Antara Malakar[1*]
Ankur Ganguly[2]
Swarnendu Kumar Chakraborty[3]

Improved Deep Learning Framework with hybrid feature selection model for Exact Identification and Categorization of Multiple Eye Diseases using Retinal Images

JES Journal of Electrical Systems

Abstract

Nowadays, different models are developed to detect and classify eye disease using Deep Learning (DL) but still issues of misclassification, complexity, and error. Also, multiple eye disease detection is challenging because of complex data. So, in this paper, design a comprehensive DL framework to detect and classify multiple eye diseases using retinal images. Initially, multiple eye disease images are collected and trained in the system. Image augmentation is performed using geometric transformation to enhance the dataset. The preprocessing technique used the Isolation Forest Algorithm (IFA) for outlier detection, BM3D for image denoising, CLAHE for contrast enhancement and Macenko's method for color normalization. To enhance the performance of ROI identification, propose a Loss and Attention-Augmented Multi-Scale U-Net++ (LAAMS-UNet++) architecture. Then extract retinal vessel features, optic disc features, texture features, and deep features in the feature engineering phase. Moreover, a Hybrid Seahorse and Beluga Whale (HSHBW) optimization is employed for the feature selection phase to enhance the performance of classification results. Finally, classification is performed using a Dense-CSPDarkNet53 + LinkNet-34 model with an EfficientNetB7 encoder. The developed model accurately detects and classifies multiple eye diseases with high accuracy. The attained results are authorized with existing techniques regarding precision, accuracy, specificity, sensitivity, etc. The model performs high accuracy of 98.521%, precision of 98.812%, sensitivity (97.102%), and specificity (97.553%) demonstrating its reliable performance over a variety of datasets.

Keywords: Multiple eye disease classification; deep learning; glaucoma; age-related macular degeneration; diabetic retinopathy, fundus images

1. INTRODUCTION

The prevalence of eye blindness in the 21st century is attributed to early interaction with electronic devices like computers and televisions, even though most eye disorders are curable if identified earlier [1]. The eyes are the visual method's organs; they take in light, use a diaphragm to control the intensity, and use a lens to create an image. By transforming the images into electrochemical impulses, the eyes send the acquired images to the brain through the optic nerve; any disruptions in this process result in visual impairment or eye disorders [2]. Consumption of electronic devices, lack of mobility to medical services, aging, and cultural groups, injuries to the face, etc. are the main causes of blindness [3]. Age-related macular degeneration (AMD), myopia, glaucoma, cataracts (GLC), diabetic retinopathy (DR), and hypertension are the most frequent eye conditions that affect people daily [4].

Diabetes mellitus, the most prevalent reason for vision loss and blindness in older people, is the cause of DR. The impact of intraocular unequal pressure, which harms the optic nerve head and results in vision loss, is the next most frequent cause of vision loss [5, 6]. AMD, which poses a serious threat in developed nations, is the third most prevalent cause of vision loss and blindness. Even if DR and GLC were more common, AMD has become increasingly common in adults over 60, and it has been estimated that AMD alone accounts for 8.7% of blindness globally, primarily in developed nations [7, 8]. One type of eye disease called glaucoma is characterized by cupping and atrophy of the optic nerve head, reduction in visual field, and frequently a rise in intraocular pressure [9]. The retina is greatly impacted by hypertension, and research on retention offers important insights for managing hypertension [10]. Although myopia is not as dangerous this disorder has increased, particularly in youngsters. It can be treated with contact lenses, glasses, and surgical [11].

With the development of computers and computing tools, the state-of-the-art treatments for retinopathy determined by the classifications of these disorders are laborious activities [12]; Experts or physicians use a variety of methods to identify eye conditions; two common ones are fundus photography & optical coherency tomography, which produces a cross-sectional image [13]. The technique of optical coherence tomography has been important in the diagnosis of several medical conditions involving not just the eye along with other organs like the brain. Many researchers have examined the cross-sectional images of the AMD, DR, and GLC affected eye using some DL techniques; but that have the drawbacks of misclassification, overfiring, high complexity, error rate, and high execution time [14, 15].

In this paper, design an effective DL architecture for accurate multiple-eye disease detection. Alao hybrid feature selection technique is employed to select the best features from the extracted results. Additionally, propose a LAAMS-UNet++ model for enhancing the segmentation process by identifying ROI. This research aims to design and develop a

[1*]PhD Scholar, Department of Computer Science and Engineering, The Assam Royal Global University, Guwahati.
*Email: antaram2010@gmail.com
[2]Professor, Department of Computer Science and Engineering, The Assam Royal Global University, Guwahati.
Email: anksjc2002@yahoo.com
[3]Associate Professor, Department of Computer Science and Engineering, National Institute of Technology, Arunachal Pradesh. Email: swarnendu@nitap.ac.in

comprehensive deep learning (DL) framework for the accurate detection and classification of multiple eye diseases using retinal images. The main objectives of the developed model are detailed below,

- To design a novel deep learning model that effectively detects and classifies multiple eye diseases using retinal images.
- To implement Advanced Preprocessing Techniques for outlier detection, image denoising, contrast enhancement, and color normalization.
- To perform image augmentation using geometric transformations to increase the robustness and diversity of the training dataset.
- To develop a Loss and Attention-Augmented Multi-Scale U-Net++ (LAAMS-UNet++) architecture for improved Region of Interest (ROI) identification.
- To employ a Hybrid Seahorse and Beluga Whale (HSHBW) optimization for effective feature selection to enhance classification performance.
- To proposed Dense-CSPDarkNet53 + LinkNet-34 model with an EfficientNetB7 encoder for the classification of eye diseases.

The key contribution of the developed model is as

- An Innovative Pre-Processing technique is developed to improve the data quality of the fundus images.
- Proposed LAAMS-UNet++ model for Sophisticated ROI Identification that integrates multi-scale convolutional blocks, attention gates, residual connections, and dense connections, along with a bottleneck.
- Extracting structural, texture, and deep features using VesselNet, LPQ, GLCM, and ResNet50V2 for detailed retinal image analysis.
- An efficient Feature Selection technique is developed such as the HSHBW model to enhance the classification performance.
- Finally, design an innovative DL model such as Dense-CSPDarkNet53 + LinkNet-34 architecture with an EfficientNetB7 encoder to achieve high accuracy in classifying retinal images.

The study/work/paper is organized as follows: sections 2 and 3 provide relevant works and a problem definition; Section 4 outlines the proposed model process; Section 5 provides the results of experiments and comparisons along with a discussion; and Section 6 concludes with a conclusion.

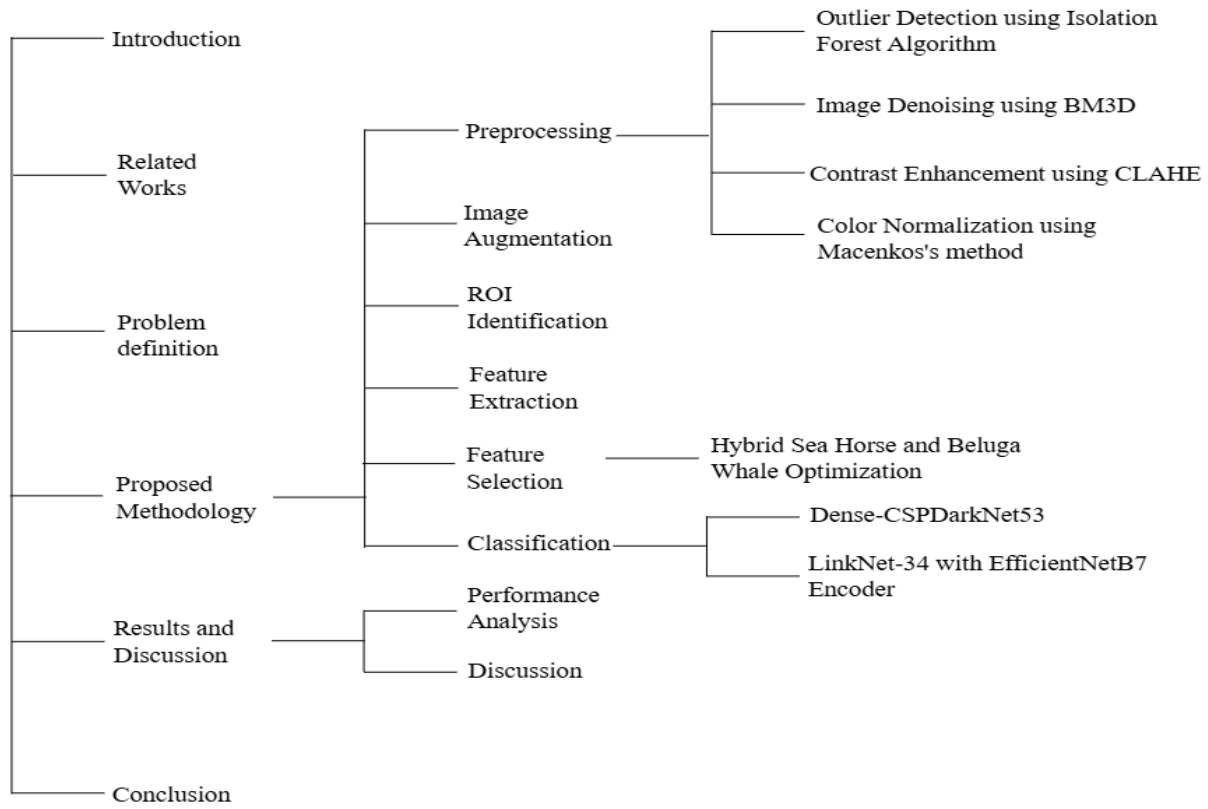


Fig. 1: Organization of the paper

2. RELATED WORKS

Xinpeng, et al [16] created CES-MVSNet, a neural network based on the MVSNet network design, taking into account the particular features of the system. The developed deep neural network (DNN), which can provide 3D reconstruction outcomes with excellent reliability and precision, was given the collected image and camera information. For 3D reconstruction, employ neural networks and the conventional multi-view geometric approach. The differences in the

outcomes of the two approaches were examined. The bionic compound-eye system's effectiveness and dependability in 3D reconstruction have been demonstrated.

Xing, et al [17] create a novel ophthalmic disease screening operation. Initially, create a new dataset, and they combine data from several sources. Then performed studies on fundus multi-disease detection using this dataset. Also suggested a strategy to enhance the training data and imitate the characteristics of various fundus cameras using a technique that utilizes style transfer. Lastly, using the merged dataset, a reliable few-shot learning-based illness screening strategy was built and evaluated against benchmark methods. Tests have demonstrated that our approach can greatly enhance the illness screening model's capability for generalization.

Kaiwen, et al [18] created a decision-analytic Markov framework that included a total of 30 1-year phases for a cohort of people 50 years of age and above. From the point of view of society, we estimated the economic viability and cost-utility of testing programs for several major eye diseases that cause blindness in China, such as pathological myopia, GLC, DR, and AMD. To evaluate the robustness of the primary results, one-way probabilistic and modeled probabilistic sensitivity assessments were performed. It is more affordable to screen for several eye conditions at the same time in China's rural and urban areas. When teleophthalmology and AI are combined, there is a chance to advance equity in eye health.

Nakhim, et al [19] employed optimal residual DNN and efficient image preprocessing methods, including data augmentation, iso-luminance plane contrast-limited adaptive equalization of histograms, and shrinking the area of interest, to improve the performance of multi-categorical categorizing of fundus photographs. By using these to classify three eye disorders using publicly accessible information as of right now, we were able to attain average and peak accuracy rates of 85.79% and 91.16%, respectively. Improved specificity results could warn patients about eye disorders early on and save them from losing their vision

Akanksha, et al [20] proposed one versus-rest approach with a transfer learning method for multiple classes of multilabel opthalmological illness prediction in fundus pictures. The Ocular Disease Intelligent Recognition (ODIR) dataset is used to identify fundus images. These images were fed into the VGG-16 CNN network, which utilized them to binary categorize each disease independently. The suggested framework significantly enhanced its capability of identifying diseases, with baseline accuracy rising from 89% to nearly 91%.

Pamula, et al [21] proposed a hybrid multilayered categorization (HMLC/CNN-VGG19) for classifying eye diseases. Using OCT, the suggested H.M.L.C. collects images from many data sources, such as Kaggle, the U.C.I. repository, etc. The H.M.L.C. uses Python to implement the CNN and VGG-19 architectures over the datasets. The correctness of the classification produced by the experiment is confirmed. Because the H.M.L.C. incorporated sophisticated features from CNN and VGG-19 theories, the classification accuracy is good.

Sungju, developed [22] creates an algorithm that uses diagnostic information to categorize and evaluate 24,000 fundus photos into four groups: cataract, normal, glaucoma, and DR. The approach is further developed into a website that would aid in fundus illness diagnosis. Convolutional neural networks (CNNs), which are experts in image processing, are utilized as learning models. The network was configured using EfficientNet. After tuning using hyperparameter optimization, the created model was eventually implemented as a publicly accessible webpage.

3. PROBLEM DEFINITION

The frequency of eye diseases is on the increase and the effects on millions of people make it a severe global health issue many of these illnesses if not addressed can lead to blindness or other forms of visual impairment. Therapy of all diseases of the eye and their treatment –AMD, cataracts, DR, glaucoma – to a significant degree depends on the early diagnosis and correct definition of a disease. However, expert manual examination is still a part of conventional diagnosis procedures, which is time-consuming, random, and can be affected by possible human imperfections. Moreover, the majority of areas are underserved with access to proper ophthalmologic attention [23]. The following are some of the problems that need a reliable, automated way of predicting the outcome of some eye disorders to be solved: For analyzing the diagnostic imaging results like the retinal images, optical coherence tomography (OCT) pictures, etc and for providing accurate and instant diagnosis this system should incorporate first-in-class deep learning algorithms. As a result, it is aimed at creating a model that would enhance the effectiveness of treatment and would not impose excessive pressure on healthcare networks while assisting medical personnel in making correct decisions [24]. This study/paper/work tries to find and develop a system that will use deep learning solutions in diagnosing various eye ailments to enhance the early detection of such diseases. Classification, segmentation, as well as image augmentation, are examples of computer vision challenges for which machine learning and deep learning have been adopted in solutions.

4. PROPOSED METHODOLOGY

In this phase, developed an efficient DL model for the accurate detection and classification of multiple eye diseases with high performance. Initially, high-resolution retinal images are collected and trained in the Python system. The designed model has six phases such as image preprocessing, image augmentation, ROI identification, feature extraction, feature selection, and classification. In the preprocessing phase, the Isolation Forest Algorithm (IFA) is used for outlier detection, BM3D for image denoising, CLAHE for contrast enhancement, and Macenko's method for color normalization. The preprocessed image data are augmented using geometric transformations. Moreover, propose LAAMS-UNet++ architecture for ROI identification. Feature extraction phase, extract structural, texture, and deep features using VesselNet, LPQ, and GLCM. The best features are selected using the HSHBW approach to improve the classification performance. Finally, classification is performed using a Dense-CSPDarkNet53 + LinkNet-34 model with an EfficientNetB7 encoder.

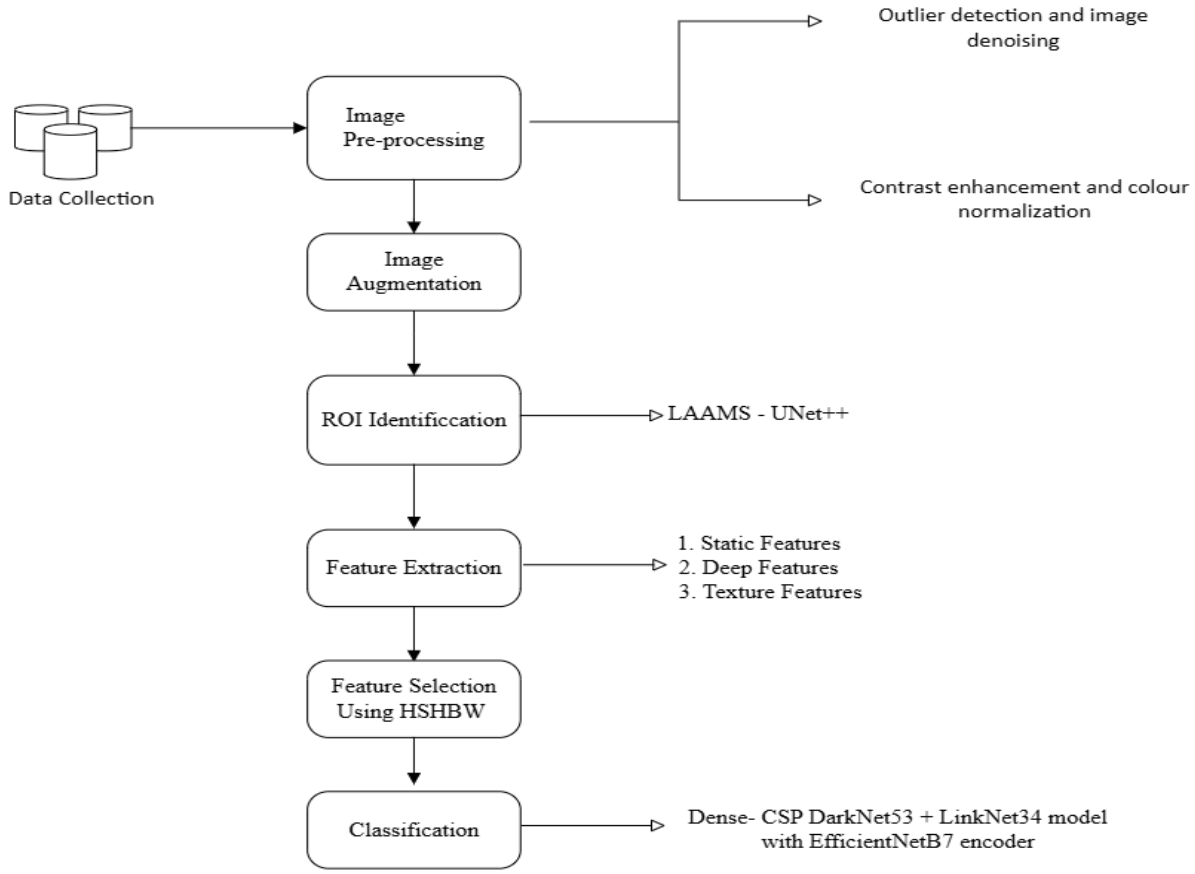


Fig.1 Developed model architecture

The developed model gained better detection and classification results using multi dataset. Also attained high accuracy and less error rate. The proposed model architecture is shown in fig.1.

4.1 Pre-processing

• Outlier Detection using IFA

The IFA is used to detect anomalies, it isolates the observing by selecting random features and splits the value randomly based on the maximum and minimum value from the selected features. The process will be repeated until it forms a tree. The anomalies are different and less, so they are more vulnerable to isolation. In the given dataset k , the isolation forest generates T trees. Then the IS algorithm selects the features s and randomly split the values. Then the outliers are detected in the IS algorithm based on the anomaly score. The anomaly score $A(k,n)$ is determined as eq. (1).

$$A(k,n) = 2^{-\frac{e(h(k))}{c(n)}} \quad (1)$$

Let, $e(h(k))$ is denoted as the median of path length for a point k ($k = \{k_1, k_2, k_3, ...k_n\}$), $c(n)$ is considered the median path length for the unsuccessful searches.

• Image Denoising using BM3D (Block-Matching and 3D Filtering)

The noisy image data l is updated to the BM3D model [25], it has three steps block matching, 3D transformation, and collaborative filtering. Initially, it identifies similar patches using block-matching and 3D transformation and loads them into a 3D array. Also, DCT is updated in the 3D array. Then collaborative filtering achieves a reduction in the changed coefficients. Lastly, the inverse transform is handled to rebuild the patches and combine them back using Eq. (2).

$$\hat{K} = BM3D(l) \quad (2)$$

• Contrast Enhancement using CLAHE

A modified version of adaptive histogram equalization is called contrast-limited adaptive equalization (CLAHE) [26]. This approach derives the transformation function by applying the enhancement function to every neighborhood pixel. This technique redistributes the histogram of a grayscale image by clipping it using the largest value. Grayscale and colorful photographs of both kinds are subjected to CLAHE. Over a noisy image, the "clip limit" technique establishes a

limit. RGB true colors are stored in the LAB color space. The CLAHE Method started by obtaining an ocular picture. then create a clipped histogram and gray level mapping. Since the number of pixels for every gray level in the surrounding region is split equally, the average amount of pixels in every gray level is given by eq. (3.)

$$X_{avg} = \frac{X_{cr-xp} * X_{cr-yp}}{X_{gray}} \quad (3)$$

Let, X_{avg} be the average number of pixels, X_{gray} be the degree of grayscale within the surrounding area, X_{cr-xp} be the number of pixels inside the surrounding area with X direction, and X_{cr-yp} be the number of pixels found in the surrounding area with Y direction.
Next, use eq. (4) to get the actual clip limit.

$$X_{cl} = X_{clip} * X_{avg} \quad (4)$$

To produce a better image, interpolate gray-level mapping. Using a four-pixel cluster, the mapping procedure is applied in this procedure. Each mapping tile then partially overlaps in the image area. An individual pixel is then retrieved, and four mappings are applied to it. Continue over an image to calculate the difference between those outcomes and obtain improved pixels.

• Color Normalization using Macenko's method

Macenko's method [27] is utilized for color normalization in histopathology images. Its goal is to attain color distribution normalization according to stain separation. Eq. (5) is used to convert the image to optical density (OD) space.

$$o_d = -\log\left(\frac{i}{i_0}\right) \quad (5)$$

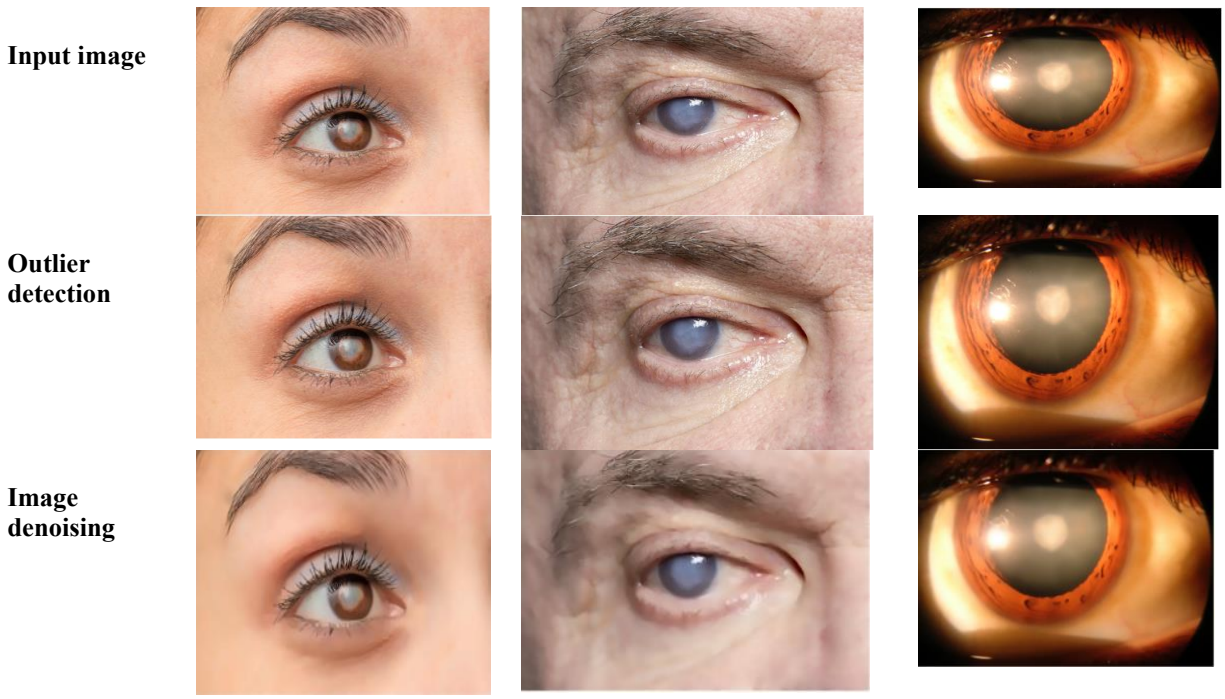
Let, i be the image intensity and i_0 be the standard intensity of white. Next, use eq. (6) to conduct singular value decomposition (SVD) on the OD matrix to determine the stain matrices.

$$o_d = u\sum v^t \quad (6)$$

Apply eq. (7) to standardize the stain intensities to the desired distribution.

$$i_{norm} = i_0 . \exp\left(-\sum_{i=1}^{p} h_n s_n\right) \quad (7)$$

where h_n is denoted as a concentration matrix, and s_n is denoted as a stain matrix. The preprocessed image is updated to the ROI identification phase for segmenting eye affected area based on the region of interest (ROI). The sample preprocessed results of the developed model are shown in fig.2.



4.2 Image Augmentation

Geometric transformations are simple image processing and computer vision operations that enable change in the spatial arrangement of images. These operations include rotation, flipping, scaling, and cropping each of which provides the appearing view at different orientations and distances.

Rotation: Rotation is concerned with rotating an image at a predetermined angle of rotation about a fixed center of rotation normally the center of the image to be rotated.

Flip: Image flipping is a process of producing a mirror image of an image along a given axis which can be either column-wise or row-wise.

Scaling: Scaling is a technique by which the size of an image is changed either uniformly or non-uniformly using a scaling factor.

Cropping: Cropping involves choosing a small region within the image and leaving the rest of the area blank.

4.3 ROI identification

The ROI identification phase is enhanced using the novel Loss and Attention-Augmented Multi-Scale U-Net++ (LAAMS-UNet++) model. It accurately segments the retinal image and enhances the classification performance. The detailed description of LAAMS-UNet++) a model with a layer-by-layer explanation is detailed below,

- **Input Layer**

Input high-resolution retinal images: In the input layer update the high-resolution retinal images for accurate segmentation of features like optic disc, lesions, and blood vessels.

- **Encoder (Downsampling Path)**

Multi-Scale Convolutional Blocks: In the encoder block, multiple convolutional filters are employed with dissimilar kernel sizes for gathering features at several scales. It will permit the network to learn coarse and fine information from the image using eq. (8)

$$O = \sum_{i=1}^{p} F_i \left(P_i * Z \right) \tag{8}$$

Let, P_i be the different sizes convolutional kernels, F_i be the activation functions, and Z be the input feature map.

Residual Connections: The residual connections are uploaded to a convolutional block input to its output. It is useful for gradient flow and permits the system to analyze more deep layers for enhancing the efficiency using Eq. (9)

$$O = Z + f\left(Z\right) \tag{9}$$

Let, $f\left(Z\right)$ be the convolutional operation.

Attention Gates: Following every convolutional block, attention gates are implemented using Eq. (10) to suppress less significant features and increase the focus on pertinent ones.

$$A_o = \delta\left(w_F * f\left(Z\right) + w_G * g\left(Z\right)\right) \tag{10}$$

Let, δ be the sigmoid function, w_F and w_G be the learnable weights, $f\left(Z\right)$ and $g\left(Z\right)$ be the feature maps.

Dense Connections: Dense connections facilitate feature reuse and effective flows of gradients using Eq. (11) by guaranteeing that each layer collects input from all previous layers.

$$O = \left[Z_0, Z_1, ..., Z_{n-1}\right] \tag{11}$$

All Z_i represents the output of the i^{th} layer.

• **Bottleneck**

Squeeze-and-Excitation (SE) Blocks: By recalibrating channel-wise feature reactions, SE blocks improve the network's capacity to use Eqs. (12) and (13), to describe channel relationships.

$$S = \delta\left(w_2 \operatorname{Re} lu\left(w_1 Pool_{avg}\left(Z\right)\right)\right) \quad (12)$$

$$O = Z.S \quad (13)$$

Let, w_1, and w_2 be the weights, S be the scale, and $Pool_{avg}$ be the global average pooling

Atrous Spatial Pyramid Pooling (ASPP): Equation (14) is used by ASPP to represent multi-scale spatial settings through the use of atrous convolutions with varying compression rates.

$$O = \sum_{R \in r} At_{conv}\left(Z, R\right) \quad (14)$$

Let, R be a group of dilation rates, and At_{conv} be the atrous convolution.

Hybrid Dilated Convolution (HDC): To prevent gridding artifacts and efficiently record spatial settings, HDC combines a series of atrous convolutions with different dilation rates.

• **Decoder (Up sampling Path)**

Transposed Convolutional Layers: By applying Eq. (15), such layers expand the feature maps and restore their spatial resolution.

$$O = Trans_{conv}\left(Z\right) \quad (15)$$

Multi-Scale Feature Fusion: By concatenating comparable encoder and decoder feature maps using Equation (16), this procedure merges features from various scales.

$$O = \left[En_{fe}, De_{fe}\right] \quad (16)$$

Let, En_{fe} be the encoder features, and De_{fe} be the decoder features.

Attention Mechanisms and skip connection: To concentrate attention on important areas of the retinal images, attention processes are performed to the fused features. By efficiently combining encoder and decoder features, attention-enhanced skip connections guarantee that significant features are kept when upsampling with eq. (17).

$$O = A_t\left[En_{fe}, De_{fe}\right] \quad (17)$$

Let, A_t be the attention mechanism

Loss Function: two kinds of loss functions are performed such as dice loss and focal loss. Using equation (18), dice loss calculates the overlap between the ground truth and anticipated masks.

$$D_L = 1 - \frac{2\left|p \cap g\right|}{\left|p\right| + \left|g\right|} \quad (18)$$

where p and g are denoted as predicted and ground truth masks. The Focal Loss uses eq. (19) to down weight simple instances and concentrate on difficult ones to fix class imbalance.

$$F_L = -\beta\left(1 - P\left(t\right)^{\alpha} \log\left(P\left(t\right)\right)\right) \quad (19)$$

where $P\left(t\right)$ is denoted as predicted probability, β, and α are considered as hyperparameters.

• **Regularization**

The regularization is performed using Dropout and Spatial Dropout**.** These dropouts prevent overfitting during the training phase by arbitrarily changing a portion of the input units to zero.

Dropout: Units are dropped randomly throughout the feature channels using eq. (20).

$$Z_{drop} = Z.B_r\left(P\right) \quad (20)$$

Spatial Dropout: Equation (21) is used to drop entire channels rather than single units.

$$Z_{sp\_drop} = Z.B_r\left(P_{ch}\right) \quad (21)$$

Where, B_r is denoted as Bernoulli. Finally, develop a model that accurately segments the retina images based on the identification of ROI. Then feature extraction process is started using the segmented results for extracting statistical, deep, and optical disc features. The sample segmented results are shown in fig.3.

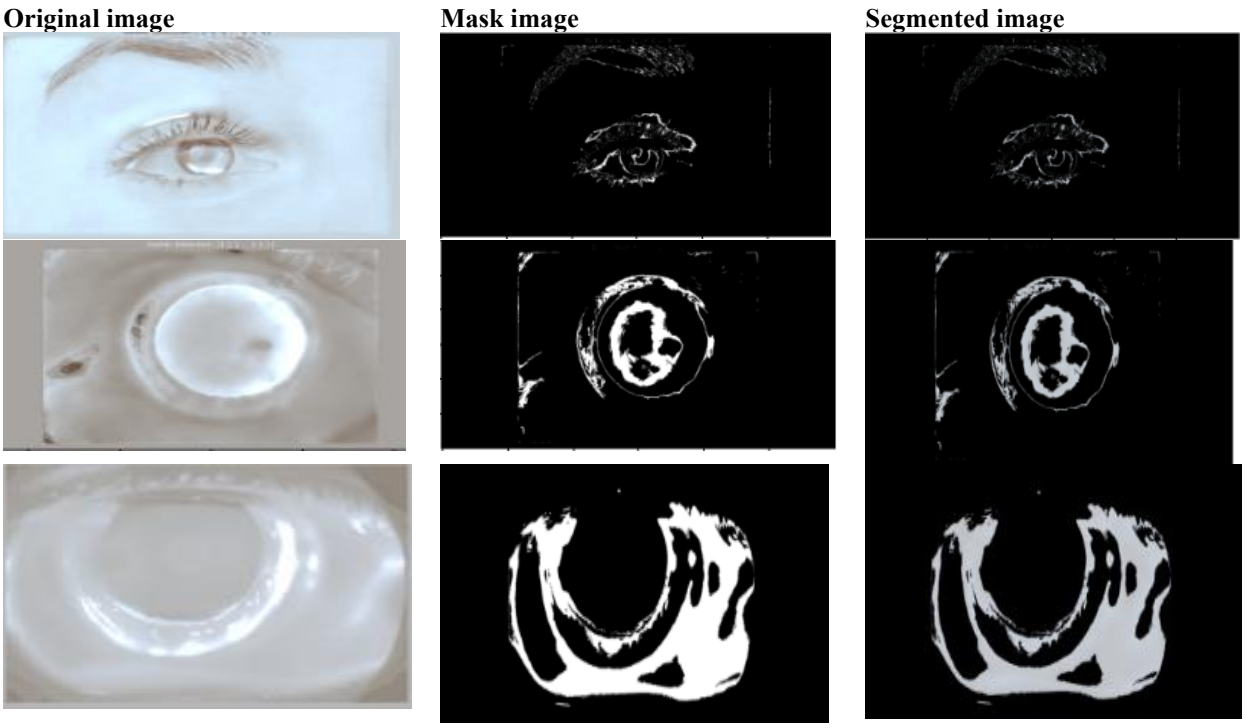


Fig.3 Sample segmented results

4.4 Feature Extraction

The structural, texture, and deep features are extracted in this phase using various techniques. The detailed description of the extracted features with techniques is detailed below,

4.4.1 Structural Features

- **Retinal Vessel Analysis:**

VesselNet, which is employed to perform vessel segmentation, extracts features associated with blood vessel characteristics, including density and tortuosity. The VesselNet is a neural network that has been developed for vessel segmentation in the fundus of the retina. It can employ convolutional layers to identify various patterns of the blood vessels in the retinal images. First, low-level processing is done to improve the contrast of the retinal images so that blood vessels stand out. Then apply VesselNet to segment the blood vessels from the retinal image R_i using the following eq. (22). The blood vessels are the essential factor to be diagnosed or analyzed for any complications in the eye and the segmented image.

$$s(a,b) = V_N\left(R_i(a,b)\right) \tag{22}$$

Let $s(a,b)$ be the segmented output where (a,b) are the coordinates of the pixels, and V_N is denoted as VesselNet. Then the features such as vessel density ($v(d)$) and tortuosity ($\vec{T}$) are calculated using eq. (23) and (24).

$$v(d) = \frac{\sum_{a,b} s(a,b)}{w \times h} \tag{23}$$

$$\vec{T} = \frac{T_l(v)}{\Sigma} \tag{24}$$

Let, w and h be the width and height of the image, $T_l(v)$ is denoted as the total length, and Σ is denoted as endpoints Euclidean distance.

- **Optic Disc and Cup:**

Cup-to-Disc Ratio: Calculate the area of the cup with the disc area which is an important parameter for glaucoma identification. It was calculated with the help of the following formula: The cup area/cup disc area. The disc segmentation and cup segmentation is done to obtain the disc and cup region separately from the disc.

Disc Area: The total area of the optic disc and is measured using eq. (25)

$$a_D = \sum_{(a,b)\in D} 1 \tag{25}$$

Cup Area: Measure the total area of the optic cup and is measured using eq. (26).

$$a_{C} = \sum_{(a,b)\in C} 1 \quad (26)$$

Let, D, and C are denoted as disc and cup.

4.4.2 Texture Features

• Local Phase Quantization (LPQ)

It acquires texture data by detecting the phase of the local frequency components of the image. The LPQ algorithm used to analyze the local phase information employs the short-term Fourier transform (STFT), and quantizes the phase values. As a result, apply STFT to the image R_i to obtain the local frequency components with the help of eq. (27).

$$F(m,n;a,b) = \sum_{u=-U}^{U} \sum_{v=-V}^{V} R_i(a+u,b+v) e^{-j2\pi(ma+nb)} \quad (27)$$

Quantize the phase of these components using eq. (28)

$$\phi(m,n;a,b) = \arg\left(f(m,n;a,b)\right) \quad (28)$$

They create a histogram of the quantized phases using eq. (29).

$$h_{lpq} = H\left(\phi(m,n;a,b)\right) \quad (29)$$

Let, H is denoted as a histogram.

• Grey Level Co-occurrence Matrix (GLCM)

GLCM describes the probabilities of getting two pixels with specific intensity values and spatial distance between them in the picture. The features comprising contrast, correlation, energy, and homogeneity are calculated using GLCM that are established using eq. (30)-(33).

$$C_n = \sum_{i,j} h(i,j)(i-j)^2 \quad (30)$$

$$C_r = \sum_{i,j} \frac{(i-\gamma_i)(j-\gamma_j)h(i,j)}{\lambda_i \lambda_j} \quad (31)$$

$$E_n = \sum_{i,j} h(i,j)^2 \quad (32)$$

$$H_g = \sum_{i,j} \frac{h(i,j)}{1+|i-j|} \quad (33)$$

1.4.3 Deep Features

All deep features are extracted using ResNet50V2 [28]. Among them is the ResNet50V2, an enhanced version of ResNet50, which has 50 layers, including the convolutional layers, batch normalization layers, ReLU activation layers, and other residual blocks. These residual blocks employ elements known as skip connections wherein one or more layers are skipped and the input to the block is added to the output of the block.

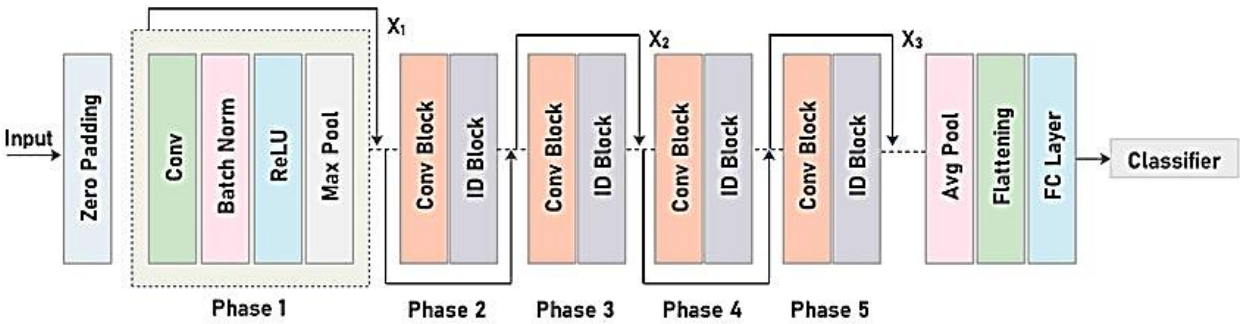


Fig.4 ResNet50V2 architecture

This technique goes a long way in addressing the degradation problem, which is when deeper architectures perform worse than shallow ones due to the challenges in training. In the context of the present work on retinal image analysis, ResNet50V2 can be employed to extract high abstract features from HR retinal images. These deep features encompass low-level attributes of images including textures, edges, and all other profound structural characteristics of an image which are very essential for analyzing images and making relevant classifications. The process includes passing the preprocessed

retinal images through the ResNet50V2 network, probably up to an intermediate layer that captures the features. These are known as the deep features representing the images and they are structured depending on the hierarchical model that is used together with the low-level features such as edges and textures and the context of the images. The ResNet50V2 architecture is depicted below in Fig 4.

4.5 Feature selection

The best qualities of the dataset are selected using hybrid sea horse and beluga whale optimization (HSHBW). It is the outcome of merging Sea Horse Optimization (SHO) [30] and Beluga Whale Optimization (BWO) [29]. The SHO algorithm simulates the movement, hunting behavior, and reproduction of sea horses. Three behaviors comprise the fundamental features of SHO. To improve the improvements of the SHO algorithm, the global and locaters approaches are tailored to the motion and hunting behaviors, respectively. Here, the seahorse's spiraling approach to its greatest advantage is portrayed by Lévy flying. Equation (34), describing the mathematical link, is given in Eq. (34).

$$Z_{new}^{1}\left(t+1\right)=Z_i\left(t\right)+Levy\left(\delta\right)\left(\left(Z_{elite}\left(t\right)-Z_i\left(t\right)\right)\times a\times b\times c+Z_{elite}\left(t\right)\right) \quad (34)$$

Let, $a,b,$ and c are represented in three dimensions as coordinate vectors. $\left(a,b,c\right)$ according to the swirling motion. Eq. (35)is employed in the computation of wave-based Brownian motion.

$$\sigma=\left(\frac{\Gamma\left(1+\delta\right)\times\sin\left(\frac{\pi\delta}{2}\right)}{\Gamma\left(\frac{1+\delta}{2}\right)\times\delta\times2^{\left(\frac{\delta-1}{2}\right)}}\right) \quad (35)$$

Equation (36), which gives the movement of the sea horse's unit size, is mimicked using Brownian motion to the left of the $\vec{r_1}$ cut-off point to help with understanding the SHO search space.

$$Z_{new}^{1}\left(t+1\right)=Z_i\left(t\right)+rand\times c\times\alpha_t\times\left(Z_i\left(t\right)-\alpha_t\times Z_{elite}\right) \quad (36)$$

Let, c is regarded as a constant coefficient, α_t is shown for Brownian motion through a random walk coefficient. The sea horse's new location at each repetition t can be found by integrating these two conditions using Equation (37).

$$Z_{new}^{1}\left(t+1\right)=\begin{cases}Z_i\left(t\right)+Levy\left(\delta\right)\left(\left(Z_{elite}\left(t\right)-Z_i\left(t\right)\right)\times a\times b\times c+Z_{elite}\left(t\right)\right) & \vec{r_1}>0\\ Z_i\left(t\right)+rand\times c\times\alpha_t\times\left(Z_i\left(t\right)-\alpha_t\times Z_{elite}\right) & \vec{r_1}\le0\end{cases} \quad (37)$$

The efficiency of beluga whale exploitation is then adjusted to improve the behavior of seahorse predation. Equation (38) displays the hybrid efficiency of the mathematical algorithm.

$$Z_{new}^{2}\left(t+1\right)=\begin{cases}R_3Z_{best}^{t}-R_4Z_i^{t}+c_1.levy.\left(Z_r^{t}-Z_i^{t}\right) & \vec{r_2}>0.1\\ \left(1-\eta\right)\times\left(Z_{new}^{1}\left(t\right)-rand\times Z_{elite}\right)+\eta\times Z_{new}^{1}\left(t\right) & \vec{r_2}\le0.1\end{cases} \quad (38)$$

Let, $Z_{new}^{1}\left(t\right)$ displays the current location of the seahorse at that time of t, Z_i^{t}, and Z_r^{t} are denoted as the current position, Z_{best}^{t} is denoted as the best position of the beluga whales, R_3, and R_4 are denoted as random numbers, c_1 is the unpredictable jump strength, which gauges how hard Levy is flying, and $\vec{r_2}$ is denoted as a random number [0, 1]. It's used to adjust the seahorse's predation phase duration, which decreases linearly with each repetition. Rather, it suggests that the prey is moving faster than the seahorse was able to track it, which lets the prey get away and prevents the seahorse from catching it. The developed model selects the best features from the dataset using the HSHBW method. It enhances the detection and classification capabilities of the multi-eye disease detection system. Subsequently, the chosen characteristics are updated for the classification stage to categorize several eye diseases. Algorithm 1 describes the HSHBW algorithm.

Algorithm 1 . Feature Selection Process Using HSHBW

1. Input: The extracted features
2. Output: Best characteristics from the dataset
3. Begin
4. Initialize Z_i
5. Compute fitness value// all seahorse
6. Determine Z_{elite} //best seahorse
7. ***While*** *(t < T)* ***do***

*8. **If (** $\overline{r_1}$ = random > 0) **then** // movement*
9. Set constant parameter values
10. Execute Brownian motion using eqn. (35)
11. Sea horse position updated using eqn.34 and 36
*12. **else***
13. Update position using eqn. (37).
*14. **End if***
15. Update beluga whale exploitation behavior in eq. (38) //Combine beluga whale optimization
16. Update the new position using eqn. (38) //new hybrid model
*17. **if** (* $\overline{r_2}$ *> 0.1)*
18. Select best features
*19. **else***
20. Continue searching
*21. **Endif***
22. Enhance prediction accuracy

1.5 Classification

The multi-disease eye classification is performed using two techniques such as Dense-CSPDarkNet53 [31] and LinkNet-34 model with EfficientNetB7 encoder [32]. Dense-CSPDarkNet53 integrates dense connections to enhance gradient flow and feature reusing, while Cross Stage Partial connections maximize performance by combining the best aspects of DenseNet and CSPDarkNet. This network recognizes complicated patterns in the input images and serves as the foundation for feature extraction. Next, LinkNet-34 is used as the decoder, which reconstructs the classification outputs by essentially merging these characteristics. This framework is further improved by the EfficientNetB7 encoder, which is renowned for its exceptional accuracy and efficiency. It does this by precisely scaling its structure to optimize the method of extracting features.

○ **Dense-CSPDarkNet53**

The three elements of Dense-CSPDarkNet53 are the head for forecasting, the neck for semantic representations of extracted features, and the backbone for feature extraction. The YOLOv4 algorithm minimizes the feature mappings in the neural network throughout eye illness detection. In the suggested architecture, each layer is linked to other levels in feed-forward mode, utilizing the DenseNet architecture to conserve significant feature maps and more effectively reuse the crucial feature information. One primary benefit of the DenseNet block is that the *n*-th layer can obtain the necessary feature data Z_n from every earlier layer $Z_0, Z_1, ...Z_{n-1}$ inputs, which have the formula eq. (39)

$$Z_n = h_n\left[Z_0, Z_1, ...Z_{n-1}\right] \quad (39)$$

Let, h_n be the function of spliced feature maps for layer n ; $\left[Z_0, Z_1, ...Z_{n-1}\right]$ is the layer feature map $Z_0, Z_1, ...Z_{n-1}$. According to the proposed framework, the primary CSPDarknet53 Cross-Stage Partial (CSP) systems convolution blocks CSP1, CSP2, CSP8, CSP8, and CSP4 were subsequently changed to D1-CSP1, D2-CSP2, D3-CSP4, D4-CSP4, and D5-CSP2 through lowering convolution blocks and incorporating dense connection sections to boost computational speed while decreasing the number of unnecessary feature processes. The suggested model makes use of the Mish activation operation, which is written as eq. (40), to improve network gradient flow equilibrium and aid in the development of more expressive attributes in the detection framework.

$$F\left(z\right) = z.\tanh\left(S_p\left(z\right) = z.\tanh\left(In\left(1+e^z\right)\right)\right) \quad (40)$$

Let, S_p be the soft plus. Because of the smoothness quality, it is also impartial about the establishment of weights and learning rate. An SPP block is tightly coupled with the final residual block (D5-CSP2) to improve the response field and distinguish significant context information during object detection. The suggested model used a maximum pool to a sliding kernel of size 5×5, 9×9 and 13×13 modifying the SPP to preserve the output spatial dimension. The backbone's receptive field is successfully increased by a reasonably large max-pooling. Additionally, an improved PANet has been incorporated into the neck portion of the suggested network framework to maintain fine-grain localized data. This shortens the path of low and high merging for a multi-scale feature pyramidal map. Furthermore, class label smoothness and drop block regularization were utilized to improve a dataset's adaptability and to develop spatially discriminating features, respectively. The detecting head was the initial YOLOv3 head. Using the uploaded picture size of $512\times512\times3$, At the detecting head, the proposed approach has predicted bounding boxes on three distinct scales: $64\times64\times24$, $32\times32\times24$, and $16\times16\times24$.

○ **LinkNet-34 model with EfficientNetB7 encoder**

The proposed system for the detection and categorization of numerous eye diseases is LinkNet-34 with EfcinetNetB7. It creates a highly effective and computationally feasible algorithm for image classification by fusing the performance of

EfcientNetB7 with the comprehensive classification powers of LinkNet-34. Deep CNN is applied in the LinkNet-34 architecture to gather high-level data from the input image using EfcientNetB7 as the encoder. EfcientNetB7's successful collection of important picture features is made possible by its efficient framework. The MBConv (Mobile Inverted Residual Bottleneck) block serves as the foundational building component for the EfcientNetB7 encoder. A deep convolution and point-wise convolution are the two successive procedures that make up a depthwise separated convolution. When employed on the input tensor Z , the depth-by-depth separable convolution method is mathematically given by Eq. (41).

$$D_w conv(Z) = P_w conv\left(D_w conv1(Z)\right) \quad (41)$$

Let, $D_w conv1$ describes the action of depthwise convolution and $P_w conv$ shows the point-wise convolution process. Responses for channel-specific features are modified by the SE module. Applying the input tensor Z , Eq. (42) describes the SE block approach.

$$S_e(Z) = \delta\left(F_c\left(G_{avgpool}(Z)\right)\right) \quad (42)$$

The mean value for every channel is determined by the GlobalAveragePooling ($G_{avgpool}$) function. The activation value is the sigmoid function, while the Fully Connected layer comprises a dense layer. Multiple MBConv blocks with varying expansion percentages, kernel sizes, and output channel sizes are stacked to create the EfcientNetB7 encoder. the encoder's general functioning $e(Z)$ for Z be used to represent EfcientNetB7. Equation (43) illustrates the series of MBConv blocks

$$e(Z) = mbConv_n\left(mbConv_{n-1}\left(.....\left(mbConv1_n\right)\right)\right) \quad (43)$$

Let, $mbConv_i$ be the i^{th} encoder's MBConv block. The high-level features that the EfcientNetB7 encoder $e(Z)$ extracted are processed by the decoder portion of LinkNet-34. The "34" in LinkNet-34 may allude to the total amount of convolutional layers, suggesting that it has 34 layers. It accomplishes the challenge of obtaining fine-grained information for accurate categorizing while increasing the sampling of the characteristics to the actual image quality. Skip links are incorporated throughout the encoder and decoder layers in LinkNet-34. The links help the model preserve and reuse low-level encoder input, which improves classification accuracy and preserves spatial information—especially for small and complicated objects. Each pixel in the final LinkNet-34 model is assigned to a certain group. A sophisticated description of EfcientNetB7 and the decoder's ability to capture complex features improve the system's classification accuracy. The decoder works by applying more convolutional layers after upsampling the features taken from the EfcientNetB7 encoder. the outcome of encoding with the EfcientB7 encoder, represented as $e(Z)$ and determined by Eq. (3), is subsequently put into the LinkNet-34 strategy decoder. Eq. (44) illustrates how this decoder interprets the input and produces the final classification.

$$D\left(e(Z)\right) = Conv_1\left(up_s\left(Conv_2\left(B_n\left(\mathrm{Re}lu\left(e(Z)\right)\right)\right)\right)\right) \quad (44)$$

Let, $Conv_1$ and $Conv_2$ be the first and later layers of convolution, D is represented as a decoder, up_s indicates the upsampling process, B_n is denoted as batch normalization, and $\mathrm{Re}lu$ denotes activation functions. LinkNet uses skip connections, which are identified as $S(Z)$, which links the encoder and decoder's matching levels. Equation (45) displays the result of the decoder using skip connections.

$$F(O) = D\left(e(Z)\right) + S(Z) \quad (45)$$

This equation, which uses EfcientB7 as the encoder, provides a fundamental mathematical representation of the LinkNet-34. When combined, these elements provide a strong, effective, and exact model that can be used for high-resolution image interpretation tasks, offering a reliable solution for accurate and detailed picture classification.

5. RESULTS AND DISCUSSION

The effectiveness of the produced model is verified using performance metrics such as accuracy, sensitivity, specificity, precision, recall, and so on. The Python tool is then used to implement the proposed technique. When the model is built with a dataset, two different learning rates—let's say 70 and 80—are used for the validation process. By comparing the developed model with various deep learning techniques, its efficacy is illustrated.

5.1 Dataset description

This study uses five different types of datasets, each of which is explained below.

- **Indian Diabetic Retinopathy Image Dataset (**IDRiD) **[35]:** The initial strategy database that is comprehensive of the Indian population is the IDRiD. Furthermore, it is the only dataset that includes both normal retinal components that have

been labeled at the pixel level and typical DR lesions. For every image, this dataset offers details on the degree of diabetic macular edema and DR.

- **Glaucoma Detection [36]:** A class of diseases known as glaucoma affects the optic nerve, the function of which is essential for clear vision. Frequently, this injury is brought on by unusually high intraocular pressure. One of the main causes of blindness in adults over 60 is glaucoma. This data set includes eye pictures and ocular CT scans.
- **Cataract dataset [37]:** This dataset was produced with real-world deep-learning applications in the medical industry in mind. The only cataract datasets now available are medical reports rather than visual photographs; yet, those datasets do not help develop applications such as cataract detectors using images of the eyes.
- **Diabetic Retinopathy Dataset [38]:** This dataset includes 2750 DR pictures along with 5 labels.
- **Eye Diseases Classification [39**]: The collection includes retinal scans from four classes: normal, DR, cataract, and glaucoma. There are roughly 1000 photos in each class. These photos were gathered from several sources, including HRF, Oculur recognition, IDRiD, and others.

5.2 Performance analysis

The suggested technique's obtained results are verified against current methods like DNN [16], VGG-16 [20], FRCNN [33] and DCNN [34]. The effectiveness of the created model was assessed using the following performance metrics: MCC, NPV, FPR, FNR, accuracy, precision, F-measure, sensitivity, and specificity. The performance results of the developed model are validated with two learning rates 70 and 80, which is detailed in the table.1 and table.2.

Table 1 Performance comparison with 70 learning rates

Model	Accuracy	Precision	F-measure	Specificity	Sensitivity	NPV	MCC	FPR	FNR
Proposed	97.352	97.094	96.851	96.608	96.085	97.582	97.285	0.0282	0.0174
DNN [16]	95.842	95.081	94.087	94.025	94.418	95.21	95.108	0.0518	0.0472
VGG-16 [20]	95.515	95.836	94.935	94.486	94.105	95.841	94.086	0.0621	0.0521
FRCNN [33]	94.043	93.084	92.034	92.058	93.409	94.086	94.117	0.0672	0.0521
DCNN [34]	94.281	94.852	93.342	93.894	93.086	93.892	94.843	0.0618	0.05821

Table 2 Performance comparison with 70 learning rates

Model	Accuracy	Precision	F-measure	Specificity	Sensitivity	NPV	MCC	FPR	FNR
Proposed	98.521	98.812	97.748	97.553	97.102	97.893	97.973	0.0182	0.00974
DNN [16]	96.663	96.348	95.559	95.396	95.537	96.951	95.968	0.0446	0.0363
VGG-16 [20]	96.086	96.064	95.539	95.553	95.325	96.473	95.667	0.0548	0.0449
FRCNN [33]	95.043	95.544	93.348	93.914	94.931	95.553	94.787	0.0529	0.04189
DCNN [34]	95.594	95.089	94.896	94.983	93.994	94.603	95.082	0.0593	0.04409

Evaluation metrics for various models in two distinct data split scenarios—70/30 and 80/20—are shown in Fig. 5 through 13. Accuracy, precision, F-measure, specificity, sensitivity, NPV, MCC, FPR, and FNR are the metrics used to evaluate each model.

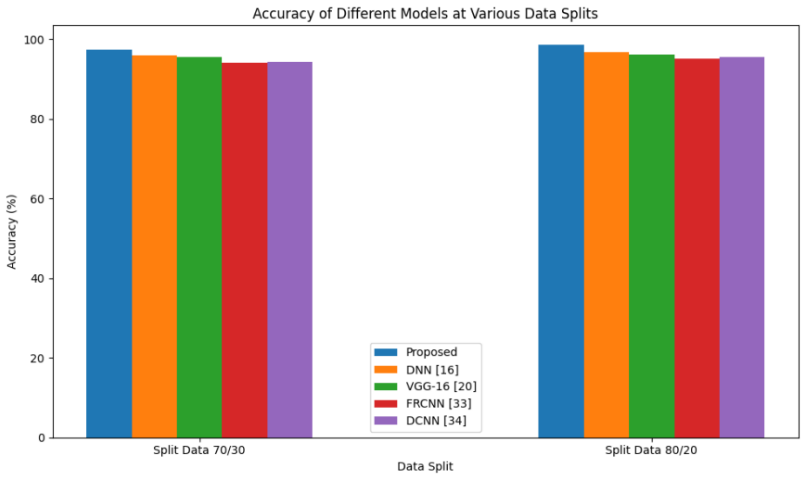


Fig.5 Accuracy comparison

As the percentage of accurate forecasts among all the predictions produced, accuracy quantifies how accurate the model's predictions are overall. The "Proposed" model obtains an accuracy of 97.352% in the 70/30 split, suggesting a high percentage of accurate classifications. The accuracy rises to 98.521% in the 80/20 split, indicating better results in the validation dataset.

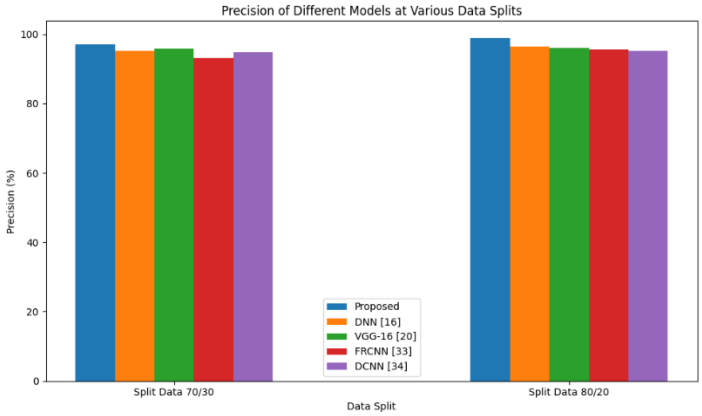


Fig.6 Precision comparison

Out of all the favorable forecasts the model makes, precision indicates the percentage of accurate positive predictions. With a precision of 97.094% in the 70/30 split, the "Proposed" approach shows a high percentage of accurately predicted positive events. The precision rises to 98.812% in the 80/20 split, indicating improved true identification accuracy.

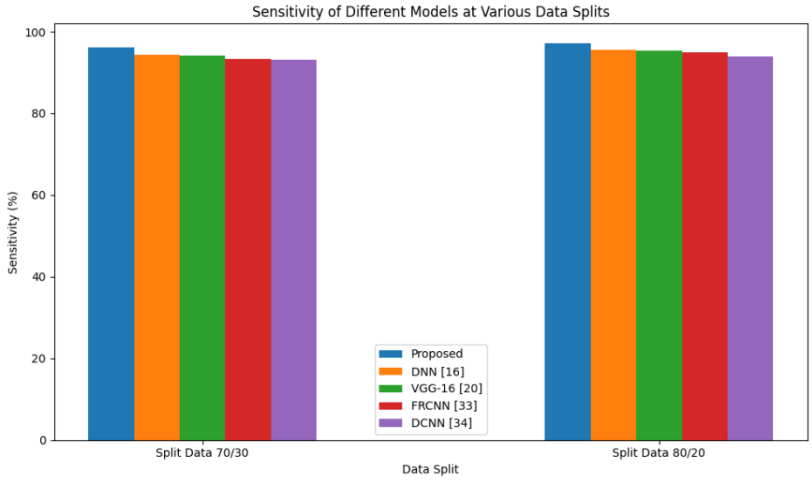


Fig.7 Sensitivity comparison

Sensitivity quantifies the percentage of real positive examples that the model accurately detects. The "Proposed" model gets a sensitivity of 96.085% in the 70/30 split, demonstrating its capacity to accurately identify true positives. Sensitivity in the 80/20 split is still high at 97.102%, indicating strong performance in accurately identifying positive cases.

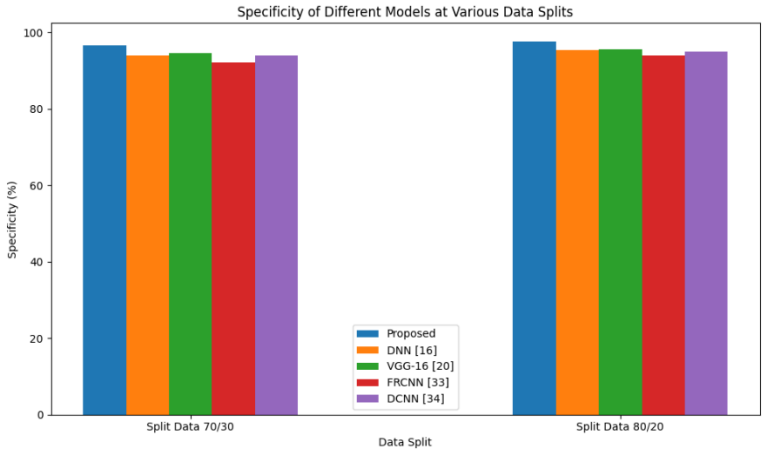


Fig.8 Specificity comparison

The percentage of real negative cases that the model accurately detects is known as specificity. The "Proposed" model obtains a specificity of 96.608% in the 70/30 split, demonstrating its capacity to accurately identify genuine negatives. Specificity is still high in the 80/20 split at 97.553%, demonstrating the model's ability to accurately identify negative events.

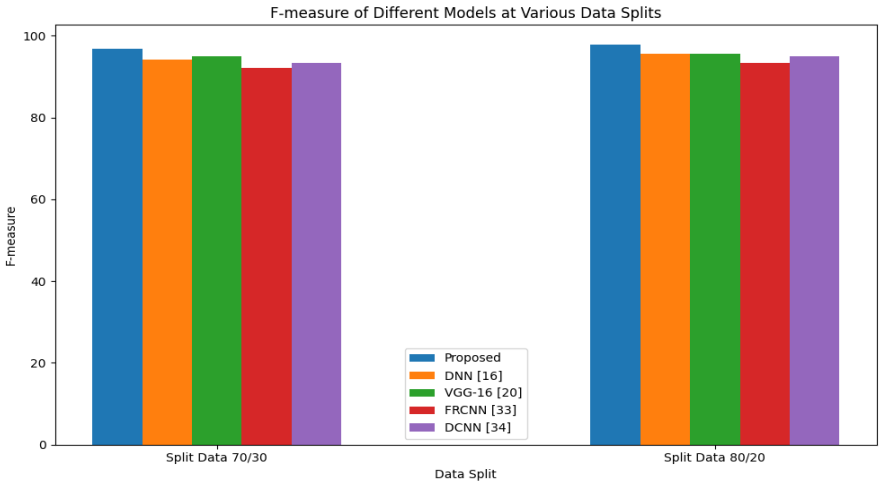


Fig.9 F-measure comparison

The F-measure, which offers a single metric to assess the efficacy of a model in binary categorization tasks, is the harmonic mean of both accuracy and recall (sensitivity). The “Proposed” model obtains an F-measure of 96.851% in the 70/30 split, indicating a performance that strikes a compromise between sensitivity and precision. The 80/20 split exhibits a high F-measure of 97.748%, suggesting a constant efficacy in attaining elevated recall and precision.

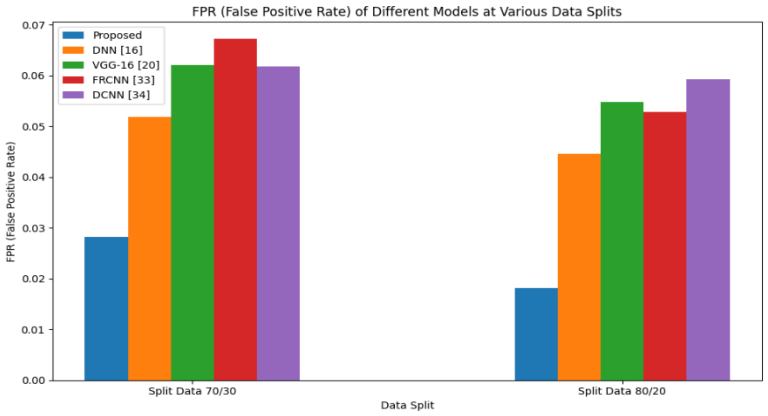


Fig.10 FPR comparison

The percentage of negative occurrences that are mistakenly projected as positive is known as the FPR. The "Proposed" approach has an FPR of 0.0282 in the 70/30 split, indicating a low rate of misclassifying negatives as positives. The FPR significantly drops to 0.0182 in the 80/20 split, suggesting better performance in reducing false positive predictions.

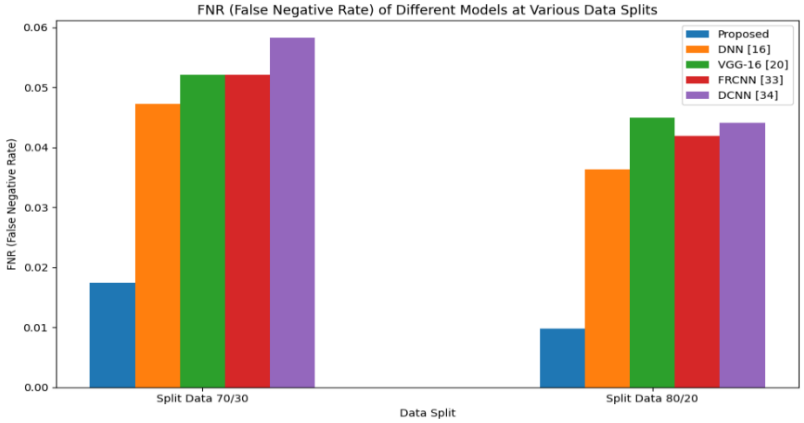


Fig.11 FNR comparison

The percentage of positive cases that are mistakenly projected as negative is known as the FNR. The "Proposed" model obtains an FNR of 0.0174 in the 70/30 split, showing a low rate of misclassifying positives as negatives. Comparably, the FNR drops to 0.00974 in the 80/20 split, indicating increased precision in properly detecting positive events.

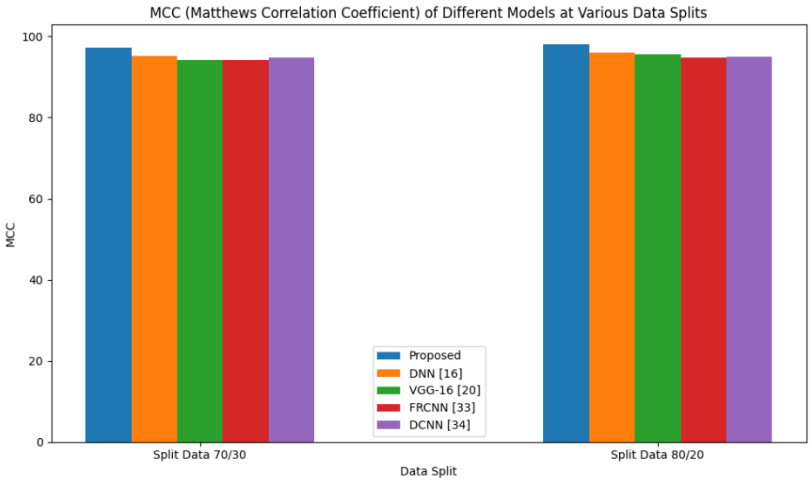


Fig.12 MCC comparison

A metric used to assess the reliability of binary classifications that accounts for false positives, false negatives, true positives, and true negatives is the MCC. The "Proposed" model obtains an MCC of 97.285% in the 70/30 split, demonstrating a very strong connection between the actual and predicted categories. Similarly, the MCC increases to 97.973% in the 80/20 split, demonstrating the model's resilience in sustaining precise predictions across various datasets.

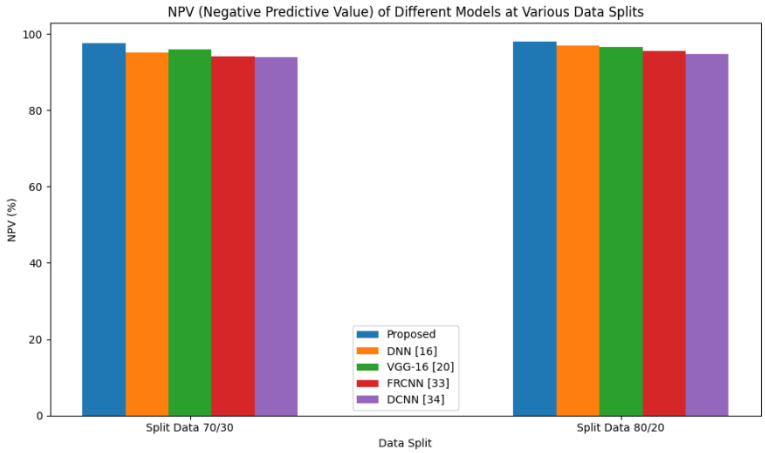


Fig.13 NPV comparison

The NPV quantifies the likelihood of a negative forecast being correct. With an NPV of 97.582% in the 70/30 split, the "Proposed" model shows that it can accurately identify genuine negatives. The NPV in the 80/20 split is still high at 97.893%, demonstrating steady performance in accurately forecasting unfavorable outcomes.

In comparison to the other models assessed, the "Proposed" model consistently performs better overall throughout both data splits, suggesting its efficacy in the classification job. The metrics demonstrate its potential to achieve high levels of precision and accuracy while preserving a decent balance between specificity and sensitivity, which is essential for the dependable deployment of the model in real-world applications.

6. CONCLUSION

The proposed framework can be considered as a substantial development regarding retinal image analysis since it employs state-of-the-art approaches at different stages of the workflow. A robust image pre-processing scheme helps in cleaning up the input images and the advanced LAAMS-UNet++ architecture aids in segmenting out areas of interest. Therefore, extracting and selecting the most required features have the maximum utilization of the crucial information for classification purposes. Furthermore, Dense-CSPDarkNet53 + LinkNet-34 with EfficientNetB7 encoder improves the classification and the performance makes the framework a strong tool for diagnosis of retinal diseases. The detection and classification of eye disease are effective in providing better prediction results and increasing performance. Strong performance is shown by the "Proposed" model in both data splits. The model performs even better in the 80/20 split, with a robust F-measure of 97.748%, accuracy of 98.521%, and precision of 98.812%. Its high sensitivity (97.102%) and specificity (97.553%) demonstrate its reliable performance over a variety of datasets.

efforts will concentrate on enhancing every element and investigating the use of extra deep learning algorithms to enhance efficiency and applicability.

REFERENCES

[1] Muthukannan, P. (2022). Optimized convolution neural network based multiple eye disease detection. Computers in Biology and Medicine, 146, 105648.

[2] Sarki, R., Ahmed, K., Wang, H., & Zhang, Y. (2020). Automated detection of mild and multi-class diabetic eye diseases using deep learning. Health Information Science and Systems, 8(1), 32.

[3] Sarki, R., Ahmed, K., Wang, H., Zhang, Y., & Wang, K. (2021). Convolutional neural network for multi-class classification of diabetic eye disease. EAI Endorsed Transactions on Scalable Information Systems, 9(4)

[4] Gour, N., & Khanna, P. (2021). Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network. Biomedical signal processing and control, 66, 102329.

[5] Pachade, S., Porwal, P., Thulkar, D., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., ... & Mériaudeau, F. (2021). Retinal fundus multi-disease image dataset (RFMiD): a dataset for multi-disease detection research. Data, 6(2), 14.

[6] Aslam, J., Arshed, M. A., Iqbal, S., & Hasnain, H. M. (2024). Deep Learning Based Multi-Class Eye Disease Classification: Enhancing Vision Health Diagnosis. Technical Journal, 29(01), 7-12.

[7] Vadduri, M., & Kuppusamy, P. (2023). Enhancing Ocular Healthcare: Deep Learning-Based multi-class Diabetic Eye Disease Segmentation and Classification. IEEE Access.

[8] Badah, N., Algefes, A., AlArjani, A., & Mokni, R. (2022). Automatic eye disease detection using machine learning and deep learning models. In Pervasive Computing and Social Networking:Proceedings of ICPCSN 2022 (pp. 773-787). Singapore: Springer Nature Singapore.

[9] Keenan, T. D., Clemons, T. E., Domalpally, A., Elman, M. J., Havilio, M., Agrón, E., ... & Chew, E.Y. (2021). Retinal specialist versus artificial intelligence detection of retinal fluid from OCT: age-related eye disease study 2: 10-year follow-on study. Ophthalmology, 128(1), 100-109.

[10] Arora, A., Gupta, S., Singh, S., & Dubey, J. (2022, July). Eye disease detection using transfer learning on vgg16. In Proceedings of Third International Conference on Computing, Communications, and Cyber-Security: IC4S 2021 (pp. 527-536). Singapore: Springer Nature Singapore.

[11] Raza, A., Khan, M. U., Saeed, Z., Samer, S., Mobeen, A., & Samer, A. (2021, December). Classification of eye diseases and detection of cataract using digital fundus imaging (DFI) and inception-V4 deep learning model. In 2021 International Conference on Frontiers of Information Technology (FIT) (pp. 137-142). IEEE.

[12] Tripathi, P., Akhter, Y., Khurshid, M., Lakra, A., Keshari, R., Vatsa, M., & Singh, R. (2022). MTCD: Cataract detection via near infrared eye images. Computer Vision and Image Understanding, 214, 103303.

[13] Adel, A., Soliman, M. M., Khalifa, N. E. M., & Mostafa, K. (2020, December). Automatic classification of retinal eye diseases from optical coherence tomography using transfer learning. In 2020 16th International computer engineering conference (ICENCO) (pp. 37-42). IEEE.

[14] Albelaihi, A., & Ibrahim, D. M. (2024). DeepDiabetic: An Identification System of Diabetic Eye Diseases Using Deep Neural Networks. IEEE Access.

[15] Bustamante-Arias, A., Ruiz Lozano, R. E., & Rodriguez-Garcia, A. (2022). Dry eye disease, a prominent manifestation of systemic autoimmune disorders. European Journal of Ophthalmology, 32(6), 3142-3162.

[16] Deng, X., Qiu, S., Jin, W. and Xue, J., 2022. Three-dimensional reconstruction method for bionic compound-eye system based on MVSNet network. Electronics, 11(11), p.1790.

[17] Han, Z.K., Xing, H., Yang, B. and Hong, C.Y., 2022. A few-shot learning-based eye diseases screening method. European Review for Medical & Pharmacological Sciences, 26(23).

[18] Liu, H., Li, R., Zhang, Y., Zhang, K., Yusufu, M., Liu, Y., Mou, D., Chen, X., Tian, J., Li, H. and Fan, S., 2023. Economic evaluation of combined population-based screening for multiple blindness-causing eye diseases in China: a cost-effectiveness analysis. The Lancet Global Health, 11(3), pp.e456-e465.

[19] Chea, N. and Nam, Y., 2021. Classification of Fundus Images Based on Deep Learning for Detecting Eye Diseases. Computers, Materials & Continua, 67(1).

[20] Bali, A. and Mansotra, V., 2021. Transfer learning-based one versus rest classifier for multiclass multi-label ophthalmological disease prediction. Transfer, 12(12).

[21] Udayaraju, P., Jeyanthi, P. and Sekhar, B.V.D.S., 2023. A hybrid multilayered classification model with VGG-19 net for retinal diseases using optical coherence tomography images. Soft Computing, 27(17), pp.12559-12570.

[22] Park, S., 2023. VisionAI-Human Eye Fundus Screening System Based on AI Deep Learning Technology. Journal of Student Research, 12(1).

[23] Kumaran, M., & Devarajan, B. (2023). eyeVarP: a computational framework for the identification of pathogenic variants specific to eye disease. Genetics in Medicine, 25(7), 100862.

[24] Xiong, Y. Z., Kwon, M., Bittner, A. K., Virgili, G., Giacomelli, G., & Legge, G. E. (2020). Relationship between acuity and contrast sensitivity: differences due to eye disease. Investigative ophthalmology & visual science, 61(6), 40-40.

[25] Dabov, K., Foi, A., Katkovnik, V. and Egiazarian, K., 2006, February. Image denoising with block-matching and 3D filtering. In Image processing: algorithms and systems, neural networks, and machine learning (Vol. 6064, pp. 354-365). SPIE.
[26] Yadav, G., Maheshwari, S. and Agarwal, A., 2014, September. Contrast limited adaptive histogram equalization based enhancement for real time video system. In 2014 international conference on advances in computing, communications and informatics (ICACCI) (pp. 2392-2397). IEEE.
[27] Liu, J., Lam, S., Lemaillet, P. and Cheng, W.C., 2021, February. Colorimetrical evaluation of color normalization methods for H&E-stained images. In Medical Imaging 2021: Digital Pathology (Vol. 11603, pp. 199-207). SPIE.
[28] Rahimzadeh, M. and Attar, A., 2020. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Informatics in medicine unlocked, 19, p.100360.
[29] Zhong, C., Li, G. and Meng, Z., 2022. Beluga whale optimization: A novel nature-inspired metaheuristic algorithm. Knowledge-Based Systems, 251, p.109215.
[30] Zhao, S., Zhang, T., Ma, S. and Wang, M., 2023. Sea-horse optimizer: a novel nature-inspired meta-heuristic for global optimization problems. Applied Intelligence, 53(10), pp.11833-11860.
[31] Roy, A.M. and Bhaduri, J., 2021. A deep learning enabled multi-class plant disease detection model based on computer vision. Ai, 2(3), pp.413-428.
[32] Sulaiman, A., Anand, V., Gupta, S., Al Reshan, M.S., Alshahrani, H., Shaikh, A. and Elmagzoub, M.A., 2024. An intelligent LinkNet-34 model with EfficientNetB7 encoder for semantic segmentation of brain tumor. Scientific Reports, 14(1), p.1345.
[33] Nazir, T., Irtaza, A., Javed, A., Malik, H., Hussain, D. and Naqvi, R.A., 2020. Retinal image analysis for diabetes-based eye disease detection using deep learning. Applied Sciences, 10(18), p.6185.
[34] Akram, A. and Debnath, R., 2020. An automated eye disease recognition system from visual content of facial images using machine learning techniques. Turkish Journal of Electrical Engineering and Computer Sciences, 28(2), pp.917-932.
[35] https://www.kaggle.com/datasets/aaryapatel98/indian-diabetic-retinopathy-image-dataset
[36] https://www.kaggle.com/datasets/sshikamaru/glaucoma-detection?select=Fundus_Train_Val_Data
[37] https://www.kaggle.com/datasets/nandanp6/cataract-image-dataset
[38] https://www.kaggle.com/datasets/sachinkumar413/diabetic-retinopathy-dataset?select=Healthy
[39] https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification