Received: 5 February 2024 | Revised: 10 June 2024 | Accepted: 16 June 2024

The Journal of Engineering

DOI: 10.1049/tje2.12410

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Computer vision for eye diseases detection using pre-trained deep learning techniques and raspberry Pi

Ali Al-Naji[1,2] | **Ghaidaa A. Khalid**[1] | **Mustafa F. Mahmood**[1] | **Javaan Chahl**[2]

[1]Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq

[2]School of Engineering, University of South Australia, Adelaide, South Australia, Australia

Correspondence
Ali Al-Naji, Electrical Engineering Technical College, Middle Technical University, Baghdad 10022, Iraq.
Email: ali_al_naji@mtu.edu.iq

Abstract

Early diagnosis of eye diseases is very important to prevent visual impairment and guide appropriate treatment methods. This paper presents a unique approach that can detect numerous eye diseases automatically. Initially, this approach used the pre-trained ImageNet models that provides various pre-trained models for training the acquired data. The existing data sets are composed of 645 data images acquired clinically, represented by two groups of subjects as healthy and others holding the proposed eye defect like cataracts, foreign bodies, glaucoma, subconjunctival haemorrhage, and viral conjunctivitis. Followed by comparisons of the pre-trained model's coefficients and prediction performance. Later, the first-class execution model is integrated within the Raspberry Pi staging and the real-time digital camera detection. The evaluation process used the confusion matrix, model accuracy, precision factor, recall coefficient, F1 score, and the Matthews Correlation Coefficient (MCC). Resulting in the performance of these pre-trained ImageNet models used in this study represented by 93% (InceptionResNetV2), 90% (MobileNet), 86% (Residual Network ResNet50), 85% (InceptionV3), 78% (Visual Geometry Group VGG19), and 72% (Neural Architecture Search Network NASNetMobile). The results show that the InceptionResNetV2 achieved the highest performance. This proposed approach shows its efficiency and strength by early detection of the subject's unhealthy eyes through real-time monitoring in the field of ophthalmology.

1 | INTRODUCTION

Eye diseases are considered one of the main factors in non-fatal disabling conditions and pose significant issues for the health of the eye globally. The World Health Organization (WHO) reported that the estimated financial losses resulting from the eye defect caused by diseases are about USD 411 billion [1]. Eye diseases include cataracts, glaucoma, foreign bodies, subconjunctival haemorrhage, and viral conjunctivitis. Cataracts occur when the eye's lenses become cloudy, over time leading to a decrease in clarity and blurry vision [2]. Ocular foreign bodies refer to objects or substances that enter the eye and can cause discomfort, pain, redness and potential harm to the cornea [3]. Glaucoma encompasses eye conditions that damage the nerve. This damage often occurs due to increased pressure within the eye and can result in vision loss if left untreated [4]. A subconjunctival haemorrhage is characterized by patches on the white part of the eye (sclera) caused by blood vessels, in the conjunctiva rupturing [5]. Viral conjunctivitis is an inflammation of the conjunctiva caused by an infection. It leads to redness, itching, and excessive tearing [6]. Traditional methods utilized for the detection of these diseases, like slit lamp biomicroscopy, optical coherence tomography (OCT) and ultrasound biomicroscopy come with limitations. These limitations are known as difficulties in accessing tools, problems associated with costs and precision in performance. Another limitation has been the need for expert input for interpretation. Additional issues include the consumption of time and the ability to monitor in real-time [7–11]. In consequence, modern, effective tactics have become a necessity to greatly improve the identification and classification process of these diseases, helping to reach a better-managed system.

The analysis and classification operation that medical images, including ophthalmic images, undergo, have been possible by

the contributions of machine learning techniques especially deep learning. Deep learning arose in the biomedical sciences as a technique that supported clinical decisions. It achieved this role by developing the sensitivity, specificity and objectivity aspects of the decision-making program, thus making a contribution in the upgrade of disease monitoring and detection to a higher standard [12, 13]. This improvement has been noticeable, particularly in recent years, where noteworthy demonstrations of potential have been expressed by the convolutional neural networks (CNNs) based model ImageNet, a pre-trained deep learning model. An instance of this significant performance was in the study by Bhowmik et al. (2019) [14] where ImageNet pre-trained models were used, including VGG16 and InceptionV3, in the classification of fundus images by utilizing datasets consisting of 4000 fundus images. The outcomes showcased an average prediction accuracy of 0.926 for InceptionV3 and 0.94 for VGG16. Another study by Diaz-Pinto et al. (2019) [15] operating with the pre-trained ImageNet models package (five pre-trained models), including the VGG16, VGG19, the InceptionV3, ResNet50, and Xception by making use of fundus images to achieve instant glaucoma detection abilities. The results based on five public databases (1707 images) showed that the best accuracy achieved was 0.96 using the Xception model. Another research by Smaida and Yaroshchak (2020) [16] used CNN, VGG16 and InceptionV3 models in the classification of eye diseases of varying classes, including diabetic retinopathy, glaucoma, and myopia, using a dataset of 2781 retinal fundus images. The end results showed that CNN achieved a prediction accuracy of 0.715 (lowest), while InceptionV3 achieved the highest accuracy of 0.877 (best). In another research by Raza et al. (2021) [17], the classification of eye diseases and sensing cataracts, glaucoma and retinal diseases was attained by Inception-V4 deep learning model and digital fundus imaging (DFI). The main results displaced the high accuracy of the suggested method, which was 0.96 from 601 fundus images. Additional research by Khan et al. (2021) [16] presented the VGG-19 pre-trained model for detecting cataract showcasing, which detects cataracts from around 800 colours of the fundus images, including 0.97 accuracies. Later, both Gour and Khanna (2021) [18] worked with four pre-trained models of ImageNet, including the ResNet, InceptionV3, MobileNet, and VGG16, that were used to detect various diseases of the eye. These diseases were represented by diabetic retinopathy and glaucoma alongside age-related macular degeneration, myopia, hypertension, and cataracts. This research has been based on the data set of fundus image pairs from Eye Disease Intelligent Recognition (ODIR), collected from 5000 patients through numerous medical centres in China. It was reported that the VGG19 model was validated with a higher accuracy represented by 0.871 compared to the ResNet model which achieved a lower accuracy represented by 0.849. Recently, there were five models represented by DarkNet19, AlexNet, SqueezeNet, Dense-Net201, and ResNet18 were trained in Choudry et al. (2022) [18] study by working on 3726 images of the fundus in order to help in diagnosis of numerous eye diseases. Diseases like age-related macular degeneration, glaucoma, cataracts, diabetic retinopathy, hypertensive retinopathy, and pathological myopia. For improving the performance of the models, techniques like image enhancement and deep features extraction where implemented. It was reported that the DarkNet19 model reached a higher accuracy with 0.936 compared to the SqueezeNet model that showed a lower accuracy represented by 0.875.

Another study by Camera et al. (2022) [19] used ten different pre-trained ImageNet models (VGG19, VGG16, Xception, ResNet50, MobileNet, DenseNet, MobileNet, InceptionResNetV2, InceptionV3, NASNetMobile) to automate the detection of glaucomatous papillae using fundus imaging. The highest accuracy achieved was 0.98 with the VGG19 model, while the lowest was 0.87 with the Xception model when evaluated on a randomized test set. Using the Madrid and Zaragoza test suite for the dataset, the highest accuracy achieved was 0.92 with VGG19, while MobileNet had the lowest accuracy of 0.77. Another study by Zhou (2022) [20] used the M-ResNet model to develop an intelligent system for diagnosing cataracts, subconjunctival haemorrhage, keratitis, and pterygium. The results demonstrated the system's effectiveness, with the best accuracy reaching 0.875. Another study conducted by Rajee et al. (2023) [21] used the InceptionV3 model for the automated prediction of diabetic retinopathy, age-related macular degeneration, glaucoma, haemorrhages, and epiretinal membrane diseases. The study achieved an accuracy of 0.891 using the InceptionV3 model on a dataset with 3294 images from the multiple disease dataset (MUD). In a recent investigation by Patil et al. (2023) [22], an ensemble model that combines Xception, InceptionV3, and Densenet201 models in parallel was proposed for the detection of multiple diseases, including cataract, pathological myopia, glaucoma, and diabetic retinopathy. By combining the datasets from the ODIR and the Joint Shantou International Eye Centre (JSIEC) and applying transfer learning, the proposed model achieved an accuracy of 0.88. While the studies mentioned above highlight the potential of artificial intelligence-based approaches for the automated detection of eye diseases and assist ophthalmologists with accurate diagnosis and treatment decisions; however, these studies mostly used fundus images rather than anterior images of the eye in diagnosing eye diseases. Relying on fundus images only, which requires specialized equipment such as a fundus camera to capture high-resolution images of the posterior segment, makes it impractical to replace traditional methods completely and presents challenges in real-time assessment due to limited access to fundus images from digital cameras or Mobile cameras. While it is technically possible to obtain fundus images from digital cameras or mobile cameras using specific accessories or adapters, this usually involves an additional cost, and the resulting images may show lower quality and resolution, leading to inaccuracies in the assessment and diagnosis of eye diseases. One notable study conducted by Siddique et al. (2022) [23] focused on diagnosing eye diseases, such as cataracts, chalazion, and strabismus, using anterior images of the eye instead of fundus images. They used six pre-trained models (VGG16, VGG19, MobileNet, Xception, InceptionV3, and DenseNet121) to recognize and classify these diseases. Among the models, MobileNet achieved the highest accuracy at 0.975. This highlights the efficacy of using

FIGURE 1 Visual examples of eye disease cases included in the dataset.

anterior eye imaging and deep learning models to accurately diagnose and classify eye diseases, offering potential advances in ophthalmic diagnosis. However, this study did not extensively explore the algorithm's applicability to other eye diseases and imaging modalities. In addition, it is important to note that the above studies did not incorporate hardware for real-time assessment, which could be a potential limitation in implementing the algorithm in real-time clinical settings. Therefore, the work of this study based on the pre-trained ImageNet models package using NASNetMobile, VGG19, ResNet50, InceptionV3, MobileNet, and InceptionResNetV2 pre-trained models is proposed for the detection of various eye diseases like the cataract, presence of the foreign body, the appeared of glaucoma, the haemorrhage subconjunctival, and lastly viral conjunctivitis from the dataset of the anterior images, when these limitations are main considerations. The performance of these models in detecting these diseases is evaluated, and a comparison is made to recognize the model with high efficiency to be incorporated within the platform of Raspberry Pi intended for real-time appraisals.

This paper is organized as follows: Section 2 provides an overview of the materials and methods used in the study, incorporating the research ethics, collection of data, the experimental environment, the hardware and software configuration recalled by the pre-processing, the performance of ImageNet models package and finally the process of evaluation and comparison of metric's performance. Section 3 presents the study results and discussion. Finally, section 4 involves the conclusion.

2 | MATERIALS AND METHODS

2.1 | Research ethics and data collection

The study was conducted according to the standards of ethical conduct in the Declaration of Helsinki, which involve humans as participants. The approved ethics were obtained from the Ethics Committee at the Training and Human Development Centre, Iraqi Ministry of Health (protocol number: 2003016-1981). Informed consent was obtained from all participants before their images were used for research purposes. The dataset used in this research was securely stored to protect the privacy and confidentiality of the participants.

The study included a dataset of around 645 subjects' eye images, where only 445 images of them referred to subjects having eye diseases and actually 200 eye images related to subjects having healthy eyes. The dataset was collected from the Balad Ruz General Hospital and Ibn Al-Haitham Teaching Eye Hospital, involving people with different eyes condition visiting the hospitals for regular check-ups starting from the second of January till the seventh of July 2023. The collected eye's images were classified under the supervision of experienced ophthalmologists by six categories: the normal or healthy group (200 images), the cataract group (100 images), the group of foreign body (91 images), the group of glaucoma (79 images), the subconjunctival haemorrhage group (84 images), and the group of viral conjunctivitis (91 images). The saved images were formatted in the JPEG file format. Some examples of each eye disease case used in this study are shown in Figure 1.

2.2 | Experimental setup

The procedure allowed the incorporation of the named models previously trained with the proposed microcomputer (Raspberry Pi), which allowed the implication of the real-time captured images. The experimental setup involved placing the subject within the front view from the camera by 0.5 m, as shown in Figure 2. This programming language, represented by Python version 3.9 and the platform of Anaconda version 2.3.2, was used in the implementation of the diagnosis process.

2.3 | Hardware

The hardware part of the experiment composed of the following components, (1) the Raspberry Pi 4 platform of Model B, (2) the efficient touch screen, and (3) the high-performance digital camera as shown in Figure 2. The Raspberry Pi 4 Model B, which is a popular microcomputer single-board type, worked

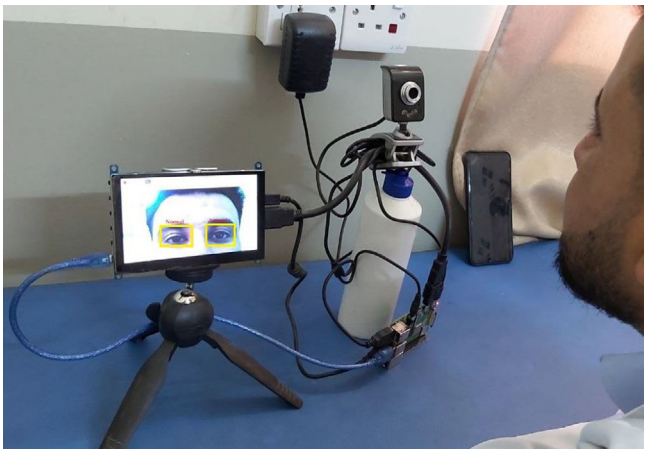

FIGURE 2 Example of patient eye disorder case from images dataset.

as the core for the proposed hardware system used to provide computational power, the required connectivity for real-time imaging processes, and finally, the interface of deep learning. In addition, this type of Raspberry Pi 4 Model B was chosen for the processing capabilities of enhancement, the ability to improve performance and the highest memory capacity [24] compared to other Raspberry Pi models. The touch screen has been integrated into the proposed system of specification represented by WaveShare HDMI Screen, and 5-inch with 800 × 480 pixel resolution. This camera is used for the subject interface and the monitoring of eye's condition. Both the Kisonli camera (Model No.: U-227, digital zoom (f = 3.85 mm)) were employed to capture images of the eye in a real-time environment and compare them with the pre-trained model in order to predict the case of the eye.

2.4 | Data pre-processing

The images taken were pre-processed as required, such as resizing, normalization, and the augmentation process, before being used to prepare and appraise the pre-trained models. The process of data augmentation was implemented through de-scaling to normalize pixel values by performing a rotation with a 5° angle to produce various angles of view, then slightly served to change the image width and height by 5% to improve the performance of the model. For lighting variations, the brightness was attuned within the range of [0.9, 1.1]. Also, a small shear and zoom (0.05) and zoom represented by 0.05 and 0.05, respectively, were utilized to introduce both the smooth skews and varies in the scale of the image correspondingly. The previously described techniques served to increase the dataset's size and improve the robustness of the pre-trained deep learning model to be robust in handling distinctions frequently located in the analysis of the image. To ensure the best and most accurate performance, the images of the eye dataset were split by 70% of them for training purposes and 30% of them for testing purposes, and using epochs were selected within 100 to secure the best performance.

The Haar cascade classifier algorithm [25, 26] was used for eye detection, in order for extracting and isolating both eyes for the process of analysis. The Haar cascade classifier is a machine learning-based approach employing Haar-like features and the AdaBoost algorithm to detect objects in images, particularly effective in identifying well-defined features, such as faces, eyes, or other objects with distinct shapes, stored in XML files. By comparing the obtained subject eyes to the deep learning pre-trained model, it allows for accurately classified and predicted eye defects or healthy eye states. The block diagram of the proposed system is shown in Figure 3.

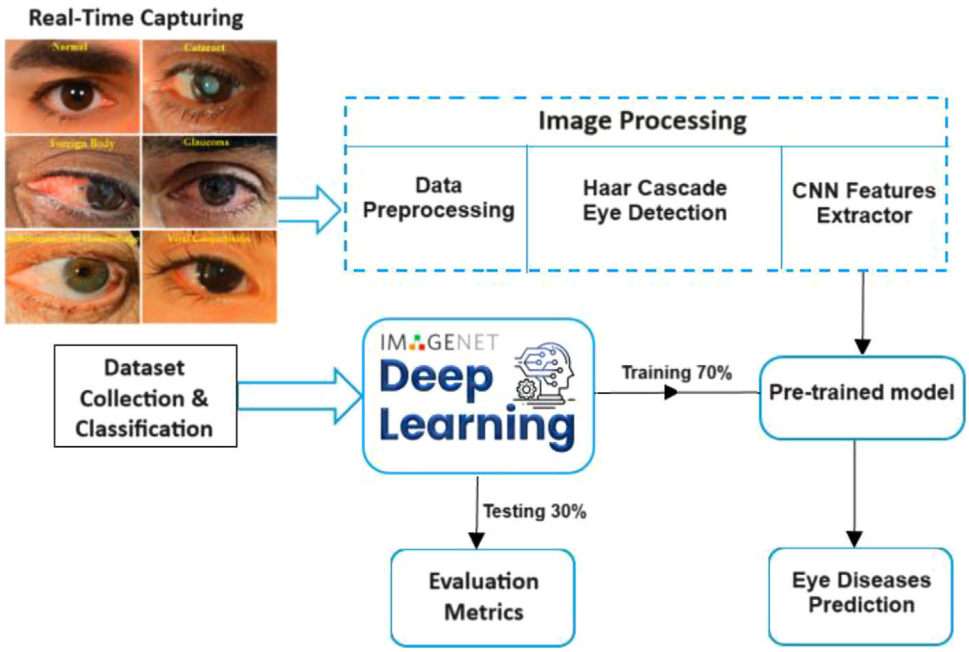

FIGURE 3 The block diagram of the proposed system.

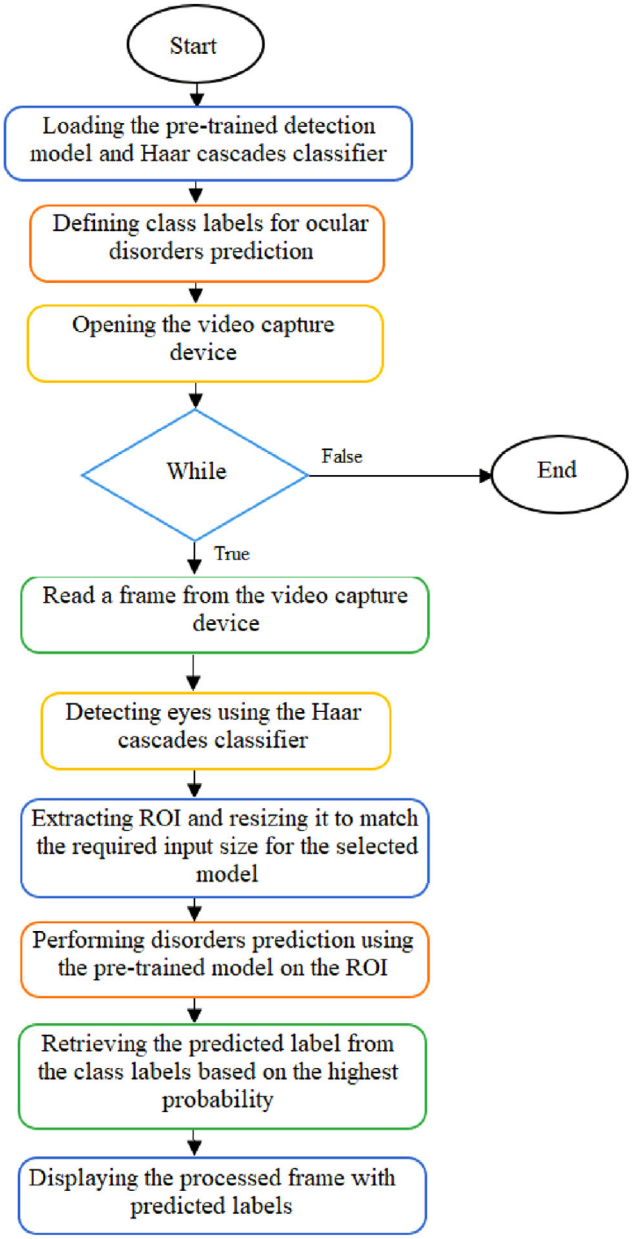

FIGURE 4 The real-time eye diseases detection system flowchart using the pre-trained deep learning model.

The flowchart of the proposed system is shown in Figure 4, which illustrates the series of strategies used with Python code. Starting by loading the pre-trained model and then Haar cascade classifier used for the process of eye detection. The category labels were outlined to stand for the various eye diseases. The capturing of video apparatus of video capturing was opened to get frames of image for the processing. Where the frame was taken continuously from the video capture device. Then for the purpose of eye detection, it was converted firstly to an RGB colour space and then to grayscale frame. Later, the Haar classifier used the cascades with the greyscale frame for the detection of the eyes. During the process of eye's detection, a rectangle was sketched about it on the displayed frame. After that, the region of interest (ROI), including the eyes, was extorted from the frames of the greyscale and RGB. By analysing, the region of interest (ROI) was resized to the required input size of the image aimed at the model's detection. The region of interest (ROI) was resized and converted to a matrix format to be used for prediction.

The prediction model was fed by the ROI based on the pre-trained detection model. The predicted label, which represents the state of the subject's detected eye, was recalled from the category labels based on the highest probability. Then, the projected label was displayed on the frame. The frame was displayed with drawn rectangles and projected labels, allowing observing the detection process in real-time. Once the round ends, the capturing process by video device is released, followed by closing of all the windows.

2.5 | ImageNet pre-trained deep learning models

The pre-trained ImageNet models package was used in this study to detect eye diseases. These pre-trained models have been trained to a large extent on the large-scale ImageNet dataset, which contains millions of annotated images of various object categories. The following models were used:

- **NASNetMobile**: The NASNetMobile architecture, introduced by Zoph et al. [27], is a developed CNN model that has gained recognition for its exceptional performance. The NASNetMobile architecture consists of a total of 53 convolutional layers. These layers are structured by numerous blocks and repeated structures of cells, creating a deep network for effective feature extraction and representation learning. The parameters of trainable were around 5.3 million. The size of the image that is referred to as an input to the NASNetMobile architecture is $224 \times 224 \times 3$. The specific configuration and architectural details of the NASNetMobile model can be found in [27].
- **VGG19**: Firstly, Simonyan and Zisserman introduced this VGG19 architecture as reported by [28], work as the CNN model that developed to be recognized for its unique performance. It consists of 19 layers, including 16 convolutional layers and three fully connected layers. The structural design of VGG19 is identified for firstly its deep structure and secondly its homogeneous design, therefore the multiple 3×3 convolutional layers have been stacked to shape their deeper demonstrations. This model has about 144 million trainable parameters. Usually, the size of the input image that is used for the model of VGG19 is $224 \times 224 \times 3$. The configuration details of the model called VGG19 described in details in [28].
- **ResNet50**: The architecture of ResNet50 model, introduced by He et al. [29], which was a CNN developed model that related to the Residual Network family (ResNet) . This ResNet50 model has had an outstanding performance in different tasks of computer vision and it has been considered as a popular option for the process of both the image classification and the image feature extraction. This deep model

totally comprises of 117 layers, compromising the convolutional layers, pooling layers, fully connected layers, and lastly the shortcut connections layers. The trainable parameters were around 25.6 million approximately. Usually, the size of the input image that used for the model of ResNet50 is 224 × 224 × 3. The configuration details of the model called ResNet50 described in details in [29].

- **InceptionV3**: The architecture of InceptionV3 model, introduced by Szegedy et al. [30], is a CNN model that developed and modified to work effectively and professionally with a complex image tasks. This deep model is composed of 316 layers, represented by the convolutional layers, the pooling layers, and the startup modules, acknowledging the representation of the robust feature. The trainable parameters of this model was around 23.9 million, and the size of its input image where represented by 299 × 299 × 3 The configuration details of the model called InceptionV3 described in details in [30].
- **MobileNet**: The architecture of MobileNet model introduced by Sandler et al. [31], is a CNN model that developed and modified to work effectively and professionally with a complex image tasks. This deep model is composed of 155 layers, represented by the depth-wise separable convolutional, the bottleneck, and the linear layers. The trainable parameters of this model were around 3.5 million, forming it as lightweight and efficient model. The size of its input image were represented by 224 × 224 × 3. The configuration details of the model called InceptionV3 described in details in [31].
- **InceptionResNetV2**: The mode design of the Inception-ResNetV2 presented by Szegedy et al. [32], characterized as a highly processed CNN model that joins the model's strength of both the Inception and the ResNet architectures. The deep learning model consists of 825 layers, presenting a design of complex network that incorporates the convolutional layers, the pooling layers, the inception modules, and the residual connections. This model has trainable parameters around 55.9 million with the size of an input image appearing as 299 × 299 × 3. The configuration details of the model called InceptionV3 are described in detail in [32].

2.6 | Evaluation metrics

Generally, the performance of the previously outlined deep learning pre-trained models was evaluated by various evaluation metrics, incorporating the confusion matrix, the accuracy, the precision, the recall, the F1 score, and the MCC. Initially, the confusion matrix used to provide a tabular demonstration of the deep learning model's performance, presenting the number of the true positive (TP), the true negative (TN), the false positive (FP), and the false negative (FN) predictions. Additionally, the accuracy factor which defined as the ratio of the correct predictions to the total number of performed predictions made. Also, the precision factor is identified as the positive predictive value, which measures the proportion of TP predictions of the model amongst all the positive predictions. In addition, the recall factor used which called as the sensitivity or the TP rate. This factor represents by the proportion of the TP predictions amongst all the actual positive cases. Furthermore, the F1 score is a harmonic way of accuracy and recovery, by capturing the balance between both accuracy and recovery. This factor served as a single metric to evaluate the model performance instantaneously in both accuracy and recall. Finally, the factor called by MCC is represented by a quality measure for the binary classification models that counts the predictions of the TP, TN, FP, and FN. By evaluating the overall agreement between both the predicted label and true label, which ranging from −1 to +1, since +1 shows a perfect prediction, while 0 indicates a random prediction, and −1 represents the complete disagreement. All the outlined parameters have been expressed as follows [33–35]:

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

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

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

$$F1\ Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \tag{4}$$

$$MCC = \frac{(TP \times TN - FP \times FN)}{\sqrt{((TP + FN)(TN + FP)(TP + FP)(TN + FN))}} \tag{5}$$

3 | EXPERIMENTAL RESULTS AND DISCUSSION

This section presents the performance of the previously illustrated pre-trained package of models (the NASNetMobile model, the model of VGG19, InceptionV3, the model of ResNet50, the MobileNet model, and the InceptionRes-NetV2 model) used for the purpose of identifying the state or situation of the left and right eye from the subject's eye image dataset collected from the hospital. The performance of these model metrics, represented by accuracy, precision, recall, F1 score and MCC, was calculated and summarized in Table 1 and Figure 5.

TABLE 1 Performance metrics for detecting eye diseases using different pre-trained models.

Model	Accuracy	Precision	Recall	F1 score	Matthews correlation coefficient
NASNetMobile	0.72	0.68	0.68	0.67	0.65
VGG19	0.78	0.75	0.74	0.74	0.72
InceptionV3	0.85	0.83	0.83	0.83	0.82
ResNet50	0.86	0.86	0.83	0.84	0.83
MobileNet	0.90	0.91	0.88	0.89	0.88
InceptionResNetV2	0.93	0.94	0.92	0.93	0.92

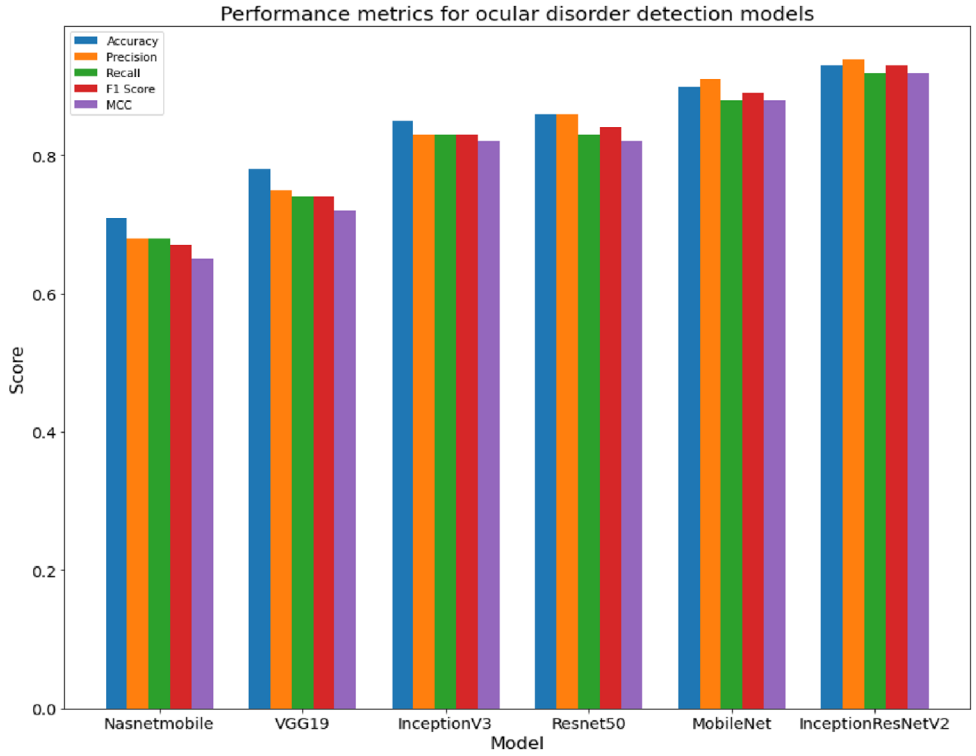

FIGURE 5 Performance metrics for eye disorder detection at different pre-trained models.

From Table 1 and Figure 5, it is clearly illustrated that the InceptionResNetV2 model showcased an outstanding performance compared to other models by consistently achieving the highest scores across all metrics. It achieved an accuracy score of 0.93, precision of 0.94, recall of 0.92 and F1 score of 0.93. According to these results, the InceptionResNetV2 model excels in accurately detecting eye diseases. Following closely behind, the MobileNet and ResNet50 models also displayed quality performances. They have proved their effectiveness in eye disease detection based on their exhibited high scores on all metrics. Furthermore, the VGG19 and InceptionV3 models showed moderately high scores, but on the other hand, the NASNetMobile model performed inadequately across all metrics.

In addition, it is observed from Table 1 and Figure 4 that the InceptionResNetV2 model achieved the most excellent normalized MCC score of 0.92, an additional teller of its brilliant eye disease detection. Its performance was almost matched by yet another strong performance from the MobileNet model, which has an MCC score of 0.88. Other models with relatively high MCC scores are ResNet50 and InceptionV3 with 0.83 and 0.82, respectively, stipulating their efficient eye disease sensing. The model with a slightly lower than higher performance with a 0.72 score, is the VGG19. The NASNetMobile model attained the lowest MCC score of 0.65, suggesting that the nature of its execution is relatively lower in its eye disease observations.

The confusion matrix's purpose is to visually depict the true positives, true negatives, false positives, and false negatives distribution for detecting eye diseases. The confusion matrix for the NASNetMobile and VGG19 models is illustrated in Figure 6.

The confusion matrix for the InceptionV3 and ResNet50 models is shown in Figure 7.

The confusion matrix for the MobileNet and InceptionResNetV2 models is shown in Figure 8.

Finally, the performance of the best model represented by InceptionResNetV2 through training and validation shows accuracy and loss trends across 100 epochs for eye diseases detection, as shown in Figure 9. From the last mentioned figure, it is evident that the accuracy of training steadily increases over the epochs, reaching a maximum accuracy of approximately 96%, compared to the accuracy of validation, which also increases, getting a maximum value of about 85%. Furthermore, the training and validation loss curves appeared to have a similar trend. The training loss gradually decreases as the model learns, while the validation loss remains consistently lower, indicating good generalization. This deep learning model, characterized by InceptionResNetV2 was learning efficiently to classify the training data of eye disease and circulates appropriately well for both the new and unseen data throughout the validation process.

The experimental results from the other deep learning models represented by the NASNetMobile, VGG19, InceptionV3, ResNet50, MobileNet, and InceptionResNetV2 models show a promising performance in eye disease detection. However, the proposed system of the outlined previous models has some limitations. One limitation is the potential impact of subject movement during imaging, image background, and low-contrast

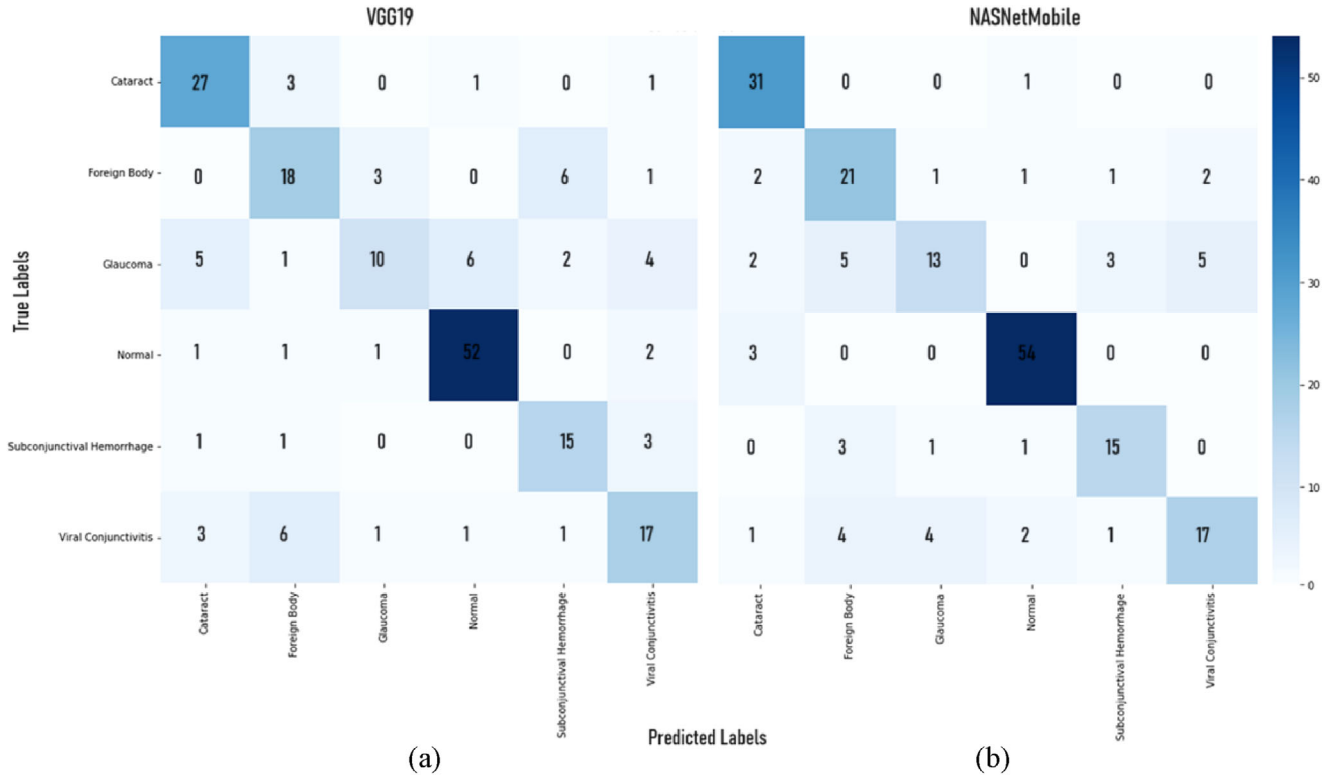

FIGURE 6 The confusion matrix illustrating the classification results for detecting eye disorder using (a) the NASNetMobile model and (b) the VGG19 model.

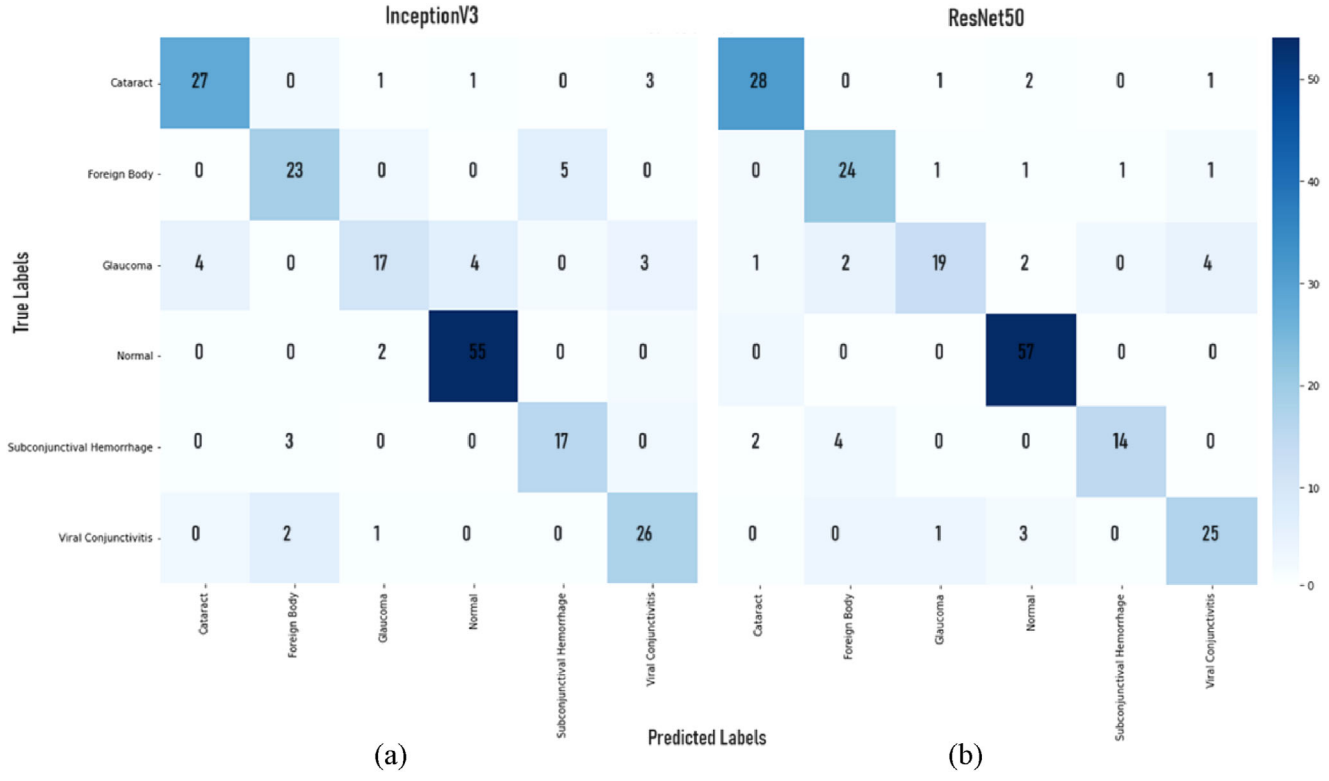

FIGURE 7 The confusion matrix illustrating the classification results for detecting eye disorder using (a) the InceptionV3 model and (b) the ResNet50 model.

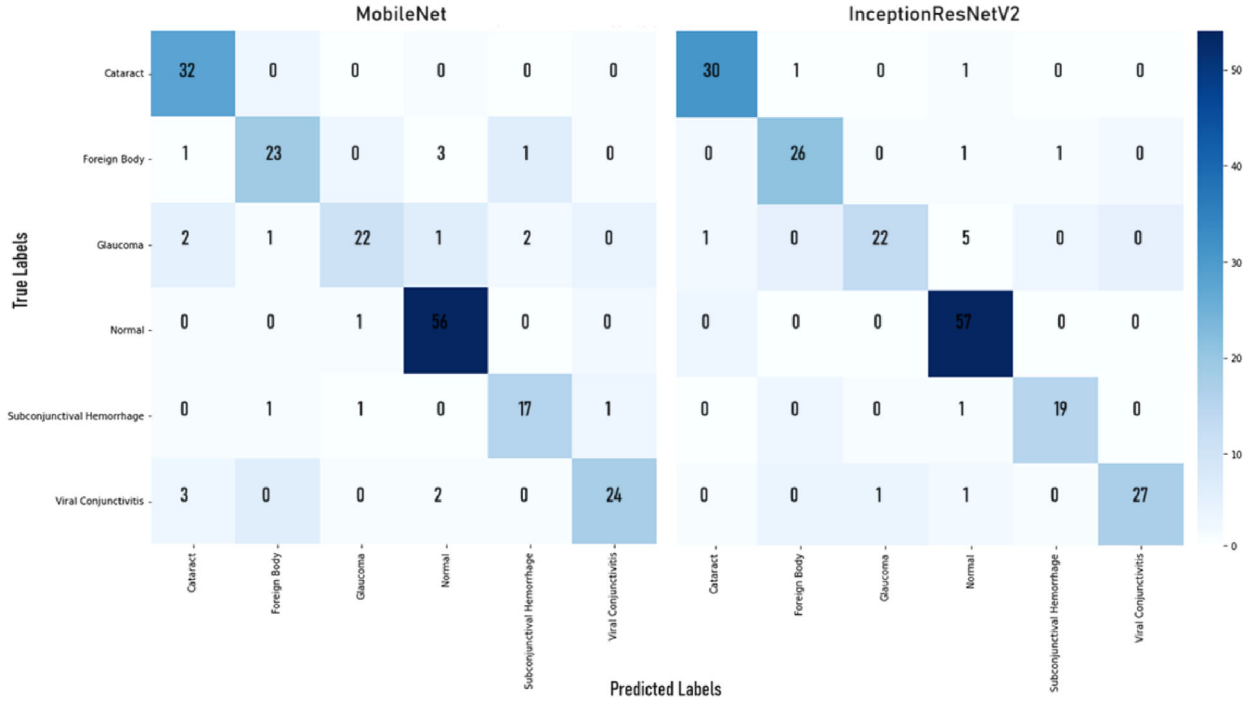

FIGURE 8 The confusion matrix illustrating the classification results for detecting eye disorder using (a) the MobileNet model and (b) the InceptionResNetV2 model.

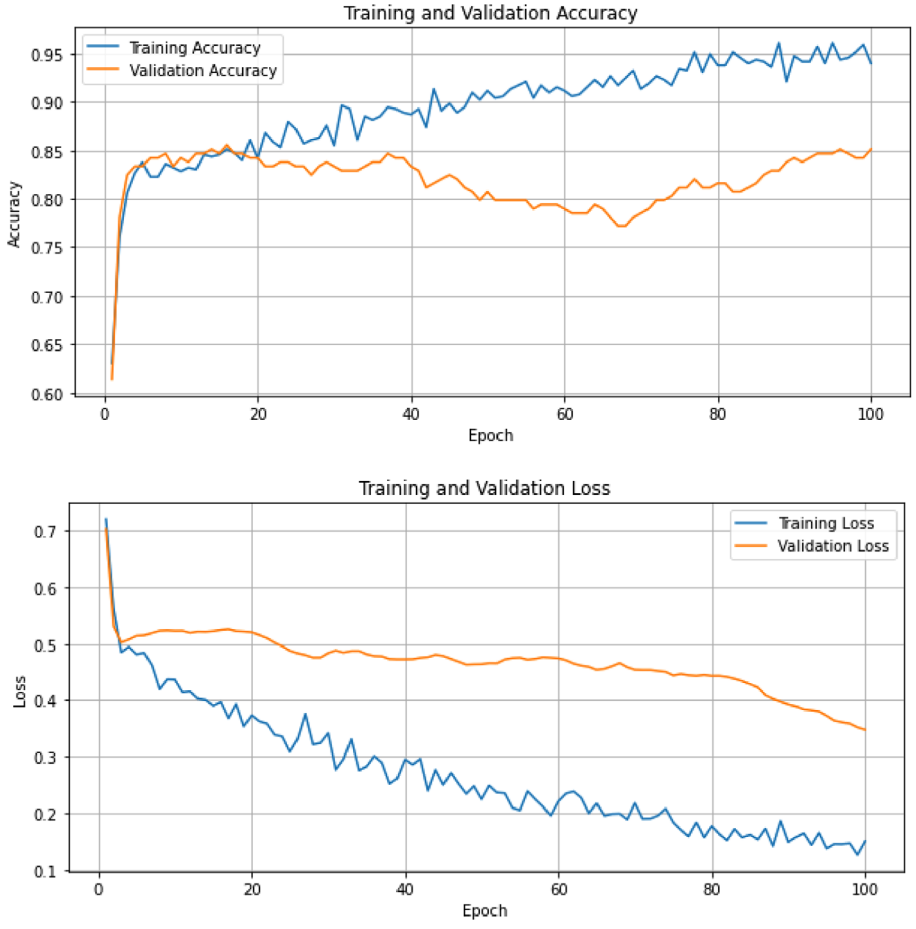

FIGURE 9 Training and validation performance of the InceptionResNetV2.

images, resulting in the introduction of artifacts and its effect on the accuracy of predictions. In addition, the reflections of eye-lighting that appeared in the captured images might pose challenges for accurate classification. Also, the limited cases in the data set of some eye models' disease categories could affect these models' generalizability. The proposed system was also restrained by its inability to assess the severity or degree of eye disease detected accurately. Thus, it cannot specifically inform the operator about the case severity. This limitation will cause difficulty in the clinical staff's composing appropriate decision-making due to insufficient deep information. Moreover, using a single-board computer like Raspberry Pi 4 could potentially have limitations in terms of scalability, expandability, and connectivity compared to traditional desktop setups.

Future studies could address all of these limitations by tackling motion artefacts, improving the techniques of image capturing, and increasing the data set to provide a more diverse representation to quantify the severity of eye diseases. Furthermore, image background and low-contrast images can be addressed based on an enhancement approach proposed in [36]. Also, the deep learning models can be broadened in their classification to other eye diseases, comprising multiple diseases within a single patient. The improvement of the classification accuracy of diseases needs extensive training and the inclusion of images from other classes. Normalizing the deep learning model by incorporating a diverse data set will improve its robustness and effectiveness in accurately classifying eye diseases.

4 | CONCLUSION

This study presented an automated system that showcased a computer vision approach. This system is used for detecting various eye diseases (cataracts, foreign bodies, glaucoma, subconjunctival for haemorrhage and viral conjunctivitis), alongside the detection of healthy eyes. This was achieved by integrating pre-trained deep learning models from ImageNet, including the NASNetMobile, VGG19, ResNet50, InceptionV3, MobileNet, and InceptionResNetV2) on the platform of Raspberry Pi. The performance of these stated deep learning algorithms and advanced embedded systems was illustrated in their experimental outcomes. The InceptionResNetV2 and MobileNet achieved the highest accuracy, represented by 93% and 90%, respectively, while the NASNetMobile model showed the lowest precision, reaching 72%. This automated diagnostic system will be capable of improving the eye healthcare system's performance.

AUTHOR CONTRIBUTIONS

Ali Al-Naji: Conceptualization; formal analysis; investigation; methodology; project administration; software; validation; visualization; writing—original draft. **Ghaidaa A. Khalid**: Data curation; formal analysis; investigation; methodology; resources; validation; writing—review and editing. **Mustafa F. Mahmood**: Data curation; formal analysis; methodology; validation; visualization; writing—review and editing. **Javaan Chahl**: Funding acquisition; methodology; visualization; writing—review and editing.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data is available on request from the corresponding author.

ORCID

Ali Al-Naji https://orcid.org/0000-0002-8840-9235
Ghaidaa A. Khalid https://orcid.org/0000-0001-6270-6445

REFERENCES

1. WHO.: Blindness and vision impairment. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment. Accessed 1st July 2023.
2. Delbarre, M., Froussart-Maille, F.: Signs, symptoms, and clinical forms of cataract in adults. J. Fr. Ophtalmol. 43, 653–659 (2020)
3. Pandey, A.N.: Ocular foreign bodies: A review. J. Clin. Exp. Ophthalmol. 08, 1â (2017)
4. Schuster, A.K., Erb, C., Hoffmann, E.M., Dietlein, T., Pfeiffer, N.: The diagnosis and treatment of glaucoma. Dtsch. Arztebl. Int. 117, 225 (2020)
5. Doshi, R., Noohani, T.: Subconjunctival hemorrhage. StatPearls Publishing LL, Treasure Island, FL (2019)
6. Solano, D., Fu, L., Czyz, C.N.: Viral conjunctivitis. StatPearls Publishing LL, Treasure Island, FL (2017)
7. Prananda, A.R., Frannita, E.L., Hutami, A.H.T., Maarif, M.R., Fitriyani, N.L., Syafrudin, M.: Retinal nerve fiber layer analysis using deep learning to improve glaucoma detection in eye disease assessment. Appl. Sci. 13, 37 (2022)
8. Cheung, C.Y., Tang, F., Ting, D.S.W., Tan, G.S.W., Wong, T.Y.: Artificial intelligence in diabetic eye disease screening. Asia-Pacific J. Ophthalmol. 8, 158–164 (2019)
9. Paradisa, R.H., Bustamam, A., Mangunwardoyo, W., Victor, A.A., Yudantha, A.R., Anki, P.: Deep feature vectors concatenation for eye disease detection using fundus image. Electronics 11, 23 (2021)
10. Abbas, Q.: Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. Int. J. Adv. Comput. Sci. Appl. 8(6), 080606 (2017)
11. Subramanian, M., Kumar, M.S., Sathishkumar, V., Prabhu, J., Karthick, A., Ganesh, S.S., Meem, M.A.: Diagnosis of retinal diseases based on Bayesian optimization deep learning network using optical coherence tomography images. Comput. Intell. Neurosci. 2022, 1–15 (2022)
12. Elangovan, P., Vijayalakshmi, D., Nath, M.K.: Covid-19net: An effective and robust approach for covid-19 detection using ensemble of convnet-24 and customized pre-trained models. Circuits Syst. Signal Process. 43, 2385–2408 (2023)
13. Vijayalakshmi, D., Elangovan, P., Nath, M.K.: Methodology for improving deep learning-based classification for CT scan COVID-19 images. In: Biomedical Engineering: Applications, Basis and Communications 2450008. World Scientific, Singapore (2024)
14. Bhowmik, A., Kumar, S., Bhat, N.: Eye disease prediction from optical coherence tomography images with transfer learning. In: Proceedings of the 20th International Conference on Engineering Applications of Neural Networks, EANN 2019, pp. 104–114. Springer, Cham (2019)
15. Diaz-Pinto, A., Morales, S., Naranjo, V., Köhler, T., Mossi, J.M., Navea, A.: CNNs for automatic glaucoma assessment using fundus images: an extensive validation. Biomed. Eng. Online 18, 1–19 (2019)
16. Smaida, M., Yaroshchak, S.: Bagging of convolutional neural networks for diagnostic of eye diseases. Proc. COLINS, 2604, 715–729 (2020)
17. Raza, A., Khan, M.U., Saeed, Z., Samer, S., Mobeen, A., Samer, A.: Classification of eye diseases and detection of cataract using digital fundus imaging

(DFI) and inception-V4 deep learning model. In: Proceedings of the 2021 International Conference on Frontiers of Information Technology (FIT), pp. 137–142. IEEE, Piscataway, NJ (2021)
18. Gour, N., Khanna, P.: Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network. Biomed. Signal Process. Control 66, 102329 (2021)
19. Camara, J., Rezende, R., Pires, I.M., Cunha, A.: Retinal glaucoma public datasets: What do we have and what is missing? J. Clin. Med. 11, 3850 (2022)
20. Zhou, H.: Design of intelligent diagnosis and treatment system for ophthalmic diseases based on deep neural network model. Contrast Media Mol. Imaging 2022, 1–10 (2022)
21. Rajee, R.J., Roomi, S.M.M., PooAnnamalai, V., Begam, M.P.: A transfer learning approach for retinal disease classification. In: Proceedings of the 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), pp. 1–6. IEEE, Piscataway, NJ (2023)
22. Patil, Y., Shetty, A., Kale, Y., Patil, R., Sharma, S.: Multiple ocular disease detection using novel ensemble models. Multimedia Tools Appl. 83, 11957–11975 (2023)
23. Siddique, M.A.A., Ferdouse, J., Habib, M.T., Mia, M.J., Uddin, M.S.: Convolutional neural network modeling for eye disease recognition. Int. J. Online Biomed. Eng. 18(9), 115–130 (2022)
24. Amsalam, A.S., Al-Naji, A., Daeef, A.Y., Chahl, J.: Automatic facial palsy, age and gender detection using a raspberry Pi. BioMedInformatics 3, 455–466 (2023)
25. Kasinski, A., Schmidt, A.: The architecture and performance of the face and eyes detection system based on the Haar cascade classifiers. Pattern Anal. Appl. 13, 197–211 (2010)
26. Cuimei, L., Zhiliang, Q., Nan, J., Jianhua, W.: Human face detection algorithm via Haar cascade classifier combined with three additional classifiers. In: Proceedings of the 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), pp. 483–487 (2017)
27. Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V.: Learning transferable architectures for scalable image recognition. In: Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8697–8710. IEEE, Piscataway, NJ (2018)
28. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015) (2015)
29. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. IEEE, Piscataway, NJ (2016)
30. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826. IEEE, Piscataway, NJ (2016)
31. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520. IEEE, Piscataway, NJ (2018)
32. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proc. AAAI Conf. Artif. Intell. 31(1), 11231 (2017)
33. Abdulrazzak, A.Y., Mohammed, S.L., Al-Naji, A.: NJN: A dataset for the normal and jaundiced newborns. BioMedInformatics 3, 543–552 (2023)
34. Zhu, Q.: On the performance of Matthews correlation coefficient (MCC) for imbalanced dataset. Pattern Recognit. Lett. 136, 71–80 (2020)
35. Chicco, D., Jurman, G.: The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21, 1–13 (2020)
36. Vijayalakshmi, D., Nath, M.K.: A systematic approach for enhancement of homogeneous background images using structural information. Graph. Models 130, 101206 (2023)

How to cite this article: Al-Naji, A., Khalid, G.A., Mahmood, M.F., Chahl, J.: Computer vision for eye diseases detection using pre-trained deep learning techniques and raspberry Pi. J. Eng. 2024, e12410 (2024). https://doi.org/10.1049/tje2.12410