# Classifying DME vs Normal SD-OCT volumes: A replicable comparison

Joan Massich, Mojdeh Rastgoo, Guillaume Lemaître, Désiré Sidibé
LE2I UMR6306, CNRS, Arts et Métiers
Univ. Bourgogne Franche-Comté
12 rue de la Fonderie
71200 Le Creusot, France

Email: joan.massich@u-bourgogne.fr

Carol Y. Cheung, Tien Y. Wong
Singapore Eye Research Institute
Singapore National Eye Center, Singapore

Fabrice Mériaudeau
UTP

***Abstract*—**
***Index Terms*—**

## I. Introduction

Eye diseases such as Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) are the most common causes of irreversible vision loss in individuals with diabetes. Just in United States alone, health care and associated costs related to eye diseases are estimated at almost 500 $M [1]. Moreover, the prevalent cases of DR are expected to grow exponentially affecting over 300 M people worldwide by 2025 [2]. Given this scenario, early detection and treatment of DR and DME play a major role to prevent adverse effects such as blindness. DME is characterized as an increase in retinal thickness within 1 disk diameter of the fovea center with or without hard exudates and sometimes associated with cysts [3].

Fundus photography and Optical Coherence Tomography (OCT) imaging are the two major screening techniques used for detection of such diseases. The latter is based on optical reflectivity and produces cross-sectional and three-dimensional images of the central retina. Spectral Domain OCT (SD-OCT) as the latest generation of OCT images offers higher resolution and faster imaging acquisition. Using the benefits of cross-sectional retinal morphology provided by SD-OCT images, to reveal and understand eye pathologies, numerous studies have been proposed in the past few years. These studies are either focused on the problem of retinal layer segmentation, which is a necessary step for retinal thickness measurements [4, 5], or classification of normal and diseased (abnormal) SD-OCT images. In this paper we review and compare the recent works dedicated to the second task.

Manual evaluation of OCT volumetric scans are not straight forward and rather time consuming, expensive and prone to errors [6]. Thus automatic feature detection and classification of these volumes is of great help to ophthalmologists. Subsequently, different methods have been proposed for automatic classification of these volumes. A fair comparison of methods developed is not possible apart from their evaluation using a common dataset. This paper presents comparison of these works while being evaluated on a common dataset. Moreover, due to the characteristics of our datasets (DME vs. normal volumes), it presents the potential of the stat-of-the-art methods for specific task of DME classification. classification of DME and normal should be mentioned.

This article is structured as follows, Sect. **??** reviews the state-of-the-art and the methods compared. This article is structured as follows: Background(Sect. **??**) offers a general idea of the methods reviwed. Materials and methods discusses data and mapping of the methodologies to our framework. Results offers (Sect. VI) (a) individual results of each methodology, as well as our strategy followed to validate that our implementation complies with the results reported by the original work (b) comparative results of the best methodlogy configurations to drive our discussion. Discussion(Sect. VII). Conclusion and Further work(Sect. VIII).[old]

## II. Data

~~Despite Venhuizen et al. tested on a public dataset [?], this dataset is intended to AMD. Srinivasan also tested on a public dataset [?], however the images are cropped and filtered etc. So thats why we collected the SERI dataset.~~ [sik]

This dataset was acquired by the Singapore Eye Research Institute (SERI), using CIRRUS TM (Carl Zeiss Meditec, Inc., Dublin, CA) SD-OCT device. The dataset consists of 32 OCT volumes (16 DME and 16 normal cases). Each volume contains 128 B-scan with resolution of 512 $\times$ 1024 pixels. All SD-OCT images are read and assessed by trained graders and identified as normal or DME cases based on evaluation of retinal thickening, hard exudates, intraretinal cystoid space formation and subretinal fluid. [old]

## III. Review

This section reviews the recent and the state-of-the-art methods in classification of SD-OCT volumes as normal vs. abnormal. These methods can be categorized into two groups, supervised and semi-supervised explained in the following.

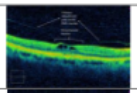
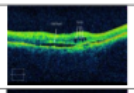
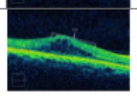
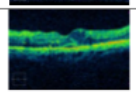
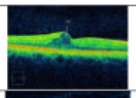
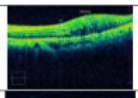
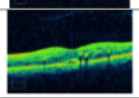
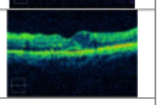
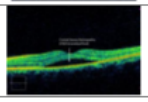
| Type of Lesions | # | | | # | |
|---|---|---|---|---|---|
| Vitreomacular Traction | 4 |  | Fluid with HE and cystoid spaces | 1 |  |
| Cystoid spaces with Hard Exudates (HE) causing central retinal thickening | 1 |  | Cystoid spaces causing parafoveal retinal thickening | 1 |  |
| Cystoid spaces causing central and parafoveal retinal thickening | 1 |  | CSR with HE causing retinal thickening | 2 |  |
| Retinal thickening | 2 |  | Cystoid spaces causing retinal thickening | 3 |  |
| CSR(subretinal fluid) causing central and parafoveal thickening | 1 |  | | | |

Figure 2. Example of the DME dataset

Fig. 1. Experimental Setup

Supervised classification is based on full annotated and labeled training set. In such methods the labeled training data is used to train the classifier function and the learned function is used for prediction. Semi-supervised classification takes advantage of both unlabeled and labeled data. This techniques are particularly useful when there is lack of annotated data moreover it has shown that use of small amount of labeled data in conjunction of unlabeled data can increase the learning accuracy.

The details implementation of the following methods and their integration to our common framework is described in Sect. *experiments*.

### A. Supervised methods

Venhuizen *et al.* proposed a supervised method for SD-OCT image classification using Bag-of-Words (BoW) models [6]. With the aim of Age-related Macular Degeneration (AMD) versus normal volumes classification, the method starts with the detection and selection of keypoints in each individual B-scan. The final keypoints are selected as the most salient points corresponding to the top $3\%$ of the vertical gradient values. Then, a texton of size $9 \times 9$ pixels is extracted around each keypoint, and Principal Component Analysis (PCA) is applied to reduce the dimension of every texton to get a feature vector of size 9. All extracted feature vectors are used to create a codebook using $k$-means clustering. Then, each volume is represented in terms of created codebook and is characterized as a histogram that captures the codebook occurrences. These histograms are used as feature vector to train a Random Forest (RF) with a maximum of $100$ trees. Using the proposed method the authors achieved an Area Under the Curve (AUC) of $0.984$ with a dataset of $384$ volumes.

Another supervised method is proposed by Srinivasan *et al.* [7] which intends to distinguish DME, AMD and normal SD-OCT volumes. In this method in order to reduce the inter-patient variations, the OCT images are pre-processed by first enhancing sparsity in a transform-domain (BM3D [**?**]), to reduce their speckle noise, and then by flattening the retinal curvature. The denoised and flatten B-scans are further cropped to reduce the size and therefore complexity of the algorithm. The Histogram of Oriented Gradients (HOG) features are then extracted from multi-resolution pyramid of each pre-processed slice of a volume. These features are classified using a linear Support Vector Machines (SVM). Note that the method classifies each individual B-scan into one of three categories, i.e. DME, AMD, and normal, and then classifies a volume based on the number of B-scans in each category. On a balanced dataset of 45 patients, this method leads to correct classification rate of $100\%$, $100\%$ and $86.67\%$ for normal, DME and AMD patients, respectively.

Replicating the method proposed by Srinivasan *et al.* [7] and adding PCA to the feature extraction step, as was proposed by Venhuizen *et al.* [6], we also compare the performance of extracted HOG and Local Binary Patterns (LBP) features (similar to [7], features are extracted from multi-resolution pyramid) for classification of B-scans, and accordingly volumes. This work was submitted for publication to a recent conference.

Lemaitre *at al.* [8] proposed another method based on extracted LBP features from OCT images and dictionary learning using BoW models [9]. Note that using BoW and dictionary learning contrary to [7] the classification is performed per volume, rather than B-scan. In this method the OCT images are first pre-processed using Non-Local Means (NLM) filtering, to reduce the speckle noise. Then the volumes are mapped into discrete set of structures namely: local, when these structures correspond to patches; or global, when they correspond to volume slices or the whole volume. According to different mapping, LBP or LBP from Three Orthogonal Planes (LBP-TOP) texture features are extracted and represented (per volume) using histogram, PCA or BoW. The final feature descriptors per volumes are classified using RF classifier. Classifying DME versus normal volumes with a balanced dataset of 32 SD-OCT volumes, they achieved a Sensitivity (SE) and Specificity (SP) of 87.50% and 75%, respectively, while using LBP-TOP features and global mapping.

Another supervised method is proposed by Liu *et al.* [10], with the specific aim of B-scan classification, rather than volume classification. In this study, the authors proposed to extract LBP and gradient information (using LBP on canny-edged B-scans) from pre-processed OCT images. The pre-processing is performed by flattening and aligning OCT images in a volume. The features are then extracted from spatial blocks of 3-level multi-scale spatial pyramid of each B-scan. All the obtained histograms are concatenated into a global descriptor whose dimensions are reduced using PCA. Finally a SVM with an Radial Basis Function (RBF) kernel is used as classifier. The method achieved good results in detection of OCT scans containing different pathology such as DME or AMD, with an AUC of $0.93$ using a dataset of $326$ OCT scans.

### B. Semi-supervised methods

An example of semi-supervised approach for SD-OCT classification is recently proposed by Sankar. *et al.* [**?**]. The proposed method is based on appearance modeling of normal OCT images using Gaussian Mixture Model (GMM) and anomaly detection. The abnormal B-scans are detected as outliers to the fitted GMM and volume classification is performed based on the number of detected outliers in the volume. This approach differs from supervised approaches since the B-scan detection method does not require a labeled training set of B-scans. This method starts by pre-processing the B-scans using resizing, flattening and denoising (NLM filter). The features are extracted by taking the intensity information of each B-scan and applying PCA to reduce their dimension. The feature space is then modeled using GMM. In the testing stage, for the new B-scan, the features are extracted in a similar way and they are classified as normal or abnormal based on their Mahalanobis distance to the GMM. Finally the volume classification is performed considering the umber of outliers (abnormal) B-scans per volume.

## IV. Common framework

All the previously mentioned methods follow a similar pipeline or framework, which consists of different steps. We categorized these steps as pre-processing, feature extraction, mapping, feature representation and finally classification, as it is shown in Fig. 2. Pre-processing of OCT volumes, as noted in Sect. III, consists of denoising, flattening the retinal curvature, aligning the B-scans through the whole volume and finally cropping or resizing the volumes. Feature extraction refers to extraction of different textural and shape information from the B-scan or the volumes. Mapping step is used to determine a discrete set of elements (structures) representing a sample (i.e. B-scan/volume). In this step either one structure is used per sample namely global-mapping or the features are extracted with reference to a set of structures, i.e. dense or sparse patches through the sample, local-mapping. In feature representation step, the representation of the final descriptor prior to classification is decided. The extracted features using different mapping techniques can be represented in lower dimensions (using PCA for instance), as a concatenated of descriptor, histogram of words (using BoW) or sparse representation (sparse coding).

Table I summarizes the aforementioned methods in terms of the proposed framework.

TABLE I
Summary of the stat-of-the-art methods based on our common framework, the dashed line separates the supervised and semi-supervised methods.

| Ref | Pre-processing | Features | Mapping | Representation | Classification |
|---|---|---|---|---|---|
| Venhuizen *et al.* [6] | | Texton | Local | BoW, PCA | RF |
| Srinivasan *et al.* [7] | De-noise Flatten Cropped | HOG | Global | | linear-SVM |
| Lemaitre *et al.* [8] | De-noised | LBP LBP-TOP | Local Global | PCA, BoW, Histogram | RF |
| Liu *et al.* [10] | Flatten Aligned | Edge, LBP | Local | PCA | RBF-SVM |
| Sankar *et al.* [**?**] | De-noised Flatten Cropped | Pixel -intensities | Global | PCA | Mahalanobis -distance to GMM |

## V. Experimental Setup~~Materials and Methods~~[sik]

### A. ~~*method comments*~~[sik]

here goes a description left to right of the modules, making remarks of the difference between the needs of each method.

First, the OCT volumes are pre-processed as presented in details in Sect. **??**. Then, LBP and LBP-TOP features are detected, mapped and represented as discussed in depth in Sect. **??**, Sect. **??**, and Sect. **??**, respectively. Finally, the classification step is presented in Sect. **??**. [old]

~~The mapping in A is computed in this manner while in B this comes on the other side bla bla bla~~[sik]

### B. Validation

All the experiments are evaluated in terms of SE and SP using the Leave-One-Patient Out Cross-Validation (LOPO-CV) strategy, in line with [8]. SE and SP are statistics driven from the confusion matrix (see Fig. **??**) as stated in Eq. (1). The SE evaluates the performance of the classifier with respect to the positive class, while the SP evaluates its performance with respect to negative class. [old]

$$SE = \frac{TP}{TP + FN} \qquad SP = \frac{TN}{TN + FP} \tag{1}$$

The use of LOPO-CV implies that at each round, a pair DME-normal volume is selected for testing while the remaining volumes are used for training. Subsequently, no SE or SP variance can be reported. However, LOPO-CV strategy has been adopted despite this limitation due to the reduced size of the dataset. [old]

### C. Management of data depending terms

~~Be aware that when computing the GMM [**?**], or the dictionary [**?**], only the training data for the current fold is used. Therefore such modules are recomputed at each fold.~~[sik] ~~Other parameter tuning such the case of XXXX and YYYYY are also carried out using only ZZZZ~~[sik]

## VI. Results

## VII. Discussion

## VIII. Conclusion and Further work

## References


[1] S. Sharma, A. Oliver-Hernandez, W. Liu, and J. Walt, "The impact of diabetic retinopathy on health-related quality of life," *Current Opinion in Ophtalmology*, vol. 16, pp. 155–159, 2005.

[2] S. Wild, G. Roglic, A. Green, R. Sicree, and H. King, "Global prevalence of diabetes estimates for the year 2000 and projections for 2030," *Diabetes Care*, vol. 27, no. 5, pp. 1047–1053, 2004.

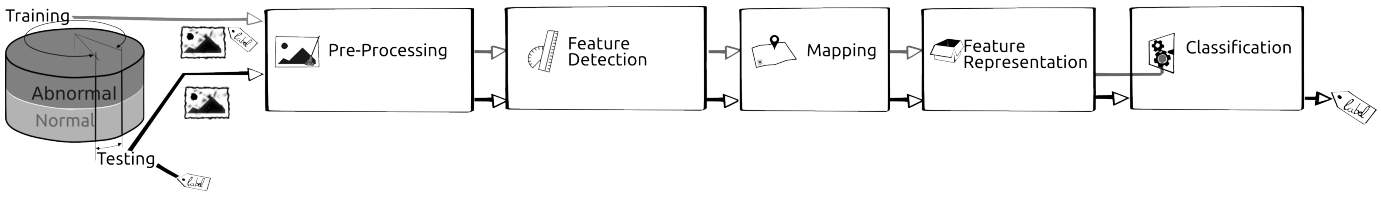



Fig. 2. Common framework


[3] Early Treatment Diabetic Retinopathy Study Group, "Photocoagulation for diabetic macular edema: early treatment diabetic retinopathy study report no 1," *JAMA Ophthalmology*, vol. 103, no. 12, pp. 1796–1806, 1985.

[4] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in sd-oct images congruent with expert manual segmentation," *Optic Express*, vol. 18, no. 18, pp. 19 413–19 428, 2010.

[5] R. Kafieh, H. Rabbani, M. D. Abramoff, and M. Sonka, "Intra-retinal layer segmentation of 3d optical coherence tomography using coarse grained diffusion map," *Medical Image Analysis*, vol. 17, pp. 907–928, 2013.

[6] F. G. Venhuizen, B. van Ginneken, B. Bloemen, M. J. P. P. van Grisven, R. Philipsen, C. Hoyng, T. Theelen, and C. I. Sanchez, "Automated age-related macular degeneration classification in OCT using unsupervised feature learning," in *SPIE Medical Imaging*, vol. 9414, 2015, p. 94141l.

[7] P. P. Srinivasan, L. A. Kim, P. S. Mettu, S. W. Cousins, G. M. Comer, J. A. Izatt, and S. Farsiu, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomedical Optical Express*, vol. 5, no. 10, pp. 3568–3577, 2014.

[8] G. Lemaitre, M. Rastgoo, J. Massich, S. Sankar, F. Meriaudeau, and D. Sidibe, "Classification of SD-OCT volumes with LBP: Application to dme detection," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI), Ophthalmic Medical Image Analysis Workshop (OMIA)*, 2015.

[9] J. Sivic and A. Zisserman, "Video google: a text retrieva approach to object matching in videos," in *IEEE ICCV*, 2003, pp. 1470–1477.

[10] Y.-Y. Liu, M. Chen, H. Ishikawa, G. Wollstein, J. S. Schuman, and R. J. M., "Automated macular pathology diagnosis in retinal oct images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding," *Medical Image Analysis*, vol. 15, pp. 748–759, 2011.