ReLU activation. Max-pooling is applied after the first two convolutional layers in this path.

The third path is responsible for global feature extraction and consists of a depth concatenation layer that combines features from the previous two paths. A 3×3 convolutional layer with 256 filters, batch normalization, and ReLU activation is then applied to the concatenated features. To further consolidate global information, max pooling is employed with a stride of 2×2, leading to the formation of a comprehensive feature representation. Subsequently, a fully connected layer with three output units is introduced for classification, followed by a SoftMax layer to estimate class probabilities. The main differences between the paths are the number of filters in each convolution layer, the size of each filter, and different hyperparameters for each layer. Finally, the classification layer maps the model's output to the three classes: papilledema, pseudo papilledema, and normal. These multi-paths architecture enables the model to capture intricate details from fundus images, understand intermediate features, and perceive the global context. The detailed structure of the multi-paths model is given in Fig. 4.

The multi-paths CNN design demonstrates a strategic arrangement that allows for the comprehensive exploration of features across different levels of abstraction, ultimately enhancing the network's ability to discern intricate patterns and structures within images, leading to improved classification performance [36].

B) LSTM model

The proposed methodology comprises a series of carefully orchestrated steps, commencing with meticulous data preparation, as elucidated in the dataset section. Subsequently, a transfer learning-based model was designed to extract deep feature representations. Specifically, the ResNet-50 model was employed to extract features from its final average pooling layer [37]. These extracted features manifest as a robust set of 2048 distinct features for each image, encompassing a diverse spectrum of feature levels.

The LSTM model [38] is primed to intake and process the sequences of extracted features, effectively discerning the presence of papilledema within the fundus images. The architecture of this LSTM model is characterized by a layer equipped with 128 hidden units. Following this, a fully connected layer and a SoftMax layer are introduced to compute the probabilities associated with each class. Subsequently, the model's performance is rigorously assessed using the dedicated testing dataset. A comprehensive visual overview of the LSTM model was illustrated in Fig. 5.

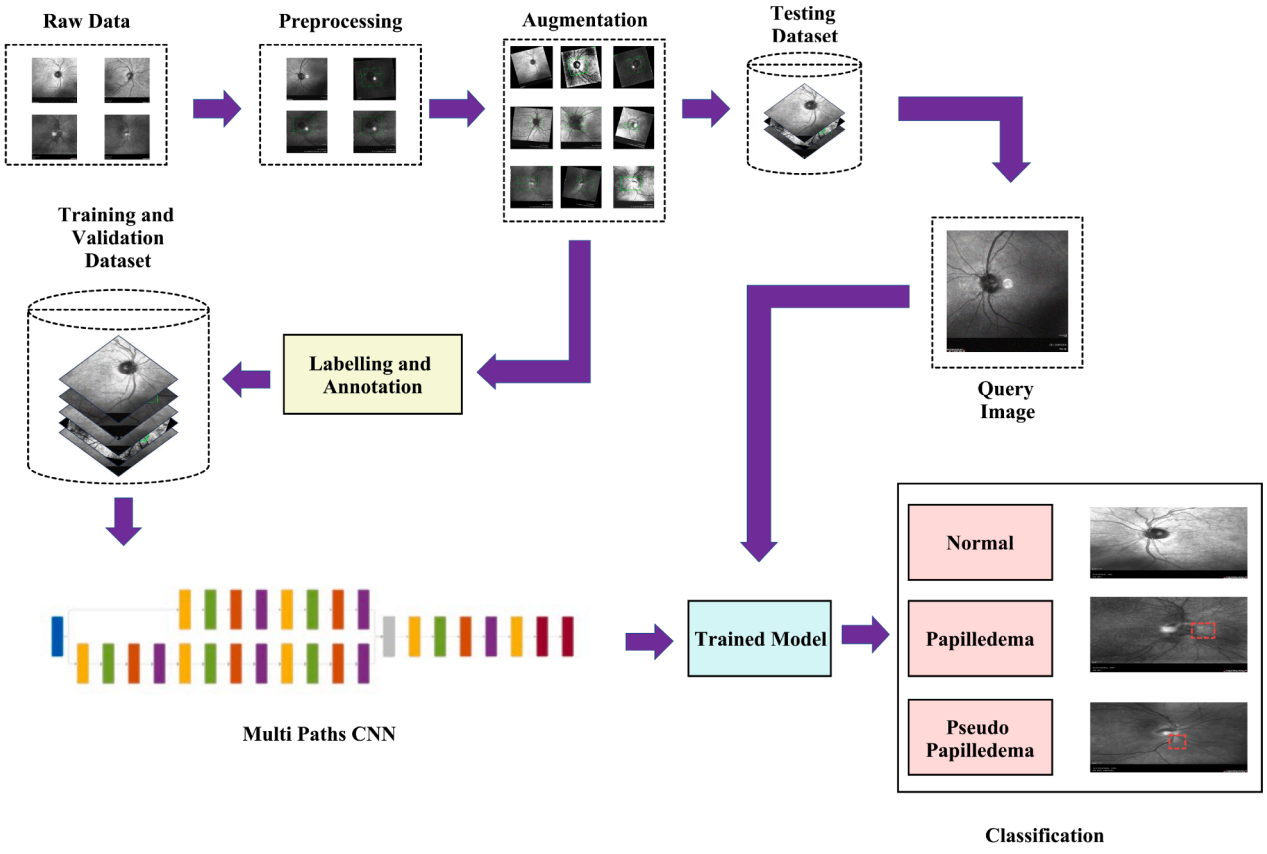


Fig. 3. The proposed multi-paths end-to-end deep learning model for papilledema detection.

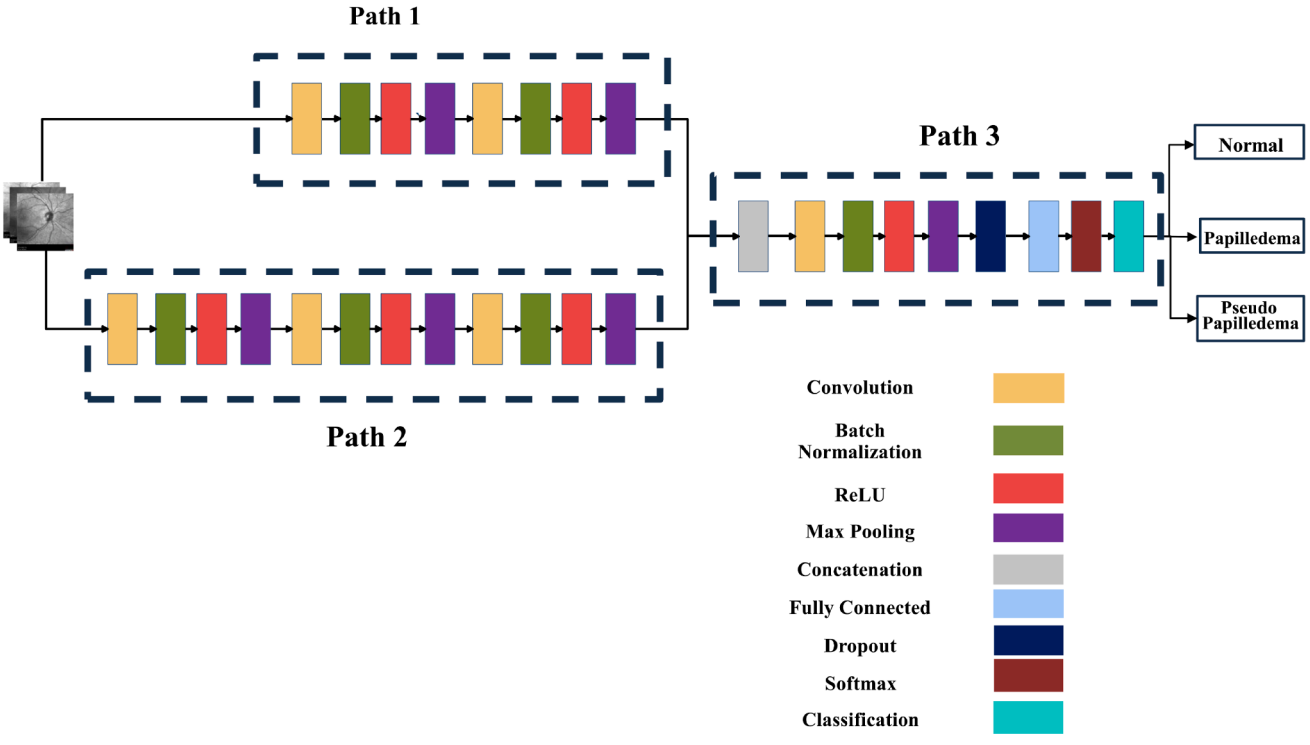


Fig. 4. Structure of the proposed multi-paths CNN model.

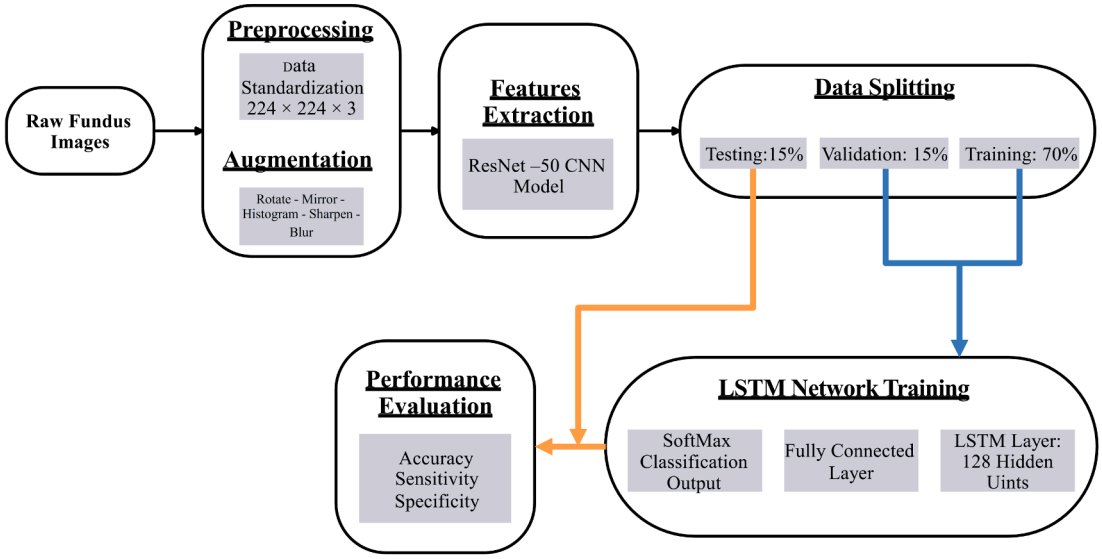


Fig. 5. An overview of the LSTM process for papilledema detection from fundus images.

4. Experimental Configuration and Performance Evaluation

4.1. Hyperparameters tuning and training

The hyperparameters for each model underwent adjustment using the stochastic gradient descent with momentum (SGDM) optimizer. Several crucial training settings were employed. The model underwent training over a span of 30 epochs, commencing with an initial learning rate of 0.0001, coupled with a piecewise learning rate schedule. We conducted periodic validation, updating the model weights after every 600 iterations. To strike a balance between efficiency and accuracy, a mini-batch size of 64 was selected. The experimental setup comprised an Intel® Core™ i9-12900KF 3.19 GHz processor, an NVIDIA GeForce RTX 3080, operating on Windows 11, 64-bit, and equipped with 64 GB of RAM. Furthermore, all experiments were meticulously validated using MATLAB R2022b software, complemented by the parallel computing toolbox, the deep learning toolbox, and the statistics toolbox.

4.2. Performance metrics and evaluation

An extensive array of performance criteria is utilized to assess the effectiveness of the proposed deep learning models. These metrics encompass classification accuracy (CA), sensitivity, specificity, precision, and F1-score. Formulas for these evaluation metrics are provided in Equations (1)–(5), with TP representing true positives, TN representing true negatives, FP representing false positives, and FN representing false negatives [39,40].

$$\text{Accuracy} = (TP + TN)/(TP + FN + FP + TN) \quad (1)$$

$$\text{Sensitivity} = TP/(TP + FN) \quad (2)$$

$$\text{Specificity} = TN/(TN + FP) \quad (3)$$

$$\text{Precision} = TP/(TP + FP) \quad (4)$$

$$\text{F1-Score} = (2 \times (\text{Sensitivity} \times \text{Precision}))/(\text{Sensitivity} + \text{Precision}) \quad (5)$$

5. Results

The study presents a promising deep learning framework for the detection of papilledema from fundus images. The proposed approach utilizes two different and innovative techniques for the detection process using the concepts of CNN and RNN. In implementation of the proposed techniques, the dataset (18,258 fundus image) was split into 70 % (12,782 images) for the training of the models, 15 % (2,738 images) for the validation and hyperparameters tuning, and 15 % (2,738 images) for the testing and performance evaluation of the models.

5.1. Performance evaluation of the multi-paths CNN model

The multi-paths CNN model has been evaluated through computation of five essential evaluation criteria. To obtain these metrics, the values of TP, TN, FP, and FN were extracted by analyzing the confusion matrix, as visually represented in Table 2. This comprehensive evaluation provides valuable insights into the model's effectiveness and its ability to accurately classify cases.

The results indicate high performance metrics for the deep learning models, with a classification accuracy of 99.97 %, sensitivity of 99.95 %, specificity of 99.97 %, precision of 99.97 %, and mean F1-Score of 99.96 %. To delve deeper into the evaluation of the model's performance, an extensive statistical analysis was conducted. The obtained results revealed a mean value of 1.9441 and a standard deviation of 0.7623. These metrics collectively demonstrate the models' proficiency in accurately detecting papilledema in retinal fundus images, underscoring their potential utility in clinical ophthalmic diagnostics. The multi-paths CNN model took 2.5 h for the training phase. The devised system achieved an image classification time of 11.3 ms. More details for the class-by-class performance are available in Table 3.

5.2. Performance evaluation of the LSTM model

In the evaluation of the LSTM model for papilledema detection, the following performance metrics are presented based on the model's predictions across three classes: normal, papilledema, and pseudo

Table 2
Confusion matrix for the multi-paths CNN model of Papilledema detection.

		Predicted Values		
		Normal	Papilledema	Pseudo Papilledema
Actual Values	Normal	876	0	0
	Papilledema	0	1138	0
	Pseudo Papilledema	0	1	723

Table 3
Detailed results for the multi-paths CNN model.

Criteria	Normal	Papilledema	Pseudo papilledema	Total Performance
TP	876	1138	723	2737
TN	1862	1599	2014	5475
FP	0	1	0	1
FN	0	0	1	1
Accuracy	100.00 %	99.96 %	99.96 %	99.97 %
Sensitivity	100.00 %	100.00 %	99.86 %	99.95 %
Specificity	100.00 %	99.94 %	100.00 %	99.97 %
Precision	100.00 %	99.91 %	100.00 %	99.97 %
F1-Score	100.00 %	99.96 %	99.93 %	99.96 %

papilledema. The confusion matrix shown in Table 4 indicates results contributed to a total accuracy of 99.81 %, sensitivity of 99.71 %, specificity of 99.84 %, precision of 99.75 %, and an overall F1-Score of 99.73 %, demonstrating the LSTM model's robust performance. In an extended examination of the model's efficacy, a thorough statistical analysis was executed, yielding a mean value of 1.6899 and a standard deviation of 0.4976. These statistical metrics offer a nuanced perspective on the model's performance, shedding light on both the average prediction and the degree of variability within the results. The LSTM model took 50 min for the process of features extraction and training. By computing the classification time of the tested images, it achieved 3.65 ms. Overall performance evaluation results are presented in Table 5.

5.3. Validation and performance assessment

To validate the proposed models and ensure their robustness, extensive testing was conducted using diverse publicly available datasets. After investigation of the literature, the utilized dataset for the external validation process is the only publicly accessible dataset including the target classes. The dataset was meticulously categorized into the primary classes relevant to the problem at hand: normal, papilledema, and pseudo-papilledema [41] including unbalanced distribution over the three classes as 295 images for papilledema, 295 images for pseudo papilledema, and 779 images for normal cases. The testing sample comprised a balanced distribution of 450 fundus images across these target classes. The comprehensive performance results of both the multi-paths CNN and LSTM models are detailed in Table 6. This rigorous validation process and balanced dataset distribution contribute to a thorough understanding of the models' capabilities and generalizability across different scenarios.

5.4. Occlusion sensitivity test

Both of the proposed models have been used to perform the occlusion sensitivity test and the results are visually represented in Fig. 6. This occlusion sensitivity test involved systematically occluding specific regions of retinal fundus images and assessing the impact on the models' performance metrics.

Table 4
Confusion matrix for the LSTM model of papilledema detection.

		Predicted Values		
		Normal	Papilledema	Pseudo Papilledema
Actual Values	Normal	870	5	0
	Papilledema	2	1137	0
	Pseudo Papilledema	0	1	724

Table 5
Detailed results for the LSTM model of Papilledema detection.

Criteria	Normal	Papilledema	Pseudo papilledema	Total Performance
TP	870	1137	724	2731
TN	1862	1594	2014	5470
FP	2	6	0	8
FN	5	2	1	8
Accuracy	99.74 %	99.71 %	99.96 %	99.81 %
Sensitivity	99.43 %	99.82 %	99.86 %	99.71 %
Specificity	99.89 %	99.63 %	100.00 %	99.84 %
Precision	99.77 %	99.48 %	100.00 %	99.75 %
F1-Score	99.60 %	99.65 %	99.93 %	99.73 %

Table 6
Performance Evaluation for the External Validation of the proposed models.

Criteria	Multi-paths CNN	LSTM
Accuracy	95.45 %	94.81 %
Sensitivity	93.16 %	92.19 %
Specificity	96.62 %	96.12 %
Precision	93.20 %	92.33 %
F1-Score	93.14 %	92.21 %

Fig. 6 illustrates the models' sensitivity to occlusion across different regions of the images. The results indicate the extent to which each model's predictive accuracy is influenced by the occlusion of specific regions. These findings shed light on the robustness and vulnerability of the models to obscured image regions, providing valuable insights into their performance in scenarios where certain areas of the fundus image may be obscured or degraded. This occlusion sensitivity test contributes to a deeper understanding of the models' strengths and weaknesses in papilledema detection under conditions of partial image obstruction, informing potential improvements and applications in real-world clinical settings.

6. Discussion

AI-based framework has been developed to detect papilledema from fundus images besides Pseudo-papilledema and normal cases. The study utilized different approaches of deep learning methods including CNN and RNN. The results obtained from this study are highly promising and reflect the potential of deep learning models, specifically the multi-paths CNN and LSTM models, in the realm of papilledema detection from retinal fundus images. Additionally, the limited number of raw fundus images has been expanded by the proposed augmentation method. As a result, the dataset has been duplicated 15 times. The multi-paths CNN model demonstrated exceptional performance across various evaluation criteria, including an impressive classification accuracy of 99.97 %, sensitivity of 99.95 %, and specificity of 99.97 %. This highlights the multi-paths CNN's potential as an invaluable diagnostic tool in clinical ophthalmology. The model's success can be attributed to its unique ability to capture intricate features at multiple levels of abstraction within fundus images. Leveraging parallel feature extraction paths, it excels in recognizing relevant patterns and structures, contributing to its high accuracy and reliability.

Similarly, the LSTM model exhibited robust performance, achieving an overall accuracy of 99.81 %, sensitivity of 99.71 %, and specificity of 99.84 %. These results affirm the LSTM model's effectiveness in distinguishing between normal cases, papilledema, and pseudo papilledema. The LSTM's strength lies in its capability to analyze sequential features extracted from retinal images. By processing the sequence of deep features obtained from the ResNet-50 model, the LSTM model gains a comprehensive understanding of image content, facilitating accurate classification. The proposed models demonstrate their adeptness

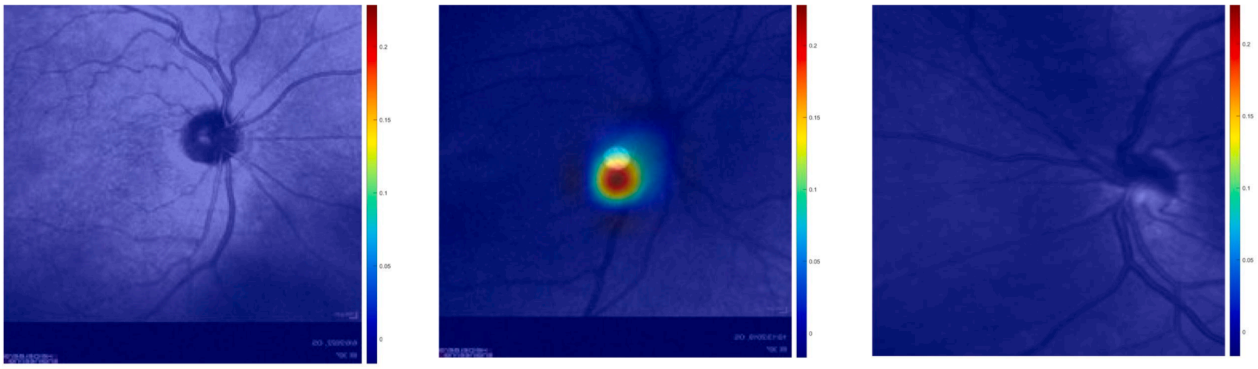


Fig. 6. Occlusion sensitivity test for (a) normal, (b) papilledema, (c) pseudo papilledema.

in swiftly analyzing and categorizing images, achieving commendable classification times for the tested samples. This underscores the efficiency of the models in handling image classification tasks with notable speed and accuracy. To further demonstrate the efficacy of temporal features, particularly LSTM approach, in our proposed models, it is essential to delve into the distinctive characteristics of fundus images. Fundus images capture dynamic information, and the temporal aspect becomes crucial for discerning subtle changes over time. In the context of papilledema detection, where pathological signs may evolve gradually, LSTM proves advantageous in capturing sequential patterns that contribute to accurate classification.

Unlike static images, fundus images may exhibit temporal variations in the appearance of critical features, and LSTM excels in learning and remembering such sequential dependencies. The ability of LSTM to retain information over extended sequences enables the model to capture nuanced variations and dependencies in fundus images that might be indicative of underlying pathologies.

While we acknowledge the potential of transformer-based approaches, we contend that the nature of fundus image data, characterized by temporal intricacies, aligns well with the strengths of LSTM. In our comparative analysis, we will elaborate on the advantages of LSTM over other temporal modeling techniques, providing a more comprehensive understanding of the rationale behind our choice.

In a comparative analysis with relevant literature, the performance of the proposed multi-paths CNN model significantly surpasses that of the approach presented by Sun et al. [23], demonstrating a substantial 6.97 % increase in classification accuracy. It is noteworthy that our model is intricately tailored for the specific problem of papilledema detection, presenting a departure from Sun et al.'s strategy of employing pre-trained models laden with a considerable number of layers. Our proposed model, comprising 29 layers, was deliberately designed to strike a balance between precision and efficiency. In stark contrast, Sun et al. utilized three transfer learning models, namely EfficientNet B7, DenseNet, and ResNet 101, boasting 813, 121, and 347 layers, respectively. This stark contrast in model complexity inevitably influenced prediction times, rendering our approach a more streamlined and effective solution with faster and more efficient outcomes in the context of papilledema detection.

Furthermore, the proposed multi-paths model demonstrated a remarkable improvement in classification accuracy by approximately 12.07 %. Sensitivity also saw a notable increase of 8.15 %, while specificity exhibited a substantial enhancement of 13.77 % over the study proposed by *Vasseneix et al.* [27]. These findings substantiate the significant contributions made by our proposed model and underscore the robustness of our investigation, particularly in employing different levels of features through parallel paths, which has evidently bolstered the overall performance of the system.

While Saba et al. [28] achieved an impressive 99.17 % classification accuracy in papilledema detection, a more in-depth examination reveals certain limitations within their methodology. Their model development relied on a relatively small dataset, comprising only 100 fundus images, which, as observed, may not provide the robustness and reliability needed for optimal model performance. In stark contrast, our study considered a substantial dataset comprising 18,258 images for the training process, facilitating greater generalization and enhancing overall model reliability and has been validated using another dataset. Moreover, despite the inherent challenges associated with limited data utilization, our proposed models not only matched but surpassed their accuracy by a notable 0.8 %.

The Multi-Paths CNN model we propose not only outperforms the accuracy achieved in He et al.'s study [14], which employed an interpretable Swin poly transformer for retinal disorders classification but also demonstrates remarkable performance across various evaluation criteria. Also, both of the proposed models attained higher accuracy than the model proposed by Song et al. [15] for retinal disorders detection including glaucoma using deep transformer mechanism which attained 88.3 % classification accuracy. This includes surpassing studies that specifically focused on integrating transfer learning approaches [23–25,28] and studies that customized CNN models [21,29,32]. Our model's efficacy is evident in its ability to excel in comparison to diverse methodologies, highlighting its versatility and superiority in retinal disorder classification.

To the best of our knowledge, this study represents a pioneering effort in leveraging RNN for the task of papilledema detection, a departure from the predominant literature which predominantly employed deep learning models, either through transfer learning or specially crafted architectures. The cascaded LSTM model introduced herein capitalizes on the strengths of both CNNs for feature extraction and RNNs for feature sequence learning, thus surpassing the performance achieved by previously proposed cascaded models by *Saba et al.* [28] and *Yang et al.* [34]. This novel combination of techniques marks a significant advancement in papilledema detection, offering enhanced capabilities compared to existing approaches.

Additionally, our proposed models demonstrated noteworthy improvements in training time, particularly in comparison to the study conducted by Saba et al. [28]. The Multi-Paths CNN model exhibited an impressive time reduction of 62.21 %, while the LSTM model achieved an even more substantial reduction of 88.66 %. It is worth noting that many existing studies in literature lack explicit discussions on training and prediction times. Recognizing the significance of these computational aspects, we emphasize the necessity of including such metrics in the evaluation criteria. This not only provides valuable insights into the efficiency of the models but also underscores their practical utility as assistive tools for neuro-ophthalmologists, aiding in the precise

diagnosis and formulation of treatment plans.

Moreover, several investigations have put forth models to address a common issue by leveraging DenseNet, primarily chosen for its efficacy in mitigating the vanishing gradient problem. The key distinctions manifest in the architectural frameworks employed: Multi-Paths CNN employs separate paths for explicit feature extraction, whereas DenseNet relies on dense connectivity. Further nuances emerge in terms of the number and dimensions of filters, hyperparameters, and the specific configurations of layers, all of which collectively shape the distinctive attributes of each model. In contrast, DenseNet employs dense connectivity, where each layer receives direct inputs from all preceding layers within a dense block. In terms of filter sizes, Multi-Paths CNN utilizes 5 $\times$ 5 and 3 $\times$ 3 convolutional layers across its paths, whereas DenseNet predominantly relies on 3 $\times$ 3 convolutions within its densely connected blocks.

The number of filters also differs significantly. For local feature extraction in Multi-Paths CNN, a 5 $\times$ 5 convolutional layer with 32 filters is followed by another layer with 64 filters. The intermediate path starts with a 3 $\times$ 3 convolutional layer with 16 filters, followed by two additional layers with 32 and 64 filters. The global path combines features from the previous paths and employs a 3 $\times$ 3 convolutional layer with 256 filters.

DenseNet, on the other hand, features dense blocks with a fixed number of filters (e.g., 128 filters in DenseNet-121) [42,43]. Each layer within a dense block contributes to the growth of feature maps by concatenating outputs. This dense connectivity helps alleviate the vanishing gradient problem and enhances feature reuse. Hyperparameters and layer configurations also differ. Multi-Paths CNN employs batch normalization and ReLU activation after each convolutional layer. In contrast, DenseNet uses batch normalization and a combination of ReLU and bottleneck layers for efficient parameter utilization.

During the external validation phase, our proposed models demonstrated results on par with existing literature, particularly aligning closely with the findings of Ahn et al. [35] and Kokulu et al. [25]. In their studies, the authors employed traditional pretrained CNN models, achieving outcomes comparable to our customized CNN model designed specifically for the targeted problem. Noteworthy is the advantage of our approach, as we utilized a more extensive dataset of 450 images for the testing phase, in contrast to Ahn et al.'s use of 219 images. This difference in dataset size may exert an influence on the overall result, further emphasizing the robustness and reliability of our model in handling a more extensive range of test samples. The majority of studies covered in the literature review relied on locally gathered datasets, which, regrettably, are not publicly accessible. In summary, a comparative analysis between our proposed models and relevant prior works is formed in Table 7, focusing on key metrics such as accuracy, sensitivity, and specificity.

7. Conclusion

This study presents pioneering models designed for the detection of papilledema within retinal fundus images. Our approach encompasses two distinctive methodologies. The first involves the development of a multi-paths CNN model, which intricately leverages diverse feature levels. The second methodology entails a cascaded model that initiates with feature extraction via ResNet-50 and culminates in training an LSTM model for the classification task. The proposed method for augmentation expanded the dataset to an adequate level for training. These innovative models exhibit remarkable performance when compared with prior research in the field, especially for the detection of papilledema using the proposed methodology. The comprehensive evaluation of the proposed models reveals their proficiency in classifying retinal fundus images, offering both medical practitioners and patients a reliable and efficient diagnostic solution. This advancement represents a substantial leap forward in the domain of papilledema detection, with the potential to enhance diagnostic accuracy and facilitate early intervention. This study introduces a freshly curated dataset explicitly assembled for the implementation of our proposed models. Furthermore, to ascertain the reliability and robustness of our models, rigorous testing has been conducted on an external dataset. The meticulous curation of our dataset ensures its relevance and specificity to the problem at hand, while the validation on an external dataset indicates to the generalizability and effectiveness of our proposed models beyond the confines of the training data. The study serves as an assistive tool for neuro-ophthalmologists to improve patient outcomes and reduce the burden on healthcare systems. The study found that papilledema cases were the dominant disease. Moreover, almost all papilledema cases were basically diagnosed with RNFL thicknesses, total peripapillary thickness, and visual field measurements. Therefore, there is a correlation between papilledema disorder and these measurements. In future work, further work is suggested to seek such a relationship. Correlating the clinical data of the patients to their acquired fundus images may lead to an interpretation of disorder causes. Additionally, more measurements could be investigated for their impact. Moreover, other techniques for the construction of the multi-paths CNN model can be implemented to further verify whether the method increases the complexity of the model and whether it is cost-effective at the same time.

Ethical standards

All procedures performed in studies were in accordance with the ethical standards.

There is no any reproduced material in the study.

There are no any experiments held on patients in the study.

Funding

The study did not receive any funding.

CRediT authorship contribution statement

Ahmed M. Salaheldin: Writing – review & editing, Writing –

Table 7
Performance comparison of the presented study and relevant works for the detection of papilledema.

Study	Accuracy	Sensitivity	Specificity	Image type	Classifier
Sun et al. [23]	93.00 %	92.22 %	99.07 %	Fundus	EfficientB7
Biousse et al. [24]	NA	84.0 %	89.6 %	Fundus	DenseNet
Kokulu et al. [25]	96.00 %	95.66 %	97.33 %	Fundus	MobileNet V2
Vasseneix et al. [27]	87.90 %	91.80 %	86.20 %	Fundus	Customized CNN
Saba et al. [28]	99.17 %	98.63 %	97.83 %	Fundus	DenseNet
Cen et al. [29]	NA	97.8 %	99.6 %	Fundus	Customized CNN
Farazdaghi et al. [30]	NA	86.00 %	88.00 %	Ultrasound	Clinical Based Study
		83.00 %	76.00 %	OCT	
Li et al. [32]	NA	85.05 %	29.15 %	Fundus	Customized CNN
Yang et al. [34]	96.1 %	95.3 %	96.7 %	Fundus	Logistic regression
Proposed study	**99.97 %**	**99.95 %**	**99.97 %**	Fundus	Multi-paths CNN
Proposed study	**99.81 %**	**99.71 %**	**99.84 %**	Fundus	LSTM

original draft, Validation, Software, Investigation, Formal analysis. **Manal Abdel Wahed:** Writing – review & editing, Supervision, Conceptualization. **Manar Talaat:** Data curation. **Neven Saleh:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] E.A. Muro-Fuentes, L. Stunkel, Diagnostic error in neuro-ophthalmology: avenues to improve, Curr. Neurol. Neurosci. Rep. 22 (2022) 243–256, https://doi.org/10.1007/s11910-022-01189-4.

[2] N. Saleh, M. Abdel Wahed, A.M. Salaheldin, Transfer learning-based platform for detecting multi-classification retinal disorders using optical coherence tomography images, Int. J. Imaging Syst. Technol. 32 (2022) 740–752, https://doi.org/10.1002/ima.22673.

[3] J.S. Xie, L. Donaldson, E. Margolin, Papilledema: a review of etiology, pathophysiology, diagnosis, and management, Surv. Ophthalmol. 67 (2022) 1135–1159. 10.1016/j.survophthal.2021.11.007.

[4] W. Chan, A.M. Flowers, B.I. Meyer, B.B. Bruce, N.J. Newman, V. Biousse, Acute central retinal artery occlusion seen within 24 hours at a tertiary institution, J. Stroke Cerebrovasc. Dis. 30 (2021) 105988.

[5] Z. Rissan, N.S. Hansen, L.N. Carlsen, J.J. Korsbæk, R.H. Jensen, S. Hamann, H. W. Schytz, Fundus imaging and perimetry in patients with idiopathic intracranial hypertension—an intermethod and interrater validity study, Eur. J. Neurol. 30 (2023) 1973–1982. 10.1111/ene.15802.

[6] U. Bhimavarapu, Automatic detection of hypertensive retinopathy using improved fuzzy clustering and novel loss function, Multimed. Tools Appl. 82 (2023) 30107–30123, https://doi.org/10.1007/s11042-023-15044-2.

[7] A.M. Salaheldin, M. Abdel Wahed, N. Saleh, Machine Learning-Based Platform for Classification of Retinal Disorders Using Optical Coherence Tomography Images, in: M. Pandit, M.K. Gaur, P.S. Rana, A. Tiwari (Eds.), Artif. Intell. Sustain. Comput., Springer Nature Singapore, Singapore, 2022: pp. 269–283, doi: 10.1007/978-981-19-1653-3_21.

[8] M.T. Mohamed, A.H. Aldghaimy, O.A. Elsoghair, W.E. Aita, Diagnosis and grading of papilledema using optical coherence tomography compared to clinical staging by Frisén scale, SVU-Int. J. Med. Sci. 3 (2020) 14–19. 10.21608/svuijm.2020.111890.

[9] L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, L. Farhan, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, J. Big Data. 8 (2021) 1–74.

[10] N. Saleh, M. Abdel Wahed, A.M. Salaheldin, Computer-aided diagnosis system for retinal disorder classification using optical coherence tomography images, Biomed. Tech. (Berl.) 67 (2022) 283–294. 10.1515/bmt-2021-0330.

[11] F. Viel, R.C. Maciel, L.O. Seman, C.A. Zeferino, E.A. Bezerra, V.R.Q. Leithardt, Hyperspectral image classification: an analysis employing CNN, LSTM, transformer, and attention mechanism, IEEE Access 11 (2023) 24835–24850, https://doi.org/10.1109/ACCESS.2023.3255164.

[12] A. Septiarini, H. Hamdani, E. Setyaningsih, E. Junirianto, F. Utaminingrum, Automatic method for optic disc segmentation using deep learning on retinal fundus images, Healthc. Inform. Res. 29 (2023) 145–151, https://doi.org/10.4258/hir.2023.29.2.145.

[13] P. Dhruv, S. Naskar, Image classification using convolutional neural network (CNN) and recurrent neural network (RNN): a review BT - machine learning and information processing, in: D. Swain, P.K. Pattnaik, P.K. Gupta (Eds.), Springer Singapore, Singapore, 2020, pp. 367–381.

[14] J. He, J. Wang, Z. Han, J. Ma, C. Wang, M. Qi, An interpretable transformer network for the retinal disease classification using optical coherence tomography, Sci. Rep. 13 (2023) 3637, https://doi.org/10.1038/s41598-023-30853-z.

[15] D. Song, B. Fu, F. Li, J. Xiong, J. He, X. Zhang, Y. Qiao, Deep relation transformer for diagnosing glaucoma with optical coherence tomography and visual field function, IEEE Trans. Med. Imaging 40 (2021) 2392–2402, https://doi.org/10.1109/TMI.2021.3077484.

[16] M. Ma, X. Zhang, Y. Li, X. Wang, R. Zhang, Y. Wang, P. Sun, X. Wang, X. Sun, ConvLSTM coordinated longitudinal transformer under spatio-temporal features for tumor growth prediction, Comput. Biol. Med. 164 (2023) 107313. 10.1016/j.compbiomed.2023.107313.

[17] R. Liu, Z.-A. Huang, Y. Hu, Z. Zhu, K.-C. Wong, K.C. Tan, Spatial-temporal co-attention learning for diagnosis of mental disorders from resting-state fMRI data, IEEE Trans. Neural Netw. Learn. Syst. (2023) 1–15, https://doi.org/10.1109/TNNLS.2023.3243000.

[18] V. Asadpour, E.J. Puttock, D. Getahun, M.J. Fassett, F. Xie, Automated placental abruption identification using semantic segmentation, quantitative features, SVM, ensemble and multi-path CNN, Heliyon 9 (2023).

[19] F. Yuan, Z. Zhang, Z. Fang, An effective CNN and transformer complementary network for medical image segmentation, Pattern Recogn. 136 (2023) 109228. 10.1016/j.patcog.2022.109228.

[20] A.E. Ilesanmi, T. Ilesanmi, G.A. Gbotoso, A systematic review of retinal fundus image segmentation and classification methods using convolutional neural networks, Healthc. Anal. 4 (2023) 100261. 10.1016/j.health.2023.100261.

[21] D. Milea, R.P. Najjar, Z. Jiang, D. Ting, C. Vasseneix, X. Xu, M. Aghsaei Fard, P. Fonseca, K. Vanikieti, W.A. Lagrèze, C. La Morgia, C.Y. Cheung, S. Hamann, C. Chiquet, N. Sanda, H. Yang, L.J. Mejico, M.-B. Rougier, R. Kho, T.H.C. Tran, S. Singhal, P. Gohier, C. Clermont-Vignal, C.-Y. Cheng, J.B. Jonas, P. Yu-Wai-Man, C.L. Fraser, J.J. Chen, S. Ambika, N.R. Miller, Y. Liu, N.J. Newman, T.Y. Wong, V. Biousse, Artificial intelligence to detect papilledema from ocular fundus photographs, N. Engl. J. Med. 382 (2020) 1687–1695, https://doi.org/10.1056/NEJMoa1917130.

[22] H.K. Yang, Y.J. Kim, J.Y. Sung, D.H. Kim, K.G. Kim, J.-M. Hwang, Efficacy for differentiating nonglaucomatous versus glaucomatous optic neuropathy using deep learning systems, Am. J. Ophthalmol. 216 (2020) 140–146. 10.1016/j.ajo.2020.03.035.

[23] G. Sun, X. Wang, L. Xu, C. Li, W. Wang, Z. Yi, H. Luo, Y. Su, J. Zheng, Z. Li, Z. Chen, H. Zheng, C. Chen, Deep learning for the detection of multiple fundus diseases using ultra-widefield images, Ophthalmol. Ther. 12 (2023) 895–907, https://doi.org/10.1007/s40123-022-00627-3.

[24] V. Biousse, R.P. Najjar, Z. Tang, M.Y. Lin, D.W. Wright, M.T. Keadey, T.Y. Wong, B. B. Bruce, D. Milea, N.J. Newman, Application of a deep learning system to detect papilledema on nonmydriatic ocular fundus photographs in an emergency department, Am. J. Ophthalmol. (2023), https://doi.org/10.1016/j.ajo.2023.10.025.

[25] M. Kokulu, H. Göker, Detection of papilledema severity from color fundus images using transfer learning approaches, Aksaray Univ. J. Sci. Eng. 7 (2023) 53–61.

[26] S.A. Badawi, M.M. Fraz, M. Shehzad, I. Mahmood, S. Javed, E. Mosalam, A. K. Nileshwar, Detection and grading of hypertensive retinopathy using vessels tortuosity and arteriovenous ratio, J. Digit. Imaging 35 (2022) 281–301, https://doi.org/10.1007/s10278-021-00545-z.

[27] C. Vasseneix, R.P. Najjar, X. Xu, Z. Tang, J.L. Loo, S. Singhal, S. Tow, L. Milea, D.S. W. Ting, Y. Liu, T.Y. Wong, N.J. Newman, V. Biousse, D. Milea, Accuracy of a deep learning system for classification of papilledema severity on ocular fundus photographs, Neurology 97 (2021) e369–e377, https://doi.org/10.1212/WNL.0000000000012226.

[28] T. Saba, S. Akbar, H. Kolivand, S. Ali Bahaj, Automatic detection of papilledema through fundus retinal images using deep learning, Microsc. Res. Tech. 84 (2021) 3066–3077. 10.1002/jemt.23865.

[29] L.-P. Cen, J. Ji, J.-W. Lin, S.-T. Ju, H.-J. Lin, T.-P. Li, Y. Wang, J.-F. Yang, Y.-F. Liu, S. Tan, L. Tan, D. Li, Y. Wang, D. Zheng, Y. Xiong, H. Wu, J. Jiang, Z. Wu, D. Huang, T. Shi, B. Chen, J. Yang, X. Zhang, L. Luo, C. Huang, G. Zhang, Y. Huang, T.K. Ng, H. Chen, W. Chen, C.P. Pang, M. Zhang, Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks, Nat. Commun. 12 (2021) 4828, https://doi.org/10.1038/s41467-021-25138-w.

[30] M.K. Farazdaghi, C. Trimboli-Heidler, G.T. Liu, A. Garcia, G.-S. Ying, R.A. Avery, Utility of ultrasound and optical coherence tomography in differentiating between papilledema and pseudopapilledema in children, J Neuro-Ophthalmology. 41 (2021) 10.aspx. https://journals.lww.com/jneuro-ophthalmology/Fulltext/2021/12000/Utility_of_Ultrasound_and_Optical_Coherence.

[31] F. Demir, B. Taşcı, An effective and robust approach based on R-CNN+LSTM model and NCAR feature selection for ophthalmological disease detection from fundus images, J. Pers. Med. 11 (2021), https://doi.org/10.3390/jpm11121276.

[32] B. Li, H. Chen, B. Zhang, M. Yuan, X. Jin, B. Lei, J. Xu, W. Gu, D.C.S. Wong, X. He, H. Wang, D. Ding, X. Li, Y. Chen, W. Yu, Development and evaluation of a deep learning model for the detection of multiple fundus diseases based on colour fundus photography, Br. J. Ophthalmol. 106 (2022) 1079–1086, https://doi.org/10.1136/bjophthalmol-2020-316290.

[33] A. Imran, J. Li, Y. Pei, F. Akhtar, T. Mahmood, L. Zhang, Fundus image-based cataract classification using a hybrid convolutional and recurrent neural network, Vis. Comput. 37 (2021) 2407–2417, https://doi.org/10.1007/s00371-020-01994-3.

[34] H.K. Yang, J.E. Oh, S.B. Han, K.G. Kim, J.-M. Hwang, Automatic computer-aided analysis of optic disc pallor in fundus photographs, Acta Ophthalmol. 97 (2019) e519–e525, https://doi.org/10.1111/aos.13970.

[35] J.M. Ahn, S. Kim, K.-S. Ahn, S.-H. Cho, U.S. Kim, Accuracy of machine learning for differentiation between optic neuropathies and pseudopapilledema, BMC Ophthalmol. 19 (2019) 178, https://doi.org/10.1186/s12886-019-1184-0.

[36] J. Tong, C. Wang, A dual tri-path CNN system for brain tumor segmentation, Biomed. Signal Process. Control 81 (2023) 104411. 10.1016/j.bspc.2022.104411.

[37] N. Saleh, A.M. Salaheldin, A benchmarking platform for selecting optimal retinal diseases diagnosis model based on a multi-criteria decision-making approach, J. Chin. Inst. Eng. Trans. Chin. Inst. Eng. A 45 (2022) 27–34, https://doi.org/10.1080/02533839.2021.1983466.

[38] Y. Yu, X. Si, C. Hu, J. Zhang, A review of recurrent neural networks: LSTM cells and network architectures, Neural Comput. 31 (2019) 1235–1270.

[39] J.R. Taylor, An introduction to error analysis, 1982.

[40] N. Saleh, K. Momtaz, A.M. Salaheldin, Machine learning-based paradigm for diagnosis of gestational diabetes, in: 3rd IEEE Int. Conf. Electron. Eng. Menoufia Univ. Egypt., 2023.
[41] U. Kim, Machine Learning for Pseudopapilledema, OSF, 2018, https://osf.io/2w5ce/.
[42] T. Liao, L. Li, R. Ouyang, X. Lin, X. Lai, G. Cheng, J. Ma, Classification of asymmetry in mammography via the DenseNet convolutional neural network, Eur. J. Radiol. Open. 11 (2023) 100502. 10.1016/j.ejro.2023.100502.
[43] H.A. Sanghvi, R.H. Patel, A. Agarwal, S. Gupta, V. Sawhney, A.S. Pandya, A deep learning approach for classification of COVID and pneumonia using DenseNet-201, Int. J. Imaging Syst. Technol. 33 (2023) 18–38. 10.1002/ima.22812.