JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 08 Issue 01
2024

Research Article
https://doi.org/10.56979/801/2024

Spatial Correlation Module for Classification of Ocular Diseases in Diabetic Retinopathy Using Color Fundus Images

Nosha Naeem[1*], Ali Hadier[2], Muhammad Irfan[3], Waqar Azeem[2]

[1]Department of Data Science, Lahore Garrison University, Lahore, Pakistan.
[2]Faculty of Computer Sciences, Lahore Garrison University, Lahore, Pakistan.
[3]Department of Information Technology Operations, Senior Programme Manager, Punjab Information Technology Board, Lahore, Pakistan.
[*]Corresponding Author: Nosha Naeem. Emails: noshanaeem9@gmail.com

Received: August 11, 2024 **Accepted:** December 01, 2024

Abstract: Early identification and therapy of ocular diseases (ODs) are crucial in avoiding irreversible vision loss. Color imaging of the fundus (CFI) is an economical and trustworthy screening tool. However, automatic and thorough diagnostic tools are required since early OD symptoms are usually modest. The traditional wisdom suggests treating the eyes independently and relying only on image-level diagnostics, without considering intraocular correlation data. Additionally, these techniques typically only target a single or a limited number of ODs simultaneously. This research presents a novel classification model called PLML_ODs. It includes patient-level multi-label OD data. Our technique integrates patient-level diagnosis, notably in diabetic retinopathy, by combining bilateral ocular and multi-label ODs classification. A feature correlation SCNet, a classification score generator, and a feature extracting backbone based on convolutional neural networks (CNN) named DenseNet-169 comprise the PLML_ODs system. The DenseNet-169 model obtains two sets of features using on both sides CFI. The SCNet then captures a correlation connecting the two collections of features pixel-wise. In order to get an embodiment at the patient level, the attributes are integrated after analysis. Using this representation, the ODs classification process is carried out. We evaluate PLML_ODs's classification performance to that of other baseline techniques using a publicly available dataset and an enveloping margin loss.

Key words: Spatial Correlation; Ocular Disease; Fundus Images; Neural Networks; DenseNet

1. Introduction

The prevalence ophthalmic illnesses (ODs) that may progress to irreversible vision loss has been skyrocketing over the past several decades. As trachoma, AMD, DR, cataracts and untreated refractive errors in this category. According to a recent World Health Organization (WHO) study, over 2.2 million people had vision impairment worldwide. Half or more of those cases could have been prevented if it works had been performed correctly [1]. Refractive disorders which are not corrected and cataracts are the most common causes of blindness and its forms (nearsightedness, farsightedness, presbyopia, trachoma) [2]. According to the World Health Organization, more than 153 million people worldwide don't get their refractive problems repaired over 18 million are blind on both sides because of cataracts and nearly one million suffer from trachoma. It is estimated that AMD causes 8.7% of all blindness worldwide, or 3 million people, and is most common in industrialized countries [3]. It is now emerging that 4.8% of the global total of 37.3 million blind individuals are blind due to DR (1.8 million) [4-6]. Some 46 percent of diabetics are unaware that they have the disease. Only about 2% of people with diabetes ultimately lose the vision, and another 10% have severe

vision loss after 15 years [8]. Moreover, over 75% of people are known to develop diabetes-related complications after twenty years [9]. You must detect and treat ODs as soon as possible to avoid any irreversible visual loss [10]. Persistently floating blood cells may harm the retina and cause permanent vision loss, or actually lead to blindness [11–13]. For the purpose of improved OD identification, numerous imaging techniques have been developed. The two most used methods are optical coherence tomography (OCT) and color fundus imaging (CFI) [14]. Cross-sectional imaging of the retina can help diagnose by determining retinal thickness [15]. The inner eye surfaces are tracked while CFI checks for potential problems. Either method can pick up OD in its early stage. Considering CFI is more efficient and economical more regular fundus evaluation with CFI should be done in patients without symptoms especially the elderly [16, 17]. Progression of common ODs (e.g., DR, cataracts, AMD etc) is often unanticipated by patients because there are no early symptoms of loss of vision [18]. Getting an accurate diagnosis is just too hard to begin with. A large amount of energy and time needs to be devoted by humans in processing large volumes of CFI data [19]. The shortage of competent radiologists in less developed areas is so severe that manual analysis is economically impossible. Convolutional Neural Networks (CNNs) have achieved amazing progress for medical imaging [21, 22].

In OD diagnosis, CNNs have excellent performance from illness categorization to object identification. Liefers et al. [9] located the fovea centers in OCT images using a pixel-wise classification. Meng et al. [14] introduced a two stage CNN approach to detect optical discs in CFI. Zekavat et al. [15] differentiated intraregional from extra retinal fluid using convolutional neural networks (CNNs) on optical coherence tomography shots.

Gu et al. [5] made a network called ReLayNet that can tell the difference between the different layers of retina and fluid buildup on optical coherence tomography (OCT) pictures. Segmentation of CFI retinal vessels was achieved using convolutional neural networks (CNNs) and linked random fields [6]. Despite promise [23], little research to date has been addressed the difficulty of multilevel OD classification using CFI. The main problem comes if more than one form of OD is found in a single patient. More importantly, since patients tend to have more than one OD, we need to optimize models for multi label OD categorization. According to a research [11], coexisting myopia increases the risk of false negative glaucoma patient categorization. Earlier studies have proven promising, though those results may not hold up in the real world, where complex situations are always going to pop up. The problems outlined above are compounded by a dearth of studies about how to diagnose OD at the patient level. Almost all studies in recent years have been about images, comparing and contrasting the CFI of each eye separately. The primary contributions of this research are outlined below:

- In order to successfully combine features collected from left and right Coloring Fundus Images (CFI), this research presents a new PLML_OD model that is built on SCNet, a custom module. Seven different ocular diseases (OD) are categorized using the suggested model, which uses a backbone CNN (DenseNet-169). In order to improve the quality from the recovered characteristics, SCNet takes their relationships into account.
- We outperform many baseline techniques that depend on direct highlight concatenation on a publicly available CFI dataset using our proposed model, which incorporates a carefully designed method of feature correlation and fusion.
- The ablation experiments show that our suggested model outperforms the current best practices.

The remainder of this article is organized as follows:

A review of the most current literature is given in Section 2. An extensive description of the suggested PLML_OD template for OD detection using CFI is provided in Section 3. In Section 4, we will go over the experimental setup, the outcomes, and how the suggested strategy stacks up against the competition. In Section 5, we draw a conclusion and provide some suggestions for further study.

2. Related work

In this section we will discuss the limitations of current approaches to the diagnosis of OD.

Also highlighted are the primary ways in which our suggested strategy tackles present-day problems. The characteristics retrieved by CFI were trained using transfer learning by Wang et al. [24]. After that, they added them using ML-C, which is problem transformation based. The authors perform histogram equalization to both color and black & white photos using a dataset that comprises eight distinct labels. After that, they experimented with two different methods to see which one worked best for classifying the two groups of images. Finally, after calculating the likelihood from two models, I average the sigmoid. Their strategy was hindered since their network performance was negatively affected by the large number of 'other diseases' ODs in their dataset. Inadequate understanding of some health issues and fuzzy feature ideas was also caused by their data's imbalance. Mayya et al. [25] used graph convolution netting (GCN) to identify eight lesions known as DR in CFI, as far as the authors are aware. In addition to ResNet 101's feature extraction, two convolutional layers were employed: a 3x3 kernel, stride 2, and adaptive max pooling. In the end, XGBoost, a completely supervised instruction approach, was used. They came up with their own to make it easier to see laser scars, onions, and bleeding sores. The model significantly improved the receivers' operational characteristics and accuracy, according to their findings. Several abnormalities, including micro aneurysms, soft or hard fluids, along and others, were undetected by their method [12]. Because the micro aneurysm was so hard to see out on the CFI backdrop—a little red dot placed against a backdrop of bigger red dots that make up the retinal capillaries—this restriction was established. The ODIR2019 dataset was used by Dipu et al. [26] to transfer-learning-based identification of eight ODs. We found that Resnet-34, EfficientNet that, MobileNetV2, and VGG-16 are the best deep learning networks. After publishing their results, the authors trained a cutting-edge algorithm on their dataset. Then, they computed the accuracy of all their models. These models were ranked according on their accuracy: In this work, we evaluate Resnet-34, MobileNetV2, VGG-16 and EfficientNet. However, they did not provide any other methods to find ODs. For example, Choi et al. [27] suggested that we should use VGG-19 with random forest transfer learning for a CAD system because it was not able to evaluate the performance of the model only based on a computational accuracy score. A small dataset was used to identify eight distinct types of OD. They were able to maintain such accuracy as long as they kept the number of categories at three. For ten categories, categorization accuracy drops 30%. Then, when attempting to apply the transfer learning ensemble classifier, I was actually able to increase the accuracy only by 5.5%. While it did improve, the writers' disappointing performance was largely due to incorrect data and inverted data that didn't help save them. While it is developed by Shaik et al. [33], we may thank them for developing this HA-Net, or Hinge Attention Network. First of all, a pre trained model, namely VGG16 is utilized to extract spatial features from the original image which helps us to generate input. Then, using a spatial attention Autoencoders, it learns lesion characteristics by producing lesion features. Then a deep neural network and channel attention are used to classify retinopathy level. Additionally, it combines the Channel attention algorithm with LSTM convolution layers in order to help it focus on the relevant spatial data inside the data with a hierarchical structure. However, evaluation of performance on the APTOS database (85.54%) and the IDRiD database (66.41%) shows that HA Net might be useful for retinopathy classification. If the models aren't working, even for the most recent reports, we can assume it is because they can't see many ODs. Here are a few limitations: The model performance is hit hard by the multi label classes when there aren't enough data points for training. Some systems are too wary, as a result of insufficient or inconsistent data sets, to be used in the real world. A new framework, Patient-Level Multi-Label Ophthalmic Disorder (PLML_OD), is introduced to address these shortcomings of previous work.

Multiple Convolutional Neural Network (CNN) model are used to train the end to end system, learning to detect various ophthalmic disorders (ODs), from Color Fundus Imaging (CFI). Since BL-SMOTE can prevent data imbalance and overfitting, the additional Borderline Synthetic Minority Oversampling Technique (BL-SMOTE) is used to supplement the training dataset. SMOTE also improves model performance. Finally, we use BL-SMOTE as we are not applying a linear transformation and therefore preserving labels. Just as important, we offer an innovative PLML_OD architecture that uses a SCNet, a spatial correlation network, to extract features quickly. Since there are seven ODs in total, the proposed method can calculate the likelihood

for each image. Six different indicators of the system's performance were assessed utilizing the system and the results then compared to those obtained from other models and systems.

3. Materials and Methods

In this section, the dataset is detailed reviewed, the technique will be used is described, and the metrics that will be used to determine performance is outlined.

3.1. Dataset Overview and Augmentation

For the training of our PLML_OD model [23], we employed the open source CFI dataset from the ODIR 2019 competition. The first group is represented by healthy people (N), the other seven groups are DB, GL, CA, AMD, HT, MP and AB. We fuse CFI with other patient data like age to build patient level labels. The crowd sourced 4,020 of the 5,000 events contained in the original CFI dataset. Distribution of 4,020 patient cases into 8 categories can be seen in Figure 1. On this CFI dataset, we evaluate our model's performance. Due to dataset size is small, we train on the first two folds while test on the third. The minor classes, e.g. GL, CA, AMD, MP, AB and HT are incorporated in the dataset by using the BL-SMOTE approach. Table 1 contains an overview of both BL-SMOTE and non-BL-SMOTE datasets. It trains six thousand, sixteen instances; validates eight hundred and eighty; and tests one thousand seven hundred and seventy-five instances. We also report the average test fold results across the 3 cross-validation splits.

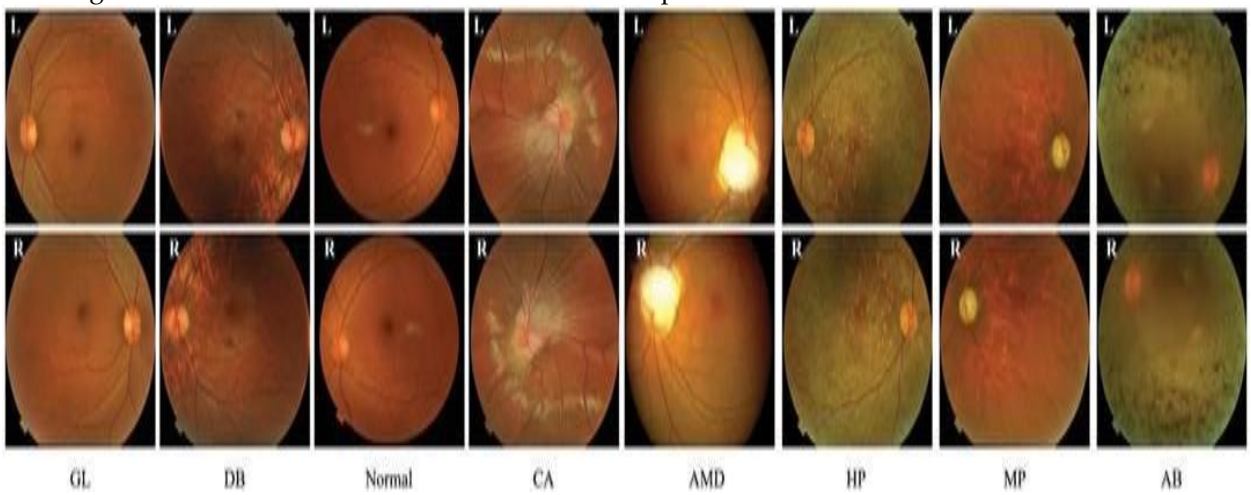

Figure 1. A selection of OD fundus pictures

Table 1. A brief look at the OD dataset

Dataset	N	DB	GL	CA	AMD	HT	MP	AB	Total
Original	1100	1100	200	200	180	150	190	900	4.020
BL-SMOTE	1100	1100	1100	1100	1100	1100	1100	1100	8.800
Training (70%)	770	770	770	770	770	770	770	770	6.160
Validation (10%)	110	110	110	110	110	110	110	110	880
Testing (20%)	220	220	220	220	220	220	220	220	1760

3.2. Proposed PLML_OD Model

The overall design of our proposed PLML_OD model (including CNN core, SCNet subsystem, and final classification layer) is shown in Figure 2.

3.2.1. CNN Backbone

Colour fundus images (CFIs) may have two sets of features extracted using a Convolutional Neural Network (CNN). Given a pair of left and right CFIs, Ll and Rr items of the OH x W x RGB space that specify

the dimensions of the input CFIs and RGB used to represent the three color channels, the CNN backbone outputs two feature maps, Dl and Dr, where Dl and Dr are elements of the 32x32 x 256 space. These feature maps have a representation Dl, Dr ∈ OH x W x F where F represents the number of extracted features. Importantly, our feature extraction procedure is performed without combining or sharing any data from the associated CFIs in any way. In other words, the matched CFIs do not need to be registered. Building our CNNs on the top of DenseNet topologies, our CNNs have none fully linked layers that are known to be helpful in handling our problem of fast time intervals from the class histogram.

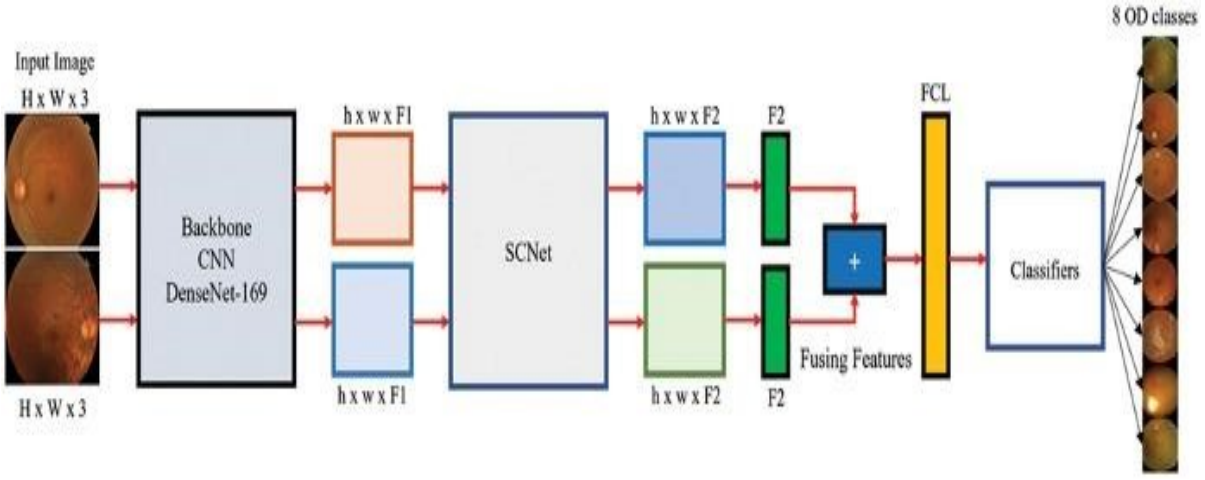

Figure 2. Architecture of proposed PLML_OD model

3.2.2. SCNet Module

The SCNet module creates two sets of features from the CNN backbone that we use as inputs for building two matching feature sets with correlation analysis.

The SCNet scheme of this study is shown in Figure 3. This module allows us to find pixel wise relationships between Ll and Rr, the individual input feature sets; at first every set of features is divided into query (Q), key (K) and value (V) features with 1×1 convolutions. To get features out of the CFI input pairs, we use Equations (1)– (3):

$$D_{lQ} = L\,(D_l; P_{lQ}),\ D_{rQ} = L\,(D_r; P_{rQ}) \quad (1)$$

$$D_{lK} = L\,(D_l; P_{lK}),\ D_{rK} = L\,(D_r; P_{rK}) \quad (2)$$

$$D_{lV} = L\,(D_l; P_{lV}),\ D_{rV} = L\,(D_r; P_{Rv}) \quad (3)$$

where L is a linear 1×1 convolution and P are the specified parameters, and they are represented by L and P, respectively. The converted Q/K and V features are of dimensions F and F′, respectively, and take empirical values of 1024.

The correlation weights for each pixel are determined by taking the inner product of the converted features and the sigmoid function (SF). The right and left CFIs' worth of data is combined by these weights:

$$C_{l\leftarrow r} \in O^{(H\times W)\,(H\times W)} = \text{Sigmoid}\,(D_{LQ}\,D^{T}_{rK}) \quad (4)$$

Similarly, correlation weights (Cr←l ∈ O(H×W)(H×W)) are calculated to combine information from the left CFI with that from the right CFI:

$$C_{r\leftarrow l} \in O^{(H\times W)\,(H\times W)} = \text{Sigmoid}\,(D_{LQ}\,D^{T}_{rK}). \quad (5)$$

We may see the interplay between every pair of CFI sites by using correlation weights. The next step is to identify two sets of features from the backbone CNN and refine the gathered weights. To create the refined feature sets, which are Dl_update ∇ OHxWx F and Dr_update Δ OHxWx F′, the weight maps linked to each set are multiplied.

$$D_{l\_update} = C_{l\leftarrow r} \times D_{rV} \quad (6)$$

$$D_{r\_update} = C_{r\leftarrow l} \times D_{lV} \quad (7)$$

Finally, the characteristics from bilateral CFI are fused to complete the SCNet procedure. As shown in Figure 3, the consequence of merging over four feature sets is fusion. Joining the updated CFI feature set on the right with the left input set, Dl_update, yields:

$$M = \{A_l, A_r\}. \quad (8)$$

$$A_l = L([D_l, D_{l\_update}]^T; P_{Al}) \quad (9)$$

$$A_r = L([D_r, D_{r\_update}]^T; P_{Ar}) \quad (10)$$

By substituting Eqs. (9) and (10) into Eq. (8), we obtain:

$$M = \{L([D_l, D_{l\_update}]^T; P_{Al}), L([D_r, Dr_{\_update}]^T; P_{Ar})\}. \quad (11)$$

Where M represents the SCNet, and Al and Ar are its outputs.

3.2.3. Classifier

Using global average pooling (GAP), the two feature sets generated by SCNet are converted into two feature vectors. A final classification module receives these vectors as a single string once they have been concatenated. There are two fully connected layers (FCLs) in the classifier, but ReLU only activates one of them. The dimensionality of the resulting features is reduced by using the first FCL. For example, in case of using DenseNet-169 backbone, the dimension of combined feature is 2048. After the first FCL, it reduces it to 1024 and, then it is equal to 1024 for all subsequent computations. A large number of classification categories are satisfied by the second FCL, which reduces features to an eighth dimension. Computing the network loss then entails comparing the eight-dimensional characteristics to the ground truth sickness categorisation labels.

3.2.4. Classifier Architecture

Two feature vectors are produced from the SCNet-generated feature sets using the global average pooling (GAP) process. Before feeding them into the last categorization module, these vectors are joined into a single string. The ReLU function activates just one of the two fully connected layers (FCLs) that make up the classifier. By using the first FCL, the dimensionality of the combined features is decreased. One example is the DenseNet-169 backbone, which uses 2048 as its starting feature dimension. It stays at 1024 after the initial FCL reduces it, and it stays there for all future computations. The characteristics are compressed to an eighth dimension in the second FCL, which corresponds to the many categorization categories. In order to find the network loss, we must next check the eight-dimensional features with the ground-truth labels for sickness classification.

3.2.5. Loss Function

To solve the problem of multi-label OD classification, a multi-label soft margin loss function is used. The mathematical expression of this function is given by Eq. (12):

$$L = -(1/G) \sum (G \text{ to } G\text{-}1)\, r[G] \log \sigma(O[G]) + (1 - r[G]) \log (1 - \sigma(O[G])). \quad (12)$$

Where G denotes the categories, σ represents the sigmoid activation, $r \in \{0, 1\}$ is the reference label, and O is the network output.

3.3. Performance Evaluation Metrics

For the official ODIR-2019 challenge website, some proposed metrics for measuring model classification performance include kappa score (KS), F1 score, the area beneath the receiver operation curve (AUC), accuracy (ACU), and mean (AVG). This data is derived from the equations (13) to (21). Here is how the KS is calculated:

$$KS = (Mo - Me) / (1 - Me) \quad (13)$$

$$Mo = \sum [TP\_h / (TP\_h + FN\_h)] \quad (14)$$

$$Me = \sum [TP\_h * (TP\_h + FN\_h) / N^2] \quad (15)$$

The F1 score is computed as:

$$F1\ score = 2 * TP / (2 * TP + FN + FP) \quad (16)$$

To get the AUC, or area under the receiver operating curve, one uses the following formula:

$$AUC = \int [TPR(FPR{-}1)]\, dx \quad (17)$$

To calculate the TPR and FPR, we use the following formulas:

$$TPR = TP / (TP + FN) \quad (18)$$

$$FPR = FP / (FP + TN) \quad (19)$$

Accuracy (ACU) is calculated as:

$$ACU = (TP + TN) / (TP + TN + FP + FN) \quad (20)$$

The average (AVG) is computed as:

$$AVG = (KS + F1\ score + AUC) / 3 \quad (21)$$

4. Results and Discussion

In this section, experimental findings of comparing the suggested PLML_OD model to some baseline models are presented with the publicly accessible OD datasets. In addition, an ablation investigation of the suggested model is performed.

4.1. Experimental Setup

We build our system on convolutional neural networks (CNNs) to explore the effects of various depth DenseNet models on feature extraction. We initialize the backbone CNNs with DenseNet models, pre-trained on ImageNet. The CFI data comes from different locations or hospitals which results in every picture scaled to 299 x 299 pixels. During training, random crops of 224×224 are taken, during test, center cropping is performed. For non convolutional network related techniques, Python is used; Keras is used for development of baseline models and proposed PLML_OD model. All tests are run on a Windows PC with 32GB of RAM and 11GB NVIDIA graphics processing unit.

4.2. Model Configuration

For maximum efficiency, the hyperparameters of the proposed PLML_OD model have been finetuned. When using the stochastic gradient descent (SGD) optimizer, Equation (11) provides the loss function for multi-label classification. The beginning rate of learning (LR) is changed from 0.001 using the poly learning rate reduction technique. The momentum is now 0.90. We conduct each experiment for 100 iterations and record the findings at the end.

4.3. Classification Results Using Various Backbone CNNs

Apart from bilateral CFI, CNNs extract strong foundation Features. We derive features with different levels of abstraction by adjusting the depth of convolutional neural networks (CNNs) in the backbone. Backbone CNN performs feature fusion via adding, pixel-wise multiplying, or concatenating when explicit fusion techniques are not available. Table 2 provides a detailed explanation of the results. The best way to fuse features is via the feature concatenation technique. Optimal performance is achieved by pixel-wise multiplication with a deep backbone CNN. Further, better CNN's with a more robust backbone perform better. DenseNet-169 models trained on feature concatenation outperform DenseNet-201 models for KS, AUC, F1-score, ACU, and final AVG by 8.8%, 2.4%, 2.4%, 4.5%, respectively. It speaks to more abstract component distinction capacity. Ultimately there seemed to be a performance ceiling, and DenseNet-169 did not make much of a difference. Research also shows that there is no linear relationship between the number of links between nodes and the performance of the network. [21]. Here are three potential reasons for this phenomenon:

- Vanishing gradient problem: With more parameters, we keep optimizing a network increasingly difficult [14].
- Inefficient feature reuse: The limitations of current training sample size lead to deep network developing too many features, which are not fully utilized [15].
- Insufficient training: Due to little number of training samples, the network is not fully trained.

The collected attributes were used to train four convolutional neural network (CNN) classifiers that optimize their Hyperparameters via grid search. The models were trained using the 5-fold cross-validation method. To test how well the models could identify various ODs, we employed a battery of performance measurements. Figure 5 shows the SCNet-based DenseNet confusion matrices. The models' ability to correctly categorize OD cases is shown by the findings. As an example, out of 220 OD occurrences, 215 were accurately detected by the DenseNet-169 model. Out of the remaining cases, 1 was misclassified as normal, 1 as DB, and 1 as HP. Every typical occurrence had its class label correctly predicted by the model. On the other hand, out

of 220 total cases, DenseNet-121 accurately recognized 199 instances of AB illness using extracted attributes for model training, while incorrectly labeling 21 cases as normal or other OD disorders. Out of 220 instances of AMD illness, 202 were properly predicted by the DenseNet-264 model, whereas 199 cases of DB disease were successfully predicted by the DenseNet-201 model.

Table 2. Use DenseNet network architecture for OD classification without SCNet.

DenseNet Architecture	Features Fusion	ACU	KS	F1-Score	AUC	AVG
	Pixel-wise addition	0.821 ± 0.019	0.401 ± 0.002	0.819 ± 0.002	0.826 ± 0.006	0.682 ± 0.005
DenseNet-121	Pixel-wise multiplication	0.801±0.025	0.391 ± 0.016	0.829 ± 0.010	0.806 ± 0.058	0.668 ± 0.079
	Pixel-wise concatenation	0.821 ± 0.019	0.412 ± 0.015	0.829 ± 0.010	0.838 ± 0.016	0.693 ± 0.071
	Pixel-wise addition	0.940 ± 0.011	0.441 ± 0.019	0.926 ± 0.055	0.927 ± 0.082	0.764 ± 0.023
DenseNet-169	Pixel-wise multiplication	0.936 ± 0.027	0.498 ± 0.039	0.919 ± 0.073	0.932 ± 0.096	0.783 ± 0.043
	Pixel-wise concatenation	0.948 ± 0.002	0.501 ± 0.042	0.945 ± 0.062	0.947 ± 0.016	0.797 ± 0.025
	Pixel-wise addition	0.821 ± 0.019	0.821 ± 0.019	0.829 ± 0.001	0.821 ± 0.019	0.692 ± 0.012
DenseNet-201	Pixel-wise multiplication	0.801±0.025	0.801±0.025	0.819 ± 0.013	0.816 ±0.025	0.678 ± 0.016
	Pixel-wise concatenation	0.821 ± 0.019	0.821 ± 0.019	0.839 ± 0.011	0.848 ± 0.017	0.709 ± 0.019
	Pixel-wise addition	0.841 ± 0.021	0.421 ± 0.004	0.839 ± 0.004	0.846 ± 0.008	0.711 ± 0.013
DenseNet-264	Pixel-wise multiplication	0.821 ± 0.027	0.422 ± 0.018	0.829 ± 0.014	0.826 ± 0.060	0.699 ± 0.042
	Pixel-wise concatenation	0.850 ± 0.025	0.432 ± 0.017	0.849 ± 0.012	0.858 ± 0.018	0.728 ± 0.015

Table 3. Classifying objects of interest using SCNet and the DenseNet architecture

DenseNet Architecture	ACU	KS	F1-Score	AUC	AVG
DenseNet-121	0.841 ± 0.024	0.422 ± 0.015	0.830 ± 0.010	0.849 ± 0.016	0.703 ± 0.071
DenseNet-169	0.968 ± 0.001	0.551 ± 0.003	0.967 ± 0.005	0.969 ± 0.007	0.850 ± 0.04
DenseNet-201	0.861 ± 0.032	0.433 ± 0.046	0.840 ± 0.089	0.859 ± 0.062	0.710 ± 0.071
DenseNet-264	0.871 ± 0.037	0.443 ± 0.059	0.850 ± 0.072	0.869 ± 0.087	0.720 ± 0.067

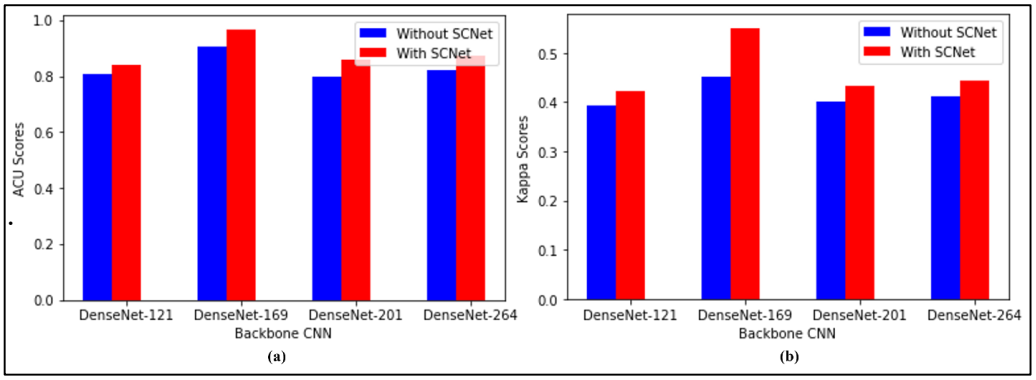

Figure 3. The SCNet's impact on Backbone CNN classification Accuracy. (a) The performance of Backbone CNN with and without SCNet is evaluated in terms of ACU(b) The performance of Backbone CNN with and without SCNet is evaluated in terms of Kappa Score

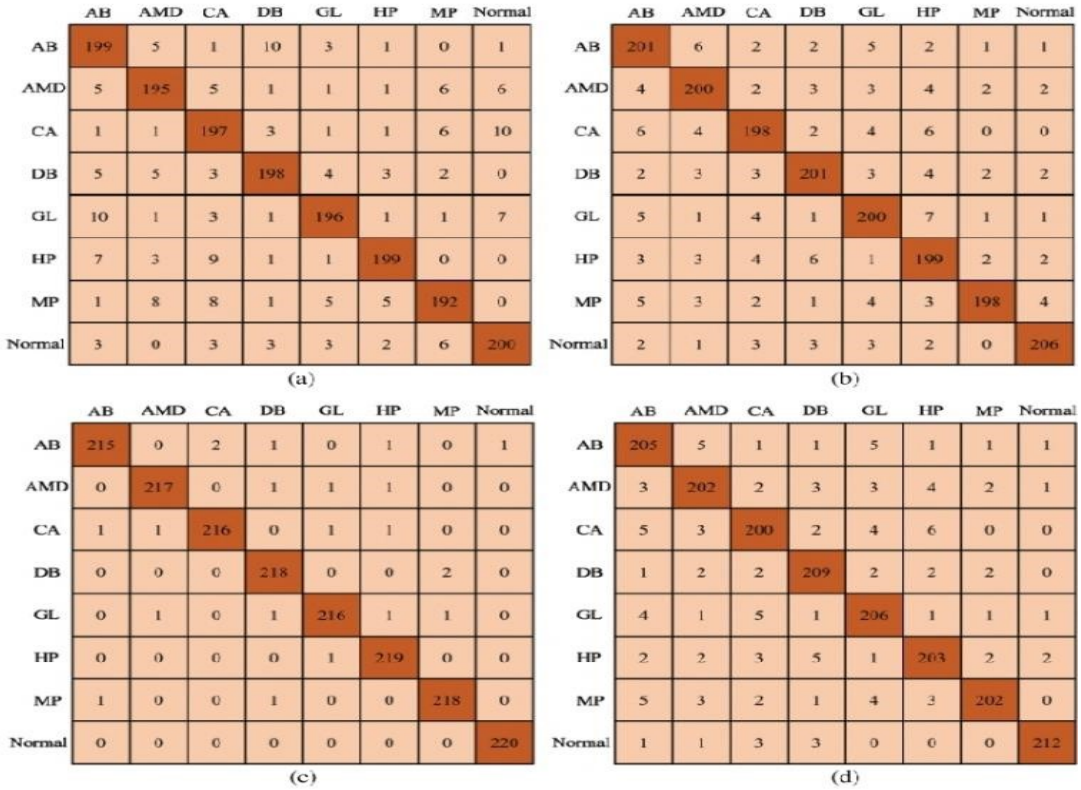

Figure 4. DenseNet-121, DenseNet-201, the proposed model, and DenseNet-264 are all shown in this confusion matrix.

4.4. Computational Complexities of DenseNet Architectures

The network complexity is described in detail in Table 4, and the classification metrics are shown in connection to FLOPs and network parameters in Fig. 6. The results of the categorization may then be compared in an impartial and objective manner.

Table 4. Differences in computational complexities between a network with () and without (×) a SCNet

DenseNet Architecture	SCNet	FLOPS	Parameters (Millions)
DenseNet-121	✓	17.12	16.9
	×	17.09	14.10
DenseNet-169	✓	32.11	30.10
	×	32.07	24.09
DenseNet-201	✓	41.11	58.05
	×	36.03	28.83
DenseNet-264	✓	71.10	77.23
	×	65.12	47.83

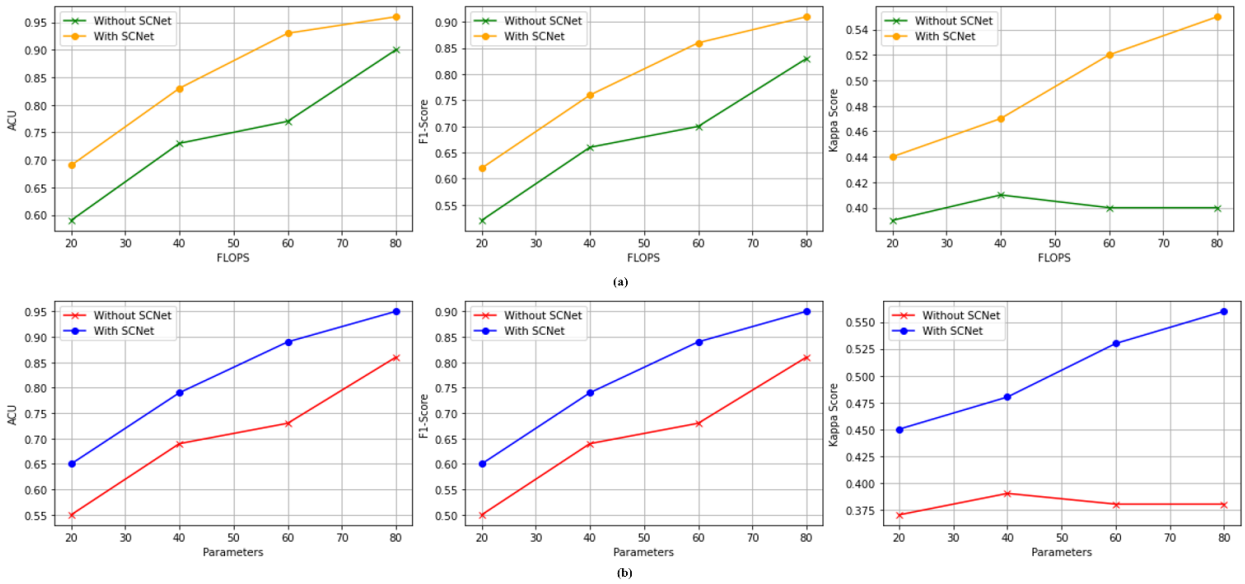

Figure 5. The results of the proposed PLML_OD model (a) FLOP classification performance (b) Model configurations for total trainable parameters

Regardless of whether standard (FLOPs) or network characteristics are considered, the proposed PLML_OD method exceeds the baseline. We also point out that, with SCNet, the DenseNet-169 backbone performs better than DenseNet-201 and DenseNet-264 baselines, using less FLOP. The results of the suggested technique reveal such a good classification result that it is likely to lead to the conclusion that network complexity alone is not decisive.

Categorization of the OD in an improved manner on a patient by patient basis requires the careful evaluation of the relationship between left and right CFI. It is of utmost relevance that patient level diagnosis that relies heavily upon clinical features (CFI) classification performance is improved with information from both the CFI [25]. SCNet does increase a small amount of FLOPs, and the number of network parameters increases significantly (particularly for baselines built on DenseNet-201 and DenseNet-264 backbones).

In Table 5, you can see which time it took to train the suggested model and the other backbone models. The results show training on the whole dataset takes 500 seconds for DenseNet-121, 433 seconds for DenseNet-169 and 693 seconds for DenseNet-201. Since the number of parameters to train is less in DenseNet-169 model, training time is lower than other models (see Fig. 6b).

Table 5. Training time of the models

Models	Training time in seconds	Testing time in seconds
DenseNet-121	500 s	95 s
DenseNet-169	433 s	26 s
DenseNet-201	693 s	62 s
DenseNet-264	725 s	87 s

4.5. Ablation Study

We wanted to make a new model by putting together SCNet, better versions of DenseNet-169, and the multi-label soft threshold loss function in this study. We were able to use the control variable method to look at the trial data across eight ODs by comparing the PLML_OD-recommended model to its control condition. A number was changed to make sure the model was correct. The research analyzed the models' accuracy and kappa score values using metrics to determine which of the modified modules improved the model's performance. The DenseNet-169 model is implemented in Experiment 2, and a comparison of it with SCNet and a multi label soft margin loss function is shown in Experiment 3. In the first two trials, we present the SCNet model. Table 6 contains all of the experimental data. Various model components, such as SCNet and DenseNet-169, are integrated in Table 6.

Table 6. Shows how the model's many parts, such as SCNet and DenseNet-169, are put together.

Experiments	SCNet	DenseNet-169	Multi-Label Soft Margin	Image Resolution	Accuracy	Kappa Score
1	✓	×	×	299 × 299 × 3	83.98 %	0.40
2	×	✓	×	299 × 299 × 3	85.13%	0.43
3	✓	✓	✓	299 × 299 × 3	96.80%	0.55

The model's average OD classification accuracy is improved by 1.15 percentage points when using DenseNet-169, according to a comparison of Experiment 1 and Experiment 2. Results from Experiment 3 show that the suggested PLML_OD model, which is a combination of SCNet, DenseNet, and a multi-label soft margin loss function, outperforms Experiment 2 in OD classification by 6.67%.

4.6. Comparison with State-of-the-Art Methods

To verify the outstanding performance of the suggested PLML_OD, we compare it with state of the art techniques here. Wei et al. [28] classify ocular disorders with an accuracy of 0.84 using their novel inception-v3 model. Convolutional neural network (CNN) models were used by an independent research [29–31] to find glaucoma on fundus pictures, scoring an F1 of 0.90, an accuracy of 0.91, and a Kappa score of 0.42. Wang et al. [32] suggested a new model MBSaNet, with accuracy 0.89. The comparison of using the suggested model with previous research are shown in Table 7, which includes a number of assessment indicators.

Table 7. Evaluation of the suggested PLML_OD model's performance in light of prior research

References	Methods	Accuracy	Kappa	F1-Score
Park et al. [30] 2022	CNN	0.7425	0.52	0.89
Wei et al. [28] 2019	Inception-v3	0.8411	0.44	0.85
Gour et al. [31] 2021	Two I/P VGG 16	0.8515	---	0..86

Wang et al. [24] 2020	EfficientNetB3	0.7356	0.50	0.88
Wang et al. [32] 2023	MBSaNet	0.8916	0.44	0.88
Ours	PLML_OD	0.9188	0.55	0.94

5. Conclusion

A SCNet for geographical correlation, a CNN core, and a classifier to provide classification scores were all part of the package. By combining the data sets acquired by the backbone CNN from the left and right Color Fundus Images (CFI), pixel-wise feature correlation is made possible by SCNet, the most innovative architectural component of our network. Next, a representation at the patient level is generated by combining the processed attributes. This is done in preparation for the final OD categorization at the patient level. The categorization capacity of the proposed model was tested using a publically available dataset that included seven distinct ODs. Our system achieved an outstanding 91.80% accuracy rate when it came to OD classification especially in diabetic retinopathy. We are hoping that our approach will assist ophthalmologists in identifying ODs. In order to circumvent the study restrictions of an imbalanced dataset and a small sample size, we used Borderline-SMOTE (BL-SMOTE). A federated learning-based collaboration architecture that can securely integrate data from several institutions without compromising patient privacy is an important area for future research.

References

1. C. Perdomo, J. Oscar and A. G. Fabio, "A systematic review of deep learning methods applied to ocular images," Ciencia e Ingeniería Neogranadina, vol. 30, no. 1, pp. 9–26, 2022.
2. W. Xiao, H. Xi, H. W. Jing, R. L. Duo, Z. Yi et al., "Screening and identifying hepatobiliary diseases through deep learning using ocular images: A prospective, multicentre study," The Lancet Digital Health, vol. 3, no. 2, pp. e88–e97, 2021.
3. Z. Wang, Z. Yuanfu, Y. Mudi, M. Yan, Z. Wenping et al., "Automated segmentation of macular edema for the diagnosis of ocular disease using deep learning method," Scientific Reports, vol. 11, no. 1, pp. 1–12,2021.
4. P. S. Grewal, O. Faraz, R. Uriel and T. S. T. Matthew, "Deep learning in ophthalmology: A review," Canadian Journal of Ophthalmology, vol. 53, no. 4, pp. 309–313, 2018.
5. H. Gu, G. Youwen, G. Lei, W. Anji, X. Shirong et al., "Deep learning for identifying corneal diseases from ocular surface slit-lamp photographs," Scientific Reports, vol. 10, no. 1, pp. 1–11, 2020.
6. K. Zhang, L. Xiyang, L. Fan, H. Lin, Z. Lei et al., "An interpretable and expandable deep learning diagnostic system for multiple ocular diseases: qualitative study," Journal of Medical Internet Research, vol. 20, no. 11, pp. e11144, 2018.
7. H. Malik, S. F. Muhammad, K. Adel, A. Adnan, N. Q. Junaid et al., "A comparison of transfer learning performance versus health experts in disease diagnosis from medical imaging," IEEE Access, vol. 8, no. 1, pp. 139367–139386, 2020.
8. [8] D. Wang and W. Liejun, "On OCT image classification via deep learning," IEEE Photonics Journal, vol.11, no. 5, pp. 1–14, 2019.
9. B. Liefers, G. V. Freerk, S. Vivian, V. G. Bram, H. Carel et al., "Automatic detection of the foveal center in optical coherence tomography," Biomedical Optics Express, vol. 8, no. 11, pp. 5160–5178, 2017.
10. T. Nazir, I. Aun, J. Ali, M. Hafiz, H. Dildar et al., "Retinal image analysis for diabetes-based eye disease detection using deep learning," Applied Sciences, vol. 10, no. 18, pp. 6185, 2020.
11. D. S. Q. Ting, R. P. Louis, P. Lily, P. C. John, Y. L. Aaron et al., "Artificial intelligence and deep learning in ophthalmology," British Journal of Ophthalmology, vol. 103, no. 2, pp. 167–175, 2019.
12. O. Ouda, A. M. Eman, A. A. E. Abd and E. Mohammed, "Multiple ocular disease diagnosis using fundus images based on multi-label deep learning classification," Electronics, vol. 11, no. 13, pp. 1966, 2022.
13. T. K. Yoo, Y. C. Joon, K. K. Hong, H. R. Ik and K. K. Jin, "Adopting low-shot deep learning for the detection of conjunctival melanoma using ocular surface images," Computer Methods and Programs in Biomedicine, vol. 205, pp. 106086, 2021.
14. X. Meng, X. Xiaoming, Y. Lu, Z. Guang, Y. Yilong et al., "Fast and effective optic disk localization based on convolutional neural network," Neurocomputing, vol. 312, no. 2, pp. 285–295, 2018.
15. S. M. Zekavat, K. R. Vineet, T. Mark, Y. Yixuan, K. Satoshi et al., "Deep learning of the retina enables phenome-and genome-wide analyses of the microvasculature," Circulation, vol. 145, no. 2, pp. 134–150,2022.
16. B. Hassan, Q. Shiyin, H. Taimur, A. Ramsha and W. Naoufel, "Joint segmentation and quantification of chorioretinal biomarkers in optical coherence tomography scans: A deep learning approach," IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1–17, 2021.
17. Y. Yang, L. Ruiyang, L. Duoru, Z. Xiayin, L. Wangting et al., "Automatic identification of myopia based on ocular appearance images using deep learning," Annals of Translational Medicine, vol. 8, no. 11, pp.7–22, 2020.
18. K. Y. Wu, K. Merve, T. Cristina, J. Belinda, H. N. Bich et al., "An overview of the dry eye disease in sjögren's syndrome using our current molecular understanding," International Journal of Molecular Sciences, vol. 24, no. 2, pp. 1580, 2023
19. L. Faes, K. W. Siegfried, J. F. Dun, L. Xiaoxuan, K. Edward et al., "Automated deep learning design for medical image classification by health-care professionals with no coding experience: A feasibility study,"
20. J. Kanno, S. Takuhei, I. Hirokazu, I. Hisashi, Y. Yuji et al., "Deep learning with a dataset created using kanno saitama macro, a self-made automatic foveal avascular zone extraction program," Journal of Clinical Medicine, vol. 12, no. 1, pp. 183, 2023.
21. P. Li, L. Lingling, G. Zhanheng and W. Xin, "AMD-Net: Automatic subretinal fluid and hemorrhage Segmentation for wet age-related macular degeneration in ocular fundus images," Biomedical Signal Processing and Control, vol. 80, no. 2, pp. 104262, 2023.

22. C. S. Lee, J. T. Ariel, P. D. Nicolaas, W. Yue, R. Ariel et al., "Deep-learning based, automated segmentation of macular edema in optical coherence tomography," Biomedical Optics Express, vol. 8, no. 7, pp. 3440–3448, 2017.
23. [23] J. He, L. Cheng, Y. Jin, O. Yu and G. Lixu, "Multi-label ocular disease classification with a dense correlation Deep neural network," Biomedical Signal Processing and Control, vol. 63, no. 24, pp. 102167, 2021.
24. J. Wang, Y. Liu, H. Zhanqiang, H. Weifeng and L. Junwei, "Multi-label classification of fundus images with efficientnet," IEEE Access, vol. 8, pp. 212499–212508, 2020.
25. V. Mayya, K. Uma, K. S. Divyalakshmi and U. A. Rajendra, "An empirical study of preprocessing Techniques with convolutional neural networks for accurate detection of chronic ocular diseases using Fundus images," Applied Intelligence, vol. 53, no. 2, pp. 1548–1566, 2023.
26. N. M. Dipu, A. S. Sifatul and K. Salam, "Ocular disease detection using advanced neural network based Classification algorithms," Asian Journal for Convergence in Technology, vol. 7, no. 2, pp. 91–99, 2021.
27. J. Y. Choi, K. Y. Tae, G. S. Jeong, K. Jiyong, T. U. Terry et al., "Multi-categorical deep learning neural Network to classify retinal images: A pilot study employing small database," PloS One, vol. 12, no. 11, pp. E0187336, 2017.
28. D. Wei, P. Ning, L. Si-Yu, W. Yan-Ge and T. Ye, "Application of iontophoresis in ophthalmic practice: An Innovative strategy to deliver drugs into the eye," Drug Delivery, vol. 30, no. 1, pp. 2165736, 2023.
29. R. Fan, A. Kamran, B. Christopher, C. Mark, B. Nicole et al., "Detecting glaucoma from fundus Photographs using deep learning without convolutions: Transformer for improved generalization," Ophthalmology Science, vol. 3, no. 1, pp. 100233, 2023.
30. K. B. Park and Y. L. Jae, "SwinE-Net: Hybrid deep learning approach to novel polyp segmentation using Convolutional neural network and Swin Transformer," Journal of Computational Design and Engineering, Vol. 9, no. 2, pp. 616–632, 2022.
31. Shah, A. M., Aljubayri, M., Khan, M. F., Alqahtani, J., Sulaiman, A., & Shaikh, A. (2023). ILSM: Incorporated Lightweight Security Model for Improving QOS in WSN. Computer Systems Science & Engineering, 46(2).
32. N. Gour and K. Pritee, "Multi-class multi-label ophthalmological disease detection using transfer learning Based convolutional neural network," Biomedical Signal Processing and Control, vol. 66, no. 3, pp. 102329,2021.
33. K. Wang, X. Chuanyun, L. Gang, Z. Yang, Z. Yu et al., "Combining convolutional neural networks and Self-attention for fundus diseases identification," Scientific Reports, vol. 13, no. 1, pp. 1–15, 2023.
34. Jena, Pradeep Kumar, Bonomali Khuntia, Charulata Palai, Manjushree Nayak, Tapas Kumar Mishra, and Sachi Nandan Mohanty. "A novel approach for diabetic retinopathy screening using asymmetric deep learning features." *Big Data and Cognitive Computing* 7, no. 1 (2023): 25.