

Image Processing, Textural Feature Extraction and Transfer Learning based detection of Diabetic Retinopathy

Anjana Umopathy,
Anusha Sreenivasan,
Divya S. Nairy
PES University, Bangalore

S Natarajan
PES University, Bangalore
natarajan@pes.edu

B Narasinga Rao
PES University, Bangalore
narsingrao@outlook.com

ABSTRACT

Diabetic Retinopathy (DR) is one of the most common causes of blindness in adults. The need for automating the detection of DR arises from the deficiency of ophthalmologists in certain regions where screening is done, and this paper is aimed at mitigating this bottleneck. Images from publicly available datasets STARE, HRF, and MESSIDOR along with a novel dataset of images obtained from the Retina Institute of Karnataka are used for training the models. This paper proposes two methods to automate the detection. The first approach involves extracting features using retinal image processing and textural feature extraction, and uses a Decision Tree classifier to predict the presence of DR. The second approach applies transfer learning to detect DR in fundus images. The accuracies obtained by the two approaches are 94.4% and 88.8% respectively, which are competent to current automation methods. A comparison between these models is made. On consultation with Retina Institute of Karnataka, a web application which predicts the presence of DR that can be integrated into screening centres is made.

CCS Concepts

•Applied computing → Bioinformatics;

Keywords

Computer Aided Diagnosis; Biomedical Image Processing; Textural Feature Extraction; Transfer Learning

1. INTRODUCTION

Diabetic Retinopathy (DR) is caused by high blood sugar levels which affect blood vessels in the retina. It is predicted that the number of people suffering from DR would grow from 126.6 million in 2010 to 191 million by 2030[1]. DR is diagnosed by ophthalmologists through the analysis of retinal fundus images, which is exacting and time-consuming. Automating the detection of DR would reduce the burden on the ophthalmologists so that they can focus on the patients in need, and will allow more patients to be screened.

Some of the manifestations in the retinal fundus image from which the pathology can be detected are hard and soft exudates, red lesions, and venous loops. The feature extraction and classification approach in this paper focuses on the detection of exudates and red lesions as these are the most prominent signs of DR. In addition to these, it also uses the textural features extracted from the image to classify it. The transfer learning approach takes into account the features of the image that the deep learning model finds to be the most significant for classification, which is determined by the model after being trained with sufficient images. The doctors in Retina Institute of Karnataka [2] were consulted for this study, and based on their inputs, an end to end web application which predicts the presence of DR given a retinal fundus image, which can be integrated into screening centres was developed. Utilising the diagnosis predicted by the application, the medical practitioners in the screening clinics can accordingly refer the patients in need to ophthalmologists.

The rest of this paper is organised as follows: Section 2 highlights the previous research done on automating the detection of DR. Section 3 details the implementation of the proposed approaches for the detection of DR. Section 4 involves the results and discussion of the methods used. Section 5 holds the conclusion of this paper.

2. RELATED WORK

Previous research and work done on automating the detection of Diabetic Retinopathy are discussed below:

Grace and Kajamohideen[3] developed a Computer-Aided Diagnostic system to classify images with DR. Image processing techniques including colour space conversion, Kirsch's template and histogram equalization were employed to detect exudates present in the image. Seven morphological features were extracted from the processed image. These features were input to an RBF-kernel SVM classifier to predict the presence of DR. This method illustrated good performance, but was restricted to the identification of exudates. According to the Diabetic Retinopathy Disease Severity Scale and International Clinical Diabetic Retinopathy Disease Severity Scale, there are many cases in Nonproliferative DR where exudates are not present[4]. These cases will go undetected by this approach.

Pratt et al. [5] proposed a CNN approach to diagnose and classify the severity of DR. The severity classes include no DR, mild DR, moderate DR, severe DR and proliferative DR. Max pooling was performed using a kernel with size 3x3 and strides 2x2. The CNN was flattened into one dimension after the last convolutional block. Overfitting was avoided using class weights relative to the number of images in the respective class. Images input to the classifier were subjected to colour normalisation and resized into 512x512

pixels. The Kaggle dataset was used to test this system and it gave an accuracy of 75%.

On interacting with doctors from multiple eye hospitals, it was found that although much research has been done on automating the detection of common medical conditions, there is no software that is currently being used by clinics or hospitals to automate the detection of DR. This paper describes an approach that could serve as a tool which can be integrated into screening clinics.

3. MATERIALS AND METHODS

3.1 Data Collection

Anonymised retinal fundus images obtained from the Retina Institute of Karnataka were used, along with images from the publicly available STARE [6] dataset for the feature extraction and classification approach. For the Transfer Learning approach, images from STARE, HRF [7] and Messidor [8] datasets were used to train the model.

3.2 Feature Extraction and Classification Approach

The feature extraction approach involved the detection of the most common manifestations of DR - exudates and red lesions, in the fundus image. The areas of these manifestations along with the textural features of the image, were input to an Information Gain Decision Tree classifier to predict the presence of DR.

3.2.1 Exudates Detection

3.2.1.1 Image Preprocessing.

The retinal image was masked around the field of view (FOV) using Hough circle detection [9]. A good contrast between the exudates and the background, as well as between the optic disc and the background is provided by the green channel. Thus, the green channel was used for the exudates detection and for making the optic disc mask.

3.2.1.2 Optic Disc mask creation.

Contrast Limited Adaptive Histogram Equalization [10] (CLAHE) was applied to enhance the contrast of the green channel of the image. To the result of this step, 2-D first derivative of Gaussian Matched Filter with dynamic thresholds corresponding to the image was applied. The maximum intensity point in the filtered image was found. This point corresponds to the optic cup, which is the brightest region of the optic disc. Using the average cup to disc ratio and size of the retina, a mask of optic disc was created using the optic cup as center. Important stages involved in the creation of optic disc mask are shown in Figure 1: For the image shown in Figure 1(a), the resultant image formed by applying CLAHE to the inversion of its green channel is shown in Figure 1(b). Figure 1(c) is the filtered image and Figure 1(d) shows the optic disc mask.

3.2.1.3 Detection of Exudates.

Illumination Equalisation was applied to the preprocessed image to overcome uneven illumination of the images and to make all the images belong to approximately the same intensity range.

$$IE(Image) = Box51(Image) - 128 \quad (1)$$

where IE represents Illumination Equalisation and Box51 represents a Box filter of size 51 x 51

A 2-D first derivative of Gaussian Matched filter was used to convolve over the illumination equalised image. Optic disc was masked from the filtered image. This is illustrated in Figure 2. Figure 2(a) shows the Illumination equalised image. Figure 2(b)

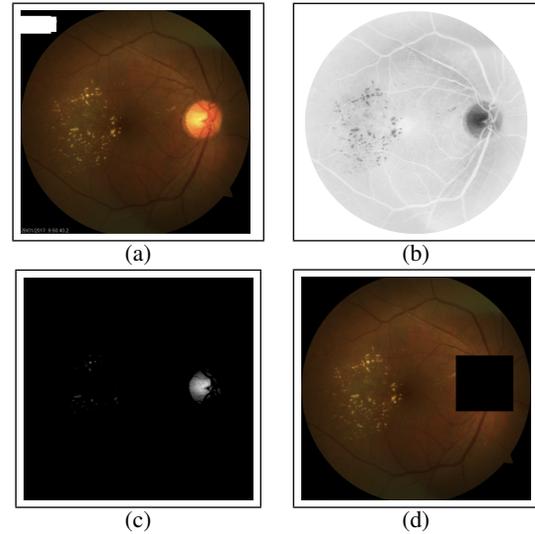


Figure 1: Optic Disc mask creation stages: For the image shown in (a) the resultant image formed by applying CLAHE to the inversion of its green channel is shown in (b). (c) is the filtered image and (d) shows the optic disc mask

shows the result of applying the exudates detection algorithm on Figure 1(a)

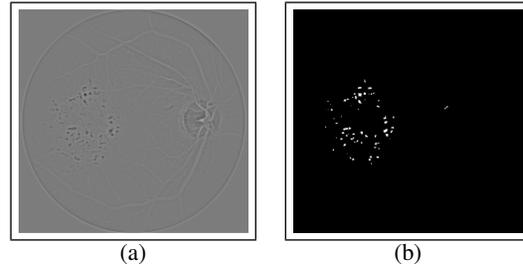


Figure 2: Exudates detection: (a) Illumination equalised image (b) Result of applying the exudates detection algorithm on Figure 1(a)

3.2.2 Red Lesions Detection

A good contrast between the red lesions and the background is provided by the green channel. To prevent the aberrant detection of bright regions (exudates and optic disc), a low intensity difference between them and the background is required, which is prominent in the red channel. To utilise the advantages of both the channels, the image formed by modifying the histogram of the green channel in accordance with that of the red channel was used. IE was employed to this image using Equation(1).

Simple image enhancement techniques like CLAHE and Contrast Stretching were used to enhance the contrast of the image while limiting the amplification of noise. This was followed by a 2-D Gaussian Matched Filter to match red lesion templates. The resulting image consisted of red lesions and blood vessels, which are similar in structure and colour, along with some noise. An opening morphological transformation with an elliptical kernel of size 7 was applied to the filtered image. Adaptive Gaussian Threshold-

ing, CLAHE and denoising were used on the filtered image to form the blood vessel mask. The blood vessel mask was applied to the morphologically transformed image to obtain the red lesions. The important stages in red lesions detection is illustrated in Figure 3: 3(a) is the original image, 3(b) is the histogram matched image, 3(c) is the filtered image 3(d) is the morphologically transformed image, 3(e) shows the blood vessel mask and 3(f) illustrates the red lesions detected.

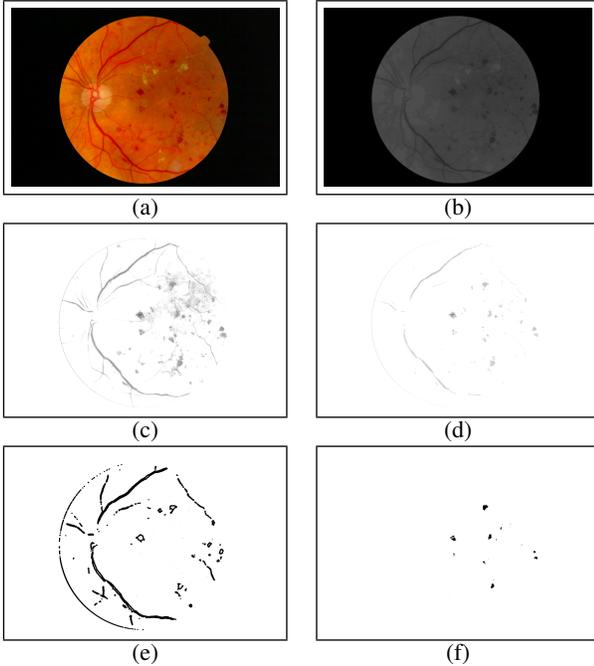


Figure 3: Important stages in red lesions detection: (a) Original image (b) Histogram matched image (c) Filtered image (d) Morphologically transformed image (e) Blood vessel mask (f) Red lesions detected

3.2.3 Textural Feature Extraction

Textural features [11] are quantifications of various textures perceived from an image. A grey level co-occurrence matrix (GLCM) was used for the calculation of the textural features of each image. To increase the accuracy of image classification and to make it less sensitive to the scale of each textural feature, the GLCM was normalised. 15 textural features were extracted, namely ASM, Energy, Dissimilarity, Entropy, Contrast, Correlation, Homogeneity, Sum of squares variance, Sum Average, Sum Variance, Sum Entropy, Difference Variance, Difference Entropy, Information Measure of Correlation - 1, and Information Measure of Correlation - 2, which are calculated using the formulae in Table 1.

3.2.4 Classification

The exudates area, red lesions area, and the 15 textural features extracted were input to Gini and Information Gain Decision Trees, Support Vector Machine with RBF Kernel and Linear Kernel and Feedforward Artificial Neural Network classifiers. The number of images of class DR was 89, while the number of images of class Normal (Healthy - No DR) was 25. To prevent class-bias and overfitting of the classifiers, stratified 5-fold cross validation was used. Table 2 shows the performance comparison of the classifiers. Since the Decision Tree with Entropy as the criterion for splitting had the

best cross validation accuracy and performance, it was the chosen as the final classifier for the feature extraction approach.

3.3 Transfer Learning

Transfer learning [12] was introduced to make machine learning systems leverage the knowledge learnt from the previous tasks for the current task. It has successfully been used to transfer knowledge and improve models from one domain with enough data to learn from, to a similar domain where not enough data is available. GoogleNet Inception-v3 [13], a deep Convolutional Neural Network (CNN) [14] with 22 layers was trained for the ImageNet challenge to classify 1,000,000 images into 1,000 classes. This pre-trained Inception-v3 deep CNN was used for image classification. The final layer of the CNN was re-trained to classify the images as DR and Normal. A total of 601 images labelled "Normal" and 761 images labelled "DR" from the online datasets HRF, STARE, and Messidor were used for training. An accuracy of 88.8% was obtained over the test dataset which was a combination of images provided by the Retina Institute of Karnataka along with images from the STARE dataset.

4. RESULTS AND DISCUSSION

The image processing, textural feature extraction and classification approach was developed using Open Source Computer Vision (OpenCV). This approach achieved an accuracy of 94.4% over a subset of the images in the STARE dataset, along with the dataset provided by the Retina Institute of Karnataka

TensorFlow was used to retrain the final layer of the pre-trained Inception-v3 deep CNN to classify the images as DR or Normal. Having an accuracy of 88.8% over the test dataset, the performance of this approach was slightly lower than the feature extraction approach.

To accommodate images from fundus cameras with varying focal length and images with varying illumination, the image processing method modifies all its images to match a baseline. During this process, some critical features of images with a different baseline may get obscured. This would reduce its performance with such images. However, it would perform very well with images having configurations similar to its baseline.

The transfer learning method uses a deep CNN which can capture all kinds of local and more abstract features of the image, and learns to adapt to different dataset configurations with increased range of datasets and images supplied for training the model, so it does better given a random image from any dataset or fundus camera, compared to the image processing method, but does not perform as well as the image processing method for images with a similar baseline.

The image processing method would be better suited for instances where the configuration of the images that will be input to it is known beforehand, and would perform well in screening clinics where the same fundus camera would be used to capture the retinal images. The transfer learning method would be of better use where the configuration of the image is not known, so it can be used in a website hosted on the internet for people to check the prediction of DR for a given retinal image from any dataset.

5. CONCLUSIONS

Two methods were used to detect the presence of DR in retinal fundus images. The first method employed image processing, textural feature extraction and classification using a decision tree with information gain classifier. The second method used Transfer Learning on the pre-trained GoogleNet Inception-v3 CNN. Both

Table 1: Textural Features Used

Texture Feature	Formula
ASM	$\sum_{i,j=0}^{levels-1} P_{i,j}^2$
Energy	\sqrt{ASM}
Dissimilarity	$\sum_{i,j=0}^{levels-1} P_{i,j} i-j $
Entropy	$-\sum_i \sum_j p(i,j) \log(p(i,j))$
Contrast	$\sum_{i,j=0}^{levels-1} P_{i,j}(i-j)^2$
Correlation	$\sum_{i,j=0}^{levels-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$
Homogeneity	$\sum_{i,j=0}^{levels-1} \frac{P_{i,j}}{1+(i-j)^2}$
Sum of squares variance	$\sum_i \sum_j (i-\mu)^2 p(i,j)$
Sum Average	$\sum_{i=2}^{2N_g} i p_{x+y}(i)$
Sum Variance	$\sum_{i=2}^{2N_g} (i-f_8)^2 p_{x+y}(i)$
Sum Entropy	$-\sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\}$
Difference Variance	$\sum_{i=0}^{N_g-1} i^2 p_{x+y}(i)$
Difference Entropy	$-\sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\}$
Information Measure of Correlation - 1	$\frac{HXY - HXY1}{\max(HX, HY)}$, where $HXY1 = -\sum_i \sum_j p(i,j) \log\{p_x(i)p_y(j)\}$ and $HXY = -\sum_i \sum_j p(i,j) \log(p(i,j))$
Information Measure of Correlation - 2	$(1 - \exp[-2(HXY2 - HXY)])^{1/2}$, where $HXY2 = -\sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}$ and $HXY = -\sum_i \sum_j p(i,j) \log(p(i,j))$

Table 2: Classifier Comparison

Classifier	Mean Accuracy	Mean Precision	Mean Recall	F1 Score
Decision Tree (Gini)	0.849	0.919	0.836	0.876
Decision Tree (Information Gain)	0.944	0.926	0.953	0.939
SVM (RBF Kernel)	0.782	0.890	1	0.942
SVM (Linear Kernel)	0.765	0.893	0.944	0.917
Feedforward Artificial Neural Network	0.776	0.889	0.993	0.938
Logistic Regression	0.761	0.874	0.930	0.901

methods demonstrated good performance.

A web application was developed which would take as input a fundus image, and output the prediction of whether the image shows signs of DR or not. The screening clinics can integrate this tool with their systems so that the technician can refer the patients with the pathology to ophthalmologists.

The authors are currently collaborating with the doctors of Retina Institute of Karnataka to integrate the web application in diabetic screening centres.

6. ACKNOWLEDGMENTS

The authors thank Dr. Hemant Murthy and Dr. Manjula V from Retina Institute of Karnataka, for their assistance in the detection of manifestations of DR, and for providing them with retinal images of patients to augment their dataset.

7. REFERENCES

- [1] Yingfeng Zheng, Mingguang He, and Nathan Congdon. The worldwide epidemic of diabetic retinopathy. *Indian journal of ophthalmology*, 60(5):428, 2012.
- [2] Retina institute of karnataka. <http://www.retinainstitute.org/>.
- [3] Annie Grace and S Kajamohideen. Diagnosis of diabetic retinopathy by extracting blood vessels and exudates using retinal color fundus images. 11:20–28, 01 2014.
- [4] Lihteh Wu, Priscilla Fernandez-Loaiza, Johanna Sauma, Erick Hernandez-Bogantes, and Marissé Masis. Classification of diabetic retinopathy and diabetic macular edema. *World journal of diabetes*, 4(6):290, 2013.
- [5] Harry Pratt, Frans Coenen, Deborah M. Broadbent, Simon P. Harding, and Yalin Zheng. Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science*, 90:200 – 205, 2016. 20th Conference on Medical Image Understanding and Analysis (MIUA 2016).
- [6] AD Hoover, Valentina Kouznetsova, and Michael Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical imaging*, 19(3):203–210, 2000.
- [7] Attila Budai, Rüdiger Bock, Andreas Maier, Joachim Hornegger, and Georg Michelson. Robust vessel segmentation in fundus images. *International journal of biomedical imaging*, 2013, 2013.
- [8] Etienne Decencière, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Béatrice Cochener, Caroline Trone, Philippe Gain, Richard Ordonez, Pascale Massin, Ali Erginay, Béatrice Charton, and Jean-Claude Klein. Feedback on a publicly distributed database: the messidor database. *Image Analysis & Stereology*, 33(3):231–234, August 2014.
- [9] Dana H Ballard. Generalizing the hough transform to detect arbitrary shapes. In *Readings in computer vision*, pages 714–725. Elsevier, 1987.
- [10] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderveld. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3):355–368, 1987.
- [11] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, Nov 1973.
- [12] Lisa Torrey and Jude Shavlik. Transfer learning. *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, 1:242, 2009.
- [13] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2818–2826, 2016.
- [14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.