Article

Computer-Aided Detection of Hypertensive Retinopathy Using Depth-Wise Separable CNN

Imran Qureshi [1,2], **Qaisar Abbas** [3], **Junhua Yan** [2], **Ayyaz Hussain** [4], **Kashif Shaheed** [5]
and Abdul Rauf Baig [3,*]

1   Department of Computer Software Engineering, Military College of Signals, National University
    of Sciences and Technology (MCS-NUST), Islamabad 44000, Pakistan
2   Key Laboratory of Space Photoelectric Detection and Perception, Ministry of Industry
    and Information Technology and College of Astronautics, Nanjing University of Aeronautics and
    Astronautics, Nanjing 211106, China
3   College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU),
    Riyadh 11432, Saudi Arabia
4   Department of Computer Science, Quaid-i-Azam University, Islamabad 44000, Pakistan
5   Department of Multimedia Systems, Faculty of Electronics, Telecommunication, and Informatics,
    Gdańsk University of Technology, 80-233 Gdańsk, Poland
*   Correspondence: abbaig@imamu.edu.sa; Tel.: +966-5633-36816

Abstract: Hypertensive retinopathy (HR) is a retinal disorder, linked to high blood pressure. The incidence of HR-eye illness is directly related to the severity and duration of hypertension. It is critical to identify and analyze HR at an early stage to avoid blindness. There are presently only a few computer-aided systems (CADx) designed to recognize HR. Instead, those systems concentrated on collecting features from many retinopathy-related HR lesions and then classifying them using traditional machine learning algorithms. Consequently, those CADx systems required complicated image processing methods and domain-expert knowledge. To address these issues, a new CAD-HR system is proposed to advance depth-wise separable CNN (*DSC*) with residual connection and a linear support vector machine (LSVM). Initially, the data augmentation approach is used on retina graphics to enhance the size of the datasets. Afterward, this *DSC* approach is applied to retinal images to extract robust features. The retinal samples are then classified as either HR or non-HR using an LSVM classifier as the final step. The statistical investigation of 9500 retinograph images from two publicly available and one private source is undertaken to assess the accuracy. Several experimental results demonstrate that the CAD-HR model requires less computational time and fewer parameters to categorize HR. On average, the CAD-HR achieved a sensitivity (SE) of 94%, specificity (SP) of 96%, accuracy (ACC) of 95% and area under the receiver operating curve (AUC) of 0.96. This confirms that the CAD-HR system can be used to correctly diagnose HR.

Keywords: retinography images; hypertensive-retinopathy; deep-neural network; transfer learning; color space transformation; residual neural network; convolutional neural network

1. Introduction

According to projections, 1.56 billion individuals will have hypertension by 2025. Furthermore, 66% of the hypertensive population resides in underdeveloped or impover-ished nations, where a lack of adequate healthcare services to identify, manage, and treat hypertension exacerbates the problem [1–3]. Hypertension is triggered by an increase in arterial pressure [4]; it is the prevalent pattern of eye illness that has lately been globally increasing. Several human organs, including the retina, heart, and kidneys, are damaged by hypertension [5]. Among these consequences, hypertensive retinopathy (HR) [6] is the main cause of cardiovascular disease, leading to death. Therefore, it has been identified as a global public health threat. Detecting and treating hypertension early can reduce the risk

of HR. HR is notoriously challenging to detect in its earliest stages because there is a dearth of modern imaging equipment and trained ophthalmologists.

Generally, HR is an abnormality in the retina triggered by an excessive rise in blood pressure. The existence of signs such as hemorrhage (HE) spots in the retina, cotton wool patches, and micro-aneurysms are potential signs of HR-related eye illness. Early detection and proper treatment of HR-related eye illness [7] is critical to saving human life. Many researchers have recently demonstrated that retinal experts use of a digital fundus camera to acquire microscopic retinal samples to assess the presence of micro-vascular alterations caused by HR is a cost-effective and non-invasive approach. Many HR patients are non-invasively screened using this fundus camera since it is inexpensive, simple to use, and most anatomical features of lesions are visible using this type of imaging. The major goal of utilizing automated systems [8] is to determine the existence of HR while relieving ophthalmologists of the burden of vast image assessment quickly and at an early stage.

Many previous research efforts have established procedures for retinal image analysis, including image enhancement, segmentation of HR lesions and retinal vessels, extracting features, and supervised machine learning classifiers for HR illness [9]. Oddly, a significant HR indication is a wideness in retinal veins, which reduces the A/V ratio (the mean artery to vein diameter). Detecting retinal vessel diameter [10] and other features such as AVR is difficult when using an image analysis system to diagnose HR-related eye disease. Furthermore, getting precise measurements of vascular diameters is quite difficult for ophthalmologists.

As stated in the above paragraph, ophthalmologists use automatic analysis of digital fundus images to identify HR by observing abnormalities in retinography photos. These abnormalities harm HR-related lesions as well as other key eye regions. If these alterations are not noticed early enough, they can progress to HR. It was found that HR-related eye disease can be reversed [11], whereas diabetic retinopathy (DR) related eye illness cannot be reversed. A normal retinal fundus image and images of HR affected fundi with different symptoms are illustrated in Figure 1. Eye specialists can use CAD systems to diagnose certain retinal diseases, such as HR-related eye illness. These technologies help academics and the health care industry by self-diagnosing. Ophthalmologists use such systems to diagnose and treat eye-related diseases, particularly HR-related diseases. The authors provided a recent survey for automated HR identification techniques, as illustrated, and compared in Section 2.

Figure 1. An instance of main HR symptoms are: (**a**) optic disk, cotton wool patches, hemorrhages and microaneurysms, (**b**) A/V ratio and (**c**) tortuosity.

A few studies suggested that hypertensive retinopathy (HR) might be automatically identified by segmenting the retina's structural features, such as the macula, optic disc (OD), blood vessels (BV), and microaneurysms (MA), as illustrated in Figure 1. Deep-learning techniques also automatically extract certain structures. These characteristics were then quantitatively assessed to determine irregularities and ultimately identify HR or non-HR cases. Extracting clinical features for HR disease diagnosis is a time-consuming and challenging operation for computerized algorithms. Instead of focusing on more serious

symptoms like cotton-wool patches or hemorrhages, researchers spent a lot of time only extracting the features. Microaneurysms were characteristic of early HR, but other lesions were critical in diagnosing malignant HR-related illnesses. In contrast, recently, deep learning methods have been deployed extensively for various computer vision applications and biological imaging analysis tasks.

Major Contributions

In this study, a CAD-HR system is designed to overcome the concerns that have been discussed above. This system makes use of a *DSC* model with residual connection and a LSVM for classification between HR and non-HR. The CAD-HR system has made several significant contributions, which are mentioned here.

1. This paper develops a new CAD-HR system to recognize HR eye-related disease based on a residual connection and *DSC* model LSVM.
2. To effectively extract features from HR and non-HR photos, a pre-trained CNN model that is lightweight and based on a *DSC* network is developed. Our utilization of a depth-wise separable CNN for feature extraction in a HR classification challenge is a first to our knowledge. To deal with a real-time context, the proposed feature extraction is more extensive.
3. For automatic HR classification, we employed a 75–25% train-test split utilizing the linear SVM machine learning classifier. The efficiency and performance of Linear SVM make it a popular choice, especially when working with limited data sets.
4. Extensive experiments were conducted by employing several statistical metrics on two publicly available and one proprietary benchmark, namely DRIVE, DiaRetDB0, and Imam-HR. A full comparative study comparing the suggested strategy to other existing DL approaches is presented.
5. The proposed CAD-HR system outperforms other transfer learning (TL) based architectures in recognizing HR.

2. Related Studies

Several automatic approaches for detecting retinal abnormalities have been proposed in the past. Despite this, few automated techniques have been implemented to identify HR, which is linked to an increase in blood vessel pressure. Many studies have previously used retinal fundus image processing methodology to automatically diagnose HR illness. Early diagnosis of HR using retinographic photo analysis saves ophthalmologists time as well as effort [12]. We have included the most up-to-date research on that topic in this section. We divided the literature into two groups: approaches built on hand-crafted features and approaches that rely on deep learning.

2.1. Handcrafted Studies

According to a survey of the literature, detecting hypertensive retinopathy (HR) from fundus images should be done using a segmentation-based methodology [12]. The scientists employed a machine-learning classification technique to identify HR from retinography photos in their experiments after first detecting distinct HR-related characteristics.

In [13–19], the artery to vein diameter ratio (A/VR), the optic disc (OD) location, mean fractal dimension (mean-D), papilledema indicators, and index of tortuosity are hand-crafted features. Several approaches used these features to measure the irregularities in the retina. Aimed at subdivision and neat tasks, the Gabor 2D or Cake-Wavelet were utilized, together with the Canny edge detector. To evaluate the performance of such systems, the DRIVE, AVR-DB, IOSTAR, STARE, DR/HAGIS, INSPIR/AVR and VICAVR datasets were employed. In [19], they used a supervised classifier to perform initial segmentation and refinement processes to find hemorrhagic lesions. The authors in [20] used a new technique for detecting HR-related eye illness. Cotton wool spots (CWS) were discovered and are among the more significant medical indications for determining HR illness. In the study, the authors enhanced candidate regions and then binarized the

image with the adaptive threshold approach. However, in paper [21], the authors achieved 82.2% of SE and 82.4% of PPV on a local dataset. They used a neural network (NN) and wavelet transform (WT) methods to detect all mentioned retinal diseases. Classification accuracies of 70% for HR were achieved. The AVR technique was used on a few color retinography pictures from ETDRS and DCCT datasets in the systems described in [22–25]. The A/V ratio was computed by segmenting the OD area and then classifying vessels as veins and arteries using the following operation gradients, morphological edge-detection, and Gabor/wavelet. An alternative technique is segmenting the OD area and classifying vessels as veins and arteries using the morphological edge-detection operation. In [26], a system with a GUI was created to perform half-automated retinal vascular recognition and assess the performance of HR illness identification. This GUI tool was also utilized for assessing the accuracy of determining the diameter of vessels anywhere of interest (ROI) to help people with HR understand their vascular risk. This tool was evaluated with the aid of the STARE and DRIVE datasets, and it performed well in classifying tasks. A clustering-based method to detect arteries and veins was used in combination with the A/V ratio in [27]. Alternatively, in [28], the moments and gray-level characteristics were retrieved as features to recognize fundus areas that fit to vessels. Differentiation between veins and arteries was achieved using color data as well as variation in intensity. On 101 images from the VICAVR dataset, various phases of HR-related eye illness were categorized using vessel width estimation. Two kinds of transforms were used, Hough as well as Radon, to segment retinal vessels, before estimation of the vessel diameter and tortuosity-index, and finally computation of the A/V ratio to detect HR from input retinography pictures [29]. Furthermore, in [30], an independent component analysis (ICA) was performed to a wavelet sub-band for identifying retina. In the study, the authors include different lesions such as optic disc, blood vessels, hemorrhages, macula, exudates and drusen. This approach was put to the test on 50 images. To categorize the different retinal blood vessel types, the authors in [31] used a several layered neural network that was supplied with invariant moments indicators. In [32], the researchers describe a nine-step automated system for extracting the different HR-related lesions. The 0.84 of AUC was attained on a collection of 74, while a 90% sensitivity and 67% specificity were achieved. According to the technique in [33], the classification of HR using DNN, and the Boltzmann Machines methodology is reviewed by determining the combined characteristics as well as changes in the location of OD in retinal pictures.

In a previous study, [34] described a method for detecting HR by figuring out characteristics from formerly handled color retinography photos. Firstly, it used CLAHE to transform the retinography photos into green-channel, resulting in a clearer view of the retina-vessels. Secondly, it used morphological closing to eliminate the optic disk. Subtraction was used to remove the backdrop. Then, utilizing zoning, characteristics were retrieved. Finally, as a classifier, a neural network with back propagation was utilized. It had a 95% accuracy rate. [35] described a method for segmenting retinal vessels from fundus pictures using the ELM classifier. Vectors of 39 local, morphological, and other characteristics were fed into the classifier as input features. On the DRIVE dataset, this approach had a 96% accuracy, a 71.4% sensitivity, and a 98.6% specificity.

2.2. Deep-Learning Based Studies

Classifying fundus images by means of any of the learning-based methods is a distinct approach. Pre-processing of the fundus is minimized in these methods. Many image-processing tasks, such as segmentation and feature extraction, may be accomplished directly with a deep-learning architecture.

In [7], an architecture of a deep learning model (DLM) for identifying HR was developed. They used 32×32 scaled consignments of grayscale transformed retina images to train a CNN. The CNN's output was either HR or regular. The accuracy of the detection was 98.6%. In [33] they used an architecture that combined the random Boltzmann machine (RBM) and deep neural network (DNN) methods to measure the alteration of arterial blood

vessels. The distinct characteristics for the deep learning method used the AVR ratio and the OD region. In a previous study [36] described a CNN-based approach for extracting the OD, retinal arteries, and fovea centralis. There were seven layers in the CNN architecture. The fovea centralis, OD, retinal arteries, and the background of the retina were represented by four nodes in the output of their CNN design. On the DRIVE dataset, a mean accuracy of classification of 92.6% was attained.

CNN was used by several researchers to segment and categorize retinal vasculature into arteries and veins [37–39]. The accuracy of these approaches was high: 88.9% for 100-photos of a poor-quality dataset and 93+% for the DRIVE data set. Finally, [40] described a CNN automated approach for detecting exudates in retinography photos. During the CNN training procedure, all the features were retrieved. The CNN was fed odd-sized patches as input, with the processed pixel at the patch's center. Convolutional layers were employed to determine whether each separate pixel belongs to an exudate or not. Exudates do not appear in the OD region. Hence, they were excluded. Aside from the input and output layers, the CNN design has four convolutional and pooling layers. Using the DRiDB dataset, the approach was found to have identical performance measures for positive predictive value (PPV) and sensitivity while having an F-score of 77%. Like in [41], the authors used different layers of CNN model to recognize diabetic retinopathy.

Table 1 summarizes the DL-based methods utilized for HR detection or at least the extraction of OD, computation of A/V and other HR aspects.

Table 1. Deep learning (DL)-based techniques for retinography photos' processing and classification.

DL Studies	Datasets	Limitations
HR Identification directly from fundus using CNN [7]	DRIVE	Very limited dataset
HR Identification using traditional feature vector and CNN-RBM classifier [33]	DRIVE STARE	No statistical results were shown
Four pretrained CNN to extract attributes from 4 diverse HR lesions then feed to a DRN to generalize those features and then classify fundus into HR or normal [35]	DRIVE, DR-HAGIS and Private dataset of 3170 fundus images	Not classifying HR into its four grades just HR or normal
Extraction of A/V and OD using CNN [36–39]	DRIVE, EPIC (100 fundus), UK Biobank (100 fundus)	Datasets used are very limited
Detect exudates in microscopic retinographs. using CNN [40]	DRiDB	Datasets used are very limited
Four pretrained CNN to extract attributes from 4 diverse HR lesions then feed to a DRN to generalize those features and then classify fundus into HR or normal [42]	DRIVE, DR-HAGIS and Private dataset of 3170 fundus images	Not classifying HR into its four grades just HR or normal

2.3. Methodoloigcal Overview

Feature extraction and recognition tasks based on deep learning models (DLMs) have been widely used in previous systems to recognize objects [40–44]. A novel active deep learning-based convolutional neural network (ADL-CNN) was implemented to overcome multilayered network architecture training challenges [45]. In practice, the ADL-CNN architecture is simple to train [46–49]. In addition, the CNN model has been applied in a great amount of research for the purpose of picture recognition on a broad scale [50]. The ADL-CNN model is unsurpassed in performance when compared to other models for use with minimal training datasets [51,52]. In previous studies, the researchers reported that the CNN model automatically learns distinguishing features from the raw samples [53].

In the past, the most common methods for identifying HR-related lesions within retinal images were conventional image processing techniques as well as machine learning algorithms. However, those traditional arts are slow and difficult to apply in real-time environments [54]. With the development of hierarchical deep learning methods, it is now possible to tackle all of the previously described issues. There are a wide variety of

deep learning approaches, including convolutional networks, deep belief networks (DBN), reinforcement learning machines (RBM), and deep reinforcement learning network (DRN) models [55,56].

The pre-training procedure is applied in a wide variety of computer vision applications to acquire characteristics for the purpose of resolving a variety of issues [57]. The transfer learning technique is an example of the pre-trained networks used in numerous computer vision systems. There are several pre-trained models that have been developed, including VGG [48], InceptionV3 [56], and ResNet50 [43]. The ImageNet dataset was utilized in the performance evaluation process of these pre-trained networks [58,59]. A substantial number of pre-trained models utilized in transfer learning are based on large CNNs. The high performance and easy training of CNN have raised CNN's popularity level in recent years [53,60]. A basic CNN model's architecture consists of two portions, including a convolutional base composed of a features map and pooling layers. Convolutional filters of various sizes were used to create the features map layer. In addition, the SoftMax classifier, which recognizes classes, is a second component. It is common for the classification layer to be fully interconnected in practice. The features extracted during the first layer of CNN's deep learning architecture are more generic, whereas the features defined during the last layer are the most specialized [61]. The process by which generic characteristics are transformed into more specific features is referred to as the features transformation [62,63].

3. Materials and Methods

3.1. Data Acquisition

In the interest of developing the CAD-HR diagnosis system, we must build a dataset with a sensible number of retinography images. Thus, we prepared 1310 HR retinography photos and 2270 regular retinography photos to test and compare the proposed CAD-HR diagnosis methodology. These retinography photos were obtained from two distinct web and one private source. An experienced ophthalmologist was involved in the development of the training dataset (manually distinguishing the HR and regular fundus photos from different datasets). To produce a gold standard, the medical professional inspected all the HR associated characteristics in a set of 3580 fundus photos, as shown in Figure 1. The explanation of the three separate datasets (used to prepare our training and testing fundus set) with varied lighting settings and dimensions exists in Table 2. All those images were re-sized to (700 × 600) pixels to conduct experiments.

Table 2. Retinal fundus datasets used with the CAD-HR-diagnosis system.

Title 1	Name	[1] HR	[2] Non-HR	Size	Fundus Images
[64]	DRIVE	100	150	(768 × 584) pixels	250
[65]	DiaRetDB0	80	80	(1152 × 1500) pixels	160
[PRV]	Imam-HR	1130	2040	(1125 × 1264) pixels	3170
		1310	**2270**	Downsized: (700 × 600) pixels	**3580**

[1] Hypertensive retinopathy (HR), [2] Non-hypertensive retinopathy (non-HR).

In addition, the proposed CAD-HR system is tested on and trained with data obtained from a private hospital known as Imam-HR. The Imam-HR dataset comprised a total number of 3170 retinal samples where 1130 images were from HR patients and 2040 represented non-HR. The resolution of all images was 1125 × 1264 and they were saved in JPEG format. These pictures were obtained as part of a standard procedure for diagnosing patients with hypertension. These images were all downsized to standard resolution (700 × 600) using data from three datasets. In addition, a professional ophthalmologist is involved in the process of preparing HR and non-HR datasets for ground truth evaluation. A sample of a fundus image from the datasets used in this study is shown in Figure 2.
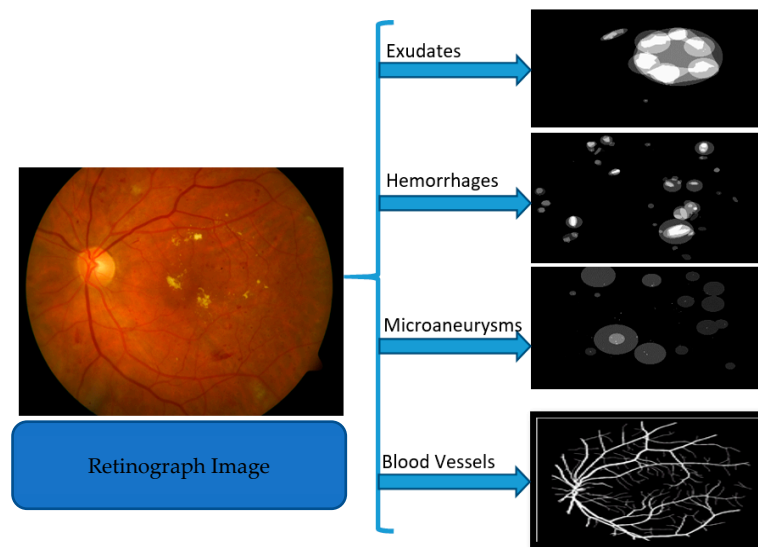
Figure 2. Example of HR fundus photo used to train the CAD-HR-diagnosis system.

3.2. Preprocessing and Data Augmentation

The goal for the initial image processing stage is to lessen the impression of disparities in lighting between photos, such as the fundus camera's brightness and angle of incidence. Using a Gaussian filter, we normalized the color balance and local illumination of each fundus picture. This was achieved in two phases: (1) color space conversion, (2) lightening and contrast correction.

Consider fundus images in RGB space with a white point chromaticity of a certain value (e.g., D65). To remove gamma or camera response nonlinearities first, RGB coordinates were linearized. Then the linearized RGB components were converted to correct colorimetric CIE-XYZ coordinates. After that, the *XYZ* axes were translated to linear *LMS* space. Equations (1) and (2) represent the transformation matrices:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.5141 & 0.3239 & 0.1604 \\ 0.2651 & 0.6702 & 0.0641 \\ 0.0241 & 0.1228 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2}$$

The data in *LMS* color space has a lot of skews, which can be reduced significantly by converting it to logarithmic space (*L*, *M*, *S*) = log (*L*, *M*, *S*) [44,48]. The lαβ color space, a visual color space based on LMS cone reactions in the human retina, is used in this paper. The logarithmic *LMS* space is then converted into lαβ using Equation (3) [44]:

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{3}$$

l: α (r + g + b), α: Yellow–blue α (r + g −b), β: Red–green α (r − g). This lαβ color space is used in the next section for fundus images color enhancement because the decorrelation process allows for disjointedly handling the three-color channels, streamlining the color enhancement process. After completing the color enhancement process, all the retina photos were transformed to RGB color space.

Most of the retina fundus images suffered from poor luminance, which significantly affected the detection of retina lesions or features and consequently affected the classifi-

cation accuracy. As a result, the luminance enhancement procedure must be completed correctly to ensure that the upgraded photos preserve the correct color information. This can be done by producing the following luminance gain matrix LG (,):

$$LG(\alpha, \beta) = \frac{r'(\alpha, \beta)}{r(\alpha, \beta)} = \frac{g'(\alpha, \beta)}{g(\alpha, \beta)} = \frac{b'(\alpha, \beta)}{b(\alpha, \beta)} \tag{4}$$

where r', g' and $b'(\alpha, \beta)$ represent the RGB values of any image-pixel, within the enhanced retinal image. As indicated, the illumination matrix is defined in Equation (5) for color invariant improvement of an RGB image.

$$LG(\alpha, \beta) = \frac{\partial'(\alpha, \beta)}{\partial(\alpha, \beta)} = \frac{\partial'(\alpha, \beta)}{\sqrt{\sum_{c \in \{r,g,b\}} c^2(\alpha, \beta)}} \tag{5}$$

where $\partial(\alpha, \beta)$ is the illumination strength of a pixel at (α, β) space and $\partial'(\alpha, \beta)$ is the boosted illumination image. The Adaptive Gamma Correction (AGC) technique is employed at this moment to increase the brightness of a given image. A quantile-based histogram equalization approach [49] was used for such cases. The q-quantiles are a set of 'q' different values that divide the intensity distribution into q equal proportions. Figure 3 illustrates the results of color space conversion and light and contrast adjustments.
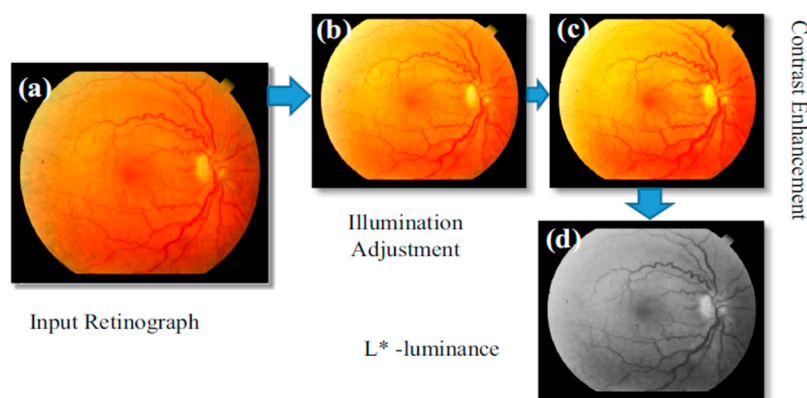

Figure 3. A visual example of the color space conversion and Light Adjustment and Contrast Enhancement, where *L** is the luminance adjusted color plane.

A substantial dataset is necessary for the CNN model to attain good performance accuracy. Moreover, the training of the CNN model using a limited dataset caused low performance due to an overfitting problem. This work applies the data augmentation approach to increase the dataset size and prevent the proposed CAD-HR system from overfitting. The augmentation strategy and parameters are shown in Table 3. Following the data augmentation, the dataset size increased to 9500 retinal samples. A visual example of the proposed data augmentation approach is shown in Figure 4.

Table 3. Details of utilized data augmentation using image processing.

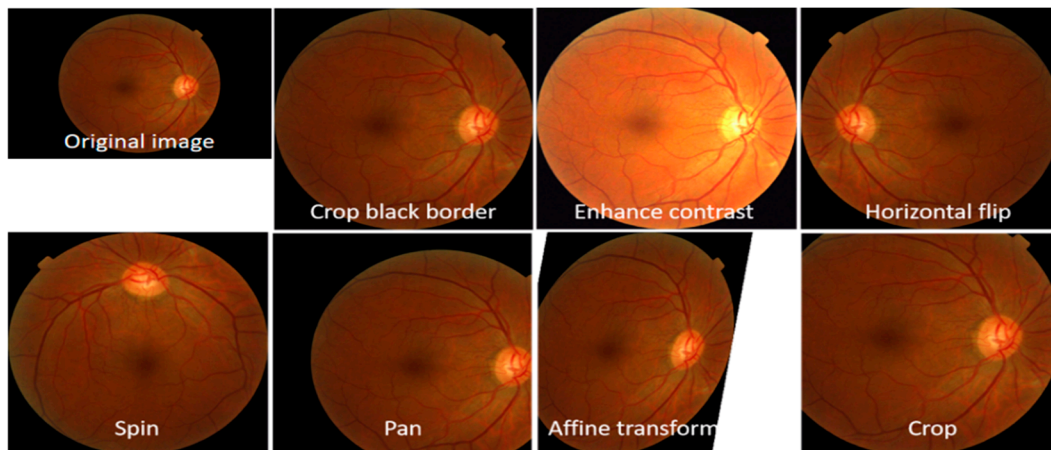| Augmentation Techniques	Values
Affine transform	True
Pan	True
Spin-range	0.2
Crop	True
Horizontal-flip	True
Vertical-flip	False
Affine transform	True

Figure 4. A visual example of the proposed data augmentation.

3.3. CNN and Transfer Learning

Recently, DL frameworks have surpassed conventional methods in various computer vision problems, such as iris recognition [66], face detection [67], classification [68], and many more. According to the findings of these studies, the CNN performs better than traditional methods. Figure 5 shows the fundamental notion of CNN. To categorize the input image into predetermined classes, the CNN uses a neural network (shown in Figure 5 as fully connected layers) that uses feature vectors to extract image features. Even though CNN has obtained remarkable results for many image-based systems, it has several problems. The two key issues associated with the CNN model are the expense in performing computations and overfitting. As CNN requires a long time to process, it is challenging to implement CNN models on a single general-purpose machine comprising fewer central processing units (CPU). Fortunately, the introduction of graphical processing units (GPU) [69] has solved this problem. By employing numerous CPUs running in parallel and GPUs, CNN can be used in real-time systems. Another issue associated with CNNs is over-fitting. The CNN, as previously discussed, is developed by learning millions of trainable parameters. Consequently, CNN-based systems often require an enormous amount of training data. Although this problem has been addressed in several techniques, including data augmentation and dropout, a massive amount of data is still used in such CNN systems for training. This problem was recently solved using transfer learning (TF) art [70]. With the TF method, we can use a CNN, trained with sufficient data for one problem to solve a different problem. This strategy has been proven useful for many problems, particularly when significant training data, such as medical images [71], are limited.
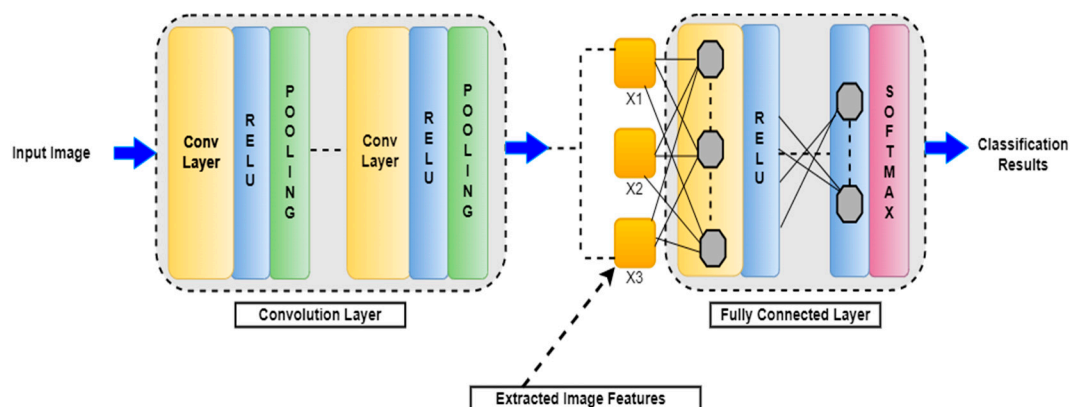

Figure 5. A visual example of the typical structure of CNN.

Since 1995, research that is based on transfer learning has seen an increase in popularity. This type of research is also well-known by the name of multitask learning [72]. Multitask learning is a method of learning numerous tasks simultaneously that is relatively equivalent to other learning schemes for the objective of knowledge transfer [73]. Figure 5 shows how a standard ML technique for various tasks differs from a transfer learning method for the same tasks. The first task in transfer learning is trained on large benchmarks, and then the second task takes advantage of the knowledge gained from the first task to learn more quickly and improve its accuracy. Based on the availability of annotated data, TF tasks are classified into mainly inductive, transductive, and unsupervised transfer learning. Usually, we employed a pre-trained deep learning model, one of the most common types of inductive TL strategy that employs source domain and job, to fine-tune different layers for the target goal. This allowed automated algorithms to get the results more precisely.

There are two ways that the target predictive function FT(.) in the target domain can be improved: first, it can be improved by combining the knowledge from the source domain DS and a task learning task LTS, and second, it can be improved by utilizing the knowledge from the target domain DT and a task learning task LTT, where DS $\neq$ DT $\neq$, or LTS $\neq$ LTT.

A comparison of the classic ML method and the TL methodology is shown in Figure 6a,b. The target task ("Target Task") in Figure 6b and knowledge (model) generated from another ML problem are the two sources from which the TL technique learns system information. In a traditional ML system, the system model is only learned for a particular task utilizing input from a single source (as shown in Figure 6). A CNN can be reused and applied to a new task using the transfer learning method. To establish the proposed CNN architecture of our work, we modified the architecture and employed the idea of the Xception model [74] for our experimental work. The training of the Xception network was completed using the ImageNet dataset. The architecture of the pre-trained model and the updated model differ significantly. The Xception model is used only in the Entry flow. Additionally, the fully connected layer was employed as the final classification layer in the pre-trained Xception model. However, to classify the two classes in our proposed modified model, we used a LSVM technique.
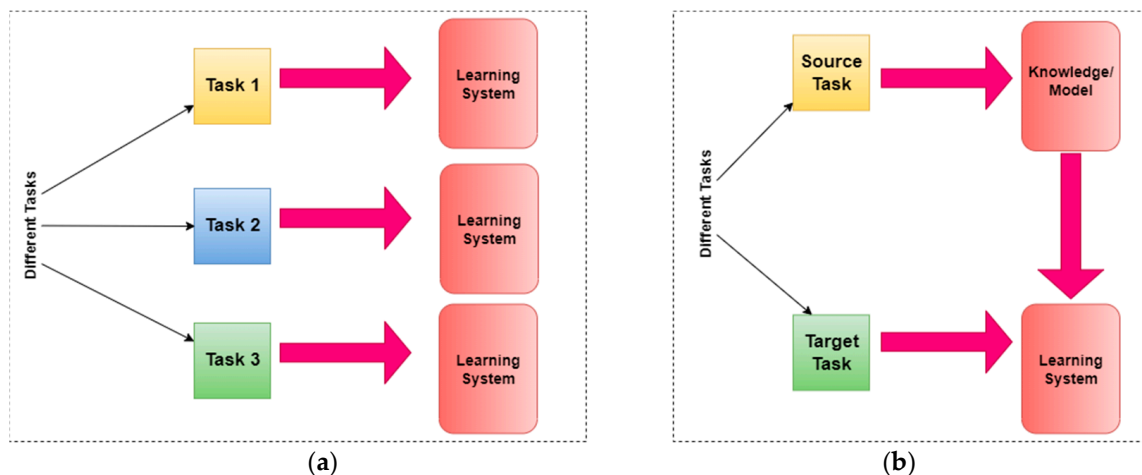

(a)　　　　　　　　　　　　　　　　　　　　　　　　**(b)**

Figure 6. An example of the difference between the traditional ML approach and the TL method [66], where Figure (**a**) shows the task and (**b**) represents the target related knowledge.

3.4. Proposed DL Architecture

A novel classification approach named *DSC*-LSVM is presented to differentiate HR-related classes, as depicted in Figure 7. To generate the feature map, this architecture uses two convolutional layers of kernel size three, three Max Pooling layers, eight Bach Normalization (BN) layers, three 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 linear activation function were used for classification.

For separable convolutional layers, the kernel size was set to three, with border padding of two and stride of two. In the pooling layers, a kernel size of three and a stride of two were also employed. Using Batch Normalization after each separable convolution layer before the activation function could improve the efficiency and stability of deep neural network training [75].
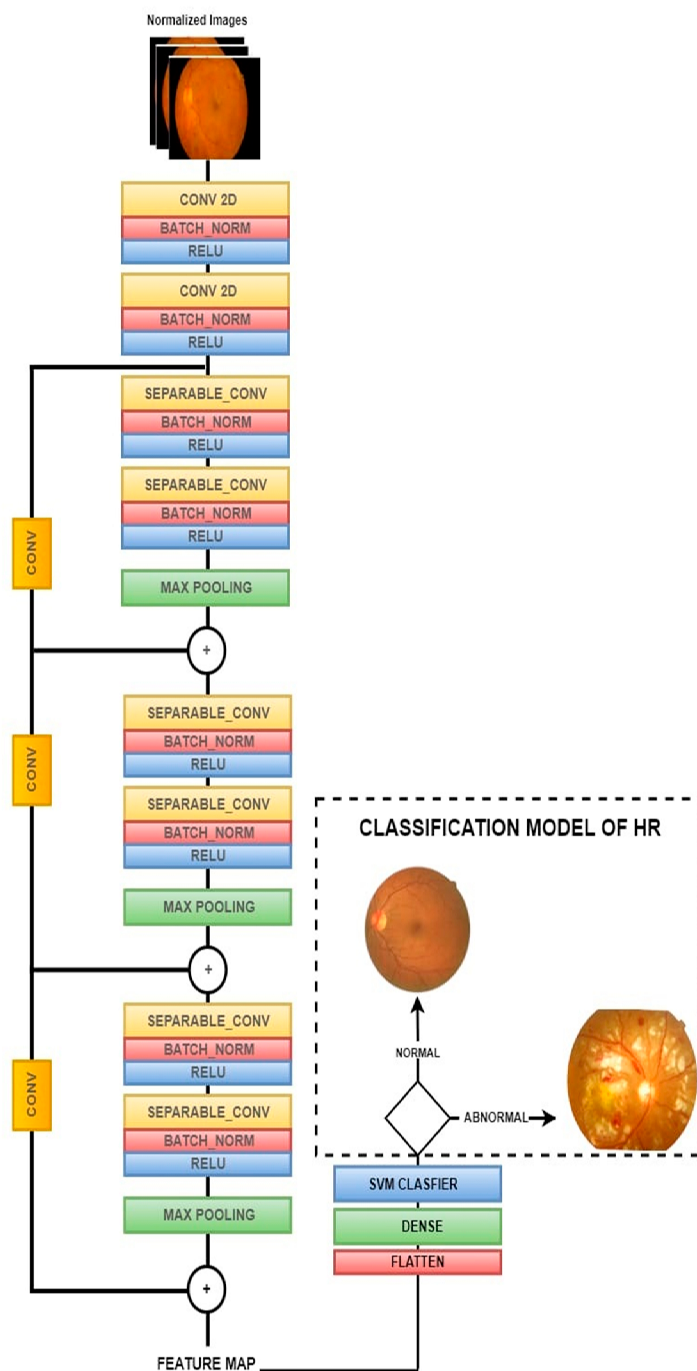
Figure 7. A visual representation of the proposed CAD-HR-diagnosis system utilizes a DSCNN, residual connections and LSVM to extract features and classify them for HR stage classification.

Additionally, the rectifier linear unit exponential linear units (RELU) function is used to define the output value of the kernel weights and convolutional layer for each separable convolutional layer. It offers a negligible value that enhances the performance of categorization. Finally, a linear SVM (LSVM) classifier was used to categorize the fundus

sample into true and false images after the feature had been flattened, densely packed, and reduced in size. Algorithm 1 provides a description of the mechanism that underpins the suggested network model for the feature extractor.

The proposed work uses the *DSC* instead of the regular convolution structure to reduce the number of calculations while keeping the CNN network's generalization. This is shown in Figure 8. The *DSC* model divides the calculation of the convolution into two stages: (1) depth-wise convolution was used to filter each channel, and (2) point-wise convolution was used as the kernel for combining. The outputs of the depth-wise convolution were then combined using a point-wise convolution. A criterion convolution generates a new output set by simultaneously filtering and merging the input.
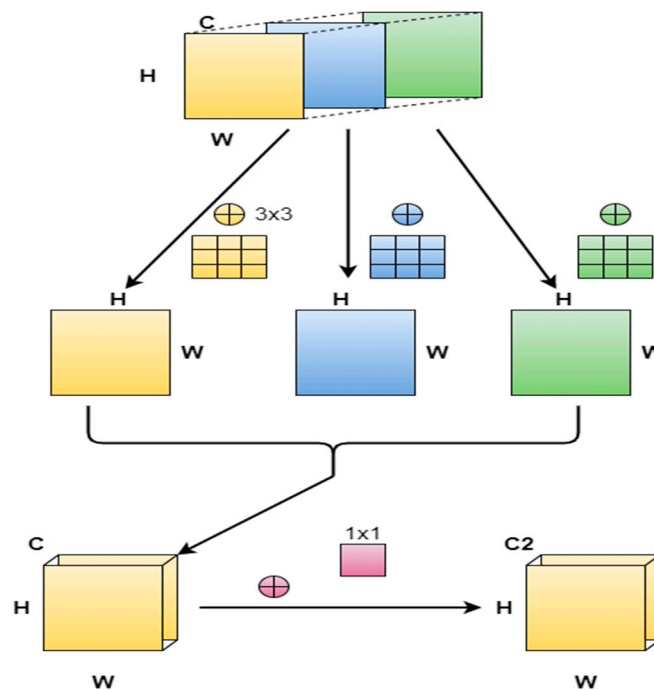

Figure 8. A visual example of Depth-wise separable CNN layer structure.

This new art reduced the computation time and parameter size. Equation (6) shows how to compute the total size of a single convolutional layer of *weights* (in bytes) [76].

$$weights = \text{C} \times \text{Nf} \times \text{Lf} \times \text{Wf} \tag{6}$$

where C represents the total number of channels, Nf indicates the total number of filters, Lf indicates the length of the filter, and Wf indicates the width of the filter. The computation of the parameters in the separable convolution layer is shown in Equation (7). By utilizing the separable convolutional layer, the number of parameters was reduced compared to the original CNN.

$$weightDSC = \text{Lf} \times \text{Wf} \times \text{CL} - 1 + \text{CL} \times 1 \times 1 \times \text{Nf} \tag{7}$$

where L indicates the convolution layer, C shows the total number of channels, Nf indicates the total number of filters, Lf represents the length of the filter, and Wf indicates the width of the filter.

The *DSC* offers two advantages for constructing a deep learning model: (1) it can help minimize parameters, and (2) it can be used to improve model generalization. Thus, it was reasonable to assume that *DSC* improved detection accuracy and training effectiveness.

The residual connection, which can also be referred to as a skipped connection, allows a network to avoid using either two or three of its levels. Figure 9 depicts the residual

single link block in the deep learning network. As can be seen in Figure 3, our suggested CNN model incorporates three residual blocks [77]. The use of residual connectivity in the deep learning model has many benefits, one of which is that it enables the preceding layers of the model network to contribute the function of the level below it to the layer above it. The last connection can be set up in a formal way.

$$O(x) = R(x) + x \tag{8}$$

| **Algorithm 1:** Implementation of the *DSC* model for feature map extraction	
Input	Array X
Output	Extraction of feature map x = (x_1, x_2 ... , x_n)
Process	
Step 1	Input raw data normalization.
Step 2	Definition of functions.

Step 3	The inputs to the Conv-batch Norm block is array X, which contains several filters, and kernel sizes.
	    a. X = Conv (X) and
	    b. X = BN (X) is then applied.

Step 4	Separable Conv2D was utilized instead of Conv2D.

Step 5	Constructing *DSC* network
	    a. The process begins with two Conv layers, each containing 32 and 64 filters. Subsequent activation of the ReLU occurs after each of them.
	    b. After that, add is utilized to use Skip Connection.
	    c. There were three different skip connections used. Following the Maxpool layer in each Skip Connection are two Separable Conv levels. The conversion factor for the skip connection is 1 to 1, and it has two strides.

Step 6	Following that, the feature map x = (x_1, x_2,..., x_n) was formed and flattened with the help of the flattened layer.
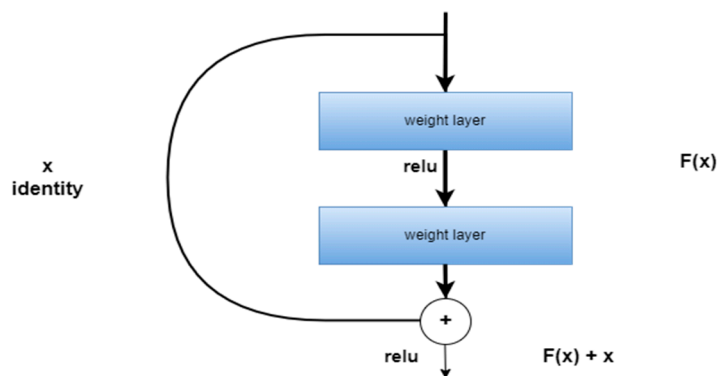
Figure 9. A visual example of residual block employed to develop the *DSC*-LSVM model.

The function $O(x)$ shows the real output value, and the residual layer learning is indicated by $R(x)$ in the network input x.

SVM is a ML classification technique that yields superior results relative to other classifier types and is regarded as an effective classifier for solving various practical situations [78]. In computer vision or image classification tasks, authors suggested a depthwise separable CNN [79] instead of deep learning or machine learning classifier. For this work, we decided to employ the Linear SVM classifier because of its ability to deal with tiny datasets and perform well in high-dimensional environments [80]. Using linear SVM was a logical choice, given that we were working with binary classification tasks. A further purpose for employing linear SVM was to improve the efficacy of our method and to identify the optimal hyperplane that divides the feature space of diseased and normal

retinal images [63]. Typically, an LSVM accepts a vector x = $(x_1, x_2..., x_n)$ and returns a value y ϵ Rn, which can be described as:

$$Y_{out} = (\text{Wei}, \text{Xiv}) + \text{b} \tag{9}$$

In the preceding Equation (4), Wei represents the weight, and b represents the offset; both Wei and b belong to R and are acquired through training. Xiv represents the input vector and is assigned to class 1 or −1 based on whether y is greater than or less than 0.

To produce the optimal hyperplane for data separation, we must minimize:

$$Wei = \frac{1}{2} ||Wei||^2 \tag{10}$$

In Algorithm 2, the mechanism of our suggested classifier is detailed.

Algorithm 2: Proposed LSVM Classifier

Input	Extracted feature map x = $(x_1, x_2,....., x_n)$ with annotations y = 0,1, Test data X_{test}
Output	Classification of normal and abnormal samples
Process	
Step 1	Initially, the classifier and Kernel Regularizer L2 parameters are defined for optimization.

Step 2	Construction of LSVM
	     a. The training process of LSVM is completed using extracted features x = $(x_1, x_2,....., x_n)$ by our Algorithm 1.
	     b For the generation of the hyperplane, use Equation (6).

Step 3	The class label is allocated for testing samples X_{test} using the decision function of the equation below. $X_{test} = (\text{Wei}, X_{iv}) + \text{b}$

---

4. Results

To assess the effectiveness of the proposed deep learning-based CAD-HR approach, in all experimental trials, datasets were split into a training set and testing set with a 3:1 split, meaning that three-quarters of the data is used for training and the remaining one-quarter is used for testing. Three separate internet and one private source provided these retina images. To execute feature extraction and classification tasks, all 3580 photos were scaled to (700 × 600) pixels. Our CAD system is created by combining the two deep learning techniques named DSCNN with residual connection, and linear SVM. A computer with a core i7 processor, 16 GB RAM, and a 4 GB Gigabyte NIVIDA GPU was utilized to implement and program our CAD-HR system. Deep learning libraries named TensorFlow (version 2.7) and Keras with windows 10 professional 64-bit edition are installed on this PC.

Various kernel dimensions are used in order to produce feature maps from the preceding stage to create/train the CNN architecture. Because the kernel dimensions (3 × 3 or 5 × 5) are generally used in this project, the convolutional layer's weights parameters are also changed. The convolutional layers are convoluted using varied window sizes and values acquired from each feature map's excitation objective function. As a convolutional layer, a similar process was utilized to create the pooling layer. There is only one difference: to optimize the characteristics gained from the previous layer, sliding steps of two and a window size of 2 × 2 are used. This phase boosts the network's overall speed while lowering the convolutional weights. The output of this average pooling is fed into a LSVM that is fully connected. This FC stage is employed for distinguishing among HR and not-HR situations. Table 4 illustrates the number of parameters in the convolution layer of the proposed CAD-HR model.

Table 4. Parametric configuration of the convolutional layer of proposed CAD-HR model.

Convolutional Layers	Parameters
Conv-32	$(3 \times 3 \times 3 + 1) \times 32$
Conv-32	$(3 \times 3 \times 32 + 1) \times 64$
Conv-64	$(3 \times 3 \times 64) + (1 \times 1 \times 64 + 1) \times 128$
Conv-64	$(3 \times 3 \times 128) + (1 \times 1 \times 128 + 1) \times 128$
Conv-128	$(3 \times 3 \times 128) + (1 \times 1 \times 128 + 1) \times 256$
Conv-128	$(3 \times 3 \times 256) + (1 \times 1 \times 256 + 1) \times 256$
Conv-256	$(3 \times 3 \times 256) + (1 \times 1 \times 256 + 1) \times 728$
Conv-256	$(3 \times 3 \times 728) + (1 \times 1 \times 728 + 1) \times 728$
Total	**87,488**

To assess the proposed CAD system's performance, statistical analysis was used to calculate the accuracy (ACC), specificity (SP), and sensitivity (SE) values. These metrics are utilized to evaluate the generated CAD system's performance and to compare it to previously developed systems [48]. This section provides a comprehensive discussion on the various experiments performed to assess the efficiency of the proposed CAD-HR model. These experiments are categorically demonstrated in the subsequent paragraphs.

Experiment 1: A cross validation testing scheme of 10 fold is utilized in the experiment 1 for comparing the resulted AUC metric to other DL methods. Overall, the AUC metric was primarily utilized to judge the categorization accuracy. The developed CAD system's performance has been quantitatively measured as shown in Table 5. The developed system achieves very good outcomes with regards to SE (94%) and SP (96%), as well as a lower training error (0.76) in identification of HR eye illness.

Table 5. Performance evaluation metrics of the implemented CAD-HR system on 9500 retinal samples.

Hypertensive Type	[1] SE	[2] SP	[3] ACC	[5] AUC	[4] E
Diabetic Hypertension (HR)	93%	96%	94%	0.95	0.76
Non-hypertension (non-HR)	95%	96%	95%	0.96	0.67
Average Result	**94%**	**96%**	**95%**	**0.96**	**0.60**

[1] SE: Sensitivity, [2] SP: Specificity, [3] ACC: Accuracy, [4] E: Training errors and [5] AUC: Area under the receiver operating curve.

Experiment 2: Several studies were carried out in experiment 2 to illustrate the efficacy of the proposed CAD system in detecting HR from input retina images. The optimum network design for retinal images HR categorization tasks was determined through experimentation. The developed CAD system shows the efficacy of the feature learned based on the optimum network design. Figure 10 shows a visual example of classification results for various input fundi. Figure 10a shows HR-classified retinography samples with various signs such as cotton wool patches and hemorrhages. A non-HR fundus picture is also shown in Figure 10b.

Experiment 3: On the retinal fundus datasets with pre-processing, the training phase is accomplished in various CNN and DRL architectures with eight to sixteen stages to perform comparisons. Table 5 summarizes the findings. It is worth noting that all the CNN and RNN deep-learning models used the same number of epochs in their training. With a validation accuracy of 59%, the highest performing network is chosen, and identical traditional Convolutional Networks were trained. Sensitivity, specificity, accuracy, and AUC metrics were utilized to compare the recital of traditional CNN, DRL, trained-CNN and trained-DRL models to the developed CAD system, as shown in Table 6.
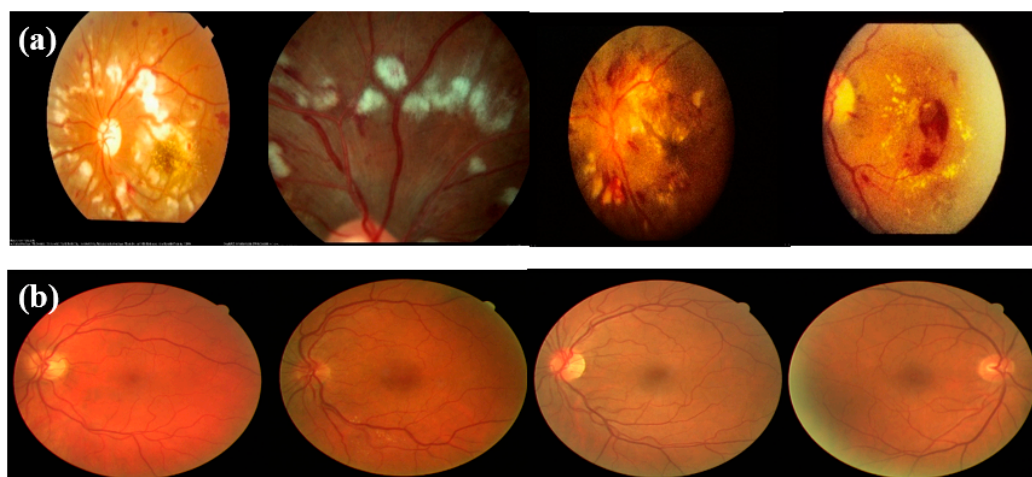
Figure 10. Output example of the CAD system (**a**) signifies HR instances, and (**b**) illustrates non-HR retinography instances.

Table 6. Assessment of the efficiency of the developed CAD-HR system vs other deep learning systems on 9500 samples.

Methods	[1] SE%	[2] SP%	[4] AUC	[3] ACC%
CNN	78.34	80.33	0.82	79.66
DRL	80.20	83.23	0.84	82.34
Pre-train CNN	81.34	83.12	0.85	81.50
Pre-train DRL	82.33	84.50	0.86	83.50
Combine Pretrain CNN and DRL	86.40	87.40	0.87	86.50
Developed CAD-HR System	**94.0**	**96.0**	**0.96**	**95.0**

[1] SE: Sensitivity, [2] SP: Specificity, [3] ACC: Accuracy, and [4] AUC: Area under the receiver operating curve

On this dataset, the trained-CNN model performed well, with the following values: 81.5%, 83.2%, 81.5%, and 0.85 for SE, SP, ACC and AUC, respectively. Metrics' results of 82.5%, 84%, 83.5%, and 0.86 for SE, SP, ACC and AUC, respectively, were attained with the DRL model. The developed CAD-HR system obtained higher results than deep-learning models by merging the capabilities of *DSC* on four annotated fundus sets with residual connections that are not disposed to overfitting difficulties.

Experiment 4: We begin to test our proposed *DSC*-CNN model on the DRIVE and DiaRetDB0 datasets using LSVM training accuracy and validation accuracy, coupled with training loss and validation loss function. As can be seen in Figure 11a,b our proposed model exhibits great performance and obtained a training accuracy and validation accuracy of over 100% while employing a total of only ten iterations of the training and validation processes. Additionally, we were able to acquire a very low loss function for both the training data and the validation data that was below 0.1, which is evidence that our proposed strategy is effective.

To adequately evaluate classification performance, we must first collect the confusion matrix. The discordance between the predicted and real labels was revealed by the confusion matrices. The column labels showed the anticipated label for each column, while the row labels showed the genuine label for each row. Consequently, based on the test set identified by our model, we were able to acquire the results of two categories, namely, HR and non-HR samples. Using 110 samples from the DRIVE and DiaRetDB0 datasets, we evaluated the performance of our model. The confusion matrix in Figure 11b makes it abundantly evident that our suggested model successfully identified every HR image utilized in the DRIVE and DiaRetDB0 datasets. Even with the limited dataset training sample for the model, it might be discovered that the p label for each category does not become

muddled. Both categories have been accurately categorized. Consequently, the confusion matrix demonstrated that our suggested model, CAD-HR, has a greater detection accuracy.
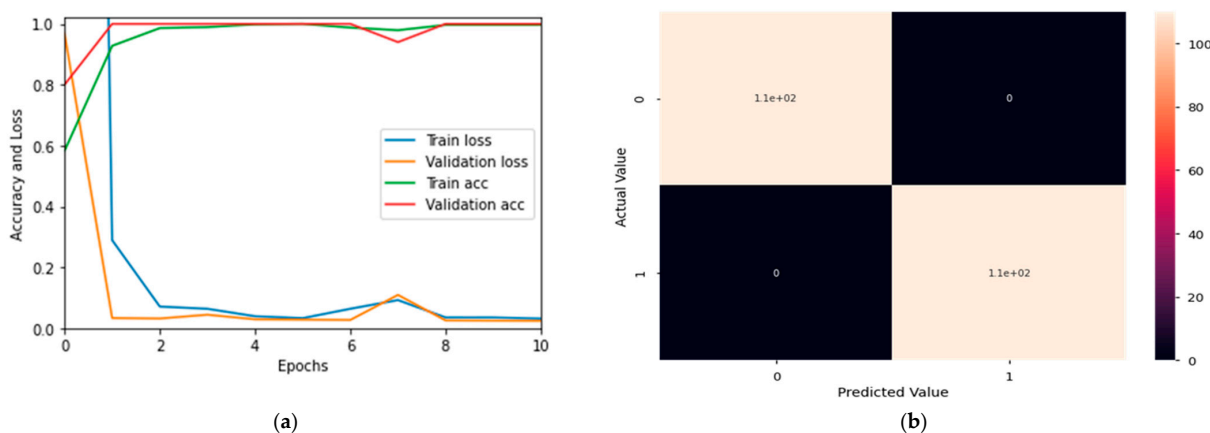

(**a**)　　　　　　　　　　　　　　　　　　(**b**)

Figure 11. (**a**) A visual representation of accuracy and loss on training and validation on DRIVE and (**b**) Confusion matrix of implemented CAD-HR on DRIVE.

Experiment 5: For determining how well our proposed CAD-HR system works, we have utilized a different dataset called Imam-HR in this experiment. First, we performed an analysis of the model's accuracy during training and validation, as well as the loss function, using both the training data and the validation data. The training and validation accuracy of the CAD-HR model employing Imam-HR is shown in Figure 12. As demonstrated in Figure 12, our model performs well in both training and validation. We achieved 100% accuracy on training and validation data, demonstrating that our technique functions well on the retina Imam-HR dataset images.
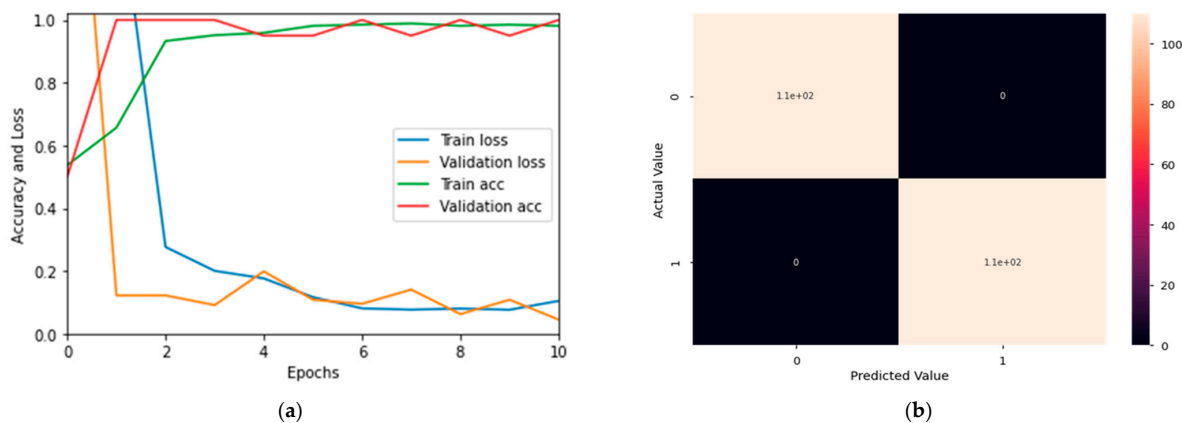

(**a**)　　　　　　　　　　　　　　　　　　(**b**)

Figure 12. (**a**) A visual representation of accuracy and loss on training and validation on DiaRetDB0 and (**b**) Confusion matrix of implemented CAD-HR on DiaRetDB0.

Next, in order to comprehend our CAD-HR algorithm's performance better. A confusion matrix that reflects the classification outcome of the suggested CAD-HR model has been obtained. It was found that our suggested approach successfully detected all the HR samples of Imam-HR, as shown in Figure 13. This indicates that every sample is accurately categorized based on the projected value. As a result, it demonstrates that our technique also has impressive Imam-HR detection accuracy.
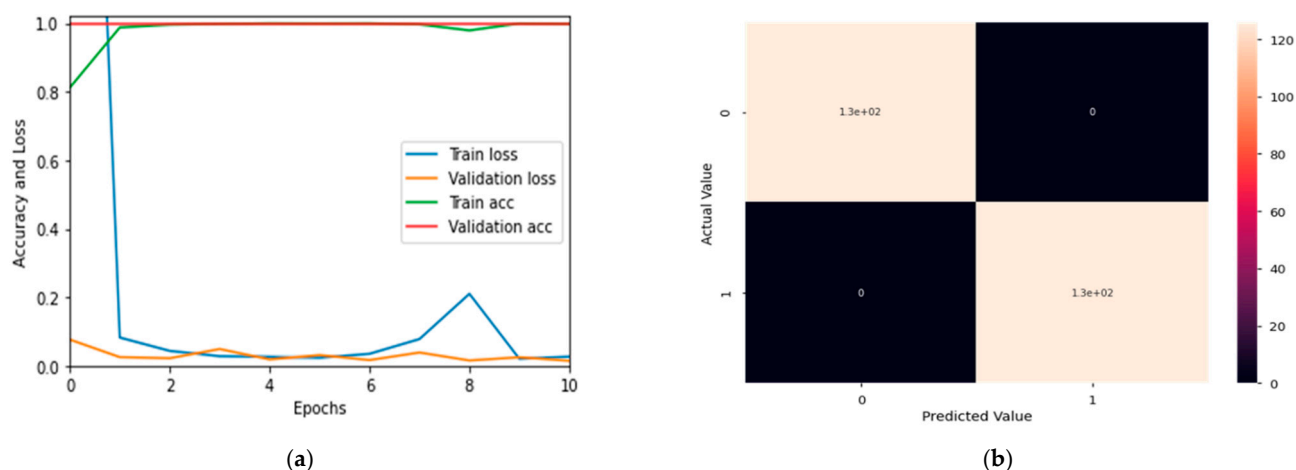
(a) (b)

Figure 13. (**a**): A visual representation of accuracy and loss on training and validation on Imam-HR and (**b**) Confusion matrix of implemented CAD-HR on Imam-HR.

4.1. State-of-the-Art Comparisons

Few research attempts used deep learning methodologies for identifying HR among retinal pictures. In this research project, two deep learning models such as Triwijoyo-2017 [7] and Pradipto-2017 [33] were selected to compare to the results of our developed system. Table 7 shows how the proposed CAD system compares to other deep learning approaches (Triwijoyo-2017 [7] and Pradipto-2017 [33]).

Table 7. Performance of the developed CAD-HR system compared to other DL systems on 9500 samples.

Methods	[1] SE	[2] SP	[3] ACC	[4] AUC
Triwijoyo-2017 [7]	78.5%	81.5%	80%	0.84
Pradipto-2017 [33]	81.5%	82.5%	84%	0.86
Proposed CAD-HR model	**94%**	**96%**	**95%**	**0.96**

[1] SE: Sensitivity, [2] SP: Specificity, [3] ACC: Accuracy, and [4] AUC: Area under the receiver operating curve

The Triwijoyo-2017 system [7] was trained using (32×32) scaled gray-level batches translated out of a big fundus set of 9500 retinal pictures, 5000 of which were non-HR and 4500 of which had HR signs. The CNN produced either HR or non-HR decisions. We implemented a similar model to that in [7]. Table 7 presents the results of several DL approaches on this 3580-image collection. The performance of the Triwijoyo-CNN-2017 and our systems were quantitively compared with regards to of SE, SP, ACC, and AUC indicators. Considering 3580 retinal pictures, the Triwijoyo-2017 system achieved 78.5%, 81.5%, 80%, and 0.84 for those metrics, respectively. The CNN architecture proposed in the Triwijoyo-CNN-2017 system fails to exert operative and broad features that can be used to distinguish between HR and non-HR.

Consequently, in comparison, the developed CAD-HR system produced better results, with of 94%, 96%, 95%, and 0.96 for SE, SP, ACC, and AUC, respectively. The authors of this research project cited the identification accuracy of 98.6% in Triwijoyo-2017 [7]. But we notice that they used a very restricted set of input fundi for training, as they only contain 40 retina photos, 20 of which are normal and 20 of which are HR which lead to the high level of precision. Accordingly, our CAD-HR system was tested and trained on a huge dataset. Therefore, we attained a classification accuracy of 95%.

The CNN-RBM model was only used as a classifier in the Pradipto-2017 [33] system. For subdivision and excavation of features, the system engages image processing algorithms. The classification stage input was a feature vector consisting of A/V Ratio and OD, more willingly than the retinal images. We used the same processes and tested them on our

dataset of 3580 photos to compare with Pradipto 2017 [33]. Table 7 shows that our system significantly outperformed the Pradipto-2017 model too. In this paper, architecture composed of one DL method called *DSC* with LSVM was constructed by means of a *DSC*-LSVM technique. Therefore, we were able to achieve greater accuracy against recent HR detection approaches. One more stage was integrated to the architecture in the improved *DSC* model, three residual connections, which contained localized and specialized features, respectively. In comparison to features recovered by the Triwijoyo-CNN-2017 [7] and Pradipto-2017 [33] models, these *DSC* with three residual attributes were more generic. We updated residual connections with skip paths to resolve the issue of extracting HR-related grazes using a fresh training technique rather than employing pre-train architectures to create features.

Moreover, we have compared it with three other transfer learning-based (TL) CNN architectures, such as VGG-16, VGG-19, ResNet-50, Inception-v3 and AlexNet. To assess the performance, the preprocessing and data augmentation techniques were first applied to these TL architectures. We employed data augmentation to get around the issue of most CNN architectures needing a lot of labeled data for training. To perform comparisons with TL algorithms, we used, by default, hyper-parameters with 200 epochs. A visual example of the performance of TL algorithms is displayed in Figure 14. The obtained results demonstrate that our CAD-HR system is outperformed when compared to other state-of-the-art TL-based CNN architectures. On the other hand, in our CAD-HR approach, we intend to deploy more dense separable CNN architectures. Furthermore, merging the deep features of many architectures is an intriguing strategy that can boost performance. The performance has significantly improved when comparing the ResNet50 to other TL models such as VGG-16, VGG-19, Inception-v3 and AlexNet. Therefore, we suggest using ResNet50 to classify HR eye-related diseases. Furthermore, we draw the conclusion that even the ResNet50 model cannot match the performance of the proposed CAD-HR model, which is trained with the optimal hyper-parameter setup. Even though TL-based approaches are often employed to improve classification task accuracy, they also significantly increase the architectural complexity of the model and might not make a meaningful difference in how well deep learning models perform when configured with the best hyper-parameters.

4.2. Computational Cost

The proposed image preprocessing phase used to adjust the lightness and remove noise from the input image took 25 s. Similarly, the feature learning and extraction step by the proposed CAD-HR system took 16.5 s on average, while the LSVM classifier took only 3 s to be created. Thus, the LSVM training to perform binary classification of hypertensive retinopathy into HR and non-HR took 5.20 s on a fixed number of 30 iterations. However, once the training is complete and the test is run, classifying the image only takes an average of 8.12 s. The proposed CAD-HR took 2.1 s higher to compute than [7,33] utilizing the convolutional neural networks (CNN) model. This is because we have implemented a novel technique by employing *DSC*, three residual blocks, and LSVM in a perceptually oriented color space.

This computational time complexity is generalized in terms of the training of the network. In the development of the CAD-HR system, there are four blocks with residual connections that can be represented as i, j, k, and l, with t training examples and n epochs. The result was $O(n \times t(i + j + k + l))$. This time complexity can be reduced by using tensor processing units (TPUs), which are provided by the Google cloud. In practice, the TPUs achieved substantial speedups for DL models and utilized less power. This point of view will be addressed in future work as well.
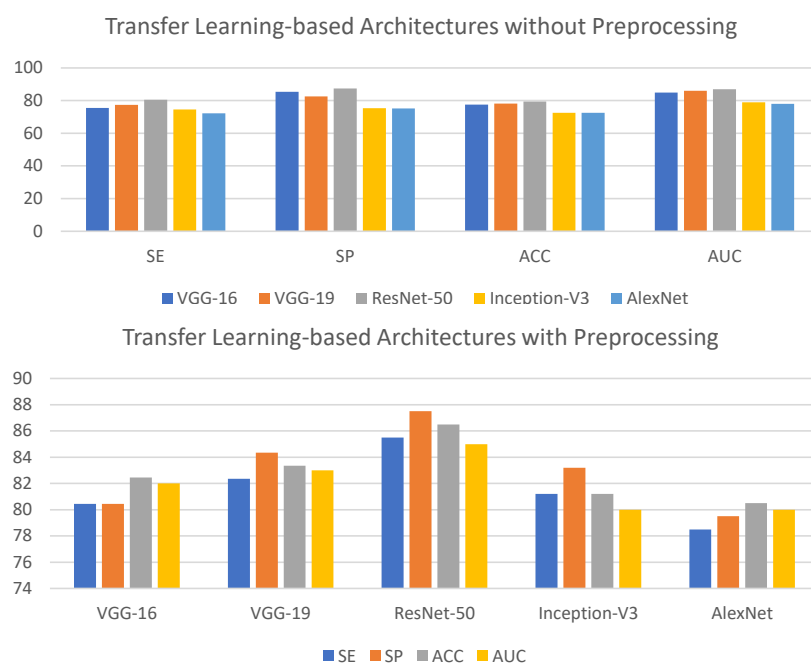
Figure 14. Comparison with other transfer learning-based (TL) CNN architectures for classification of HR based on data augmentation and preprocessing steps.

5. Discussion

CAD-HR was implemented by utilizing a trained CNN model as an input to *DSC* architectures to classify HR using three consecutive residual blocks and a linear SVM. The learning process of specialized features was completed by introducing a multilayered hierarchical framework without utilizing complex image processing and feature selection methods. This multi-layer architecture learnt features directly from the input image with the learning methods, eliminating the need for human intervention. The Xception model is modified to include *DSC* and residual blocks to create more generalizable features to develop CAD-HR architecture. Based on a scratch-based training technique, the depth-wise convolutional layer obtains localized and trained features from four HR-related lesions. The CNN model for learning deep features is made up mostly of convolutional, pooling, and fully connected layers. To construct this model, those layers must be trained and found to be effective in extracting useful information. These features are not optimal for recognizing HR in retinography pictures. As a result, deep residual connections were integrated, which provided highly specialized features rather than feature-based categorization methods that required human intervention.

The success in detecting HR was made possible by an autonomous features learning method. For the diagnosis of HR disease, however, the hand-crafted based classification techniques rely on the pre-processing, segmentation, and localization of HR-related data, which are computationally expensive algorithms. Compared to other crucial symptoms like cotton wool spots or hemorrhage recognition, researchers only exerted considerable attention to extract the necessary elements.

In the past, a few classification systems for HR-related and non-HR-related eye diseases were developed, as described in Section 2. These systems use deep-learning methods instead of traditional machine learning arts. No datasets with clinical expert annotations defining these HR-related lesion patterns are available for training and evaluating the network. Therefore, it is challenging for computerized systems to identify these illness traits. According to the available literature, the authors used manually constructed features to train the network and assess the performance of standard and cutting-edge deep learning models. Consequently, an automatic technique is necessary to identify optimal features. Compared to the conventional approach, deep-learning models produce superior outcomes. However,

other models utilized trained models built from scratch to automatically learn features, but they all shared the same weighting technique at each level. Subsequently, it may be difficult for layers to transfer accurate decision-making weights to deeper network levels.

The CAD-HR system is created in this study to overcome the problems by classifying images into HR and non-HR using two multi-layer deep learning techniques rather than concentrating on image processing algorithms. The CAD-HR system's significant contributions are listed here. This research developed two new deep learning arts using a convolutional neural network (CNN) and residual blocks. Four distinct HR lesions were used to train the first CNN model, which was then used to determine the hierarchy of features. To enhance the efficiency of the learning procedure, the second residual blocks were used to identify the feature maps with the most useful information. It is the first system built in this study for the purpose of classifying HR, and it is based on perceptually oriented color space. The deep features are classified utilizing the *DSC* model and LSVM classifier. According to our knowledge, this is the first attempt to identify HR illness automatically. To construct the CAD-HR system described in this paper, the multilayer deep learning network must be trained on many samples to provide greater feature generalization. Using a DSCNN with three residual blocks, a new deep learning-based technique was created to learn features automatically. However, the proposed CAD-HR system misclassifies a small number of samples. In Figure 13, a visual illustration is shown. It was an instance of hypertensive retinopathy (HR) with a severe level, and we will address this problem in future research.

For HR recognition, Tables 6 and 7 indicate that the CAD-HR system obtained higher accuracy than the [7] and the [33] systems. This is because the CAD-HR system is built with trained features employing DSCNN architecture and deep residual learning techniques. Furthermore, the DRL design was modified by the addition of three residual blocks that extract localized and specialized information. The authors updated deep residual network blocks with three shortcuts to solve the challenge of recovering HR-related lesions through the from scratch training technique rather than employing off-the-shelf pre-trained models to define specialized features.

The CAD-HR system for HR recognition can be enhanced in the future by offering a larger collection of retinography images that will be gathered from various sources. Instead of merely employing deep features, it would be possible to include hand-crafted features to improve the model's classification accuracy. Since several research groups used the saliency maps technique for segmenting DR-related lesions, those lesions were then extracted from retinography pictures using a train classifier. Only the segmentation step was done in those studies. Future integration of these saliency maps will improve HR eye-related disease classification precision. Furthermore, the various HR severity levels will be evaluated in the future. Much recent research has found that clinical characteristics are crucial indicators for determining the severity level of HR. The extraction of those HR-related lesions with varied thresholds, on the other hand, will be used to detect the disease level of HR. As a result, clinicians may find it advantageous to use them to tackle the problem of hypertension.

We have developed this CAD-HR system to recognize only two classes such as HR and normal. However, the CAD-HR system has not been tested on the five-stages of the HR system. In addition, the optimization of hyper-parameters is required to fine-tune this deep-learning model. Apart from the time complexity, the overall computational time can be decreased by adding a block-based fine-tuning strategy. Also, the preprocessing step can be enhanced in terms of utilizing different color space models. The ConvMixer architecture should also be tested against the proposed model, which should be one of the future works.

6. Conclusions

In this research, a unique multi-layer deep residual learning (DRL) with features training system CAD-HR is designed to detect HR to address these issues. The CAD-HR system is made up of a multilayer architecture comprising DSCNN-trained features and deep residual learning blocks that extract features from retinal fundus pictures and identify

them as HR- or non-HR related disease. Furthermore, the CAD-HR system learns and categorizes features automatically by modifying the DRL network architecture via deep residual connections. This research introduces three residual connections to determine optimal and deep features for the categorization of HR eye-related disease. The CAD-HR system can identify HR, and it will help the ophthalmologist to take an accurate decision. It also helps with screening large groups of people. The results of the experiments show that the CAD-HR system can be used to accurately diagnose HR. In future works, we will use this CAD-HR system to recognize five classes of the HR system. In addition, we would like to reduce or optimize hyper-parameters to fine-tune this deep-learning model.

Author Contributions: Conceptualization, I.Q.; Q.A.; A.R.B. and K.S.; methodology, Q.A.; I.Q.; K.S. and A.H.; software, I.Q.; validation, Q.A.; A.R.B.; J.Y.; and I.Q.; resources, Q.A.; K.S. and I.Q.; data curation, K.S.; I.Q.; and Q.A.; writing—original draft preparation, Q.A.; I.Q.; J.Y.; and I.Q.; writing— review and editing, Q.A.; J.Y. and I.Q.; visualization, Q.A. and I.Q.; supervision, I.Q.; funding acquisition, Q.A.; K.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mozaffarian, D.; Benjamin, E.J.; Go, A.S.; Arnett, D.K.; Blaha, M.J.; Cushman, M.; Das, S.R.; de Ferranti, S.; Després, J.-P.; Fullerton, H.J. Executive Summary: Heart Disease and Stroke Statistics-2016 Update: A Report from the American Heart Association. *Circulation* **2016**, *133*, 447–454. [CrossRef] [PubMed]
2. Rosendorff, C.; Lackland, D.T.; Allison, M.; Aronow, W.S.; Black, H.R.; Blumenthal, R.S.; Gersh, B.J. Treatment of hypertension in patients with coronary artery disease: A scientific statement from the American Heart Association, American College of Cardiology, and American Society of Hypertension. *J. Am. Coll. Cardiol.* **2015**, *65*, 1998–2038. [CrossRef] [PubMed]
3. Raghavendra, U.; Fujita, H.; Bhandary, S.V.; Gudigar, A.; Tan, J.H.; Acharya, U.R. Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Inf. Sci.* **2018**, *441*, 41–49. [CrossRef]
4. Akagi, S.; Matsubara, H.; Nakamura, K.; Ito, H. Modern treatment to reduce pulmonary arterial pressure in pulmonary arterial hypertension. *J. Cardiol.* **2018**, *72*, 466–472. [CrossRef]
5. Gamella-Pozuelo, L.; Fuentes-Calvo, I.; Gomez-Marcos, M.A.; Recio-Rodriguez, J.I.; Agudo-Conde, C.; Fernández-Martín, J.L.; Martínez-Salgado, C. Plasma cardiotrophin-1 as a marker of hypertension and diabetes-induced target organ damage and cardiovascular risk. *Medicine* **2015**, *94*, 30. [CrossRef]
6. Suryani, E. The review of computer aided diagnostic hypertensive retinopathy based on the retinal image processing. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *620*, 012099.
7. Triwijoyo, B.K.; Budiharto, W.; Abdurachman, E. The Classification of Hypertensive Retinopathy using Convolutional Neural Network. *Procedia Comput. Sci.* **2017**, *116*, 166–173. [CrossRef]
8. García-Floriano, A.; Ferreira-Santiago, Á.; Nieto, O.C.; Yáñez-Márquez, C. A machine learning approach to medical image classification: Detecting age-related macular degeneration in fundus images. *Comput. Electr. Eng.* **2017**, *75*, 218–229. [CrossRef]
9. Asiri, N.; Hussain, M.; Aboalsamh, H.A. Deep Learning based Computer-Aided Diagnosis Systems for Diabetic Retinopathy: A Survey. *Artif. Intell. Med.* **2018**, *99*, 101701. [CrossRef]
10. Abbas, Q.; Ibrahim, M.E.; Jaffar, M.A. A comprehensive review of recent advances on deep vision systems. *Artif. Intell. Rev.* **2018**, *52*, 39–76. [CrossRef]
11. Abbas, Q.; Celebi, M.E. DermoDeep-A classification of melanoma-nevus skin lesions using multi-feature fusion of visual features and deep neural network. *Multimed. Tools Appl.* **2019**, *78*, 23559–23580. [CrossRef]
12. Sengupta, S.; Singh, A.; Leopold, H.A.; Gulati, T.; Lakshminarayanan, V. Ophthalmic diagnosis using deep learning with fundus images—A critical review. *Artif. Intell. Med.* **2020**, *102*, 101758. [CrossRef]
13. Akbar, S.; Akram, M.U.; Sharif, M.; Tariq, A.; Yasin, U. Arteriovenous ratio and papilledema based hybrid decision support system for detection and grading of hypertensive retinopathy. *Comput. Methods Programs Biomed.* **2018**, *154*, 123–141. [CrossRef] [PubMed]

14. Abbasi-Sureshjani, S.; Smit-Ockeloen, I.; Bekkers, E.J.; Dashtbozorg, B.; Romeny, B.M. Automatic detection of vascular bifurcations and crossings in retinal images using orientation scores. In Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016; pp. 189–192.

15. Abbas, Q.; Qureshi, I.; Yan, J.; Shaheed, K. Machine Learning Methods for Diagnosis of Eye-Related Diseases: A Systematic Review Study Based on Ophthalmic Imaging Modalities. *Arch. Comput. Methods Eng.* **2022**, *29*, 3861–3918. [CrossRef]

16. Akbar, S.; Akram, M.U.; Sharif, M.; Tariq, A.; Khan, S.A. Decision support system for detection of hypertensive retinopathy using arteriovenous ratio. *Artif. Intell. Med.* **2018**, *90*, 15–24. [CrossRef] [PubMed]

17. Grisan, E.; Foracchia, M.; Ruggeri, A. A Novel Method for the Automatic Grading of Retinal Vessel Tortuosity. *IEEE Trans. Med. Imaging* **2008**, *27*, 310–319. [CrossRef]

18. Holm, S.I.; Russell, G.; Nourrit, V.; McLoughlin, N.P. DR HAGIS—A fundus image database for the automatic extraction of retinal surface vessels from diabetic patients. *J. Med. Imaging* **2017**, *4*, 014503. [CrossRef]

19. Tramontan, L.; Ruggeri, A. Computer estimation of the AVR parameter in diabetic retinopathy. In Proceedings of the World Congress on Medical Physics and Biomedical Engineering, Munich, Germany, 7–12 September 2009; Dössel, O., Schlegel, W.C., Eds.; IFMBE Proceedings, 25 11. Springer: Berlin/Heidelberg, Germany, 2009.

20. Goswami, S.; Goswami, S.; De, S. Automatic Measurement and Analysis of Vessel Width in Retinal Fundus Image. In Proceedings of the Springer 1st International Conference on Intelligent Computing and Communication, Singapore, 21 November 2016; pp. 451–458.

21. Ortiz, D.; Cubides, M.; Suarez, A.; Zequera, M.L.; Quiroga, J.; Gómez, J.L.; Arroyo, N. Support system for the preventive diagnosis of Hypertensive Retinopathy. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, 31 August–4 September 2010; pp. 5649–5652.

22. Muramatsu, C.; Hatanaka, Y.; Iwase, T.; Hara, T.; Fujita, H. Automated selection of major arteries and veins for measurement of arteriolar-to-venular diameter ratio on retinal fundus images. *Comput. Med. Imaging Graph. Off. J. Comput. Med. Imaging Soc.* **2011**, *35*, 472–480. [CrossRef]

23. 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 Proceedings of the 2011 E-Health and Bioengineering Conference (EHB), Iasi, Romania, 24–26 November 2011; pp. 1–6.

24. Saez, M.; González-Vázquez, S.; Penedo, M.G.; Barceló, M.A.; Pena-Seijo, M.; Tuero, G.C.; Pose-Reino, A. Development of an automated system to classify retinal vessels into arteries and veins. *Comput. Methods Programs Biomed.* **2012**, *108*, 367–376. [CrossRef]

25. Narasimhan, K.; Neha, V.C.; Vijayarekha, K. Hypertensive Retinopathy Diagnosis from Fundus Images by Estimation of AVR. *Procedia Eng.* **2012**, *38*, 980–993. [CrossRef]

26. Noronha, K.; Navya, K.T.; Nayak, K.P. Support System for the Automated Detection of Hypertensive Retinopathy using Fundus Images. *IJCA Spec. Issue Int. Conf. Electron. Des. Signal Process. ICEDSP* **2013**, *1*, 7–11.

27. Nath, M.; Dandapat, S. Detection of changes in color fundus images due to diabetic retinopathy. In Proceedings of the 2012 2nd National Conference on Computational Intelligence and Signal Processing (CISP), Guwahati, India, 2 March 2012; pp. 81–85.

28. Agurto, C.; Joshi, V.; Nemeth, S.C.; Soliz, P.; Barriga, E.S. Detection of hypertensive retinopathy using vessel measurements and textural features. In Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 26–30 August 2014; pp. 5406–5409.

29. Khitran, S.A.; Akram, M.U.; Usman, A.; Yasin, U. Automated system for the detection of hypertensive retinopathy. In Proceedings of the 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), Paris, France, 14–17 October 2014; pp. 1–6.

30. Irshad, S.; Akram, M.U.; Salman, M.S.; Yasin, U. Automated detection of Cotton Wool Spots for the diagnosis of Hypertensive Retinopathy. In Proceedings of the 2014 Cairo International Biomedical Engineering Conference (CIBEC), Giza, Egypt, 11–13 December 2014; pp. 121–124.

31. Irshad, S.; Akram, M.U. Classification of retinal vessels into arteries and veins for detection of hypertensive retinopathy. In Proceedings of the 2014 Cairo International Biomedical Engineering Conference (CIBEC), Giza, Egypt, 11–13 December 2014; pp. 133–136.

32. 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 Res. Int.* **2015**, *2015*, 752957. [CrossRef] [PubMed]

33. Triwijoyo, B.K.; Pradipto, Y.D. Detection of Hypertension Retinopathy Using Deep Learning and Boltzmann Machines. *J. Phys. Conf. Ser.* **2017**, *801*, 012039. [CrossRef]

34. Syahputra, M.F.; Amalia, C.; Rahmat, R.F.; Abdullah, D.; Napitupulu, D.; Setiawan, M.I.; Albra, W.; Nurdin; Andayani, U. Hypertensive retinopathy identification through retinal fundus image using backpropagation neural network. *J. Phys. Conf. Ser.* **2018**, *978*, 012106. [CrossRef]

35. Zhu, C.; Zou, B.; Zhao, R.; Cui, J.; Duan, X.; Chen, Z.; Liang, Y. Retinal vessel segmentation in colour fundus images using Extreme Learning Machine. *Comput. Med. Imaging Graph.* **2017**, *55*, 68–77. [CrossRef] [PubMed]

36. 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. *J. Comput. Sci.* **2017**, *20*, 70–79. [CrossRef]

37. AlBadawi, S.; Fraz, F.F. Arterioles and Venules Classification in Retinal Images Using Fully Convolutional Deep Neural Network. In Proceedings of the 15th International Conference on Image Analysis and Recognition (ICIAR'18), Póvoa de Varzim, Portugal, 27–29 June 2018; pp. 659–668.

38. Welikala, R.A.; Foster, P.J.; Whincup, P.; Rudnicka, A.R.; Owen, C.G.; Strachan, D.P.; Barman, S. Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort. *Comput. Biol. Med.* **2017**, *90*, 23–32. [CrossRef]

39. Yao, Z.; Zhang, Z.; Xu, L. Convolutional Neural Network for Retinal Blood Vessel Segmentation. In Proceedings of the 9th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 10–11 December 2016; Volume 1, pp. 406–409.

40. Prentasic, P.; Loncaric, S. Detection of exudates in fundus photographs using convolutional neural networks. In Proceedings of the 2015 9th International Symposium on Image and Signal Processing and Analysis (ISPA), Zagreb, Croatia, 7–9 September 2015; pp. 188–192.

41. Abbas, Q.; Fondon, I.; Sarmiento, A.; Jiménez, S.; Alemany, P. Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features. *Med. Biol. Eng. Comput.* **2017**, *55*, 1959–1974. [CrossRef]

42. Abbas, Q.; Ibrahim, M.E. DenseHyper: An automatic recognition system for detection of hypertensive retinopathy using dense features transform and deep-residual learning. *Multimed. Tools Appl.* **2020**, *79*, 31595–31623. [CrossRef]

43. Abbas, Q.; Qureshi, I.; Ibrahim, M.E. An automatic detection and classification system of five stages for hypertensive retinopathy using semantic and instance segmentation in DenseNet architecture. *Sensors* **2021**, *21*, 6936. [CrossRef]

44. 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, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

45. Qureshi, I.; Ma, J.; Abbas, Q. Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. *Multimed. Tools Appl.* **2021**, *80*, 11691–11721. [CrossRef]

46. Wu, S.; Zhong, S.; Liu, Y. Deep residual learning for image steganalysis. *Multimed. Tools Appl.* **2018**, *77*, 10437–10453. [CrossRef]

47. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity mappings in deep residual networks. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherladns, 8–16 October 2016; pp. 630–645.

48. Liu, C.; Gardner, S.J.; Wen, N.; Elshaikh, M.A.; Siddiqui, F.; Movsas, B.; Chetty, I.J. Automatic segmentation of the prostate on CT images using deep neural networks (DNN). *Int. J. Radiat. Oncol. Biol. Phys.* **2019**, *104*, 924–932. [CrossRef] [PubMed]

49. Abbas, Q.; Ibrahim, M.E.A.; Jaffar, M.A. Video scene analysis: An overview and challenges on deep learning algorithms. *Multimed. Tools Appl.* **2019**, *77*, 20415–20453. [CrossRef]

50. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition; Computing Research Repository (CoRR). *arXiv* **2014**, arXiv:1409.1556.

51. Liu, S.; Deng, W. Very deep convolutional neural network based image classification using small training sample size. In Proceedings of the IEEE 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 3–6 November 2015; pp. 730–734.

52. Sahoo, A.K.; Pradhan, C.; Das, H. Performance Evaluation of Different Machine Learning Methods and Deep-Learning Based Convolutional Neural Network for Health Decision Making. In *Nature Inspired Computing for Data Science*; Springer: Berlin, Germany, 2020; pp. 201–212.

53. Zhao, M.; Kang, M.; Tang, B.; Pecht, M. Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes. *IEEE Trans. Ind. Electron.* **2017**, *65*, 4290–4300. [CrossRef]

54. 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 Gener. Comput. Syst.* **2019**, *101*, 14–28. [CrossRef]

55. Liang, G.; Hong, H.; Xie, W.; Zheng, L. Combining Convolutional Neural Network with Recursive Neural Network for Blood Cell Image Classification. *IEEE Access* **2018**, *6*, 36188–36197. [CrossRef]

56. Vaghefi, E.; Yang, S.; Hill, S.; Humphrey, G.; Walker, N.; Squirrell, D. Detection of smoking status from retinal images; a Convolutional Neural Network study. *Sci. Rep.* **2019**, *9*, 1–9. [CrossRef]

57. Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29*, 2352–2449. [CrossRef]

58. Deng, J.; Dong, W.; Socher, R.; Li, L.; Li, K.; Li, F. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.

59. Canziani, A.; Paszke, A.; Culurciello, E. An Analysis of Deep Neural Network Models for Practical Applications. *arXiv* **2017**, arXiv:abs/1605.07678.

60. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* **2018**, *2018*, 7068349. [CrossRef] [PubMed]

61. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS 2014), Montreal, QC, Canada, 8–13 December 2014; pp. 3320–3328.

62. Sun, W.; Zhang, X.; He, X. Lightweight image classifier using dilated and depthwise separable convolutions. *J. Cloud Comput. Adv. Syst. Appl.* **2020**, *9*, 55. [CrossRef]

63. Pasquet, J.; Chaumont, M.; Subsol, G.; Derras, M. Speeding-up a convolutional neural network by connecting an SVM network. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 2286–2290.

64. Staal, J.; Abràmoff, M.D.; Niemeijer, M.; Viergever, M.A.; Ginneken, B.V. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging* **2004**, *23*, 501–509. [CrossRef] [PubMed]

65. Kauppi, T.; Kalesnykiene, V.; Kamarainen, J.K.; Lensu, L.; Sorri, I.; Raninen, A.; Voutilainen, R.; Uusitalo, H.; Kalviainen, H.; Pietilä, J. The diaretdb1 diabetic retinopathy database and evaluation protocol. In Proceedings of the 17th British Machine Vision Conference (BMVC), Warwick, UK, 10–13 September 2007; Volume 1, pp. 1–10.

66. Nguyen, K.; Fookes, C.; Ross, A.; Sridharan, S. Iris recognition with off-the-shelf CNN features: A deep learning perspective. *IEEE Access* **2017**, *6*, 18848–18855. [CrossRef]

67. Mahmood, Z.; Muhammad, N.; Bibi, N.; Ali, T. A review on state-of-the-art face recognition approaches. *Fractals* **2017**, *25*, 1–19. [CrossRef]

68. Kassani, S.H.; Kassani, P.H.; Khazaeinezhad, R.; Wesolowski, M.J.; Schneider, K.A.; Deters, R. Diabetic retinopathy classification using a modified xception architecture. In Proceedings of the 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, United Arab Emirates, 10–12 December 2019; p. 5. [CrossRef]

69. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.

70. Anbarasi, A.; Ravi, S.; Vaishnavi, J.; Matla, S.V.S.B. Computer aided decision support system for mitral valve diagnosis and classification using depthwise separable convolution neural network. *Multimed. Tools Appl.* **2021**, *80*, 21409–21424. [CrossRef]

71. Stolte, S.; Fang, R. A survey on medical image analysis in diabetic retinopathy. *Med. Image Anal.* **2020**, *64*, 101742. [CrossRef]

72. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]

73. Ribani, R.; Marengoni, M. A survey of transfer learning for convolutional neural networks. In Proceedings of the 32nd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), Rio de Janeiro, Brazil, 28–31 October 2019; pp. 47–57. [CrossRef]

74. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807. [CrossRef]

75. Ioffe, S.; Christian, S. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 730–743.

76. Iandola, F.N.; Ashraf, K.; Moskewicz, M.W.; Keutzer, K. FireCaffe: Near-linear acceleration of deep neural network training on compute clusters. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1–13.

77. Mujtaba, H. What is Resnet or Residual Network | How Resnet Helps? *2021.* Available online: https://www.mygreatlearning.com/blog/resnet/ (accessed on 11 August 2022).

78. Tang, Y. Deep Learning using Linear Support Vector Machines. *arXiv* **2013**, arXiv:abs/1306.0239.

79. Lu, Y.; Jiang, M.; Wei, L.; Zhang, J.; Wang, Z.; Wei, B.; Xia, L. Automated arrhythmia classification using depthwise separable convolutional neural network with focal loss. *Biomed. Signal Process. Control* **2021**, *69*, 102843. [CrossRef]

80. Basly, H.; Ouarda, W.; Sayadi, F.E.; Ouni, B.; Alimi, A.M. *CNN-svm Learning Approach Based Human Activity Recognition*; Springer International Publishing: Cham, Switzerland, 2020. [CrossRef]