

Received 24 January 2024, accepted 15 February 2024, date of publication 20 February 2024, date of current version 27 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3367977

RESEARCH ARTICLE

A Detectability Analysis of Retinitis Pigmetosa Using Novel SE-ResNet Based Deep Learning Model and Color Fundus Images

RUBINA RASHID¹, WAQAR ASLAM², ARIF MEHMOOD²,
DEBORA LIBERTAD RAMÍREZ VARGAS^{3,4,5},
ISABEL DE LA TORRE DIEZ⁶,
AND IMRAN ASHRAF⁷

¹Department of Computer Science and IT, Government Post Graduate College for Women, Bahawalnagar 62300, Pakistan

²Department of Information Security, The Islamia University of Bahawalpur, Punjab 63100, Pakistan

³Universidad Europea del Atlántico, 39011 Santander, Spain

⁴Universidad Internacional Iberoamericana de Campeche, Campeche 24560, Mexico

⁵Universidad Internacional Iberoamericana de Arecibo, Arecibo, PR 00613, USA

⁶Department of Signal Theory, Communications and Telematics Engineering, University of Valladolid, 47011 Valladolid, Spain

⁷Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, South Korea

Corresponding authors: Waqar Aslam (waqar.aslam@iub.edu.pk) and Imran Ashraf (ashrafimran@live.com)

This work was supported by the European University of Atlantic.

ABSTRACT Retinitis pigmentosa (RP) is a group of genetic retinal disorders characterized by progressive vision loss, culminating in blindness. Identifying pigment signs (PS) linked with RP is crucial for monitoring and possibly slowing the disease's degenerative course. However, the segmentation and detection of PS are challenging due to the difficulty of distinguishing between PS and blood vessels and the variability in size, shape, and color of PS. Recently, advances in deep learning techniques have shown impressive results in medical image analysis, especially in ophthalmology. This study presents an approach for classifying pigment marks in color fundus images of RP using a modified squeeze-and-excitation ResNet (SE-ResNet) architecture. This variant synergizes the efficiency of residual skip connections with the robust attention mechanism of the SE block to amplify feature representation. The SE-ResNet model was fine-tuned to determine the optimal layer configuration that balances performance metrics and computational costs. We trained the proposed model on the RIPS dataset, which comprises images from patients diagnosed at various RP stages. Experimental results confirm the efficacy of the proposed model in classifying different types of pigment signs associated with RP. The model yielded performance metrics, such as accuracy, sensitivity, specificity, and f-measure of 99.16%, 97.70%, 96.93%, 90.47%, 99.37%, 97.80%, 97.44%, and 90.60% on the testing set, based on GT1 & GT2 respectively. Given its performance, this model is an excellent candidate for integration into computer-aided diagnostic systems for RP, aiming to enhance patient care and vision-related healthcare services.

INDEX TERMS Retinitis pigmentosa, pigment signs, retinal diseases, ResNet, deep learning, data augmentation.

I. INTRODUCTION

Retina is ranked as one of the most energetically active tissues in the human body, and various illnesses can lead

The associate editor coordinating the review of this manuscript and approving it for publication was Ze Ji¹.

to alterations in its structure. These alterations can be detected to aid in diagnosis. Fundus images and optical coherence tomography (OCT) are valuable methods for examining eye conditions like retinitis pigmentosa (RP), diabetic retinopathy, macular degeneration, glaucoma, macular edema, and macular dystrophy [1]. RP is a prevalent group

of hereditarily retinal disorders. It occurs due to gene abnormalities that result in the degeneration of photoreceptor cells. RP typically emerges during childhood, and unfortunately, there is currently no cure available to halt the disease's progression [2]. The photoreceptor cells constitute a network of interconnected neurons that comprise the light-sensitive tissue present in the inner area of the eye. These tissues and cells play a critical part in the initial stages of visual processing. Retinopathy and the resulting impairment of retinal function are prevalent causes of blindness. The initial indicators of RP are usually loss of sight at night and loss of vision in the mid-peripheral areas, which can eventually lead to full loss of sight [3]. In the initial steps, the color fundus images may exhibit varying characteristics, ranging from mild deterioration to more prevalent pigment epithelium dystrophy. PSs can be observed in the middle peripheral regions initially and later extend to the next pole of the retina [4].

RP can be visualized in retinal images as pigmented areas located on the posterior part of the retina, which may enlarge and migrate over time along with accompanying symptoms. Both OCT and color fundus imaging have been instrumental in analyzing the state of RP [5]. While visual analysis by doctors is often relied upon for disease diagnosis, it may not always yield optimal results due to factors such as lack of experience, fatigue, variations in shape and texture, and poor image quality [6]. Already, the healthcare industry is benefiting from artificial intelligence-based algorithms and semantic segmentation. Consequently, deep learning approaches have emerged as a valuable technology for disease detection and examination [7]. Based on the information available in the related work, there is a significant scarcity of research regarding the automated detection of presumed diagnoses solely through the analysis of color fundus images. In contrast, the diagnosis performed using a fundus camera offers several advantages, such as enhanced convenience, reduced invasiveness, ease of repetition, and quicker results [8].

Various advanced models of CNN are widely used in medical image processing or industrial images such as complementary adversarial network-driven SDD (CASDD) comprises segmentation and an inventive complementary discriminator mechanism for encoding and decoding, featuring a specially crafted loss measurement to capture defect boundaries and enhance feature representation [9]. Another advanced model used a transfer approach relying on Clinical Prior Experience and Sample Analysis factors that were further confirmed through the probability distribution of sample images. Subsequently, a fusion attention block structure delivers an advanced non-uniform sparse representation of images [10]. Furthermore, a comprehensive depth domain adaptive network known as DDANet incorporates integration gradient CAM and attention guidance based on priori experience that enhances grading performance interpretability by incorporating the high magnification innovative to alleviate the overfitting problem [11]. A novel approach combining the

ViT and AMC blocks based on Hoeffding's inequality with adaptive model fusion and multiobjective optimization. This approach resolved the issue of concurrent optimal feature representation in ViT and AMC blocks. Finally, an adaptive model fusion technique incorporates the metrics block and the fusion block, aiming to enhance differences between feature representations and mitigate redundancy [12]. Another Swin-Transformer algorithm devised an auxiliary diagnostic, incorporating Focal Loss during training. It assessed the diagnostic precision of the Swin-Transformer in comparison to assessments by pathologists [13]. Furthermore, another proposed adaptive model fusion method comprises multi-objective optimization, adaptable feature representation metrics, and flexible feature fusion. This approach markedly enhances the fusion capabilities of the model [14]. The motivation behind the proposed model is the following consideration

- Retinitis Pigmentosa is a degenerative eye disease that affects the retina, leading to vision loss. The motivation may be to contribute to the improvement of medical diagnosis and early detection of RP using advanced technology.
- Traditional methods of diagnosis may be challenging due to the complexity and diversity of retinal abnormalities associated with the disease.
- Deep learning models have shown significant success in image classification tasks, especially in medical imaging. The motivation could be to leverage the capabilities of such models to enhance the accuracy and reliability of RP detection in color fundus images.
- The goal may be to create an automated screening tool that can assist healthcare professionals in quickly and accurately identifying signs of RP.
- Successful implementation of an accurate deep learning model for RP detection has the potential to positively impact clinical practices, leading to earlier diagnoses, improved patient outcomes, and more efficient use of healthcare resources.

This study presents an approach for automatic and accurate classification of RP disease using specialized neural network architecture, the squeeze-and-excitation ResNet (SE-ResNet) block [15]. The SE-ResNet block is an extension of the ResNet [16] architecture, augmented with a powerful squeeze and excitation mechanism to enhance the learning that performs segmentation and analysis of PS associated with degeneration in RP. The network efficiently identifies PS on the retina and precisely fragments these pigments, enabling examination and determination of the disease's growth rate. Handling with the small number of pixels is a challenging task such as retinal pigment signs are very small in RP. The proposed method embraces a related deep feature concatenation approach using SE block with a residual learning model. By combining the SE block with the residual blocks, the network can effectively capture and recalibrate the channel-wise dependencies, allowing it to focus on the most informative features while suppressing noise. In the

squeezing step, the block reduces the spatial dimensions of the feature map to a solitary value for each channel, typically through global average pooling. This process aggregates the channel-wise information, creating a compact representation of the feature map. In the excitation step, the SE block uses one or multiple fully connected layers to model the channel dependencies. It generates a set of channel-wise scaling factors that represent the relevance position of each channel. In the end, the weight file is deployed and applied to get the feature map that highlights the important channels and restrains the less relevant ones. This study stands out from existing methods in three significant ways:

- The proposed model enhances the expressive power of effective learning of hierarchical features and residual learning. This learning architecture helps address the vanishing gradient problem and ensures accurate PS detection even with limited pixels.
- It employs an SE block that introduces a channel-wise attention mechanism. It adaptively recalibrates the importance of different channels and suppresses less relevant ones that enhance the feature quality with deep feature extraction using concatenation.
- The proposed introduces only a small number of additional parameters, making the training process computationally efficient. The objective is to contribute to the field of RP detection and analysis using color fundus imaging. It enhances making it feasible for deployment in practical application and treatment of this disease.
- The proposed model emphasizes the color fundus imaging dataset for RP analysis, detecting abnormalities in the vessel, and distinguishing it from OCT-based RP detection algorithms.
- The proposed network achieves fine segmentation results even on inferior-quality color fundus images, aiding ophthalmologists and medical practitioners in detecting and analyzing RP growth in patients.

The remaining parts of the paper are organized as follows. In Section II, related work in the field of automated RP detection and deep learning-based frameworks is presented. Section III describes the architecture of our proposed architecture, including the SE-ResNet block and the dataset used for training. Section IV presents the experimental details, as well as, the discussion on results and Section V presents the conclusions of our work.

II. RELATED WORK

Retinal analysis encompasses retinal vessel division and optical disk identification, both of which have been particularly imperative tasks and have been extensively explored in the existing literature. However, there remains relatively scarce research dedicated to analyzing retinal RP. Most existing approaches tend to focus on analyses of optical coherence tomography (OCT), fundus photography, autofluorescence imaging, and electroretinography (ERG). The study [17]

has assessed the capacity of a deep convolutional neural network (CNN) that can distinguish between ultrawide field pseudocolor and autofluorescence color fundus images in the case of RP. The model encompasses both RP-affected and normal retina images that generated remarkable sensitivity and specificity for producing a high-quality heatmap of RP.

A U-Net hybrid approach composed of primary segmentation, coupled with a gliding window, is designed to refine and rectify the segmentation error made by the U-Net [18]. This model underwent training to classify entirely pixel values within the tested B-scans. For comparison of layer boundary lines, Bland-Altman, and correspondence analysis were performed to segment photoreceptor outer length. A CNN model was applied to differentiate between two visual acuities with a notable degree of sensitivity and specificity in [19]. It successfully recognized visual damage using a visual acuity threshold that demonstrates the capability to anticipate the correlation between retinal and visual function in persons tormented by RP.

An algorithm accompanied the segmentation of eight retinal sheets within RP data by releasing limitations on the thickness and smoothness of each layer [20]. A random forest classifier is trained using the RP dataset to assess edge probabilities. Then a graph search algorithm to identify the optimal set of nine surfaces that best align with the data [20]. An alternative combined method, called U-Net, evaluated the effectiveness of the group transformer for segmenting pigment signs in fundus images in [21]. This architecture incorporates multi-head self-attention hunks within the convolutional layers of both the narrowing and escalating tracks of the network, thus altering the classical U-Net framework. Moreover, within another hybrid model of CNN transformation, the group transformer U-Net was employed and explored to analyze the influence of different loss functions and parameter adjustments on the learning process. Interestingly, the results indicated significant enhancements, demonstrating a clear superiority of this hybrid architecture [22].

A dense U-Net segmentation model was specifically designed to accurately delineate the choroid layer and vessels within the RP dataset in [23]. So, this automated system to quantify choroidal factors in OCT images of RP-affected eyes was comprised of the overall choroidal area and thickness, stromal area, and luminal. The study [24] developed a two-stage model that was trained on independent x-ray photos from microperimetry. This approach used two techniques: manual and artificial alignment. For manual alignment, proprietary software was utilized to label six significant points situated at the vessel. The identification of cone boundaries was achieved through a deep learning segmentation technique. Subsequently, cone centers were determined using the extracted cone boundaries. To achieve this, a custom-built adaptive optics scanning light was used to capture non-confocal retinal fundus images. This imaging process was conducted on two patients diagnosed with confirmed RHO-associated RP [25].

TABLE 1. Literature overview of studies that addressed different issues using different methods and techniques to support the diagnosis of RP disease.

Reference	Methods & Techniques	Contributions
[17]-2019	Generated the heat map by distinguishing ultrawide-field pseudocolor and autofluorescence images using VGG-19.	Discrimination of the capacity of CNN for RP.
[18]-2021	A hybrid model comprises U-Net and SW in testing B-scans to classify all pixel values.	Evaluation of a hybrid model with different architectures for automatic segmentation.
[19]-2022	ResNet-152 shared identical exclusion criteria involving visually significant media centered on central macula opacities for a given dataset.	A method proposal to correlate the structure and function of cSLO parameters with visual acuity(VA) in RP.
[20]-2017	The random forest performs segmentation of retinal layers in RP by comforting limits on thickness.	Delineation of eight retinal layers in RP by loosening constraints related to the thickness of layers.
[21]-2022	A combination of bottleneck multihead self-attention blocks is employed for pigment identification in fundus images.	A transformer U-Net proposal (a hybrid method) designed for segmenting pigments.
[22]-2022	A hybrid Group Transformer U-Net investigated the effects on the learning process.	Tuning the hyperparameters for enhanced performance.
[23]-2022	The U-Net segmentation model used for segmenting of RP dataset.	An architecture proposal, ChoroidNet to enhance the segmenting process of the choroid layer and choroidal vessels for the RP dataset.
[24]-2023	Manually aligning overlaid infrared images obtained from a scanning laser ophthalmoscope.	Determining retinal multimodal labeled key points at the vessel of RP in fundus images.
[25]-2021	Custom-built adaptive optics (AO) scanning is used to acquire split detection.	A method proposal to derive cone boundaries to generate area-based cone size.
[26]-2020	CDSS employs ML and Chromatic Pupillometry to aid in diagnosing inherited diseases in pediatrics.	A technique proposal to assess the outer and the inner retina functions.
[27]-2021	A hypothesis regarding RP comprises four stages with two significant key factors.	A quantitative assessment by OCTA technology.
[28]-2023	AutoML platform designed for training to segregate between RVO, RP, and RD from normal fundus images.	An evaluation of the performance of AutoML in retinitis pigmentosa (RP).

A novel CD support system was introduced, utilizing machine learning and cross-validation to assist in the identification of inherited diseases in pediatric patients. This innovative technique combined a specialized device, a pupillometer with a custom-designed support system. Two individual SVMs were employed per eye to categorize data obtained from pupillometrics. The CDSS was specifically applied for the identification of RP in pediatric individuals [26]. OCT angiography is a modern technology that allows non-invasive visualization and quantitative evaluation of the vascular network within the retinal and choroidal regions. The progress of OCTA has furnished valuable insights into the fundamental mechanisms driving the pathogenesis of RP. A hypothesis was proposed with four phases and pinpointing two significant aspects: vascular dysfunction and microglia stimulation [27]. A multiclass model was generated using saliency maps to gauge the design of AutoML models using a publicly available image dataset with labels. Google Cloud AutoML vision platform underwent to identify and remove low-quality and mislabeled images from the dataset before training. Multi-binary models were created to distinguish between diseased and normal fundus images [28]. Different techniques and issues studied in the literature review are described in Table 1.

III. DESIGN OF PROPOSED MODEL

The detection and classification task serves as the foundation for various computer visualization tasks such as object detection, classification, recognition, segmentation, and more. CNNs conduct image classification by first detecting fundamental features like edges and curves at

lower levels and progressively combining them to form more intricate concepts through successive convolutional layers. The proposed network investigates a different aspect of squeeze and excitation [15] with residual learning that achieves precise pixel-wise classification by using a pixel classification layer to mark the detection pixels and improves the quality of representation by using interdependencies between the channels of its convolutional filters. The proposed method benefits from deep feature concatenation, and feature recalibration, and enables the net to import and combine extraordinary info from various layers. This design is highly effective in learning global information and suppresses less useful ones for accurately making saliency maps and classifications even in challenging scenarios. Notably, the proposed net directly gets the original color fundus images as input and outputs the detected retinal saliency mask with classification results. Figure 1 depicts the flowchart of the proposed model, illustrating the capturing of color fundus images, model training, validation, and testing phases.

A. SE-RESIDUAL LEARNING FOR PS SELECTION

PS region displays elevated values for both the pronounced darkness degree and well-defined edge strength. It can be deduced that the corresponding values surpass the average darkness degree and edge entropy values exhibited by typical color fundus regions. To effectively perform PS detection, neural networks need to be deep, often containing numerous convolutional layers. CNN employs the low-level pixel-wise features extracted in the initial stages to generate a potential PS, candidate of higher-level features, enabling it

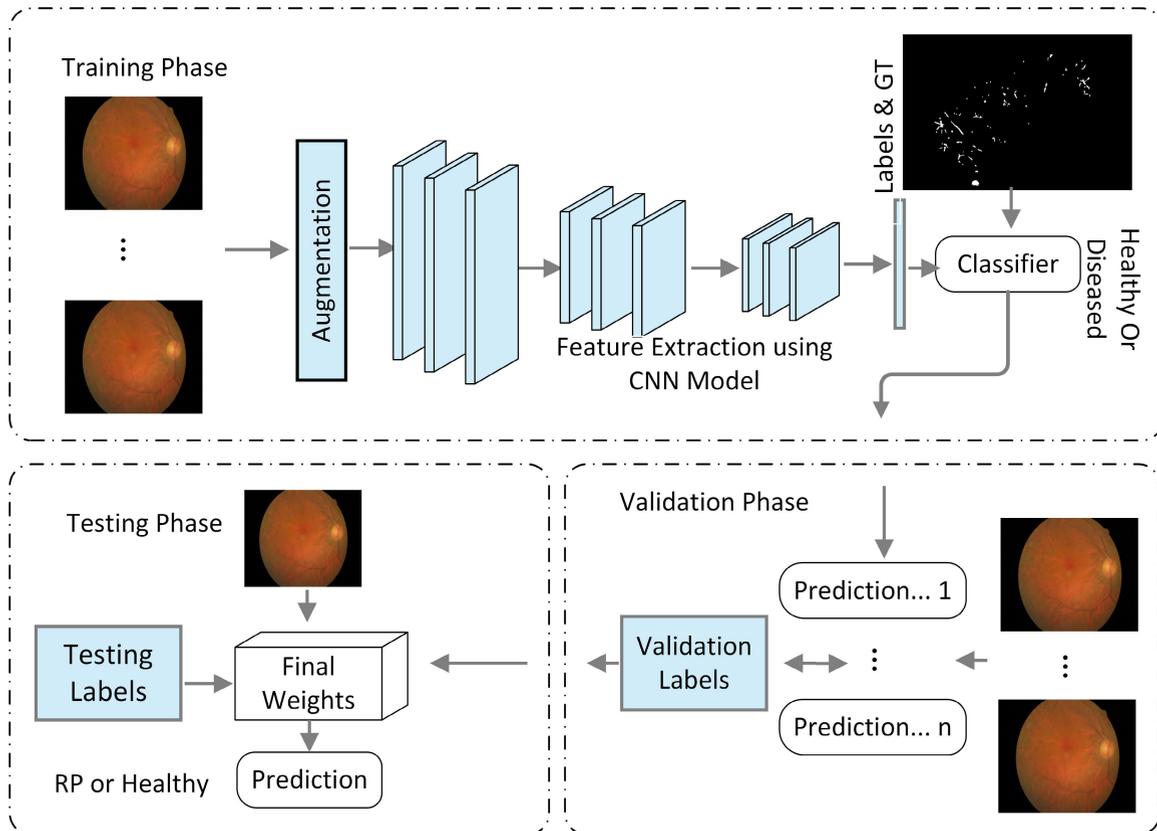


FIGURE 1. Visualization of the flowchart of the proposed model, illustrating the capturing of input color fundus images, training, validation, and testing phases.

to recognize and identify pixel signs in the color fundus diseased image. CNN's layers leverage their convolutional filters to extract spatial and temporal information from the color fundus images. However, as these layers process data, they may gradually lose spatial information at each operation. The widely recognized method for addressing this problem is through feature empowerment using ResNet. Figure 2 presents the schematic representation of the proposed model having residual and SE blocks.

This approach incorporates a skip connection based on summation that allows the input data to flow directly to the output information, bypassing one or more layers. That aids in alleviating the vanishing gradient problem and allows the network to learn residual features, resulting in significant performance improvement in PS recognition tasks. We begin by considering the implementation of SE blocks with residual networks which can directly employ transformation to the non-identity branch of a residual block. In this context, residual learning with SE block provides deep feature concatenation and ensures an efficient flow of information throughout the network for pixel-wise detection in RP.

The SE blocks aim to improve the pixel-wise detection process by introducing an adaptive content-aware mechanism that assigns weights to each channel based on its PS contextual relevance. It is a functional component that can

be constructed based on transformation (Ftr), which maps color fundus images to feature maps. In the context of this notation, Ftr is considered a convolutional operator. In addition to this, the SE blocks consider the importance of each channel when calculating the saliency maps of PS by summation through all channels. SE block performs two actions. First, modeling channel interdependencies explicitly develops the relationships between channels within layers to better understand the connections between different features. Second, feature recalibration selectively enhances useful PS features and suppresses less relevant ones, allowing the network to focus on the most informative aspects of the pixel-wise PS region. These two strategies with residual learning discussed above are developed in the design, which performs classification and saliency generation of RP.

The SE block can be integrated with residual block information to enhance the feature representation power of CNNs by adaptively recalibrating the importance of each channel in a feature map and improving the performance of the network. The SE block can be inserted after the convolutional layers of the residual block which makes it easier for the network to learn identity mapping for the optimization. The residual block processes the input through one or more convolutional layers, capturing various features. Through the utilization of four residual block operations

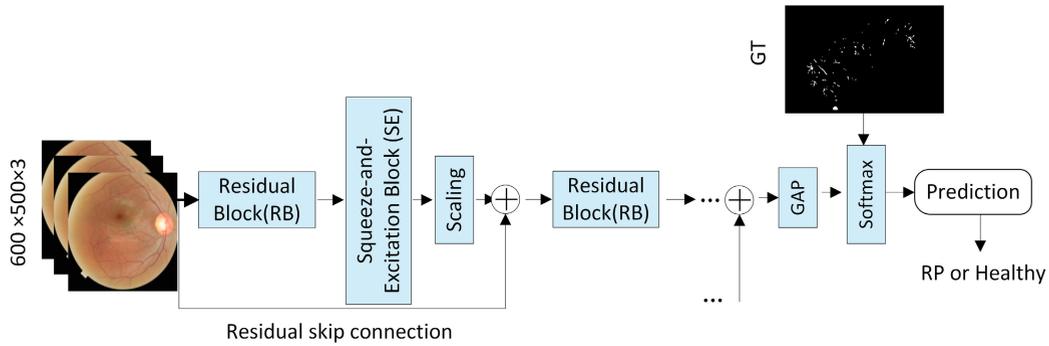


FIGURE 2. Schematic representation of the suggested model with residual and SE blocks.

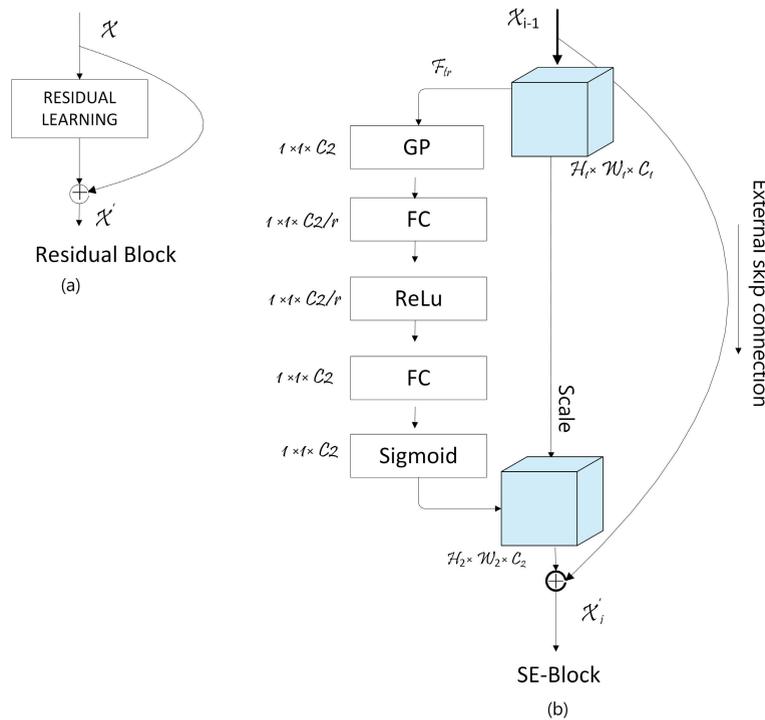


FIGURE 3. The scheme of the novel residual block (a), detail layers architecture with residual block and SE block with multiple layers.

comprising 34 convolutional layers and four SE blocks with fully connected layers.

Figure 3 shows the scheme of the original ResNet block. Figure 3a shows detailed layers architecture with residual block, and SE block with multiple layers is shown in Figure 3b. The SE block performs global pooling, typically global average pooling (GAP), on every single channel of the feature map. This averages the values of each channel across the entire spatial dimension. The pooled feature information is then passed through two fully connected (FC) layers, called the squeeze and excitation steps. The first FC layer decreases the channel dimensionality, capturing the global context of each channel. The second FC layer renovates the original channel dimensionality and produces channel-wise weights or importance scores for each channel. The channel-wise weights are then used to reweight the original feature maps

generated by the residual block. The feature maps are scaled by the corresponding channel-wise weights to highlight the important channels and restrain the less relevant ones. The re-weighted feature map information is then fed to the next layer of the neural network for further processing or to the next residual block for additional feature extraction. After that, the ultimate feature maps are generated at a size 25×32 when applied to $500 \times 600 \times 3$ representing a color image from the SE-ResNet block. Table 2 depicts the SE-ResNet with 34 depth layers. Layer-specific settings of the proposed model are also listed in the table.

B. DATA ACQUISITION

This study focuses on detecting retinal pigments using color fundus images to aid medical practitioners in the early diagnosis of the rare disease RP. The experiments

TABLE 2. SE-ResNet with 34 depth layers. Layer exact settings of the proposed model are listed inside the table.

Output Size	SE-ResNet Model									
256×256	Conv, 11×11, 64, stride 1									
112×112	Max pool, 7×7, stride 1									
	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr> <td><i>conv</i></td> <td>1 × 1</td> <td>128</td> </tr> <tr> <td><i>conv</i></td> <td>3 × 3</td> <td>128</td> </tr> <tr> <td><i>FC</i></td> <td>16</td> <td>128</td> </tr> </table> × 3	<i>conv</i>	1 × 1	128	<i>conv</i>	3 × 3	128	<i>FC</i>	16	128
<i>conv</i>	1 × 1	128								
<i>conv</i>	3 × 3	128								
<i>FC</i>	16	128								
56×56	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr> <td><i>conv</i></td> <td>1 × 1</td> <td>256</td> </tr> <tr> <td><i>conv</i></td> <td>3 × 3</td> <td>256</td> </tr> <tr> <td><i>FC</i></td> <td>32</td> <td>256</td> </tr> </table> × 4	<i>conv</i>	1 × 1	256	<i>conv</i>	3 × 3	256	<i>FC</i>	32	256
	<i>conv</i>	1 × 1	256							
	<i>conv</i>	3 × 3	256							
<i>FC</i>	32	256								
28×28	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr> <td><i>conv</i></td> <td>1 × 1</td> <td>512</td> </tr> <tr> <td><i>conv</i></td> <td>3 × 3</td> <td>512</td> </tr> <tr> <td><i>FC</i></td> <td>64</td> <td>512</td> </tr> </table> × 6	<i>conv</i>	1 × 1	512	<i>conv</i>	3 × 3	512	<i>FC</i>	64	512
	<i>conv</i>	1 × 1	512							
	<i>conv</i>	3 × 3	512							
<i>FC</i>	64	512								
14×14	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr> <td><i>conv</i></td> <td>1 × 1</td> <td>1024</td> </tr> <tr> <td><i>conv</i></td> <td>3 × 3</td> <td>1024</td> </tr> <tr> <td><i>FC</i></td> <td>128</td> <td>1024</td> </tr> </table> × 3	<i>conv</i>	1 × 1	1024	<i>conv</i>	3 × 3	1024	<i>FC</i>	128	1024
	<i>conv</i>	1 × 1	1024							
	<i>conv</i>	3 × 3	1024							
<i>FC</i>	128	1024								
7×7	GAP									
	FC									
	Softmax									

were conducted using the RIPS dataset, the only publicly available real dataset taken by a Canon CR4-45NM Retinal camera. It comprises 120 color fundus photos with a pixel resolution of 1440 × 2160. In three different sessions, a total of five photos per eye were captured at different sections of the retinal color fundus images. These images were captured from four patients over the course of three sessions (3 sessions × 5 images). The intermission between two successive sessions was six months and the total period was one year. Among the 120 images, 90 depict cases of RP, while 30 were of the healthy ones. The dataset exhibits inconsistency in terms of color, balance, focus, sharpness, and contrast. Two manual 'Binary Masks' were created by two ophthalmologists for each color fundus image, marking foreground regions representing pigment signs [29].

C. DATA AUGMENTATION

CNN models possess formidable capabilities when trained on extensive datasets; the greater the dataset size, the higher the CNN's performance. However, in the case of RP, obtaining an abundant number of color fundus images can be challenging due to their scarcity and privacy concerns associated with acquiring them from hospitals. Regrettably, there is a limited availability of publicly accessible datasets for RP. Consequently, CNN models face difficulties, leading to overfitting when working with medical image datasets that lack sufficient samples for training. To solve the potential issues of overfitting or underfitting, we used the cross-validation technique to make a more robust model and

enhance the generalization of models. In the conducted experiments, specifically, the proposed net underwent training using three-fold sample images employing a 3-fold cross-validation approach to evaluate the efficacy of the proposed technique and was subsequently tested with the fourth fold, all four folds were encompassing different patients. The approach of four-fold with different patients was chosen because images acquired from the same patient in different sessions often exhibited considerable similarity. Including such images in either the training or validation fold could potentially lead to erroneous performance improvement.

Before applying the augmentation, a Gaussian blur filter was used to smoothen, and randomly change the sharpness of the images and help to reduce the high-frequency noise. The proposed net underwent training using three-fold images employing a 3-fold cross-validation approach and was subsequently tested with the fourth fold, all four folds encompassing different patients. As the three folds originally comprised only 90 images including healthy and diseased images of RP, that proved insufficient for effective training of the proposed model. Artificial images were generated through a data augmentation process similar to that used by [30, 31]. Specifically, the 90 images from the three folds were combined, resulting in 90 images that underwent horizontal flips, creating an additional 90 images. This process yielded a total of 180 images (original three folds images = 90, H-flip = 90). Subsequently, these 180 images underwent $X = 5, Y = -5$, conversion with a horizontal flip, resulting in 360 images. Following this, the 360 images were subjected to another round of $X = -5, Y = 5$, conversion with vertical flip, producing a total of 720 images. The 720 images from the previous stage underwent XY conversion at $X = 10, Y = 10$ with a horizontal flip, resulting in a total of 1440 images. In the final step, underwent images were at $X = 10, Y = 10$ conversion with vertical flip, resulting in a grand total of 2880. Finally, we get 2880 diseased images of RP and 960 healthy images.

TABLE 3. Provide a breakdown of the training dataset before and after the application of data augmentation.

Augmentation Process	Diseased RP	Healthy
Original images	90	30
Rotation at +5 to -5	180	60
V-flip	360	120
Cropping (bounding box)	720	240
Resizing	1440	480
H-flip	2880	960

Table 3 provides a breakdown of the training datasets before and after the application of data augmentation. The representative images are shown in Figure 4.

IV. RESULTS AND DISCUSSION

A. EXPERIMENTAL SETUP

The training process was developed using Python programming language which is a powerful machine learning platform. We used Python 2.7.0, tflearn, sklearn,

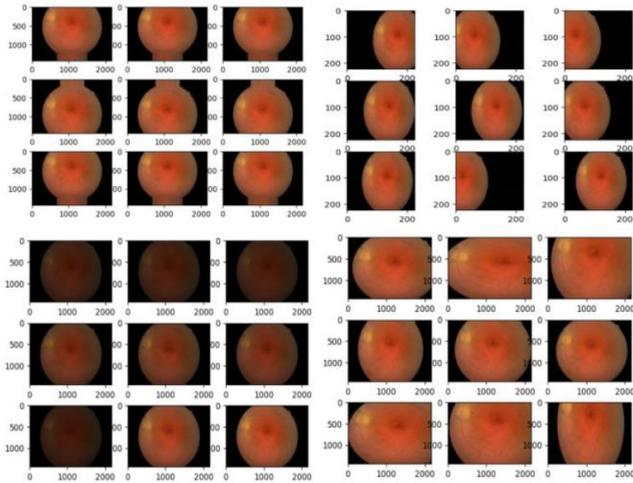


FIGURE 4. Augmented sample images.

tensorflow, numpy, matplotlib to customize our proposed model. Therefore, we prepared an NVIDIA Titan GPU (K80) system with 12GB memory for training the trained detector. The entire dataset consumed two hours for the training of the proposed model.

TABLE 4. A breakdown of training dataset after the application of data augmentation.

Augmentation Process	3-Fold	4th-Fold
Original images (Encompassing different patients)	90	30
Original images + H-flip	180	
Conversion at X= 5, Y= -5 (H-flip)	360	
Conversion at X= -5, Y= 5 (V-flip)	720	
Conversion at X= 10, Y= 10 (H-flip)	1,440	
Conversion at X= 10, Y= 10 (V-flip)	2,880	

B. MODEL TRAINING

The proposed model is designed using residual and SE block connectivity for the transfer of immediate information between layers. This structure facilitates the rapid convergence of the network with rich features for accurate detection of PS. The training process was conducted from scratch on augmented color fundus images excluding any weight sharing from external networks. Given the original dimensions of the image, which are 1440×2160 , the training network poses a challenge due to the constraints of GPU memory. As a solution to this issue and to facilitate training and testing of the model, the input color fundus images were altered in size to 500×600 dimensions to compare with it on GT1 & GT2. To tackle the problem of uneven class distribution, we explore different approaches, both within individual patches and across patches. The motivation for incorporating GT maps could be rooted in improving the model's performance by providing additional guidance during the training process. GT maps typically represent the true distribution of certain features or characteristics in the input data. By combining GT maps with Softmax,

the model may be better equipped to focus on relevant regions or features that are crucial for the specific task at hand. As for the ablation experiment, it is a common practice in research to conduct ablation studies to evaluate the contribution of specific components or techniques to the overall performance of a model. The ablation experiment used GT1 and GT2 maps demonstrates that the addition of GT maps significantly improves performance and for a comprehensive understanding of the findings. Specifically, we employ a cross-entropy loss function with input-balancing strategy.

Information about how long the proposed model was trained. The criteria for stopping training were convergence, performance metrics, or a fixed number of epochs. Details about the hyperparameters used during training, such as learning rate, batch size, and optimizer choice. The specific methodology for hyperparameter optimization systematically searches through a predefined grid of hyperparameter values. Randomized search with cross-validation is used to evaluate random sample hyperparameter combinations and select the best-performing set. The entire dataset was split into batches, and each batch covered every sample within it. Once all data samples from all batches completed this process, one epoch was concluded. An epoch represents the total number of times the entire datasets are processed. In this study, the training process involved 50 epochs and a batch size of 15 images in each epoch. Notably, the proposed model began to converge after 46 epochs and reached stability at 50 epochs. The ultimate features extracted from the last fully connected layer generated 128×1024 outputs which served as input for the softmax layer. The model derived decision weights based on accurate probabilities computed by the softmax layer. These weights were then utilized to make the detection decision. The cross-entropy loss function produces the established output results, representing the posterior probability of the target class. In order to preserve the advantages of the Adam optimizer in comparison to the conventional stochastic gradient descent, we opted for Adam as an optimizer, setting the starting learning rate to 0.001. The constant learning rate during the training with an epsilon value was 0.00001. Hyperparameter optimization is carried out using the GridSearchCV method. The detail of stopping hyperparameters is shown in Table 5.

TABLE 5. The details of hyperparameters that stop the training process.

Hyperparameter	Values
Optimizer	Adam optimizer
Learning rate	0.001
Epochs	50
Minibatch size	15
Epsilon Value	0.00001

C. EVALUATION METRICS

Various evaluation metrics are utilized to evaluate the performance of the proposed model. Sensitivity quantifies the ratio

of actual positive cases correctly identified by an analytical test. Specificity gauges the ratio of correctly identified actual negative cases by a diagnostic classifier. These metrics rely on true positive (TP), signifying the number of images correctly categorized as RP disease, true negative (TN), indicating the number of images accurately identified as normal color fundus images, false positive (FP), indicating the number of images incorrectly assigned as RP, and false negatives (FN), showing the number of images erroneously categorized as healthy color fundus. In detail: accuracy, sensitivity, specificity, and f-measure are calculated using

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

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

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

D. RESULTS

The training set, consisting of 2880 images with RP and 960 images without RP disease, was employed for model training. The dataset was arbitrarily divided into training and testing sets using an 80%-20% split, resulting in 2304 RP and 768 normal color fundus images used in training and validation. Additionally, a separate test dataset, comprising 576 RP and 192 normal color fundus images, was utilized to assess the trained model. The division ensured that the dataset's quality remained intact while achieving improved recognition results. In training, to validate the effectiveness of our model, a 5-fold cross-validation technique is employed on the 80% training dataset, which is usually pretty accurate. The 5-fold cross-validation technique decreases the size variance between the training set and the resampling subset and leads to a smaller bias. To assess the efficacy of our model, we employed the extensively augmented version of the RIPS dataset.

The results indicate a strong and consistent classification performance and showcase the model's ability to accurately identify even the rare conditions of RP (PS) in retinal color fundus images. Evaluation of dataset performance is depicted in the confusion matrix and ROC based on training and testing. The confusion matrix's statistics are used to describe the classification assessment of the proposed model on the RIPS dataset with accuracy 100%, 99.89%, 99.16%, and 99.37% based on training and testing with true values of GT1 & GT2. The accuracy of the model is found to be a confusion matrix that is particularly valuable when dealing with an imbalanced dataset or when different types of errors have varying levels of significance. The receiver operating characteristic (ROC) values indicate the model's performance in discriminating between the two classes and serve as a measure of the model's overall accuracy and effectiveness in retinal PS segmentation of RP.

Figure 5 (1st row) confusion matrix and Figure 5 (2nd row) ROC curves illustrate the proposed method performance based on the training and testing dataset, respectively, with reverence to GT1 and GT2. As a result, the confusion matrix and other evaluation metrics of the proposed model are presented. The values of ROC AUC curves were found to be 98.50%, 97.70%, and 98.00%, 97.80% based on training and testing with known GT1 and GT2. Although the ROC AUC for the test set is marginally lower than that of the training dataset, it remains above 97%. ROC curves plot the relationship between FPR and TPR. The linear blue line represents the line of equal error rate (EER), where the TPR is equal to the FPR that determines the balance between these two rates. In this work, we utilized performance metrics such as accuracy, sensitivity, specificity, and f-measure that can evaluate the performance of diagnostic tests or classifiers in RP disease. These metrics provide quantitative measures of the test to detect the RP disease and specificity to rule out individuals without the disease. It can help in assessing the reliability and clinical utility of a diagnostic tool or classifier and can aid in decision-making regarding diagnosis, treatment, and patient management. In the context of RP diagnosis, the AUC summarizes the overall performance of a discriminatory power; meaning that the test can effectively identify individuals with the disease.

Numerous experiments were conducted, involving the training of the proposed model with different learning parameters to comprehend the involvement of various features to the ultimate identification outcomes. In the initial investigation, we select a hybrid method that incorporates the entire set of features to evaluate the importance of individual features by analyzing them within the SE-ResNet network. We conducted a comparative evaluation of the model with five other models. Notably, from Table 6, random forest and AdaBoost.M1 [29] exhibit enhancement in accuracy (99.45%, 99.44%, 99.35%) and specificity (99.65%, 99.51%) on more than 250 trees component based training. Although 250 trees represent the best selection, it escalates computational expenses without proportionally advancing performance. On the other side, the proposed model determines the optimal number of SE-ResNet 34 layers required for achieving a balance between performance metric values and computational costs, encompassing execution time and memory. As far as our understanding extends, standard and normal validation [30] was the first approach regarding the automated identification of PS in color fundus images in the literature and yields significant outcomes (88.58%).

Table 6 reports the qualitative results according to the standard indexes of accuracy, specificity, sensitivity, and f-measure for the identification of PS using GT1 & GT2 as a basis on the same dataset. In accordance with the data presented in Table 6, the proposed model showcases higher levels of sensitivity, specificity, f-measure, and the 2nd highest accuracy compared to state-of-the-art methodologies, but the achieved accuracy by the proposed model is 0.29%

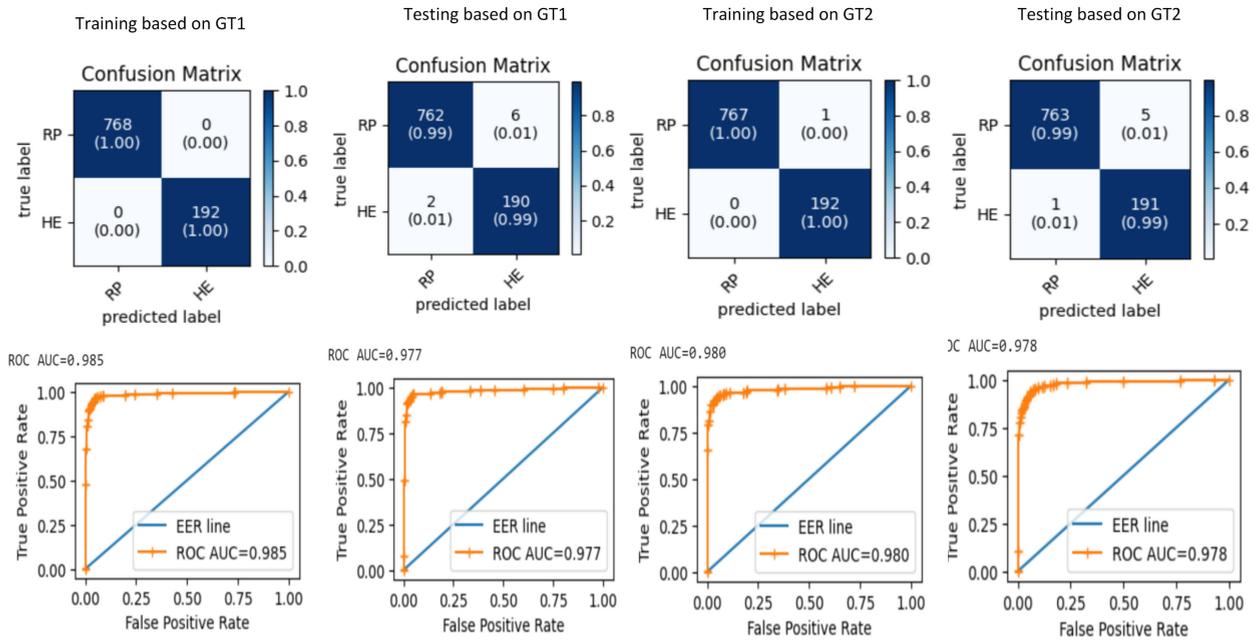


FIGURE 5. Overall dataset performance presents in confusion matrix and AUROC curves. Confusion matrix of the proposed model based on training and testing dataset (1st row). ROC AUC curves for the suggested model on training and testing dataset (2nd row).

TABLE 6. Effectiveness of the proposed model and other five state-of-the-art methodologies with different variations on RIPS datasets. Among the colors used, red signifies the top-performing outcome, followed by blue representing the second-best, and green indicating the third-best.

Dataset			RIPS – GT1				RIPS – GT2			
Year	Algorithm	Method	AC	SE	SP	FM	AC	SE	SP	FM
2018 [29]	Random Forest	Pixel-based	99.14	58.26	99.46	47.93	99.11	56.20	99.48	49.29
	AdaBoost.MI		99.01	64.29	99.30	46.76	98.99	61.79	99.33	48.30
	Random Forest	Component-based	99.45	72.85	99.65	63.80	99.44	69.47	99.65	64.69
	AdaBoost.MI		99.35	79.40	99.51	62.48	99.35	75.66	99.51	63.70
2018 [30]	Standard validation		97.90	74.43	98.44	59.04	97.83	69.64	98.46	58.42
	Normal validation		98.57	88.58	98.84	75.01	98.63	85.65	98.97	76.61
2020 [31]	RPS-Net	Deep-Features	99.15	80.54	99.60	61.54	99.51	78.09	99.62	62.62
2019 [32]	Deep CNN		96.00	72.00	97.00	62.00	96.00	72.00	97.00	61.00
2019 [33]	U-Net	48 × 48	99.00	55.70	99.40	50.60	99.00	55.70	99.40	50.60
		72 × 72	99.00	62.60	99.30	52.80	99.00	62.60	99.30	52.80
		96 × 96	99.20	55.20	99.60	55.10	99.20	55.20	99.60	55.10
Proposed	Channel-based Testing		99.16	97.70	96.93	90.47	99.37	97.80	97.44	90.60

lower than that of the preceding component based on random forest [29]. Furthermore, results indicate superior performance of the proposed model over the earlier methods RPS-Net [31], deep CNN [32], and U-Net [33], all applied on the same dataset. FN carries greater significance than FP, and the erroneous pixels classified as FN are assessed using the sensitivity. Based on the information, presented in Table 6, it is obvious that the proposed model demonstrates a high sensitivity, indicating its ability to minimize the occurrence of FN pixels. Training the proposed network involved utilizing three out of the four folds for training, while the remaining fold was reserved for testing. The results of testing data are shown in Table 6.

E. DISCUSSION

RP is a gradual retinal deteriorating condition marked by the gradual decline of photoreceptors and the subsequent loss

of these cells. During the intermediate phase of the disease, distinctive pigment signs become evident on color fundus examination. Fundus color photos are further categorized by a multitude of distortions and posing complexity for segmentation and classification analysis. In the case of RP, the classification of pigment signs becomes notably arduous due to the significant variability in color and shape that it can manifest. In later stages, color fundus examination shows extensive retinal degeneration. Moreover, RP exhibits genetic heterogeneity that complicates the diagnosis. Due to these challenges, the proposed method aims to automate the finding of RP by SE-ResNet learning-based techniques that identify PS in retinal color fundus images. The objective of this study is to determine if this approach can effectively tackle the thought-provoking task of identifying small numbers of pixels amidst a highly variable background. In most cases, PS manifests as compact regions containing a limited number of pixels, this trait has also been observed in other types

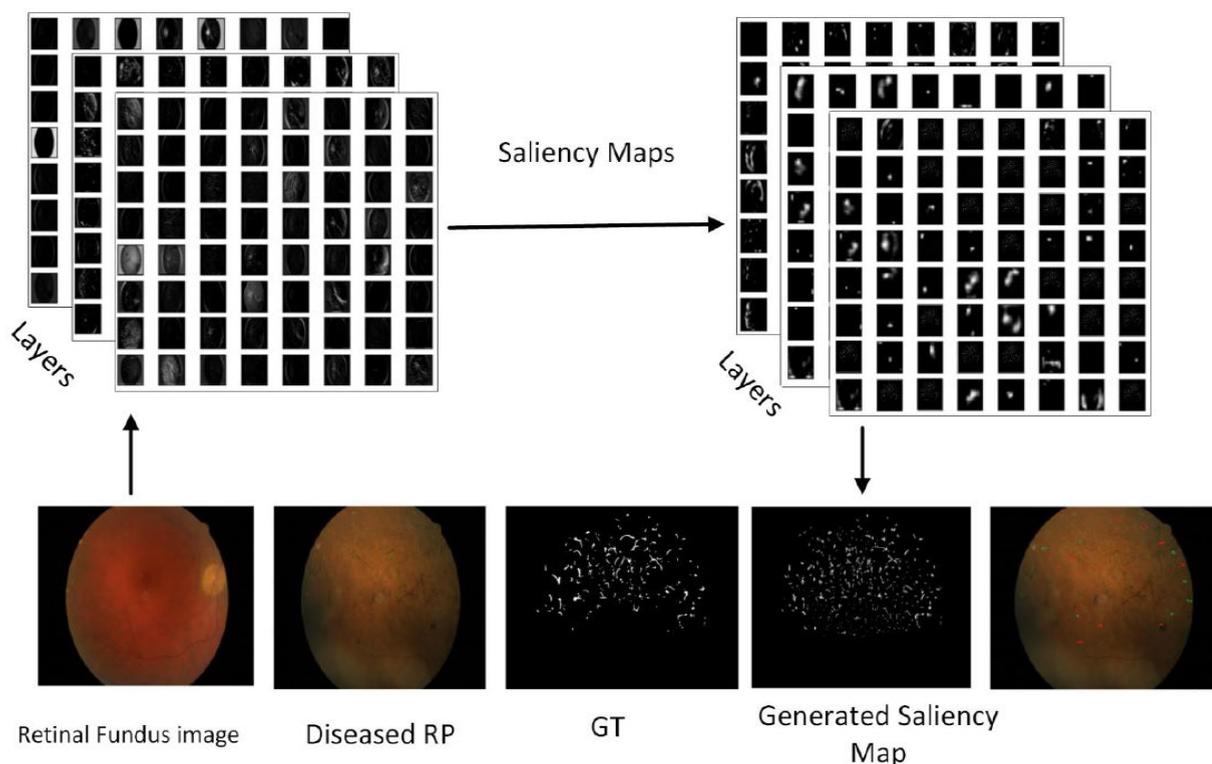


FIGURE 6. Visual outcomes and saliency maps of retinal PS at different layers with respect to GT, were achieved using the proposed model, where FP is indicated in green and FN is indicated in red.

of indicators, such as exudates in diabetic retinopathy. When dealing with such small regions as PS regarding the utilization of pixel-level validation for assessing that are accurately detected, border pixels might be inaccurately labeled as FP or FN.

Visualization techniques similar to Grad-CAM can be employed to gain insights into the predictive regions corresponding to true positive (TP) and true negative (TN) outcomes. By comparing these predictive regions with findings in medical literature, it becomes possible to assess the alignment between observed predictive regions and existing knowledge documented in the literature. This comparative analysis aims to validate and corroborate the identified predictive regions with information already established in the medical field. Another notable benefit of consolidating small regions based on their visual attributes is the inherent resilience of our approach to channel fluctuations, as numerous existing approaches in the literature require the implementation of multi-scale processing to achieve comparable robustness.

Figure 6 displays the visual outcomes of retinal pigment signs at the grey level achieved using the proposed model. It is also important to highlight that in cases concerning a healthy color fundus image, the f-measure and sensitivity metrics hold less importance due to the balanced relationship between TP and FN, resulting in comparable values for accuracy and specificity. Consequently, assessments were

performed separately for both healthy color fundus images and images exhibiting PS to ensure accurate evaluation. To differentiate between PS and normal fundus area in RGB color fundus images, we examined regional features or attributes that are derived solely from the spatial organizes of pixels in the whole image region. In the end, the grey-scale information can be extracted from a solo channel, potentially yielding up to three changed feature levels. Nonetheless, the blue color typically lacks valuable information for the classification of PS in RP. Furthermore, the grayscale values in images reflect the original pixel intensities, while the corrected grayscale values in the image offer more insightful information concerning the contrast between the region and its neighboring regions. Several grey-level features involve computation utilizing either the boundary or vicinity of the image regions. This approach contributes to the robustness of the method as evidenced by the achieved results. Notable, qualitative assessment reveals that even trivial pigment signs are effectively classified by the proposed method.

Identifying the regions that lead to FP and FN is a critical step in understanding model errors and improving their performance. Generating saliency maps can highlight regions of interest and regions influencing the decision. Ground truth annotations are overlaid on the images to compare the predicted regions with the actual disease locations. This step is crucial for identifying regions where the model made errors.

The limited number of patients and the potential lack of generalization to real-world scenarios is a valid one. Firstly, generalizability is a crucial aspect, but the proposed model trained on diverse datasets by getting the data augmentation techniques that perform well on unseen data beyond the training set in real-world scenarios. Secondly, intra-patient variability is considered by acquiring the images from multiple sessions and different times for each patient. To assess the clinical relevance and effectiveness of the ML model, it is crucial to compare its outputs with the assessments made by human experts or clinicians. This helps in understanding the model's performance in a real-world context and it aligns with the expertise of medical professionals. It may be a limitation of the study and other compared models are also used in similar scenarios.

F. LIMITATIONS OF STUDY

Analysis of RP using a SE-ResNet-based deep model and fundus images is a promising approach to developing an accurate detectability system, but still, it has some limitations which are described as.

- We utilized the variant form of CNN learning models that provide an accurate assessment, but other machine learning algorithms need to be investigated to analyze their performance for PS detection in comparison to the proposed model.
- In a clinical setting, a complete system may be necessary with the deployed model for real-time processing of fundus images.

V. CONCLUSION

This study introduces a deep CNN architecture tailored explicitly for the detection of retinitis pigmentosa, a retinal eye disease, using color fundus images. By incorporating residual learning deep features with SE blocks, the proposed model enhances its representational capacity. The SE blocks, in particular, enable dynamic channel-wise recalibration and address feature dependencies, thus improving the segmentation of retinal pigment signs, even those with a limited pixel presence in the pigment area. By reducing the number of residual blocks and incorporating SE blocks, we enhanced feature quality while optimizing memory usage and minimizing information loss. The proposed network adeptly estimates the saliency of retinal pigment signs through RP analysis. With its notable sensitivity and specificity in detecting eyes affected by RP, this model holds the potential to aid medical practitioners in the timely and detailed assessment of the disease's progression and severity.

REFERENCES

- [1] J. K. Oh, Y. Nuzbrokh, J. R. L. de Carvalho, J. Ryu, and S. H. Tsang, "Optical coherence tomography in the evaluation of retinitis pigmentosa," *Ophthalmic Genet.*, vol. 41, no. 5, pp. 413–419, Sep. 2020.
- [2] A. Fahim, "Retinitis pigmentosa: Recent advances and future directions in diagnosis and management," *Current Opinion Pediatrics*, vol. 30, no. 6, pp. 725–733, 2018.
- [3] M. F. Dias, K. Joo, J. A. Kemp, S. L. Fialho, A. da Silva Cunha, S. J. Woo, and Y. J. Kwon, "Molecular genetics and emerging therapies for retinitis pigmentosa: Basic research and clinical perspectives," *Prog. Retinal Eye Res.*, vol. 63, pp. 107–131, Mar. 2018.
- [4] S. S. Mondal, N. Mandal, K. K. Singh, A. Singh, and I. Izonin, "EDLDR: An ensemble deep learning technique for detection and classification of diabetic retinopathy," *Diagnostics*, vol. 13, no. 1, p. 124, Dec. 2022.
- [5] T.-C. Chen, W. S. Lim, V. Y. Wang, M.-L. Ko, S.-I. Chiu, Y.-S. Huang, F. Lai, C.-M. Yang, F.-R. Hu, J.-S.-R. Jang, and C.-H. Yang, "Artificial intelligence-assisted early detection of retinitis pigmentosa—The most common inherited retinal degeneration," *J. Digit. Imag.*, vol. 34, no. 4, pp. 948–958, Aug. 2021.
- [6] Y. Ye, J. Mao, L. Liu, S. Zhang, L. Shen, and M. Sun, "Automatic diagnosis of familial exudative vitreoretinopathy using a fusion neural network for wide-angle retinal images," *IEEE Access*, vol. 8, pp. 162–173, 2020.
- [7] Y. Patil, A. Shetty, Y. Kale, R. Patil, and S. Sharma, "Multiple ocular disease detection using novel ensemble models," *Multimedia Tools Appl.*, vol. 83, no. 4, pp. 11957–11975, Jan. 2024.
- [8] S. Muhammad Saiful Islam, M. Mahedi Hasan, and S. Abdullah, "Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images," 2018, *arXiv:1812.10595*.
- [9] S. Tian, P. Huang, H. Ma, J. Wang, X. Zhou, S. Zhang, J. Zhou, R. Huang, and Y. Li, "CASDD: Automatic surface defect detection using a complementary adversarial network," *IEEE Sensors J.*, vol. 22, no. 20, pp. 19583–19595, Oct. 2022.
- [10] P. Huang, X. Tan, X. Zhou, S. Liu, F. Mercaldo, and A. Santone, "FABNet: Fusion attention block and transfer learning for laryngeal cancer tumor grading in P63 IHC histopathology images," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, pp. 1696–1707, Apr. 2022.
- [11] P. Huang, X. Zhou, P. He, P. Feng, S. Tian, Y. Sun, F. Mercaldo, A. Santone, J. Qin, and H. Xiao, "Interpretable laryngeal tumor grading of histopathological images via depth domain adaptive network with integration gradient CAM and priori experience-guided attention," *Comput. Biol. Med.*, vol. 154, Mar. 2023, Art. no. 106447.
- [12] P. Huang, P. He, S. Tian, M. Ma, P. Feng, H. Xiao, F. Mercaldo, A. Santone, and J. Qin, "A ViT-AMC network with adaptive model fusion and multiobjective optimization for interpretable laryngeal tumor grading from histopathological images," *IEEE Trans. Med. Imag.*, vol. 42, no. 1, pp. 15–28, Jan. 2023.
- [13] Y. Wang, F. Luo, X. Yang, Q. Wang, Y. Sun, S. Tian, P. Feng, P. Huang, and H. Xiao, "The swin-transformer network based on focal loss is used to identify images of pathological subtypes of lung adenocarcinoma with high similarity and class imbalance," *J. Cancer Res. Clin. Oncol.*, vol. 149, no. 11, pp. 8581–8592, Sep. 2023.
- [14] H. Pan, H. Peng, Y. Xing, L. Jiayang, X. Hualiang, T. Sukun, and F. Peng, "Breast tumor grading network based on adaptive fusion and microscopic imaging," *Opto-Electron. Eng.*, vol. 50, no. 1, 2023, Art. no. 220158.
- [15] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [17] H. Masumoto, H. Tabuchi, S. Nakakura, H. Ohsugi, H. Enno, N. Ishitobi, E. Ohsugi, and Y. Mitamura, "Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultrawide-field images," *PeerJ*, vol. 7, p. e6900, May 2019.
- [18] Y.-Z. Wang, W. Wu, and D. G. Birch, "A hybrid model composed of two convolutional neural networks (CNNs) for automatic retinal layer segmentation of OCT images in retinitis pigmentosa (RP)," *Transl. Vis. Sci. Technol.*, vol. 10, no. 13, p. 9, Nov. 2021.
- [19] T. Y. A. Liu, C. Ling, L. Hahn, C. K. Jones, C. J. Boon, and M. S. Singh, "Prediction of visual impairment in retinitis pigmentosa using deep learning and multimodal fundus images," *Brit. J. Ophthalmol.*, vol. 107, no. 10, pp. 1484–1489, Oct. 2023.
- [20] A. Lang, A. Carass, A. K. Bittner, H. S. Ying, and J. L. Prince, "Improving graph-based oct segmentation for severe pathology in retinitis pigmentosa patients," *Proc. SPIE*, vol. 10137, pp. 426–433, Aug. 2017.

- [21] M. Sangiovanni, N. Brancati, M. Frucci, L. Di Perna, F. Simonelli, and D. Riccio, "Segmentation of pigment signs in fundus images with a hybrid approach: A case study," *Pattern Recognit. Image Anal.*, vol. 32, no. 2, pp. 312–321, Jun. 2022.
- [22] M. Sangiovanni, M. Frucci, D. Riccio, L. Di Perna, F. Simonelli, and N. Brancati, "Exploring a transformer approach for pigment signs segmentation in fundus images," in *Proc. Int. Conf. Image Anal. Process.* Cham, Switzerland: Springer, 2022, pp. 329–339.
- [23] T. T. Khaing, T. Okamoto, C. Ye, M. A. Mannan, G. Miura, H. Yokouchi, K. Nakano, P. Aimmancee, S. S. Makhanov, and H. Haneishi, "Automatic measurement of choroidal thickness and vasculature in optical coherence tomography images of eyes with retinitis pigmentosa," *Artif. Life Robot.*, vol. 27, no. 1, pp. 70–79, Feb. 2022.
- [24] S. H. Yassin, Y. Wang, W. R. Freeman, A. Heinke, E. Walker, T. Nguyen, D.-U. G. Bartsch, C. An, and S. Borooah, "Efficacy and accuracy of artificial intelligence to overlay multimodal images from different optical instruments in patients with retinitis pigmentosa," *Clin. Experim. Ophthalmol.*, vol. 51, pp. 446–452, 2023.
- [25] J. Giannini, J. Liu, T. Liu, N. Aguilera, L. A. Huryn, and J. Tam, "Adaptive optics retinal imaging with deep learning cone segmentation enables area-based analysis in rho-associated retinitis pigmentosa," *Investigative Ophthalmol. Vis. Sci.*, vol. 62, no. 8, p. 1805, 2021.
- [26] E. Iadanza, F. Goretti, M. Sorelli, P. Melillo, L. Pecchia, F. Simonelli, and M. Gherardelli, "Automatic detection of genetic diseases in pediatric age using pupillometry," *IEEE Access*, vol. 8, pp. 34949–34961, 2020.
- [27] B.-W. Lu, G.-P. Wu, and L.-K. Xie, "In depth understanding of retinitis pigmentosa pathogenesis through optical coherence tomography angiography analysis: A narrative review," *Int. J. Ophthalmol.*, vol. 14, no. 12, pp. 1979–1985, Dec. 2021.
- [28] F. Antaki, R. G. Coussa, G. Kahwati, K. Hammamji, M. Sebag, and R. Duval, "Accuracy of automated machine learning in classifying retinal pathologies from ultra-widefield pseudocolour fundus images," *Brit. J. Ophthalmol.*, vol. 107, no. 1, pp. 90–95, Jan. 2023.
- [29] N. Brancati, M. Frucci, D. Gragnaniello, D. Riccio, V. Di Iorio, L. Di Perna, and F. Simonelli, "Learning-based approach to segment pigment signs in fundus images for retinitis pigmentosa analysis," *Neurocomputing*, vol. 308, pp. 159–171, Sep. 2018.
- [30] N. Brancati, M. Frucci, D. Gragnaniello, D. Riccio, V. Di Iorio, and L. Di Perna, "Automatic segmentation of pigment deposits in retinal fundus images of retinitis pigmentosa," *Computerized Med. Imag. Graph.*, vol. 66, pp. 73–81, Jun. 2018.
- [31] M. Arsalan, N. R. Baek, M. Owais, T. Mahmood, and K. R. Park, "Deep learning-based detection of pigment signs for analysis and diagnosis of retinitis pigmentosa," *Sensors*, vol. 20, no. 12, p. 3454, Jun. 2020.
- [32] G. Ravichandran, P. Elangovan, and M. K. Nath, "Diagnosis of retinitis pigmentosa from retinal images," *Int. J. Electron. Telecommun.*, vol. 65, p. 519, Jul. 2019.
- [33] N. Brancati, M. Frucci, D. Riccio, L. Di Perna, and F. Simonelli, "Segmentation of pigment signs in fundus images for retinitis pigmentosa analysis by using deep learning," in *Proc. Int. Conf. Image Anal. Process.*, Trento, Italy. Cham, Switzerland: Springer, 2019, pp. 437–445.



RUBINA RASHID received the M.Sc. degree in computer science from Quaid-i-Azam University, Islamabad, Pakistan, and the M.S.C.S. degree from the University of Sargodha, Lahore Campus, Pakistan. She is currently pursuing the Ph.D. degree in computer science with The Islamia University of Bahawalpur, Pakistan, under the supervision of Dr. Waqar Aslam. She is also a Lecturer with the Department of Computer Science and IT, Government Post Graduate College for

Women, Bahawalnagar. She has more than 15 years of teaching experience with the college and university levels in Pakistan. She has comprehensive programming skills in Python, Android Studio, and Arduino Controller. Her research interests include machine/deep learning models, the Internet of Things, and computer vision.



WAQAR ASLAM received the M.Sc. degree in computer science from Quaid-i-Azam University, Islamabad, Pakistan, and the Ph.D. degree in computer science from Eindhoven University of Technology, The Netherlands. He is a Professor of computer science and IT, The Islamia University of Bahawalpur, Pakistan. He received the Overseas Scholarship from HEC, Pakistan, for the Ph.D. degree. His research interests include performance modeling and QoS of wireless/computer networks,

performance modeling of (distributed) software architectures, radio resource allocation, the Internet of Things, fog computing, effort/time/cost estimation and task allocation in (distributed) Agile setups based software development, social network data analysis, DNA/chaos-based information security, and machine/deep learning models.



ARIF MEHMOOD received the Ph.D. degree from the Department of Information and Communication Engineering, Yeungnam University, South Korea, in November 2017. He is currently an Assistant Professor with the Department of Computer Science and IT, The Islamia University of Bahawalpur, Pakistan. His current research interests include data mining, mainly working on AI and deep learning-based text mining and data science management technologies.

DEBORA LIBERTAD RAMÍREZ VARGAS is currently a Professor with Universidad Europea del Atlántico, Santander, Spain. She is also affiliated with Universidad Internacional Iberoamericana de Campeche, Mexico; and Universidad Internacional Iberoamericana de Arecibo, PR, USA. Her research interests include project engineering, advanced mathematics, and fuzzy logic.



ISABEL DE LA TORRE DIEZ is currently a Professor with the Department of Signal Theory and Communications and Telematic Engineering, University of Valladolid, Spain, where she is also the Leader of the GTE Research Group (<http://sigte.tel.uva.es>). Her research interests include design, development, and evaluation of telemedicine applications; services and systems; e-health; m-health; electronic health records (EHRs); EHRs standards; biosensors; cloud and fog computing; data mining; quality of service (QoS); and quality of experience (QoE) applied to the health field.



IMRAN ASHRAF received the M.S. degree in computer science from Blekinge Institute of Technology, Karlskrona, Sweden, in 2010, and the Ph.D. degree in information and communication engineering from Yeungnam University, Gyeongsan, South Korea, in 2018. He was a Post-doctoral Fellow with Yeungnam University. He is currently an Assistant Professor with the Information and Communication Engineering Department, Yeungnam University. His research interests

include indoor positioning and localization, advanced location-based services in wireless communication, smart sensors (LIDAR) for smart cars, and data mining.

...