<u>ARTICLE</u>

Deep Transfer Learning Models for Mobile-Based Ocular Disorder Identification on Retinal Images

Roseline Oluwaseun Ogundokun[1,2]**, Joseph Bamidele Awotunde**[3]**, Hakeem Babalola Akande**[4]**, Cheng-Chi Lee**[5,6,*] **and Agbotiname Lucky Imoize**[7,8]

[1]Department of Multimedia Engineering, Kaunas University of Technology, Kaunas, 44249, Lithuania

[2]Department of Computer Science, Landmark University, Omu Aran, 251103, Nigeria

[3]Department of Computer Science, Faculty of Information and Communication Sciences, University of Ilorin, Ilorin, 240003, Nigeria

[4]Department of Telecommunication Science, University of Ilorin, Ilorin, 230003, Nigeria

[5]Department of Library and Information Science, Fu Jen Catholic University, New Taipei City, 24205, Taiwan

[6]Department of Computer Science and Information Engineering, Fintech and Blockchain Research Center, Asia University, Taichung City, 41354, Taiwan

[7]Department of Electrical and Electronics Engineering, Faculty of Engineering, University of Lagos, Akoka, Lagos, 100213, Nigeria

[8]Department of Electrical Engineering and Information Technology, Institute of Digital Communication, Ruhr University, Bochum, 44801, Germany

*Corresponding Author: Cheng-Chi Lee. Email: cclee@mail.fju.edu.tw

ABSTRACT

Mobile technology is developing significantly. Mobile phone technologies have been integrated into the healthcare industry to help medical practitioners. Typically, computer vision models focus on image detection and classification issues. MobileNetV2 is a computer vision model that performs well on mobile devices, but it requires cloud services to process biometric image information and provide predictions to users. This leads to increased latency. Processing biometrics image datasets on mobile devices will make the prediction faster, but mobiles are resource-restricted devices in terms of storage, power, and computational speed. Hence, a model that is small in size, efficient, and has good prediction quality for biometrics image classification problems is required. Quantizing pre-trained CNN (PCNN) MobileNetV2 architecture combined with a Support Vector Machine (SVM) compacts the model representation and reduces the computational cost and memory requirement. This proposed novel approach combines quantized pre-trained CNN (PCNN) MobileNetV2 architecture with a Support Vector Machine (SVM) to represent models efficiently with low computational cost and memory. Our contributions include evaluating three CNN models for ocular disease identification in transfer learning and deep feature plus SVM approaches, showing the superiority of deep features from MobileNetV2 and SVM classification models, comparing traditional methods, exploring six ocular diseases and normal classification with 20,111 images post-data augmentation, and reducing the number of trainable models. The model is trained on ocular disorder retinal fundus image datasets according to the severity of six age-related macular degeneration (AMD), one of the most common eye illnesses, Cataract, Diabetes, Glaucoma, Hypertension, and Myopia with one class Normal. From the experiment outcomes, it is observed that the suggested MobileNetV2-SVM model size is compressed. The testing accuracy for MobileNetV2-SVM, InceptionV3, and MobileNetV2 is 90.11%, 86.88%, and 89.76% respectively while MobileNetV2-SVM, InceptionV3, and MobileNetV2 accuracy are observed to be 92.59%, 83.38%, and 90.16%,

respectively. The proposed novel technique can be used to classify all biometric medical image datasets on mobile devices.

KEYWORDS

Retinal images; ocular disorder; deep transfer learning; disease identification; mobile device

1 Introduction

Numerous applications, including image classification, object identification, and segmentation, have seen considerable advancements because of the deep learning (DL) model [1,2]. DL-based methods have been effectively translated into mobile phones for precise user authentication thanks to the revolution in mobile technology and the increasing use of biometric technologies, deeply linked auto-encoders [3]. Among other DL-based techniques, convolutional neural networks (CNN) and deep sparse filtering-based feature representation [4] have both been suggested for user identification using biometrics in mobile settings [5]. The study aims to use deep transfer learning, which is most typically used to evaluate visual imagery in camera-based mobile biometrics. Various hidden layers comprise a conventional CNN-based model consisting of input nodes and output neurons, among others. Among these hidden layers are polling layers, network normalization, convolutional layers, and fully linked levels of the CNN-based model. Existing CNN models, 1.2 million photos from the ImageNet collection, were used to train algorithms like VGG [6] and ResNet [7], employing transfer learning, and have been repurposed to gather features and authenticate mobile users [8].

Transfer learning is the process of transferring knowledge from one task to another. By applying the learned feature layers from the startup of one job to a pre-trained CNN, transfer learning may be used for CNNs. In most of these studies [5,9,10], training and validation are conducted on the same participants when doing restricted validation or recognition. Classifiers relying on softmax, sparsity, or machine learning are used for decision-making. Hence, the network can be tailored for specific dataset issues. Similarly, prior studies have never investigated the computational costs of projected mobile device application models [11]. For example, pre-trained models, like VGG, are too large and need too many operations to be used effectively on a mobile device.

Model size and computing cost are essential for frequent and real-time mobile device biometric authentication. This vast model may use that memory with a high computational price, running more slowly than a small model with a low computational cost. For the best user experience, anything that can be used frequently, like unlocking a phone, is necessary. These features are required for learnable layers like convolutional and connected layers. With more parameters in the learnable layers, the model grows in size. Additionally, multiply-add (MAdd) operations form the foundation for operations like convolution and matrix multiplication in learnable layers.

The challenges associated with ocular disorder identification include the need for accurate and timely diagnosis of various eye conditions such as age-related macular degeneration (AMD), Cataracts, Diabetes, Glaucoma, Hypertension, Myopia, and Normal, which require specialized expertise and equipment. Additionally, the interpretation of retinal fundus images for disease identification can be complex and subjective, often requiring extensive training and experience. The clinical motivation driving our research lies in addressing these challenges by developing a robust and efficient

deep-learning model for the identification of ocular disorders from retinal images. By leveraging advanced machine learning techniques, we aim to provide healthcare practitioners with a reliable tool for early detection and classification of eye conditions, facilitating timely intervention and treatment decisions. This research endeavor ultimately seeks to improve patient outcomes and reduce the burden on healthcare systems by enabling more efficient and accurate diagnosis of ocular disorders.

This study will present a mobile-based ocular disorder identification method using retinal images and assess its usefulness for mobile biometric user verification. Additionally, we demonstrate and benchmark our original, compact DL-based model, indicating that it performs similarly to other DL-based models while using a smaller model size and less computational cost.

Therefore, the study contributions are as follows:

a) The study evaluates the performance of 3 CNN models for ocular disease identification in transfer learning and deep feature plus SVM approach.
b) The statistical analysis results show that the deep features of MobileNetV2 and SVM classification models are superior to others.
c) A comparative analysis uses all classification models based on CNN and traditional methods.
d) The study is conducted on six octal diseases and a normal with a dataset of 20,111 after data augmentation.
e) The proposed model requires fewer trainable parameters due to using pre-trained weights across all layers of the MobileNetV2 architecture. This enhances its suitability for implementation on devices with limited resources, such as mobile devices.
f) Our proposed model's end-to-end training and deployment eliminates the need for distinct feature extraction and classification phases, which are present in conventional machine learning approaches.

2 Related Works

Several research studies have been conducted in computer vision, image processing, classification, and disease detection [12]. Some researchers have researched in some areas, such as the role of a transcriptional factor in attenuating a specific medical condition [13] and object detection [14–16]. Researchers [17] developed a method for estimating and tracking surgical instrument posture. In this study, we have discussed some related works in the ocular eye disease detection field in this section.

Li et al. [18] employed the DL approach to fine-tune the ImageNet-pre-trained VGG-16 system and examined its effectiveness on a test dataset. The receiver-operating characteristic (ROC), predictive accuracy, sensitivity, and specificity were then determined. The findings from experiments demonstrated that the suggested technique exhibited outstanding outcomes in detecting macular OCT photographs, with an estimated accuracy of 98.6%, a sensitivity of 97.8%, a specificity of 99.4%, and an ROC of 100%.

Pin et al. [19] investigated a practical DL framework that recognizes three main optical conditions: diabetic retinopathy (DR), AMD, and glaucoma (GLC). The procedure recommended consists of two stages. First, a new approach involving a preliminary qualitative assessment in the algorithm for classification to separate low-quality images and the technique's efficiency is recommended. The transfer learning approach is also implemented to train different CNN algorithms that naturally discover multiple attributes in the digital retinal image and then employ those traits to diagnose eye disorders.

Khan et al. [20] sought to build a DL strategy capable of predicting structural characteristics from fundus photographs and identifying the most effective approach for building models for this endeavor. Within March 2020 and March 2021, data was gathered from a group of individuals undertaking the DR examination. Building on the DenseNet201 structure, two distinct techniques were developed for every single one of 12 systemic medical attributes: one using TL with data obtained from ImageNet and a different approach with 35,126 fundus photographs. One thousand two hundred seventy-seven fundus photographs have been employed to develop AI algorithms in this study.

Guo et al. [21] projected the use of a lightweight DL strategy called MobileNetV2 and TL to differentiate four prevalent eye disorders, including Glaucoma, Maculopathy, Pathological Myopia, and Retinitis Pigmentosa, from typical conditions employing barely noticeable training samples. To make the framework further explicable, the researchers additionally use an imaging method that emphasizes regions with the most robust connections with the condition's classifications. The stressed area selected by the technique may provide clues to additional fundus photograph research. Five separate evaluations of their experiments indicate that the approach obtains an overall accuracy of 96.2%, a sensitivity of 90.4%, and a specificity of 97.6% on the test data, outperforming the two additional DL-powered strategies concerning accuracy and effectiveness.

Cen et al. [22] established DL strategies for identifying several typically identifiable fundus conditions and disorders (39 classes) utilizing 249,620 fundus photographs annotated with 275,543 annotations from different databases. The recommended DL methodologies obtained an intensity-weighted typical F1-score of 0.923, sensitivity of 0.978, specificity of 0.996, and ROC of 0.9984 for multiple labels in the primary evaluation dataset. They attained the requirements of retina professionals on average.

Devi et al. [23] suggested a TL methodology with data augmentation approaches and Gaussian-blur, circle-crop preprocessing procedures to recognize each phase of DR utilizing ResNet50 with topmost tiers.

Chen et al. [24] proposed a computer-based classification of retinal surface conditions using highly interconnected CNN architecture and mobile device imagery. They acquire normal and abnormal medical pictures using multiple mobile phones and customize end-to-end tightly coupled CNN utilizing a combination of units to discover a broader range of attributes, thus substantially decreasing the algorithm's dimension, the aggregate number of variables, and float computation. The testing outcomes indicate that the approach the investigated suggested is efficient yet prospective for effective screening for ocular surface abnormalities. More specifically, the recommended method for classifying smartphone images utilizing extensive machine learning obtained an average automatic identification accuracy of 90.6%.

In 2023, Sajid et al. [25] created a MobileNet architecture using transfer learning (TL) and dense blocks to enhance the network's performance in diagnosing HR eye-related diseases. The researchers developed a compact method for diagnosing eye diseases connected to human resources, called Mobile-HR, by combining a pre-trained model with dense blocks. The authors used a data augmentation strategy to expand the size of the training and test datasets. The results of the trials demonstrate that the proposed method was surpassed in several instances. The Mobile-HR system showed a 99% accuracy and an F1-score of 0.99 across several datasets. An experienced ophthalmologist validated the findings. The findings suggest that the Mobile-HR CADx model achieves favourable outcomes and surpasses state-of-the-art HR systems in accuracy.

Wahab Sait [26] sought to create a low-resource deep-learning (DL)-based DR-severity assessment system. The model used picture preprocessing to reduce fundus image noise and artefacts. Yolo V7

feature extraction was recommended. It provided feature sets. The author selected attributes using a customized quantum marine predator algorithm (QMPA). The hyperparameter-optimized MobileNet V3 model predicted severity levels from photos. The author used APTOS and EyePacs datasets to generalize the model. EyePacs has 35,100 fundus pictures, whereas APTOS has 5590. The suggested model has an accuracy of 98.0 and 98.4 and an F1-score of 93.7 and 93.1 in the APTOS and EyePacs datasets, respectively. The DR model used fewer parameters, FLOPs, learning rate, and training time to learn fundus picture key patterns.

El Harti et al. [27] suggested a hybrid method integrating four machine-learning models with a traditional neural network to categorize these eye diseases. Our method extracts features using the ImageNet dataset's pre-trained VGG16 model. The retrieved characteristics are then fed into the ML models. Ocular recognition, retinal datasets, the Indian Diabetic Retinopathy Image Dataset, and other freely accessible Kaggle databases were used to conduct this study.

To overcome these issues, Jeny et al. [28] developed an automated deep learning-based ensemble technique to identify and categorize eye disorders using fundus pictures. The ensemble technique suggests a 20-layer CNN-based model with activation, optimization, and loss functions. Preprocessing uses contrast-limited adaptive histogram equalization (CLAHE) and Gaussian filter to improve picture clarity and reduce noise. Training overfitting is avoided using augmentation procedures. Three pre-trained CNN models—VGG16, DenseNet201, and ResNet50—are used to evaluate the proposed CNN model. Experimental findings show that the ensemble strategy outperforms current state-of-the-art approaches in the ODIR dataset.

Thiagarajan et al. [29] suggested an intelligent system to aid healthcare workers in this endeavour to make it more efficient. To detect eye diseases, this research indicates a system that uses trained Neural Networks to classify retinal fundus pictures (RFI). In testing, ODDS attains a 93% accuracy rate with good precision, recall, and F1-scores for all classes using an EfficientNetB3 model with a dataset of 5600 pictures across five categories. Community screening programs, particularly in underserved regions, may benefit from this approach; however, to make it more reliable and accessible, more data augmentation and investigation into smartphone integration are recommended.

Table 1 summarizes the related works and demonstrates the limitations of the studies that motivated this present study.

Table 1: Summary of related works

Authors	Year	Methods	Limitations
Li et al. [18]	2019	VGG16	They only tested a single test dataset and did not access their technique with existing techniques.
Pin et al. [19]	2022	TL techniques with many CNN approaches	The investigators only considered three distinct eye disorders, limiting the generalization to other eye defects.

(Continued)

Table 1 (continued)

Authors	Year	Methods	Limitations
Khan et al. [20]	2022	DenseNet201	They only considered forecasting structural attributes from fundus scans, limiting the suggested approach's applicability to other areas of ophthalmology. They did not perform a comparative analysis with existing studies.
Guo et al. [21]	2021	MobileNetV2	Here, the scholars only investigated four eye conditions, limiting the approach's generalization to other eye diseases.
Cen et al. [22]	2021	Many DL approaches	The study considered recognizing particular fundus diseases, which limited its generalization to other domains of ophthalmology, and they did not evaluate the present research and existing ones.
Devi et al. [23]	2022	ResNet50	The investigator researched to recognize every phase of DR, limiting its applicability to other ears of ophthalmology.
Chen et al. [24]	2021	CNN technique with mobile devices	The authors only offered to categorize retinal surface conditions, limiting its applicability to other ears of ophthalmology.
Sajid et al. [25]	2023	MobileNet with transfer learning	Lack of discussion on potential biases in dataset selection.
Wahab Sait [26]	2023	Yolo V7, quantum marine predator algorithm (QMPA), MobileNetV3	There is limited discussion on the scalability and generalizability of the proposed model beyond the datasets used.
El Harti et al. [27]	2024	VGG16 model, four machine-learning models	Insufficient details on the integration process of different ML models and potential challenges in model fusion.
Jeny et al. [28]	2023	20-layer CNN-based model, contrast-limited adaptive histogram equalization (CLAHE) and gaussian filter	Limited discussion on computational resources required for training and inference of the proposed ensemble model.

(Continued)

Table 1 (continued)

Authors	Year	Methods	Limitations
Thiagarajan et al. [29]	2023	EfficientNetB3 and neural network	Lack of comparison with alternative ensemble techniques or single-model approaches to evaluate relative performance advantage.

The presented related works illustrate numerous research that use deep learning (DL) algorithms to diagnose eye problems, each with constraints such as lack of generalization, data inefficiencies, interpretability difficulties, challenges in real-world deployment, and ethical implications. The research concentrates on particular eye disorders, demonstrating good accuracy but lacking general relevance and scalability due to dataset constraints. Therefore, the motivation for additional research includes developing DL models that generalize well across diverse conditions, increasing data efficiency through transfer learning and data augmentation techniques, improving interpretability and explainability, refining models for real-world deployment, and addressing ethical concerns such as bias and privacy issues in healthcare AI applications.

Hence, we decided to solve the problem of data inefficiencies and data imbalances. The study suggested a model that combined the use of a mobile pre-trained model, which was hybridized with a machine learning technique SVM for the classification process. The study involved the use of transfer learning and data augmentation to enhance data efficiency. The study also used a dataset with seven classes against the related works: four, three, etc. The effectiveness of the proposed approach was also evaluated using measures like accuracy, specificity, recall, F1-score, etc.

The rationale for employing mobileNetV2 PCNN is that it is intriguing for mobile and embedded vision services owing to its lightweight design, high accuracy, and swift rapidity. SVM was utilized with MobileNetV2 since it offers enhanced accuracy, more effective handling of small datasets, more accurate comprehensibility, and decreased computation time, making it suitable for mobile phone and tablet implementation with only a few resources.

3 Materials and Methods

3.1 Dataset

The effectiveness of DL-based techniques depends on the availability of a valid and appropriate dataset. The research in question makes use of the ODIR5K dataset. The ocular disease dataset consists of 6392 images. The information was collected from the Kaggle repository. It comprises seven classes of ocular diseases, which include AMD with 266 photographs, cataract with 293 photos, diabetes with 1608 images, glaucoma with 284 illustrations, hypertension with 128 images, myopia with 232 images and Normal (without any ocular diseases) with 2873 images.

Considering the unbalanced data from these eight classifications, 6392 photos of ocular disorders were added to the Kaggle library to address the class imbalance problem. The images were then enhanced using various data augmentation methods like horizontal flipping, resizing, network shifting, width shift, and rotation that were transformed to 20,111. The class imbalance issue was addressed in the dataset using 2873 images across all classes. All of the images of the ocular classes were randomly

chosen from the entire collection of images. Fig. 1 presents samples of ocular eye diseases, and Table 2 provides class information.
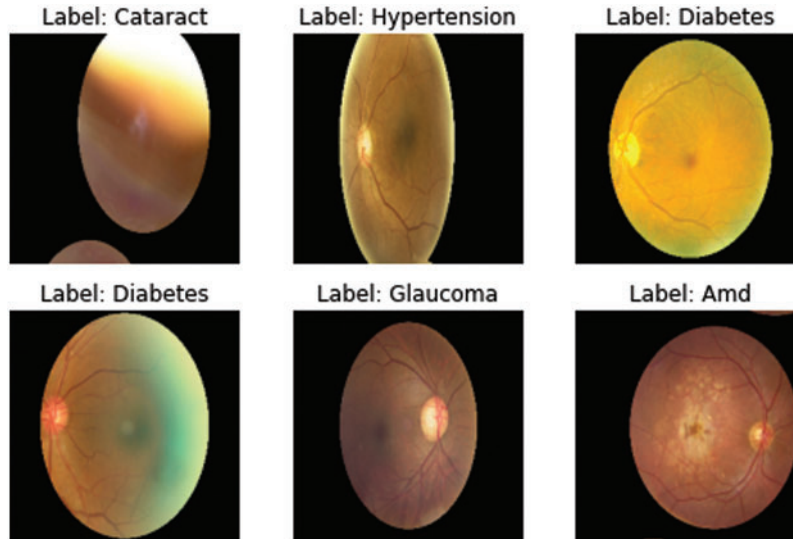

Figure 1: Samples of ocular eye diseases

Table 2: Summary of the ocular disorder dataset

Classes	Before augmentation	After augmentation
AMD	266	2873
Cataract	293	2873
Diabetes	1608	2873
Glaucoma	284	2873
Hypertension	128	2873
Myopia	232	2873
Normal	2873	2873
Total	**6392**	**20,111**

3.2 Image Preprocessing

Image preprocessing is used for all input images of the ocular condition to get more consistent findings with better characteristics. A large-scale picture dataset was needed for the PCNN approach's recurrent training requirements to avoid the risk of over-fitting. The original ocular disorder dataset contains all images in various sizes, including 1956×1934, 2592×1728, 1444×1444, and others. Hence, the dataset is resized to 224×224. The model performance will drastically reduce, and the processing time will increase.

3.2.1 Image Augmentation

With the help of the Keras library's Python image data generation method, various picture augmentation techniques have been used in training sets to combat overfitting and boost the dataset's heterogeneity [30]. The scale transformation was performed using a lower computing cost, and pixel values were of the same range. Therefore, using the (1/255) value parameter, each pixel's value ranged from 0 to 1. By utilizing rotations, photos were turned to a particular angle revolution. Hence, 40 pixels were picked. The dimension shift ranging transition was used to set the width change to 0.2, and images can be spontaneously moved to the right or left. The training images were vertically shifted with a setting of 0.02 for the altitude shifting range variable. One dimension of the picture is constant using the shear transformation technique, while the other axis is strained at a 0.2 shear viewpoint.

A value greater than 1.0 image was zoomed in using the zoom range option to do the random zoom transformation, and if it was less than 1.0, they were zoomed out. Hence, the image was enlarged using a 0.2 zoom range. The image was horizontally flipped using the flip command. The zoom range was set to 0.5–1.0 using brightness transformation if 1.0 represents the brightest possible light and 0.0 represents no light, and the horizontal flip setting was set to true. Because channel shift transformation uses a 10-channel shift range, a value chosen from the provided range is applied randomly to alter the channel values and then reflect fill mode, as shown in Table 3.

Table 3: Image augmentation approaches

Approach	Setting
Rotation_range	40
Width_shift	0.2
Height_shift_range	0.2
Shear_range	0.2
Zoom_range	0.2
Horizontal_flip	True
Channel_shift_range	10
Fill_mode	Reflect

3.2.2 Training, Validation, and Testing

The dataset for ocular disorders was separated into three stages: training, confirmation, and assessment components. The introduced MobileNetV2-SVM model's performance was assessed using the validation and test datasets, while the training set was used to train the PCNN models. Therefore, the authors divided the dataset into training, testing, and validation, with weights of 80%, 10%, and 10%. The 10% of the dataset for the test set was present across the board (20,111). The remaining dataset (18,100) was applied to the models' learning, and 10% of the validation set was used. To train the MobileNetV2-SVM model, the dataset was used. In addition, 16,290, 1811, and 2012 images were employed for the ocular dataset training, validation, and testing, respectively.

3.3 Deep Transfer Learning Methods

TL Once DCNNs in 2012, TL became the foundation of state-of-the-art (SOTA) object identification models by recycling and perfecting visual information gained in DCNNs (CV) [31,32]. Most

DL-based models tested CNNs trained on the vast ImageNet are used in this work [33]. 1.2 million training photos make up this dataset, which is typical for large-scale image categorization.

- **InceptionV3**

There are 42 layers in the Inception module that [34] has proposed. InceptionV3 is the third version of Google Brain's Convolution layer suggested, comprising 159 layers. The primary goal of the Inception module is to merge tiny kernels with large kernels to learn multi-scale models, simplify computation, and use fewer parameters overall.

- **MobileNetV2**

A CNN design for mobile devices called MobileNetV2 was proposed by [35] in 2018. Its initial design also included facial, developed and validated using Google's internal datasets for feature identification [36]. In addition, they balance inference time and performance for popular benchmarks like ImageNet by introducing linear constraints and inversion residuals, and they obtain cutting-edge results, COCO [37], and VOC [38]. Three hundred million MACs and 3.47 million parameters make up our version of MobileNetV2.

- **Xception**

Francois Chollet expanded the Inception architecture by introducing the Xception architecture [39]. This design comprises a linear stacking of depth-wise separated convolution layers with feedback connections. The depthwise separable convolution attempts to lower memory and computational costs. All 14 modules in Xception's 36 convolutional layers have linear residual relationships except for the first and last modules—the detachable separate the learning of stream and space-wise characteristics convolution in Xception. The residual connection developed by authors in [40] also provides a sequencing network approach for dealing with vanishing gradients and symbolic constraints. This approach connection uses a summing operation rather than a concatenation to render an initial layer's result accessible as data to a subsequent layer.

3.4 The Proposed Model

The custom model is based on MobileNetV2 and SVM architectures. The MobileNetV2 was used as a feature extractor to process the ocular biometric disorder datasets. Next, the deep features returned were classified using a Support Vector Machine. The proposed model has only 34,374,727 total parameters. Our suggested model has the fewest variables of the other three PCNN models. Fig. 2 shows the architectural diagram for the proposed MobileNetV2-SVM with their layers. First, the ocular dataset is uploaded or input, after which the MobileNetV2-SVM model is used for training, and the classification result is then output, as seen in Fig. 2.

Four transfer learning techniques are put forth in this work to classify images of ocular diseases. The pre-trained CNN models are fed the image using the first technique, which then uses the outputs from various layers to extract image features. Once these features have been classified, the SVM classification method is used. The second strategy alters the classification layer of the current designs and trains the neural network's filter weights. Finally, the pre-trained CNN is used to initialize the filter settings in each layer except for the final classification stage, whose parameters are set using Gaussian-distributed random integers. Table 4 shows the parameters for training, validating, and testing the suggested approach. Fig. 3 shows the proposed approach flow diagram. Integrating MobileNetV2 and SVM capabilities can result in a robust, cost-effective, reliable classification framework, as shown in Fig. 3.
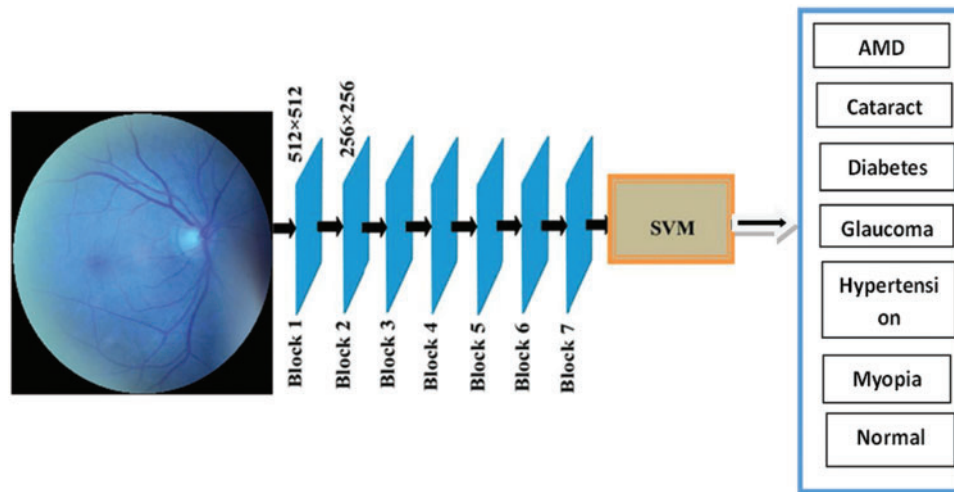
Figure 2: MobileNetV2-SVM architecture

Table 4: Hyperparameter for proposed model

MobileNetV2	SVM
a) Input size $224 \times 224 \times 3$	a) Kernel type: In this, we used linear higher dimensional feature space
b) Depth multiplier: This is set between 0 and 1	b) C parameter: used to penalize misclassifications during training
c) Number of classes: 7	c) Gamma parameter: In this study, RBF kernel was used
d) Learning rate: 0.001	
e) Optimal: Adam	

The research methodically chooses techniques by taking into account various crucial factors, including the limitations of mobile devices in medical imaging, the need for efficient computations that do not compromise precision, and the indispensability of dependable performance across a wide range of biometric image datasets. The selection of MobileNetV2 as the foundational CNN architecture is based on its lightweight, which is accomplished via depthwise separable convolutions. This guarantees smooth functionality on mobile platforms without compromising the ability to differentiate. The classification capabilities of MobileNetV2 are improved by integrating a Support Vector Machine (SVM), which capitalizes on the SVM's ability to generate high-dimensional decision boundaries. By leveraging the feature extraction capabilities of MobileNetV2 and the decision-making capabilities of SVM, this integration strikes a harmonious equilibrium among computational efficiency, discriminatory capability, and interpretability. In essence, this astute decision is the foundation for the effectiveness and adaptability of the MobileNet-SVM method suggested for classifying biometric images on mobile devices; it effectively tackles the obstacles of mobile medical imaging.

The suggested technique incorporates the Quantizing pre-trained CNN (PCNN) MobileNetV2 and a Support Vector Machine (SVM) into our deep transfer learning architecture. MobileNetV2 was chosen for its efficiency in mobile-based applications, which uses depthwise separable convolutions to minimize computational complexity while retaining feature extraction capacity. Quantization methods are used to reduce model size and computing overhead to improve efficiency further. The SVM model,

a popular classification technique, enhances MobileNetV2 by creating high-dimensional decision limits for reliable classification. The SVM is equipped with a radial basis function (RBF) kernel and tweaked using parameters such as C (regularization parameter) and gamma (kernel coefficient). These parameters are carefully calibrated to ensure peak performance in eye disease detection activities on mobile devices. By merging PCNN MobileNetV2 with SVM, the proposed model provides interoperability with mobile platforms while improving suitability for real-time classification. This integration is a cornerstone, allowing for the quick and reliable diagnosis of ocular illnesses across various mobile imaging datasets.
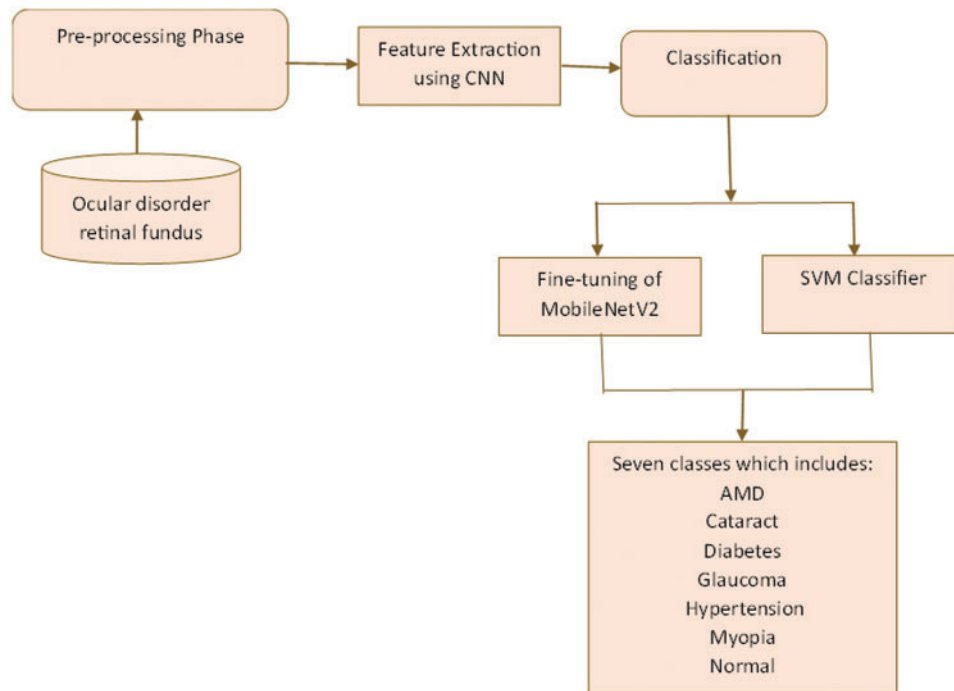
Figure 3: Proposed MobileNetV2 with SVM framework

The proposed method flow diagram is also presented in Fig. 4. The following procedure demonstrates how this can be accomplished:

Step One: Import the dataset of ocular fundus images into the system.

Step Two: Perform preprocessing duties on the dataset, such as normalization.

Step Three: Apply the MobileNetV2 method to extract features from the input images.

Step Four: Fine-tune MobileNetV2 to improve system performance.

Fifth Step: The output of MobileNetV2 is a set of feature vectors.

Sixth Step: The feature vectors are then sent to an SVM classifier trained on the features obtained to acquire knowledge of the decision boundary between various data classes.

Step Seven: During inference, new images are first processed by the MobileNetV2 method to extract features, and then the SVM classifier is used to predict the class labels of the input image.
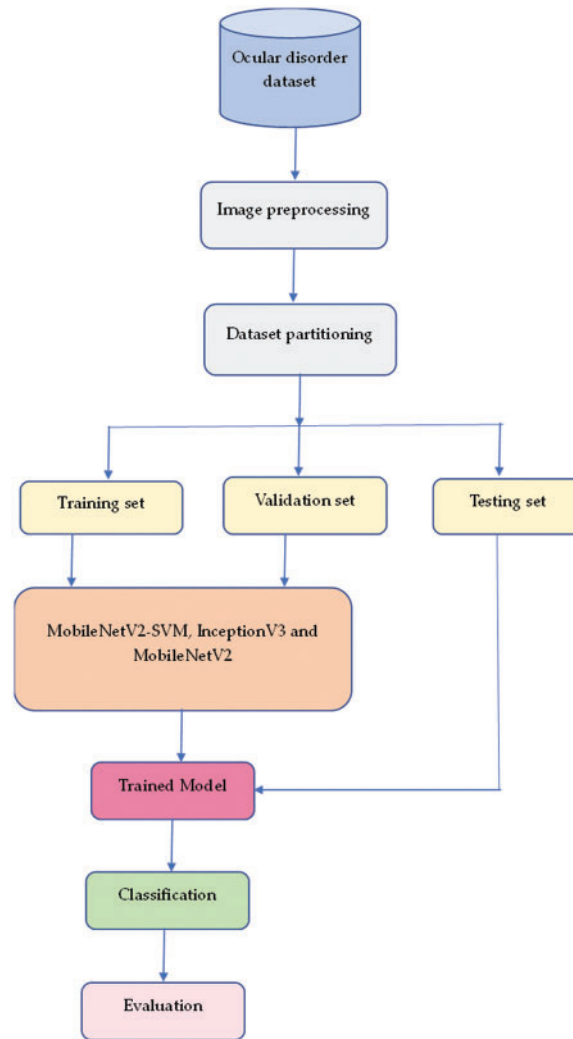
Figure 4: Proposed method flow diagram

3.5 *Fine Tuning*

When fine-tuning, we suggested training the filter weights of the most recent few layers and fixing the filter parameters from earlier layers. There are two factors at play in this learning adaptation. To begin with, the dataset used for transfer learning is modest compared to ImageNet, where the pre-trained models were developed. Therefore, fine-tuning the entire neural network may not be the best action given the tiny dataset and limited data. Secondly, the shallow portion of CNN designs typically learns filters that correlate to more generic visual properties, including edges and corners. They are acknowledged as fixed characteristics for photographs from many application domains. The parameter used to implement the proposed model is shown in Table 5.

Table 5: Parameters used for the implementation

Parameters	Values
The proposed model	MobileNetV2-SVM
Transfer type	Freezing the top layers of the pre-trained model
Loss function	CategoricalCrossentropy
Batch size	64
Generate layers	All
algorithm for learning	Adam
Learning rate	Default alpha rate
Function of activation	Relu & softmax
Epochs	10

3.6 The Performance Evaluation Metrics for the Proposed Models

The proposed technique was assessed following the training procedure on the testing dataset. Various performance metrics, such as accuracy, recall, precision, F1-score, and ROC-AUC score, were used to test the proposed architecture's performance. The following is a detailed analysis of the performance metrics used in this study [41]. Here are the definitions and equations for the performance metrics: TP signifies the True Positives, TN represents the True Negatives, FN signifies False Negatives, and FP signifies the False Positives.

- **Accuracy**

According to Eq. (1), classification accuracy is calculated as the number of precise forecasts per accurate estimate.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

- **Precision**

This indicator shows the proportion of data instances that the model correctly predicts to be typical. Divide the overall number of true positives by the whole number of true positives plus false positives, and it is calculated. It is computable mathematically by (2):

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

- **Recall**

Another crucial metric is recall; this represents how the input samples were divided into the classes that the algorithm correctly predicted. Recall eligibility is decided by (3):

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

● **F1-score**

A popular statistic that combines recall and precision measurements is the F1-score. The F1-score is determined by (4):

$$F1 - score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \tag{4}$$

● **AUC score and ROC Curve**

A reliable method that balances sensitivity and specificity is using ROC curves, which are based on TP and FP rates.

4 Results and Discussion

On Jupiter Notebook, an experiment using the proposed MobileNetV2-SVM architecture was conducted. The proposed model used the Python programming language, the open-source Keras tools, and the Tensor-Flow platform to develop the MobileNetV2-SVM approach. It uses the Adam optimizer for training with a CategoricalCrossentropy loss function and a preset learning rate. The computational details used to implement the suggested strategy involve utilizing a laptop system with an intel CORE i7 vPro, 16 GB RAM, and 1Terabyte hard disk.

4.1 Performance of Proposed Model on the Ocular Disorders Database

To assess the effectiveness of the suggested model, various experiments were performed on the recently released MobileNetV2-SVM framework. As can be seen in Table 6, the experiment employed the Categorical Cross entropy loss function, the Adam optimizer, 64 batches, ten iterations, and the default alpha rate. Table 6 displays the research findings introduced approach had 92%, 99%, 65%, 92%, 98%, 98%, and 84% accuracy for AMD, cataract, diabetes, glaucoma, hypertension, myopia, and normal, respectively. Myopia did the best of all the ocular disorders, scoring 100% accuracy, 97% recall, and 98% on the F1-score.

Table 6: Performance metrics of the proposed MobileNetV2-SVM on the ocular dataset

Classes	Precision	Recall	F1-score
AMD	0.99	0.92	0.95
Cataract	0.98	0.99	0.98
Diabetes	0.79	0.67	0.73
Glaucoma	0.98	0.93	0.95
Hypertension	0.99	0.97	0.98
Myopia	1.00	0.97	0.98
Normal	0.66	0.85	0.74

Fig. 5 depicts the outcomes of the planned framework during training and validation for each epoch for the accuracy and losses. The findings show that the training accuracy is high, and validation grew speedily after the first epoch and stabilized after almost eight. However, after two epochs, training and validation losses started to drop off quickly, and after four, they stabilized. When the training set

was subjected to the data augmentation approaches, the findings indicated that the proposed technique obtained superior classification scores on the ocular dataset.
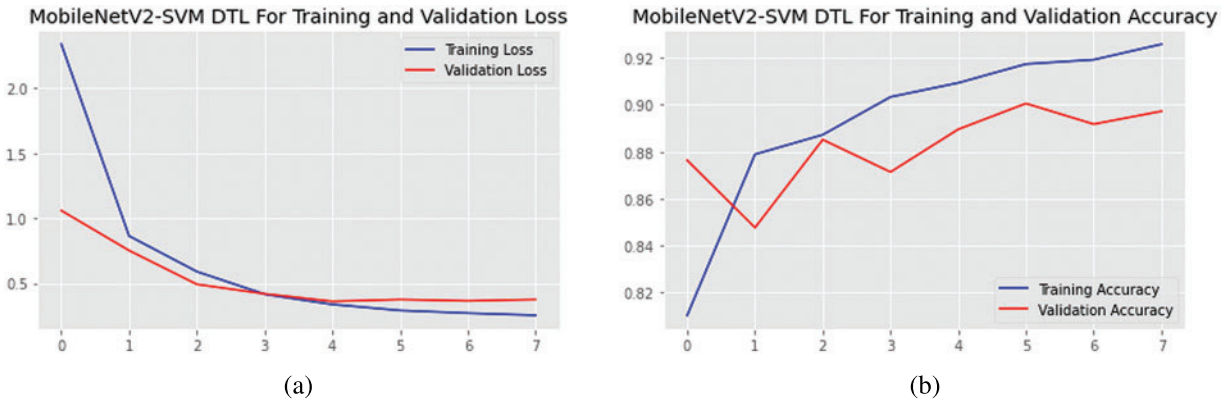

(a)   (b)

Figure 5: (a) Accuracies and (b) loss graphs of the proposed MobileNetV2-SVM model

The confusion matrix is a proper DL technique that assesses a model's precision, recall, accuracy, and ROC curve. The categorization accuracy was measured graphically using a confusion matrix. Darker hue denoted the MobileNetV2-SVM's more accurate categorization of the correct class, while lighter colour denoted erroneously recognized samples. In the confusion matrix, accurate forecasts were shown diagonally, and inaccurate guesses were shown off-diagonally. The outcomes in Fig. 6 show that the current MobileNetV2-SVM architecture performed better when the ocular dataset underwent data augmentation methods. It showed 263 AMD images out of 286; 271 Cataract images out of 275; 184 Diabetes images out of 275; 270 Glaucoma images out of 290; 278 Hypertension images out of 287; 292 Myopia images out of 300; and 255 Normal images out of 299 were accurately classified by the MobileNetV2-SVM model. The introduced MobileNetV2-SVM model was generalized based on the presented MobileNetV2-SVM system's 90% overall accuracy and 10% inaccuracy.
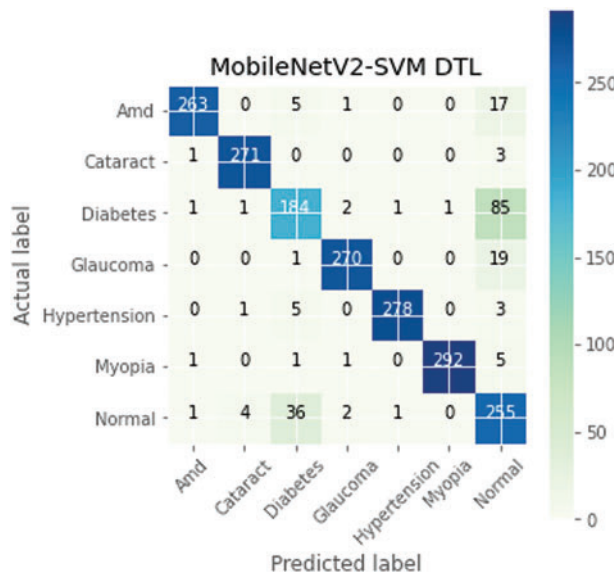

Figure 6: MobileNetV2-SVM model confusion matrix

The proposed model has 94.01% using area under the curve for the validation and test set classes, and the MobileNetV2-SVM model demonstrated remarkable classification performance. The ROC curve shown in Fig. 7 was used to gauge the effectiveness of the proposed model. Asthma, cataracts, diabetes, glaucoma, hypertension, myopia, and normal are represented by classes 0, 1, 2, 3, 4, 5, and 6. Table 7 shows that, with an AUC value of 0.9401, the suggested MobileNetV2-SVM outperformed the other three baseline PCNN approaches.
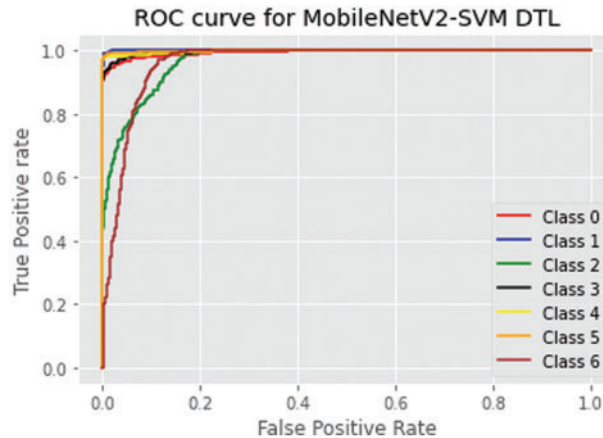
Figure 7: MobileNetV2-SVM ROC curve on ocular dataset

Table 7: AUC values for the proposed MobileNetV2-SVM and other models

Models	AUC
MobileNetV2-SVM	0.9401
InceptionV3	0.8852
MobileNetV2	0.9354
Xception	0.9271

Various performance metrics were used on the proposed model with the data augmentation techniques, and they performed remarkably well on the ocular disorders dataset. In the earlier investigations, the duration of the training and validation phase execution is indicated in Table 6, which is used to evaluate the effectiveness of the proposed model. Train the suggested model over ten epochs. It took about 80 min. InceptionV3 and MobileNetV2 require significantly less computing time than MobileNetV2-SVM. However, compared to other performance progress indicators like sensitivity, specificity, and accuracy, MobileNetV2-SVM showed a higher prediction accuracy. As demonstrated in Table 8, the computational time of the proposed MobileNetV2-SVM is respectably good, making it possible to implement the technology to run over computationally lightweight devices (mobile devices). By retaining the essential features required for the faster and more accurate classification of the dataset for ocular disorders, the SVM module will help with more rapid convergence.

Early stopping was introduced in this study to allow the definition of many training epochs arbitrarily, and the model's training was terminated when the validation dataset performance stopped advancing. In this study, the early stopping callback was set to specify performance measures on which validation accuracy was monitored, and training stopped when the validation accuracy stopped

improving. A decay called patience argument was added to the early stopping approach to have the best model performance before training stops. The early stopping epochs tables are provided for each model (See Table 9).

Table 8: The computational time of MobileNetV2-SVM and other approaches

Model	Computational time
MobileNetV2-SVM	**80 min**
InceptionV3	215 min
MobileNetV2	83 min
Xception	387 min

Table 9: Early stopping approach epochs for each model implemented

Model	Epochs
MobileNetV2-SVM	8
InceptionV3	10
MobileNetV2	7
Xception	8

4.2 Performance Evaluation with Some DTL Models

Table 9 shows the proposed MobileNetV2-SVM technique's training, validation, and testing accuracies, along with the other three baseline PCNN techniques that have been applied. For each of the four PCNN approaches, the losses of training, validating, and testing procedures employed in this study are also shown in the table. The proposed MobileNetV2-SVM performed better than the other baseline techniques, as shown in Table 10, with training, validation, and testing accuracies of 92.59%, 89.73%, and 90.11%, respectively. Regarding training, validation, and testing losses, the suggested method also produced the best results, with values of 0.2535, 0.3056, and 0.3020.

Table 10: The performance evaluation of the proposed models

Models	Training accuracy	Validation accuracy	Testing accuracy	Training loss	Validation loss	Testing loss
MobileNetV2-SVM	**0.9259**	**0.8973**	**0.9011**	**0.2535**	**0.3056**	**0.3020**
InceptionV3	0.8336	0.8752	0.8688	0.4169	0.3241	0.3392
MobileNetV2	0.9016	0.8912	0.8976	0.2545	0.3741	0.3674
Xception	0.9224	0.8835	0.8862	0.2540	0.4119	0.3927

4.3 Comparative Analysis with State-of-the-Arts

We compared the performance of the offered model with cutting-edge methods to represent the generalization of the introduced methodology. It was found that the deep transfer learning system

offered performed better than cutting-edge methods. When comparing the suggested strategy to cutting-edge techniques, the misclassification varied a bit—comparing the efficiency of the developed technique to other previously published classification strategies for ocular disorders allowed for assessment. Table 9 displays the study findings and compares the suggested model to other recent investigations, and the most accurate method was the one that was introduced.

The proposed MobileNetV2-SVM model performed better than the previous research, such as research by authors in [42], which obtained 84.17% accuracy in classifying diabetes diseases on the Pima Indian Diabetes. In [43], the authors used Gradient-weighted Class Activation Mapping on 2012 retinal photographs to detect hypertension diseases and obtained an accuracy of 60.94%. The authors in [44] reported 90.0% accuracy using the InceptionV3 model on an OCT dataset to classify Myopia diseases.

Similarly, authors in [45] used DL techniques, performed with 87.53% accuracy, and used the Advanced Clinical Center for Myopia. The authors in [46] used ResNet18, ResNext50, EfficientNetB0, and EfficientNetB4 for Myopia classification using 367 eyes of patients (Custom dataset) datasets and had a 95% accuracy. We may state that the envisioned MobileNetV2-SVM model outperformed the already-in-use approaches and attained accuracy values for training, validation, and testing of 92.59%, 89.73%, and 90.11%, respectively. It obtained the highest accuracy compared to other models, as shown in Table 11.

Table 11: Comparative analysis with state-of-the-art models

Authors	Method	Disease	Dataset	Accuracy
Kumar et al. [42]	Multi-layer feed forward neural networks (MLFNN)	Diabetes	Pima Indian diabetes	84.17%
Dai et al. [43]	Gradient-weighted class activation mapping	Hypertension	2012 retinal photographs	60.94%
Choi et al. [44]	InceptionV3	Myopia	OCT images	90.0%
Du et al. [45]	DL techniques	Myopia	Advanced clinical center for myopia	87.53%
Park et al. [46]	ResNet18, ResNext50, EfficientNetB0, EfficientNetB4	Myopia	367 eyes of patients (Custom dataset)	95%
Proposed model	**MobileNetV2-SVM**	**AMD, cataract, diabetes, glaucoma, hypertension, myopia, and normal**	**Ocular dataset from the Kaggle repository**	**90.11%**

5 Conclusion

One of the limitations of this work is the lack of sufficient computational resources, which constrained the extent to which we could explore specific aspects of the research, including the number of training epochs and the complexity of the model architecture. With little computing time, power, or effort, the suggested model based on the MobileNetV2-SVM technique effectively classified and

detected ocular disorder datasets. The results show promise, with an accuracy of 90.11% when tested against other approaches using real-time photos from the Kaggle library. A mobile device with a stride2 mechanism can be used with the MobileNetV2-SVM architecture. The model is computationally efficient; combining it with MobileNetV2 might improve forecast accuracy by preserving historical timestamp data. The model would become more reliable by adding information about the present state through weight optimization. It is compared with other widely used models, including InceptionV3, MobileNetV2, Xception, and the base model. Based on the textured-based data shown in the results and discussion section, it is evident that the suggested model has done better than other models in classifying and assessing the progression of ocular illnesses. The model's performance may be enhanced much further by the SVM. When the planned model is used, it requires a lot of work to integrate either model, which links the front end created with the help of the Android Studio/SSDLite/DeepLabv3+ and the business model created using Kaggle. However, there are now many issues that will need to be fixed in subsequent work. When the model's accuracy is tested against a set of images taken in less-than-ideal lighting circumstances that are different from those used during testing, it is noticeably reduced to just under 80%. Ultimately, the proposed technique supports rather than replaces current disease-diagnosis methods. Visual inspection alone frequently poses problems to early identification, and laboratory test results are always more reliable than diagnoses made only based on visual symptoms. In prospect work, we will enlarge the range of ocular disorders research and use data from other institutions to demonstrate the model's external validity.

Further investigation into the proposed method could be fruitful in employing advanced techniques such as attention mechanisms to enhance the model's performance by enabling it to focus on relevant features within the retinal images. Additionally, integrating eXplainable Artificial Intelligence (XAI) techniques will allow us to provide deeper insights into the model's decision-making process, thereby improving its interpretability and trustworthiness in clinical settings. Furthermore, we would test more architectures for detecting ocular disorders by training the pre-trained CNN models on new datasets. In the future, the authors have proposed implementing the proposed model on another set of datasets and doing a cross-dataset analysis assessment to confirm the study's outcome. We also plan to experiment to compare the results between before and after data augmentations.

Author Contributions: The manuscript was written with contributions from all authors. **Roseline Oluwaseun Ogundokun:** Conceptualization, Methodology, Software, Data curation, and Writing–Original draft preparation. **Joseph Bamidele Awotunde:** Data Curation, Investigation, and Methodology. **Hakeem Babalola Akande:** Software, Investigation, and Writing–Reviewing and Editing. **Cheng-Chi Lee:** Supervision, Investigation, and Writing–Reviewing and Editing. **Agbotiname Lucky Imoize:** Methodology, Supervision, and Writing–Reviewing and Editing. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data supporting this study's findings are available at the Kaggle repository with a link: https://www.kaggle.com/datasets/tanjemahamed/odir5k-classification. (accessed on 25/03/2024).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] R. Chen *et al.*, "Dissecting causal associations of type 2 diabetes with 111 types of ocular conditions: A mendelian randomization stud," *Front. Endocrinol.*, vol. 14, pp. 109118, 2023. doi: 10.3389/fendo.2023.1307468.

[2] A. Adegun, S. Viriri, and R. O. Ogundokun, "Deep learning approach for medical image analysis," *Comput. Intell. Neurosci.*, vol. 2021, no. 2, pp. 1–9, 2021. doi: 10.1155/2021/6215281.

[3] R. Raghavendra and C. Busch, "Learning deeply coupled autoencoders for smartphone-based robust periocular verification," in *2016 IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, IEEE, 2016, pp. 325–329.

[4] K. B. Raja, R. Raghavendra, and C. Busch, "Collaborative representation of deep sparse filtered features for robust verification of smartphone periocular images," in *2016 IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, IEEE, 2016, pp. 330–334.

[5] K. Ahuja, R. Islam, F. A. Barbhuiya, and K. Dey, "Convolutional neural networks for ocular smartphone-based biometrics," *Pattern Recognit. Lett.*, vol. 91, no. 1, pp. 17–26, 2017. doi: 10.1016/j.patrec.2017.04.002.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[7] M. Shafiq and Z. Gu, "Deep residual learning for image recognition: A survey," *Appl. Sci.*, vol. 12, no. 18, pp. 8972, 2022. doi: 10.3390/app12188972.

[8] C. Chen, D. Han, and C. C. Chang, "MPCCT: Multimodal vision-language learning paradigm with context-based compact transformer," *Pattern Recognit.*, vol. 147, no. 8, pp. 110084, 2024. doi: 10.1016/j.patcog.2023.110084.

[9] W. Tang, B. He, Y. Luo, and X. Duan, "Morphology and microcirculation changes of the optic nerve head between simple high myopia and pathologic myopia," *BMC Ophthalmol.*, vol. 23, no. 1, pp. 208, 2023. doi: 10.1186/s12886-023-02949-7.

[10] W. Wu, H. Zhu, S. Yu, and J. Shi, "Stereo matching with fusing adaptive support weights," *IEEE Access*, vol. 7, pp. 61960–61974, 2019. doi: 10.1109/ACCESS.2019.2916035.

[11] M. Wang, H. Li, and F. Wang, "Roles of transepithelial electrical resistance in mechanisms of retinal pigment epithelial barrier and retinal disorders," *Discov. Med.*, vol. 34, no. 171, pp. 19–24, 2022.

[12] H. Sheng *et al.*, "Cross-view recurrence-based self-supervised super-resolution of light field," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 12, pp. 7252–7266, 2023. doi: 10.1109/TCSVT.2023.3278462.

[13] H. Jin *et al.*, "Development of proliferative vitreoretinopathy is attenuated by chicken ovalbumin upstream promoter transcriptional factor 1 via inhibiting epithelial-mesenchymal transition," *Discov. Med.*, vol. 34, no. 172, pp. 103–113, 2022.

[14] H. Zhang, G. Luo, J. Li, and F. Y. Wang, "C2FDA: Coarse-to-fine domain adaptation for traffic object detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 12633–12647, 2021. doi: 10.1109/TITS.2021.3115823.

[15] D. Yang, "An occlusion and noise-aware stereo framework based on light field imaging for robust disparity estimation," *IEEE Trans. Comput.*, vol. 73, no. 3, pp. 764–777, 2023. doi: 10.1109/TC.2023.3343098.

[16] B. Cao, J. Zhao, Y. Gu, Y. Ling, and X. Ma, "Applying graph-based differential grouping for multiobjective large-scale optimization," *Swarm Evol. Comput.*, vol. 53, no. 3, pp. 100626, 2020. doi: 10.1016/j.swevo.2019.100626.

[17] S. Lu *et al.*, "Surgical instrument posture estimation and tracking based on LSTM," *ICT Express*, vol. 10, no. 3, pp. 465–471, 2024. doi: 10.1016/j.icte.2024.01.002.

[18] F. Li, H. Chen, Z. Liu, X. Zhang, and Z. Wu, "Fully automated detection of retinal disorders by image-based deep learning," *Graefes Arch. Clin. Exp. Ophthalmol.*, vol. 257, no. 3, pp. 495–505, 2019. doi: 10.1007/s00417-018-04224-8.

[19] K. Pin, J. H. Chang, and Y. Nam, "Comparative study of transfer learning models for retinal disease diagnosis from fundus images," *Comput. Mater. Contin.*, vol. 70, no. 3, pp. 5821–5834, 2022. doi: 10.32604/cmc.2022.021943.

[20] N. Khan et al., "Predicting systemic health features from retinal fundus images using transfer-learning-based artificial intelligence models," *Diagnostics*, vol. 12, no. 7, pp. 1714, 2022. doi: 10.3390/diagnostics12071714.

[21] C. Guo, M. Yu, and J. Li, "Prediction of different eye diseases based on fundus photography via deep transfer learning," *J. Clin. Med.*, vol. 10, no. 23, pp. 5481, 2021. doi: 10.3390/jcm10235481.

[22] L. P. Cen, "Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks," *Nat. Commun.*, vol. 12, no. 1, pp. 4828, 2021. doi: 10.1038/s41467-021-25138-w.

[23] Y. S. Devi and S. p. Kumar, "A deep transfer learning approach for identification of diabetic retinopathy using data augmentation," *IAES Int. J. Artif. Intell.*, vol. 11, no. 4, pp. 1287, 2022. doi: 10.11591/ijai.v11.i4.pp1287-1296.

[24] R. Chen et al., "Automatic recognition of ocular surface diseases on smartphone images using densely connected convolutional networks," in *2021 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Mexico, IEEE, 2021, pp. 2786–2789.

[25] M. Z. Sajid et al., "Mobile-HR: An ophthalmologic-based classification system for diagnosis of hypertensive retinopathy using optimized MobileNet architecture," *Diagnostics*, vol. 13, no. 8, pp. 1439, 2023. doi: 10.3390/diagnostics13081439.

[26] A. R. Wahab Sait, "A lightweight diabetic retinopathy detection model using a deep-learning technique," *Diagnostics*, vol. 13, no. 19, pp. 3120, 2023. doi: 10.3390/diagnostics13193120.

[27] M. El Harti, S. Zaamoun, S. J. Andaloussi, and O. Ouchetto, "Classification of eye disorders using deep learning and machine learning models," in *Int. Conf. Adv. Comput. Res.*, Cham, Springer Nature Switzerland, 2024, pp. 184–194.

[28] A. A. Jeny, M. S. Junayed, and M. B. Islam, "Deep neural network-based ensemble model for eye diseases detection and classification," *Image Anal. Stereol.*, vol. 42, no. 2, pp. 77–91, 2023. doi: 10.5566/ias.2857.

[29] P. Thiagarajan and M. Suguna, "Deep learning ocular disease detection system (ODDS)," in *Int. Conf. Min. Intell. Knowl. Explor.*, Cham, Springer Nature Switzerland, 2023, pp. 213–224.

[30] R. O. Ogundokun, R. Maskeliūnas, and R. Damaševičius, "Human posture detection using image augmentation and hyperparameter-optimized transfer learning algorithms," *Appl. Sci.*, vol. 12, no. 19, pp. 10156, 2022. doi: 10.3390/app121910156.

[31] L. Peng, H. Liang, G. Luo, T. Li, and J. Sun, "Rethinking transfer learning for medical image classification," arXiv preprint arXiv:2106.05152, 2022.

[32] R. O. Ogundokun, R. Maskeliūnas, S. Misra, and R. Damasevicius, "A novel deep transfer learning approach based on depth-wise separable CNN for human posture detection," *Information*, vol. 13, no. 11, pp. 520, 2022. doi: 10.3390/info13110520.

[33] O. Russakovsky et al., "ImageNet large-scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015. doi: 10.1007/s11263-015-0816-y.

[34] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.

[35] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2018, pp. 4510–4520.

[36] A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

[37] T. Y. Lin et al., "Microsoft COCO: Common objects in context," in *Eur. Conf. Comput. Vis.*, Zurich, Switzerland, Cham, Springer, 2014, pp. 740–755.

[38] M. Everingham et al., "The pascal visual object classes challenge: A retrospective," *Int. J. Comput. Vis.*, vol. 111, no. 1, pp. 98–136, 2015. doi: 10.1007/s11263-014-0733-5.

[39] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern, Recogn.*, 2017, pp. 1251–1258.

[40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[41] R. O. Ogundokun, R. Maskeliūnas, S. Misra, and R. Damasevicius, "Hybrid InceptionV3-SVM-based approach for human posture detection in health monitoring systems," *Algorithms*, vol. 15, no. 11, pp. 410, 2022. doi: 10.3390/a15110410.

[42] S. Kumar, B. Bhusan, D. Singh, and D. Choubey, "Classification of diabetes using deep learning," in *2020 Int. Conf. Commun. Signal Process. (ICCSP)*, Chennai, India, IEEE, 2020, pp. 0651–0655.

[43] G. Dai *et al.*, "Exploring the effect of hypertension on retinal microvasculature using deep learning on the east asian population," *PLoS One*, vol. 15, no. 3, pp. e0230111, 2020. doi: 10.1371/journal.pone.0230111.

[44] K. J. Choi *et al.*, "Deep learning models for screening of high myopia using optical coherence tomography," *Sci. Rep.*, vol. 11, no. 1, pp. 1–11, 2021. doi: 10.1038/s41598-021-00622-x.

[45] R. Du *et al.*, "Deep learning approach for automated detection of myopic maculopathy and pathologic myopia in fundus images," *Ophthalmol. Retina*, vol. 5, no. 12, pp. 1235–1244, 2021. doi: 10.1016/j.oret.2021.02.006.

[46] S. J. Park, T. Ko, C. K. Park, Y. C. Kim, and I. Y. Choi, "Deep learning model based on 3D optical coherence tomography images for the automated detection of pathologic myopia," *Diagnostics*, vol. 12, no. 3, pp. 742, 2022. doi: 10.3390/diagnostics12030742.