BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF INFORMATION TECHNOLOGY
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF INTELLIGENT SYSTEMS
ÚSTAV INTELIGENTNÍCH SYSTÉMŮ

GENERATION OF SYNTHETIC RETINAL IMAGES WITH HIGH RESOLUTION

GENEROVÁNÍ SYNTETICKÝCH SNÍMKŮ SÍTNICE VE VYSOKÉM ROZLIŠENÍ

TERM PROJECT
SEMESTRÁLNÍ PROJEKT

AUTHOR **Bc. TOMÁŠ AUBRECHT**
AUTOR PRÁCE

SUPERVISOR **Prof. Ing., Dipl.-Ing. MARTIN DRAHANSKÝ, Ph.D.**
VEDOUCÍ PRÁCE

BRNO 2020

Brno University of Technology
Faculty of Information Technology

Department of Intelligent Systems (DITS) Academic year 2019/2020

Master's Thesis Specification

21968

Student: **Aubrecht Tomáš, Bc.**
Programme: Information Technology Field of study: Information Systems
Title: **Generation of Synthetic Retinal Images with High Resolution**
Category: Computer Graphics

Assignment:

1. Study the literature in the area of processing and recognition of human retinal images, especially in ophthalmology sources, incl. generation of synthetic retinal images.
2. Propose your own method for generation of synthetic retinal images in high resolution (over 12 Mpix).
3. Implement the proposed method from the previous point. Generate a database of at least 1,000 images without pathological damage.
4. Test the implemented solution from the previous point, compare your results with real retinal images (e.g. vessels distribution) and summarize the achieved results.

Recommended literature:

- BONALDI, Lorenza, et al. Automatic generation of synthetic retinal fundus images: vascular network. *Procedia Computer Science*, 2016, 90: 54-60.
- WONG, Tien Yin; TING, Daniel Shu Wei. Generative Adversarial Networks (GANs) for Retinal Fundus Image Synthesis. In: *Computer Vision-ACCV 2018 Workshops: 14th Asian Conference on Computer Vision, Perth, Australia, December 2-6, 2018, Revised Selected Papers*. Springer, 2019. p. 289.
- FIORINI, Samuele, et al. Automatic Generation of Synthetic Retinal Fundus Images. In: *Eurographics Italian Chapter Conference*. 2014. p. 41-44.

Requirements for the semestral defence:

- Items 1 and 2.

Detailed formal requirements can be found at https://www.fit.vut.cz/study/theses/

Supervisor:	**Drahanský Martin, prof. Ing., Dipl.-Ing., Ph.D.**
Consultant:	Biswas Sangeeta, Ph.D., UITS FIT VUT
Head of Department:	Hanáček Petr, doc. Dr. Ing.
Beginning of work:	November 1, 2019
Submission deadline:	May 20, 2020
Approval date:	October 31, 2019

Abstract
The main goal of this thesis is to design and implement a program that will be able to generate synthetic retinal images high resolution that will be indistinguishable from real images of the retina of the human eye. The next step is to use this program to create a database with at least 1000 images of healthy retinas, i.e. without pathological findings. This database can then be used to develop various retina applications. These may be, for example, applications in the field of medical or biometric systems. The algorithm is based on a system of two neural networks, where the first network generates retinal images and the second network classifies whether the images are real or synthetic.

Abstrakt
Cílem této práce je navrhnout a implementovat program, který bude schopný generovat syntetické snímky sítnice ve vysokém rozlišení, které budou nerozeznatelné od reálných snímků sítnice lidského oka. Dalším krokem je pomocí tohoto programu vytvoření databáze s alespoň 1000 snímky zdravých sítnic, tedy bez patologických nálezů. Tato databáze může být následně použita při vývoji různých aplikací pracujících se sítnicemi. Může se jednat například o aplikace pohybující se v oblasti medicínských nebo biometrických systémů. Algoritmus je založený na systému dvou neuronových sítí, kde první síť generuje snímky sítnic a druhá síť provádí klasifikaci, zdali jsou dané snímky reálné či syntetické.

Keywords
human eye, eye retina, synthetic retinal images, image processing, image generation, machine learning, neural networks, GAN, high resolution, Python, TensorFlow

Klíčová slova
lidské oko, sítnice oka, syntetické snímky sítnice, zpracování obrazu, generování obrazu, strojové učení, neuronové sítě, GAN, vysoké rozlišení, Python, TensorFlow

Reference
AUBRECHT, Tomáš. *Generation of Synthetic Retinal Images with High Resolution*. Brno, 2020. Term project. Brno University of Technology, Faculty of Information Technology. Supervisor Prof. Ing., Dipl.-Ing. Martin Drahanský, Ph.D.

Generation of Synthetic Retinal Images with High Resolution

Declaration

I hereby declare that this Master's thesis was prepared as an original work by the author under the supervision of prof. Ing., Dipl.-Ing. Martin Drahanský, Ph.D. The supplementary information was provided by Biswas Sangeeta, Ph.D. I have listed all the literary sources, publications and other sources, which were used during the preparation of this thesis.

. . . .
Tomáš Aubrecht
February 11, 2020

Acknowledgements

I would like to thank Mr. Martin Drahanský for his thorough guidance and Miss Biswas Sangeeta for her help during the diploma work, valuable advice and time spent on consultations.

Contents

Chapter 1

Introduction

Eyesight allows us to interpret the surrounding environment using the light in its visible spectrum. Thanks to this, we can perceive contrast, contours of objects and their distance from us. It also contributes to perception of spatial orientation. For this reason, it is important to protect our eyesight, as the loss of it leads to a significant deterioration in the quality of life.

We begin to see when the cornea, together with the lens of the eye, focuses light from our surroundings on the light–sensitive membrane in the back of the eye, which is called the retina. It contains specialized light–sensitive cells: rods that allow perception of contrast and cones that allow perception of color. These cells convert the light into electrical signals that are transmitted to the visual cortex of the brain by the optic nerve. Therefore, the retina is the most sensitive and most important part of the human eye and its diseases or slightest mechanical damage can lead to loss of vision.

Machine learning is an application of artificial intelligence that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. One of the most popular areas of machine learning today is deep learning. It has been inspired by the human brain and it generally consists of a large number of parameters with multiple nonlinear layers. Generative models are an example of deep learning, more specifically generative adversarial networks (GAN).

Synthesizing realistic images of the eye fundus is a challenging task. Currently, the advances in technology have brought high computational power leading machine learning to neural networks with deep architectures. Considering recent progresses in deep learning algorithms, GAN represents a valuable framework. The rapid enhancement of GANs facilitated the synthesis of realistic-looking images, leading to slightly anatomically consistent and reasonable visual quality colored retinal fundus images [19].

1.1 Goals

In my thesis I will focus on the generation of retinal images using the generative adversarial network mentioned above. The main goal is to design and implement an algorithm that allows automatic generation of high resolution digital images of the retina. The next step is to compare the results of this algorithm with actual retinal images. In the case of high accuracy of this algorithm, a database of synthetic retinal images will be created, which could be used in practice for the development of various medical or biometric systems.

1.2 Contents

In the chapter 2 I will focus on the anatomy of the human eye, which is an important basis for understanding its physiology and the risks posed by various diseases. This chapter also describes the most common methods of examination of the eye and the individual eye diseases together with a description of their symptoms and possible methods of treatment. Information about the human eye was taken from my previous work [2]. The following chapter 3 deals with machine learning on which the proposed algorithm is based.

Chapter 2

Human eye

Human eyes are paired organs of the visual system, which provide us with vision, the ability to perceive the surrounding world and to orient ourselves in space thanks to the light in the visible spectrum reflected by the objects in the environment. Up to 80 % [6] of information from the external environment is perceived by sight. Therefore, the eye is the most important sensory organ. It has an approximately spherical shape and it is made up of three layers, enclosing various anatomical structures. The outermost layer is composed of the cornea and sclera. The middle layer consists of the choroid, ciliary body, pigmented epithelium and iris and the innermost layer is the retina.

2.1 Vision

When looking at an object, light rays reflect from that object and enter the cornea. The light rays are refracted and concentrated in one place through the cornea, lens and vitreous humor. Of these three structures, only the lens can change its optical power, thus ensuring that the rays are concentrated to the point of sharpest vision. The resulting image on the retina is turned upside down. Photons of light falling on the light-sensitive cells of the retina are converted into electrical signals that are transmitted to the brain by the optic nerve. These signals are interpreted as the resulting image in the visual cortex of the brain [6].

2.2 Anatomy

The human eye is a very complex system made up of many parts that must work together perfectly. The most important parts are described below. You can see them in the picture 2.1.

- **Cornea** is a transparent dome-shaped layer covering the anterior portion of the eyeball. The cornea, with regard to its optical power, is the most important component of the optical system of the eye and is the largest contributor to quality vision. Its main function is to refract light. It is responsible for focusing most of the light that enters the eye. The cornea is colorless, completely transparent and without blood vessels, which may prevent it from refracting light properly and may adversely affect vision. Since there are no nutrient-supplying blood vessels in the cornea, tears and the aqueous humor in the anterior chamber provide the cornea with nutrients. It represents a mechanical and chemically impermeable barrier between the inner and outer environment together with the conjunctiva, sclera and tear film.

- **Conjunctiva** is the clear, thin membrane that consists of two segments: *bulbar conjunctiva*, which covers the anterior part of the sclera, and *palpebral conjunctiva*, which covers the inner surface of both the upper and lower eyelids. The conjunctiva has many small blood vessels that provide nutrients to the eye and lids. Its main function is to keep the eye moist and lubricated by producing mucus and tears. It also contributes to the protection from dust, debris and microorganisms that can cause an infection.

- **Sclera**, also known as the white of the eye, is the protective, opaque, outer layer of the human eye. The whole sclera is white, contrasting with the coloured iris. It is continuous with the cornea offering resistance to internal and external forces to protect sensitive eye structures stored inside. The sclera also provides a sturdy attachment for the extraocular muscles that control the movement of the eyes. It is perforated by many nerves and vessels passing through its posterior part, where the hole is formed by the optic nerve.

- **Choroid**, also known as the choroidea, is another layer surrounding the eyeball that lies between the sclera and the retina. It provides oxygen and nourishment to the outer layers of the retina and maintains the temperature and volume of the eye.

- **Anterior chamber** of eyeball is the space inside the eye that is behind the cornea and in front of the iris. It is filled with a clear, watery fluid known as the aqueous humor. This is where the excess fluid can normally flow out. If the normal outflow of aqueous humour is blocked, the intraocular pressure is increased and glaucoma usually develops. This can lead to progressive damage to the optic nerve head, and eventually blindness.

- **Iris** is a thin, circular structure located behind the anterior chamber, that is usually strongly pigmented. The color of our eyes is determined by the amount of pigment in the iris. It contains a circular opening in the center called a pupil. The iris is responsible for controlling the diameter and size of this pupil, which regulates the amount of light reaching the retina. The larger the pupil, the more light can enter the eye and reach the retina.

- **Lens** is composed of transparent, flexible tissue and is located directly behind the iris and the pupil. It is important for refraction of light and its accommodation. The accommodation is a process of changing the curvature of the lens allowing closer objects to be brought into better focus by changing the optical power of the lens.

- **Posterior chamber** of eyeball is the second chamber consisting of small space directly behind the iris and in front of the lens. Like the anterior chamber of the eye, it is also filled with the aqueous humor. This fluid normally passes into the posterior chamber from where it flows into the anterior chamber. There the excess fluid can flow out of the eye.

- **Vitreous humor**, also known simply as the vitreous, is a clear, colorless fluid that fills the space behind the lens and in front of the retina in the eye. It has a firm gelatinous consistency and it makes up most of the volume of the eyeball. The vitreous helps to hold the shape of the eye and its pressure helps to keep the retina in place.

- **Optic nerve** connects the eye to the visual cortex of the brain. It is the nerve that transmits visual information in the form of impulses formed by the retina. These impulses are dispatched through the optic nerve to the brain, which interprets them as images. Glaucoma is a disease which result in damage to the optic nerve and cause vision loss. It is caused by high intraocular pressure, which compresses the optic nerve and causes its cells to die. It is referred to as the atrophy of the optic nerve.

- **Retina** is the most important part of this work, so it is described separately and in more detail in the following section 2.3.

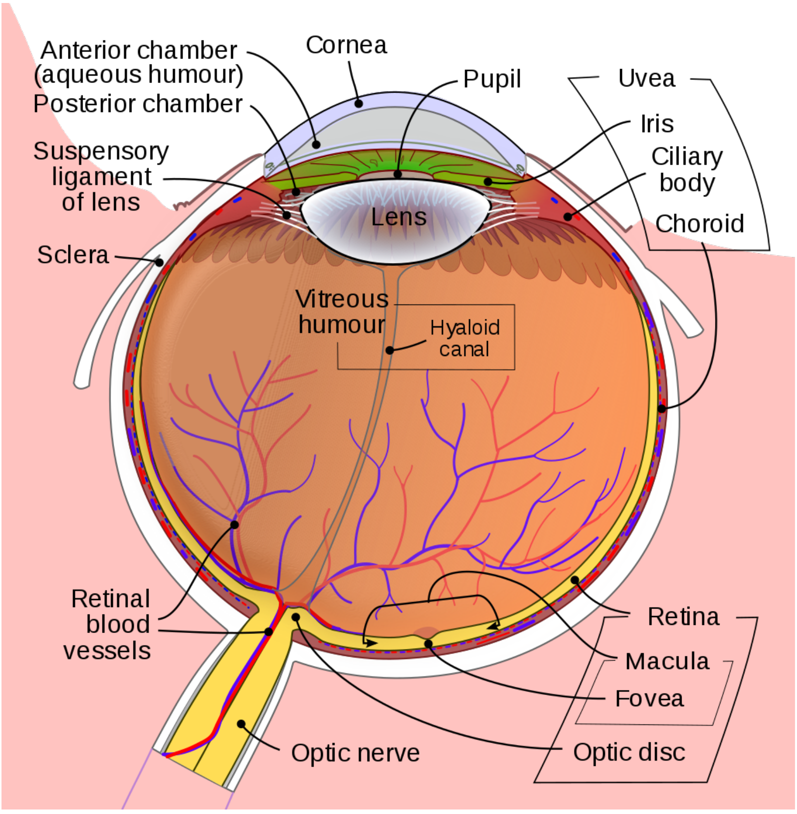

Figure 2.1: Schematic diagram of the human eye [16].

2.3 Retina

The retina is the most important and also the most sensitive part of our eye. It is a thin layer of tissue that lines the inner surface of the back of the eyeball. The retina processes light through a layer of light-sensitive cells, responsible for detecting qualities such as color and light-intensity. These specialized cells are called photoreceptors. The retina captures the light falling on these photoreceptors and converts the light into neural signals that are transmitted through the optic nerve to the visual cortex of the brain for visual recognition.

2.3.1 Photoreceptors

A photoreceptor is a specialized light-sensitive cell found in the retina that is responsible for converting light into signals that can stimulate biological processes. The photoreceptor absorbs photons that are striking the retina, which triggers a change in the membrane potential of the cell. There are two types of photoreceptor cells in the human retina: **rods** and **cones**. There are major functional differences between the rods and cones. Rod cells are much more sensitive than cone cells. At very low light levels, visual experience is based solely on the rod signal, so they are responsible for night vision. However, they do not mediate color vision, which is the main reason why colors are much less apparent in dim light, and not at all at night. The rods are concentrated at the outer edges of the retina and are used in peripheral vision. Cones require significantly larger number of photons to produce a signal. They are responsible for the perception of color and for high spatial acuity used for tasks such as reading. Cones are most concentrated in the center of the retina in an area called the macula and their density gradually decreases towards the outer edges of the retina [9].

2.3.2 Macula

The macula is a yellow oval-shaped area near the the center of the retina where the light is focused by the cornea and lens. The macula is responsible for the central, high-resolution and color vision. Therefore, the macula provides us with the ability to read and see in great detail. In the very center of the macular region is the fovea that has a very high concentration of photoreceptor cells, more specifically a high density of cones and low density of rods.

2.3.3 Optic disc

The optic disc, also called the optic nerve head, is located at the optic papilla, where the optic nerve fibres leave the eye. There are no photoreceptors in this area, so it is sometimes called the blind spot. The optic disc appears as an approximately oval area and it is the entry point for the blood vessels that supply the retina. These structures can be seen in the picture 2.2.

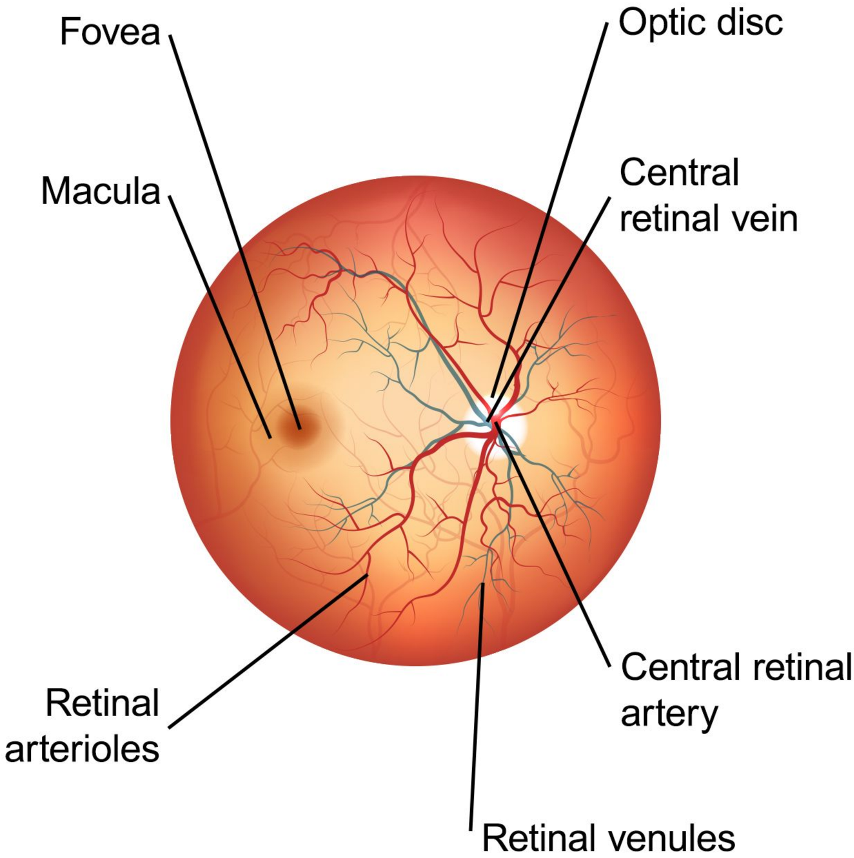

Figure 2.2: Retina of the human eye [18].

2.4 Eye examination

2.4.1 Ophthalmology

Ophthalmology is a branch of medicine dealing with anatomy and physiology of the eye and with the diagnosis, treatment and prevention of diseases of the whole visual system. This is a very specialized field, especially since the eye is a very complicated apparatus. An ophthalmologist is a medical doctor who specializes in diagnosing and treating eye-related conditions. In other words, an ophthalmologist is a specialist in ophthalmology.

2.4.2 Ophthalmoscopy

Ophthalmoscopy is an examination of the back part of the eye. This part of the eye is called the fundus, and consists of: retina, optic disc, choroid and blood vessels. Ophthalmoscopy may also be called funduscopy or retinal examination. Through ophthalmoscopy an eye doctor can find evidence of many kinds of eye problems including, but not limited to, glaucoma, high blood pressure damage, retinal detachment, diabetes, eye tumors, and many other problems. The dilation of the pupils, also known as mydriasis, is a simple and effective way to better observe the structures behind them. This is often done with eye drops before the examination. There are three different types of examinations: direct, indirect and slit-lamp examination.

Direct ophthalmoscopy produces an upright image of approximately 15× magnification. The handheld instrument that our primary care physician uses to look into our eyes is called a direct ophthalmoscope. You can see one in the picture 2.3. It is about the size of a small flashlight and it consists of a concave mirror and a battery-powered light. The doctor looks through a single monocular eyepiece into the eye of a patient in a darkened room. The ophthalmoscope is equipped with a rotating disc of lenses to permit the eye be examined at different depths and magnifications. It provides good, but limited visualization of the back of the eye. This type of ophthalmoscope is most commonly used during a routine physical examination.

Indirect ophthalmoscopy provides a wider view of the inside of the eye and produces an inverted image of 2 to 5× magnification using an indirect ophthalmoscope(picture 2.4). An indirect ophthalmoscope can be either monocular or binocular. It constitutes a bright light attached to a headband positioned on forehead of the eye doctor and magnifying lenses. The eye doctor holds the eye open while shining a very bright light into the eye using this indirect ophthalmoscope and views the back of the eye through the lens held close to the eye. Some pressure may be applied to the eye using a small, blunt probe. The pupil must be fully dilated for a satisfactory result. This examination is usually used for peripheral viewing of the retina and to look for detached retina.

The slit lamp is the most widely used ophthalmic device. It will have a place for us to rest our chin and forehead. This will help keep our head steady. This procedure gives us the same view of the eye as an indirect examination, but with greater magnification. A microscope is connected to a lamp, which is a high-intensity light source that can be focused to shine a thin ray of light into the eye. The doctor directs the light right into the eye of the patient, thus illuminating the area accurately. During the examination, the tissues are illuminated either by a thin ray of light or by reflected light. By examining the illuminated eye with the microscope, the ophthalmologist then obtains a magnified image of the observed area, allowing the detection of very subtle changes and symptoms of eye diseases.

Figure 2.3: Direct ophthalmoscope [10].

Figure 2.4: Indirect ophthalmoscope [17].

2.4.3 Fundus photography

Fundus photography uses a fundus camera to record images of the condition of the interior surface of the eye, also known as the fundus. Ophthalmologists use these retinal photographs for detailed evaluation as well as to document clinical observations and possible diagnosis of eye diseases. The fundus camera (figure 2.5) is a device that replaces the ophthalmoscope. It is a specialized low power microscope with an attached camera and it is based on the principle of monocular indirect ophthalmoscopy. The optics of a fundus camera are similar to those of an indirect ophthalmoscope in that the observation and illumination systems follow dissimilar paths. Fundus cameras are described by the angle of view and provide an upright, magnified view of the back of an eye. A typical camera captures images between $30°$ and $50°$ of retinal area with a magnification of $2.5\times$. This relation can be modified using zoom or auxiliary lenses. Wide angle fundus cameras capture images between $45°$ and $140°$ and provide proportionately less retinal magnification. For a better inspection, dilating eye drops are applied prior to examination to enlarge the pupil, thus increasing the angle of observation [3].

Figure 2.5: Fundus camera [14].

2.5 Retinal diseases

Retinal diseases vary widely, but most of them cause visual symptoms. Retinal diseases can affect any part of the retina and they are always very serious, often irreversible and can lead to severe vision loss or blindness. Treatment is available only for some retinal diseases. Depending on the retina condition, treatment goals may be to stop or slow the disease and preserve, improve or restore your vision. Common retinal diseases and conditions are described below.

2.5.1 Macular degeneration

Macular degeneration, also known as age-related macular degeneration (AMD or ARMD), is a macular disease that occurs in patients over age 50 and is the most common cause of practical blindness in economically developed countries. With the increasing number of seniors, it becomes a major societal health problem. Several factors influence the origin and development of this disease. In addition to increasing age, it can also be high blood pressure, smoking, poor eating habits and the associated obesity and genetic predisposition. Patients describe its symptoms so that visual acuity gradually decreases, they are complaining about image distortions and, in more advanced stages, a blurred or sometimes even black spot appears in the center of the field of view. Color vision also deteriorates. There is currently no known cure for macular degeneration, but there are options to reduce the risk and possibly slow the progression of the wet form. Vision will no longer improve and only the current quality of vision will stabilize [7].

AMD is divided into 2 forms: dry (atrophic, nonexudative) and wet (exudative). Up to 90 % of patients are affected by the dry form, but it causes severe visual damage in only 12–21 %. Fewer patients suffer from the wet form, but it is far more dangerous than the dry form because it progresses very quickly. Both forms can be combined during disease. In the macular area of the patients we find changes and loss of retinal pigment epithelium and drusen. Drusen are divided according to their appearance and size into hard and soft. Their comparison can be seen in the pictures 2.6 and 2.7. Hard drusen are small bounded deposits of yellowish color under the retina. On the contrary, soft drusen have no sharp boundaries and may even coalesce, they are associated with a significantly higher risk of the formation of the wet form of AMD [8].

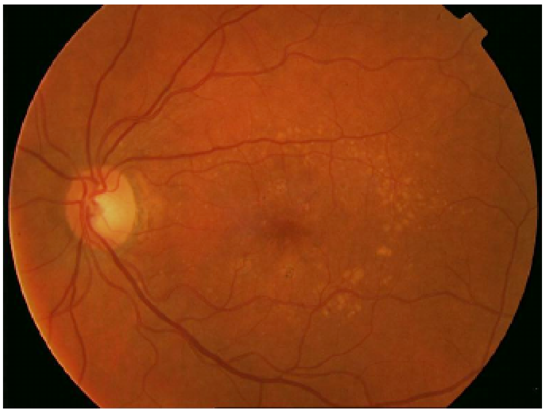

Figure 2.6: Hard drusen [8].

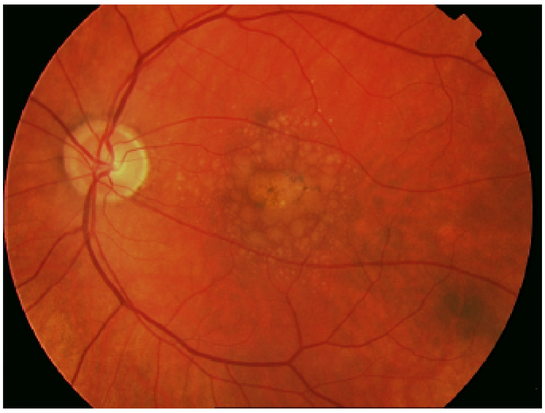

Figure 2.7: Soft drusen [8].

Dry AMD starts with the build up of drusen in the retina. Vision is usually good or only slightly reduced at this stage. Most of these patients with mild dry AMD can continue to read and drive, although it may not be as easy as it was when they were younger. Some patients, but not all, progress to a more advanced stage of dry AMD called geographic atrophy(figure 2.8). This can result in severe loss of central vision and loss of the ability to read and drive. Even in these severe cases, patients almost always retain normal peripheral vision, enough to see where they are going. Unfortunately there is no treatment for dry AMD. However, supplementation of antioxidant vitamins C, E and minerals zinc, selenium and essential omega-3 fatty acids may have a beneficial effect on preventing or slowing its progression. A diet rich in fish, vegetables and fruits also has a supporting role [8].

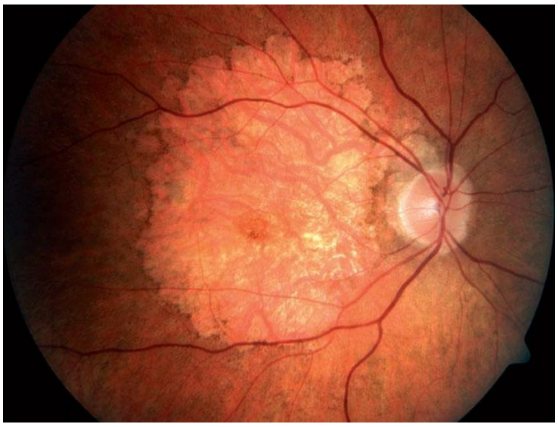

Figure 2.8: Geographic atrophy [11].

Wet AMD occurs when abnormal new blood vessels grow into the retina and start leaking fluid. Macular edema is the build-up of this fluid in the macula. This causes the retina to swell and the longer it is swollen, the more the retinal fibres deteriorate. Because these blood vessels are abnormal, they are more fragile than typical blood vessels and can bleed into the retina. This bleeding can cause irreversible damage to the photoreceptors and rapid vision loss if left untreated. You can see the characteristic image of a retina with the macular edema in the figure 2.9. It is usually, but not always, preceded by the dry form of AMD. The wet form progresses faster compared to the dry form and the loss of vision is more significant. Rapid deterioration occurs within a few weeks and practical blindness within a few months. Treatment previously consisted of destruction of the neovascular membrane by a photocoagulation or thermotherapeutic laser. However, treatment results were variable. The starting point should be a more targeted so–called photodynamic therapy, in which the intravenously injected substance is absorbed by the target tissue and then activated by laser [8].

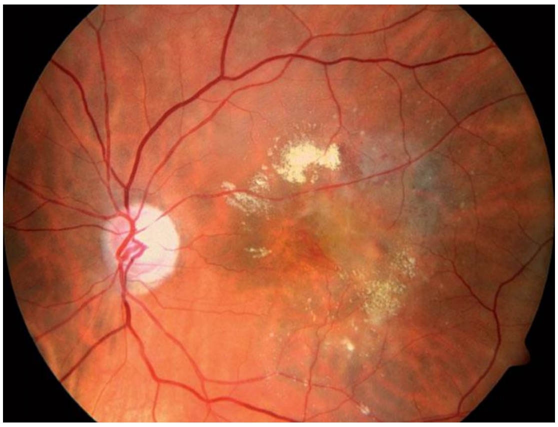

Figure 2.9: Wet form with the macular edema [12].

2.5.2 Diabetic retinopathy

Diabetic retinopathy is a diabetes complication that affects eyes. Retinopathy occurs when high blood sugar levels lead to the blockage of the tiny blood vessels that nourish the retina, cutting off its blood supply. The weakened blood vessels leak fluid into the retina. Some of the blood vessels break and bleed into the retina. This is called retinal haemorrhage and can be seen in the figure 2.10. As the disease becomes more advanced, new abnormal blood vessels may grow and these new blood vessels can bleed, cause cloudy vision, and destroy the retina. This condition can develop in anyone who has type 1 or type 2 diabetes. The longer the patient have diabetes and the less controlled his blood sugar is, the more likely he is to develop this eye complication. Diabetic retinopathy begins before the patient have any symptoms, but as the problem gets worse the patient may have: blurred vision, temporary or permanent blindness or distortion of vision. Early treatment is the key to reduce vision loss. A laser is used to seal leaking blood vessels or destroy abnormal blood vessels [15].

Figure 2.10: Retinal haemorrhage [4].

2.5.3 Retinal detachment

A retinal detachment is defined by the presence of fluid under the retina. If a hole develops in the retina, then the suction force is lost and the fluid that normally fills the inside of the eye passes through the hole and enters the space underneath the retina. As more fluid passes under the retina, the retina gradually detaches from the inner wall of the eye. If the retina remains detached, it will slowly deteriorate and lose function permanently, but if the retina can be reattached with surgery quickly enough, it is possible to recover some function and to avoid permanent vision loss [15].

2.5.4 Retinal vein occlusion

A retinal vein occlusion is a blockage of one of the veins draining blood from the eye. Retinal vein occlusions is divided into categories, based on the size of the vein which is blocked. A branch retinal vein occlusion is a blockage of one branch only and affects only part of the retina and a central retinal vein occlusion is a blockage of the main vein and affects the whole retina. If there is a very severe blockage and the blood flow stops altogether, the retinal cells die due to lack of oxygen. This is called ischaemia and there is no treatment that can bring the cells back to life. The increased pressure in the small vessels in the eye results in fluid leaking into the retina, making it swollen. A swollen retina does not see as well, and the longer the retina remains swollen, the more the vision deteriorates with time. Possible treatment options are intravitreal injections to reduce the swelling or laser surgery. If the blood supply is not restored, new blood vessels can grow into the retina. These new vessels are very fragile and can bleed, which can dramatically reduce the vision. In some cases this bleeding will require surgery to remove the blood in order to restore vision [15].

2.5.5 Retinitis pigmentosa

Retinitis pigmentosa is a group of rare, genetic disorders that involve a breakdown and loss of cells in the retina. The rods are more severely affected than cones in the early stages and people have difficulty seeing at night and lose the peripheral vision. The loss of rods eventually leads to a breakdown and loss of cones. In the late stages, people tend to lose more of the visual field, developing tunnel vision. Retinitis pigmentosa is diagnosed by an examination of the retina, which typically reveals abnormal, dark pigment deposits that streak the retina. There is currently no cure for this disorder [13].

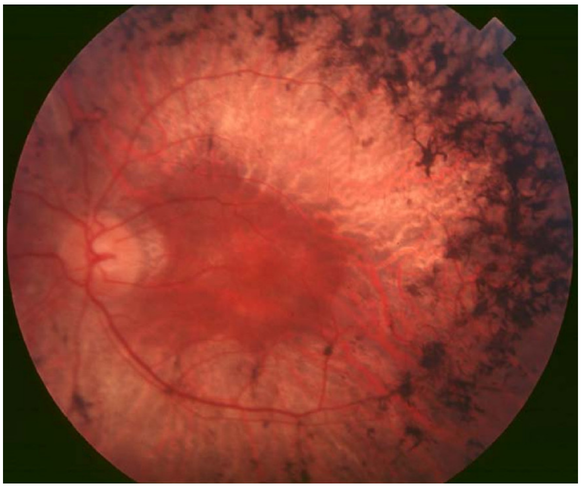

Figure 2.11: Fundus of a patient with retinitis pigmentosa [5].

Chapter 3

Machine learning

To solve a problem on a computer, we need an algorithm. An algorithm is a sequence of instructions that should be carried out to transform the input to output. For some tasks, however, we do not have an algorithm. Therefore, we do not know how to transform the input to output. What we lack in knowledge, we make up for in data. With advances in computer technology, we currently have the ability to store and process large amounts of data, as well as to access it from physically distant locations over a computer network. There are certain patterns in the data. Such patterns may help us better understand the data, or we can use those patterns to make predictions. Assuming that the future, at least the near future, will not be much different from the past when the sample data was collected, the future predictions can also be expected to be right. Application of machine learning methods to large databases is called data mining. Its application areas are abundant. In finance, banks analyze their past data to build models to use in credit applications, fraud detection, and the stock market. In manufacturing, learning models are used for optimization, control, and troubleshooting. In medicine, learning programs are used for medical diagnosis. In telecommunications, call patterns are analyzed for network optimization and maximizing the quality of service and in science, large amounts of data in physics, astronomy, and biology can only be analyzed fast enough by computers.

Machine learning is not just a database problem, it is also a part of artificial intelligence. To be intelligent, a system that is in a changing environment should have the ability to learn. If the system can learn and adapt to such changes, the system designer does not need to foresee and provide solutions for all possible situations. Machine learning also helps us find solutions to many problems in vision, speech recognition, and robotics. One example of the pattern recognition is the face recognition. This is a task we do effortlessly. Every day we recognize family members and friends by looking at their faces or from their photographs, despite differences in pose, lighting, hair style, and so forth. But we do it unconsciously and are unable to explain how we do it. Because we are not able to explain our expertise, we cannot write the computer program. At the same time, we know that a face image is not just a random collection of pixels. A face has structure. It is symmetric. There are the eyes, the nose, the mouth, located in certain places on the face. Each face of a person is a pattern composed of a particular combination of these. By analyzing sample face images of a person, a learning program captures the pattern specific to that person and then recognizes by checking for this pattern in a given image [1].

Machine learning is programming computers to optimize a performance criterion using example data or past experience. We have a model defined up to some parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or past experience. The model may be predictive to make predictions in the future, or descriptive to gain knowledge from data, or both. Machine learning uses the theory of statistics in building mathematical models, because the core task is making inference from a sample. The role of computer science is twofold. First, in training, we need efficient algorithms to solve the optimization problem, as well as to store and process the massive amount of data we generally have. Second, once a model is learned, its representation and algorithmic solution for inference needs to be efficient as well. In certain applications, the efficiency of the learning or inference algorithm, namely, its space and time complexity, may be as important as its predictive accuracy. Machine learning algorithms are typically classified into three broad categories: supervised learning, unsupervised learning, and reinforcement learning [20].

3.1 Supervised learning

For supervised learning problems, the training data comprises examples of the input vectors along with their corresponding target vectors. When the target vectors are categorical, the problems are known as classification or pattern recognition, and when the target vectors are real-valued, the problems are known as regression. Loss or distance functions are defined between the current output vector and the target vector for each input vector, and optimization is performed to minimize the loss over all training examples. By teaching the system with known input and target pairs, we expect to respond correctly even if unseen data are presented to the trained system. In other words, the goal of supervised learning is to learn a mapping from the input to an output whose correct values are provided by a supervisor.

3.1.1 Classification

Classification is a process of identifying to which of a set of categories a new observation belongs, on the basis of a training set of data containing observations whose category membership is known. Often, the individual observations are analyzed into a set of quantifiable properties, known as features. These properties may variously be categorical, ordinal, integer-valued or real-valued. Other classifiers work by comparing observations to previous observations by means of a similarity or distance function.

A bank that calculates risk based on customer information is an example of a classification problem where, for simplicity, there are two classification classes: low-risk and high-risk customers. The information about the customer includes his or her income and savings, and makes up the input to the classifier whose task is to assign the input to one of the two classes. After training with the past data, a classification rule learned may be of the form

$$\text{IF income} > \alpha \text{ AND savings} > \beta \text{ THEN low-risk ELSE high-risk} \tag{3.1}$$

for suitable values of α and β (see figure 3.1). This is an example of a discriminant. The discriminant is a function that separates the examples of different classes. Having a rule like this, the main application is prediction. Once we have a rule that fits the past data, if the future is similar to the past, then we can make correct predictions for novel instances.

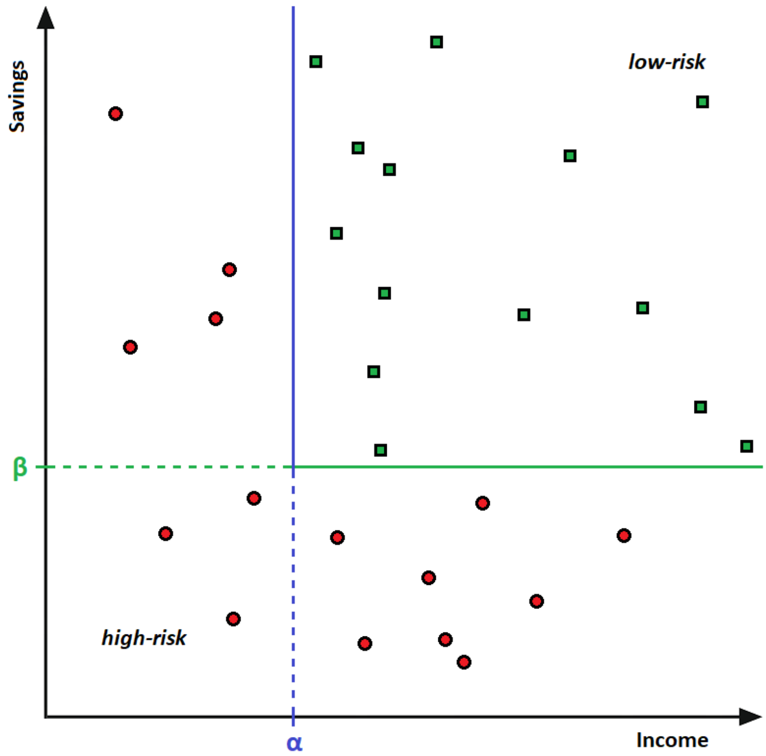

Figure 3.1: Example of a training dataset where each point corresponds to one data instance with input values in the corresponding axes. Red circles indicate the high-risk class and green squares the low-risk class.

3.1.2 Regression

Regression is a statistical measurement that estimates the relationship between one dependent variable and one or more independent variables.

Let us say we want to have a system that can predict the price of a used car. Inputs are car attributes: brand, year, engine capacity, mileage and other information that affects the worth of the car. The output is the price of the car. Let X denote the attributes of the car and Y denote the price of the car. From the past transactions, we can collect a training data and the machine learning program fits a function to this data to learn Y as a function of X. An example is given in the figure 3.2 where the fitted function is of the form

$$y = w_1 x + w_0 \tag{3.2}$$

for suitable values of w_1 and w_0. For simplicity, mileage is taken as the only input attribute and a linear model is used. In cases where the linear model is too restrictive, a quadratic

$$y = w_2 x^2 + w_1 x + w_0 \tag{3.3}$$

or a higher-order polynomial, or any other nonlinear function of the input can be used.

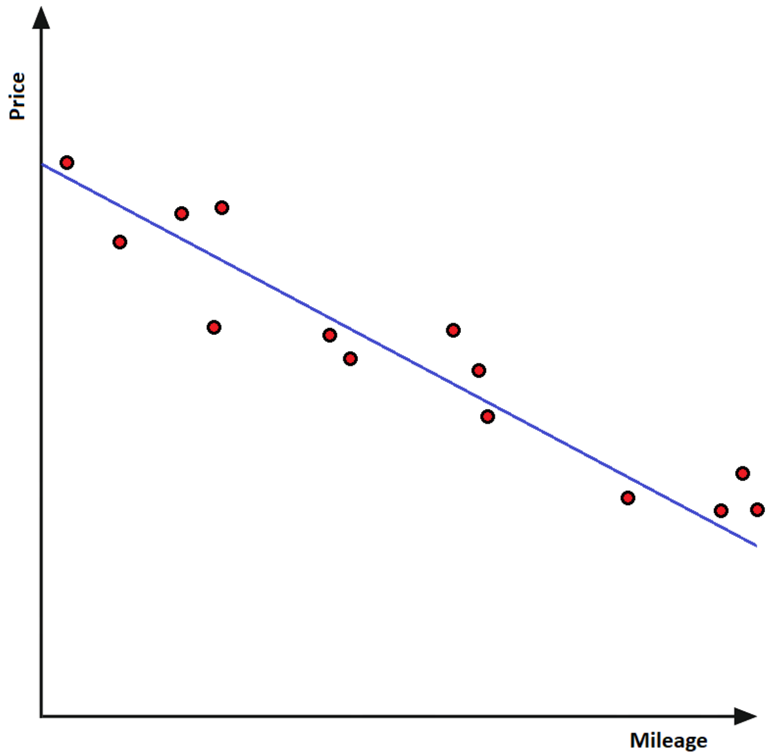

Figure 3.2: A training dataset of used cars and the fitted function.

Another example of regression is navigation of a mobile robot, for example, an autonomous car, where the output is the angle by which the steering wheel should be turned at each time, to advance without hitting obstacles and deviating from the route. Inputs are provided by sensors on the car, for example, a video camera and GPS. Training data can be collected by monitoring and recording the actions of a human driver [1].

3.2 Unsupervised learning

In unsupervised learning, there is no supervisor, no targets are defined so that the training data consist of only a set of input vectors. The goal is to find the regularities in the input data. There is a structure to the input space such that certain patterns occur more often than others, and we want to see what generally happens and what does not. In statistics, this is called density estimation. One method for density estimation is clustering. Therefore, a variety of clustering algorithms are canonical examples of unsupervised learning.

3.2.1 Clustering

The goal of clustering is to find clusters or groupings of input. In a case of a company with a data of past customers, the customer data contains the demographic information as well as the past transactions with the company, and the company may want to see the distribution of the profile of its customers, to see what type of customers frequently occur. In this case, a clustering model allocates customers similar in their attributes to the same group, providing the company with natural groupings of its customers. Once the groups are found, the company may decide strategies specific to different groups. This grouping also allows identification of outlying values, namely, those who are different from other customers, which may imply a niche in the market that can be further exploited by the company.

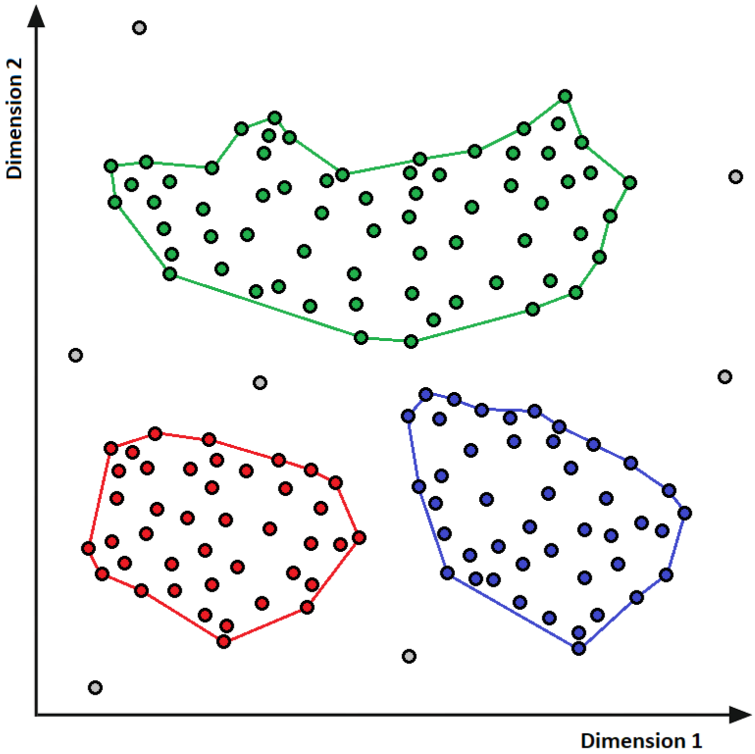

Figure 3.3: Example of density-based clustering that connects areas of high input data density into clusters.

An interesting application of clustering is in image compression. In this case, the input instances are image pixels represented as RGB values. A clustering program groups pixels with similar colors in the same group, and these groups correspond to the colors occurring frequently in the image. If there are only shades of a small number of colors in an image and if we code those belonging to the same group with one color, then the image is quantized. If the pixels are 24 bits, which represent 16 million colors, but there are only 64 shades of the main color, we need 6 bits instead of 24 for each pixel. For example, if the scene has various shades of blue in different parts of the image, and if we use the same average blue for all of them, we lose the details in the image, but gain space in storage and transmission.

Machine learning methods are also used in bioinformatics. DNA in our genome is a sequence of bases, namely, A, G, C, and T. RNA is transcribed from DNA, and proteins are translated from the RNA. Just as a DNA is a sequence of bases, a protein is a sequence of amino acids. One application area of computer science in molecular biology is alignment, which is matching one sequence to another. This is a difficult string matching problem, because strings may be quite long, there are many template strings to match against, and there may be deletions, insertions, and substitutions. Clustering is used in learning motifs, which are sequences of amino acids that occur repeatedly in proteins. Motifs are of interest because they may correspond to structural or functional elements within the sequences they characterize. The analogy is that if the amino acids are letters and proteins are sentences, motifs are like words, namely, a string of letters with a particular meaning occurring frequently in different sentences [1].

3.3 Reinforcement learning

The goal of reinforcement learning is to learn how to act or behave in a given situation for given reward or penalty signals. In this type of learning, a state for current status is defined and environment, usually a criterion function, evaluates the current state to generate a proper reward or penalty action through a set of policies. Instead of having exact target values, it learns with critics [20].

A good example is game playing where a single move by itself is not that important. It is the sequence of right moves that is good. A move is good if it is part of a good game playing policy. Game playing is an important research area in both artificial intelligence and machine learning. This is because games are easy to describe and at the same time, they are quite difficult to play well. A game like chess has a small number of rules but it is very complex because of the large number of possible moves at each state and the large number of moves that a game contains. Once we have good algorithms that can learn to play games well, we can also apply them to applications with more evident economic utility.

A robot navigating in an environment in search of a goal location is another application area of reinforcement learning. At any time, the robot can move in one of a number of directions. After a number of trial runs, it should learn the correct sequence of actions to reach the goal state from an initial state, doing this as quickly as possible and without hitting any of the obstacles.

One factor that makes reinforcement learning harder is when the system has unreliable and partial sensory information. For example, a robot equipped with a video camera has incomplete information and thus at any time is in a partially observable state and should decide taking into account this uncertainty. For example, it may not know its exact location in a room, but only that there is a wall to its left. A task may also require a concurrent operation of multiple robots that should interact and cooperate to accomplish a common goal [1].

3.4 Deep learning

Deep learning has been inspired by the human brain and has been proving its powerful ability in detection, classification, segmentation, key point estimation and activity classification. It generally consists of a large number of parameters with multiple nonlinear layers. Deep learning architectures include two popular categories: convolutional neural networks (CNN) for automatic feature extraction and recurrent neural networks (RNN) for sequence estimation. They have been applied to computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, and bioinformatics with outstanding performances. In addition, generative models such as variational encoders and generative adversarial networks are also becoming popular with their artificial sample generation capability [20].

3.4.1 Generative adversarial networks

Generative adversarial network (GAN) is an unsupervised deep learning machine, introduced by Ian Goodfellow. It is based on two models: a generator and a discriminator. The generative model learns to capture the data distribution, taking random samples of noise, and generates plausible images from that distribution. The discriminative model estimates the probability that a sample comes from the data distribution rather than generator distribution, and therefore is tasked to discriminate between real and fake images [19]. The general GAN scheme is shown in figure 3.4 below.

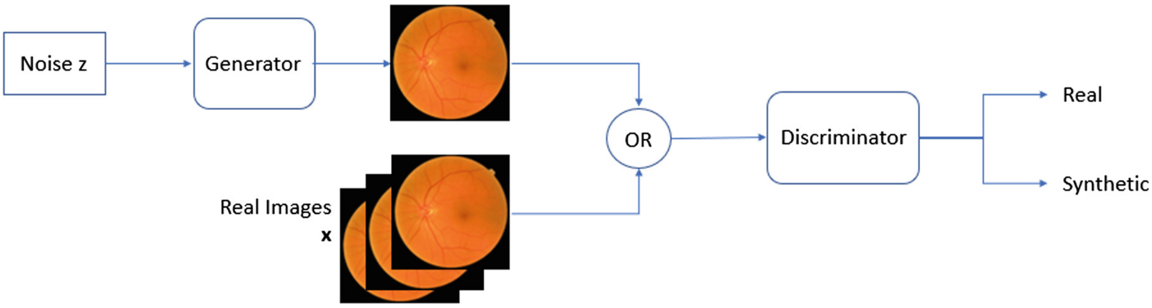

Figure 3.4: General scheme of a generative adversarial network [19].

Bibliography

[1] ALPAYDIN, E. and BACH, F. *Introduction to Machine Learning*. 2nd ed. Cambridge: MIT Press, 2014. ISBN 978-0-262-02818-9.

[2] AUBRECHT, T. *Detekce onemocnění ve snímku sítnice oka*. Brno, 2017. 40 p. Bachelor's thesis. Vysoké učení technické v Brně. Fakulta informačních technologií. Vedoucí práce Ing. Lukáš Semerád.

[3] BENEŠ, P. *Přístroje pro optometrii a oftalmologii*. 1st ed. Brno: Národní centrum ošetřovatelství a nelékařských zdravotnických oborů, 2015. ISBN 978-80-7013-577-8.

[4] CALHOUN, J. S. *Retinal Hemorrhage*. 2013. [Online; visited 20.11.2019]. Available at: https://imagebank.asrs.org/file/10070/retinal-hemorrhage.

[5] HAMEL, C. *Fundus of patient with retinitis pigmentosa*. 2006. [Online; visited 20.11.2019]. Available at: https://en.wikipedia.org/wiki/Retinitis_pigmentosa#/media/File:Fundus_of_patient_with_retinitis_pigmentosa,_mid_stage.jpg.

[6] HRVOLOVÁ, B. *Biofyzika*. 1st ed. Ostrava: Vysoká škola báňská - Technická univerzita Ostrava, 2013. ISBN 978-80-248-3105-3.

[7] HYCL, J. and TRYBUČKOVÁ, L. *Atlas oftalmologie*. 2nd ed. Praha: Triton, 2008. ISBN 978-80-7387-160-4.

[8] KOLÁŘ, P. et al. *Věkem podmíněná makulární degenerace*. 1st ed. Praha: Grada, 2008. ISBN 978-80-247-2605-2.

[9] KVAPILÍKOVÁ, K. *Anatomie a embryologie oka*. 1st ed. Brno: Institut pro další vzdělávání pracovníků ve zdravotnictví, 2000. ISBN 80-7013-313-9.

[10] MEDISAVE.CO.UK. *Direct ophthalmoscope*. 2017. [Online; visited 11.11.2019]. Available at: https://www.medisave.co.uk/heine-mini3000-2-5v-ophthalmoscope.html.

[11] MF.CZ. *Geografická atrofie*. 2016. [Online; visited 20.11.2019]. Available at: https://img.mf.cz/060/617/b.jpg.

[12] MF.CZ. *Vlhká forma s edémem*. 2016. [Online; visited 20.11.2019]. Available at: https://img.mf.cz/062/617/c.jpg.

[13] NEI. *Retinitis Pigmentosa* [online]. 2014. Updated July 10, 2019 [cit. 20. November 2019]. Available at: https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/retinitis-pigmentosa.

[14] NIDEK INTL.COM. *Fundus camera*. 2015. [Online; visited 11.11.2019]. Available at: http://www.nidek-intl.com/product/ophthaloptom/diagnostic/dia_retina/afc-330.html.

[15] PAŠTA, J. *Základy očního lékařství*. 1st ed. Praha: Univerzita Karlova, Nakladatelství Karolinum, 2017. ISBN 978-80-246-2460-0.

[16] RHCASTILHOS. *Schematic diagram of the human eye*. 2007. [Online; visited 11.11.2019]. Available at: https://en.wikipedia.org/wiki/Human_eye#/media/File:Schematic_diagram_of_the_human_eye_en.svg.

[17] WELCHALLYN.COM. *Indirect ophthalmoscope*. 2017. [Online; visited 11.11.2019]. Available at: https://www.welchallyn.com/en/products/categories/physical-exam/eye-exam/ophthalmoscopes--binocular-indirect/binocular_indirect_ophthalmoscope.html.

[18] WONG, G. *Retina of the human eye*. 2016. [Online; visited 11.11.2019]. Available at: https://d3b3by4navws1f.cloudfront.net/177901919.jpg.

[19] WONG, T. Y. and TING, D. S. W. *Generative Adversarial Networks (GANs) for Retinal Fundus Image Synthesis. In:* CARNEIRO, *G. and* (EDS.), *S. Y., ed.* Computer Vision-ACCV 2018 Workshops: 14th Asian Conference on Computer Vision. *Perth, Australia: Springer, 2019, p. 289–302. ISBN 978-3-030-21073-1.*

[20] ZHANG, *Y.-Q. and* RAJAPAKSE, *J. C. Machine Learning and Biometrics. 1st ed. Hoboken, N.J.: Wiley, 2008. ISBN 978-0-470-11662-3.*

Appendix A

Contents of the attached CD

The directory structure on the attached CD is shown and described below.

- **databases** – contains individual databases of retinal images