

Automatic Detection Of Diabetic Retinopathy From Retinal Fundus Images Using Deep Learning



Zeyad Ezzat Abdallah
Mostafa Ibrahim Tohamy
Nour El-Din Mostafa Fayez
Aya Amin Hamed Elkhatib
Alia Mohamed Maamoun

Computer and Communication Department

Alexandria University

Supervisor

Dr. Ahmed El-Tarras

In partial fulfillment of the requirements for the degree of

B.Sc. Computer and Communication Engineering

August 13, 2020

Acknowledgement

All the thanks and gratitude go to **Almighty ALLAH** who is the creator of this vast universe and the source of all knowledge and intelligence, the most merciful, gracious, who gave us the strength, guidance, patience and ability to complete the thesis.

Our heartfelt gratitude to our thesis supervisor *Dr. Ahmed El-Tarras* for his generous guidance and continued support and inspiration throughout our work, without his help and assistance, we wouldn't be able to finish our research successfully.

Finally, we are extremely thankful to our parents, family members and friends for their support and encouragement. The journey would be much harder if they weren't present for us in every moment whenever we needed them.

Abstract

Diabetic Retinopathy (DR) is a human eye disease among people with diabetes that causes damage to the retina of the eye and may eventually lead to complete blindness. Detection of diabetic retinopathy in the early stage is essential to avoid complete blindness. Effective treatments for DR are available though it requires early diagnosis and the continuous monitoring of diabetic patients. In this thesis, two approaches are presented for diagnosing DR. First, Deep Learning approach to diagnose DR from fundus images using Convolutional Neural Networks (CNN) architectures and data augmentation. In Addition to the Computer Vision approach, using a combination of Image Assessment and Features Extraction techniques. An open source dataset is used of 35,126 images with different levels of severity of DR provided by Kaggle.

Table Of Contents

| | |
|------------------------------------------------------|----|
| Chapter I | 7 |
| Introduction | 7 |
| 1.1 Background | 7 |
| Eye and diabetes | 8 |
| 1.1 Structure and function of the eye | 8 |
| 1.2 Diabetic eye diseases | 10 |
| 1.2.1 Diabetic retinopathy | 11 |
| STAGES OF DIABETIC RETINOPATHY AND MASCULOPATHY | 13 |
| 1.2.2 Cataract | 13 |
| 1.2.3 Neovascular glaucoma | 15 |
| 1.2.4 Diabetic neuropathies | 15 |
| 1.3 Diagnosing diabetic retinopathy | 15 |
| Clinical eye examination | 17 |
| Eye fundus photography | 17 |
| Alternate diagnostic modalities | 19 |
| 1.4 Screening diabetic retinopathy | 20 |
| 1.5 Automatic detection of diabetic retinopathy | 21 |
| Microaneurysms and haemorrhages | 21 |
| Hard and soft exudates | 24 |
| 1.6 Diabetic retinopathy detection | 27 |
| 1.6 Summary | 28 |
| Chapter II | 29 |
| Feature-based approach | 29 |
| 2.1 Image Quality Assessment | 29 |
| 2.1.1 Step 1: Extract Natural Scene Statistics (NSS) | 31 |
| Mean Subtracted Contrast Normalization (MSCN) | 32 |
| Pairwise products for neighborhood relationships | 33 |
| 2.1.2 Step 2: Calculate Feature Vectors | 33 |
| 2.1.3 Step 3: Prediction of Image Quality Score | 34 |
| 2.2 Microaneurysms Detection | 35 |
| 2.2.1 Stage I (Preprocessing) | 35 |
| 2.2.2 Stage II (Extraction of Mask) | 35 |
| 2.2.3 Stage III (Optic Disk Localization) | 36 |
| 2.2.4 Stage IV (Extraction of Blood Vessels) | 37 |
| 2.2.5 Stage V (Detection of Lesions (Microaneurysm)) | 38 |

| | |
|---------------------------------------------------------------------------------|----|
| 2.3 Exudates Detection | 38 |
| 2.4 Euclidean Distance Between Fovea and The Optic disk and the disk's diameter | 40 |
| 2.4.1 Training | 40 |
| 2.4.2 Predicting OD and Fovea Location on a Single Image | 40 |
| Chapter III | 42 |
| 3.1 Introduction | 42 |
| 3.2 What is Machine Learning | 42 |
| 3.3 Neural Networks and Deep learning | 42 |
| 3.3.1 Neural Network | 42 |
| 3.3.2 Neuron | 42 |
| 3.3.3 The Activation Function (f): | 43 |
| 3.3.4 Feedforward Neural Network | 44 |
| 3.3.5 Single-layer perceptron network | 44 |
| 3.3.6 Multi-layer perceptron network | 45 |
| 3.4 Convolutional Neural Networks | 46 |
| 3.4.1 Convolution | 46 |
| 3.4.2 Non-linearity: (ReLU) | 48 |
| 3.4.3 Pooling or Sub-sampling | 48 |
| 3.4.4 Fully-connected layer | 49 |
| 3.5 What Is Transfer Learning? | 50 |
| 3.5.1 How to Use Pre-Trained Models | 50 |
| 3.5.2 Models for Transfer Learning | 51 |
| Chapter IV | 52 |
| 4.1 Dataset | 52 |
| 4.2 Classical Machine Learning Results | 54 |
| 4.3 Neural Networks Results | 56 |
| 4.3.1 First CNN: Inception | 56 |
| 4.3.2 Second CNN: ResNet | 58 |
| 4.3.3 Comparing two CNN's results | 59 |
| Chapter V | 60 |
| Conclusion and Future work | 60 |
| 5.1 Conclusion | 60 |
| 5.2 Future work | 60 |
| Bibliography | 62 |
| References | 62 |

Chapter I

Introduction

1.1 Background

Diabetes, which can be characterized as a chronic increase of glucose in the blood, has become one of the most rapidly increasing health threats worldwide [55][56]. There are an estimated 150 to 200 million people diagnosed with diabetes, of which approximately 50 million within Europe alone. Moreover, a large number of people remain undiagnosed. In Finland, which has a population of around 5 million, there are 280 000 people under diabetes care of which insulin production in the pancreas is permanently damaged for 40,000 people (type 1 diabetes), and resistance to insulin is increased for 240,000 people (type 2 diabetes). In addition, the current estimates predict that there exist 200,000 undiagnosed patients and that the number of people receiving diabetes care will double every 12 years. These alarming facts promote prevention strategies and screening over a large population since proper and early treatment of diabetes is cost-effective.

Digital imaging technology has developed into a versatile non-invasive measurement tool which enables a wealth of applications also in medical sciences. Imaging the eye fundus with modern techniques is a current practice in many eye clinics, and it is becoming even more important as the expected lifetime and the costs of health care increase. Since the retina is vulnerable to microvascular changes of diabetes and diabetic retinopathy is the most common complication of diabetes, eye fundus imaging is considered a non-invasive and painless route to screen and monitor such diabetic eyes.

Since diagnostic procedures require attention of an ophthalmologist, as well as regular monitoring of the disease, the workload and shortage of personnel will eventually exceed the current screening capabilities. To cope with these challenges, digital imaging of the eye fundus, and automatic or semi-automatic image analysis algorithms based on image processing and computer vision techniques provide a great potential. By automating the analysis process, more patients can be screened and referred for further examinations, and the ophthalmologists have more time for patients that require their attention since most of the eye fundus images are not leading to any medical action.

Eye and diabetes

In this chapter, the diabetic complications in the eye and their implications to vision are discussed. The chapter contains the physiological background concerning the structure and function of the eye, and the description of diabetes-related eye diseases and their symptoms. For the most common diabetic eye disease, diabetic retinopathy, the diagnostic procedures and modalities are presented, and the current and future prospects of early detection are discussed. The discussion includes the shortcomings of the current diagnosis and the potential benefits of automated eye fundus image analysis.

1.1 Structure and function of the eye

In the optical sciences, the human eye is often compared to a camera [50]. Light reflected from an object is focused on the retina after passing through the cornea, pupil and lens, which is similar to light passing through the camera optics to the film or a sensor. In the retina, the incoming information is received by the photoreceptor cells dedicated for detecting light. From the retina, the information is further transmitted to the brain via the optic nerve, where the sensation of sight is produced. During the transmission, the information is processed in the retinal layers. A cross-section of the eye and the structures involved in the image formation are presented in Fig. 1.1.

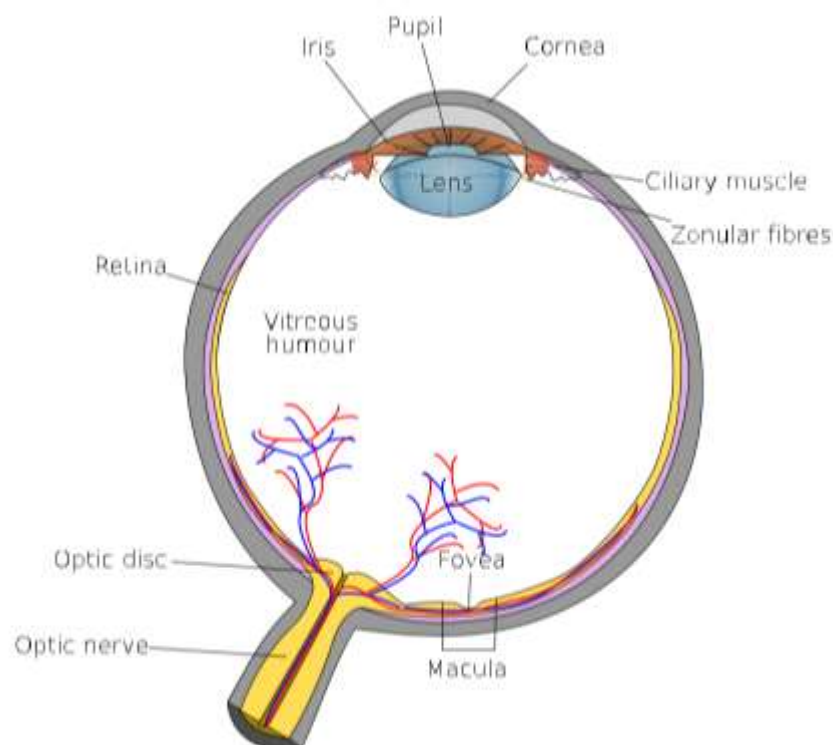


Figure 1.1: Cross-section of the eye.

There are three important features in the camera which can be seen analogous to the function of the eye: aperture, camera lens, and the camera sensor. In the eye behind the transparent cornea, the colored iris regulates the amount of light entering the eye by changing the size of the pupil. In the dark, the pupil is large allowing the maximum amount of light to enter, and in the bright the pupil is small preventing the eye to receive an excess amount of light. In the same way, the camera regulates the amount of light entering the camera with the aperture. In order of the eye to focus on objects at different distances, the ciliary muscle reshape the elastic lens through the zonular fibres. For objects in short distances, the ciliary muscle contracts, zonular fibres loosen, and the lens thickens into orb shaped which results high refractive power. When the ciliary muscle is relaxed, the zonular fibres stretch the lens into thin shaped and the distant objects are in focus. This corresponds to the function of focal length, i.e. the distance between the lens and sensor, when focusing the camera. If the eye is properly focused, the light passes through the vitreous gel to the camera sensor of the eye, that is the retina.

The retina is the inner surface of the eye and consists of transparent tissue of several layers of cells designated to absorb and convert the light into neural signals [68]. The order of the retinal layers is peculiar since the conversion is carried out by the light detecting photoreceptor cells on the layer which is in the back of the retina and furthest from the light. Thus, the light has to travel through the retinal layers before it reaches the photoreceptor cells. Once the light is detected, converted and the neural signals collected to the optic nerve, the impulses are finally transmitted to the brain. During transmission from the photoreceptor cells to the optic nerve the electric impulses are further processed in the inner layers of the retina.

The detailed central vision is formed in the macula which is a highly light sensitive area 5 to 6 mm in diameter in the central region of the retina. In the centre of the macula is a round shaped area known as fovea, where the cones are almost exclusively found. The cones are photoreceptor cells selectively sensitive to different wavelengths of light. Next to the macula is the beginning of optic nerve (optic nerve head or optic disc), from where the main artery and vein emerge in the retina. There are no normal retinal layers

in this region and therefore the absence of photoreceptor cells results in a blind spot in the retina. The nutritional support to the retina is provided by the choroid and the two main capillary networks: the nerve fibre layer network and the connecting neuron layer network. The capillary density increases towards the centre region of the retina and the most dense network is found in the macula, but the fovea itself is absent of capillaries. Therefore, the fovea is dependent on the choroidal blood supply from the vascular layer behind the retina (choroid). The presented anatomical parts (macula, fovea, capillaries, and optic nerve head) highlighted in Fig. 1.2 are the relevant structures of the retina in terms of retinal diseases and this thesis.

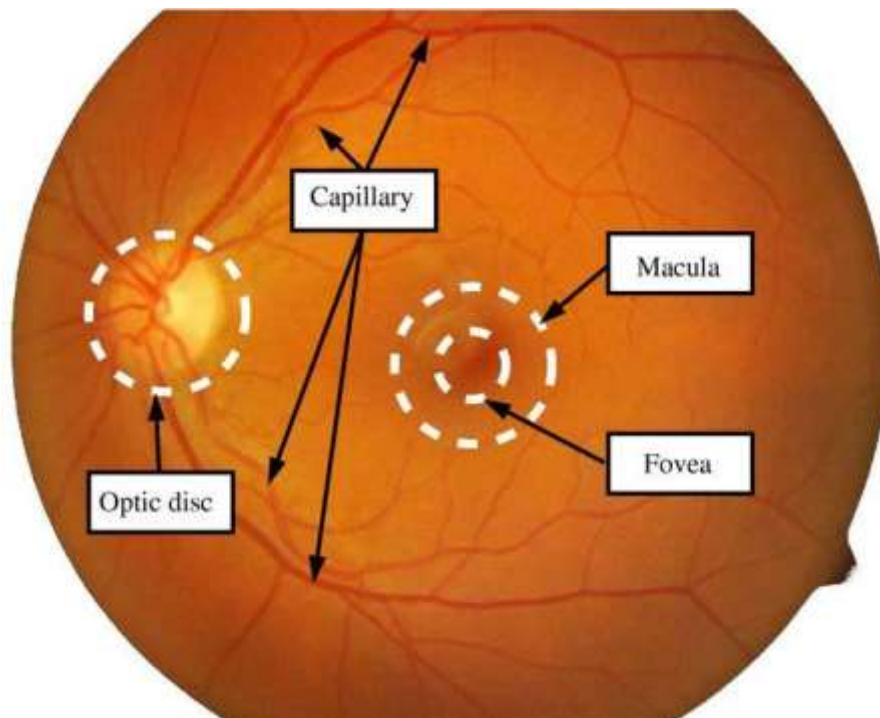


Figure 1.2: Normal physiological parts of the eye fundus.

1.2 Diabetic eye diseases

There are a number of reasons that can cause reduced visual acuity, visual impairment, and blindness. In diabetic eye diseases, the cause of visual disturbances is in most cases related to those vascular changes diabetes is causing to the eye. The discussion in this section concentrates on the diabetic eye diseases that encompass a group of eye problems, such as diabetic retinopathy, cataract, neovascular glaucoma and diabetic neuropathies. The section discusses how the symptoms of the diabetic eye diseases emerge and how they affect the vision. The effect of the diabetic eye diseases on vision is illustrated in Fig. 1.3.



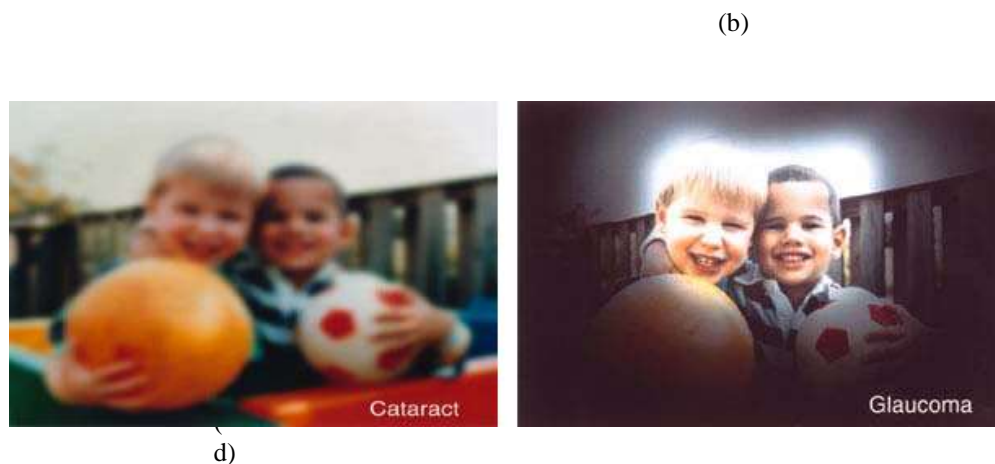


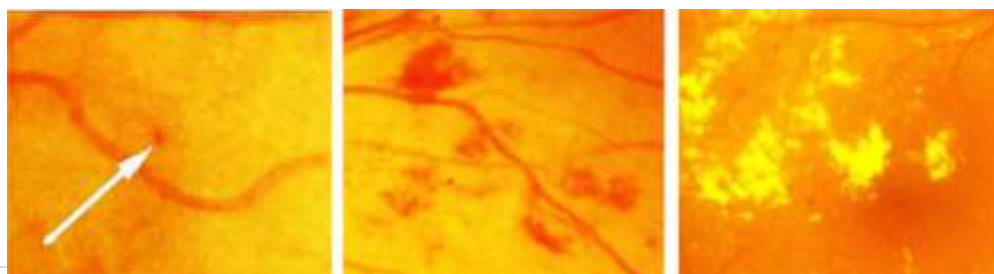
Figure 1.3: Influence of diabetes on vision: (a) normal vision; (b) diabetic retinopathy; (c) cataract; (d) neovascular glaucoma (Courtesy: National Eye Institute, National Institutes of Health).

1.2.1 Diabetic retinopathy

Diabetic retinopathy is a microvascular complication of diabetes, causing abnormalities in the retina. Typically, there are no salient symptoms in the early stages, but the number and severity predominantly increase in time. In the following, the progress of the disease is described in detail.

The diabetic retinopathy typically begins as small changes in the retinal capillaries. The smallest detectable abnormalities, microaneurysms (MA), appear as small red dots in the retina and are local distensions of the weakened retinal capillary (Fig. 1.4(a)). Due to these damaged capillary walls, the small blood vessels may rupture and cause intraretinal hemorrhages (HA). In the retina, the hemorrhages appear either as small red dots indistinguishable from microaneurysms or larger round-shaped blots with irregular outline

(Fig. 1.4(b)). The diabetic retinopathy also increases the permeability of the capillary walls which results in retinal edema and hard exudates (HE). The hard exudates are lipid formations leaking from the weakened blood vessels and appear yellowish with well-defined borders (Fig. 1.4(c)). If the local capillary circulation and oxygen support fail due to obstructed blood vessels, pale areas with indistinct margins appear in the retina.



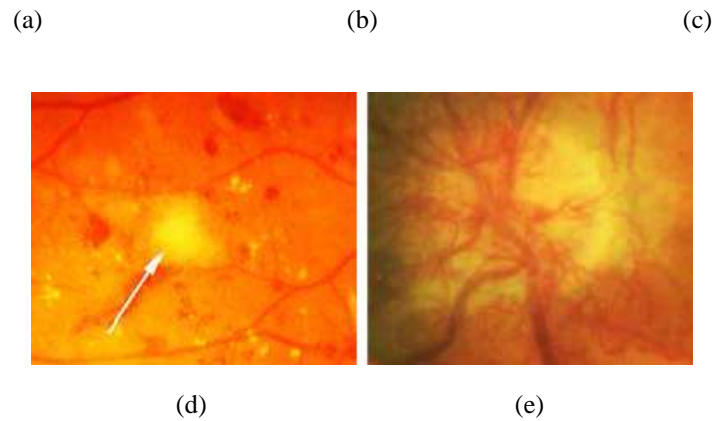


Figure 1.4: Symptoms of diabetic retinopathy (images processed for better visualisation): (a) microaneurysm; (b) hemorrhages; (c) hard exudates; (d) soft exudate; (e) neovascularization in optic nerve head.

These areas are small microinfarcts known as soft exudates (Se) (Fig. 1.4(d)). Intra-retinal microvascular abnormalities (IRMA) and venopathy are signs of a more severe stage of diabetic retinopathy, where intraretinal microvascular abnormalities appear as dilation in the capillary system and venopathy as shape changes in artery and veins. An extensive lack of oxygen and obstructed capillary in the retina lead to the development of new fragile vessels. These new vessels attempt to grow towards the suffering tissue to supply nutrition and oxygen. However, the new vessels are fragile and tend to grow into the space between the retina and vitreous humour, or directly to the vitreous humour, which can lead to preretinal hemorrhage and a sudden loss of vision. The growth of these new vessels is called neovascularization. (Fig. 1.4(e)).

STAGES OF DIABETIC RETINOPATHY AND MASCULOPATHY

The severity of diabetic retinopathy is divided into two stages: nonproliferative (background retinopathy) and proliferative retinopathy. The nonproliferative retinopathy indicates the presence of diabetic retinopathy in the eye and consist of microaneurysms, hemorrhages, exudates, retinal edema, IRMA and venopathy [186, 29]. The microaneurysms and especially hard exudates typically appear in the central vision region (macula) which predicts the presence of macular swelling (macular edema). The symptoms of nonproliferative retinopathy and the macular swelling characterize the maculopathy which is the most common cause of visual disability among the diabetic people. [186, 29]. Although, the maculopathy may occur at any stage of the diabetic retinopathy, it is more likely in the advanced stages of the disease. In the worst case, it can result irreversible damage to the fovea [47]. A retina with nonproliferative retinopathy is illustrated in Fig. 1.5 and a retina with maculopathy is illustrated in Fig. 1.6.

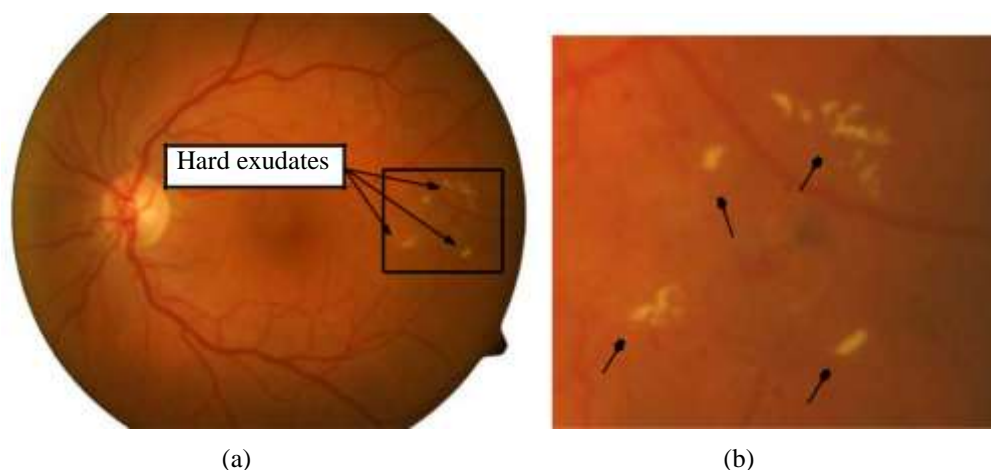


Figure 1.5: Example of nonproliferative diabetic retinopathy: (a) eye fundus image showing hard exudates; (b) close up image of the hard exudates.

If the nonproliferative retinopathy is untreated or undiagnosed it will turn into proliferative retinopathy which is also an eye-sight threatening condition. The proliferative diabetic retinopathy may cause sudden loss in visual acuity or even a permanent blind-ness due to vitreous hemorrhage or tractional detachment of the central retina. This stage is considered if neovascularization or vitreous/preretinal hemorrhage is present in the retina [186, 29]. A retina with proliferative retinopathy is illustrated in Fig. 1.7.

1.2.2 Cataract

Cataract is defined as a decrease in the clarity of the lens which gradually degrades the visual quality. In hyperglycemia, the opacification in the posterior pole of the

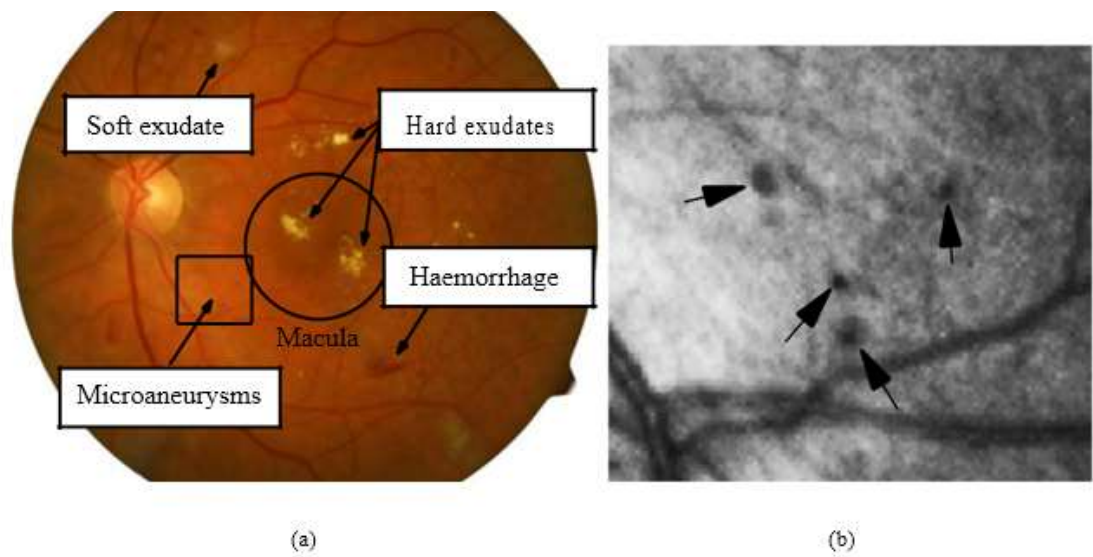


Figure 1.6: Example of maculopathy: (a) eye fundus images with maculopathy showing hemorrhages, microaneurysms, exudates (soft and hard); (b) close-up image of microaneurysms.

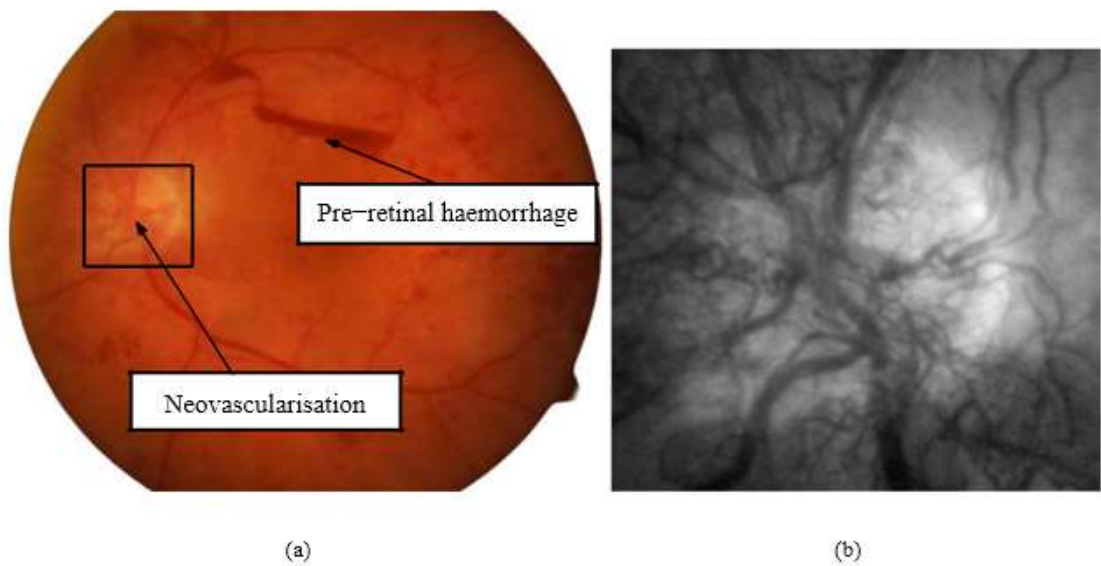


Figure 1.7: Example of proliferative diabetic retinopathy: (a) eye fundus image showing pre-retinal hemorrhage and neovascularization; (b) close up image of neovascularization in the optic nerve head (zoomed from contrast stretched green channel).

lens is caused by the changed metabolism of the lens epithelial cell (posterior subcapsular cataract). Since the lens is responsible for focusing light to the retina, the cataract blocks and distorts the light passing through the lens making the imaging of eye fundus difficult. Therefore, a cataract is a common annoyance in the diagnosis of diabetic retinopathy. Typical visual effects are decreased sensitivity to the light, blurred vision, difficulty with glare and dulled colors (Fig. 1.3(c)). The disease is common for older people since it is usually related to aging and develops gradually in time. In rare occasions, the disease is present at birth or in early childhood, but there are several reasons for the disease to occur earlier in life, such as severe eye trauma, uveitis and diabetes. It is approximated that cataract occur 10-15 years earlier in people with diabetes which is related to the fluctuation of the blood sugar levels [47].

1.2.3 Neovascular glaucoma

The failure of microcirculation in the eye can cause the growth of new vessels in the iris and the chamber angle resulting acute raise in intraocular pressure [162, 153]. This condition is neovascular glaucoma which may occur in people with diabetes due to the ischemic nature of the proliferative retinopathy. The neovascular glaucoma may develop without symptoms, but many experience pain, red eye, light sensitivity or decreased vision. However, the growth of new vessels in the iris and the chamber angle is considered as highly advanced stage of diabetic eye disease which is difficult to cure and often results serious vision loss and therefore the early treatment is essential. The effect of neovascular glaucoma on vision is illustrated in Fig. 1.3(d).

1.2.4 Diabetic neuropathies

Diabetes can also temporarily affect the optic nerve and nerves controlling the eye movement such as nervus oculomotorius (III), trochlearis (IV) and abducens (VI) [47]. The diabetic neuropathies typically cause temporary cross-eyedness that can be alarming for the patient, but it does not indicate permanent damage.

1.3 Diagnosing diabetic retinopathy

Diabetic retinopathy is the most common complication of diabetes and the primary cause for visual impairment and blindness in adults. In this section, the diagnosis of diabetic retinopathy is discussed and the main diagnostic modalities are briefly described.

The diagnosis of diabetic retinopathy is based on clinical eye examination and eye fundus photography. The self-diagnosis of diabetic retinopathy is extremely difficult if diabetes is not suspected, verified from the blood samples or visual impairment is not present. Thus, making diabetic retinopathy a devious eye disease. The eye fundus photography is the preferred diagnostic modality for the primary health care, and for the cases where retinal fundus photographs are ungradable or unavailable, the clinical eye examination is used. Alternate modalities, such as fluorescein angiography and optical coherence tomography, are typically utilized to reinforce the eye examination. If the retina is unreachable and light cannot traverse in the eye, the condition of the retina can be inspected using ophthalmic ultrasound. However, the ultrasound cannot

directly detect diabetic retinopathy, but it can detect if retinal detachment is present due to proliferative retinopathy. It is important to note that it is not possible to diagnose diabetic retinopathy using laboratory tests.

In the screening of diabetic retinopathy, the primary health care doctor use either retinal photography (the first eye fundus photograph evaluation) or direct ophthalmoscopy to investigate the state of the retina. Patients having either no or mild changes are monitored in the primary health care. If the symptoms are in the more advanced stage or the eye fundus images are ungradable, the patient is referred to an ophthalmologist, preferably specialized in diabetic retinopathy. The ophthalmologist re-evaluates or take new eye fundus images (the second eye fundus photograph evaluation), or conduct clinical examinations to diagnose the severity of the disease. Depending on the diagnosis, the patients are appointed for further examinations or treatment. A flowchart of diagnostic procedures is illustrated in Fig. 1.8.

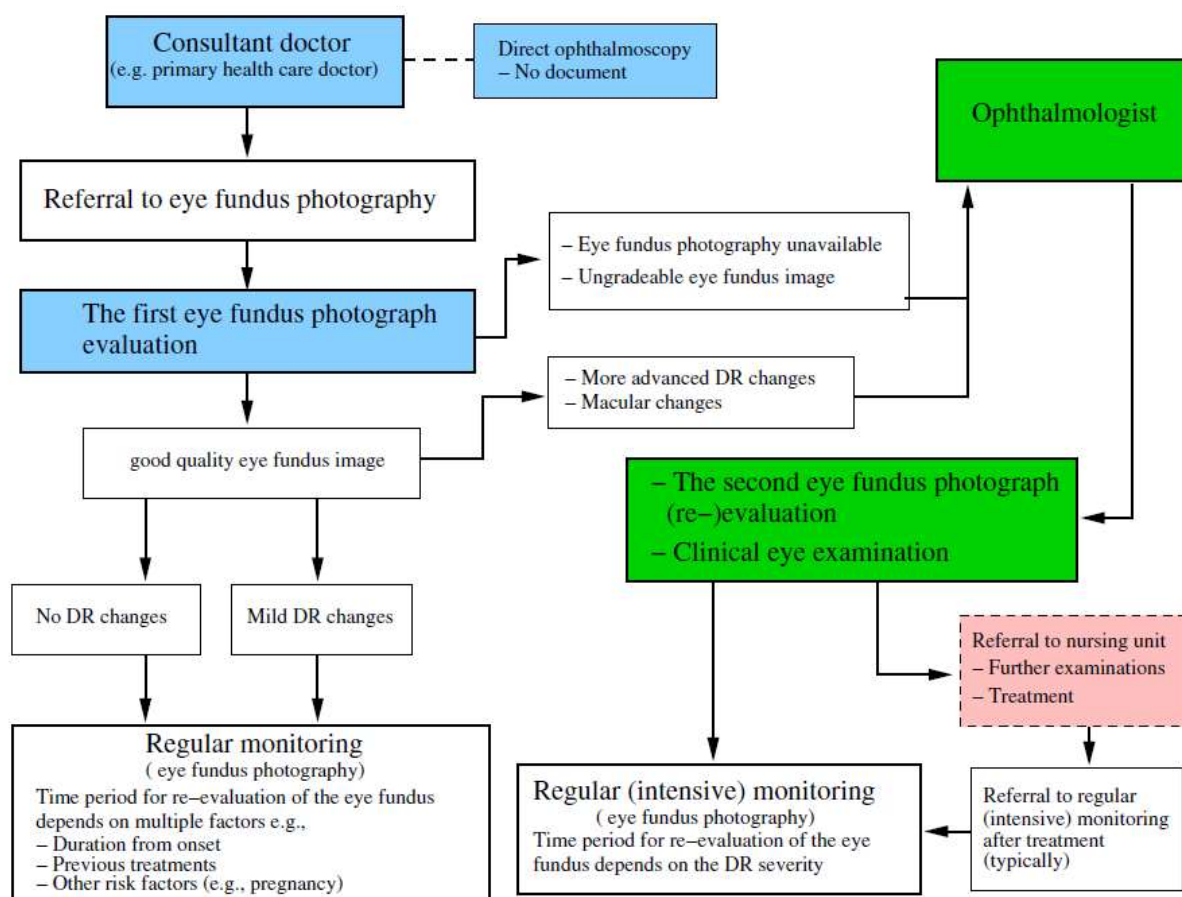


Figure 1.8: A flowchart for diagnosing and monitoring diabetic retinopathy. Blue and green filled boxes with solid borders denote the phases in the eye examination and the responsible medical expert, and the red filled box with the dashed border the treatment phase (modified from [4]).

Clinical eye examination

Main tools in clinical eye examination are direct and indirect ophthalmoscopes, and biomicroscope with indirect lenses. A direct ophthalmoscope is a hand held apparatus through which a medical expert can observe the patient's eye. The apparatus consists of the illumination source and corrective lenses, where the light beams are reflected into to the patient's eye using a mirror or prism. In the indirect ophthalmoscopy, the patient's eye is examined from an arm's length by focusing high intensity light through a hand-held condensing lens to the patient's eye and examining the reflected light (stereoscopic image) with the binocular lenses. The illumination source and the binocular lenses are mounted in a medical expert worn headband. The biomicroscope comprises an observation system and illumination system, where the observation system is a biomicroscope capable of wide range of magnifications and the illumination system emits focal light into the patient's eye that can be controlled with slit mechanism and apertures. Combined with wide field retinal lenses, large areas of the retina can be visualized.

Eye fundus photography

As mentioned, eye fundus photography is considered the preferred diagnostic modality if available since it is reliable, non-invasive and easy to use. In contrast to traditional ophthalmoscopy, it allows to record diagnostic data and enable the expert consultation afterwards, and more importantly the eye fundus photography results in a better sensitivity rate, that is, a better detection rate of abnormal eye fundus. Due to the rapid development of digital imaging, the eye fundus cameras also provide easy to file images in portable format that enable automatic diagnosis of diabetic retinopathy using image analysis algorithms. An eye fundus camera is illustrated in Fig. 1.9.

Eye fundus cameras are divided into two groups: mydriatic and non-mydriatic cameras, where the prefix denotes the requirement for dilation of the pupils with eye drops. The prefix is misleading since in practice the dilation is employed in both fundus camera types. Non-mydriatic fundus cameras are smaller and suitable for screening purposes, but at the same time the image quality is worse and the field-of-view smaller. Thus, mydriatic cameras are used when a more accurate diagnosis is needed.

The patient is seated in front of the fundus camera and the head is positioned into the instrument's head rest. A flash lamp produced light is emitted into patient's eye using optical mirrors and lenses and the reflected light is captured by the camera sensor. The captured images are typically color images (Fig. 1.10(a)), but since the retina is transparent and the penetration depth of the emitted light depends on the wavelength, the desired retinal structures can be emphasized using optical filters. A typical alternative for color images for diagnosing diabetic retinopathy are the red-free eye fundus images (Fig. 1.10(b)). The recommendation in the case of diabetic retinopathy diagnosis is to use the both red-free and color images. Where two images are captured by focusing the 45o field-of-view fundus camera to macula and optic disc (two-field 45o fundus photography). For long-term diabetic patients, two-field 60o photography is recommended since the neovascular changes that need treatment are typically found in the periphery, even if the changes in the central areas of the retina are slight [49].



Figure 1.9: An eye fundus camera.



Figure 1.10: Examples of eye fundus images: (a) color image of an eye fundus; (b) corresponding red-free image.

Alternate diagnostic modalities

In addition to clinical eye examination and eye fundus photography, the fluorescein angiography and optical coherence tomography play an important role in the diagnosis of diabetic retinopathy. In the fluorescein angiography [27], a fluorescent dye (sodium fluorescein) is injected in the systemic circulation of a patient and by emitting light into patient's eye in specific wavelength the fluorescent properties of the dye are activated. The emitted light excites the dye molecules into the higher energy level and as the molecules return to the original state, they emit lower energy light that is captured using eye fundus photography. The obtained image is called angiogram. Since the dye circulates in the ocular vasculature, the fluorescein angiography provides valuable information for the diseases pertaining retinal vasculature such as microaneurysms, capillary nonperfusion and vessels leaking exudate in diabetic retinopathy. Minor disadvantages are the injection and in rare occasions side effects such as nausea. It is worth noting that automatic image analysis algorithms can be applied to the fluorescein angiograms obtained using digital eye fundus photography. A fluorescein angiogram is shown in Fig.1.11.

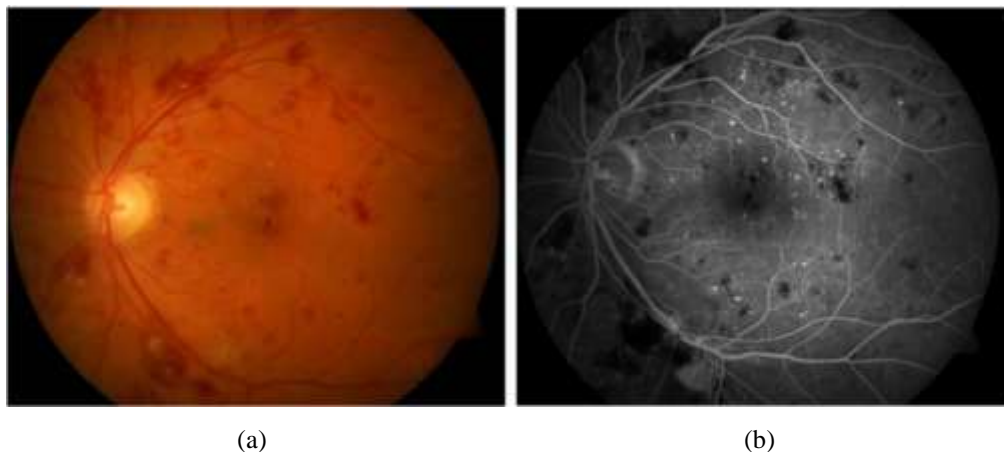


Figure 1.11: Examples of eye fundus images: (a) color image of an eye fundus; (b) corresponding fluorescein angiogram.

Optical coherence tomography (OCT) produces a two-dimensional cross-sectional image of ocular tissue structures, where the dimensions are propagation direction of the light and the perpendicular spatial direction. A broadband beam of light (laser) is scanned across the ocular tissue and due to transparent structures of the retina the time of propagation is longer from the deeper tissue layers. Optical coherence tomography collects the reflected light and use the low-coherence interferometer to measure the time-of-flight delay[27]. The final optical coherence tomography image is composed from several axial scans and using several OCT images a computational three-dimensional reconstruction of

the retina can be devised. In diabetic retinopathy it is mainly used to provide accurate information about macular swelling and its type.

Other modalities used in eye examination: adaptive optics ophthalmoscopy, color Doppler imaging, computed tomography, confocal laser scanning microscope, magnetic resonance, ophthalmic ultrasound, retinal thickness analyzer and scanning laser polarimetry.

1.4 Screening diabetic retinopathy

The prevention of diabetic retinopathy concentrates on controlling the complications of diabetes in the eye through lifestyle and early treatment. These preventive actions can severely delay or stop the progression of the disease, prevent blindness and improve the quality of life. Since there are no salient symptoms in the early stages of diabetic retinopathy, and the number of symptoms and severity predominantly increase with time, a cost-effective screening over large populations is required [21] [41]

Screening is a secondary preventative action which aims to find and treat conditions that have already occurred, but which have not reached a stage that require medical attention. For successful screening the following criteria should be met:

- (P1) “The condition sought should be an important health problem.”
- (P2) “There should be an accepted treatment for patients with recognized disease”
- (P3) “Facilities for diagnosis and treatment should be available.”
- (P4) “There should be a recognizable latent or early symptomatic stage.”
- (P5) “There should be a suitable test or examination.”
- (P6) “The test should be acceptable to the population”
- (P7) “The natural history of the condition, including development from latent to declared disease, should be adequately understood.”
- (P8) “There should be an agreed policy on whom to treat as patients.”
- (P9) “The cost of case finding (including the diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as whole.”
- (P10) “Case-finding should be a continuing process and not a “once and for all” project.”

To shortly review the current state of screening diabetic retinopathy, the Finnish diabetic population is used as a case study for the screening criteria. In Finland, which has a population of around 5 million, there are 280 000 people under diabetes care and this is expected to double in every 12 years if major preventive actions are not undertaken. In addition, it has been estimated that 200 000 people are undiagnosed **Error! Reference source not found.** Since each of the diabetic patients may lose their sight due to diabetic retinopathy, it can be considered an important health problem (P1). The early signs (P4) and progressive nature (P7) of diabetic retinopathy is well documented as well as the severity scales (P8) (Section 1.2.1), and for the diagnosis there are available (P3), acceptable (P6) and non-invasive diagnostic modalities (Section 1.3). If diabetic retinopathy is diagnosed, it can be treated with laser treatment, medical therapy or surgical intervention (P2, P3). However, diabetes contributes to 15% of the total health care costs of which

90% is spent on treating complications such as diabetic retinopathy. Clearly, the cost is not in balance with the expenditure on Finnish medical care as a whole due to the late treatment of diabetes (P9) and as a result the screening cannot be continuous (P10) if the number of diabetic patients doubles as it is estimated.

The problem lies in the process of grading the eye fundus images which is time consuming and repetitive, and requires attention of an ophthalmologist. The grading time is mostly spent on eye fundus images that are not leading to any medical action that is tedious and makes the diagnosis prone to errors. Moreover, the disease progress is reviewed at least once in 1-3 years which results an increasing amount of information for the examination. Thus, a large amount of the costs is tied to the people conducting the diagnosis.

Digital imaging of the eye fundus and automatic or semi-automatic image analysis algorithms provide a potential solution. By automating the grading process more patients could be screened and referred for further examinations, and the ophthalmologists would have more time for patients which require their attention.

1.5 Automatic detection of diabetic retinopathy

The interest towards automatic detection of diabetic retinopathy has been increasing along with the rapid development of digital imaging and computing power. However, the single most important event that attracted the wider attention of medical research community has been the decision to recognize digital imaging as an accepted modality to document eye fundus. Since then, a considerable amount of effort has been spent on automated detection and diagnosis of diabetic retinopathy from digital eye fundus images. The relevant research is well documented in three recent surveys which encapsulate the main algorithms used in the field during the past 10-15 years. In this section, the methodology behind these existing approaches are shortly reviewed.

The review is divided into three parts of which the first two parts discuss the automatic detection of diabetic retinopathy from the lesion point of view (i.e. detecting lesions indicative of diabetic retinopathy) while the discussion in the third part concentrates on algorithms that attempt to detect the presence or even the severity of diabetic retinopathy. The review provides a short description for each method and the reported performances are gathered into summary tables. It should note, however, that the interpretation of the used performance measures varied between publications and different data sets were employed in the evaluation. Therefore, the direct quantitative comparison should not be made. Moreover, the performance measures are restricted to sensitivity, specificity, positive predictive value and average number of false positives.

Microaneurysms and hemorrhages

Some of the first automated detection methods for diabetic retinopathy were published by Baudoin et al. [1] to detect microaneurysms from fluorescein angiograms. By using a morphological top-hat transform with linear structuring element at different orientations small round shaped microaneurysms were distinguished from connected elongated

structures such as vessels. Although the top-hat transform was very sensitive to microaneurysms, it introduced too many false alarms. Spencer et al. [45] exploited this feature and used the top-hat transform to produce candidate microaneurysms. The true microaneurysms were then pruned by using post-processing based on their earlier work [46] and classification. The candidate microaneurysm segmentation was conducted using a combination of top-hat transform and matched filtering with region growing. To improve the sensitivity of the candidate, search a shade correction and dynamic range normalization steps were introduced in the pre-processing. After detection and segmentation of the candidate microaneurysms, the true microaneurysms were pruned from the spurious responses using a rule-based classifier with a number of shape and intensity-based features. By using the computer vision-based detection concept (i.e. image acquisition, pre-processing, candidate object segmentation and classification) Spencer et al. achieved a better control over the problem and allowed the easier development of variant methods. The main difference between the method proposed by Spencer et al. and the variant methods lay in the classification step, where different classifiers and features were used. The feature and classification selection were also studied by Ege et al. [5].

The intravenous use of the fluorescein restricts the use of fluorescein angiography in large scale screening that turned the interest of researchers towards the red-free and color eye fundus photography. Unlike in the fluorescein angiograms, the microaneurysms appear dark in the red-free and color eye fundus images, and have lower contrast. Otherwise the detection task is, however, quite similar. A version of the top-hat transform based method was presented for red-free images, and for color eye fundus images. The top-hat approach was also studied in detection of hemorrhages by Fleming et al. [10].

An alternative mathematical morphology-based approach to overcome a shortcoming of the top-hat based methods: the linear structuring element at discreet orientations tended to detect tortuous vessels as candidate microaneurysms. Instead of using linear structuring element, a bounding box closing was applied with the top-hat transform. Since the candidate object detection of the top-hat transform based methods also missed number of true microaneurysms, Niemeijer et al. [32] proposed a red lesion (microaneurysm and hemorrhage) detection algorithm by introducing a hybrid method to relax the strict candidate object size limitations. A combination of the top-hat based method and a pixel-based classification scheme was proposed to produce a more comprehensive set of candidates. After detecting candidates, the true red lesions were pruned in k-nearest-neighbour classification.

There are also number of approaches for microaneurysm detection published in literature that are not based on morphological operations. One of the first approaches applied in detection diabetic retinopathy was proposed by Gardner et al. **Error! Reference source not found.** who conducted preliminary experiments to study whether neural networks can be used in screening diabetic retinopathy. The neural network and supervised learning were utilized on red-free eye fundus images to extract the microaneurysm and hemorrhage characteristics from set of image patches. Using the trained neural network, the microaneurysms and hemorrhages were located from previously unseen set of test images. Kamel et al. [83] proposed a similar method for fluorescein angiograms to substitute the slow intermediate

filtering sequences in the previously described top-hat transform based methods.

The former approaches assumed that dark areas in the color eye fundus images consist of vessels, microaneurysms and hemorrhages. By excluding the vessels and vessel segments, the remaining objects could be identified as microaneurysms and hemorrhages. Sinthanayothin et al. [43] and Usher et al. [48] used recursive region growing to cluster the dark areas in the image and classified the vessel and vessel segments from the region growing result using a neural network. Instead of using region growing, Grisan and Ruggeri [6] detected dark objects by clustering similar pixels with high local spatial density. The method was further improved by Garcia et al. [12] who added an automatic feature selection and classification step (neural networks) to prune the true red lesions.

Pallawala et al. [33] proposed a different approach, where the microaneurysm detection was defined as a segmentation problem between micro-regions (i.e. microaneurysms) and the background regions (i.e. other eye fundus structures). The graph theory-based segmentation procedure exploited the similarity and spatial proximity of image pixels to cluster small tightly grouped areas into one category and more loosely connected larger areas into the other. In the post-processing step, the true microaneurysms were pruned from the small tightly grouped areas using the intensity difference between the area and the immediate surrounding.

Xiaohui and Chutatape [60][199] extracted the characteristic features of hemorrhages from image templates using the principal component analysis (PCA). The extracted features were used with the support vector machine to classify the image patches of previously unseen color eye fundus image. To detect the different sized hemorrhages, a pyramid of images was generated by computationally changing the image resolution. Quellec et al. [35][131] also used image templates in detection of microaneurysms by conducting template matching in wavelet domain. Bhalerao et al. [2] modelled the circular shape and locally darker appearance of microaneurysms by combining the Laplacian of Gaussian and circular-symmetry operator in filtering the candidate microaneurysms. A template from each candidate microaneurysm location was extracted to find the true microaneurysms using a PCA-based region analysis.

An alternative approach was proposed by introducing the use of temporal changes between color eye fundus images to detect the symptoms of diabetic retinopathy. First, the images were pre-processed by registering the images with the dual-bootstrap iterative closest point algorithm, correcting the illumination using the iterative homomorphic surface fitting, and removing the dust particles using the ratio of the green and red reflectance components. The difference of two pre-processed eye fundus images were then classified to change and no-change regions, and further to lesion categories including microaneurysms and hemorrhages based on the difference ratio values

Hard and soft exudates

The early work in the automated detection of hard and soft exudates generally investigated possibilities of thresholding techniques. Ward et al. **Error! Reference source not found.** proposed a semi-automatic exudate detection system based on shade correction and thresholding, where the user interaction was required in the thresholding. By introducing a dynamic thresholding procedure, it was provided a considerable improvement to the previous system. Philips et al. [34] detected large high intensity areas from red-free eye fundus images using a global thresholding scheme, whereas a block-wise local thresholding was applied to segment the smaller exudates.

The method was able to produce relatively good results in detecting the exudate pixels, but at the same time an unacceptable number of false positives were generated. As a result, the ophthalmologic expertise was required in the interpretation of the detection results. To counter the false positives, Zheng et al. [62] introduced the use of local neighborhood in the dynamic block-wise local thresholding procedure.

Simultaneously with the early thresholding, Goldbaum et al. [15][16] proposed a bright lesion detection approach based on template matching and edge detection. By computationally resampling the image resolution and applying the template matching, the bright lesions of all sizes were located. The outlines of each located lesion were refined using edge detection. Most importantly, Goldbaum et al. [17] took into account the sub-classes of bright lesions, i.e. hard and soft exudates, that should be differentiated for the diagnosis of diabetic retinopathy. For the identification, a spherical color space was introduced.

From template matching and thresholding, the method development turned towards supervised statistical pixel-based lesion classification. The first pixel- and block-based classification approaches utilized a simple minimum distance discriminant classifier to classify image pixels into two classes (i.e. to bright lesion and background) according to their pixel value. Wang et al. [53] applied the minimum distance discriminant classifier in the color space proposed by Goldbaum et al. [17] to detect candidate hard exudate pixels. The true hard exudate pixels were then pruned using the contrast information of the local neighborhood.

Sánchez et al. [40] further developed the method using alternative approaches for the non-uniform illumination correction and hard exudate pixel pruning. Goh et al. [14] applied the minimum distance discriminant classifier directly in RGB color space to detect hard exudate pixels.

Niemeijer et al. [31] proposed a pixel classification scheme based on k-nearest neighbor classification to detect and differentiate hard and soft exudates, and drusen. The system searched for candidate bright lesion pixels according to the features selected in the training stage. By using the density of classified lesion pixels among the neighboring pixels in the feature space, a lesion probability was assigned for each pixel in the test image. The high probability regions were pruned to find the true bright lesions by extracting descriptive features for each region and applying the KNN classification once again. Finally, a linear classifier was used to classify the detected true bright lesions to hard and soft exudates and drusen. To improve the hard exudate detection, Sanchez et al. **Error! Reference source not found.** introduced the use contextual information, i.e. the spatial relations of the surrounding anatomical structures and similar lesions.

A feature-based classification scheme was also proposed by Xu and Luo [57], where a feature combination based on stationary wavelet transform and gray level co-occurrence matrix was used to characterize the textural properties of hard exudates. The pixel-level classification was conducted using a support vector machine. Another support vector machine-based approach was suggested by Silberman et al. [42], where the local textural properties were extracted using a scale-invariant feature transform (SIFT).

Given the irregular properties of hard exudates in shape and scale, the pixel- and block-based clustering methods provided another strong approach for local segmentation of hard exudates. Hsu et al. [20] utilized dynamic clustering after shade correction to group the image pixels into (bright) lesion and non-lesion areas. The hard exudate clusters were differentiated from other bright lesions based on the contrast difference between the (bright) lesion and non-lesion classes. A more elaborate clustering technique was proposed where the candidate object detection was based on Gaussian-smoothed histogram analysis and fuzzy c-means (FCM) clustering. The idea was to use the Gaussian-smoothed histogram analysis to coarsely cluster the image according to the significant extreme points in the image histogram. The FCM clustering was then used to assign the remaining unclassified pixels to the previously determined clusters. To identify the hard exudate regions from the non-hard exudate regions image features were extracted from the clustered regions and then classified using a neural network approach. Zhang and Chutatape [59][61] further improved the FCM method to exploit information in the neighboring pixels to decrease the effect of the image noise.

Ram and Sivaswamy [37] studied a clustering scheme, where two feature spaces were composed by combining color channels of different color spaces. The objective was to construct two divergent feature spaces descriptive for different eye fundus structures. First, the k-means clustering was used on both feature spaces to group similar image pixels and then by using reasoning and the complementary behaviour of the feature spaces, a set of candidate hard exudate regions were selected from the clustering results. To suppress false positives, the color information of the candidate regions before and after the color channel decorrelation procedure was exploited.

Region growing is another clustering technique applied in detection of hard and soft

exudates. Luo et al. [26] extracted candidate bright lesions using a block-wise region-growing scheme after emphasising bright areas in each image block by manipulating histogram properties of L and u color components in LUV color space. The difference between the mean pixel value inside and along the region contour was applied in the search of true bright lesions. The method was later adopted by Li and Chutatape [24][25]. The widely cited research of Sinthanayothin et al. [43] proposed another recursive region-growing procedure to identify similar pixels in the detection of bright lesions. After the pixels were merged into regions, the bright lesions were differentiated by thresholding. Prior to the region-growing procedure a local color contrast enhancement step was performed. A more recent region-growing approach was presented by Esrawan et al [8], where the bright lesions were detected from edge magnitude image using a marker driven watershed transform. To emphasise the regions of interest, an average filtering and contrast stretching preprocessing steps were performed.

In addition to the classification and clustering methods, a number of mathematical morphology-based approaches have been published in the literature. Unlike the previously described methods, Walter et al. [51][52] attempted to morphologically reconstruct the eye fundus image exclusive of bright lesions. The method was based on the idea that the difference between the reconstructed image and the original image would ultimately express the bright lesion locations. Firstly, the bright candidate lesions were coarsely located based on their local contrast variation after morphologically suppressing the blood vessels. Secondly, the candidate bright lesion areas were removed and morphologically reconstructed to correspond the appearance of the retinal background. By thresholding the difference of the original image and the reconstructed image, the final lesion areas were obtained. The same idea was later applied and refined.

Ravishankar et al. [38] applied morphological operations to suppress the blood vessel network and to emphasise the bright lesion boundaries by dilating the image at two different scales and subtracting the outcomes. By using dynamic thresholding and morphologic filling, the candidate exudates were extracted from the dilation result by finding the connected regions within the emphasized lesion boundaries. In the final stage, brightness and edge properties of the candidate regions were used with the rule-based classification to determine the true hard exudate regions. Kumari et al. [22] had a very similar approach without the final classification step and with the dynamic thresholding step substituted to edge detection with the Laplacian of Gaussians.

Fleming et al. [11] utilized their earlier experiences with the microaneurysm detection and introduced an inversion of the morphological top-hat technique to detect hard exudates. A linear structuring element at different orientations and scales were applied with the top-hat transform to emphasise bright regions in the image. By using a dynamic thresholding scheme, the candidate exudates were extracted from the top-hat result and finally classified to exudates, drusen and background according to their color, shape, brightness and the contrast difference to the local background. The use of contextual information in the form of distance to the closest microaneurysm was also studied.

An alternative hard exudate detection approach was proposed where the statistical properties of an eye fundus image were captured into a mixture of Gaussians. By estimating the probability density function of each image and by analysing the concavities of the approximated density function the candidate bright regions were

thresholded from the background. False positives were eliminated in the postprocessing step, where the edge strength of the candidate regions was made use of. A similar idea was introduced where the mixture model was substituted to a global and several local image histogram, and the true hard exudate search to automatic feature selection and neural network classification.

1.6 Diabetic retinopathy detection

The microaneurysm count was one of the first criteria's proposed to differentiate healthy and diabetic images. The idea exploited the knowledge that the number of microaneurysms correlates with the severity of diabetic retinopathy. Since the idea was simple, it was considered as a promising candidate to detect the presence of diabetic retinopathy in eye fundus images which boosted the development of microaneurysm detection algorithms.

Since other lesions indicative of more severe disease state can occur prior to microaneurysms, the automatic detection of diabetic retinopathy using microaneurysm count had room for improvement. To overcome this problem, different lesion counts were experimented e.g., Sinthanayothin et al. [44] devised the presence of diabetic retinopathy in eye fundus images using the hard exudate count, Larsen et al. [23] and Hansen et al. [19] used microaneurysm and hemorrhage counts, and Usher et al. [48] used microaneurysm, hemorrhage and hard exudate counts. Reza and Eswaran [39] also attempted to devise the severity of diabetic retinopathy using a lesion count-based approach, whereas Estabridis and Figueiredo [7] used the spatial lesion locations.

Alternatively, machine learning is utilized in the literature. Goldbaum et al. [16] proposed a method, where an image-description language was utilized to symbolise the properties of lesions and normal structures in eye fundus images. By using a neural network and a set of representative images of known diagnosis, the connection between the image-description information and the known diagnosis was learned. Based on the training data, the trained network was able to detect images with diabetic retinopathy using the image-description information derived from the result of automatic eye fundus image analysis system. Instead of using image descriptions, Neural networks with wavelet coefficients are used, Anitha et al. **Error! Reference source not found.** trained the network using cluster centroids obtained from channel-wise fuzzy c-means clustering, and the work in [3][28][58] utilized shape related features extracted from the segmented vessel and lesion regions.

Another feature-based classification scheme based on machine learning was proposed where the normal and pathological structures were characterised according to their amplitude and frequency content, conceptually, describing the textural properties. After decomposing the image using the amplitude modulation - frequency modulation (AF-FM) decomposition and extracting the textural features for each image from several image locations, k-means clustering was used to cluster the extracted features into several groups. The number of textural features in each of the group defined the final descriptor for each image. Using the described descriptors and a partial least square classifier, the presences of diabetic retinopathy was detected in eye fundus images.

Goh et al. [13] combined several classifiers to detect diabetic retinopathy from eye fundus images by following the main principal of classifier ensembles, i.e. while the single best performance is achieved using the best classifier, the misclassified samples by the best classifier are not necessarily misclassified by rest of the classifiers. A set of nine ensembles containing several classifiers were constructed, where three classifier ensembles were used for the background, four for the blood vessel and three for the bright lesion detection. The classifiers within each ensemble were based on neural networks. By using decision rules, the classifier outputs within each ensemble and the resulting ensemble outputs were combined into single label representing background, blood vessel or bright lesion. Each sub-image in the image was assigned with such label and based on the image patches containing bright lesions the presence of diabetic retinopathy was finally detected.

Niemeijer et al. [29] also studied the use of multiple information sources and their fusion to achieve more robust detection for screening diabetic retinopathy. The proposed system analysed two images per eye and four images in total to produce a subject-wise score value that indicated the likelihood that the test subject needs to be examined by a human observer. Each of the four images were separately analysed with complex computer aided detection or diagnosis (CAD) system to determine the image quality and the diabetic lesion locations (red and bright lesions). The subject-wise score value was then combined from the outputs of the CAD systems and based on the score value the diabetic retinopathy was detected in eye fundus images. Fleming et al. [9] had a similar idea, where eye fundus images of both eyes were first analysed for microaneurysms, hemorrhages, hard exudates and image quality and by using the output several image-based lesion measures were computed and combined into overall subject-wise score. Similarly, as previously, the score value was used to detect the patients with referable diabetic retinopathy.

1.6 Summary

In this chapter, diabetes-related eye diseases and their symptoms were discussed. For the most common diabetic eye disease, diabetic retinopathy, the diagnostic procedures and modalities were described. Since the eye fundus photography is non-invasive, efficient, easy to use and enable the documentation of the diagnostic data, it was considered the preferred tool for diagnosing diabetic retinopathy. The diagnosis of diabetic retinopathy from an eye fundus image was based on identifying the described symptoms of nonproliferative, proliferative retinopathy and maculopathy,

The chapter also addressed that future prospects of diabetic retinopathy are alerting and preventive actions, such as a large-scale screening of high-risk individuals, are required. The number of diabetic patients at the current increase rate will eventually exceed the current screening capabilities due to the time consuming and tedious grading process of the eye fundus images. Fortunately, the development of digital imaging in the diagnosis of diabetic retinopathy, i.e., digital eye fundus photography, has provided a possibility to automate the grading process and thereby improve the efficiency of the screening. The fields of image processing and computer vision has shown a great interest towards automatic and semi-automatic detection of diabetic retinopathy, and number of publications are already presented in the literature as stated in the literature review. The need for automatic image analysis system to detect the early signs of diabetic retinopathy (hemorrhages, microaneurysms, hard and soft exudates) is the main motivation for this thesis.

Chapter II

Feature-based approach

2.1 Image Quality Assessment

Image Quality Assessment (IQA) algorithms take an arbitrary image as input and output a quality score as output. There are three types of IQA:

1. Full-Reference IQA: Here you have a ‘clean’ reference (non-distorted) image to measure the quality of your distorted image. This measure may be used in assessing the quality of an image compression algorithm where we have access to both the original image and its compressed version.
2. Reduced-Reference IQA: Here you don’t have a reference image, but an image having some selective information about it (e.g. watermarked image) to compare and measure the quality of the distorted image.
3. Objective Blind or No-Reference IQA: The only input the algorithm gets is the image whose quality you want to measure. This is thus called, No-Reference or Objective-Blind.

No-Reference IQA

In this post, we will discuss one of the No-Reference IQA Metrics, called Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE). Before we go deeper into the theory, let’s first understand two basic terms:

1. Distorted Image: As the name suggests, a distorted image is a version of the original image that is distorted by blur, noise, watermarking, color transformations, geometric transformations so on, and so forth.

| № | Type of distortion (four levels for each distortion) | Correspondence to practical situation | Accounted HVS peculiarities |
|----|-----------------------------------------------------------------------------------------------------|------------------------------------------|------------------------------------------|
| 1 | Additive Gaussian noise | Image acquisition | Adaptivity, robustness |
| 2 | Additive noise in color components is more intensive than additive noise in the luminance component | Image acquisition | Color sensitivity |
| 3 | Spatially correlated noise | Digital photography | Spatial frequency sensitivity |
| 4 | Masked noise | Image compression, watermarking | Local contrast sensitivity |
| 5 | High frequency noise | Image compression, watermarking | Spatial frequency sensitivity |
| 6 | Impulse noise | Image acquisition | Robustness |
| 7 | Quantization noise | Image registration, gamma correction | Color, local contrast, spatial frequency |
| 8 | Gaussian blur | Image registration | Spatial frequency sensitivity |
| 9 | Image denoising | Image denoising | Spatial frequency, local contrast |
| 10 | JPEG compression | JPEG compression | Color, spatial frequency sensitivity |
| 11 | JPEG2000 compression | JPEG2000 compression | Spatial frequency sensitivity |
| 12 | JPEG transmission errors | Data transmission | Eccentricity |
| 13 | JPEG2000 transmission errors | Data transmission | Eccentricity |
| 14 | Non eccentricity pattern noise | Image compression, watermarking | Eccentricity |
| 15 | Local block-wise distortions of different intensity | Inpainting, image acquisition | Evenness of distortions |
| 16 | Mean shift (intensity shift) | Image acquisition | Light level sensitivity |
| 17 | Contrast change | Image acquisition, gamma correction | Light level, local contrast sensitivity |

Fig. 2.1 Distortions used in TID 2008 Database

2. Natural Image: An image directly captured by a camera, with no post-processing, is a natural image in our context. Here is an example of a natural image and a distorted image.



Fig. 2.2 Natural Image (left) and Noisy Image (distorted, right)

As you can imagine, it is not always clear-cut whether an image is distorted or it's natural. For example, when a video is smartly rendered with motion blur, the algorithm may get confused about its quality because of the intentional blur. So, one has to use this quality measure in the right context.

Image Quality Assessment (IQA) Dataset

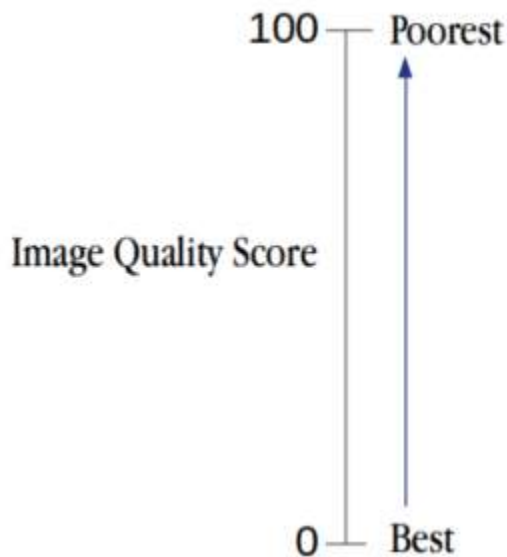


Fig. 2.3 TID2008 Image Quality Score Scaling (0 to 100): lesser the score, better the subjective quality.

Quality is a subjective matter. To teach an algorithm about good and bad quality, we need to show the algorithm examples of many images and their quality score.

Who assigns the quality score for these training images? Humans, of course. But we cannot rely on the opinion of just one human. So, we need opinions of several humans and assign the image a mean score between 0 (best) and 100 (worst). This score is called the Mean Quality Score in academic literature.

Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)

In this section, we will describe the steps needed for BRISQUE algorithm used for No-Reference IQA.

Fig. 2.4 shows the steps involved in calculating BRISQUE.

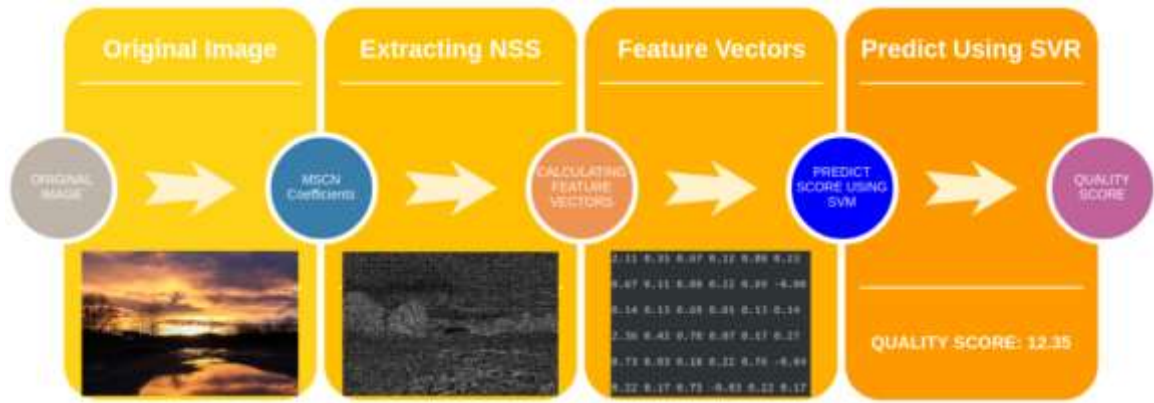


Fig. 2.4 Steps to Calculate Image Quality Score using BRISQUE Model

2.1.1 Step 1: Extract Natural Scene Statistics (NSS)

The distribution of pixel intensities of natural images differs from that of distorted images. This difference in distributions is much more pronounced when we normalize pixel intensities and calculate the distribution over these normalized intensities. In particular, after normalization pixel intensities of natural images follow a Gaussian Distribution (Bell Curve) while pixel intensities of unnatural or distorted images do not. The deviation of the distribution from an ideal bell curve is therefore a measure of the amount of distortion in the image.

We have shown this in the Figure below.

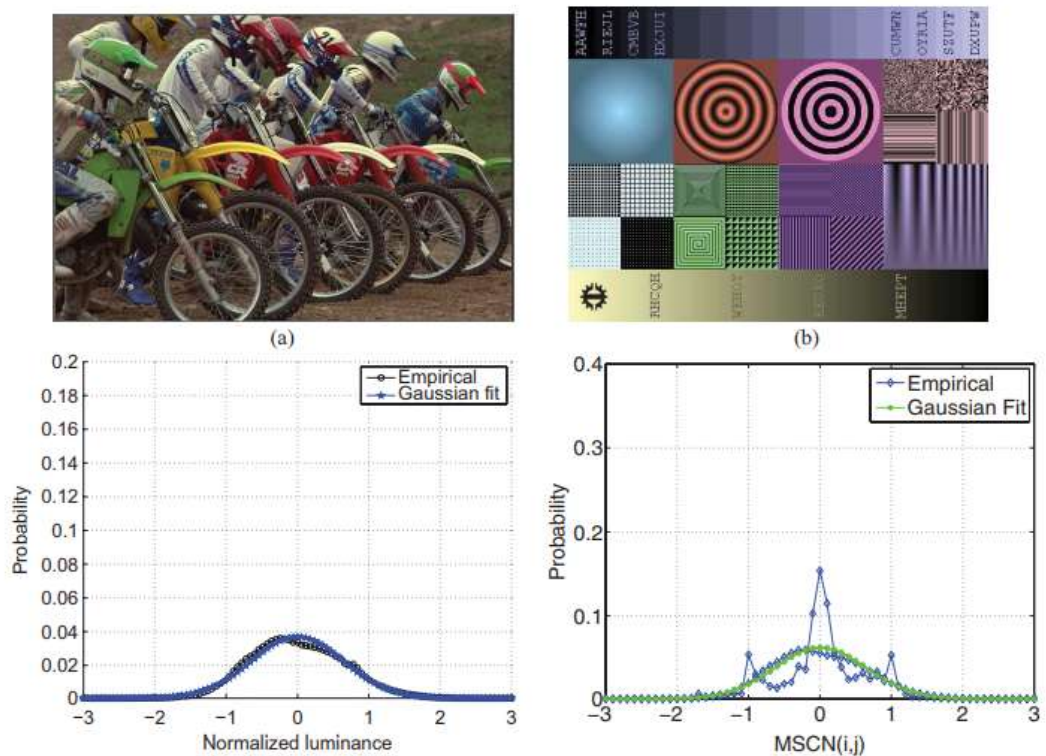


Fig. 2.5 On the left: shows a natural image with no artificial effects added, fits Gaussian distribution. On right: An artificial image, doesn't fit the same distribution well.

Mean Subtracted Contrast Normalization (MSCN)

There are a few different ways to normalize an image. One such normalization is called Mean Subtracted Contrast Normalization (MSCN). The figure below shows how to calculate MSCN coefficients.

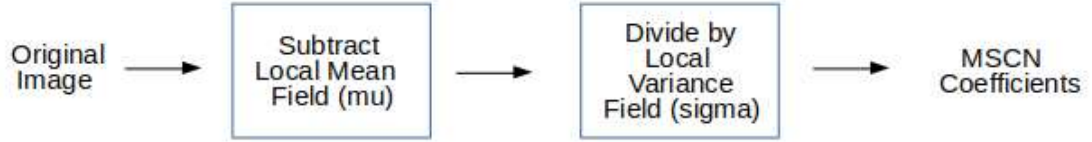


Fig. 2.6 Steps to calculate MSCN Coefficients
This can be visualized as:

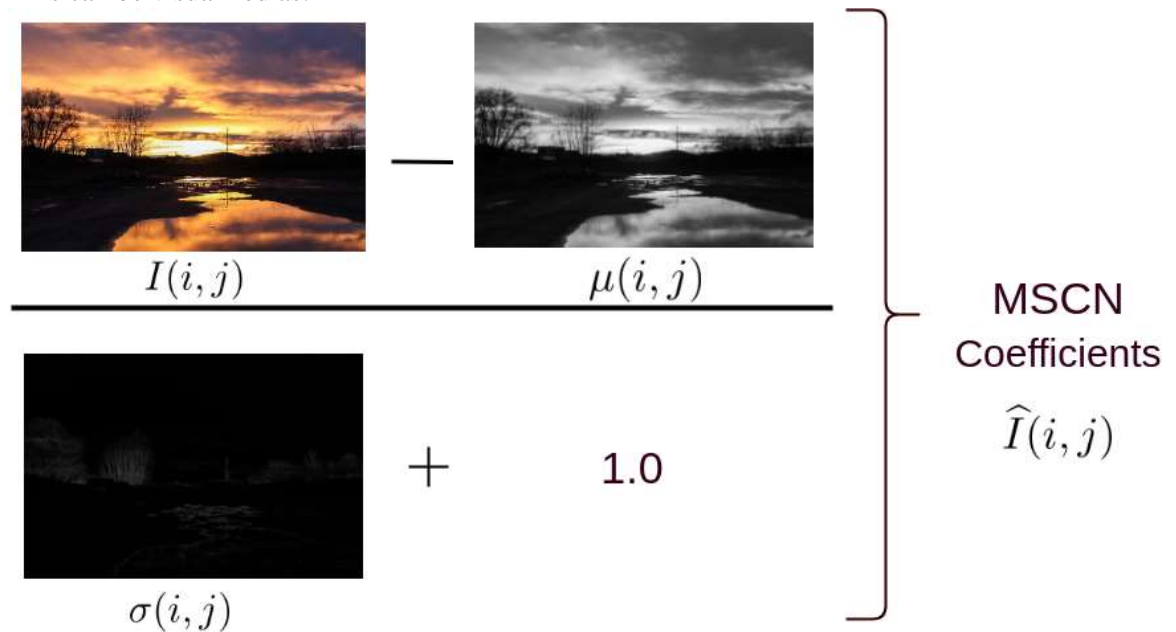


Fig. 2.7 Calculation of MSCN Coefficients

Don't let the math below scare you. The code after the math is much simpler and easy to understand!

To calculate the MSCN Coefficients, the image intensity $I(i, j)$ at pixel (i, j) is transformed to the luminance $\hat{I}(i, j)$.

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + C} \quad (1)$$

Where $i \in 1, 2, \dots, M, j \in 1, 2, \dots, N$ (M and N are height and width respectively). Functions $\mu(i, j)$ and $\sigma(i, j)$ are local mean-field and local variance field respectively. Local Mean Field (μ) is nothing but the Gaussian Blur of the original image, while Local Variance Field (σ) is the Gaussian Blur of the square of the difference of original image and μ . In the equation below \mathbf{W} is the Gaussian Blur Window function.

$$\begin{aligned} \mu &= \mathbf{W} * \mathbf{I} \\ \sigma &= \sqrt{\mathbf{W} * (\mathbf{I} - \mu)^2} \quad (2) \end{aligned}$$

We use the GaussianBlur function in Python to calculate MSCN Coefficients.

Pairwise products for neighborhood relationships

MSCN provides a good normalization for pixel intensities. However, the difference in natural vs. distorted images is not limited to pixel intensity distributions, but also the relationship between a pixel and its neighbors.

To capture neighborhood relationships the authors used **pair-wise products** of MSCN image with a shifted version of the MSCN image. Four orientations are used to find the pairwise product for the MSCN coefficients, namely: Horizontal (H), Vertical (V), Left-Diagonal (D1), Right-Diagonal (D2).

$$\begin{aligned}
 H(i, j) &= \hat{I}(i, j) \hat{I}(i, j + 1) \\
 V(i, j) &= \hat{I}(i, j) \hat{I}(i + 1, j) \\
 D1(i, j) &= \hat{I}(i, j) \hat{I}(i + 1, j + 1) \\
 (3) D2(i, j) &= \hat{I}(i, j) \hat{I}(i + 1, j - 1)
 \end{aligned}$$

2.1.2 Step 2: Calculate Feature Vectors

Until now, we have derived 5 images from the original image — 1 MSCN image and 4 pairwise product images to capture neighbor relationships (Horizontal, Vertical, Left Diagonal, Right Diagonal).

Next, we will use these 5 images to calculate a feature vector of size 36×1 (i.e. an array of 18 numbers). Note the original input image could be of any dimension (width/height), but the feature vector is always of size 36×1 .

The first two elements of the 36×1 feature vector are calculated by fitting the MSCN image to a **Generalized Gaussian Distribution** (GGD). A GGD has two parameters — one for shape and one for the variance.

Next, an **Asymmetric Generalized Gaussian Distribution** (AGGD) is fit to each of the four pairwise product images. AGGD is an asymmetric form of Generalized Gaussian Fitting (GGD). It has four parameters — shape, mean, left variance, and right variance. Since there are 4 pairwise product images, we end up with 16 values.

Therefore, we end up with 18 elements of the feature vector. The image is downsized to half its original size and the same process is repeated to obtain 18 new numbers bringing the total to 36 numbers.

This is summarized in the table below.

| Feature Range | Feature Description | Procedure |
|---------------|--------------------------------------------|------------------------------------------------|
| 1 - 2 | Shape and Variance. | GGD fit to MSCN coefficients. |
| 3 - 6 | Shape, Mean, Left Variance, Right Variance | AGGD fit to Horizontal Pairwise Products |
| 7 - 10 | Shape, Mean, Left Variance, Right Variance | AGGD fit to Vertical Pairwise Products |
| 11 - 14 | Shape, Mean, Left Variance, Right Variance | AGGD fit to Diagonal (left) Pairwise Products |
| 15 - 18 | Shape, Mean, Left Variance, Right Variance | AGGD fit to Diagonal (Right) Pairwise Products |

Table (2.1): Feature Ranges and Description of QA algorithm

2.1.3 Step 3: Prediction of Image Quality Score

In a typical Machine Learning application, an image is first converted to a feature vector. Then the feature vectors and outputs (in this case the quality score) of all images in the training dataset are fed to a learning algorithm like Support Vector Machine (SVM).

LIBSVM is used to predict the final quality score by first loading the trained model and then predicting the probability using the support vectors produced by the model. It's important to note that the feature vectors are first scaled to -1 to 1 and are then used for prediction.

We have performed the metric on 4 types of distortions. Here is the final quality score, for each distortion:


| Original Image | JPEG2K Compression | Heavy Compression | Gaussian Noise | Median Blur |
|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
|  |  |  |  |  |
| 26.8286 | 30.7417 | 33.0692 | 79.8751 | 72.7044 |

Fig. 2.8: QA scores of different types of distortion

2.2 Microaneurysms Detection

2.2.1 Stage I (Preprocessing)

The preprocessing is done to remove noise (pixel whose color is distorted) from the background and to enhance the image. In the first stage of preprocessing green channel is taken out, because the green channel shows high intensity as compared to red and blue. For enhancement of green channel, histogram equalization is used. Mathematical formula for finding the green channel is as follows:

$$g = \frac{G}{(R + G + B)} \quad (1)$$

$$h(v) = \text{round} \left(\frac{\text{cdf}(v) - \text{cdf}_{\min}}{(M \times N) - \text{cdf}_{\min}} \times (L - 1) \right) \quad (2)$$

Where g is a Green channel and R , G and B are Red, Green, and Blue respectively. In the green channel, all minute details of the image can be viewed. Using red channels only boundary is visible, and in the blue channel, the image shows lots of noise. Due to these reasons, the green channel is used in the proposed system. For finding histogram equalization of an image following formula is used.

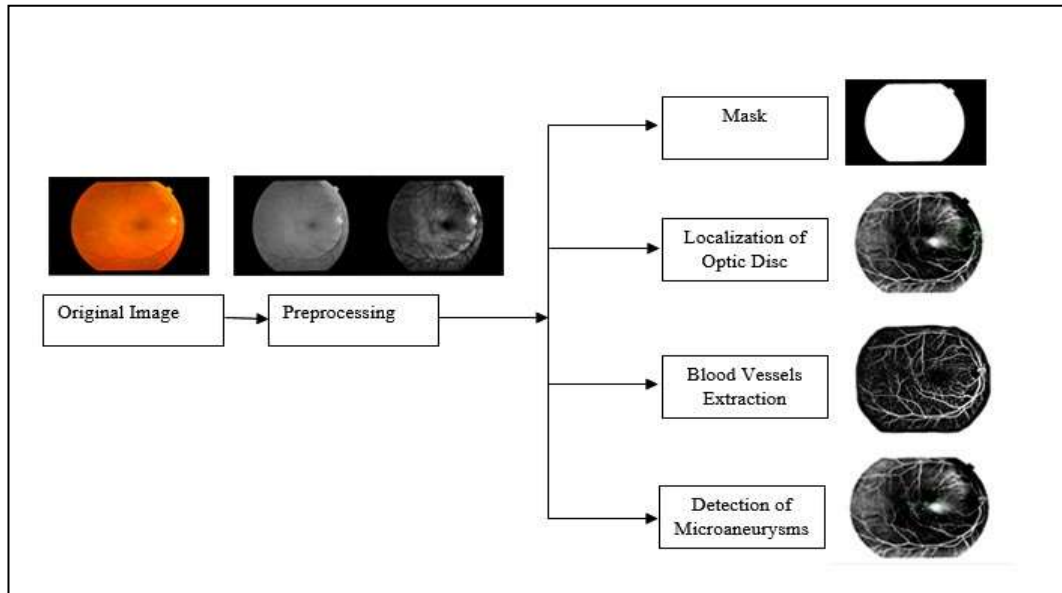


Fig. 2.9” Flow chart for “Automated Diabetic Retinopathy Detection System”

Here cdf_{\min} is the minimum value of the cumulative distribution function, $M \times N$ gives the image's number of pixels and L is the number of grey levels. In histogram equalization frequencies of images are spread out, thus images get enhanced

2.2.2 Stage II (Extraction of Mask)

Mask detection is necessary because it displays the exact structure of the boundaries of an image. If boundaries are not proper then the image can be discarded to avoid further

processing. To extract the mask of fundus image red channel is considered because it is used to detect boundaries. Once the red channel is taken out, thresholding operation is performed. For finding threshold function following formula is used:

$$T = \frac{1}{2(N1 + N2)} \quad (3)$$

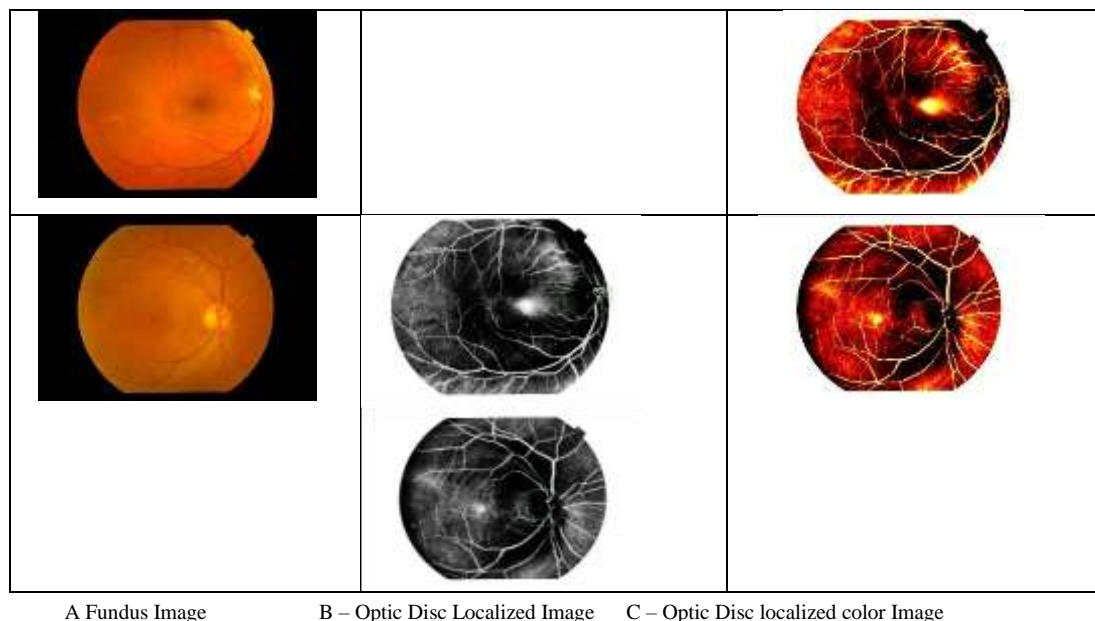
Where $m1$ & $m2$ are the Intensity Values. Threshold is a type of segmentation where the required object is extracted from the background

2.2.3 Stage III (Optic Disk Localization)

A ganglion cell is a cell found in a ganglion, which is a biological tissue mass, most commonly a mass of nerve cell bodies. The optic nerve head in a normal human eye carries from 1 to 1.2 million neurons from the eye towards the brain. The optic disc or optic nerve head is the location where ganglion cell axons exit the eye to form the optic nerve. The optic disc is also the entry point for the major blood vessels that supply the retina. Pattern of these blood vessels near the optic disc region plays an important role in diagnosis for eye disease. An Ophthalmologist checks this region to detect normal and abnormal vessels. Abnormal vessels are called as tortuous vessels. If blood vessels get tortuous then the chances of leaking the blood is higher, which in turn can damage Retina. For the localization of optic disc, the green channel is used as it shows high intensity for the pixel values compared to red and blue. After applying some image processing techniques seed-up robust features are used. Following is the formula for Speed Up Robust Features

$$I_{\Sigma}(x, y) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y) \quad (4)$$

Given an input image I and a point $(x; y)$ the integral image I_{Σ} is calculated by the sum of the values between the point and the origin. Figure 2 shows the original images in column A and localized optic disk images in columns B and C.



A Fundus Image

B – Optic Disc Localized Image

C – Optic Disc localized color Image

Fig. 2.10 Optic Disc Localization

2.2.4 Stage IV (Extraction of Blood Vessels)

For extraction of blood vessels, Image processing operations are performed on the green channel image. Figure 3 shows original images in column A and Extracted blood vessels of images are shown in column B and C. Histogram equalization function is used for enhancing the green channel image followed by 2-D median filter [11 to 14]. Morphological structuring element is applied for highlighting the blood vessels of the retina.

$$I_{dilated}(i, j) = \max_{f(n, m)=true} I(i + n, j + m) \quad (5)$$

$$I_{eroded}(i, j) = \min_{f(n, m)=true} I(i + n, j + m) \quad (6)$$

Morphological open function is used for thickening the retinal

$$A \circ B = (A \ominus B) \oplus B \quad (7)$$

Where $A \circ B$ is morphological opening, \ominus is Erosion and \oplus is Dilation.

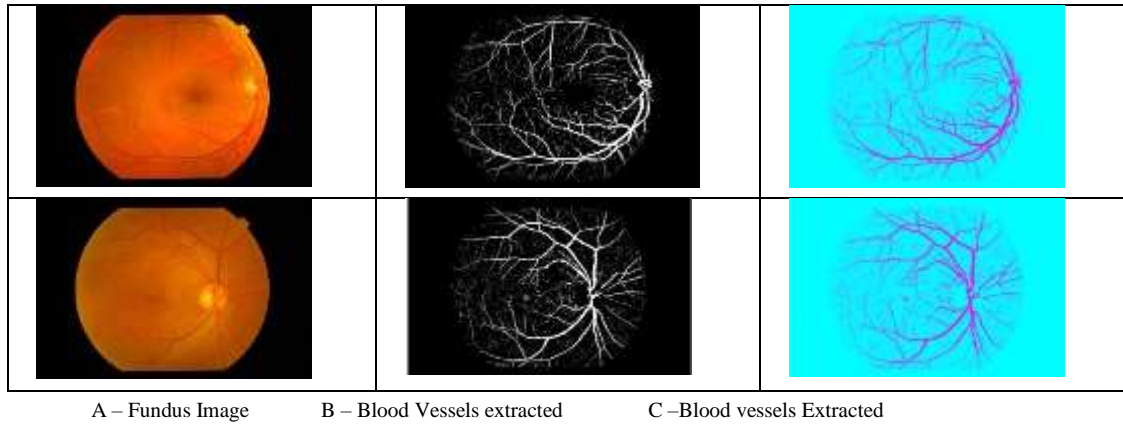


Fig. 2.11 Blood Vessels Extraction

To remove noise 2D median filter is used.

$$y[m, n] = \text{median}\{x[i, j], (i, j) \in \omega\} \quad (8)$$

Where ω Represents a neighborhood centered around location (m, n) in the image. In the last Threshold function is used for extracting the retinal blood vessels.

$$T = \frac{1}{2(m1 + m2)} \quad (9)$$

Where $m1$ & $m2$ are the Intensity Values.

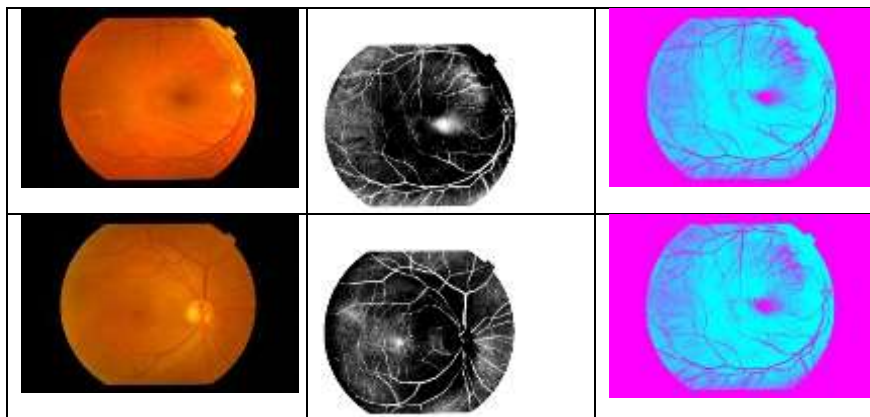
After using image processing techniques in the end intensity-transformation function is used. Following is the formula for Intensity Transformation Function.

$$s = T(r) \quad (10)$$

Where T is Transformation and r is Intensity.

2.2.5 Stage V (Detection of Lesions (Microaneurysm))

Microaneurysms are the first clinically detected lesions. It is Tiny swelling in the wall of a blood vessel. It appears in the retinal capillaries as a small, round, red spot. Usually, they are located in the inner nuclear layer of the retina. To begin with the detection of microaneurysm, Green channel is taken out, which is followed by compliment function, and histogram equalization function. In figure 4 original images are shown in column A and Microaneurysm detected images are shown in column B and C.



A – Fundus Image B – Microaneurysms C – Microaneurysms

Fig. 2.12 Detection of Microaneurysms

2.3 Exudates Detection

The machine learning algorithm is based on our earlier work using retinal pixel and lesion classification. To perform detection and differentiation of bright lesions, if any, in a previously unseen image, the following steps were performed:

1. The optical disk is masked first so that it doesn't count as an exudate.
2. Each pixel was classified, resulting in a so-called lesion probability map that indicates the probability of each pixel to be part of a bright lesion.
3. Pixels with high probability were grouped into probable lesion pixel clusters.
4. Based on cluster characteristics each probable lesion pixel cluster was assigned a probability indicating the likelihood that the pixel cluster was a true bright lesion.

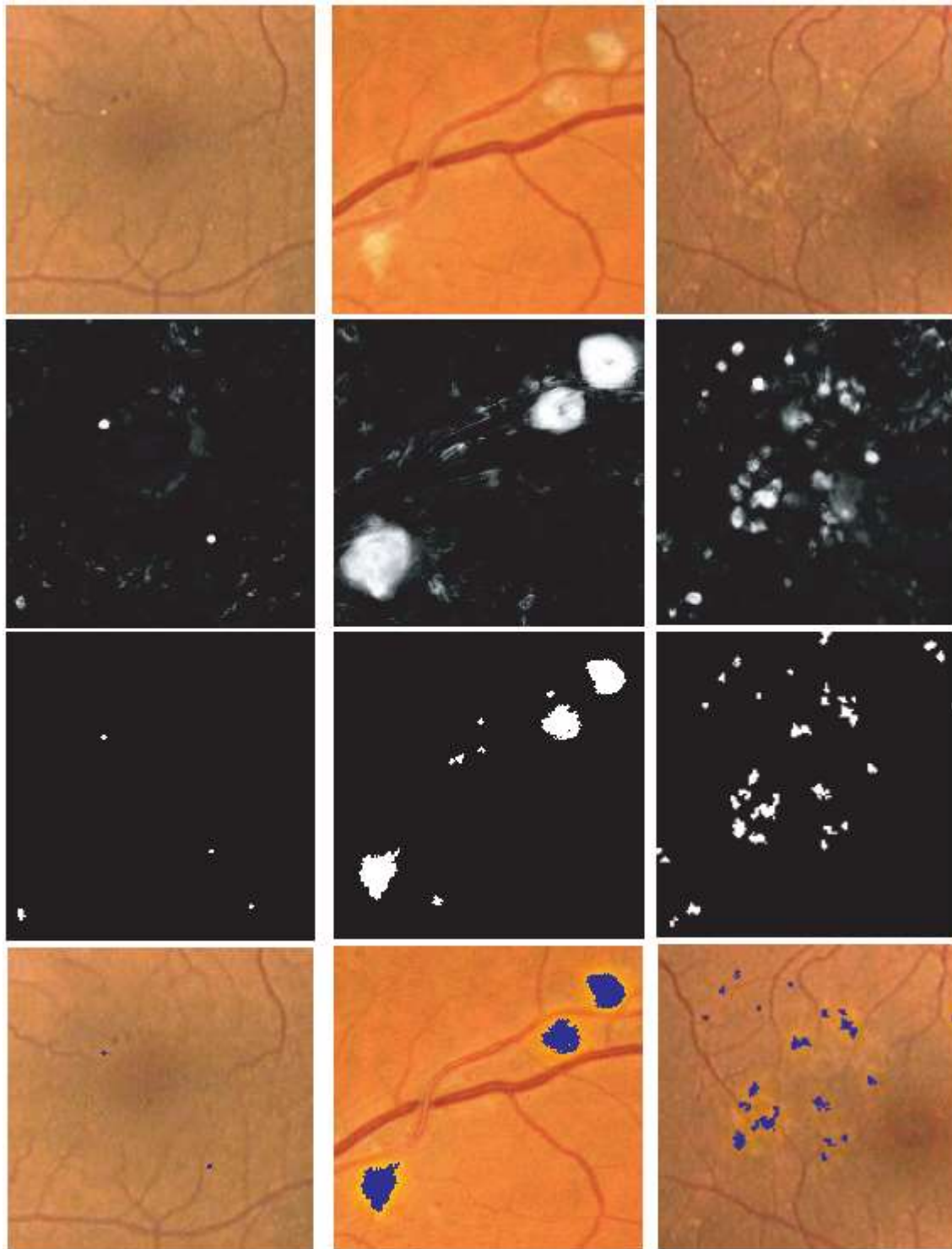


Fig. 2.13: Machine learning algorithm steps performed to detect and differentiate “bright lesions”. From left to right column, exudates, cotton-wool spots, and drusen. From top to bottom, the first row shows the relevant region in the retinal color image (all at the same scale), the second row shows the posterior probability map after the first classification step, the third row shows the pixel clusters that are probable bright lesions (second step), and the bottom row shows those objects which the system classified as true bright lesions overlaid on the original image.

4. Each bright lesion cluster likely to be a bright lesion was classified as exudate, Cotton-wool spot or drusen.

Figure 2.13 illustrates the above steps. The system performs the classification steps (1, 3, and 4) using a statistical classifier that, given a training set of labeled examples, can differentiate among different types of pixels or clusters based on so-called features or numerical characteristics, such as pixel color, cluster area, etc. The features used depend on the type of objects that are to be classified. For each classification step, several statistical classifiers were tested on the training set, and the one demonstrating the best performance was used on the test set. A k-nearest neighbor classifier was selected for steps 1 and 3, and a linear discriminant classifier for the third classification step [2]. When presented with an unseen image in the test set, the automated algorithm gave two outputs: whether bright lesions were present or not, and which class(es) of lesions, exudates, or cotton-wool spots, or drusen.

2.4 Euclidean Distance Between Fovea and The Optic disk and the disk's diameter

In contrast with previous techniques, the proposed method does not attempt to directly detect only OD and fovea centers. Instead, the distance to both locations is regressed for every pixel in a retinal image. This regression problem can be solved by means of a Fully-Convolutional Neural Network. This strategy poses a multi-task-like problem, on which information of every pixel contributes to generate a globally consistent prediction map where the likelihood of OD and fovea locations are maximized.

In particular, we make use of the U-net architecture, while using a loss function suitable to perform pixel-wise distance regression (L2 Loss).

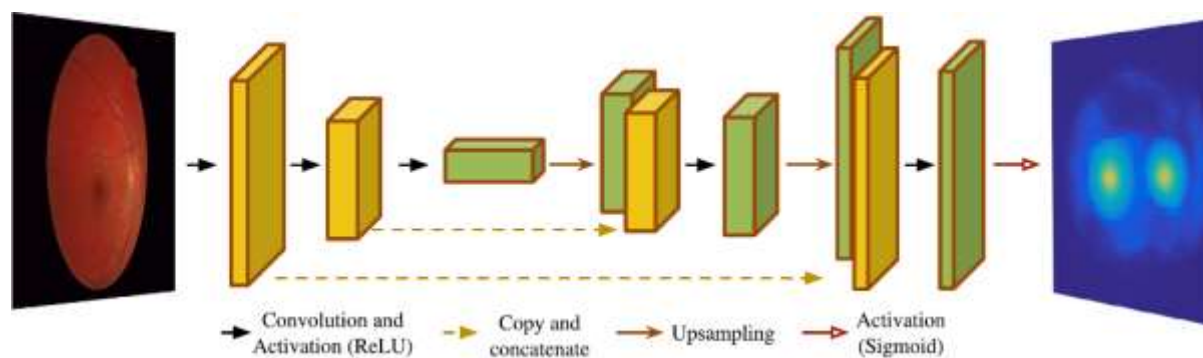


Fig. 2.14

2.4.1 Training

The method was trained and validated on the Messidor dataset.

2.4.2 Predicting OD and Fovea Location on a Single Image

The pre-trained weights of the model are in the file `best_weights.h5`. Please be aware that this model was trained on a particular random split of the data as specified above. If you use it on a Messidor subset without re-training, some of the images may belong to the original training set. The example image in the `images/` folder is from the test set of the particular split this model was trained on.

NOTE:

When possible, input images should be cropped around the FOV, or a mask of the FOV provided as specified above. If no FOV mask is provided, this implementation will still return the optic disc and fovea locations, although results may be suboptimal.

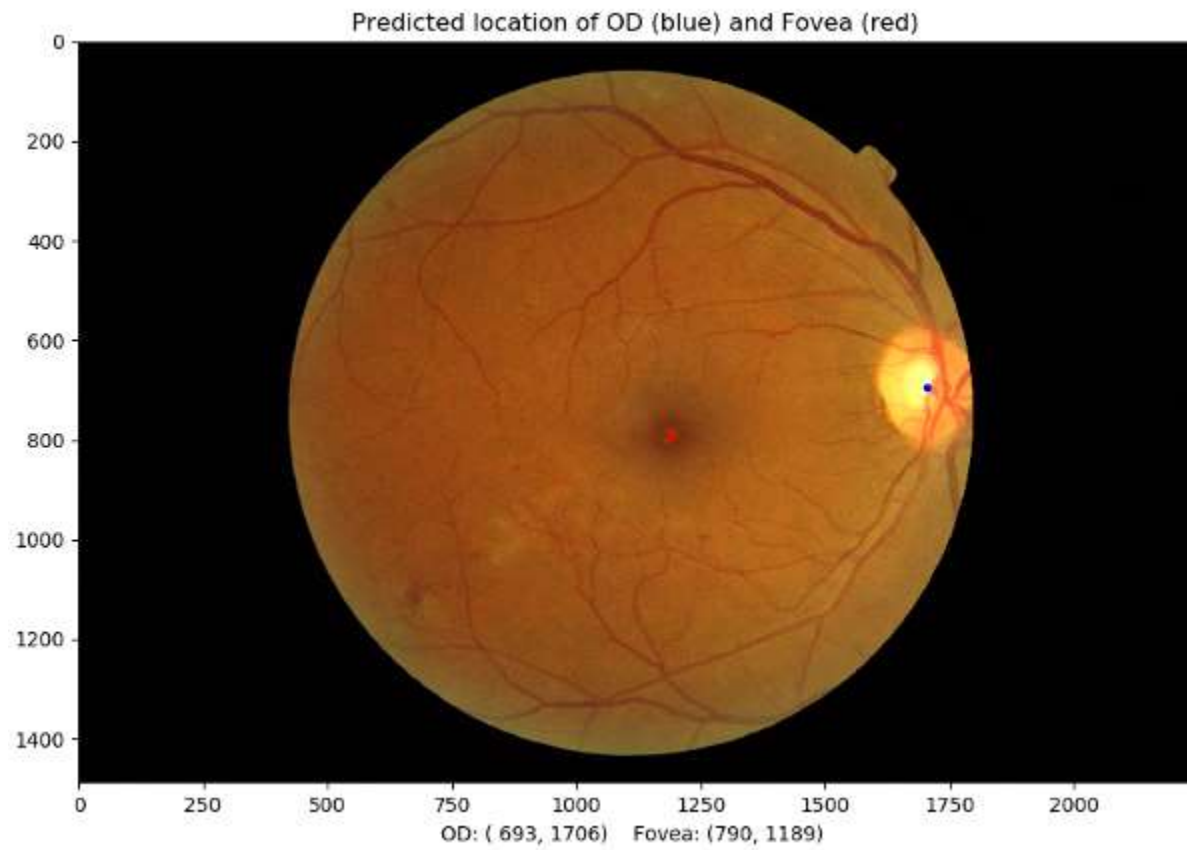


Fig. 2.15: Predicted location of OD (blue) and Fovea (red).

Chapter III

3.1 Introduction

In this chapter, we introduce some of the basic concepts that are merely useful for understanding the underlying concepts of the used algorithms throughout this book. This chapter can also be considered as a concise handy reference for beginners in the field of image classification and processing using Deep learning and computer vision.

3.2 What is Machine Learning

Machine learning (ML) is a category of algorithm that allows software applications to become accurate in predicting results and classifying data without being explicitly programmed. The basic premise of machine learning is to build algorithms that can receive some input data and use statistical analysis to predict the output while updating its outputs as new data becomes available. The processes involved in machine learning are similar to that of predictive modeling. That is, the model searches through data to look for patterns and adjust program actions accordingly. One of the most sophisticated algorithms of machine learning is called Neural Networks (NNs) which are non-linear models that are in their basic premise a trial for modeling the humanoid brain cells.

3.3 Neural Networks and Deep learning

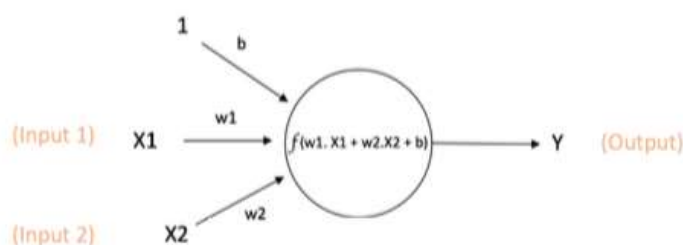
Deep learning is a subset of machine learning where artificial neural networks, algorithms inspired by the human brain, learn from large amounts of data. Similarly, to how we learn from experience, the deep learning algorithm would perform a task repeatedly, each time tweaking it a little to improve the outcome

3.3.1 Neural Network

An Artificial Neural Network (ANN) or a Neural Network, is a computational model that has drawn inspiration from the way neurons in the human brain process information. With the groundbreaking results in computer vision and natural language processing, ANNs have catered a lot of interest in the Machine Learning industry.

3.3.2 Neuron

A neuron, also called a node or a unit, is the simplest computation unit in a neural network. A neuron receives information or inputs from other nodes in the network or from external sources. Based on its relative importance to other inputs, every input has a weight w , assigned to it. The neuron then computes the output by applying a function to the weighted sum of its inputs.



$$\text{Output of neuron} = Y = f(w_1.X_1 + w_2.X_2 + b)$$

Figure 3.1: Representation of a Neuron

As shown in Figure 3.3.2 **X1** and **X2** with associated weights **w1** and **w2** are passed as inputs to the network. There is another input to the neuron called the **Bias (b)**. The primary function of bias is to offer each node a constant value that is trainable, along with the inputs that the node gets normally.

3.3.3 The Activation Function (f):

The Activation Function (f), is non-linear and is used to add non-linearity to the neuron's output. Introducing non-linearity is very much needed as most of the data in the real world is non-linear and it is very important that the neurons learn these nonlinear representations.

There are several activation functions:

- **Sigmoid** Figure 3.2: A sigmoid function takes a real-valued number as input and produces an output between zero and one.

$$\sigma(x) = 1 / (1 + \exp(-x))$$

Figure 3.3: Sigmoid function

- **tanh** Figure 3.4: A tanh function takes a real-valued number as input and produces an output between positive and negative one.

$$\tanh(x) = 2\sigma(2x) - 1$$

Figure 3.4: tanh function

- **ReLU** Figure 3.4: ReLU stands for Rectified Linear Unit. This function takes a real-valued number as input and produces as output the maximum of zero or the input number. That is, it swaps negative numbers with zeros.

$$f(x) = \max(0, x)$$

Figure 3.4: tReLU function

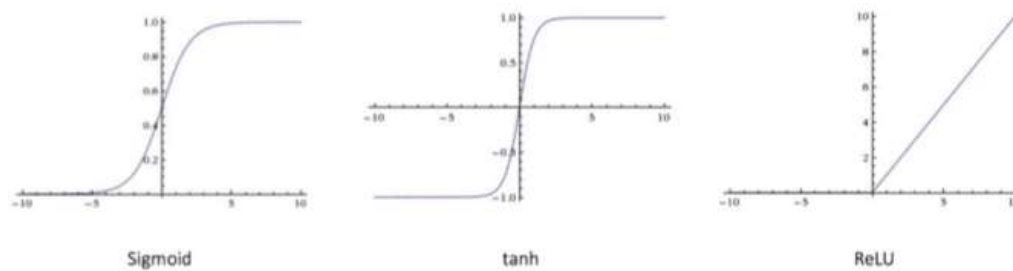


Figure 3.5: Represents the graphs of the three activation functions

3.3.4 Feedforward Neural Network

One of the very first neural networks devised is the Feedforward neural network. It contains multiple layers of nodes or neurons. Nodes from adjacent layers have edges or connections with certain weights associated between them. The information in a feed-forward neural network flows only in the forward direction. It starts at the input layer, flows through the hidden layer and then reaches the output layer. Feed-forward neural network Figure 3.6 has no loops or cycles in it.

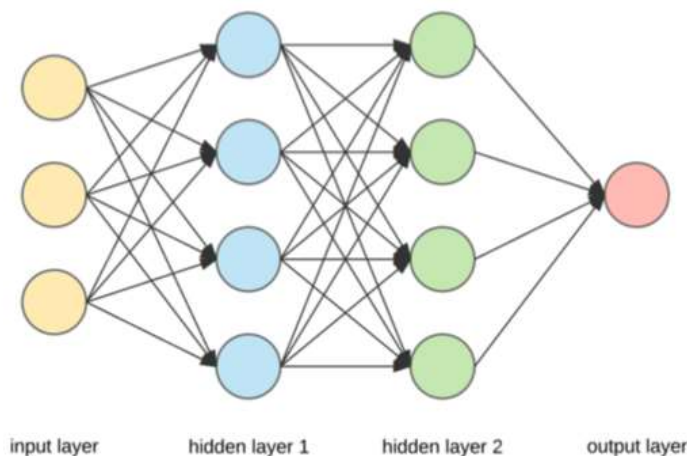


Figure 3.6 represents a simple feedforward network.

3.3.5 Single-layer perceptron network

A Single-layer perceptron network Figure 3.7 is one of the simplest kinds of neural networks. A single-layer perceptron network consists of only one layer of output nodes and the inputs with associated weights are directly fed to the output nodes. Each node computes the sum of the product of the weights and inputs and if the computed result is greater than a fixed threshold value (> 0), the neuron takes the activated value (generally 1), otherwise takes the deactivated value (generally -1).

Learning is achieved through a simple learning technique called the delta rule. The error between the expected output and the predicted output is calculated and the weights are adjusted based on this error.

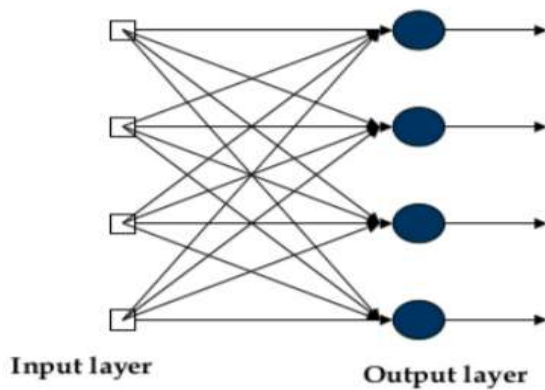


Figure 3.7 Represents a single-layer perceptron network.

3.3.6 Multi-layer perceptron network

Multi-layer perceptron networks as the name suggests, consists of multiple layers of neurons. Neurons in all the layers are connected to each other. Multi-layer networks use several different learning techniques and back propagation is the most used one among them.

In this learning technique, the predicted outputs are compared to the ground truth values and the error is calculated according to an error function that is predefined. This error is then fed into the system, using which, the weights of each input are adjusted with the aim to decrease the error by a small fraction. The same process is repeated for many iterations after which the model converges to a state where the error is small. With this, we could say that the model has *learned* a specific target function.

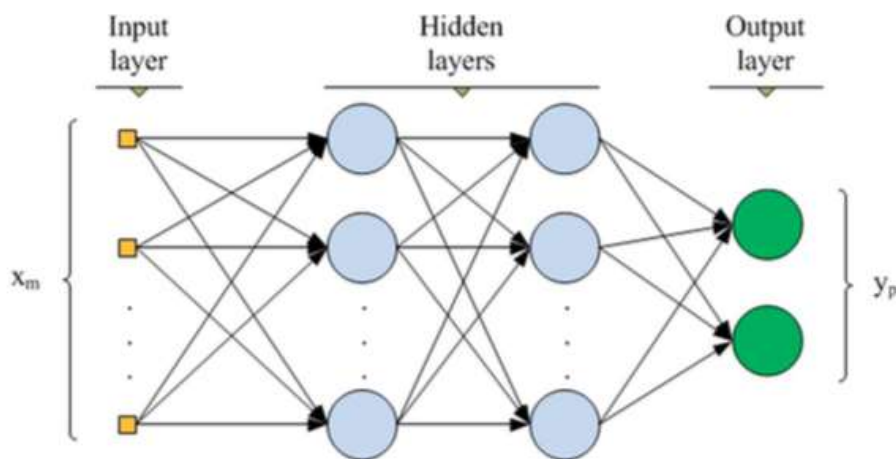


Figure 3.8 Represents a multi-layer perceptron network.

In the next section, we study in detail about convolutional neural networks, their workings and how they are used in our model.

3.4 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a specific class of neural networks that perform exceptionally well in image classification and recognition and other computer vision tasks. Convolutional Networks have been very effective in identifying objects, human faces and traffic signals apart from rendering vision in self driving cars and robots.

In the early 1980's, Fukushima in his work on Neocognitron model [a], started experimenting with convolutional layers. He worked on character recognition problems with a convolutional layer along with a subsampling layer called the S-layer. He then classified images into various classes using this architecture. In the year 1998, LeCun et al. used a few convolutional layers along with a few fully connected layers to identify digits that were handwritten. This model was named LeNet5 and was trained using complex algorithms such as backpropagation.

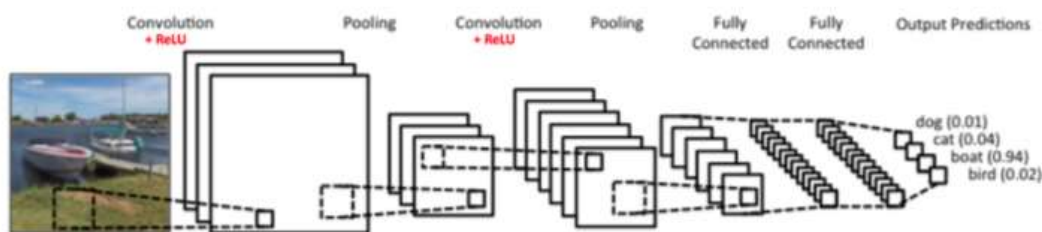


Figure 3.8 A convolutional neural network model.

All of the convolutional models will consist of the following four operations, which can be considered as the building blocks of CNNs:

1. 1. Convolution
2. Non-Linearity (ReLU)
3. Pooling or Sub Sampling
4. Fully-connected Layer

3.4.1 Convolution

Any picture taken with a standard camera will represent the picture in three channels – blue, green and red. These channels can be imagined as a 2-D matrix with pixel values ranging from 0 to 255. In the case of a grey-scale image, we will have only one channel with pixel values ranging between 0 and 255, where values closer to 0 represent blacks and values closer to 255 represent whites. The main purpose of a convolution operation in convolutional neural networks is to collect image features from a given image. It learns the image features by preserving the spatial relationship between the image pixels.

Consider a 5x5 image with pixel values 0 and 1 Figure 3.10(a) and Consider a 3x3 matrix Figure 3.10(b)

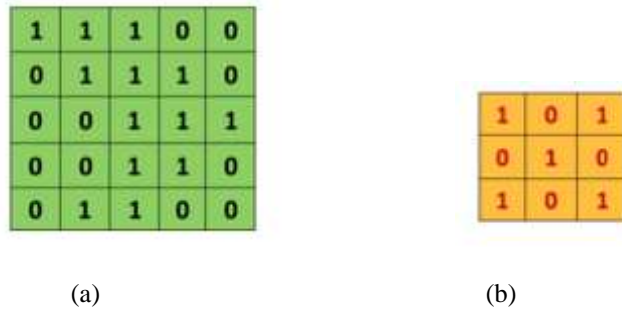


Figure 3.10 (a)Represents 5x5 image with pixel values 0 and 1 (b) Represents a 3x3 matrix

Now, the convolution of the 5x5 image with the 3x3 matrix can be calculated as Figure 3.11:

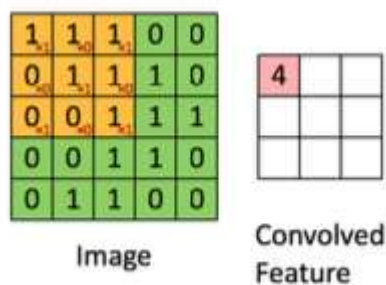


Figure 3.11 calculation of convolved image using 3x3 filter

The 3x3 matrix is made to slide over the input image pixel by pixel and at each pixel, multiplication is done element-wise, and all the multiplication outputs are summed together to get a value which is added as a value in the resultant matrix. The 3x3 matrix is called a 'filter' or a 'kernel'. And the matrix that is the result of sliding the filter over the input image is called the 'Feature Map' or 'Convolved Feature'. Different filter matrices generate different feature maps for the same input. That is different filters can extract different image features such as color information, boundaries and edges, pattern, etc.

The values of the filter are learned by the CNN model on its own during the training process, but the parameters such as filter size, number of filters, etc. should be specified before the training. As the filters are increased, more image features are extracted and more effective recognition of patterns is achieved.

Three parameters are used to tweak the size of the generated feature map: stride, zero-padding and depth.

- **Stride:** Stride corresponds to the count of pixels on which we move the filter over the input image.
- **Zero-padding:** Zero-padding lets us change and modify the size of the filter. To apply filter to the bordering elements of the image matrix, the border of the input matrix is padded with zeros.
- **Depth:** Depth is the count of filters used for the convolution operation.

3.4.2 Non-linearity: (ReLU)

ReLU stands for Rectified Linear Unit. ReLU is majorly used to add non-linearity in the network as the majority of the real-world information that our convolutional network would learn would be non-linear. ReLU is applied per pixel and replaces every non-negative pixel value in the image by a zero. Figure 9 shows a graph that represents a ReLU function.

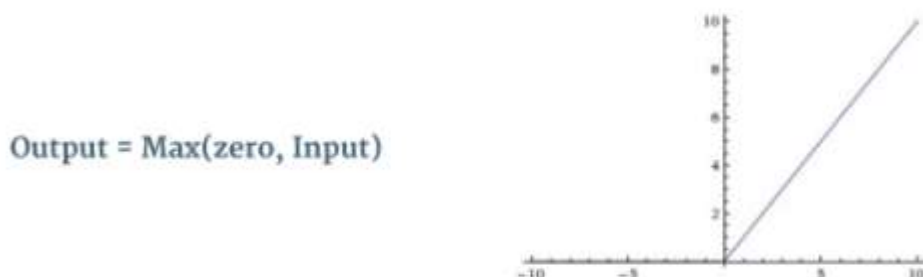


Figure 3.12 Graph Representing ReLU function

3.4.3 Pooling or Sub-sampling

Pooling or Sub-sampling retains the most important information from a feature map while reducing its dimensionality. It can be of various types: max pooling, average pooling, sum pooling, etc.

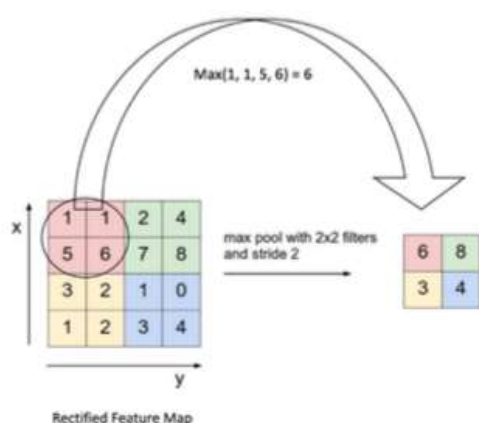


Figure 3.13 Represents the max pooling operation.

In max pooling, a spatial region is defined and the maximum value from the feature map within that spatial region is considered. Rather than considering the maximum element, we could consider the average of all the elements in that region (average pooling) or take the sum of all the elements in that region (sum pooling). Max

pooling performs best when compared to the rest.

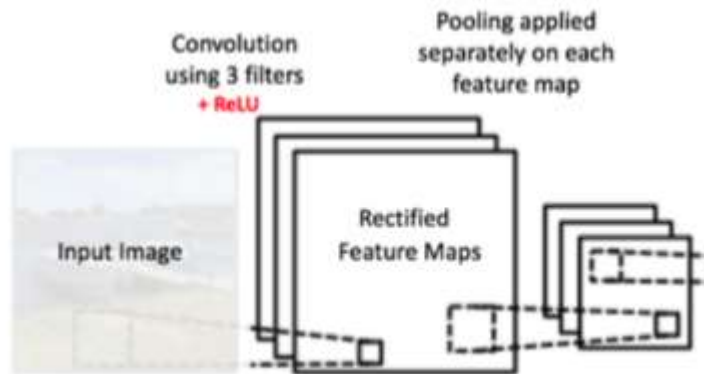


Figure 3.14 Represents a max pooling operation on a given input image.

In addition to reducing the size of the input representation, pooling,

- controls overfitting by reducing the number of parameters and computations needed for processing.
- turns the model invariant to minor distortions and transformations in the image as we only consider either the max values or average values from the local neighborhood.
- outputs a scale-invariant representation of the image so that objects can be detected anywhere in the image.

3.4.4 Fully-connected layer

The results from the previous convolutional and spatial pooling layers represent the features of the image from a high level. These features are passed as input to the fully-connected layer which classifies the image into various predefined categories depending on the training data. The features that are extracted from the convolutional and pooling layers might work well for classification, but a combination of those features will give even better results. Adding a fully-connected layer is an effective way of learning these non-linear combinations.

By choosing to apply a softmax function in the last layer of the fully-connected layer, it is ensured that the sum of probabilities from the fully connected layer is one. A sigmoid function takes a real-valued number as input and produces an output between zero and one, that sum to one. Figure 3.15 shows the layers in a convolutional neural network.

By summarizing all the operations and layers, the training phase in a convolutional neural network can be described in the following steps:

- **Step 1:** Use random values to initialize the weights, filters and parameters.
- **Step 2:** For a given training image as input, compute the output probabilities of each class with the forward propagation step (convolutional layer + ReLU + pooling layer + fully-connected layer).
- **Step 3:** Compute the final error in the last layer.
- **Step 4:** Compute the gradients of the error with respect to all the weights in the model by using back propagation. Then adjust the weights, parameter and filter values to reduce the error. By this time, the model should be able to output the correct class of the given image.
- **Step 5:** Reiterate steps 2-4 with the complete training set.

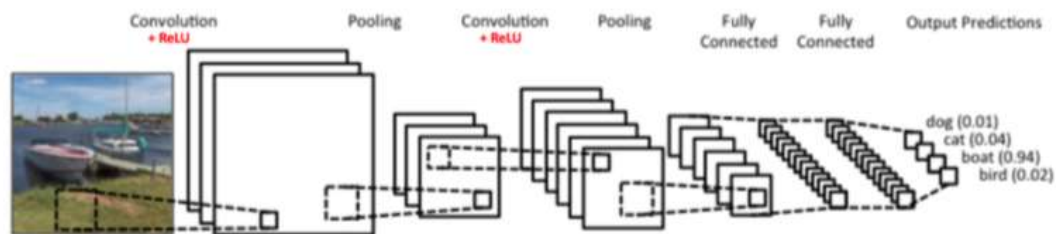


Figure 3.15 shows the layers in a convolutional neural network.

Given a training set that is considerably huge, the convolutional network will work well even with unseen images and will be able to put them correctly into their respective categories.

Deep convolutional neural network models may take days or even weeks to train on very large datasets.

A way to short-cut this process is to re-use the model weights from pre-trained models that were developed for standard computer vision benchmark datasets, such as the ImageNet image recognition tasks. Top performing models can be downloaded and used directly, or integrated into a new model for your own computer vision problems.

In the next section, **Transfer learning** is discussed.

3.5 What Is Transfer Learning?

Transfer learning generally refers to a process where a model trained on one problem is used in some way on a second related problem.

In deep learning, transfer learning is a technique whereby a neural network model is first trained on a problem similar to the problem that is being solved. One or more layers from the trained model are then used in a new model trained on the problem of interest.

Transfer learning has the benefit of decreasing the training time for a neural network model and can result in lower generalization error.

The weights in re-used layers may be used as the starting point for the training process and adapted in response to the new problem. This usage treats transfer learning as a type of weight initialization scheme. This may be useful when the first related problem has a lot of more labeled data than the problem of interest and the similarity in the structure of the problem may be useful in both contexts.

3.5.1 How to Use Pre-Trained Models

Models may be downloaded and used as feature extraction models. Here, the output of the model from a layer prior to the output layer of the model is used as input to a new classifier model.

Recall that convolutional layers closer to the input layer of the model learn low-level features such as lines, that layers in the middle of the layer learn complex abstract features that combine the lower level features extracted

from the input, and layers closer to the output interpret the extracted features in the context of a classification task.

Armed with this understanding, a level of detail for feature extraction from an existing pre-trained model can be chosen. For example, if a new task is quite different from classifying objects in photographs (e.g. different to ImageNet), then perhaps the output of the pre-trained model after the few layers would be appropriate. If a new task is quite similar to the task of classifying objects in photographs, then perhaps the output from layers much deeper in the model can be used, or even the output of the fully connected layer prior to the output layer can be used.

The pre-trained model can be used as a separate feature extraction program, in which case input can be pre-processed by the model or portion of the model to a given an output (e.g. vector of numbers) for each input image, that can then use as input when training a new model.

Alternatively, the pre-trained model or desired portion of the model can be integrated directly into a new neural network model. In this usage, the weights of the pre-trained can be frozen so that they are not updated as the new model is trained. Alternatively, the weights may be updated during the training of the new model, perhaps with a lower learning rate, allowing the pre-trained model to act like a weight initialization scheme when training the new model.

3.5.2 Models for Transfer Learning

There are perhaps a dozen or more top-performing models for image recognition that can be downloaded and used as the basis for image recognition and related computer vision tasks.

Perhaps three of the more popular models are as follows:

- VGG (e.g. VGG16 or VGG19).
- GoogLeNet (e.g. InceptionV3).
- Residual Network (e.g. ResNet50).

These models are both widely used for transfer learning both because of their performance, but also because they were examples that introduced specific architectural innovations, namely consistent and repeating structures (VGG), inception modules (GoogLeNet), and residual modules (ResNet).

Keras provides access to a number of top-performing pre-trained models that were developed for image recognition tasks.

Chapter IV

4.1 Dataset

Kaggle, a subsidiary of Google LLC, is an online community of data scientists and machine learning practitioners. Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

Kaggle got its start in 2010 by offering machine learning competitions and now also offers a public data platform, a cloud-based workbench for data science, and Artificial Intelligence education. Its key personnel were Anthony Goldbloom and Jeremy Howard. Nicholas Gruen was the founding chair succeeded by Max Levchin. Equity was raised in 2011 valuing the company at \$25 million. On 8 March 2017, Google announced that they were acquiring Kaggle.

So, using such a competitive platform is typically a “Diabetic Retinopathy Detection” dataset. This project was carried out.

Dataset contained 35,000 pictures of a total of 40 gigabytes. 4,000 pictures were extracted for validation and another 4,000 for testing purposes.

The following paragraph describes the dataset as stated from its collector.

A clinician has rated the presence of diabetic retinopathy in each image on a scale of 0 to 4, according to the following scale:

- 0 - No DR
- 1 - Mild
- 2 - Moderate
- 3 - Severe
- 4 - Proliferative DR

Your task is to create an automated analysis system capable of assigning a score based on this scale.

The images in the dataset come from different models and types of cameras, which can affect the visual appearance of left vs. right. Some images are shown as one would see the retina anatomically (macula on the left, optic nerve on the right for the right eye). Others are shown as one would see through a microscope condensing lens (i.e. inverted, as one sees in a typical live eye exam). There are generally two ways to tell if an image is inverted:

It is inverted if the macula (the small dark central area) is slightly higher than the midline through the optic nerve. If the macula is lower than the midline of the optic nerve, it's not inverted.

If there is a notch on the side of the image (square, triangle, or circle) then it's not inverted. If there is no notch, it's inverted.

Like any real-world data set, you will encounter noise in both the images and labels. Images may contain artifacts, be out of focus, underexposed, or overexposed. A major aim of this competition is to develop robust algorithms that can function in the presence of noise and variation.

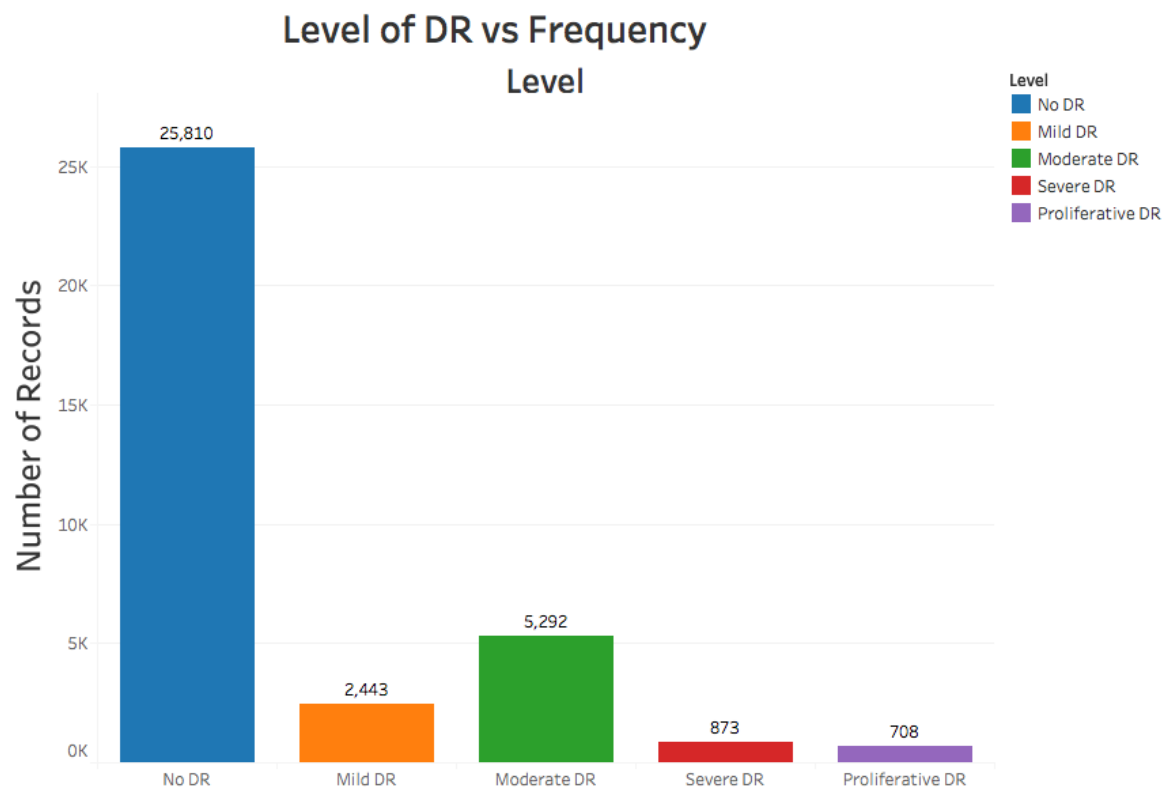


Figure 4.1 Represents distribution of images among different stages of DR

However, in this project. Two classes only were compared and classified and there were people with diabetic retinopathy and healthy people with no diabetic.

Hence, a new graph was better to represent actual data used as below:



Figure 4.2 Distribution of images after Combining images with DR

4.2 Classical Machine Learning Results

Given the techniques in former chapters. The following graphs are obtained and the comparison is performed.

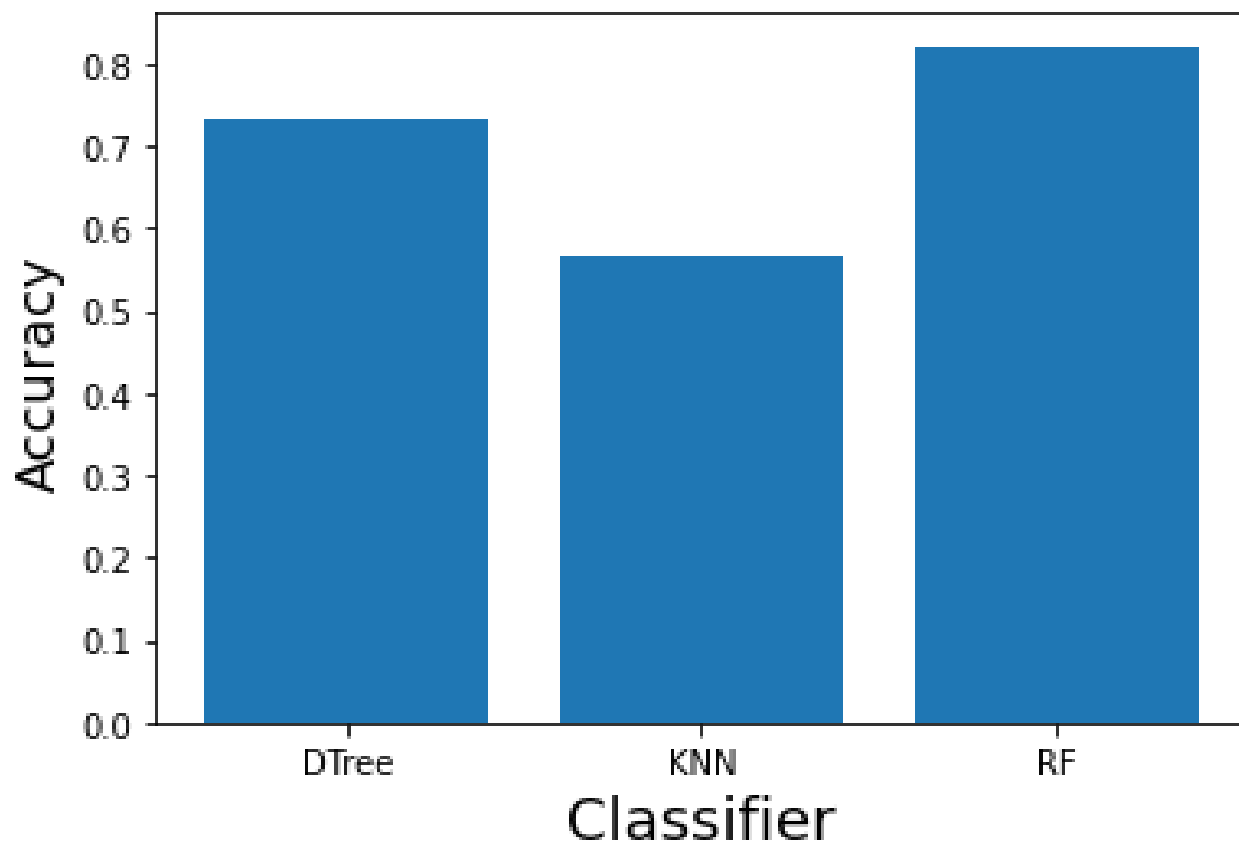


Figure 4.3 Comparison between Decision Tree, KNN and Random forest classifiers' Accuracy

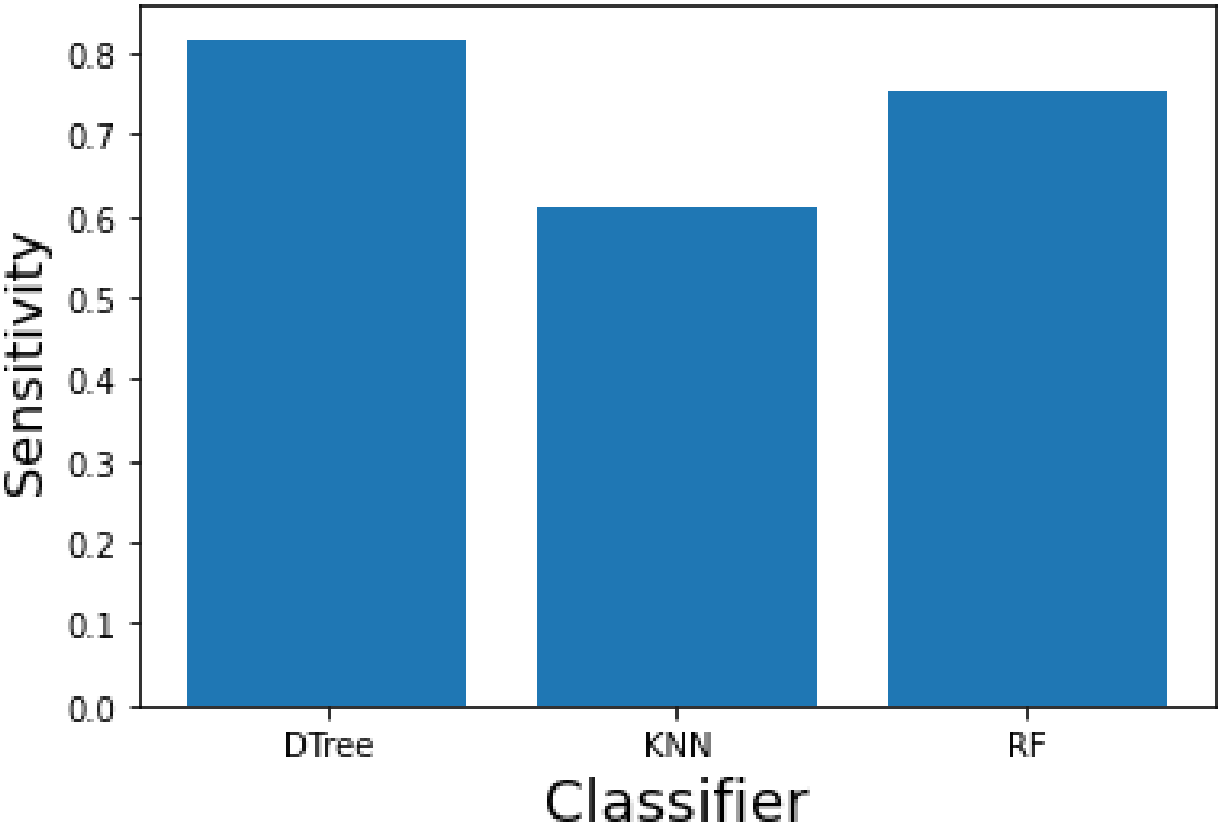


Figure 4.4 Comparison between Decision Tree, KNN and Random forest classifiers’ Sensitivity

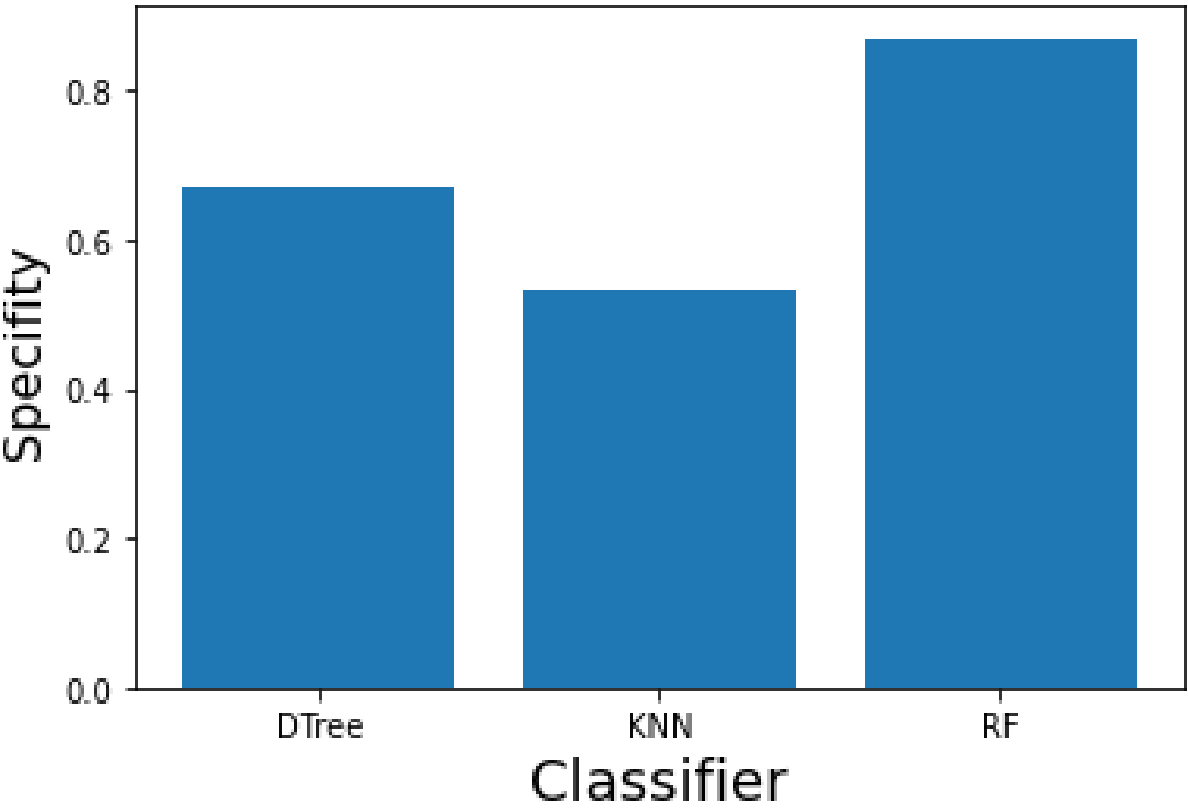


Figure 4.5 Comparison between Decision Tree, KNN and Random forest classifiers’ Specificity

As shown above, only best classifiers were selected for comparison and usage which are Decision tree (DTree), K-nearest-neighbor (KNN) and Random forest (RF). All of their metrics including accuracies, sensitivity and specificity with precision are included in table (4.1) below.

| Classifier | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---------------|--------------|-----------------|-----------------|
| Decision Tree | 73.3 | 81.7 | 67.1 |
| KNN | 59 | 61 | 75.4 |
| RF | 82 | 75 | 86.9 |

Table (4.1) Summarizes the results of the three classifiers

4.3 Neural Networks Results

Two neural networks are being trained and tested for such purposes.

Also, for Increasing training speed in both architectures an Optimizer called “Adam” was used.

This optimizer is meant to change the learning rate α after each epoch in order to avoid reaching the local minimum.

Before feeding the CNN with data. Data was preprocessed, data augmentation was done, and a resize to all photos as well to a unified size of 256x256 pixels.

512 workers were used to speed up the training process too, these workers helped transfer data from RAM and feed it to the GPU.

25 Successful epochs on inception were done. and 15 for the Resnet.

With a batch size of 64 photographs.

4.3.1 First CNN: Inception

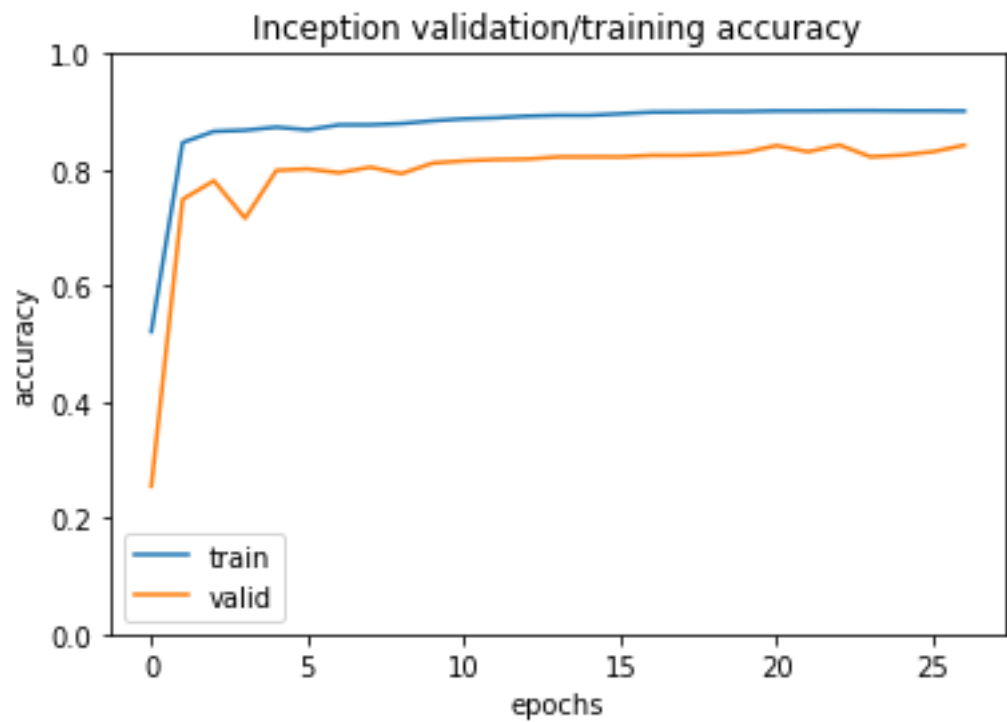


Figure 4.6 learning curves for training and validation accuracy for Inception model

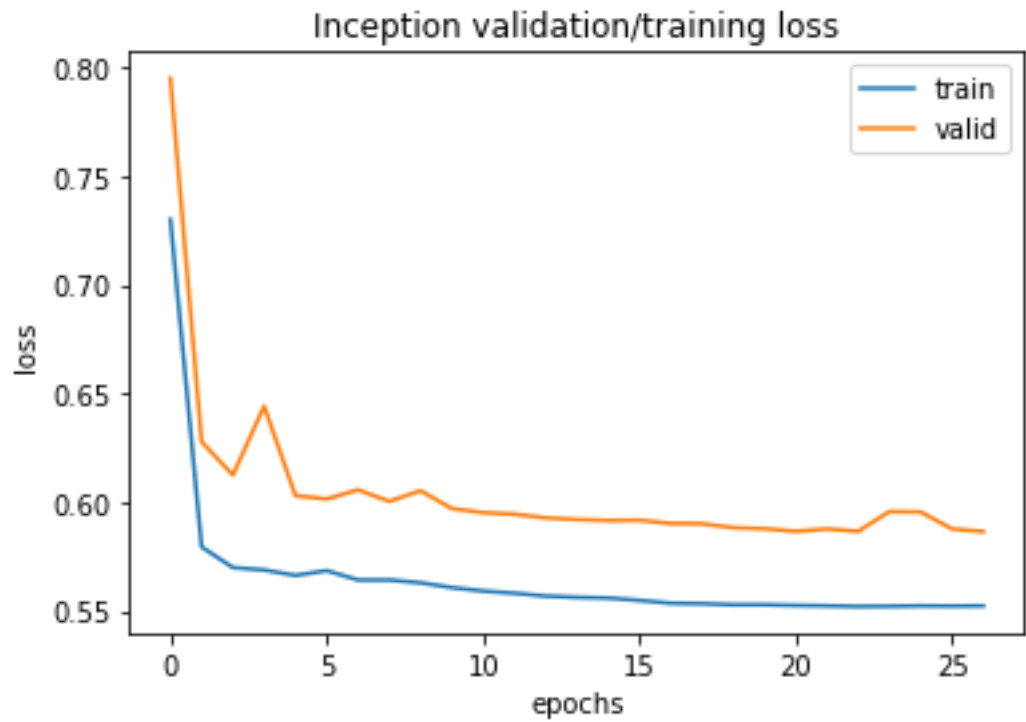


Figure 4.7 learning curves for training and validation loss for Inception model

4.3.2 Second CNN: ResNet

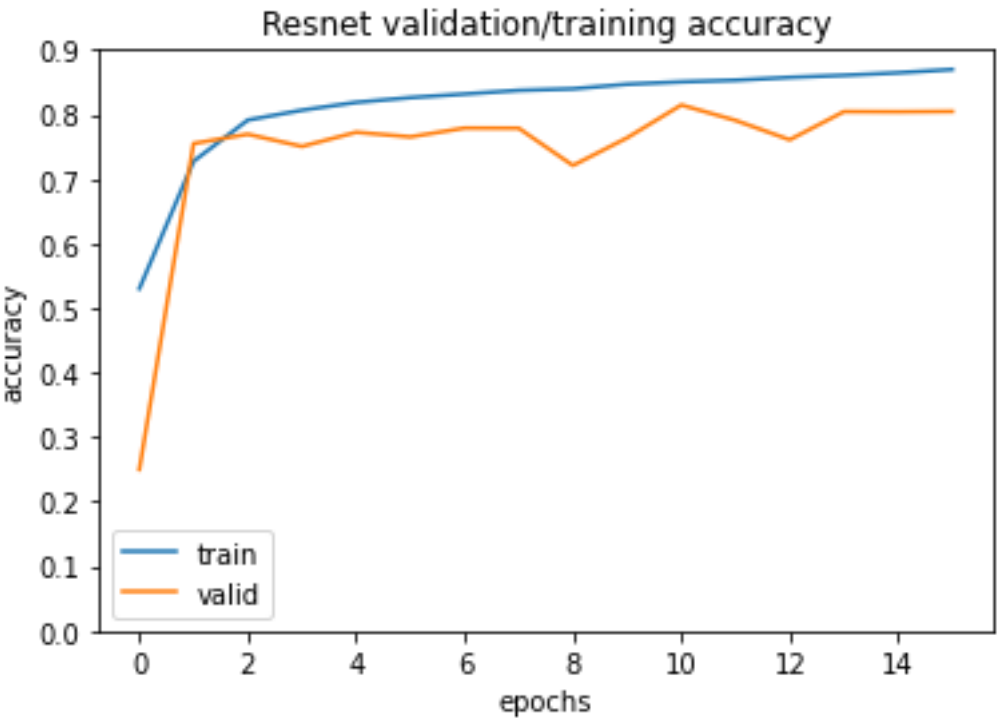


Figure 4.8 learning curves for training and validation accuracy for ResNet model



Figure 4.9 learning curves for training and validation loss for ResNet model

Below are the confusion matrices of both architectures Table 4.2 and Table 4.3

| | ND | DR |
|----|------|-----|
| ND | 2668 | 332 |
| DR | 355 | 645 |

Table 4.2 Inception model confusion matrix

| | ND | DR |
|----|------|-----|
| ND | 2535 | 465 |
| DR | 328 | 672 |

Table 4.3 ResNet model confusion matrix

4.3.3 Comparing two CNN's results

| | Accuracy | Precision | Recall (Sensitivity) | Specificity |
|-------------|----------|-----------|-------------------------|-------------|
| InceptionV3 | 82% | 88.2% | 88.3% | 88.2% |
| ResNet | 70% | 88.5% | 84.5% | 86.5% |

Table 4.4 Shows Comparison in Accuracy, Precision, Recall and Specificity between Inception and ResNet Models.

As shown above in Table 4.4, Resnet architecture is less accurate by 2%. So, the Inception model was preferred and hence confirmed.

Below, you can see each Table (4.5) and Table (4.6) with significant changes in accuracies of inception model after each modification we made as follows:

| | |
|----------------------------|-----|
| Data without preprocessing | 75% |
| Data with preprocessing | 76% |
| Data with augmentation | 78% |
| After hyper tuning | 83% |

Table 4.5 Shows the effect of preprocessing, augmentation and hyperparameter tuning on Inception model's accuracy

Then when applying the above data to the ResNet architecture CNN yields:

| | |
|---------------------|-----|
| ResNet architecture | 80% |
|---------------------|-----|

Table 4.6 Shows ResNet Accuracy after applying preprocessing, augmentation and hyperparameter tuning.

Chapter V

Conclusion and Future work

5.1 Conclusion

This thesis described novel methods for automatic analysis of retinal images with the goal of automated diabetic retinopathy screening. In automated screening a computer system analyzes retinal images before an ophthalmologist does.

To achieve this objective, state of the art algorithms for the automatic analysis of fundus images are presented, with particular attention to the aspects related to diabetic retinopathy and retinal morphology. An introduction to the required medical terms and concepts was given, together with the common protocols to manually diagnose and grade diabetic retinopathy with fundus images. Different retinal imaging techniques were described and a rationale for the choice fundus images was given.

An automatic quality assessment algorithm able to generate a quality score in a short amount of time was introduced. Its performances were evaluated with a small dataset, but more importantly in a live screening environment of a teleophthalmology network.

Two new lesion segmentation algorithms were developed, one to target exudates, one microaneurysms. Both of them have been compared with similar methods and proved to have competitive performance. As far as the computation time is concerned, they are able to estimate the lesions in a short amount of time.

The Deep convolutional Neural Network proposed approach is a wholesome approach to diagnose diabetic retinopathy. The used CNN architectures with dropout techniques yielded significant classification accuracy.

5.2 Future work

Overall, this work has presented several advancements in retinal image processing geared toward improving the state-of-the-art in automated screening, from the acquisition phase, through quality assessment/enhancements, lesion detection and disease classification. Hence, there are various directions that can be pursued for further studies such as:

- **Combining Handcrafted Image features with CNN to improve accuracy and performance.**

- **Implement Application for Diabetic Retinopathy diagnosis:** WHO believes that 80% of all visual impairment caused by DR can be cured with regular screening and early detection. Diabetes mellitus (DM) and its complications including DR has become one of the 21st century's major health challenges. Emerging healthcare technologies focus on reducing unnecessary visits to medical specialists, minimizing the overall cost of treatment and optimizing the number of patients seen by each doctor. Major vision loss due to DR is preventable with timely remedial intervention like eye exams (including visual acuity testing, tonometry and pupil dilation) and regular screening at the earlier stages by using a DR diagnostic and grading screening tool. This DR tool can assist the doctor with patient fundus image analysis, which in turn helps to quickly inform the next steps in the patient's treatment. Also, doctors can attend to more patients that need attention. Note: The tool is not a replacement for a doctor.
- **Extend the project to predict cardiovascular diseases:** Recent studies showed that cardiovascular risk can be extracted from retinal images, which can be obtained quickly, cheaply and non-invasively in an outpatient setting. Markers of cardiovascular disease, such as hypertensive retinopathy and cholesterol emboli, can often manifest in the eye. Furthermore, because blood vessels can be non-invasively visualized from retinal fundus images, various features in the retina may reflect the systemic health of the cardiovascular system as well as future risk. These multiple cardiovascular risk factors from retinal images can be extracted using deep learning.
- **Collect different Datasets to extend the project to predict different diseases:** Since Medical Images diagnosis is a very complex problem it needs large amounts of data to train well. Therefore, the main challenge to extend the project to diagnose different diseases is to collect amounts of suitable data while considering patients' confidentiality and privacy.

Bibliography

- [1] Baudoin, C. E., Lay, B. J., and Klein, J. C. Automatic detection of microa-
 - i. neurysms in diabetic fluorescein angiography. *Revue DEpid´emiologie et de Sante Publique* 32 (1984), 254 – 261.
 - ii. neurysms in diabetic fluorescein angiography. *Revue DEpid´emiologie et de Sante Publique* 32 (1984), 254 – 261.
- [2] Bhalerao, A., Patanaik, A., Anand, S., and Saravanan, P. Robust detection of microaneurysms for sight threatening retinopathy screening. In *Proceedings of the Sixth Indian Conference on Computer Vision, Graphics and Image Process-ing (ICGVIP2008)* (2008), pp. 520 – 527.
- [3] David, J., Krishnan, R., and Kumar, A. S. Neural network based retinal image analysis. In *Proceedings of the International Congress on Image and Signal Processing (CISP2008)* (2008), vol. 2, pp. 49 – 53.
- [4] Diabetic Retinopathy (online). Current Care Guideline. Working group set up by the Finnish Medical Society Duodecim and the Finnish Respiratory Society. Helsinki: Finnish Medical Society Duodecim, 2006 [referred 19.1.2010]. Available online at: www.kaypahoito.fi.
- [5] Ege, B. M., Hejlesen, O. K., Larsen, O. V., Mller, K., Jennings, B., Kerr, D., and Cavan, D. A. Screening for diabetic retinopathy using computer-based image analysis and statistical classification. *Computer Methods and Programs in Biomedicine* 62 (2000), 165 – 175.
- [6] Enrico Grisan, A. R. Segmentation of candidate dark lesions in fundus images based on local thresholding and pixel density. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS2007)* (2007), pp. 6735 – 6738.
- [7] Estabridis, K., and de Figueiredo, R. J. P. Automatic detection and diagnosis of diabetic retinopathy. In *Proceedings of the International Conference on Image Processing (ICIP2007)* (2007), vol. 2, pp. 445 – 448.
- [8] Eswaran, C., Reza, A. W., and Hati, S. Extraction of the contours of optic disc and exudates based on marker- controlled watershed segmentation. In *Proceedings of the International Conference on Computer Science and Information Technology (ICCSIT2008)* (2008), pp. 719 – 723.
- [9] Fleming, A. D., Goatman, K. A., Philip, S., Williams, G. J., Prescott,
 - a. J., Scotland, G. S., McNamee, P., Leese, G. P., Wykes, W. N., Sharp,
 - b. F., and Olson, J. A. The role of hemorrhage and exudate detection in automated grading of diabetic retinopathy. *British Journal of Ophthalmology* 94 (2009), 706 – 711.
- [10] Fleming, A. D., Goatman, K. A., Williams, G. J., Philip, S., Sharp,
 - a. F., and Olson, J. A. Automated detection of blot hemorrhages as a sign of referable diabetic retinopathy. In *Proceedings of the Medical Image Understanding and Analysis (MIUA2008)* (2008), pp. 235 – 239.
- [11] Fleming, A. D., Philip, S., Goatman, K. A., Williams, G. J., Olson, J. A., and Sharp, P. F. Automated detection of exudates for diabetic retinopathy screening. *Physics In Medicine And Biology* 52 (2007).

-
- [12] Garc'ia, M., Sanchez, C. I., Lopez, M. I., D'iez, A., and Hornero, R. Automatic detection of red lesions in retinal images using a multilayer perceptron neural network. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS2008)* (2008), pp. 5425 – 5428.
 - [13] Goh, J., Tang, L., Saleh, G., Turk and Yu Fu, L. A., and Browne, A. Filtering normal retinal images for diabetic retinopathy screening using mul-tiple classifiers. In *Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine (ITAB2009)* (2009), pp. 1 – 4.
 - [14] Goh, K. G., Hsu, W., Lee, M. L., and Wang, H. Adris: an automatic diabetic retinal image screening system. In *Medical Data Mining and Knowledge Discovery*, vol. 60 of *Studies in Fuzziness and Soft Computing*. 2001.
 - [15] Goldbaum, M., Boyd, S. M. A. T. S. C. J., Hunter, E., and Jain, R. Automated diagnosis and image understanding with object extraction, object clas-sification, and inferencing in retinal images. In *Proceedings of the International Conference on Image Processing (ICIP1996)* (1996), vol. 3, pp. 695 –. 698.
 - [16] Goldbaum, M. H., Katz, N. P., Chaudhuri, S., and Nelson, M. Image understanding for automated retinal diagnosis. In *Proceedings of the Annual Symposium on Computer Application in Medical Care* (1989), vol. 8, pp. 756 – 760.
 - [17] Goldbaum, M. H., Katz, N. P., Nelson, M. R., and Haff, L. R. The discrimination of similarly colored objects in computer images of the ocular fundus. *Investigative Ophthalmology & Visual Science* 31 (1990).
 - [18] Hanley, J. A., and McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology* 143 (1982), 39 – 36.
 - [19] Hansen, A. B., Hartvig, N. V., Jensen, M. S., Borch-Johnsen, K., Lund-Andersen, H., and Larsen, M. Diabetic retinopathy screening using digital non-mydratic fundus photography and automated image analysis. *Acta Ophthalmologica Scandinavica* 82 (2004).
 - [20] Hsu, W., Pallawala, P. M. D. S., Lee, M. L., and Eong, K.-G. A. The role of domain knowledge in the detection of retinal hard exudates. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2001)* (2001), vol. 2, pp. 246 – 251.
 - [21] Kristinsson, J. K., Hauksdottir, H., Stefansson, E., Jonasson, F., and Gislason, I. Active prevention in diabetic eye disease. *Acta ophthalmologica Scandinavica* 75, 3 (June 1997), 249–54.
 - [22] Kumari, V. V., Suriyahanayagam, N., and Saranya, C. T. Feature extrac-tion for early detection of diabetic retinopathy. In *Proceedings of the International Conference on Recent Trends in Information, Telecommunication and Computing (ITC2010)* (2010), pp. 359 – 361.
 - [23] Larsen, N., Godt, J., Grunkin, M., Lund-Andersen, H., and Larsen, M. Automated detection of diabetic retinopathy in a fundus photographic screening population. *Investigative Ophthalmology & Visual Science* 44, 2 (2003).
 - [24] Li, H., and Chutatape, O. A model-based approach for automated feature extraction in fundus images. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV2003)* (2003), vol. 1, pp. 394 – 399.

-
- [25] Li, H., and Chutatape, O. Automated feature extraction in color retinal images by a model-based approach. *IEEE Transactions On Biomedical Engineering* 51, 2 (2004), 246 – 254.
 - [26] Luo, G., Chutatape, O., Lei, H., and Krishnan, S. M. Abnormality detection in automated mass screening system of diabetic retinopathy. In *Proceedings of the IEEE Symposium on Computer-Based Medical Systems (CBMS2001)* (2001), pp. 132 – 137.
 - [27] Margolis, R., and Kaiser, K. Diagnostic modalities in diabetic retinopathy. In *Diabetic retinopathy*, E. Duh, Ed. Humana Press, 2008, pp. 109–133.
 - [28] Nayak, J., Bhat, P. S., U, R. A., Lim, C. M., and Kagathi, M. Automated identification of diabetic retinopathy stages using digital fundus images. *Journal of Medical Systems* 32 (2008), 107 – 115.
 - [29] Niemeijer, M., Abramoff, M. D., and van Ginneken, B. Information fusion for diabetic retinopathy cad in digital color fundus photographs. *IEEE Transactions on Medical Imaging* 28, 5 (2009), 775 – 785.
 - [30] Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., and Abramoff, M. D.
 - a. Comparative study of retinal vessel segmentation a new publicly available database. In *Proceedings of the SPIE Medical Imaging* (2004), pp. 648–656.
 - [31] Niemeijer, M., van Ginneken, B., Russell, S. R., and Abramoff, M. D. Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. *Investigative Ophthalmology and Visual Science* 48, 5 (2007), 2260–2267.
 - [32] Niemeijer, M., van Ginneken, B., Staal, J., Suttorp-Schulten, M. S. A.,
 - i. and Abramoff, M. D. Automatic detection of red lesions in digital color fundus photographs. *IEEE Transactions on Medical Imaging* 24, 5 (May 2005), 584 – 592.
 - [33] Pallawala, P. M. D. S., Hsu, W., Lee, M. L., and Goh, S. S. Auto-mated microaneurysm segmentation and detection using generalized eigenvectors. In *Proceedings of the Seventh IEEE Workshop on Applications of Computer Vision (WACV/MOTION2005)* (2005), vol. 1, pp. 322 – 327.
 - [34] Phillips, R., Forrester, J., and Sharp, P. Automated detection and quantification of retinal exudates. *Graefe's Archive for Clinical and Experimental Ophthalmology* 231, 2 (1993), 90 – 94.
 - [35] Quellec, G., Lamard, M., Josselin, P. M., Cazuguel, G., Cochener, B., and Roux, C. Detection of lesions in retina photographs based on the wavelet transform. In *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2006)* (2006), pp. 2619 – 2621.
 - [36] Quellec, G., Lamard, M., Josselin, P. M., Cazuguel, G., Cochener, B., and Roux, C. Optimal wavelet transform for the detection of microaneurysms in retina photographs. *IEEE Transactions On Medical Imaging* 27, 9 (September 2008), 1230 – 1241.
 - [37] Ram, K., and Sivaswamy, J. Multi-space clustering for segmentation of exudates in retinal color photographs. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS2009)* (2009), pp. 1437 – 1440.
 - [38] Ravishankar, S., Jain, A., and Mittal, A. Automated feature extraction for early detection of diabetic retinopathy in fundus images. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2009)* (2009), pp. 210 – 217.

-
- [39] Reza, A. W., and Eswaran, C. A decision support system for automatic screen-ing of non-proliferative diabetic retinopathy. *Journal of Medical Systems* 4 (2009), 1–8.
- [40] Sanchez, C., Hornero, R., Lopez, M., and Poza, J. Retinal image analysis to detect and quantify lesions associated with diabetic retinopathy. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS2004)* (2004), vol. 1, pp. 1624 – 1627.
- [41] Savolainen, E., and Lee, Q. Diabetic retinopathy - need and demand for photocoagulation and its cost-effectiveness: evaluation-based services in the United Kingdom. *Diabetologia* 23, 2 (August 1982), 138–140.
- [42] Silberman, N., Ahrlich, K., Fergus, R., and Subramanian, L. Case for automated detection of diabetic retinopathy. In *Proceedings of the AAAI Spring Symposium Series* (2010).
- [43] Sinthanayothin, C., Boyce, J. F., Williamson, T. H., Cook, H. L., Men-sah, E., Lal, S., and Usher, D. Automated detection of diabetic retinopathy on digital fundus images. *Diabetic Medicine* 19 (2002), 105 – 112.
- [44] Sinthanayothin, C., Kongbunkiat, V., Phoojaruenchanachai, S., and Singalavanija, A. Automated screening system for diabetic retinopathy. In *Proceedings of the International Symposium on Image and Signal Processing and Analysis (ISPA2003)* (2003), vol. 2, pp. 915 – 920.
- [45] Spencer, T., Olson, J. A., Mchardy, K. C., Sharp, P. F., and Forrester, a. V. An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. *Computers And Biomedical Research* 29 (1996), 284 – 302.
- [46] Spencer, T., Phillips, R. P., Sharp, P. F., and Forrester, J. V. Auto-mated detection and quantification of microaneurysms in fluorescein angiograms. *Graefe's Archive Ior Clinical and Experimental Ophthalmology* 230 (1992), 36 – 41.
- [47] Taylor, R. *Handbook of Retinal Screening in Diabetes*. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chishester, West Sussex PO10 8SQ, England, 2006.
- [48] Usher, D., Dumskyj, M., Himaga, M., Williamson, T. H., Nussey, S., and Boyce, J. Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. *Diabetic Medicine* 21 (2003), 84 –90.
- [49] von Wendt, G., Heikkila, K., and Summanen, P. Detection of retinal neovas-cularizations using 45° and 60° photographic fields with varying 45° fields simulated on a 60° photograph. *Acta Ophthalmologica Scandinavica* 80, 4 (2002), 372 – 378.
- [50] Wade, N. J. Image, eye, and retina (invited review). *Journal of the Optical Society of America. A, Optics, image science, and vision* 24, 5 (2007), 1229–1249.
- [51] Walter, T., and Klein, J.-C. A computational approach to diagnosis of diabetic retinopathy. In *Proceedings of the 6th Conference on Systemics, Cybernetics and Informatics (SCI2002)* (2002), pp. 521 – 526.

-
- [52] Walter, T., Klein, J.-C., Massin, P., and Erginay, A. A contribution of image processing to the diagnosis of diabetic retinopathy – detection of exudates in color fundus images of the human retina. *IEEE Transactions On Medical Imaging* 21, 10 (2002), 1236 – 1243.
 - [53] Wang, H., Hsu, W., Goh, K. G., and Lee, M. L. An effective approach to detect lesions in color retinal images. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2000)* (2000), vol. 2, pp. 181 – 186.
 - [54] Wilkinson, C. P., Ferris, F. L., Klein, R. E., Lee, P. P., Agardh, C. D., Davis, M., Dills, D., Kampik, A., Pararajasegaram, R., and Verdaguer, J. T. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology* 110, 9 (September 2003), 1677 – 1682.
 - [55] World Health Organization and the International Diabetes Federation. Diabetes action now: An initiative of the world health organization and the international diabetes federation, 2004. ISBN 92 4 159151.
 - [56] Xie, X., and Mirmehdi, M. Localising surface defects in random color textures using multiscale texem analysis in image eigenchannels. In *Proceedings of the International Conference on Image Processing (ICIP2005)* (Genova, Italy, 2005), vol. 3, pp. 837–840.
 - [57] Xu, L., and Luo, S. Support vector machine-based method for identifying hard exudates in retinal images. In *Proceedings of the IEEE Youth Conference on Information, Computing and Telecommunication (YC-ICT2009)* (2009), pp. 138 – 141.
 - [58] Yun, W. L., Acharya, U. R., Venkatesh, Y., Chee, C., Min, L. C., and Ng, E. Identification of different stages of diabetic retinopathy using retinal optical image. *Information Sciences* 178 (2008), 106 –121.
 - [59] Zhang, X., and Chutatape, O. Detection and classification of bright lesions in color fundus images. In *Proceedings of the International Conference on Image Processing (ICIP2004)* (2004), vol. 1, pp. 139 – 142.
 - [60] Zhang, X., and Chutatape, O. A svm approach for detection of hemorrhages in background diabetic retinopathy. In *Proceedings of International Joint Conference on Neural Networks* (2005), vol. 4, pp. 2435 – 2440.
 - [61] Zhang, X., and Chutatape, O. Top-down and bottom-up strategies in lesion detection of background diabetic retinopathy. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2005)* (June 2005), vol. 2, pp. 422 – 428.
 - [62] Zheng, L., Chutatape, O., and Shankar, M. K. Automatic image analysis of fundus photograph. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS1997)* (1997), vol. 2, pp. 524 – 525.
 - [63] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600–612.
 - [64] Ponomarenko, N., Lukin, V., Zelensky, A., Egiazarian, K., Carli, M., & Battisti, F. (2009). TID2008-a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10(4), 30–45.

-
- [65] Vasanthi Satyananda, K V Narayanaswamy, K Karibasappa, "Extraction of Exudates from the Fundus Images A Review".
- [66] Patwari Manjiri, Manza Ramesh, Rajput Yogesh, Saswade Manoj, Deshpande Neha, "Automated Localization of Optic Disk, Detection of Microaneurysms and Extraction of Blood Vessels to Bypass Angiography", Springer, Advances in Intelligent Systems and Computing. ISBN: 978-3-319-11933-5, DOI: 10.1007/978-3-319-11933-5_65. 2014
- [67] Meyer M.I., Galdran A., Mendonça A.M., Campilho A.. A Pixel-Wise Distance Regression Approach for Joint Retinal Optical Disc and Fovea Detection. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, LNCS, vol 11071, pp 39-47, 2018. doi:10.1007/978-3-030-00934-2_5
- [68] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S. (2017), "Dermatologist-level classification of skin cancer with deep neural networks", Vol. 542 No. 7639, pp. 115–118.
- [69] Negi A, Vernon SA. An overview of the eye in diabetes. *J R Soc Med.* 2003 Jun;96(6):266-72. doi: 10.1258/jrsm.96.6.266. PMID: 12782689; PMCID: PMC539505.
- [70] Razzak, M.I., Naz, S. and Zaib, A., 2018. Deep learning for medical image processing: Overview, challenges and the future. In *Classification in BioApps* (pp. 323-350). Springer, Cham.