ACTA IMEKO
ISSN: 2221-870X
September 2023, Volume 12, Number 3, 1 - 7

Optical coherence tomography based diabetic – ophthalmic disease classification and measurement using bilateral filter and transfer learning approach

K. Yojana[1], L. Thillai Rani[1]

[1] *Department of Electronics and Instrumentation Engineering, Annamalai University, Tamilnadu, India*

ABSTRACT
Optical Coherence Tomography (OCT) is a smooth application of low coherence interferometer with high air resolution and highly sensitive heterodyne detection technology to tomographic image measurement of living organisms. Currently, clinical applications are becoming more widespread in ophthalmology, cardiovascular system, dermatology, and dentistry. The problem with OCT is that the measurement area is as narrow as a few mm compared to other tomographic image measurement techniques, and it was initially applied to ophthalmology. Since then, various researches and developments have been carried out to expand clinical applications. Michelson type fiber optic interferometer is used for image acquisition. This paper presents a classification of ophthalmic diseases caused by diabetes. Bilateral filter is used for image pre-processing and noise removal. A transfer learning approach is implemented which uses AlexNet and Support vector machine (SVM) to classify the images. The AlexNet model is used to extract the features form the images and these features are classified using SVM model. The novelty of the proposed model lies in the use of image denoising using bilateral filter and then classification of the AlexNet features using SVM thereby achieving better classification accuracy with less training data. This also leads to better resource utilization. The ailments under study are Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), DRUSEN, and NORMAL. The proposed approach produced a higher classification accuracy of 99 % when compared to other deep learning algorithms like CNN, AlexNet and GoogleNet. The precision, sensitivity and specificity are recorded as 0.98, 0.99, and 0.99 respectively.

Section: RESEARCH PAPER

Keywords: Optical Coherence Tomography (OCT); AlexNet; SVM; ophthalmic diseases; transfer learning; bilateral filter; Michelson type fiber optic interferometer; tomographic image measurement

Citation: K. Yojana, L. Thillai Rani, Optical coherence tomography based diabetic – ophthalmic disease classification and measurement using bilateral filter and transfer learning approach, Acta IMEKO, vol. 12, no. 3, article 56, September 2023, identifier: IMEKO-ACTA-12 (2023)-03-56

Section Editor: Francesco Lamonaca, University of Calabria, Italy

Received July 14, 2022; **In final form** June 3, 2023; **Published** September 2023

Corresponding author: K. Yojana, e-mail: kyojana55@gmail.com

1. INTRODUCTION

Optical Coherence Tomography (OCT) is a non-invasive medical imaging technique used to produce high-resolution cross-sectional images of the retina and other parts of the eye. OCT works by using light waves to produce detailed images of the internal structures of biological tissues, such as the retina. The basic principle of OCT is low-coherence interferometry, in which light from a source is split into two beams: one is directed into the eye and the other serves as a reference. The light that enters the eye is partially reflected by various internal structures, and the returning light is combined with the reference light. The interference pattern produced by this combination is analyzed to determine the distance and reflectivity of the internal structures, which can then be used to create cross-sectional images of the tissue (e.g., Figure 1). The technology provides high-resolution images of tissue structure without the need for dyes or radiation, making it a non-invasive and safe diagnostic tool.

OCT is a safe and painless option for patients as it does not require the use of dyes, needles, or radiation. It can detect changes in the tissue structure early on, which can help in the early diagnosis and treatment of various eye diseases, such as age-related macular degeneration and glaucoma. The images can be repeated as often as needed, allowing for ongoing monitoring and comparison of changes over time. OCT images can provide important information about the structure and thickness of the

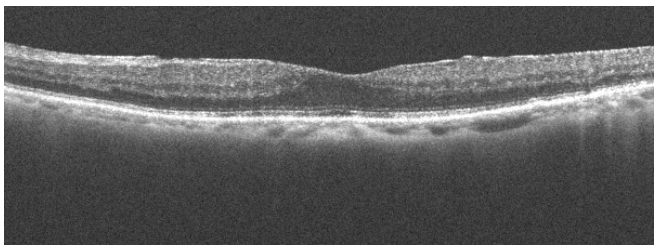


Figure 1. Example of fundus image obtained by OCT

retina, which can be used to guide treatment decisions for conditions such as macular holes and retinal detachments.

Choroidal Neovascularization (CNV) appears as a hyper-reflective (bright) lesion in the choroid on Optical Coherence Tomography (OCT) images. The lesion typically appears as an irregular, poorly defined area with increased thickness compared to the surrounding choroid. The choroidal neovascularization can cause fluid and blood to leak into the retina, leading to vision loss.

Diabetic Macular Edema (DME) appears as a thickened, hyper-reflective (bright) area in the center of the retina on Optical Coherence Tomography (OCT) images. The thickening is caused by accumulation of fluid in the retina, which can lead to swelling and vision loss. In DME, the retina appears thicker than normal, and there is often an increased level of reflectivity in the retina due to the accumulation of fluid. On OCT, the foveal center may appear blurred and the retinal layer structure may be distorted. In severe cases, the fluid accumulation can lead to the formation of cystoid spaces within the retina, which appear as fluid-filled cavities on OCT images.

Drusen can be detected in optical coherence tomography (OCT) images of the retina. OCT is a non-invasive imaging technique that uses light waves to create high-resolution cross-sectional images of the retina. In OCT images, drusen appear as hyper-reflective or bright spots between the retina and the underlying layers. They can be classified based on their size, shape, and location within the retina. Hard drusen are usually smaller and more uniform in shape, while soft drusen are larger and have a more irregular shape.

In the proposed model, the bilateral filter is used to denoise the image without burring the edges. Conventional denoising techniques like non-local means, median, mean etc produces blurring during the denoising process. The blurring in the input images reduces the training accuracy. The results of denoising obtained by the bilateral filter are fed to the AlexNet model for feature extraction. The presence of noise effects the training process as features cannot be extracted and accurate filters cannot be selected. The use of transfer learning reduces the training time and the resourced needed. The filters in the convolution layers extract useful features for classification of the images by a machine learning model. Here, Support Vector Machine (SVM) has been used to classify the features extracted by the AlexNet model. This increases the training accuracy while optimally utilizing the resources. Section 1 contains the introduction to the paper. Section 2 depicts the literature survey. Section 3 presents the proposed model description. Section 4 presents the experimental results followed by conclusion and references.

2. LITERATURE

This section describes the state of art works in the field of ophthalmic disease classification. In [1], the optical coherence tomography images with classification were presented based on a capsule network using an OCT dataset with a training dataset of 83,484 images. From each category, a total number of 250 images include in the test dataset. The dataset consists of four classes: choroidal neovascularization (CNV), Diabetic Macular Edema (DME), DRUSEN and normal. A capsule network was used to establish a proposed model that improves classification accuracy. The training dataset was used to train the capsule network while the test dataset was considered for evaluation of trained model.

In [2], the author was proposed an automatic technique based on deep learning for detection of Age-related Macular Degeneration (AMD) and DME lesions. The small datasets problem was solved by implementing the network model based on reuse feature and adoption of two available OCT datasets of retina. It would improve the difference in the different datasets of approach. However, different network models were realized by comparing with the reusable feature and the transfer learning with pre-trained models on networks. Within the scope of this study, the proposed model is evaluated using data from two separate datasets. The dataset consists of three classes: AMD, AME and normal. The first dataset has a total of 3231 images, 1101 of which have DME, 723 of which have AMD, and 1407 of which are normal. The second dataset has a total of 5,084 images, 1185 of which have DME, 1524 of which have AMD, and 2375 of which are normal.

In [3], authors were proposed a weakly supervised deep learning technique based on uncertainty estimation for addressing the issue of macula-related disease classification from available OCT images using the volume-level label. The proposed uncertainty-driven deep multiple instances learning scheme uses for refining the convolutional neural network (CNN) based instance-level classifier iteratively that could enable detecting the abnormal suspicious instances and the relevant deep embedding is abstracted based on the higher capability of representation. From the same bag, the instance features are considered as input using a recurrent neural network (RNN) and the final bag-level prediction is generated based on the individual local instance data and representation of global aggregated bag-level information. However, more resources are consumed due to the large architecture. There are just two classes included in this data set: DME and normal. The authors constructed two large DME-OCT datasets using a variety of devices and imaging techniques in order to test the usefulness of the proposed model.

In [4], a surrogate-assisted classification technique was proposed using CNN. Primarily, image denoising performs for minimizing the noise. The masks extract using the morphological dilation and thresholding as well. More surrogate images are generated using the masks and denoised images that could be utilized for training the CNN model. However, the proposed technique evaluates using different databases. The higher accurate results were not obtained using the CNN model. In order to evaluate the efficiency of the proposed method, authors two databases were used: a private database and a public database. The data sets are consisting three classes: normal, AMD, and DME.

In [5], a pre-trained CNN was used to classify the normal and abnormal OCT (Optical Coherence Tomography). In this work, authors used a small dataset which consist 32 OCT images are

classified into two classes: DME and Normal. At different network layers, features have been extracted based on VGG16 (Visual Geometry Group). For example, after each fully connected layer and before fully connected layer. Different classifiers were used to perform classification of features. The simulation results show the accuracy for same dataset as 87.5% with specificity of 81% and sensitivity of 93.5%. The lower accuracies were achieved using the VGG16 model.

In [6], a step of image segmentation was proposed that could be categorized the optical coherence tomography (OCT) for detection of drusen shape in RPE layer and determination of retinal pigment epithelium (RPE) layer. In RNFL complex, the RPE layer uses to determine the retinal nerve fiber layer (RNFL) and identify the bubble of blood area. A binary classification is used in this technique for classification of characteristics of two diseases like DME and AMD. The desired results wouldn't generate by the layer segmentation of OCT images and disease classification is made at an image level.

In [7], the authors were used different automated methods based on deep learning while demonstrating the utilization of deep learning for diagnosing the diabetic retinopathy based on OCT images. For this task, existing deep learning models were retrained and investigated the optimization of a retrained model. It's effective to use the optimized pre-trained model as a feature extractor and trained conventional classifier for diagnosis of diabetic retinopathy.

In [8], an automated system was presented to discover and grade the non-proliferative DR from the OCT images. Three sequential images were used in the proposed CAD system. Based on earlier segmentation approach using the integrated joint model, twelve distinct retina layers have been localized primarily. Here, the features like spatial data, intensity, and shape were combined. From the segmented layers, the features of thickness, curvature, and reflectivity were determined. The database has 109309 images which consists of four classes: Choroidal Neovascularization (CNV), DME, DRUSEN, and NORMAL pictures. To classify the subject as DR, or normal, a two-stage deep fusion classification network (DFCN) that was trained by the stacked non-negativity constraint autoencoder (SNCAE) has been utilized. Further, DR grade is also analyzed in terms of either mild/moderate or early stage.

From the literature survey, most of the research works are focused to classify the DME and DRUSEN deceases. In the conventional models, on the OCT images results in lower accuracies because of the noise present in the images. This noise prevents the deep learning models to select accurate filters for feature extraction during the training stage. During the training process, the deep learning models require large data for accurate results. The use of transfer learning approach can produce higher accuracies with much less resources.

3. PROPOSED METHOD

The proposed model is based on a transfer learning approach. The pre-processing is performed using bilateral filter. The AlexNet model and SVM are used for classification.

Bilateral Filter

Filters such as the Gaussian filter have the disadvantage that the contours are also blurred when trying to remove noise as much as possible. The input image is represented with f(i,j), the sequence of the processed image data is represented as g(i,j). Here, i and j are the pixel positions along the x and y axis.

Table 1. Network layers

Layers	Performance
Input:	3-channel color image normalized to 224 x 224 (color whitened)
Conv1:	Convolution layer – [11 x 11] kernel x 96 channels, stride = 4, padding = 2
Activation function:	ReLU + local response normalization
P1	Pooling layer – Overlapping maximum pooling, [3 × 3] kernel, stride = 2
Conv2	Convolution layer – [5 x 5] kernel x 256 channels, stride = 1, padding = 2
Activation function:	ReLU + local response normalization (Section 3.2.2)
P2	Pooling Layer – Overlapping Maximum Pooling, [3 × 3] Kernel, stride = 2
Conv3	Convolution Layer – [3 x 3] Kernel x 96 channels, stride = 1, padding = 1
Conv4	Convolution Layer – [3 x 3] Kernel x 96 channels, stride = 1, padding = 1
Conv5	Convolution Layer – [3 x 3] Kernel x 96 channels, stride = 1, padding = 1
P3	Pooling Layer – Overlapping Maximum Pooling, [3 × 3] Kernel, stride = 2 (Only when learning: Dropout)
FC6	Fully coupled layer – 9216 (= 256 × 6 × 6) vs 4096
Activation function	ReLU (Only when learning: Dropout)
FC7	Fully coupled layer – 4096 vs 4096 Activation function: ReLU
FC8	Fully coupled layer – 4096 vs 1000
Output	1000-dimensional vector probabilized by softmax function (output probability of each dimension).

$$g(i,j) = \frac{\sum_{n=-w}^{w}\sum_{m=-w}^{w} f(i+m,j+n)E_1E_2}{\sum_{n=-w}^{w}\sum_{m=-w}^{w} E_1E_2}, \quad (1)$$

where

$$E_1 = \exp\left(-\frac{m^2+n^2}{2\sigma_1^2}\right) \quad (2)$$

$$E_2 = \exp\left(-\frac{(f(i,j)-f(i+m,j+n))^2}{2\,\sigma_2^2}\right) \text{ and} \quad (3)$$

where

w is the kernel size,

σ_1 controls the Gaussian filter,

σ_2 controls the brightness difference.

The input OCT images contain noise which is removed using bilateral filter without causing structural damage. The edges in the image remain intact thereby the resulting denoised images have higher clarity.

AlexNet structure

AlexNet is a type of CNN for large-scale image class classification. When CNN model is used, features are automatically extracted as a multi-resolution hierarchical convolutional kernel layer group even for a dataset.

AlexNet is a class identification CNN with a structure in which each layer is connected in series in the following Table 1:

LeNet has 60K (60,000) learning parameters, while AlexNet has 60 million parameters. AlexNet was a fairly huge model not only in comparison with LeNet but also in comparison with machine learning up to that point. In AlexNet, in order to ensure image recognition accuracy in ImageNet, it is necessary to learn

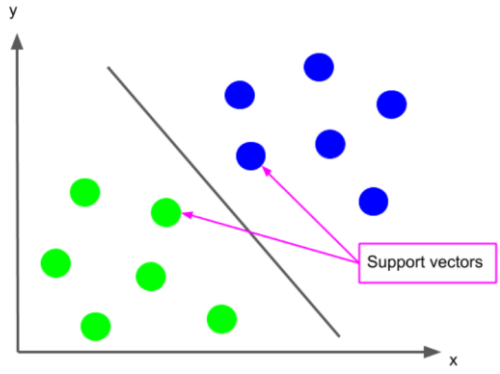


Figure 2. Support vectors.

well while suppressing overfitting because it is a network with a huge number of parameters compared to LeNet, which is the representative of CNN until then. rice field. Therefore, various ideas were newly introduced into CNN.

The proposed model uses Rectified Linear Unit (ReLU) as the activation function:

$$f(x) = \max(0, x). \tag{4}$$

The sigmoid function and tanh (also used in LeNet) were used as activation functions in traditional neural networks. However, in these saturated functions, the differential value becomes smaller as the number of insertions is increased, so the deeper the CNN layer, the more likely it is that the gradient disappears.

Therefore, AlexNet decided to use only the unsaturated type ReLu for the deactivation function. As a result, the vanish gradient was avoided and the learning speed was increased (by increasing the gradient value). By replacing it with ReLu, the accuracy can be maintained while achieving about 5 times faster speed in each layer than other activation functions such as sigmoid function and tanh function.

Since then, it has become standard for CNNs to use the ReLU-based activation function. As mentioned in the first section, how to suppress and reduce overfitting for AlexNet, which is an overparametrised model, was the key to scaling Deep CNN to support ImageNet learning.

Therefore, AlexNet has adopted the following three measures against overfitting:

1. Data expansion
2. Local Response Normalization
3. Drop Out

1. Data extensions

Data Augmentation, which is very important in the current CNN, is used in AlexNet for the purpose of reducing overfitting. ImageNet used for learning is a large-scale class identification data set for image identification problems. ImageNet image data is annotated with a class label (one of 1000 classes) corresponding to the image (256 x 256) in which only one central object is shown. In ImageNet, the number of images for each class is not so large (at most 500 images / class). Therefore, if possible, we would like to artificially increase the number of images to improve the generalization performance of learning. The easiest way to increase data and improve generalization performance is to use "data expansion" to edit the image slightly by arbitrary affine transformation or change the color value while keeping the label value fixed.

Therefore, AlexNet first expands the original image into the following two data, and as a result, learns AlexNet from the expanded ImageNet dataset whose total number of samples has increased:

2. Geometric extension

Create a flip image that is flipped horizontally from the image (256 x 256). Expansion of object position by translation: From the original image and flip image, a 224 x 224 area is randomly cropped, and these are used as data-expanded input images. These steps 1 and 2 increase the size of the dataset by 2048 times.

3. Local response normalization (LRN)

Another regularization is the neighbourhood with respect to the pixel values in the feature map. Same spatial location on the feature map of channels (5 channels before and after by default) Local response normalization (LRN), which normalizes from the luminance value group in, was proposed.

Since ReLU is an unbound function, it is a trick used to adjust the function form by performing local response normalization on the output after ReLU. In AlexNet, local response normalization was performed after each ReLU layer. The dropout, which was the standard method for deep neural network regularization at the time, was also used on AlexNet. AlexNet significantly reduced overfitting by using dropouts with a drop probability of 0.5 for neurons in the final two fully connected layers (FC6 and FC7).

The "average value pooling" used in previous CNNs such as LeNet acts like an average filter, so there was a problem that the image of the feature map after pooling was blurred. Therefore, AlexNet adopted Max Pooling as the pooling layer. AlexNet also used overlapped max pooling with a kernel size of 3 x 3 and a stride size of 2. In previous CNNs such as LeNet, tile-type operations such as kernel size 2 x 2 and stride size 2 were common so that pooling areas did not overlap.

3.1. Support vector machine (SVM)

SVM is basically available for both classification and regression tasks. It is a fast and reliable classification algorithm, and it can expect good performance even with a small amount of data. It comes with an example code.

For simplicity, consider a dataset of 2D data with only two features. Then, in SVM, the basic idea is to find the straight line that divides the target data set most appropriately according to the class as shown in Figure 2. Vector, which is also mentioned in the name of the Support vector algorithm, is the data point closest to the straight line that divides the dataset.

$$y = a\,x + b\,. \tag{5}$$

The classifier is represented as a straight line with slope a and intercept b. In SVMs, this Support Vector plays a major role in determining the straight lines that divide the dataset. More on this in the next section. Once this dividing straight line is determined, the class is determined based on which side of this straight line the data to be determined is located. In the graph above, if there is data to the left of the straight line, it is a yellow-green class, and if it is to the right, it is a blue class.

To create a classifier that classifies whether the data is in the blue or yellow-green class based on the two features x and y. First, consider the training data that has already been classified on the xy plane. As mentioned earlier, SVM finds a straight line that best separates this training data, taking into account the class of each data. Then, with this straight line as the boundary, all the data located on one side is classified as blue, and all the data located on the other side are classified as yellow-green. For this

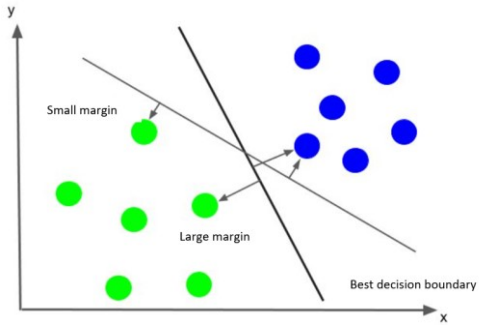


Figure 3. Support vectors.

reason, this straight line is sometimes called the decision boundary.

In SVM, in each class, the data point (Support Vector) closest to the straight line is considered, and the straight line is determined so that the distance (margin) between the data point and the straight line is as large as possible. The slope and intercept are together termed as

$$w = (a, b). \tag{6}$$

The minimization problem is defined as

$$\min_{w,b} \frac{1}{2} \|w\|^2 . \tag{7}$$

Figure 3 shows the determine the decision boundary in SVM.

4. EXPERIMENTAL RESULTS

This section presents the experimental results carried out eye dataset.

The number of images in the dataset are 3200. The total dataset is divided into 4 categories which are Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), DRUSEN, and NORMAL. Each category has 800 images from the total dataset. From these images, 500 images are used for training and 300 images are used for testing. Figure 4, Figure 5, Figure 6 and Figure 7 shows the input images.

True positive (TP) = count of correctly classified images
False positive (FP) = count of incorrectly identified images
True negative (TN) = The count of negative images correctly identified
False negative (FN) = the count of incorrectly classified negative cases
Accuracy: Accuracy is defined as the number of images correctly classified by the model. It is represented as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Sensitivity: Sensitivity is defined as the number of model's ability to identify the diseases in the images correctly. It is represented as:

$$Sensitivity = \frac{TP}{TP + FN}$$

Specificity: Sensitivity is defined as the number of model's ability to identify the healthy cases in the images correctly. It is represented as:

$$Specificity = \frac{TN}{TN + FP}$$

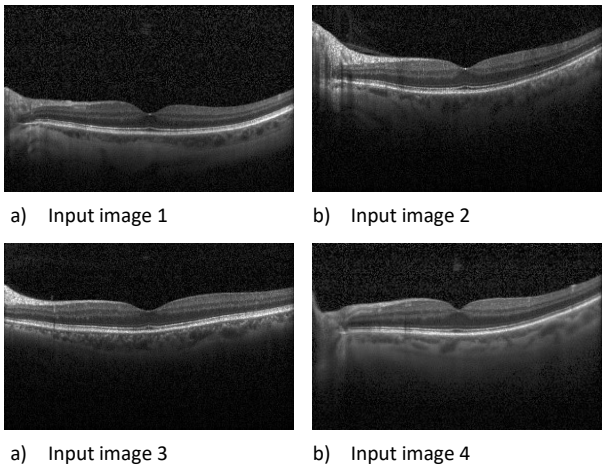


Figure 4. Normal Images.

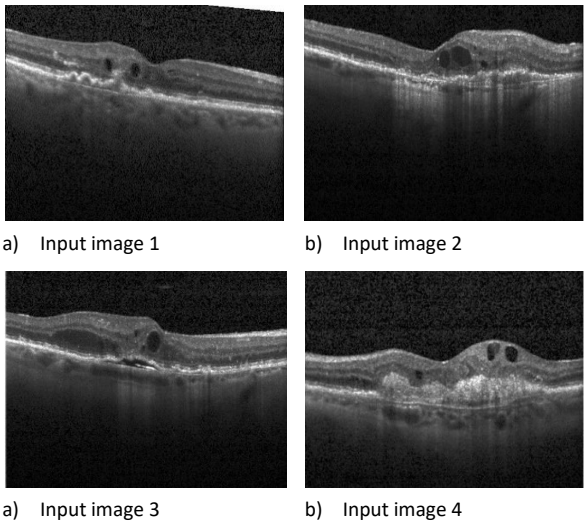


Figure 5. CNV Images

AlexNet is a model that has already been trained on a million of photos. The network has learnt rich visual features for a broad range of pictures. The network accepts an input images and outputs a label for each category in the picture, as well as probabilities for each type of object. Transfer learning is commonly used in deep learning applications. As a starting point, a pretrained network may be used to learn a new task. Transfer learning is typically faster and easier than building a network from scratch with randomly initialized weights. With fewer training photographs, it's easier to rapidly apply what you've learned to a new task. Using the representational capabilities of pretrained deep networks is as simple as extracting features.

Table 2 shows the confusion matrix of training dataset. Training completed using transfer learning and produces results with the accuracy of 100% and almost 0% loss. Table 2 represents a confusion matrix or contingency table, which compares the predicted class labels with the actual class labels in a classification task. Each column represents the correctly predicted classes, while each row represents the actual classes. The values in the cells indicate the total number of instances belonging to specific combinations of actual and predicted classes. In this particular scenario, the classification task involves four classes: CNV, DME, DRUSEN, and NORMAL. The "Actual Class" column

Table 2. Training confusion matrix

		Predicted Class			
		CNV	DME	DRUSEN	NORMAL
Actual Class	CNV	180	0	0	0
	DME	0	180	0	0
	DRUSEN	0	0	180	0
	NORMAL	0	0	0	180

Table 3. Test confusion matrix

		Predicted Class			
		CNV	DME	DRUSEN	NORMAL
Actual Class	CNV	119	0	1	0
	DME	1	118	1	0
	DRUSEN	0	2	118	0
	NORMAL	0	0	0	120

Table 4. Parameters comparison for train dataset.

Method	Precision	Sensitivity	Specificity	Accuracy
CNN	0.69	0.68	0.68	0.68
AlexNet	0.95	0.95	0.95	0.95
GoogleNet	0.97	0.96	0.96	0.97
Proposed	1.0	1.0	1.0	1.0

displays the actual labels of the instances, while the "Predicted Class" row displays the anticipated labels. For example, the cell at the intersection of the CNV row and CNV column contains the number 180. This indicates that there were 180 instances correctly classified as CNV (true positives). Similarly, the cell at the intersection of the DME row and DME column also contains 180, representing the accurate classification of 180 instances as DME (true positives). The diagonal cells from the top-left to the bottom-right of the table indicate correct predictions, where the predicted class matches the actual class. These values correspond to true positives for each class.

Table 3 shows the results of confusion matrix for test data set. In testing 480 images are categorized into 4 class with each class have 120 images. For test data set the results produces with the accuracy 97.5%.

Table 4 provides a comparison of multiple approaches, including CNN, AlexNet, GoogleNet, and the Proposed approach, based on various performance metrics evaluated on a train dataset. These metrics provide insights into the effectiveness and accuracy of each approach in classifying cases. Starting with the CNN technique, it achieves a precision of 0.69, sensitivity of 0.68, specificity of 0.68, and accuracy of 0.68. Based on these metrics, the CNN model demonstrates moderate levels of precision, sensitivity, specificity, and accuracy. It correctly identifies approximately 69 % of positive cases, 68 % of true positive instances, and 68 % of true negative instances. Overall, the CNN model achieves an accuracy of 68 %.

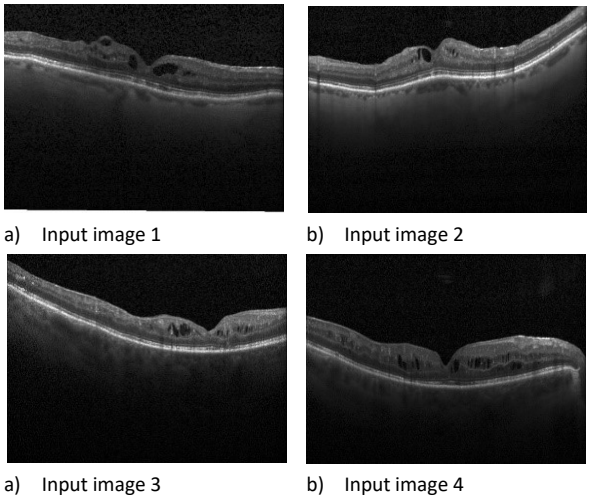
Moving to the AlexNet approach, it exhibits a precision of 0.95, sensitivity of 0.95, specificity of 0.95, and accuracy of 0.95. These values indicate that the AlexNet model outperforms the CNN approach significantly. It achieves high levels of precision, sensitivity, specificity, and accuracy, correctly identifying around 95 % of positive cases, 95 % of true positive instances, and 95 % of true negative instances. The overall accuracy of the AlexNet model is 95 %.

Next, the GoogleNet approach demonstrates an accuracy of 0.97, precision of 0.97, sensitivity of 0.96, and specificity of 0.96. The performance of the GoogleNet model is slightly better than that of the AlexNet model, with slightly higher accuracy and sensitivity. It accurately identifies approximately 97 % of positive cases, 96 % of true positive instances, and 96 % of true negative instances. The overall accuracy of the GoogleNet model is 97 %.

In conclusion, the Proposed approach exhibits exceptional performance across all metrics. It achieves a precision, sensitivity, specificity, and accuracy of 1.0, indicating perfect classification on the train dataset. The Proposed approach correctly identifies all positive cases, true positive instances, and true negative instances, resulting in a 100 % accuracy.

Table 5 shows the results of parameters comparison for train dataset with existing methods. Parameters used for comparison are precision, sensitivity, specificity and accuracy. Precision for CNN, Alexnet and googlenet are 0.36, 0.93 and 0.96. Sensitivity for CNN, Alexnet and Googlenet are 0.35, 0.92 and 0.95. Specificity for CNN, Alexnet and Googlenet are 0.33, 0.93 and 0.95. Accuracy for CNN, Alexnet and Googlenet are 0.35, 0.93 and 0.95. Precision, sensitivity, specificity, and accuracy for proposed method are 0.98, 0.97, 0.99 and 0.99.

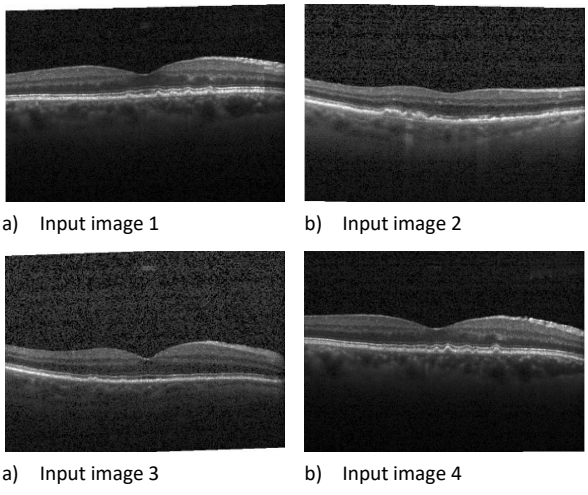
The proposed model is compared with existing models in Table 6 in terms of accuracy. The proposed model obtained better accuracy when compared to the existing models. Table 6 provides



a) Input image 1 b) Input image 2

a) Input image 3 b) Input image 4

Figure 6. DME images.



a) Input image 1 b) Input image 2

a) Input image 3 b) Input image 4

Figure 7. DRUSEN images.

Table 5. Parameters comparison for train dataset

Method	Precision	Sensitivity	Specificity	Accuracy
CNN	0.36	0.35	0.33	0.35
AlexNet	0.93	0.92	0.93	0.93
GoogleNet	0.96	0.95	0.95	0.95
Proposed	0.98	0.97	0.99	0.99

Table 6. Accuracy

Method	Accuracy
A. A. Saraiva [16]	93%
D. Wang [17]	94%
S. P.K. Karri [18]	95%
T. Tsuji et al., [19]	93%
Proposed model	97%

a comparison of different methodologies based on their accuracy in forecasting outcomes or categorizing occurrences. The "Accuracy" column in the table corresponds to each row, representing the precision achievable with a particular approach. Compared to the original approach developed by A. A. Saraiva [9], it achieves an accuracy of 93 %. This indicates that the model devised by A. A. Saraiva accurately predicts class labels with a 93 % accuracy rate.

The second approach by D. Wang [10], it attains an accuracy of 94%, a slight improvement. This suggests that the model developed by D. Wang performs marginally better than A. A. Saraiva's approach, accurately categorizing instances with a 94 % accuracy rate. The third approach, known as S. P. K. Karri [11], achieves a success rate of 95 %. Outperforming both A. A. Saraiva and D. Wang methods, this indicates that the model established by S. P. K. Karri achieves a higher accuracy rate of 95 % in its applications. The accuracy of the fourth approach, developed by T. Tsuji et al. [12], is 93 %, similar to A. A. Saraiva's method. This suggests that the model created by T. Tsuji and colleagues has a comparable accuracy rate of 93% in properly predicting class labels.

The last row of the Table 6 represents a proposed model with an accuracy of 97 %. This indicates that the model proposed by the authors of the table achieves the highest accuracy among all mentioned approaches, correctly identifying instances with a 97 % accuracy rate. In summary, the table presents a comparison of various methodologies in terms of their accuracy in forecasting outcomes or categorizing occurrences. The proposed model stands out with the highest accuracy rate of 97 %, surpassing the accuracy rates of other methods ranging from 93 % to 95 %.

5. CONCLUSIONS

OCT called optical coherence tomography (OCT) is based on a low coherence interferometer. Initially, this low coherence interferometer was applied to the evaluation of optical waveguides, etc., and its sensitivity was examined. In the field, it was applied to the measurement of axial length and angular film thickness. The proposed approach consisted of a transfer learning model using AlexNet for feature extraction, and SVM for the classification. The proposed model achieved an accuracy of 99 %, while other deep learning models, CNN, AlexNet, Googlenet are lower. Feature extraction is the easiest and fastest way to use the representational power of pretrained deep networks. When compared to existing deep learning algorithms such as CNN, AlexNet, and GoogleNet, the suggested technique achieved a classification accuracy of 99 percent. 0.98, 0.99, and 0.99 are the precision, sensitivity, and specificity, respectively.

REFERENCES

[1] T. Tsuji, Y. Hirose, K. Fujimori, T. Hirose, A. Oyama, Y. Saikawa, T. Mimura, Classification of optical coherence tomography images using a capsule network, BMC ophthalmology 20(1) (2020), pp. 1-9.
DOI: 10.1186/s12886-020-01382-4

[2] D. Wang, L. Wang, On OCT image classification via deep learning, IEEE Photonics Journal 11(5) (2019), pp. 1-14.
DOI: 10.1109/JPHOT.2019.2934484

[3] X. Wang, F. Tang, H. Chen, L. Luyang, Z. Tang, A. Ran, C. Cheung, H. Pheng-Ann, UD-MIL: uncertainty- driven deep multiple instances learning for OCT image classification, IEEE journal of biomedical and health informatics 24(12) (2020), pp. 3431-3442.
DOI: 10.1109/JBHI.2020.2983730

[4] Y. Rong, D. Xiang, W. Zhu, K. Yu, F. Shi, F. Zhun, and C. Xinjian, Surrogate-assisted retinal OCT image classification based on convolutional neural networks, IEEE journal of biomedical and health informatics 23(1) (2018), pp. 253-263.
DOI: 10.1109/JBHI.2018.2795545

[5] M. Awais, H. Müller, TB. Tang, F. Meriaudeau, Classification of sd-oct images using a deep learning approach, IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuching, Malaysia, 12-14 September 2017, pp. 489-492.
DOI: 10.1109/ICSIPA.2017.8120661

[6] S. Naz, A. Ahmed, M. U. Akram, S. A. Khan, Automated segmentation of RPE layer for the detection of age macular degeneration using OCT images, IEEE Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), Oulu, Finland, 12-15 December 2016, pp. 1-4, 2016.
DOI: 10.1109/IPTA.2016.7821033

[7] S. Aumann, S. Donner, J Fischer, F. Müller, Optical coherence tomography (OCT): principle and technical realization, High Resolution Imaging in Microscopy and Ophthalmology, Springer, Cham, pp. 59-85, 2019.
DOI: 10.1007/978-3-030-16638-0_3

[8] K. T. Islam, S. Wijewickrema, S. Leary, Identifying diabetic retinopathy from oct images using deep transfer learning with artificial neural networks, 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba, Spain, 05-07 June 2019, pp. 281-286.
DOI: 10.1109/CBMS.2019.00066

[9] A. A. Saraiva, D. B. S. Santos, P. Pimentel, J. V. M. Sousa, M. N. Ferreira, J. D. E. S. Batista, S. Soares, A. Valente, Classification of optical coherence tomography using convolutional neural networks, Proceeding of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2020), pp. 168-175, 2020.
DOI: 10.5220/0009091001680175

[10] D. Wang, L. Wang, On OCT image classification via deep learning, IEEE Photonics Journal 1(5) (2019), pp 1-14.
DOI: 10.1109/JPHOT.2019.2934484

[11] S. P. K. Karri, D. Chakraborty, J. Chatterjee, Transfer learning-based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration, Biomedical optics express 8(2) (2017), pp 579-592.
DOI: 10.1364/BOE.8.000579

[12] T. Tsuji, Y. Hirose, K. Fujimori, T. Hirose, O. Asuka, Y. Saikawa, T. Mimura, Classification of optical coherence tomography images using a capsule network, BMC ophthalmology 20(1) (2020), pp 1-9.
DOI: 10.1186/s12886-020-01382-4