

CNN with Multiple Input for automatic glaucoma assessment using Fundus Images

ABDELALI ELMOUFIDI

Department of computer science, Faculty of Sciences and Technics, Sultan Moulay Slimane University
B.P: 523 Beni Mellal Morocco, Tel: 212 (0) 523 48 51 12 /22 /82, Fax: +212 (0) 523 48 52 01, a.elmoufidi@usms.ma

SAID JAI-ANDALOUSSI

Department of Mathematics and computer science, Faculty of science Ain Chock, Hassan II University, B.P 5366 Maarif 20100, Casablanca Morocco.

In the area of ophthalmology, glaucoma affects an increasing number of people. It is a major cause of blindness. Early detection avoids severe ocular complications such as glaucoma, cystoid macular edema, or diabetic proliferative retinopathy. Intelligent artificial has been confirmed beneficial for glaucoma assessment. In this paper, we describe an approach to automate glaucoma diagnosis using funds images. The setup of the proposed framework is, in order: The Bi-dimensional Empirical Mode Decomposition (BEMD) algorithm is applied to decompose the Regions of Interest (ROI) to components (BIMFs + residue). CNN architecture VGG19 is implemented to extract features from decomposed BEMD components. Then, we fuse the features of the same ROI in a bag of features. These last are very long; therefore, Principal Component Analyses (PCA) are used to reduce features dimensions. Obtained bags of features are the input parameters of the implemented classifier based on the Support Vector Machine (SVM). To train the built models, we have used two public datasets, which are ACRIMA and REFUGE. For testing our models, we have used a part of ACRIMA and REFUGE plus four other public datasets, which are RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF. The overall accuracy of 98.31%, 98.61%, 96.43%, 96.67%, 95.24%, and 98.60% are obtained on ACRIMA, REFUGE, RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF datasets, respectively, by using the model trained on REFUGE. Against an accuracy of 98.92%, 99.06%, 98.27%, 97.10%, 96.97%, and 96.36% are obtained on ACRIMA, REFUGE, RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF datasets, respectively, using the model training on ACRIMA. Obtained experimental results from different datasets demonstrate the efficiency and robustness of the proposed approach. A comparison with some recent previous work in the literature has shown a significant advancement in our proposal.

Keywords: Artificial intelligence; Ophthalmology; Glaucoma; Deep Learning; Bag of features; Image classification; Computer Aided Diagnosis.

1. Introduction

Glaucoma is an acute or chronic visual neuropathy, characterized by the progressive death of the ganglion cells of the retina, causing damage to the optic nerve and

the visual field. The final consequence of glaucoma is permanent blindness without recovery [1]. It is a serious and most important eye disease in ophthalmology, being the leading cause of blindness and low vision for adults with the incurable disease worldwide. Glaucoma is one of the most pressing problems in ophthalmology today. It is a progressive chronic disease that affects between 8-10% of people worldwide. Generally, the disease affects people over than 40 age, although it occurs in young people and even in childhood. In numbers, 65 million people worldwide living with glaucoma in 2010. This number rises to 79.6 million in 2020, increasing to 111.8 million individuals will have glaucoma in 2040 [2], [3]. As well as glaucoma, blindness goes from 6.7 million people in 2000 to 8.5 million people in 2010, then by estimate about 11 million in 2020 [3]. The constant increase in blindness and morbidity due to this disease in the context of socio-economic changes, determine the urgency of the glaucoma problem and improve the system for its detection, treatment, and monitoring of patients. This applies in particular to the most frequently diagnosed open-angle primary form [4-9]. To increase the effectiveness of prevention, early detection, and treatment of glaucoma, reduce disability, it is necessary to study automatic and weak decision support tools. In this context, several scientific pieces of research emphasize this subject as a main research axis. In the last decade, artificial intelligence (AI) has become a promising and essential field for the development of methods for health diagnostics, treatments, and interpretation of databases and medical images [2]. Machine learning is a class of artificial intelligence methods used to analyze complex data and find patterns and interdependencies without explicit programming [5, 6].

Machine learning algorithms analyze the signs of the input data and, through a series of repetitive operations, spread over several layers produce a linear and nonlinear prediction of the models that define the signals, classify the models, and predict the results [3-6]. There are many classic automatic learning methods, namely : random forest, learning associative rules, decision tree, inductive logic programming, nearest neighbor method, vector method support, artificial neural networks, etc. [5, 7-9]. In the last few years, convolutional neural networks take an important place, and its improvements increase from year to year. Typically, a neural network is made up of layers of interconnected nodes. The node contains a weighted sum, which is passed through the output link to the activation function and then to the next node. The values of the weighted amount change dynamically during the training phase. There are three types of layers:

- Input layer: The reception of input data;
- Hidden layers: highlighting patterns among the data;
- Output layer: transmission of information processing results.

For overcoming diagnosis difficulties, many automated methods have developed for glaucoma diagnosis. To know, the methods based on extraction and selection for glaucoma assessment [10-13] and the methods used deep learning approach for glaucoma diagnosis[14-19]. These developed methods have improved the diagnos-

tic quality, but the performance is not yet reaching the desired accuracy, as well as is not yet stable. As a result, researchers are still developing new techniques to improve diagnosis. In this work, we propose a new method to automate glaucoma assessment using fundus images. The setup of the proposed method is, in order: The BEMD algorithm is applied to decompose the image to components (BIMFs levels + residue). Then, CNN architecture VGG19 is called to extract features from decomposed BEMD components. Next, we fuse the extracted features of the same image in a bag of features. This last is very long; therefore, Principal Component Analyses (PCA) are used to reduce features dimensions. Selected bags of features are the input parameters of the implemented classifier based on the SVM. To train the built models, we have used two public datasets, which are ACRIMA and REFUGE. For testing our models, we have used a part of ACRIMA and REFUGE plus four other public datasets, which are RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF. The proposed method has shown a significant advancement.

The rest of this paper is organized as follows: in section 2, we present an overview of some previous related work. Section 3 introduces the used materials and the proposed methods. The experiments are introduced in section 4. The discussion is given in section 5. finished by a conclusion in section 6.

2. Related Previous Works

In recent years, machine vision has made spectacular leaps forward, notably thanks to recent breakthroughs in optimization and the explosion in the computing power of computers. What has been done for facial recognition, for example, the research community is trying to replicate in the medical field. Artificial intelligence, also known as AI, can process thousands of images in seconds and detect with great precision important information that a radiologist would have taken months to find. For this reason many computer-aided diagnoses has been developed to help physician improving the diagnosis results in several medical fields, such as breast cancer [20], [21] [22], brain cancer [23], retinopathies diabetic [24], [25], etc. In this paper, we have implemented the BEMD and a CNN architecture to automate glaucoma assessment using fundus images.

2.1. *Bi-dimensional Empirical Mode Decomposition*

Empirical Mode Decomposition (EMD)[26] is an approach adapted to non-stationary time-frequency analysis and has been applied in several fields, namely, biological [27], marine environment [28], structural diagnosis [29], mechanical diagnosis faults [30], [31], and other domains based on one dimensional signal. In addition, Nunes et al. [32], [33] have Introduced EMD in image processing (two-dimensional signal) and they have developed the algorithm for two-dimensional decomposition in empirical mode (BEMD). This algorithm has caught the attention of several researchers and has been applied in image denoising [34], image compression [35],[36], image segmentation [37], [38], scaling the image [39], extraction of image

characteristics [40], texture synthesis [41], and classification of image texture [42]. There are many methods of decomposition signals and images (DCT, Wavelets, Fourier, etc.) which assume decomposition on a basis given a priori. The major advantage of empirical modal decomposition (EMD) is to define decompositions of the images which do not depend on the choice of a particular base. In addition, the EMD is particularly well suited for the study of non-stationary signals. We will present the original EMD algorithm whose efficiency is recognized by the applications it allows to process. Similar to the Empirical Mode Decomposition (EMD) of the one-dimensional signal, the BEMD for two-dimensional images based on the extreme that exists in the original image or obtained from the first derivative of the original or higher-order derivative, to achieve the decomposition of the image signal. Distances between extrema may provide the information describing the image on intrinsic length scales. The Empirical Mode Decomposition (EMD) method is an adaptive decomposition that allows decomposing any signal into a redundant set of signals intrinsic mode functions denoted BIMF [43]. In this paper, we have implemented the BEMD to describe regions in order to make the classification phase very easy given the use of several components in the same region.

2.2. *Glaucoma diagnosis*

The following section introduces the related recently previously developed methods of retinal abnormalities classification. Numerous promising approaches are coming up for solving the problems of retinal abnormalities classification in fundus images. In the previous few years, hard efforts had been made to develop automated systems to assist practitioners to detect and diagnose retinal abnormalities using fundus images. The use of computer-aided diagnosis has become increasingly relevant to medical decision helping; recently artificial intelligence methods have demonstrated significant advances in the medical area. In this section, we introduce a simple overview of previous and current work developed artificial intelligence tools and their applications in the retinopathy field, to help the practitioner's ophthalmologists early diagnose retinal abnormalities generally and glaucoma specifically, also understand their potential impact on glaucoma care. Diaz-Pinto, Andres, et al [14] have developed an approach for automatic glaucoma assessment using five different ImageNet-trained models (VGG16, VGG19, InceptionV3, ResNet50, and Xception), five different databases based on funds image. In this work, the authors applied directly CNN with different architecture and without passed by the segmentation method and the extracted features to automate glaucoma assessment. Their proposal gives good results generally (an average AUC of 0.9605, an average specificity of 0.8580, and an average sensitivity of 0.9346), but gives modest results and no stability when they change testing set (an AUC of 0.7678, an accuracy of 0.7021, a Sensitivity of 0.6893, and a Specificity of 0.7020) by using ACRIMA dataset. Chakravarty, Arunava, and Jayanthi Sivaswamy [10] have published an article that trait the subject of glaucoma classification by using the classical method,

regions of interest segmentation, feature extraction then classification. The novelty in this work that the authors used different features and fusing them for a more ROIs description. Obtained results in the test set are too modest (a Sensitivity of 80%, a Specificity of 60%, an accuracy of 76.77%, and an AUC of 78%) using DRISHTI-GS1. Maheshwari, Shishir [11] Have presented a method to automate glaucoma diagnosis using Empirical Wavelet Transform (EWT) based on fundus images. In this paper, the authors have applied The EWT to decompose the image, and then features are extracted from decomposed EWT components. Selected features have ranked based on the t-value feature selection algorithm. Obtained features are used to classify images as normal or glaucoma using Least Squares Support Vector Machine (LS-SVM) classifier. The LS-SVM is implemented with Morlet wavelet, Mexican-hat wavelet kernels, and Radial Basis Function (RBF). The proposed method is achieving a classification accuracy of 98.33% and 96.67% using three-fold and ten-fold cross-validation respectively. This work achieved high performance, except that used databases are privates, therefore, we cannot compare their proposal this the existing in the literature.

3. Materials and methods

In this section and their subsection, we introduce the used materials and methods. First, we present the used datasets (ACRIMA, REFUGE, RIM-ONE, Drishti-GS1, ORIGA-light, and sjchoi86-HRF), next by the proposed approach.

3.1. Materials

3.1.1. Used databases

To train, validate and testing the proposed approach, we have used six publicly available databases. Witch are: ACRIMA [14], REFUGE [15], RIM-ONE [44], Drishti-GS1 [45], ORIGA-light [46], and sjchoi86-HRF [47]. Which are represent database reference for the majority of research work in the field. Table 1 below presents a description and statistics about using databases.

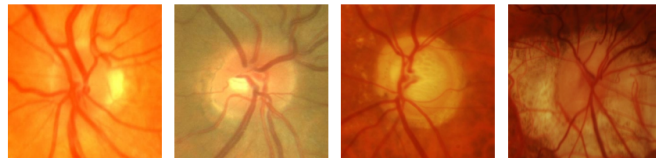


Fig. 1 Example of the used fundus images from ACRIMA database

Table 1 List of the public available databases with glaucoma labels, used to train, validate and test the proposed method.

Datasets	Glau	Nor	Total	Digitizer	Resolutions	Source	Availa
ACRIMA	396	309	705	Topcon TRC	JPEG 2048x1536	Diaz-Pinto et al. [14]	Online
REFUGE	120	1080	1200	Train: Zeiss Visucam 500 Val/Test: Canon CR-2	JPEG 2124x2056 JPEG 1634x1634	Orlando et al. [15]	Online
RIM-ONE	194	261	455	Kowa WX 3D	JPEG 2144x1424	Fumero et al. [44]	Online
Drishti-GS1	70	31	101	NM/FA	PNG 2896x1944	Sivaswamy et al. [45]	Online
ORIGA-light	168	482	650	NR*	JPEG 3072x2048	Zhang et al. [46]	Online
sjchoi86-HRF	101	300	401	NR*	JPEG 2592x1728, 2464x1632 or 1848x1224	[47]	Online
Total	1049	2463	3512	***	***	***	***

NR*: Not Reported; Glau:Glaucoma Nor: Normal Val : Validation
Test: Testing Availa:Available

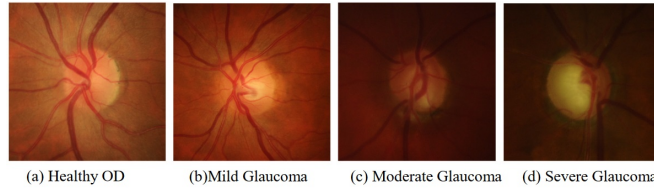


Fig. 2 Healthy OD and glaucoma diseases stages in retinal fundus images: (a) Healthy Disc; (b) Mild Glaucoma; (c) Moderate Glaucoma (d) severe glaucoma.

3.1.2. *K-fold cross-validation*

Learning the parameters of a prediction function and testing it on the same data is a methodological error: a model that would simply repeat the labels of the samples we have just seen would have a perfect score but could not predict anything useful on unseen data. This situation is called overfitting. To avoid this, it is common during the execution of a machine learning experience (supervised) to keep some of the available data in the form of a test set. Therefore, in this work, we have used the 5-fold cross-validation technique to avoid overfitting.

3.2. *Proposed Methodology*

The proposed method's framework is shown in Figure 2. The proposed method steps are: first, we apply the BEMD algorithm to decompose ROI to components (BIMFs+Residue), these components are the inputs of CNN architecture VGG19. Then, we extract the features of the input components. Next, we fuse the features of the components of the same ROIs in a bag of features. This list is very long; therefore, PCA is used to reduce the feature dimension and select optimal features. Finally, we apply binary classification by implementing SVM to classify ROI on

funds images to normal or glaucoma. The following subsections discuss the proposed method in detail.

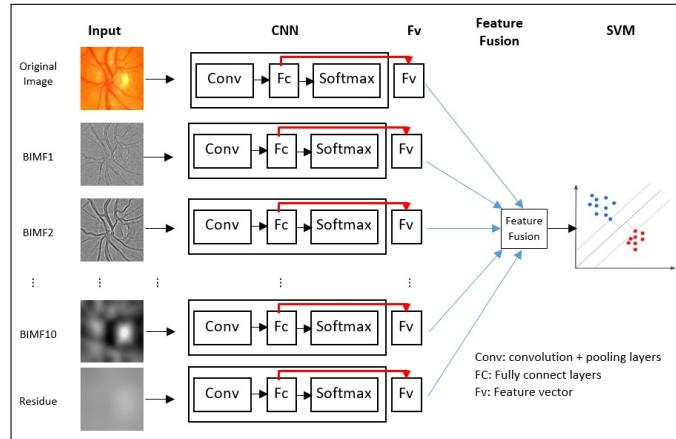


Fig. 3 Architecture description of the proposed approach for glaucoma diagnosis, by using (1) CNN-based deep learning to conduct feature learning, and (2) an SVM for the final prediction (Normal or glaucoma)

3.2.1. BEMD applied

Due to glaucoma disease, obviously optic disk structure, cup disk structure are changing, therefore the corresponding images change, which results in intensity, texture, and even shape changes. Sometimes these changes are not too clear and readable. Some details are hidden in digital images; hence, the use of BEMD is useful in order to extract more detail. First, the BEMD algorithm was applied to decompose ROIs to a list of BIMF according to different intensity levels plus a residue. The BIMF and residue will name components. Each layer is considered as a CNN input. Figure 3 to Figure 6 present four (4) different examples. Figure 3 and figure 4 using normal fundus image; figure 5 and figure 6 using glaucoma fundus images.

3.2.2. Feature Extraction by CNN

After BEMD algorithm applied to decompose image into components (BIMF +Residue). CNN architecture VGG19 is used to extract feature from input data.

Feature extraction and selection to describe candidates regions in order to classify them is a crucial task, and it affects directly classification accuracy. Recently, the convolutional neural network (CNN) architecture is very useful to design powerful filters to extract the sensitive image features for the classification task. Therefore, is resolved this problem almost definitively. In this paper, we implement the CNN

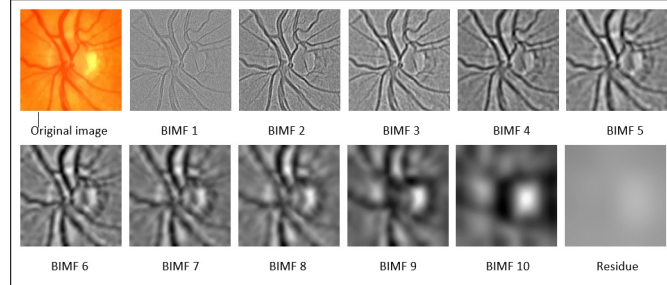


Fig. 4 Example 1: Normal retina image: Image decomposed BEMD components in the order of high to low level intensities from left to right in each row

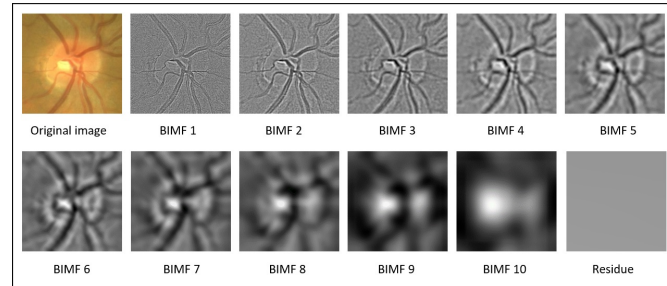


Fig. 5 Example 2: Normal retina image : Image decomposed BEMD components in the order of high to low level intensities from left to right in each row

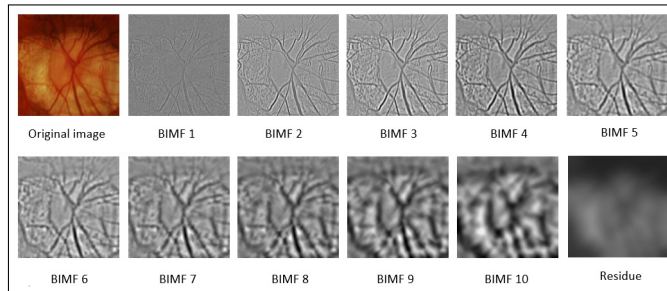


Fig. 6 Example 3: Glaucoma retina image : Image decomposed BEMD components in the order of high to low level intensities from left to right in each row

architecture VGG19, it is a CNN model with 19 layers. This model is widely used to extract image features and solve image classification problems.

To classify healthy and glaucomatous eyes. We build a list of parallel models in a total of 12 VGG19 models with different inputs. The input of the first model is the original candidate region; the inputs of model 2 to model 11 are the BIMFs of ROI: BIMF1 to BIMF10 respectively, and the input of the last model is the ROI

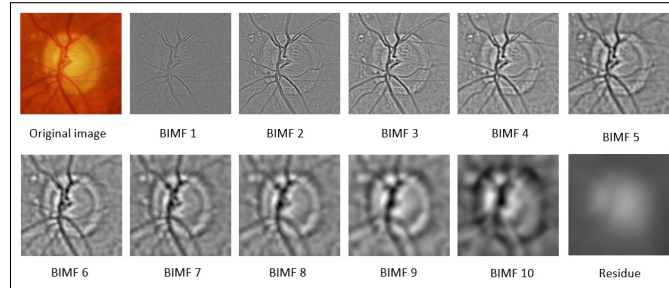


Fig. 7 Example 4 : Glaucoma retina image : Image decomposed BEMD components in the order of high to low level intensities from left to right in each row

residue. We have used 10 intensities-scale (BIMFs) because in the three last BEMD components (BIMF8, BIMF9, and BIMF10) we observed a change between healthy and glaucomatous eyes). Then the full-connected layers of each VGG19 model have extracted, and fusion into a bag of features. Data augmentation and dropout have used during training.

3.2.3. Feature Fusion and Dimensionality Reduction

After features extraction of input data, feature vectors of the same region are concatenated to form a higher dimensional bag of features. Generally, there is some correlation and many information redundancies among these feature and higher dimensional features lead to higher computation complexity. Therefore, we applied principle component analysis (PCA) for dimensionality reduction and selection of features. In this paper, we define the eigenvalue statistical rates as the ratio of the number of principal components (eigenvalues) retained by PCA to the number of all components. Extracting features from the image helps to fully learn the attributes for the detailed description of their rich internal information, except that a few are unnecessary or redundant, resulting in additional computational complexity. Therefore a study for dimensionality reduction is desirable. Through this dimensionality reduction operation, we can achieve a compact, optimal description of the built-in features, resulting in lower computational complexity and better image detection and classification performance in a stress-free environment.

3.2.4. Binary Classification Using SVM

The traditional SVM is a non-probabilistic linear classifier. In practice, it goes through two phases: a learning phase and a test phase. The learning phase constructs a model that will classify future unknown data. In the test phase, you will use the structure created previously to classify your unknown samples and finally, you can calculate the accuracy of the classifier.

4. Experimental result on the used dataset

Early glaucoma diagnosis remains a challenging task actually. Although CAD systems for early glaucoma diagnosis using fundus images have been developed aiming to assist ophthalmologists to distinguish between glaucoma and normal retina. To assess the effectiveness of the proposed approach on large-scale datasets, we applied it to many different datasets.

4.1. Train model on the used dataset

In this section, we present the trained models using the ACRIMA and REFUGE datasets. We have trained the proposed models using 50 and 100 epochs respectively, and 5-folds cross-validation to resolve the overfitting problem. Randomly 60% images of the used dataset are selected for training the model, 20% images for the validation phase, and for the remaining 20% images are used for testing. The used datasets are not too large. Therefore, to enlarge them rotation data augmentation technique is employed. Below, we introduce the training model detail by using the ACRIMA dataset. [Figure 9](#) and [figure 11](#) show the last parts of training processes of the proposed model with 50 and 100 epochs respectively. [Figure 8](#) and [Figure 10](#) show obtained accuracies and losses in different cases. Therefore, the same way we have retrained the model used the REFUGE datasets. The goal is to verify the robustness of our proposed approach. [Table 2](#) summarises the obtained quantitative performances by using two models which are: the first one trained using the ACRIMA dataset, and the second one trained using the REFUGE dataset.

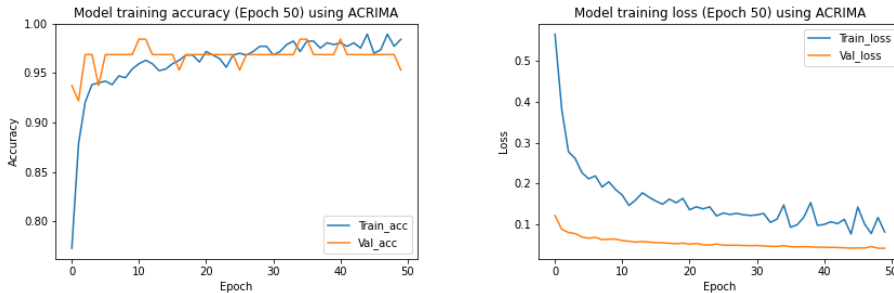


Fig. 8 Model training performance using ACRIMA dataset with 50 epochs : 1) Model accuracy performance (left), 2) Model loss performance(right).

4.2. Test model on the used dataset

[Tables 2](#) and [3](#) below show the obtained quantitative results in the testing phase using the two training models formed. Which are the model trained on the ACRIMA

```

568/568 [=====] - 23s 41ms/step - loss: 0.0936 - bag_accuracy: 0.9824 - val_loss: 0.0472 - val_bag_accuracy: 0.9219
Epoch 00047: val_loss improved from 0.04731 to 0.04719, saving model to Saved_model/_Batch_size_1epoch_best.hd5
Epoch 48/50
568/568 [=====] - 23s 41ms/step - loss: 0.0927 - bag_accuracy: 0.9824 - val_loss: 0.0463 - val_bag_accuracy: 0.9844
Epoch 00048: val_loss improved from 0.04719 to 0.04630, saving model to Saved_model/_Batch_size_1epoch_best.hd5
Epoch 49/50
568/568 [=====] - 23s 41ms/step - loss: 0.0930 - bag_accuracy: 0.9806 - val_loss: 0.0459 - val_bag_accuracy: 0.9844
Epoch 00049: val_loss improved from 0.04630 to 0.04591, saving model to Saved_model/_Batch_size_1epoch_best.hd5
Epoch 50/50
568/568 [=====] - 23s 41ms/step - loss: 0.0948 - bag_accuracy: 0.9824 - val_loss: 0.0464 - val_bag_accuracy: 0.9844

```

Fig. 9 Screenshot represents a last part of the training process using the proposed model in the ACRIMA dataset with 50 epochs.

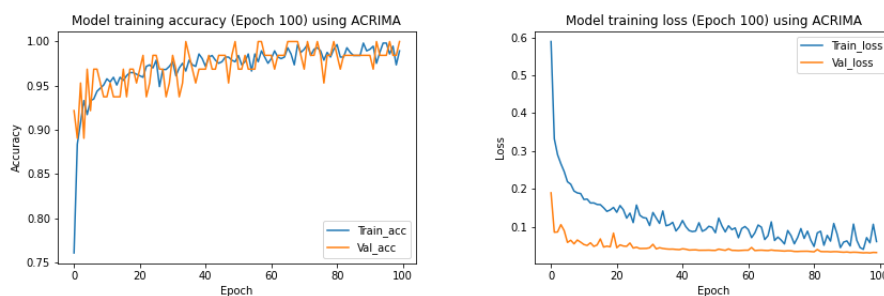


Fig. 10 Model training performance using ACRIMA dataset with 100 epochs : 1) Model accuracy performance (left), 2) Model loss performance(right).

```

Epoch 97/100
568/568 [=====] - 13s 23ms/step - loss: 0.0664 - bag_accuracy: 0.9912 - val_loss: 0.0332 - val_bag_accuracy: 0.9219
Epoch 00097: val_loss did not improve from 0.03315
Epoch 98/100
568/568 [=====] - 13s 23ms/step - loss: 0.0449 - bag_accuracy: 0.9965 - val_loss: 0.0325 - val_bag_accuracy: 0.9531
Epoch 00098: val_loss improved from 0.03315 to 0.03254, saving model to Saved_model/_Batch_size_1epoch_best.hd5
Epoch 99/100
568/568 [=====] - 13s 23ms/step - loss: 0.0399 - bag_accuracy: 0.9982 - val_loss: 0.0319 - val_bag_accuracy: 0.9531
Epoch 00099: val_loss improved from 0.03254 to 0.03185, saving model to Saved_model/_Batch_size_1epoch_best.hd5
Epoch 100/100
568/568 [=====] - 13s 23ms/step - loss: 0.0901 - bag_accuracy: 0.9789 - val_loss: 0.0328 - val_bag_accuracy: 0.9219

```

Fig. 11 Screenshot represents a last part of the training process using the proposed model in the ACRIMA dataset with 100 epochs

dataset, and the model trained on the REFUGE dataset, respectively. In this work, we have trained the proposed method using two different datasets to study the stability of our proposal.

Table 3 reveals the obtained results using the model trained on the ACRIMA dataset, then testing on ACRIMA, REFUGE, RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF datasets.

Table 4 reveals the obtained results using the model trained on the REFUGE dataset, then testing on ACRIMA, REFUGE, RIM-ONE, ORIGA-light, Drishti-

Table 2 Glaucoma diagnosis performance - The training average accuracy and minimal loss using different databases

Epochs	ACRIMA				REFUGE			
	Accuracy	loss	V_{acc}	V_{loss}	Accuracy	loss	V_{acc}	V_{loss}
50	0.9824	0.0926	0.9844	0.0459	0.9890	0.0812	0.9853	0.0406
100	0.9905	0.0514	0.9489	0.0310	0.9911	0.442	0.9517	0.0290

V_{acc} : Validation accuracy

V_{loss} : Validation loss

GS1, and sjchoi86-HRF datasets.

Table 3 Glaucoma diagnosis performance: Obtained performance, using REFUGE dataset to train and others datasets to test.

	Training set: ACRIMA			
Test set	Sen	Spe	Acc	AUC
ACRIMA	0.9840	0.9826	0.9831	0.9830
REFUGE	0.9857	0.9874	0.9862	0.9860
RIM-ONE	0.9856	0.9867	0.9861	0.9859
DRISTI-GS1	0.9612	0.9680	0.9643	0.9675
ORIGA-light	0.9665	0.9671	0.9667	0.9672
sjchoi86-HRF	0.9542	0.9508	0.9524	0.9538

5. Discussion

We tested the proposed approach using 3512 retina images, whose 1049 are glaucomatous and 2463 are normal images. The images were obtained from six publicly available datasets, which are ACRIMA, REFUGE, RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF dataset. We have trained and on-line tested two models, the first one by using the ACRIMA dataset and the second one using the REFUGE dataset. We have built two different models to confirm the stability of our proposed approach. In the testing phase, we have used different datasets, to ensure that the result of the glaucoma assessment is not biased due to the change of the dataset in the training and testing phases. The receiver operating characteristic (ROC) curve, sensitivity, specificity, and balance accuracy were used as the performance metrics. The two built models are trained with, 50 and 100, epochs. They, to demonstrate that obtained results in the validation phase with 50 epochs

Table 4 Glaucoma diagnosis performance: Obtained average test accuracy and standard error, using ACRIMA dataset to train and others datasets to test.

Test set	Training set: REFUGE			
	Sen	Spe	Acc	AUC
ACRIMA	0.9887	0.9898	0.9892	0.9898
REFUGE	0.9904	0.9909	0.9906	0.9918
RIM-ONE	0.9837	0.9819	0.9827	0.9842
DRISTI-GS1	0.9717	0.9705	0.9710	0.9725
ORIGA-light	0.9708	0.9688	0.9697	0.9700
sjchoi86-HRF	0.9627	0.9641	0.9636	0.9658

Sen: Sensitivity; Spe: Specificity; Acc: Accuracy
AUC: Area Under the ROC Curve

are best. Therefore, the implemented approach return training phase very easy and very quick. [Table 3](#) summarizes the obtained performance using the model trained with the ACRIMA dataset. In the testing phase, an accuracy of 98.31%, 98.61%, 96.43%, 96.67%, 95.24%, and 98.60% are obtained on ACRIMA, REFUGE, RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF datasets, respectively. The sensitivity, specificity, and AUC are figured in [table 3](#). In another word, [table 4](#) summarizes the obtained performance using the model trained with the REFUGE dataset. In the testing phase, an accuracy of 98.92%, 99.06%, 98.27%, 97.10%, 96.97%, and 96.36% are obtained on ACRIMA, REFUGE, RIM-ONE, ORIGA-light, Drishti-GS1, and sjchoi86-HRF datasets, respectively. The sensitivity, specificity, and AUC are figured in [table 4](#). We have found that the proposed method gives a similar performance with different trained models. Thanks to the diversified data, the proposed method has shown excellent performance in various training, testing, and validation datasets. Therefore, it is stable and robust.

Many similar studies on CAD system for glaucoma diagnosis has recently been published [[10](#)],[[11](#)], [[14-19](#)], [[47](#)], [[48](#)] these studies allow us make a comparison in order to assess the performance of the proposed appoche. [Table 5](#) below shows a comparison of the obtained results in this work and the results of previous works. Our approach is slightly better.

6. Conclusion

In this proposal, we have addressed the problem of glaucoma disease diagnosis using digital fundus images. The objective is to automate the diagnosis based on the opposed method and using digital fundus images only and without having to resort to additional examinations. The framework of our method is: first, the BEMD algorithm is applied to decompose ROI under consideration to the intensity scale

Table 5 Quantitative results comparison of the proposed method and related methods on test sets.

Dataset	Methodology	Classifier	Sensitivity	Specificity	Accuracy	AUC
RIM-ONE	Maheshwari et al. (2016) [11]	SVM	100%	96.67%	98.33%	NR*
	Gómez-Valverde et al. (2019) [17]	VGG19 TL	87.01%	89.01%	88.05%	94.20%
	Diaz-Pinto et al. (2019) [14]	Xception architecture	79.31%	79.90 %	71.21 %	85.75%
	Sreng, Syna, et al. (2020) [19]	CNNs + transfer learning	NR*	NR*	97.37%	100%
		CNNs features + SVM	NR*	NR*	92.11%	99.04%
The proposed method	CNNs features + SVM	98.37%	98.19%	98.27%	98.42%	
DRISTI-GS1	Chakravarty et al.(2017) [10]	SVM	80%	60%	76.77%	78%
	Orlando et al. (2017) [18]	CNNs	NR*	NR*	NR*	76.26%
	Diaz-Pinto et al. (2019) [14]	Xception architecture	74.19%	71.43%	75.25%	80.41%
	Sreng, Syna, et al. (2020) [19]	CNNs + transfer learning	NR*	NR*	83.33%	85.19%
		CNNs features + SVM	NR*	NR*	90%	92.06%
The proposed method	CNNs features + SVM	96.12%	96.80%	96.43%	96.75%	
ORIGA	GUO et al. (2018) [48]	RF+SMOTE	79.9%	73.8%	76.90%	83.10%
	Bajwa et al.(2019) [16]	RCNN)	71.17%	NR*	NR*	87.40%
	Sreng, Syna, et al. (2020) [19]	CNNs + transfer learning	NR*	NR*	83.59%	88.86%
		CNNs features + SVM	NR*	NR*	80.00%	85.26%
	The proposed method	CNNs features + SVM	96.65%	96.71%	96.67%	96.67%
sjchoi86-HRF	Diaz-Pinto et al. (2019) [14]	Xception architecture	70.33%	70.30%	70.82%	77.39%
	The proposed method	CNNs features + SVM	95.42%	95.08%	95.24%	95.38%
ACRIMA	Diaz-Pinto et al. (2019) [14]	Xception architecture	68.93%	70.20%	70.21%	76.78%
	Sreng, Syna, et al. (2020) [19]	CNNs + transfer learning	NR*	NR*	99.53%	99.98%
		CNNs features + SVM	NR*	NR*	96.23%	99.01%
The proposed method	CNNs features + SVM	98.87%	98.98%	98.92%	98.98%	
REFUGE	Orlando et al. (2020) [15]	Deep learning approaches	97.52%	NR*	NR*	98.85%
	Sreng, Syna, et al. (2020) [19]	CNNs + transfer learning	NR*	NR*	95.59%	95.10%
		CNNs features + SVM	NR*	NR*	95.75%	94.32%
The proposed method	CNNs features + SVM	99.04%	99.09%	99.06%	99.18%	

NR*= Not Reported; CNNs= Convolutional Neural Network

(BIMFs + residue). Second, the CNN architecture VGG19 is implemented to feature extraction, then concatenated them with a bag of features. Third, the ACP method is adopted to reduce features dimensionality. Finally, the SVM classifier called for the classification task. The proposal has shown robustness despite the diversity of the dataset used either for training or for testing. As well as the results obtained have compared with others recently published in the literature and they showed a significant improvement in the diagnostic performance of glaucoma retinal abnormality.

Nomenclature

AI	Artificial Intelligence
BEMD	Bidimensional Empirical Mode Decomposition
BIMF	Bidimensional Intrinsic Mode Functions;
CAD	Computer Aided Diagnosis;
CNN	Convolutional Neural Network ;
PCA	Principal Component Analyses
ROI	Regions of Interest
SVM	Support Vector Machine
VGG	Visual Geometry Group;

References

1. Weinreb, Robert N., Tin Aung, and Felipe A. Medeiros. "The pathophysiology and treatment of glaucoma: a review." *Jama* 311.18 (2014): 1901-1911.
2. Tham, Yih-Chung, et al. "Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis." *Ophthalmology* 121.11 (2014): 2081-2090.
3. Quigley, Harry A., and Aimee T. Broman. "The number of people with glaucoma worldwide in 2010 and 2020." *British journal of ophthalmology* 90.3 (2006): 262-267.
4. Pesapane, Filippo, Marina Codari, and Francesco Sardanelli. "Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine." *European radiology experimental* 2.1 (2018): 35.
5. Lakhani, Paras, et al. "Machine learning in radiology: applications beyond image interpretation." *Journal of the American College of Radiology* 15.2 (2018): 350-359.
6. Kapoor, Rahul, Stephen P. Walters, and Lama A. Al-Aswad. "The current state of artificial intelligence in ophthalmology." *Survey of ophthalmology* 64.2 (2019): 233-240.
7. Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas. "Supervised machine learning: A review of classification techniques." *Emerging artificial intelligence applications in computer engineering* 160.1 (2007): 3-24.
8. Liaw, Andy, and Matthew Wiener. "Classification and regression by random forest." *R news* 2.3 (2002): 18-22.
9. Rosenblatt, Frank. "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review* 65.6 (1958): 386.
10. Chakravarty, Arunava, and Jayanthi Sivaswamy. "Glaucoma classification with a fusion of segmentation and image based features." 2016 IEEE 13th international symposium on biomedical imaging (ISBI). IEEE, 2016.
11. Maheshwari, Shishir, Ram Bilas Pachori, and U. Rajendra Acharya. "Automated diagnosis of glaucoma using empirical wavelet transform and correntropy

- features extracted from fundus images." *IEEE journal of biomedical and health informatics* 21.3 (2016): 803-813.
12. Acharya, U. Rajendra, et al. "Automated diagnosis of glaucoma using texture and higher order spectra features." *IEEE Transactions on information technology in biomedicine* 15.3 (2011): 449-455.
 13. Dua, Sumeet, et al. "Wavelet-based energy features for glaucomatous image classification." *Ieee transactions on information technology in biomedicine* 16.1 (2011): 80-87.
 14. Diaz-Pinto, Andres, et al. "CNNs for automatic glaucoma assessment using fundus images: an extensive validation." *Biomedical engineering online* 18.1 (2019): 29.
 15. Orlando, José Ignacio, et al. "Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs." *Medical image analysis* 59 (2020): 101570.
 16. Bajwa, Muhammad Naseer, et al. "Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning." *BMC medical informatics and decision making* 19.1 (2019): 136.
 17. Gómez-Valverde, Juan J., et al. "Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning." *Biomedical optics express* 10.2 (2019): 892-913.
 18. Orlando, José Ignacio, et al. "Convolutional neural network transfer for automated glaucoma identification." *12th international symposium on medical information processing and analysis*. Vol. 10160. International Society for Optics and Photonics, 2017.
 19. Sreng, Syna, et al. "Deep Learning for Optic Disc Segmentation and Glaucoma Diagnosis on Retinal Images." *Applied Sciences* 10.14 (2020): 4916.
 20. Elmoufidi, Abdelali, et al. "Anomaly classification in digital mammography based on multiple-instance learning." *IET Image Processing* 12.3 (2017): 320-328.
 21. Elmoufidi, Abdelali, et al. "Detection of regions of interest's in mammograms by using local binary pattern, dynamic k-means algorithm and gray level co-occurrence matrix." *2014 International Conference on Next Generation Networks and Services (NGNS)*. IEEE, 2014.
 22. Elmoufidi, Abdelali. "Pre-Processing Algorithms on Digital X-ray Mammograms." *2019 IEEE International Smart Cities Conference (ISC2)*. IEEE, 2019.
 23. El-Dahshan, El-Sayed A., et al. "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm." *Expert systems with Applications* 41.11 (2014): 5526-5545.
 24. Skouta, Ayoub, et al. "Automated Binary Classification of Diabetic Retinopathy by Convolutional Neural Networks." *Advances on Smart and Soft Computing*. Springer, Singapore 177-187.
 25. Gulshan, Varun, et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs." *Jama*

- 316.22 (2016): 2402-2410.
26. Wu, Zhaohua, and Norden E. Huang. "A study of the characteristics of white noise using the empirical mode decomposition method." *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 460.2046 (2004): 1597-1611.
 27. Huang, Wei, et al. "Use of intrinsic modes in biology: examples of indicial response of pulmonary blood pressure to \pm step hypoxia." *Proceedings of the National Academy of Sciences* 95.22 (1998): 12766-12771.
 28. Song, Haibin, et al. "Analysis of ocean internal waves imaged by multichannel reflection seismics, using ensemble empirical mode decomposition." *Journal of Geophysics and Engineering* 9.3 (2012): 302-311.
 29. Garcia-Perez, Arturo, et al. "Fused empirical mode decomposition and wavelets for locating combined damage in a truss-type structure through vibration analysis." *Journal of Zhejiang University SCIENCE A* 14.9 (2013): 615-630.
 30. Zheng, Jinde, Junsheng Cheng, and Yu Yang. "Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis." *Mechanical Systems and Signal Processing* 40.1 (2013): 136-153.
 31. Zhu, Keheng, Xigeng Song, and Dongxin Xue. "Incipient fault diagnosis of roller bearings using empirical mode decomposition and correlation coefficient." *Journal of Vibroengineering* 15.2 (2013): 597-603.
 32. Nunes, Jean Claude, et al. "Image analysis by bidimensional empirical mode decomposition." *Image and vision computing* 21.12 (2003): 1019-1026.
 33. Nunes, Jean Claude, Steve Guyot, and Eric Deléchelle. "Texture analysis based on local analysis of the bidimensional empirical mode decomposition." *Machine Vision and applications* 16.3 (2005): 177-188.
 34. Zhou, Yi, and Hongguang Li. "Adaptive noise reduction method for DSPI fringes based on bi-dimensional ensemble empirical mode decomposition." *Optics express* 19.19 (2011): 18207-18215.
 35. Linderhed, Anna. "2D empirical mode decompositions in the spirit of image compression." *Wavelet and Independent Component Analysis Applications IX*. Vol. 4738. International Society for Optics and Photonics, 2002.
 36. Linderhed, Anna. "Compression by image empirical mode decomposition." *IEEE International Conference on Image Processing 2005*. Vol. 1. IEEE, 2005.
 37. Qiao, Lihong, et al. "Perfect reconstruction image modulation based on BEMD and quaternionic analytic signals." *Science China Information Sciences* 54.12 (2011): 2602-2614.
 38. Chen, Ying, et al. "Fusion of color microscopic images based on bidimensional empirical mode decomposition." *Optics express* 18.21 (2010): 21757-21769.
 39. LINDERHED, Anna. *Adaptive image compression with wavelet packets and empirical mode decomposition*. Univ., 2004.
 40. Yang, Bo Suk, Fengshou Gu, and Andrew Ball. "Thermal image enhancement using bidimensional empirical mode decomposition in combination with rele-

- vance vector machine for rotating machinery fault diagnosis." *Mechanical Systems and Signal Processing* 38.2 (2013): 601-614.
41. Liu, Zhongxuan, and Silong Peng. "Boundary processing of bidimensional EMD using texture synthesis." *IEEE Signal Processing Letters* 12.1 (2004): 33-36.
 42. He, Zhi, et al. "Multivariate gray model-based BEMD for hyperspectral image classification." *IEEE Transactions on Instrumentation and Measurement* 62.5 (2013): 889-904.
 43. Huang, Norden E., et al. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences* 454.1971 (1998): 903-995.
 44. Leppänen, Tomi, et al. "Augmenting Microsurgical Training: Microsurgical Instrument Detection Using Convolutional Neural Networks." *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2018.
 45. Sivaswamy, Jayanthi, et al. "Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation." *2014 IEEE 11th international symposium on biomedical imaging (ISBI)*. IEEE, 2014.
 46. Zhang, Zhuo, et al. "Origa-light: An online retinal fundus image database for glaucoma analysis and research." *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE, 2010.
 47. sjchoi86: sjchoi86-HRF Database. GitHub. 2017. Accessed 2 Feb 2017.
 48. Guo, Fan, et al. "Yanbao: a mobile app using the measurement of clinical parameters for glaucoma screening." *IEEE Access* 6 (2018): 77414-77428.