IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 18 October 2023, accepted 26 November 2023, date of publication 30 November 2023,
date of current version 6 December 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3338362*

RESEARCH ARTICLE

# Prediction of Postoperative Visual Acuity in Rhegmatogenous Retinal Detachment Using OCT Images

**SINDA HOSNI[1], HAJER KHACHNAOUI[1], HSOUNA MEHDI ZGOLLI[2], SONYA MABROUK[2], DÉSIRÉ SIDIBÉ[3], HEDI TABIA[3], AND NAWRES KHLIFA[1]**

[1]Laboratory of Biophysics and Medical Technology, Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, Tunis 1006, Tunisia
[2]Hedi Raies Institute of Ophthalmology, Tunis 1006, Tunisia
[3]IBISC Laboratory, Univ. Evry, Université Paris-Saclay, 91020 Evry, France

Corresponding author: Hajer Khachnaoui (hajerkhachnaoui69@gmail.com)

This work was supported in part by the OCTIPA Project, as part of the PHC-Utique Program managed by CMCU of the French Ministry of Europe and Foreign Affairs, under Grant CMCU 23G1418; and in part by the Tunisian Ministry of Higher Education and Scientific Research.

**ABSTRACT** Deep Learning (DL) methods, such as Convolution Neural Networks (CNNs), have shown great potential in diagnosing complex diseases. Among these diseases, Rhegmatogenous Retinal Detachment (RRD) stands out as a critical condition necessitating precise diagnosis and postoperative Visual Acuity (VA) prediction. This research introduces a DL-based Computer-Aided Diagnosis (CAD) system that utilizes Optical Coherence Tomography (OCT) images for both the diagnosis of RRD and the prediction of postoperative VA. The CAD system utilizes DL techniques and a diverse dataset, including OCT images of patients with RRD from the Hedi Raies Ophthalmology Institute of Tunis and a large public dataset of normal subjects OCT. Preprocessing steps, such as image cropping, enhancement, denoising, and resizing, are applied to the tomographic images. Data oversampling and augmentation techniques address class imbalance and improve the dataset by generating additional samples. Various DL models, including pre-trained CNN models (VGG-16, Inception-V3, Inception-ResNet-V2), Bilinear (BCNN) (BCNN $(VGG-16)^2$ and BCNN $(Inception-V3)^2$), and a custom CNN architecture, are implemented for RRD diagnosis and postoperative VA prediction. The experimental outcomes demonstrate the effectiveness of the proposed CAD system in accurately diagnosing RRD and predicting postoperative VA. The system achieves high accuracy, with 99.87% for diagnosing RRD and 98.06% for predicting postoperative VA using the BCNN $(VGG-16)^2$ model. The developed CAD system represents a significant advancement in the field of RRD and postoperative VA prediction. By combining DL and OCT imaging, the system provides automated and accurate diagnosis, showing potential in improving patient care and treatment decisions.

**INDEX TERMS** Convolutional neural networks, deep learning, CAD system, rhegmatogenous retinal detachment, OCT imaging.

## I. INTRODUCTION

The remarkable advancements in artificial intelligence have brought about transformative changes in various industries, and healthcare is certainly no exception [1]. In particular, the emergence of machine learning (ML) techniques, with a specific focus on deep learning (DL), has opened up new horizons in the field of medical imaging analysis [2]. One area that has seen significant progress is the development of computer-aided diagnosis (CAD) systems [2], [3], [4]. These systems harness the power of DL algorithms to assist healthcare professionals in the detection and diagnosis of a wide range of diseases. Within the realm of ophthalmology, the integration of CAD based on DL holds significant promise

The associate editor coordinating the review of this manuscript and approving it for publication was Okyay Kaynak.

for addressing a critical condition known as Retinal Detachment (RD) [5], [6]. This serious impairment occurs when the neurosensory retina becomes separated from the underlying retinal pigment epithelium due to the accumulation of subretinal fluid [6]. Rhegmatogenous Retinal Detachment (RRD) is due to retinal tear, this is a surgical emergency and carries severe visual impairment that can lead to blindness [7], [8]. Early diagnosis is crucial for reattachment's success rate and improving visual outcomes. However, identifying RRD in early stages presents challenges, as it often begins asymptomatically and progresses from the peripheral retina. Moreover, final visual outcome depends on surgery delay which is the most important preoperative prognosis visual factor [7]. Among the valuable imaging techniques in ophthalmology, optical coherence tomography (OCT) stands out for providing high-resolution cross-sectional images of the retina [9]. The integration of DL-driven CAD systems, along with OCT imaging capabilities, has the potential to revolutionize the diagnosis and management of RRD [10], [11], [12], [13]. These systems automate the detection process and leverage the computational power of DL algorithms to accurately identify RRD and retinal layers evolution. Furthermore, they hold promise in predicting postoperative Visual Acuity (VA), aiding in treatment planning and improving patient outcomes. The primary objective of this research is to develop a CAD system that utilizes DL on OCT images to identify RRD stage and predict postoperative VA. This addresses the significant challenge of predicting visual recovery during the preoperative examination in the management of RRD.

## II. RELATED WORKS

In the literature, there has been an increasing amount of research dedicated to exploring the utilization of DL methods in the field of retinal image analysis. By leveraging the power of DL, researchers aim to enhance the accuracy and efficiency of diagnosing retinal diseases through automated analysis of retinal images. For instance, Feng et al. [14] introduced a DL-based approach for the automated diagnosis of Choroidal NeoVascularization (CNV), Diabetic Macular Edema (DME), Drusen, and normal subjects in OCT images. The proposed method leverages an ensemble of four classification models, each employing an improved ResNet-50 architecture, to analyze retinal OCT images. The final predictions are obtained by calculating the average of class probabilities estimated by the constituent networks. The proposed method was evaluated on a dataset comprising 21357 retinal OCT images from the Shanghai Zhongshan Hospital and the Shanghai First People's Hospital. A 10-fold cross-validation technique was employed for the evaluation process. The experimental results highlight an impressive classification accuracy of 97.3%, along with a sensitivity of 96.3% and a specificity of 98.5%. In addition, Amit et al. [15] proposed a DL model aimed at classifying OCT images of the retina into four distinct categories of retinal diseases (CNV, DME, Drusen, and normal). To accomplish this, they employed the VGG-19 network, leveraging pre-trained weights derived from the ImageNet dataset through Transfer learning (TL) technique. Specifically, they modified the VGG-19 architecture by removing its top layers and substituting them with a flatten layer, a dropout layer, and an output dense layer. Importantly, all layers within the proposed CNN model were made trainable. To assess the performance of their CNN model, the researchers utilized the publicly available dataset known as "Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-ray Images." This dataset comprises a total of 84568 OCT retinal scan images and was used for training, validation, and testing purposes. Furthermore, they also validated their proposed method on an independent testing dataset obtained from Duke University, Harvard University, and the University of Michigan. This additional dataset consisted of OCT scanned retinal images from 45 patients. The results demonstrated the effectiveness of the proposed model, achieving an impressive classification accuracy of 99.17%. Moreover, the model exhibited high specificities of 99.5% and sensitivity of 99%.

Several research studies have been dedicated to focusing specifically on the utilization of DL for the diagnosis of RD. For instance, Hideharu et al. [10] conducted a study utilizing DL technique to automate the detection OF RRD from Ultra-Widefield Fundus (UWF) images. The dataset consisted of 831 UWF images, with 420 images from non-RRD patients and 411 images from RRD patients, obtained from the clinical database of the Ophthalmology Department of Tsukazaki Hospital. Among these images, 665 were used for training the model, while 166 images were used for model validation. To enhance the training dataset, data augmentation technique was applied, resulting in a total of 11970 images. The researchers employed a custom CNN model, composed of three convolutional layers followed by ReLU activation layers, pooling layers, and two fully connected layers. The final output layer employed a softmax function for binary classification. The custom CNN model outperformed the Support Vector Machine (SVM) method, achieving a sensitivity of 97.6%, specificity of 96.5%, and an Area Under the Curve (AUC) of 98.8%. Then, Zhongwen et al. [11] proposed a cascaded DL system to automate the detection of RD and differentiate between macula-on and macula-off RD based on UWF images. The cascaded DL system consisted of two models, both trained using the InceptionResNetV2 architecture. In their methodology, the first model was trained to identify RD in the input images. Subsequently, the second model was employed to further discern macula-on RD from the RD images detected by the first model. The proposed method was developed using a dataset of 11087 UWF images. To evaluate the performance of the DL models, the eligible images were randomly divided into three sets: 70% for the training set, 15% for the validation set, and 15% for the test set. The first DL model, used for identifying RD, achieved an impressive AUC of 98.9%, with a sensitivity of 96.1% and specificity of 99.6%. On the other hand, the second DL

model, employed for discerning macula-on RD from macula-off RD, achieved an AUC of 97.5%, with a sensitivity of 93.8% and specificity of 90.9%. In addition, Chenxi et al. [12] carried out a study aimed at developing a combined DL system for the detection of lattice degeneration, retinal break, and RD in tessellated eyes using Optos images. The system consisted of three binary classification CNN models, each designed to detect one of the aforementioned conditions. For training the binary classification models, the seResNext-50 architecture was utilized. The outputs of these three models were analyzed to generate the final prediction. The proposed method was evaluated on a dataset of Optos images consisting of 911 samples obtained from Beijing Hua'er Eye Hospital. The dataset included images of lattice degeneration (267), retinal breaks (49), RD (44), and normal peripheral retina (609). It was divided into three subsets, with 60% of the data used for training (employing data augmentation techniques), 20% for model validation, and 20% for performance assessment. The results demonstrated remarkable performance for the detection of RD, with an AUC value of 100%, sensitivity of 87.5%, and specificity of 100%. The overall accuracy of the system was found to be 79.8%. Recently, Sonal et al. [13] conducted a study aimed at diagnosing RD from non-RD images using various pre-trained CNN models and the TL technique on color fundus images. The study employed several well-known CNN models, including AlexNet, Inception-V3, GoogleNet, VGG-19, DenseNet, and ResNet-50. These models were trained and tested using publicly available datasets of RD and non-RD fundus images. Specifically, the training set consisted of 1227 images, whereas the testing set comprised 400 images. Through their experimentation, the researchers found that the ResNet50 framework, implemented through TL, exhibited the best classification performance. It achieved remarkable values for accuracy (99.50%), sensitivity (99.00%), specificity (99.99%), precision (99.99%), and F1 score (99.49%). This model outperformed other learning models in accurately detecting RD and non-RD fundus images.

The application of DL techniques in diverse imaging modalities, such as UWF images and Optos images, has greatly enhanced the precision and efficiency of diagnosing retinal diseases, with a particular focus on RD [10], [11], [12], [13], [14], [15]. These studies have showcased significant advancements, leading to improved patient care. However, it is worth noting that the number of studies specifically targeting RD diagnosis is limited. Moreover, a dedicated method for diagnosing RD and accurately predicting the final VA outcome using OCT images is currently absent, highlighting the need for further research in this area.

## III. MATERIALS AND METHODS

### A. PROPOSED METHOD

#### 1) MAIN CONTRIBUTIONS

The proposed method addresses a specific and critical medical condition, RRD, which necessitates precise diagnosis and accurate postoperative VA prediction. Our research introduces a DL-based CAD system meticulously tailored to RRD diagnosis and postoperative VA prediction, harnessing the potential of OCT images. The uniqueness of the proposed research study lies in several key contributions. Firstly, we have painstakingly designed and developed a CAD system, employing cutting-edge DL techniques, to automate RRD identification and predict postoperative VA based on preoperative OCT images. This system serves as a valuable complement to the diagnostic expertise of medical professionals, offering automated and precise insights. The most notable accomplishment of our study lies in the establishment of a DL-powered framework for predicting postoperative VA outcomes. Importantly, this system stands as the first of its kind to leverage the sophisticated capabilities of DL techniques for precise and reliable prediction of VA results. This advancement holds substantial potential for enhancing patient care and facilitating treatment decisions in the realm of RRD.

Furthermore, we extend our gratitude to the Hedi Raies Institute of Tunis for their invaluable support, providing OCT images from RRD patients. These images were meticulously prepared and annotated by ophthalmologists, drawing from postoperative patient follow-up spanning several months.

While the application of DL in medical imaging is not novel, our distinct focus on RRD, coupled with the integration of DL for VA prediction, stands as the hallmark of our unique contributions to the field.

#### 2) THE STEPS OF THE PROPOSED METHOD

The proposed method encompasses a multi-step process. Initially, various image enhancement techniques are applied to preprocess OCT images, involving image cropping to focus on the relevant region, image quality enhancement, noise reduction via a non-local mean filter, and resizing to a standardized $224 \times 224$ dimension. To address class imbalance, data oversampling is performed in the dataset. Furthermore, data augmentation techniques such as rotation, shear, zooming, and flipping are applied to augment the dataset. Subsequently, the dataset is partitioned into training and validation sets. Concurrently, different DL methods are employed for RRD diagnosis and postoperative VA prediction. Pre-trained CNN models, including VGG-16, Inception-V3, and Inception-ResNet-V2, are utilized using TL. Additionally, a Bilinear (BCNN) approach is applied using VGG-16 and Inception-V3 model combinations; BCNN $(\mathrm{VGG}-16)^2$ and BCNN $(\mathrm{Inception}-\mathrm{v3})^2$. Furthermore, a custom CNN architecture is developed and employed. The process for each CNN model involves two distinct phases: training and validation, followed by testing and evaluation on separate datasets. Initially, it is trained and evaluated for diagnosing RRD, functioning as the primary model. Subsequently, the identical model is reutilized for the subsequent training and evaluation phase, designated as the secondary model, with a specific emphasis on predicting postoperative VA. The overall methodology and workflow of the proposed method are

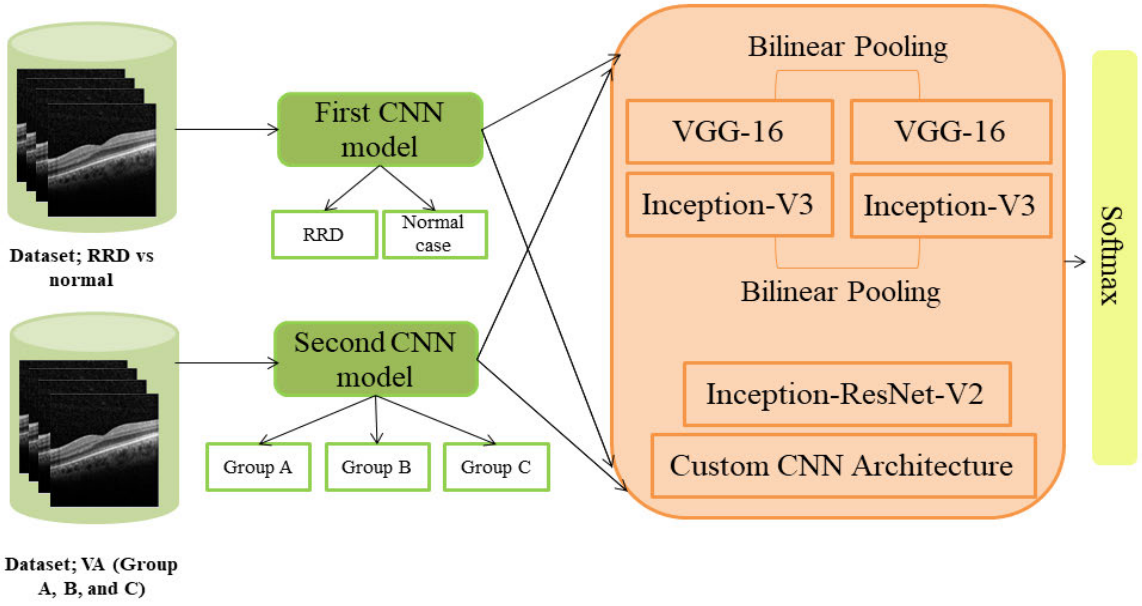



**FIGURE 1.** Block diagram of the proposed method.

depicted in Figure 1, providing a visual representation of the different steps involved.

We developed our models by utilizing an NVIDIA GeForce GTX 1080 Ti equipped with 11 GB of graphics memory. The development process involved Python 3 along with a range of libraries, including TensorFlow, Keras, Sci-Kit-Image, OpenCV, NumPy, Pandas, Matplotlib, and Sci-Kit-Learn, among others.

### B. PREPARATION OF DATASETS

#### 1) DATA DESCRIPTION

In this study, we utilized two datasets of OCT images representing both diseased and non-diseased patients. The first dataset, obtained from a local Tunisian database, consists of 468 OCT images from patients diagnosed with RRD who sought medical attention between 2013 and 2020 at the Ophthalmology Department of Institut Hédi Raies in Tunis. The labeling process for this dataset involved capturing OCT images after the surgery for each patient and categorizing them into three groups based on the postoperative final VA outcome, which were recorded as labels. These labels, representing the postoperative VA, were then assigned to the preoperative images to build a model capable of predicting the postoperative VA from preoperative images. The preoperative images were categorized into three distinct classes, namely Group A, Group B, and Group C, each representing different ranges of final VA outcomes. Group A signifies poor visual recovery, Group B denotes average visual recovery, and Group C indicates good visual recovery.

The second dataset used in this study is the ''Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images'' [16]. It is publicly accessible and encompasses a wide range of OCT and chest X-ray images [16]. From this extensive collection, we specifically extracted a subset of 486 OCT images that depict normal retinas.

These images are used in the development, which includes both the training and validation phases, of the proposed methodology. Additionally, two other datasets were specifically created for testing the models of the proposed method. The first dataset comprises 94 patients diagnosed with RRD and 104 individuals without the condition, and it was created for the initial CNN models. Furthermore, a second dataset, containing 20 images from Group A, 22 from Group B, and 23 from Group C, was created to test the second CNN model.

Thus, the dataset utilized for developing the first CNN models consisted of a total of 1152 images, comprising 562 RRD and 590 normal images. Within the RRD subset, 468 OCT images were used for training and validation, with 94 allocated for testing. Similarly, out of the 590 normal images, 486 underwent training and validation, while 104 were specifically reserved for testing. For the development of the second CNN models, a dataset of 533 images was categorized into Groups A, B, and C (257 in Group A, 204 in Group B, and 72 in Group C). In this dataset, 468 OCT images were utilized for training and validation purposes: 237 from Group A, 182 from Group B, and 49 from Group C. The remaining 65 images (22 from Group A, 22 from Group B, and 23 from Group C) were designated for testing.

A total of 562 RRD patients were part of the development for the first CNN models, yet only 533 were involved in constructing the second CNN models. Notably, 29 RRD images weren't considered for the second CNN models due to their

**TABLE 1. Summary of image datasets for proposed method development and testing.**

| | RRD | Normal | Group A | Group B | Group C |
|---|---|---|---|---|---|
| Number of images | 562 | 590 | 257 | 204 | 72 |

classification solely as RRD without the specific visual acuity labels corresponding to Groups A, B, and C.

Table 1 provides an overview of the total number of images used in constructing the proposed method, encompassing both the training and validation phases, as well as the subsequent testing phase.

Figure 2 illustrates examples of images from each group of the Tunisian database, as well as an example of the normal class from the public database.

#### 2) DATA PREPARATION AND PREPROCESSING

Prior to applying the CNN model, data preprocessing plays a crucial role in enhancing DL performance [17]. The dataset underwent several preprocessing steps as follows:

Firstly, image cropping was performed to eliminate irrelevant parts of the OCT images, focusing solely on the pertinent retinal region. This step is essential to optimize the analysis. Secondly, image enhancement techniques were employed to improve image quality. By increasing the overall contrast of the images, the impact of external factors such as light occlusion and patient movements, which can hinder retinal pathology detection, was mitigated. Next, noise reduction was addressed using the non-local mean filter [23]. OCT imaging often introduces significant noise due to coherent beam reflection. Reducing this noise is crucial as it can degrade image quality and impair the performance of subsequent image processing algorithms. Following noise reduction, the image size was standardized to $224 \times 224$ pixels to meet the requirements of the pre-trained CNN models utilized in the study.

These preprocessing steps were consistently applied to all images within the datasets, both during their development and testing phases. In the development phase, which included datasets containing RRD and normal classes, as well as datasets with three distinct groups (Group A, Group B, and Group C), a random split was executed. This split allocated 80% of the images for training purposes, while the remaining 20% were set aside for validation. This division was essential to ensure the effective training of the developed CNN models. As For the test set, a separate and independent preparation process was initiated from the outset. This involved meticulous steps to ensure that the test dataset remained distinct from any influence of the training or validation phases.

Tables 2 and 3 present a comprehensive breakdown of dataset composition, encompassing training, testing, and validation sets, along with the patient counts within each class. These tables cover both the dataset containing RRD and normal classes and the dataset comprising Group A, B, and C.

**TABLE 2. Training, validation, and testing data for RRD and normal cases.**

| | RRD | Normal case |
|---|---|---|
| Number of Training images | 374 | 389 |
| Number of validation images | 94 | 97 |
| Number of testing images | 94 | 104 |

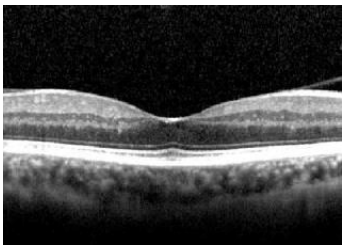
**TABLE 3. Training, validation, and testing data for Groups A, B, and C.**

| | Group A | Group B | Group C |
|---|---|---|---|
| Number of Training images | 190 | 146 | 39 |
| Number of validation images | 47 | 36 | 10 |
| Number of testing images | 20 | 22 | 23 |

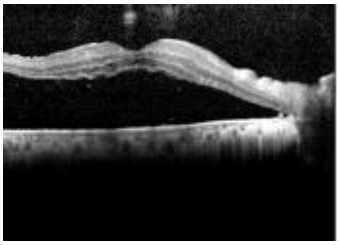
Data oversampling and data augmentation techniques were exclusively employed in the datasets used for the development phase of the proposed method. This choice was driven by the need to address data imbalances and the limited quantity of data, both of which could significantly affect the construction of robust CNN models. These techniques were intentionally omitted from the testing set. This decision was made on the basis that the testing set serves a distinct purpose: evaluation rather than model construction. Thus, to address the class imbalance within the dataset, which initially comprised Group A with 237 samples, Group B with 182 samples, and Group C with 49 samples, data oversampling was implemented to create a more balanced dataset. This approach ensured that each class within the dataset contained 237 images, resulting in a total of 711 images. In contrast, the dataset containing RRD and normal cases already had a relatively balanced number of images in each class, making data oversampling unnecessary for this dataset. Moreover, data augmentation techniques were applied. These data augmentation techniques included image rotation, a shear factor of 0.2, zooming, as well as vertical and horizontal flipping. As a result of this augmentation strategy, the dataset size significantly expanded to a total of 5724 samples, compared to the initial count of 954 for the dataset containing RRD and normal cases, and 4266 for the dataset comprising Group A, B, and C, a substantial increase from the initial 711. Extensive care was dedicated to guaranteeing the absence of any overlap between the training and validation subsets. Our approach involved taking deliberate measures to prevent instances where data samples generated from a single image would inadvertently find their way into both sets. This precaution was fundamental to maintaining the independence and integrity of the training and validation data, ensuring the robustness of our model development and evaluation process. Figure 3 presents various preprocessing steps applied to the OCT images.

### C. AN ANALYSIS OF EMPLOYED CNN ARCHITECTURE

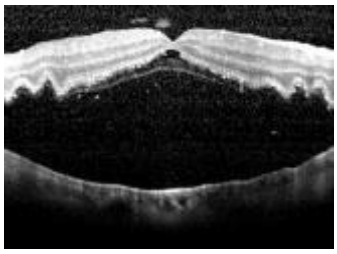
This study proposes a cascaded DL model consisting of two CNN models. The first model focuses on diagnosing RRD in OCT images, while the second model aims to classify the identified RRD cases based on their final VA outcomes. To achieve this, we harnessed the potential



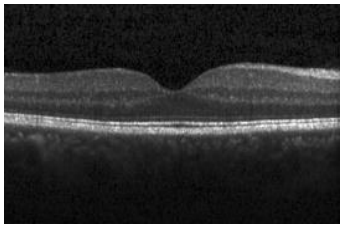
a) Representative images from Group A in the Tunisian database.



b) illustrative images showcasing Group B within the Tunisian database.



c) Depiction of images representing Group C as found in the Tunisian database.



d) Representative image of a normal case in the public database.

**FIGURE 2.** **Visual examples from tunisian database groups (Group A, Group B, and Group C) and comparative image from the normal class in the public database: A visual exploration of data diversity and representation.**

of TL by utilizing three pre-trained CNN models: VGG-16, Inception-V3, and Inception-ResNet-V2 [18], [19], [20]. These models were initially trained on the extensive ImageNet dataset, enabling them to learn valuable generic image features. Additionally, we explored two variations of BCNN architectures by combining the aforementioned pre-trained models: BCNN $(VGG-16)^2$ and BCNN $(Inception-v3)^2$ [21], [22]. Furthermore, we developed a custom CNN model tailored specifically to address our research problem. The categorical_crossentropy loss metric was selected to optimize the models' performance, and the Adam optimizer with a learning rate of 0.001 was used during training. The proposed method underwent 50 epochs of training, with a batch size of 32, to achieve the optimal performance.

### 1) TRANSFER LEARNING TECHNIQUE

TL is a powerful technique in DL that leverages the knowledge acquired by pre-trained CNN models like AlexNet, VGG, ResNet, and Inception. These models are trained on extensive datasets and excel at extracting important features from images [18], [19], [20], [21], [22]. By employing TL, we can take advantage of these pre-trained models' ability to recognize common patterns and structures in images, even for different tasks. This approach saves computational resources and time by building upon the existing knowledge of the pre-trained models and adopting them for our specific task. By transferring the learned representations, we can achieve faster convergence and potentially improve our model's performance. In this study, we specifically chose three popular pre-trained CNN models, namely VGG-16, Inception-V3, and Inception-ResNet-V2, known for their exceptional performance and feature extraction capabilities. These models have been pre-trained on large-scale ImageNet datasets, enabling them to learn generic visual features that can be applied to various image recognition tasks. Through TL, we utilize the learned representations from these models to address our specific problem of classifying RRD in OCT images.

Figure 4 illustrates the TL of the three pre-trained CNN (VGG-16, Inception-V3, and Inception-ResNet-V2) models

### 2) PRE-TRAINED VGG-16 ARCHITECTURE

VGG16 is a widely recognized and influential CNN architecture in the field of DL [18]. It was introduced by the Visual Geometry Group (VGG) at the University of Oxford and has gained significant popularity and success in various computer vision tasks. VGG16 is characterized by its deep structure, consisting of 16 weight layers, which enables it to capture intricate features and patterns in images. With its uniform architecture and small convolutional filters, VGG-16 achieves impressive performance in image classification and object recognition tasks. Due to its effectiveness and versatility, VGG-16 has become a go-to choice for many researchers and practitioners when working on image-related problems. Nevertheless, this pre-trained CNN model suffers from an architectural imbalance where more convolution layers are sequentially stacked in the upper layers than in the bottom layers due to resource constraints. This imbalance can result in inefficiencies in memory usage and computation, as the bottom layers, which possess higher spatial resolution, dominate the available resources. To overcome these limitations, advanced CNN architectures like the pre-trained Inception model have been developed. These architectures adopt a different approach by initially reducing the spatial resolution and subsequently introducing convolution layers.

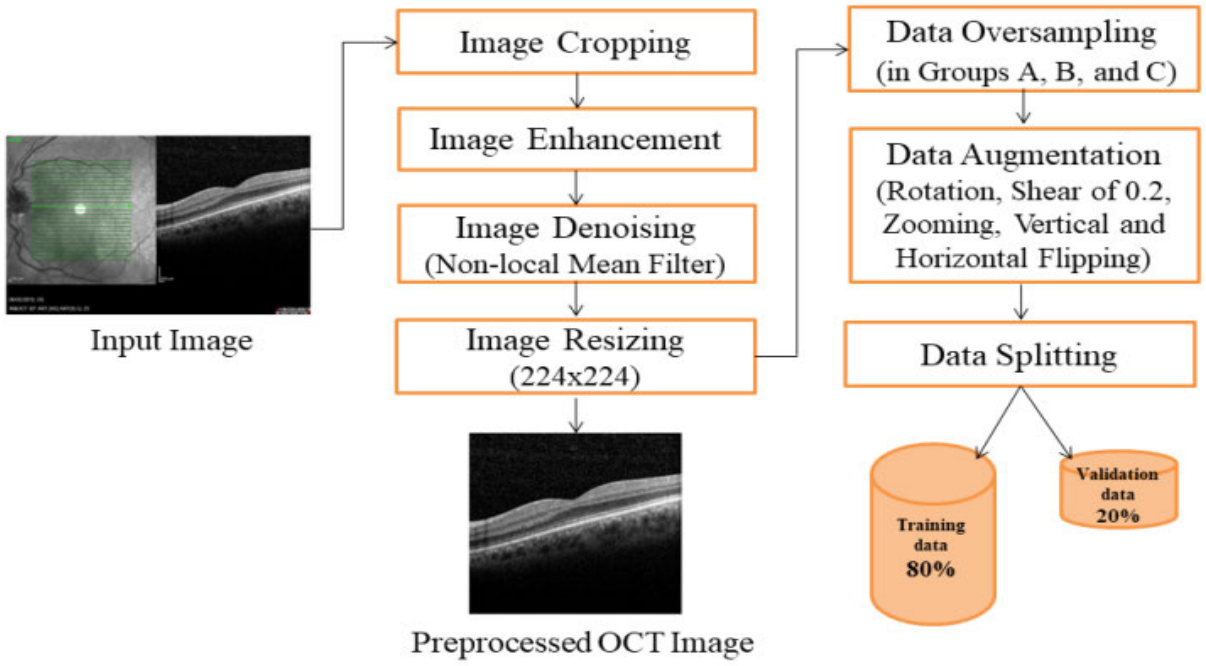



**FIGURE 3.** Preprocessing steps for OCT images.

This design choice aims to optimize resource allocation and improve overall performance.

### 3) PRE-TRAINED INCEPTION-V3 ARCHITECTURE

The Inception-V3 architecture, created by Google's research team, is a popular pre-trained CNN model extensively utilized in computer vision applications [19]. It surpasses its predecessors, Inception-V1 and V2, with its 48-layer structure and a larger input size of 299 × 299. By enabling deeper networks without a disproportionate increase in parameters, Inception-V3 effectively balances complexity and depth. This pre-trained CNN architecture excels at extracting features from images while simultaneously minimizing computational complexity. Its key characteristic is the use of inception modules, which employ parallel convolutional layers with different filter sizes to capture information at multiple scales and abstraction levels simultaneously. This allows the model to learn rich representations by capturing both low-level and high-level features. Inception-V3 also incorporates techniques like factorized convolutions to reduce parameters and computations while maintaining efficiency. It utilizes auxiliary classifiers to encourage the learning of discriminative features and overcome the vanishing gradient problem.

### 4) PRE-TRAINED INCEPTION-RESNET-V2 ARCHITECTURE

The Inception-ResNet-V2 architecture is a powerful pre-trained CNN model that combines the strengths of the Inception and ResNet architectures [20]. It was developed by the Google Brain team as an extension of the original Inception and ResNet models. Inception-ResNet-V2 is renowned for its exceptional performance in image recognition tasks, demonstrating state-of-the-art results on various benchmark datasets. The architecture incorporates residual connections, which help alleviate the vanishing gradient problem and enable efficient training of very deep networks. It also utilizes inception modules, similar to Inception-V3, to capture multi-scale features and enhance the model's ability to extract rich representations. In addition, Inception-ResNet-V2 employs batch normalization and factorized convolutions to improve computational efficiency while maintaining high accuracy.

### 5) BILINEAR CNN MODEL DESCRIPTION

The Bilinear CNN (BCNN) architecture has gained significant attention in computer vision for its ability to capture spatial interactions between features, making it effective in tasks like image classification and object recognition [21], [22]. BCNN utilizes bilinear pooling as a key component, combining the output vectors of two separate CNN architectures through an outer product function. This pooling operation captures pairwise feature interactions, enabling BCNN to encode higher-order feature correlations and extract fine-grained details and complex relationships in images. By leveraging both low-level and high-level features simultaneously, BCNN models enhance their discriminative power. These BCNN architectures provide a robust framework for extracting informative features from images and have the potential to improve our classification task. The bilinear model in BCNN is defined by the formula, $M = F(f_A, f_B, P, C)$ where $f_A$ and $f_B$ are feature extractors, $P$ is the pooling function, and $C$ is the classification function. Bilinear

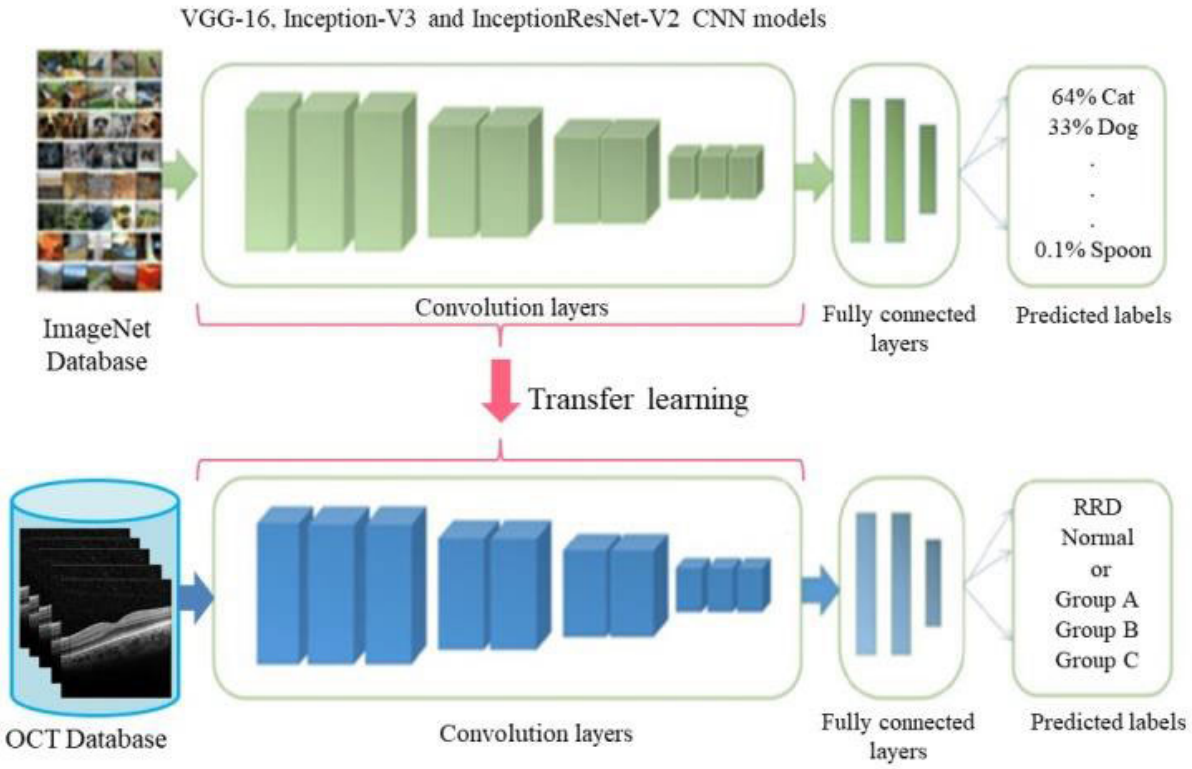



**FIGURE 4.** Transfer learning of pre-trained CNN Models (VGG16, Inception-V3, InceptionResNet-V2).

features are obtained by performing the outer product of $f_A$ and $f_B$, followed by pooling and normalization steps. The bilinear form allows for end-to-end training and simplifies gradient computation, making it an effective approach.

In this study, we explore two BCNN variants: BCNN $(\text{VGG}-16)^2$ and BCNN $(\text{Inception}-\text{v3})^2$, which combine pre-trained VGG16 and Inception-v3 models using bilinear pooling. These variants involve the combination of pre-trained VGG16 and Inception-v3 models using the bilinear pooling technique. By leveraging the complementary strengths of these well-established CNN architectures, we aim to harness the power of their learned representations and the rich spatial interactions captured by bilinear pooling. The BCNN $(\text{VGG}-16)^2$ variant merges two instances of the VGG16 model, while the BCNN $(\text{Inception}-\text{v3})^2$ variant combines two instances of the Inception-v3 model. Through this exploration, we seek to leverage the benefits of both architectures and exploit their synergistic effects in improving the performance of our classification task.

Figure 5 presents the BCNN architecture employed in this study.

### 6) CUSTOM CNN ARCHITECTURE DESCRIPTION

Given the complexity of the BCNN model, we have developed a custom CNN architecture from scratch. This custom CNN model is specifically developed with a reduced number of layers compared to pre-trained CNN models. Our goal is to simplify the architecture while preserving its effectiveness. The constructed CNN architecture consists of two blocks, each comprising a convolutional layer with 32 filters and a kernel size of $3 \times 3$. Following the convolutional layer, a ReLU activation function is applied to introduce non-linearity. Additionally, after each convolutional layer, we have incorporated a MaxPooling layer with a kernel size of $2 \times 2$. This pooling layer helps in downsampling the spatial dimensions of the input. Following the block structure, a Flatten layer was included to convert the output into a one-dimensional tensor. Subsequently, two dense layers are employed. The first dense layer consists of 128 nodes and employs the ReLU activation function to introduce non-linearity. Finally, the second dense layer is composed of 2 nodes, and the Softmax activation function is applied for the purpose of classification. Figure 6 presents the custom CNN Architecture employed in this study.

## IV. RESULTS

During this study, we have developed two cascaded models using DL techniques. The first model was specifically designed to detect RRD, while the second model focused on the prediction of postoperative VA. Each model was trained and evaluated using appropriate datasets.

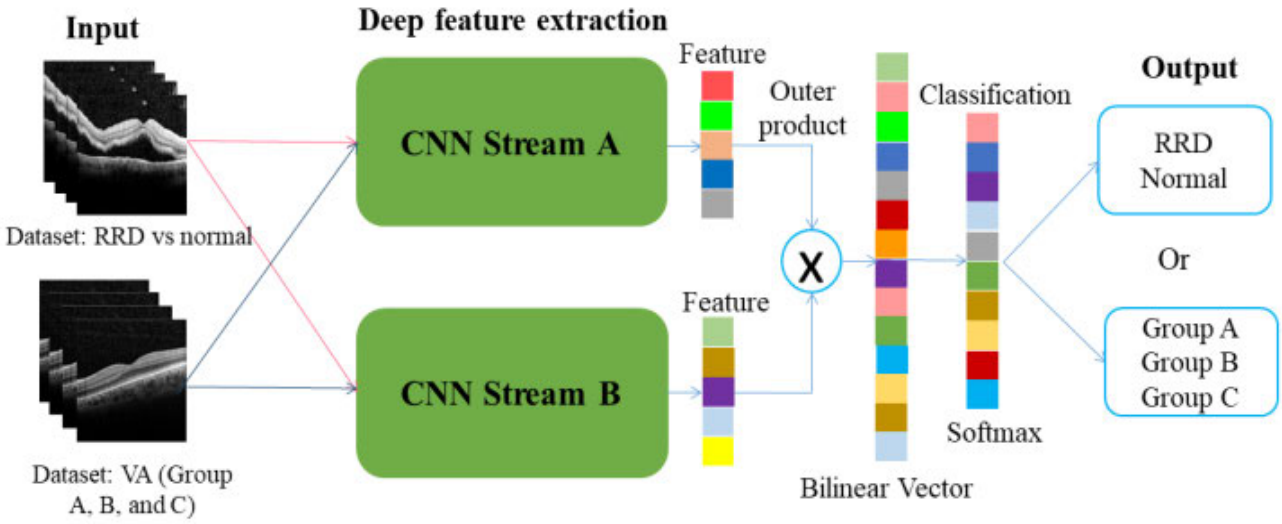



**FIGURE 5.** Bilinear pooling CNN architecture.

**TABLE 4.** Results of various CNN models FOR RRD diagnosis.

| Models | Accuracy | Precision | Recall | AUC | Sensitivty | Specificity |
|---|---|---|---|---|---|---|
| VGG16 | 97.32% | 97.32% | 97.32% | 97.90% | 100% | 95.19% |
| Inception-V3 | 96.79% | 96.97% | 96.97% | 99.76% | 91.35% | 95.74% |
| Inception-ResNet-V2 | 90.90% | 90.90% | 90.90% | 91.29% | 97.83% | 84.21% |
| BCNN$(\text{VGG}-16)^2$ | 99.87% | 99.87% | 99.87% | 99.99% | 100% | 95.19% |
| BCNN$(\text{Inception}-\text{V3})^2$ | 98.94% | 98.94% | 98.94% | 99.54% | 100% | 95.19% |
| Custom CNN model | 98.93% | 98.93% | 98.93% | 99.76% | 89.19% | 100% |

**TABLE 5.** Results of various CNN models for va prediction.

| Models | Accuracy | Precision | Recall | AUC | Sensitivty | Specificity |
|---|---|---|---|---|---|---|
| VGG16 | 88.37% | 88.37% | 88.37% | 91.27% | 77.36% | 87.31% |
| Inception-V3 | 90.76% | 90.76% | 90.76% | 98.08% | 93.99% | 96.83% |
| Inception-ResNet-V2 | 91.53% | 92.46% | 92.46% | 97.38% | 90.95% | 95.16% |
| BCNN$(\text{VGG}-16)^2$ | 98.06% | 98.58% | 98.58% | 98.58% | 86.25% | 92.6% |
| BCNN$(\text{Inception}-\text{V3})^2$ | 96.40% | 96.48% | 96.35% | 98.85% | 85.02% | 91.70% |
| Custom CNN model | 92.30% | 92.30% | 92.30% | 97.99% | 92.55% | 96.07% |

**TABLE 6.** Global performance evaluation of different CNN models.

| Models | Accuracy | Precision | Recall | AUC | Sensitivty | Specificity |
|---|---|---|---|---|---|---|
| VGG16 | 86.00% | 86.00% | 86.00% | 89.35% | 77.36% | 83.11% |
| Inception-V3 | 87.85% | 88.01% | 88.01% | 97.84% | 85.86% | 92.70% |
| Inception-ResNet-V2 | 83.20% | 84.05% | 84.05% | 88.90% | 88.98% | 80.13% |
| BCNN$(\text{VGG}-16)^2$ | 97.93% | 98.45% | 98.45% | 98.57% | 86.25% | 88.14% |
| BCNN$(\text{Inception}-\text{V3})^2$ | 95.38% | 95.45% | 95.33% | 98.39% | 85.02% | 87.29% |
| Custom CNN model | 91.31% | 91.31% | 91.31% | 97.75% | 82.54% | 96.07% |

These two models were developed using six CNN models: pre-trained VGG-16, pre-trained Inception-V3, pre-trained Inception-ResNet-V2, BCNN $(\text{VGG}-16)^2$, BCNN $(\text{Inception}-\text{V3})^2$, and a custom CNN model.

The performance and results of all six CNN models, are presented in Table 4 and Table 5. These tables provide an overview of the models' accuracy, precision, recall, sensitivity, specificity, and AUC.

The AUC represents the model's ability to make accurate classification decisions across various thresholds, without delving into the specific probability values. However, in the context of this study, our central emphasis is on the clas-

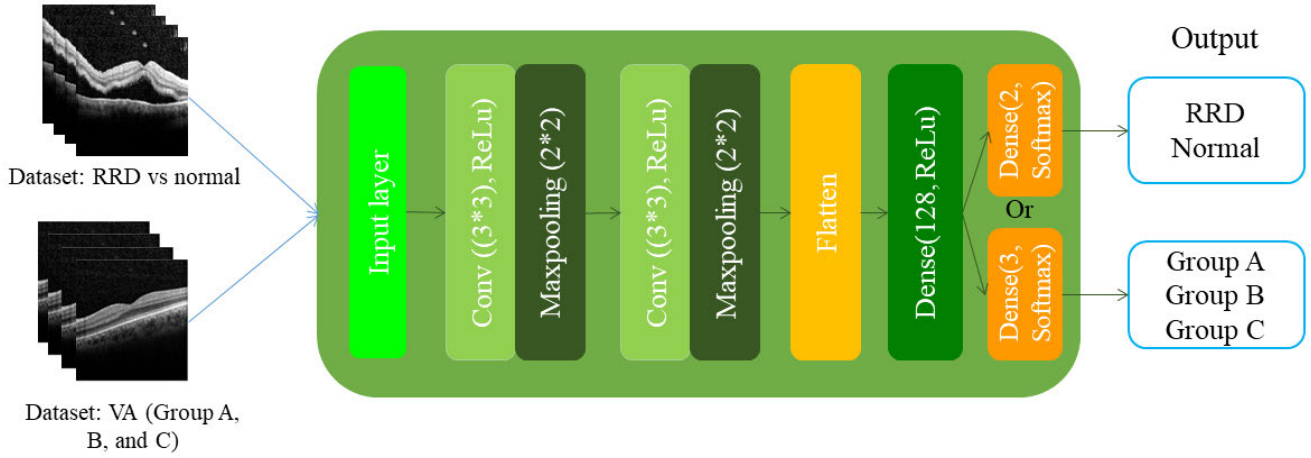



**FIGURE 6.** Custom CNN architecture.

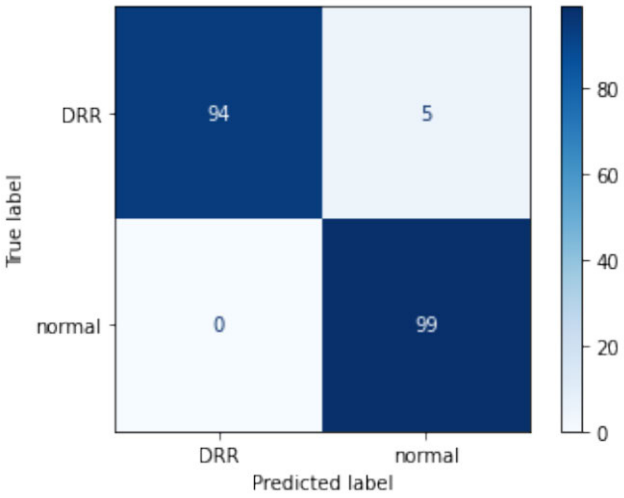



**FIGURE 7.** Confusion matrix: BCNN $(VGG-16)^2$ model on RRD and normal cases dataset.

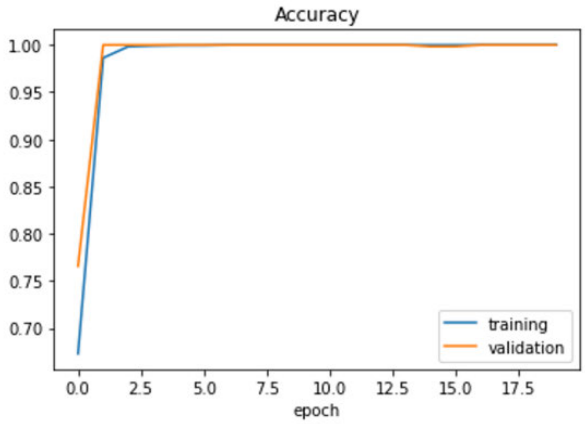



**FIGURE 8.** Training and validation curves for accuracy metric: BCNN $(VGG-16)^2$ model on RRD and normal cases dataset.

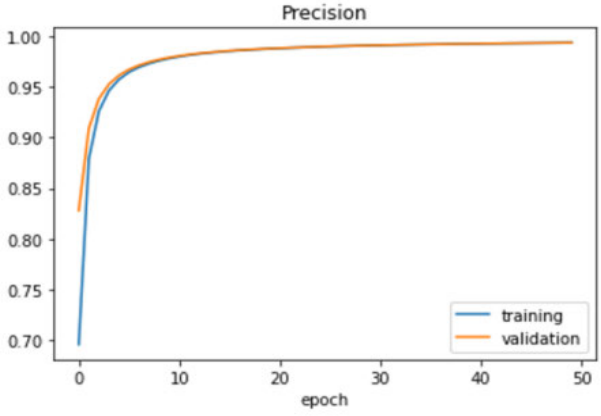



**FIGURE 9.** Training and validation curves for precision metric: BCNN $(VGG-16)^2$ model on RRD and normal cases dataset

sification outcomes rather than the specific probabilities. Therefore, we have opted to prioritize the selection of the highest AUC in this study.

To comprehensively assess performance, we combined the results of two models by multiplication: the classification of RRD and the prediction of VA. This approach allowed us to obtain a more precise measure of the overall system's reliability. The collective performance of the various CNN models is presented in Table 6.

Figures 7 to 16 offer insights into the performance of two CNN models through the presentation of confusion matrices, learning curves for key metrics such as Accuracy, Precision, Recall, and AUC. These figures specifically highlight the performance of two CNN models: BCNN $(VGG-16)^2$ trained on a dataset containing RRD and normal cases, and BCNN $(Inception-V3)^2$ trained on a dataset comprising data from Group A, Group B, and Group C.

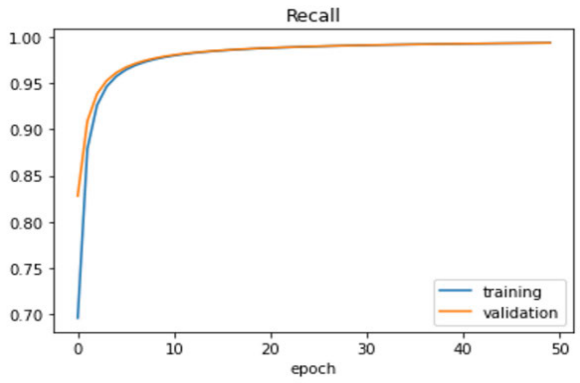



**FIGURE 10.** Training and validation curves for recall metric: BCNN $(VGG-16)^2$ model on RRD and normal cases dataset.

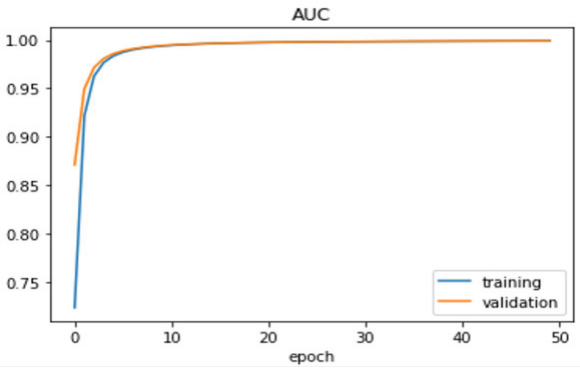



**FIGURE 11.** Training and validation curves for AUC metric: BCNN $(VGG-16)^2$ model on RRD and normal cases dataset.

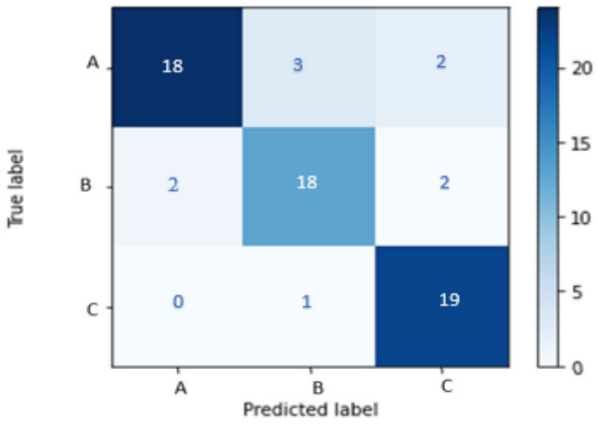



**FIGURE 12.** Confusion matrix: BCNN $(Inception-V3)^2$ model on Group A, Group B, and Group.

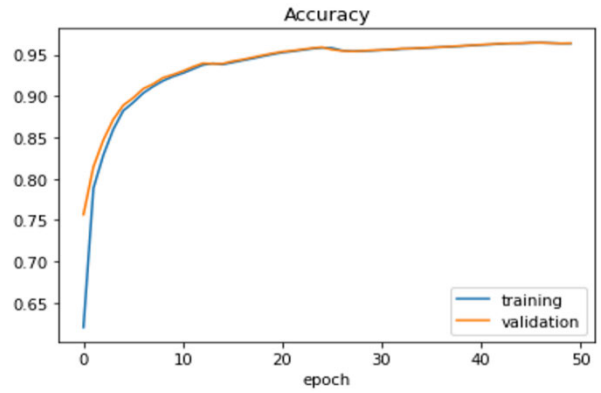



**FIGURE 13.** Training and validation curves for accuracy metric: BCNN $(Inception-V3)^2$ model on Group A, Group B, and Group.

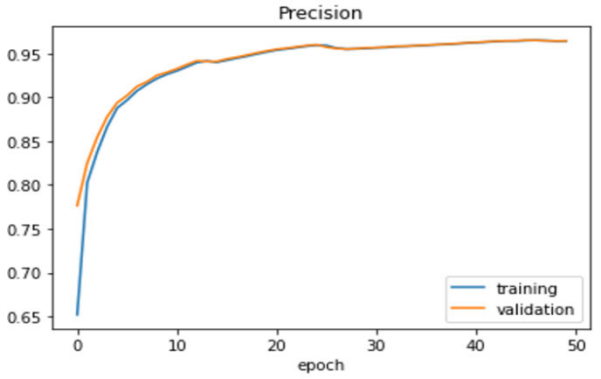



**FIGURE 14.** Training and validation curves for precision metric: BCNN $(Inception-V3)^2$ model on Group A, Group B, and Group.

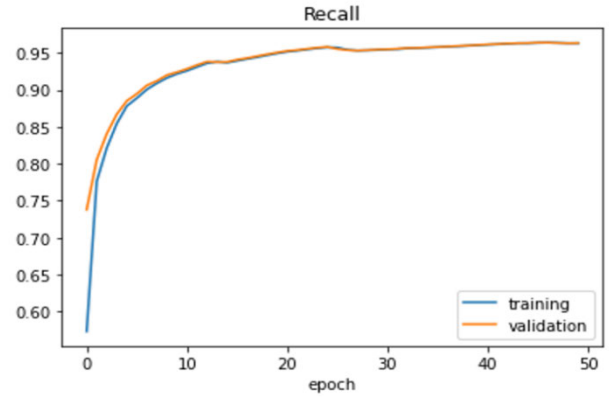



**FIGURE 15.** Training and validation curves for recall metric: BCNN $(Inception-V3)^2$ model on Group A, Group B, and Group.

## V. DISCUSSION

In this study, we have developed CNN cascaded model using DL. The first model was specifically designed for the identification of RRD, while the second model is dedicated to predicting the postoperative VA. The results of our analysis revealed notable insights. For RRD diagnosis, BCNN $(VGG-16)^2$ emerged as the top-performing model, achieving an accuracy of 99.87% and an outstanding AUC of 99.99%. Inception-V3 also exhibited strong performance with an accuracy of 96.79% and an AUC of 99.76%. Other models, including VGG16, Inception-ResNet-V2, BCNN $(Inception-V3)^2$, and a custom CNN architecture, demonstrated competitive results, highlighting

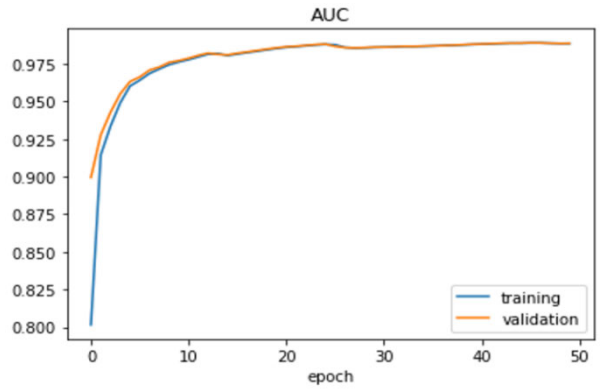



**FIGURE 16.** Training and validation curves for AUC metric: BCNN (Inception − V3)$^2$ model on Group A, Group B, and Group.

their potential in RRD diagnosis. Shifting to VA prediction, BCNN $(VGG-16)^2$ continued to excel, delivering an accuracy of 98.06% and an impressive AUC of 98.58%. Inception-V3 displayed robust performance with an accuracy of 90.76% and an AUC of 98.08%, establishing its effectiveness in VA prediction. Inception-ResNet-V2 and the custom CNN model also showed promise in this task. In a comprehensive evaluation spanning both RRD diagnosis and VA prediction, BCNN $(VGG-16)^2$ remained a standout model, with an overall accuracy of 97.93% and an AUC of 98.57%. Inception-V3 maintained strong performance with an overall accuracy of 87.85% and an AUC of 97.84%. Other models consistently delivered competitive results, underscoring their versatility in addressing both tasks. These findings illuminate the capabilities of CNN models in diagnosing RRD and predicting VA, with BCNN $(VGG-16)^2$ standing out as a high-performing and reliable choice across both domains.

These diverse CNN model results were attained through a meticulous process involving training, validation, and testing on both a public dataset and a carefully curated local database. The local database was carefully prepared and labeled by physicians based on post-operative patient follow-up over several months. Specifically, the A, B, and C groups were established by ophthalmologists according to post-operative visual acuity levels, with specific thresholds for each group. Group A includes individuals with visual acuity ranging from 2 to 1 LogMAR. Group B comprises individuals with visual acuity ranging from 0.9 to 0.4 LogMAR. Group C consists of individuals with visual acuity ranging from 0.3 to 0 LogMAR. Furthermore, both the local and public databases underwent the same preprocessing process. This consistency in data preprocessing ensures data comparability and consistent construction of the DL model. Regardless of potential differences in the populations of the two databases, the primary goal of this approach is to develop a robust model. In other words, the emphasis is on creating a model capable of generalizing and providing accurate predictions regardless of the original dataset.

The outstanding performance achieved by the various CNN architectures employed in our methodology results from the combination of several crucial factors. This approach encompasses various elements, including the adoption of cutting-edge CNN models alongside our meticulously tailored custom CNN model, precisely configured to align with the specific task requirements. Additionally, the preprocessing steps for image enhancement are thoughtfully selected, ensuring that the input data provided to the models is not only of optimal quality but also of an appropriate quantity. Lastly, the utilization of a diverse dataset, encompassing samples from a variety of sources and exhibiting a broad spectrum of characteristics, plays a pivotal role in augmenting the variability within the training data. This diversity strengthens the model's ability to effectively handle diverse situations, thereby solidifying its exceptional performance.

There is a certain degree of redundancy in the evaluation metrics among the different CNN models deployed for two distinct tasks: diagnosing RRD and predicting postoperative VA. This redundancy can be attributed to the dataset's meticulous structuring and equilibrium, resulting in consistent model performance across a range of metrics. The balanced nature of the dataset, where class samples are comparably distributed, often leads to consistent scores in accuracy, precision, recall, specificity, sensitivity, and AUC across various models. Moreover, the chosen models, such as VGG16, Inception-V3, and BCNN, are established architectures recognized for their strong performance in diverse tasks. Leveraging their inherent design and capabilities contributes to the uniformity of outcomes across distinct evaluation metrics, particularly when applied to a well-processed and balanced dataset.

Several DL-based methods have been proposed in the literature for diagnosing RD. However, direct comparison with these models on the same database is challenging due to the use of different image modalities in these studies. Nevertheless, it is evident that the proposed method has achieved commendable performance. Table 7 presents a comparison between the proposed method and selected previous methods from the literature, showcasing the effectiveness of the proposed approach.

Despite the notable findings and robustness of the present research, it is essential to address the limitations associated with dataset size and the potential of combining DL with Hand-Crafted ML methods. The limited size of the database is a prevalent challenge in OCT imaging research due to constraints in image availability and data collection complexities. Additionally, it is crucial to acknowledge that DL methodologies may not encompass all pertinent image features. This emphasizes the continual requirement for integrating Hand-Crafted ML methods to complement and augment the performance of deep learning approaches. Moreover, the application of the proposed method should be used judiciously, considering the specific clinical context. It's important to note that RD is a complex condition consisting of various subtypes, each with its unique characteristics and

**TABLE 7. Comparison of the proposed method with previous studies.**

| Study | DL model | Dataset | Accuracy % |
|---|---|---|---|
| Hideharu et al. [10] | Custom CNN architecture | 831 UWF images | Sensitivity: 97.6% Specificity: 96.5% AUC: 98.8% |
| Zhongwen et al. [11] | Inception-ResNet-V2 architecture | 11087 UWF images | Sensitivity: 96.1% Specificity: 99.6% AUC: 98.9% |
| Chenxi Zhang et al. [12] | seResNext50 architecture | 911 Optos images | Accuracy: 79.8%. Sensitivity: 87.5% Specificity: 100% AUC: 100% |
| Sonal Yadav et al. [13] | AlexNet, InceptionV3, GoogleNet, VGG19, DenseNet, and ResNet50 architectures | 1227 color fundus images | Accuracy: 99.50% Sensitivity: 99.00% Specificity: 99.99% Precision: 99.99% F1 score: 99.49% |
| Proposed method (RRD diagnosis) | BCNN$(\text{VGG}-16)^2$ | 953 OCT images | Accuracy: 99.87% Precision: 99.87% Recall: 99.87% AUC: 99.99% |
| Proposed method (VA prediction) | BCNN$(\text{VGG}-16)^2$ | 467 OCT images | Accuracy: 98.06% Precision: 98.58% Recall: 98.58% AUC: 98.58% |

clinical implications. However, it's crucial to underscore that the proposed method is primarily tailored for the accurate diagnosis of RRD. Consequently, there exists a potential risk of misclassification when dealing with other forms of RD. Thus, misidentifying non-RRD abnormal cases as RRD could result in significant clinical consequences, potentially leading to misguided treatment decisions.

## VI. CONCLUSION

This study makes a significant contribution to the development of a CAD system using DL techniques for the automatic RRD diagnosis and the prediction of postoperative VA using OCT images. By utilizing various preprocessing steps and implementing different CNN models, the proposed CAD system demonstrates high accuracy in diagnosing RRD and predicting VA. The experimental results validate the effectiveness of the system, achieving a diagnosis accuracy of 99.87% for RRD and a VA prediction accuracy of 98.06% using the BCNN $(\text{VGG}-16)^2$ model.

The findings of this study underscore the promising prospects of DL-based approaches in improving patient care and treatment decisions for RRD, including postoperative VA prediction. Future research should focus on exploring alternative DL models to enhance performance metrics and investigating hybrid approaches that integrate DL with Hand-Crafted ML methods. Moreover, due to the scarcity of limited labeled data and the absence of publicly shared datasets, coupled with the paramount importance of generalizability in any medical study, we intend to conduct further validation of the proposed model using data sourced from various independent medical centers. This approach will enable us to assess the robustness of our model and affirm its utility across a range of clinical contexts.

## REFERENCES


[1] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya, "Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022)," *Comput. Methods Programs Biomed.*, vol. 226, Nov. 2022, Art. no. 107161.

[2] K. K. Patro, J. P. Allam, B. C. Neelapu, R. Tadeusiewicz, U. R. Acharya, M. Hammad, O. Yildirim, and P. Pławiak, "Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images," *Inf. Sci.*, vol. 640, Sep. 2023, Art. no. 119005.

[3] U. Raghavendra, A. Gudigar, A. Paul, T. S. Goutham, M. A. Inamdar, A. Hegde, A. Devi, C. P. Ooi, R. C. Deo, P. D. Barua, F. Molinari, E. J. Ciaccio, and U. R. Acharya, "Brain tumor detection and screening using artificial intelligence techniques: Current trends and future perspectives," *Comput. Biol. Med.*, vol. 163, Sep. 2023, Art. no. 107063.

[4] F. Jerbi, N. Aboudi, and N. Khlifa, "Automatic classification of ultrasound thyroids images using vision transformers and generative adversarial networks," *Sci. Afr.*, vol. 20, Jul. 2023, Art. no. e01679.

[5] C. Cao, "Deep learning and its applications in biomedicine," *Genomics, Proteomics Bioinform.*, vol. 16, no. 1, pp. 17–32 2017.

[6] S. Yadav, N. K. Roy, N. Sharma, and R. Murugan, "Classification of retinal detachment using deep learning through retinal fundus images," in *Proc. IEEE India Council Int. Subsections Conf. (INDISCON)*, Jul. 2022, pp. 1–6.

[7] Z. N. Sultan, E. I. Agorogiannis, D. Iannetta, D. Steel, and T. Sandinha, "Rhegmatogenous retinal detachment: A review of current practice in diagnosis and management," *BMJ Open Ophthalmol.*, vol. 5, no. 1, Oct. 2020, Art. no. e000474.

[8] M. H. Qureshi and D. H. W. Steel, "Retinal detachment following cataract phacoemulsification—A review of the literature," *Eye*, vol. 34, no. 4, pp. 616–631, Apr. 2020.

[9] E. E. Christou, P. Stavrakas, G. Batsos, E. Christodoulou, and M. Stefaniotou, "Association of OCT–A characteristics with postoperative visual acuity after rhegmatogenous retinal detachment surgery: A review of the literature," *Int. Ophthalmol.*, vol. 41, no. 6, pp. 2283–2292, Jun. 2021.

[10] H. Ohsugi, H. Tabuchi, H. Enno, and N. Ishitobi, "Accuracy of deep learning, a machine-learning technology, using ultra–wide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment," *Sci. Rep.*, vol. 7, no. 1, p. 9425, Aug. 2017.

[11] Z. Li, C. Guo, D. Nie, D. Lin, Y. Zhu, C. Chen, X. Wu, F. Xu, C. Jin, X. Zhang, H. Xiao, K. Zhang, L. Zhao, P. Yan, W. Lai, J. Li, W. Feng, Y. Li, D. S. Wei Ting, and H. Lin, "Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images," *Commun. Biol.*, vol. 3, no. 1, p. 15, Jan. 2020.

[12] C. Zhang, F. He, B. Li, H. Wang, X. He, X. Li, W. Yu, and Y. Chen, "Development of a deep-learning system for detection of lattice degeneration, retinal breaks, and retinal detachment in tessellated eyes using ultra-wide-field fundus images: A pilot study," *Graefe's Arch. Clin. Experim. Ophthalmol.*, vol. 259, no. 8, pp. 2225–2234, Aug. 2021.

[13] S. Yadav, S. Das, R. Murugan, S. Dutta Roy, M. Agrawal, T. Goel, and A. Dutta, "Performance analysis of deep neural networks through transfer learning in retinal detachment diagnosis using fundus images," *Sādhanā*, vol. 47, no. 2, p. 49, Jun. 2022.

[14] F. Li, H. Chen, Z. Liu, X.-D. Zhang, M.-S. Jiang, Z.-Z. Wu, and K.-Q. Zhou, "Deep learning-based automated detection of retinal diseases using optical coherence tomography images," *Biomed. Opt. Exp.*, vol. 10, no. 12, pp. 6204–6226, 2019.

[15] A. Choudhary, S. Ahlawat, S. Urooj, N. Pathak, A. Lay-Ekuakille, and N. Sharma, "A deep learning-based framework for retinal disease classification," *Healthcare*, vol. 11, no. 2, p. 212, Jan. 2023.

[16] M. Paul, *Retinal OCT Images (Optical Coherence Tomography)*. Kaggle Dataset, 2018.

[17] K. U. Ahamed, M. Islam, A. Uddin, A. Akhter, B. K. Paul, M. A. Yousuf, S. Uddin, J. M. W. Quinn, and M. A. Moni, "A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images," *Comput. Biol. Med.*, vol. 139, Dec. 2021, Art. no. 105014.

[18] T. Kaur and T. K. Gandhi, "Automated brain image classification based on VGG-16 and transfer learning," in *Proc. Int. Conf. Inf. Technol. (ICIT)*, Dec. 2019, pp. 94–98.
[19] V. Choudhary, P. Guha, K. Tripathi, and S. Mishra, "Automatic detection of cowpea leaves using image processing and Inception-V3 model of deep learning," in *Proc. 6th Int. Conf. Signal Process., Comput. Control (ISPCC)*, Oct. 2021, pp. 314–318.
[20] A. Demir and F. Yilmaz, "Inception-ResNet-v2 with LeakyReLU and averagepooling for more reliable and accurate classification of chest X-ray images," in *Proc. Med. Technol. Congr. (TIPTEKNO)*, Nov. 2020, pp. 1–4.
[21] N. Aboudi, H. Khachnaoui, O. Moussa, and N. Khlifa, "Bilinear pooling for thyroid nodule classification in ultrasound imaging," *Arabian J. Sci. Eng.*, vol. 48, no. 8, pp. 10563–10573, Aug. 2023.
[22] R. Mastouri, N. Khlifa, H. Neji, and S. Hantous-Zannad, "A bilinear convolutional neural network for lung nodules classification on CT images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 16, no. 1, pp. 91–101, Jan. 2021.
[23] F. Rousselle, C. Knaus, and M. Zwicker, "Adaptive rendering with non-local means filtering," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 1–11, Nov. 2012.

**SINDA HOSNI** received the B.S. and M.S. degrees in intelligent systems and IoT from the Faculty of Sciences of Tunis, University of Tunis El Manar, in 2019 and 2022, respectively.

**HAJER KHACHNAOUI** received the B.S. and M.S. degrees in medical image processing from the Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, in 2016 and 2018, respectively, where she is currently pursuing the Ph.D. degree with the BTM Laboratory, Department of Medical Image Processing. Her research interests include medical image processing and artificial intelligence.

**HSOUNA MEHDI ZGOLLI** received the dual Diploma degree in vitreo-retinal surgery and optical coherence tomography in ophthalmology and in digital pedagogy and image processing from the University of Bordeaux. He is an Associate Professor of ophthalmology with the National Institute of Ophthalmology of Tunis, the Head of the Department of Vitreoretinal Surgery, and a member of the research laboratory with the Institute of Ophthalmology of Tunis. His current research interests include tomography and AI in vitreoretinal surgeries and medical pathologies, and has published more works in above interests. He was a Winner of the Besrour Doctoral Thesis Prize in Medicine for the work entitled: Interests of Spectral Domain OCT in the Surgery of Rhegmatogenous Retinal Detachment; and the Ophthalmology Research Prize for the work "Spectral Domain Optical Coherence Tomography in Surgical Ophthalmology," in 2016.

**SONYA MABROUK,** photograph and biography not available at the time of publication.

**DÉSIRÉ SIDIBÉ,** photograph and biography not available at the time of publication.

**HEDI TABIA** received the M.S. degree in computer science from the INSA of Rouen Public School of Engineers, France, in 2008, and the Ph.D. degree in computer science from the University of Lille, in 2011. From October 2011 to August 2012, he was a Postdoctoral Research Associate with the IEF Laboratory, University of Paris-Sud. From 2012 to 2019, he was an Associate Professor with ENSEA. Since September 2019, he has been a Professor with Université Paris-Saclay.

**NAWRES KHLIFA** received the Engineer and Ph.D. degrees from the National School of Engineers of Tunis (ENIT). She is currently a Professor with the Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar. She coordinates the TIMEd Team: Medical Image Processing of the BTM Laboratory. Her research work focuses on artificial intelligence and CAD design in medical imaging, emotion recognition, and gaze tracking.