

Chapter 4

Red Lesion Segmentation and DR Detection

4.1 Introduction

DR is a vision-threatening disease, which progresses silently in a patient with prolonged diabetes. It starts with the weakening of the blood vessels nourishing the retinal layer. These fragile vessels may leak fluids like blood, fats, lipids, etc., in their surroundings. Later, these vessels divide into multiple vulnerable vessels with increased leakage. The term ‘silent’ means that in the early stages, DR is asymptomatic. The disease symptoms start occurring in the late stages when the patient’s vision is blurred, and he starts observing the floating patches in his sight. Such severe cases of DR lead into a permanent damage to the retina, which cannot be reversed [4]. The only solution to combat this disease is regular monitoring and an early diagnosis.

DR has become one of the most significant causes of blindness amongst working-age adults worldwide. In India itself, we have over 13 million people suffering from DR [27], which is a matter of grave concern. Its manual screening demands much workforce and social awareness. Developing countries lag in both of these factors. With a vast population, we need automated screening tools which can help in community-level testing. Also, we need these tools simple enough to be used both by medical and non-medical experts for the initial screening of the disease. Such automated diagnostic

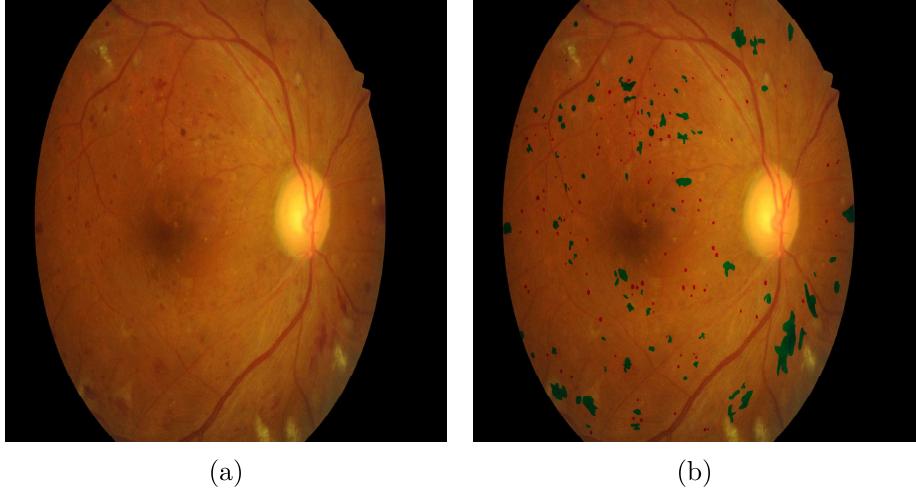


Figure 4.1: *A typical example of presence of red lesions on the retina. MAs and HEMs are marked in red and green color, respectively.*

systems can be a significant help in saving numerous visions.

The design of image-based automated diagnostic systems began with conventional image processing, moving towards deep learning with recent developments. These systems generally start with processing the input image for abnormality enhancement, followed by their localization and segmentation for prediction purposes. As discussed in chapter 1, DR progresses in four stages from mild, moderate, and severe NPDR to PDR [51]. The clinical symptoms of DR, which can be captured using the 2-D fundus images, include MAs, HEMs, HEs, SEs, neovascularisation, venous beading, etc. Out of these symptoms, the earliest symptoms of DR include MAs and HEMs, which are collectively termed as red/dark lesions [86].

On an RGB fundus image, the MAs appear as sharp red dots covering just a few pixels whereas HEMs are comparatively more significant red patches caused due to leakage of blood in the retinal layer [54]. Figure 4.1 shows a typical example of MAs and HEMs on the retina. Figure 4.1 (a) presents a DR-affected RGB fundus image, and Figure 4.1 (b) shows the same image with superimposed labeling for MAs and HEMs in red and green color, respectively.

In order to target a robust DR detection, we focus this work towards a precise extraction of red lesions only, leaving aside other DR-specific abnormalities which appear

in later stages. Here, robust means a DR detection in any stage, especially the early ones. This lesion segmentation is supposed to serve as a communicative, preliminary step towards DR screening.

Here, we present a brief literature review on the detection of DR-specific red lesions (MAs or HEMs or both) using 2-D fundus images. Red lesion detection using fundus images started with conventional image processing techniques like [78], [63], [54] etc. These methods are generally pipeline-based, where the images are pre-processed then analyzed manually to derive handcrafted features for detecting the specific abnormalities. Lazar et al. [78] proposed the MA detection by studying the directional cross-sectional profile of the candidate pixel. Fleming et al. [63] compared various contrast normalization methods and used the watershed transform for MA detection. Kar et al. [54] proposed a lesion detection algorithm in which curvelet transform is used to enhance the red lesions. Quellec et al. [81] used template matching based on adaptive wavelets for detection of MAs on the fundus. Rocha et al. [83] made use of a visual word dictionary, marked by experts, for locating various DR-lesions. The major drawback of these methods is subjectivity. Most of these methods are less generic and tend to fail with new datasets.

After the era of conventional methods, machine learning methods came into existence and became immensely popular within a short span. These methods use exclusively derived (either manually or automated) features to train the machine learning model with the help of labeled or unlabeled data. Niemeijer et al. [67] pre-processed the fundus images by removing the vessels and then used the k-NN classifier for red lesion detection. Antal et al. [66] proposed an ensemble-based MA detection algorithm, which uses adaptive weights for MA ensembles. Derwin et al. [?] presented a Local Neighborhood Differential Coherence Pattern (LNDCP) technique to extract features corresponding to MA segmentation using neural networks. With the increasing popularity of deep learning in diagnostics, state-of-the-art methods are mostly CNN-based. These models are fully automated i.e., the machine learns the features by itself and then classifies the data. Guo et al. [68] proposed the L-seg model, which segments the four kinds of lesions using a single deep framework. Zago et al. [86] proposed a patch-based

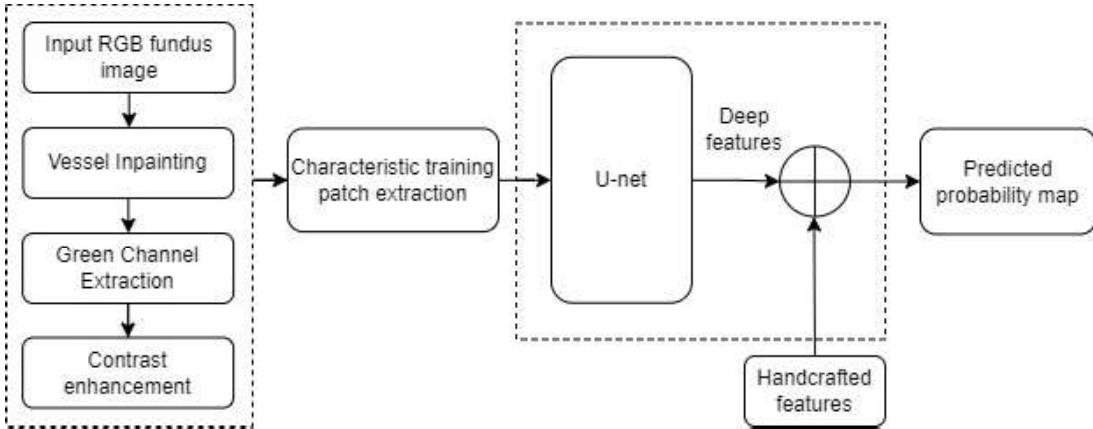


Figure 4.2: *Block diagram of the proposed red-lesion segmentation algorithm illustrating three major blocks: pre-processing, training patch extraction, integration of handcrafted features to the deep model.*

red lesion localization algorithm which further predicts DR. Lahmiri et al. [43] used features extracted by deep CNN and selected best ten features using the Student t-test. These features are fed to SVM for the detection of HEMs. Yan et al. [70] combined the local and global features to extract the local details without losing the contextual information. Tan et al. [65] proposed a ten layer deep architecture for automated segmentation of HEMs, MAs, and exudates. Biswal et al. [61] segmented exudates using a deep M-CapsNet using Expectation-Maximization routing. These methods are more generic but need a lot of data for efficient model training.

For DR detection, the state-of-the-art algorithms are mostly CNN-based image-classification methods [84], [85], etc. Professionals find such diagnosis methods less communicative as there is no lesion localization. There is another class of algorithms which focuses on locating or segmenting single or all kinds of lesions. Such methods are more complex, as there is a need to understand the minute differences in very similar structures (lesions). Rest of the approaches, which are very few in number are based on abnormality localization, segmentation and then disease screening [86], [87], [88], which can be more communicative for both medical professionals and patients.

In the proposed work, our focus lies in a robust DR detection, which means the detection even in the early stages. Eventually, our algorithm targets towards segmentation of early symptoms of DR which are red lesions, i.e. MAs and HEMs combined.

We do not differentiate between MAs and HEMs, which helps avoid unnecessary complexity. The main challenges while targeting the red lesion segmentation are: poor contrast; less sharp boundaries; no fixed size (can cover one to several pixels); no specific region of presence i.e., can be present anywhere on the retina. We target this segmentation by combining deep learning with some expert manual analysis in this work. We use five intensity-based, handcrafted features along with the deep features to improve the model’s performance. Our focus lies in developing a preliminary step for early detection of DR. Eventually, our algorithm does not differentiate between MAs and HEMs, which helps avoid unnecessary complexity. This work also proposes a novel ‘characteristic patch-based’ approach, which uses red lesion-specific patches for model training. Moreover, we inpaint the vessel pixels into the fundus background in data pre-processing. This step brings the network’s focus towards the vessel-free fundus background. The proposed algorithm has performed quite promisingly on IDRiD and DDR datasets. The block diagram of the proposed lesion segmentation algorithm is shown in Figure 4.2. It shows the pipeline of pre-processing block, the characteristic training patch extraction, concatenation of handcrafted features to the U-net derived deep features, and the final prediction.

4.2 Red lesion segmentation and DR screening

Any diagnosis is worthwhile if the disease is still in the curable stage. In the case of DR, where the damage becomes irreversible in the later stages, we require early detection. In this work, we target the detection and segmentation of the DR-specific red lesions, which are reported to be the earliest symptoms of the disease [86].

We start with enhancing the abnormalities in the color fundus image by a few pre-processing steps, followed by a novel ‘characteristic patch based’ training of a U-net model for deep feature extraction. Further, we concatenate five handcrafted intensity-based features to the deep-model for the proposed red lesion segmentation.

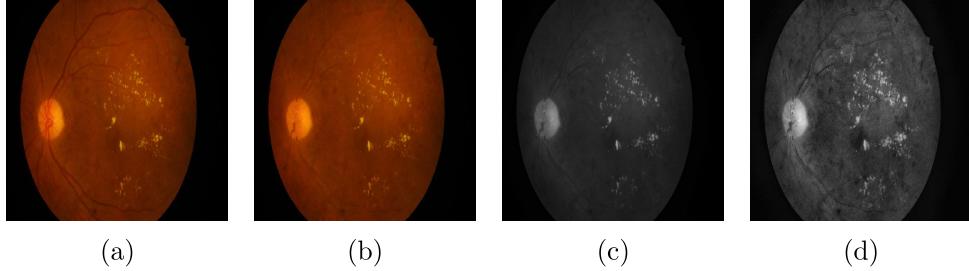


Figure 4.3: *Illustration of pre-processing steps applied for input image enhancement :* (a) Input RGB fundus image, (b) Vessel map inpainting to enhance red lesions, (c) Green channel extraction for best visualization of red lesions on fundus background, (d) Contrast enhancement using CLAHE to obtain final pre-processed image.

4.2.1 Pre-processing

In the RGB fundus images, the targeted red lesions appear dark in gray-scale intensities and, thus, share a similar pixel intensity range as that of the blood vessels. Therefore, we start with a vessel-removal pre-processing step to enhance the dark abnormalities and minimize the false positives. We mask-out the retinal blood vessels and suppress these pixels to facilitate further processing. This is the first pre-processing step. We use the algorithm-III proposed in chapter-3 for the vessel detection. Next, we use an image-inpainting technique [75] to remove these segmented vessel pixels from the fundus background. After removing the retinal blood vessels, we proceed with the extraction of the green channel, as it provides the best contrast in dark pixels and the fundus background. Further, we improve the contrast of the extracted green channel by applying CLAHE algorithm [44]. This step is essential as the red lesions are, generally, found in poor contrast with less-sharp boundaries in the fundus images [68]. Figure 4.3 illustrates the pre-processing steps applied in the proposed algorithm.

4.2.2 Patch extraction

After pre-processing the fundus image, we proceed towards the patch extraction. This work adopts patch-based training despite the global image-based training to handle the data scarcity. The training patches are the sub-images of size $a \times a$ pixels, extracted from the global image to multiply the data using a single image. We extract ' n ' number

of patches per image. Choosing appropriate patches for training the model is a crucial step that influences the segmentation performance. Thus, while targeting the red lesion extraction, we focus on the following two typical properties:

1. The red lesions appear dark in the green channel;
2. These lesions have no fixed location on the retina, i.e., they can be present anywhere on the fundus.

Thus, in order to avoid a poor training of the deep model by using only the random patches, we propose a simple yet effective, lesion-specific patch extraction. We term it as ‘characteristic’ training patch extraction. Keeping in mind the first property, we extract half of the total number of patches from the lower intensity regions of the pre-processed fundus image, which are supposed to have a higher probability of the presence of red lesions. We term these low intensity or dark patches as ‘specific’ patches, denoted by $I_{spec}(x, y)$. The patches with a mean intensity value less than the mean intensity of the green channel qualify for this class, i.e.,

$$\underset{(x,y) \in a \times a}{\text{mean}} \{I_{spec}(x, y)\} < \underset{(x,y) \in H \times W}{\text{mean}} \{I_{green}(x, y)\} \quad (4.1)$$

where, $I_{green}(x, y)$ denotes the green channel image, H and W represent the height and width of the image. The rest half are extracted randomly but inside the FoV of the fundus image, satisfying the second property. These random patches are denoted by $I_{rand}(x, y)$.

Let the set of characteristic patches be denoted by $S_{char}(x, y)$, which is composed of sets of random and specific patches given as,

$$S_{char}(x, y) = S_{rand}(x, y) \cup S_{spec}(x, y) \quad (4.2)$$

where, $S_{rand}(x, y)$ and $S_{spec}(x, y)$ denote the sets of random and specific patches, respectively and given as,

$$S_{rand}(x, y) = \{I_{rand}^1(x, y), I_{rand}^2(x, y), \dots, I_{rand}^{n/2}(x, y)\} \quad (4.3)$$

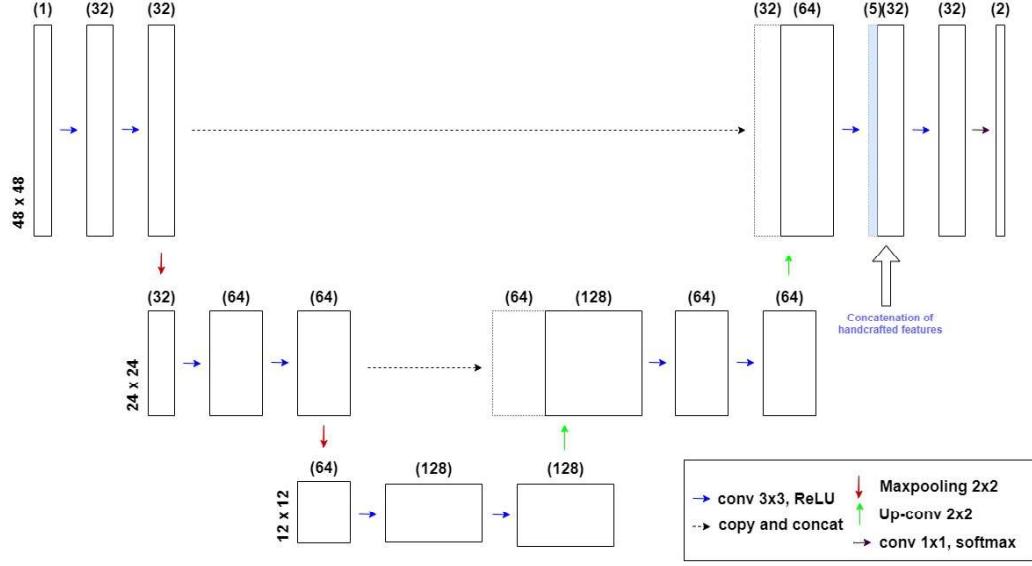


Figure 4.4: Illustration of the deep model used for the proposed red lesion segmentation. Here, we use an encoder decoder based U-net architecture with concatenation of handcrafted feature map in the end.

$$S_{spec}(x, y) = \{I_{spec}^1(x, y), I_{spec}^2(x, y), \dots, I_{spec}^{n/2}(x, y)\} \quad (4.4)$$

We make the set of characteristic patches corresponding to each training image and combine them to train our model. With equal number of random and specific patches, the model understands the fundus images better with a tuned-in focus towards lesion-prone regions. Now, for ‘ M ’ number of training images, we have $M \times n$ number of patches, increasing the data by n folds.

4.2.3 Handcrafted feature extraction

The deep models are supposed to learn the high-level patterns from the data, which are not visible even to the expert medical practitioners. However, a deep learning model may sometimes overlook some crucial physical features that only an expert can see. Thus, we combine a set of handcrafted features with the extracted deep features for the best results.

On the color fundus image, the MA’s appear as tiny red dots with sharp edges.

HEMs appearing because of blood leakage in the retina are red patches of significant size on color fundus images. These lesions appear dark as they acquire a lower intensity range in all the three red, green and blue color channels. The handcrafted features considered here are inspired by these regional intensity variations of the targeted lesions w.r.t. the neighborhood. In the green channel, we observe the difference in minimum, maximum, mean, standard deviation and entropy of neighborhood. The intensity variations in their neighborhood can characterize the presence of red lesions in the fundus image. These features are inspired by the vessel-specific features proposed in a previous work [32]. We experimentally derive this set of five intensity-based features, which are further combined with deep features for red lesion segmentation. These features are discussed in brief as follows,

1. $I_{min}(x, y)$: This feature extracts the intensity-difference between the current pixel and the darkest pixel lying in the neighborhood ‘ N ’. It is expressed as follows,

$$I_{min}(x, y) = S_{char}(x, y) - \min_{(p,q) \in N} \{S_{char}(p, q)\}$$

2. $I_{max}(x, y)$: It extracts the intensity-difference between the current pixel and the brightest pixel of neighborhood ‘ N ’. It is expressed as follows,

$$I_{max}(x, y) = \max_{(p,q) \in N} \{S_{char}(p, q)\} - S_{char}(x, y)$$

3. $I_{mean}(x, y)$: It extracts the intensity-difference between the current pixel and the mean of neighborhood ‘ N ’. It is expressed as follows,

$$I_{mean}(x, y) = S_{char}(x, y) - \text{mean}_{(p,q) \in N} \{S_{char}(p, q)\}$$

4. $I_{std}(x, y)$: It extracts the standard deviation of neighborhood ‘ N ’. It is expressed as follows,

$$I_{std}(x, y) = \underset{(p,q) \in N}{std} \{S_{char}(p, q)\}$$

5. $I_{entropy}(x, y)$: This feature calculates the entropy in the neighborhood ‘ N ’, expressed as,

$$I_{entropy}(x, y) = \underset{(p,q) \in N}{entropy} \{S_{char}(p, q)\}$$

These five handcrafted features are evaluated corresponding to each extracted patch and concatenated with the corresponding deep features for the final prediction (i.e., segmentation).

4.2.4 Training

After the extraction of ‘characteristics patches’ and the corresponding handcrafted features, we proceed towards the deep learning of lesions. Here, we use the baseline deep U-net model [71] for lesion segmentation, with the proposed modifications. U-net is a famous architecture, popularly used with medical images [74], [70], [72], [73] etc. It is an encoder-decoder network with skip connections. The major motivation for choosing U-net is its ‘less’ data requirement. It works well, even with less amount of training data. Also, the skip connections used here, help in retaining the contextual details of the image. It uses multiple 3×3 convolutional layers with ReLU activation function. Further, in the proposed work, while encoding, we down-sample the input using 2×2 max-pooling layers. On the other hand, up-sampling is done while decoding, using 2×2 convolutional layers. Also, we use dropout layers to prevent the overfitting of the model. In the second last layer, we concatenate the above discussed handcrafted feature vector of size $a \times a \times 5$ to the U-net extracted deep features of size $a \times a \times 32$, followed by 1×1 convolutional layer with softmax activation function. Figure 4 illustrates the U-net architecture for the proposed lesion extraction.

Finally, we train the deep U-net architecture with characteristic patches and concatenate the lesion-specific handcrafted features. Here, again, we use Stochastic Gra-

dient Descent (SGD) as the optimizer and binary cross-entropy as the loss function, same as used in chapter 3, given by equation 3.9.

4.2.5 DR Screening

After the prediction of the red lesion probability map, we proceed towards the DR screening. In this thesis, we are targeting the DR screening for which we look for the appearance of any red lesions in the fundus image. In order to decide the presence of red lesions, we apply a gray-scale thresholding step to the predicted probability map. We calculate this threshold value using the conventional, rule-based Otsu's method [47]. This binary map is further used for DR screening. As the proposed method will facilitate a disease screening, we consider maximizing the sensitivity to minimize the false negatives.

Our ultimate aim is a robust ‘DR screening’, which means that we need to classify the input fundus images into two classes: ‘DR’ and ‘No DR’, even in the difficult, early stages. Here, we denote these classes by C_1 which includes healthy cases, and C_2 , which includes the rest of cases like mild, moderate, severe, and PDR. Moreover, in order to infer a prediction from the obtained binary lesion map, we require a deciding parameter. This parameter can be the number of lesions, the area covered in the fundus, regions occupied, etc. Here, in the proposed work, as our target is a preliminary DR screening, we avoid any complexity and calculate the mean area occupied by the lesion pixels. It is denoted by A_{lesion} . This area value is calculated morphologically. As already discussed, the appearance of red lesions is an early symptom of DR, and these can cover one to several pixels in the area. Thus, any non-zero value of the area parameter will classify the input as ‘DR’ (C_2). Only zero value of A_{lesion} will mean ‘No DR’ (C_1).

$$A_{lesion} = 0 \Rightarrow C_1 \quad (4.5)$$

$$A_{lesion} > 0 \Rightarrow C_2 \quad (4.6)$$

The proposed DR screening is based on the presence of any red lesions. Here, we

target the appearance of lesions with an area as small as one pixel wide. This step is supposed to minimize the false negatives, which is crucial for a preliminary disease diagnosis.

4.3 Material and Metrics

4.3.1 Dataset

Today we have several fundus image datasets available in the public domain. However, this work demands pixel-level annotations for MAs and HEMs. As per the best of our knowledge, such labeling is provided only in IDRiD [77], and DDR [53] datasets. Thus, we use these two databases to experiment and validate the proposed work.

IDRiD has a total of 516 color fundus images divided into 413 training and 103 testing images in ‘.jpg’ format. These images are graded for DR severity in five levels. This dataset also provides the pixel-level annotations for four DR-specific lesions on 81 fundus images. This data was made available for a grand challenge organized by the International Symposium on Biomedical Imaging conference in 2018. The other dataset, DDR, has 13673 color fundus images with 4105 test images, graded for DR severity levels. It also provides 757 images with pixel-level annotations for DR-specific lesions.

For both the datasets, we use the given labels for MAs and HEMs to prepare the combined label for the proposed red lesion segmentation. The new label corresponding to red lesions is denoted by $I_{label}^{RL}(x, y)$, given as,

$$I_{label}^{RL}(x, y) = I_{label}^{MA}(x, y) \cup I_{label}^{HEM}(x, y) \quad (4.7)$$

where, $I_{label}^{MA}(x, y)$ and $I_{label}^{HEM}(x, y)$ denote the binary labels corresponding to MAs and HEMs, respectively. Moreover, for screening of DR, we convert the given five-level severity gradation (0-4) into two classes: No DR (0) and DR (1-4).

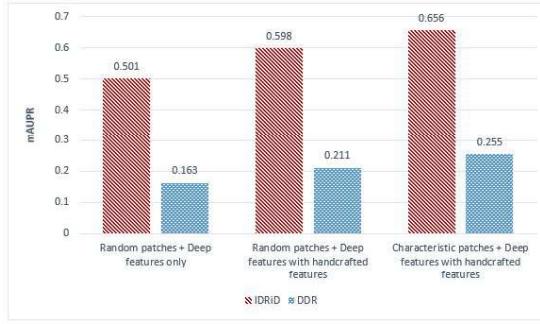


Figure 4.5: *An Ablation study to understand the individual contribution of different steps in the proposed approach. It considers three steps: 1. Random patch-based training of baseline U-net, 2. Random patch-based training of U-net with concatenated hand-crafted feature map, 3. Characteristic patch-based training of U-net with concatenated hand-crafted feature map.*

4.3.2 Performance metrics

Lesion segmentation is eventually a pixel-level classification in which each pixel is classified into lesion or non-lesion classes, which are skewed. Keeping in mind the major data imbalance, we consider the area under the precision-recall (PR) curve, denoted by $AUPR$, for a fair evaluation of the proposed algorithm. Precision measures the fraction of pixels correctly predicted as lesion out of all the lesion-predicted pixels. On the other hand, recall (sensitivity) quantifies the correct lesion predictions out of all the lesion pixels. The PR curve is a plot between precision (Pr) and recall (Re) values, which are given as follows,

$$Pr = \frac{(TP)}{(TP + FP)}$$

$$Re = \frac{(TP)}{(TP + FN)}$$

where, TP , FP , and FN denote true positives, false positives, and false negatives, respectively. We skip the evaluation of metrics like accuracy, specificity, etc., as they will only mislead the analysis in this case.

In this work, we have performed the combined segmentation of MAs and HEMs under the category of red lesions. Thus, for comparison with other methods, we eval-

ate the mean area under the precision-recall curve, denoted by $mAUPR$ values for the MAs and HEMs [68]. It allows us to compare our results with the other state-of-the-art methods.

In addition, to evaluate the performance of the proposed DR screening, we use Sensitivity (or Recall) and area under the ROC curve, denoted by Re and AUC , respectively.

4.4 Experiments and Discussions

This work proposes a segmentation-based DR screening, which is a less followed path in the state-of-the-art. Most of the research groups are primarily working towards gradation of fundus images to diagnose DR. In these methods, the retinal images are classified into various stages of DR. The medical community finds such gradation-based diagnosis comparatively less reliable, as there is no abnormality localization. It is also less communicative for the patient. Thus, segmentation-based screening is supposed to be more helpful.

In order to develop a sensitive DR screening algorithm, this proposal work towards a precise segmentation of red lesions on the retina. These lesions include MAs and HEMs. We use 2-D color fundus images to proceed with the proposed algorithm. For various experimentation, we use the images from IDRiD and DDR datasets. As discussed, these sets have publicly available labels corresponding to lesion segmentation and DR gradation. IDRiD has images of uniform dimension 4288×2848 , whereas DDR has images of variable sizes. We downsample and normalize the images of both the datasets to the size 2144×1424 (by the factor of 2 for IDRiD). Further, as an important pre-processing, we extract the blood vessels using the algorithm-III proposed in Chapter 3. These vessels are then fused in the fundus background by employing an image inpainting technique [75]. The vessel-free fundus background is processed for red-lesion extraction. In order to train the deep U-net architecture, we extract the patches of size 48×48 from the pre-processed image. These patches are extracted randomly as well as specifically from the dark regions of the fundus. Here, dark regions imply the

patches with mean gray-scale intensity less than the mean gray-scale intensity of the pre-processed global image. Next, to account for the appearance and texture of the red lesions, we experimentally find five suitable, intensity-based handcrafted features. For these features, we use a 9×9 neighborhood. These features of size $48 \times 48 \times 5$ are concatenated at the second last layer of the U-net with size $48 \times 48 \times 32$ to generate a layer of size $48 \times 48 \times 37$. For parameter optimization, we use a learning rate of 0.0001 with an SGD optimizer.

In order to analyze the step-wise contribution and significance of various proposed steps, we perform an ablation study to observe the algorithm’s performance. First, we train the baseline U-net model with random patches and observe its mAUPR. Next we consider training the modified model (i.e., model with concatenated handcrafted feature layer) with random patches. Lastly, we train the modified model with characteristic patches. Figure 4.5 presents this study of step-wise contribution via a bar graph. It can be clearly seen that with each step, the segmentation performance improves for both datasets. It signifies that the proposed algorithm has successfully made

Table 4.1: Performance comparison for proposed segmentation of red lesions with other recent methods

Method	mAUPR(IDRiD)	mAUPR(DDR)
Xie et al. [57]	0.468	0.121
Mo et al. [55]	0.379	0.069
Yu et al. [58]	0.423	0.181
Yan et al. [70]	0.614	NA
Chen et al. [56]	0.314	0.206
Guo et al. [68]	0.550	0.232
Guo et al. [82]	0.602	NA
Proposed method	0.656	0.255

Table 4.2: Performance comparison of the proposed method with finalists of IDRiD grand challenge (*Source : <https://idrid.grand-challenge.org/Leaderboard/>*)

Rank	Team Name	mAUPR
1	VRT	0.588
2	PATech	0.561
3	iFLYTEK-MIG	0.530
4	SOONER	0.470
7	SDNU	0.434
NA	Proposed method	0.656

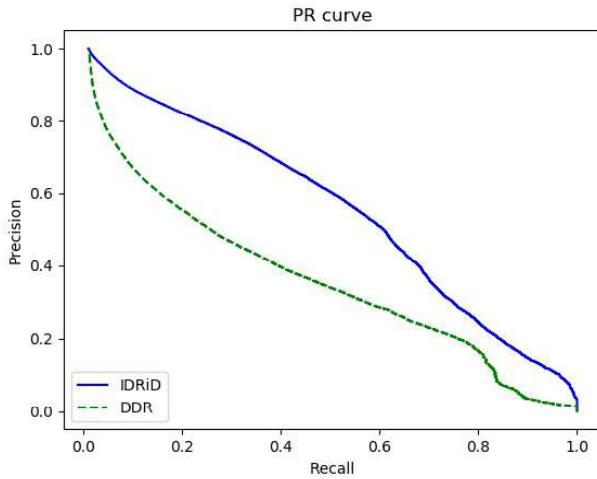


Figure 4.6: The Precision-Recall (PR) curve for red lesion segmentation on IDRiD and DDR datasets.

the baseline U-net model learn the DR-specific red lesions in a better manner. Incorporating the external features improves the model’s prediction by a significant amount (approximately by 19.36% in the case of IDRiD and by 29.44% in the DDR dataset). Moreover, training the model with the characteristic patches further enhances the performance by 9.70% and 20.85% for IDRiD and DDR, respectively.

Table 4.1 compares the performance of the proposed red lesion segmentation method with the other state-of-the-art methods on IDRiD and DDR datasets. It is worth mentioning that most of the other lesion segmentation methods focus on individual segmentation of different retinal lesions, which is not required for an early DR screening. In this work, to aim the minimum complexity while an initial disease screening, we have considered MAs and HEMs, combined as red lesions. Thus, we consider the mean of *AUPR* in the case of other methods to compare with our combined segmentation [68]. The proposed approach improves the *mAUPR* score by nearly 7% and 10% with respect to the previous best performances for IDRiD and DDR datasets, respectively. In another comparison, we consider the segmentation performance of the finalists of the IDRiD challenge. Table 4.2 presents this comparison. It is clear that in terms of *mAUPR*, we have outperformed the other existing methods. The main two factors behind the significant performance of the proposed method are as follows,

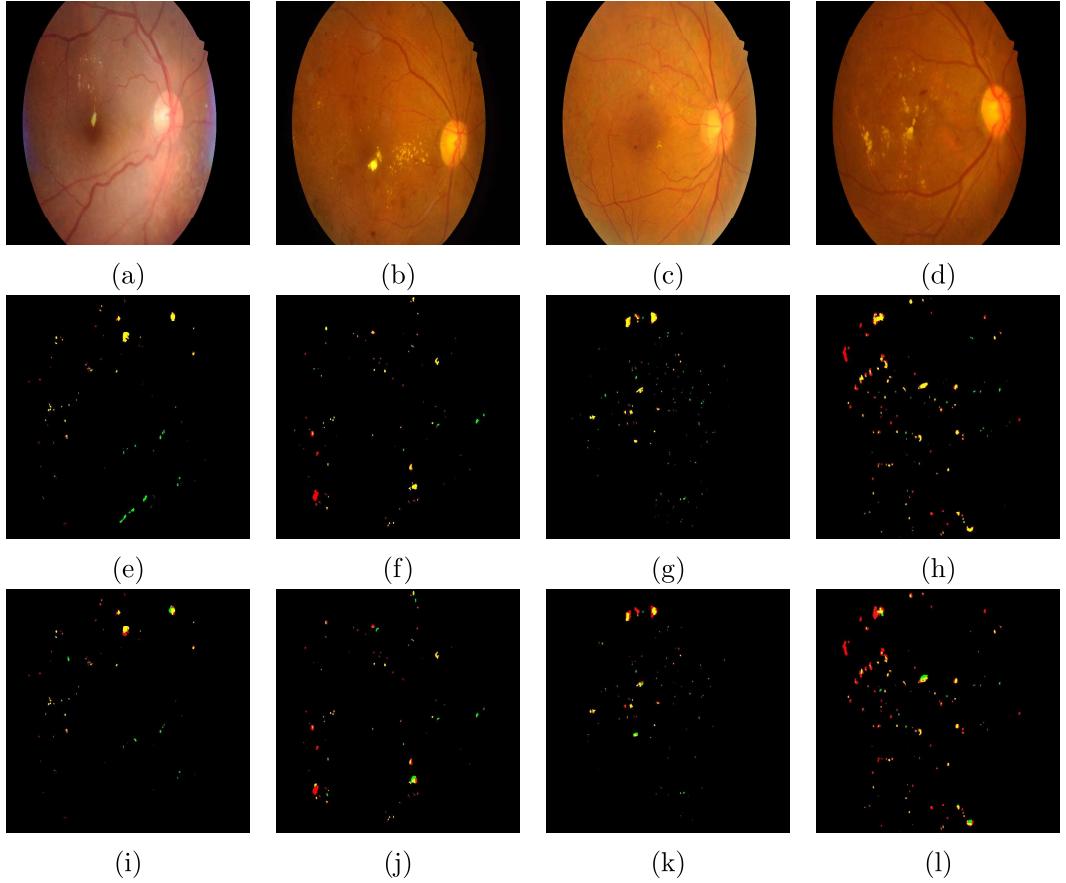


Figure 4.7: Illustration of red lesion segmentation on sample images from IDRiD database. (a, b, c, d): Input color fundus images; (e, f, g, h): Segmentation results for proposed method; (i, j, k, l): Segmentation results for Guo et al. [82]. In the segmentation results, the predicted lesions are shown in green, and ground truth in red color. The yellow pixels show the obtained true positives. The red pixels are the false negatives, i.e. the missed lesions. The proposed lesion segmentation is clearly outperforming the nearest best performing method.

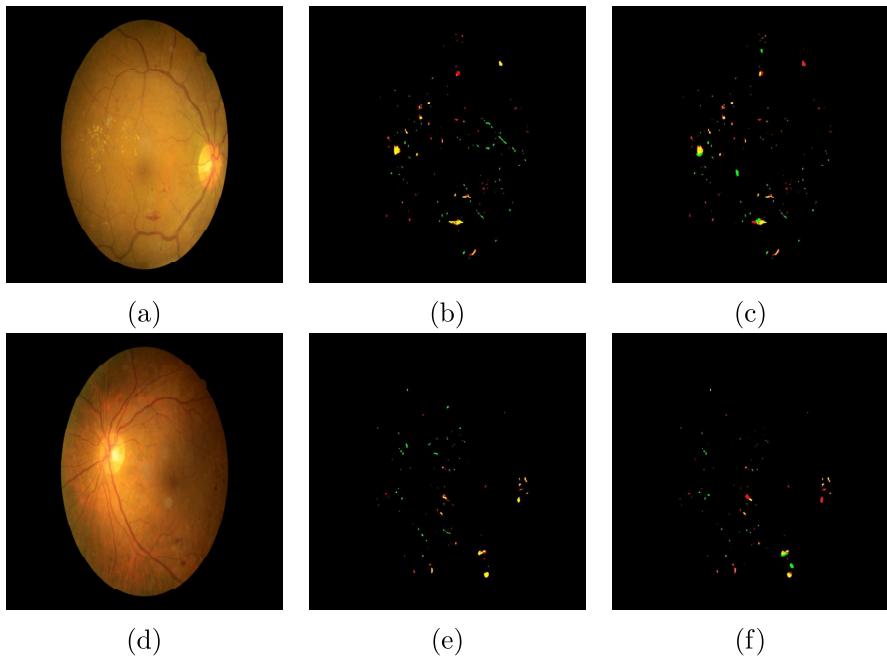


Figure 4.8: Illustration of red lesion segmentation on sample images from DDR database. (a, d): Input color fundus images; (b, e): Segmentation results for proposed method; (c, f): Segmentation results for Guo et al. [68]. In the segmentation results, the predicted lesions are shown in green, and ground truth in red color. The yellow pixels show the obtained true positives. The red pixels are the false negatives, i.e. the missed lesions.

- The characteristic patch-based training makes the model learn over the challenging regions of the fundus. Also, random patch training would have been inefficient and costly in the case of the lesion segmentation, where the target is found in a scattered manner, without any fixed location on the retina.
- Incorporation of lesion-specific handcrafted features to the deep features gives an edge to the model for understanding the region of interest compared to the methods completely relying only on the deep features.

Figure 4.6 presents the obtained PR curves for IDRiD and DDR datasets. Figures 4.7 and 4.8 correspond to red lesion segmentation using the proposed method on IDRiD and DDR datasets, respectively. In the figure, we use an overlap of the segmented lesions over the corresponding ground truth to visualize the obtained false positives and false negatives, along with the true positives. The green channel plots the segmented lesions, and the red channel plots the ground truth. Eventually, the true positives appear yellow (overlap of red and green), the missed lesions or the false negatives as red, and the over-segmentation or false positives as green pixels.

The lesion segmentation or localization is crucial for reliable DR detection. The abnormality localization makes it more communicative for both the medical professionals and the patients. It is quite a tedious job if performed manually by an expert. Thus, such an automated algorithm is supposed to have the potential to serve as a preliminary screening of DR, which can detect DR in the early stages.

In this work, we have successfully segmented the red lesions obtaining 0.656 and 0.255 score values for the mean area under the precision-recall curve for IDRiD and DDR datasets, respectively. With these values, the proposed algorithm outperforms the other best state-of-the-art methods by 7 and 10%, respectively. This improvement can be attributed to the inclusion of handcrafted features and the characteristic patch training in the learning process, which in turn alleviates the issue of data limitations.

After segmentation of red lesions, we take the non-zero value of their mean area as the deciding parameter for DR screening. In literature, there are very few lesion-segmentation-based methods for DR screening. Table 4.3 compares the obtained results

Table 4.3: Comparison of the proposed DR screening with other recent methods

Dataset	Method	Sen	AUC
IDRiD	Orlando et al. [87]	0.912	0.886
	Zago et al. [86]	0.841	0.818
	Proposed method	0.934	0.895
DDR	Orlando et al. [87]	0.900	0.865
	Zago et al. [86]	0.891	0.848
	Proposed method	0.897	0.872

in terms of *Sen* and *AUC* measures with two similar segmentation-based state-of-the-art DR screening methods. The proposed method performs better DR screening than both of these methods on IDRiD datasets. For the DDR dataset, the proposed algorithm has outperformed the other two methods in terms of AUC.

The lesion segmentation or localization is crucial for reliable DR detection. This is quite a tedious job if performed manually by an expert. Thus, such an automated tool is supposed to have the potential to serve as a preliminary screening of DR, which can detect DR in the early stages.

4.5 Chapter Summary

DR-specific red lesions include MAs and HEMs, which are reported to be the early symptoms of DR. Targeted DR screening based on detection of these early symptoms is supposed to be more robust as it enables us to diagnose DR in any stage, including the difficult early ones. This chapter proposes a red lesion segmentation algorithm which is further used for DR screening. While dealing with highly biased data, our method proposes a pipeline to alleviate this imbalance at each step. In the pre-processing stage, we inpaint the retinal blood vessels, which share a similar intensity range as the targeted lesions. It brings the model’s focus on the other dark features of the retina, which are now primarily abnormalities. In another novel procedure, we combine the U-net-based deep features to a set of handcrafted, intensity-based features. This step facilitates the model with some expert manual hints in the right direction. Moreover, we adopt a novel training strategy to train our deep model with ‘characteristic patches’. These patches explicitly train the model to understand the dark regions of the fundus, which

are more prone to either have red lesions or may lead to false positives. The proposed method has intuitively targeted the red lesion extraction, and it has outperformed many state-of-the-art lesion segmentation methods. The extracted lesions have the potential to screen out the cases of DR in the early stages.