Texila International Journal of Public Health
ISSN: 2520-3134
DOI: 10.21522/TIJPH.2013.13.02.Art056

A Stacking Ensemble Federated Deep Learning Model with Optimization for the Efficient Ocular Pathology Detection

S. Geethamani*, L. Sankari
Department of Computer Science, Sri Ramakrishna College of Arts & Science for Women, Coimbatore

Abstract

One major challenge in healthcare is utilizing Fundus Images (FI) to diagnose ocular pathology (OP). An ocular disorder disrupts the eye's regular functioning or adversely impacts the eye's visual acuity. Almost everyone experiences eye-sight issues throughout their lives, ranging from minor problems that can be managed at home to more severe conditions requiring specialized medical care. While certain kids require specialized care, others are minors who do not show up to support requests or who can be handled at home with ease. Ocular pathology detection approaches depend on Stacking Ensemble Federated (DL)Deep Learning (SEFDL), which was suggested in this work. First, an adaptive weight (AW)-based median filter (MF) is applied to image resizing and removing noise. Then, the data augmentation, coupled with the Synthetic Minority Over-sampling Technique (SMOTE), z, is employed to address data imbalance, a common issue in medical datasets. Finally, SEFDL is proposed for disease detection (DD). Adaptive TSO (Tuna-Swarm Optimization) Technique adjusted hyperparameters (HP) for 4 pre-trained models: CNN, VGG16, Inceptionv2, and ResNet50. DL models trained centrally have been compared with the enhanced algorithms in a federated framework. The proposed SEFDL model demonstrates superior accuracy and robustness when benchmarked against existing methods, highlighting its potential as a reliable diagnostic tool. Finally, the result must be compared with existing approaches to improve ocular pathology detection while addressing data privacy concerns in healthcare applications.

Keywords: *Cropping, Data Augmentation (DA), Handling Data Imbalance, Ocular pathology, Preprocessing, Stacking Ensemble Federated Deep Learning (SEFDL).*

Introduction

Globally, the disease rate for a number of ocular pathologies is increasing. The ageing population and imbalanced diet contribute to their rising prevalence [1]. Diabetic Retinopathy (DR), glaucoma, Age-related Macular Degeneration (AMD), and other conditions are among the many that cause blindness. Regardless of the OP, certain late stages represent permanent chronic disease. Therefore, ophthalmologists encourage patients to undergo screenings regularly to identify potential diseases. The automatic Detection of ocular pathologies is of great interest to multiple research initiatives in this setting. Higher performance detection was possible with Machine Learning (ML) techniques, particularly with DL techniques. Several retinal components, including the macula, optic nerve head, and blood vessel tree, have affine forms and morphologies, as demonstrated by the FIs [2]. Retinal components may regenerate, or lesions may occur due to ocular pathologies.

The size, form, contrast, and other characteristics of these lesions differ. They also frequently have characteristics in common with other retinal regions or clinical abnormalities.

Received: 30.01.2025 ***Accepted: 25.02.2025*** ***Published on: 30.06.2025***
****Corresponding Author: geethagokul2013@gmail.com***

The surgical specialty of ophthalmology focuses on diagnosing and treating eye diseases [3]. Ophthalmology has a long history of leading medical research and improving eye care. This subject entails thoroughly examining and analyzing a range of illnesses, utilizing pertinent technologies to ensure successful therapy. Anatomy [4], physiology, and eye disorders are the main areas of study for the medical specialty of ophthalmology. Medical experts in this profession who specialize in surgery and medicine are called ophthalmologists. They can identify, treat, and prevent disorders affecting the vision (Figure 1) and visual system. Retina and optic nerve prompt the Detection of early symptoms for disorders like glaucoma or cataracts [5]. Eye disorders, infections, and other conditions are skillfully diagnosed and treated by ophthalmologists. There is a noticeable lack of thorough studies examining AI's potential uses in this subject, despite its potential. A scientific examination is necessary to understand the applicability, restrictions, and DL, AI, and ML's potential in ophthalmology.

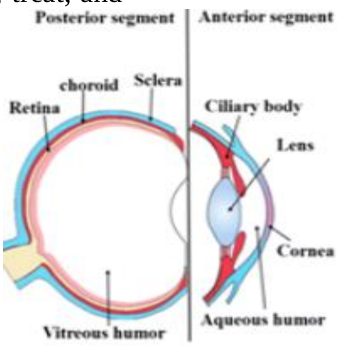


Figure 1. Cross-Sectional View of the Eye

In human history, AI represents the 4th industrial revolution. In recent years, DL, a class of State of Art (SOTA) ML algorithms, has gained much attention worldwide [6]. DL processes incoming data automatically, recognizing complex structures in (HD) High-Dimensional information by projecting it onto a lower-dimensional manifold, using representation-learning techniques with several layers of abstraction. Human FE (Feature Engineering) is no longer required as a result. In a number of fields, including (SR) Speech Recognition, (CV) Computer Vision, and NLP (Natural Language Processing), it has been shown that DL may achieve significantly higher accuracies than conventional techniques [7]. The primary use of DL in healthcare and medicine is Medical Imaging Analysis (MIA).

Superior diagnosing ability in detecting a variety of clinical diseases has been shown by DL techniques, including metastatic lymph nodes from (BC) Breast Cancer from parts of tissues, skin images showing malignant melanoma, and chest X-rays showing TB. In addition, DL has been utilized in ocular imaging [8], primarily in Optical Coherence Tomography (OCT) and FI.

AMD, glaucoma, DR, and Retinopathy of Prematurity (ROP) are among the major ocular pathologies for which DL approaches have been employed [9]. Then, Refractive Error (RE) and Cardiovascular Risk Indicators (CV) (e.g., age, BP, smoking status, BMI Mass Index) have been computed using DL [10]. The primary benefit of DL in ophthalmology is possible in screening for disorders with well-defined criteria, such as DR and ROP. Other conditions like AMD and glaucoma may also require screening tests and continual monitoring.

However, screening demands significant time and money from medical facilities in wealthy and developing nations [11].

Screening and monitoring patients in primary vision care centres may be accomplished in the long run by combining the application of DL with telemedicine. Problems with vision are something that almost everyone has to deal with periodically. Many youngsters can be readily handled at home or do not show up on claims, while others require the care of an expert. SEFDL-based techniques for detecting ocular pathologies are suggested in this study.

The study is structured as follows: Section 2 examines a few modern methods for detecting disorders related to ocular pathology. The proposed methodology's approach is presented in Section 3. The results and comments are presented in Section 4. Section 5 discusses future work and the conclusion.

Literature Review

Various ML and DL methods that have been utilized recently to diagnose disorders related to ocular pathology and it is discussed in this section.

A robust Fuzzy Recognition (FR) approach was suggested by Sujatha et al. [12]. The High-Resolution Fundus (HRF) Image Database is used in this analysis to illustrate how ocular illnesses affect the iris.

A sequence of image processing (IP) techniques is applied to the eye images, beginning with methods for noise reduction (NR) with Blind de-convolution. Wavelet-based Feature Extraction (FE) and Principal Component Analysis (PCA) for dimensionality reduction come next. For classifying the healthy and infected eyes, fuzzy C-means clustering (FCM) deduction method is used. The recommended approach completes the process in under two minutes, demonstrating exceptional efficiency. With sensitivity and specificity scores ranging from 94% to 98%, this suggested method delivers remarkable accuracy.

For the Detection of cataracts, a Deep Convolution Neural Network (DCNN)-based technique was presented by Hossain et al. [13]. It consists of 2 modules: training and testing. (RFI) Retinal Fundus The proposed DCNN structure is trained, validated, and tested using images. The experiment's outcomes indicate that the proposed technique may accurately identify eye cataracts.

An ML method for detecting the presence of pathological myopia depends on FI was created by Rauf et al. [14]. For this, the CNN DL approach is employed. In Spyder, a CNN structure is created. Initially, the FI undergo preprocessing. The CNN structure that has been created is then given the preprocessed images. The CNN framework effectively distinguishes between normal and pathological myopia in the input images by FE from the images. With an AUC of 0.9845, the effective performance of the CNN framework was determined. 0.1457 is the best validation loss that was found. Based on the FI, the framework has been demonstrated to detect pathological myopia effectively.

A DL-based algorithm that can be implemented on a smartphone was introduced by Pérez et al. [15] to evaluate the quality of ocular FI. A public eye fundus dataset containing 2 annotations was employed to assess the algorithm. In the case of the three-class classification work and the binary classification (BC) task, the accuracy of the suggested technique was 0.911 and 0.856, respectively. In addition, the technique that has been provided offers fewer parameters than previous SOTA models, making it a viable substitute for a mobile-based device for classifying the quality of the eye fundus.

Junayed et al. [16] introduced a unique (DNN) Deep NN called CataractNet for automatic cataract diagnosis in FI. For training the network with small kernels, fewer training parameters, and layers, the loss and (AF) Activation Functions are adjusted. Therefore, CataractNet's computation cost and average operating time are much lower than those of other pre-trained CNN frameworks. The (AO) Adam Optimizer is employed to optimize the suggested network. For training the model, 1130 FI, both cataract and non-cataract, are gathered and enhanced to create 4746 images. Before training the framework, the dataset gets larger by augmentation to prevent the overfitting issue. With an average accuracy of 99.13%, experimental data demonstrate that the suggested technique executes better than the SOTA cataract detection techniques.

To extract significant features from the image, Sharma et al. [17] used DR of the image and channel contraction, which allowed simply the high-level (HL) features required for reconstructing the segmented feature image. The current approach performs well when detecting glaucoma from retinal OCT Angiography (OCTA) images (area covered by the receiver-operator characteristic curve, or AUC ~ 0.81). The manual and DL-based segmentation parameters showed a statistically significant Pearson's Correlation Coefficient ($p = 0.9969$) and a minimal bias (~ 0.00185) in Bland-Altman analysis. Commercial software exhibits a larger bias (~ 0.0694, $p = 0.8534$) on similar datasets. Depending on the examination of 3670 OCTA images, the current model is 10 times lighter and has superior segmentation accuracy and model training reliability than Unet, which is employed for Biomedical Image Segmentation (BTS).

For computer-aided categorization of Diabetic Macular Edema (DME), drusen, and Choroidal Neo Vascularization (CNV), Sunija et al. [18] presented a DNN-based classifier using standard retinal OCT images. Serious retinal diseases can also be successfully diagnosed and classified from OCT images using a DCNN with six convolutional blocks.

An approach for (CAM) Class Activation Mapping based on gradients is employed to clarify the classification outcomes. High-quality images were produced with minimal assistance from a human operator, as suggested by Teikari et al.'s [19] "active acquisition" embedded DL technique. The enhanced (IQ) Image Quality should result in more reliable DL-based clinical diagnostics in clinical settings. The low cost and continuously rising hardware performance will make embedded DL viable. Pathologists will quickly review potential calculation techniques for more extensive clinical systems. In short, applications can be part of a 3-layer architecture consisting of Fog Layer, cloud layers, and Edge Layer (EL) with the device-level implementation of the former. Improved DL-based clinical (DM) Data Mining is made possible via enhanced EL execution through "active acquisition", It results into higher-quality data saved in cloud and electronic health records (EHR) and serves as an independent data curation operator.

Using porcine eyes that had been treated using Ringer's lactate solution as samples, Song et al. [20] presented an initial automated technique for measuring the volume of subretinal blebs. The Duke Eye Center employed microscope-integrated OCT to provide a 3-D image of the subretinal blebs in porcine eyes. Of 37 eyes, 15 (30 volumes) were chosen, and 2 distinct injection stages were captured and evaluated. The team developing and testing the method didn't impact on the inclusion/exclusion criteria. A unique lightweight technique based on DL was created to segment subretinal bleb borders. Selection bias was prevented by using a cross-validation (CV) technique. An ensemble-classifier approach was employed to produce the outcomes for the test dataset.

A DL method for (MF) Macula Fovea detection on ultra-Widefield Fundus (UWF) images was presented by Wang et al. [21]. 2300 UWFI from China's Shenzhen Aier Eye Clinic were taken for this investigation. The pre-amplifying of 1800 training FI, 400 validation images, and 100 test images is done using methods that utilize U-shape networks (Unet) and Fully Convolutional Networks (FCN). Three qualified ophthalmologists were invited to mark the fovea. An approach is examined from the perspective of anatomy. The spatial link between the MF and the optic disc centre in UWF is the basis for this method. This technique's parameters are established and confirmed to be effective according to the expertise of ophthalmologists. The UWF-OCT technique detects the accurate grounded standard, from which the outcomes are determined by computing the (ED) Euclidean Distance among the suggested techniques and this standard. After comparing the proposed approaches, researchers demonstrate that Unet's DL technique performed better on MF detection tests than the other approaches, yielding results similar to those of the grounded conventional technique.

On the basis of a hierarchical coarse-to-fine deep regression neural network, Xie et al. [22] introduced a novel end-to-end fovea localisation technique. A Gaussian-shift-cropping technique for effective training (DA) Data Augmentation, a multi-field-of-view (multi-FOV) Feature Fusion (FF) approach for context-aware feature learning, and a Multi-Scale FF (MSFF) approach and self-attention approach for leveraging location, semantic, and contextual data in a combined structure are some of the new features of the new approach. Through demonstration, it has been indicated that the innovative strategy obtained SOTA performances by presenting substantial test outcomes on 2 public databases. To further emphasize the overall structure's effectiveness and technical soundness and its different component parts, researchers also provide extensive ablation research and analysis.

Proposed Methodology

This work proposes ocular pathology classification approaches based on Stacking Ensemble Federated Deep Learning (SEFDL). First, an adaptive weight-based median filter is applied to image resizing and removing noise. Then, data augmentation and handling data imbalance using SMOTE are applied. Finally, Stacking Ensemble Federated Deep Learning (SEFDL) is proposed for disease detection. Researchers have used four pre-trained frameworks such as CNN, VGG16, Inceptionv2, and ResNet50 and implemented hyperparameter tuning with the help of the Adaptive Tuna-Swarm Optimization Algorithm.

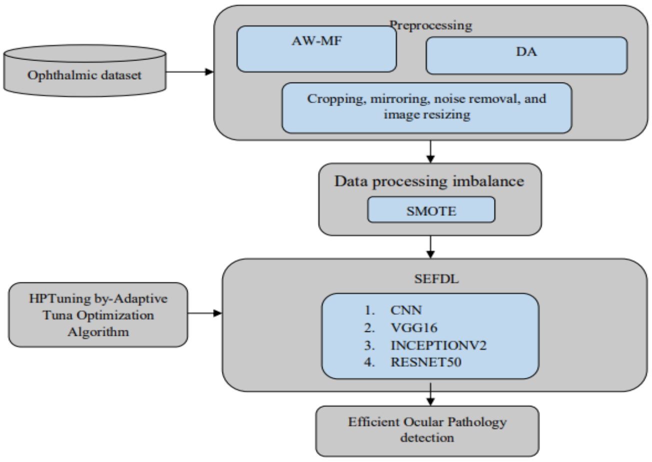


Figure 2. The Overall Process of the Suggested Procedure

Preprocessing using AWF

In order to prepare image data for model input, preprocessing [23] is necessary. In this work, preprocessing is used to enhance fundus image quality. In this study, popular retinal imaging methods are fully explained. The picture is magnified to help detect ocular disease stages.

Noise Filtering

Extraneous or redundant data in images is referred to as image noise. Theoretically, a random error which can be unpredictable in the natural environment is called noise. Gaussian White Noise (GWN), Poisson noise, Salt and Pepper Noise (SPN), and Multiplicative Noise (MN) are prevalent forms of noise in digital IP.

The noise breaks up the clear image.Thus, this technique aims to boost the applicability for many image types; the AWF is suggested to avoid the noises in the image.

Cropping

One technique for editing images is called image cropping to enhance the frame, adjust the aspect ratio, or highlight a subject in an image. The area to crop the image to is indicated by 4 dimensions. The cropping process yields a modified image with the dimensions of the cropping rectangle by removing all details outside of it. The proportion of the image's height to width. 2 numbers split by a colon, such as 4:3 or four-to-three, denote this ratio. Cropping reduces the number of pixels and resizes the image by removing portions.

Mirroring

A reversed item oriented perpendicular to the surface of the mirror and reflected almost exactly is called a mirror image (in a plane mirror).The reflection of materials like water or mirrors causes it to appear as an optical effect.

Image resizes

IQ and file size are impacted when an image is resized since it changes its dimensions. The most common reason to scale an image is to reduce the size and complexity of enormous files. Picture interpolation occurs when you scale or distort a picture transferred from one grid of pixels to another. To increase or decrease the overall amount of pixels in an image, resizing is required. In contrast, remapping compensates for lens distortion or rotates an image.

Adaptive Weiner Filter (AWF)

Based on the local variance of an image, AWF delivers superior results by modifying the filter's output. The (MSE) Mean Square Error among the original and restored images must be as small as possible for an approach to be effective [24]. This method provides an improved overall filtering effect and maintains the edges and (HF) high-frequency regions of the image better than other filters. It has also made a few modifications. One way to handle such conditions is to automatically utilize various window options and choose the best one. However, in the moving window, the centre sample needs to be used efficiently in rough parts but disregarded in smooth regions in order to suppress subjectively annoying singularities. The issue of filtering images corrupted via signal-intended noise can be represented as follows as equation (1)

$$y(i,j) = x(i,j) + n(i,j) \quad (1)$$

Here, the noise-free image can be denoted by $x(i,j)$, the noisy measurement is represented by $y(i,j)$, and the additive noise is represented by $n(i,j)$. Eliminating noise $y(i,j)$or "denoise" are considered to be the main objective in this work. To determine the final processing window for a given pixel in the image, the mean and variance of the pixel in various (WS) Window Sizes, like (3+2i)2, i = 0, 1, 2, 3, are compared. Multiple locations can be used to adaptively select the filter template.

To enhance efficiency and preserve edges and texture, the detail region uses a minimum (WF) window filter, while the smooth area uses the broad WF. The pixels are processed using the formula as follows in equation (2) and (3),

$$r(i,j) = \mu + (1 - q + \Delta) * (s(i,j) - \mu) \quad (2)$$

$$q = \frac{\sigma_{avg}}{\sigma_{var}+1} \quad (3)$$

$$\Delta = \frac{\sigma_{var}}{\sigma_{avg}+\sigma_{max}+1} \quad (4)$$

In this case, σ_{avg} can be used to represent the average of the total variance of every pixel in the selected window.The variance of the current pixel can be symbolized as σ_{var}, and the representation for the largest variance of each pixel in the picture is σ_{max} in equation (4). the original pixel can be denoted as $s(i,j)$, the output pixel. This significantly improves and allows the inner window's output to be transferred to a bigger external window for computing.

Preprocessing for DA

The provided data is insufficient to build robust and highly accurate classifiers in many classification scenarios. A popular method for dealing with the small quantity of data provided in comparison to the number of variables that may be changed in a classifier is to use DA [25]. DA entails employing label-preserving transformations to change existing samples into new ones described in Figure 3. It is commonly known, for example, that sufficiently tiny affine transformations of data preserve an image's label. DA has also been proven to improve the resilience of classifiers against various perturbations, including additive adversarial noise and geometric distortions. Despite dealing with a tiny transformation group, this strategy produces results comparable to the previous research and considerable gains in accuracy and resilience over non-DA schemes.

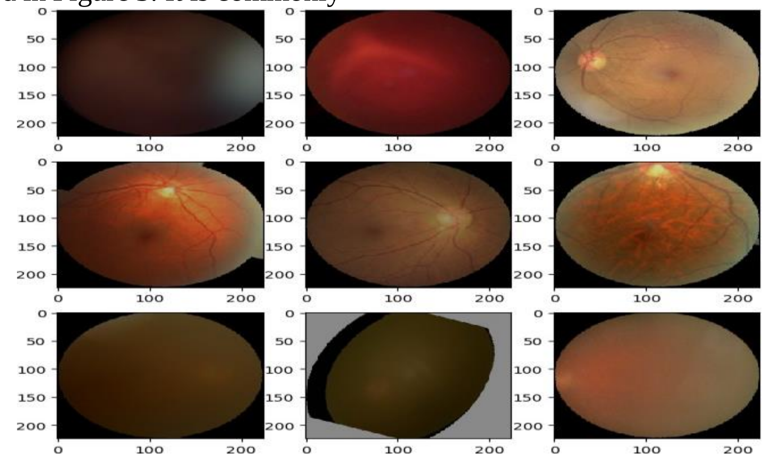


Figure 3. Retinal Fundus Images for Analysis

To increase the amount of data collected and improve the resilience of the FDL algorithm, the DA procedure entails changing the FI. To achieve this objective, FI rotation is utilized multiple times at a random angle. DA is another way to invert or mirror FI. In addition to rotating the images left, right, and down, the camera angles are chosen at random. Image cropping also entails adjusting the structure of size of the FI.FI is subjected to random pixel adjustments ranging from 0% to 10% or 15^% in another DA technique called shape adjustment. Furthermore, the process alters FI's brightness and light intensity by noising, blurring, and adjusting them.

Handling Data Imbalance

Considering that most machine learning techniques for classification were created assuming that each class may contain an equal amount of data, imbalanced classifications present a barrier to Predictive Modelling (PM). For the minority class in particular, this leads to models with minimal predictive accuracy. The minority class is usually more significant than the majority class; hence, this makes the minority class more vulnerable to classification errors. Most classification algorithms strive for true samples to learn and establish as precise a border as possible for every class to improve the prediction performance.

Most classifiers determine that classifying synthetic cases near the border presents a significant learning challenge, while examples far from the border are easier to categorize. Present a novel enhanced technique (A-SMOTE) for preparing imbalanced training sets based on these facts. The goal of this method is to accurately determine the borderline and generate pure synthetic samples using SMOTE generalisation. The recommended procedure consists of the following two steps:

To create the synthetic sample according to the following equation, the first stage initially utilizes the SMOTE method [26].

$N = 2 * (r - z) + z$ (5)

Here, the amount of minority class samples can be denoted as z, the amount of majority class samples can be denoted as r, and the initial synthetic instances amount (currently created) can be symbolized as N in equation (5).

The goal of this method is to accurately determine the borderline and generate pure synthetic samples using SMOTE generalisation. The recommended procedure consists of the following two steps:

After cleaning the given data, the Feature Selection (FS) process is carried out, as mentioned in the section below.

Stacking Ensemble Federated Deep Learning

The performance of the SEDL technique is frequently either better than or equal to that of a single (BC) Base Classifier since it incorporates several separate learning techniques. This has led to its increasing popularity and effectiveness in the area of ML. Various BCs and Meta Classifiers (MC) covered in this section are utilized in the suggested optimum BC for stacking. In this case, the FV (Feature Vector) sets f, b, and d are produced from the dataset; the set is denoted by b, which is applied to various feature sets, and the trained BC set can be denoted as d.

Here, $f = [f_1, f_2., f_n]$ is the representation of the n input FV. The k-train BC is represented by the matrix $b = [b_1, b_2., b_k]$. The output of the m -train base classifiers should now be represented as $d = [d_1, d_2., d_m]$.

The ith trained BC produced the following results, as an instance in equation (6),

$d_i = b(f_i)$ (6)

The MC is selected from the b vector among the BCs, and as a result, the final output p = $[p_1, p_2,., p_n]$ is expressed as follows: equation (7),

$p_i = m_c(d_i)$ (7)

Here, $y = [y_1, y_2,., y_n]$ indicates the dataset sample's actual value and applies the most accurate trained BC to create predictions. More accurate BC that have been trained are more efficient. The OSEL framework allows for generating k additional MC as an additional to the MC. The following standards were employed to determine which BC was the most effective

$$minj(\theta) = \frac{1}{t}\sum_{i=1}^{m}(\|y_i - p_i\|)^2 \; s.t \begin{cases} 1 \leq i \leq m \\ b_i \in b \\ m_i \in b \end{cases} \quad (8)$$

Here, equation (8) shows p_i stands for the predicted value and y_i for the actual value.

CNN

DL with CNNs is the most effective and efficient technique. One type of artificial neural network (ANN) is CNN. In simple terms, (HL) Hidden Layers, OL (Output Layers), and IL (Input Layers) constitute ANNs. Artificial neurons that resemble brain neurons make up each layer. Weights have been attached to every neuron, passing information via HL from the IL to the OL [27]. These weights are repeatedly modified using an (AF) Activation Function that accepts the total input weights as input. Networks make iterations to reduce errors. A (DNN) Deep Neural Network is produced by including more HL. However, unlike conventional ANNs, CNNs scale effectively and may accept whole images as input (Figure 4).

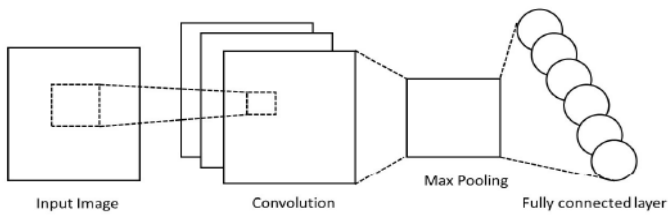


Figure 4. CNN

The majority of a CNN's design is made up of the IL, CL, Rectified Linear Units (ReLU) layer, pooling layer, and (FC) Connected layer. Usually, the IL receives an input image of (height _ width _ the number of channels). Three colour channels, are present in an RGB image. These are the basic layers of a CNN since CL contains convolutional filters, also referred to as kernels. Each of these filters joins the entire image to produce an AF that responds to certain elements, such as colors and edges. Next, a ReLU layer is typically used since it expedites the training procedure using the ReLU AF. Because the pooling layer eliminates redundant data, it gradually reduces the input image's sample size to avoid overfitting. Lastly, several CL and pooling layers are followed by the FC layers. Each neuron in these layers is linked to each AF in the layer before it, for identifying intricate patterns.

VGG 16

The conventional CNN structure is called VGG or VGGNet. VGG was created to enhance the technique's efficiency and these CNNs. A traditional, multi-layered DCNN structure is called Visual Geometry Group (VGG) [28]. Using VGG-16 or VGG-19, which comprises 16 and 19 CL, the term "deep" refers to the number of layers. Figure 5 demonstrates how significant OD (object detection) methods are based on the VGG framework. Additionally, the VGGNet outperforms standard techniques on extensive tasks and datasets beyond ImageNet, but it has been created as a DNN.

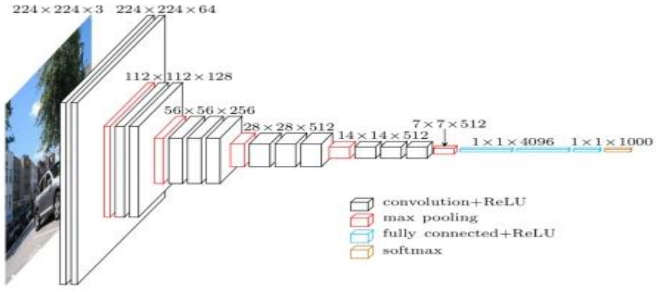


Figure 5. Architecture of VGGnet

VGG Architecture

The fundamental components of CNN serve as the foundation for VGGNets. Very tiny convolutional filters are employed in constructing the VGG network. 13 CLs and 3 FC layers constitute the structure of VGG-16. Analyze the VGG architecture in simple terms:

Input: A 224 x 224 image gets input into the VGGNet. The model's developers divided every image's middle 224x224 patch for the ImageNet competition to retain the image's consistent input size.

Convolutional Layers (CL): Utilizing a minimal receptive field of 3x3, the smallest size that retains up/down and left/correct capture, the CL of VGG utilizes this. Furthermore, 1x1 convolution filters are included, which function as an input (LT) Linear Transformation. The next important AlexNet concept that reduces training times is a ReLU unit. If the input is positive, the ReLU, a piecewise Linear Function (PLF) returns input, or else it returns 0. Preserving the (SR) Spatial Resolution following convolution requires setting the convolution stride, or the number of pixel shifts on the input matrix, at 1.

Hidden Layers (HL): ReLU is employed by every HL in the VGG network. Local Response Normalization (LRN) is usually not employed by VGG since it lengthens training times and uses more memory. Moreover, it generally does not affect accuracy.

Fully-Connected (FC) Layers: There are 3 FC layers in the VGGNet. There are three layers with a total of 1000 channels, one for each class. The first two layers have 4096 channels each.

VGG 16

The CNN framework is also known as VGG16. It is also known as the VGG framework, or VGGNet, which has 16 layers. In ImageNet, the VGG16 framework obtains around 92.7% top-5 test accuracy. One dataset is ImageNet. It substantially surpasses AlexNet by eventually substituting multiple 3×3 kernel-sized filters for the larger kernel-sized filters. Nvidia Titan Black GPUs were implemented for several weeks to train the VGG16 architecture. As mentioned, the 16-layer VGGNet-16 can identify images into 1000 distinct object categories, including pencils, mouse, keyboards, and animals. Additionally, 224x224 images can be entered into the structure, as shown in Figure 6.

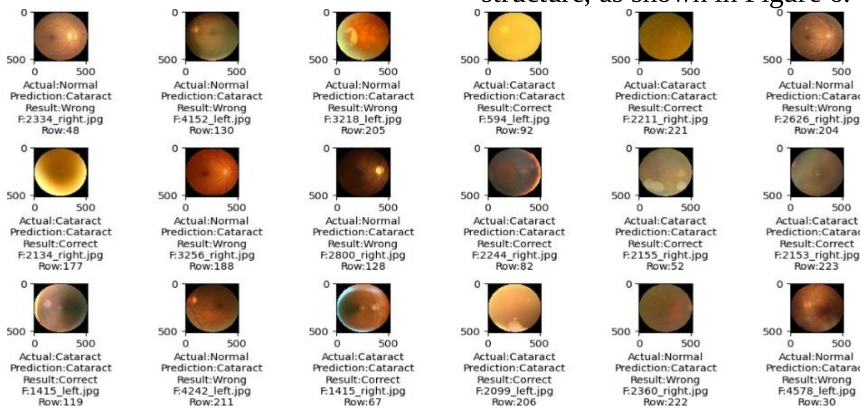
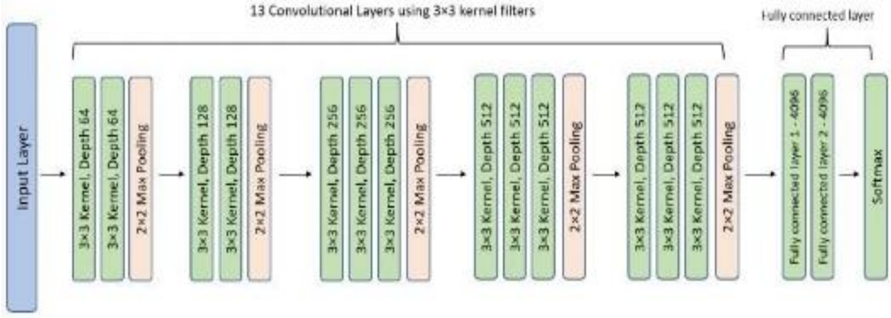


Figure 6. Classification Results of Retinal Fundus Images: Actual vs Predicted Outcomes

The statement that the VGGnet is a 16-layer DNN is indicated by the number 16. This indicates that VGG16 contains over 138 million parameters, making it a relatively large network. It's a massive network, even by today's standards. Figure 7 indicates how the network's simplicity of VGGNet16 structure enhances its appeal. The structure suggests that it is more homogeneous. A few CLs are followed by a pooling layer that reduces the height and width.

It is possible to increase the number of filters employed from the approximately 64 currently accessible to approximately 128 and, ultimately, 256. Utilize 512 filters in the final layers.



\]

Figure 7. Architecture of VGG16

Several VGG16 issues with FE from image datasets were brought to light by the previous analysis. Initially, 1000 categories were recognized from 1 million images using VGG16 in an HD dataset. Consequently, a noticeable underfitting was applied to tiny datasets with fewer training characteristics. Due to the excessive homogeneity of image features, frameworks have difficulty capturing pertinent features, making it challenging to determine if a disease is present or absent in images. As a result, recognition errors are frequently produced during training using deeper network layers. This problem is resolved using the regularization in the layers of the VGG16 model. This work introduces the regularization-based Adaptive factor (RA) to solve the underfitting problem.

Regularization based Adaptive Factor (RA)

Further strategies are employed in addition to the regularized target to prevent underfitting. After every boosting step, the RA function is comprised of freshly added weights through a factor RAw_i^*. Adaptive learning lessens the impact of every distinct layer and creates possibilities for future predictions to enhance the framework, comparable to the LR (Learning Rate) in optimization.

Determine the ideal weight. RAw_j^* of layers j for a fixed structure q(x) as given below in equation (9),

$$RAw_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (9)$$

An approximation factor is denoted by ∈. This suggests that approximately selected points are automatically selected. In this case, the weight of each data point is represented by hi in equation (10) and (11),

$$\sum_{i=1}^{n} \frac{1}{2} h_i (f_t(x_i) - g_i/h_i)^2 + \Omega(f_t) \quad (10)$$

As the instance set of layer j, define $I_j = \{i | q(x_i) = j\}$. It can also expand Ω in the following ways:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i f_t^{\,2}(x_i)] + \Omega(f_t)$$
(11)

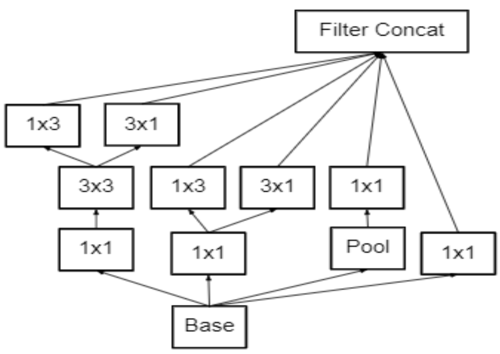
$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i f_t^{\,2}(x_i)] + \frac{1}{2} \lambda \sum_{j=1}^{T} w_j^2 \quad (12)$$

It has labels g_i/h_i and weights h_i, and it is accurately weighted squared loss. It is difficult to identify candidate splits that demand huge datasets in equation (12).

Inspection v2

Batch Normalization (BN) is a notable feature of Inception v2, the 2nd generation of Inception CNN structures. Due to the advantages of BN, additional modifications include the removal of Local Response Normalization and the reduction of dropout. Applying convolutions like 5*5 in Inception V1 occasionally results in a significant reduction in the input dimensions. As a result, the NN employs a fall in accuracy [29]. The NN may lose some of its data if the input dimension drops too much. Additionally, using larger convolutions 5×5 instead of 3×3results in a decrease in complexity.

Factorization can be extended further, meaning that a 3×3 (C) Convolution can be divided into an asymmetric 1×3 C. Subsequently, a 3×1 C. Compared to a 3×3C, and this is similar to sliding a two-layer network with the same receptive field, but at a rate of 33% less. This factorization only functions well when the input size is mxm (12 to 20), not when the input dimensions are large for early layers. According to the Inception V1 architecture, the auxiliary classifier improves the network's convergence. Moving the beneficial gradient to earlier layers (to lower the loss), researchers suggest that it may support and mitigate the effects of the Vanishing Gradient Problem (VGP) in deep networks.



\

Figure 8. Inspection V2 Architecture

In the Inception V2 architecture, as shown in figure 8. Two 3×3 C take the place of the 5×5 C. A 5×5 C takes 2.78 times more than a 3×3 C; implementing this reduces computing time and boosts speed. Accordingly, the architecture performs better when it employs two 3×3layers rather than five 5×5 layers. Additionally, 1xn and nx1 factorization are produced from nXn factorization using this structure. As previously mentioned, a 3×3 C can be transformed into a 1×3 C, which is followed by a 3×1 C, which has a computational complexity 33% lower than that of a 3×3. Rather than exploring deeper, the module's feature banks were enlarged to address the representational bottleneck issue. Doing this would avoid the knowledge loss that results from going deeper.

Resnet 50

Gradient explosion/dissipation is the initial issue with increasing depth. This is because as the amount of layers rises, the GBP (Gradient Back Propagation) in the network becomes unstable due to multiplications, growing either very high or very little. Gradient Dissipation (GD) is one of the frequent issues. Many methods have been discovered for overcoming GD, including Xavier initialization, utilizing BN, and altering the AF to ReLU. GD has been effectively solved, it may be stated [30]. Degradation, or the worsening of the network's performance with increasing depth, is another issue with deepening networks.

The network can do more intricate structure FE when the amount of layers is raised so that deeper models may offer superior outcomes. The deep network was discovered to be degenerating, nevertheless, by the test. The network's accuracy becomes saturated or declines as network depth increases. The accuracy of the training set is declining. Overfitting doesn't seem to be the reason behind this. Since overfitting should result in high training set accuracy, ResNet solves this difficulty, and several orders of magnitude are added to the network's depth once it is resolved.

The final output is y = F(x)+x. ResNet presented two types of mapping: The "curved curve" in Figure 9 is referred to as Identity Mapping (IM) and the component that is not the "curved curve" is referred to as Residual Mapping (RM). In the formula, IM refers to x because, as its name implies, it pertains to "difference," or y − x. On the other hand, F(x) is what RM means. ResNet-50 initially executed convolution operations on the input, then four residual blocks, and final FC operations to complete classification tasks. With 50 Conv2D procedures, ResNet-50's network structure is presented in Figure 10.

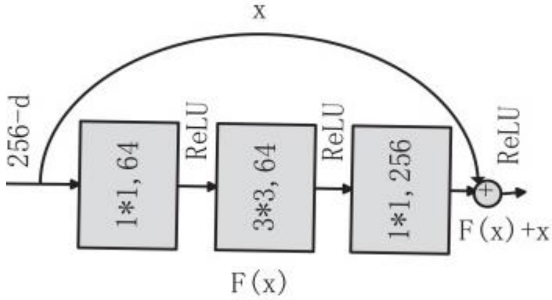


Fig. 9. Residual Block

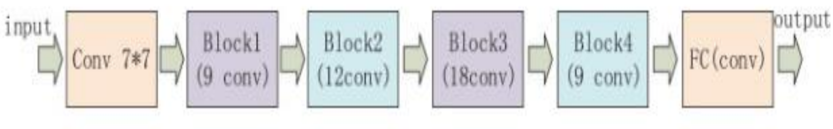


Fig. 10. ResNet-50

To summarize the features of the preceding layers, the FC layer often appears at the final stage of the CNN. The last FC layer can be viewed as feature weighing as it considers the preceding convolution and pooling to be the feature engineering, local amplification, and local FE processes. The FC layer's structure typically facilitates rapid learning of the nonlinear combination of sophisticated features produced via the CL. The potential nonlinear function will be taught to the FC layer. The following is the fundamental process of learning. The image is initially compressed into CV (Column Vectors) and put back into the FFNN (Feed Forward NN) after being transformed into a format that can be utilized with an MLP. Every training process subsequently utilizes the compressed information. Using classification approaches like Softmax, the model can effectively discriminate between the image's primary features and a few low-level features.

Parameter Optimization using Adaptive Tuna Swarm Optimization (TSO)

The most prominent hunter in the water is the tuna. Some little predators are more flexible than tuna despite the tuna's high swimming speed. As a result, when engaging in predatory behaviour, tuna frequently prefer to cooperate in groups to catch targets. The spiral and (PF) Parabolic Foraging techniques are the 2 effective predatory tactics the TS employs [31]. Each tuna closely follows the preceding fish as the swarm utilizes the PF technique. The tuna swarm encircles the target in a parabola. The tuna swarm will group in spiral formations and push fish toward shallow water areas using the SF (Spiral Foraging) method. There is a higher chance of capturing prey. Through observing these 2 (TS) Tuna Swarm foraging behaviours, researchers introduced TSO, a novel (SI) Swarm Intelligence Optimization.

Population Initialization

A TS comprisesNP tunas. During the early phase of the swarm, the TSO approach generates the initial swarm is randomly distributed throughout the search space. The following are the formulae that are employed to initialize tuna individuals in equation (13) and (14),

$$X_i^{int} = rand.(ub - lb) + lb \quad (13)$$

$$= [x_i^1 x_i^2 \dots \dots x_i^j] \begin{cases} i = 1,2, \dots., NP \\ j = 1,2, \dots., Dim \end{cases} \quad (14)$$

Here, ub and lb represent the upper and lower bounds of the tuna exploration range; the i-th tuna is denoted by X_i^{int}, and random variables with uniform distributions between 0 and 1 can be denoted as $rand$. Specifically, every member of the TS, X_i^{int}, denotes a potential TSO solution. Every single tuna comprises a collection of (Dimensional)Dim numbers.

Parabolic Foraging Strategy

The main foods of tuna are eel and herring. They frequently adjust their swimming direction to avoid predators by using their speed advantage. For predators, catching them is extremely difficult. The TS will attack their prey together since they are less nimble than their target. The prey will serve as a point of reference for the TS as it pursues other prey [32]. As they hunt, the TS creates a parabola around the prey, with each fish following the previous one. The TS additionally employs a spiral foraging technique. The following is a mathematical representation of the TS's PF, assuming a 50% chance of selecting either strategy:

$$X_i^{int} = \begin{cases} X_{best}^t + rand.(X_{best}^t - X_i^t) + TF.p^2.(X_{best}^t - X_i^t), if\ rand < 0.5 \\ TF.p^2.X_i^t, if\ rand \geq 0.5 \end{cases} \quad (15)$$

$$p = \left(1 - \frac{t}{t_{max}}\right)^{\frac{t}{t_{max}}} \quad (16)$$

Here, equation (15) and (16) shows the maximum amount of preset repetitions denoted as t_{max} and the repetition that is being executed at the moment can be represented as t. Random values of 1 or −1 make up TF.

Spiral Foraging Strategy

The SF approach is an additional effective cooperative foraging technique besides the PF. Most tuna cannot make the correct direction decisions when pursuing their prey, but a select few can steer the swarm of fish in the appropriate direction. When this small group of fish begins to pursue the prey, the adjacent tuna will follow.

The entire swarm of tuna will finally come together in a spiral pattern to capture their prey. The TS employs an SF method, and members of the swarm will communicate with each other and the best people to follow or nearby members. Even the greatest individual may not always be able to direct the swarm efficiently to catch prey. A random member of the swarm will then be chosen by the tuna to be followed. The SF strategy's mathematical equation is given below:

$$X_i^{t+1} = \begin{cases} \alpha_1(X_{rand}^t + \tau.|X_{rand}^t - X_i^t| + \alpha_2.X_i^t), \\ \alpha_1(X_{rand}^t + \tau.|X_{rand}^t - X_i^t| + \alpha_2.X_{i-1}^t), if\ rand < \frac{t}{t_{max}} \\ \alpha_1(X_{best}^t + \tau.|X_{best}^t - X_i^t| + \alpha_2.X_{i-1}^t), i = 2,3, \ldots, NP \end{cases} \quad (17)$$

Procedure 1. Pseudocode of Adaptive TSO

```
Initialization: Set parameters NP, Dim, a, z and TMax
Initialize the tuna's position    X_i (i = 1, 2, …, NP) in (13)
Counter    t = 0
  while T < TMax do
Calculate every tuna's fitness value (FV)
Update the best tuna's position and value.    X_best^t
    for (each tuna) do
Update        α1, α2, p through (18), (19), (20)
        if (rand < z) then
Update            X_i^{t+1} via (13)
        else if (rand < z) then
          if (rand < 0.5) then
Update              X_i^{t+1} in (17)
          else if (rand <0.5) then
Update              X_i^{t+1} in (14)
t = t + 1
Update
α_1i(t) by (21)
α_2i(t) by (22)
  return the best FV f (Xbest) and the best tuna Xbest
```

Every tuna randomly selects whether to use the PF or SF strategies during the repetitive TSO method procedure. Following probability Z, tuna will also produce additional individuals

Here, X_{best}^t denotes the current best individual. The i-th tuna in the t+1iteration is indicated by X_i^{t+1} . The reference point chosen at random within the TS is denoted by X_{rand}^t in equation (17).

The trend weight coefficient is denoted by α. The tuna individual swims towards the ideal individual or randomly chosen nearby individuals, and α is used to control this behaviour.The trend weight coefficient, denoted as α_2, is employed to regulate the tuna individual swimming in the direction of the individual. The parameter denoted as τas determines the distance that separates the optimal or randomly chosen individual position from the current individual. The following computational framework is expressed as equation (18) and (19),

$$\alpha_1 = \alpha + (1 - a).\frac{t}{t_{max}} \quad (18)$$

$$\alpha_2 = (1 - a) - (1 - a).\frac{t}{t_{max}} \quad (19)$$

$$\tau = e^{bl}.\cos(2\pi b) \quad (20)$$

$$l = e^{3\cos\left(\left(\left(t_{max}+\frac{1}{t}\right)-1\right)\pi\right)} \quad (21)$$

Here, the random value is distributed uniformly in the interval [0, 1] can be denoted as b, and a is a constant that determines the unit of tuna in equation (20) and (21).

inside the search area. TSO will thus select various methodologies depending on Z when creating novel individual locations. In the population, every tuna individuals are adapted continuously as the TSO method runs up until the specified amount of repetitions is reached. The TSO approach produced the optimal individual and its optimal value within the population.

The first method explains the following benefits of TSO: (1) The TSO method can be implemented more easily because it has fewer configurable parameters. (2) This approach will save the best tuna individual's position for every repetition; the optimal value position will remain unchanged irrespective of the superiority of the candidate solution regressions. (3) By choosing 2 foraging methods, the TSO method can maintain the exploration and exploitation equilibrium. Moreover, the weight coefficient of tuna pursuing behaviour in ATSO is modified by introducing a nonlinear adaptive weight operator. Global exploration and local exploitation have a balanced connection in ATSO's adaptive approach.

Adaptive Weight Coefficient Factor

It is evident from Equations (5) and (6) that there are linear changes in the weight parameters $\alpha1$ and $\alpha2$. Nevertheless, the TSO optimization procedure is intricate, and the linear adjustments of weight parameters W1 and W2 cannot represent the algorithm's real optimization process. Recently, many researchers have applied (NAW) Nonlinear Adaptive Weights to enhance SI optimization techniques, overcoming the limitations of linear control weights. Extensive testing suggests that the NAW technique overtakes the linear weight strategy by optimization. Thus, this study introduces 2 better nonlinear weight parameters, α_{1i} and α_{2i}. It is expressed as follows:

$$\alpha_{1i}(t) = \alpha_{1ini} - (\alpha_{1ini} - \alpha_{1fin}).\sin\left(\frac{t}{\mu.T_{ma}}.\pi\right) \quad (22)$$

$$\alpha_{2i}(t) = \alpha_{2ini} - (\alpha_{2ini} - \alpha_{2fin}).\sin\left(\frac{t}{\mu.T_{ma}}.\pi\right) \quad (23)$$

Here, $\alpha1$'s beginning value with μ= 2 can be denoted asα_{1ini}, $\alpha1$'s final value can be denoted as α_{1fin}, $\alpha2$'s initial value was represented by α_{2ini}, and $\alpha2$'s final value was denoted as α_{2fin} in equation (22) and (23). The original weight parameters, $\alpha1$ and $\alpha2$, were compared with (α_{1i} and α_{2i})the enhanced weight parameters.

Results and Discussion

A real-world ophthalmic dataset is employed for the experiments. In this case, a GeForce GTX 1080 Ti GPU is employed to run Python and the Keras library for this research.

Dataset Details

From the Shanggong Medical Technology Co., Ltd., a real ophthalmic dataset is acquired named OIA-ODIR. Medical information and retinal FI from 5,000 patients in China are included in this.

The patients in this dataset are categorised in Table 1 based on their conditions. Annotations are also documented by a certified ophthalmologist. Because glaucoma, cataracts, and AMD detection are the main goals of this investigation, from the original dataset, only four FI classes are used. The AMD, cataract, and glaucoma all remain normal. All of the AMD, cataract, and glaucoma images in the collection are used. For training, 1,000 normal samples were randomly selected. Table 1 shows the total number of images in each class.

Table 1. Framework Parameters for the 3 Disease Dataset

Disease	**Model**	**Parameters**				
		N	n	α	γ	σ
AMID	EfficientNet-B3 FL- EfficientNet-B3	600	-	0.25	2	0.80

	BCL- EfficientNet-B3 FCL- EfficientNet-B3	600	200	0.25	2	0.80
Glaucoma	EfficientNet-B3 FL-EfficientNet-B3	500	-	0.30	3	0.80
	BCL- EfficientNet-B3 FCL- EfficientNet-B3	500	150	0.30	3	0.80
Cataract	EfficientNet-B3 FL-EfficientNet-B3	50	-	0.25	2	0.70
	BCL- EfficientNet-B3 FCL- EfficientNet-B3	50	15	0.25	2	0.70

Performance Metrics

The efficacy of the proposed method is assessed using the metrics of accuracy (Acc), precision (P), recall (R), and f-measure. The suggested Federated Deep Learning (FDL) method is compared with the existing Lesion-Localization Convolution Transformer (LLCT) and CNN. The following is the calculation of the performance metrics based on this confusion matrix.

Precision measures how many successfully detected positive observations there are out of all shown positive observations, shown in equation (24).

$Precision = TP/(TP + FP)$ (24)

Sensitivity or recall is the proportion of correctly identified positive observations to all observations in equation (25).

$Recall = TP/(TP + FN)$ (25)

The weighted average of accuracy and recall is referred to as the F-measure. It involves both false positives and false negatives as an equation (26) outcome.

$F - measure = 2 * (Recall * Precision)/(Recall + Precision)$ (26)

In terms of positives and negatives, $Accuracy$ is computed in the following way of equation (27),

$Accuracy = (TP + FP)/(TP + TN + FP + FN)$ (27)

Where TP- True Positive, FP- False Positive, TN-True Negative, FN- False Negative.

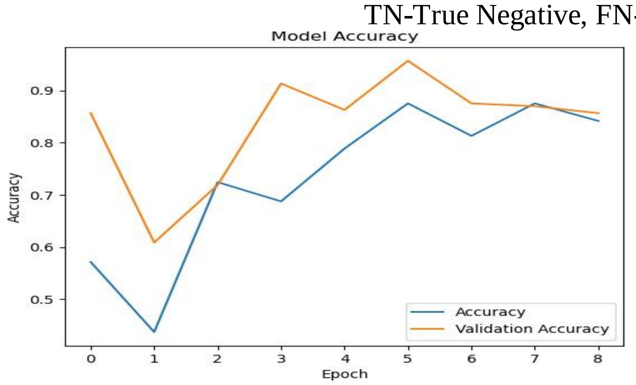


Figure 11. Model Accuracy

Figure 11 represents the provided image that illustrates a line graph comparing a taining of a model and validation precision in a series of epochs (E). Here, the x-axis are plotted with a number of E, and on the y-axis, the model's accuracy is presented. Subsequently, the orange line represents the validation accuracy and the blue line represents the training accuracy. Initially, validation accuracy is higher than training accuracy but fluctuates significantly. Towards the final epochs, both accuracies converge near 0.9, suggesting improved model generalization with minor fluctuations.

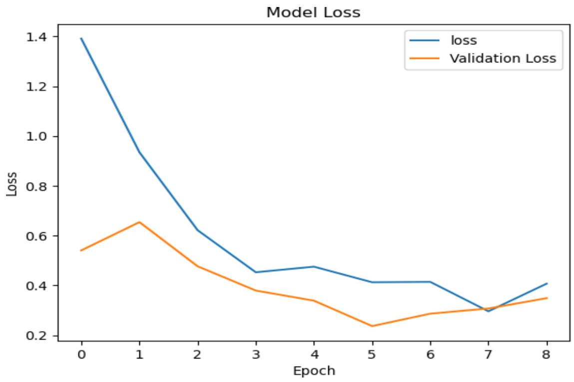


Figure 12. Model Loss

Figure 12 shows the image depicting a line graph comparing model loss and validation loss over epochs during training.). Here, the x-axis are plotted with a number of E, and on the y-axis, the model loss is presented. Following that, the orange line shows the validation loss, while the blue line shows the training loss. In the beginning, both losses drop down rapidly, with training loss higher than validation loss. As training progresses, the losses converge around epoch 7, though validation loss rises slightly, indicating potential overfitting.

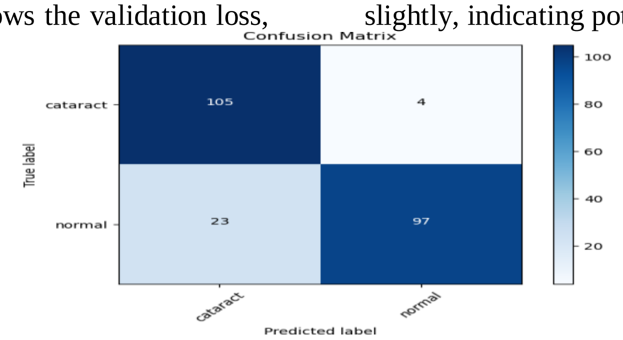


Figure 13. Confusion Matrix (CM)

Figure 13 describes the graph as a CM that assesses a classification model's performance in two categories: "cataract" and "normal." The model correctly predicted 105 cataract cases and 97 normal cases, demonstrating good overall accuracy. However, it misclassified 4 cataract cases as normal (false negatives (FN)) and 23 normal cases as cataract (false positives (FP)). The diagonal values, representing correct classifications, dominate, indicating the model’s effectiveness. While the model performs well, reducing FP and FN can further improve its reliability.

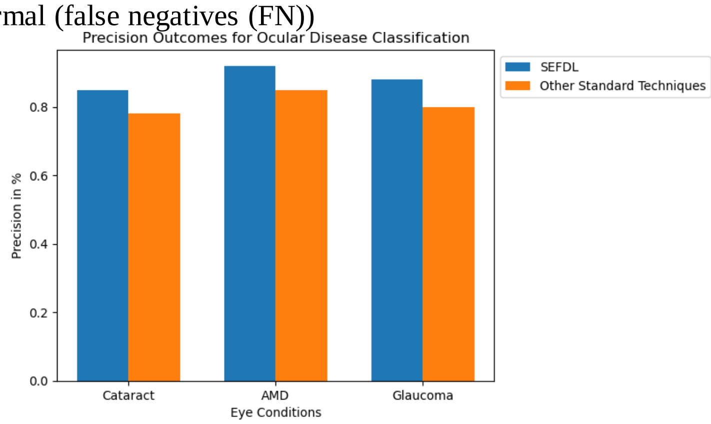


Figure 14. Results of a Precise Comparison between the Recommended and Current Techniques for Ophthalmic Disease Data Classification

Figure 14 shows the precision results for comparison between the recommended SEFDL and the current method for categorizing the information on ocular diseases. According to those outcomes, the FDL algorithm performed better than other standard techniques out of all the DL frameworks tested to identify 3 eye conditions: cataract, AMD, and glaucoma. Overall, the outcomes showed that the SEFDL model performed better on the provided datasets than the ML techniques under consideration. These outcomes align with the error rate previously obtained and can be linked to the rule sets that the suggested SEFDL classification framework produced. The results show that the suggested SEFDL approach performs more accurately than the present classification techniques.

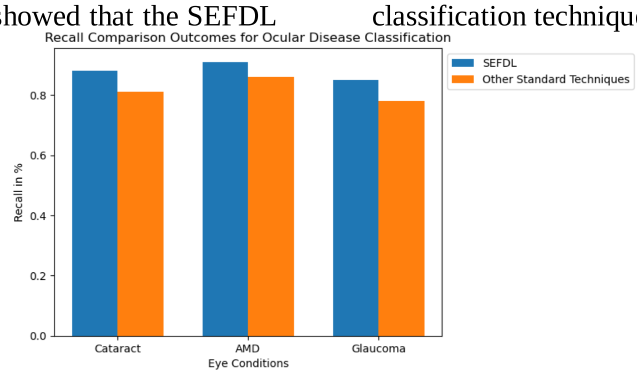


Figure 15. Recall Comparison Outcomes among the Suggested and Current Technique for the Ophthalmic Disease Data Classification

Figure 15 shows the recall comparison results between the recommended and current methods for categorizing the information on ocular diseases. The primary focus of the data utilized in this research deals with diagnosing people exhibiting signs of ocular diseases, taking into account multiple variables that often impact the diagnosis outcome. Therefore, the PM is considered a classification issue arising from if ocular disease is present or absent. As an outcome, the suggested SEFDL frameworks were employed for the study, and they examined at and evaluated the results.

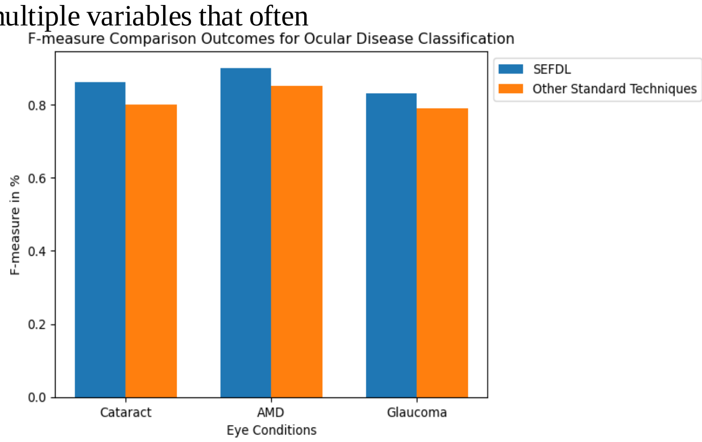


Figure 16. F-measure Comparison Outcomes among the Suggested and Current Techniques for the Ophthalmic Disease Data Classification

Figure 16 displays the outcomes of the F-measure comparison between the recommended and current techniques for categorizing the information on ocular disease. Due to the transparency of the eye's lenses being affected by cataracts, blood vessels usually disappear, and the retina appears blurry in cataract FI. It assumes that frequently low contrast leads to the cataract, which is simple to identify. If a condition is not severe, it is difficult to classify. As can be seen from the comparison above, the suggested frameworks have the highest f-measure rate assessment in each database when compared to the other DL frameworks. The used database produced the best f-measure outcomes when compared to alternative methods.

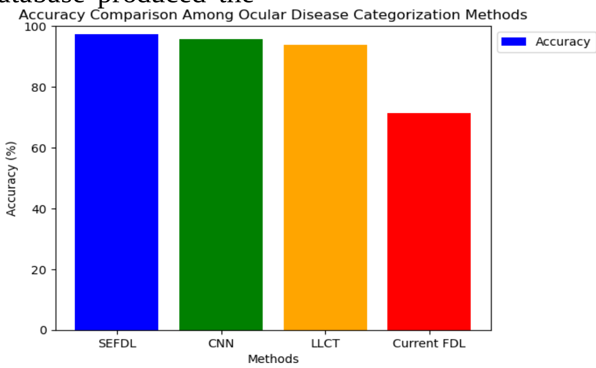


Figure 17. Accuracy Comparison Outcomes among the Suggested and Current Techniques for the Ophthalmic Disease Data Classification

Figure 17 compares the accuracy of the recommended and recent methods for categorizing the information on ocular disease. An ensemble FDL is considered effective if it accurately predicts its objective and generalizes predictions to additional cases. Sensitivity and specificity are two subcategories of accuracy that are typically used to measure a framework's validation. In contrast to the CNN, LLCT, and current FDL models, which have respective accuracy of 95.67%, 93.9%, and 71.41%, the simulation results show that the suggested SEFDL technique has a high accuracy of 97.27%. The investigation shows that the suggested SEFDL strategy performs more accurately than the existing classification methods.

Conclusion

The primary objective of this study is to develop a reliable and effective diagnostic technique to detect ocular pathologies. This work suggests that the Detection of ocular pathology methods depends on SEFDL. An adaptive weight-based median filter is initially applied to image resizing and removing noise. Then, data augmentation and handling of data imbalance using SMOTE are applied. Finally, Stacking Ensemble Federated Deep Learning (SEFDL) is suggested for disease detection. 4 pre-trained models, including CNN, VGG16, Inceptionv2, and ResNet50, were employed, and the TSO Technique was utilized to adjust the hyperparameters. For the chosen SEFDL approach, the Adaptive TSO, ReLU, and sigmoid AFs demonstrated the most effective performance. The results of the recommended approach are compared with those of different SOTA methods. The comparison's accuracy and precision values demonstrate that the suggested SEFDL approach outperforms the others. This research could include the diagnostic investigation of additional eye-related disorders, such as retinopathy and glaucoma. Moreover, it may be applied to investigate the effectiveness of this method and other approaches for this kind of situation in the FDL study.

Acknowledgement

Author Contribution

Mrs. S. Geethamani: Methodology, Writing – original draft, Conceptualization, Supervision

Dr. L. Sankari: Data collection and analysis, Writing – review &amp, Data interpretation

Conflict of Interest

The authors of this study state that they have no conflicts of interest.

Funding

"This research received no external funding."

Ethical Approval Statement

This research did not use human or animal subjects, or any methods that would call for ethical approval under current norms, so this study doesn’t need ethical approval.

References

[1]. Elsawy, A., Abdel-Mottaleb, M., and AbouShousha, M., 2020, Diagnosis of corneal pathologies using deep learning. In *Ophthalmic Technologies XXX*, vol. 11218, 150-160.

[2]. Rauf, N., Gilani, S. O., and Waris, A., 2021, Automatic Detection of pathological myopia using machine learning. *Scientific Reports*, col. 11, no. 1, 16570.

[3]. Ting, D. S., Peng, L., Varadarajan, A. V., Keane, P. A., Burlina, P. M., Chiang, M. F., and Wong, T. Y., 2019, Deep learning in ophthalmology: the technical and clinical considerations. *Progress in retinal and eye research*, vol. 72, 100759.

[4]. Akil, M., Elloumi, Y., and Kachouri, R., 2021, Detection of retinal abnormalities in fundus image using CNN deep learning networks. In *State of the Art in Neural Networks and their Applications*, 19-61.

[5]. Perdomo Charry, O. J., and González, F. A., 2020, A systematic review of deep learning methods applied to ocular images. *Ciencia e IngenieriaNeogranadina*, vol. 30, no. 1, 9-26.

[6]. Gargeya, R., andLeng, T., 2017, Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*, vol. 124, no. 7, 962-969.

[7]. Liu, T. A., Wei, J., Zhu, H., Subramanian, P. S., Myung, D., Paul, H. Y., and Miller, N. R., 2021, Detection of optic disc abnormalities in color fundus photographs using deep learning. *Journal of Neuro-Ophthalmology*, vol. 41, no. 3, 368-374.

[8]. Kazi, K. S., 2024, Computer-Aided Diagnosis in Ophthalmology: A Technical Review of Deep Learning Applications. *Transformative Approaches to Patient Literacy and Healthcare Innovation*, 112-135.

[9]. Grewal, P. S., Oloumi, F., Rubin, U., and Tennant, M. T., 2018, Deep learning in ophthalmology: a review. *Canadian Journal of Ophthalmology*, vol. 53. No. 4, 309-313.

[10]. Balyen, L., and Peto, T., 2019, Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. *The Asia-Pacific Journal of Ophthalmology*, vol. 8, no. 3, 264-272.

[11]. Moraru, A. D., Costin, D., Moraru, R. L., and Branisteanu, D. C., 2020, Artificial intelligence and deep learning in ophthalmology-present and future. *Experimental and Therapeutic Medicine*, vol. 20, no. 4, 3469-3473.

[12]. Sujatha, K., Thivya, K. S., Anand, M., Jayachitra, N., Durgadevi, G., Bhavani, N. P. G., and Srividhya, V., 2020, Vision machine learning for Detection of ocular pathologies from iris images. *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 23, no. 1, 145-155.

[13]. Hossain, M. R., Afroze, S., Siddique, N., and Hoque, M. M., 2020, Automatic Detection of eye cataract using deep convolution neural networks (DCNNs). In *2020 IEEE region 10 symposium (TENSYMP)*, 1333-1338.

[14]. Rauf, N., Gilani, S. O., and Waris, A., 2021, Automatic Detection of pathological myopia using machine learning. *Scientific Reports*, vol. 11, no. 1, 16570.

[15]. Pérez, A. D., Perdomo, O., and González, F. A., 2020, A lightweight deep learning model for mobile eye fundus image quality assessment. In *15th international symposium on medical information processing and analysis*, vol. 11330, 151-158.

[16]. Junayed, M. S., Islam, M. B., Sadeghzadeh, A., and Rahman, S., 2021, CataractNet: An automated cataract detection system using deep learning for fundus images. *IEEE access*, vol. 9, 128799-128808.

[17]. Sharma, P., Ninomiya, T., Omodaka, K., Takahashi, N., Miya, T., Himori, N., and Nakazawa, T., 2022, A lightweight deep learning model for automatic segmentation and analysis of ophthalmic images. *Scientific reports*, vol. 12, no. 1, 8508.

[18]. Sunija, A. P., Kar, S., Gayathri, S., Gopi, V. P., and Palanisamy, P., 2021, Octnet: A lightweight cnn for retinal disease classification from optical coherence tomography images. *Computer methods and programs in biomedicine*, vol. 200, 105877.

[19]. Teikari, P., Najjar, R. P., Schmetterer, L., and Milea, D., 2019, Embedded deep learning in ophthalmology: making ophthalmic imaging smarter. *Therapeutic advances in ophthalmology*, vol. 11, 2515841419827172.

[20]. Song, Z., Xu, L., Wang, J., Rasti, R., Sastry, A., Li, J. D., and Farsiu, S., 2021, Lightweight learning-based automatic segmentation of subretinal blebs on microscope-integrated optical coherence tomography images. *American journal of ophthalmology*, vol. 221, 154-168.

[21]. Wang, H., Yang, J., Wu, Y., Du, W., Fong, S., Duan, Y., and Wu, F., 2021, A fast lightweight based deep fusion learning for detecting macula fovea using ultra-widefield Fundus images.

[22]. Xie, R., Liu, J., Cao, R., Qiu, C. S., Duan, J., Garibaldi, J., and Qiu, G., 2020, End-to-end fovea localization in colour fundus images with a hierarchical deep regression network. *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, 116-128.

[23]. Kotsiantis, S. B., Kanellopoulos, D., and Pintelas, P. E., 2006, Data preprocessing for supervised leaning. *International journal of computer science*, vol. 1, no. 2, 111-117.

[24]. Jin, F., Fieguth, P., Winger, L., and Jernigan, E., 2003, Adaptive Wiener filtering of noisy images and image sequences. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, vol. 3, III-349.

[25]. Shorten, C., andKhoshgoftaar, T. M., 2019, A survey on image data augmentation for deep learning. *Journal of big data*, vol. 6, no. 1, 1-48.

[26]. Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P., 2002, SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, vol. 16, 321-357.

[27]. Zhu, H., Zhang, H., andJin, Y., 2021, From federated learning to federated neural architecture search: a survey. *Complex and Intelligent Systems*, vol. 7, no. 2, 639-657.

[28]. O'Shea, K., and Nash, R., 2015, An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

[29]. Iqbal, H., Chawla, H., Varma, A., Brouns, T., Badar, A., Arani, E., and Zonooz, B., 2022, AI-Driven Road Maintenance Inspection v2: Reducing Data Dependency and Quantifying Road Damage. *arXiv preprint arXiv:2210.03570*.

[30]. Koonce, B., andKoonce, B., 2021, ResNet 50. *Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization*, 63-72.

[31]. Xie, L., Han, T., Zhou, H., Zhang, Z. R., Han, B., and Tang, A., 2021, Tuna swarm optimization: a novel swarm-based metaheuristic algorithm for global optimization. *Computational intelligence and Neuroscience*, vol. 2021, 1-22.

[32]. Obayya, M., Arasi, M. A., Alruwais, N., Alsini, R., Mohamed, A., andYaseen, I., 2023, Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model. *IEEE Access*.