

# A Novel Weakly Supervised Multitask Architecture for Retinal Lesions Segmentation on Fundus Images

Ravi Shankar Mishra<sup>a</sup>, Thameem Ansari<sup>b</sup>, Pabba Vamshi<sup>c</sup> and Mulkupogu Praveen<sup>d,\*</sup>

Department of Electronics & Communications Engineering,  
Guru Nanak Institute of Technology, (GNIT)  
Ibrahimpatanam Rangareddy- 501506, India

**Abstract:-** The first step towards an automated diagnosis tool for retinopathy that is interpretable in its decision-making. However, the limited availability of ground truth lesion detection maps at a pixel level restricts the ability of deep segmentation neural networks to generalize over large databases. In this paper, we propose a novel approach for training a convolutional multitask architecture with supervised learning and reinforcing it with weakly supervised learning. The architecture is simultaneously trained for three tasks: segmentation of red lesions and of bright lesions, those two tasks done concurrently with lesion detection. In addition, we propose and discuss the advantages of a new preprocessing method that guarantees the color consistency between the raw image and its enhanced version. Our complete system produces segmentations of both red and bright lesions.

**Keywords----**Retina; Image segmentation; Feature Extraction; Task analysis; lesions

## I. INTRODUCTION

In the field of ophthalmology, the development of deep learning, as well as the ease of acquisition of retinal fundus images, has already led to the approval of an automatic diabetic retinopathy (DR) grading system by US health authorities [1]. Fundus images give access to a highly detailed 2D representation of the surface of the retina, generally centered on the macula or the optic disc and containing several diagnostically relevant biomarkers. An abnormality in the retina can either be a manifestation of an eye disease (e.g. age-related macular degeneration, DR), a systemic disease (e.g. hypertensive retinopathy) or even a trauma (e.g. traumatic macular hole).

From a clinical point of view, the presence or absence of a given lesion, as well as its shape, texture and location are essential indicators for assessment of the progression and the severity of diseases. For example, early detection of microaneurysms (A) is essential. as their presence in the early stages of DR is the criterion to distinguish between non-healthy and healthy fundi.

Therefore, the detection of A has been the object of an online challenge [2]. As A are usually only a few pixels wide, the submitted approaches focus on detection rather than segmentation and share the same steps: pre-processing, candidates extraction, feature selection and candidates classification. Due to the circular

shape of the A, candidates extraction can be done using morphological operations [3], [4], followed by match-filtering with a prior on the Gaussian profile of A as introduced in [5], or with a multi-orientation gradient weighting scheme as in [6]. However, this method remains sensitive to background noise, so an approach based on local contrast independent of the size and shape of the regional minimum was proposed in [7]. Features are then extracted for each candidate (morphological, intensity based and/or gradient based in [3]– [6], dynamic in [7]) and classified (Random Forest in [7], KNN in [3]). In addition to match-filtering, template-matching approaches with generalized Gaussian functions were also proposed in [8].

In order to counter-balance image illumination variations and noise, the authors propose to work in the wavelet domain. In [9], a set of features is extracted from the profile of dark candidates and a K-nearest neighbor is used to recognize A among them. Nonetheless, all those approaches rely on hyper-parameters that need to be tuned, which might compromise the generalization capacity of the system. Therefore, fully trainable approaches have also been used [10], [11]. In [10], pixel-wise classification is done using a CNN. The network is trained on image patches and results are shown to improve with the use of non-uniform sampling and foveation. In [11], candidates are extracted by analyzing the intensity profile along a spiral reading of a slicing window. A convolutional network is then used for automatic feature extraction and classification of each candidate patch, trained with reinforcement sample learning. Concerning bright lesion segmentation, the literature mostly focuses on detection and/or segmentation of exudates.

As for A, a common approach is to identify candidates and refine the results through further analysis. For example, in [12]–[14], candidates extraction is achieved with morphological operations and the refinement by applying maximum entropy thresholding [13] or Random Forest classification [14]. In [15], an alternative approach is proposed in which a Random Forest classifies each patch depending of the presence or absence of lesions. Positive patches are then fed to a segmentation module that refines the boundaries based on local variance, size and local contrast.

The reviews in [16], [17] give exhaustive overviews of the state of the art for exudate segmentation and in addition

proposes a generalized method to improve on the previous works. Their approach relies on adaptive k-means clustering followed by dynamic decision thresholding. Nonetheless, the proposed method requires optic disc segmentation and vessels removal. CNN-based approaches have also been investigated, as in [18]. Interestingly, to be able to successfully train their network, the authors pre-trained it on a different but related task, namely the segmentation of A. Transfer-learning and finetuning were then applied to exudates segmentation. The work in [19] proposes an approach based on two steps: the first step identifies potential bright lesions, which are then classified according to their type (drusen, exudates, or cotton-wool spots) in a second step. This was the first work on joint segmentation of all types of bright lesions. The first classifier used is a k-NN whereas the second step uses a linear discriminant classifier.

This work was further extended in [20] to be integrated in a screening program able to identify the presence or absence of drusen and exudates in retinal images, as both are related to different diagnoses.

## II. IMAGE PROCESSOR:

An image processor does the functions of image acquisition, storage, preprocessing, segmentation, representation, recognition and interpretation and finally displays or records the resulting image. The following block diagram gives the fundamental sequence involved in an image

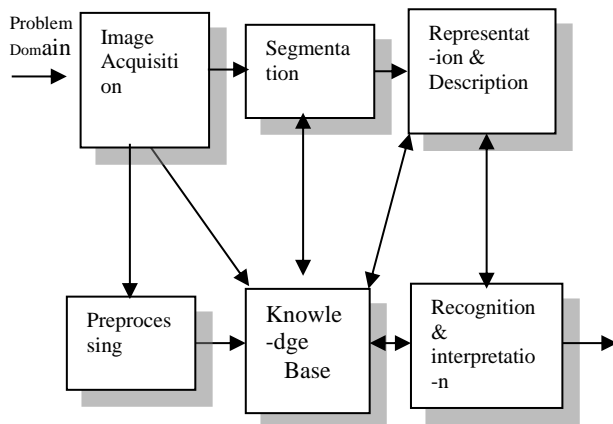


Fig 1. Block Diagram Of Fundamental Sequence Involved In An Image Processing System

As detailed in the diagram, the first step in the process is image acquisition by an imaging sensor in conjunction with a digitizer to digitize the image. The next step is the preprocessing step where the image is improved being fed as an input to the other processes. Preprocessing typically deals with enhancing, removing noise, isolating regions, etc. Segmentation partitions an image into its constituent parts or objects.

The output of segmentation is usually raw pixel data, which consists of either the boundary of the region or the pixels in the region themselves. Representation is the process of transforming the raw pixel data into a form useful for subsequent processing by the computer. Description deals

with extracting features that are basic in differentiating one class of objects from another. Recognition assigns a label to an object based on the information provided by its descriptors. Interpretation involves assigning meaning to an ensemble of recognized objects.

The knowledge about a problem domain is incorporated into the knowledge base. The knowledge base guides the operation of each processing module and also controls the interaction between the modules. Not all modules need be necessarily present for a specific function. The composition of the image processing system depends on its application. The frame rate of the image processor is normally around 25 frames per second.

## III. IMAGE PROCESSING TECHNIQUES:

This section gives various image processing techniques.

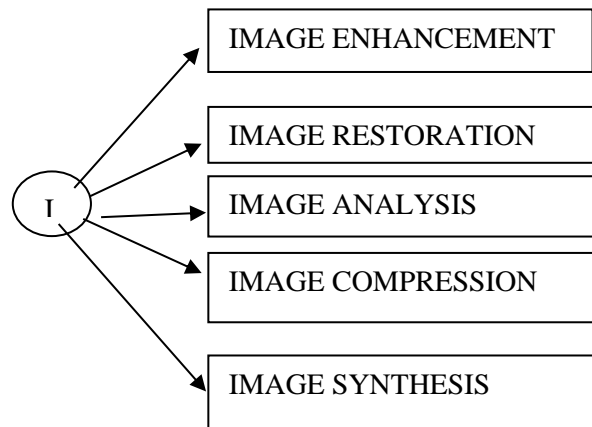


Fig2: Image Processing Techniques

### IMAGE ENHANCEMENT:

Image enhancement operations improve the qualities of an image like improving the image's contrast and brightness characteristics, reducing its noise content, or sharpen the details. This just enhances the image and reveals the same information in more understandable image. It does not add any information to it.

### IMAGE RESTORATION:

Image restoration like enhancement improves the qualities of image but all the operations are mainly based on known, measured, or degradations of the original image. Image restorations are used to restore images with problems such as geometric distortion, improper focus, repetitive noise and camera motion. It is used to correct images for known degradations.

### IMAGE ANALYSIS:

Image analysis operations produce numerical or graphical information based on characteristics of the original image. They break into objects and then classify them. They depend on the image statistics. Common operations are

extraction and description of scene and image features, automated measurements, and object classification. Image analyze are mainly used in machine vision applications.

**IMAGE COMPRESSION:**

Image compression and decompression reduce the data content necessary to describe the image. Most of the images contain lot of redundant information, compression removes all the redundancies. Because of the compression the size is reduced, so efficiently stored or transported. The compressed image is decompressed when displayed. Lossless compression preserves the exact data in the original image, but Lossy compression does not represent the original image but provide excellent compression.

**IMAGE SYNTHESIS:**

Image synthesis operations create images from other images or non-image data. Image synthesis operations generally create images that are either physically impossible or impractical to acquire.

micro controller board which is a low power, cheaper, highly flexible with ATmega328p microcontroller is designed [12]. It's a great platform for interfacing with many devices like laptops, Smart phones at the same time. The integrated wireless sensor network system transmits data based on wireless communication using Bluetooth module technology, and performs



**Retinal image of eye is taken as a input image**

It consists in applying to each input image a random combination of one or more transformations among the following: rotation, shearing, elastic distortion, horizontal flipping, scaling, gamma, brightness, saturation, and contrast transform. Each individual transform is parameterized by a random amplitude (for example, the angle of rotation or the multiplicative coefficient for contrast), and re-sampled at each training iteration. With this strategy, the network is guaranteed to never see the exact same image twice.

**Patches extraction :**

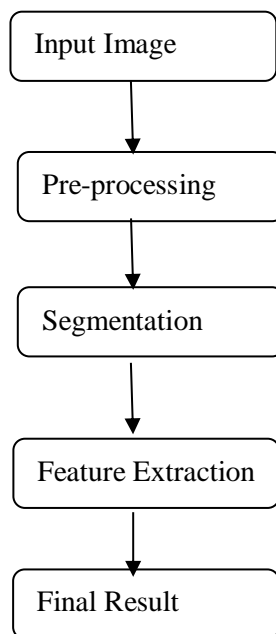


**Simplified Representation Of The Proposed Architecture**

The distribution of the patches' centers is conditioned by the ground truth: A prior favors patches centered on lesions (which helps counterbalancing the high imbalance of classes, the background being the predominant class). The prior is stronger on red lesions than bright ones, as they tend to be smaller and sparser.

- (a) An Encoder Block is shared among two decoders (b), each of which is specialized in its class of lesions (red or bright).
- (b) An additional block (c) is trained on detecting whether the presented patch is pathological or not.

**IV. BLOCK DIAGRAM**



**Preprocessing Method:**

Our methodology relies on a fully convolutional architecture trained following a novel procedure that combines training at a pixel level and weakly supervised training with label at an image level. The proposed system takes as an input an RGB image and outputs two probability maps for red and bright lesions segmentation. We detail each step of the proposed methodology in the following section

**V. CONCLUSION:**

The proposed approach provides a fully trainable model for joint segmentation of bright and red lesions in retinal images and proposes a novel methodology for exploiting image-level labels to improve segmentation performance. The fast execution time makes the model suitable for clinical deployment. Consequently, future work will assess how it can accelerate the process of labelling data by a clinician, by providing lesion pre-annotations. Subsequent research will also focus on classifying the individual segmented lesions. This will serve two purposes: 1) building a detailed atlas of the retinal lesions, and 2) providing an automatic diagnosis established using a set of rules based on the detected lesions. Both of these developments will contribute to a computer-aided diagnosis system that is transparent in its decision-making.

## V. ACKNOWLEDGEMENT

This work was supported by Department of Electronics and communication Engineering(ECE) Of Guru Nanak Institute Of Technology(GNIT).

## REFERENCES

- [1] FDA, "Fda permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems," FDA News Release, Apr. 2018
- [2] M. Niemeijer, B. van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuguel, J. You, A. Mayo, Q. Li, Y. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita, and M. D. Abramoff, "Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 185–195, Jan 2010.
- [3] M. J. Cree, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "A fully automated comparative microaneurysm digital detection system," *Eye*, vol. 11, no. 5, p. 622, 1997.
- [4] M. Haloi, "Improved microaneurysm detection using deep neural networks," *CoRR*, vol. abs/1505.04424, 2015. [Online]. Available: <http://arxiv.org/abs/1505.04424>
- [5] U. Budak, A. Sengur, Y. Guo, and Y. Akbulut, "A novel microaneurysms detection approach based on convolutional neural networks with reinforcement sample learning algorithm," *Health Information Science and Systems*, vol. 5, no. 1, p. 14, Nov 2017