Republic of Iraq
Ministry of Higher Education and Science Research
University of Babylon
College of Science for Women
Department of Computer Science

Multi Class Detection of Cataract in the Retina Based on Convolution Neural Network

A Thesis

Submitted to the council of The Science College for Women of Babylon University in Partial Fulfillment of The Requirement for the Degree of Master of Science in Computer Science

Submitted by

Hind Hadi Ali Al -Saadi

Supervised by

Dr. Ali Yakoob Al-Sultan

Dr. Enas Hamood Al-Saadi

2022 A.D. *1444 A.H.*


بسم الله الرحمن الرحيم
يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنكُمْ وَالَّذِينَ
أُوتُوا الْعِلْمَ دَرَجَاتٍ

Dedication

To

The most beautiful women,

a woman who stayed up, tired and raised called

My Mother

To

That who yanked a living from the cruelty of the rocks

My dear Father

To

My companion and friend in my life,

and a mixture of father, brother and friend

My Husband

To

My Children

To all those, I dedicate my work.

Supervisor Certification

I certify that project entitled "**Multi Class Detection of Cataract in the Retina Based on Convolution Neural Network**" was prepared at the Department of Computer Sciences / College of Science for Women/ University of Babylon, by (**Hind Hadi Ali Al-Saadi**) as partial fulfilment of the requirements for the degree of Master in Computer Science.

Signature:

Name: **Dr. Ali Yakoob Al-Sultan**

Date: / / 2022

Address: University of Babylon/College of Science for Women.

Signature:

Name: **Dr. Enas Hamood Al-Saadi**

Date: / /2022

Address: Department of Mathematical, College of Education for Pure Sciences, University of Babylon, Iraq.

The Head of the Department Certification

In view of the available recommendations, I forward the research entitled " **Multi Class Detection of Cataract in the Retina Based on Convolution Neural Network"** for debate by the examination committee.

Signature:

Name:

Date: / / 2022

Address: University of Babylon / College of Science for Women.

Acknowledgments

First of all, I am grateful to the Almighty Allah, the owner of grace and favor whowas and is still my support in preparing and achieving this thesis. My deep love and massive thanks go to the masters of Creatures, the Greatest prophet Mohammed and his progeny (peace be upon them).

Praise is to Allah, who illuminated me the way of science and gave me the patience to continue.

Also, sincere appreciation and love go to all my family members whose have encouraged me to go through this study and provides me with love and pure affection.

I wish to express my thanks and deep gratitude to my supervisors Dr. Ali Yakoob Al-Sultan and Dr. Enas Hamood Al-Saadi for their advice, support, encouragement, and supervision throughout this work to be in the best manner.

I would also like to acknowledge University of Babylon - Collage of Science for Women-Department of Computer Science for providing support in conducting this thesis.

Finally, I would like to thank the people who gave me help and advice.

Abstract

A cataract is a leading eye disease that gradually evolves and has no immediate impact on vision. One of the most common vision problems is cataracts, which causes visual distortion and the late stage of this disease can lead to blindness. It is considered a silent disease that can occur without the appearance of symptoms. Therefore, the most effective way to detect cataracts is through accurate and timely detection to avoid painful and costly surgeries and prevent blindness.

The purpose of this thesis is to propose an automated system based on the deep learning approach (Convolution Neural Network CNN) for the detection of cataract and classify cataract fundus images ito three stages.

The proposed system contains three main phases, the first phase is the pre-processing phase which begins with converting fundus images into gray level fundus images with size equal, and the "Contrast Limited Adaptive Histogram Equalization (CLAHE)" technology was applied to enhance contrast and demonstrate the feature of eye's lens.

Data Augmentation methods are used to avoid overfitting problems and to enhance the performance of the system. In the second phase, the CNN based on multi- Image (fundus images) augmentation technique was applied as a deep feature extraction technique to identify fundus samples. In the third phase, (Softmax function) is employed for the classification of cataract patients and its stages (Mild, Moderate, and Sever). Specificity, Sensitivity, Accuracy, F1-score, and Area Under Curve (AUC) were used as criteria to estimate the efficiency of the classification.

The proposed system using two publicly available datasets: Kaggle and ODIR dataset .

The results and the test of all the details of the proposed system and all the stages were discussed in-depth, the results showed high accuracy of cataract detection and reached (99%) and (96.9 %) for (Normal and cataract) and (Mild, Moderate, and Severe) classification respectively in the test ODIR data set by using CNN-Soft Max classifier.

The proposed method is very fast in diagnosis, the fundus image prediction time takes approximately (0.06) millisecond for one image.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
Acc	Accuracy
AHE	Adaptive Histogram Equalization
AI	Artificial Intelligence
AUC	Area Under Curve
BP	Back propagation
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolution Neural Network
DA	Data Augmentation
DL	Deep Learning
DNN	Deep Neural Network
FC	Fully Connected
FN	False Negative
FP	False Positive
GPU	Graphical Processing Unit
MLP	Multi-Layer Perceptron
ODIR	Ocular Disease Recognition
Pre	Precision
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Networks
Sen	Sensitivity
Spe	Specificity
TN	True Negative
TP	True Positive

List of Symbols

Symbol	*Meaning*
•	Morphological Closing Operation
o	Morphological Opening Operation
$\oplus$	Dilation
$\ominus$	Erosion
Γ	Gamma
Θ	Theta
$\mathscr{M}$	learning rate
$\odot$	Hadar product
D_{out}	The output size of the Activation map
D_{in},	Represent the input size of fundus images
D_k	Represent the kernel size
S	a stride value
p	Padding value

List of Algorithms

Chapter One

General Introduction

1.1 Introduction

People are increasingly relying on their eyes as a result of the widespread usage of electronic devices in modern life. As a result, eye protection is becoming more vital in everyday life. The most prevalent eye disorders are cataracts, diabetic retinopathy, conjunctivitis, glaucoma, and other eye diseases. However, common cause of blindness is cataracts [1].

Cataract is the loss of crystalline lens transparency caused by the clumping of protein inside the lens [2]. Most of the time, the lens converges light onto the retina. Then the existence of cloud structure (a cataract) prevents light from reaching the lens, resulting in impaired clarity of vision. Cataract is a common illness of eye which gradually grows and doesn't damage vision until later in life. Cataracts can develop without causing symptoms, and while they rarely cause pain, they can cause significant vision loss and even blindness[2].

Cataract is divided into three categories based on their causes: Age-related cataract, Pediatric cataract (PC), and Secondary cataract [3][4].They are classified as Nuclear Cataract (NC), Cortical Cataract (CC), and Posterior Sub capsular Cataract (PSC) based on the location of the lens opacity [5]. Age, diabetes, and smoking are just a few of the major causes of these three different forms of cataracts [6].

Depending on the severity of the cataract, early detection may help to avoid painful and costly procedures as well as avert blindness [7]. According to the World Health Organization (WHO), around 285 million people worldwide are visually impaired. There are (39) million people with vision limitations among them, while the rest have impaired vision. Cataract was found to be accountable for (33%) of impaired vision and (51%) are blind [5].

Early detection of cataracts is critical in therapy and drastically reducer the risk of becoming blind. However an autonomous system for detection the cataract is a challenging task. There are few indications of illness in the early stages, automated and effective diagnostic algorithms based on color funds images are desperately needed.

Deep neural networks are emergencies a tool to solve many problems such as: object detection, speech recognition, and image classification, which have seen significant and rapid growth in recent years. Convolution neural networks (CNN's) with the Data Augmentation technique, in particular, produced exceptional images classification results [3].

To address this issue, this thesis proposes a system that detects Cataract automatically based on retinal funds images. The proposed method should provide an acceptable level of accuracy in determining the condition of the eye and whether the eye is healthy or has Cataract, as well as the grading of Cataract. Also, the proposed system should use fewer parameters than the current works to increase the accuracy of the deep learning model, and thus the computational power will be less.

1.2 Problem Statement

Cataracts form gradually and do not damage vision until later in life. The risk of blindness increases with the length of time a patient has cataracts or with delayed treatment. On the other hand, the current approach of detecting and assessing cataracts is manual, costly, and necessitates the use of a skilled ophthalmologist. Therefore, it is necessary to design a good an automatic detection system as a quick alternative diagnosis option to detected Cataract at different stages as (Mild, Moderate, Sever) and to solving the main drawbacks in existing CNN models in the literature.

The previously proposed systems used manual methods as traditional image processing techniques for feature extraction, but features extraction at Deep Learning carried out by the model itself. Then the number of parameters increases and leads to an increase in the training time because of arithmetic operations. The selection of the size of the feature map in each layer of CNN structure is critical in reducing the numbers of parameters.

1.3 Aim of the study

1. Building a prediction system to detect cataract disease with grading cataract severity automatically by utilizing the retinal funds image based on deep learning techniques specifically a CNN model.
2. Find and perform good image enhancement method and noise removal for the input image without distortion of image detail which concede as a key step in the prepossessing stage.
3. Automatic features extraction using Deep Convolution Neural Network (DCNN) to reach accurate results to classify funds images.
4. Improve the accuracy of detection of Cataract in funds images using deep learning and data augmentation technique.
5. Reduce the required cataract detection time.

1.4 Contributions of the Thesis

1. The propose system focussed on developing a deep learning - based model (CNN-Softmax) to detect the cataract and its stages (Mild, Moderate, and Sever) by using funds images.

2. Adopting fewer parameters than the most modern deep learning models (Convolution Neural Network), leading to fast prediction time where it takes approximately (0.06) millisecond for one image, and this allows the system to be used in real-time.

3. The proposed system is highly accurate using CNN-Softmax when compared with previous researchers , the proposed system achieved excellent results for diagnosing the disease in the terms of accuracy of the classification .

1.5 Literature Review

In the related work, several studies have been proposed regarding the discovery of cataract and its stages. These studies are mainly related to extracting characteristics through funds images. As for the techniques that have been used to extract the characteristics, most of the researches used CNN models such as VGG Net and Res Net, and the other used image processing techniques. As for the techniques that have been used to classify Cataract, most of these researches used machine learning methods. In the following, the description of this research:

Zhang, et al. in 2017 [8] Introduced a way of classification of cataract disease using Deep Convolution Neural Network (DCNN) to detect and grad cataract automatically. The best accuracy, this method achieved, is 93.52% and 86.69% in cataract detection and grading tasks seperatey.

Islam, et al. In 2019 [9] Introduced new technique for identifying malignant tum-ours. The convolution neural network (CNN) has been used to diagnose eight different kinds of eye disorders, and the performance of the CNN's has been assessed. Some standard prepossessing is carried out before the data is transmitted to the network for rigorous categorization to be carried out. The greatest level of accuracy was obtained with an F-score of (0.85), (0.31) of a Kappa score, and the value (0.80) of AUC .

Pratap, et al. in 2019 [10] Introduced a way by utilizing a transfer learning technique with an entrained CNN for the automatic classification of cataracts. Utilizing feature extraction and an SVM classifier, the final classification was completed. The fourth stage classification (Normal, Mild, Moderate, Sever) accuracy obtained is 92.91%.

Wang, et al. in 2020 [11] At this method one or more funds diseases may be diagnosed based on CNN-style model imaging of funds images that does not need any extra labeling information. The first half of the solution relies on an efficient net-based feature extraction network, while the second half is a customized classification neural network that is suitable for multi-label classification scenarios. Finally, in order to determine the final recognition result, multiple models' output probabilities are merged. Then producing satisfactory results. (0.89) Accuracy, Recall is (0.58). AUC is (0.73). and Precision is (0.63).

Hossain, et al. in 2020 [7] Suggested a Deep Convolution Neural Network (DCNN)-based method for detecting eye cataracts that has two modules: training and testing.The results of the experiments demonstrate that the suggested approach is highly accurate (95.77%) at detecting eye cataracts.

Ram, et al., In 2020 [12] Used the deep convolution neural network topology with N-Way fully connected layers. This investigation's main emphasis was on the classification of normal, cataract, AMD, and myopia. As the network's feature extraction component (i.e. the convolution net) is trained, the feature mapping component (i.e. the linear net) of the network is also trained to different specifications. The greatest level of accuracy obtained was (0.819), and the highest level of specificity was (0.663), with a sensitivity of (0.714), and a specificity of (0.663).

Syarifah, et al. in 2020 [13] Introduced cataract detection system by classifying the funds image of cataract using CNN and optimize it using Look ahead optimizer. The proposed algorithm can classify the data into two classes. The classes are normal funds images and cataract funds images with accuracy 97.5%.

Sudarsono, et al. in 2020 [14] Introduced cataract detection system by classifying the funds image of cataract using CNN and optimize it using diffGrad optimizer. the proposed algorithm can classify the data into two classes. The classes are normal funds images and cataract funds images with accuracy 97.5%.

Cao Lvchen, et al. in 2020 [15] Offered an automated cataract detection method using the Haar wavelet which has improved according to the characteristics of retinal images. Retinal images of non-cataract, as well as mild, moderate, and severe cataracts, are automatically recognized using the improved Haar wavelet. The accuracies of the two-class classification (cataract and non-cataract) and four-class classification are 94.83% and 85.98%, respectively.

Jayachitra, et al. in 2021 [16] Introduced cataract disease classification using U-Net to detect and grad cataract automatically. The obtained accuracy is 93.5%.

Khan, et al. in 2021 [17] Used a model VGG19 which is a convolution neural network model to detect the cataract by using color funds images. The funds images are classified into (cataract and non-cataract) with accuracy 97.47%.

Sirajudeen, et al. in 2021 [18] Suggested that the diagnosis of cataract disease be made using an image processing technology. This study proposed a Novel Angular Binary Pattern (NABP) for the extraction of texture features. And after the extraction of features, the images are subjected to classification through the implementation of the proposed novel Kernel Based convolution Neural Networks. The accuracy rate of the suggested system is 0.9739.

Junayed, et al. in 2021 [5] Introduced newly developed deep neural network, called Cataract Net. This method is introduced for detecting cataract automatically of funds images into (cataract , non-cataract) with accuracy 98%.

Weni I., et al. in 2021 [2] Utilized a convolution neural network (CNN) for deep learning which is utilized for pattern recognition, which can aid in automating the classification of images. When the epoch value becomes 50 then reached the highest value with a value of 95%.

The following table summarizes the Literature review (Table (1.1)).

Table (1.1) Summary of the Literature Review.

References	Dataset	Model	Accuracy %
Islam, M. T., et al, In 2019 [9]	ODIR dataset	(CNN) has been used to diagnose eight different kinds of eye disorders.	86%
Ram, A., et al, In 2020 [12]	ODIR dataset	DCNN topology with N-Way fully connected layers.	81%
Wang, J., Yang, et al. in 2020 [11]	ODIR dataset	CNN-style model imaging of fundus images that does not need any extra labeling information.	89%
Junayed, M. S., et al. in 2021 [5]	Kaggle dataset	newly developed deep neural network, called (Cataract Net) is introduce for diagnosis cataract automatically.	98%
Sudarsono, E., et al. in 2020 [14]	Kaggle dataset	CNN and optimize it using diff grad optimizer into cataract and non-cataract.	97.5%
Syarifah, M. A., et al. in 2020 [13]	Kaggle dataset	CNN and optimize it using Look ahead optimizer into cataract and non-cataract.	97.5%
Cao Lvchen, et al. In 2020 [15]	The retinal images were provided by Beijing Tongren Hospital.	Utilizing (the improved Haar wavelet).	94.8%
Hossain, M. R., et al. In 2020 [7]	Several eye hospitals in Bangladesh.	This research introduces an eye cataract detection system using (DCNNs)	95.7%
Jayachitra, S., et al. in 2021 [16]	Retinal fundus images from open access data set.	U-Net to detect and grad cataract automatically.	93.5%
Khan, M. S. M., et al. in 2021 [17]	Several eye hospitals in Bangladesh.	used a model VGG19 which is a convolution neural network model to detect the cataract	97.4%
Pratap, T., et al. In 2019 [10]	Several open access datasets.	CNN for feature extraction. Then the classification using SVM	92.2%

Sirajudeen, A., et al. in 2021 [18]	Several eye hospitals in Bangladesh.	for the extraction of texture features, a Novel Angular Binary Pattern (NABP) has been proposed.	97%
Weni, I., et al. in 2021 [2]	Several eye clinics in Bangladesh.	using a convolution neural network (CNN) which is used for pattern recognition	95%
Zhang, L., et al. in 2017 [8]	Several eye facilities in Bangladesh.	using Deep Convolution Neural Network (DCNN) to detect and grad cataract automatically	93.52%, 86.69%

1.6 The Layout of Thesis

The reset of this thesis includes four chapters in addition to chapter one, which is described briefly and have been arranged as bellow:

- **Chapter two, "Theoretical Background":** It describing image prepossessing techniques (Enhancement, Resize, Data Augmentation), the technique of classification for the extraction of funds images features by Deep learning (convolution Neural Network), and criteria for evaluation applied to this thesis.
- **Chapter three, "The Suggested System":** An overview of the structure and execution of classification algorithms that are used in the proposed diagnosis cataract system is demonstrated in this chapter.
- **Chapter four, "Experimental Results and Evaluation":** A description of the various experiments of each step of the work was introduced in this chapter. Moreover, a discussion of the evaluations and results obtained from the execution of the suggested method.
- **Chapter five, "Conclusions and Future Works":** A summary of the study project is presented in this section. Besides, this section shows the future works to be undertaken in this respect.

Chapter Two

Theoretical Background

2.1 Introduction

This chapter offers a summary of various techniques and processes that have been used in the cataract detection system that has been suggested. This chapteris arranged as follows: **Firstly**, an overview of cataract disease with its stages. Image processing techniques. The data augmentation technique is used to increase the training data to prevent the over fitting problem. **Secondly,** a thorough explanation of artificial neural networks is provided, along with Equations for convolutional neural networks. An explanation of the performance metrics such as Specificity, Sensitivity, accuracy, F1-score, AUC and precision are also covered.

2.2 Cataract Disease

One of the most frequent eye conditions that distorts vision is cataract.When a cataract is present, the lens that normally converges light to the retina is blocked, impairing vision. A cataract is simply the accumulation of protein-specific cloud materials that make it challenging to see objects or other entities [2]. Cataracts can occur without the appearance of symptoms. Cataract rarely causes pain but can make central vision loss and it leads to blindness. One of the initial complaints that patients not resistance to bright light. Other complaints that can arise include foggy vision, vision is unclear colors, or double vision [2]. Cataract is the most common eye condition in the world, and it affects vision progressively over time. Cataract is the highest cause of blindness. According to recent studies provided by the WHO-World Health Organization, more than 40,000,000 individuals are anticipated to experience blindness during the next ten years [19]. Based on the reason for the occurrence of cataracts, it can be categorized into three common types like:

1-Age related cataracts (most commonly occurring).

2- Pediatric cataracts.

3- Additional secondary types of cataracts that develop for a number of different reasons.

The first type (Age related cataracts) can be categorized into three stages (Early , Intermediate , and Late stage) that depend on cataract severity [3][4].

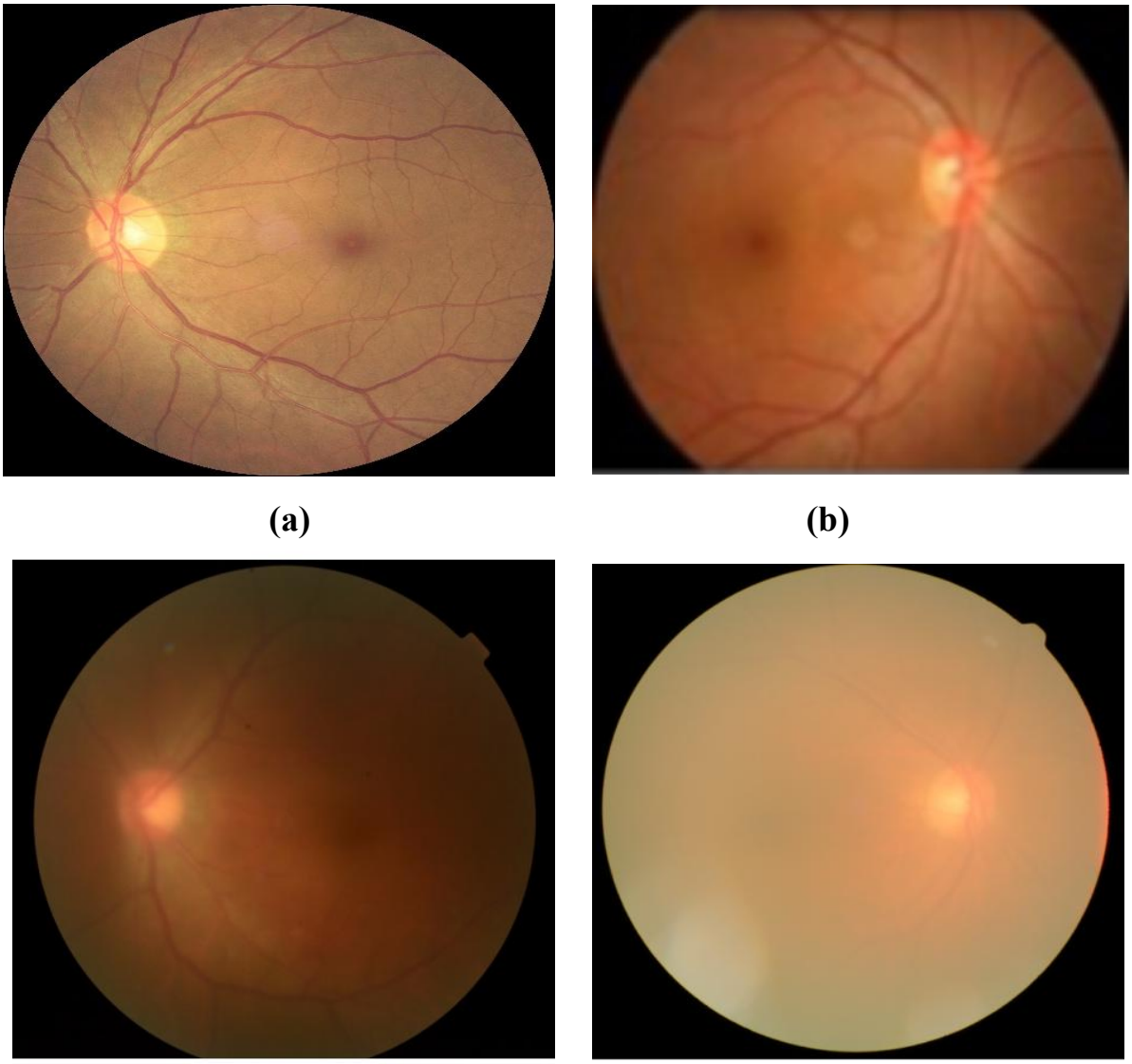
The three stages of cataract can be seen as follows [15]:

- **Early stage - a mild cataract**:when the choroid and capillary vessels are barely discernible, but the major veins and optic disk are clearly apparent.
- **Intermediate stage - moderate cataract**: There are only the optic disc and principal blood arteries are visible.
- **Late stage - severe cataracts**: Retinal structures are not visible.

With more severe cataract, it might be argued that fewer retinal structures can be seen. Figure (2.1) shows the stages of a cataract disease, Figure (2.1) (a) such that normal eye and demonstrates a healthy retina, with the ability to clearly see the main veins, optic disc, choroid, and even capillary vessels. Figure (2.1) (B) shows an eye with cataract in Early stage. Where the choroid and capillary vessels can barely be seen, but the major veins and optic disk are clearly apparent. Figure (2.1) (C) shows Intermediate stage in cataract disease, where just the optic disc and major blood arteries are visible. While Figure (2.1) (D) shows the Late stage, where it is impossible to see any retinal structures. With more severe cataracts, it might be inferred that fewer retinal structures can be seen.

2.3 Image Processing Methods

This paragraph covers the steps for image processing that have been utilized in this thesis for prepossessing and improve the funds images.



(a) **(b)**

(c) **(d)**

Figure (2.1): Cataract disease stages: (a) Normal retina, (b) early Stage, (c) Intermediate stage, (d) Late stage (sever)

2.3.1 Converting Image from Color(RGB) to Grayscale

One of the benefits of transforming images from color to grayscale domain is for reduction of data, because the grayscale image has one channel only. While color images contain three channels (RGB), which lead to speed the processing. This transforming can be done by the following Equation :

$$Gray(i,j) = (0.2989 * R) + (0.5870 * G) + (0.1140 * B) \quad (2.1)[20]$$

Where R = read, G = green, B = blue.

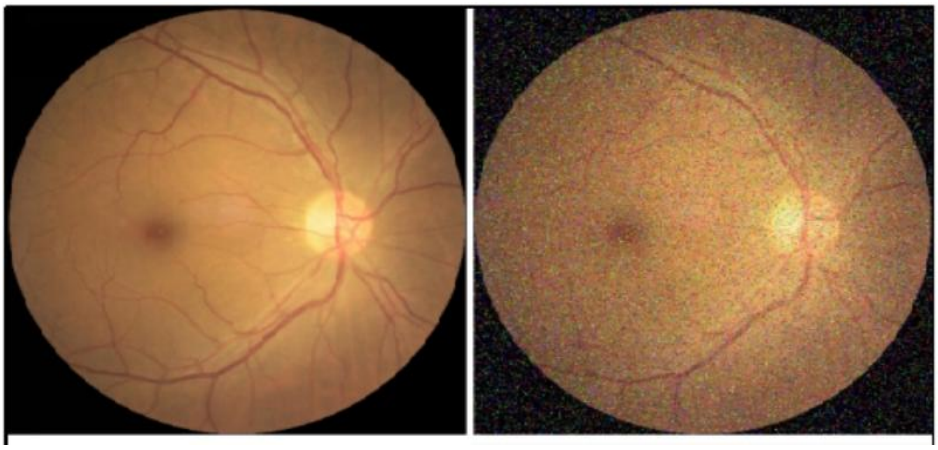
2.3.2 Remove Noise (Denoise)

At this process, the images are cleared from noises and defects. During the capture and transmission of images, there is noise. Filters were used, and it was very effective in reducing image noise. [23]

The noise in image is usually a variance in brightness and color intensity among the pixels of the image. It occurs usually during the process of capturing the image or scanning it, due to artificial effects (such as flashlight) or natural effects (such as high sun bright)[6][23]. These defects carry a large impact on the process of extracting features from the images and classifying them. Some of the most common noises are:

- **Salt and Pepper Noise**: Sometimes, it's called impulse noise, spike noise, random noise, or independent noise. This noise can be caused by sharp and sudden disturbances in the image signal. It presents itself as separated white and black pixels [22].As shown in Figure (2.2).

On the left side, there is normal retina image, and on the right side the same image with salt and pepper noise.

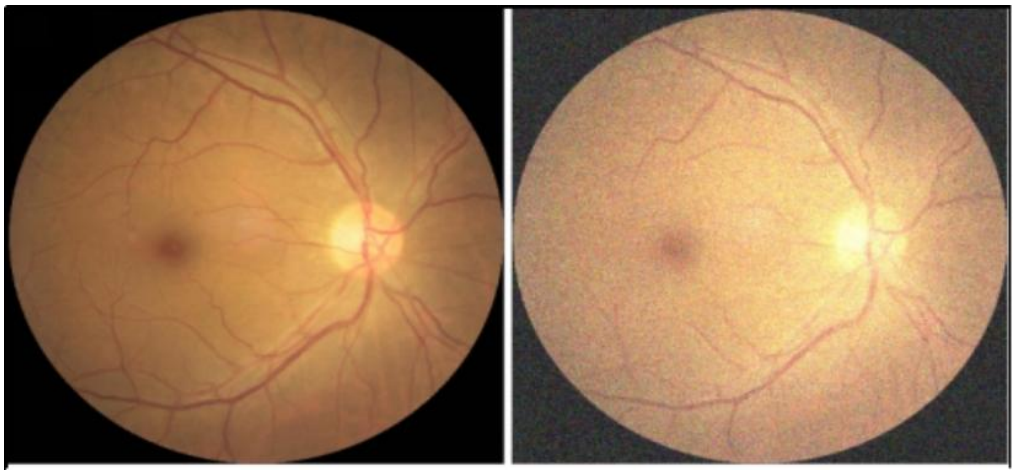


(a) **(b)**

Figure (2.2): Salt and Pepper Noise: (a) before noise, (b) after noise[22].

- **Gaussian Noise**: is also called white noise, or normal noise. This is because the noise has similar density to the normal distribution (i.e., Gaussian distribution).

In practice, the occurrence of this type of noise is usually caused by poor illumination [23]. Figure (2.3) shows the effect of Gaussian noise. On the left side, one normal retina image, and on the right side, the same image with Gaussian noise.



(a) **(b)**

Figure (2.3): Gaussian Noise: (a) before noise, (b) after noise[23].

- **Noise filters**: in order to enhance the quality of the images, and reduce the noises effect, filters (sometimes-called kernel) must be applied to the images. Filter is relatively small matrix, which is multiplied with the image itself to reduce the noise.
- **Average filter (mean filter)**: One of the main defects in images is intensity variation between the pixels in the images.This problem is addressed by applying the Average (Mean Filter). A linear type filtering technique called the Mean Filter is used to smear the image data. An average filter is one that separates out individual pixels by averaging the performance of each pixel mask; this is why it is so named. With the help of this filter, grain noises are primarily removed from photographic images (i.e., financial photographic photos).

  Equation (2.2) shows the Average filter mathematical bases [24].

$$\widehat{f}(x,y) = \frac{1}{m.n}\sum_{(s,t)\in S_{xy}} g(s,t) \qquad (2.2)$$

Where **m** and **n** represent the dimension of the kernel. The **g(s,t)** represents the pixel value of the original image. **s** and **t** represent the coordinates of the kernel.

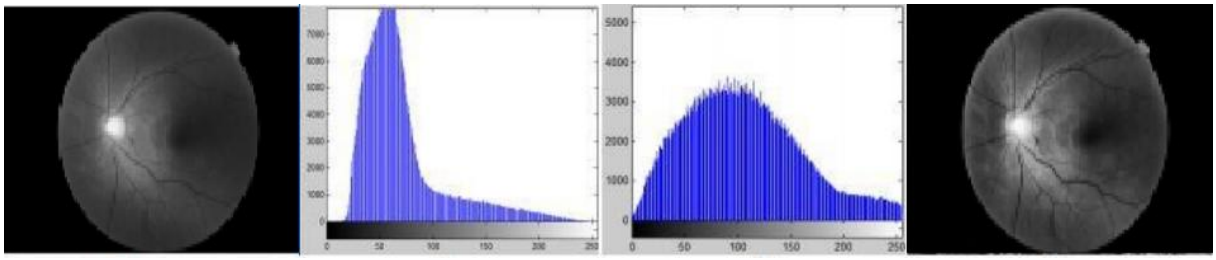
2.3.3 Contrast Limited Adaptive Histogram Equalization (CLAHE)

The enhancement of images is one of the most significant steps in the field of medical image discovery and study. Image enhancement improves the precision and quality of images from a human point of view, eliminates blurring and noise, improves contrast, and shows the image's details. CLAHE is the extended version of the adaptive histogram equalization algorithm (AHE). It was developed to eliminate the over- amplification of noises that occurs in the AHE method. CLAHE adopts a technique for reducing contrast amplification that is implemented for each adjacent pixel, which then forms a change mechanism to minimize the noise [25].

The CLAHE technique applies two parameters; the first is clips limit (CL), which is a numerical value specifying the noise amplification. The second is the number of tiles (NT), a numerical value specifying the number of non-overlapping sub-regions. The computation of CLAHE is performed in equation (2.3).

$$p = (p_{max} - p_{min}) * P(f) + p_{min} \qquad (2.3)$$

Where, ***p*** represents pixel value after applying CLAHE, p_{max}, p_{min} represents the maximum and minimum pixel value of an image respectively and ***P (f)*** represents cumulative probability distribution function after the clip limit [26][27]. When the CLAHE technique is implemented to gray scale image, it will significantly increase the gray scale image's optical characteristics. As a result, more information will be visible in the improved image, as shown in Figure (2.4) [28].



(a) (b) (c) (d)

Figure (2.4) CLAHE stages (a) Gray scale image (b) Original image histogram. (c)Enhanced image histogram, (d) The enhanced image.[28]

2.3.4 Image Normalization

Normalization is a method used in image processing that change the range of Pixel intensity values that may be seen photos with poor contrast due to glare are one example of this phenomenon.

For a variety of applications, the goal of dynamic range expansion is often to place an image or other kind of signal into a more familiar or normal range for the senses, thus the term normalization. When dealing with a collection of data, signals, or images the aim is often to preserve the dynamic range of the collection in order to avoid mental distraction or fatigue. For example, a newspaper will make an effort to ensure that all of the photographs in an issue have a grayscale range that is similar [29]. Equation (2.4) shows how the normalization works in mathematical formula.

$$X_{Normlization} = \frac{X - \min(X)}{max(X) - \min(X)} \qquad (2.4)$$

Where: $\mathbf{X}_{\mathbf{Normalization}}$ is the new normalized image, **X** is the original image, **max(X)** is the upper value in the new normalized image, **min(x)** is the lower value in the new normalized image [2].

2.4 Data Augmentation

The imbalanced datasets can be regarded collectively as a problem when concerning computer vision and image classification issues. The problems of under fitting and over fitting may occur if there is a deficiency of images in each category. This will exceedingly affect the efficiency and performance of deep learning, particularly CNN. To avoid the above problems, a technique for augmenting data within cataract datasets has been suggested to improve the classifier's performance. Data augmentation is a method that increases the number of images used for training the neural network. That is, creating new data for categories that have fewer numbers in the data set. This process succeeds the constraining effect on data to prevent an asymmetric representation and successfully escape over fitting complications. Appropriate data augmentation techniques can help improve deep learning model strength [30]. The Deep Convolutional Neural Network (DCNN) models need a tremendous amount of data for sufficient training and have seen improved results on larger data sets [31].

CNN's are extremely complicated, with so many million weights that have to be balanced and fit to the training sample. Moreover, the Data set in question must be of good quality and wide with a sufficient number of instances. However, in deep learning applications, the number of parameters sometimes surpasses the number of training samples [32].

There are various techniques to augment classic data such as zooming, flipping, shifting, rotation, and transformation to be implemented to the original images.

1. Horizontal and Vertical Flip

Both methods of vertical and horizontal flipping are very effective and popular for data augmentation. The augmented data is one of the simplest applications in terms of execution, which has been proven to be useful in the data set. At the same time, the flipping of the horizontal axis is considered to be more common compared

to the flipping of the vertical axis, as it is very convenient for most of the projects due to its high probability for the existenceof images flipped horizontally as shown in Figure (2.5) (b) [33].

2. Rotation

Another method of data augmentation is rotation, and it is considered as an efficient method. Rotations are performed by rotating the right or left image on an axis between (1 and 359) as shown in Figure (2.8) (c). The maximal angle of rotation is adjusted by the degree of rotation. It is useful whenever it is slight because the network must recognize the object in any direction in the image.

This precise rotation of the image can cause problems for certain applications. As the rotation angle increases, the data label is no longer retained after conversion, and background noise is being introduced. If the background noise was too various compared to the other areas of an image, then the networks can learn incorrect features [32].

The background noise problem can be solved by converting the resulting blank space after rotating into white pixels using the fill mode (contrast) technique [34].

3. Zooming

A random zoom can be applied as a useful processing level for image data with different width and height dimensions. In addition to that, random zoom can also implement a very comparable influence to translations. But the difference between them is that zooming will reduce the size of the image if the zoom range argument was less than one and increase the size of the image if it is greater than one, as shown in Figure (2.8)(d). At the same time, the translations conserve the spatial measurements of the image [33].

To retain the images quality, there are many choices for filling the empty spaces generated by data augmentation techniques (such in Zoom, Rotation) with various pixel values to preserve the image's quality.

There are many methods for filling the empty space, including:

A. Nearest: in this method, all empty space is filled with the values of the nearest pixel.

B. Reflect: in this method, all empty space is filled with a reflection of the original images, in the opposite order of the known pixels.

C. Constant: in this method all empty pixels in the image are filled with the constant value (such as 0 or 255).

D. Wrap: this method is somewhat similar to the reflect mode. However instead of reflecting the pixels in opposite order, it will copy the pixels in their normal direction so it maintains consistency of the arrangement [34].

(a) Original **(b) Flip horizontally** **(c) Rotation** **(d) Zoom**

Figure (2.5) Different data augmentation techniques [34].

2.5 Machine Learning Models

2.5.1 General Introduction to Machine Learning

Machine learning is one of the artificial intelligence sections that depends on computer science, statistics, and mathematics. The basic goal of machine learning is to enable computer programs to learn from data and then make appropriate decisions based on the information that has been learned by a prior experience or prior skills. The machine learns directly from the fundamental input data structure and becomes more intelligent [32].

Machine learning methods are generally classified into three major categories: supervised, unsupervised, and reinforcement.

1. Supervised Learning

Models are trained in supervised learning based on input data (X) and output data (Y) to forecast the future outputs of the unseen input, where it is supervised or monitored learning based on output data or label the ground truth. Depending on this label, the algorithm frequently operates forecasts until it reaches the level of acceptable performance. General algorithms in supervised learning are divided into two main categories:

A. **Classification:** is a group of data related to each other where the intention is to divide the points of this data into a collection of predefined categories based on some characteristics of the data such as Neural Network, Support Vector Machines, k-nearest-neighbors, Random Forest, Naive Bayes.

B. **Regression:** is used to forecast real values that are referred to as continuous values such as Decision Trees, Linear Regression, Assembly methods .

Supervised learning technology will used in this thesis. The classification is a

deep learning algorithm .CNN model will be used to diagnose the images of cataract.

2. Unsupervised Learning

Common algorithms within unsupervised learning are the clustering problem algorithm such as k-mean, the algorithm of A priority for association rule learning, and the dimensional reduction algorithm. Models are trained in unsupervised learning based on input data only, without these inputs being labeled i.e., the model is not given the ground truth label during training. It is the exact opposite of supervised learning, where input data is divided into a group of elements that share the same attribute [35].

3. Reinforcement Learning

Reinforcement learning is defined as action-based learning. An agent who, in a given situation, takes steps to optimize rewards. The agents are expected to determine the fastest possible route to obtain the reward. This type of algorithm does not require any data for learning. In reinforcement learning, the input required is instead a function that can calculate the reward. One of the most important reinforcements learning applications is the path exploration function, which is used in computer vision to locate a particular room or location. In video games, another application is to find the right movements to gain the game [36].

2.5.2 The Activation Functions

The activation function is fundamental in transmitting information to a neural network to learn and handle any complicated tasks. There are various kinds of activation functions, some linear and others nonlinear, and the output value ranges between (0, 1) for unipolar activation function or (-1, 1) for bipolar activation function. In the hidden layers, the sigmoid function as illustrated in Equation (2.5) and the Rectified Linear Unit (ReLU) function, shown in Equation (2.6), are the most widely used activation functions in these layers, as shown in Figure (2.6) [35].

$$A = \frac{1}{1+e^{-x}} \qquad (2.5)$$

$$g(z) = \max(z, 0) \qquad (2.6)$$

ReLU is a non-linear function that substitutes all image pixels whose value is negative in the activation map with zero value. This leads to the advantage of calculation speed and reduces the occurrence of over fitting [37].

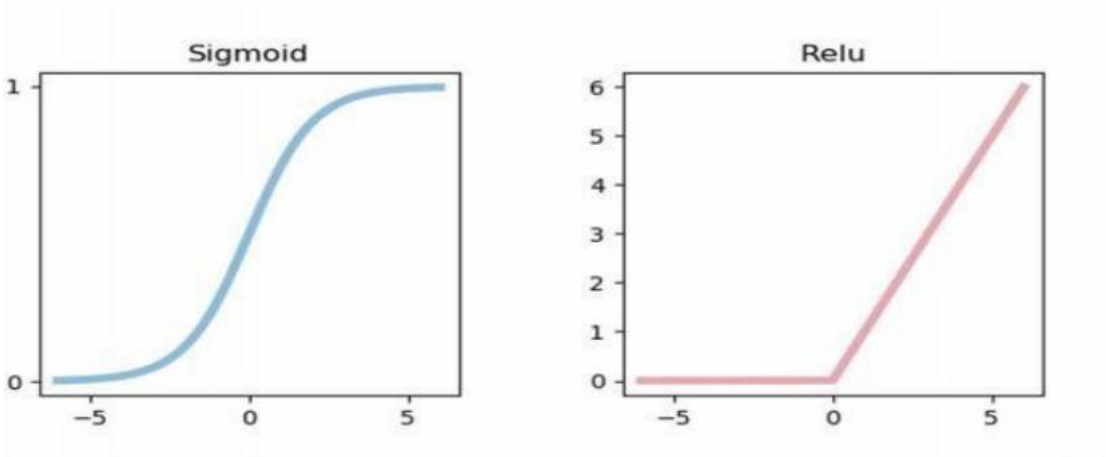


Figure (2.6) illustrates activation functions.[24]

Whereas in the output layer, the **soft max** function may be used as an activation function that is widely utilized to measure probabilities and execute multi-class classifications because ranges of soft max are in the range of 0 and 1, and they have a value of 1 if all the features are used according to the following Equation (2. 7) [37].

$$softmax\,(x_i) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad for \;\; j = 1 \ldots K\,. \qquad (2.7)$$

Where K is the number of classes, **Zj** is the production corresponding to class j

2.5.3 Loss Function

The loss function (also referred to as the cost function) measures the neural network's ability to predict the presented date's right output. Several loss functions are used in machine learning, such as the mean squared error (MSE) and cross-entropy [38].The cross-entropy is the most specific loss function of classification since it determines the classification algorithms performance depending on the probability of the class falls within the range (0, 1) and it is described as following:

$$loss = -\sum_{i=1}^{n} y_i \log\,(\hat{y}_i) \qquad (2.8)$$

Symbol n is for how many classes there are, y represents the true value, and $\hat{y}$ represents the predicted value [35].

The cross-entropy cost and the sigmoid function are called the binary cross-entropy loss and are used when there are only two classes of classifications (0,1). In addition to the Soft-max function, the cross-entropy cost is called categorical cross- entropy loss and is used where two or more label classes exist (labels are given in the one-hot expression). The Sparse Categorical Cross-entropy loss function is used if the labels are presented as integers [32].

2.6 Deep Learning Technique:

A subset of machine learning known as "Deep Learning Technique" which simulates complex abstract concepts in data by using a multi-layer architectural design. The most commonly from of DL models is neural networks, and non-linear transforms in its algorithms. The goal of using such techniques is to achieve "real" artificial intelligence, whichmeans that a machine can learn how to perform extremely complex tasks in a way comparable to how the human brain works through layers of neurons. Computationally, construction and training, deep learning are expensive. Recent developments in applications based on general-purpose graphic processing units (GPU s), the rapid advancement of machine learning algorithms, processing of signals and information, and the growing amount of data that is used in training are all reasons that have increased the popularity and success of deep learning [39]. Figure (2.7) demonstrates a deep nervous system structure consisting of several layers that make it deep [40].

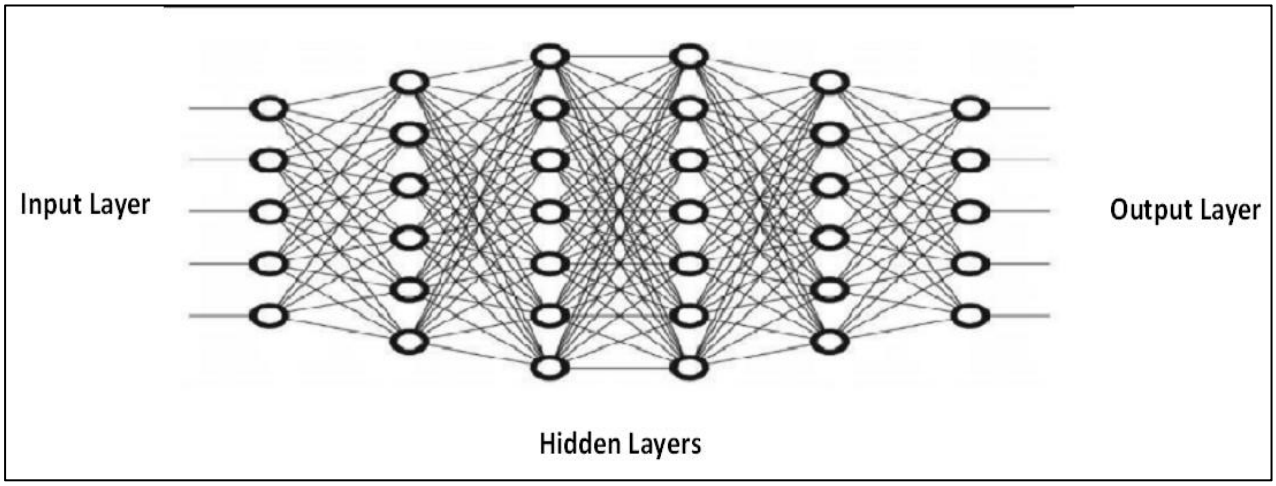


Figure (2.7) Deep network architecture

2.6.1 Deep Neural Networks (DNN)

The DNN is an artificial neural network (ANN) that is made up of multiple hidden layers. MLP is the most general ANN architecture that is used for DNN. Neurons that are coupled at different levels together to form neural networks. DNN needs long calculation periods to implement and many data feeds at training samples, so the number of weights in these networks will reach thousands or equal millions. There are different architectures for deep learning networks, the most important of which are: Convolutional Neural Network (CNN), which is used to classify images, and Recurrent Neural Network (RNN) which is used with texts and continuous data [41].

2.6.2 Convolution Neural Networks (CNN)

CNN is one of the kinds of DNN's that are most commonly used in machine vision fields. CNN is analogous to multi-layer Perceptron, except that the variation lies in its ability to combine many regionally interconnected layers are utilized to extract features, accompanied by some completely interconnected layers are employed for categorization [42]. It is an effective technique for image recognition, visualization of images, and identification of objects. When it is related to Deep Learning and medical image processing, CNN's are the most commonly used neural networks in AI because this network can process a large amount of data and does not need to extract features manually[32].

In convolution neural networks, the input data is in the form of a matrix ofpixel values instead of feed forward neural networks (FFNs) in which the input is a vector of pixel values [35].

One of the most powerful features of convolution neural networks is their use of shared weights,i.e. group of connections that share the same weights instead of

using different weights for each connection. The other feature is that CNN has the local connection where each neuron does not contact all the neurons in the previous layer. Still, rather it only contacts a specific group of neurons to see if they contain the object's feature instead of contacting all cells. This produces strong responses to obtain local characteristics in an image input (such as ridges, edges, curves). These two features exceedingly reduce the number of parameters in the network, and thus the training time will be reduced [37].

Three distinct layers make up the CNN's standard architecture, as shown in Figure (2.8), and any convolution Neural Network model is built from these different layers:

1. Convolution layer.

2. Max pooling layer (or Sub Sampling layer).

3. Fully Connected Layer (Classification layer).

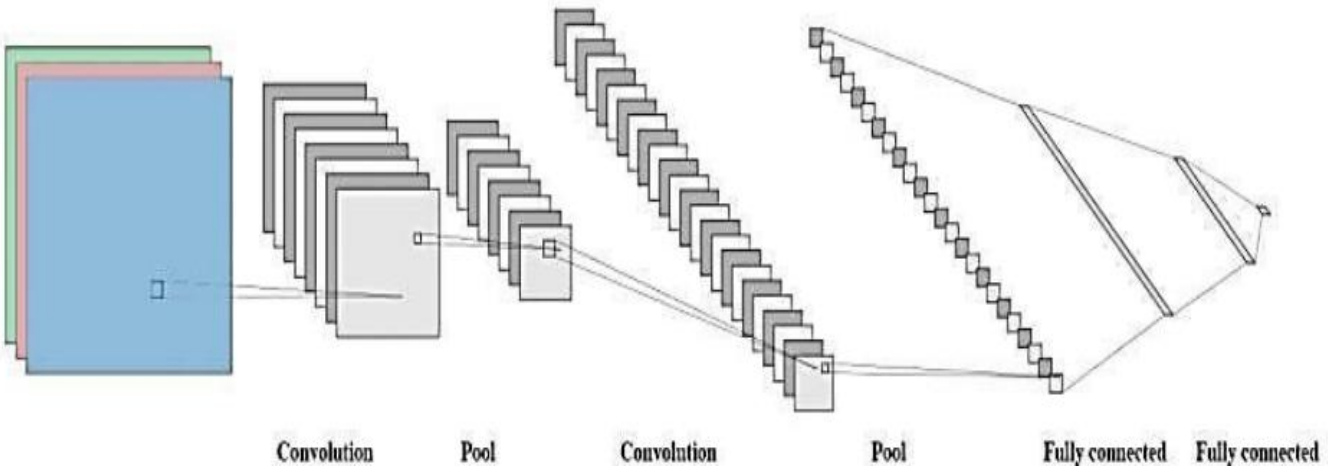


Figure (2.8) The structure of a CNN [43].

1. The Convolution Layer:

It is the fundamental layer to construct a CNN model. This layer's main purpose is to extract features from the original input image using the mathematical operation that is known as "convolution," .This layer contains three matrices: the first is the

input image, where it is transformed into a matrix (be three-dimensional or grayscale two- dimensional) and the second is called the filters matrix, also named "feature detector" or "kernel." The third matrix is the result of moving the filter matrix with horizontal and vertical steps over the input image to compute the dot product called "Feature Map" or "Activation Map" this is the convolution process as shown in the Figure (2.9) [44].

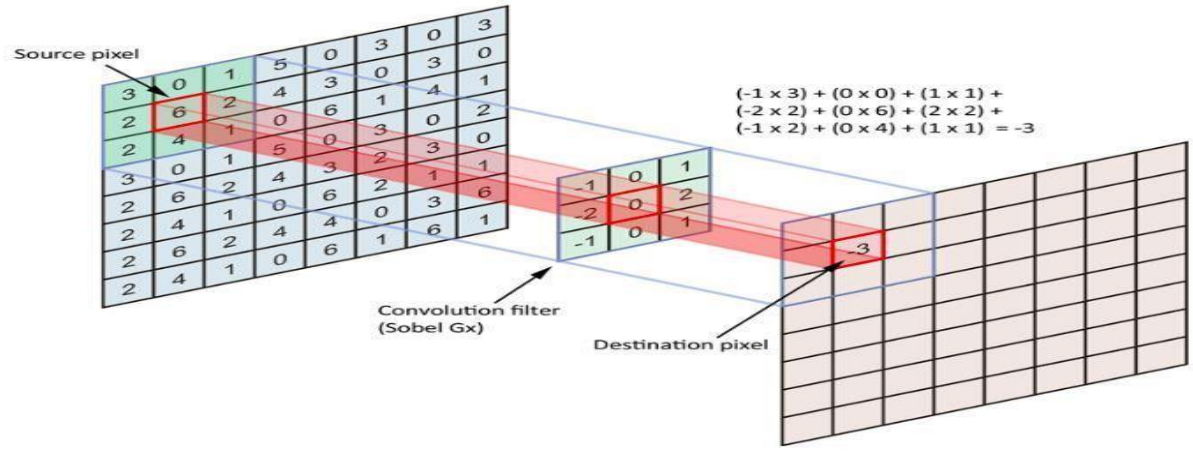


Figure (2.9) convolution operation between input X and the kernel K in CNN convolution layer [44].

Multiple convolution filters are applied for a single input. The resulting activation maps are then combined to obtain the final result for a single convolution layer, where this final result represents input data to the next layer.

Each value of the filter matrix is a weight that is given by default[44]. These values must be different from one filter to another to give different characteristics or features to each matrix from the feature map matrices [45].

The following equation (2.9) describes a general convolution layers processfor images.

$$Z_{ij} = (X * K)_{(i,j)} = \sum_{l=1}^{k1} \sum_{a=1}^{k2} \sum_{b=1}^{C} X_{(i+l,j+a,b)} K_{(l,a,b)} \qquad (2.9)$$

In the above equation, the convolution process between the input X and the kernel weights K in convolution layer l to produce the feature map Z and $K1$ is the height of the kernel, $K2$ is the width of the kernel, C is the number of channels [46].

Following every convolution layer, the results pass through the activation function (ReLU) non-linear as in equation (2.6) [37].

There are three parameters which determine the size of the feature map which are:

1. **Depth:** the depth represents the number of filters that are implemented in the convolution process. If the original image was convoluted using three filters, then the activation maps 'depth' will be equal to three[.
2. **Stride:** represents the number of pixels in the filter matrix that leaves during the convolution process. If a stride is equal to 1, the kernel will move by 1. If it equals 2, the kernel moves by 2. As the number of steps increases, the function maps diminish.
3. **Zero-padding:** surrounds the input matrix with zero values so that the number of rows and columns increases by 2; if p=1. It is possible to make larger padding if p=2, then the number of rows and columns increases by 4, and so on. This technique has been used to process gradual image fading across multiple layers in the CNN. Another problem is that the pixels on the outside of the image are not used in the convolution operation except very few times, which reduces the efficiency of the model [47] [48].

There are two types of convolutions: a valid convolution, which is a convolution without padding, and therefore the size of the matrix will gradually decrease; zero or "same" convolution, the convolution in which the image size remains and does not change before or after the convolution, this type uses the technique of padding [44].

The following equation illustrates how to calculate the feature map's size in the same convolution.

$$\mathrm{D_{out}} = \frac{D_{in} - D_k + 2p}{\mathrm{S}} + 1 \qquad (2.10)$$

Where $\boldsymbol{D_{out}}$ is the output size of the Activation map, $\boldsymbol{D_{in}}$, $\boldsymbol{D_k}$ represent the inputsize and kernel size, respectively, **S** is a stride, and **p** is the zero padding and can be calculated by ($\frac{D_k-1}{2}$) [37].

2. Max pooling layer (also called subsampling or downsampling).

It is the operation of down sampling of a group of adjacent pixels into a single pixel. This layer was employed after the convolution layer to reduce the size of the image activation maps[49]. There are common types used of Spatial layers: (Average, Max, etc.) as shown in Figure (2.10). The Max-Pooling determines a geographical area (sub-region) like (2×2 window) and selects the biggest element from the rectified activation maps within each window. The activation maps image size was lessened from (4×4) to (2×2) by this down sampling. Whereas average pooling returns the mean value for each sub-region [49].

The max-pooling down sampling process is described as in equation (2.11). Pooling layers can solve the problem of over fitting, and Max Pool has proven stronger:

$$s_i = \underset{i \in R_j}{max}\, h_i \qquad (2.11)$$

Where h represents some pixel in the window (or sub-region) R_j from the rectified Activation Maps [50].

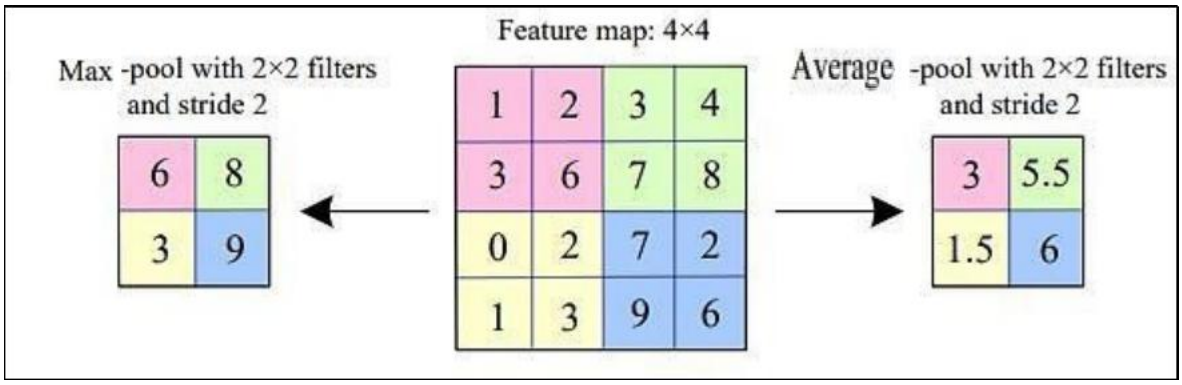


Figure (2.10) Pooling operations [49].

3. Fully Connected Layer (Classification layer).

This layer works in a similar way to the traditional MLP network as each node has a complete connection to all the nodes in the following layer ("Fully Connected")[51]. It's used as a classifier in the last layers of CNN structure to assess the probability of the object in the image by using the Soft max function in the output layer illustrated in section (2.5.2).

2.6.3 Training the Convolution Neural Networks:

Back propagation is the core component of ANN, and it is more difficult in CNN's because it contains various types of layers[48]. This technique calculates gradient descent for all network weights in two stages: the first has been called forward propagation, as the training starts from the inputs of the first layer in the network to the last layer in which the error between the outputs and the desired value is calculated through the loss function shown in the section (2.5.3). The second stage is called backward propagation, which starts from the last layer in the network to the first layer [48].

1- Forward Propagation:

Training samples are spread across all network layers from input to output using various equations that depend on the layer type to output a prediction value.

- The convolution layer in the forward propagation implements a convolution process between all its inputs and filters using equation (2.8). Then the output is passed to the ReLU function using the equation (2.5).
- Max pooling layer in the forward propagation is explained in the equation (2.10).
- A fully Connected Layer in the forward propagation works in a similar way to the MLP network that is illustrated in equation (2.12).

$$s = \bar{X}\bar{W} = \sum^{n} x_i \, w_i + b \qquad (2.12)$$

Where a neuron computes the dot multiplication between the input vector X=[x1, x2,..., xn] with their identical weights W=[w1,w2,...,wn] and then adds the bias value (b) as in the formula following (2.11). Finally, the value (s) is transferred to the activation function. The SoftMax function which is shown in equation (2.6) was used to assess the probability of the object in this layer. After that, the error between the desired value and the output is calculated using the loss function that has been defined in equation (2.7) [49][44].

2- Backward Propagation:

Backward propagation starts from the last layer (output) in the network to the first layer (input), which operates as follows depending on the layer type:

- The Fully Connected Layer in the backward propagation works in a similar way to the MLP network. The delta (error) is calculated in the output layer using the following equation .

$$\delta^L = \nabla_x f \odot \acute{g}(z^L) \qquad (2.13)$$

Where δ^L is a matrix of all the neurons' deltas in the layer, L indicates the final layer of the network, $\nabla_x f$ is the scale of loss for x, where x is the activation function's output in this case. ⊙ is the Hadar item (product of matrices).

The equation (2.14) is used to determine the errors for hidden layers:

$$\delta^l = ((w^{l+1})\ \delta^{l+1}) \odot \acute{g}(z^l) \qquad (2.14)$$

Here l represents a concealed layer's layer count, (w^{l+1})represents the weight matrix of the next layer. Modify all weights in the network by computing the weight adjustment $\Delta w_{i,\mathrm{j}}$ in equation (2.15).

$$\Delta w_{i,} = -\mu \frac{\partial \mathcal{L}}{\partial w_{\mathrm{i,j}}} = -\mu \delta_{\mathrm{j}} z_i \qquad (2.15)$$

Where μ represents the learning rate, δ_j represents the delta of neuron j [48].

The convolution layer in backward propagation follows the same rulesas the Fully Connected Layer (explained above), except that the convolution filter shares weights for the whole layer.

Equation (2.16) is used for calculating the effect of filter weights on the Loss function that is used to update the weights:

$$\frac{\partial \mathcal{L}}{\partial k_{a,b}} = \sum_{i=0}^{n-m} \sum_{j=0}^{n-m} \frac{\partial \mathcal{L}}{\partial z_{i,j}^l} \frac{\partial z_{i,j}^l}{\partial k_{a,b}} \qquad (2.16)$$

According to the Equation (2. 9), $\frac{\partial z_{i,j}^l}{\partial k_{a,b}} = x_{(i+a),(j+b)}^{l-1}$, and from Equation (2. 14), $\frac{\partial \mathcal{L}}{\partial z_{i,j}^l} = \frac{\partial \mathcal{L}}{\partial y_{i,j}^l} \acute{g}(z_{i,j}^l)$, The input and output feature map notation (p, q) is appended to the formula, resulting in the following Equation (2. 17).

$$\frac{\partial \mathcal{L}}{\partial k_{p,q,a,b}} = \sum_{i=0}^{n-m} \sum_{j=0}^{n-m} \frac{\partial \mathcal{L}}{\partial y_{q,i,j}^{l}} \acute{g}(z_{i,j}^{l}) x_{p,(i+a),(j+b)}^{l-1} \qquad (2.17)$$

Since the phrase ($\frac{\partial \mathcal{L}}{\partial y_{q,i,j}^{l}}$) is assumed to come from layer l+1,then the weights would be adjusted using Equation (2. 15). The last step is to propagate error into l-1 layer , which is computed by Equation (2.18):

$$\frac{\partial \mathcal{L}}{\partial y_{p,i,j}^{l-1}} = \sum_{q \epsilon F^{l}} \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \frac{\partial \mathcal{L}}{\partial z_{q,(i+a),(j+b)}^{l}} k_{p,q,a,b}^{l} \qquad (2.18)$$

The result looks similar to convolution process and can be understood as a convolution of error with a reversed (flipped) kernel [48][52].

2.6.4 Regularization Techniques:

Several techniques have been proposed to prevent the problem over fitting called regularization.

- **Dropout**

The dropout is the technique of randomly dropping a set of neurons with a predetermined probability value at each training iteration. This technique greatly improved neural networks' performance to solve the over fitting problem [25], as shown in figure (2.11). The dropout layer Equation can be executed as follows:

$$_{j}r^{i} = \sim(\text{bernoulli(p)}) \qquad (2.19)$$

$$\gamma\sim^{l} = r^{l} \odot \gamma^{l} \qquad (2.20)$$

Where r^{l} represents a vector of independent Bernoulli random variables, which each have a probability of being 1 of p, the result of *l* layer, and ($\gamma\sim^{l}$) is decreased output that will be as input X in Equation (2.11) [49][34].

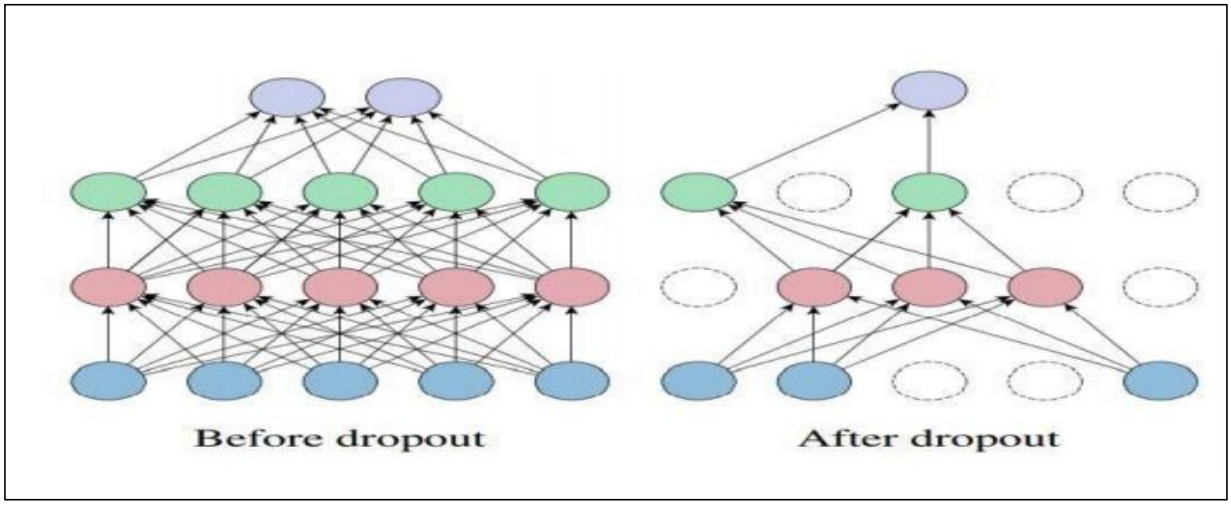


Figure (2.11) The dropout influence in a network.[49]

- **Early Stopping**

Early stopping is a powerful regularization technique that is used to find the best number of epochs in deep learning network training. With the increase in the number of epochs, when the loss increases in the validation sample and reduces in the training sample, the training process must be stopped after few subsequent periods using early stopping as shown in the figure (2.12), because if the number of epochs increases, it will lead to over fitting, and if the number of epochs reduces, it will lead to Under fitting [53].

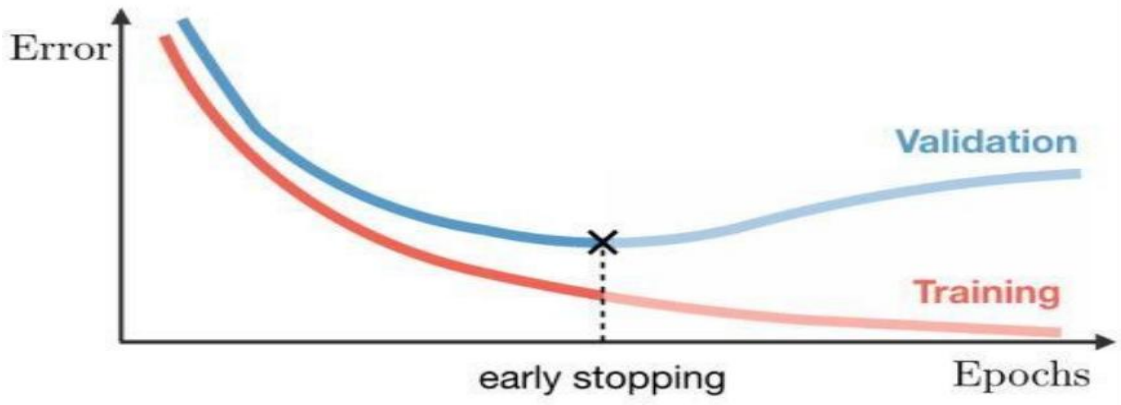


Figure (2.12) The early stopping technique.[53]

- **Data Augmentation**

This method is also one of the most efficient techniques in handling over fitting and has been illustrated in section (2.4).

2.6.5 Optimization Algorithm (Adaptive moment estimation (Adam)).

By altering the weights and bias values in the model, the optimizer aids in lowering the loss function's output error. There are many performance optimization tools used in deep neural networks, the most important of which is the Adam Optimizer.

"Adaptive Moment Estimation" is what Adam means, atool that calculates each parameter's adaptive learning rate and conserves the previous square gradients' average exponential decay rate as it maintains the average exponential decline of the previous scales. Adam performs better than any other optimization technique, making it appropriate for the demand for rapid convergence and highly complex neural networks [21].

2.7 Performance Measures

Several criteria have been employed to assess the performance of the classification algorithms:

1. **The accuracy:** The number of correctly classified cases is used to measure accuracy, whether negative or positive instances.

$$\textit{Accuracy}\ (\text{Acc}) = (TP+TN/TP+TN+FP+FN) \quad \ldots.(2.21)$$

2. **Sensitivity:** is the rate of identification of positive samples rightly.

$$\textit{Sensitivity}\ (\text{Sen}) = (TP/TP+FN) \quad \ldots..(2.22)$$

3. **The precision:** of the model's performance is determined by testing the true

positive from the expected positives.

Precision (Pre) **= (TP/TP+FP)** …..(2.23)

4. Specificity: is the proportion of accurate identification of negative cases.

Specificity **(Spe) = (TN/TN+FP)** …..(2.24)

5. F1-score: illustrates a combination of accuracy and sensitivity for calculating a balanced mean output.

F1- Score **= (2*TP)/(2*TP+FP+FN)** …..(2.25)

6. Area Under the Curve (AUC):

AUC **= (Sensitivity + Specificity)/2** …..(2.26)

Matrix of classifier system is shown in Table (2.1).

Table (2.1): Confusion Matrix of Classifier System.

		Predict	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

The prediction error is recorded by four parameters:

- The positive states that are accurately classified as positive states are known as True Positives (TP).
- False Positive (FP) is a negative state that has been mistakenly classified as a positive state.
- The appropriate category for a negative diagnosis is True Negative (TN).
- False Negative (FN) is a term that refers to mistakenly positive cases.

Chapter Three

The Proposed System

3.1 Introduction

In this chapter, all the methods and techniques used to build the proposed system for diagnosing cataract using funds images will be described. The introduced (proposed) system include three phases, The First phase, is the prepossessing phase which begins with converting funds images into gray level images of equal size, the Contrast Limited Adaptive Histogram Equalization (CLAHE) technology was applied to enhance contrast and demonstrate the feature of retina. The Second phase, the (CNN) based on multi-Image which augmented by applied augmentation technique as a deep feature extraction technique to extract specific feature form funds images then used this feature for the next phase. And in the Final-phase**,** Soft max function is employed for the classification of cataract disease and its stages (Mild, Moderate, Sever).

In this thesis, 70% of the database was used to train the system through passing in the three phases of the above. As the remaining database, (30%) was used to test the model after fully trained. The proposed system is tested in order to evaluate its accuracy and performance, the evaluation criteria such as Specificity, Sensitivity, Accuracy, F-score, Area Under Curve (AUC) were utilized as key criteria to evaluate the efficiency of the introduced model.

3.2 The Cataract Diagnosis System

The system has many stages that are performed include funds image prepossessing operations, the features extraction process, classification process, and finally evaluation of model performance using multiple criteria. The proposed system is demonstrated in Figure (3.1) and (3.2).

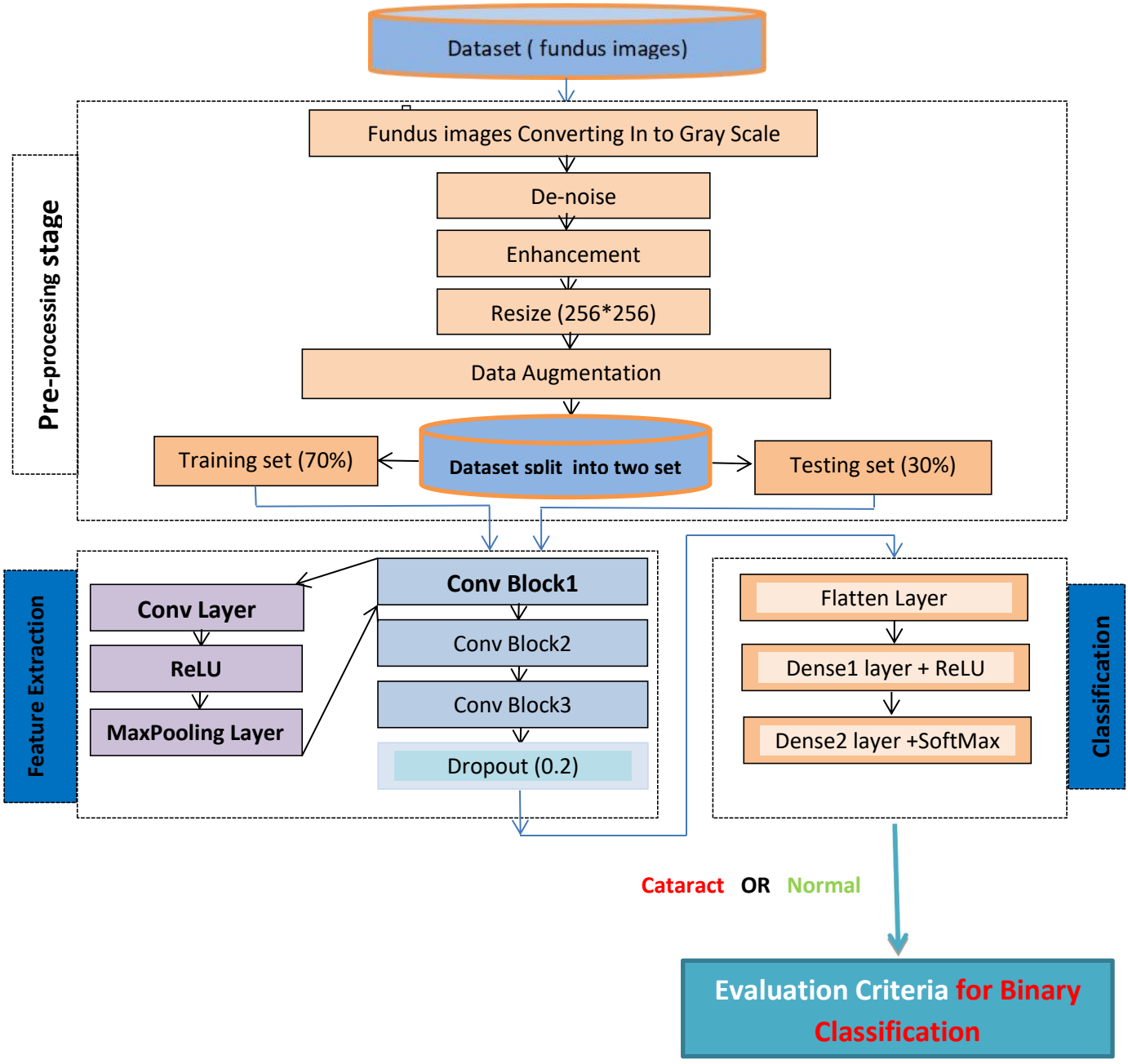


Figure (3.1) Block diagram of the proposed system for Binary Classification.

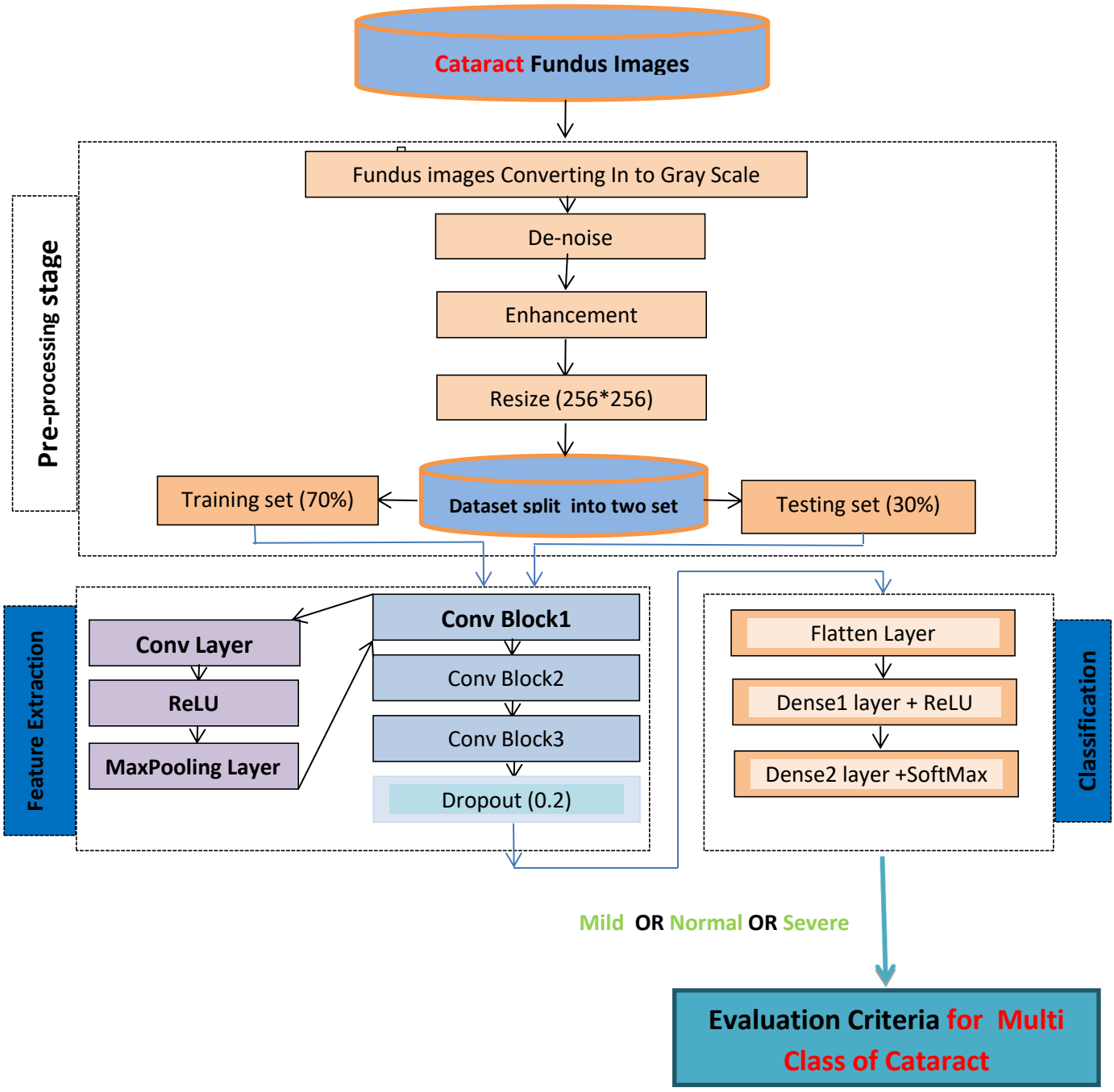


Figure (3.2) Block diagram of the proposed system for Multi Classes of Cataract.

3.3 Prepossessing Phase

Operations of prepossessing are a series of fundamental steps in the cataract classification system to initialize data for the next stage (the feature extraction stage)and in order to focus on disease-specific features to increase the accuracy of the system. It consists of five steps: Convert funds images to Gray-Scale Images, Resize funds images, Denoise funds images by applying Mean filter, funds Images enhancement using CLAHE technology, and Data Augmentation on prepossessing dataset. The prepossessing phase include the following steps:

1. Convert Input funds Images from color fundus images to Gray-Scale .
2. Resize Gray-Scale Images to 256×256 grayscale 8-bit image.
3. Removed the Noise from funds images by applying Average (Mean) Filter using the Equation (2.1).
4. Enhanced images by applying CLAHE technology using theEquation (2.2).
5. Implemented the Data Augmentation technique on the enhanced images

3.3.1 Funds Images Converting

Transforming color funds images into a gray scale funds image would reduce the time of processing in the following step and simplify processing of the pixels. Which means, instead of dealing with three channels in colored images (RGB), the proposed system will deal with only one channel.

3.3.2 Funds Images Resizing

All funds images are scaled at a fixed (256×256) size to be suitable for more processing in the deep learning systems. This step would benefit to reduse the computational cost; re-sizing the funds images to a (256×256) pixel scale would enhance the efficiency of processing .

3.3.3 Funds Images Denoising

At this stage, funds image denoising is performed using Average (Mean) Filter, as kernel size of (9*9) is specified. Mean Filter is an effective filter to enhance the quality of the funds images and to remove the types of noise described previously, Average filter (Mean filter) is used (described in the previous chapter). It is a linear type filtering approach. The image data is smoothed as a result of this. The effectiveness of each pixel mask is averaged together to create a pixel distinct from other pixels, and hence, it is named as an Average filter.

3.3.4 Funds Images Enhancement

At this stage, contrast limited adaptive histogram equalization (CLAHE) technique, as Clip Limit is equal to 2.0 and Tile Grid Size is equal to (8,8) are specified. CLAHE is an effective contrast enhancement method that effectively raises the image's contrast, removes blur and noise, to make the features more prominent, and displaying the image's details such as edge and boundary detection without changing the natural structure of the funds image.

3.3.5 funds Images Augmentation

Three types of DA techniques will be implemented in this thesis to address the issue of overfitting exposure and increase the efficiency of the proposed system for precise evaluation, and the network will be able to generalize to invisible data. The observed influence of four common DA techniques on the original dataset has been presented and these techniques are :

1. Horizontal Flip.
2. Zoom _range.
3. Rotation _range , and Fill _mode.

3.4 Features Extraction Phase

In order to extract the feature from the funds images by the proposed system , CNN uses self-learnable values of the filters used in the extraction of features. The extracted feature will serve as the starting point for the classification step. The extracted features in the proposed structure of a convolution Neural Network technique consists of three succeeded Conv blocks. The structure for extracting the features is presented in tables (3.1) and (3.2) from the first row to the seventh row, while the remaining rows of the table represent the classification phase.

Each block comprises from : a convolution layer with activation function Rectified Linear Unit (ReLU), Max Pooling layer. An explanation of these layers based on their type as follows:

1) The convolution Layer

Each of convolutional Neural Network has three convolution layers comprises a set of feature maps, the number and size of which are mentioned in the “output size” column in Table (3.1) and (3.2). Feature map is a collection of 3D matrix of units (neurons), with each neuron created by convoluting and summation one learnable filter of weights showing in the “Filter Size” column of Tables (3.1) and (3.2) with all feature maps in the previous layer or input image in the situation of layer1. Rectified Linear Unit (ReLU) is applied as the activation function after each convolution layer.

2) The Max-Pooling Layer

Max-pooling is placed after the convolution layers and ReLU, and these layers (convolution and ReLU, Max Pooling) , forming one in three blocks in the proposed structure design for CNN. Three blocks followed by dropout layer (the dropout layer is used with a small ratio of (0.2).The main task of Max-Pooling layer is to reduce the dimensions by reducing the dimension of the feature map to a quarter while preserving the most important detail.

3) The Fully Connected Layer and Vectorization

The output of the Max-Pooling layer in the third block is 3D matrix after it passes through the dropout layer with a ratio (0.2),then the output (feature map)passes to fully connected layer (flatten layer) to reshape by column scan to construct a vector of one dimension with the length of 32× 32× 64= 65536 neuron , which it passes as input to a fully Connected layer (Dense1).

The first steps (from step 1 to step 3) of the convolution Neural Network building Algorithm (3.1) represent the feature extraction stage. The proposed structure is shown in Tables (3.1) and (3.2), and they comprises of 10 consecutive layers divided into two stages based on their work, which are: the feature extraction phase and the classification phase.

The input and output sizes for the proposed CNN structure shown in Tables (3.1) and (3.2), as described in (row× column× filter no.), the filter is specified as (row× column× stride).

Table (3.1) the proposed CNN structure for binary classification.

Num	Layer name	Input Layer	Output Layer	Filter Size	Param #
1	Conv 2D + ReLU	(256, 256, 1)	(256, 256, 16)	(3×3), 1	160
2	Max-Pooling 2D	(256, 256, 16)	(128, 128, 16)	(2×2), 2	0
3	Conv 2D + ReLU	(128, 128, 16)	(128, 128, 32)	(3×3), 1	4640
4	Max-Pooling 2D	(128, 128, 32)	(64, 64, 32)	(2×2), 2	0
5	Conv 2D + ReLU	(64, 64, 32)	(64, 64, 64)	(3×3), 1	18496
6	Max-Pooling 2D	(64, 64, 64)	(32, 32, 64)	(2×2), 1	0
7	Dropout (0.2)	(32,32,64)	(32,32,64)	/	0
8	Flatten	(32, 32, 64)	65536	/	0
9	Dense1 + ReLU	65536	128	/	8388736
10	Dense2 + Soft-max	128	2	/	258
					8,412,290

Table (3.2) the proposed CNN structure for multi stage classification.

Num	Layer name	Input Layer	Output Layer	Filter Size	Param #
1	Conv 2D + ReLU	(256, 256, 1)	(256, 256, 16)	(3×3), 1	160
2	Max-Pooling 2D	(256, 256, 16)	(128, 128, 16)	(2×2), 2	0
3	Conv 2D + ReLU	(128, 128, 16)	(128, 128, 32)	(3×3), 1	4640
4	Max-Pooling 2D	(128, 128, 32)	(64, 64, 32)	(2×2), 2	0
5	Conv 2D + ReLU	(64, 64, 32)	(64, 64, 64)	(3×3), 1	18496
6	Max-Pooling 2D	(64, 64, 64)	(32, 32, 64)	(2×2), 1	0
7	Dropout (0.2)	(32,32,64)	(32,32,64)	/	0
8	Flatten	(32, 32, 64)	65536	/	0
9	Dense1 + ReLU	65536	128	/	8388736
10	Dense2 + Soft-max	128	3	/	387
					8,412,419

Algorithm (3.1) demonstrates the process of funds images feature extraction and classification using the proposed convolution Neural Network structure.

Algorithm (3.1): Building a CNN structure.

Input: ***X*** # funds images (gray scale (256×256) pixel).
K # For each convolution layer, a collection of kernel weights has been initialized.
W # For each last layer Fully Connected (FC), a collection of weights has been initialized.
M # Size of the feature map, (for first layer = 256 size of input image).
y # All funds' images labels (desired output). continue

Output: # CNN model.

Begin

Step 1: $XC = xc$ #The input funds images will be used as the input feature map, with layer L = 0.

$F_1=16$ #The number of kernels in the first layer is equal to 16 in this thesis, and this number increases in the subsequent layers (multiplied by 2), so the filters are (16, 32, 64), respectively.

Step 2: For **L** = **1** to 3 # L represents the **convolution layer** number.

For q = **0** to F^L # F^L represents the number of feature maps Layer (L)

For i = **0** to M # M represents the **dimension** of q feature map (M=256).

For j = **0** to M

Pass XC^{L-1} and K_L into equation (2.8) for compute zc^L # 2D Convolution L.q,i,j

Pass zc^L into equation (2.5) to compute yc^L # ReLU equation q,i,j

u = int (i/2), v = int (i/2) # indexes for max-pooling feature-map-size

Pass yc^L to equation (2.10) to compute xc^L # 2D Max Pooling L.q,(i+a),(i+b) q,u,v

Next j

Next i

Next q

$F^{l+1} = F^l \times 2$ # duplicated the number of kernels based on the CNN structure.

$M = M / 2$ # reduction the size parameter to be equal to the size of the Max pooled feature map.

Next L

Pass X c^L into equation (2.19) to compute **DC** # apply Dropout layer (0.2). q,u,v

Step 3: FL = Vector ({xc^L} q =1, 2, .., FL; i,j = 1, 2 ,.., M)

implement vectorization to convert (3D) Matrix to (1D) vector with the size of $F^l \times M \times M$ which it = 65536.

Step 4: Build Dense layer1 (128).

Step 5: Build Dense layer2 (2 or 3). # Save weights of features.

Step 6: Pass output of Dense layer 2 into the equation (2.6) for computing S. # Soft-max function

Step 7: Pass (S and Y) into the equation (2.7) to compute the value of the loss.

#The error of the convolution neural network prediction is represented by one value.

End.

3.5 Classification Phase

For the diagnosis of Cataract disease, features of funds images are utilized to correctly categorize patients based on whether they are infected or not. At tenth layer of CNN the out put will be (2 class) which represents (Cataract or Normal class). After that the Cataract class can be classified into three classes (Mild, Moderate, Sever) depending on the extracted attribute, and Softmax classifier is used in this study.

✧ Soft Max Classifier.

After the Flatten layer (step 3) of the convolutional Neural Network building algorithm (3.1), in step 4, the Dense1 architecture has a Fully Connected layer with an activation function (ReLU). Finally, achieve the diagnosis of Cataract in the output layer (Dense layer 2 in the step 5) is a fully connected layer. The output of the Dense layer 2 is passed into the Soft max activation function in step 6, which calculates the probability for every category depending on the input funds image. The probabilities from step 6 are transmitted into the loss function equation (2.7) to calculate the error value in step 7, which will be applied to adjust (update) the weights through the back propagation process of training the proposed architecture.

In the training phase of CNN, funds image classification uses the Adam optimization algorithm, loss function (sparse categorical cross entropy) to evaluate the network, learning rate reduction to improve the value of the validation loss when model performance stops increasing (reduce on the plateau), and early stopping to find the best number of epochs in the training of CNN. The train is performed on the augmented data; the train algorithm begins with prepossessing as described in algorithm (3.3) that illustrates the steps of training CNN.

The following step is to pass the funds images on the convolution neural network structure in the forward direction and the backward direction within a number of epochs depending on an early stopping to decrease the error between the predicted output of the

CNN model and the actual label of the training sample, and adjust the weights to be used in the testing phase.

The output of a convolution neural network training algorithm is a trained set of weights and kernels for all layers of the network architecture. These trained weights and kernels are stored for use by the network later in the testing process.

Algorithm (3.2): convolution Neural Network training phase

Input: ***X*** # all funds images in database (augmented funds images).
Y # all funds images labels (desired output)
epoch # a number of iterations required for training, 30 epochs defined but executed according to the use of early stopping.
μ = 0.001 # learning rate .

Output: Trained sets of kernels(***K)*** and weights (***W)*** . #Trained sets of kernels values (***K)*** for all " convolution layers" , and trained sets of weights (***W)*** for all " fully connected layers" .

Begin

Step 1: For ***epoch*** = 1 to ***30***

For k =1 to num # (number of the funds images in the Training sample).

 # In the Forward direction

Pass X, Y, K, W into algorithm (3.1) to produce loss.

Kernel (K) in convolution layers, and weight (W) in fully connected layers.

 # In the Backward direction

Update Wl using equation (2.12) and equation (2.14) #last layer of FC- loss ,w.

Update Wh using equation (2.13) and equation (2.14)#hidden layer of FC-erro,w

Update K using equation (2.16) and equation (2.17) # in convolution layer.

If the validation loss is not reduced for three epochs, **then**. continue

$\mu = \mu \times 0.1$ #minimum = 0.0000

If the validation loss is not reduced by 0.005 for fifteen epochs, **then**

early stopping

Next k

Next epoch

END.

In the test phase of convolution neural network CNN, the test is performed on the invisible test data, the test algorithm begins with prepossessing steps. The next step is to pass the funds images on the convolution neural network structure in the forward direction only to extract the features and then classify these images into (Cataract, Normal)by using the trained weights in the fully connected layers and the trained kernel in the convolution layers that were stored in the training phase and applied later in the test phase.

Algorithm (3.3) describes the implementation of CNN in testing stage.

Algorithm (3.3): System testing phase
Input: ***X*** # all funds images in the testing sample database. ***W*** # trained sets of weights in fully connected layers. ***K*** # trained sets of kernel weights in convolution layers.
Output: Prediction of funds image class # classified into (Cataract, Normal) and cataract stages.
Begin **Step 1:** Pass X, W, K into algorithm (3.1) to Create predication model. # K is the trained kernel in convolution layers, and W is the trained weight in Fully Connected layers. **Step 2:** Return the class of the funds image. **End.**

Chapter Four

Experimental Results and Evaluation

4.1 Introduction

The previous chapter described the steps of the Cataract classification system. The system's performance results will be discussed in this chapter, which is divided into two phases: training and testing. There are three steps in each phase: results of the prepossessing steps, results of the proposed convolution Neural Network structure steps (features extraction steps), and results of the classification steps (CNN-Soft max). The major subjects that will also be covered in this section are the fundus image database and their division, detailed outputs with descriptions for each step in the proposed system, and the evaluation of the proposed system by calculating performance using the measurement tools. These results played an important role in supporting this thesis which was done using the python programming language version 3.8.3.

4.2 Hardware and Software Requirements

The proposed Cataract Detection System is implemented by using personal computer Think Pad (Intel (R) Core i7- 8565U @ 1.80 GHz for CPU, 8 GB windows 10 of RAM, and 64-bit OS), and also Google Co-lab (Collaborators) is an open Jupiter notebook environment that works completely in the cloud.

This system is implemented by using free GPU s from Google Co-lab. It is a Google application that enables programmers to run Python code directly from their web browser. It is an ideal platform for deep learning (DL) tasks as it allows us to train a large database and build complex models very quickly on Google computing resources (GPU).

TensorFlow open-source framework (a Google open-source programming library that is concentrated on supporting tensors to produce successful work), K eras (a neural network library based on TensorFlow written in Python, and also open-source), and Python are used for the implementation of the CNN code.

The system relies on open-source libraries such as Open, Sci-kit Learn, and Pandas. These libraries are dedicated to deal with machine learning and data analysis.

4.3 Fundus Images Dataset Preparations

In this thesis, fundus images were taken from two datasets are: Kaggle and ODIR.

1. **Kaggle dataset:** is a collection of fundus images of retina which examinations by professional clinicians. This dataset contains about 400 colored fundus images of both left and right eyes. Kaggle dataset contains (300 normal) and (100 cataract) images for both healthy and cataract images. Kaggle dataset will be divided into two samples: the training data sample is 70% of total dataset (280 fundus images), and the testing data sample is 30% (120 fundus image). Table (4.1) shows the Statistics of the fundus image of Kaggle dataset.

Table (4.1) Statistics of fundus images divide of Kaggle.

	Cataract	Normal	Total
Train	70	210	280
Test	30	90	120
Total	**100**	**300**	**400**

The fundus image dataset described in table (4.1) is not sufficient for model development because systems that use deep learning need sufficient data for the training sample. Therefore, we will apply the Data Augmentation technique on this dataset. DA is a powerful and important way to train any algorithm, as well as the network's ability to generalize to invisible data. Therefore, the total dataset will be after using the data augmentation technique 1200 (840 to train, 360 to test) because two new variants labeled with the same original label were used. Table (4.2) shows the Statistics of the fundus image of Kaggle dataset after

Augmentation.

Table (4.2) Statistics of fundus images divide for Kaggle dataset after Augmentation.

	cataract	Normal	Total
Train	210	630	840
Test	90	270	360
Total	300	900	1200

2. The ODIR dataset stands for Ocular Disease Intelligent Recognition: is a well-organized database of ophthalmology contain 5,000 Color fundus images of the left and right eyes in patients of various ages , as well as diagnostic terms taken from doctors. ODIR dataset has information about many patients, it represents a set of "real-life" which gathered by Shandong Medical Technology Co., Ltd. from various Chinese medical facilities and hospitals. Fundus images are recorded in these institutions using on the market different cameras , such as Zeiss, Kowa, and Canon, as a result of which image resolutions vary. Trained human readers labeled the annotations using management of quality control. The patients are categorized into many (eight) categories as:

- Hypertension (H)
- Glaucoma (G)
- **Cataract (C)**
- Age-related Macular Degeneration (A)
- Diabetes (D)
- Pathological Myopia (M)
- Other diseases/abnormalities (O)
- **Normal (N)**

This investigation focuses on a specific type of eye illness, Cataract. This database contains 512 images (212 Cataract images and 300 Normal images).

ODIR dataset is divided into two samples: the training data sample is 70% of total dataset (358 fundus images), and the testing data sample is 30% (154 fundus image). Table (4.3) shows the Statistics of the fundus image divide of ODIR dataset.

Table (4.3) Statistics of fundus images divide of ODIR.

	Cataract	Normal	Total
Train	148	210	358
Test	64	90	154
Total	**212**	**300**	**512**

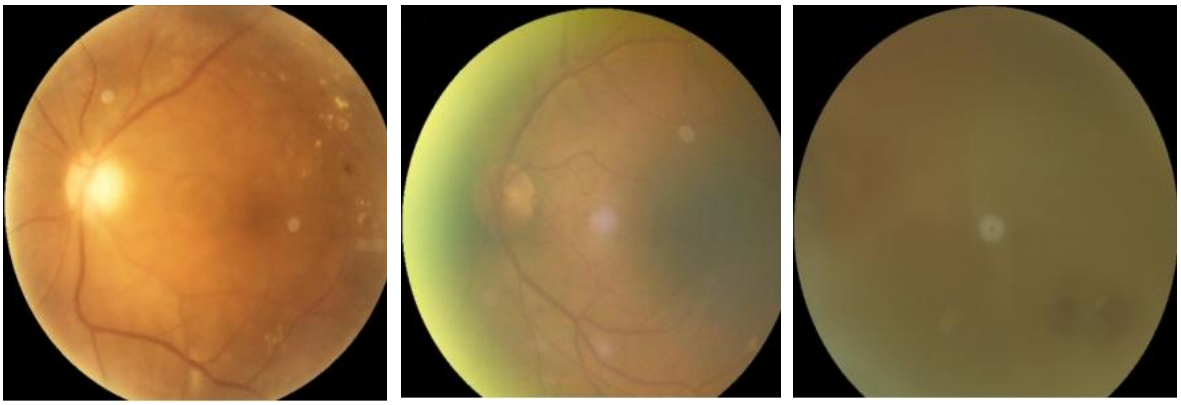
The fundus image dataset described in table (4.3) is not sufficient for model development. So, we will apply the Data Augmentation technique of dataset.

Therefore, the total dataset will be after using the data augmentation technique is 1024 (716 to train, 308 to test) because **one** new variant labeled with the same original label was used. Table (4.4) shows the Statistics of the fundus image divide of ODIR dataset after Augmentation.

Table (4.4) Statistics of fundus images divide for ODIR dataset after Augmentation.

	cataract	Normal	Total
Train	296	420	716
Test	128	180	308
Total	**424**	**600**	**1024**

Figure (4.1) shows original images of Cataract and non-cataract before applying the Data Augmentation .



(a) Samples of the Cataract dataset.

(b) Samples of the Non -Cataract dataset.

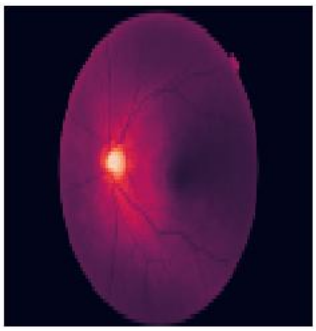
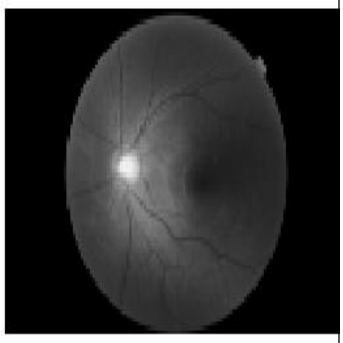
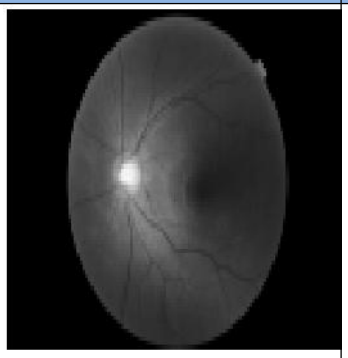
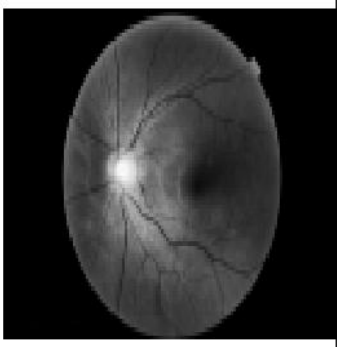
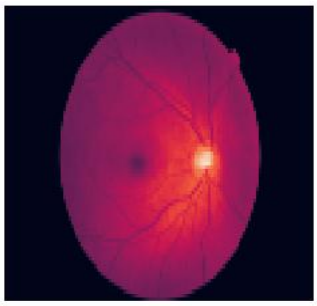
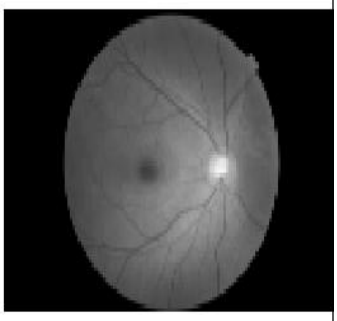
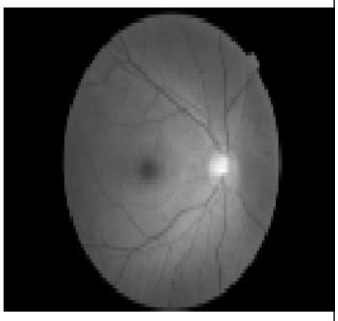
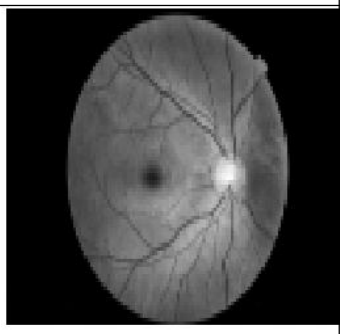
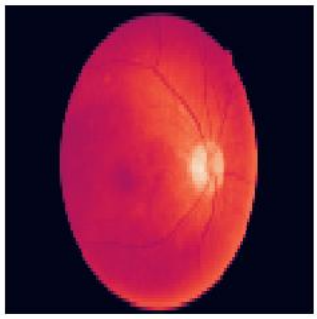
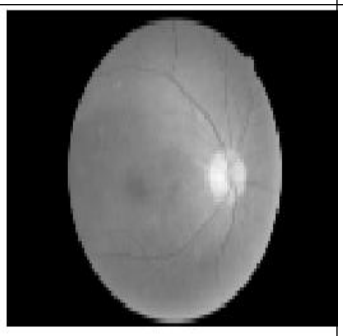
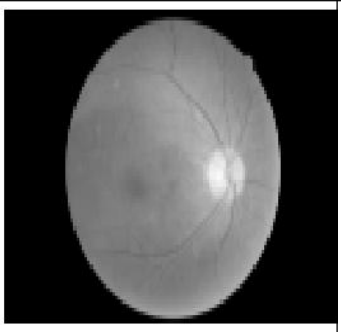
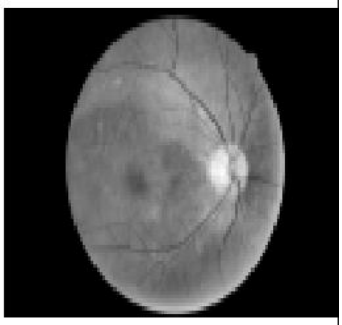
Figure (4.1) some samples of the ODIR dataset (the original images).

4.4 Results of Applying Prepossessing on Fundus Images Datasets

The prepossessing consists of three main steps. The results of applying prepossessing on fundus image databases are described in Table (4.5) based on the following main steps:

a) Transforming fundus image into a gray level format of (8-bit).
b) Remove noise by using Mean Filter.
c) Fundus images enhancement using CLAHE technique.

Table (4.5) The fundus images after applying prepossessing operations.

Original image	Gray-Scale Image	Remove noise (Mean Filter)	Image Enhancement
			
			
			

4. Fundus Image Augmentation

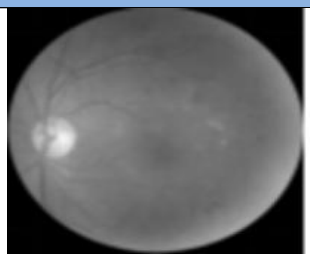
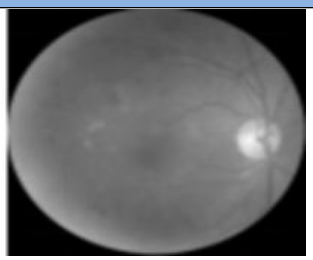
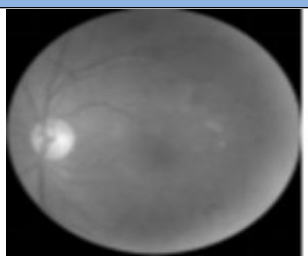
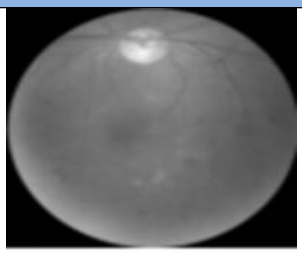
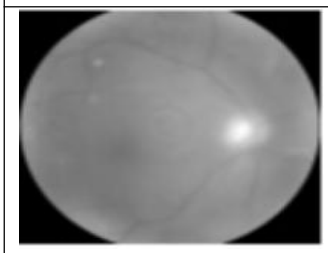
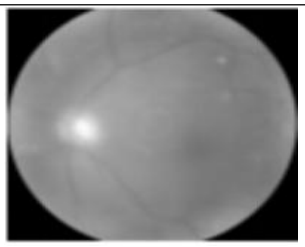
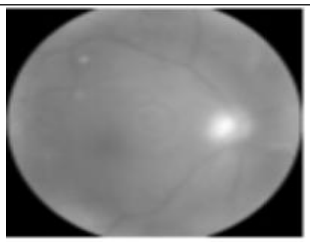
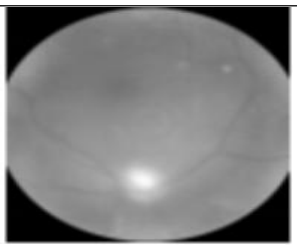
In table (4.6), the observed effect of four common data augmentation techniques applied on fundus image dataset was presented as follows:

1. **Horizontal Flip:** The fundus images will be reversed horizontally by giving the boolean statement its true value in the horizontal flip parameter as illustrated in

the second column of table (4.6).

2. **Zoom Range**: The zoom parameter is applied to determine the value of zoom size (0.01). The image size will decrease by 0.01, as shown in the third column of table (4.6).
3. **Rotation range:** a small angle (10 degrees) is specified so that the shape of the fundus image will not be affected by the rotation range variable as shown in the fourth column of table (4.6).
4. **Fill mode:** used a constant value with the **cval** argument (**cval** = 255 pixels) to fill all the pixels extending outside the input's borders in DA techniques (Zoom range, Rotation range).

Table (4.6): Applying DA techniques on ogininal fundus images .

The Original fundus image.	Horizontally Flip.	Zoom Range (0.01)	Rotation of 10 degrees
			
			

4.5 Results of Proposed System on the Original Datasets

This section describes the detailed output of the steps of the Cataract proposed system. The outputs are organized into two phases: the outputs of the training phase, the outputs of the test phase.

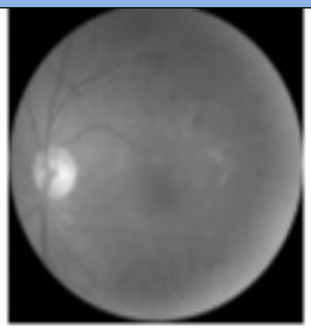
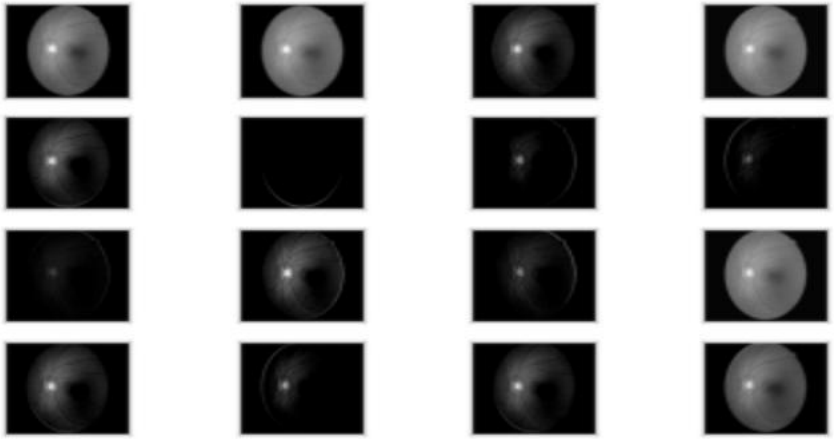
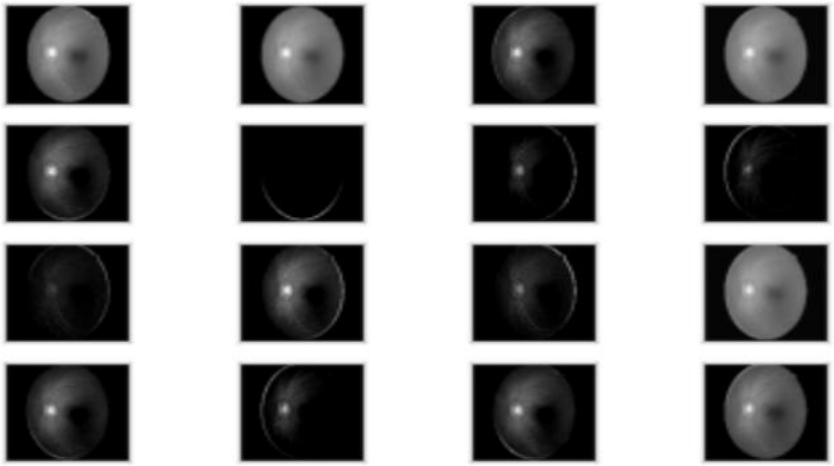
4.5.1 Training Phase Outputs before Augmentation

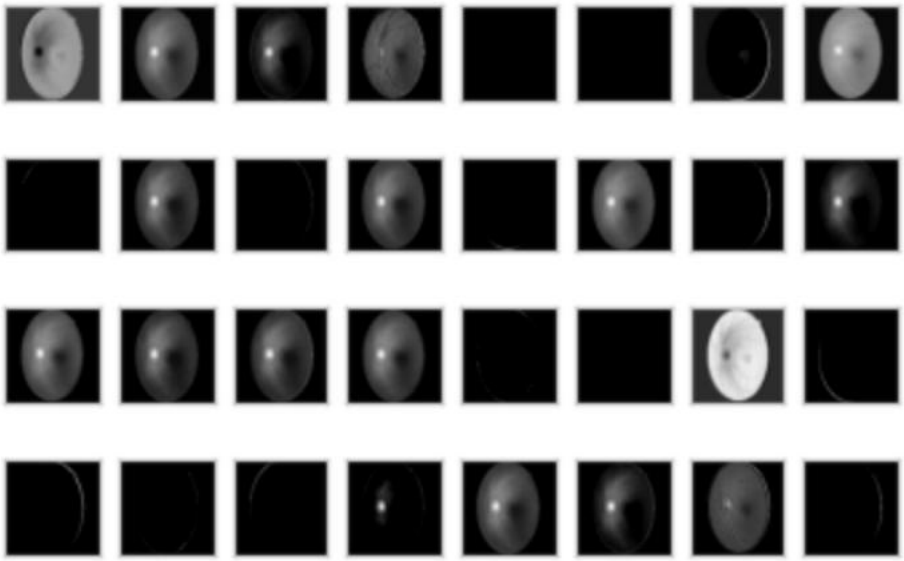
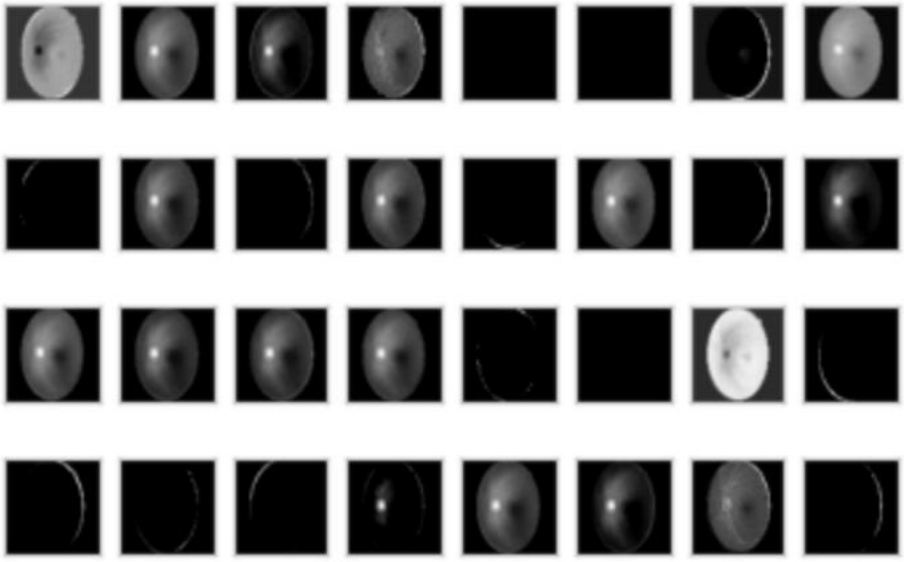
In the training phase, all training samples (280 of Kaggle) and (358 of ODIR) fundus images with their labels must pass through the system simultaneously to be trained. In the following ,a result and output fundus images of features extraction steps, and classification steps.

4.5.1.1 Results of the Features Extraction Steps

The feature extraction steps are performed using CNN that contains of three succeeded blocks, and each block comprises of : convolution layer with "Rectified Linear Unit (Relu)" activation function, and Max-Pooling layer. Table (4.7) illustrates the output of each layer in the three-block layers.

Table (4.7) the outputs of the feature maps for the Cataract sample.

Block / Layer	Feature Maps of the Fundus Images
 **Block 1** Convolution layer 1 (256×256 ×16)	
Max pooling layer 1 (128×128×16)	

Block 2 Convolution layer 2 (128×128×32)	
Max pooling layer 2 (64×64×32)	

Block 3 Convolution layer 3 (64×6432×64)	
Max pooling layer 3 (32×3232×64)	

The convolution layer in CNN model, as shown in table (4.7), extracts the features of the Cataract fundus images by passing many different filters such as edges detection, vertical, horizontal and diagonal lines, curves, corners, color contrast sensitive filters and so on over fundus image. Each of these filters produces a feature map that indicates spatial patterns where the feature is located. Then, feature map passes through the activation function (ReLU) non-linear.

The size of each filter is (3×3), and the number of filters in each layer is (16, 32, 64), respectively, as shown in the table (3, 1) in the chapter three.

The Max Pooling layer reduces computational power by reducing the dimensions of features map to half because the kernel size is equal to (2×2), as shown in the table (4.7). This contributes to reducing the over fitting problem.

Then The dropout layer used with a ratio (0.2) to regulate and prevent the network from over fitting that happens all that during training process of the model.

4.5.1.2 Results of the Classification Steps before Augmentation

The classifier CNN-Soft Max is used in this study to detect Cataract disease :

- **Results of fundus image classification using Soft max function**

Soft-Max is a method (classifier) for classifying fundus images, and it consists of the remaining of three layers of convolution neural network architecture CNN:

1. Flatten layer: converting the 3D feature maps (three dimensions) such that the result is a 1D vector (one dimension) 32 ×32×64 = 65536.
2. Fully Connected layer of the Dense1 architecture with an activation function (ReLU), and there are (128) units.
3. The output layer is a Dense 2 of a fully connected layer with number of units (2). The output of this layer is passed into the Soft max activation function to compute probabilities for each class according to the input image. The probabilities are passed to the loss function (sparse categorical cross entropy) to calculate the error value which will be applied to adjust (update) the weights through the back propagation process.
4. The classification of input fundus image into two classes (Normal, Cataract) by using Soft max activation function.

In the training phase, fundus image classification uses the Adam optimization, learning rate reduction to improve the value of the validation loss when model performance stops increasing (reduce on the plateau), batch size=64, and early stopping to find the best number of epochs (maximum= 30). The learning curves of the accuracy and the loss for (training sample and validation sample) in the figures (4.2) and (4.3) illustrate the over fitting condition when the model is trained on the original ODIR and Kaggle datasets sequentially (without Data Augmentation) .Where there is a large gap that can be noticed between the values of training accuracy and validation accuracy and also between the values of loss. If the value of accuracy for training or validation did not change and stabilized at a certain value after 15 epochs then the training was stopped.

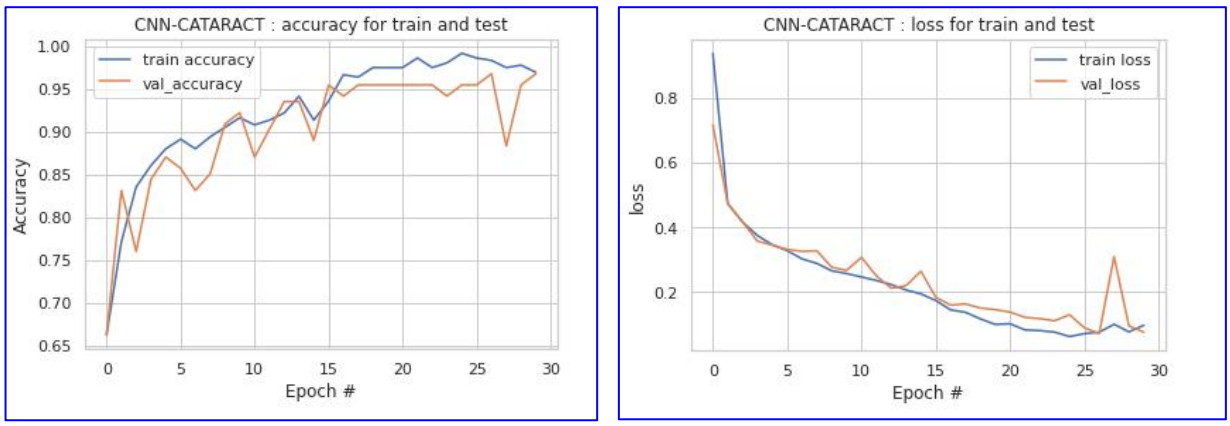


Figure (4.2) the **accuracy** and the **loss** learning curves of train sample and validation sample (the **original data of ODIR** dataset).

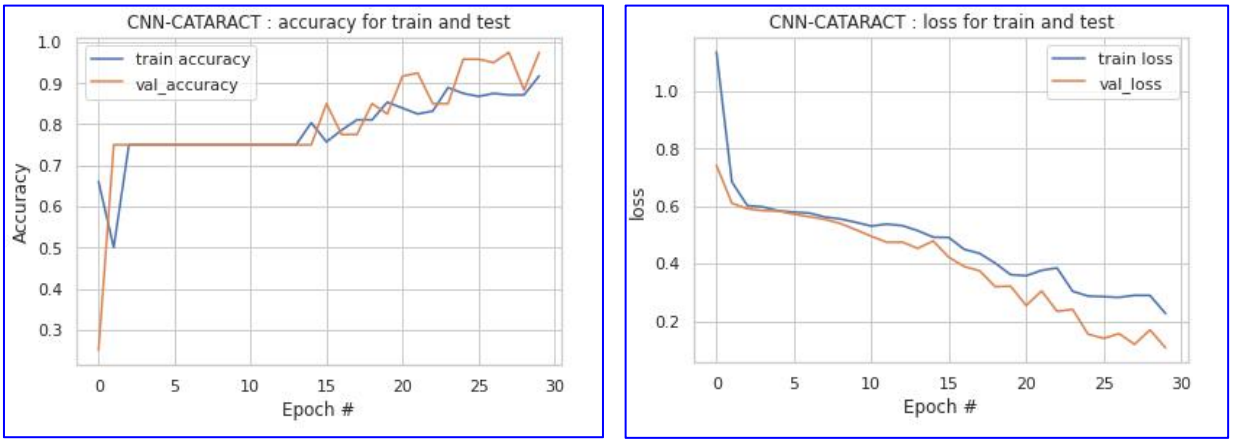


Figure (4.3) the **accuracy** and the **loss** learning curves of train sample and Validation sample (the **original data of Kaggle** dataset).

4.5.2 Testing Phase Results before Augmentation

In the testing phase, all test samples that are 30% of the database (120 fundus images of Kaggle) and (154 fundus images of ODIR) will be passed to the system without their labels to be tested. The features are extracted in the testing phase in the same way as in the training phase, but only in the forward direction. The test images pass over the three blocks dedicated to extracting features in the structure of convolution neural networks and then classifying these images into (Normal and Cataract) by using the trained weights in the fully connected layers and the trained kernel in the convolution layers that were stored in the training phase to be applied in the test phase.

In the classification stage, the results of the proposed system output will be presented using classification method applied to the test sample: SoftMax function to detect cases infected with Cataract.

✧ Result of fundus Image Classification using Soft-max Function

After the feature extraction stage, the extracted features are passed to the layers for the classification that was explained in section (4.5.1.2) to labeled a fundus image as Normal or Cataract status.

Figure (4.4) and Figure (4.5) represent the confusion matrix using "Soft Max function" classifier when the model is trained on the original data of ODIR dataset and Kaggle datasets, where True Positive (TP) is a right classification of positive prognosis such as Cataract categorized as Cataract. “False Positive”(FP) is a wrong classification of positive diagnosis such as Cataract categorized as non-Cataract. "True Negative" (TN) is the correct designation of a negative diagnostic such as Cataract categorized as Cataract. “False Negative” (FN) isa wrong classification of negative diagnosis such as non-Cataract categorized as Cataract.

Table (4.8) illustrates the performance measures results when the system is trained on the original data of ODIR and Kaggle dataset at Sensitivity, Area

Under Curve (AUC), F1-score, Accuracy, and Specificity are the key metrics.

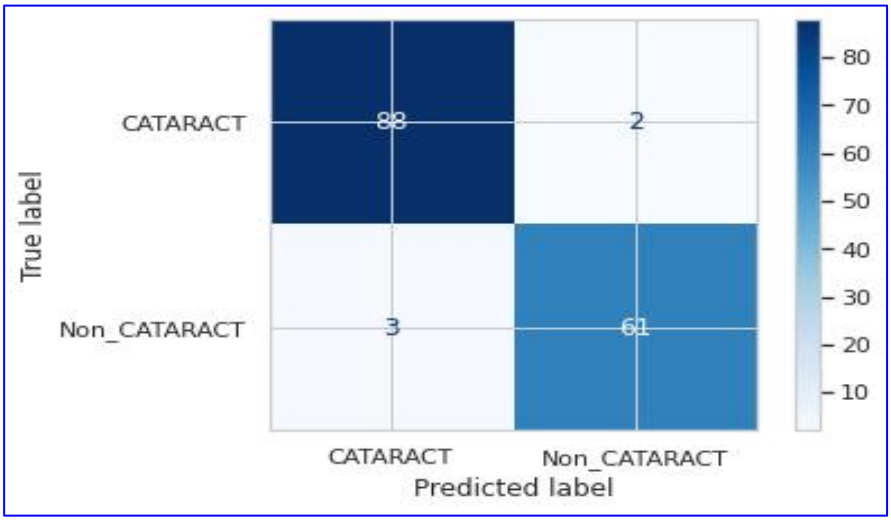


Figure (4.4) the confusion matrix of the test sample of original ODIR dataset

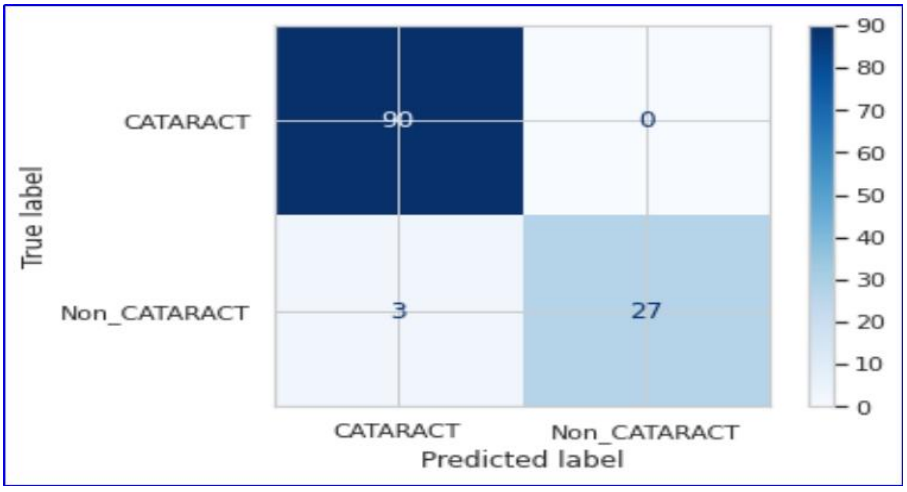


Figure (4.5) the confusion matrix of the test sample of original Kaggle dataset

Table (4.8) Performance Measures results of Original ODIR and Kaggle datasets.

Data type	Sensitivity	Specificity	F1-score	AUC	Accuracy
Kaggle	**100%**	**90.00 %**	**98. 36 %**	**95. 00 %**	**97.5 %**
ODIR	**97.78 %**	**95. 31 %**	**97. 23 %**	**96. 55 %**	**96.7 %**

4.6 Result of the Proposed System After Data Augmentation

This section explains how the proposed system performed after using the Data Augmentation technique. It described the detailed output of the steps of the Cataract proposed system. The outputs are organized into two phases: the outputs of the training phase, the outputs of the test phase.

4.6.1 Training Phase Output

In the training phase, all training samples after Augmentation (840 of Kaggle) and (716 of ODIR) fundus images with their labels must pass through the system simultaneously to be trained. In the following a resultand output fundus images of features extraction steps, and classification steps.

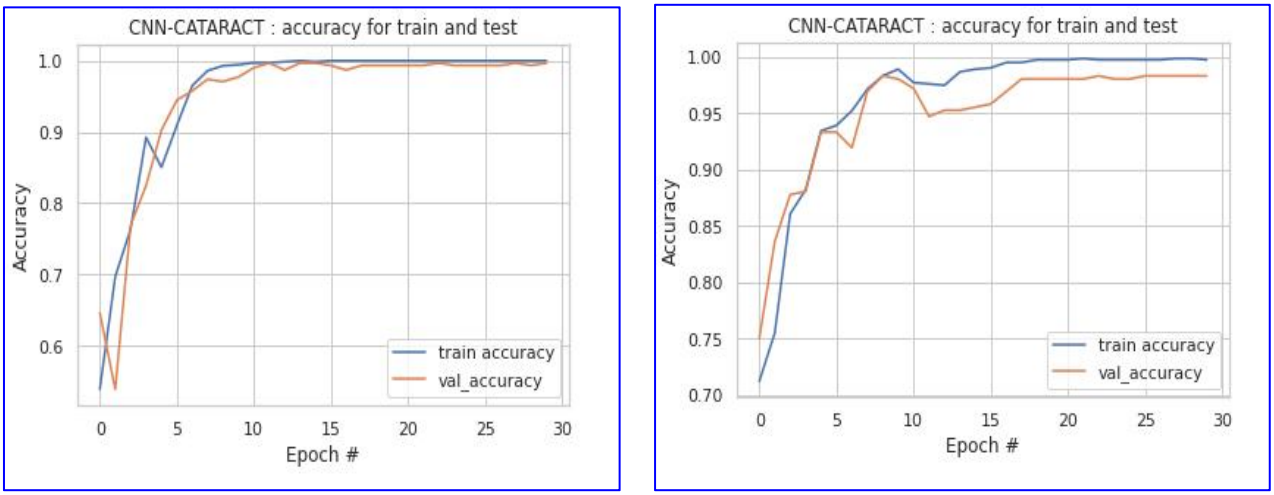
4.6.1.1 Results of the Features Extraction Steps

The feature extraction steps are performed using CNN that comprises of three succeeded Conv blocks, and each block comprises of : convolution layer with "Rectified Linear Unit (ReLU)" activation function, Max-Pooling layer. This phase is similar to results of the features extraction steps on original dataset explained in the section 4.5.1.1, the difference is that now passed Augmented data to CNN for feature extraction and classification.

4.6.1.2 Results of the Classification Steps

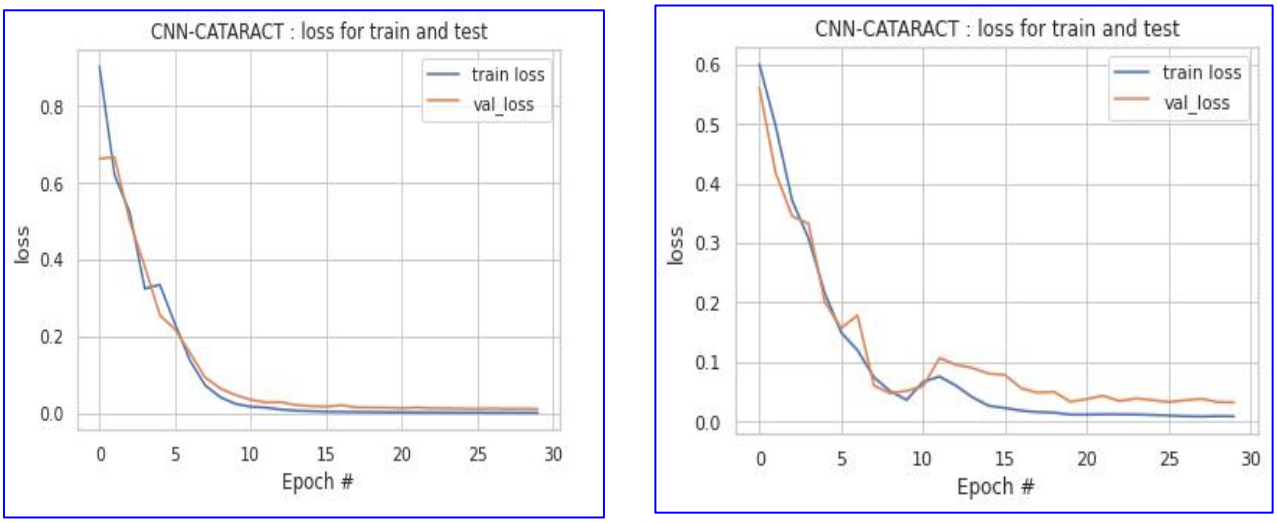
The classifier CNN-Soft Max is used in this study to classify augmented fundus image into (Cataract and normal) cases. In the training phase, Augmented fundus image classification uses the Adam optimization, learning rate reduction to improve the value of the validation loss when model performance stops increasing (reduce on the plateau), batch size = 64, and early stopping to find thebest number of epochs

(maximum=30). The train is performed on the augmented data. Figure (4.6) illustrates the evaluation of CNN performance in accuracy score is the total amount of errors the model predicted of augmented ODIR and Kaggle datasets respectively. If the value of accuracy for training or validation did not change and stabilized at a certain value after 15 epochs then the training was stopped.



(a) **(b)**

Figure (4.6) Train and Validation accuracy:(a) Augmented ODIR, (b)Augmented Kaggle



(a) **(b)**

Figure (4.7) Train and Validation Loss: (a) Augmented ODIR, (b) Augmented Kaggle dataset.

The curves of learning (loss of training and validation) plot in figures (4.7) illustrates the good fit condition when the model is trained on the augmented data of ODIR and Kaggle datasets respectively, because of training loss and the loss of approval (loss of validation) depreciate to a point of balance, and the gap is very little between the two-loss values. It is potentially persistent training of a good fit that would produce a problem of over fitting so an early stopping was used when training the model.

4.6.2 Testing Phase Results

In the testing phase, all test samples that are 30% of the database (360 fundus images of Kaggle) and (306 fundus images of ODIR) will be passed to the system without their labels to be tested.

The features are extracted in the testing phase in the same way as in the training phase, but only in the forward direction, where the test images pass over the three blocks dedicated to extracting features in the structure of convolution neural networks and then classifying these images into (Normal and Cataract) by using the trained weights in the fully connected layers and the trained kernel in the convolution layers that were stored in the training phase to be applied in the test phase.

In the classification stage, the results of the proposed system output will be presented using classification method applied to the test sample: Soft Max function to detect cases infected with Cataract.

✧ Result of fundus Image Classification Using Softmax Function.

Figure (4.8) and (4.9) represent the confusion matrix using "Soft Max function" classifier when the model is trained on the augmented ODIR and Kaggle datasets.

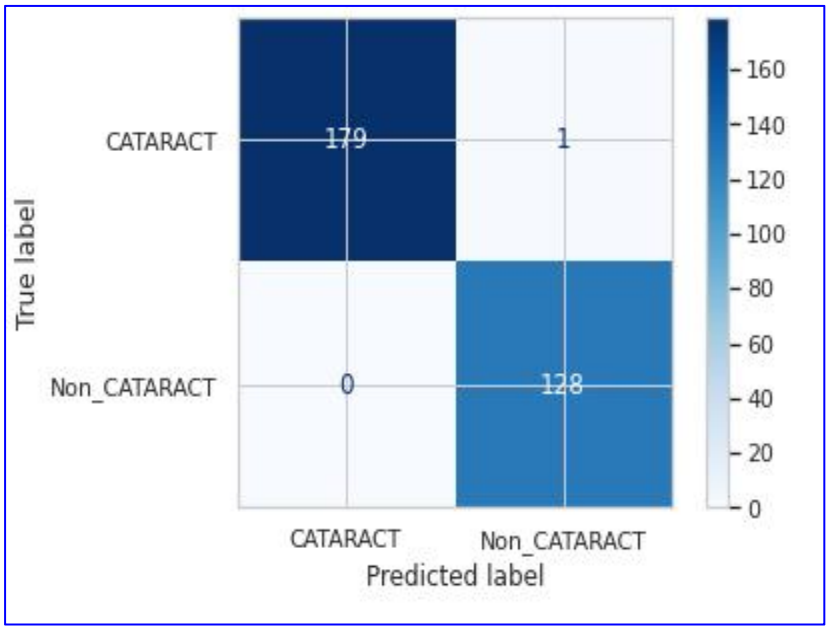


Figure (4.8) the test sample's confusion matrix of Augmented ODIR dataset

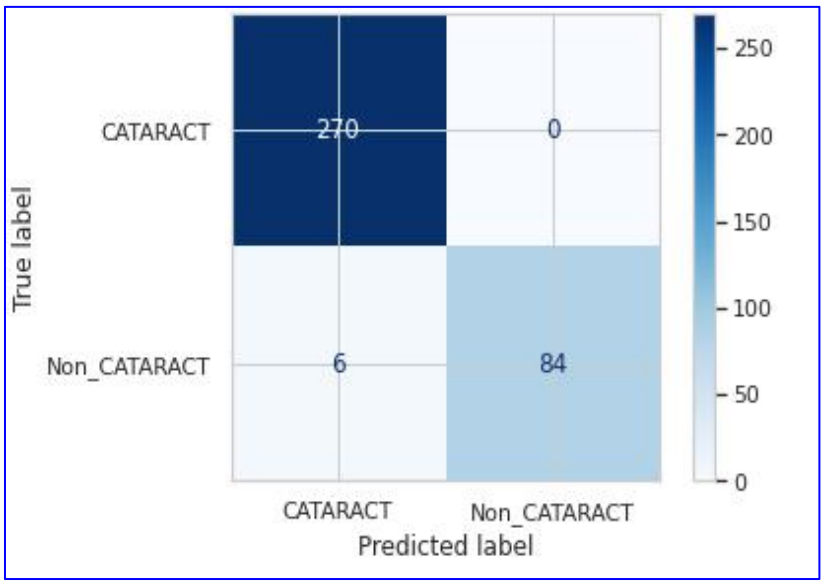


Figure (4.9) The test sample's confusion matrix of Augmented Kaggle dataset

Table (4.9) illustrates the performance measures results when the system is trained on the augmented data of ODIR and Kaggle dataset at Sensitivity, Area Under Curve (AUC), F1-score, Accuracy, and Specificity are the key metrics.

Table (4.9) performance measures results of Augmented ODIR and Kaggle datasets.

Data type	Sensitivity	Specificity	F1-score	AUC	Accuracy
Kaggle	100 %	93. 33 %	98.90%	96.67 %	98.5 %
ODIR	99. 44 %	100 %	100. 72 %	99. 72 %	99.6 %

4.7 Results after Dataset Augmentation

In this section there is a comparison between the System training and performance on the original dataset, and dataset after using Data Augmentation technique. Table (4.10) shows the dataset name, accuracy, and loss value when training the system on the original dataset, and when training it on the augmented dataset. Obviously, the training performance for the Augmented dataset is quite better than the original dataset.

Table (4.10) Model training performance on original and augmented data

Dataset		Accuracy	Loss
Kaggle	Original data	97. 5 %	0.1068
	Augmented data	98. 5 %	0.0322
ODIR	Original data	96. 7%	0.0754
	Augmented data	99. 6 %	0.0098

Table (4.11) shows the evaluation criteria when training the CNN model on both the Original and Augmentation datasets, by using the metrics of Sensitivity, Accuracy, F1-score, Specificity, and (AUC) Area Under Curve.

Table (4.11) Evaluation criteria for model testing on Original and Augment dataset.

Dataset		Accuracy	Sensitivity	Specificity	F1-score	AUC
Kaggle	Original data	97.5 %	100%	90 %	98.36%	95%
	Augmented data	98.5 %	100 %	94. 33%	98.90%	96.67%
ODIR	Original data	96.7%	97.78 %	95.31%	97.23 %	96.55%
	Augmented data	99.6%	99.44 %	100%	99. 72%	99.72%

4.8 Comparison of the Proposed System with Related Works

Table (4.12) incorporate a comparison of several deep learning-based Cataract diagnostic approaches with the suggested system performance that some of them utilizes the same datasets and others are not. It should be noted that the introduced system outperformed the other systems (existing systems).

Table (4.12) Comparison of the proposed System with another research

Authors	Dataset	Accuracy	AUC	Specificity	Sensitivity
Masum S. Junayed, et al **2021[5]**	Kaggle	98.00%	97.00%	99.00%	96.00%
Ely Sudarsono. et al **2020 [12]**	Kaggle	97.00%	-	-	-
Mas A. Syarifa, et al **2020 [13]**	Kaggle	97.00%	-	-	-
Proposed method	**Kaggle**	**98.5%**	**96.67%**	**94.33%**	**100%**
Proposed method	**ODIR**	**99.6 %**	**99.72%**	**100%**	**99.44%**

4.9 Cataract Stages Detection

Proposed system firstly can classify fundus images into two classes (Normal, Cataract). After that the system take Cataract fundus images of ODIR dataset to classify them into three stages of cataract based on their severity by sequence of operations. ODIR database has fundus images of Cataract disease (212 images) with three stages. It includes (51 Mild, 61 Moderate, and 100 Sever stage). This dataset is split in to two categories: 70% training data and 30% test data. Table (4.13) shows the number and percentage of each class (Mild, Moderate, Sever) in each part of the dataset (train and test datasets).

Table (4.13) The number of cataract images in training and testing ODIR dataset.

	Mild	Moderate	Sever	Total
Train	35	42	70	147
Test	16	19	30	**65**
Total	**51**	**61**	**100**	**212**

The fundus images dataset described in table (4.13) is not sufficient for model development because systems that use deep learning need sufficient data for the training sample, so the Data Augmentation technique will apply on the dataset. DA is a powerful and important way for training many algorithms. One of DA technique will be implemented in this project (DA techniques were described in Section (2.4).

Table (4.14) The number of Cataract images in training and testing ODIR dataset after Augmentation

	Mild	Moderate	Sever	Total
Train	70	84	140	294
Test	32	38	60	130
Total	**102**	**122**	**200**	**424**

As the same of binary classification, after the feature extraction stage, the extracted features are passed to the layers for the classification which labeled a fundus image into 3 classes as (Mild, Moderate, Sever) status.

Network design (shown in Table (3.2)) has some change as "Softmax function" used with three parameters instead of two then the number of parameters will be larger , and training process is developed in order to get better result. Figure (4.10) shows the training process of the system during the different epochs compared with accuracy.

While Figure (4.11) shows the training process of the system during the different epochs compared with loss function value.

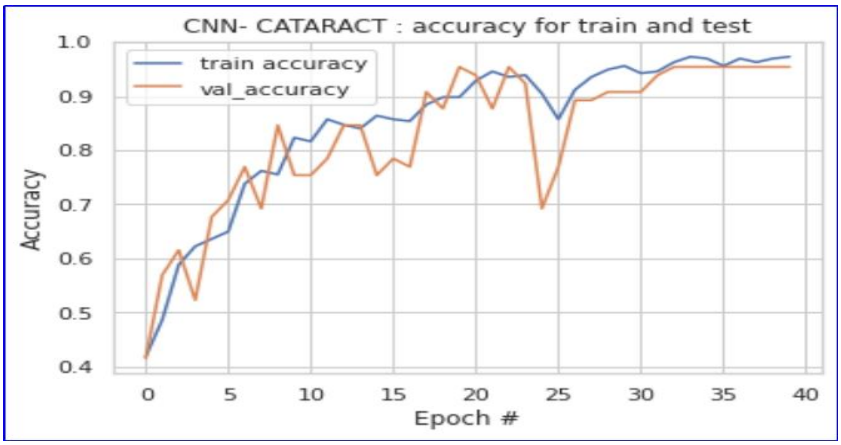


Figure (4.10) The accuracy of train and validation for cataract stages classification

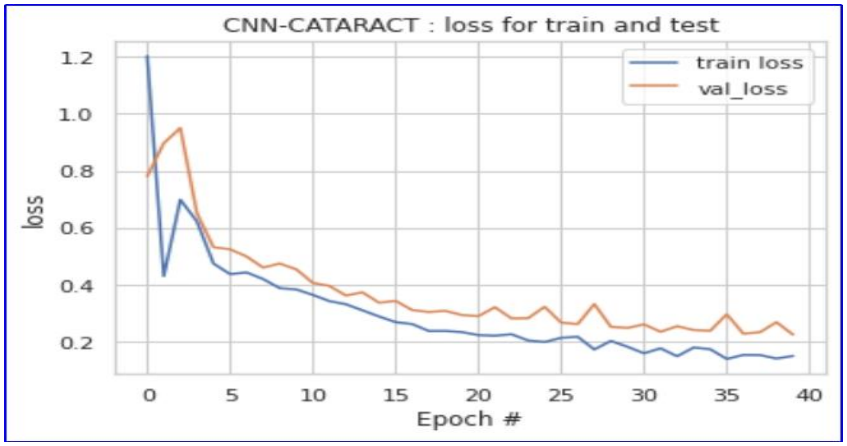


Figure (4.11) The loss of train and validation for cataract stages classification

Figure (4.12) represents the confusion matrix using the "Soft Max function" classifier when the model is trained on the ODIR dataset after Augmentation.

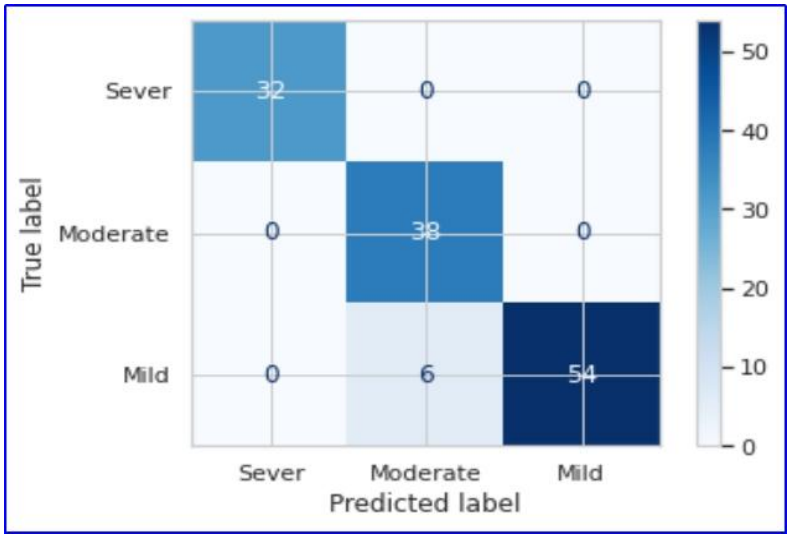


Figure (4.12) the confusion matrix of the test sample of ODIR dataset after Augmentation .

Table (4.15) illustrates the performance measures results when the system is trained on the Augmented data of ODIR dataset in key metric of Specificity, Area Under Curve (AUC), F1-score, Accuracy, and Sensitivity when proposed system classified fundus images into 3 classes.

Table (4.15) Performance measures results of ODIR dataset After Augmentation.

Dataset	Sensitivity	Specificity	F1-score	AUC	Accuracy
ODIR	91.2 %	100%	95.5 %	95.6 %	96.9 %

Table (4.16) includes a comparison of different deep learning-based Cataract diagnostic approaches with the suggested system performance that some of them utilizes many datasets (the same datasets) and the other are not, all of them classify cataract image into three stages (Mild, Moderate, Sever). It should be observed that the suggested system reached higher performance when compared

to other systems .

Table (4.16) comparison of the proposed System with another research(multi class of cataract)

Authors	Dataset	Accuracy	AUC	Specificity	Sensitivity
Ram and Reyes et al., 2020 [14]	**ODIR**	81.9%	-	66%	71.4%
Islam et al., 2019 [17]	**ODIR**	87.6%	80.5%	-	-
Jing et al.., 2020 [15]	**ODIR**	89%	73%		
Proposed method	**ODIR**	**96.9%**	**95.6%**	**100%**	**91.2%**

The main reason for obtaining these achieved results:

Firstly, the structure of the CNN architecture shown in Tables (3.1) and (3.2), in which identical blocks of layers were stacked to capture discriminative features and the dropout layer was used after Max Pooling layer of the third block to improve accuracy and generalization.

Secondly, the overfitting problem was prevented through the use of data augmentation technology and early stopping during a train of CNN and Dropout layer, which is the problem that most CNN models suffer.

Thirdly, the network was trained on preprocess fundus images and this gives higher accuracy results because it focused on retina where the disease is present.

Fourthly, the parameters number which is derived from the model (8,412,290) for binary classification and (8,412,419) for multi class of cataract , this is small in comparison to the number of parameters in other literature models.

4.10 An Experimental Database (Case Study)

In order to further validate the results of this study, real life fundus images of Cataract are tested on our proposed system. Real patients' data are collected from Imam Al- Sadiq General Hospital in Babylon city, in collaboration with the medical doctors and technicians. While collecting these images the privacy of the patients is preserved. The total number of 60 cataract fundus images are collected from 47 cases (patients) with size (512, 512), which include three stages of cataract. Table (4.17) shows cataract fundus images stages with their numbers which are collected for case study.

Table (4.17) shows the cataract fundus images divide for case study.

Cataract Stages	Mild	Moderate	Sever	Total
Images Numbers	10	10	40	60

The fundus images are tested by our system. Figure (4.13) shows the confusion matrix of the case study. The performance is almost perfect, with only two fundus image is miss classified.

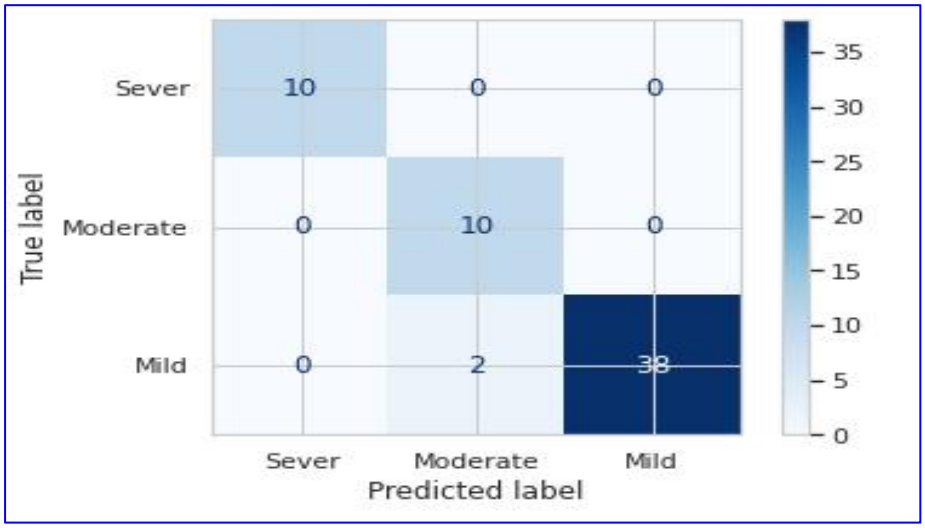


Figure (4.13) The confusion matrix of the selected case study.

Moreover, table (4.18) shows different evaluation metrics of the results for CNN-SoftMax function classifier when the system was tested on collected fundus images of case study in key metrics of Specificity, F1- score, Area Under Curve (AUC), Accuracy, and Sensitivity.

Table (4.18) shows the performance measures for the CNN-SoftMax of case study.

	Sensitivity	Specificity	F1-score	AUC	Accuracy
Case study	**96.8%**	**94.4%**	**97.4 %**	**95.6 %**	**96.6 %**

When using experimental data, the proposed system was tested on the original case study fundus images. These fundus images pass through preprocessing phase without augmentation techniques to find out the efficiency of the proposed system when it is tested on real data.

The proposed method is very fast ,the Fundus image prediction time takes approximately (0.06) millisecond for one image when using CNN-Softmax which reduces the problems of handling more cataract cases.

Chapter Five

Conclusions and Future Works

5.1 Conclusions

The essential conclusions of the results obtained from the utilization of the proposed system for Cataract detection and classification in the funds images of retina are as follows:

1. The prepossessing stage is a very important process, at which the images are optimized to be suitable for the following stages. Also, it is the most important and fundamental step in image classification. It is utilized to improve the image quality .
2. The proposed system used multi-images augmentation methods. This system was implemented on these augmented funds images to decrease the problem of over fitting and to increase the proposed system's efficiency. As a high level of categorization accuracy 96.9 % was acquired for the augmented funds images for ODIR datasets (classification of funds images to three stage of cataract).The accuracy was 94% when applying the method on the original funds images of ODIR data set .
3. The proposed system is fast, the funds image prediction time takes approximately (0.06) millisecond for one image when using CNN-Soft max which reduces the problems of handling more cataract cases and this allows the system to be used in real-time.
4. The proposed system achieved higher accuracy with a fewer number of parameters and requiring less computational power.
5. Compare the performance of our architecture and training parameters with other researches by experimental results, the performance of CNN algorithm is more accurate than researcher on the same datasets (Kaggle & ODIR).

5.2 Suggestions for Future Works

The following is recommended for future works:

1. Develop the proposed system further to include more eye diseases (such as eye hypertension, Age Macular Degeneration (AMD), glaucoma, and others) based on funds images.
2. Using another deep learning technique instead of Convolution Neural Network such as VGG19 , U-Net .
3. Extend the suggested system's capabilities by using a big data set of funds images (thousands of photos).
4. Develop the proposed system to classify cataract into three types as Nuclear Cataract (NC), Cortical Cataract (CC), and Posterior Sub capsular Cataract (PSC) based on the location of the lens opacity.

References

References

[1] Radomska-Leśniewska, D. M., Osiecka-Iwan, A., Hyc, A., Góźdź, A., Dąbrowska, A. M., & Skopiński, P. (2019). Therapeutic potential of curcumin in eye diseases. *Central European Journal of Immunology*, *44*(2), 181-189.

[2] Weni, I., Utomo, P. E. P., Hutabarat, B. F., & Alfalah, M. (2021). Detection of Cataract Based on Image Features Using Convolution Neural Networks. Indonesian Journal of Computing and Cybernetics Systems), 15(1), 75-86.

[3] Heidari, M., Mirniaharikandehei, S., Khuzani, A. Z., Danala, G., Qiu, Y., & Zheng, B. (2020). Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. *International journal of medical informatics*, *144*, 104284.

[4] Qiao, L., Zhu, Y., & Zhou, H. (2020). Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. *IEEE Access*, *8*, 104292-104302.

[5] Junayed, M. S., Islam, M. B., Sadeghzadeh, A., & Rahman, S. (2021). CataractNet: An Automated Cataract Detection System Using Deep Learning for fundus Images. IEEE Access, 9, 128799-128808.

[6] Flaxman, S. R., Bourne, R. R., Resnikoff, S., Ackland, P., Braithwaite, T., Cicinelli, M. V., ... & Zheng, Y. (2017). Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *The Lancet Global Health*, *5*(12), e1221-e1234.

[7] Hossain, M. R., Afroze, S., Siddique, N., & Hoque, M. M. (2020, June). Automatic Detection of Eye Cataract using Deep Convolution Neural Networks (DCNNs). In 2020 IEEE Region 10 Symposium (TENSYMP) (pp. 1333-1338). IEEE.

[8] Zhang, L., Li, J., Han, H., Liu, B., Yang, J., & Wang, Q. (2017, May). Automatic cataract detection and grading using deep convolution neural network. In 2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC) (pp. 60-65). IEEE.

[9] Islam, M. T., Imran, S. A., Arefeen, A., Hasan, M., & Shahnaz, C. (2019, November). Source and camera independent ophthalmic disease recognition from fundus image using neural network. In *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)* (pp. 59-63). IEEE.

[10] Pratap, T., & Kokil, P. (2019). Computer-aided diagnosis of cataract using deep transfer learning. Biomedical Signal Processing and Control, 53, 101533.

[11] Wang, J., Yang, L., Huo, Z., He, W., & Luo, J. (2020). Multi-label classification of fundus images with efficientnet. *IEEE Access*, *8*, 212499-212508.

[12] Ram, A., & Reyes-Aldasoro, C. C. (2020). The relationship between Fully Connected Layers and number of classes for the analysis of retinal images. *arXiv preprint arXiv:2004.03624*.

[13] Syarifah, M. A., Bustamam, A., & Tampubolon, P. P. (2020, November). Cataract classification based on fundus image using an optimized convolution neural network with lookahead optimizer. In AIP Conference Proceedings (Vol. 2296, No. 1, p. 020034). AIP

[14] Sudarsono, E., Bustamam, A., & Tampubolon, P. P. (2020, November). An optimized convolution neural network using diff grad for cataract image classification. In AIP Conference Proceedings (Vol. 2296, No. 1, p. 020090). AIP Publishing LLC.

[15] Cao, L., Li, H., Zhang, Y., Zhang, L., & Xu, L. (2020). Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. Information Fusion, 53, 196-208.

[16] Afolabi, O. J., Mabuza-Hocquet, G. P., Nelwamondo, F. V., & Paul, B. S. (2021). The use of U-Net lite and Extreme Gradient Boost (XGB) for glaucoma detection. *IEEE Access*, *9*, 47411-47424.

[17] Khan, M. S. M., Ahmed, M., Rasel, R. Z., & Khan, M. M. (2021, May). Cataract Detection Using Convolution Neural Network with VGG-19 Model. In 2021 IEEE World AI IoT Congress (AIIoT) (pp. 0209-0212). IEEE.

[18] Sirajudeen, A., & Karthikeyan, S. (2021). Detection of Cataract Through Feature Extraction by the Novel Angular Binary Pattern (NABP) and Classification by Kernel Based convolution Neural Networks.

[19] Dong, Y., Zhang, Q., Qiao, Z., & Yang, J. J. (2017, October). Classification of cataract fundus image based on deep learning. In 2017 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 1-5). IEEE.

[20] Abdul Aziz Saleh, M. B. W. (2018). Spoken Arabic digits recognition using deep learning/AbdulAziz Saleh Mahfoudh Ba Wazir (Doctoral dissertation, University of Malaya).

[21] Saravanan, C. (2010,March). Color image to grayscale image conversion. In 2010 Second International conference on computer engineering and applications (Vol. 2, 196-199).IEEE.

[22] Fan, L., Zhang, F., Fan, H., & Zhang, C. (2019). Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, *2*(1), 1-12.

[23] Azzeh, J., Zahran, B., & Alqadi, Z. (2018). Salt and pepper noise: Effects and removal. *JOIV: International Journal on Informatics Visualization*, *2*(4), 252-256.

[24] Popowicz, A., Pigulski, A., Bernacki, K., Kuschnig, R., Pablo, H., Ramiaramanantsoa, T., ... & Zwintz, K. (2017). BRITE Constellation: data processing and photometry. *Astronomy & Astrophysics*, *605*, A26.

[25] Rubini, C., & Pavithra, N. (2019). Contrast enhancement of MRI images using AHE and CLAHE techniques. International Journal of Innovative Technology and Exploring Engineering, 9(2), 2442-2445.

[26] Campos, G. F. C., Mastelini, S. M., Aguiar, G. J., Mantovani, R. G., Melo, L. F. D., & Barbon, S. (2019). Machine learning hyperparameter selection for contrast limited adaptive histogram equalization. EURASIP Journal on Image and Video Processing, 2019(1), 1-18.

[27] Siddhartha, M., & Santra, A. (2020). COVIDLite: A depth-wise separable deep neural network with white balance and CLAHE for detection of COVID-19. arXiv preprint arXiv:2006.13873.

[28] Sultan, D. A., & Yonis, A. A. (2019). Contrast Enhancement in Gray Level Images. JOURNAL OF EDUCATION AND SCIENCE, 28(2), 259-281.

[29] Koo, K. M., & Cha, E. Y. (2017). Image recognition performance enhancements using image normalization. *Human-centric Computing and Information Sciences*, *7*(1), 1-11.

[30] Nisa, M., Shah, J. H., Kanwal, S., Raza, M., Khan, M. A., Damaševičius, R., & Blažauskas, T. (2020). Hybrid malware classification method using segmentation-based fractal texture analysis and deep convolution neural network features. Applied Sciences, 10(14), 4966.

[31] Nayak, S. R., Nayak, D. R., Sinha, U., Arora, V., & Pachori, R. B. (2021). Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study. Biomedical Signal Processing and Control, 64, 102365.

[32] Heidari, J. (2019). Classifying Material Defects with Convolutional Neural Networks and Image Processing (Doctoral dissertation, Uppsala University).

[33] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of big data, 6(1), 1-48.

[34] Grover, M. S., Kumar, Y., Sarin, S., Vafaee, P., Hama, M., & Shah, R. R. (2020). Multi-modal automated speech scoring using attention fusion. arXiv preprint arXiv:2005.08182.

[35] Kaliyugarasan, S. K. (2019). *Deep transfer learning in medical imaging* (Master's thesis, The University of Bergen).

[36] Panov, A. I., Yakovlev, K. S., & Suvorov, R. (2018). Grid path planning with deep reinforcement learning: Preliminary results. Procedia computer science, 123, 347-353.

[37] Dou, X., Hasa, I., Saurel, D., Vaalma, C., Wu, L., Buchholz, D., ... & Passerini, S. (2019). Hard carbons for sodium-ion batteries: Structure, analysis, sustainability, and electrochemistry. Materials Today, 23, 87-104.

[38] Calli, E. (2017). Faster Convolutional Neural Networks., M.Sc. Thesis, Faculty of Social Sciences, Radboud University, Nijmegen.

[39] Mattsson, N., Zetterberg, H., Janelidze, S., Insel, P. S., Andreasson, U., Stomrud, E., & Blennow, K. (2016). Plasma tau in Alzheimer disease. Neurology, 87(17), 1827-1835.

[40] Legaard, C. M., Schranz, T., Schweiger, G., Drgoňa, J., Falay, B., Gomes, C., & Larsen, P. G. (2021). Constructing neural network-based models for simulating dynamical systems. *arXiv preprint arXiv:2111.01495*.

[41] Serra, X., & Castán, J. (2017). Face recognition using Deep Learning. Catalonia: Polytechnic University of Catalonia, 78.

[42] Al-Waisy, A. S., Qahwaji, R., Ipson, S., Al-Fahdawi, S., & Nagem, T. A. (2018). A multi-biometric iris recognition system based on a deep learning approach. Pattern Analysis and Applications, 21(3), 783-802.

[43] LeNail, A. (2019). NN-SVG: Publication-Ready Neural Network Architecture Schematics. J. Open Source Softw., 4(33), 747.

[44] Tran, T. V., Nguyen, H. C., Pham, L. V., Nguyen, M. H., Nguyen, H. C., Ha, T. H., ... & Van Duong, T. (2020). Impacts and interactions of COVID-19 response involvement, health-related behaviors, health literacy on anxiety, depression and health-related quality of life among healthcare workers: a cross-sectional study. BMJ open, 10(12), e041394.

[45] Xia, M., Li, T., Xu, L., Liu, L., & De Silva, C. W. (2017). Fault diagnosis for rotating machinery using multiple sensors and convolution neural networks. IEEE/ASME transactions on mechatronics, 23(1), 101-110.

[46] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[47] Karn, U. (2016). An intuitive explanation of convolution neural networks. The data science blog.

[48] Moravcik, I., Cizek, J., Zapletal, J., Kovacova, Z., Vesely, J., Minarik, P., ... & Dlouhy, I. (2017). Microstructure and mechanical properties of Ni1, 5Co1, 5CrFeTi0, 5 high entropy alloy fabricated by mechanical alloying and spark plasma sintering. Materials & Design, 119, 141-150.

[49] Gong, W., Chen, H., Zhang, Z., Zhang, M., Wang, R., Guan, C., & Wang, Q. (2019). A novel deep learning method for intelligent fault diagnosis of rotating machinery based on improved CNN-SVM and multichannel data fusion. Sensors, 19(7), 1693.

[50] Alradad, M. (2015). *Robust classification with convolutional neural networks* (Doctoral dissertation, University of Missouri--Columbia).

[51] *Dao, H. (2020). Image classification using convolutional neural networks.*

[52] Alzubaidi, L., Fadhel, M. A., Oleiwi, S. R., Al-Shamma, O., & Zhang, J. (2020). DFU_QUTNet: diabetic foot ulcer classification using novel deep convolutional neural network. *Multimedia Tools and Applications*, *79*(21), 15655-15677.

[53] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Regularization for deep learning. *Deep learning*, 216-261.

[54] Jayachitra, S., Kanna, K. N., Pavithra, G., & Ranjeetha, T. (2021, June). A Novel Eye Cataract Diagnosis and Classification Using Deep Neural Network. In Journal of Physics: Conference Series (Vol. 1937, No. 1, p. 012053). IOP Publishing.

المستخلص

إعتام عدسة العين هو أحد أمراض العيون الرئيسية التي تتطور تدريجياً وليس لها تأثير فوري على الرؤية. من أكثر مشاكل الرؤية شيوعًا إعتام عدسة العين، والذي يسبب تشوهًا بصريًا ويمكن أن تؤدي المرحلة المتأخرة من هذا المرض إلى الإصابة بالعمى. يعتبر مرضًا صامتًا يمكن أن يحدث دون ظهور الأعراض. لذلك، فإن الطريقة الأكثر فعالية للكشف عن إعتام عدسة العين هي من خلال الكشف الدقيق وفي الوقت المناسب لتجنب العمليات الجراحية المؤلمة والمكلفة والوقاية من العمى.

الغرض من هذه الرسالة هو اقتراح نظام آلي قائم على نهج التعلم العميق (الشبكة العصبية التلافيفية CNN) لتحديد مرضى الساد .

يحتوي النظام المقترح على ثلاث مراحل رئيسية، المرحلة الأولى هي مرحلة المعالجة الابتدائية والتي تبدأ بتحويل صور قاع العين إلى صور قاع ذات مستوى رمادي بحجم متساوٍ ، وتم تطبيق تقنية "التباين المحدود لمعادلة الرسم البياني التكيفي (CLAHE)" لتعزيز التباين وإظهار ميزة عدسة العين.

تُستخدم طرق زيادة البيانات Augmentation لتجنب مشاكل التجهيز الزائد Overfitting ولتحسين أداء النظام. في المرحلة الثانية، تم تطبيق شبكة CNN (نموذج الشبكة البسيطة) على أساس تقنية تكبير الصور المتعددة (صور قاع العين) كأسلوب لاستخراج الميزات العميقة لتحديد عينات قاع العين. في المرحلة الثالثة تم استخدام دالة (Softmax) لتصنيف مرضى الساد ومراحلها (خفيف، معتدل، شديد). تم استخدام الخصوصية والحساسية والدقة ودرجة F1 والمنطقة تحت المنحنى (AUC) كمعايير لتقييم كفاءة التصنيف.

يستخدم النظام المقترح مجموعتي بيانات متاحتين للجمهور: مجموعة بيانات Kaggle و ODIR.

نوقشت النتائج والاختبارات كافة تفاصيل النظام المقترح وجميع المراحل بعمق وأظهرت النتائج دقة عالية في الكشف عن مرض عتمة العين وبلغت النسبة (99٪) و (96.9٪) لـ (العادي والساد) و (معتدل، متوسط، وحاد) على التوالي في مجموعة بيانات اختبار ODIR بواسطة المصنف المستخدم CNN-Softmax.

الطريقة المقترحة سريعة جدًا في التشخيص، يستغرق وقت التنبؤ بصورة قاع العين حوالي (0.06) ملي ثانية لصورة واحدة.

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة بابل ـ كلية العلوم للبنات
قسم علوم الحاسوب

اكتشاف مرض عتمة العين متعدد الاصناف في الشبكية بناءً على الشبكة العصبية التلافيفية

رسالة

مقدمة الى مجلس كلية العلوم للبنات ـ جامعة بابل

وهي جزء من متطلبات نيل درجة الماجستير في العلوم/علوم الحاسوب

من قبل

هند هادي علي السعدي

بإشراف

د.علي يعكوب يوسف السلطان

د.ايناس حمود محيسن السعدي

1444هـ 2022م