JOURNAL OF OPTOELECTRONICS LASER ISSN: 1005-0086

Volume 41 Issue 8, 2022

Retinal OCT images classifications using SVM and DenseNet

Dr. Muddamalla Naresh*[1]

[1]Iraq Assistant Professor, Department of ECE, Matrusri Engineering College, Hyderabad, Telangana state, India.

ABSTRACT

Given Deep Learning's consistent success, it is now widely used in the diagnosis of diseases via biomedical image classification tasks. This is because manual classification is time-consuming and prone to error, and Deep Learning eliminates this problem. In that case, artificial intelligence has the potential to make a significant contribution to the solution of this problem. OCT, the most extensively used imaging tool for capturing retinal abnormalities such age-related macular degeneration, diabetic retinopathy, and vitreous traction, provides a cross-sectional picture of the human eye. Early detection of these diseases using an OCT image can save a person's life by preventing them from developing various eye-related problems. Manual image classification, on the other hand, is prone to error because it necessitates a large number of image processing tasks as well as time. As a result, in this paper, it is proposed to develop a deep learning-based image classification method that can accurately identify diseases from optical coherence tomography (OCT) images. This is why we have employed pre-trained Convolutional neural networks (CNNs), namely DenseNet and Bayesian SVM, as a feature extractor from the OCT image, rather than training our own. As a result, we used a publicly available dataset of OCT images that contains images of four different types of diseases (namely CNV, DME, Drusen, and Normal) and contains approximately 80,000 images. The SVM classifier was employed in the classification process. Our experimental results revealed that the proposed DenseNet model outperformed any other handcrafted feature extraction model in terms of classification accuracy, as demonstrated in the literature.

Keywords: CNN, OCT, DenseNet, SVM, Drusen, deep learning

I. INTRODUCTION

An imaging technology known as optical coherence tomography (OCT) employs coherent light to acquire high-resolution images of biological tissues in real-time. Optical coherence tomography (OCT)[1] is frequently used to get high-resolution pictures of the retina (OCT). The retina of the eye operates significantly better than a film in a camera. Using optical coherence tomography (OCT) pictures [2], several retinal illnesses can be diagnosed.

Non-invasive, high-resolution imaging of highly scattering tissues is provided by optical coherence tomography (OCT). When doing diagnostic imaging on the retina and the front segment of the eye, ophthalmologists employ this technique to get the most accurate readings. If AMD is diagnosed and treated early, vision loss can be avoided. In ophthalmology, optical OCT is the most often used imaging modality to detect these conditions. Screening for these disorders can be done using OCT pictures; this can lead to an increased workload for eye doctors, who must interpret the images. The strain on ophthalmologists' time has been reduced with the development of an automated diagnostic screening method.

It is possible to aid doctors in their diagnosis and treatment of patients by automating the detection and diagnosis of retinopathy. Patients benefit from cheaper treatment costs and less medical resource consumption when diseases are detected and treated earlier. The use of existing image analysis and processing technology to automatically analyze medical images has grown in popularity in recent years. It is difficult to analyze retinal OCT images because of two difficulties. OCT retinal image datasets for study are smaller than those utilized in typical image analysis domains, such as MRI, and this is a significant difference. The quality and size of retinal OCT pictures might vary greatly depending on the tissue in which they are obtained (as well as significant differences in the quality and size of retinal OCT images encountered by physicians practicing clinical medicine).

Classifying medical photos using deep learning [4-5] in the medical industry OCT imaging will be used for categorization, it has been agreed.

1. Age-related diseases account for four out of the top ten causes of blindness, with three of them affecting the retina.
2. An important part of clinical diagnosis is the examination of specific retinal layer features as structural changes have been found to be connected with the most common eye illnesses' manifestations.

 ISSN: 1005-0086

Volume 41 Issue 8, 2022

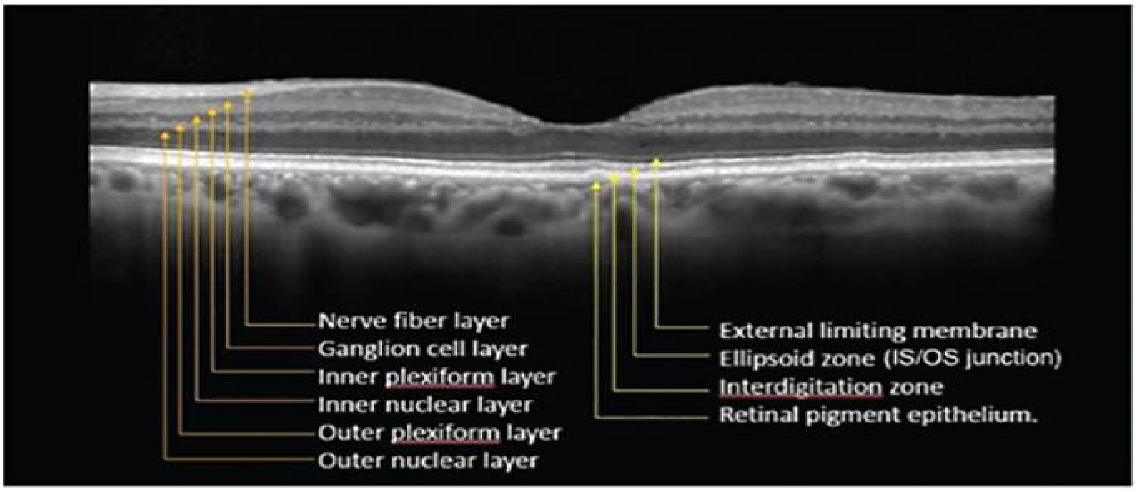

Figure 1. optical coherence tomography (OCT) images of a normal eye's macula.

In some algorithms Adaptive Genetic Algorithm (AGA)[6] is proposed to identify the features that can be better represent the image are suggested. The primary goal of this project is to identify the underlying cause of eye disease. In this study, an attempt is made to improve classification accuracy by using a model based on a DenseNet and a Bayesian SVM[7-10] on an OCT dataset that was used for research.

II. METHODOLOGY

Datasets

Optical coherence tomography (OCT) of the retina is a noninvasive imaging technique that can be used to obtain high-resolution cross-sections of the retina in live patients. About 30 million people have OCT scans each year, and the time and resources required to analyze and interpret these pictures are considerable.

Each of the image categories has its own subfolder in the dataset's three main folders (train, test, and validation)

Sub retinal fluid and the white membranes of neovascularization are clearly visible in this picture, which shows choroidal neovascularization (arrows). (Clockwise from top left) The presence of both retinal thickening and intraretinal fluid (RTEF) is a hallmark of diabetic macular edema, or DME (arrows). (Middle-Right Multiple drusen are seen in the early stages of AMD (arrowheads). (To the left) One may see a healthy retina without any fluid or edema in the area surrounding the fovea.

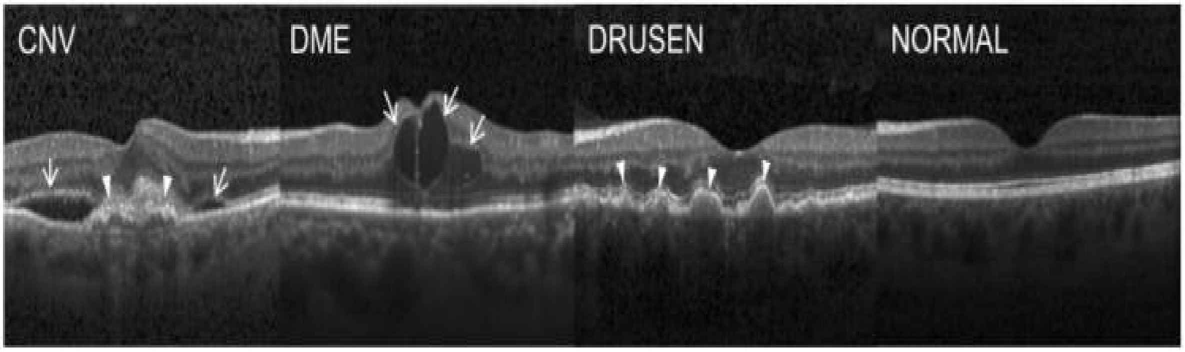

Fig 2. Dataset Sample

1. Choroidal neovascularization (CNV):

Choroidal neovascularization refers to the development of new blood vessels in the choroid layer of the eye. Myopia, malignant myopia, and other age-related changes might worsen choroidal neovascularization, which in turn leads to neovascular degenerative maculopathy (also known as "wet" macular degeneration").

 DOI: 10050086.2022.08.68

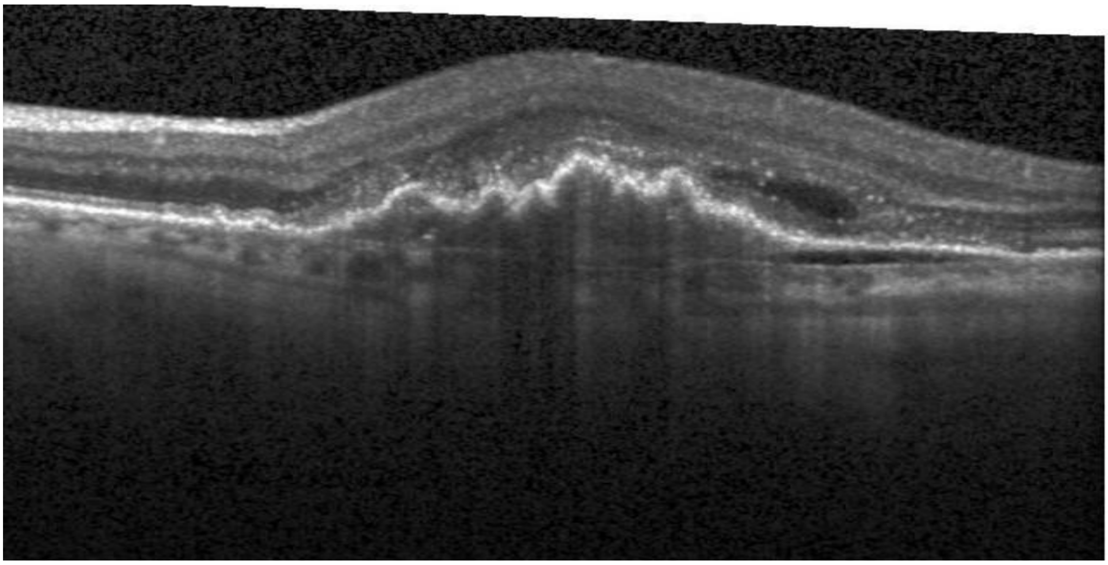

Fig3. Choroidal neovascularization

2. Diabetic Macular Edema (DME):

DME is a consequence of diabetes that occurs when fluid accumulates in the macula, resulting in damage to the fovea. Diabetic macular edema is another name for it. With its location at the back of the eye and sharpness of vision, the macula is the central portion of the retina, which is the sharpest part to see with. When DME is present, vision loss can progress over a period of months, making it impossible to concentrate clearly.

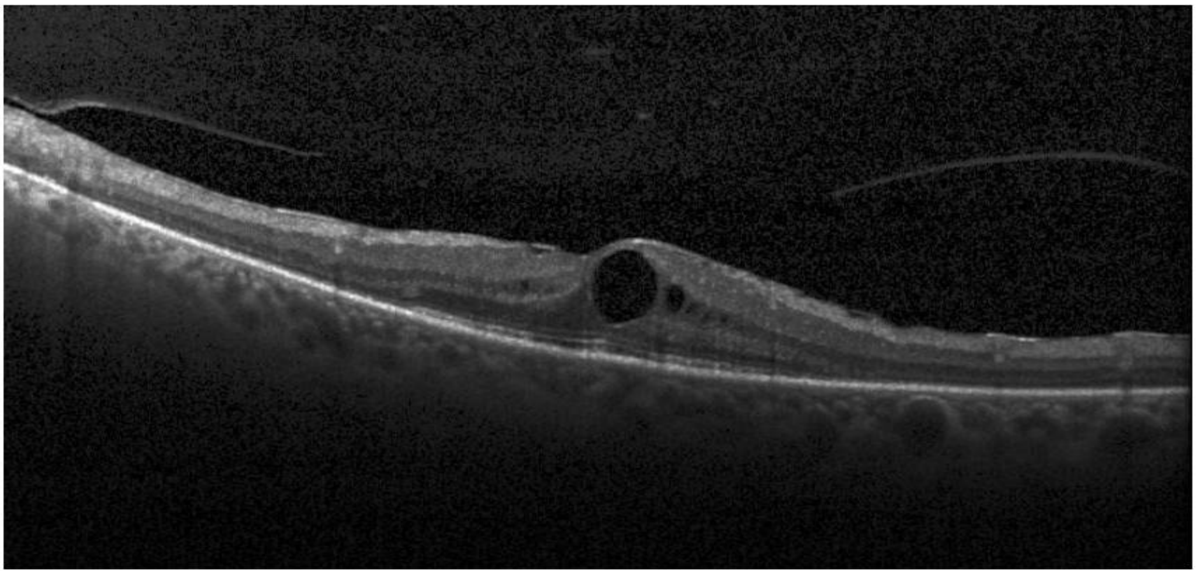

Fig 4. Diabetic Macular Edema (DME)

3. Druse

Drusen are deposits that are difficult to notice and are yellow in color. They are located under the retina. Drusen are composed of lipids, which are a type of fatty protein. Drusen are not believed to be a cause of age-related macular degeneration (AMD).

The presence of drusen, on the other hand, is associated with an increased likelihood of developing AMD. In most cases, drusen can be seen in both eyes due to the fact that they are made up of protein and calcium salts, both of which can be found in both eyes.

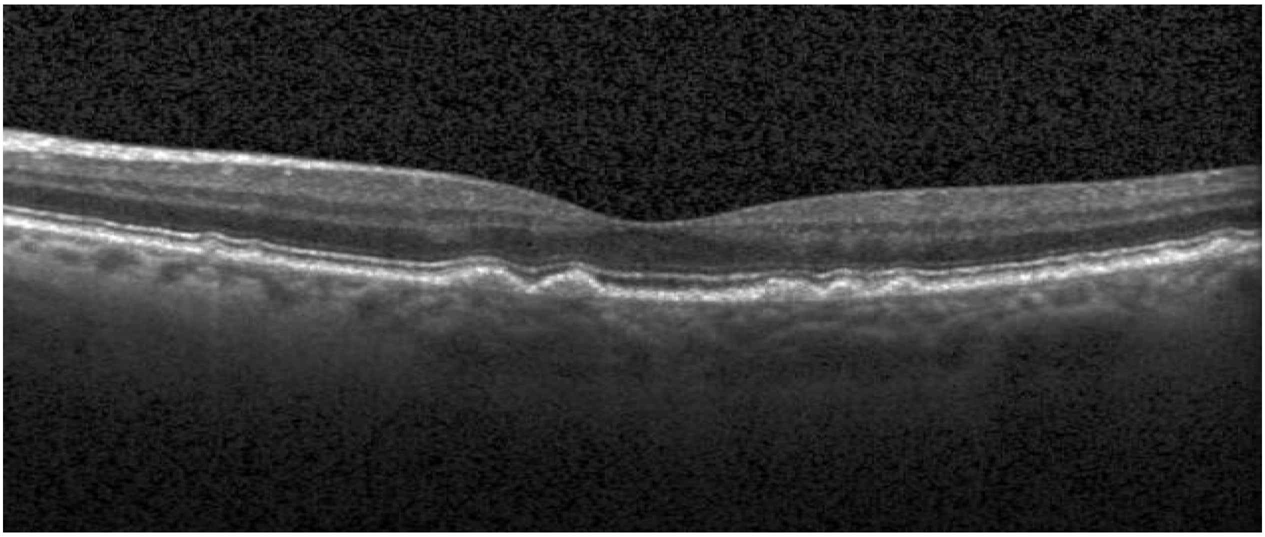

Fig 5. Drusen

4. Normal:

In normal vision, the retina receives direct light rather than the light that is reflected from in front of or behind it. A person with normal vision is able to distinguish between objects that are close and far away.

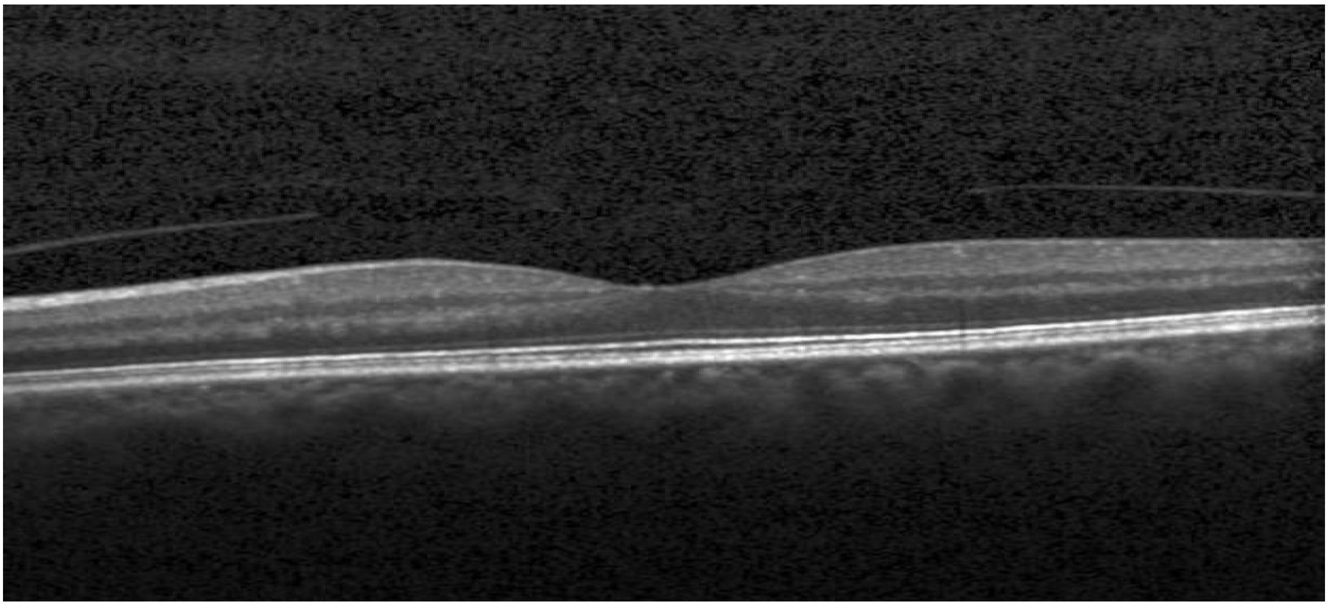

Fig 5. Normal

B. Data Preparation

The pre-processing or cleansing of data is an essential step in the job of a Machine Learning Engineer, and the vast majority of Machine Learning Engineers put in a significant amount of effort in this portion of the process before beginning to construct a model from scratch. A few examples of data pre-processing procedures include the detection of outliers, the treatment of missing values, and the elimination of data that is either undesirable or noisy.

Images at the lowest level of abstraction are referred to as image pre-processing, which is the same as image processing. According to entropy as an information metric, these processes do not increase image information content, but rather decrease it. In pre-processing, the goal is to make the image data better by suppressing unwanted distortions and enhancing some visual properties that are important for the task of subsequent processing and analysis after it has been captured. Pre-processing procedures are classified into two categories, which are listed below.

JOURNAL OF OPTOELECTRONICS LASER ISSN: 1005-0086

Volume 41 Issue 8, 2022

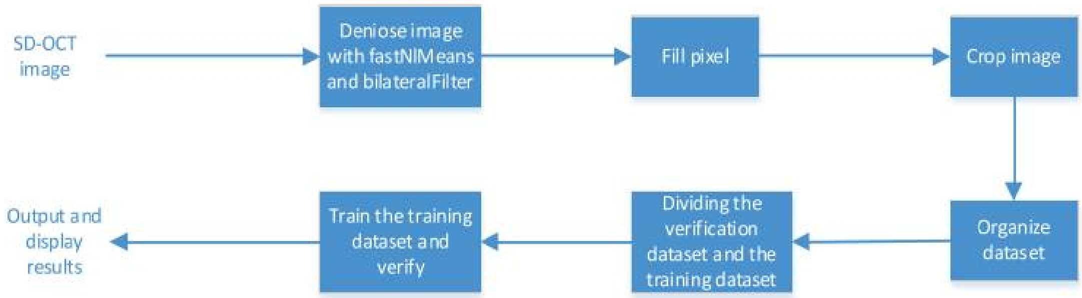

Fig 6. Data processing

Before the original OCT images from 596 by 1264 pixels can be used as input for a CNN[3], it is required to first eliminate all of the extraneous components that were included in those images. This will result in RGB images that have dimensions of 380 by 764 pixels.

After that, the images that had been cropped to a size of 380 x 764 pixels were then down sampled to RGB images that had a size of 224 × 224 pixels before being fed into the CNN algorithm. Images with a resolution of 224 × 224 pixels in RGB are a common image standard for use in classification models such as Logistic Regression and Resnet-18. These models also make use of other image standards. In order to prevent the model from being overfit by using too many of the input photos, a data augmentation technique is required.

The data augmentation procedure included a number of random manipulations of the image, including random horizontal flips, random brightness adjustments between 0.7 and 1.3, and random rotations of the image up to 15 degrees. During the training phase, this particular data augmentation approach was the only one that was used.

C. Modeling

DenseNet[11-15] and Bayesian support vector machine approaches were utilised during the development of the model. The final stage involved validating the proposed model by applying it to the test dataset.

I. DenseNet

DenseNet is a framework that aims to increase the depth of deep learning networks while also making it more cost-effective to train them[7] . Shorter connections between layers of the network are used to accomplish this. For example, the first layer of a DenseNet convolutional neural network is connected to every layer deeper in the network, and so on, and so forth; the second layer, and so forth. So that all of the information can flow freely between different network layers, this is how it's being done. Feature maps from each layer are transmitted to the following levels in the hierarchy, taking into account inputs from the preceding layers. That way, the feed-forward nature of the system can be maintained.

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully-connected, softmax			

Fig 7a.various architectures of DenseNets

ISSN: 1005-0086

Volume 41 Issue 8, 2022

Resnets do not combine features by summarizing them; rather, it concatenates the features to generate a new feature. 'ith' layer inputs and feature maps come from all previous convolution blocks. To each of the successive 'I-i' layers, these feature maps are sent along with them. Instead of just 'I' connections, the network now has '(I(I+1))/2' connections as in earlier deep learning models. 'I' connections are not sufficient.

$$\mathbf{x}_\ell = H_\ell(\mathbf{x}_{\ell-1}) + \mathbf{x}_{\ell-1}.$$

There is no need to train features that are irrelevant to the task at hand, so it uses far fewer parameters than standard convolution neural networks.

$$\mathbf{x}_\ell = H_\ell([\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}]),$$

Another key component of DenseNet is its two convolutional and pooling layers. There are two kinds of layers: Dense Blocks and Transitions.

As a foundation for the network, DenseNet uses a simple convolution and pooling layer. This is followed by a dense block, followed by a transition layer, followed by another dense block, and finally a dense block, followed by a classification layer.

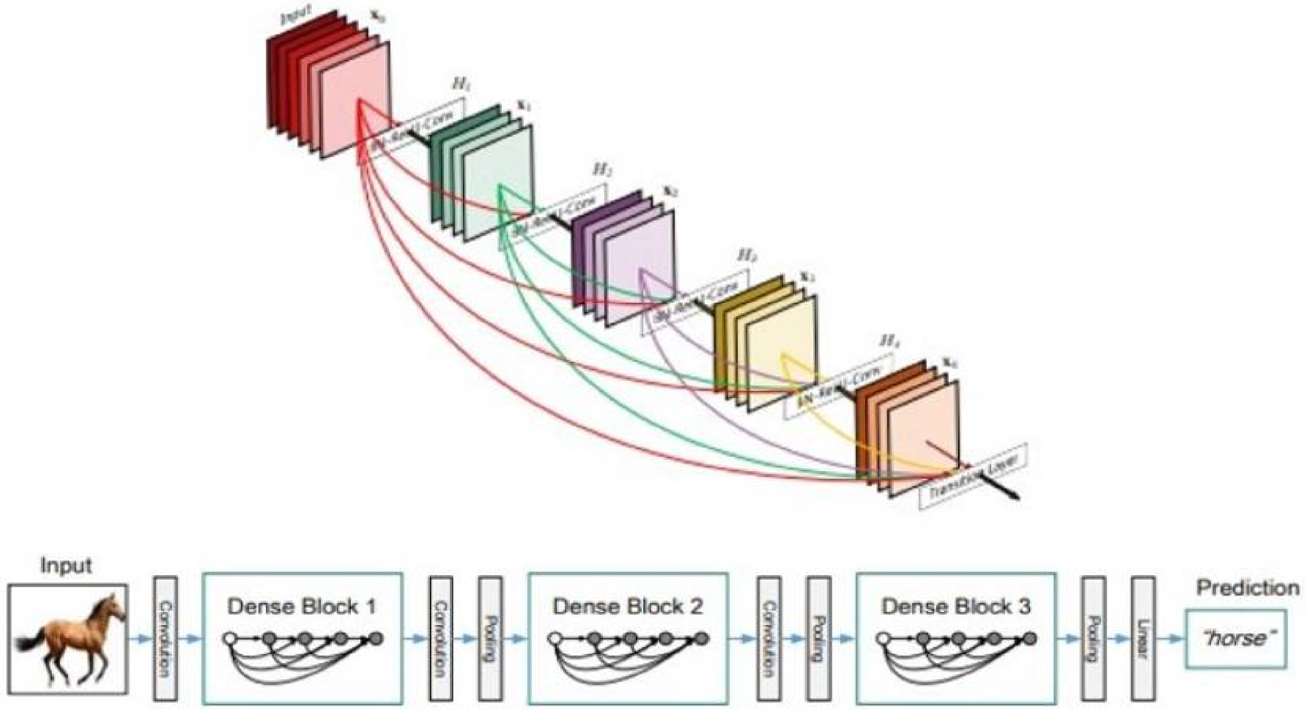

Fig 7b. DenseNet's various blocks and layers

Convolutional neural networks such as DenseNet have multiple layers, and each layer is linked to all of the layers below it. The first layer connects every one of the layers below it, and so on. DenseNet is a convolutional neural network.

Multiple convolutions are used to extract high-level features from a picture in the Standard Convolution layer.

As each layer is added to the DenseNet architecture, it receives new inputs, including feature maps, from the layers that came before it. As you can see, we're making use of concatenation here. As the layers build upon one another, they share "collective wisdom."

Network channels can be reduced because each layer receives feature maps from previous layers. The number of channels introduced to each layer determines the growth rate k. The following are the steps taken by each composition layer: With the output feature maps of the K channels in 3-3 Conv, for example, you can use the following example

DOI: 10050086.2022.08.68

ISSN: 1005-0086

Volume 41 Issue 8, 2022

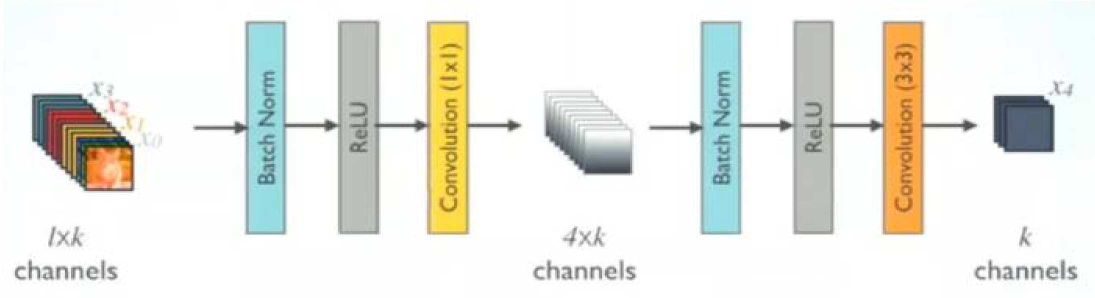

Fig 8: Steps Performed in DenseNet composition layer

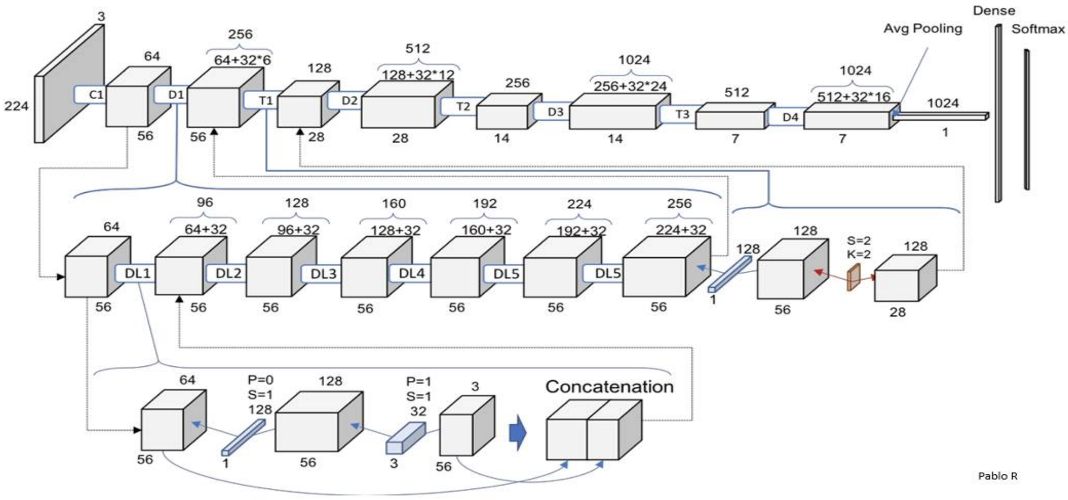

Fig 8a. Full architecture of DenseNet

As the number of Dense Blocks changes, so does the size of the channel. Helps generalize the lth layer's growth rate (k). It determines how much data is added to each layer.

$$k^{[l]} = \left(k^{[0]} + k(l-1) \right)$$

II. Bayesian SVM

Classification and regression problems can be solved using the "Support Vector Machine" (SVM) algorithm. The "support vectors" theory is the foundation of this strategy. The most common application of this method that may be found in practice is categorization. Each piece of data is represented as a point in an n-dimensional space, and the value of each feature corresponds to the value of a specific coordinate in that space (where n is the number of features you have). SVM stands for support vector machine. After then, classification is accomplished by locating the hyper-plane that provides the clearest demarcation between the two classes (look at the below snapshot).

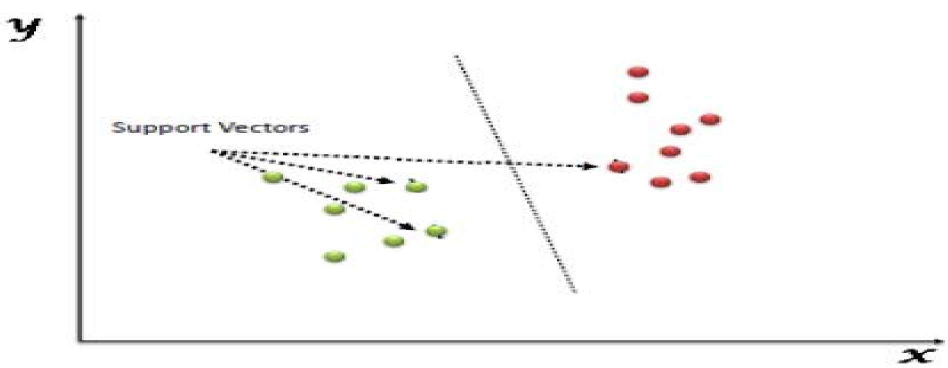

Fig 9. Vector Support Machine

JOURNAL OF OPTOELECTRONICS LASER
ISSN: 1005-0086

Volume 41 Issue 8, 2022

According to the most recent studies, SVM methods have recently been improved in order to accomplish variable selection through the use of frequents regularization or Bayesian shrinkage. It is possible to explicitly measure uncertainty by employing the Bayesian framework, which offers a probabilistic interpretation for the support vector machine.

D. Validation Method

It is possible to utilize a confusion matrix to keep track of the number of times a predicted class was placed in the incorrect classification bin after being compared to the actual classes. The number of valid classifications for each class is represented by the number of diagonal elements, which is 3, 1, and 3 for classes 0, 1, and 2 accordingly. The misclassifications are caused by the off-diagonal elements; for instance, two of the class 2 components were misclassified as 0, whereas none of the class 0 elements were misclassified as 2. When added together, the "All" subtotals in both y true and y pred produce a number of categories that correspond to each class in both variables.

IV. RESULTS

The accuracy, recall, precision, and AUC of each DenseNet neural network were used as evaluation criteria in order to display the confusion matrix of each DenseNet neural network.

Following the testing, DenseNet receives the highest overall rating, with a score of90%.

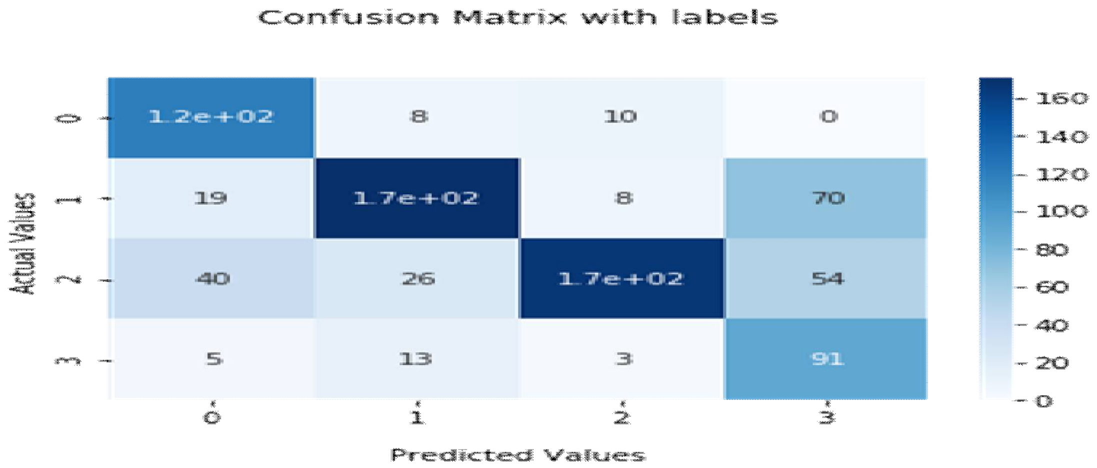

Fig 10a. Confusion matrix of DenseNet

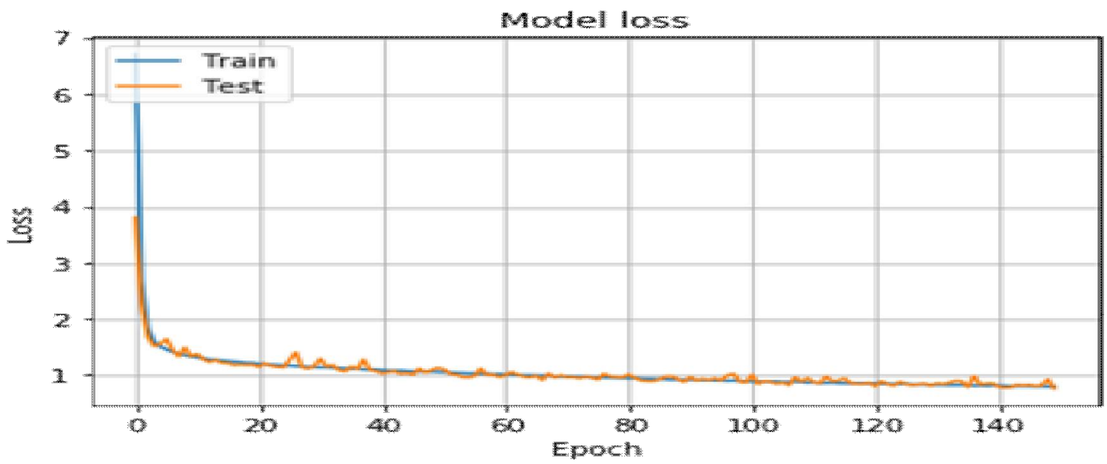

Fig 10b. Training and model loss during training

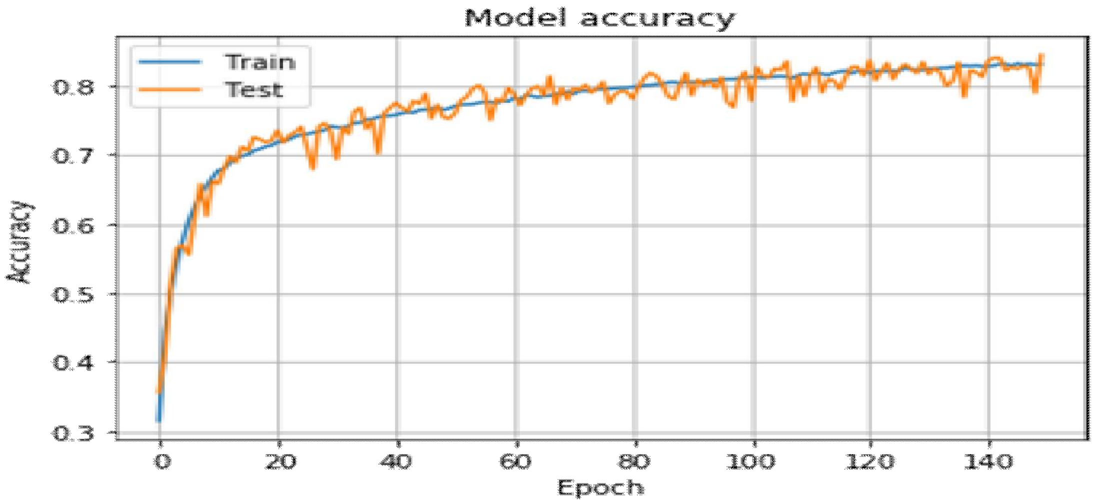

Fig 10c. Training and model accuracy

In the course of this research, a Bayesian support vector machine (SVM) model that possessed the necessary features for feature reusing was constructed and evaluated. In addition, the use of the SVM in the diagnosis of AMD and DME by making use of datasets from retinal OCT was examined, tested, and evaluated. In addition to the area under the curve (AUC), the ROC curve and confusion matrix were shown for each neural network. Accuracy, recall, and precision were the evaluation criteria that were applied.

Following the testing, Bayesian SVM is found to have the highest accuracy, with a score of 70 percent.

```
accuracy_score(pred, ytest)
0.7077114427860697
```

Fig 10d. Bayesian SVM model accuracy

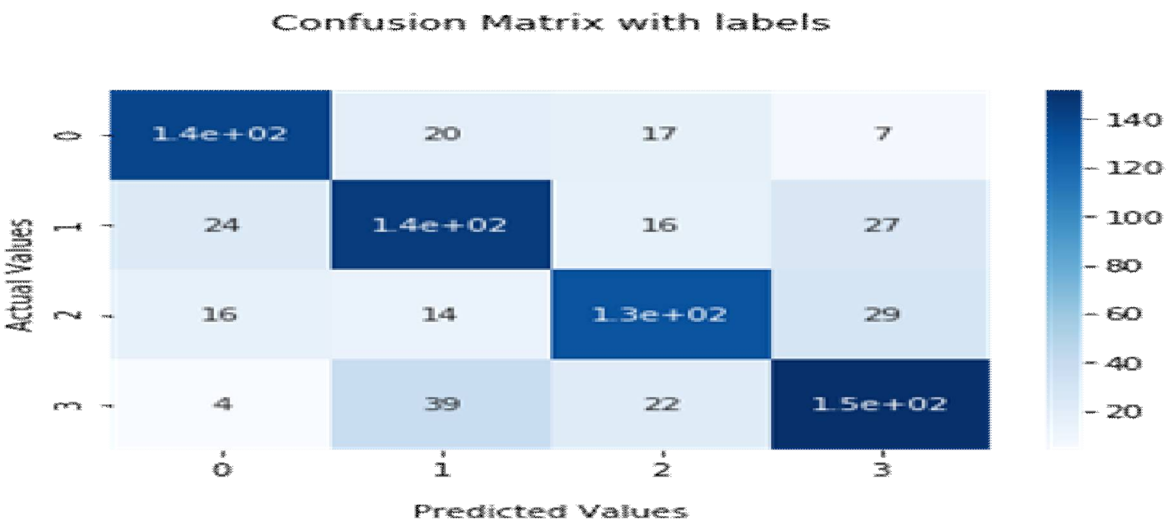

Fig 10e. Bayesian SVM model Confusion Matrix

As criteria for evaluating the neural network's confusion matrix, accuracy, recall, precision, and AUC were used. DenseNet is the top performer on the dataset, with a three-type AUC of 0.83, a three-type recall rate of 0.85, an average precision of 0.85 for three types, and an accuracy rate of 0.83.

V. CONCLUSION AND FUTURE WORK

Neural network models can benefit from the advantages of feature reuse in terms of improving their performance and adaptability to a wide range of datasets, according to experiments. Neural network models can benefit from these characteristics, as evidenced by their improved performance.

DenseNet training takes a significant amount of time, but it yields greater results. The Bayesian support vector machine (SVM) model's overall performance, as well as its adaptability to changes between different categories of OCT, can be improved by using the same features over and over again.

When dealing with medical datasets that have a limited amount of data, it is feasible to take into consideration various neural network models that have the useful attribute of reuse. Additionally, it is possible to take into consideration data amplification by means of data augmentation and generative adversarial networks (GAN).

REFERENCES

[1] D. Wang and L. Wang, "On OCT Image Classification via Deep Learning," in *IEEE Photonics Journal*, vol. 11, no. 5, pp. 1-14, Oct. 2019, Art no. 3900714, doi: 10.1109/JPHOT.2019.2934484.

[2] C. Kolluru, J. Lee, Y. Gharaibeh, H. G. Bezerra and D. L. Wilson, "Learning With Fewer Images via Image Clustering: Application to Intravascular OCT Image Segmentation," in *IEEE Access*, vol. 9, pp. 37273-37280, 2021, doi: 10.1109/ACCESS.2021.3058890.

[3] Tallapragada, VV Satyanarayana, et al. "Improved Atrial Fibrillation Detection using CNN-LSTM." 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2022.

[4] Tallapragada, V. V., et al. "Greek Handwritten Character Recognition Using Inception V3." Smart Systems: Innovations in Computing. Springer, Singapore, 2022. 247-257.

[5] Tallapragada, VV Satyanarayana. "Deep Learning and Its Applications in Biomedical Image Processing." Handbook of Deep Learning in Biomedical Engineering and Health Informatics. Apple Academic Press, 2021. 281-302.

[6] Tallapragada, VV Satyanarayana, V. Sireesha, and GV Pradeep Kumar. "Telugu Character Recognition Using Adaptive Fuzzy Membership Functions With Adaptive Genetic Algorithm Based Techniques." Vol.8, pp. 3092-3097, 2019.

[7] J. Li, Y. Zhao, X. Gao and W. Zhao, "Feature Extraction and SVM Parameters Optimization Based on ICQPSO," 2020 International Conference on Computer Network, Electronic and Automation (ICCNEA), 2020, pp. 329-332, doi: 10.1109/ICCNEA50255.2020.00074.

[8] Q. Li, H. Ma, Z. Zhang and C. Feng, "An Effective Nuclear Extraction Mask Method for SVM Classification," 2020 5th International Conference on Computer and Communication Systems (ICCCS), 2020, pp. 104-108, doi: 10.1109/ICCCS49078.2020.9118498.

[9] Tallapragada, VV Satyanarayana, et al. "Image Denoising Using Low Rank Matrix Approximation in Singular Value Decomposition." REVISTA GEINTEC-GESTAO INOVACAO E TECNOLOGIAS 11.2 (2021): 1430-1446.

[10] Tallapragada, V. V., et al. "Significance of Internet of Things (IoT) in Health Care with Trending Smart Application." Smart Systems: Innovations in Computing. Springer, Singapore, 2022. 237-245.

[11] Q. Wang, B. Yang, W. Liu and G. Chen, "X-ray Images Detection of COVID-19 Based on Deepwise Separable DenseNet," 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP), 2021, pp. 294-298, doi: 10.1109/ICSIP52628.2021.9688876.

[12] M. T. Rahman and A. Dola, "Automated Grading of Diabetic Retinopathy using DenseNet-169 Architecture," 2021 5th International Conference on Electrical Information and Communication Technology (EICT), 2021, pp. 1-4, doi: 10.1109/EICT54103.2021.9733431.

[13] V. Lalitha, G. Raghul and A. R. Premkumar, "Leukocyte Counting and Reporting Using Densenet - Deep Learning," 2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), 2020, pp. 1-6, doi: 10.1109/ICPECTS49113.2020.9337022.

[14] X. Gao, T. Chen, R. Niu and A. Plaza, "Recognition and Mapping of Landslide Using a Fully Convolutional DenseNet and Influencing Factors," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7881-7894, 2021, doi: 10.1109/JSTARS.2021.3101203.

[15] Y. Zhang and C. Wang, "SIIM-ISIC Melanoma Classification With DenseNet," *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, 2021, pp. 14-17, doi: 10.1109/ICBAIE52039.2021.9389983.