An Efficient Deep Learning-Based Classification Framework for Hypertensive Retinopathy

By

Muhammad Zaheer Sajid

Supervisor: Dr. Nauman Ali Khan

A thesis submitted to the faculty of Computer Software Engineering Department, Military College of Signals, National University of Sciences and Technology, Islamabad, Pakistan in partial fulfilment of the requirements for the degree of MS in Computer Science (Software) Engineering

April 2023

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by Mr. **Muhammad Zaheer Sajid,** Registration No. **0000329298**, of **Military College of Signals** has been vetted by undersigned, found complete in all respect as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial, fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the student have been also incorporated in the said thesis.

Signature: _____

Supervisor: Asst Prof Dr. Nauman Ali Khan, PhD

Date: _____

Signature (HoD): _____

Date: _____

Signature (Dean/Principal): _____

Date: _____

Declaration

I hereby declare that no portion of work presented in this thesis has been submitted in support of another award or qualification either at this institution or elsewhere.

Dedication

“In the name of Allah, the most Beneficent, the most Merciful"

I dedicate this thesis to my family, friends, and teachers who supported me each step of the way.

Abstract

Hypertensive retinopathy (HR) is a well-known eye disease that is caused by high blood pressure (hypertension). In this illness, symptoms typically develop later. The AV nicking, cotton wool patches, constricted veins in the optic nerve, and blood pouring into the eye's optic nerve all contribute to the appearance of the HR symptoms. HR disease may have different types of serious complications, including retinal artery blockage, destruction of the visual nerves, and maybe vision loss. The automated early detection of this illness can be aided by AI and deep learning models. In this research, a novel dataset for HR is collected from Pakistani hospitals (Pak-HR) and internet sources. Second, a brand-new methodology (Incept-HR) is developed to evaluate hypertensive retinopathy using InceptionV3 and residual blocks. 6,000 digital fundus images from the collected datasets were used to train the Incept-HR system. The proposed classification method, Incept-HR, has 99% classification accuracy and an f1-score of 0.99. The results show that this model produces useful outcomes and can be applied as a diagnostic testing tool. The system is not intended to replace optometrists; rather, it aims to assist professionals. The proposed methodology outperforms both the cutting-edge models VGG19 and VGG16 in terms of classification accuracy.

Acknowledgements

All praises to Allah for the strengths and His blessing in completing this thesis.

I would like to convey my gratitude to my supervisor, Asst Prof Dr. Nauman Ali Khan, PhD, for his supervision and constant support. His invaluable help of constructive comments and suggestions throughout the experimental and thesis works are major contributions to the success of this research. Also, I would thank my committee members; Assoc Prof Dr. Ehtisham ul Islam, and Asst Prof Dr. Faheem Arif for their support and knowledge regarding this topic.

Last, but not the least, I am highly thankful to my family, friends and teachers. They have always stood by my dreams and aspirations and have been a great source of inspiration for me. I would like to thank them for all their care, love and support through my times of stress and excitement.

Table of Contents

List Figures

List of Tables

Chapter 1

Introduction

The World Health Organization (WHO) reported that the number of people who have hypertension increased from 594 million in 1975 to 1.13 billion in 2015, with the majority of the increases taking place in low- and middle-income countries. The main cause of this development is the prevalence of diseases that already exist and increase the likelihood that hypertension will develop. 1.56 billion people are expected to have hypertension by the year 2025. Additionally, over 66% of people with hypertension reside in developing or underdeveloped countries, where the lack of resources for proper health care to diagnose, track, and treat hypertension exacerbates the issue[1]. The most widespread type of eye disease (HR) that has spread over the entire world is caused by hypertension, which is brought on by rising vascular resistance. [2]. Hypertension damages several human tissues, including vascular disease, the eyes, and the heart. [3]. In addition to all of these harms, hypertensive retinopathy (HR) is the most notable cause of cardiovascular disease and the one that ultimately results in mortality [4].Hypertensive retinopathy (HR) is an abnormality of the retina caused by hypertension. Significant signs of HR-related eye illness include the development of arteriolar narrowing, arteriovenous nicking, retinal hemorrhage , microaneurysms, Cotton wool spots , papilledema and in extreme situations, optic disc, and macular edema [5]. For medical intervention and the protection of human life, early diagnosis of HR-related eye illnesses is essential [6]. Clinical experts (optometrists) are using retinal fundus image recognition technology to determine the presence of eye illness linked to HR in a non-invasive and cost-effective way. The major goal of employing an automated detection mechanism is to relieve the burden of an image assessment by optometrists by offering a crucial step for determining and treating the existence of retinopathy [7]. In earlier studies, the researchers developed a number of methods for processing retinal images, such as image enhancement through pre-processing, categorization of HR-related regions and arteries, feature extraction, and finally, a monitored machine-learning automated system to classify HR-related eye problems. [4]. Extremely large veins with a low percentage of the typical size of arteries to veins are a significant indicator of HR-related diseases. However, getting accurate measurements of vascular widths is quite difficult for optometrists. Micro fundus images captured with an optical camera have recently been shown in a number of studies to be

useful for observing retinal microvascular diseases. Due to its low cost, ease of use, and ability to clearly depict most clinical lesion forms in its fundus pictures, this fundus camera is frequently and safely used to check the eyes of a large number of HR patients [4]. Optometrists use automated evaluation of digital retinal images to spot alterations in the images and identify HR. The HR-related areas as well as other crucial eye areas are damaged by these alterations, as was previously described. If these alterations are not recognized at an early stage, hypertensive retinopathy develops. Optometrists are helped by computerized methods to identify various retinal diseases, such as eye illnesses associated with HR [8]. By enabling self-diagnosis, these technologies benefit academia and the worldwide medical community. Optometrists employ these tools to treat and diagnose eye-related disorders, particularly those that are HR-related. There are a few research papers that really on feature extraction to automatically detect hypertensive retinopathy (HR). These feature extraction techniques involves segmenting the structural elements of the retina, such as the macular, optic nerves, arteries, and vasculature, as shown in Fig 1.

Deep-learning techniques were also developed to extract certain structures. These characteristics were statistically evaluated in order to detect anomalies and, ultimately, to identify HR or non-HR illness. For automated systems to find these pathological conditions and use them to diagnose HR illness, they have to go through a long and difficult process. Researchers spent a lot of time just pulling out the characteristics, as opposed to other symptoms like cotton-wool patches or bleeding. They demonstrated that whereas other lesions were crucial for determining the presence of malignant HR-related illnesses, aneurysms defined the initial phases of HR. On the other hand, a lot of computer vision algorithms [9] and biological imagery behavior analytics [10] have heavily relied on deep learning techniques in recent years.

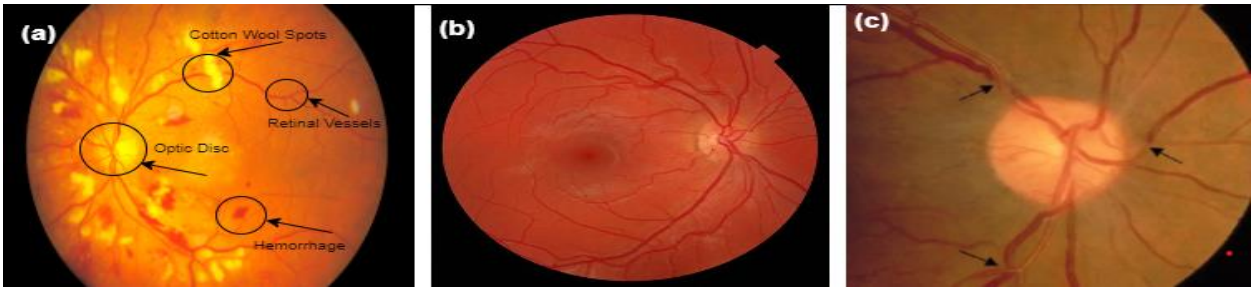

Fig. 1: Illustration of vascular system: (a) optic disk, cotton wool spots, hemorrhages (b) tortuosity (c) A/V ratio

1.1. Research Motivation

Although a few methods have been established to identify images of the optical fundus as HR or non-HR, there are still some major challenges.

- Even with the use of sophisticated pre- or post-image processing technologies, defining HR features from retinography pictures is challenging because to the difficulty in locating and retrieving HR-related lesion features.
- A few datasets with no expert medical annotations that detail a wide range of HR-related harm features are publicly available. This makes it difficult for computerized systems to identify the symptoms of some diseases.

Consequently, the primary objective of the research is to create a dataset for HR classification (PAK-HR). the second objective is to develop a multilayered deep learning model for completely autonomous processing of retina images, particularly in the situation of HR-related eye disease. In this research, a multilayered InceptionV3 system is used to construct an Inception-HR (Incept-HR) system that integrates residual blocks. A fully trained Incept-HR system can identify HR with accuracy on various HR-related retina images that include anatomic components. An experienced optometrist has identified certain HR-related diseases and anatomic features. a training technique developed by the CNN model to recognize abnormalities from fundus images. This study is very important in real world as it proposes a new system for hypertensive retinopathy classification.

1.2. Research Contribution

In this framework we propose a novel deep learning model to solve the problem of hypertensive retinopathy. we also propose a new dataset that was collected from PAKISTAN hospitals PAK-HR. In this study, the Inception-HR (Incept-HR) model is created to address the issues raised above by classifying data into HR and non-HR using deep learning architectures with multiple layers and image processing methods. Here we highlight the major contributions of the Incept-HR system.

- In this study the authors gathered a huge dataset from Pakistani hospitals name (Pak-HR) and internet sources. The 6,000 photos in this dataset allowed the trained model to achieve very high accuracy.
- In this study, residual blocks and trained InceptionV3 convolutional neural networks (CNN) are used to make a multi-layer architecture for the Incept-HR system.

- To identify HR-eye-related disease, four layers are added to the system architecture of the Incept-HR model. To build the trained features layer, the CNN model is first utilized to extract characteristics of HR-related lesions. The residual block technique is then used to extend those characteristics further.
- The building color space's deep properties and phases serve as the foundation for the algorithm developed in this work to classify HR. The author claims that this was the first attempt to create an automated system that was more effective than existing approaches for identifying HR diseases.
- Our systems achieved very high accuracy 99% that is higher than proposed methodologies in literature.

1.3. Thesis Outline

This thesis is divided into seven chapters:

- **Chapter 1**: This chapter includes the basic introduction, establish the research motivation and research contribution.
- **Chapter 2**: This chapter describes the literature survey of articles related to this research.
- **Chapter3**: This chapter describes the scientific background of the research.
- **Chapter 4**: This chapter describes the proposed architecture.
- **Chapter 5**: This chapter describes the experiments results.
- **Chapter 6**: This chapter compares our work with state-of-the-art research.
- **Chapter 7**: This chapter presents the results of the research.
- **Chapter 8**: This chapter concludes the report and highlights the direction for future work.

Chapter 2

Literature Review

In this work the authors propose deep learning model to investigate the classification problem of HR illness. Deep learning (DL) is a subset of Machine learning algorithms where human designed rules are not required. DL can be used in different applications. In [11] A survey on the application of deep learning to free-hand sketch problem is presented. The paper provides a review on different datasets and algorithms that can be applied for this problem. In [12] Enhancing sketch-based image retrieval by CNN sematic re-ranking is proposed. Numerous automated methods have already been proposed to detect retinal conditions. Despite this, there aren't many automated methods for detecting HR eye illness. To automatically recognize the HR state, several researchers have already used a fundus image processing technique. Optometrists may be able to save a lot of time and effort by adopting fundus image analysis to detect HR illnesses early, according to [11]. The most recent research in the area is provided in this section. We divided the articles we looked at into two categories: learning-based techniques with handcrafted features and learning based techniques with deep features.

2.1. Hand-crafted feature-based methodology

According to the composition research, retinal fundus images might be used to identify hypertensive retinopathy [11] by using a feature extraction method. The following manually created features are used in automated mechanisms to identify optic anomalies such as arterial branching and grade HR: the ocular disc (OD) position, the mean fractal dimension (mean-D), the arteriolar-venular equivalent diameter (AVR), the papilledema symptoms, and the compressibility index [12–18]. The Gabor 2D, Cake Wavelet and Canny edge detection algorithms were used for segmented and low-level applications. The following datasets were utilized to assess the efficacy of such systems: INSPIR-AVR, AVRDB, VICAVR, STARE, DRIVE, DR-HAGIS, and IOSTAR.

While in [18], they used a supervised classifier to conduct initial segment and refinement procedures in order to find hemorrhagic lesions. To identify eye disease associated with HR, a different method was used in [19]. The authors identified cotton wool spots (CWS) as one of the

critical clinical indicators for identifying HR-related optical disease. The candidate areas were improved using the Gabor filter bank, and the picture was then binarized using the thresholding approach. For such a localized dataset, the researchers have reported sensitivity and PPV of 82.21% and 82.38%, respectively. Including diabetic retinopathy (DR), hypertensive retinopathy (HR), macular degeneration (MD), vein branch occlusion (VBO), vitreous hemorrhage (VH), and normal retina, the full system of five categories of retina abnormalities was recognized in [20]. They used a wavelet-based neural network approach with pre-processed photos to find all of these retinal diseases. Through a features-based classification technique, they reported that the %50, %70, %83, %90, %93 and %95 for testing five retinopathy instances, respectively. The systems developed in [21–24]employed the AVR method on a small subset of color fundus images obtained from sizable datasets like DCCT and ETDRS. Either the OD region is segmented, and the blood is divided into veins and arteries, or the arterial diameters are calculated using gradients, morphological edge detection, and the Gabor wavelet. In order to implement semi-automatic recognition and measuring of the optic disc for the detection of eye-related HR disease, a graphical user interface (GUI) was created in [25]. This interactive GUI can be used to measure the diameters of any regions of interest (ROI) to help people with HR figure out their vascular risk. This method demonstrated improved classification outcomes when it was evaluated on two datasets, namely DRIVE and STARE.

The AVR ratios were used in combination with a clustering method in [26] to identify veins and arteries. As opposed to [27], which classified blood vessels into arteries and veins using intensity variation and color data after extracting the shade of grey and the time characteristics to identify pixels that belonged to blood vessels. On 101 photos gathered from the VICAVR dataset, different phases of HR-related eye illness are identified by estimating vessel width. In the VICAVR dataset, they used 25 for normal and 76 for hypertensive retinopathy. The transformations of Radon and Hough were used by the authors in [28] to segment blood vessels, after which the tortuosity index and vessel diameter were evaluated, and lastly the AVR was computed using fundus images, from which one may identify hypertensive retinopathy. Similar to this, the authors of [29]used harmonic sub-bands for individual component analysis to identify the many alterations that may be seen in the retina fundus, including blood vessels, the optic disc, exudates, hemorrhages, the macula, and drusen. On 50 photos, this approach was tried, and it correctly identified the alterations. To categorize the various retinal blood vessel types, the authors in [30] used a deep convolutional

neural network technique that was supplied using moment-based element characteristics and Gabor wavelets. The authors achieved noticeably better segmentation results on the DRIVE dataset. [15] describes the development of a nine-step, automatically generated system for extracting the OD region, segmenting the vessels, detecting color features, computing the AVR ratio, classifying the vessels as arteries or veins, computing the mean red intensity, computing the AM-FM features, and finally classifying the patients as having HR-related or non-HR-related eye disease. The ROC-AUC of 0.84 was attained on a collection of 74 color fundus pictures with sensitivity and specificity of 90% and 67%, respectively. Last but not least, Table 1 showed a comparison between standard machine learning approaches and state-of-the-art automated detection strategies for recognizing HR based on features taken from pictures of the retinal fundus. [31]proposed a method for extracting characteristics from color fundus pictures before HR recognition. The vessels could be clearly seen after first using CLAHE to convert the fundus picture to the green channel. Second, it used morphological closure to eliminate the optic disc. Third, the backdrop was removed by subtracting. Then, utilizing zoning, characteristics were retrieved. As a classifier, an ANN with feed-forward activation was utilized at the end. It had a 95% accuracy rate. The ELM classifier was used to demonstrate a method for segmenting features of the image from fundus pictures [32]. 39 morphological features in a feature vector and other characteristics were fed to the classifier. On the DRIVE dataset, this approach has a 96.07% accuracy rate, a 71.4% sensitivity rate, and a 98.68% specificity rate.

Table 1: Handcrafted feature-based HR classification

Methodologies	CWS	H	T	Results
AVR and Optic neuritis are used in an automated CAD approach to discriminate between blood vessels and arteries using fundus images[14]	No	No	No	Accuracy = 95%
Semi-automatic CAD method for examining fundus images to determine retinal vascular shape [15]	No	No	No	Accuracy = 96%
AVR and OD to identify HR with various degrees. [14]	No	No	No	Accuracy = 95%
A technique of density analysis was used to analyze A/V ratio to diagnose arteriolar constriction. [16]	No	No	No	Accuracy = 98%

Methods used for image pre-processing and classification evaluate the AVR and the OD.[17]	No	No	No	Accuracy = 98%

2.2. Models-based on deep learning

The second method described in the literature uses deep learning models to categorize fundus images into HR or non-HR eye disorders. Preprocessing time required by manually created feature-based techniques is decreased by such models. A deep learning-based model (DLM) was used to identify HR, and it was published in [6]. Several image processing tasks, including segmentation and feature extraction, are handled directly by a deep-learning architecture. The authors employed scaled batches of 32 by 32 tiles that had been converted into grayscale fundus images in order to train a CNN. The identification accuracy rate was 98.6%. In [33], the authors used random Boltzmann machine (RBM) and deep neural network (DNN) methods to assess the change in arterial blood vessels. To define the features for the deep-learning system, they combined the OD region and AVR ratio. The authors achieved much greater accuracy in that research. In [34], a technique for extracting the center fovea of the optical disc and retinal arteries was presented. CNN's seven levels comprise its architecture. The output of their CNN design consists of four nodes that represent the fovea centralis, optic disc, retinal vasculature, and retinal backdrop. An average classification accuracy of 92.68 percent was reached in the DRIVE dataset. Many research employed CNN for the classification and segmentation of the retinal vasculature into arteries and veins [35–37].

For the Drive dataset, these methods yielded outstanding classification accuracy rates of 93.5%, and 88.89% for a collection of 100 low-quality pictures. In [38], a CNN-based automated approach was presented to identify exudates in microscopic retinography. Throughout the CNN training procedure, all of the characteristics were retrieved. The modified pixel was situated at the patch's center, which was CNN's input. Convolutional layers were employed to calculate the likelihood that each particular pixel would exude fluid or not. Since there are no exudates in the optic disc region, they are eliminated. The input and output layers are not the only ones; the CNN design also includes four convolutional and pooling layers. Using the DRiDB dataset for evaluation, the method demonstrated equal sensitivity, a 77% F-score, and a positive predictive value (PPV).

In [39] a technique for classifying arteriovenous nicking using the patient's retinal pictures that was developed using deep learning. Using the suggested model, the dataset from the Structured Analysis of the Retina project (STARE) has been utilized to train and then classify the presence and absence of arteriovenous nicking.

A performance measure evaluation of the proposed model, aggregate residual neural network (ResNeXt), showed that it was more accurate than 94% on the test dataset. A comparison of recent research projects using DL approaches for either HR identification or the extraction of A/V, OD, and other HR symptom aspects is shown in Table 2.

CHAPTER 3

Background Information

There are two sorts of systems that have been established in the past for the automated recognition system of HR. One kind is connected to approaches based on complicated image processing [13–17, 19–21, 23, 27, 31, [32, 40]. The second group, however, is concerned with deep learning models [6, 15, 34–37], whose traits are automatically extracted without concentrating on image processing techniques. Tables 1and 2 compare the systems in terms of datasets, models, and accuracy. The systems based on deep learning outperformed those that extracted characteristics via complex image processing in terms of accuracy. However, these deep learning model networks require huge datasets to be trained without overfitting. when it comes to deep learning models There are two ways to train a CNN in the literature: either from scratch, or via transfer learning to tune an already learned model. Pre-trained models are a useful representation of transfer learning in a wide variety of computer vision applications. A few examples include the pre-trained models ResNet50 [9], InceptionV3 [41], and VGG [42]. The ImageNet dataset was used to assess how well these pre-train models performed [43, 44]. Large CNN models are a foundational component of many pre-trained transfer learning algorithms [45]. Two of the primary aspects influencing CNN's appeal over the last several years are its great performance and simplicity of training [46]. The convolutional base, which combines features.

Table 2: Deep learning model-based HR classification

Methodologies	DL Models	Dataset	Limitations
HR detection directly from greyish fundus images using CNN [6]	CNN	DRIVE	limited image dataset
A method that detection HR characteristics using CNN-RBM classifier[33	CNN-RBM	DRIVE STARE	Without statistical results
A technique for employing CNN to identify the retinal vessels and the optic disc's fovea centralis. [34]	CNN	DRIVE	Just normal or HR, not classify four stages of HR

A/V and OD classification using a CNN based method [35]	CNN	100 images Subset of EPIC	limited dataset
CNN-based method for identifying and segmenting arteries and veins (A/V)[36]	CNN	100 images UK Biobank	limited image dataset
An autonomous vessels segmentation technique using CNN. [37]	CNN	DRIVE	Just normal or HR, not classify four stages of HR
An automated technique based on CNN to generalize the features and classify fundus into HR or non-HR [38]	CNN	DRiDB	Limited dataset

maps and pooling layers, is one of the two sections of the architecture of a conventional CNN model. With the help of convolutional filters of different sizes, the feature map layer was created. The softmax classifier, which recognizes classes, is a second component. The categorization layer is completely linked in real life. The characteristics that are extracted in the first layer of the CNN deep-learning architecture are generic, but the features that are specified in the final layer are the most specialized [45]. ”Features transform” refers to the change from generic to more focused features. In Fig. 2, this standard CNN model is shown. There have been systems for identifying objects utilizing feature extraction and identification tasks in the past, and they frequently used deep learning models[47].

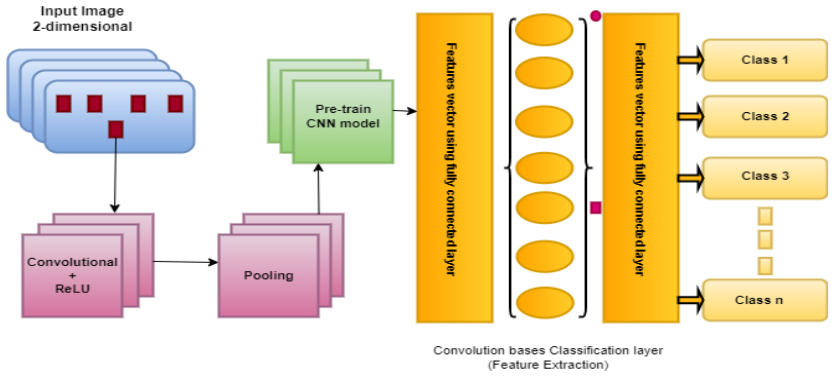

Fig. 2: CNN architecture

CHAPTER 4

Inception V3, VGG16 and VGG19 Architecture

4.1. Inception V3

The Inception v3 model, which was introduced in 2015, has 42 layers overall and a reduced mistake rate than its forerunners. Let's examine the many improvements that the Inception V3 model has received. The Inception V3 model has changed significantly, including:

- Factorization into Smaller Convolutions
- Spatial Factorization into Asymmetric Convolutions
- Utility of Auxiliary Classifiers
- Efficient Grid Size Reduction

Let's examine the implementation of each of these improvements and how they benefited the model.

4.1.1. Factorization into Smaller Convolutional

The extensive dimension reduction was one of the Inception V1 model's key advantages. The model's bigger Convolutions were factorized into smaller Convolutions to make it even better.

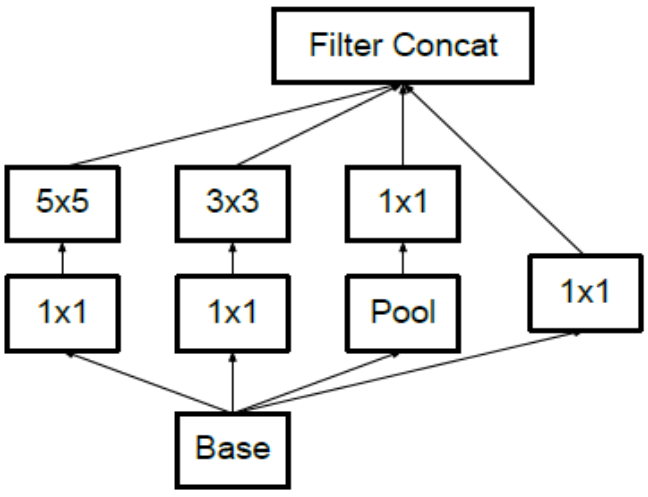

Fig. 3: Take the conception V1 module's basic module as an illustration.

It contains a 55 convolutional layer, which as was previously said was computationally costly. The 55 convolutional layer was therefore replaced with two 33 convolutional layers in order to lower the computational cost, as seen below.

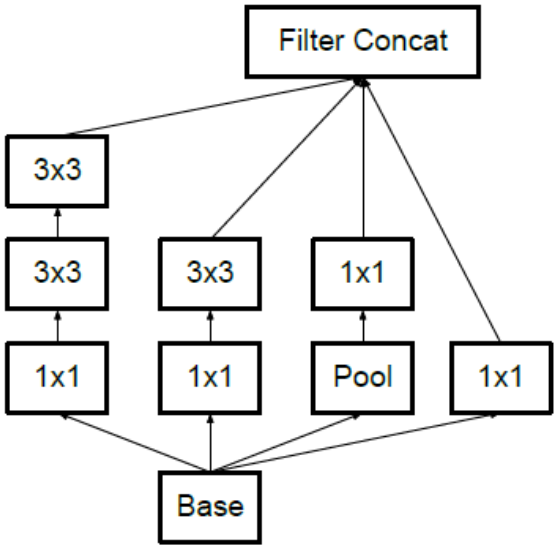

Fig. 4: 33 Convolutional layers in order

See how the use of two 33 convolutions minimizes the amount of parameters to have a better understanding of it.

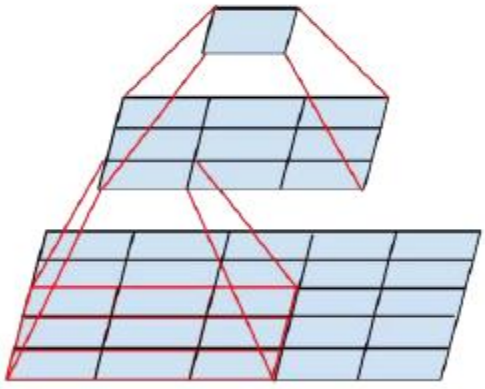

Fig. 5: Two 33 convolutions minimizes

The computing expenses also decrease as a result of the fewer parameters. A relative gain of 28% was obtained as a result of factorizing bigger convolutions into smaller convolutions.

4.1.2. Spatial Factorization into Asymmetric Convolutions

Despite the fact that bigger convolutions are split into smaller convolutions. You might be curious what would happen if we were to factorize any more, perhaps to a 22 convolution. But asymmetric convolutions were a superior option to improve the model's efficiency.

Asymmetric convolutions are of the form n×1

They therefore substituted a 13 convolution followed by a 31 convolution for the 33 convolutions. This is equivalent to sliding a two-layer network with a 33 convolution's receptive field.

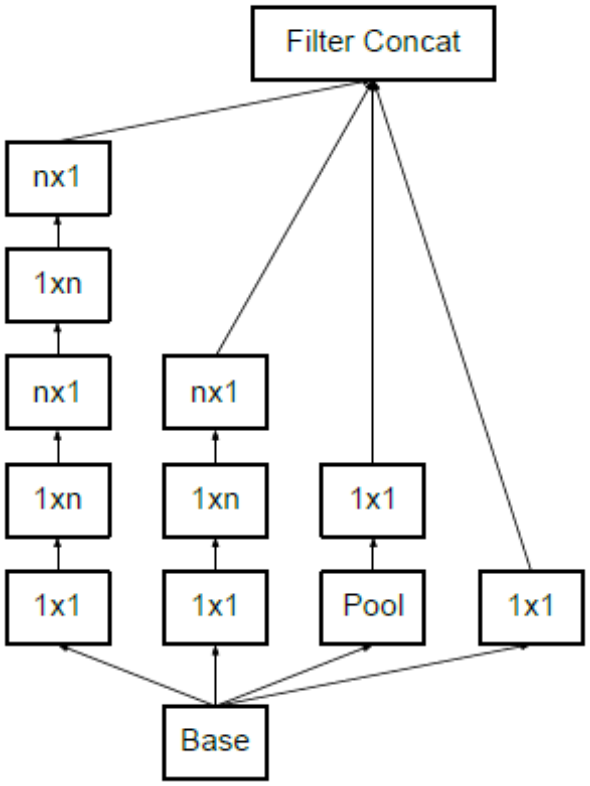

Fig. 6: *Asymmetric convolutions*

Structure of Asymmetric Convolutions

If both the input and output filter counts are the same, the two-layer approach is 33% less expensive. This is how the inception module appears following the use of the first two optimization approaches.

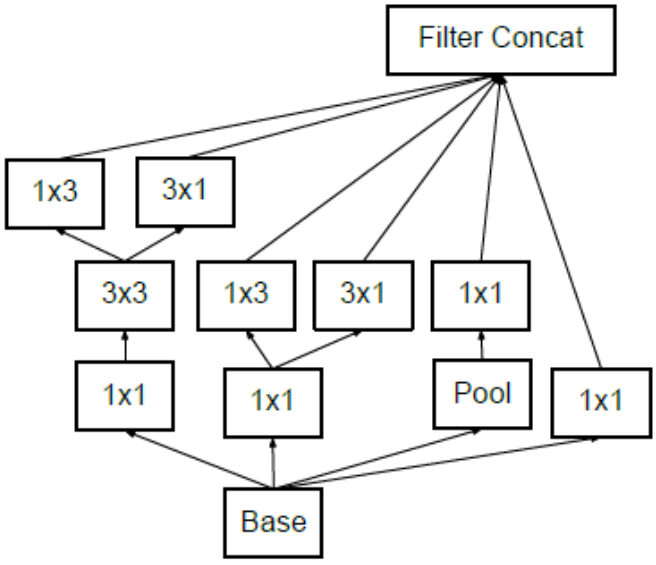

Fig. 7: *Structure of Asymmetric Convolutions*

4.1.3. Utility of Auxiliary classifier

Auxiliary classifiers are used to accelerate the convergence of extremely deep neural networks. In very deep networks, the vanishing gradient problem is mostly addressed by the auxiliary classifier.

Early on in the training, there was no improvement as a result of the auxiliary classifiers. However, as the experiment progressed, the network with auxiliary classifiers outperformed the network without auxiliary classifiers in terms of accuracy.

As a result, the Inception V3 model architecture's auxiliary classifiers function as a regularizer.

4.1.4. Efficient Grid Size Reduction

Traditionally, the grid size of the feature maps was decreased using average and maximum pooling. The activation dimension of the network filters is enhanced in the Inception V3 model to more effectively minimize the grid size.

For instance, reduction produces a d/2 d/2 grid with 2k filters from a dd grid with k filters.

And two concurrent blocks of convolution and pooling that were subsequently concatenated are used to accomplish this.

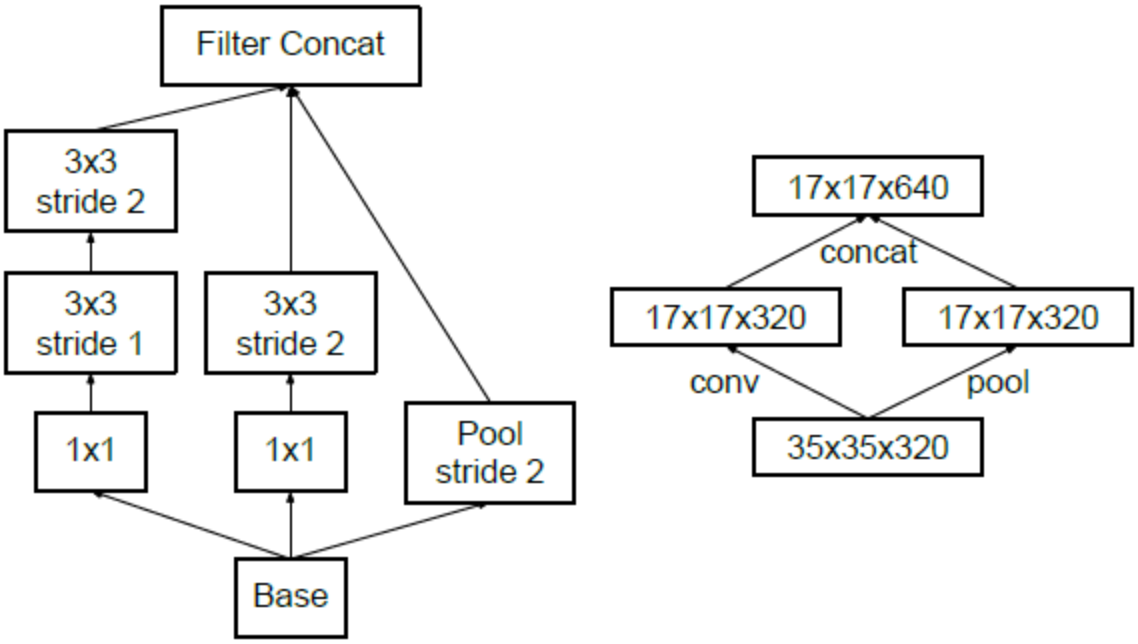

Fig. 8: Efficient Grid Size Reduction

The aforementioned figure demonstrates how the grid size is effectively decreased while the filter banks are expanded.

4.1.5. Inception V3 Explanation

The final Inception V3 model after all the improvements seems like this:

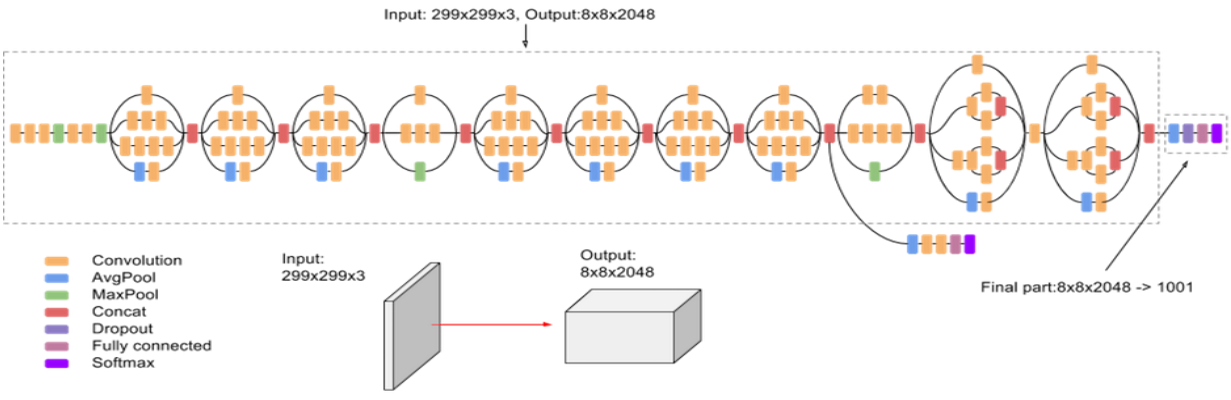

Fig. 9: Inception V3 Architecture

The inception V3 model has 42 layers in total, which is a little more than the inception V1 and V2 models. However, this model's effectiveness is absolutely remarkable. We'll get to it shortly, but let's first take a closer look at the parts that make up the Inception V3 model.

TYPE	PATCH / STRIDE SIZE	INPUT SIZE
Conv	3×3/2	299×299×3
Conv	3×3/1	149×149×32
Conv padded	3×3/1	147×147×32
Pool	3×3/2	147×147×64
Conv	3×3/1	73×73×64
Conv	3×3/2	71×71×80
Conv	3×3/1	35×35×192
3 × Inception	Module 1	35×35×288
5 × Inception	Module 2	17×17×768
2 × Inception	Module 3	8×8×1280
Pool	8 × 8	8 × 8 × 2048
Linear	Logits	1 × 1 × 2048
Softmax	Classifier	1 × 1 × 1000

Fig. 10: Inception V3 Model Layers

The inception V3 model's general structure is shown in the accompanying table. Each module's output size serves as the subsequent module's input size in this situation.

4.2. VGG16 Model

VGG16 is a VGG model variation with 16 convolution layers, and we have thoroughly investigated the VGG16 architecture.

VGGNet-16 has 16 convolutional layers and is desirable because to its relatively consistent architecture. It, like AlexNet, contains just 3x3 convolutions but a large number of filters. It may be trained on four GPUs for two to three weeks. It is now the most popular option for extracting features from photographs in the community. The VGGNet weight setting is freely accessible and has been utilised as a baseline feature extractor in many different applications and challenges.

It can be a little difficult to manage VGGNet's 138 million parameters, though. Transfer Learning can help people reach VGG. When the parameters are updated for greater precision and the model

is pretrained on a dataset, the parameter values may be used.

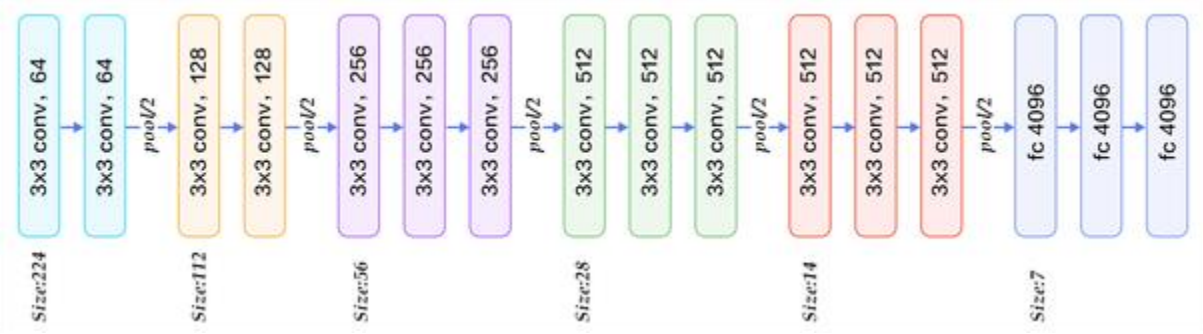

Fig. 11: VGG16 Model Layers

4.2.1. 16 Layers of VGG16

16 Layers of VGG16 are as following:

1.Convolution using 64 filters

2.Convolution using 64 filters + Max pooling

3.Convolution using 128 filters

4. Convolution using 128 filters + Max pooling

5. Convolution using 256 filters

6. Convolution using 256 filters

7. Convolution using 256 filters + Max pooling

8. Convolution using 512 filters

9. Convolution using 512 filters

10. Convolution using 512 filters+Max pooling

11. Convolution using 512 filters

12. Convolution using 512 filters

13. Convolution using 512 filters+Max pooling

14. Fully connected with 4096 nodes

15. Fully connected with 4096 nodes

16. Output layer with Softmax activation with 1000 nodes.

4.2.2. VGG 16 Architecture

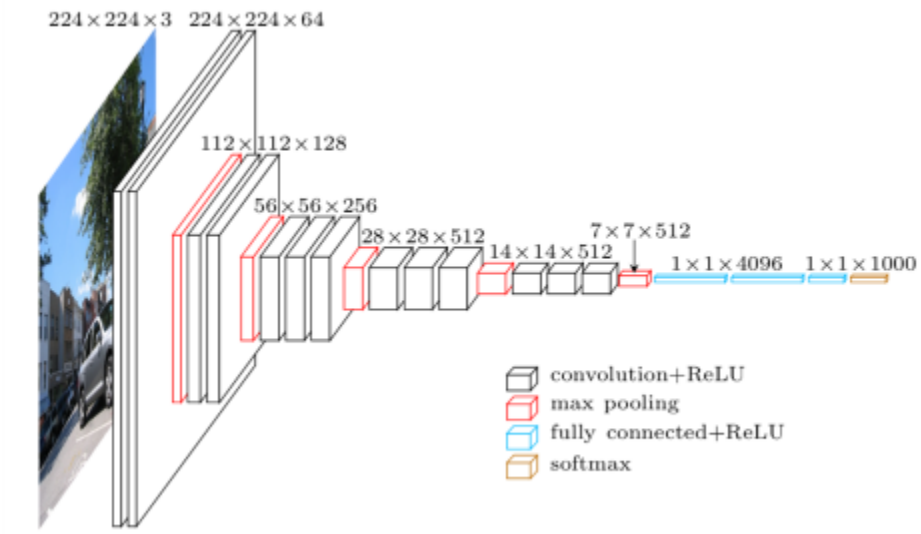

Fig. 12: VGG 16 Architecture

4.3. VGG19 Model

VGG19 is a variation of the VGG model that, in essence, has 19 levels (16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer). Other VGG variations include VGG11, VGG16, and more. FLOPs in VGG19 total 19.6 billion.

AlexNet, which was released in 2012 and improved on conventional Convolutional neural networks, can be seen as the successor to AlexNet. However, VGG was developed by a different team at Oxford University, hence the name Visual Geometry Group. It takes some concepts from its predecessors and builds on them while using deep Convolutional neural layers to increase accuracy.

Let's examine VGG19, compare it to previous iterations of the VGG architecture, and examine some of the VGG architecture's beneficial and real-world applications.

Before getting started, let's look into what the VGG19 Architecture is and have a look at ImageNet and CNN in general.

4.3.1. Convolutional Neural Network(CNN)

Let's first examine what ImageNet is. It is an image database with 14,197,122 pictures that are arranged in accordance with the WordNet hierarchical structure. This program aims to support image and vision researchers, students, and other stakeholders.

The ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), one of the competitions it hosts, tasked researchers from around the world with developing solutions that would produce the lowest top-1 and top-5 error rates (the top-5 error rate would be the percentage of images where the correct label is not one of the model's five most likely labels). A validation set of 50,000 photos, a test set of 150 000 images, and a 1,000 class training set of 1.2 million images are provided for the competition.

4.3.2. VGG 19 Layers and Architecture

Here's the VGG Architecture, which in 2014 outperformed other cutting-edge models and is still favored for many difficult tasks today.

VGG is a deep CNN that is used to identify pictures, to put it simply. The VGG19 model's layers are as follows**:**

- Conv3x3 (64)
- Conv3x3 (64)
- MaxPool
- Conv3x3 (128)
- Conv3x3 (128)
- MaxPool
- Conv3x3 (256)
- Conv3x3 (256)
- Conv3x3 (256)
- Conv3x3 (256)
- MaxPool

- Conv3x3 (512)
- Conv3x3 (512)
- Conv3x3 (512)
- Conv3x3 (512)
- MaxPool
- Conv3x3 (512)
- Conv3x3 (512)
- Conv3x3 (512)
- Conv3x3 (512)
- MaxPool
- Fully Connected (4096)
- Fully Connected (4096)
- Fully Connected (1000)
- SoftMax

4.3.3. VGG 19 Architecture

- This network received a fixed-size (224 * 224) RGB picture as input, indicating that the matrix was shaped (224,224,3).The only preprocessing that was done is that they subtracted the mean RGB value from each pixel, computed over the whole training set.
- They covered the entire idea of the image by using kernels of (3 * 3) size with a stride size of 1 pixel.
- To maintain the image's spatial resolution, spatial padding was applied.
- Using sride 2, max pooling was carried out over a 2 * 2 pixel frame.
- Rectified linear unit (ReLu) was used after this to add non-linearity to the model in order to enhance classification accuracy and computation time. As opposed to earlier models that employed tanh or sigmoid functions, this one performed far better.

- Developed three fully linked layers, the first two of which had a size of 4096, followed by a layer with 1000 channels for classification using the 1000-way ILSVRC, and the third layer being a softmax function.

The VGG net was primarily created with the intention of winning the ILSVRC, but it has been used to many other uses as well.

- Used simply as a good classification architecture for several different datasets, and since the authors made the models accessible to the public, they may be used without modification for more jobs that are comparable.
- Transfer learning is also applicable to facial recognition tasks.
- With other frameworks, like as Keras, weights are conveniently accessible and may be adjusted and utilised anyway the user sees fit.
- Loss of style and content while utilising VGG-19 network

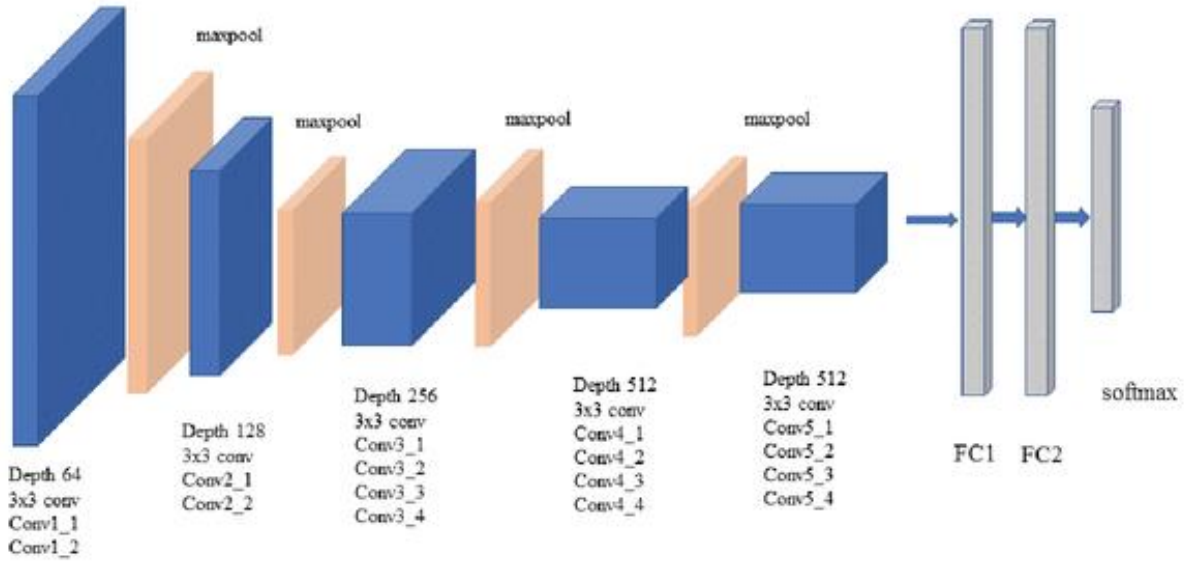

Fig. 13: VGG 19 Architecture

CHAPTER 5

Proposed Approach

In this study, the CNN (InceptionV3) and residual blocks are combined to create the Incept-HR system. Incept-HR system is used to categorize eye images as either non-HR infected eyes or HR infected eyes. In the Incept-HR system, valuable characteristics are extracted using the residual blocks method. To train on HR-related lesions, transform learning of residual blocks is used. The Incept-HR uses key processes to diagnose retinal fundus pictures and identify HR. Figure 3 graphically illustrates the phases in a systematic diagram. The characteristics that were retrieved from CNN (InceptionV3) and the residual block are combined. The Residual Blocks' settings are continuously changed throughout training. Next, a feature transform layer is created, which makes advantage of combining characteristics using element-wise multiplication Finally, the SVM classifier with a liner activation function were used to enhance classification outcomes.

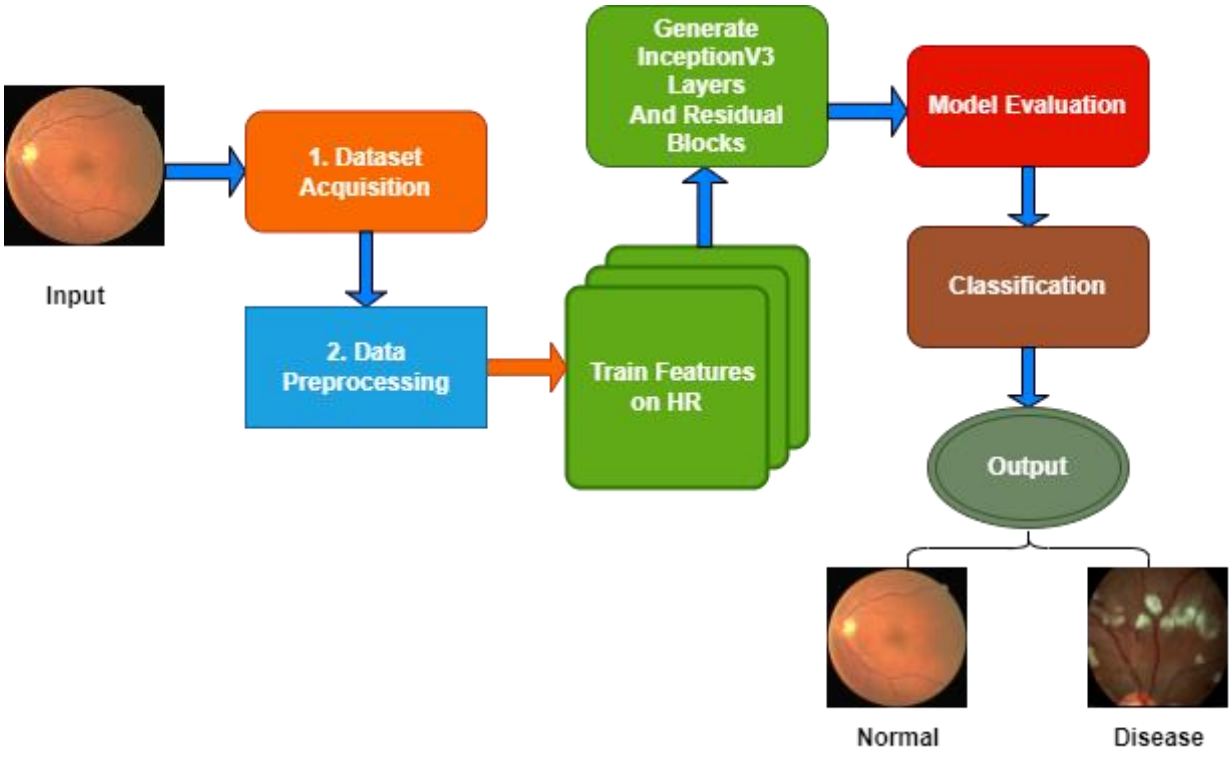

Fig. 14: An organized flow diagram of the Incept-HR system for classifying eye diseases

5.1. Dataset Acquisition and Pre-processing

Data acquisition is the process of capturing signals that gauge actual physical circumstances and transforming the resulting samples into computer-handling digital numeric values. To train and evaluate the Incept-HR model, a 6,000-image dataset Pak-HR as well as from well-known internet sources. The training dataset was created with the help of a professional ophthalmologist (manual separation of HR and non-HR fundus images from dataset collected). The doctor examined 6000 fundus pictures for HR-related features to create a benchmark, as shown in Figure1. In Table 3, we can see the breakdown of the three datasets (with different dimension settings) that were utilized to construct our testing and training fundus set. This experimentation required resizing all of these photographs to (700 x 600) pixels. After processing, binary labels are produced from these pictures. The dataset consists of 6,000 images in total, of which 3000 were utilized to test the system. The initial dataset converts the dataset of two classes into a binary dataset. By balancing the total of images with and without the disease, this is done to ensure that the dataset is objective. Images from the dataset are reduced in size to (700 x 600) pixels for pre-processing before being delivered to an algorithm designed especially for the Incept-HR model. Images are standardized to lower the variance between data points.

Table 3: Image dataset of the retina for the Incept-HR system.

Reference	**Name**	**HR**	**Non-HR**	**Image-size**	**Fundus Images**
[41]	DRIVE	100	150	(768 × 584) pixels	250
[48]	DiaRetDB0	80	80	(1152 × 1500) pixels	160
Private	Pak-HR	2100	3490	(1125 x 1264) pixels	5,590
		2,280	3,720	(700 × 600) pixels	6,000

Furthermore, data from Pak-HR is used to train and evaluate the proposed Incept-HR system. The dataset included 5590 retinal samples; 2100 were all from HR patients and the remaining 3490

were from non-HR patients. All of the JPEG files were saved at a resolution of 1125 x1264. The photographs were taken as part of routine testing for hypertension. Using information from three different sources, these pictures were reduced to a more usual resolution of 700 x 600 pixels. Also, expert ophthalmologists are engaged in the process of producing both HR and non-HR data for ground truth examination. Figure 4 is a fundus picture taken from one of the datasets utilized in this investigation.

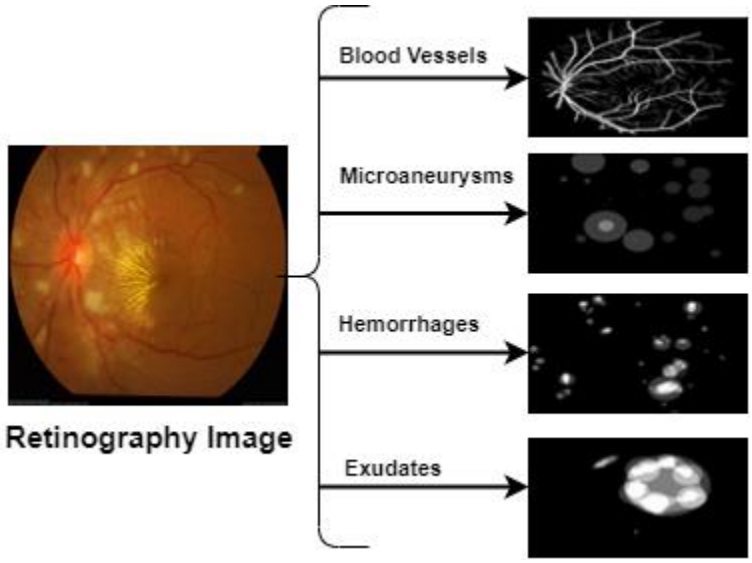

Fig. 15: An HR fundus image used for Incept-HR system training

5.2. Data Augmentation

Data augmentation is the technique of creating additional data points from current data in order to artificially increase the amount of data. In order to amplify the dataset, this may include making small adjustments to the data or utilizing machine learning models to produce new data points in the latent space of the original data. The dataset table makes clear that the dataset that was gathered is uneven. As a result, the model might be biased in favor of one class during training. In order to train a model with more diverse data without collecting new data, a procedure known as ”data augmentation” can be used. Data augmentation is the technique of creating additional data points from current data in order to artificially increase the amount of data. It helps to improve and stabilize the model’s performance. Aids in avoiding the problem of over-fitting as well. The Keras

deep learning package has a feature for data augmentation. We used the ImageDataGenerator class to add to our data. We used rotation, width shift, shear, zoom, crop, horizontal-flip, vertical-flip, and fill mode. some of these augmentation techniques results are shown in Fig 5. Also the parameters used for images augmentation are given in Table 4 As a result of the augmentation procedure, more image data were generated. We only used the data augmentation strategy for the training dataset. And for our model evaluation, we used the original photos rather than the augmented ones. To shift an image, all of its pixels must be moved in the same direction. Two different types of shifting exist (width shift and height shift). Flipping of an image means reversing the columns or rows of pixels in case of When a picture is flipped, the pixels' rows or columns are reversed, depending on whether the image is horizontal or vertical. It resembles turning an item up or down or left or right.

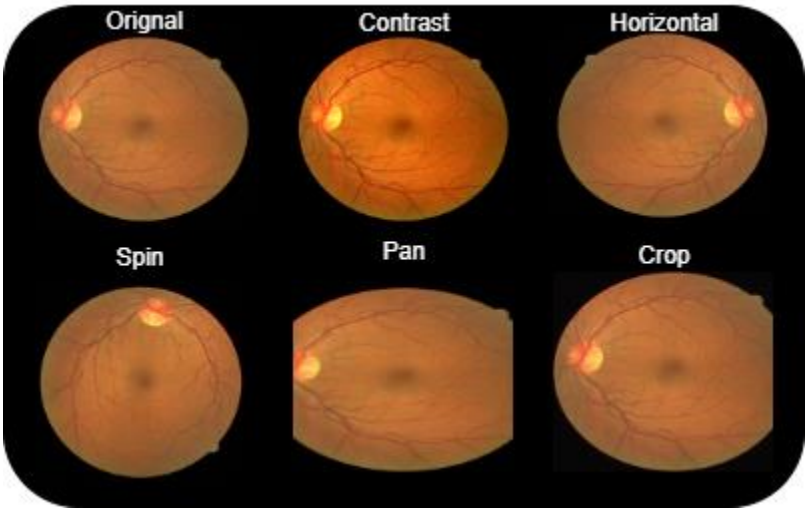

Fig. 16: Image augmentation techniques

Table 4: Data Augmentation table

Techniques for Augmentation	Values
Rotation range	15
Width shift range	0.2
Shear range	0.2
Zoom range	0.2
Crop	Ture
Horizontal-flip	True
Vertical-flip	False
Fill mood	Nearest

5.3. Incept-HR Architecture

The suggested Incept-HR system's architecture is designed to detect HR from retinography images. The design was built on the specifications of the InceptionV3 and Residual blocks. It focuses on utilizing fewer processing resources and is more computationally effective. This is a multi-level feature extraction. The Incept-HR model is made up of residual blocks, symmetrical building blocks, and asymmetrical building blocks. It contains completely linked layers, convolution, maximum pooling, average pooling, and is concatenated. Below Fig. 6, a schematic diagram is shown. We are using several fundus images to apply our approaches. Using an enhanced training dataset, we trained the Incept-HR model. To generate the feature map, this architecture uses 1 input layer, 93 Cov2d layers, 93 batch normalization layers, 92 activation layers, 10 mixed layers, 9 average pooling layers, 3 maximum pooling levels, 1 concatenate layers, and residual blocks with two separable convolutional layers and one convolutional layer of kernel size one. Following that, one flattened layer, one dense layer, and an SVM classifier with a liner activation function were used for classification. Test-fundus photos were used to assess our model. Our algorithm will also be able to distinguish between fundus pictures that are normal or diseased.

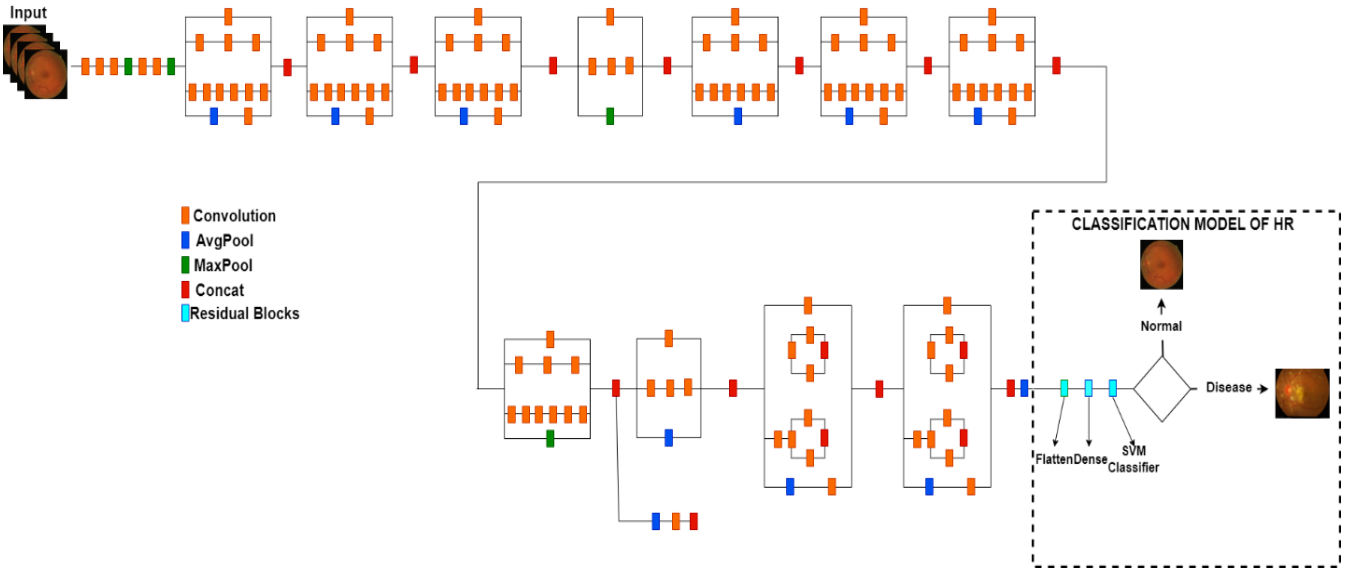

Fig. 17: Schematic diagram of Incept-HR

5.4. Recognition of HR

Computer-aided diagnostic (CAD) systems have a difficult time automatically identifying HR from retinal fundus images. The residual blocks are added to the InceptionV3 model to address this issue. The multi-layer design of the Incept-HR process is utilized to categorize retinal fundus images. Additionally, the network introduces the skip connection to quicken the learning process. Fig. 3 provides a graphic representation of these actions. Incept-HR is the name of the system as a whole. There are several convolutional layers in the case of residual blocks, which are followed by the activated ReLU function, a maxpooling layer, a batch normalization layer, which are dense layers. All of these layers are linked together with a skip connection so that the residual network can be trained well. The three Residual Blocks (RBs) are further economical for creating trained three-CNN-feature maps with deep depths in order to produce the dense block. Every RB's input and output are of identical size in the feature learning stage. Down sampling from each RB by a factor of two is used for the feature maps. Three RB were created for this model. The last fully linked output of the network is expanded by a layer to carry out the categorization operation. This layer is used to divide the characteristics of the residual learning sub-network into two groups. To guarantee that this network performs to its full potential, the number of neurons is fixed at 850. As a preprocessing step, InceptionV3 inserts a batch normalization layer between the auxiliary classifier and the fully connected layer. The batch stochastic gradient approach may be used in the batch normalization model to accelerate deep neural network training and model completion. The batch normalization formulas are described as follows:

$$B = \{ X_{1.....m} \}, \gamma, \beta \quad (1)$$

$$\{ y_i = BN_{\gamma\beta}(X_i) \} \quad (2)$$

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^{m} Xi \quad (3)$$

$$\sigma_B \leftarrow \frac{1}{m} \sum_{i=1}^{m} Xi - \mu_B \quad (4)$$

$$X_i \leftarrow \frac{X_i - \mu_B}{\sigma_B^2 + \epsilon} \quad (5)$$

$$y_i \leftarrow YX_i + \beta = BN_{\gamma\beta}(X_i) \quad (6)$$

Table 5: Notation table

Techniques for Augmentation values	Values
B	Batch
X	batch minimum activating value
μB	mini-batch mean
σ2B	mini-batch variance
ϵ	constant added for numerical stability
β	learning parameter
γ	learning parameter

Here, x is the batch B minimal activating value, m is the number of activating values, and are accessible variables (is responsible for modifying the value distribution's volatility and is in charge of moving the averaged value's location), depict the overall average in a single dimension, are the point differences for each feature map dimension, and are a constant. InceptionV3 introduces batch normalization to address the standard deep neural network's distribution discrepancy among inputs and outputs, which poses significant challenges to feature selection. The learning impact is improved by normalizing the inputs through each layer.

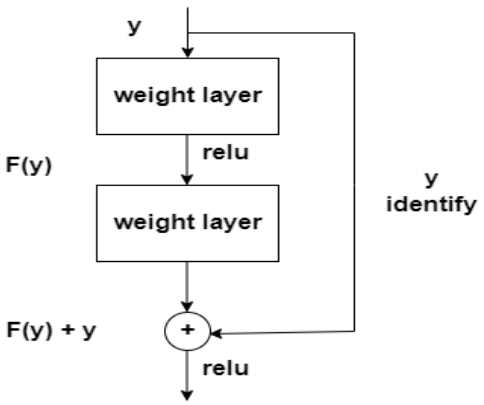

Fig. 18: Residual Blocks

$$RB(y) = Output - Input = H(y) - y \qquad (7)$$

$$H(y) = RB(y) + y \qquad (8)$$

The residual block as a whole is analyzing the real output to improve its performance H(y). If you examine Fig.7 attentively, you'll see that because of the y-identity link, the residual that the layers are attempting to learn, RB(y). So, the conventional network layers are gaining an accurate understanding of the output (H(y)), while a residual network's layers are acquiring knowledge about the residual RB(y).

4.5. Train Feature on HR

To efficiently extract features from HR or Non-HR photos, a lightweight pre-trained CNN model based on a depth-wise separable convolution neural network is used. This is the first time, to our knowledge, that a depth-wise separable convolution neural network for feature extraction is employed in an HR detection system. The depth-wise separable convolution divides the ordinary convolution into two components. The first component is depth-wise convolution, which is used to extract features from each input channel individually. The second phase is the pointwise convolution, which combines the result of the depth-wise convolution using 1 1 convolution. When compared to ordinary convolution, depth-wise separable convolution reduces the number of parameters and the computational cost of the convolution layer dramatically.

Table 6: Algorithm 1: Implementation of the Incept-HR model for feature map extraction.

Input Array X	**Array X**
output procedure	Feature map extraction x = (x1, x2 . . . , xn)
step1	Input normalization of raw data
step2	Function definition
Step3	Kernel sizes and array X, which comprises a number of filters, are the inputs to the Conv-batch Norm. a. X = Conv (X) and b. bX = BN (X) then applied
Step4	Depthwise Conv2D was used rather than Conv2D.
Step5	Establishing the network a. 14 Conv layers, each comprising 32,64,128,256,512,1024 make up the first step of the procedure After each of them, the ReLU is subsequently activated. b. The next step is to use Add to use Skip Connection.

	c.Three distinct skip connections were utilized. Each Skip Connection has two Depthwise Conv layers after the Maxpool layer The skip connection has two strides and a conversion ratio of 1 to 1.
Step6	After that, using the flattened layer, the feature map x = (x1, x2,., xn) was created and flattened.

4.6. LSVM

The linear SVM machine learning classifier is used for the automatic classification of HR using a train test splitting strategy of 75% to 25%. Linear SVM is used because of its effectiveness in dealing with small datasets and as well as its performance. SVM is a machine learning classification strategy that outperforms other types of classifiers and is commonly used to tackle real-world issues. For computer vision or image classification challenges, the authors developed a depth-wise separable CNN instead of a deep learning or machine learning classifier. We picked the Linear SVM classifier for our research because it can handle small datasets while still performing well in high-dimensional settings. Given that we were working with binary classification issues, it seemed logical to use linear SVM. Another reason for employing linear SVM was to improve the efficacy of our approach and to identify the optimum hyperplane that divides the feature space of ill and normal cells in retinal pictures. An LSVM generally accepts a vector Y = (z1, z2,..., zn) and produces a value y c Rn,
which may be written as:

$$Y\ out = (Weig, Ziv) + c \quad (9)$$

Weig represents the weight and c represents the offset in Equation (9); both Weig and c belong to R and are learned during training. Ziv is the input vector, and it is allocated to class 1 or class 1 depending on whether y is larger than or less than 0.

Table 7: Algorithm 2: Proposed LSVM classifier

Input	Extracted feature map x = (z1, z2,. , zn) with annotations x = 0,1, Test data Ztest
output	Classification of normal and abnormal samples
step1	Initially, the classifier and Kernel Regularize L2 parameters are defined for optimization.
Step2	Construction of LSVM a. The training process of LSVM is completed using extracted features x = (z1, z2,. , zn) by our Algorithm 1. b For the generation of the hyperplane, use Equation (6).
Step3	The class label is allocated for testing samples ztest using the decision function of the equation below. Ztest = (Weig, Ziv) + c

CHAPTER 6

Results

Incept-HR is trained using a dataset of 3000 retinal images, both HR and non-HR. These retinography images were obtained from a number of respected hospitals in Pakistan (Pak-HR) as well as from well-known internet sources. To perform feature extraction and categorization tasks, all 6,000 photos were downsized to (700 x 600) pixels. The Incept-HR system is developed using the InceptionV3 and residual blocks. The Incept-HR model has been trained for 100 epochs; the best model, with a f1-score of 0.99, was found in the 20th epoch. Accuracy (ACC), specificity (SP), and sensitivity (SE) scores were calculated using statistical analysis to evaluate the proposed Incept-HR system's efficacy. The created Incept-HR system's performance is measured against these metrics and compared to that of other systems.

A computer with an HP-i7 processor, 8 cores, 16 GB of RAM, and a 2 GB Gigabyte NIVIDA GPU was utilized to construct and develop the Incept-HR system. Windows 11 Professional 64-bit is installed on this machine.

6.1. Experiment 1

In the first experiment, we use a 10-fold cross validation testing strategy to evaluate the AUC metric obtained against those obtained from different DL techniques. The area under the curve (AUC) was the primary parameter used to evaluate the classification performance. Table 5 displays the results of our quantitative analysis of the produced Incept-HR system's performance. The developed approach has a low training error (0.1) and high AUC (99%) for detecting HR eye disease.

Table 8: Performance metrics of Incept-HR

Hypertensive Type	SE	SP	ACC	AUC	Error Rate
Diabetic Hypertension (HR)	98%	99%	99%	0.99	.01

Non-hypertension (non-HR)	99%	98%	99%	0.99	.01
Average Result	**99%**	**99%**	**99%**	**0.99**	**.01**

6.2. Experiment 2

In this experiment we train deep learning models(VGG16, VGG19) to compare them with the proposed system Incept-HR. It is worth mentioning that these deep-learning models employed the same number of epochs in their training. After selecting the best performing network with a validation accuracy, two identical Deep neural Networks were trained. Table 6 and Figure 8 display the results of a comparison between the Incept-HR system and the VGG16, VGG19 models in terms of sensitivity, specificity, accuracy, and area under the curve (AUC). The comparison shows the superior performance of Incept-HR over VGG16 and VGG19.

Table 9. Comparison of the effectiveness of the Incept-HR system developed with other deep learning systems using a dataset of 6000 examples.

Methodologies	SE	SP	AUC	ACC
VGG19	87.0	88.0	0.89	91
VGG16	84.0	87.0	0.87	90
Developed Incept-HR System	**99.0**	**99.0**	**0.99**	**99**

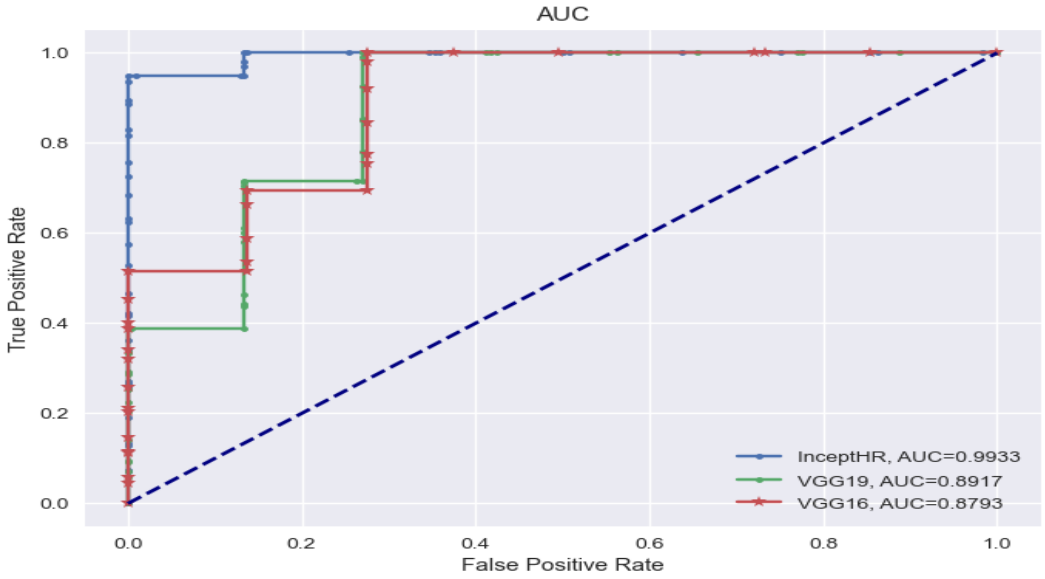

Fig. 19: Compares the proposed Incept-HR system with the area under the receiver operating curve (AUC) of several deep learning methods on a chosen dataset.

6.3. Final Experiment

In this experiment, we use a new dataset Pak-HR collected from Pakistani hospitals to evaluate the efficacy of our proposed Incept-HR system. We started by comparing the model's performance on the training and validation sets, as well as the loss function, by utilizing both sets of data. Figure 9 displays the training and validation accuracy of the dataset-using Incept-HR model. Figure 9 shows that our model is effective in both training and validation settings. Using the retina dataset, we were able to obtain perfect accuracy on both the training and validation sets.

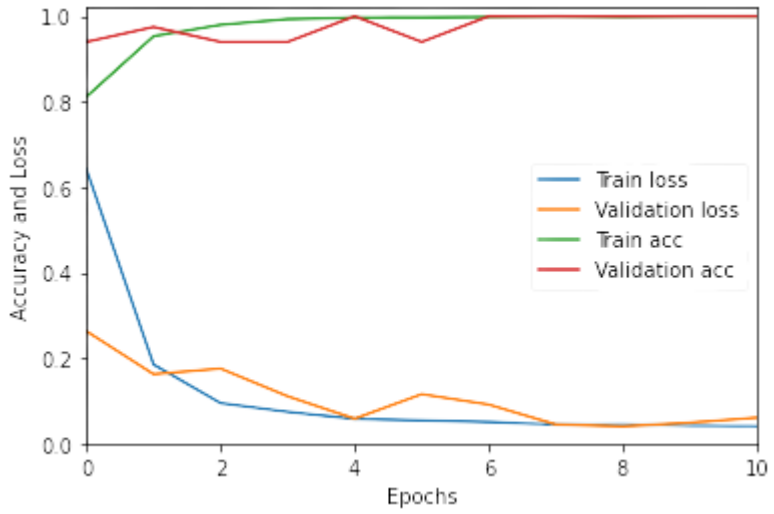

Fig. 20: Accuracy and loss on training and validation

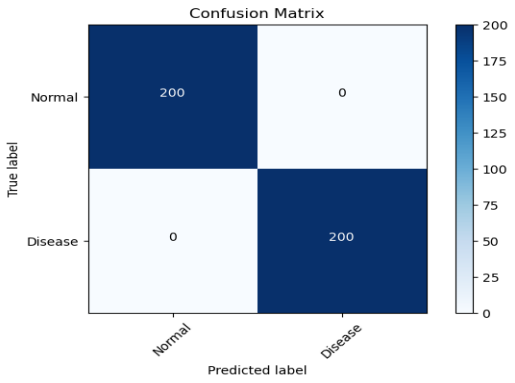

Fig. 21: Confusion matrix of implemented Incept-HR.

CHAPTER 7

State of the art comparison

Only a few researchers have attempted to use deep learning methods to find HR in retinal pictures. The most recent research that uses DL for detecting HR in retinal pictures are Triwijoyo-2017 [6] and CAD-HR [49]. CAD-HR is state of the art deep learning model used for detecting HR. It can be seen from Table 7 that Incept-HR has a superior performance over CAD-HR.

Table 10: performance comparison between Incept-HR, Trivijoyo-2017 and CAD-HR

Methods	SE	SP	ACC	AUC
Triwijoyo-2017 [7]	78.5%	81.5%	80%	0.84
CAD-HR [14]	94%	96%	95%	0.96
Proposed Incept-HR	**99%**	**99%**	**99%**	**0.99**

Consequently, in comparison, the developed Incept-HR system produced better results, with 92%, 97%, 99%, and 0.99 for SE, SP, ACC, and AUC, respectively. The identification accuracy in CAD-HR was 95%, according to the authors of this study [49]. In Triwijoyo-2017 [6] we observe that they employed a very restricted collection of input fundus for training, as they only comprise 40 retina images, of which 20 are normal and 20 are HR, which led to the high degree of precision and accuracy but the same their dataset is not approved from expert optometrists. Therefore, with the approval of expert optometrists, our Incept-HR system was tested and trained on a balanced 6000 images dataset. As a result, we achieved 99% accuracy in our classification, which is considered a big improvement over state of the artwork. As a result, we achieved 99% accuracy in our classification which is considered a big improvement over state of the artwork.

CHAPTER 8

Discussion

The Incept-HR system uses a trained CNN model called InceptionV3 as an input to classify the HR using four successive residual blocks and an output layer that is fully linked. A multilayered hierarchical architecture was built to learn enhanced properties. Without the aid of a specialist, this multilayer architecture used learning processes to automatically extract features from the input picture. The remaining blocks are added to the original model to add more universal outcomes and features for the development of the Incept-HR architecture. The CNN model, which is used to learn deep features, consists mostly of convolutional, pooling, and fully connected layers. Before being utilized to build the model, these layers must be trained and shown to be effective in extracting valuable features. An independent feature learning technique allowed for the success in detecting HR. This makes our method superior to handcrafted-based classification systems that rely on the pre-processing, segmentation, and localization of HR-related data, which are highly time-consuming and complicated algorithms, for the diagnosis of HR sickness.

There were a number of significant issues when HR automated systems were developed using conventional methods. The first issue is that it is extremely challenging to identify and extract crucial HR-related lesion features when using advanced pre- or post-image processing techniques from retinography photographs. The network cannot be trained or tested since there are no datasets with clinical expert labelling to describe specific HR-related lesion patterns. As a result, it is difficult for automated systems to Such disease related characteristics are difficult for automated systems to recognize. Our system solves both issues first, we propose using deep learning models for the extraction of important features from the eye. Second, we introduce a new dataset PAK-HR. To learn feature extraction, several models, however, used trained models created from scratch, although they all used the same weighting technique at each stage. When this happens, it may be challenging for layers to transmit weights to the deeper network layer for exact choices. This study develops the Incept-HR system, which uses two multi-layer deep learning approaches to distinguish between HR and non-HR without concentrating on techniques for image processing to overcome the aforementioned problems. The following are some of the Incept-HR system most significant contributions. Incept-HR system to the best of our knowledge is the first attempt of

using InceptinV3 and residual blocks in HR classification. To determine the most useful feature maps, establish the precedence of features, and enhance the effectiveness of the learning process, the first model was used to acquire four distinct HR-related injuries. The initial system created for this HR classification uses a perceptual-oriented color space, and the deep features are categorized using InceptionV3 and Residual blocks. In this research, the multilayer deep learning network needs to be trained on a large number of samples before it can be used to make the Incept-HR system. This is so that the learned features will be more general. Future versions of the Incept-HR system for HR identification might include a greater selection of retinography images that have been obtained from diverse sources. To enhance the model's classification performance, it might additionally incorporate handmade features in addition to deep features. Since much research used the saliency maps approach [50– 52] to segregate HR-related injuries, those injuries were then extracted from retinography pictures using a classifier. These studies merely segmented the population. In the future, these saliency maps will be put together to improve how well HR eye-related sickness is categorized. Also In the future, classification of the degree of the HR illness will be considered. Numerous studies conducted recently suggest that clinical characteristics are crucial criteria for determining the degree of HR intensity. However, in order to determine the degree of HR illness, those HR-related lesions with varying thresholds will be extracted which is a challenging task.

Chapter 9

Conclusion

The current method focuses on collecting dataset for HR classification from different sources and classifying HR-related traits (such as the ratio of arteries to veins, cotton wool patches, microaneurysms, vascularity, and hemorrhages), using deep learning techniques. Therefore, this hypertensive retinopathy identification system is constructed using feature selection and image processing skills. A few research publications classify HR diseases in the eye using deep learning, according to the presented literature review. primarily due to their inability to locate a suitable dataset for training such models. Due to this, categorization accuracy was similarly poor. Due to these shortcomings, it is deemed inappropriate to use these approaches as a screening tool for HR detection. An innovative computerized HR system called Incept-HR has been created to address these issues. The proposed system uses InceptionV3 and residual blocks to classify images. Adding residual blocks to the model makes increased its performance even though dynamic acceleration and compression of the system should reduce the accuracy of classification. The system achieves high accuracy around 99%.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

[1] Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., Das, S.R., De Ferranti, S., Despr´es, J.-P., Fullerton, H.J., et al.: Heart disease and stroke statistics—2016 update: a report from the american heart association. circulation 133(4), 38–360 (2016)

[2] Akagi, S., Matsubara, H., Nakamura, K., Ito, H.: Modern treatment to reduce pulmonary arterial pressure in pulmonary arterial hypertension.

[3] Gamella-Pozuelo, L., Fuentes-Calvo, I., Gomez-Marcos, M.A., RecioRodriguez, J.I., Agudo-Conde, C., Ferniandez-Martin, J.L., CannataAndia, J.B., Lopez-Novoa, J.M., Garcia-Ortiz, L., Martinez-Salgado, C.: Plasma cardiotrophin-1 as a marker of hypertension and diabetes-induced target organ damage and cardiovascular risk. Medicine 94(30) (2015)

[4] Suryani, E., et al.: The review of computer aided diagnostic hypertensive retinopathy based on the retinal image processing. In: IOP Conference Series: Materials Science and Engineering, vol. 620, p. 012099 (2019). IOP Publishing

[5] Qureshi, I., Ma, J., Abbas, Q.: Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. Multimedia Tools and Applications 80(8), 11691–11721 (2021)

[6] Triwijoyo, B.K., Budiharto, W., Abdurachman, E.: The classification of hypertensive retinopathy using convolutional neural network. Procedia Computer Science 116, 166–173 (2017).

[7] Garc´ıa-Floriano, A., Ferreira-Santiago, A., Camacho-Nieto, O., Y´a˜nez- ´ M´arquez, C.: A machine learning approach to medical image classification: Detecting age-related macular degeneration in fundus images. Computers & Electrical Engineering 75, 218–229 (2019)

[8] Asiri, N., Hussain, M., Al Adel, F., Alzaidi, N.: Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey. Artificial intelligence in medicine 99, 101701 (2019)

[9] Abbas, Q., Ibrahim, M.E., Jaffar, M.A.: A comprehensive review of recent advances on deep vision systems. Artificial Intelligence Review 52(1), 39– 76 (2019)

[10] Abbas, Q., Celebi, M.E.: Dermodeep-a classification of melanoma-nevus skin lesions using multi-feature fusion of visual features and deep neural network. Multimedia Tools and Applications 78(16), 23559–23580 (2019)

[11] Sengupta, S., Singh, A., Leopold, H.A., Gulati, T., Lakshminarayanan, V.: Ophthalmic diagnosis using deep learning with fundus images–a critical review. Artificial Intelligence in Medicine 102, 101758 (2020)

[12] Abbasi-Sureshjani, S., Smit-Ockeloen, I., Bekkers, E., Dashtbozorg, B., ter Haar Romeny, B.: Automatic detection of vascular bifurcations and crossings in retinal images using orientation scores. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 189–192 (2016). IEEE

[13] Akbar, S., Akram, M.U., Sharif, M., Tariq, A., Khan, S.A.: Decision support system for detection of hypertensive retinopathy using arteriovenous ratio. Artificial intelligence in medicine 90, 15–24 (2018)

[14] Akbar, S., Akram, M.U., Sharif, M., Tariq, A., ullah Yasin, U.: Arteriovenous ratio and papilledema based hybrid decision support system for detection and grading of hypertensive retinopathy. Computer methods and programs in biomedicine 154, 123–141 (2018)

[15] Cavallari, M., Stamile, C., Umeton, R., Calimeri, F., Orzi, F.: Novel method for automated analysis of retinal images: results in subjects with hypertensive retinopathy and cadasil. BioMed research international 2015 (2015)

[16] Grisan, E., Foracchia, M., Ruggeri, A.: A novel method for the automatic grading of retinal vessel tortuosity. IEEE transactions on medical imaging 27(3), 310–319 (2008)

[17] Holm, S., Russell, G., Nourrit, V., McLoughlin, N.: Dr hagis—a fundus image database for the automatic extraction of retinal surface vessels from diabetic patients. Journal of Medical Imaging 4(1), 014503 (2017)

[18] Tramontan, L., Ruggeri, A.: Computer estimation of the avr parameter in diabetic retinopathy. In: World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany, pp. 141–144 (2009). Springer

[19] Goswami, S., Goswami, S., De, S.: Automatic measurement and analysis of vessel width in retinal fundus image. In: Proceedings of the First International Conference on Intelligent Computing and Communication, pp. 451–458 (2017). Springer

[20] Ort´ız, D., Cubides, M., Su´arez, A., Zequera, M., Quiroga, J., G´omez, J., Arroyo, N.: Support system for the preventive diagnosis of hypertensive retinopathy. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, pp. 5649–5652 (2010). IEEE

[21] Manikis, G.C., Sakkalis, V., Zabulis, X., Karamaounas, P., Triantafyllou, A., Douma, S., Zamboulis, C., Marias, K.: An image analysis framework for the early assessment of hypertensive retinopathy signs. In: 2011 EHealth and Bioengineering Conference (EHB), pp. 1–6 (2011). IEEE

[22] Muramatsu, C., Hatanaka, Y., Iwase, T., Hara, T., Fujita, H.: Automated selection of major arteries and veins for measurement of arteriolar-tovenular diameter ratio on retinal fundus images. Computerized medical imaging and graphics 35(6), 472–480 (2011)

[23] Narasimhan, K., Neha, V., Vijayarekha, K.: Hypertensive retinopathy diagnosis from fundus images by estimation of avr. Procedia engineering 38, 980–993 (2012)

[24] Saez, M., Gonz´alez-V´azquez, S., Gonz´alez-Penedo, M., Barcel´o, M.A., Pena-Seijo, M., de Tuero, G.C., Pose-Reino, A.: Development of an automated system to classify retinal vessels into arteries and veins. Computer methods and programs in biomedicine 108(1), 367–376 (2012)

[25] Noronha, K., Navya, K., Nayak, K.P.: Support system for the automated detection of hypertensive retinopathy using fundus images. In: International Conference on Electronic Design and Signal Processing (ICEDSP), pp. 7–11 (2012)

[26] Nath, M.K., Dandapat, S.: Detection of changes in color fundus images due to diabetic retinopathy. In: 2012 2nd National Conference on Computational Intelligence and Signal Processing (CISP), pp. 81–85 (2012). IEEE

[27] Agurto, C., Joshi, V., Nemeth, S., Soliz, P., Barriga, S.: Detection of hypertensive retinopathy using vessel measurements and textural features. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5406–5409 (2014). IEEE

[28] Khitran, S., Akram, M.U., Usman, A., Yasin, U.: Automated system for the detection of hypertensive retinopathy. In: 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1–6 (2014). IEEE

[29] Khitran, S., Akram, M.U., Usman, A., Yasin, U.: Automated system for the detection of hypertensive retinopathy. In: 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1–6 (2014). IEEE

[30] Irshad, S., Akram, M.U.: Classification of retinal vessels into arteries and veins for detection of hypertensive retinopathy. In: 2014 Cairo International Biomedical Engineering Conference (CIBEC), pp. 133–136 (2014). IEEE

[31] Irshad, S., Akram, M.U.: Classification of retinal vessels into arteries and veins for detection of hypertensive retinopathy. In: 2014 Cairo International Biomedical Engineering Conference (CIBEC), pp. 133–136 (2014). IEEE

[32] Zhu, C., Zou, B., Zhao, R., Cui, J., Duan, X., Chen, Z., Liang, Y.: Retinal vessel segmentation in colour fundus images using extreme machine. Computerized Medical Imaging and Graphics 55, 68–77 (2017)

[33] Triwijoyo, B., Pradipto, Y.: Detection of hypertension retinopathy using deep learning and boltzmann machines. In: Journal of Physics: Conference Series, vol. 801, p. 012039 (2017). IOP Publishing

[34] Tan, J.H., Acharya, U.R., Bhandary, S.V., Chua, K.C., Sivaprasad, S.: Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. Journal of Computational Science 20, 70– 79 (2017)

[35] Tan, J.H., Acharya, U.R., Bhandary, S.V., Chua, K.C., Sivaprasad, S.: Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. Journal of Computational Science 20, 70– 79 (2017)

[36] Welikala, R., Foster, P., Whincup, P., Rudnicka, A.R., Owen, C.G., Strachan, D., Barman, S., et al.: Automated arteriole and venule classification using deep learning for retinal images from the uk biobank cohort. Computers in biology and medicine 90, 23–32 (2017)

[37] Yao, Z., Zhang, Z., Xu, L.-Q.: Convolutional neural network for retinal blood vessel segmentation. In: 2016 9th International Symposium on Computational Intelligence and Design (ISCID), vol. 1, pp. 406–409 (2016). IEEE

[38] Prentaˇsi´c, P., Lonˇcari´c, S.: Detection of exudates in fundus photographs using convolutional neural networks. In: 2015 9th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 188–192 (2015). IEEE

[39] Kriplani, H., Patel, M., Roy, S.: Prediction of arteriovenous nicking for hypertensive retinopathy using deep learning. In: Computational Intelligence in Data Mining, pp. 141–149. Springer, ??? (2020)

[40] Wu, J., Yang, S., Xiao, Z., Zhang, F., Geng, L., Cui, N., Zhang, D., Song, S.: Measurement of arteriolar-to-venular diameter ratio based on hessian matrix and multi-scale analysis. Journal of Medical Imaging and Health Informatics 8(1), 38–44 (2018)

[41] Staal, J., Abr`amoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. IEEE transactions on medical imaging 23(4), 501–509 (2004)

[42] Simonyan, K., Zisserman, A.: Very deep convolutional networks for largescale image recognition. arXiv preprint arXiv:1409.1556 (201

[43] Simonyan, K., Zisserman, A.: Very deep convolutional networks for largescale image recognition. arXiv preprint arXiv:1409.1556 (201

[44] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255 (2009). Ieee

[45] Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? Advances in neural information processing systems 27 (2014)

[46] Keshavarzian, A., Sharifian, S., Seyedin, S.: Modified deep residual network architecture deployed on serverless framework of iot platform based on human activity recognition application. Future Generation Computer Systems 101, 14–28 (2019)

[47] 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 (2016)

[48] Kauppi, T., Kalesnykiene, V., Kamarainen, J.-K., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., K¨alvi¨ainen, H., Pietil¨a, J.: The diaretdb1 diabetic retinopathy database and evaluation protocol. In: BMVC, vol. 1, pp. 1–10 (2007)

[49] Qureshi, I., Abbas, Q., Yan, J., Hussain, A., Shaheed, K., Baig, A.R.: Computer-aided detection of hypertensive retinopathy using depth-wise separable cnn. Applied Sciences 12(23), 12086 (2022)

[50] Gao, Y., Yu, X., Wu, C., Zhou, W., Lei, X., Zhuang, Y.: Automatic optic disc segmentation based on modified local image fitting model with shape prior information. Journal of healthcare engineering 2019 (2019)

[51] Niu, D., Xu, P., Wan, C., Cheng, J., Liu, J.: Automatic localization of optic disc based on deep learning in fundus images. In: 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), pp. 208–212 (2017). IEEE

[52] Zou, X., Zhao, X., Yang, Y., Li, N.: Learning-based visual saliency model for detecting diabetic macular edema in retinal image. Computational intelligence and neuroscience 2016 (2016)