Two Stages Approach for Detection and Classification of Retinal Images Using Deep Neural Network and Deep-wavelet

Authors Names

Huda Dhari[a]

Nidhal K. El Abbadi [b]

Publication date: 14 /12 /2024

Keywords: *Retina image, Glaucoma, deep wavelet network, deep learning, KNN, SVM, classifiers.*

ABSTRACT

Retina is the deepest layer in the inner surface regarding the eyeball and the sensory layer of the eye. This sensitive part of the eye is susceptible to several diseases. Realizing that the state regarding the retina represents one of the primary causes of severe vision loss and blindness globally, retinal disease is receiving major attention. Early diagnosis related to retinal pathology is now based on an analysis of the geometric features of retinal blood vessels, including branch lengths, widths, angles, tortuosity, branching patterns, and vessel diameters. Various diseases can be diagnosed by specialists with using such features. Since the year 1982, computer science has demonstrated its ability to effectively contribute to the diagnosis and detection regarding disease for biomedical sciences. The two stages of the method we present in this research are for the classification and detection of diseases connected to the eyes. Level one uses a parallel convolution neural network (CNN) for detecting the presence of any disease. The Ocular Disease Intelligent Recognition Dataset and the Ocular Disease Intelligent Recognition Dataset are two datasets used in level two for classifying four eye-related diseases with the use of Deep Wavelet as a technique for feature extraction (FE). We have decided that our approach performs well and produces good results based on the comparative estimate.

1. Introduction

Here "The eye, which has quite often been called the window of the soul, is the chief organ whereby the senso comune could have the most complete and magnificent view of infinite works of nature," wrote Leonardo da Vinci, demonstrating his keen awareness of the eye's primary role. Vision is a complex process that necessitates the cooperation of multiple components of the human brain and eye [1]. One of the primary sense organs intended for vision is the eye. It is a special organ that allows us to sense danger, stay aware of our surroundings, and keep our thoughts sharp [2]. The deepest layer of the eye, the sensory membrane lining the inner surface of the eyeball, is the retina [3]. The OD, macula, and blood vessels are the three primary anatomical components regarding the retina, which is a light-sensitive tissue made up of photosensitive cells (photoreceptors) which convert light to nerve signals that pass down optic nerves to the brain [4, 5]. Retina is a delicate part of the eye that could be affected by many different diseases [6]. The state of retina is a primary cause of severe vision loss and blindness worldwide, thus real eye disease is given great consideration [7]. Both internationally and regionally, the prevalence regarding eye diseases has increased recently [8]. Diagnosis is just as crucial in medicine as treatment [4]. Early detection could aid in the development of prevention measures, which in turn might result in impact reduction through appropriate treatment [9]. Since there might not be any symptoms at first, a patient could not be aware of its existence, yet as it worsens over time, the patient's condition would deteriorate [10]. Approximately 95% of blindness could be avoided with early detection, prompt treatment, and suitable follow-up measures [11]. Fundus photography is a non-invasive method of examining and interpreting retinal images and blood vessels. It helps ophthalmologists assess common eye diseases like diabetic retinopathy, glaucoma, hypertension, and

[a]Informatics Institute for Postgraduate Studies, Iraqi Commission for Computers & informatics Baghdad, Iraq
E-Mail: phd202110681@iips.edu.iq
[b]Al Mustaqbal Center for AI Applications, Al Mustqbal University, Babylon, Iraq, E-Mail: nidhal.elabbadi@uokufa.edu.iq

macular degeneration, along with various other diseases like ophthalmopathy and cardiovascular disease. It also provides easily accessible information about the vascular health of the body, brain, and eye [12], [13], and [14]. Early diagnosis of retinal pathology is now based on an analysis of the geometric features of the retinal blood vessels, including branch lengths, angles, widths, tortuosities, branching patterns, angles, and vessel diameters [15]. Every person has a different pattern of fundus blood vessels, which makes them suited for biometric identification [16]. Various diseases can be diagnosed by specialists with the help of these features [17]. For instance, diabetic retinopathy could lead to neovascularization, or creation of new blood vessels, whereas hypertension might change the angle of branching or vessel curvature [4]. Telemedicine has pioneered the automatic diagnosis and classification regarding cardiovascular as well as ophthalmologic diseases through the examination of retinal images [18]. It could identify a huge number of fundus images faster and more precisely than a medical practitioner. While image classification aims to classify all image pixels into one of predetermined classes, the artificial selection of features has severely limited the classification of performance [19] with the rising number of samples and parameters. While classification comes naturally to humans, an automated system finds it far more difficult [20]. Age-Related Macular Degeneration (AMD), Retinoblastoma, Diabetic Retinopathy (DR), Retinitis Pigmentosa, Macular Bunker, and Retinal Detachment are only a few of the retinal diseases for which computer-based systems are useful for early diagnosis and real-time classification.

Since the year 1982, computer science has demonstrated its ability to effectively contribute to diseases diagnosis and detection for biological sciences seeking to improve detection accuracy and efficiency [21]. Additionally, a number of NN-based techniques, both unsupervised and supervised, were used to analyze retinal images. MLP, ANN, SVM, and decision tree (DT) classifiers were incorporated into numerous supervised approaches. With the use of a combination of two methods—CNN and deep-wavelet—this study aims to develop an automated diagnosis method which will aid in the detection as well as classification regarding retinal diseases depending on two levels. Features of typical retina images are extracted, and multiple image classifiers utilize such features to determine whether or not there is an indication of a disease in the subsequent step. The suggested approach has no trouble learning to detect healthy eye image and is capable of concurrently detecting early and illness indicators. This study's remaining sections are arranged as follows: relevant work in the second section. The theoretical foundation for the techniques we employed in the detection as well as classification approach is provided in section three. In Section four, the study's approach is explained in full. Section five after that shows the dataset that was utilized and the operation that was carried out on them as well as, the acquired outcomes were presented and deliberated. Finally, section six provides a summary and wraps up the work.

2. Related work

Wavelet Scattering Network implementation for Retina disease diagnosis is lacking in the many studies conducted in Retina image classification based upon deep learning (DL) and machine learning (ML). For classifying X-ray images into four different classes—COVID-19, opacity, normal, and pneumonia cases. Mostapha Al Saidi et al. (2022) suggested a DL-based method which involves fine-tuning pre-trained CNNs (VGG19 and VGG16) and end-to-end training regarding a developed CNN model. For the experiment, a dataset including over 20,000 X-ray images was obtained via Kaggle. For comparing with the one-shot classification approach, a two-stage method was put into practice. Contrary to expectations, the VGG16 achieved 95% accuracy across five times the training set with the use of a one-shot technique [22]. Maira Araujo de Santana and Wellington Pinheiro dos Santos (2022), suggested a deep hybrid architecture depending on six-layer deep-wavelet NNs for supporting digital mammography region-of-interest imaging diagnosis. The architecture extracts attributes regarding the regions of interest from the mammograms and uses a support vector machine (SVM) with a kernel 2nd-degree polynomial for final classification. Tests were conducted on traditional classifiers, which include SVMs, single hidden layer multi-layer perceptrons, decision trees (DTs), random forests (RFs), and Bayesian classifiers. Based on the findings, injuries may be identified and classified with an average kappa of 0.91 and 94% accuracy when a 6-layered deep-wavelet network and a 2-degree polynomial kernel SVM are used as final classifier [23]. Deep-Wavelet Neural Networks (DWNN) are convolutional architectures that are based upon general theory of wavelets that are suggested by V. A. de Freitas Barbosa etal. (2023) for extracting features from images for the thermographic image classification. A hybrid architecture depending on deep networks for FE, RFs for selecting the most statistically relevant features, and linear

kernel SVMs for final layer classification have been suggested. A total of 336 thermographic images that have been classified as malignant, benign, cystic, and healthy (no lesion) make up the dataset that was employed. The 6-layered DWNN attained sensitivity, accuracy, kappa, specificity, and precision above 98%, according to experimental results. These findings demonstrate that competitive deep architectures, such as DWNN, are useful for improving clinical adoption and thermographic image analysis [24].

A unique date fruit type classification and grading system is presented by Arivazhagan Selvaraj and Newlin Shebiah Russel (2024). It is achieved through feature-level fusion of wavelet scattering characteristics and DL features. Reliable information extraction from a variety of channels is made possible by the extraction of wavelet scattering features at different stages of decomposition. This work uses pre-trained architectures such as Googlenet, Alexnet, Resnet, and MobileNet-V2 for extracting deep features. The nine-class Date Fruit in Controlled Environment data-set was used to empirically evaluate the suggested method, and the results showed a 95.90% accuracy for date species classification. TU-DG data-set included a variety of date fruit species that were evaluated; the accuracy for Ajwa species was 97.8%, the Sukkary species was 99.5%, and the Mabroom species was 92.6% [25].

3. Theoretical Background

3.1. Convolution neural network

CNNs are mostly used for tasks involving speech recognition and computer vision. They could be used for tasks involving spatially related datasets (image data, for example), in which the rows and columns are not interchangeable. Depending on the specific modeling task, a sequence of steps in their network design enable hierarchical feature learning. For instance, the initial network layers are in charge of the extraction of fundamental properties like corners and edges when it comes to object recognition in images. In the final few layers, these are after that gradually combined to create more intricate features that resemble the real items of interest, such houses, vehicles, or animals. The auto-generated features are then utilized to new images in order to identify objects of interest through prediction [26]. The CNN is employed at several layers:

• Input Layer: The "input layer," which is the initial layer in every CNN utilized, resizes images before sending them to further FE layers.

• Convolution Layer: The following layers are called "convolution layers," which function as filters for images to extract features. They are utilized to determine the points at which an image matches a feature throughout the testing.

• Pooling Layer: The "pooling layer" receives extracted feature sets after that. This layer reduces the size of large images without sacrificing the most crucial details. It maintains the highest value from every window and the optimal fits for every feature inside the window.

• Rectified Linear Unit Layer (ReLU): Every negative integer in pooling layer is substituted by 0 in the following ReLU layer. This prevents learnt values from becoming stuck close to 0 or from exploding up toward infinity, which aids in CNN's mathematical stability.

• Fully Connected Layer: The last layer consists of completely connected layers that convert high-level filtered images to labeled classes [27].

3.2. Deep-wavelet (DW)

Mallat introduced this algorithm [24]. From the time domain to the feature domain, there is a transformation [28]. It is produced through applying wavelet decomposition repeatedly at different levels in an iterative manner [29]. It is a DL approach for FE whose architecture as well as filters are predefined wavelets rather than learnt; it is a model for the automatic extraction of input signal features, involving convolutional, non-linear, and average cascade operations [30]. With this method, more consistent features are extracted from images. It is a useful tool that works well with the majority of classification algorithms on the market for FE and accurate data representation. Through using the images as inputs

and encoding them as a collection of coefficients which are locally stable to small deformations, which we will use to classify the signals, wavelet scattering transform enables to construct dependable features [31]. Low-pass and high-pass filters are applied to an image during wavelet decomposition, producing a set of different images. Approximations and details, respectively, are the terms used to describe the images produced by low-pass and high-pass filters. The original image's softness is emphasized in the approximations, whilst edges (or discontinuity regions) are underlined in the details. This technique, which permits picture analysis in the spatial as well as the frequency domains, is applied in pattern recognition [24]. Convolution, non-linearity transform, and convolution averaging employing complex wavelets, modulus operation, and scale function, in that order, are the three processes that make up the deep wavelet function [32]. Stated differently, the wavelet handles the convolution task, the modular operator handles the non-linearization, and the wavelet low-pass filter's filtering function finishes the pooling operation [33], (**see Fig.1.)**. As a result, a neuron is produced by fusing a certain filter with a technique of image size reduction. The typical low pass filter—which is widely employed in wavelet transformations in multi-resolution analyses to decay the image into several degrees of details—is not the same as the low pass scaling filter [24].

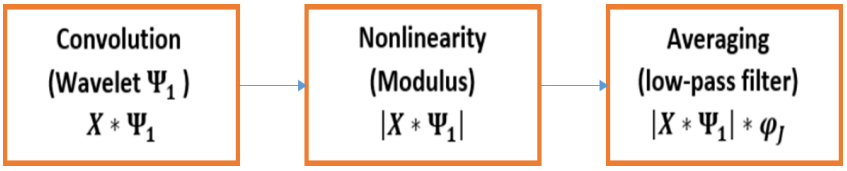

Fig. 1 - wavelet scattering transform operations

To obtain the zeroth-order scattering coefficients, as shown in Equation 1, the data have been first convolved with scaling function.

$$S[0] = i * \varphi \quad (1)$$

In which i denotes input data; φ, scaling function, and S [0], zeroth-order scattering coefficients. This removes the small details in the image and creates a basic feature [32].

As wavelet functions, or bank filters, one could utilize dilated mother wavelets and Morlet mother wavelets. The wavelet filters covered a variety of image sizes and rotations, allowing for a thorough examination of the input data. Details in a certain orientation will be highlighted by each filter. The filter bank (G) can be expressed mathematically as follows [24]:

$$G = \{g_1; g_2; g_3; \ldots; g_n\} \quad (2)$$

Whereas g2 is a horizontal high-frequency filter that highlights vertical edges, g1 is a high-frequency vertical filter that highlights horizontal edges, and g3 and g4 represent the diagonal filters that highlight the image. A framework featured two filter banks, indicating a three-depth structure (**see Fig. 2**).

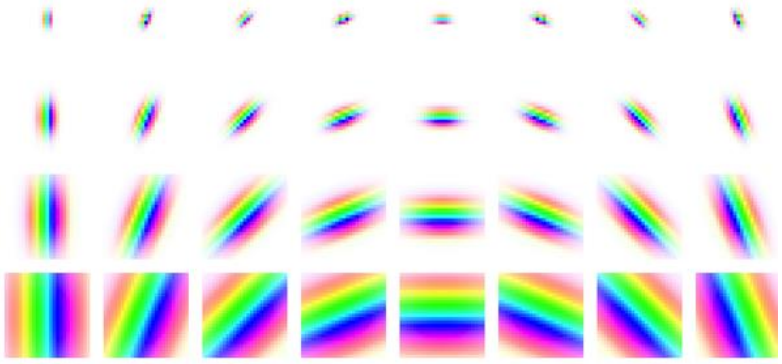

Fig. 2. Wavelets at different scales 8 orientations [30].

The modulus of every one of the filtered outputs was after that computed after applying the wavelet transform to input data with the use of every wavelet filter in the first filter bank. The 1st-order scattering coefficients were obtained by the averaging of these moduli using the scaling filter [32].

$$\mathbf{S[1] = |i* \psi j1]* \varphi} \quad \mathbf{(3)}$$

In which ψj1 denotes the wavelet, and S [1], is the 1st-order scattering coefficient. Through iterating the above steps, the 2nd-order scattering coefficients were computed

$$\mathbf{S[2] = ||i* \psi j1]* \psi j2|* \varphi} \quad \mathbf{(4)}$$

Where S [2] represents 2nd-order scattering coefficients.

Deep-wavelet is a viable option for applications where labeled data is scarce because of its low computing cost, high classification rate, and requirement for fewer training instances than DL techniques. A substantially less quantity of training data is required [6]. Significant features can be recovered at various scales with the use of Deep Wavelt [29]. The outputs of the Deep Wavelet should be utilized as input data to train a classifier, as the Deep Wavelet cannot do classification on its own [34].

3.3 Support Vector Machine

SVMs are a group of similar supervised learning techniques that are applied to regression and classification. They are a member regarding the generalized linear classification family [35]. Through converting the original training data, which could be utilized as SVM input, into multidimensional space and building a hyper-plane in higher dimensions for observing the hyperplane between 2 different classes in high dimensional feature space, which could be utilized as classification, the SVM, also referred to as SVM, is a potent classification technique [36] [37].

3.4 K-Nearest Neighbors (KNN)

One popular classification method is KNN [38]. It is considered to be one of the most straightforward approaches in DM and ML and is instance-based, non-parametric, or lazy [39]. Its short computation time and ease of interpretation make it a popular choice [38]. The KNN method operates on the premise that samples that are most comparable to one another and belong to the same class have a higher probability [39]. Generally, KNN method predicts the query with the major class in k nearest neighbors after first locating the query's k nearest neighbors in training data-set. KNN is hence susceptible to the choice of k value. In the KNN algorithm, determining the k value is still quite difficult [38], [39]. Two parameters that must be accessible on distinct k values are the training error rate and the validation error rate [40].

3.5 Evaluation criteria

The criteria used for evaluating the classifiers in this paper is the accuracy which can be calculated by the equation below:

$$\boldsymbol{Acc = \frac{TN+TP}{TN+FP+TP+FN}} \quad \mathbf{(5)}$$

Whereas: TP = True Positive, TN = True Negative, FN = False Negative and FP = False Positive [41].

4 .Methodology

In this paper a levels of detection and classification methods for eye-related diseases have been accomplished: level one is implemented by using a parallel convolution neural network to detect whether there is any disease or not regardless of type. Level two classifies four eye-related diseases. Level one was implemented using a convolution neural network, while, level two was implemented by using a deep wavelet network (**see Fig. 3**). Each level will be explained in detail.

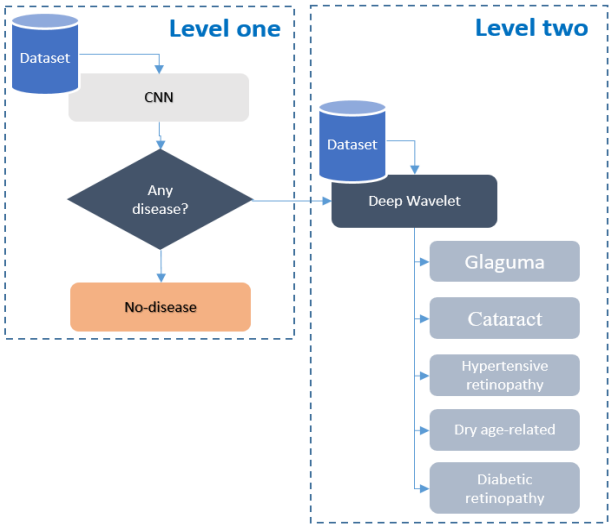

Fig.3. The diagram of the two- levels.

4.1 Level –One Detection

Level one detects whether there is any disease or not. The answer for the model will be either: healthy or infected. A proposed model is designed using a parallel convolution neural network by using a dataset with two classes (see Fig. 4).

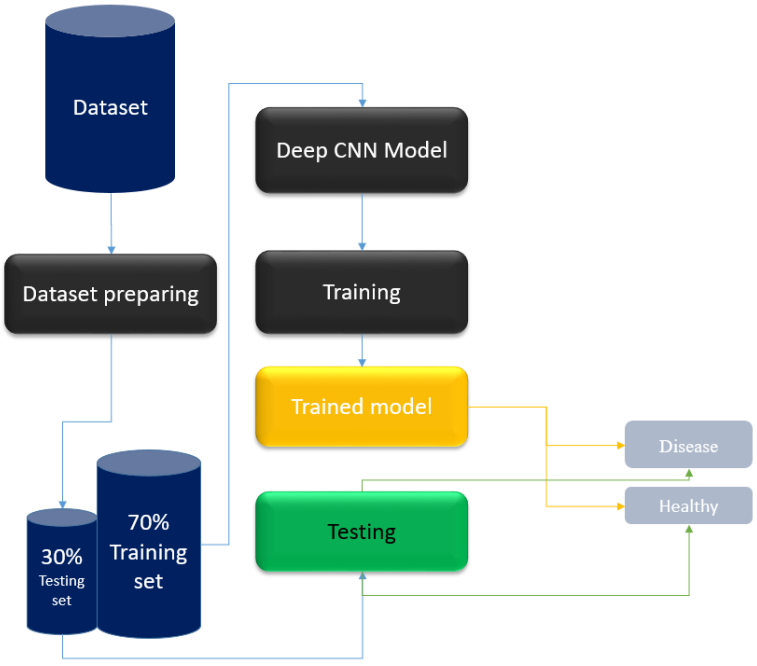

Fig.4. Diagram of the level one detection model.

4.1.1 Level –One Detection Model Model Building

Our suggested model design combines CNN) parallel and sequential layers. The Autoencoder: Upsampling and Downsampling technique serves as its foundation. Layers for activation, convolution,

batch normalization, and max-pooling are used. The flattened layer was next applied, then Softmax and the classification layer. The suggested model architecture will be covered thoroughly in the upcoming sections. The first layer of CNN is image input layer, as is customary. Its size, [200×200×1], is adjusted to match the size of training dataset samples. The input image passes via two parallel convolutional layers during the downsampling process. To identify features in the photos, such layers apply filters that move over them. Each convolutional layer adjusts with a filter size of [3×3], which concentrates on particular details within prominent features since the blood vessels are slightly smaller. A total of 32 filters are employed in such layers to create 32 feature maps, ensuring that all important information is extracted from the input images. The batch normalization layer, which comes after such layers, normalizes the output of the preceding layer, speeds up training, uses higher learning rates, and facilitates learning. The network is after that able to discover intricate patterns and relationships in the data through the introduction of non-linearity through the application of ReLu activation function. Finally, the Max pooling layer is used to preserve the most significant features while decreasing spatial dimensions of feature maps that the convolution layer outputs. Next, a concatenation layer that concatenates inputs along a certain dimension is used to connect the two parallel sets of layers. The preceding layer group was succeeded by an additional group consisting of a single convolutional layer with a filter size of [3, 3] and 16 filters employed. ReLu, max pooling layer, and batch normalization layer come next. The output was after that linked to two sets of parallel convolutional layers, each consisting of eight filters and having a filter size of [3,3]. Both then moved on to the maxpooling layer, ReLu, and batch normalization layer. Next, a concatenation layer is utilized in order to join the two parallel groups of layers. The previous configuration of layers with the identical filter numbers and sizes was repeated in the upsampling, but in reverse order. Lastly, a dense layer, sometimes referred to as a fully connected layer, is employed. It is in charge of extracting high-level features from the layers' output. Since there are two classes in the dataset, this layer's output size is two. After that, to provide outputs that are probability-like for the class to which it belongs, the softmax layer is employed. In the case when the input image is healthy or has any diseases, the classification layer is employed to provide the final class predictions (**see Fig.5**).

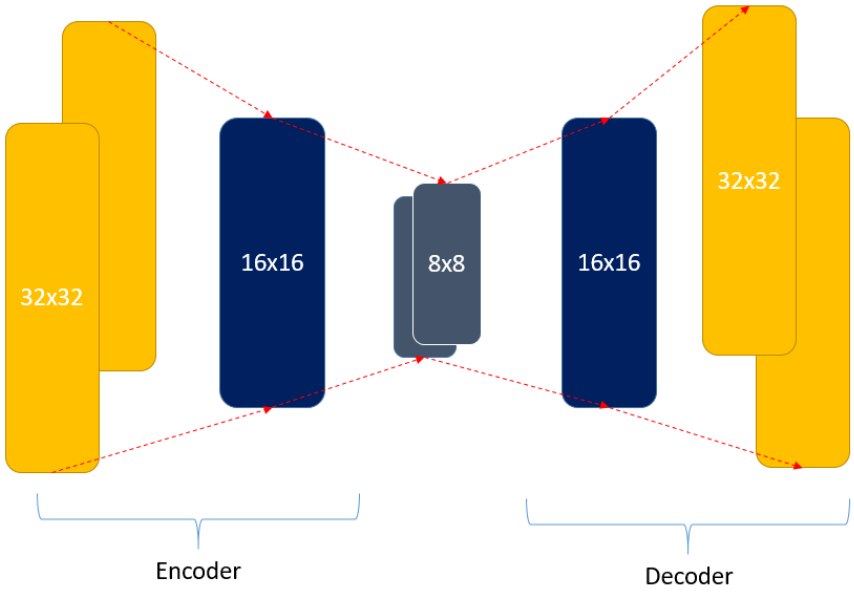

Fig.5. the proposed model Autoencoder for encoding and decoding

4.2 Level –Two Classification Model

Level two classifies five eye-related diseases. The output for the model will be either one of these diseases: Glaguma, Cataract, Hypertensive retinopathy, Dry age-related, or Diabetic retinopathy. This model was designed using a Deep wavelet network by using a dataset with five classes (**see Fig. 6**).

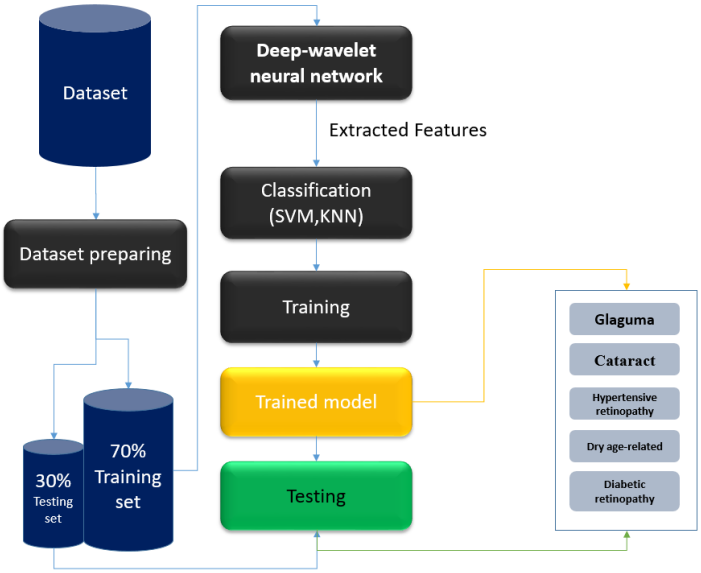

Fig.6. The block diagram of the level two classification model

The two primary components of this classification level are feature extraction and classification. Key elements of the image have been extracted using a two-layered deep wavelet as a feature extraction method.

1- First, the input image is filtered using a scaled Gaussian (low-pass) function φ using convolving. In doing so, the image's finer details are eliminated, and a fundamental feature known as 0th-order scattering coefficients is produced. The input for the following level will be the 0th-order scattering coefficients.

2- In layer one: We employed the Morlet wavelet bank, a kind of complex wavelet transform with an easy mathematical representation, to apply every one of the wavelet filters in the first filter bank to the image. Those filters covered various image rotations and scales, allowing for a thorough examination of the supplied data. We set the quality factor (the number of wavelet filters per octave) to 8 for this level, the number of rotations to 5, and the Invariance Scale of the filter to 56. The modulus of every filtered output was then computed. The first-order scattering coefficients are obtained by combining such moduli with the scaling filter. This layer examines the lost details in the zero layer in greater detail.

3-Layer Two: Repeat the process for layer one, where the input for each new layer is the output of the previous layer. The second-order scattering coefficients were computed. This process will be more flexible and will help capture complex patterns better because the previous processes lead to nonlinearity **(see Fig.7).**

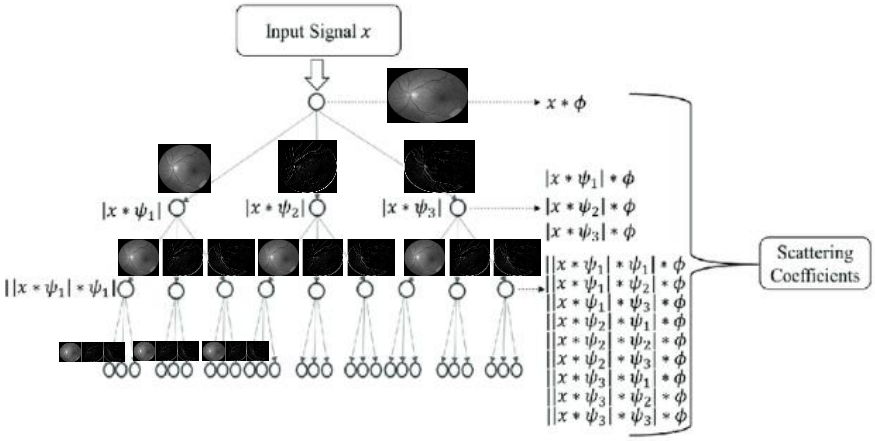

Fig.7. Deep wavelet scattering network layers of level–two classification model

After the features were extracted from the input images, two types of classifiers were used: SVM and KNN.

5. Results and Discussions

This section describes the stages of the suggested model's results, which include detection and classification when applied to two distinct datasets.

5.1 Dataset

The data is the fuel in the learning process, in this paper we have used three different datasets one for each level:

A. The Ocular Toxoplasmosis Dataset

The dataset was downloaded from the website kaggle.com, and it is freely available on the website [42].This dataset consists of an eye image collected from two hospitals both de Clínicas Medical Center and Niños de Acosta Ñú General Pediatric. Two classes are included in this dataset: 132 images for the Healthy class and 279 for the diseased with different sizes. The images have been preprocessed by Alam et al [43]. This dataset has many issues which have been described and solved as follows:

1- There are two classes in the dataset: diseased and healthy. The augmentation procedure was used on the original images since CNN may be used to attain great accuracy on larger datasets. New fake samples can be created and mixed with the existing data. Data augmentation is the term for this procedure [44]. The samples were turned and flipped to three distinct angles: 90°, 180°, and 270°. For improving accuracy and lessen overfitting throughout the DNN's learning process, the dataset is expanded with newly generated images. (**See Table 1**) which listed the entire number of dataset images, both prior to and following augmentation.

Table 1. Dataset images samples number before and after augmentation

Classes	No. of samples before augmentation	No. of samples after augmentation
Healthy	132	528
Diseased	279	1116

2- Class Imbalance: the classes of the dataset are imbalanced which means there is an unequal distribution of labeled images in different classes. CNN's classification performance will be impacted by the class imbalance. Our solution involved the use of under-sampling. This is a method for modifying a data set's class distribution. Under-sampling techniques function by minimizing the bulk of class samples (**see Fig. 8**).

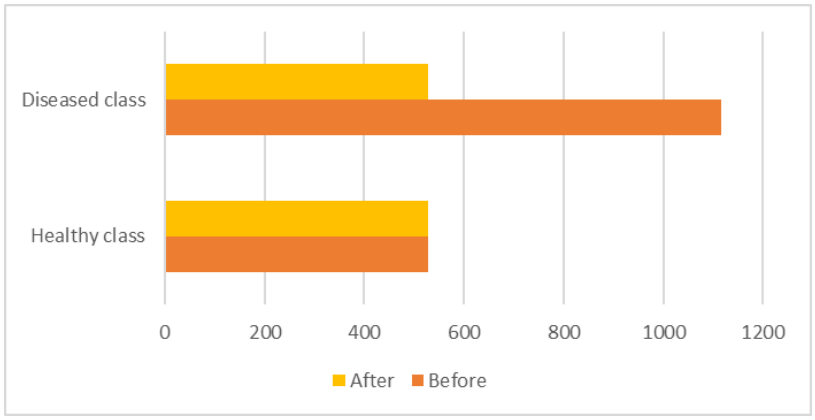

Fig.8. Class imbalance before and after the Downsampling technique.

3- Dataset images are of different sizes. The image size has been unified within [200×200].

4- Lastly, a training set and a test set are created from the data-set. 30% has been used for testing and 70% was used for training the original dataset. Randomly chosen examples from every class are given to each set.

B. Ocular Disease Intelligent Recognition (ODIR) Dataset

One of the most extensive datasets for identifying eye diseases that is publicly accessible on Kaggle is this one [45]. The fundus images are of a patient's left and right eyes, and there are 5000 patients in total. Zeiss, Canon, and Kowa cameras are utilized in order to take the fundus images. The sizes of these images vary. Eight categories for the classification of ocular diseases make up the multi-class multi-label database known as ODIR. The seven disease classifications that fall under such categories are: age-related macular degeneration (A), glaucoma (G), cataract (C), myopia (M), normal (N), diabetes (D), hypertension (H), and other abnormalities/diseases (O). Several operations have been took place to prepare the dataset for the next process (**see Figure 9**) which summarizes all these operations.

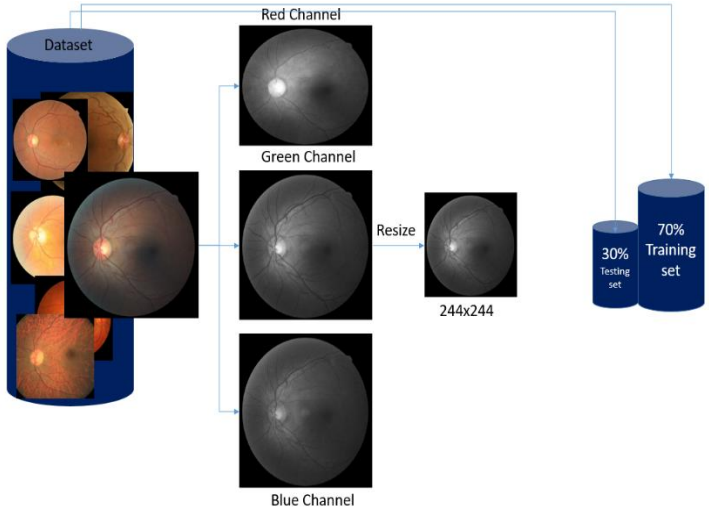

Fig. 9. Dataset preparation.

5.2 Level One detection model

Because CNN is established on the concept of trial and error. Two experiments have been examined in which different numbers of filters as well as sizes have been tried.

1- The image of size 200x200x1 convolved with [parallel 32 filters, 16 filters, parallel 8 filters to form encoder] Then [16 filters, parallel 32 filters to form decoder] all with size of 3x3.

2- Image of size 200x200x1 convolved with [parallel 32 filters, 16 filters, parallel 8 filters to form encoder] Then [16 filters, parallel 32 filters to form decoder] all with size of 5x5.

(**See Table 2**) lists the accuracy, the epoch number of training, and the training time taken in minutes of each experiment.

Table 2. Results of the two experiments.

Experiments No.	Accuracy	Epoch Number	Training Time (Min)
Experiment 1	94.2%	691	2315
Experiment 2	93.2%	100	297

From table 2, we could notics that expermint 1 with filter size of 3x3 has the highest accuracy that means the size filter could affect on the final results, as well as, the selection of filter size depend on the input images itself.

State of Art Results Comparison

(**See Table 3**) represents a comparison between the results of the previous and the results obtained by the level one detection model. As shown, the proposed obtained the highest results.

Table 3. Comparison between the results.

Method	Accuracy %
Rodrigo PARRA et. al.[46]	91.3
Adithi D. Chakravarthy et. al. [47]	92.5
Daniel Milada et. al. [48]	93.0
Bambang Krismono et. al.[49]	80.93
Level one detection model (Experiment 1)	94.2
Level one detection model (Experiment 2)	93.2

5.3 Level –Two Classification Model

Cross-validation technique used, in which every training process a new class added and evaluate the accuracy of the model. In addition, the data downsampling performed in each experiment depends on the total number of the minor and major classes in that experiment (**see Table 4**). As well as , all the classifier's accuracy for the four experiments listed (**see Table 5**)

Table 4. details classes' total number and downsampling process.

Experiment No.	No. of classes	Classes used (diseases)	Class Samples	No. of samples (by downsampling)
Experiment 1	2	Diabetic retinopathy	1098	1038
		Cataract	1038	
Experiment 2	3	Diabetic retinopathy	1098	1038
		Cataract	1038	

		Dry age-related	1272	
Experiment 3	4	Diabetic retinopathy	1098	1007
		Cataract	1038	
		Dry age-related	1272	
		Glaguma	1007	
Experiment 4	5	Diabetic retinopathy	1098	808
		Cataract	1038	
		Dry age-related	1272	
		Glaguma	1007	
		Hypertensive retinopathy	808	

Table 5. Classifiers accuracy for the four experiments.

No. of Experiment	KNN	SVM
Experiment 1	98.5%	98.6%
Experiment 2	95.1%	96.2%
Experiment 3	84.9%	86.9%
Experiment 4	83.6%	80.1%

State of Art Results Comparison

Table 6 represents a comparison between the results of the previous and the results obtained by the level one detection model. As shown, the proposed work obtained the highest results

Table 6. Comparison between the results.

Methods	Methods	Dataset	Accuracy
Zainoor AhmadChoudhry et. al. [50]	SVM	ODIR	93.8%
Ajna Ram et. al. [51]	CNN	ODIR	88%
Md. Tariqul Islam et. al.[52]	CNN	ODIR	80.5%
Experiment 1	DWNN + SVM	ODIR	95.5%
	DWNN + KNN	ODIR	98.6%
Experiment 2	DWNN + SVM	ODIR	95.1%
	DWNN + KNN	ODIR	96.2%
Experiment 3	DWNN + SVM	ODIR	84.9%
	DWNN + KNN	ODIR	86.9%

Experiment 4	DWNN + SVM	ODIR	83.6%
	DWNN + KNN	ODIR	80.1%

6. Conclusions

We have presented in this research a two-stage method for the classification and detection of diseases connected to the eyes that is implemented in two levels: level one uses a parallel CNN for the identification of the presence or absence of any disease, irrespective of its form. With the use of two classifiers, which are KNN and SVM, and two datasets, the Ocular Toxoplasmosis Dataset and Ocular Disease Intelligent Recognition (ODIR), level two identifies four eye-related disorders. Deep-wavelet is used as a technique for FE. We came to the conclusion that a large number of samples are typically required for the training when using the DL method. As a result, overfitting may result from a small number of images. At the same time, the deep wavelt needs less training examples compared to DL techniques, which makes it a strong option for labeled applications. The deep wavelet network cannot classify images on its own; instead, its outputs should be utilized as training data for a classifier. As a result, when features are extracted, the classifier selection is crucial for classifying images and finding abnormalities in images. We have decided that our approach performs well and produces decent results based on the comparative estimate.

References

[1] I. D. of E. & E. E. KIET Group of Institutions (Ghāziabād, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, and Institute of Electrical and Electronics Engineers., 2nd IEEE International Conference on Innovative Applications of Computational Intelligence on Power, Energy and Controls with Their Impact on Humanity (CIPECH-16).

[2] O. Sule and S. Viriri, "Enhanced Convolutional Neural Networks for Segmentation of Retinal Blood Vessel Image," in 2020 Conference on Information Communications Technology and Society, ICTAS 2020 - Proceedings, Institute of Electrical and Electronics Engineers Inc., Mar. 2020. doi: 10.1109/ICTAS47918.2020.233996.

[3] T. J. Jebaseeli, C. A. Deva Durai, and J. D. Peter, "Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM," Optik (Stuttg), vol. 199, Dec. 2019, doi: 10.1016/j.ijleo.2019.163328.

[4] E. Uysal and G. E. Güraksin, "Computer-aided retinal vessel segmentation in retinal images: convolutional neural networks," Multimed Tools Appl, 2020, doi: 10.1007/s11042-020-09372-w.

[5] H. Boudegga, Y. Elloumi, M. Akil, M. Hedi Bedoui, R. Kachouri, and A. Ben Abdallah, "Fast and efficient retinal blood vessel segmentation method based on deep learning network," Computerized Medical Imaging and Graphics, vol. 90, Jun. 2021, doi: 10.1016/j.compmedimag.2021.101902.

[6] Z. Baharlouei, H. Rabbani, and G. Plonka, "Wavelet scattering transform application in classification of retinal abnormalities using OCT images," Sci Rep, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-46200-1.

[7] U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunović, "Artificial intelligence in retina," Nov. 01, 2018, Elsevier Ltd. doi: 10.1016/j.preteyeres.2018.07.004.

[8] M. Zhao and G. Hamarneh, "Retinal Image Classification via Vasculature-guided Sequential Attention."

[9] M. M. Islam, T. N. Poly, B. A. Walther, H. C. Yang, and Y. C. Li, "Artificial intelligence in ophthalmology: A meta-analysis of deep learning models for retinal vessels segmentation," J Clin Med, vol. 9, no. 4, Apr. 2020, doi: 10.3390/jcm9041018.

[10] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," Sensors, vol. 21, no. 11, Jun. 2021, doi: 10.3390/s21113704.

[11] Y. Jiang, H. Zhang, N. Tan, and L. Chen, "Automatic retinal blood vessel segmentation based on fully convolutional neural networks," Symmetry (Basel), vol. 11, no. 9, 2019, doi: 10.3390/sym11091112.

[12] D. Chen, Y. Ao, and S. Liu, "Semi-supervised learning method of U-net deep learning network for blood vessel segmentation in retinal images," Symmetry (Basel), vol. 12, no. 7, Jul. 2020, doi: 10.3390/SYM12071067.

[13] S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds., Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, vol. 9901. in Lecture Notes in Computer Science, vol. 9901. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-46723-8.

[14] M. R. K. Mookiah et al., "A review of machine learning methods for retinal blood vessel segmentation and artery/vein classification," Feb. 01, 2021, Elsevier B.V. doi: 10.1016/j.media.2020.101905.

[15] T. A. Soomro et al., "Deep Learning Models for Retinal Blood Vessels Segmentation: A Review," 2019, Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ACCESS.2019.2920616.

[16] K. Irsch and D. L. Guyton, "Anatomy of Eyes," in Encyclopedia of Biometrics, Springer US, 2009, pp. 11–16. doi: 10.1007/978-0-387-73003-5_253.

[17]. Das, K. Kharbanda, S. M, R. Raman, and E. D. D, "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy," Biomed Signal Process Control, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102600.

[18] M. Badar, M. Haris, and A. Fatima, "Application of deep learning for retinal image analysis: A review," Feb. 01, 2020, Elsevier Ireland Ltd. doi: 10.1016/j.cosrev.2019.100203.
[19] Institute of Electrical and Electronics Engineers and IEEE Instrumentation and Measurement Society, IST 2017 : IEEE International Conference on Imaging Systems and Techniques : Beihang University, Beijing, China, October 18-20, 2017, Beijing China : 2017 conference proceedings.
[20] F. Y. Shih and H. Patel, "Deep Learning Classification on Optical Coherence Tomography Retina Images," Intern J Pattern Recognit Artif Intell, vol. 34, no. 8, Jul. 2020, doi: 10.1142/S0218001420520023.
[21] P. Rajpurkar, E. Chen, O. Banerjee, and E. J. Topol, "AI in health and medicine," Nat Med, vol. 28, no. 1, pp. 31–38, Jan. 2022, doi: 10.1038/s41591-021-01614-0.
[22] M. Alsaidi, A. S. Altaher, M. T. Jan, A. Altaher, and Z. Salekshahrezaee, "COVID-19 Classification Using Deep Learning Two-Stage Approach," Nov. 2022, [Online]. Available: http://arxiv.org/abs/2211.15817
[23] M. A. de Santana and W. P. dos Santos, "A deep-wavelet neural network to detect and classify lesions in mammographic images," Research on Biomedical Engineering, vol. 38, no. 4, pp. 1051–1066, Sep. 2022, doi: 10.1007/s42600-022-00238-8.
[24] V. A. de Freitas Barbosa, A. Félix da Silva, M. A. de Santana, R. Rabelo de Azevedo, R. de C. Fernandes de Lima, and W. P. dos Santos, "Deep-Wavelets and convolutional neural networks to support breast cancer diagnosis on thermography images," Comput Methods Biomech Biomed Eng Imaging Vis, vol. 11, no. 3, pp. 895–913, May 2023, doi: 10.1080/21681163.2022.2118174.
[25] N. S. Russel and A. Selvaraj, "Wavelet scattering transform and deep features for automated classification and grading of dates fruit," J Ambient Intell Humaniz Comput, vol. 15, no. 6, pp. 2909–2923, Jun. 2024, doi: 10.1007/s12652-024-04786-y.
[26] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," Electronic Markets, vol. 31, no. 3, pp. 685–695, Sep. 2021, doi: 10.1007/s12525-021-00475-2.
[27] N. Sharma, V. Jain, and A. Mishra, "An Analysis of Convolutional Neural Networks for Image Classification," in Procedia Computer Science, Elsevier B.V., 2018, pp. 377–384. doi: 10.1016/j.procs.2018.05.198.
[28] A. Shamaei, J. Starčuková, and Z. S. Zenon Starčuk, "A wavelet scattering convolutional network for magnetic resonance spectroscopy signal quantitation," in BIOSIGNALS 2021 - 14th International Conference on Bio-Inspired Systems and Signal Processing; Part of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2021, SciTePress, 2021, pp. 268–275. doi: 10.5220/0010318502680275.
[29] W. Ghezaiel, L. Brun, and O. Lezoray, "Wavelet Scattering Transform and CNN for Closed Set Speaker Identification," in IEEE 22nd International Workshop on Multimedia Signal Processing, MMSP 2020, Institute of Electrical and Electronics Engineers Inc., Sep. 2020. doi: 10.1109/MMSP48831.2020.9287061.
[30] E. Oyallon and S. Mallat, "Deep Roto-Translation Scattering for Object Classification." [Online]. Available: http://www.di.ens.fr/data/software.
[31] B. Soro and C. Lee, "A wavelet scattering feature extraction approach for deep neural network based indoor fingerprinting localization," Sensors (Switzerland), vol. 19, no. 8, Apr. 2019, doi: 10.3390/s19081790.
[32] L. R. Dsilva et al., "Wavelet scattering- and object detection-based computer vision for identifying dengue from peripheral blood microscopy," Int J Imaging Syst Technol, vol. 34, no. 1, Jan. 2024, doi: 10.1002/ima.23020.
[33] C. Shi et al., "Wavelet Scattering Convolution Network-Based Detection Algorithm on Nondestructive Microcrack Electrical Signals of Eggs," Agriculture (Switzerland), vol. 13, no. 3, Mar. 2023, doi: 10.3390/agriculture13030730.
[34] S. Agerhäll, A. Gallouët, G. Lacharme, H. Leterme, F. Rubes, and V. Perrier, "The Wavelet Scattering Network for Image Classification Project Report," 2019.
[35] H. Bhavsar and M. H. Panchal, "A Review on Support Vector Machine for Data Classification," 2012.
[36] St. Peter's College of Engineering and Technology, Institute of Electrical and Electronics Engineers, Council of Scientific & Industrial Research (India), and Indian Council of Medical Research, Proceedings of 2017 Third IEEE International Conference on Sensing, Signal Processing and Security (ICSSS 2017) : May 4th and 5th, 2017 : St. Peter's College of Engineering and Technology, Avadi, Chennai, Tamil, Nadu, India-600054.
[37] M. A. Chandra and S. S. Bedi, "Survey on SVM and their application in image classification," International Journal of Information Technology (Singapore), vol. 13, no. 5, Oct. 2021, doi: 10.1007/s41870-017-0080-1.
[38] ICIT 2017 : the 8th International Conference on Information Technology : Internet of Things IoT : conference proceedings : May 17th - 18th, 2017, Amman, Jordan. IEEE, 2017.
[39] S. Zhang, D. Cheng, Z. Deng, M. Zong, and X. Deng, "A novel kNN algorithm with data-driven k parameter computation," Pattern Recognit Lett, vol. 109, pp. 44–54, Jul. 2018, doi: 10.1016/j.patrec.2017.09.036.
[40] W. Du, Z. Zhan, and W. Du Zhijun Zhan, "Building Decision Tree Classifier on Private Data," 2002. [Online]. Available: https://surface.syr.edu/eecshttps://surface.syr.edu/eecs/8
[41] Institute of Electrical and Electronics Engineers, 2018 2nd International Conference on Engineering Innovation (ICEI) : July 5-6, 2018, Bangkok, Thailand.
[42] O. Cardozo et al., "Dataset of fundus images for the diagnosis of ocular toxoplasmosis," Data Brief, vol. 48, p. 109056, Jun. 2023, doi: 10.1016/j.dib.2023.109056.
[43] S. S. Alam, S. B. Shuvo, S. N. Ali, F. Ahmed, A. Chakma, and Y. M. Jang, "Benchmarking Deep Learning Frameworks for Automated Diagnosis of Ocular Toxoplasmosis: A Comprehensive Approach to Classification and Segmentation," May 2023, [Online]. Available: http://arxiv.org/abs/2305.10975
[44] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, "Retinal image quality assessment using deep learning," Comput Biol Med, vol. 103, pp. 64–70, Dec. 2018, doi: 10.1016/j.compbiomed.2018.10.004.
[45] "Ocular disease recognition. https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k ".
[46] R. Parra et al., "Automatic diagnosis of ocular toxoplasmosis from fundus images with residual neural networks," in Public Health and Informatics: Proceedings of MIE 2021, IOS Press, 2021, pp. 173–177. doi: 10.3233/SHTI210143.
[47] A. D. Chakravarthy et al., "An approach towards automatic detection of toxoplasmosis using fundus images," in Proceedings - 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering, BIBE 2019, Institute of Electrical and Electronics Engineers Inc., Oct. 2019, pp. 710–717. doi: 10.1109/BIBE.2019.00134.

[48] D. Milad, F. Antaki, A. Bernstein, S. Touma, and R. Duval, "Automated Machine Learning versus Expert-Designed Models in Ocular Toxoplasmosis: Detection and Lesion Localization Using Fundus Images," Ocul Immunol Inflamm, pp. 1–7, Feb. 2024, doi: 10.1080/09273948.2024.2319281.
[49] B. K. Triwijoyo, B. S. Sabarguna, W. Budiharto, and E. Abdurachman, "Deep learning approach for classification of eye diseases based on color fundus images," in Diabetes and Fundus OCT, Elsevier, 2020, pp. 25–57. doi: 10.1016/B978-0-12-817440-1.00002-4.
[50] Z. AhmadChoudhry, H. Shahid, S. Z. H. Naqvi, S. Aziz, and M. U. Khan, "DarkNet-19 based Decision Algorithm for the Diagnosis of Ophthalmic Disorders," in 2021 International Conference on Innovative Computing (ICIC), IEEE, Nov. 2021, pp. 1–6. doi: 10.1109/ICIC53490.2021.9693030.
[51] A. Ram and C. C. Reyes-Aldasoro, "The relationship between Fully Connected Layers and number of classes for the analysis of retinal images," Apr. 2020, [Online]. Available: http://arxiv.org/abs/2004.03624
[52] Md. T. Islam, S. A. Imran, A. Arefeen, M. Hasan, and C. Shahnaz, "Source and Camera Independent Ophthalmic Disease Recognition from Fundus Image Using Neural Network," in 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), IEEE, Nov. 2019, pp. 59–63. doi: 10.1109/SPICSCON48833.2019.9065162.