

Segmentation-Driven Morphology-Guided Attention Mechanism for Rhegmatogenous Retinal Detachment Classification

Husna Banu¹, Balaji Prabhu Baluvaneralu Veeranna²

¹Department of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur, Visvesvaraya Technological University, Belagavi, India

²Department of Computer Science and Engineering, Malnad College of Engineering, Hassan, India
E-mails: husna.hafeez@gmail.com (Corresponding author) balajitiptur@gmail.com

Abstract: Rhegmatogenous Retinal Detachment (RRD) is a vision-threatening condition that requires accurate classification to prevent vision loss. The existing Deep Learning (DL) models for RRD primarily rely on generic feature extraction from fundus images and often overlook subtle morphological variations across RD subtypes. To address this issue, an effective integrated model was proposed by combining EfficientNetB3-based DeepLabV3+ segmentation with morphological feature extraction and a morphology-guided attention mechanism referred to as Morpho-RRDNet. The segmentation stage isolates detachment zones by enabling RDD-relevant morphometric descriptors, including area percentage, centroid location, and convexity. These features were fused with deep spatial embeddings using a Morphological-Guided Attention Mechanism (MGAM) to focus mainly on RRD-specific regions while suppressing irrelevant regions. Finally, a Light Gradient Boosting Model (LightGBM) was introduced to classify the RRD and non-RRD classes. The proposed Morpho-RRDNet achieved competitive results, including accuracy (99.22%), sensitivity (100%), specificity (99.20%), and F1-score (85.71%), when compared to the existing optimized Ensemble Network.

Keywords: EfficientNetB3, Light gradient boosting model, Morphometric descriptors, Morphological-guided attention mechanism, Rhegmatogenous retinal detachment.

1. Introduction

Rhegmatogenous Retinal Detachment (RRD) is a critical retinal condition characterized by the separation of the retinal neurosensory layer due to the

thickness of the retinal tear from the underlying retinal pigment epithelium [1]. Most people experience permanent vision loss due to retinal detachment even with successful intervention [2]. This retinal tear allows vitreous fluid to accumulate beneath the retina, which often leads to rapid vision loss if left untreated [2]. Timely diagnosis of RRD is critical, as it significantly improves the outcomes of surgical repair when performed within an optimal time window [3, 4]. Fundus photography is considered the primary imaging modality for RRD classification, whereas Deep Learning (DL) models are widely used for automated screening. Over the years, different surgical approaches such as Scleral Buckling (SB) and Pars Plana Vitrectomy (PPV) have been widely used according to the retinal break location, posterior vitreous and lens status [5, 6]. Despite these advancements, DL models for generic performance in RD classification often lack lesion-specific feature representation in RRD classification [7]. Current DL models, such as Convolutional Neural Networks (CNNs) and segmentation models that mostly exist in RRD classification, still treat the problem as a generic classification task, mainly focusing on global image-level features [8, 9]. Some models primarily concentrate on intensity patterns while ignoring subtle morphological variants in RRD [10, 11]. The different descriptors included differences in tear shape, detachment extent, and location of the lesion, which are clinically important. This significant limitation reduces the effectiveness of differentiating RRD from other retinal abnormalities and recognizing its subtypes [12].

Furthermore, most models lack interpretability, making them less suitable for domain adoption, where transparency in decision-making is essential to support diagnosis and treatment [13, 14]. There is a need for an RRD classification model that works beyond conventional image-level classification and addresses the limitations of current DL methods. The proposed work is motivated by the need for an RRD classification framework that captures both the global visual context and the clinically relevant morphological descriptors of a detachment region. To achieve this requirement, an integrated model that combines EfficientNetB3-based DeepLabV3+ segmentation with morphological feature extraction and a morphological-guided attention mechanism (Morpho-RRDNet) is proposed. The segmentation combination is used for precise lesion isolation while capturing morphometric descriptors such as area percentage, centroid location, and convexity. Sequentially, a Morphological-Guided Attention Mechanism (MGAM) is introduced to suppress irrelevant lesion channels by enhancing morphological relevance in detachment regions. Finally, the classification stage effectively handles nonlinear feature interactions while maintaining transparency to understand the

reasons for each prediction. Using this integrated capability, the proposed model resolves the current limitations of subtype differentiation.

The contributions are listed as follows:

- The EfficientNetB3-based DeepLabV3+ precisely isolates the detachment regions while preserving the lesion boundaries and global spatial context. This segmentation network provides clinically interpretable morphometric descriptors directly from a binary mask for creating a clinically grounded feature space.
- Morphological feature extraction effectively captured RRD-aware descriptors, including the lesion area percentage, centroid location, and convexity from the segmented region. This feature extraction enables the classification model to capture subtle morphological variations that are crucial for an accurate RRD classification.
- A parameterized channel attention block is introduced as MGAM to embed clinical descriptors through a learnable mapping and concatenate with globally pooled deep channels for producing per-channel weights.

The organization is structured as follows: Section 2 describes related work on RRD classification; Section 3 describes the methodology with an architecture; Section 4 demonstrates the experimental results; Section 5 presents the discussion, and Section 6 presents the conclusion.

2. Related works

Anitha and Aravindhar [15] recommended efficient retinal detachment classification using hybrid Machine Learning (ML) with Levy flight-based optimization. This model uses data augmentation to enhance diversity during segmentation and mainly focuses on the Region Of Interest (ROI). The recommended model was employed to select the most discriminative features for the RD classification. The recommended model achieved better results compared to traditional ML models; however, it did not provide clear insights into the retinal-specific features. Subsequently, Yadav et al. [16] presented RD classification using a Graph Neural Network (GNN) through fundus images. The input images utilized a GNN to classify the RD using the Euclidean distance with simple linear iterative clustering. Graph features were extracted using the GGRD-Net model, encoded for prediction, and trained on limited datasets with minimal memory requirements. Nevertheless, the model relies heavily on a specific dataset, limiting its generalizability to diverse scenarios.

Yadav, Murugan and Goel [17] demonstrated an optimized ensemble deep neural network for RD classification using fundus images. The framework

identified RD through five stages, in which an ensemble ResNet50 was employed with extreme learning and a random vector functional-link network. However, the demonstrated model depends on specific datasets and image qualities that require computational resources to ensure generalizability and broader applicability. Huimin Li et al. [18] presented Artificial Intelligence assisted (AI-assisted) management of RD from ultra-widefield fundus images based on a weakly supervised approach. The proposed model localizes and outlines RD regions using 24,308 ultra-widefield fundus images. The process involved training the model to detect RD lesions and to generate customized preoperative postural guidance. Nevertheless, the presented model may not provide detailed information for complex cases, and further refinement of diverse datasets is required to ensure applicability.

Lake et al. [19] demonstrated a retinal shape-based classification of RD in eyes with posterior vitreous detachment. The model used Optical Coherence Tomography (OCT) images to quantify retinal irregularities as biomarkers for RD. This process analyzed images from 264 individuals to calculate retinal irregularities in the frequency domain. The demonstrated model utilized features such as retinal irregularity and axial length. However, the classifier's moderate performance identified only 1/3 of the RD, with potential variability in image quality. Zhongwen Li et al. [20] developed a cascaded deep learning system for automated RD detection and macula-on/off classification using ultra-widefield fundus images. The developed model involves macula on/off classification and automated guidance for preoperative and repair urgency. However, potential limitations include dependence on RDD-related morphological features to ensure reliable performance in real-world applications.

Guo et al. [21] demonstrated an Artificial Intelligence (AI) model to predict postoperative functional outcomes in patients with RRD undergoing PPV and gas tamponade. A fusion method was employed to capture the interactions between different modalities to enhance the predictive accuracy for better patient outcomes. The demonstrated AI model typically utilizes Convolutional Neural Networks (CNNs), which are constrained by RRD-aware features that reduce performance.

Research gap. Despite advancements in RRD classification, most existing studies rely on CNN backbones or segmentation approaches, which do not explicitly focus on RRD morphological characteristics. The conventional model extracted hierarchical features but was constrained by multi-scale contextual details, which made it critical for understanding the extent of detachment, location, and shape. To address these limitations, there is a need for a hybrid framework that does not simply combine the existing methods but establishes a new integration between

segmentation, morphology, and machine learning. Unlike previous works, which used morphological descriptors or attention mechanisms separately, the proposed model integrates segmentation-driven morphological extraction, MGAM for descriptor-to-feature fusion, and a Light GBM classifier for non-linear decision-making. This integration enables the proposed model to ensure both clinical interpretability and improved generalization across datasets.

3. Methodology

In this section, the proposed Morpho-RRDNet is introduced for automated RRD classification by uniquely integrating morphological domain knowledge into deep learning. Unlike conventional CNN-based approaches, the proposed framework employs EfficientNetB3-based DeepLabV3+ to segment detachment zones and extract relevant RRD morphometric features, such as area percentage, centroid, and shape descriptors. An MGAM block fuses these descriptors with deep spatial embeddings to enable the model to focus on the RRD-specific regions. Finally, the Light Gradient Boosting Model (LightGBM) classifier captures complex feature interactions and classifies them into RRD and non-RRD classes. The architecture of the proposed Morpho-RRDNet is illustrated in Fig. 1.

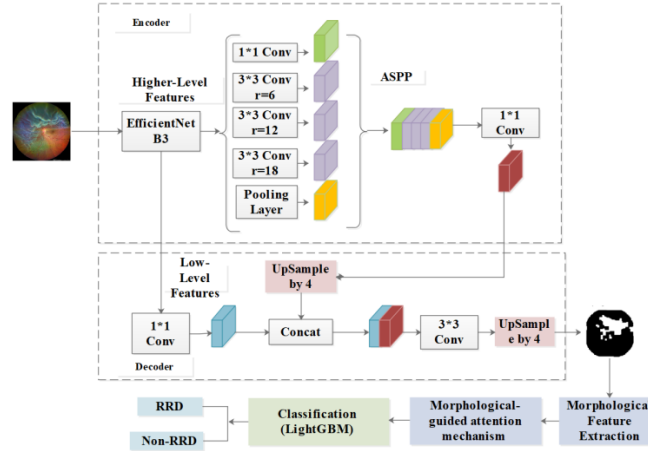


Fig. 1. Architecture of the proposed segmentation-driven morphology-guided framework for RRD classification including ASPP module

3.1. Dataset

In this section, the RFMiD 2.0 dataset is introduced for RRD classification as RRD or non-RRD [22]. This dataset contains approximately 860 retinal fundus images covering multiple ocular diseases. The dataset images, such as retinal detachment, retinal holes, tears, and other retinal conditions, contain 19 fundus images grouped

to form the RRD class. The remaining 841 fundus images, considered as the non-RRD class, include retinal images without detachment, covering other retinal conditions such as diabetic retinopathy and normal cases. The non-RRD class consists of a representative subset of other retinal conditions, including normal cases and diseases such as diabetic retinopathy, age-related macular degeneration, and other abnormalities. The subset was restricted to maintain class balance with approximately 30% normal and 70% pathological non-RRD images. For this study, images corresponding to retinal detachment were selected as the RRD class. In contrast, a representative subset of other retinal conditions was grouped as the non-RRD class to ensure a balanced binary classification setup. The data were split into 70% training, 15% validation, and 15% testing sets to ensure class balance across splits. An ARMD retinal image dataset is considered for cross-data validation [23]. This dataset comprises 511 fundus images, which are sourced from four different publicly available datasets with 39 categories. The original input resolution of the input image is 4288×2848 pixels, which varies slightly across samples, and all images were resized to a resolution of 224×224 pixels before processing. These ARMD dataset images are collected from multiple sources and exclusively represent non-RRD cases, covering a wide spectrum of retinal pathologies. The dataset does not contain any RRD samples and is therefore used only for evaluating the false positives behavior and model specificity under domain shift.

3.2. Data preprocessing

The dataset has an imbalance; to address this issue, data augmentation has been introduced for only the minority class. Geometric and photometric augmentations have been employed for rotation and flipping with brightness adjustment [23]. The augmented image is obtained using the following mathematical expression:

$$(1) \quad I' = C_{\gamma} \left(B_{\beta} \left(F_h \left(R_{\theta}(I) \right) \right) \right),$$

where I represents the input image, I' denotes the augmented image, C_{γ} , B_{β} refers to the contrast scaling and brightness adjustment, F_h denotes the horizontal flip, and R_{θ} specifies the rotation. This augmentation improves generalization without altering the RRD-relevant morphology. The use of traditional GAN-based methods to address imbalance issues is complex and requires additional training. The augmentation technique is computationally efficient and does not require additional training to avoid synthetic artifacts. Finally, the data augmentation increased the minority class from 19 images up to 190 images. This augmentation improves feature diversity and reduces severe imbalance while avoiding excessive duplication

that could lead to overfitting. The dataset remains moderately imbalanced to reflect real-world class distribution.

Sample images for data augmentation are shown in Fig. 2.

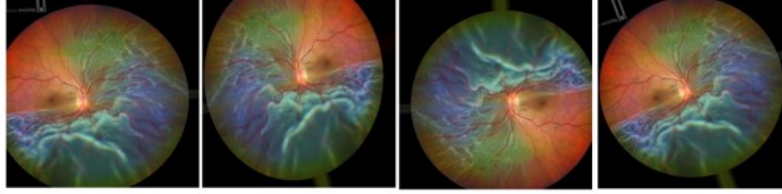


Fig. 2. Preprocessed RFMID 2.0 images with data augmentation for class imbalance

3.3. RRD-aware segmentation with EfficientNetB3-based DeepLabV3+ segmentation

From the segmentation, the DeepLabV3+ model is utilized with EfficientNet B3 as an encoder to extract hierarchical features from the fundus image. Traditional segmentation models, such as UNet, extract images with convolutional filters and are constrained by hierarchical and multi-scale features. The proposed segmentation network effectively extracts hierarchical features using EfficientNet B3 and multi-scale features using Atrous Spatial Pyramid Pooling (ASPP). The EfficientNetB3 [24] as a backbone in segmentation is introduced to adapt the RRD classification by using morphological post-processing by linking the segmentation mask explicitly for clinically interpretable descriptors. This model is a convolutional neural network architecture that scales depth, width, and resolution using a compound scaling method. In this work, the model was initialized with ImageNet-pretrained weights and fine-tuned on the RFMID 2.0 datasets. The final layers were adapted to the segmentation task by integrating them within the DeepLabV3+ framework, and all the layers were retrained using AdamW optimization with a learning rate of 0.001 to capture RRD-specific features. This strategy effectively ensures that the segmentation output is not only accurate but also clinically meaningful for the classification. Initially, the preprocessed image was passed through the convolutional blocks of EfficientNetB3, where multiple blocks were used to extract hierarchical features. The initial layers extract low-level features such as vessel edges and small texture details. Consequently, the intermediate layers extract the shape outlines of the RRD detachment zones, tear shapes, and folds. Finally, the deeper layers extract high-level features, which include abstract lesion representations, such as detachment morphology and macula patterns. The mathematical expression for the multi-scale feature representation of the encoder is given in the next equation

$$(2) \quad F_{\text{enc}} = \{f_1, f_2, \dots, f_k\}, \quad f_i \in \mathbb{R}^{h_i \times w_i \times c_i}.$$

Here, f_1 denotes the low-level edge and texture features, and f_k represents the features from the deeper layers that provide highly abstract lesion representations. Moreover, $h_i \times w_i$ specifies the dimensions of the progressively reduced feature map. In addition, c_i denotes the i -th feature channel extracted from the encoder corresponding to different levels of abstraction. Consequently, these hierarchical features, including the intermediate- and deep-layer features, are passed through the ASPP module, which applies dilated convolutions with different dilation rates, r to capture local and global structures without losing spatial resolution in the equation

$$(3) \quad F_{\text{ASPP}} = \bigcup_{r=1}^m \text{Conv}_{d_r}(f_k),$$

where, Conv_{d_r} denotes the convolution operation with dilation rate r , and m denotes the number of parallel atrous convolutions. A small dilation rate captures fine details, such as local retinal breaks or retinal tears, while a large dilation rate captures the global context in wide RRD areas. The ASPP outputs from different scales were concatenated to obtain a final feature map. F_{multi} with the local and global contexts of the RRD image in the equation

$$(4) \quad F_{\text{multi}} = \text{Concat}(F_{\text{ASPP}}).$$

The concatenated feature map is passed through the decoder, which generates a binary mask for the input image using the features from both the ASPP and EfficientNetB3. F_{multi} is upsampled by a factor of four to restore the spatial resolution,

$$(5) \quad F_u = \text{Upsample}_4(F_{\text{ASPP}}).$$

Here, F_u denotes the upsampled feature map, and the low-level hierarchical feature map from the encoder is defined as F_{low} which contains the vessel boundaries and texture details, and was passed through a 1×1 convolution to reduce the spatial dimensions. The upsampled ASPP feature map was concatenated with the processed low-level feature map F_{low} to produce a feature map with a combined representation in the equation

$$(6) \quad F_{\text{fuse}} = \text{Concat}(\text{Upsample}_4(F_{\text{ASPP}}), \text{Conv}_{1 \times 1}(F_{\text{low}}),$$

where the F_{fuse} refers to the fused feature map, which contains the global semantic context with local details in precision, and a 3×3 convolution was applied to this fused map to refine the spatial consistency and smooth boundaries. Finally, the refined feature map is again upsampled by a factor of four to match the original input resolution. Following a 1×1 convolution with a sigmoid activation function to generate a binary mask for the feature map in the equation

$$(7) \quad S = \sigma \left(\text{Conv}_{1 \times 1}(\text{Upsample}_4(F_{\text{re}})) \right),$$

where the produced binary mask $S \in [0, 1]^{H \times W}$, and each pixel value indicates the probability of belonging to an RD region and background. Unlike a generic

segmentation mechanism for RRD classification, this research explicitly refines the masks using morphological opening with erosion followed by dilation to ensure clean lesion boundaries. This binary mask is morphologically cleaned to remove unwanted artifacts and isolated pixels to enhance the reliability of feature extraction. The noise from the binary mask is removed using morphological opening, which is erosion followed by dilation in the equation

$$(8) \quad S' = (S \ominus K) \oplus K.$$

Here, S is the binary mask and S' denotes the cleaned mask, and K signified as the structuring element. Here K is the disk-shaped kernel of radius 3 pixels, which is a size of 7×7 with one iteration each of erosion and dilation. This mechanism effectively isolates the noise pixels while preserving the clinically relevant lesion boundaries. After cleaning the binary mask, morphological features related to RRD were extracted, and the area percentage was computed to determine the proportion of retinal tears for RRD in relation to the total retinal area in the equation

$$(9) \quad \text{RRD}_{\text{area}}(\%) = \frac{|R_{\text{tear}}|}{|R_{\text{total}}|} \times 100\%,$$

where, $|R_{\text{tear}}|$ denotes the number of pixels that belong to the retinal tear region and $|R_{\text{total}}|$ refers to the total retinal area in the binary mask. A higher $\text{RRD}_{\text{area}}(\%)$ specifies the severe detachment area. Consequently, the centroid position of the lesion was determined using the center-of-mass calculation in the equation

$$(10) \quad C_x = \frac{1}{N} \sum_{i=1}^N x_i, \quad C_y = \frac{1}{N} \sum_{i=1}^N y_i.$$

Here: x_i, y_i denotes the coordinates of pixels in the lesion region; N refers to the total number of pixels; (C_x, C_y) represents the centroid. This center-of-position helps localize RRD as superior, inferior, or inferior, based on the Lincoff rules. Additionally, the convexity and irregularities were determined by comparing the area of the convex hull of the contour with the actual area in the equation

$$(11) \quad \text{Convexity} = \frac{\text{Area}_R}{\text{Area}(\text{ConvexHull}_R)},$$

where: Area_R specifies the area of the lesion region; ConvexHull_R defines the convex hull which encloses the lesion boundary. A lower convexity ratio indicates irregular or multiple tear shapes.

Morphological-guided attention mechanism. The attention mechanism fuses morphologically extracted features with deep spatial feature maps from the segmentation encoder to enable the network to learn the most relevant RRD regions. The encoder produces a feature map as $F \in \mathbb{R}^{H \times W \times C}$, morphological features such as RRD_{area} , centroid coordinates (C_x, C_y) and convexity are fused and organized as a vector $m \in \mathbb{R}^d$. The attention mechanism transforms m with a

learnable mapping $W_m \in \mathbb{R}^{d \times C}$ to align it with the channel dimension in the equation

$$(12) \quad m' = W_m m + b_m, \quad m' \in \mathbb{R}^C, \quad W_m \in \mathbb{R}^{C \times d}, \quad b_m \in \mathbb{R}^C.$$

Here, m' denotes the descriptor embedding in channel space, which is concatenated with a global average-pooled channel $g \in \mathbb{R}^C$ to form a channel attention in the next equation:

$$(13) \quad \alpha = \sigma(W_a[g \oplus m'] + b_a),$$

where, $\alpha \in [0, 1]^C$ refers to channel attention, $W_a \in \mathbb{R}^{C \times 2C}$ denotes the weight, and $b_a \in \mathbb{R}^C$ specifies the bias term. Then, the reweighted feature map is obtained as in the equation

$$(14) \quad F' = \alpha \odot F.$$

Above: $\oplus$ denotes channel-wise multiplication; $\odot$ specifies the concatenation; σ denotes the sigmoid activation function to generate attention weights. The final refined feature map F' is fed to the LightGBM classification head to allow the classification model to focus mainly on RRD-relevant patterns with prior knowledge of morphological features, instead of relying on CNN-learned features.

3.4. Classification with LightGBM

The refined feature map $F' \in \mathbb{R}^{H \times W \times C}$ was converted into a tabular vector with global statistics to make the input format suitable for the classification model. Unlike the conventional fully-connected deep classifiers, the proposed model utilized LightGBM as a classifier. This classification model is introduced to tabularize the deep embeddings with morphological descriptors. LightGBM is selected due to the need for interpretability in the clinical context, non-linear decision-making, and efficiency in avoiding overfitting through leaf-wise growth. The refined feature map produced by MGAM is global average-pooled to a vector as $g \in \mathbb{R}^C$. This vector is then concatenated with the morphological descriptors to form a fused feature vector in the equation

$$(15) \quad \hat{x} = g \oplus [\text{RRD}_{\text{area}}, C_x, C_y, \text{Convexity}] \in \mathbb{R}^{C+4}.$$

This fused representation explicitly represents the derived embeddings with an attention mechanism with clinically interpretable descriptors that are then passed to LightGBM for classification. These vectors are used with LightGBM to learn an additive model of the decision tree T in the equation

$$(16) \quad f(z) = \sum_{t=1}^T \eta f_t(\hat{x}),$$

where η specifies the learning rate that minimizes the regularised logistic loss for binary classification of RRD in the equation

$$(17) \quad \mathcal{L} = \sum_i [-y_i \log \sigma(f(z_i)) - (1 - y_i) \log(1 - \sigma(f(z_i)))] + \Omega,$$

where $\sigma(x) = \frac{1}{1+e^{-x}}$, Ω denotes the tree regularization, and LightGBM discards each feature into B histograms to speed up the splitting. For each sample i , the gradient is computed as $g_i = \partial \mathcal{L} / \partial f$ and Hessian $h_i = \partial^2 \mathcal{L} / \partial f^2$ at each boosting step t under the classification model. For a candidate split produced in the Left (L) or Right (R) sets, aggregated as in the equation

$$(18) \quad G_L = \sum_{i \in L} g_i, \quad H_L = \sum_{i \in L} h_i.$$

From this candidate split, the split gain is calculated to make the model make impactful decisions and reduce the training error in the equation

$$(19) \quad \text{Gain} = \left(\frac{G_L^2}{(H_L + \lambda)} + \frac{G_R^2}{(H_R + \lambda)} - \frac{(G_L + G_R)^2}{(H_L + H_R + \lambda)} \right) - \gamma.$$

Here, λ and γ denote the regularization terms that reduce the leaf penalty, and the classification model grows leaf-wise, which fits the complex boundaries from z . Finally, a threshold value is set for the prediction model, and the prediction is computed as in the equation

$$(20) \quad \hat{p} = \sigma(f(z)),$$

where, $\hat{p}$ denotes the final prediction, and a threshold value $\tau = 0.5$ is considered; if $\hat{p} \geq \tau$ then the class is classified as RRD else non-RRD. The classification of RRD is based on its morphological features, which helps classify RRD more effectively and accurately. For the effective reproducibility of the proposed Morpho-RRDNet, pseudocode is described in Algorithm 1.

Algorithm 1. The Pseudo Code for the Proposed Morpho-RRDNet for Retinal Detachment

Input: Fundus image dataset $D = \{I_1, I_2, \dots, I_n\}$

Output: Classification results {RRD, Non-RRD}

for each image $I \in D$:

 Let image I of size $M \times N$,

 Extract hierarchical features F_{enc} using the Equation (2)

 Apply ASPP with dilations and padding. $r \rightarrow$ multi-scale map using the Equations (3) and (4)

 Fuse encoder + ASPP $\rightarrow$ feature map F using the Equation (6)

 Decode and upsample $\rightarrow$ segmentation mask S using the Equation (7)

 Clean (erosion+dilation) a binary mask S using the Equation (8)

 End for

for each S' :

 Compute lesion area percentage. $\text{RRD}_{\text{area}}(\%)$ using the Equation (9)

 Compute centroid (C_x, C_y) using the Equation (10)

 Compute *convexity* using the Equation (11)

```

End for
Form descriptor vector  $m = \{A (\%), C, \text{Conv}\}$ 
for each segmentation feature  $F$ :
    Embed  $m$  into channel space  $\rightarrow m'$  using the Equation (12)
    Concatenate  $m'$  with global pooled channels using the Equation (13)
    Compute channel attention weights  $\alpha$ 
Refine features:  $F' = \alpha \odot F$  using the Equation (14)
End for
For each  $F'$ :
    Global average pool  $\rightarrow$  vector  $\hat{x}$  using the Equation (15)
    Concatenate  $\hat{x}$  with  $m \rightarrow$  fused vector  $f(z)$ 
    Train LightGBM with leaf-wise splitting using the Equations (16)-(19)
    Predict final class:
    if  $\hat{p} \geq \tau \rightarrow \text{RRD}$ 
    else  $\rightarrow$  non-RRD using the Equation (20)
Return predicted labels  $\{\text{RRD}, \text{Non-RRD}\}$  with probabilities.

```

4. Experimental results

The proposed Morpho-RRDNet was implemented on a system equipped with standard computational resources to ensure efficient training and evaluation. The specifications were 8 GB RAM, Ubuntu 22.04, and an Intel Core i3 and GPU setup for training and evaluation. A PyTorch framework was utilized, and the attention mechanism utilized. $d = 4$ Morphological descriptors. The channel dimension C is aligned with encoder features. The weights W_m and W_a are initialized using Xavier initialization with zero bias. Further, the model is trained using the AdamW optimizer with a learning rate of 0.001, a batch size of 8, and mixed precision (AMP) on a GPU with more than 12 GB VRAM. The model was trained for 50 epochs with early stopping based on validation loss to prevent overfitting. A weight decay of $1e-4$ was applied, and the learning rate was reduced using a cosine annealing scheduler. Data augmentation was applied during training, while validation and testing were performed on unseen data without augmentation. The morphological operation was performed using OpenCV 4.7 to ensure reproducibility. The evaluation metrics, including Accuracy (Acc), Sensitivity (Sen), Specificity (Spe), F1-score, and Area Under the Curve (AUC), are given as in the next equations:

$$(21) \quad \text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}},$$

$$(22) \quad \text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

$$(23) \quad \text{Spe} = \frac{\text{TN}}{\text{TN} + \text{FP}},$$

$$(24) \quad \text{F1-score} = \frac{\text{TP}}{\text{TP} + \frac{1}{2}(\text{FP} + \text{FN})},$$

$$(25) \quad \text{AUC} = \int_a^b F(t) dt.$$

Here, TP, TN denote: True Positive, True Negative; FP, FN indicate False Positive, False Negative; F_t represents the Cumulative Distribution Function (CDF) as well as a, b define the integration limits. In these metrics, high sensitivity ensures timely intervention for detachment while the specificity avoids unnecessary regions. The F1-score addresses class imbalance by effectively balancing the precision and recall.

4.1. Performance analysis

The performance of the proposed Morpho-RRDNet model was compared to that of the baseline model with other combinations. The considered performance models are DeepLabV3+ with Channel Block Attention Module (DeepLabV3+-CBAM), DeepLabV3+ with Self-Attention-based Random Forest Classification (DeepLabV3+-SA-RF), and DeepLabV3+ with Multi Attention and Decision Tree (DeepLabV3+-MDT) for RRD classification. To account for stochastic variation, each performance model was trained and evaluated over five runs with different kernels using a fixed RFMiD test set. The results are presented in Fig. 3.

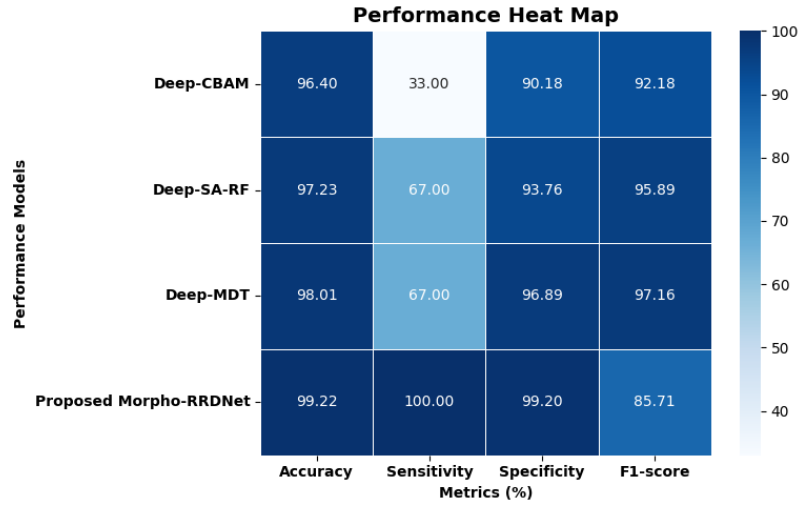


Fig. 3. Performance analysis of the proposed Morpho-RRDNet through a heat map over five runs

As shown in Fig. 3, these results were calculated using the reported metrics based on the total number of correctly classified samples over the total number of evaluated samples in each class. The proposed Morpho-RRDNet model achieved better results than the baseline model with other attention and classification

combinations for the RRD classification. The proposed Morpho-RRDNet demonstrated better results in terms of Accuracy (99.22%), Sensitivity (100%), Specificity (99.20%), and F1-score (85.71%). The total samples in the test set include 129, where 126 refer to non-RRD and three images refer to RRD samples. For the test set, no augmentation is performed, and the performance results were derived from the values as accuracy (128/129), sensitivity (3/3), and specificity (125/126), respectively. Here, to mitigate overfitting issues, data augmentation and AdamW optimization with weight decay were introduced while training. The dataset is split 70% training, 15% testing, and 15% for validation, ensuring no data leakage across sets.

4.2. False-positive stress test under domain shift

The ARMD dataset is intentionally used as a negative-only external stress test to evaluate the model's behavior under diverse non-RRD pathological conditions. This setup specifically isolates and measures the model's tendency to produce false positives instead of overall classification performance. The model is trained on the RFMID 2.0 dataset and validated with the ARMD dataset. The results are given in Table 1. Additionally, the trained model is directly tested on the independent ARMD dataset without retraining to evaluate generalization, and the corresponding performance metrics are reported separately. Here, the ARMD dataset does not explicitly include RRD cases; 511 cases are related to non-RRD samples, which include a heterogeneous set of retinal pathologies. Therefore, the ARMD dataset is utilized to evaluate the model's robustness in distinguishing RRD from a wide range of non-RRD conditions and its ability to minimize false positives under dataset shift.

Table 1. False positive stress test of the proposed Morpho-RRDNet under domain shift by training on RFMID 2.0 and testing on the ARMD dataset (non-RRD samples only)

Cross-data validation	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)
DeepLabV3+-CBAM	88.90	N/A	86.90	N/A
DeepLabV3+-SA-RF	90.34	N/A	89.45	N/A
DeepLabV3+-MDT	95.12	N/A	94.17	N/A
Proposed Morpho-RRDNet	99.82	N/A	99.82	N/A

From Table 1, the proposed Morpho-RRDNet attained better results than the traditional models when tested on the ARMD dataset. The proposed model attained an Accuracy of 99.82% (510/511) and a specificity of 99.82% (510/511). Sensitivity and F1-score are marked as N/A because the ARMD dataset contains only non-RRD samples, making these metrics undefined. Therefore, this cross-data validation with ARMD demonstrates the robustness of the proposed Morpho-

RRDNet in minimizing false positives and generalizing effectively across diverse non-RRD conditions. Nevertheless, the ARMD dataset does not contain RRD samples; the diversity across 39 retinal conditions makes it a challenging benchmark for evaluating false positive suppression. A rationale for this evaluation is that a model that generalizes poorly would incorrectly classify unfamiliar pathologies as RRD.

4.3. Statistical analysis

Furthermore, to validate the robustness of the proposed Morpho-RRDNet, a paired t -test was conducted between the proposed model and the baseline models (DeepLabV3+-CBAM, DeepLabV3+-SA-RF, and DeepLabV3+-MDT). All the performance metrics were evaluated on the fixed RFMiD 2.0 test set and computed. p -values for Accuracy, Sensitivity, Specificity, and F1-score are summarized in Table 2.

Table 2. Paired t -test results between the proposed Morpho-RRDNet and baseline models

Comparison	Accuracy (p)	Sensitivity (p)	Specificity (p)	F1-score (p)
Morpho-RRDNet vs DeepLabV3+-CBAM	< 0.001	< 0.001	< 0.001	< 0.001
Morpho-RRDNet vs DeepLabV3+-SA-RF	< 0.001	< 0.001	< 0.001	< 0.001
Morpho-RRDNet vs DeepLabV3+-MDT	0.005	0.004	0.006	0.008

From Table 2, the Morpho-RRDNet achieved statistically significant enhancements over all baseline models. To account for stochastic variation during training, each model was independently retrained five times using different random initializations while maintaining the same fixed RFMiD 2.0 dataset split across all runs. The evaluation was performed on the same test set, and the reported performance values correspond to mean performance obtained across these independent runs. The minor variations across runs arise due to the stochastic optimization effects, including weight initialization and mini-batch sampling. As the fixed test set contains a limited number of RRD samples, the variability in sensitivity-related metrics remained minimal across runs. Statistical significance between the proposed model and baseline models was assessed using paired t -tests across repeated experimental runs. The dataset split (70% training, 15% validation, and 15% testing) was considered fixed across all runs, and the evaluation was consistently performed on the same test set. The paired t -test conducted and shown improvements against baseline DeepLabV3+-CBAM and DeepLabV3+-SA-RF were highly significant ($p < 0.001$), while the results over DeepLabV3+-MDT also

demonstrated a significant result ($p < 0.05$). These results confirm that the better performance of the proposed Morpho-RRDNet is not attributed to random variation.

4.4. Ablation study

The ablation study is carried out to quantify the contribution of each component in the proposed Morpho-RRDNet for RRD classification. Mainly, 4 variants have been considered for the ablation study, which are described as follows in Table 3.

Table 3. Ablation study of the proposed Morpho-RRDNet

Model	Accuracy (%)
(Baseline DeepLabV3+) + LightGBM	90.23
(DeepLabV3+) + Morphology descriptors (No MGAM)	94.83
(DeepLabV3+)+MGAM	97.23
(DeepLabV3+)+ MGAM+LightGBM	99.22

From Table 3, the results of the ablation study show that the addition of morphological descriptors improved interpretability and raised accuracy (+4.60). The MGAM further enhanced the performance (+3.40), while the classification with LightGBM achieved the best accuracy of 99.49%, respectively.

4.5. Comparative analysis

A comparative analysis was conducted to validate the performance of the proposed Morpho-RRDNet with existing models for RRD classification. The existing models include LFB-ASO [17] and GRD-Net [18]. To ensure a fair comparison, all baseline models were simulated on the same preprocessed RFMID 2.0 dataset with identical splits for training and testing as explained in Table 4.

Table 4. Comparative analysis of proposed Morpho-RRDNet

Comparative model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)
LFB-ASO [17]	98.75	N/A	97.62	97.01
GRD-Net [18]	92.91	91.00	93.18	91.50
Proposed Morpho-RRDNet	99.22	100	99.20	85.71

As shown in Table 4, the proposed Morpho-RRDNet was compared with the existing model for RRD classification, and it demonstrated better results. These results were obtained on the RFMiD 2.0 test set, and the existing LFB-ASO [17] established results in terms of Accuracy (98.75%), Specificity (97.62%), and F1-score (97.01%) owing to the utilization of ASO. However, it demonstrated improved results compared to GRD-Net [18]; it did not provide clear insight into the features of RRD. Nevertheless, this model is heavily dependent on a specific dataset, which requires more computational resources to ensure generalizability. The proposed Morpho-RRDNet achieved competitive results compared to existing

models, including Accuracy (99.22%), Sensitivity (100%), Specificity (99.20%), and F1-score (85.71%). The results of the proposed model use the same test set as mentioned in Fig. 3, and all reported counts are identical to ensure consistency and eliminate ambiguity. While the numerical improvement over existing Opti-En-net [19] appeared marginal, the paired t -test results in statistical analysis confirmed the differences with a statistical significance of $p < 0.05$. Moreover, the proposed framework also considered the AUC metric to evaluate how the model concentrated on the true positives with respect to other existing models is shown in Fig. 4.

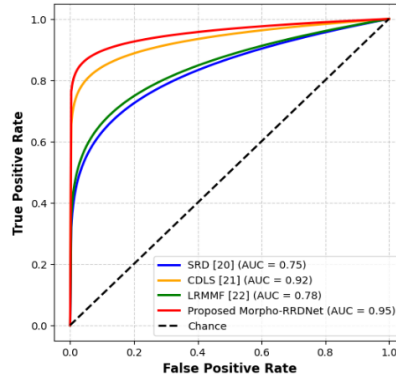


Fig. 4. Comparison of AUC for the proposed Morpho-RRDNet with existing models

Fig. 4 clearly shows that the existing model used for RRD classification achieved lower results than the proposed Morpho-RRDNet. The proposed Morpho-RRDNet attained an AUC of 0.92, which is better than that of existing models because of the utilization of the morphological features of RRD for classification.

5. Discussion

In this section, the research implications are discussed as follows: The proposed Morpho-RRDNet model carries several important research implications for retinal detachment classification. Initially, the proposed model directly addresses a key limitation in existing deep learning models, which often ignore explicit lesion morphology by integrating RRD-aware segmentation with morphological descriptors such as detachment area percentage, centroid location, and convexity. This integration ensures that classification is not entirely based on standard feature extraction but instead focuses on clinically relevant and interpretable features. Then, the proposed model improved its ability to generalize across different datasets by fusing deep spatial embeddings with morphology-driven descriptors. Consequently, the proposed model achieved high sensitivity by enabling the reliable detection of subtle RRD cases, which is essential for timely intervention for

preventing irreversible vision loss. Additionally, the MGAM integration into the proposed model enhanced specificity by suppressing irrelevant background features such as the optic disc and others, reducing false positives and unnecessary interventions. Finally, the integration of morphology-based descriptors introduces an inherent level of interpretability by linking model decisions to clinically meaningful features. Unlike conventional black-box CNN models, the proposed framework allows the contribution of each descriptor to be explicitly analyzed through MGAM attention weights and LightGBM feature importance scores. A limitation of this study is the small number of RRD samples in the RFMiD test set, which results in sensitivity being computed on a limited sample size. However, the exact count-based evaluation is reported, and future work will focus on validating the model using larger and more diverse datasets containing additional RRD samples. These components enable clinicians to understand which morphological characteristics influence the classification outcome. Nevertheless, qualitative representation and attention mechanisms provide a transparent decision pathway by improving trust and potential clinical applicability.

6. Conclusion

In this study, the effectiveness of the proposed Morpho-RRDNet for automated RRD classification with an effective integration of the models was evaluated. The main limitation of existing models is their reliance on CNN and other segmentation models that do not focus on RRD morphological characteristics. This limitation led to ineffective classification of RRD owing to subtle differences that are unique to the RD subtypes. Therefore, the proposed Morpho-RRDNet utilizes effective integration by combining EfficientNetB3 with DeepLabV3+ for segmentation with morphological feature extraction. This process effectively adds a unique variation in the RRD to the model, and the MGAM module combines deep spatial embeddings with clinically relevant morphometric descriptors. This attention mechanism preserves only the channels that belong to RRD while suppressing irrelevant channels. Finally, the LightGBM classifier enhances decision-making by modelling the nonlinear relationships between features. The experimental results showed that the proposed Morpho-RRDNet achieved competitive results in terms of accuracy (99.22%), sensitivity (100%), specificity (99.20%), and F1-score (85.71%). In the future, the proposed Morpho-RRDNet will be extended to classify not only binary RRD, but also other RD subtypes, including Tractional RD and Exudative RD, with multimodal imaging to further improve diagnostic accuracy.

References

1. Gan, F., Q. Wan, H. Fan, W. Chen, W. Qin, X. Li, T. Liu, J. Cao, Z. You. Automatic Classification of Retinal Detachment Using a Deep Transfer Learning Algorithm Based on Multimodal Fundus Images: A Multicenter Study. – *Ophthalmic Epidemiology*, 2024.
2. Lake, S. R., M. J. Bottema, T. Lange, K. A. Williams, K. J. Reynolds. Swept-Source OCT Mid-Peripheral Retinal Irregularity in Retinal Detachment and Posterior Vitreous Detachment Eyes. – *Bioengineering*, Vol. **10**, 2023, No 3, 377.
3. Al-Dwairi, R., O. Saleh, H. Mohidat, S. Al Beiruti, A. Alshami, L. El Taani, A. Sharayah, A. H. Al Sharie, A. Aleshawi. Characteristics, Risks, and Prevention of Rhegmatogenous Retinal Detachment in the Contralateral Eye. – *Journal of Clinical Medicine*, Vol. **14**, 2025, No 1, 222.
4. Sharmin, S., M. R. Rashid, T. Khatun, M. Z. Hasan, M. S. Uddin. A Dataset of Color Fundus Images for the Detection and Classification of Eye Diseases. – *Data in Brief*, Vol. **57**, 2024, 110979.
5. Ferrara, M., M. Al-Zubaidy, A. Song, P. Avery, D. A. Laidlaw, T. H. Williamson, D. Yorston, D. H. Steel. The Effect of Age on Phenotype of Primary Rhegmatogenous Retinal Detachment. – *Eye*, Vol. **37**, 2023, No 6, pp. 1114-1122.
6. Motta, L., R. Frisina, M. Ripa, I. Gius, A. Greggio, L. Tozzi, G. De Salvo, A. Meduri. Postoperative Complications after Successful Primary Rhegmatogenous Retinal Detachment Repair. – *BMC Ophthalmology*, Vol. **23**, 2023, No 1, 77.
7. Choovuthayakorn, J., P. Rajsirisongsri, D. Patikulsila, O. Nanegrungsunk, V. Chaikitmongkol, S. Seetasut, Y. Tantivit, N. Krisanuruks, A. Sangkaew, P. Tangkitchot, A. Apivatthakakul, P. Kunavisarut. Characteristics and Surgical Outcomes of Giant Retinal Tear Associated Rhegmatogenous Retinal Detachment. – *Scientific Reports*, Vol. **14**, 2024, No 1, 19943.
8. Oquendo, P. L., G. S. Sodhi, S. C. Naidu, I. M. Melo, A. Pecaku, S. Demian, P. J. Belin, W. W. Lee, P. G. Christakis, H. Hamli, A. Bansal. Optical Coherence Tomography Features in Fovea-Off Exudative vs Rhegmatogenous Retinal Detachment. – *American Journal of Ophthalmology*, Vol. **268**, 2024, pp. 212-221.
9. Pecaku, A., S. C. Naidu, S. Demian, M. C. Pimentel, I. M. Melo, R. H. Muni. Morphologic Features of Regulated vs. Dysregulated Rhegmatogenous Retinal Detachment Using Swept-Source Optical Coherence Tomography. – *American Journal of Ophthalmology*, Vol. **268**, 2024, pp. 19-28.
10. Barioulet, L., M. Rueter, V. Sentis, P. Fournié, S. Mrozek, D. Mélite, G. Moulis, A. Sommet, V. Soler. Meteorological Factors and Rhegmatogenous Retinal Detachment in Metropolitan France. – *Scientific Reports*, Vol. **14**, 2024, No 1, 18857.
11. Schöneberger, V., J. Q. Li, L. Menghesha, F. G. Holz, F. Schaub, T. U. Krohne. Outcomes of Short-Versus Long-Acting Gas Tamponades in Vitrectomy for Rhegmatogenous Retinal Detachment. – *International Journal of Retina and Vitreous*, Vol. **10**, 2024, No 1, 16.
12. Hosni, S., H. Khachnaoui, H. M. Zgolli, S. Mabrouk, D. Sidibé, H. Tabia, N. Khelifa. Prediction of Postoperative Visual Acuity in Rhegmatogenous Retinal Detachment Using OCT Images. – *IEEE Access*, Vol. **11**, 2023, pp. 135435-135448.

13. Garneau, J., M. Hébert, E. You, A. Lachance, S. Bourgault, M. Caissie, É. Tourville, A. Dirani. Outcomes of Surgical Repair of Retinoschisis-Associated Retinal Detachment Compared to Rhegmatogenous Retinal Detachment. – BMC Ophthalmology, Vol. **22**, 2022, No 1, 10.
14. Kovács, L., G. Szabó. B-Morpher: Automated Learning of Morphological Language Characteristics for Inflection and Morphological Analysis. – Cybernetics and Information Technologies, Vol. **22**, 2022, No 4, pp. 111-128.
15. Anitha, E., D. J. Aravindhara. Efficient Retinal Detachment Classification Using Hybrid Machine Learning with Levy Flight-Based Optimization. – Expert Systems with Applications, Vol. **239**, 2024, 122311.
16. Yadav, S., R. Murugan, S. Mondal, P. Bhattacharjee. GRD-Net: Graph Neural Network Model for Retinal Detachment Classification through Fundus Image. – In: Proc. of IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES'24), 2024, pp. 35-40.
17. Yadav, S., R. Murugan, T. Goel. Opti-EN-Net: Optimized Ensemble Deep Neural Network for the Classification of Retinal Detachment through Fundus Images. – Biomedical Signal Processing and Control, Vol. **91**, 2024, 105999.
18. Li, H., J. Cao, K. You, Y. Zhang, J. Ye. Artificial Intelligence-Assisted Management of Retinal Detachment from Ultra-Widefield Fundus Images Based on a Weakly Supervised Approach. – Frontiers in Medicine, Vol. **11**, 2024, 1326004.
19. Lake, S. R., M. J. Bottema, K. A. Williams, T. Lange, K. J. Reynolds. Retinal Shape-Based Classification of Retinal Detachment and Posterior Vitreous Detachment Eyes. – Ophthalmology and Therapy, Vol. **12**, 2023, No 1, pp. 155-165.
20. Li, Z., C. Guo, D. Nie, D. Lin, Y. Zhu, C. Chen, X. Wu, F. Xu, C. Jin, X. Zhang, H. Xiao. Deep Learning for Detecting Retinal Detachment and Discerning Macular Status Using Ultra-Widefield Fundus Images. – Communications Biology, Vol. **3**, 2020, No 1, 15.
21. Guo, H., C. Ou, G. Wang, B. Lu, X. Li, T. Yang, J. Zhang. Prediction of Visual Outcome after Rhegmatogenous Retinal Detachment Surgery Using Artificial Intelligence Techniques. – Translational Vision Science & Technology, Vol. **13**, 2024, No 5, 17.
22. Panchal, S., A. Naik, M. Kokare, S. Pachade, R. Naigaonkar, P. Phadnis, A. Bhanghe. Retinal Fundus Multi-Disease Image Dataset (RFMiD) 2.0: A Dataset of Frequently and Rarely Identified Diseases. – Data, Vol. **8**, 2023, No 2, 29.
23. ARMD Dataset (Accessed on July 2025).
24. <https://www.kaggle.com/datasets/saketlad/armd-combined-dataset-fundus-and-oct>
25. Tan, M., Q. Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. – In: Proc. of International Conference on Machine Learning (PMLR'19), 2019, pp. 6105-6114.

*Received: 19.03.2026; First Revision: 12.05.2026; Second revision: 20.05.2026;
Accepted: 25.05.2026*

© 2026 by the authors. Licensee IICT-BAS, Sofia, Bulgaria. This article is an open access article distributed under the terms and conditions of the **Creative Commons Attribution-NonCommercial-NoDerivatives** license.