

Date: June 14, 2021, date of current version June 14, 2021.

Digital Object Identifier xxxx

Retinal Lesions Segmentation in Fundus Color Images

HUSAM NUJAIM¹, TEWELE TAREKE¹, ANASTASIYA ROZHYN¹, AROJ HADA¹

¹University of Cassino, Erasmus Mundus Joint Master in Medical Imaging and Applications (MAIA)

Supervisors: (1) Prof. Alessandro Bria (e-mail:a.bria@unicas.it) (2) Prof. Claudio Marrocco (e-mail:c.marrocco@unicas.it).

ABSTRACT Diabetic retinopathy (DR) is the leading cause of complete or partial blindness among people with diabetes mellitus. DR can be diagnosed early with the identification of Retinal Lesions on the surface of the eye. However, early diagnosis faces multiple challenges including lack of skilled manpower, knowledge on the disease and complexity in identification because most cases of diabetic retinopathy are identified only when they become severe. Therefore, automatic DR detection from fundus images can offer strong clinical value. We trained our models on the publicly available labeled dataset for DR, the Indian Diabetic Retinopathy Image Dataset (IDRiD). It presents retinal fundus images with pixel-level annotations of four distinct lesions: microaneurysms (MA), hemorrhages (HE), soft exudates (SE), and hard exudates (EX). In this project, we implemented a Machine Learning (ML) and a Deep Learning (DL) framework on Retinal fundus images from the IDRiD dataset to segment the retinal lesions. For ML, the images were first processed using advanced image analysis methods in order to detect candidate lesions and then segmented using ML models including SVM, KNN, LGBM, and Random Forest. Among them, the best classifier was found to be Random Forest with an F1 score of 0.85 for Red Lesions (MA,HE) and an F1 Score of 0.91 for the Exudates (EX,SE). The AUPR Curve for the whole test images of the four lesions in the IDRiD dataset was calculated using the Trapezoidal rule. We achieved AUPR = 0.45, 0.10, 0.06, and 0.21 for EX, SE, MA, and HE, respectively. In the Deep Learning model, we have implemented a HEDNet edge detector to solve the semantic segmentation task on the same dataset, and then segment the DR lesions by incorporating HEDNet into a Conditional Generative Adversarial Network (cGAN). In the DL pipeline, we achieved AUPR = 0.86, 0.40, 0.45, 0.47 for EX, SE, MA, and HE, respectively. The comparison between the two frameworks shows that both pipelines were able to segment the retinal lesions well but as expected, the DL pipeline was able to achieve quantitatively better results among the two.

INDEX TERMS Deep Learning, Machine Learning, Image Segmentation, Diabetic Retinopathy, Retinal Fundus Images

I. INTRODUCTION

DIABETIC Retinopathy (DR), one of the complications of Diabetes Mellitus (DM), is a vasculopathy that affects the fine vessels in the eye and one of the most common causes of irreversible blindness. The world health organization (WHO) estimated that the global prevalence of Diabetes mellitus was approximately 8.8% in 2017 and is expected to further rise to 9.9% by the year 2045. People with diabetes need to have a full eye exam every year by an ophthalmologist or optometrist who is a specialist with diabetic patients. Performing retinal screening examinations on all diabetic patients is of crucial importance in today's public health. However, there are many undiagnosed and untreated cases of DR patients. The demand of automated

analysis of DR images has increased not only due to the scarcity of ophthalmologist, but also to increase the accuracy and efficiency of the diagnosis. The main motivation for this work is to implement an automatic segmentation of 4 types of retinal lesions that are the major signs of DR, microaneurysms (MAs), hemorrhages (HEs), soft exudates (SEs) and hard exudates (EXs). Fig.1 illustrates retinal image (in center) by highlighting normal structures (blood vessels, optic disc and fovea center) and abnormalities associated with DR: Enlarged regions (in left) MAs, and HEs and (in right) SEs, and EXs.

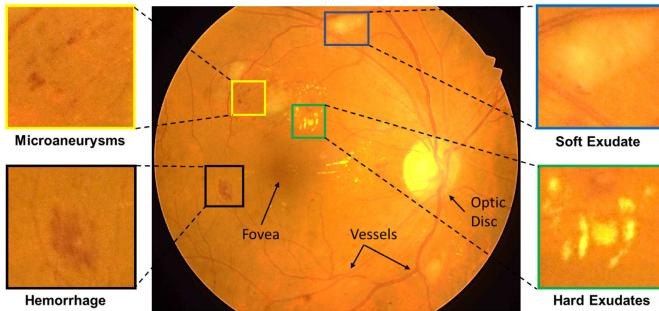


FIGURE 1. Illustration of retinal fundus image and the four typical DR lesions

A. DATASET

The Indian Diabetic Retinopathy Image Dataset (IDRiD) [17] is a well-known dataset created from real clinical exams acquired at an eye clinic located in Nanded, India. The IDRiD dataset provides expert markups of typical DR lesions and normal retinal structures. Fig.2 illustrates some sample images from the IDRiD dataset.

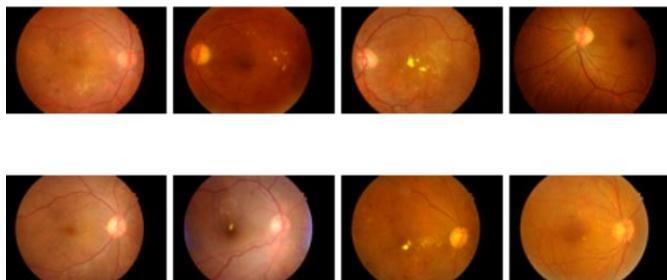


FIGURE 2. Retinal image samples from the IDRiD dataset

The dataset for this work consisted of 81 fully-annotated images (54 images for training set and 27 images for testing). The dataset provides expert annotations for the typical DR lesions and normal retinal structures with pixel level annotations for the four typical DR lesions: microaneurysms (MA), hemorrhages (HE), hard exudates (HEx), and soft exudates (SEx). Fig.3 shows a retinal images with its corresponding ground truth masks for each lesion.

II. LITERATURE REVIEW

The analysis of retinal fundus images using automatic image processing techniques has proven to be an emerging choice to future eye care. Automated techniques in DR screening programs have been introduced, and interesting outcomes were achieved by the rapidly growing deep learning technology which is considered as a potential of AI to revolutionize the future of eye care. The analysis of retinal images was initially kicked off by non-deep learning methods. Handcrafted features based methods involve many stages: a preprocessing stage for contrast enhancement, segmentation, feature extraction, and finally classification. In 2006, Patton et al. [3] presented principles on which the analysis of retinal images is based and explained initial approaches used to detect retinal lesions and landmarks associated with DR. Most

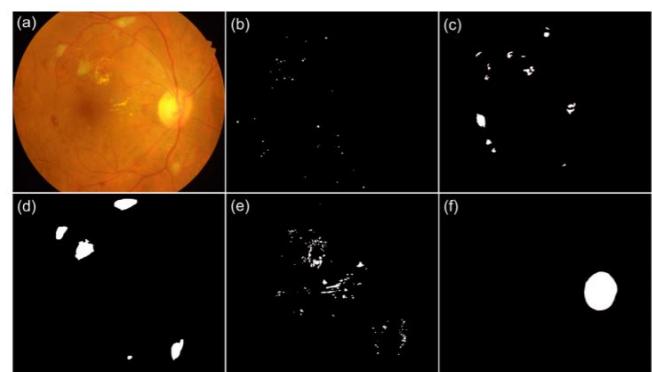


FIGURE 3. Retinal image and different pixel-level annotations: (a) sample fundus image from the IDRiD dataset; sample ground truths for (b-f) MAs, HEs, SExs, EXs and OD respectively

of the segmentation approaches have focused on exclusive segmentation of only one type of lesions, either HEs, MAs, EXs or SExs, from a Fundus image.

Joshi and Karule [19] proposed mathematical morphology for lesion segmentation. Fleming et al. [2] proposed region growing, Wu et al. [13] came up with supervised methods to segment retinal lesions. Apart from these approaches, for the purpose of MAs lesion segmentation, Das et al. [8] used entropy thresholding approach. However, for exclusive segmentation of HEs, Romero-Oraáet al. [18] proposed using super-pixel based method and declared that it is effective. Moreover, for EXs lesions, a clustering approach was presented by Osareh et al. [4]. Furthermore, many techniques were implemented for multiple lesion detection such as machine learning as was proposed by Roychowdhury et al. [6]. The aforementioned techniques provided an automated DR screening outcome since they take advantage of the anatomical structures of the inner side of the eye (i.e. OD and fovea) with lesion detection.

On the other hand, methods based on deep learning have shown more promising results recently. Deep learning refers to multi-layered neural networks able to learn higher-level parameters and low-level representation from data. The most popular methodological variants of deep learning in the field of medical image analysis are Convolutional Neural Networks (CNNs). There are several models of CNN's available in the literature, the most popular models are AlexNet (Krizhevsky et al.), VGG (Simonyan and Zisserman), GoogLeNet (Szegedy et al.) and ResNet (He et al.). Off-the-Shelf CNNs features were utilized by the majority of the approaches in the literature as complementary information channels to other hand-crafted features for detection of lesions associated with DR (Chudzik et al. [14]). Moreover, two-dimensional (2D) image patches were used as an input instead of full-sized images for lesion detection (Lam et al. [16]). García et al. [12] trained the CNN model from scratch and compared it with transfer learning results based on two other existing architectures. In recent years, the automated identification of DR using CNN models has

been improved. However, there are many limitations of these models to achieve a higher performance in the detection of DR due to the imbalance classification. Furthermore, the limited number of samples in the publicly available datasets is another challenge that needs to be tackled .

III. PROPOSED WORK

In this section, we will elaborate on the methods utilized for the segmentation of the Retinal Lesions. The first pipeline includes the usage of handcrafted methods for Image Analysis followed by Machine Learning algorithms for segmentation. The second involves a Deep Learning approach. Fig.4 shows the work flow of the AIA/ML pipeline.

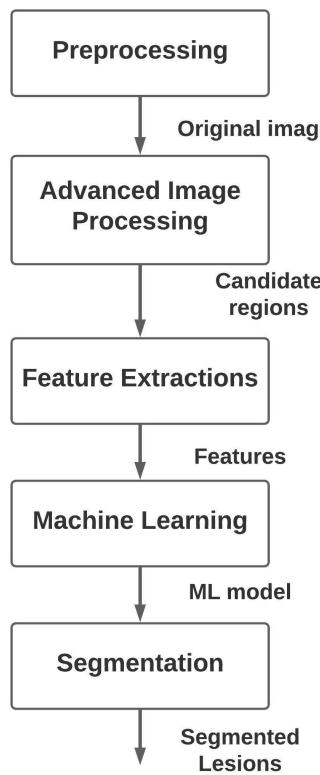


FIGURE 4. Main framework for AIA + ML pipeline

A. IMAGE ANALYSIS AND MACHINE LEARNING

The segmentation workflow starts from preprocessing of the raw images which is vital for image preparation to obtain better segmentation. We have followed two different approaches to extract candidates of the lesions, Red lesions (MA,HE) and Exudates (EX,SE), based on similar characteristics such as color and texture. This pipeline is composed of 2 steps. Firstly candidates extraction using advance image analysis approaches. Secondly, features extraction from these candidates that will be the source of our machine learning dataset (training set).

1) Candidates Extraction

In this step, we mainly used image filtering, thresholding, and grayscale morphology with different kernel sizes to extract candidate objects and therefore partially segmented images from the retinal fundus images. Fig.5,6 illustrate the detailed workflows for the Exudates (EX,SE), and Red Lesions(MA,HE) candidate region extraction using Image Analysis, respectively.

- Contrast Enhancement using Adaptive Histogram Equalization CLAHE: Contrast enhancement ensures the pixel intensities cover a wide range of values, which can make details more readily apparent.
- Denoising: We assume that the images contain Gaussian white noise, and apply the Non-local Means Denoising algorithm.
- OTSU Thresholding: is used to perform automatic image thresholding. In the simplest form, the algorithm returns a single intensity threshold that separate pixels into two classes, foreground and background.
- Morphological Operations: dilation, erosion, opening were used to remove or highlight some details in the images with different structuring elements.
- Other operations such as subtracting, bitwise_or, and bitwise_and were used as well.

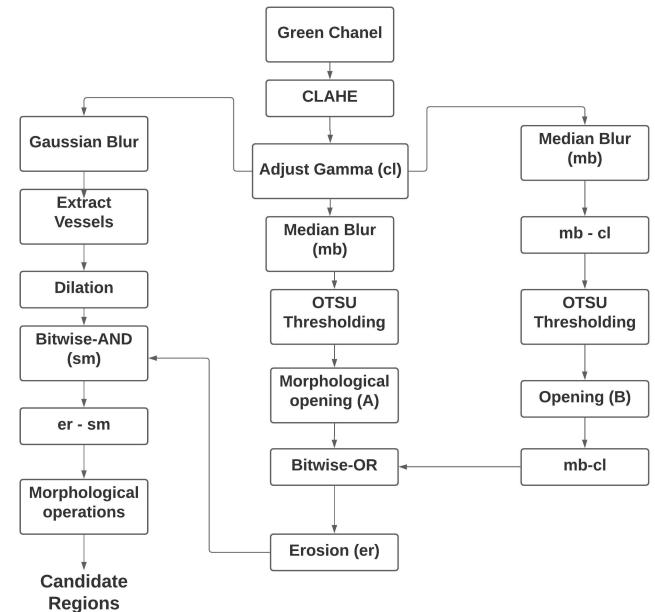


FIGURE 5. Workflow of the Exudates (EX,SE) candidate regions extraction using Image Analysis

Most of the steps for the candidates extraction in the Red Lesions and the Exudates are identical. That is because we were targeting the bright objects. For instance, in the case of Red Lesions segmentation, we used the complementary image after applying Gamma correction to convert the color of the red lesions from dark to bright. Then we followed the

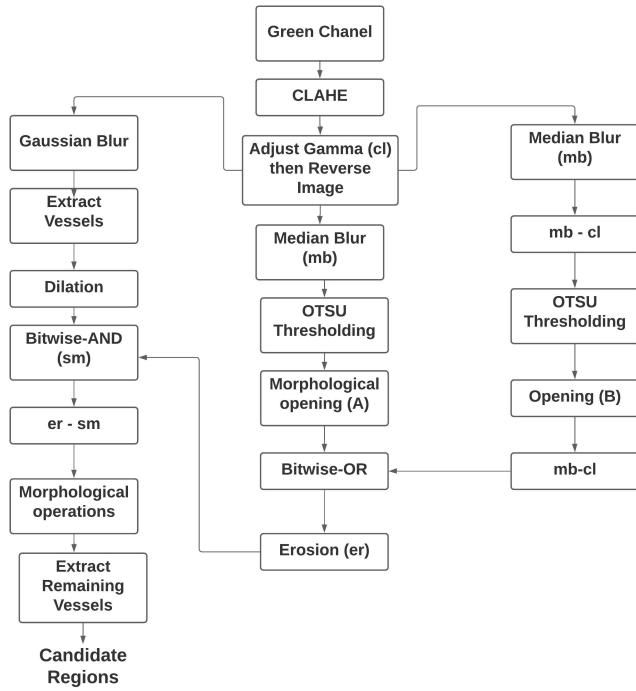


FIGURE 6. Workflow of the Red Lesions (MA,HE) candidate regions extraction using Image Analysis

same steps as with Exudates segmentation. However, the big vessels have to be removed utilizing their size in the last step.

2) Features Extraction

The candidate regions extracted from the original image were then used to extract 338 features from each candidate region. These features include the following:

- Grey Level Co-occurrence Matrix (GLCM): Given an image composed of pixels each with an intensity (a specific gray level), the GLCM is a tabulation of how often different combinations of gray levels co-occur in an image or image section. Texture feature calculations use the contents of the GLCM to give a measure of the variation in intensity (a.k.a. image texture) at the pixel of interest. We measured the Energy, Correlation, Dissimilarity, Homogeneity, and Contrast for different pixels in different positions and angles.
- Local Binary Patterns: labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes

it possible to analyze images in challenging real-time settings.

- Geometric Shape features: The main goal of this method is to find a set of representative features of geometric form to represent an object by collecting geometric features from images. We measured the Area, the Perimeter, and the Eccentricity of each region.
- Geometric Moments: used to provide information about the orientations of a region. We used all the geometric moments with the normalized cenetal moments and the seven Hu moments as well.

3) Machine Learning

Machine learning can used to solve a variety of classification problems, including the automatic classification of DR. [1] The so-obtained features for each candidate patch were then used to train 2 different multiclass ML classifiers, one for the red lesions and another one for the exudates Based on these features, each classifier was trained to classify each region of interest into three classes. For instance, the red lesions classifier classifies each region into either MA, HE, or background (non-lesion). However, the exudates classifier classifies each region into EX, SE, or background (non-lesion).

- Feature Engineering: The dataset that we created which include all the required features was clean and consistent. Therefore, we did not need to perform many feature engineering methods. We just used a Standard Scalar to standardize the observations by removing the mean and scaling to unit variance. This will transform our data such that its distribution will have a mean value 0 and standard deviation of 1.
- Validation Set: We have tried to use K-fold cross validation but we got poor classification performance particularly in the Soft exudates since there were a significant class imbalance. A solution for the class imbalance was to use stratifiedK-fold instead of k-fold but the model failed to detect Soft exudates again. We ended up using 25% of the observations as validatein set and 75% as training set. However, we got poor performance as well since the background observations were tens of thousands yet we had less than a thousand of observations for other lesions. Most of the regions were classified as a background class. We tackled this issue by using under sampling and over sampling approaches.
- Under Sampling: we randomly downsampled the background observations by 20
- Over sampling: we then oversampled the lesions using a minority sampling strategy. Therefore, we achieved a significantly better classification performance.

We trained the model using four different classifiers:

- LGBM (Dart LightGBM): a gradient boosting framework that uses tree based learning algorithms.
- SVM (Support Vector Machine): a set of supervised learning methods used for classification, regression and

outliers detection.

- Random Forest: a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. The sub-sample size is controlled with the `max_samples` parameter if `bootstrap=True` (default), otherwise the whole dataset is used to build each tree.
- KNN: a non-parametric classification method used for classification and regression. In both cases, the input consists of the k closest training examples in data set. The output depends on whether k-NN is used for classification or regression.

B. DEEP LEARNING

In DL part of our project, we built upon an algorithm proposed by Xiao et.al [4]. A Holistically-Nested Edge Detection (HEDNet) network was used to compute a segmentation map from a fundus image. To enhance HEDNet segmentation performance, this model was incorporated into a conditional Generative Adversarial Network (GAN) with a standard PatchGAN discriminator. The addition of an adversarial loss was then used to improve the lesion segmentation performance.[15]

1) Related Work

- **HEDNet in Semantic Segmentation:** Holistically-Nested Edge Detection (HEDNet) [9] is a state-of-art algorithm proposed to solve image-to-image problems with a deep convolutional neural network. Unlike traditional edge detectors, HEDNet can generate semantically meaningful edge maps that identify object contours. Experiments on Berkeley Segmentation Dataset show that HEDNet performs much better than traditional edge detection algorithms like Canny edge detection, and it also outperforms patch-based edge detection algorithms in terms of speed and accuracy [9]. Considering the effectiveness of HEDNet for edge detection and semantically meaningful contour maps, we choose to base our work on top of this architecture.
- **GAN in Semantic Segmentation:** Classification algorithms, such as those for semantic segmentation, perform well when the task has a clearly defined objective. In practice, however, the objective function used often incorporates hidden assumptions that can be overly simplistic. For instance, the classification setting might assume that each pixel belongs to precisely one class, but in reality, a pixel could represent presence of both soft exudates (which occur in the Nerve Fiber Layer of the Retina) and hard exudates (which occur deeper in the retina). Semantic segmentation tasks have been framed as adversarial generative modeling problems[11], where the generative model's objective function is learned. While a HEDNet is used to solve a straight-forward classification problem with the pixel-wise ground truth labels from IDRiD, we can also evaluate how realistic

the HEDNet annotations are. Therefore, the authors present a semantic segmentation on IDRiD as a generative modeling task, and train HEDNet to both classify pixels correctly and generate realistic segmentation maps of typical DR lesions.

2) Methodology

In this section we start by showing how we preprocess the retinal images, then we explain how we use an image-to-image network to segment DR lesions and how we combine the GAN loss to further refine the segmentation results. An overview of the model structure is shown in Figure 8.

3) Processing

The raw images were first subjected to the following pre-processing steps before feeding into the network to enhance the images.

- Uniform Brightness: All images, test and train must have an average, uniform pixel intensity.
- Contrast Enhancement: Contrast enhancement ensures the pixel intensities cover a wide range of values, which can make details more readily apparent. We applied CLAHE (Contrast Limited Adaptive Histogram Equalization) for contrast enhancement.
- Denoising: We assume that the images contain Gaussian white noise, and apply the Non-local Means Denoising algorithm[5]. In addition to this, we apply a bilateral filter to the image, which replaces the intensity of each pixel with a weighted average of intensity values from nearby pixels, thus preserving edge information while the noise is minimized.

4) Segmentation using HEDNet Edge Detector

The lesion segmentation is achieved using a HEDNet (Holistically-Nested Edge Detection Network) architecture. HEDNet builds on VGGNet [7] by adding side outputs to the last convolutional layer in each stage and by removing the last stage and all fully connected layers. The VGGNet structure is initialized with weights pre-trained on ImageNet and then fine tuned. The side outputs are fused together via a trainable weighted-fusion layer.

Diabetic Retinopathy lesions typically make up a very small proportion of a diseased fundus image and do not exist for images of healthy eyes. As a result, the ground truth is highly imbalanced. We use a class-balancing weight , which differentiates cross-entropy loss for positive and negative samples:

$$Loss_{weight-BCE} = (\text{beta} * y \log p + (1 - y) \log(1 - p))$$

where y is the binary indicator 0 or 1 and p is the predicted probability of observation of the positive class.

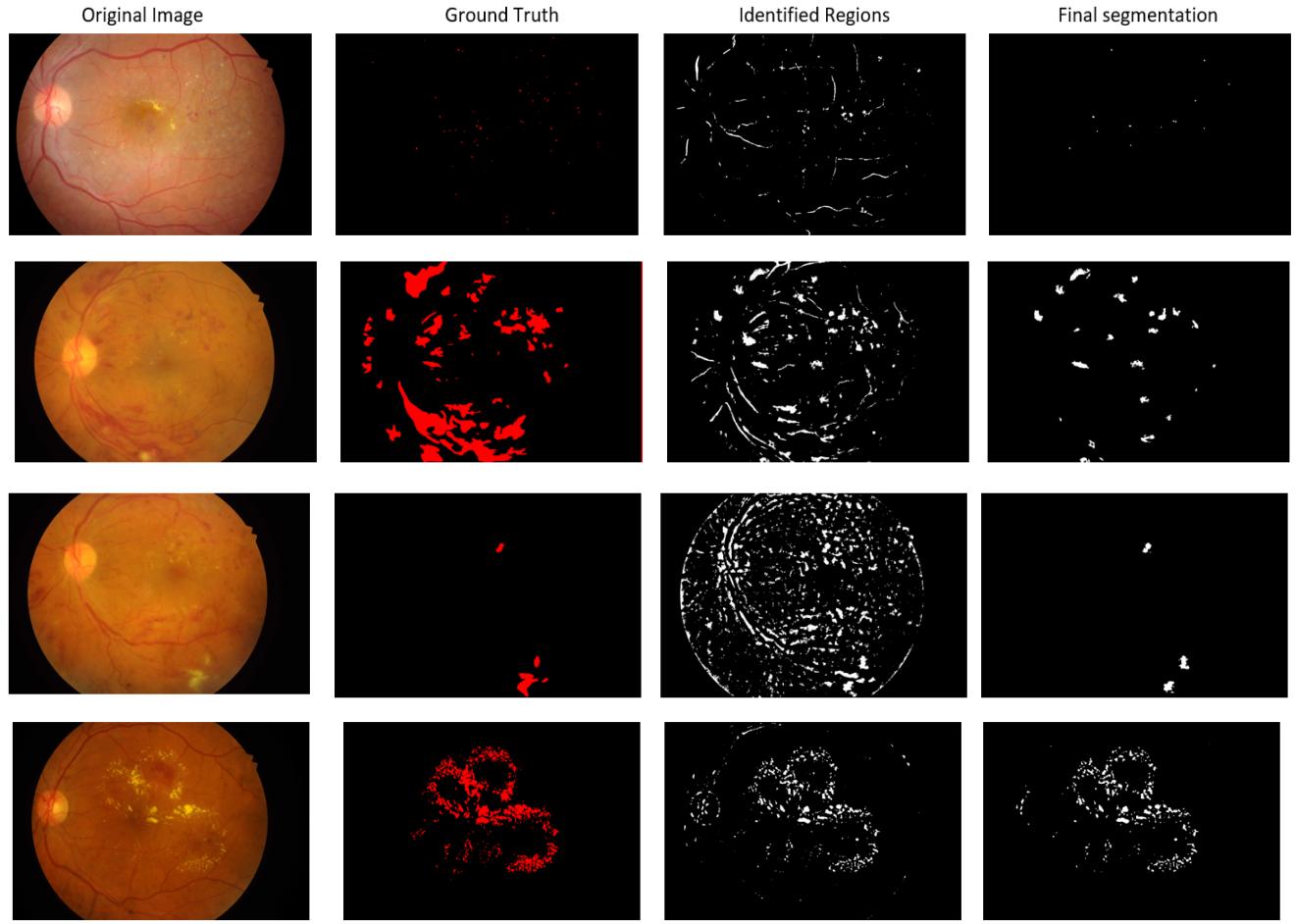


FIGURE 7. Segmentation results using AI and ML lesions for the four lesion types 1) MA 2) HE 3) SE 4) Ex. (a) original image (b) ground truth (c) candidate regions (d) final segmentation

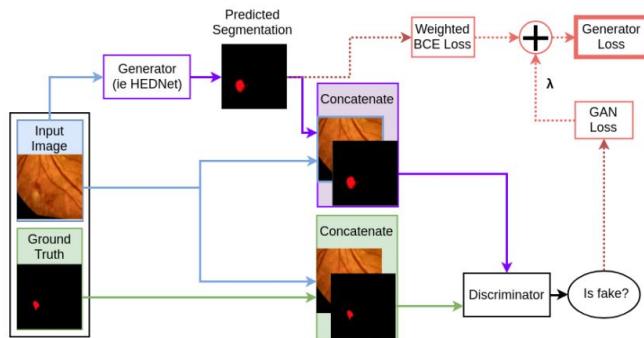


FIGURE 8. Main framework of conditional generative adversarial network

5) Incorporation of HEDNet into a Conditional Generative Adversarial Network (cGAN)

The generator of SSNet is a Global Convolutional Network (GCN), which is inherently an image-to-image fully convolutional network with a large receptive field. Since we want to output an image of the same size with the input image, here we need an equivalent kernel size, therefore we propose to use HEDNet to replace the GCN for diabetic retinopathy.

A conditional GAN is used to discriminate the output, whose architecture is the same as infoGAN. The discriminator utilizes the framework of PatchGAN [10], where the input image is split into smaller patches, and each small image patch is applied with a cross entropy loss to decide whether that patch is fake or real. The input to the discriminator is the concatenation of the original image patch and the corresponding segmentation output from the generator.

6) Loss Function

The generator loss term is a weighted average of binary cross-entropy loss and GAN loss:

$$Loss_{Generator} = Loss_{weighted-BCE} + lambda * Loss_{GAN}$$

The objective of network is to produce good segmentation as well as to make segmentation consistent such that the segmentation result seems real to the discriminator. Therefore, it is used to further refine the segmentation results.

7) Implementation

- Hyperparameters: We use a pixel value in $[0, 1]$ for each lesion image and ground truth segmentation. We

use a patch size of 128 for the SE, EX and HE models and patch size of 64 for the MA model. We set the weight beta() in BCE loss to 10 to balance the positive and negative labels. We set the weight of the GAN loss lambda() = 0:01: We use SGD as our optimizer for both HEDNet and the discriminator with an initial learning rate of 0.001 in both cases. The momentum factor of the optimizer is 0.9 and helps to determine the direction to go. The training and validation batch size is 4 and the testing batch size is 1. For all experiments, the model is trained for 5000 epochs.

- Preprocessing: As mentioned above, For contrast enhancement, we apply the CLAHE technique with tiles of 8 X 8 pixels and a default contrast limit of 40. We have normalized the channel of the images and applied Non-local Means Denoising algorithm with filter 10 to minimizing the noise.
- Data Augmentation: First we crop the image to 512 X 512 pixels and rotate each image using angle of 20°.
- Hardware Used: Experiments for the Deep Learning pipeline were run on a workstation with GPU Nvidia V100 with access from UniCAS (University of Cassino).

IV. RESULTS AND DISCUSSION

A. EVALUATION

Performance evaluation of the models was made using the Area Under the Precision Recall Curve (AUPRC). The AUPRC is calculated as the area under the PR curve. A PR curve shows the trade-off between precision and recall across different decision thresholds. Calculation of AUPRC involve the following steps:

- Assign confidence scores to findings: For each object found in the image, a confidence score was assigned. The model outputs of the ML and DL pipeline were used as the confidence score.
- Match findings with Groundtruth: For each finding/object in the groundtruth, values for true positive ,false positive, and false negative was assigned using the criterion of IoU ≥ 0.5
- Generate detection file : Step 2 is repeated for all the images in the dataset and a detection file is generated that stores all detections as pairs of label and confidence scores.
- Obtain precision-recall curve : A precision-recall curve is plotted with Recall at the x-axis and Precision at Y-axis from the descending sorted detection file generated.
- Calculate the average precision as the area under the precision-recall curve (AUPRC) with trapezoidal method.

B. IMAGE ANALYSIS AND MACHINE LEARNING

Among the four classifiers used, the best segmentation result was achieved by the Random Forest classifier. Table 1 represents the F1 scores from the four different classifiers obtained on the test images in the IDRiD dataset.

Classifier	F1 score (Red Lesions)	F1 score (Exudates)
SVM	0.79	0.84
Random Forest	0.85	0.91
KNN	0.61	0.83
LGBM DART	0.79	0.88

TABLE 1. Results obtained from different classifiers on the train dataset

Random Forest was the best achieving classifier in both red lesions and exudates classification. Fig.10,11 show the confusion matrix of the random forest classifier for the red lesions and the exudates, respectively, where 0 label represents the background, 1 label represents (EX/MA), and 2 label represents (SE/HE).

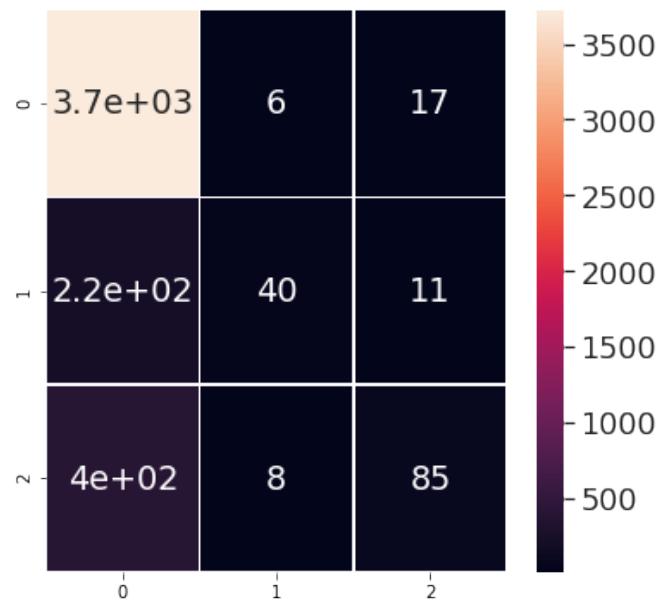


FIGURE 9. the confusion matrix of the random forest classifier (Red Lesions)

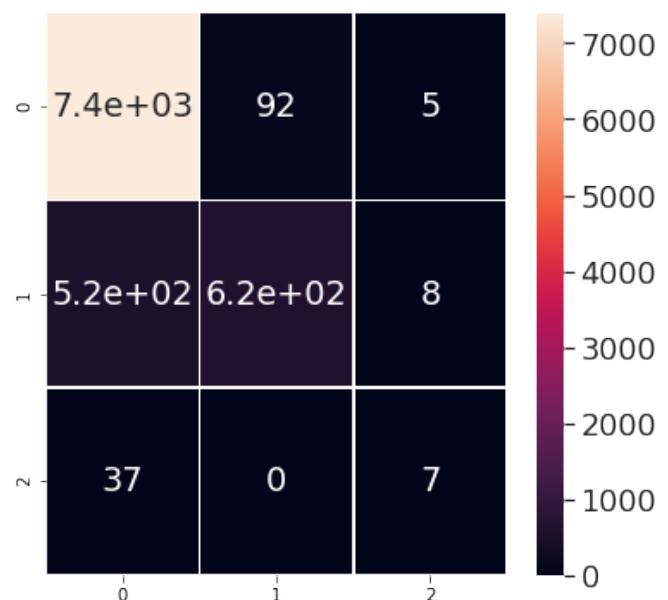


FIGURE 10. Confusion matrix for the random forest classifier (Exudates)

Using the best model obtained, i.e. Random Forest classifier, we then evaluated the test images in the dataset. The metrics achieved are shown in table 2.

Lesion Type	AUPRC
Microaneurysms(MA)	0.06
Hemorrhages(HE)	0.21
SoftExudates(SEx)	0.10
HardExudates(HEx)	0.45

TABLE 2. AUPRC scores achieved on the RF model on test data

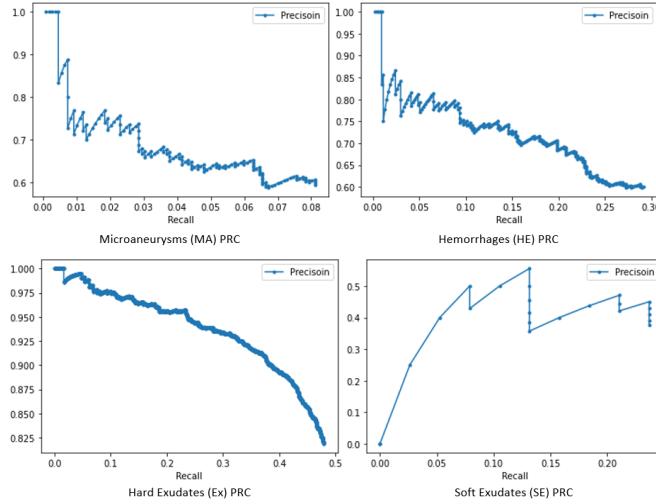


FIGURE 11. Precision Recall Curve of the random forest classifier for the 4 different lesion types

The AUPRC obtained show that the pipeline is not able to segment the lesions very well. The highest score was obtained for Hard exudates (HEx) at 0.45 while the least score was obtained for Microaneurysms (MA) at 0.06. This may be due to the fact that micro-aneurysms are very small, have a lower contrast than other lesions and share a high similarity to blood vessels. Figure 9 presents the Precision Recall Curve of the random forest classifier for the 4 different lesion types.

C. DEEP LEARNING

We compute the score of AUPCR (Area Under the Precision Recall Curve), and plot the PRC for each model on the test images. The proposed model gives a very good average precision result for the Hard Exudates segmentation. The scores obtained for the other three types of lesions were similar with AUPRC score ranging between 0.4 to 0.47.

The results show that the model performs best on hard exudates, where it achieves the highest AUPCR scores. Hard Exudates are small shiny white or yellowish white deposits which appear distinct from the retinal vessels with sharp margins and thus seemed to be easily segmented by the model.

Scores for the four lesions using HEDNet+cGAN model are shown in table 3.

Figure 9 presents the Precision-Recall curve for the four lesion types from the DL framework.

Lesion Type	AUPRC
Microaneurysms(MA)	0.45
Hemorrhages(HE)	0.47
SoftExudates(SEx)	0.40
HardExudates(HEx)	0.86

TABLE 3. AUPRC scores achieved from the DL model on test data

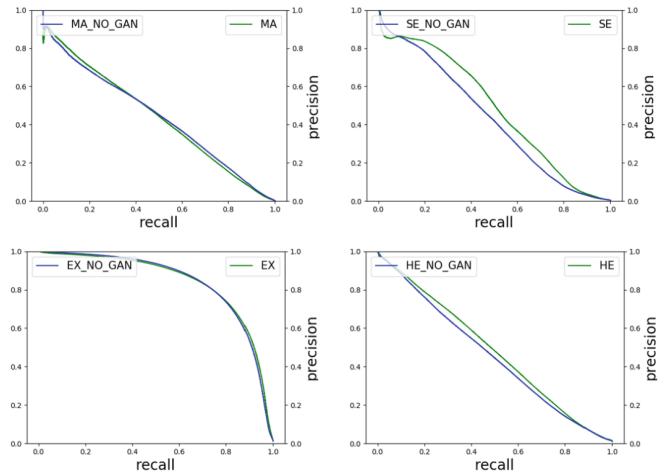


FIGURE 12. Precision-Recall curve for the four lesion types from the DL framework.

D. COMPARISON BETWEEN THE TWO PIPELINES

Finally, we present the results of a comparison between Machine learning and Deep Learning segmentation for the four different lesion types. A quick qualitative comparison show good segmentation results from both the pipelines. Quantitative comparison is made on the basis of Area Under the Precision-Recall curve (AUPR) obtained from the two pipelines; Image processing/analysis and machine learning with Random Forest as the best classification model and Deep Learning model. Table 4 represents the AUPR scores obtained from the two pipelines for the four different lesion types. The scores show significant improvement in segmentation from the DL pipeline over the IP+ML pipeline.

Lesion Type	Image Processing + ML	DL (HEDNet+cGAN)
Microaneurysms(MA)	0.06	0.45
Hemorrhages(HE)	0.21	0.47
SoftExudates(SEx)	0.10	0.40
HardExudates(HEx)	0.45	0.86

TABLE 4. Average Precision on the test dataset for the four lesion types

V. CONCLUSION

In this project, we have implemented two different methodologies for the segmentation of the retinal lesions from color fundus image of the eye ; a image analysis with Machine Learning approach and a Deep Learning approach. As expected, the Deep Learning based method utilising HEDNet incorporated with a cGAN achieved the best results between the two. Among the lesions, best performance was achieved for Hard Exudates with an AUPRC score of 0.86. Usage

of handcrafted methods of image analysis yielded below satisfactory results and lagged behind compared to the DL methods.

References

- [1] Meindert Niemeijer et al. “Comparative study of retinal vessel segmentation methods on a new publicly available database”. In: *Medical imaging 2004: image processing*. Vol. 5370. International Society for Optics and Photonics. 2004, pp. 648–656.
- [2] Alan D Fleming et al. “Automated microaneurysm detection using local contrast normalization and local vessel detection”. In: *IEEE transactions on medical imaging* 25.9 (2006), pp. 1223–1232.
- [3] Niall Patton et al. “Retinal image analysis: concepts, applications and potential”. In: *Progress in retinal and eye research* 25.1 (2006), pp. 99–127.
- [4] Alireza Osareh, Bita Shadgar, and Richard Markham. “A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images”. In: *IEEE Transactions on Information Technology in Biomedicine* 13.4 (2009), pp. 535–545.
- [5] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. “Non-Local Means Denoising”. In: *Image Processing On Line* 1 (2011). https://doi.org/10.5201/ipol.2011.bcm_nlm, pp. 208–212.
- [6] Sohini Roychowdhury, Dara D Koozekanani, and Keshab K Parhi. “DREAM: diabetic retinopathy analysis using machine learning”. In: *IEEE journal of biomedical and health informatics* 18.5 (2013), pp. 1717–1728.
- [7] Karen Simonyan and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition”. In: *arXiv preprint arXiv:1409.1556* (2014).
- [8] Vineeta Das, Niladri B Puhan, and Rashmi Panda. “Entropy thresholding based microaneurysm detection in fundus images”. In: *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*. IEEE. 2015, pp. 1–4.
- [9] Saining Xie and Zhuowen Tu. “Holistically-nested edge detection”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 1395–1403.
- [10] Xi Chen et al. “Infogan: Interpretable representation learning by information maximizing generative adversarial nets”. In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*. 2016, pp. 2180–2188.
- [11] Pauline Luc et al. “Semantic segmentation using adversarial networks”. In: *arXiv preprint arXiv:1611.08408* (2016).
- [12] Gabriel Garcia et al. “Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images”. In: *International Conference on Artificial Neural Networks*. Springer. 2017, pp. 635–642.
- [13] Bo Wu et al. “Automatic detection of microaneurysms in retinal fundus images”. In: *Computerized Medical Imaging and Graphics* 55 (2017), pp. 106–112.
- [14] Piotr Chudzik et al. “Microaneurysm detection using fully convolutional neural networks”. In: *Computer methods and programs in biomedicine* 158 (2018), pp. 185–192.
- [15] Mingyun He et al. “Image Edge Detection Base on Conditional Generative Adversarial Nets”. In: *2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. IEEE. 2018, pp. 18–21.
- [16] Carson Lam et al. “Retinal lesion detection with deep learning using image patches”. In: *Investigative ophthalmology & visual science* 59.1 (2018), pp. 590–596.
- [17] Prasanna Porwal; Samiksha Pachade; Ravi Kamble; Manesh Kokare; Girish Deshmukh; Vivek Sahasrabuddhe; Fabrice Meriaudeau. *Indian Diabetic Retinopathy Image Dataset (IDRiD)*. 2018. DOI: 10.21227/H25W98. URL: <https://dx.doi.org/10.21227/H25W98>.
- [18] Roberto Romero-Oraá et al. “Entropy rate superpixel classification for automatic red lesion detection in fundus images”. In: *Entropy* 21.4 (2019), p. 417.
- [19] Shilpa Joshi and PT Karule. “Mathematical morphology for microaneurysm detection in fundus images”. In: *European journal of ophthalmology* 30.5 (2020), pp. 1135–1142.