

RESEARCH

Open Access



Wavelet image scattering based glaucoma detection

Hafeez Alani Agboola* and Jesuloluwa Emmanuel Zaccheus

Abstract

Background The ever-growing need for cheap, simple, fast, and accurate healthcare solutions spurred a lot of research activities which are aimed at the reliable deployment of artificial intelligence in the medical fields. However, this has proved to be a daunting task especially when looking to make automated diagnoses using biomedical image data. Biomedical image data have complex patterns which human experts find very hard to comprehend. Against this backdrop, we applied a representation or feature learning algorithm: Invariant Scattering Convolution Network or Wavelet scattering Network to retinal fundus images and studied the efficacy of the automatically extracted features therefrom for glaucoma diagnosis/detection. The influence of wavelet scattering network parameter settings as well as 2-D channel image type on the detection correctness is also examined. Our work is a distinct departure from the usual method where wavelet transform is applied to pre-processed retinal fundus images and handcrafted features are extracted from the decomposition results. Here, the RIM-ONE DL image dataset was fed into a wavelet scattering network developed in the Matlab environment to achieve a stage-wise decomposition process called wavelet scattering of the retinal fundus images thereby, automatically learning features from the images. These features were then used to build simple and computationally cheap classification algorithms.

Results Maximum detection correctness of 98% was achieved on the held-out test set. Detection correctness is highly sensitive to scattering network parameter setting and 2-D channel image type.

Conclusion A superficial comparison of the classification results obtained from our work and those obtained using a convolutional neural network underscores the potentiality of the proposed method for glaucoma detection.

Keywords Wavelet scattering network, Scattering features, Glaucoma detection, Retinal fundus image

Introduction

Glaucoma constitute a weighty eye problem plaguing the human race. It is a group of diseases in which the optic nerve is impaired resulting in unalterable loss of vision. In majority of cases, this impairment is due to an increased intra ocular pressure within the eye. The ciliary body in the eye secretes a fluid known as aqueous humour into the space between the iris and the lens (i. e. the posterior

chamber). The fluid then flows through the pupil into the space between the iris and the cornea (i. e. the anterior chamber) from where it drains through the trabecular meshwork, a spongy like structure at the base of the eye. A steady flow is established in a healthy eye as the rate of secretion balances the rate of drainage. In contrast, rate of drainage does not match the secretion rate in the unhealthy eye giving rise to accumulation of fluid in the anterior chamber. As accumulation increases, pressure builds up within the eye and the optic nerve which carries visual signals to the brain gets damaged leading to a permanent loss of vision. There are two main types of glaucoma. The open angle glaucoma and the angle closure glaucoma. In open angle glaucoma, the drainage

*Correspondence:

Hafeez Alani Agboola
hafeezagboola@abuad.edu.ng
Afe Babalola University, Ado-Ekiti, Nigeria



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

canal (i.e. trabecular meshwork) is open while in angle closure glaucoma, the iris bulges forward and closes the drainage canal. In either case, the progression of the disease can be stopped with medical intervention but part of vision already lost cannot be restored even with surgery. This is why it is important to detect the signs of glaucoma early through regular eye examination particularly, eye pressure measurement. Although, the procedure carried out during a typical eye examination for glaucoma is fairly simple and non-invasive, the cost of the examination might constitute a barrier to regular and unlimited access to examinations most especially in the developing world. An important fact to note about glaucoma is that it can be present in patients with normal eye pressure and in like manner cause serious sight loss [24]. This fact proves that full understanding of this disease is still lacking. Therefore, many of its salient features, signs and manifestations are still left undiscovered. Many attempts have been made towards understanding the pathophysiology of glaucoma through mathematical modelling where more often than not, mechanical response of the optic nerve head is characterized as certain physiological parameters are varied. Prominent amongst these parameters are the intraocular pressure, cerebrospinal fluid pressure, scleral tension etc. However, little insights have only been drawn from the modelling exercise as glaucoma is dubbed multi-factorial disease thus developing a single model that will capture the essence of several factors such as ethnicity, diabetic status, gender, obesity, age etc. that contribute to glaucoma [25] has proved very difficult. The lack of good understanding of glaucoma might explain the reason why eyes with glaucomatous field loss have high false negative responses during medical examinations [32]. Although we cannot rule out the influence of human error as fatigue, state of mind and carelessness may also contribute significantly to high false negative responses.

The associated issues with early detection and correct diagnosis of glaucoma raised in the foregoing can largely be tackled by considering automation of the glaucoma diagnosis process. Here, an intelligent system is envisaged. The system takes as input the retinal fundus image of a patient, performs a predefined mathematical operation on the image and produces an output encoding the health status of the eye. The predefined operation which is usually derived from techniques in biomedical image processing and machine learning determines the overall performance of the system. Several approaches focus on image segmentation methods including sparse dissimilarity constraint coding [10], super pixel classification [6, 9], and adaptive thresholding [18, 28] to delineate the optic disc and optic cup in retinal fundus images towards the estimation of features such as cup to disk ratio CDR [3],

vertical cup and disc diameters (VCD and VDD) [23] and neuro-retinal rim (NRR) [15]. These features have been suggested [23] to have certain correlation to the presence of glaucoma. Another commonly explored approach is to process the entire retinal fundus image using time [21], frequency [22] or joint time-frequency analysis [20] in order to obtain new representation of the image data that are helpful for discriminatory or classification tasks. The use of Fourier analysis [13], Gabor transform [1] and wavelet transform [19] have been reported for glaucoma detection. Also, many classification algorithms that learn discriminatory features and classify data into different classes have been deployed [4, 34]. Specifically, in recent years, Convolutional Neural Networks (CNNs) [29] which are a class of Deep Neural Networks (DNNs) have gained tremendous traction in glaucoma detection research. This is largely due to the fact that they don't require manual feature extraction as they are designed in a way that allows them to automatically obtain good data representation for classification task, learn the useful discriminatory features in the representation and finally classify the data into different classes [36]. Another reason is that CNNs have consistently displayed superior performance in several other time series [39] and image data [16] classification tasks. Although deep learning algorithms have been successfully used for glaucoma detection task [29], the computational resources required for their implementation is usually very huge. This stems from the fact that they usually have lots of tunable hyperparameters that can only be tuned properly with large amount of training data which is not always readily available in biomedical data analysis problems [12]. Secondly, their optimal network architecture and hyperparameters configuration are not well understood [27]. Therefore, a simpler alternative method that will produce the same or even better result than CNNs will be desirable.

Wavelet Image Scattering Network (WISN) is a representation or feature learning scheme described as very computationally cheap, well understood and very efficient in learning desirable data representation or automatic feature extraction (i.e. within class low-variance and between class high-variance features) [7]. The learned features can then be used downstream to build very simple and computationally cheap classification algorithms. Convolution, nonlinearity and pooling are the major operations carried out in the upstream section of all CNN architectures, these same operations are carried out efficiently in WISN by the successive application of wavelet filters, modulus operator and scaling filters respectively to the data. An important distinction between WISN and CNNs is that filters weights are learned in CNNs while they are fixed in WISN. As we write, we are not aware of any research

work documenting a detailed performance of wavelet scattering features from retinal fundus image data on glaucoma detection. We note that wavelet scattering features were used in [17] however, the study is conspicuously silent on the effect of scattering transform parameters. Therefore, as far as we know, this paper presents the first research work studying the influence of wavelet image scattering network parameters on the scattering features obtained from retinal fundus image data for glaucoma detection. The aim is to study the potentiality of WISN as a reliable feature learning scheme for an envisioned fully automated glaucoma detection/diagnosis system. The rest of the paper is arranged as follows. Section [Wavelet Scattering](#) gives a concise theoretical background to wavelet scattering. Materials and methods are described in section [Materials and Methods](#) while the results and discussion are contained in sections [Results](#) and [Discussion](#) respectively. Lastly, we give our conclusion in section [Conclusion](#).

Wavelet Scattering

In this section, we provide a brief description of the theory behind 2-D wavelet scattering transform or wavelet image scattering. The advantage of using wavelets for biomedical signals analysis is discussed in section [Biomedical Signal Analysis](#) followed by the definition of the 2-D wavelet transform in section [2-D Wavelet Transform](#). Section [Wavelet Scattering Network](#) explains how wavelet scattering transform is derived from wavelet transform. A more detailed treatment of wavelet scattering transform can be found in [7].

Biomedical Signal Analysis

Biomedical data or signals frequently exhibit slowly changing trends or oscillations punctuated with transients [37]. In particular, biomedical images usually consist of smooth regions interrupted by edges or abrupt changes in contrast. Generally, these abrupt changes are the most interesting part of the data both perceptually and in terms of the information they provide. The canonical Fourier Transform is a powerful signal analysis tool however, it does not represent abrupt changes in signals efficiently [35] in that it represents signals as a sum of sine waves which are not localized in time and space. In contrast, wavelet transform represents signals as sum of wavelets which are well localized in time and space [35]. This makes wavelet transform suitable for the analysis of most real world signals. The 2-D wavelet transform supplies the basic theory for wavelet scattering transform network to learn discriminatory features from image data.

2-D Wavelet Transform

A 2-D wavelet transform of an image signal $f(\vec{x})$, $f \in \ell^2(\mathbb{R}^2)$ (i.e. finite energy signal) is given as

$$c(\vec{b}, a, \theta) = \int \psi(a^{-1}, r_{-\theta}(\vec{x} - \vec{b})) f(\vec{x}) d\vec{x} \tag{1}$$

Equation (1) defines a convolution operation $f \otimes \psi$ where ψ , the analysing wavelet or convolution kernel is dilated by $a > 0$, translated by $\vec{b} \in \mathbb{R}^2$ and rotated by angle θ ($r_{-\theta}$ denotes the rotation operator). The analysing wavelet ψ satisfies the admissibility condition which in most applications may be interpreted as ψ having a zero mean (i.e. $\int \psi(\vec{x}) d\vec{x} = 0$). When the admissibility condition is coupled with the localization capability of ψ and its Fourier transform (i.e. band pass characteristic in the frequency domain) it becomes obvious that wavelet transform implements a local filtering in space (\vec{x}) and scale (a). This local filtering operates in constant relative bandwidth, $\Delta\omega/\omega$. Consequently, wavelet transform is more efficient at small scales or high frequencies particularly, in scanning singularities (transients or high frequency components in signals). The low-frequency components of the signal not captured are then decomposed through a separate function known as the scaling function (ϕ) whose partial derivatives result in the analysing mother wavelet functions [33]. Even though wavelet transform is good at scanning localized features in signals, it suffers from the fact that it is translation covariant [14]. This means similar signals at dissimilar locations get mapped into separate signal classes, and thus making learning of discriminatory signal features by classification algorithms more complicated. Direct application of engineered features obtained directly from 2-D wavelet transform of retinal fundus images for glaucoma detection can be found in [26, 31]

Wavelet Scattering Network

Fortunately, It can be shown [7] that the modulus of a wavelet transform, $|f \otimes \psi|$ makes it translation invariant. By taking this idea further, [7] created the scattering propagator (V_p), a path-ranked iteration operator on the modulus and applied the scaling function (ϕ) to the propagator in order to capture the slowly varying features in the signal. Thus, introducing what they call the Wavelet scattering transform ($S_p f$)

$$V_p f = || |f \otimes \psi_1| \otimes \psi_2 | \cdots \otimes \psi_m | \tag{2}$$

$$S_p f = || |f \otimes \psi_1| \otimes \psi_2 | \cdots \otimes \psi_m | \otimes \phi \tag{3}$$

where $1, 2, \dots, m$ denote the scattering stages. A wavelet image scattering network is produced by continued stage-wise evaluation of the convolution operation in Eq. (3). There are many mother wavelet functions in literature however, the choice of the analysing wavelet for a particular problem depends on the nature of the problem. Since we aim to detect relevant information such as segments and edges in biomedical images, anisotropic or directional wavelets are preferred. These wavelets are sensitive to rotations or directions and as such they can track oriented features such as segments and edges in images. Specifically, we chose the Morlet wavelet as the analysing wavelet in this work because it is the most widely used anisotropic wavelet in literature. The Morlet wavelet is obtained by tapering a sine wave by a Gaussian as shown in Eq. (4).

$$\psi_m = \exp(i2\pi ft) \exp(-t^2/(2\sigma^2)) \tag{4}$$

Materials and Methods

This section presents detailed explanation on the methods and materials used in this work. The workflow is depicted in Fig. 1.

Data Acquisition

The RIM-ONE DL image dataset [5] was downloaded and used for the study. Three hundred and thirteen and

one hundred and seventy-two 3-D retinographies from healthy and glaucoma patients respectively are contained in the dataset. The images were captured in PNG format from subjects in three different Spanish hospitals. Additionally, the dataset is partitioned into training and test sets in two different ways: random partitioning and partitioning by hospital. In the former, the training and test sets are assembled randomly from all images in the dataset. However, in the latter, the partitioning was carried out with respect to hospital. That is, images taken in one hospital are used for the training set while those taken in the other two hospitals are used for the test set. The number of retinal fundus images in the training/test set of the randomly partitioned dataset is 339/146 out of which 219/94 images belong to the healthy or normal class while 120/52 belong to the glaucoma class. Similarly, the number of retinal fundus images in the training/test set of the hospital partitioned dataset is 311/174 out of which 195/118 images belong to the normal class and 116/56 images belong to the glaucoma class. The dataset is available for free download at <https://bit.ly/rim-one-dl-images>. Figure 2 shows retinal fundus images randomly selected from healthy and glaucoma classes in the dataset.

Data Preprocessing

To be amenable to the 2-D wavelet transform, the 3-D images needed to be translated into their 2-D versions.

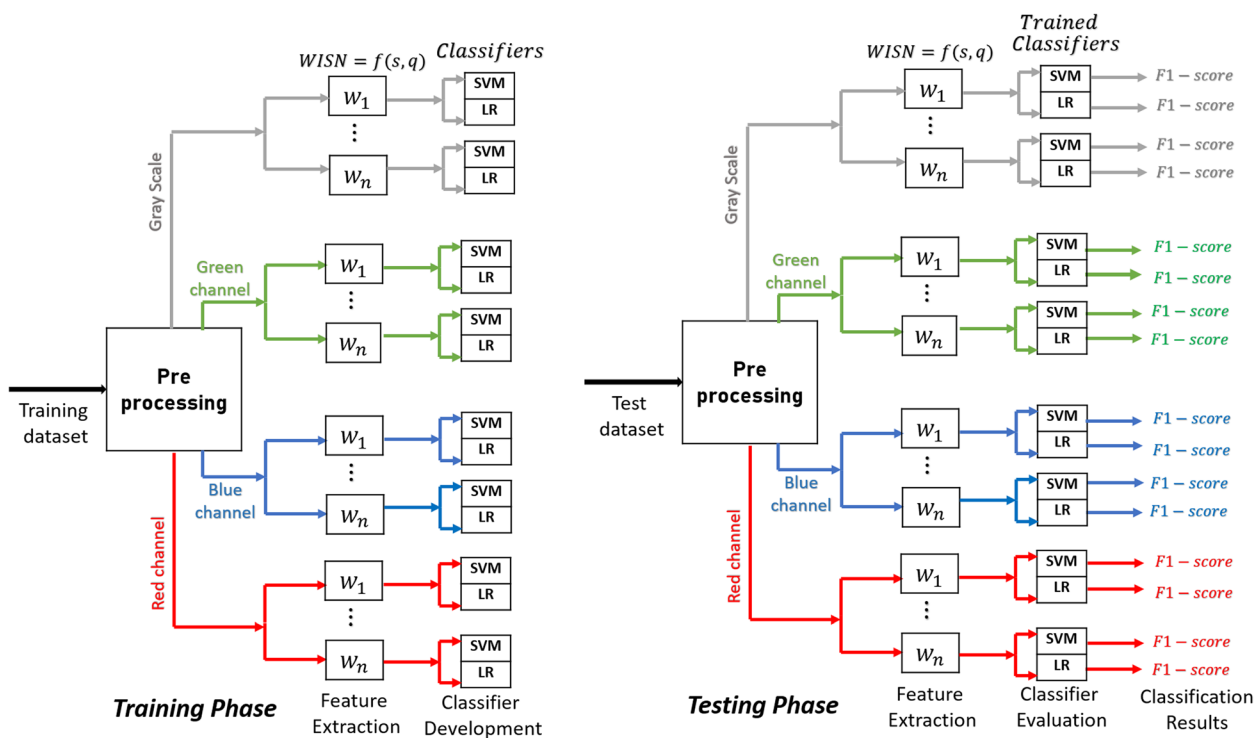


Fig. 1 Training and testing workflows for the proposed method

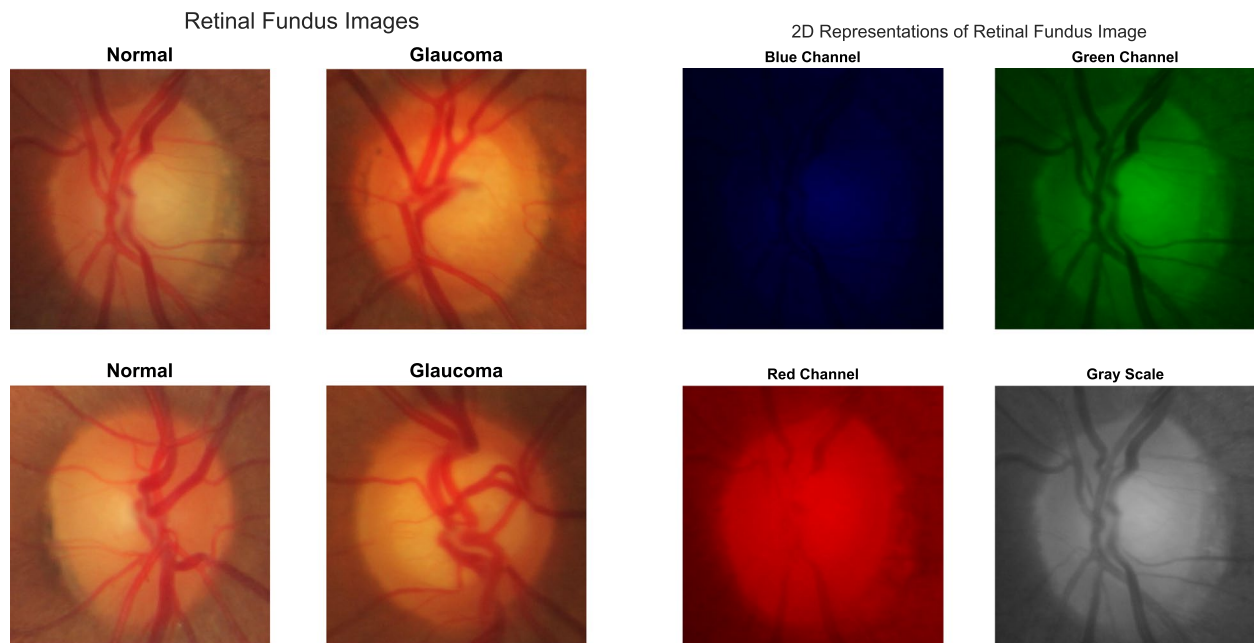


Fig. 2 Retinal fundus images (left panel) and 2-D channel representations of a retinal fundus image (right panel)

Consequently, individual 2-D channels that is red, green and blue channels were extracted from all images. In addition, all images were converted to the 2-D gray scale format. Therefore, each image in the dataset has four different 2-D representations. These are blue channel (BC), green channel (GC), red channel (RC), and gray scale (GS) representations. Lastly, the 2-D images were resized to 300-by-300 in order to have a uniform image size as sizes of the original images vary significantly. All preprocessing steps were carried out in the Matlab environment using the Image Processing Toolbox. Representations for randomly selected images from normal and glaucoma image classes are displayed in Fig. 2.

Development of our Wavelet Image Scattering Network

The wavelet scattering framework used for the wavelet image scattering decomposition was implemented in Matlab using the Image Processing and Wavelet toolboxes. The framework uses 2 complex-valued 2-D Morlet filter banks (i.e. 2 scattering stages). The scattering decomposition result depends on parameter setting in the framework. The parameters include Quality Factors (q), Invariance Scale (s) and Number of Rotations (r). In order to determine how sensitive the scattering features are to changes in individual parameter values, a t-test was conducted. The test gave information regarding the significance of difference in means of a particular feature as a given parameter is varied. For most of the features, there was a significant difference in the means when the scale invariance and quality factor parameters were

varied. However, none of the features have significance difference in means with changes in the number of rotation parameter. Therefore, only the scale invariance and quality factors parameters were varied in the scattering framework. Quality Factors control the number of wavelets per octave in each of the filter banks. Although the flexibility of allowing wavelets within each octave may be desirable as it supports fine scale analysis, it could also escalate the computational complexity of the framework if many wavelets are used. Therefore, a balance must be struck somewhere in between. A maximum of four and three wavelets per octave in the first and second filter banks respectively was experimented in this study. The Invariance Scale parameter determines the spatial support of the scaling and wavelet filters (i.e. the spatial support of the scaling and wavelet filters can only take on values not exceeding the one specified for the invariance scale parameter). The default Invariance Scale parameter value in the framework is one-half the lower of the number of rows and columns in the image rounded to the nearest whole number. Therefore, in our case the default value for the Invariance Scale parameter is 150. As this default value does not imply the optimal value for the parameter, we experimented Invariance Scale parameter values in the range [25 150] with an incremental step of 25 (i.e. 25, 50, 75, 100, 125 and 150). Finally, the Number of Rotations parameter sets the number of rotations of each wavelet in each filter bank in the scattering framework. As stated earlier, varying this parameter showed an insignificant effect on the scattering features. Therefore,

six clockwise rotations (with linearly spaced angles between 0 and π radians) per wavelet per filter bank which corresponds to the framework default parameter setting was used in the study.

Wavelet Scattering Learned Features

Regardless of the WISN parameter configuration used, the dimension of the resulting feature space for each retinal fundus image is $x \times y \times z$. That is there are x scattering paths and individual scattering path gives a scattering coefficient matrix of dimension $y \times z$. Subsequently, we obtained the mean along the 2nd and 3rd dimensions (i.e. y and z respectively) of the scattering coefficient matrix to arrive at x - element feature vector for individual image in the training and test datasets. This resulted in a significant data reduction from 90,000 (i.e. 300×300 image size) elements to x where x varies between 100 and 700 in this study.

Design of Classification Algorithms

The workflow for the classifier design include training and hyperparameter optimization. The classification algorithms explored are the binomial Logistic Regression (LR) and binary Support Vector Machine (SVM). To serve as a check, a simple Convolutional Neural Network (CNN) classifier was also considered. The CNN was constructed to have a convolution layer with 25 10-by-10 filters with 1-by-1 strides. This is followed by a RELU activation, max pooling layer and a fully connected layer. Furthermore, a softmax layer was deployed in order to normalize the output of the fully connected layer into probabilities. Lastly, a cross entropy loss was used as the loss function. Each classifier type was designed using individual 2-D image data representation and feature vectors from the training datasets were used exclusively for training and optimizing the classifiers. Optimized hyperparameters include regularization penalty and learning rate for LR and Kernel function, Kernel scale and box constraint for SVM. A 5 - fold cross validation scheme was used for training the LR/SVM classifiers. The best hyperparameter setting was the one that returned the minimum cross-validated classification loss for the LR/SVM model and was eventually used for classifying the test sets. The Bayesian optimization procedure was used and it was implemented in Matlab.

Performance Metric

Since there is an imbalance in the number of samples in the two data classes (see section [Data Acquisition](#)), the appropriate performance metrics for classifier evaluation is the F1-score. F1-score describes a classifier's performance on individual data class. To obtain

the F1-score, one needs to first calculate precision and recall for the classifier. Precision measures the proportion of correct positive predictions in all positive predictions while recall measures the proportion of correct positive predictions in all positive data samples. In other words, precision answers the question: out of all positive predictions made by the classifier how many are truly positive? However, recall answers the question: out of all positive data samples how many are classified as positive? The harmonic mean of precision and recall is the F1-score. That is the F1-score reports the classifier's performance in terms of both precision and recall. It is worthy of note that precision and recall are both defined in terms of relevance, meaning that they are defined relative to what the experimenter designates as the positive data class. Here, we refer to the glaucoma class is the positive class.

Results

We present the results of the retina fundus image classification experiments in the boxplots shown in Fig. 3. Classification results for the test set in the case of data partitioning by hospital are displayed in Fig. 3c and d while those for random partitioning are displayed in Fig. 3a and b. The boxplots summarizes the influence of WISN parameter setting, data representation scheme and classification algorithm type on the classification accuracy of the test sets. Results in Fig. 3a and c were obtained using support vector machine classifier while those in Fig. 3b and d were obtained using logistic regression classifier. Each boxplot within a Figure reveals the distribution of the classification accuracy results and hence the effect of WISN parameters setting on the classification accuracy results for each data representation scheme (i.e. BC, GC, RC, and GS). The influence of WISN parameter setting on F1-score is shown more explicitly in Fig. 4 where we have plotted F1-scores against individual parameter configuration for each of the 2-D channel representations and classification algorithm. There are thirty-nine (39) distinct parameter settings explored in this work (see section [Development of our Wavelet Image Scattering Network](#)). In Table 1 we have shown specifically, the best set of hyperparameter values for each of the classifiers with respect to individual 2-D channel representations and data partitioning scheme. The best set of hyperparameter values are those that produced the maximum F1-scores. Lastly, in Table 2 we have compared the classification results obtained from our proposed method (i.e. combination of WISN features and simple classifiers) and an ordinary convolutional neural network (CNN) classifier (section [Design of Classification Algorithms](#)) for both data partitioning schemes.

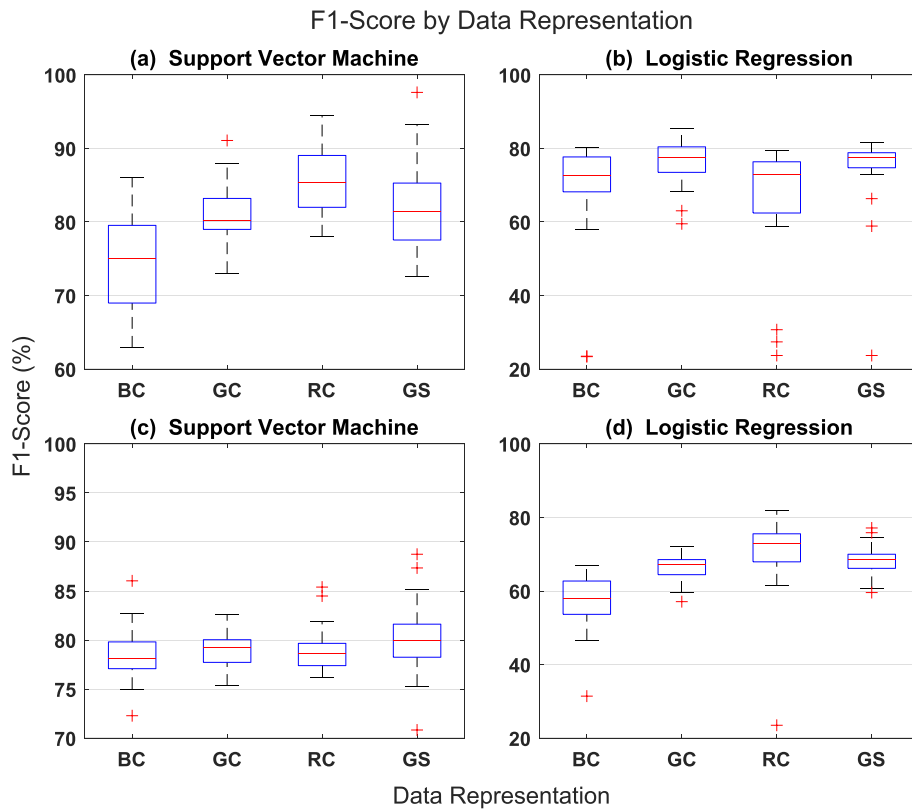


Fig. 3 Boxplots summarizing the influence of WISN parameter setting on classification result for data partitioning by hospital (lower panel) and random partitioning (upper panel) schemes



Fig. 4 Plots of F1-score against parameter setting for data partitioning by hospital (lower panel) and random partitioning (upper panel) schemes

Table 1 Best hyperparameter setting

Data	WISN			F1-score (%)
Partitioning	Classification	Parameter	Image	
Scheme	Algorithm	Setting	Representation	
Random	Support	$s = 125$	Gray Scale	98
	Vector	$q = [1, 1]$		
	Machine			
Hospital	Logistic	$s = 125$	Green Channel	85
	Regression	$q = [1, 1]$		
	Support	$s = 125$	Gray Scale	89
	Vector	$q = [1, 1]$		
	Machine			82
	Logistic	$s = 125$	Red Channel	
Regression	$q = [3, 2]$			

Discussion

Generally, it is observed from Fig. 3 that highest F1-scores were obtained by the support vector machine classifier in both data partitioning schemes. For the random partitioning, highest F1-score of 98%/85% was achieved by the support vector machine/logistic regression classifier on the gray scale/green channel data representation. For the partitioning by hospital however,

F1-score of 89%/83% was achieved by the support vector machine/logistic regression classifier on the gray scale/red channel data representation. It is quite reasonable to expect that the highest F1-score value would come from the random partitioning dataset. In the random partitioning scheme, there are chances that some subjects would have their data placed in both the training and test datasets. As patient-specific features have been found to be present in glaucoma [38], learning algorithms can pick up these features during training and this will definitely make them perform well during testing. Conversely, in the partitioning by hospital scheme, the training and test datasets contain data from different subjects.

Furthermore, F1-score is highly sensitive to WISN parameter setting for each of the 2-D channel representations and classification algorithms. This is clearly highlighted by Fig. 4. Figure 4 further reveals the superiority of support vector machine to logistic regression in correctly classifying the test datasets. It is observed that there are a number of missing F1-score values in the plots. This is more pronounced in the plots for logistic regression where all the data representations (gray level, red, blue and green channels) have missing F1-score values. Missing F1-score values occur as a result of NaNs. NaNs

Table 2 Classification accuracy results for SVM, LR and CNN

Data	WISN		Maximum F1-score (%)	
Partitioning	Classification	Parameter		
Scheme	Algorithm	Setting		
Random	Support	$s = 125, q = [1, 1]$ #8	98	
	Vector	$s = 100, q = [4, 1]$ #5	94	
	Machine	Green Channel	$s = 75, q = [1, 1]$ #33	91
		Blue Channel	$s = 125, q = [3, 1]$ #10	82
	Logistic	$s = 125, q = [1, 1]$ #8	85	
	Regression	Gray Scale	$s = 100, q = [3, 2]$ #4	82
		Blue Channel	$s = 125, q = [1, 1]$ #8	80
		Red Channel	$s = 125, q = [4, 1]$ #12	79
	Convolutional	NA	Red Channel	82
	Hospital	Support	$s = 125, q = [1, 1]$ #8	89
Vector		$s = 125, q = [4, 1]$ #37	86	
Machine		Red Channel	$s = 100, q = [4, 1]$ #5	85
		Green Channel	$s = 125, q = [2, 1]$ #9	83
Logistic		$s = 125, q = [3, 2]$ #11	82	
Regression		Gray Scale	$s = 125, q = [4, 3]$ #14	77
		Green Channel	$s = 150, q = [4, 2]$ #31	72
		Blue Channel	$s = 150, q = [3, 2]$ #29	67
Convolutional		NA	Blue Channel	83
Neural				
Network				

are in turn occasioned by the occurrence of zero positive predictions by the classification algorithm at a particular hyperparameters setting. This means that the classification algorithm failed to classify none of the retinal fundus images in the glaucoma class as positive when certain hyperparameters values are used. This set of hyperparameters value is the worst for the classification algorithm. Table 1 clearly shows that a scale invariance of 125 and quality factor [1,1] (i.e. one wavelet per octave in each of the filter banks) which corresponds to hyperparameter setting number 8 (see Fig. 4) appears to be the best hyperparameter setting out of the thirty-nine (39) explored in the classification problem. Furthermore, the gray scale representation proved to contain the most discriminatory features between healthy (normal) and glaucomatous retinal fundus images amongst the four 2-D channel representations. The gray scale representation gave the best F1-score result in the two data partitioning schemes.

As evident from Table 2, the new method proved to be a better alternative to CNN, at least for this problem and the particular CNN configuration used. For the random partitioning scheme, the CNN trailed both the support vector machine and logistic regression classifiers. However, the CNN trailed only the support vector machine classifier but led the logistic regression classifier in the data partitioning by hospital scheme. The best 2-D channel representation for the CNN in both cases is the red channel.

Regarding ease of algorithmic implementation, an examination of the inference time was conducted to access the ease of implementation of our method. Five batches of 25 retinal fundus images were drawn randomly from each class for the inference analysis. The inference time (in milliseconds) is the average of the time taken to classify all 25 images in each batch. We used a CPU device with 32 Gigabyte of memory and Intel Core i9 (8cores) execution unit. The inference time included the time it took to read the re-sized 2D channel images from file, obtain the wavelet scattering features and classify the images. The inference time for each scheme explored is shown in Table 3. The wavelet scattering network schemes have the least inference time and the scheme with the support vector machine is about 2 times faster than that with logistic regression.

Finally, we compare our results with those reported in literature. The RIM-ONE DL (RIM-ONE for Deep Learning) dataset was released in 2020 as a refined version of the three initially released RIM ONE datasets (RIM-ONE v1, v2, and v3). The dataset was specifically optimized for deep learning applications [5]. Different CNN architectures have been utilized for the classification of both hospital and random partitioned datasets. Table 4 compares the classification accuracy results from these CNN architectures and our method. To the best of our efforts,

Table 3 Inference Time (per image) comparison

Model	Time (ms)
WISN + SVM	894.55
WISN + LR	1562.06
CNN	1725.21

only one publication [17] was found in literature to have applied wavelet scattering network features for glaucoma detection and the dataset used in the article is the initial dataset (i.e. RIM-ONE v3). The results from the work is also included in Table 4. This indicates that our work is the first to apply wavelet scattering features for glaucoma detection using the RIM-ONE DL dataset. From the table, it is obvious that our method gave better accuracy values in both hospital and random partitioned datasets.

Conclusion

This work exploits wavelet image scattering to obtain within class low-variance representations from 2-D channel representations of retinal fundus images for glaucoma detection. Utilizing the 2-D scattering transform with fixed filter weights and simple classification algorithms, we were able to attain a maximum of 98% and 89% correct classification on a held-out test sets. Wavelet Image Scattering Network proved to be a robust and effective feature extractor for glaucoma detection requiring only a minimal set of user-specified parameter values. On the same problem but with a simple Convolutional Neural Network whose filters were learned, we achieved a maximum of 82% and 83% correct classification. It is important to reiterate that our work is not intended as an absolute comparison of WISN and CNNs but rather, to

Table 4 Wavelet scattering network versus CNN architectures on RIM-ONE DL dataset. Left: Randomly partitioned, Right: Partitioned by hospital. Accuracy values for the CNN Architectures were copied from <https://github.com/miag-ull/rim-one-dl>

Network	Accuracy(%)	Accuracy(%)
VGG19	93	85
VGG16	92	85
Xception	91	79
ResNet50	91	83
MobileNetV2	90	53
DenseNet	90	78
MobileNet	93	82
InceptionResNetV2	91	76
InceptionV3	89	80
NASNetMoile	75	79
WISN (RIM-ONE DL)our work	98	89
WISN (RIM-ONE v3)[18]	93	NA

demonstrate the potentiality of Wavelet Image Scattering for producing robust and efficient representations of retinal fundus image data for glaucoma detection learning task. WISN and CNNs have different hyperparameter and architectural changes that can significantly influence the classification results. For instance, highly optimized CNN architectures such as VGG19 [8], VGG16 [30], Xception [11], and ResNet50 [2] have been reported to have achieved 93%, 92%, 91% and 91% correct classification respectively on the random data partitioning test set and 85%, 85%, 84% and 79% respectively on the data partitioning by hospital test set. A future research task of our group is to apply our method to all publicly available glaucoma fundus datasets while also exploring different retina fundus image pre-processing modalities. The results from such task will further give insight into the versatility and reliability of our method.

Abbreviations

BC	Blue channel
CDR	Cup to disc ratio
CNN	Convolutional Neural Network
DNN	Deep Neural Network
GC	Green Channel
GS	Gray Scale
LR	Logistic Regression
NRR	Neuro-Retina Rim
RC	Red Channel
RELU	Rectified Linear Unit
SVM	Support Vector Machine
VCD	Vertical Cup Diameter
VDD	Vertical Disc Diameter
WISN	Wavelet Image Scattering Network

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42490-023-00067-5>.

Additional file 1. Glaucoma Random. This is an excel workbook containing classification results for the glaucoma class in the test set of the randomly partitioned dataset. The workbook contains three worksheets. Each worksheet contains results for individual classification algorithm (i.e. SVM, LR and CNN).

Additional file 2. Glaucoma Hospital. This is an excel workbook containing classification results for the glaucoma class in the test set of the dataset partitioned by hospital. The workbook contains three worksheets. Each worksheet contains results for individual classification algorithm (i.e. SVM, LR and CNN).

Additional file 3. Healthy Hospital. This is an excel workbook containing classification results for the healthy or normal class in the test set of the dataset partitioned by hospital. The workbook contains three worksheets. Each worksheet contains results for individual classification algorithm (i.e. SVM, LR and CNN).

Additional file 4. Healthy Random. This is an excel workbook containing classification results for the healthy or normal class in the test set of the randomly partitioned dataset. The workbook contains three worksheets. Each worksheet contains results for individual classification algorithm (i.e. SVM, LR and CNN).

Additional file 5. Wavelet Image Scattering Paper Code. This file contains the codes used for the project.

Acknowledgements

We thank Francisco Jose Fumero Batista, Tinguaro Diaz-Aleman, Jose Sigut, Silvia Alayon, Rafael Aray, and Denisse Angel-Pereira for making the dataset used for our research available to the public on <https://bit.ly/rim-one-dl-images>.

Authors' contributions

H.A: Conceptualization, Methodology, Software, and Writing (Original draft preparation, final review, and editing). J.E: Data Curation and Writing (final review and editing). All authors have read and approved the final manuscript.

Funding

The authors received no funding for the research.

Availability of data and materials

Data generated during the work is available in the related information files. Furthermore, the dataset used for the study is available for free download at <https://bit.ly/rim-one-dl-images>. We have also made codes used for the project available at https://drive.google.com/file/d/1KjxdRHzeK6BK2uqHtCDaXC3ub7AXG8bb/view?usp=share_link. Any additional information relating to the work can be sought from the authors through reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 21 July 2022 Accepted: 6 February 2023

Published online: 02 March 2023

References

1. Acharya UR, Ng EYK, Eugene LWJ, Noronha KP, Min LC, Nayak KP, Bhandary SV. Decision support system for the glaucoma using gabor transformation. *Biomed Signal Proc Control*. 2015;15:18–26.
2. Ali N, Quansah E, Köhler K, Meyer T, Schmitt M, Popp J, Niendorf A, Bocklitz T. Automatic label-free detection of breast cancer using nonlinear multimodal imaging and the convolutional neural network resnet50. *Transl Biophoton*. 2019;1(1–2):e201900003.
3. Ali R, Sheng B, Li P, Chen Y, Li H, Yang P, Jung Y, Kim J, Chen CLP. Optic disk and cup segmentation through fuzzy broad learning system for glaucoma screening. *IEEE Trans Ind Informat*. 2020;17(4):2476–87.
4. Barros D, Moura JCC, Freire CR, Taleb AC, Valentim RAM, Morais PSG. Machine learning applied to retinal image processing for glaucoma detection: review and perspective. *Biomed Eng Online*. 2020;19(1):1–21.
5. Batista JF, Diaz-Aleman T, Sigut J, Alayon S, Aray R, Angel-Pereira D. Rim-one dl: A unified retinal image database for assessing glaucoma using deep learning. *Image Anal Stereology*. 2020;39(3):161–7.
6. Bechar MA, Settouti N, Barra V, Chikh MA. Semi-supervised superpixel classification for medical images segmentation: application to detection of glaucoma disease. *Multidim Syst Sign Process*. 2018;29(3):979–98.
7. Bruna J, Mallat S. Invariant scattering convolution networks. *IEEE Trans Pattern Anal Mach Intell*. 2013;35(8):1872–86.
8. Carvalho T, De Rezende ERS, Alves MTP, Balieiro FKC, Sovat RB. Exposing computer generated images by eye's region classification via transfer learning of vgg19 cnn. In: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE; 2017. p. 866–70.
9. Cheng J, Liu J, Xu Y, Yin F, Wong DWK, Tan N-M, Tao D, Cheng C-Y, Aung T, Wong TY. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE Trans Med Imaging*. 2013;32(6):1019–32.

10. Cheng J, Yin F, Wong DWK, Tao D, Liu J. Sparse dissimilarity-constrained coding for glaucoma screening. *IEEE Trans Biomed Eng.* 2015;62(5):1395–403.
11. Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. p. 1251–8
12. Deperlioglu O, Kose U, Gupta D, Khanna A, Giampaolo F, Fortino G. Explainable framework for glaucoma diagnosis by image processing and convolutional neural network synergy: Analysis with doctor evaluation. *Futur Gener Comput Syst.* 2022;129:152–69.
13. Essock EA, Sinai MJ, Fechtner RD, Srinivasan N, Bryant FD. Fourier analysis of nerve fiber layer measurements from scanning laser polarimetry in glaucoma: emphasizing shape characteristics of the double-hump pattern. *J Glaucoma.* 2000;9(6):444–52.
14. Farge M, et al. Wavelet transforms and their applications to turbulence. *Annu Rev Fluid Mech.* 1992;24(1):395–458.
15. Fortune B. Optical coherence tomography evaluation of the optic nerve head neuro-retinal rim in glaucoma. *Clin Exp Optom.* 2019;102(3):286–90.
16. Hussain M, Bird JJ, Faria DR. A study on cnn transfer learning for image classification. In: UK Workshop on computational Intelligence. Springer; 2018. p. 191–202
17. Hussain S, Guo F, Shi X, Li W, Shen Z. Compact scattering features for glaucoma detection, vol 2253. IOP Publishing; 2022. p. 012031.
18. Issac A, Sarathi MP, Dutta MK. An adaptive threshold based image processing technique for improved glaucoma detection and classification. *Comput Methods Prog Biomed.* 2015;122(2):229–44.
19. Kirar BS, Agrawal DK, Kirar S. Glaucoma detection using image channels and discrete wavelet transform. *IETE J Res.* 2022;68(6):4421–8.
20. Kundra H, Park JC, McAnany JJ. Comparison of photopic negative response measurements in the time and time-frequency domains. *Documenta Ophthalmologica.* 2016;133(2):91–8.
21. Leung CK-S, Chiu V, Weinreb RN, Liu S, Ye C, Yu M, Cheung CY-I, Lai G, Lam DS-C. Evaluation of retinal nerve fiber layer progression in glaucoma: a comparison between spectral-domain and time-domain optical coherence tomography. *Ophthalmology.* 2011;118(8):1558–62.
22. Li S, Wang X, Li S, Wu G, Wang N. Evaluation of optic nerve head and retinal nerve fiber layer in early and advance glaucoma using frequency-domain optical coherence tomography. *Graefes Arch Clin Exp Ophthalmol.* 2010;248(3):429–34.
23. Manjunath TC, Lamani D, Chandrappa S, Ranjan Kumar HS. Different clinical parameters to diagnose glaucoma disease: a review. *Int J Comput Appl.* 2015;115(23):42–6.
24. Mansouri K, Weinreb RN. Ambulatory 24-h intraocular pressure monitoring in the management of glaucoma. *Curr Opin Ophthalmol.* 2015;26(3):214–20.
25. McMonnies CW. Glaucoma history and risk factors. *J Optom.* 2017;10(2):71–8.
26. Mookiah MRK, Acharya UR, Lim CM, Petznick A, Suri JS. Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features. *Knowl-Based Syst.* 2012;33:73–82.
27. Mvoulana A, Kachouri R, Akil M. Fine-tuning convolutional neural networks: a comprehensive guide and benchmark analysis for glaucoma screening. In: 2020 25th International Conference on Pattern Recognition (ICPR). IEEE; 2021. p. 6120–7
28. Nawaldgi S, Lalitha YS, Reddy M. A novel adaptive threshold and isnt rule based automatic glaucoma detection from color fundus images. In: Data Engineering and Intelligent Computing. Springer; 2018. p. 139–47.
29. Phan S, Satoh S, Yoda Y, Kashiwagi K, Oshika T. Evaluation of deep convolutional neural networks for glaucoma detection. *Jpn J Ophthalmol.* 2019;63(3):276–83.
30. Qassim H, Verma A, Feinzimer D. Compressed residual-vgg16 cnn model for big data places image recognition. In: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC). IEEE; 2018. p. 169–75.
31. Raja C, Gangatharan N. Glaucoma detection in fundal retinal images using trispectrum and complex wavelet-based features. *Eur J Sci Res.* 2013;97(1):159–71.
32. Rao HL, Yadav RK, Begum VU, Addepalli UK, Choudhari NS, Senthil S, Garudadri CS. Role of visual field reliability indices in ruling out glaucoma. *JAMA Ophthalmol.* 2015;133(1):40–4.
33. Schmeelk J. Wavelet transforms and edge detectors on digital images. *Math Comput Model.* 2005;41(13):1469–78.
34. Serener A, Serte S. Transfer learning for early and advanced glaucoma detection with convolutional neural networks. In: 2019 Medical technologies congress (TIPTEKNO). IEEE; 2019. p. 1–4.
35. Sifuzzaman M, Islam MR, Ali MZ. Application of wavelet transform and its advantages compared to fourier transform. Midnapore: Vidyasagar University; 2009.
36. Wang ZJ, Turko R, Shaikh O, Park H, Das N, Hohman F, Minsuk Kahng, and Duen Horng Chau. Cnn 101: Interactive visual learning for convolutional neural networks. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020. p. 1–7
37. Xie H-B, Chen W-T, He W-X, Liu H. Complexity analysis of the biomedical signal using fuzzy entropy measurement. *Appl Soft Comput.* 2011;11(2):2871–9.
38. Yokoyama Y, Maruyama K, Konno H, Hashimoto S, Takahashi M, Kayaba H, Kokubun T, Nakazawa T. Characteristics of patients with primary open angle glaucoma and normal tension glaucoma at a university hospital: a cross-sectional retrospective study. *BMC Res Notes.* 2015;8(1):1–8.
39. Zhao B, Lu H, Chen S, Liu J, Wu D. Convolutional neural networks for time series classification. *J Syst Eng Electron.* 2017;28(1):162–9.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

