Khulna University Studies **Volume 17(1 & 2): 41-55: 2020**

ENGINEERING & TECHNOLOGY

DIABETIC RETINOPATHY LESION DETECTION FROM MULTISPECTRAL RETINAL IMAGES THROUGH NEURAL NETWORK

G M Atiqur Rahaman*, S.M. Riasat Ali, Soma Paul

Computational Color and Spectral Image Analysis Lab,
Computer Science & *Engineering Discipline, Khulna University, Khulna 9208, Bangladesh*

KUS: 20/01-05022020

Manuscript submitted: January 20, 2020 Accepted: March 21, .2021

Abstract : Diabetic Retinopathy (DR) is one of the fastest growing dysfunctions of human retina. Significant research has been conducted using RGB fundus imaging for automatic detection of retinal lesions affected by DR. However, due to only three imaging bands, the accuracy from RGB fundus images is unlikely to improve any further. In contrast to RGB imaging, multispectral imaging has the key advantage of multiple narrow wavelength bands that can be used as spectral features to improve the detection accuracy. Nevertheless, the inter and intra-retinal variation of color, contrast, and illumination is a challenge to process the multispectral images. In this study, a complete framework is proposed to develop and evaluate methods for automatic detection of DR lesions. A multispectral retinal image database, DIARETSPECDB1, is investigated in order to detect the most common DRs such as Microaneurysms (MA), Hard Exudates (HE) and Hemorrhages (HEM). The reflectance values of the spectral bands are used as features of a three-layer basic neural network (NN) to determine the baseline performance of multispectral data instead of any advanced model. According to the results, the model outperforms existing technique producing overall accuracy 94.5%, and the obtained specificity/sensitivity is 0.95/0.89, 0.97/0.89, and 0.88/0.84 for MAs, HEs and HEMs, respectively.

Keywords: Detection, Classification, Multispectral Image, Neural network, Diabetic Retinopathy

Introduction

Diabetic Retinopathy (DR) refers to all the problems introduced by diabetes mellitus in human retina. It is a progressive disease that is often responsible for loss of vision. Diabetes mellitus results in high blood sugar which gradually makes the vascular system fragile (Gardner *et al.*, 1996). The unexpected or abnormal lesions on retinal surface thus appear due to abnormal permeability, leaking fluids (protein, lipid), and bleeding vessels (Mookiah *et al.,* 2013). Fig.1 demonstrates some common DR such as Microaneurysms (MA) that are small red dots of blood, Hemorrhages (HEM) that are blood depositions over a relatively large area, and Hard Exudates (HE) that are fluids of protein, lipid etc. of yellowish/ yellowish-white color. These abnormal lesions are significant to be detected at an early stage to prevent the vision loss.

Conventionally, the ophthalmologists investigate the retina through ophthalmoscopes and try to visualize if there exists any of such DR lesions. Obviously, the color of the lesions is a decisive factor in this case. This diagnosis needs time and precise skills and experience of medical experts.

*Corresponding Author: <gmatiqur@gmail.com>

However, the number of expert ophthalmologists, especially in developing countries is insufficient as the number of population get affected by DR rapidly increases (Mookiah et al., 2013). To this end, computer-aided diagnosis through imaging system is considered an effective non-invasive solution to this problem (Styles et al., 2006). Most of the literatures dealing with automatic detection and classification of DRs are relevant to conventional 3-channel RGB color imaging (Acharya *et al.,* 2009).

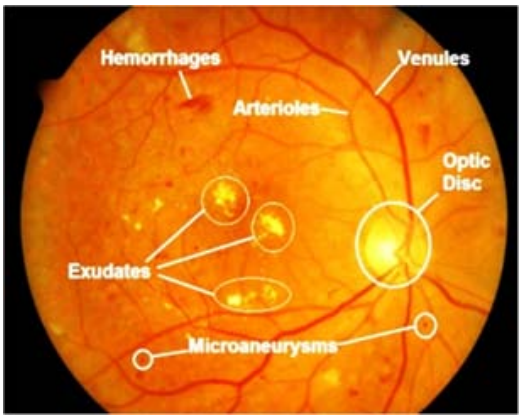

Figure 1. Some common diabetic retinopathies: HEM, MAs, and HEs (Source: Muhammad, Fraz& Barman, 2016).

The RGB images need substantial processing to use shape, color and morphological features to detect the DR lesions. Image segmentation, edge detection, shape and texture analysis are the common operations (Vallabha *et al.,* 2004). For texture analysis, several filters from Gabor filter bank were used after transferring the image in Fourier domain. In (Sánchez *et al.,* 2004) the authors detected HEs in fundus RGB images using the color and sharpedges of the affected lesions. They subsequently detected the optic disc, yellowish objects, and objects with sharp edges and finally combined with previous findings to detect the affected area. After processing the RGB image with a normalization function and contrast enhancements, the MAs were also detected.

In addition to histogram equalization in (Nayak *et al.,* 2008) the authors applied morphological filters to detect the HEs. The filter set consists of diamond and disc elements. A machine learning technique was applied to detect HEs, Cotton-wool spots and Drusen in (Niemeijer *et al.,* 2007). Although Drusen is not a DR lesion, they classified Drusen separately as to avoid confusion with HEs. HEMs and HEs were detected by (Gardner, Keating, Williamson, & Elliott, 1996) concluding that different kinds of processing should be applied for each type of retinal lesions to achieve a desirable accuracy. However, in most of the studies either bright lesions such as HEs (Sánchez *et al.,* 2004) or red lesions such as MAs and/or HEMs (García *et al.,* 2010) are detected. Very few studies have addressed both bright and red lesions together (Ravishankar, Jain, & Mittal, 2009). The main reason is the limited color information of RGB images and the image processing techniques that is non-trivial to generalize for all the DR lesions.

However, due to the limited color information of RGB imaging the detection accuracy cannot be optimal, and thus multispectral imaging is proposed as an alternative technology (Kauppi et. al., 2013; Kauppi, 2010). The inherent limitation of RGB imaging is addressed by a series of narrow and contiguous band that captures more details in each pixel of the scene. Fig. 2 illustrates that in multispectral imaging usually the range of visible wavelength (400-700 nm) is divided into a series of multiple narrow bands and captured the scene for each band, whereas RGB imaging records only

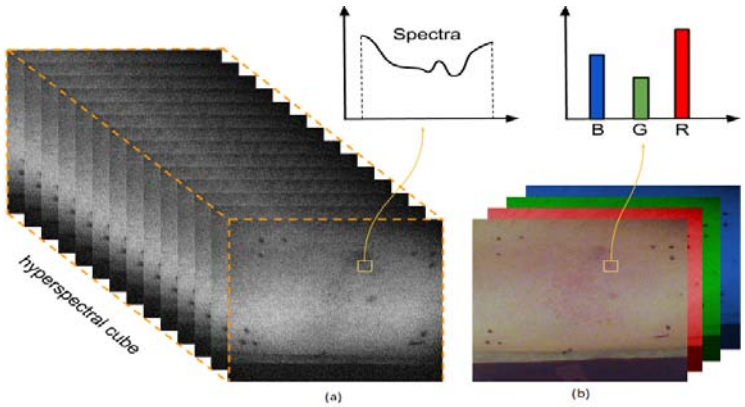

Figure 2. (a) Multispectral information of a pixel (b) RGB information of the same pixel (Source: Liu & Zerubia, 2015)

three aggregated values representing Red, Green and Blue color channels. Note that biological lesion absorbs light based on the composition of the lesion and reflects accordingly. So the reflectance varies in some wavelength bands as the lesion type changes. In Fig. 3, three spectra of three different lesions of the retina are presented. Notice that MAs and HEMs exhibit similarities in color in the RGB image, but, the reflectance spectra are different.

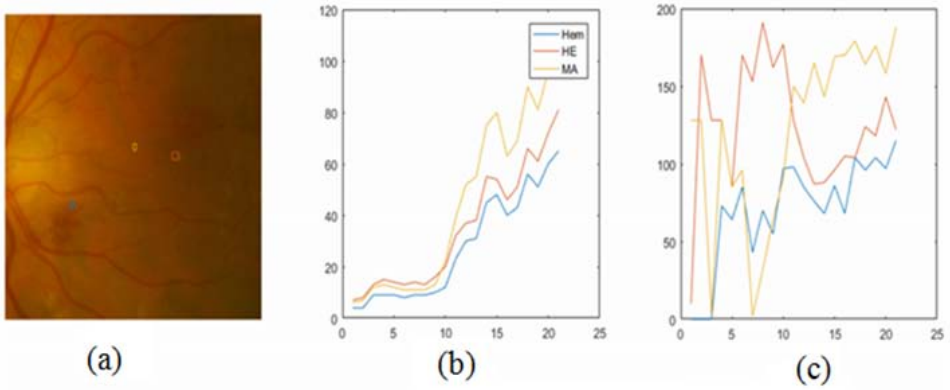

Figure 3. Reflectance spectra of the pixels of DR lesions: (**a**) the corresponding RGB image (**b**) the spectra of the marked regions before pre-processing, (**c**) the spectra after pre-processing

Considering the rich information, and non-invasive characteristics, multispectral technique has been applied in different areas of medical science (Yu *et al.,* 2008; Lu, & Fei, 2014; Koronyo-Hamaoui *et al.,* 2011; Soliz, Truitt, & Nemeth, 2001), however, very few have experimented with retinal imaging to detect and classify DR lesions (Styles et al., 2006; Rahaman *et al.,* 2013). The very initial attempt (Soliz *et al.,* 2001) reported that multispectral bands clearly represent the patterns of a particular retinal tissue. The scope of applying multispectral imaging in retina was extensively described in the report (Styles *et al.,* 2006), who attempted to detect HEMs. In (Rahaman *et al.,* 2013) the authors used multispectral images to detect and classify six types of DR lesions. A stochastic approach was applied addressing the fact of overlapping reflectance spectrum of different class of DR lesions.

Here, we present a technique to detect various DRs within the same framework using the multispectral reflectance spectra. The proposed approach eliminates the laborious feature engineering and/or image processing tasks, and yields high classification accuracy. Through this study we also explore the effectiveness of color reflectance spectrum as the single feature type to detect the DR lesions. We show that without applying the core image processing techniques, the method can be generalized for different classes of DRs. We chose to apply a basic Neural Network (NN) instead of advanced learning model to investigate the robustness or baseline performance of machine learning techniques using reflectance spectra as the input features. MAs, HEMs, and HEs are the abnormal lesions of the retina that are addressed with equal importance, and following the same framework.

Materials and Methods

Image Database: The studied retinal multispectral image database named DIARETSPECDB1 was created by a joint collaboration project between University of Eastern Finland and Kuopio University Hospital (Fält *et al.,* 2009). The database contains 71 multispectral and respective RGB images of human retina of both diabetic and non-diabetic patients. Among them 55 images contain retinal lesions that were affected by diabetes. The original spatial resolution of each image was 1024 x 1024 pixels. The number of spectral bands was 30 starting wavelength at 400 nm and ending at 700 nm with a bandwidth of 10 nm. Three expert ophthalmologists from Kuopio University Hospital annotated the areas of DR lesions on the images (Rahaman *et al.,* 2013). The annotated images were considered as ground truths for training and testing phase of this study. The full details of the dataset including the description of multispectral camera are available at (Fält *et al.,* 2009).

There exists some noise or dead pixels around the spatial borders of the images; therefore, an equal area was discarded from each side of the images. As an illustration Fig. 4 shows one random image of the database after discarding the original borders. The final spatial resolution was 801 x 801 pixels of each image. Also note that due to substantial amount of noises and very low contrast the first 06 spectral bands (400 nm - 460 nm) and last 03 bands (680 nm -700 nm) were ignored. As a result, in total 21 spectral bands (470 nm - 670 nm) are considered in this study.

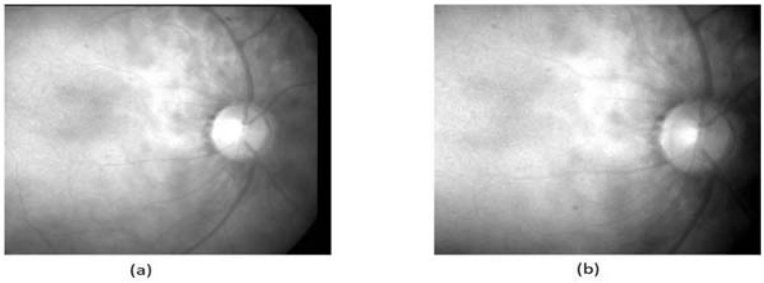

Figure 4. Cropping spatial boundary area: (a) Original image; (b) after cropping.

Image Pre-processing: Image pre-processing was applied globally in order to reduce the intra and inter-retinal color differences, to make both local and global contrast uniform, and to correct uneven illuminations. At first, a median filter was applied with a kernel size of 3 x 3 pixels to remove salt and pepper nose if exists. Due to this filtering, the images became smoothed as well (see Fig. 5). Each image was divided into small sub-images/patches of size of 267 x 267 pixels. To make intensity range equal for all images, intensity adjustment was applied on every image. This intensity adjustment makes intensity range of all images between 0 to 255. One percent (1%) of pixels of each image was

made saturated at both low and high intensities of input image to increase the contrast of the output image.

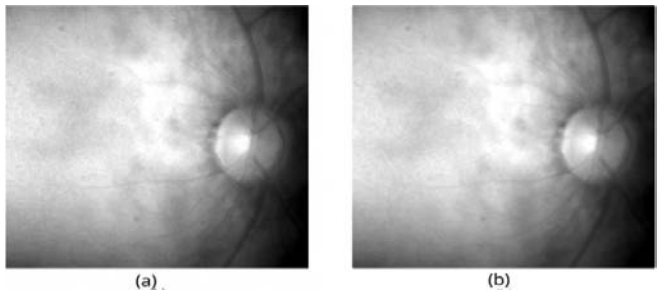

Figure 5. Result of median filtering: (a) before applying; (b) after applying.

After some trials, an ideal sub-image was selected containing samples of all DRs of interest. All other sub-images were further processed with reference to the selected ideal sub-image. Histogram specification was accomplished to make all the sub-images almost equal in contrast. Let H_x and H_y is the histogram of a random sub-image and the reference image, respectively; and H_{cx} and H_{cy} represents their cumulative histogram. The equations for H_{cx} and H_{cy} are as follows;

$$H_{cx}[j] = \sum_{i=0}^{j} H_x[i] \quad (1)$$

$$H_{cy}[j] = \sum_{i=0}^{j} H_y[i] \quad (2)$$

A lookup table was constructed for mapping the intensity of an input image. For every intensity of an input image, the best match with the reference image was found using respective H_{cx} index and the index of H_{cy}. The mapping was accomplished using the following equation:

$$H_{cx}[j] = H_{cy}[i] \textbf{such that} |H_{cy}[i] - H_{cx}[i]| \text{ is } \textbf{minimum} \quad (3)$$

Through the pre-processing operations, both local and global color and contrast variations were minimized. For contrast enhancement, the lesions become more visible that helps more accurate detection. To demonstrate the effect, two patches from two different retinal images are presented in Fig.6.They exhibit a clear difference in contrast in Fig. 6(a) and Fig. 6(b), but after pre-processing, this difference was minimized as shown in Fig. 6(c) and Fig. 6(d). Fig. 3(c) shows the effect of image pre-processing on the reflectance spectra of the pixels originated from the DR lesions.

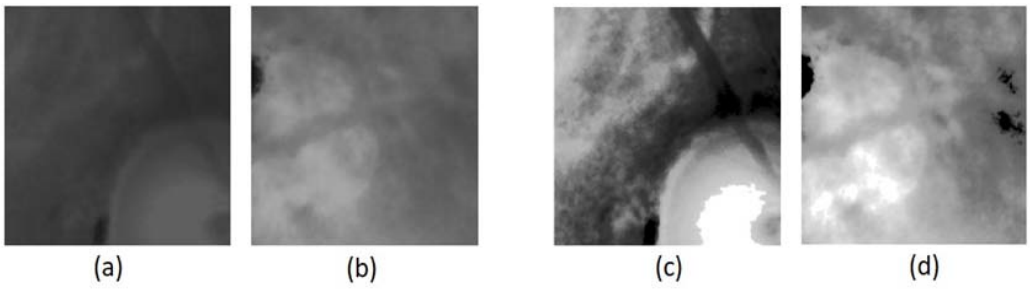

Figure 6. Example images: (a, b) before pre-processing; (c, d) after pre-processing; (a) and (b) are sub-images of two different images. They show quite different contrast, but after pre-processing contrast difference is minimized.

Training and Testing Image Data: Half of the number of available images containing a particular retinopathy lesion was randomly chosen for training and the rest half for testing the method. Among the 43 HEMs images,we selected 21 images for training and 22 images for testing; among the 11 HEs images, 06 images were selected for training and 05 images for testing; 44 images contain MAs among which 22 images were selected for training and 22 images for testing. From each of the training image, 50% pixels were randomly chosen for a particular retinopathy class. For the healthy background, pixels were randomly collected from a variety of images. Later the pixels were merged and formed the training data set, while all others were reserved for the test set. A fraction of the training data set was reserved to be used for validation. Through the validation procedure, the training performance was evaluated after a regular interval, and data over-fitting or under-fitting was examined and resolved. The reason was that the number of pixels in the target lesion class was extremely imbalanced. There are 55 images containing DR lesions, and every image is of size 801 x 801 pixels, so 55*801*801= 3,52,88,055 pixels were available. Among these pixels, almost 97.26% were healthy background pixels, 2.10% of pixels were of HEMs, 0.55% pixels were of HEs, and 0.09% pixels were of MAs. Thus to avoid the biasing problem during model training, random over-sampling was performed for every class except the healthy background class.

Neural Network Architecture: In this experiment, we applied a basic artificial Neural Network (NN) model as a classifier to detect the DRs. It should be noted that as one of the objectives of this report was to study the strength of spectral reflectance color as a discriminating feature, we applied a general three-layer NN model, rather than an advanced model to detect retinopathy lesions. In the NN model, a learning method and an error measurement function were embedded for training. After training, a network was built to test the data. A *scaled conjugated gradient* function was used as a training method and *cross-entropy* was used for error calculation. In this approach, the best parts of the model-trust region known from the *Levenberg-Marquardt algorithm (see Appendix I)*, were combined with a conjugated gradient function (Møller, 1993). Table 1 contains the basic parameters of the model.

Table 1. Specifications of Neural Network Parameters

Parameter	**Description**
Layer	3
Number of hidden neurons	100
Network type	Feed forward back propagation
Training function	Conjugated scaled gradient
Performance	Cross entropy
Transfer function	Sigmoid

As input, the reflectance values from 21 bands are fed for each pixel with the class label defining the corresponding DR lesion. The model can be divided into three parts:

1. Input: Pixel wise spectral reflectance. So there were total 21 nodes in the input layer. When a pixel of a class is trained or tested, these 21 nodes receive input of 21 respective imaging bands.
2. Hidden layer: In the hidden layer of this architecture *sigmoid* activation function is used. Total number of nodes in this layer was 100. The hidden layer makes a binary decision on the output depending on the value of the activation function. The formula for *sigmoid* function is given on Eq. 4.

$$F(x) = \frac{1}{1+e^{-x}} \quad (4)$$

3. Output layer: The output layer has 04 nodes for each of the classes: background, HEM, HE and MA lesions. The logistic function *softmax* is used in the nodes of this layer. *Softmax* is basically a probabilistic function that calculates the probability of a lesion to be one of the four predefined classes using Eq. 5. The calculated probability is in the range 0 to 1 and total sum of the probability is 1.

$$F(x_i) = \frac{e^{x_i}}{\left(\sum_{j=1}^{k} e^{x_j}\right)}, i = 0,1,2,\ldots\ldots,k \quad (5)$$

Experiments: A complete pipeline of the proposed approach is shown in Fig. 7. We consider any retinal image has four classes of lesions: healthy background, HEM, HE and MA. The NN model was first trained by feeding the labeled training dataset. To avoid over- or under-fitting a performance curve was monitored. At a regular interval, the validation data was used to evaluate the training performance. After few trials during the training a curve was generated that show consistent performance as presented in Fig. 8. Training was accomplished by completing a total of 1000 epochs. Initially, the training performance was not satisfactory, but it improved as the epochs increased. After approximately 700 epochs, the performance became stable, but we continued 1000 epochs to ensure high level of confidence. The training accuracy was represented by the Receiver Operating Characteristics (ROC) curve (Fig. 9.) that was generated by plotting the true positive rate against the false positive rate for each class**.**

After the training, the test data were fed to the trained module in order to classify the unknown pixels. The results of the NN model were then compared against the ground truths in order to calculate different performance metrics as described in the following subsection. As an example, Fig. 10 illustrates three sub-images in RGB format with the ground truths and the corresponding image with predicted lesions: MAs are marked by Blue color, HEs are marked by Green color and HEMs are marked by Red Colors. To compare the performance of the multispectral images with the conventional RGB images the same procedure was followed for corresponding RGB images as separated in training and test image data set.

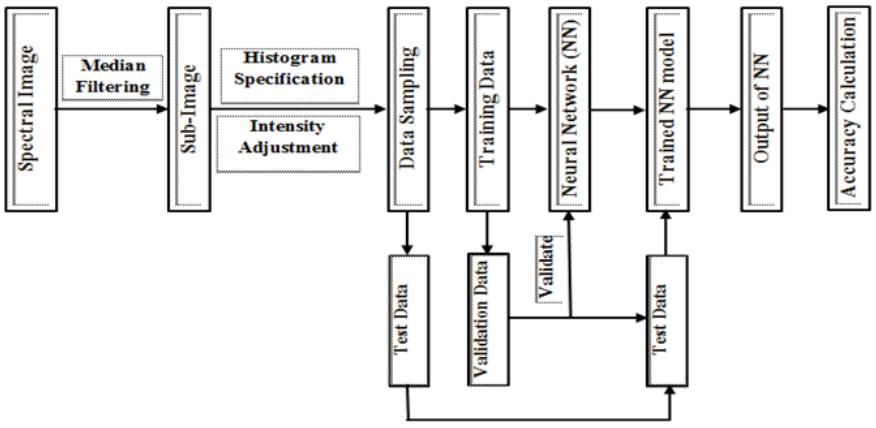

Figure 7. The proposed framework to detect DR from multispectral images

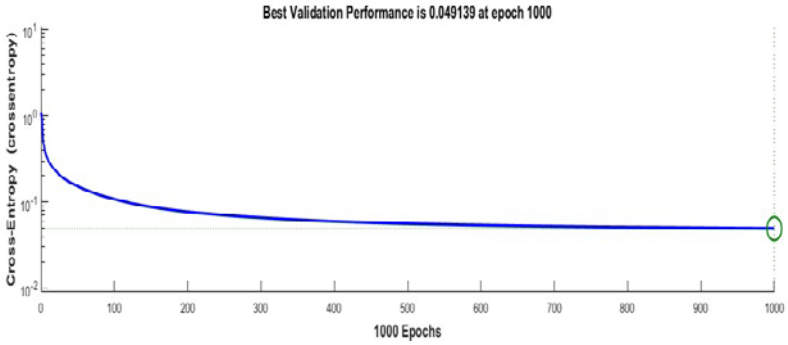

Figure 8. Neural network training performance curve.

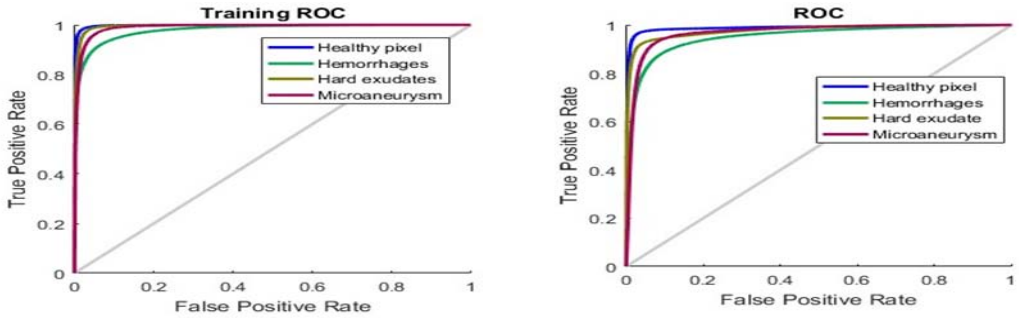

Figure 9. Training and testing ROC curves.

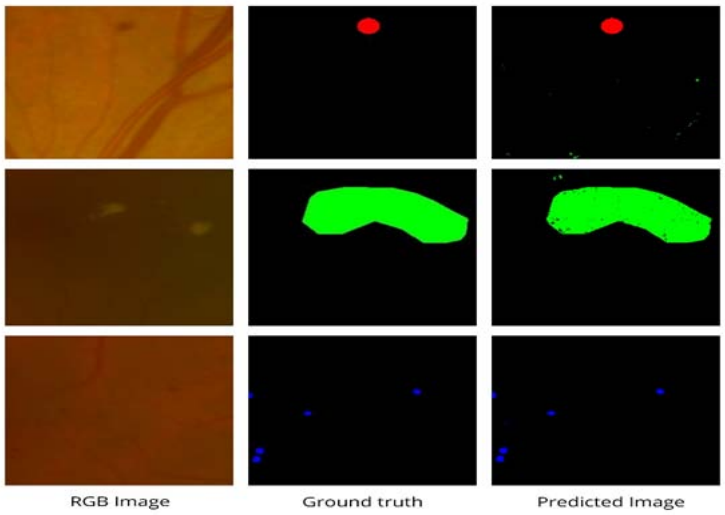

Figure 10. Sample patches of a retinal multispectral image with ground truth and test result: (from left to right) a patch of retina of traditional RGB image, ground truth, and the predicted result after testing. In ground truth and predicted image color Red is assigned to HEM, Green is assigned to HEs and Blue is for MAs for clarity.

Performance/ Accuracy Calculation: Precision and Recall are most familiar and frequently used performance metrics used by computer vision research community (Kauppi *et al.,* 2013). In medical research, the ROC-based analysis is a perfect method for medical decision making (Kauppi *et al.,* 2013). For image-based (patient-wise) evaluation, the accuracy terms Sensitivity (or Recalls) and Specificity are calculated from the ROC curve in the range [0%, 100%] or [0, 1]. The ROC curve originates from a binary classifier where four possible cases can be occurred that are summarized in the confusion matrix (Table 2).

Table 2. Confusion matrix of a binary classifier

	Predicted Class		
Actual Class	True Positive, TP	False Negative, FN	Actual Positives, P
	False Positive, FP	True Negative, TN	Actual Negatives, N
	Predicted Positive, P	Predicted Negative, N	Total Sample, S = P+N

In this study, the main results were evaluated and reported in the metrics terms: Precision and Recall. In addition, for a perfect comparison with existing literature (Rahaman *et al.,* 2013), the results were also calculated in terms of Sensitivity (or Recalls) and Specificity. Following are short descriptions of the basic terms relevant to a binary classifier:

1) True Positive (TP): If actual class and predicted class are both positive.
2) True Negative (TN): If actual class and predicted class are both negative.
3) False Positive (FP): If actual class is negative but predicted class is positive.
4) False Negative (FN): If actual class is positive but predicted class is negative.

Recall/Sensitivity/True positive rate: It is the ratio of the number of correctly predicted positive cases and the number of real positive cases in the data, i.e.,

$$\text{Recall/Sensitivity} = \frac{TP}{TP+FN} \quad (6)$$

Precision: It is the ratio of the number of correctly predicted positive pixels and the number of total predicted positive pixels, i.e.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

Clearly, high Recall indicates low false negative (FN) prediction where high Precision indicates low false positive prediction.

Specificity /True Negative Rate: It is the ratio of the number of correctly predicted negative cases and the number of actual negative cases, i.e.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (8)$$

Overall Accuracy (in %):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} * 100 \quad (9)$$

Results

The detection accuracy of diabetic retinopathies of both multispectral and RGB images is presented in Table 3. Evidently, detection in multispectral images was more accurate than the conventional RGB images. For the multispectral image dataset, in cases of MAs, HEs, and HEMs, respectively; the

Precision (0.86, 0.93, 0.83) and *Recall* (0.89, 0.89, 0.84) are much higher than RGB images. The average *Precision* and *Recall* is 0.87 and 0.87 for multispectral images; whereas the quantities are only 0.54 and 0.50 for the RGB images. The ROC curve for the test dataset is shown previously in Fig. 9. This curve demonstrates that all three types of abnormal lesions were classified with almost equal performance. From Fig. 9, it is also clear that every line is significantly above the diagonal line and close to the upper-left corner meaning that their accuracy is highly consistent and reliable (Fawcett, 2006) (Kauppi *et al.,* 2013).

Table 3. Accuracy comparison between RGB color images and multispectral images

DR Lesions	RGB Image		Multispectral Image	
	Precision	*Recall*	Precision	Recall
Microaneurysms(MA)	0.52	0.46	0.86	0.89
Hard Exudates(HE)	0.57	0.59	0.93	0.89
Haemorrhages(HEM)	0.53	0.45	0.83	0.84

Table 4, Fig. 11 and Fig. 12 show the *Sensitivity (/ Recalls)* and *Specificity* of the proposed method and the method proposed by (Rahaman *et al.,* 2013). The bars in both graphs noticeably show the higher accuracies achieved in this study. The average *Precision* and *Recall* is 0.69 and 0.69 in case of the method reported in (Rahaman *et al.,* 2013). The overall accuracy is 94.5% resulted in this study whereas the overall accuracy is 65.3% at the best cutoff point in (Rahaman *et al.,* 2013). It should be noted the same multispectral dataset was used to develop and evaluate the methods proposed by (Rahaman *et al.,* 2013). Therefore, the comparison of the results is uniform and establishes the supremacy of the method proposed in this paper.

Table 4. Accuracy comparison between proposed method and (Rahaman *et al.,* 2013)

DR Lesions	Proposed Method		(Rahaman et al., 2013)	
	Sensitivity (/Recall)	Specificity	Sensitivity (/Recall)	Specificity
Microaneurysms(MA)	0.89	0.95	0.78	0.78
Hard Exudates(HE)	0.89	0.96	0.70	0.70
Hemorrhages(HEM)	0.84	0.88	0.58	0.58
	Overall Accuracy: 94.5%		Overall Accuracy: 65.3%	

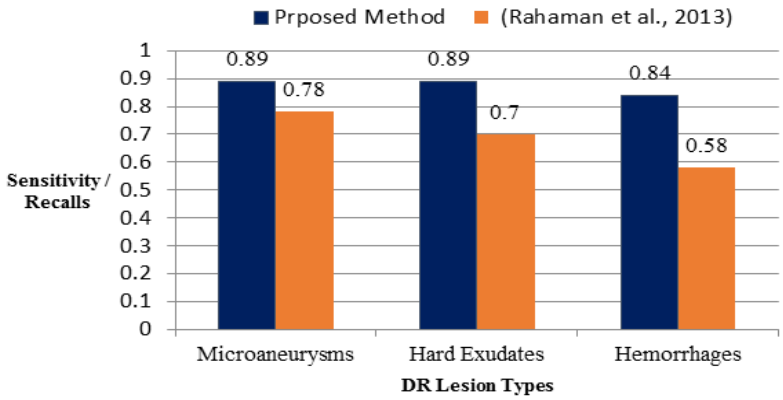

Figure 11. Performance evaluation of the proposed method based on sensitivity/recall

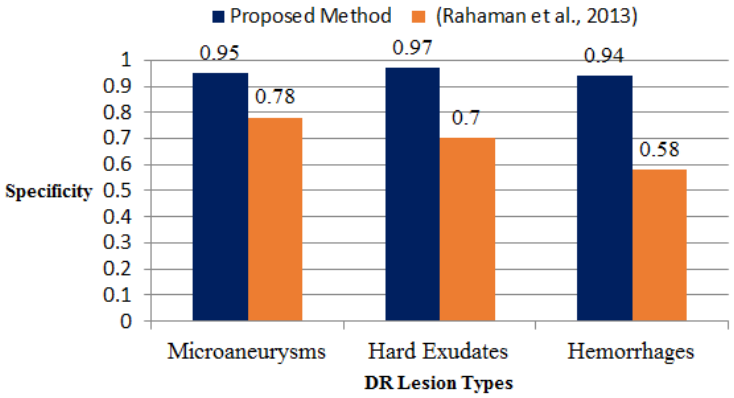

Figure 12. Performance evaluation of the proposed method based on specificity

Discussion

If not diagnosed in time, DR may cause permanent loss of vision. There is a plurality of works for automatic analysis of DR from conventional RGB fundus photographs. However, due to inherent limitation of color information, the RGB imaging does not offer the optimal performance. In multispectral imaging technique, the fundus is photographed in a series of narrow and contiguous reflectance bands removing the limitation. But conventional computer vision techniques developed for RGB imaging are challenging to apply in case of multispectral image. In this context, the artificial NN as the state of the art machine learning technique should be effective in utilizing redundant amount of spectral color information of multispectral images. But to the best of the authors' knowledge, literatures that use a basic 3 layer NN as a baseline method to evaluate the performance of DR detection technique from multispectral images is not available. Therefore, before applying an advanced NN model (e.g. Convolutional NN (CNN)) a basic 3 layer NN model was applied in this study to explore the robustness of the multispectral imaging in order to automatically detect and classify the DR lesions. Tuning the parameters of an advanced NN model should increase the accuracy; however, in that case, the effectiveness of multispectral data would be merged with the power of the learning model. Moreover, in case DIARETSPECDB1 image database any deep learning technique is not effective due to the requirements of a large amount of training data for each class. It would be interesting to conduct such experiment and should be reported in a new study. In this study, use of the multispectral reflectance information shows high accuracy when used with a machine learning technique.

In Fig. 9, the ROC curves demonstrate that all three types of DR lesions were classified with almost similar accuracy. HEMs and MAs have similar color characteristics. Moreover, the spectral colors of healthy tissues like vessels and retinal background sometimes also match with them. But, HEs are quite dissimilar in color compared to the reddish tissues. As a result, the pixels of HEs do not get confused with HEMs, MAs, Vessels and the healthy background. Therefore, the precision is highest in case of detection of HEs. When a DR lesion overlaps with another DR lesion, it creates confusion for the classifier and the classifier makes random decision sometimes. Some preliminary image pre-processing tasks were applied on the images to remove or minimize noises, uneven illuminations and intra and inter retinal spectral color variations. The removal of optic discs and blood vessels as a pre-processing task should increase the detection accuracy. Moreover, an image dataset containing almost equal amounts of instances of each class of DR lesions would make the learning and prediction more efficient.

It should be noted that as a recent imaging technology there are still some constraints such as the expenses, expertise, special medical setup etc. to apply commercial multispectral camera for retinal fundus imaging. In addition, availability of the human subjects with a particular retinopathy is also essential and time dependent. The most recent work (He, Y. *et al,* 2020) has experimented with 50 multispectral images among which 40 are of healthy and only 10 are of unhealthy subjects. Nourrit *et al.* (2010) reported collection of total 28 multispectral images including healthy and unhealthy subjects. In (Calcagni, A. *et al,* 2011) a non-invasive tool was developed based on multispectral images of 16 healthy subjects and one subject with retinal disorders. Hence, even though multispectral retinal imaging is highly prospective, automatic analysis of DR has not been widely explored yet due to difficulties in collecting images of a high number of subjects. The results of the method proposed in this paper are based on maximum available 71 multispectral images among which 55 are of unhealthy subjects, and the results are compared with results reported in (Rahaman *et al.,* 2013) that used the same dataset.

Conclusion

A complete framework to detect DR from multispectral image is demonstrated. A basic NN machine learning technique is used to detect various retinal lesions affected by diabetes mellitus. The reflectance spectrum of each labeled pixel of training dataset is used as the only input feature for learning and prediction. The use of multispectral information enables even a simple NN to detect the retinal disorders more accurately than the conventional RGB images. The proposed approach is straightforward and does not need any laborious feature engineering and/or image processing tasks that vary with the lesion types. The global image pre-processing tasks applied in this study are to remove or minimize the effect of noise, uneven illumination and inter/intra retinal color variations. The results clearly outperform existing stochastic technique based on *Kullback-Leibler* information divergence. In future, advanced machine learning technique can be deployed within the proposed framework to improve the detection accuracy of HEM, HE and MA as well as other DRs including normal retinal structures such as optic disc, blood vessel and fovea.

Acknowledgements

The authors thank and acknowledge for sharing the DIARETSPECDB1 multispectral image database by the Spectral Color Research Lab. of School of Computing of University of Eastern Finland, and the partners of the *Image Ret* project (http://www2.it.lut.fi/project/imageret/).

References

Acharya, U. R., Lim, C. M., Ng, E. Y. K., Chee, C., & Tamura, T. (2009). Computer-based detection of diabetes retinopathy stages using digital fundus images. *Proceedings of the institution of mechanical engineers, part H: journal of engineering in medicine*, *223*(5), 545-553. doi: 10.1243/09544119JEIM486

Calcagni, A., Gibson, J. M., Styles, I. B., Claridge, E., &Orihuela-Espina, F. (2011). Multispectral retinal image analysis: a novel non-invasive tool for retinal imaging. *Eye*, *25*(12), 1562–1569. doi: 10.1038/eye.2011.202

Drake, L. (2007). Prevention of Blindness from Diabetes Mellitus--Report of a WHO Consultation in Geneva, Switzerland, 9-11 November 2005. *Nursing Standard*, *21*(32), 30-31.

Fawcett, T. (2006). An Introduction to ROC Analysis. *Pattern Recognition Letters*. 27 (8): 861–874.

Fält, P., Hiltunen, J., Hauta-Kasari, M., Sorri, I., Kalesnykiene, V., & Uusitalo, H. (2009, June). Extending diabetic retinopathy imaging from color to spectra. In *Scandinavian*

Conference on Image Analysis (pp. 149-158). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-02230-2_16

García, M., López, M. I., Álvarez, D., & Hornero, R. (2010). Assessment of four neural network based classifiers to automatically detect red lesions in retinal images. *Medical engineering & physics*, *32*(10), 1085-1093. doi: 10.1016/j.medengphy.2010.07.014

Gardner, G. G., Keating, D., Williamson, T. H., & Elliott, A. T. (1996). Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *British journal of Ophthalmology*, *80*(11), 940-944. doi: 10.1136/bjo.80.11.940

He, Y., Jiao, W., Shi, Y., Lian, J., Zhao, B., Zou, W., … Zheng, Y. (2020). Segmenting Diabetic Retinopathy Lesions in Multispectral Images Using Low-Dimensional Spatial-Spectral Matrix Representation. *IEEE Journal of Biomedical and Health Informatics*, *24*(2), 493–502. doi: 10.1109/jbhi.2019.2912668

Kauppi, T. (2010). Eye fundus image analysis for automatic detection of diabetic retinopathy. *ISBN: 9789522650160*, *PhD Thesis 176, Lappeenranta University of Technology*.

Kauppi, T., Kämäräinen, J. K., Lensu, L., Kalesnykiene, V., Sorri, I., Uusitalo, H., &Kälviäinen, H. (2013). Constructing benchmark databases and protocols for medical image analysis: Diabetic retinopathy. *Computational and Mathematical Methods in Medicine, V*ol. 2013, Article ID 368514, https://doi.org/10.1155/2013/368514.

Koronyo-Hamaoui, M., Koronyo, Y., Ljubimov, A. V., Miller, C. A., Ko, M. K., Black, K. L., ... & Farkas, D. L. (2011). Identification of amyloid plaques in retinas from Alzheimer's patients and noninvasive in vivo optical imaging of retinal plaques in a mouse model. *Neuroimage*, *54*, S204-S217. doi: 10.1016/j.neuroimage.2010.06.020

Lu, G., & Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, *19*(1), 010901. doi: 10.1117/1.JBO.19.1.010901

Liu, Z., &Zerubia, J. (2015). Skin image illumination modeling and chromophore identification for melanoma diagnosis. *Physics in Medicine & Biology*, *60*(9), 3415.

Mookiah, M. R. K., Acharya, U. R., Chua, C. K., Lim, C. M., Ng, E. Y. K.,& Laude, A. (2013). Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in Biology and Medicine*, 43(12), 2136-2155. doi: 10.1016/j.compbiomed.2013.10.007

Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, *6*(4), 525-533. doi: 10.1016/S0893-6080(05)80056-5

Abdullah, M., Fraz, M. M., & Barman, S. A. (2016). Localization and segmentation of optic disc in retinal images using circular Hough transform and grow-cut algorithm. *PeerJ*, *4*, e2003.

Nayak, J., Bhat, P. S., Acharya, R., Lim, C. M., &Kagathi, M. (2008). Automated identification of diabetic retinopathy stages using digital fundus images. *Journal of Medical Systems*, *32*(2), 107-115. doi: 10.1007/s10916-007-9113-9

Niemeijer, M., Van Ginneken, B., Staal, J., Suttorp-Schulten, M. S., &Abràmoff, M. D. (2005). Automatic detection of red lesions in digital color fundus photographs. *IEEE Transactions on Medical Imaging*, *24*(5), 584-592. doi: 10.1109/TMI.2005.843738

Niemeijer, M., van Ginneken, B., Russell, S. R., Suttorp-Schulten, M. S., & Abramoff, M. D. (2007). Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. *Investigative Ophthalmology & Visual Science*, *48*(5), 2260-2267. doi: 10.1167/iovs.06-0996

Nourrit, V., Denniss, J., Muqit, M., Schiessl, I., Fenerty, C., Stanga, P., & Henson, D. (2010). High-resolution hyperspectral imaging of the retina with a modified fundus camera. *Journal Français D Ophtalmologie*, *33*(10), 686–692. doi: 10.1016/j.jfo.2010.10.010

Rahaman, G. A., Parkkinen, J., Hauta-Kasari, M., & Norberg, O. (2013). Retinal spectral image analysis methods using spectral reflectance pattern recognition. *Lecture Notes in Computer Science* 7786 (pp. 224-238). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-36700-7_18

Ravishankar, S., Jain, A., & Mittal, A. (2009, June). Automated feature extraction for early detection of diabetic retinopathy in fundus images. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 210-217). IEEE. doi: 10.1109/CVPR.2009.5206763

Sánchez, C. I., Hornero, R., Lopez, M. I., & Poza, J. (2004, September). Retinal image analysis to detect and quantify lesions associated with diabetic retinopathy. In *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 1, pp. 1624-1627). IEEE. doi: 10.1109/IEMBS.2004.1403492

Soliz, P., Truitt, P. W., & Nemeth, S. C. (2001, November). Spectrally-based fundus imaging: implications for image enhancement and diagnosis of retinal diseases. In *Conference Record of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers (Cat. No. 01CH37256)* (Vol. 2, pp. 1268-1272). IEEE. doi: 10.1109/ACSSC.2001.987694

Styles, I. B., Calcagni, A., Claridge, E., Orihuela-Espina, F., & Gibson, J. M. (2006). Quantitative analysis of multi-spectral fundus images. *Medical Image Analysis*, *10*(4), 578-597. doi: doi.org/10.1016/j.media.2006.05.007

Vallabha, D., Dorairaj, R., Namuduri, K., & Thompson, H. (2004). Automated detection and classification of vascular abnormalities in diabetic retinopathy. In Conference Record of the Thirty-Eighth Asilomar Conference on *Signals, Systems and Computers, 2004*. (Vol. 2, pp. 1625-1629). IEEE

Yu, C. C., Lau, C., O'Donoghue, G., Mirkovic, J., McGee, S., Galindo, L., & Dasari, R. R. (2008). Quantitative spectroscopic imaging for noninvasive early cancer detection. *Optics Express*, *16*(20), 16227-16239. doi: 10.1364/OE.16.016227

APPENDIX-I

The-Levenberg-Marquardt-nonlinear-least-squares-algorithm

Input: A vector function $f : \mathcal{R}^m \rightarrow \mathcal{R}^n$ with $n \geq m$, a measurement vector $\mathbf{x} \in \mathcal{R}^n$ and an initial parameters estimate $\mathbf{p}_0 \in \mathcal{R}^m$.
Output: A vector $\mathbf{p}^+ \in \mathcal{R}^m$ minimizing $||\mathbf{x} - f(\mathbf{p})||^2$.
Algorithm:

$k := 0$; $\nu := 2$; $\mathbf{p} := \mathbf{p}_0$;
$\mathbf{A} := \mathbf{J}^T\mathbf{J}$; $\epsilon_\mathbf{p} := \mathbf{x} - f(\mathbf{p})$; $\mathbf{g} := \mathbf{J}^T\epsilon_\mathbf{p}$;
stop:=$(||\mathbf{g}||_\infty \leq \varepsilon_1)$; $\mu := \tau * \max_{i=1,\ldots,m}(A_{ii})$;
while (not stop) and $(k < k_{max})$
    $k := k + 1$;
    repeat
        Solve $(\mathbf{A} + \mu\mathbf{I})\delta_\mathbf{p} = \mathbf{g}$;
        if $(||\delta_\mathbf{p}|| \leq \varepsilon_2(||\mathbf{p}|| + \varepsilon_2))$
            stop:=true;
        else
            $\mathbf{p}_{new} := \mathbf{p} + \delta_\mathbf{p}$;
            $\rho := (||\epsilon_\mathbf{p}||^2 - ||\mathbf{x} - f(\mathbf{p}_{new})||^2)/(\delta_\mathbf{p}^T(\mu\delta_\mathbf{p} + \mathbf{g}))$;
            if $\rho > 0$
                stop:=$(||\epsilon_\mathbf{p}|| - ||\mathbf{x} - f(\mathbf{p}_{new})|| < \varepsilon_4||\epsilon_\mathbf{p}||)$;
                $\mathbf{p} = \mathbf{p}_{new}$;
                $\mathbf{A} := \mathbf{J}^T\mathbf{J}$; $\epsilon_\mathbf{p} := \mathbf{x} - f(\mathbf{p})$; $\mathbf{g} := \mathbf{J}^T\epsilon_\mathbf{p}$;
                stop:=(stop) or $(||\mathbf{g}||_\infty \leq \varepsilon_1)$;
                $\mu := \mu * \max(\frac{1}{3}, 1 - (2\rho - 1)^3)$; $\nu := 2$;
            else
                $\mu := \mu * \nu$; $\nu := 2 * \nu$;
            endif
        endif
    until $(\rho > 0)$ or (stop)
    stop:=$(||\epsilon_\mathbf{p}|| \leq \varepsilon_3)$;
endwhile
$\mathbf{p}^+ := \mathbf{p}$;