FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

ASSIGNMENT OF BACHELOR'S THESIS

Title:	Detecting diabetic retinopathy and related diagnoses using Neural Networks
Student:	Vitalii Tokarchyn
Supervisor:	Ing. Jakub Žitný
Study Programme:	Informatics
Study Branch:	Knowledge Engineering
Department:	Department of Applied Mathematics
Validity:	Until the end of summer semester 2020/21

Instructions

Research current state-of-the-art techniques that are used for prediction, classification and segmentation tasks in the medical imaging domain, focus on retina images.
Implement your own prototype model that will work on one of the open datasets provided by the supervisor.
Compare the performance of your model with reference results from literature or existing models and discuss the pros and cons.
Publish your prototype code and make sure your results are reproducible.

References

Will be provided by the supervisor.

Ing. Karel Klouda, Ph.D.
Head of Department

doc. RNDr. Ing. Marcel Jiřina, Ph.D.
Dean

Prague February 5, 2020

Acknowledgements

I wish to acknowledge the guiding and great advises from my supervisor, Ing. Jakub Žitný. Also, I would like to commend EyePACS and Kaggle for large public dataset of retina fundus photos we use in this thesis.

Declaration

I hereby declare that the presented thesis is my own work and that I have cited all sources of information in accordance with the Guideline for adhering to ethical principles when elaborating an academic final thesis.

I acknowledge that my thesis is subject to the rights and obligations stipulated by the Act No. 121/2000 Coll., the Copyright Act, as amended. In accordance with Article 46 (6) of the Act, I hereby grant a nonexclusive authorization (license) to utilize this thesis, including any and all computer programs incorporated therein or attached thereto and all corresponding documentation (hereinafter collectively referred to as the "Work"), to any and all persons that wish to utilize the Work. Such persons are entitled to use the Work in any way (including for-profit purposes) that does not detract from its value. This authorization is not limited in terms of time, location and quantity. However, all persons that makes use of the above license shall be obliged to grant a license at least in the same scope as defined above with respect to each and every work that is created (wholly or in part) based on the Work, by modifying the Work, by combining the Work with another work, by including the Work in a collection of works or by adapting the Work (including translation), and at the same time make available the source code of such work at least in a way and scope that are comparable to the way and scope in which the source code of the Work is made available.

In Prague on June 4, 2020 . . . . . .

Czech Technical University in Prague
Faculty of Information Technology

Citation of this thesis

Tokarchyn, Vitalii. *Detecting Diabetic Retinopathy and Related Diagnoses Using Neural Networks*. Bachelor's thesis. Czech Technical University in Prague, Faculty of Information Technology, 2020.

Abstrakt

Diabetická retinopatie je závažné oční onemocnění, které může vést k úplné slepotě. Včasná diagnostika významně zvyšuje šance na úspěšnou léčbu. V této bakalářské práci jsme vybudovali konvoluční neuronovou síť pro klasifikaci fotografií očního pozadí do 5 fáz diabetické retinopatie. Také jsme představili další studie týkající se automatické diagnostiky diabetické retinopatie a teorie, která za ní stojí.

Klíčová slova diabetická retinopatie, klasifikace, konvoluční neuronové sítě, neuronové sítě, strojové učení, oční fundus

Abstract

Diabetic retinopathy is a serious eye disease that can lead to total blindness. Early diagnostic significantly increases the chances of successful treatment of it. In this bachelor's thesis, we built a convolutional neural network to classify photos of eye fundus into 5 stages of diabetic retinopathy. Also, we introduced other studies related to automatic diabetic retinopathy diagnostic and the theory behind it.

Keywords diabetic retinopathy, classification, convolutional neural networks, neural networks, machine learning, fundus

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This bachelor thesis is focused on constructing self-taught model to classify different stages of diabetic retinopathy disease and on making overview of methods used to automatic classification in the medical imaging domain. But before moving directly into this topics we will make brief introduction how artificial intelligence in medicine evolves, current state of applying AI in the medical imaging domain and why it is important to diagnose diabetic retinopathy at early stage.

1.1 Artificial Intelligence in Medicine

Since the invention of the computer, researchers have tried to find applications for it in areas such as diagnostics of different diseases, processing of medical data, and guidance of treatment decisions. One of the notable projects was Dendral [1], which was used for chemical analysis. It was the first developed expert system and it laid the basis for other ones. Such as MYCIN [2] developed at Stanford University in the 1970s and ONCOCIN [3] in the 1980s. MYCIN could diagnose and recommend treatment of blood infections. ONCOCIN was used to assist doctors in the treatment of cancer. Despite the so called artificial intelligence (AI) winters in the 1970s and 1990s when interest to AI has essentially halted, research continued. With appearance of the Internet, more powerful hardware, new algorithms for data processing, the amount of data started rapidly increasing. This, in turn, contributed to an increase in machine learning popularity, as can been seen in the fig. 1.1, which shows the number of publications about machine learning and data mining in journal of *Artificial Intelligence in Medicine* since 1985.

Machine learning is one of the areas of artificial intelligence. The main principle of it is that machines self-train based on input data, their own experience and mistakes, instead of being explicitly programmed. Nowadays, it is rapidly evolving, actively used in business in many fields including healthcare.

Research theme	1985	1987	1989	1991	1993	1995	1997	1999	2001	2003	2005	2007	2009	2011	2013	Total	%
Knowledge engineering	8	16	21	15	40	22	11	14	11	5	5	9	1	4	4	186	25
Ontologies and terminologies	0	0	0	0	5	5	4	2	5	8	11	8	8	11	13	80	11
Natural language processing	0	1	1	0	0	1	4	5	4	5	7	8	9	6	8	59	8
Guidelines and protocols	0	0	0	0	4	1	5	5	3	10	16	8	4	9	5	70	10
Temporal information management	0	1	2	0	2	4	3	7	7	7	11	8	5	6	5	68	9
Case based reasoning	0	1	0	0	2	3	2	3	1	3	2	0	2	0	0	19	3
Planning and scheduling	0	0	0	0	0	0	3	1	3	1	1	4	4	2	1	20	3
Distributed and cooperative systems	0	0	0	2	2	2	3	3	4	1	4	7	2	0	2	32	4
Uncertainty management	0	4	3	3	4	4	8	2	11	7	4	8	5	3	6	72	10
Machine learning, data mining	0	1	3	2	11	11	15	14	18	12	23	26	25	11	17	189	26
Image and signal processing	0	0	2	1	6	2	9	4	5	2	6	8	10	4	3	62	8
Bioinformatics	0	0	0	0	0	0	0	0	2	2	4	2	5	3	4	22	3
Total accepted papers	15	26	36	26	60	60	58	50	63	52	70	66	62	45	45	734	

Figure 1.1: Frequency of occurrence of the twelve main research themes over the years.
Source: [4]

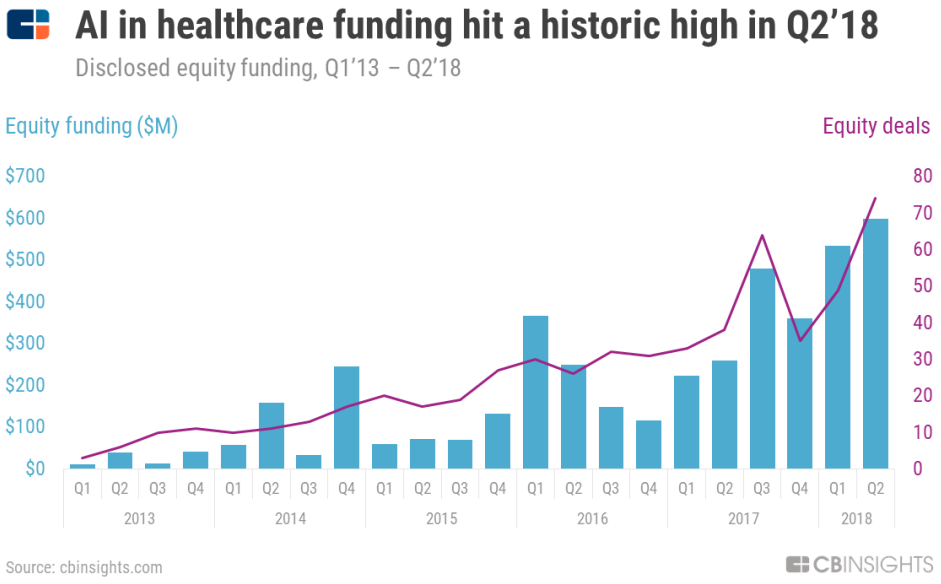

Figure 1.2: AI in healthcare funding.
Source: [6]

For example [5] which provides a remote patient monitoring using analyzing patients' gestures with machine learning.

The increase in popularity of AI in healthcare can be seen not only in numbers of publications on this topic but also in investment. According to annual reports by CBInsights [6] the funding of startups focused on AI in healthcare increased from several million USD in the first quarter in 2013 to something about 600 million USD in the second quarter in 2018 (fig. 1.2). Another re-

port by MarketsandMarkets [7] expects that value of artificial intelligence in healthcare market will reach 36.1 billion USD by 2025. So it seems this trend will continue in the next years.

1.2 Artificial Intelligence in Medical Imaging

Computer vision is a discipline that deals with problems related to how computers understand this world through visual data such as videos, images, etc. This discipline is a subset of artificial intelligence and dates back to the late 1950s when the article "Receptive fields of single neurones in the cat's striate cortex" [8] was published. At first glance, it seems to be not related to computer science, but it describes such core principles of image processing that it is considered as one of the most influential works in computer vision. Then there were lots of discoveries and studies that significantly moved computer vision to what we can observe now. The common applications for computer vision in medicine are those which work with CT, MRI or X-Ray scans [9, 10, 11]. Over the last several years many promising projects appeared in this area. Some of them even outperformed human specialists. For example, in the recent article "International evaluation of an AI system for breast cancer screening" [12] where authors compare AI systems for diagnostic breast cancer to human specialists. They conclude that their system works better than average specialist by 11.5% of AUC-ROC.

1.3 Diabetic Retinopathy

Diabetic retinopathy is one of the complications of diabetes that causes serious problems with eyes, and can even lead to total blindness. This disease is caused by high blood sugar that blocks vessels. Which in turn worsens the retina nutrition, so it tries to grow new vessels to compensate the blocked ones. But unfortunately, these new vessels often have leakage of blood and other fluids. The retina is a light-sensitive area in the eyes, so when it is damaged, a person may see everything in a blur, some objects could be invisible to the human eye or may be just represented as dark spots (fig. 1.3) and so on.

Diabetic retinopathy is the main cause of vision loss and the most common cause of preventable blindness. Due to the high prevalence of DR among people with diabetes mellitus, which is about 30% [13], and the fact that a person may not even notice any difference in vision during the firsts stages, it is very important for them to visit ophthalmologist from time to time. Early diagnosis can significantly improve the effectiveness of treatment and prevent complications that would emanate from DR. One of the problem we face here is the difficulty to provide quality treatment to all who needs it. Many countries have only one ophthalmologist per million population [14]. Also, the number of people with DR rises, in 2015 there were 2.6 million people worldwide and

Figure 1.3: (a) Normal vision (b) A simulation of what someone sees with advanced diabetic retinopathy.
Source:
`www.researchgate.net/figure/a-Normal-vision-b-A-simulation-of-what-someone-see-n-with-advanced-diabetic_fig1_261022372`

in 2020 it is expected to be 3.2 million [15]. Therefore, it would be really helpful to have an automatic system to make diagnostics faster, cheaper and with equally high accuracy for everyone.

Chapter 2

Diabetic Retinopathy and Other Retina Diseases

This chapter describes existing abnormalities on the retina, diseases they cause, possible classification of diabetic retinopathy, how information we require to make diagnostic is obtained.

2.1 Abnormalities on The Retina

Ophthalmologists search for signs on the retina that could provide information about the presence of different systematic diseases. The following is a list of some abnormalities that could be found.

- **Enhanced arteriolar light reflex**
  It is caused by the thickening of atherosclerotic vessel walls and could be a sign of hypertensive retinopathy or other diseases. On the photo, it differs from other healthy blood vessels by remarkable glowing (fig. 2.1).

- **Hard exudates**
  They appear due to leakage of proteins, lipids and blood fiber. They look like white-yellow discrete accumulations and could indicate the presence of diabetic or hypertensive retinopathy, Coat's disease and other (fig. 2.1).

- **Cotton-wool spots, or soft exudates**
  They are white-yellow or gray-white patches of discoloration in the nerve fiber layer with fimbriated borders as the result of local ischemia. Cotton-wool spots are often surrounded by microaneurysms and hemorrhages (fig. 2.1).

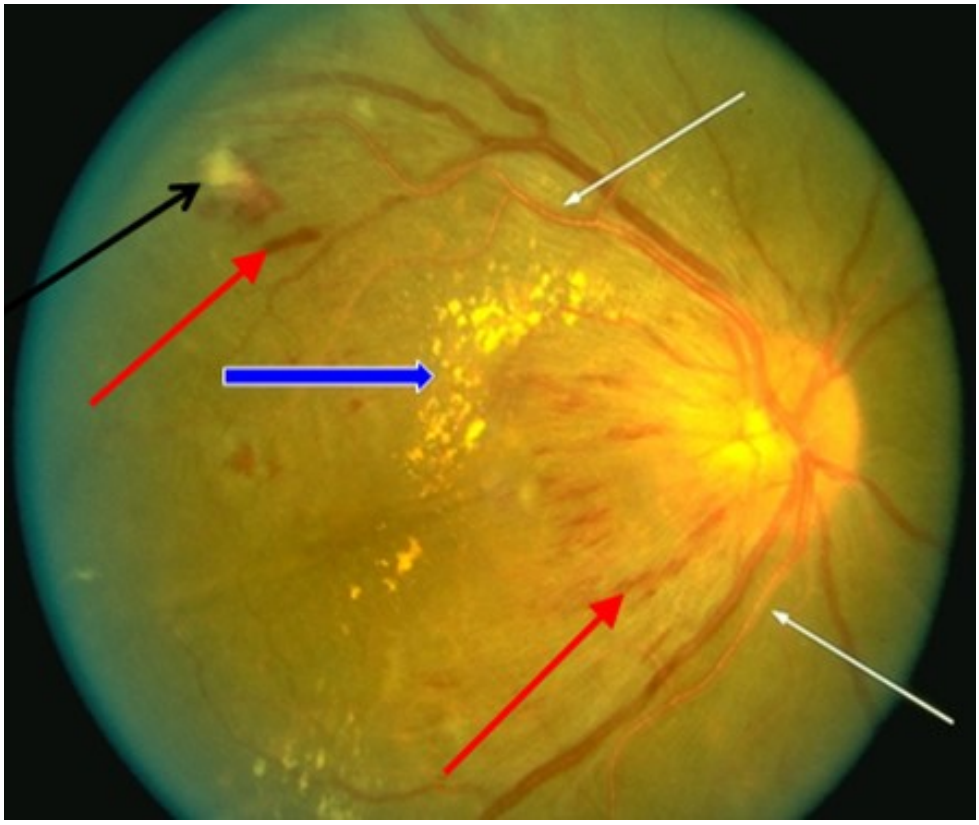

Figure 2.1: Arteriolar narrowing and increased retinal arteriolar light reflexes (thin white arrows), retinal hard exudates (blue arrow), cotton wool spots (black arrow), and flame retinal hemorrhages (red arrows).
Source: [16]

- **Hemorrhages**
  Hemorrhages are the bleeding from blood vessels in the retina. The type of hemorrhage when necrotic vessels bleed into either the nerve fiber layer called flame retinal hemorrhage (fig. 2.1); or the inner retina, called dot blot hemorrhage (fig. 2.2). Also, there are other types. The first one is subhyaloid and preretinal hemorrhages located on the retina's surface. The second is Subretinal and subretinal pigment epithelium (RPE) hemorrhages located under the neurosensory retina. Hemorrhages could be caused by various conditions, for example head trauma, high blood pressure, leukemia or diabetes.

- **Microaneurysms**
  Microaneurysm is a tiny swelling inside the retinal blood vessels. Usually, it is caused by diabetes mellitus but also could be caused by high blood pressure or by various vascular diseases. Sometimes, it is quite dif-

Figure 2.2: Multiple blot hemorrhages (blue arrows) of various sizes together with microaneurysms (red arrow).
Source: [16]

ficult to tell the difference between dot hemorrhages and microaneurysms (fig. 2.2).

- **Neovascularization**
  It is the process when new blood vessels start growing inside the eye. These new vessels could cause the appearing of hemorrhages and retinal detachment (fig. 2.3).

2.2 DR Classification

The most popular type of diabetic retinopathy classification and the one we use is presented by *American Academy of Ophthalmology* [17]. According to this classification, there are 5 levels of disease, depending on symptoms and their number (fig. 2.4):

1. **No diabetic retinopathy**

2. **Mild non-proliferative diabetic retinopathy (NPDR)**
   Patients with this stage have at least one microaneurysm but no other

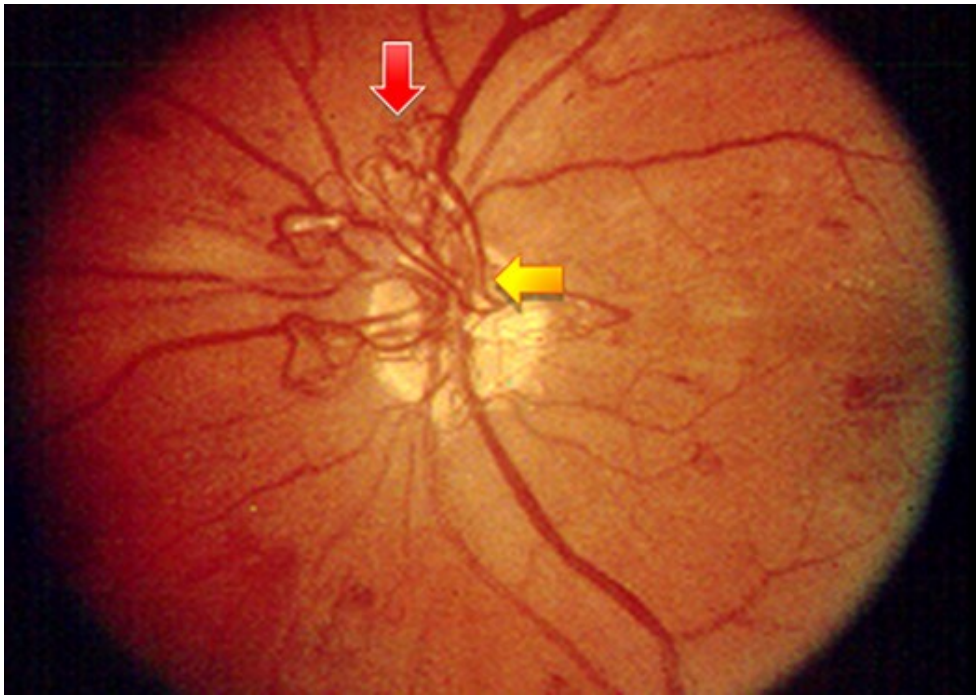

Figure 2.3: New blood vessels due to neovascularization process on the surface of the retina (red arrow) and optic nerve (yellow arrow) of the left eye.
Source: [16]

abnormalities. Because this is an early stage, other abnormalities can be very inconspicuous, therefor closer inspection is required.

3. **Moderate non-proliferative diabetic retinopathy**
   Additionally to abnormalities from the previous stage other symptoms such as hemorrhages, soft exudates, venous beading or intraretinal microvascular abnormalities can be presented.

4. **Severe non-proliferative diabetic retinopathy**
   These patients have venous beading in several retina quadrants, intraretinal hemorrhages and microaneurysms in each of four quadrants. In case there is neovascularization it is not severe NPDR, but already proliferative diabetic retinopathy (PDR).

5. **Proliferative diabetic retinopathy**
   This stage of DR is considered as proliferative if neovascularization or vitreous/preretinal hemorrhage is present.

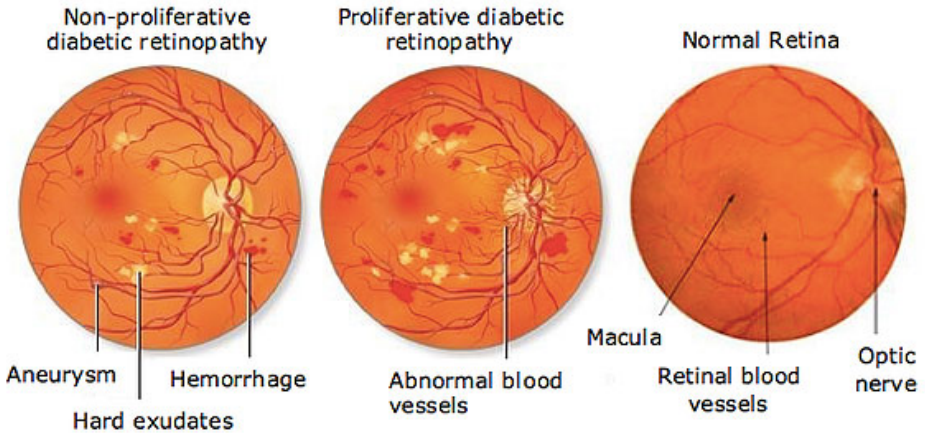

Figure 2.4: Stages of diabetic retinopathy.
Source: [18]

2.3 Eye Examination

Ophthalmologists can use several instruments to examine the eye's inner backside. The fundamental principle of almost all these instruments is using an additional light source, that highlights the inner eye's area through the pupil. During this process, the special drops could be used to widen the pupils of the eyes in order to improve fundus visibility. The following are some of them.

- **Direct ophthalmoscope.**
  Direct ophthalmoscope uses light that is passed through the aperture wheel, several lenses, mirror, and then through the pupil from xenon–halogen bulb or LED in order to highlight the small part of the eye's fundus (fig. 2.5). The aperture wheel allows to tune shape and light's color, so the visibility of specific structures in the fundus could be improved. For example, by applying green filter we can enhance the visibility of blood vessels. After fundus being highlighted, it is projected through the compensation lens onto the ophthalmologist's retina. This instrument has one downside, though. It has a tiny field of view (FOV) which is equal to only 2 millimeters.

- **Indirect ophthalmoscope.**
  This tool (fig. 2.6) works on similar principles as a direct ophthalmoscope. The main difference is so-called pupil matching when patient's and doctor's pupils are imaged on the same positive lens. Thanks to this technique the problem of a tiny field of vision is solved. The field of vision is approximately 6 times larger than using a direct ophthalmoscope. Other advantages of this technique are possibilities to make an

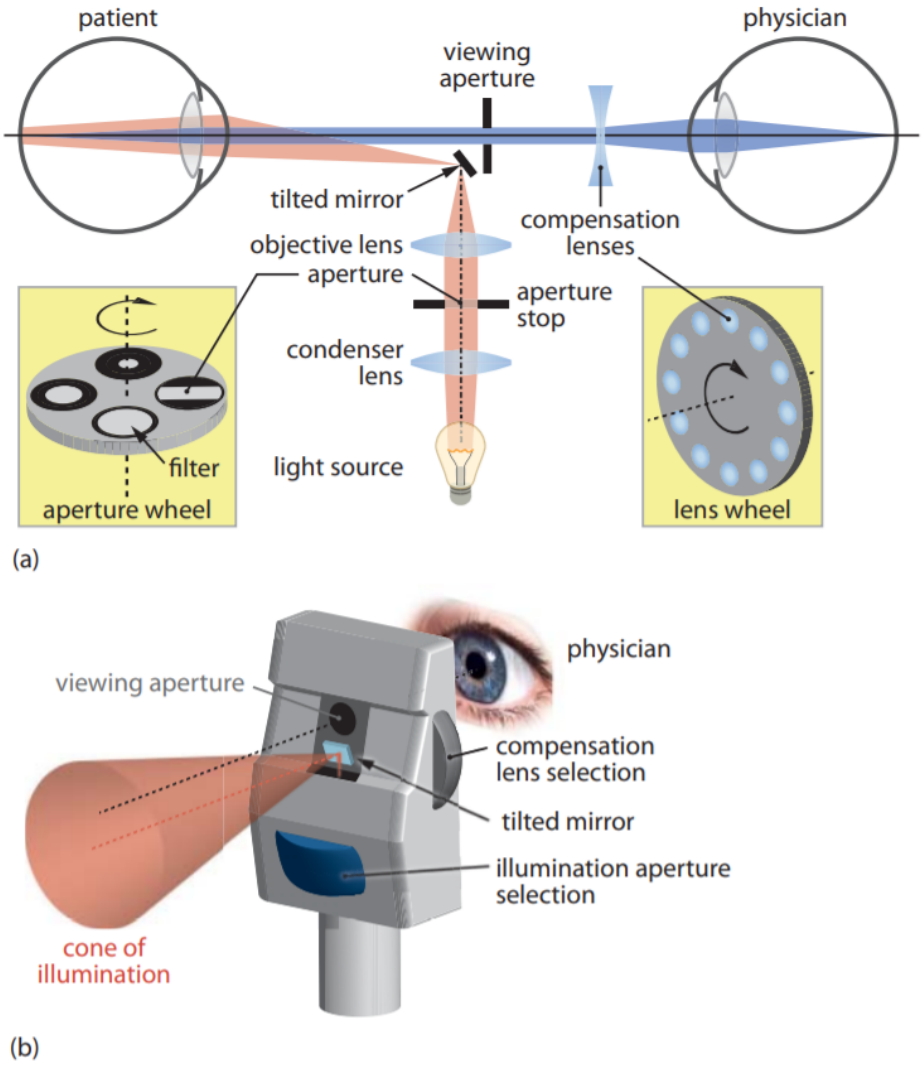

Figure 2.5: Principle of a direct ophthalmoscope. (a) Optical design of a direct ophthalmoscope. The light of a xenon–halogen bulb or LED is collected by a condenser lens and passes through an aperture wheel (left inset) with which different aperture shapes can be selected. With an objective lens, the light is then focused onto a tilted mirror and projected onto the patient's eye. There, it illuminates the eye fundus. With a suitable compensation lens (chosen from a lens wheel; right inset), the fundus is then imaged into the physician's eye. Adapted from [30]. (b) Scheme of a typical direct ophthalmoscope which shows the cone of illumination and the selection wheels.
Source: [19]

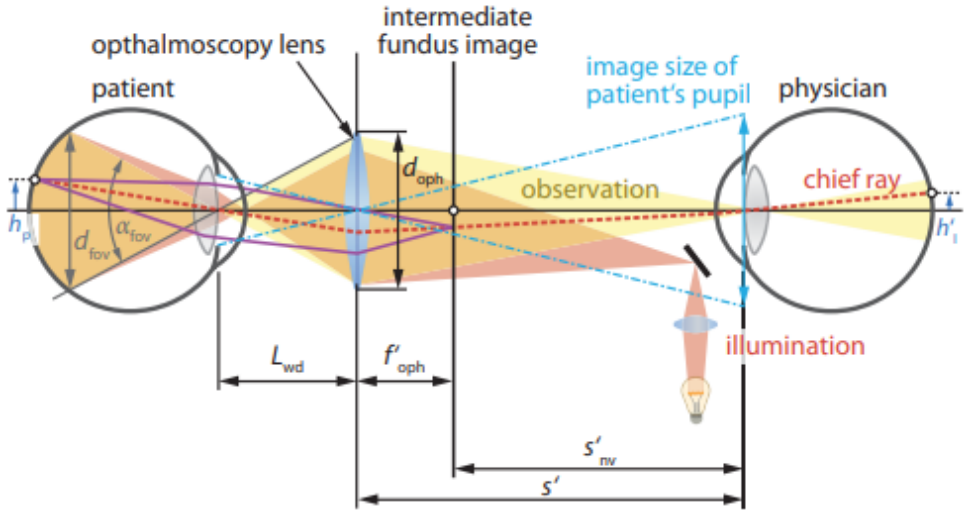

Figure 2.6: Ray diagram of an indirect ophthalmoscope which shows the illumination beam path (red) and the observation path (yellow). L_{wd} is the working distance of the ophthalmoscopy lens, s' the distance between the ophthalmoscopy lens and the physician, f'_{oph} the image-side focal length of the ophthalmoscopy lens, and s'_{nv} the typical near viewing distance between the intermediate image plane and the physician. h_p is the object size on the patient's fundus and h'_I the corresponding image size formed on the physician's fundus. α_{fov} denotes the angle of the retinal field of view, d_{fov} the corresponding diameter of the field of vision, and d_{oph} the free diameter of the ophthalmoscopy lens.
Source: [19]

observation reflection-free with less effort and to implement stereoscopic viewing, which gives the doctor ability to see the thickness of abnormalities. The downside of an indirect ophthalmoscope is its magnification power, it is 5 times smaller than in a direct ophthalmoscope.

- **Fundus camera.**
  The fundus camera (fig. 2.7) is a more complex instrument than previously discussed. Thanks to the different techniques used in it, it has many advantages comparing to direct and indirect ophthalmoscope. Stereoscopic viewing, different image modes, field of view up to 200°, digitalization of fundus image are several of them. The last one is especially useful, as it gives the ability to the doctor to see the progression of the disease. Photos of the fundus taken by the fundus camera you can see on figs. 2.1 to 2.3.

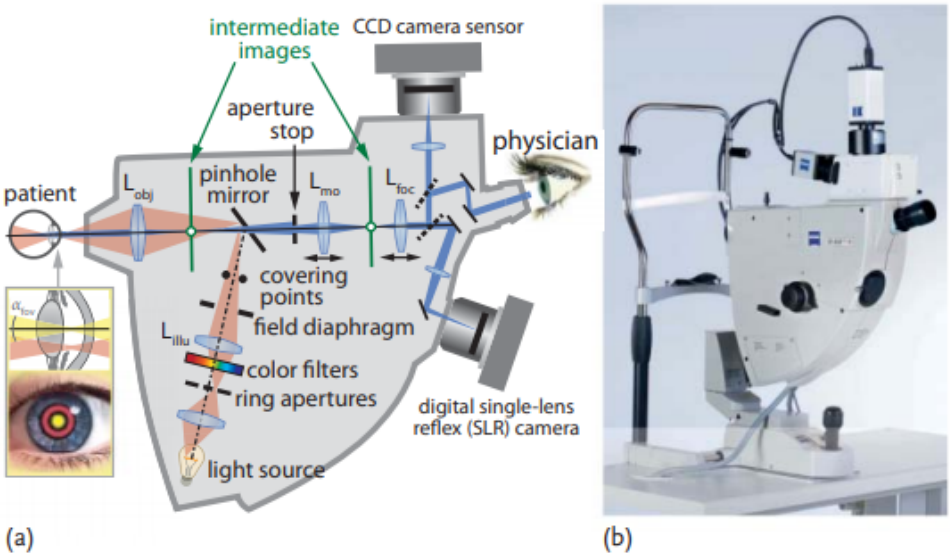

Figure 2.7: (a) Simplified schematic beam path of the ZEISS FF 450$^{\text{plus}}$ fundus camera. L_{obj} denotes the objective lens which corresponds to the ophthalmoscopy lens of an indirect ophthalmoscope. L_{foc} and L_{mo} are the focusing and main objective lenses, respectively. L_{illu} denotes the illumination lens. The individual optical components and their function are explained in the text. Top inset: Beam path separation in the anterior eye segment to avoid disturbing reflections in the image. Between the corneal front surface and the anterior lens surface, the illumination beam path (red) is separated from the observation path represented by the field of vision (yellow). Bottom inset: Front view of the examined eye in which the ring-shaped enter pupil of the illumination path (red) and the exit pupil of the observation path (yellow) are marked. (b) Photograph of the ZEISS FF 450$^{\text{plus}}$ fundus camera. Courtesy of Carl Zeiss.

Source: [19]

2.4 Optical Coherence Tomography (OCT)

OCT helps the doctor to see a cross-sectional image of the retina tissues. The method it uses is similar to ultrasound imaging, but instead of ultrasound, the light is used. This tool is relatively new but already has widespread clinical adoption. OCT is a noninvasive tool, but despite this can penetrate tissues up to 5mm of depth. One of the many other advantages of this tool is almost real-time imaging.

2.5 Fundus Camera using Smartphone

The fundus camera is commonly used in medical ophthalmic practice as it is very important for the diagnosis of many eye disease. But despite its spreading, there are many people who can not afford treatment using this tool, because its cost is high especially for rural regions and because of the lack of professionals who can work with it. Thanks to progress in smartphone cameras there are successful implementations that could take photos of fundus using it [20, 21]. These implementations significantly reduce the cost of diagnostic and also reduce the expertise required to use it. However, it is not replacing existing professional instruments, as the photo quality of lasts is much better. In the fig. 2.8 you can see the examples of photos taken by Devrim Toslak, MD, Ali Ayata, MD, and other [20].

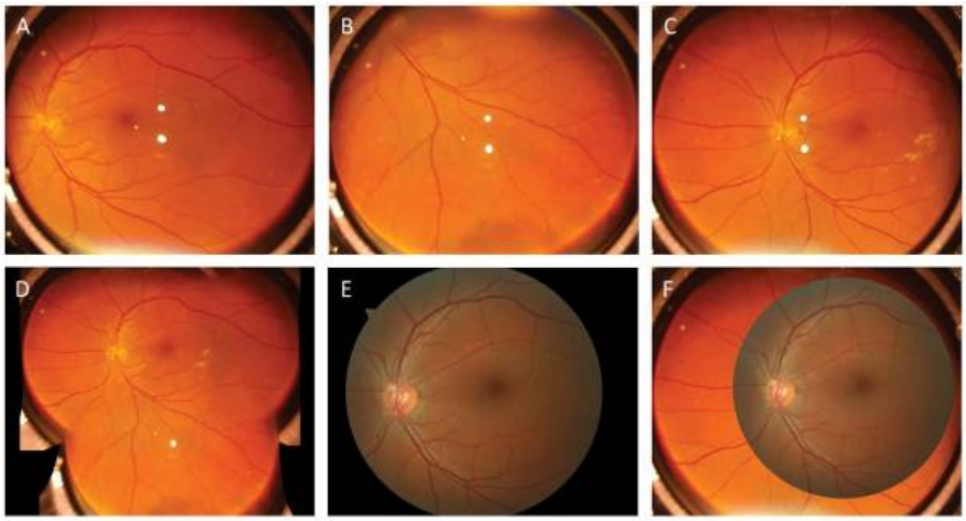

Figure 2.8: (A–C) Representative single-shot images captured from a 41 years old subject. The subject has no reported eye diseases. (D) Montage of single-shot images in A–C. (E) Representative fundus image from the same subject collected with a clinical fundus camera (Zeiss, Cirrus Photo 800), which has a single-shot FOV of 45°external angle, corresponding to 67.5°. (F) Overlap of images of C and E for FOV comparison.

Source: [20]

Chapter 3

Theory

In this chapter we will provide theory related to machine learning (ML), its problems and methods to tackle with them. For the following brief introduction we used the following sources [22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33].

3.1 Machine Learning

Machine learning is a fast-growing subset of AI. The main difference between ML and other computer systems counted as AI is the ability to modify itself automatically based on input data. Thanks to reducing human factor, it could process more information and find new patterns in data that previously were not obvious for people. The machine learning algorithms could be divided into 3 main types based on the input data handling, specifically supervised, unsupervised, and reinforcement learning.

3.1.1 Supervised Learning

In this type of algorithm, the developer should prepare samples with various features (also called independent variables) and relative expected result (or dependent variables). Then during the training process, the algorithm tries to find the best parameters for the function that map features to the result labels. The best parameters are those which allow the algorithm to predict labels not only for samples that it was trained on but also for unseen ones.
Using algorithms of supervised learning type the tasks of classification and regression could be solved.

- The classification task is when the samples should be categorized into some known set of classes based on features input. For example, input could be the human's height, weight, age and as a result of prediction could be human's sex.

- In contrast to classification, the regression task should predict continuous value. It could be for example the price of company shares or how much of the distance the ball will make until fall on the ground.

3.1.2 Unsupervised Learning

An unsupervised learning algorithm does not use any information about what result the developer expects from it. Instead, this algorithm learns the underlining structures of data, their similarities, and difference based on features provided to it.
The tasks that unsupervised learning algorithms attempt to solve could be divided into clustering and principal components analysis.

- The clustering task is when the dataset should be divided into different subgroups, where samples in each subgroup are similar to each other and at the same time differ from samples from other subgroups.

- The goal of principal components analysis is to reduce the dimension of input data. This in turn helps to find non-obvious relation between variables and to reduce the size of the dataset without significant information loss.

3.1.3 Reinforcement Learning

The fundamental characteristics of the reinforcement learning algorithm are interaction through the agent with the environment in which it operates in order to maximize (or minimize) some score defined by the developer. The agent is not explicitly taught but it learns based on its previous steps. The agent makes steps considering current state of the environment, its state and previous steps it made, and also experience what it gained previously. The system rewards agent for each step considering how successful it was.

Thanks to this trial and error learning process the overall performance of the program could reach amazing results in various areas. For example system for planning treatment strategy for sepsis, which is the main cause of mortality in hospitals [34], or human-like robot hand which can manipulate physical objects with great dexterity [35].

3.2 Artificial Neural Networks (ANNs)

Artificial Neural Networks are the mathematical models for processing information. Those models are the set of units called neurons and their connections between each other. Even though we do not know much how the human brain works, ANN was introduced and it is based just on already existed knowledge about the brain's neural network.

The processing done by neuron can be described by function:

$$f(x) = \alpha\left(\sum_{i=1}^{n} x_i * w_i + b\right),$$

where each value x_i from vector x is multiplied by weight w_i, then they are summed, the bias b is added to this sum and the result is then through activation function α is passed to the output. Multiple neurons are grouped into layers. To solve real-world problems multiple layers are required, typically those are fully-connected layers, which are layers where each neuron is connected to each neuron from the previous layer.

3.2.1 Activation Function

Almost all activation functions are non-linear as it leads to better performance in the complex problems. The are many activation functions, each with its advantages and disadvantages, the most common among them are the following:

- Sigmoid:
$$\alpha(x) = \frac{1}{1 + e^{-x}}$$
- Tanh:
$$\alpha(x) = tanh(x)$$
- Rectified linear unit (ReLU):
$$\alpha(x) = max(0, x)$$
- Leaky rectified linear unit:
$$\alpha(x) = max(c * x, x),$$
where the c is the small constant. This constant is an attempt to fix "dying ReLU" problem [36].
- Softmax:
$$\alpha(y_i) = \frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_j}},$$
where y_i represent the result of the weighted sum of inputs computed in neuron i. n is a size of vector y that contains all values from each neuron before passing to the activation function. This vector also called logits. The softmax function is frequently applied at the last layer of the neural network of classification tasks, the result of it is a normalized vector with probabilities for each class.

3.2.2 Training

Training is an iterative process of extracting knowledge from a training dataset to make better predictions. Specifically, the knowledge of ANN consists of learnable parameters, such as weights, biases or filters values in convolutional neural networks (CNN).

For each output the objective function is calculated, the result of this function represents how good the model makes predictions. This function is calculated based on the desired output for corresponded input and the predicted one. Depending on the objective function either minimizing or maximizing should be done. When the objective function should be minimized, it is also called the loss function. For example, the one we used in this thesis called categorical cross-entropy. This loss function is suitable for classification task and is calculated as follows:

$$l(y, p) = -\sum_{c}^{n} y_{i,c} log(p_{i,c}),$$

where n is classes' number, y is binary indicator if class c is the correct classification for input i, and i is predicted probability that i is of class c.

After computing the objective function, randomly initialized learnable parameters can be changed by optimization algorithm in order to achieve better predictions in the future. Optimization algorithms that are based on gradient descent, such as Stochastic gradient descent (SGD), Adagrad, RMSprop and Adam are the most used ones. However, for networks with multiple layers, the technique called backpropagation is required, as it allows to propagate result of objective function from the output to the input layer. This in turn allows to adjust weights of all layers, not only of the last one.

3.3 Convolutional Neural Networks

The convolutional neural network is almost the same network as ordinary ANN except for the fact that it has some additional types of layers. This type of ANN shows really good results for data with a spatial correlation between neighborhood data points, for example, images, text or time-series data. The basic idea is to make use of mathematical convolution operation, thanks to this the number of parameters that should be stored is far fewer, which affects overall training speed and memory requirements.

The typical CNN architecture consists of convolution blocks and several fully-connected layers at the end. Those blocks in turn, in simple implementation could be divided in 3 stages: convolution, non-linear activation (or detector) and pooling stages (fig. 3.1).

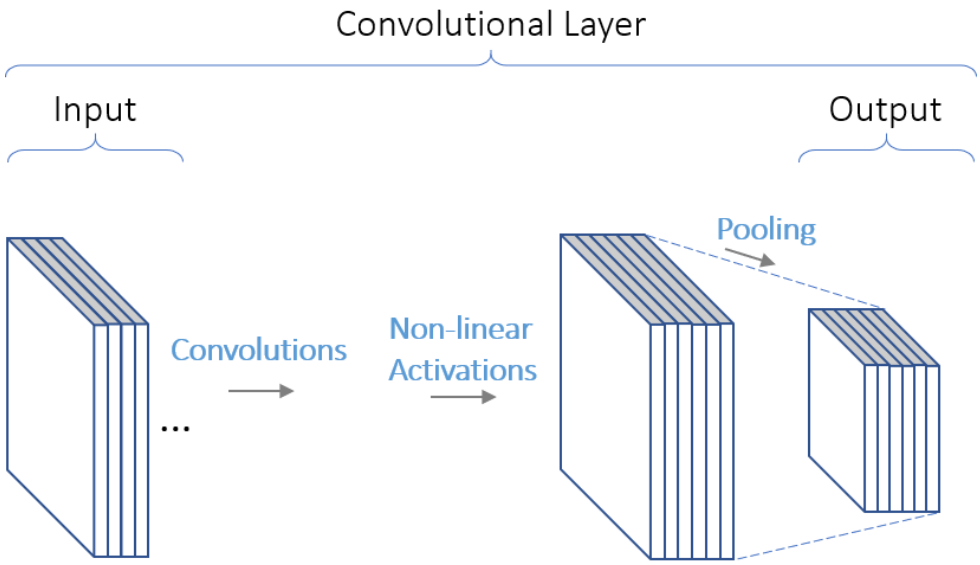

Figure 3.1: The three stages of a convolutional layer.
Source: `https://towardsdatascience.com/visualizing-the-fundamentals-of-convolutional-neural-networks-6021e5b07f69`

3.3.1 Convolution Stage

Convolution stage involves convolution operation on one or more kernels. The number of kernels, their size and the number of data points shift over the input are hyperparameters, that are parameters initialized before the training. The convolution operation is in general operation on two functions of real values, but in the case of CNN, the discrete convolution is used. The convolution is often applied to images. In a simple case where image is gray-scale, the convolution operation can be defined by the formula:

$$C(i,j) = \sum_w \sum_h I(w,h)K(i-w, j-h),$$

where I is our image, i and y are pixels coordinates in a column and row respectively, K is kernel (or also filter), w and h are width and height of kernel respectively. But in many machine learning libraries the convolution nevertheless is implemented as cross-correlation instead:

$$C(i,j) = \sum_w \sum_h I(i+w, j+h)K(w,h).$$

3.3.2 Pooling Stage

The aim of this stage is size reducing the input using the pooling operation. This, in turn, improves computational efficiency and makes the model more

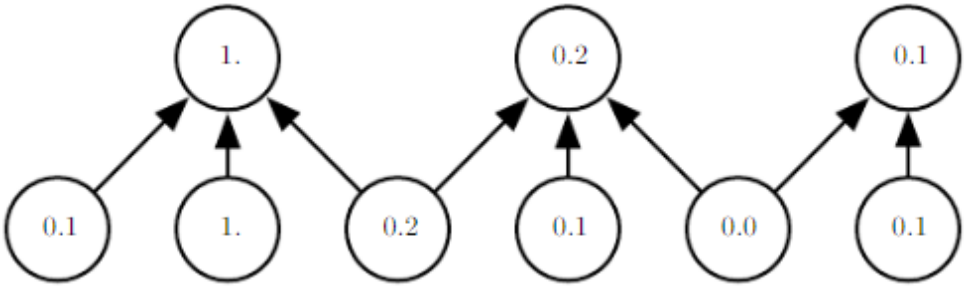

Figure 3.2: Example of max pooling. For each second data point get maximum among pool values. The pool size in this particular example is equal to 3 and the stride is equal to 2.
Source:

stable to small translations of the input. In general, it is done by calculating statistics of data points in the input and its neighborhood. Such operation could be applied for each k-th data point, where $k \in \mathbb{N}^+$. This value k is called stride. There are various methods to do pooling, such as max (fig. 3.2) or average pooling, global max or average pooling. The global pooling differs from usual pooling by the fact that the pool size is equal to the size of input data. The pool size represents the number of data point neighborhood including itself.

3.4 Class Imbalance Problem

The problem of class imbalance occurs when the numbers of each class in the dataset are significantly different. This problem is very common in machine learning, especially in healthcare datasets. For example, in the binary classification of the presence of chronic illness, the number of samples of healthy people is much greater than of people with chronic illness. These two classes are also referred to as majority and minority classes respectively.

There are many methods to overcome dataset imbalance [37], here are some of them:

- **Oversampling, undersampling and their combination**
  These methods imply changing samples' numbers using randomly deleting, duplication and also more sophisticated techniques, for example, Synthetic Minority Over-sampling Technique (SMOTE) [38]. However, they could lead to other problems, such as overfitting minority class or the loss of information.

- **Cost-sensitive learning**
  Cost-sensitive learning uses a cost matrix to define the attention each class gets during training. There are several methods of how to set up

a cost matrix, the simplest one is computing costs for each class by dividing total dataset size by class size.

3.5 Metrics in Machine Learning

Metric functions provide developers with an understanding of model performance, and contrary to the loss functions, are not used in model optimization. The usage of metrics should be considered for each project separately, as there are not such a metric that would fit every task. Below we will take a look at some of the common metrics.

3.5.1 Mean Absolute Error (MAE) and Mean Squared Error(MSE)

These metrics represent average difference between true value and value predicted by the model. MSE is different in that the average is taken from squared errors instead of just absolute errors like in MAE. These metrics could be also used in loss function, where MSE has some advantages such as easier differentiability and more penalising for large errors. In contrast to MSE, MAE is more understandable to humans. MAE and MSE have the following definitions:

$$MAE = \frac{1}{n}\sum_{i=0}^{n-1}|y_i - y_i'|$$

$$MSE = \frac{1}{n}\sum_{i=0}^{n-1}(y_i - y_i')^2,$$

where n is number of predictions, y is true value and y' is predicted one.

3.5.2 Confusion Matrix

Confusion matrix is an NxN table, where predicted labels made by the model are on the one axis and true labels on another (fig. 3.3). Thanks to this table developers are aware of how good model is at predicting each class and distinguish them. There are 4 important terms related to confusion matrix, specifically true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Using them, it is easy to calculate various metrics such as sensitivity (also recall), specificity, precision, F1 score. The true or false in the above-mentioned terms represents correctness of predictions, and positive or negative represents two classes in binary classification, for example, positive could be the class of samples with presented diabetic retinopathy and negative - samples without diabetic retinopathy.

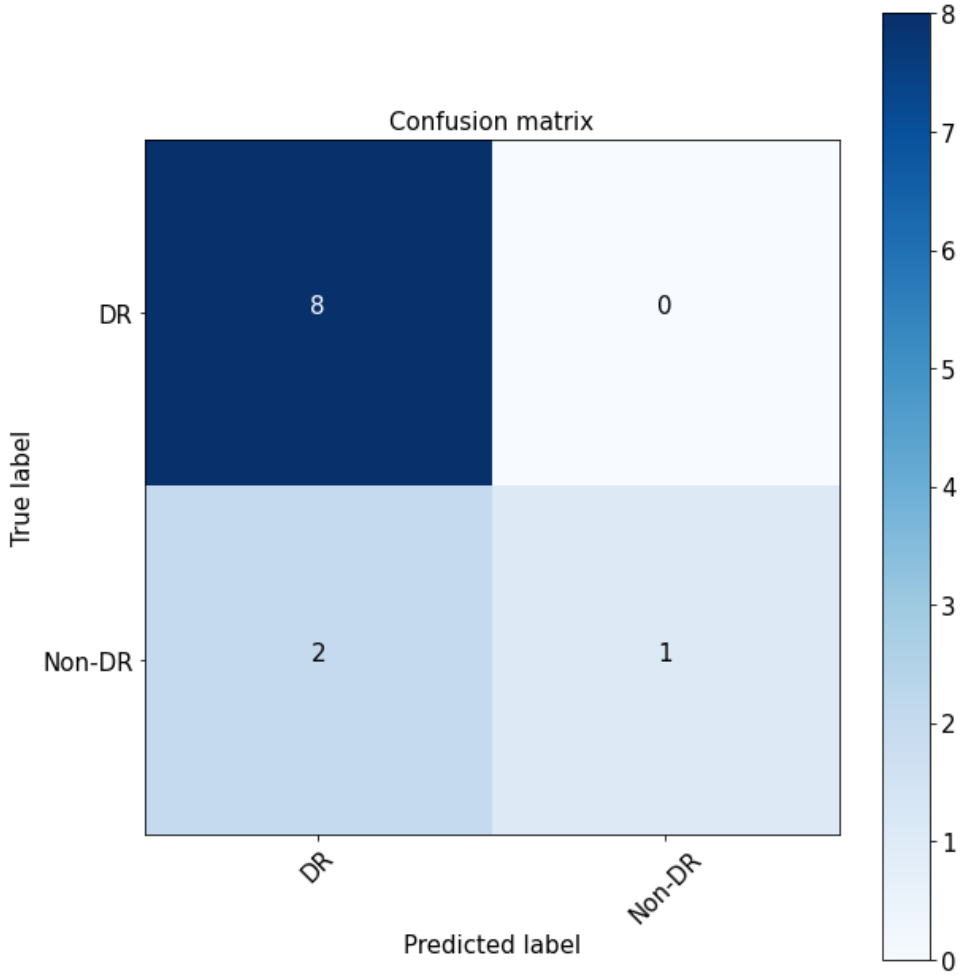

Figure 3.3: Confusion matrix example of size 2x2.

3.5.3 Precision, Recall and F1 Score

Precision represents the ratio between correctly predicted positives and all samples predicted as positives. Recall, in turn, represents the ratio between correctly predicted positives and all actual positives. They are defined as follows:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

There is a tradeoff between these two metrics, in most cases, it is not possible to achieve ideal precision and also recall due to noise in the dataset. If we tune the model to increase precision, then recall will decrease. Often, the model is tuned to the desired ratio between them, but in the case when

both metrics are important we can use F1 score. This metric is harmonic mean between recall and precision, and is calculated as follows:

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

3.5.4 Classification Accuracy

Classification accuracy is a simple commonly used metric to evaluate the performance of classification models. It is calculated by the formula:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

However, this metric is not suitable for an imbalanced dataset. For example, if we take a look at fig. 3.3 and compute accuracy:

$$Accuracy = \frac{8+1}{8+1+2+0} \approx 0.82\%,$$

we obtain quite high value. But in reality, it is not as good as this metric, because the model was able to correctly identify diabetic retinopathy only in a third of samples with DR. What leads to the problem that many people would not receive a proper treatment.

3.5.5 Weighted Cohen's Kappa

Weighted Cohen's kappa metric is used not only in machine learning but in various other areas. This metric indicates how similar are the classifications done by one rater and by another one. In machine learning one of the raters is machine learning model which makes classification, and another rater could be one person as well as a group of people or another model. Weighted Cohen's kappa is very helpful in an imbalanced dataset, especially if classes are ordered. Weighted Cohen's kappa is given by the formula:

$$K_w = 1 - \frac{\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} w_{i,j}p_{i,j}}{\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} w_{i,j}e_{i,j}},$$

where n is the number of classes, $w_{i,j}$ are the weights, $p_{i,j}$ are the observed probabilities and $e_{i,j}$ are the expected probabilities.

Chapter 4

Related Studies

Much work has been done to develop a system that can make the classification of DR automatic and more robust. The majority of studies were aimed at binary classification with DR and no DR classes. Back in 1996, the team from Tennent Institute of Ophthalmology in Glasgow made use of a neural network and photos of fundus divided into small fragments. They achieved sensitivity and specificity of 88% and 83% respectively [39]; and given that model was trained only on approximately 200 photos, the result is really impressive.

Another newer study was done by Jorge de la Calleja et al. [40]. They used Local Binary Patterns for feature extraction. Then to make a classification the features were passed into three different structures as follows: Artificial Neural Networks (ANN), Random Forest (RD), and Support Vector Machine (SVM). The best result was obtained with RD with 97% of accuracy in binary classification.

In 2017 Enrique V. Carrera et al. [41] published their paper, where they used SVM and various operations on images to extract blood vessels, microaneurysms and hard exudates. For example, in the fig. 4.1 you can see the steps to extract hard exudates. Using their methodology they were able to achieve sensibility of almost 95% with a specificity of 66% in binary classification if a patient has or does not have NPDR.

In recent years many studies that use CNN for feature extraction were published with the advent of new large datasets and cheaper training processes. Feng Li et al. [42] using a private dataset of almost 20,000 eye fundus images, transfer learning and Inception-v3 architecture [43] achieved an accuracy of 93%, with a sensitivity and specificity of 97% and 93% respectively. The images were preprocessed with contrast-limited adaptive histogram equalization (CLAHE) [44] and non-local algorithm for image denoising [45].

Marco Alban and Tanner Gilligan in their report [46] compared GoogLeNet [47] and AlexNet [48] models trained on the same dataset maintained by EyePACS as we used in our work. Those models were pretrained on another

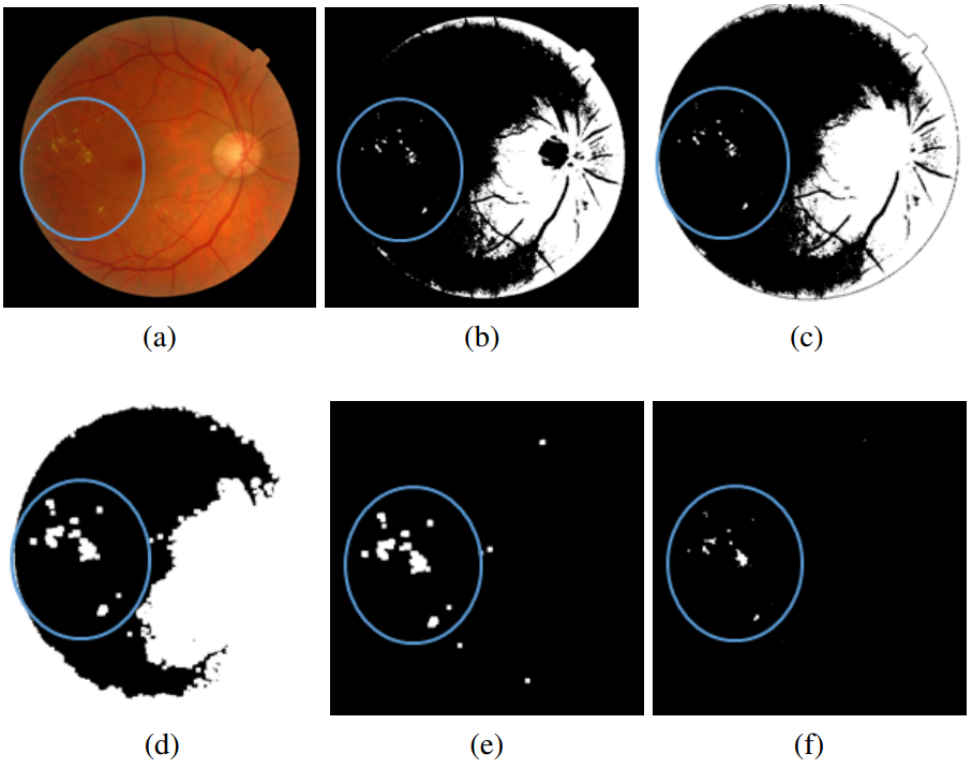

Figure 4.1: Detection of hard exudates (*hE*)—a blue oval highlights region of interest: (a) Original image, (b) Color segmentation, (c) Background removal and optical-disc mask is added, (d) Dilated image, (e) Pre-selected *hE* (f) Finally selected *hE*.
Source: [41]

dataset called ImageNet [49] and then only the last layers of models were trained on retina fundus photos. The classification was done for 2-class, 3-class, and 5-class datasets. GoogLeNet model performed in their study slightly better than AlexNet, its accurate results relative to class numbers are 71%, 58% and 42% respectively.

The most accurate model as far as we know, is developed by researchers from Google Inc and other organizations [50]. They used a large dataset with over 128,000 images, each image was graded by ophthalmologists 5 times on average, the images with low quality were not used in training. The quality of their dataset made a great contribution to model's performance as we can see on the graphs fig. 4.2. They used Inception-v3 architecture with preinitialized weights from ImageNet dataset, ensemble of 10 networks. What is also interesting, probably thanks to the dataset size and quality, they were able to achieve great results almost without preprocessing. The model had a sensitivity of 90.3% and specificity of 98.1% for EyePACS dataset with tuning

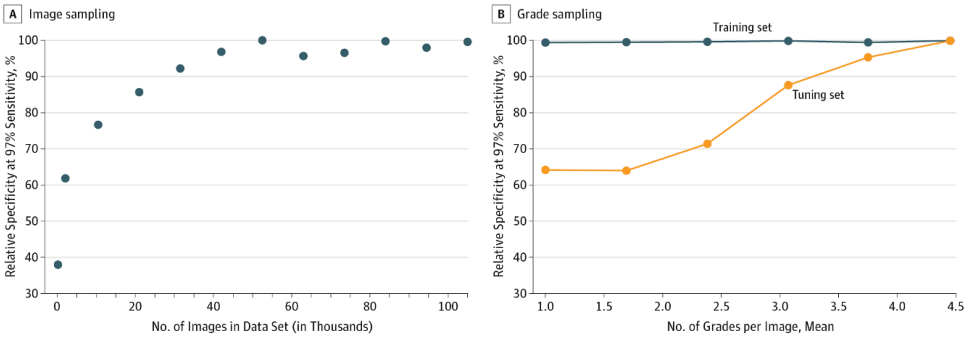

Figure 4.2: Model performance for detection of diabetic retinopathy and all-cause referable diabetic retinopathy as a function of the number of images and grades used in the development set.
Source: [50]

for high specificity. In high sensitivity tuning the model had a sensitivity of 97.5% and specificity of 93.4%. Also, their model was build to classify if diabetic macular edema was developed. In this task they achieved a sensitivity of 90.8% and specificity of 98.7%. However, it is difficult to compare with other studies as this model outputs multiple binary predictions rather than just a single predicted class.

All previous papers were focused only on classification, but for practical use it is more useful to not only see the stage of DR, but also understand why the model reached this conclusion. This type of task is more difficult, but there are also interesting studies with promising results. For example, in the challenge IDRiD report [51], different teams show up to 89% of area under precision-recall curve (AUPR) in hard exudates segmentation and up to 68% in hemorrhages segmentation.

Chapter 5

Implementation

This chapter describes how we build our model for the classification task, different techniques and architecture we tried, the dataset and its problems.

5.1 Dataset and Related Decisions

We used the dataset with retinal fundus photos provided by EyePACS on platform Kaggle. It contains 25810 photos labeled as no DR, 2443 as mild NPDR, 5292 as moderate NPDR, 873 as severe NPDR, and 708 as PDR. And that is the first problem, because of highly imbalanced classes the model pays more attention to the major ones, and therefore we will get the model which is prone to classify every photo as no DR. Another problem with this dataset is a large variation in the quality of photos, according to Mike Voets et al. [52] there are up to 20% ungradable photos (fig. 5.1).

5.1.1 Imbalance Problem Solving

To overcome so-called imbalanced data we have tried different methods such as oversampling, undersampling, weight loss function, and their combination in varying proportions. The oversampling minority classes have been chosen as this approach shows the best performance. This is done by ran-

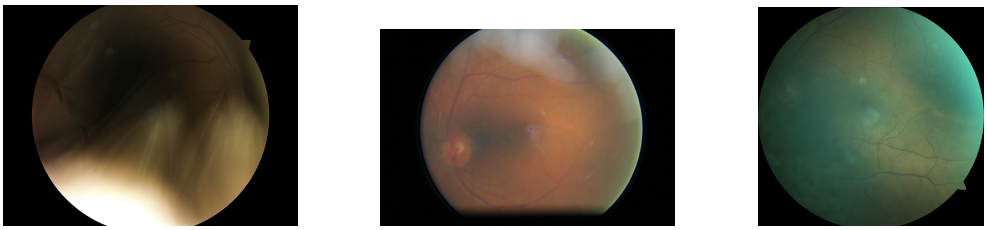

Figure 5.1: Ungradable photos of eye fundus from the EyePACS dataset.

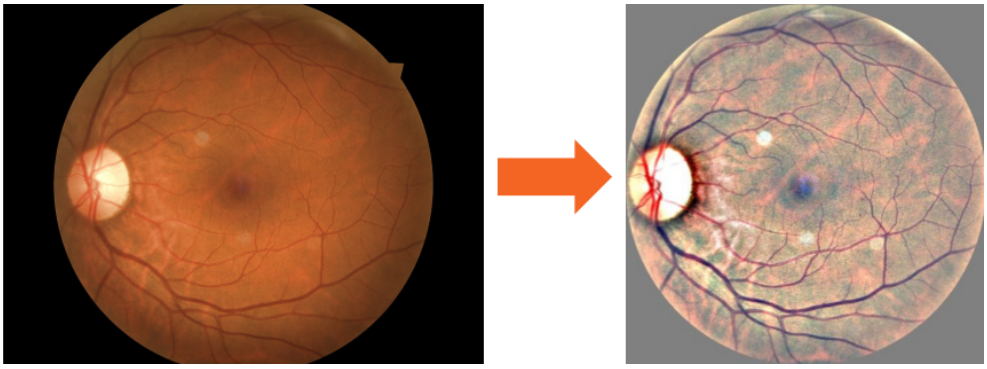

Figure 5.2: Original image and result of preprocessing

domly selecting with samples replacing from minority classes until the number reaches majority class.

5.1.2 Image Preprocessing

The preprocessing was applied on every image to solve variations in images and to enhance them for better feature extraction (fig. 5.2). Each image was preprocessed in several steps:

- Resize image to 900 pixels in width to improve speed of preprocessing
- Substract Gaussian filtered image from original
- Fill every pixel out of the eye with gray
- Shrink borders

Also, we tried to apply CLAHE, and to train without preprocessing at all, but these options were not as good as preprocessing described above.

5.1.3 Augmentation

Augmentation helps to increase the number of data in the dataset using already existed ones. This, in turn, leads to better generalization and therefore to better model performance [53]. We tried several options, parameters, and finally stopped at this:

- Rotations by 90, 180 and 270 degrees
- Horizontal and vertical flips
- Random zooms in 50% images
- Random changes in hue, brightness, saturation and contrast by $\pm$ 10%

5.2 Convolutional Neural Network Architecture

After many experiments with InceptionV3 and transfer learning the best result we achieved was Cohen's kappa of 43%. That is why we chose another architecture called VGG [54], as it was used by many teams in the competition "Diabetic Retinopathy Detection" on the Kaggle platform.

The structure of our network is described below:

1. 2 convolution layers with 32 filters
2. Max pooling
3. Batch normalization
4. 3 convolution layers with 64 filters
5. Max pooling
6. Batch normalization
7. 3 convolution layers with 256 filters
8. Max pooling
9. Batch normalization
10. 2 convolution layers with 512 filters
11. Max pooling
12. Batch normalization
13. 2 convolution layers with 512 filters
14. Max pooling
15. Batch normalization
16. Flatten
17. Fully-connected layer with 256 units
18. Dropout with rate of 50%
19. Fully-connected layer with 256 units
20. Dropout with rate of 50%
21. Fully-connected layer with 5 units

The leaky ReLU was used as an activation function for the entire network except for the last fully-connected layer, where we used softmax function to make classification. The batch size of 16 images was chosen, where each image was normalized and reshaped to 512 by 512 pixels, as this size gave us the best ratio between performance and computation speed. Every convolution layer had a kernel size of 3 by 3 pixels, a stride of 1 pixel and padding around images. Each max pooling layer had the pooling window of 3 by 3 pixels and stride of 2. Also, we used L2 regularization for both, kernel and bias, for each layer that supports it.

We tried to use SGD with momentum, RMSprop, and Adam as our optimizer, finally, the last one was chosen. We used cyclical learning rates [55] for training, as it simplifies the process of finding the right learning rate and also leads to fewer training iterations.

Entire implementation was done using Python programming language, TensorFlow platform for machine learning, libraries such as Pandas, NumPy and OpenCV.

Chapter 6

Results and Discussion

In our best model, we achieved Cohen's kappa of 52%, accuracy of 80% and average F1 score of 55% after 442 training epochs. As you can see in the table 6.1, the model did a good job in predicting samples without any signs of DR, but had problems with other stages of DR, especially with mild NPDR.

DR stage	F1 score in percent
No DR	90
Mild NPDR	33
Moderate NPDR	56
Severe NPDR	44
PDR	49
Average	55

Table 6.1: F1 scores for each DR stage.

The result is not as good as the ones from the competition mentioned above but is comparable to many other studies where a similar dataset was used for 5-class problem [56, 57, 46].

According to competitors' reports, ensemble learning could be used to slightly improve performance [58, 59, 60]. Also, as was assumed by Mike Voets [52] the number of grades per image in Gulshan et al's study [50] could be the main reason for its exceptional result. Therefore, it would reasonable to put more effort into a dataset improvement in the next works.

Conclusion

At the beginning of this thesis, we briefly introduced the application of artificial intelligence in healthcare, diabetic retinopathy disease, and why it is important to develop automatic systems to classify it. After this, we made an overview of the theory required for a better understanding of the following chapters, and of papers that are related to applying machine learning in diabetic retinopathy problem, their approaches, techniques and achieved results. This thesis mainly focused on the building model to classify diabetic retinopathy into 5 stages using a convolutional neural network. The model we built showed promising results, but further improvements are needed to make this model useful in practice.

Bibliography

1. LEDERBERG, Joshua. *Dendral-64 - a System for Computer Construction, Enumeration and Notation of Organic Molecules as Tree Structures and Cyclic Graphs. Part I- Notational Algorithm for Tree Structures.* 1964. Available also from: https://ntrs.nasa.gov/search.jsp?R=19650003557.
2. SHORTLIFFE, Edward. *Computer-Based Medical Consultations: MYCIN.* Elsevier, 1976.
3. SHORTLIFFE, Edward; SCOTT, A.; BISCHOFF, Miriam; CAMPBELL, A.; MELLE, William; JACOBS, Charlotte. ONCOCIN: An Expert System for Oncology Protocol Management. In: 1981, pp. 876–881.
4. PEEK, Niels; COMBI, Carlo; MARIN, Roque; BELLAZZI, Riccardo. Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes. *Artificial Intelligence in Medicine.* 2015, vol. 65, no. 1, pp. 61–73. ISSN 0933-3657. Available also from: http://www.sciencedirect.com/science/article/pii/S0933365715000871.
5. *Somatix.* Available also from: https://somatix.com/.
6. CBINSIGHTS. *The AI Industry Series: Top Healthcare AI Trends To Watch.* Available also from: https://www.cbinsights.com/research/report/ai-trends-healthcare/.
7. MARKETSANDMARKETS. *Artificial Intelligence in Healthcare Market worth $36.1 billion by 2025.* 2018. Available also from: https://www.marketsandmarkets.com/PressReleases/artificial-intelligence-healthcare.asp.
8. HUBEL, D. H.; WIESEL, T. N. Receptive Fields of Single Neurones in the Cat's Striate Cortex. *The Journal of physiology.* 1959, vol. 148, no. 3, pp. 574–591. ISSN 0022-3751. Available also from: https://pubmed.ncbi.nlm.nih.gov/14403679.

9. KHAN, M. Attique; RUBAB, S.; KASHIF, Asifa; SHARIF, Muhammad Imran; MUHAMMAD, Nazeer; SHAH, Jamal Hussain; ZHANG, Yu-Dong; SATAPATHY, Suresh Chandra. Lungs Cancer Classification From CT Images: an Integrated Design of Contrast Based Classical Features Fusion and Selection. *Pattern Recognition Letters*. 2020, vol. 129, pp. 77–85. ISSN 0167-8655. Available also from: http://www.sciencedirect.com/science/article/pii/S0167865519303289.

10. ESTEVA, Andre; KUPREL, Brett; NOVOA, Roberto A.; KO, Justin; SWETTER, Susan M.; BLAU, Helen M.; THRUN, Sebastian. Dermatologist-level Classification of Skin Cancer With Deep Neural Networks. *Nature*. 2017, vol. 542, no. 7639, pp. 115–118. ISSN 1476-4687. Available from DOI: 10.1038/nature21056.

11. *AI-based CT Imaging Company MaxQ*. Available also from: https://www.maxq.ai/.

12. MCKINNEY, Scott Mayer et al. International Evaluation of an Ai System for Breast Cancer Screening. *Nature*. 2020, vol. 577, no. 7788, pp. 89–94. ISSN 1476-4687. Available from DOI: 10.1038/s41586-019-1799-6.

13. LEE, Ryan; WONG, Tien Y.; SABANAYAGAM, Charumathi. Epidemiology of Diabetic Retinopathy, Diabetic Macular Edema and Related Vision Loss. *Eye and vision (London, England)*. 2015, vol. 2, pp. 17. ISSN 2326-0254. Available from DOI: 10.1186/s40662-015-0026-2.

14. ZHENG, Yingfeng; HE, Mingguang; CONGDON, Nathan. The Worldwide Epidemic of Diabetic Retinopathy. *Indian journal of ophthalmology*. 2012, vol. 60, no. 5, pp. 428–431. ISSN 1998-3689. Available from DOI: 10.4103/0301-4738.100542.

15. FLAXMAN, Seth R et al. Global Causes of Blindness and Distance Vision Impairment 1990–2020: a Systematic Review and Meta-analysis. *The Lancet Global Health*. 2017, vol. 5, no. 12, pp. e1221–e1234. ISSN 2214-109X. Available from DOI: https://doi.org/10.1016/S2214-109X(17)30393-5.

16. ANDREW A. DAHL, MD. *Retina Abnormalities: 14 Signs of Systemic Disease*. 2013. Available also from: https://reference.medscape.com/features/slideshow/retina.

17. AMERICAN ACADEMY OF OPHTHALMOLOGY RETINA-VITREOUS PANEL. *Preferred Practice Pattern® Guidelines. Diabetic Retinopathy*. San Francisco, CA: American Academy of Ophthalmology, 2016. Available also from: https://www.aao.org/ppp.

18. NEW ENGLAND RETINA ASSOCIATES. *Stages of Diabetic Retinopathy*. Available also from: https://www.retinamd.com/diseases-and-treatments/retinal-conditions-and-diseases/diabetic-retinopathy/.

19. KASCHKE, Michael. *Optical Devices in Ophthalmology and Optometry: Technology, Design Principles and Clinical Applications*. Weinheim: Wiley-VCH, 2013. ISBN 978-3-527-41068-2.

20. TOSLAK, Devrim; AYATA, Ali; LIU, Changgeng; EROL, Muhammet Kazim; YAO, Xincheng. Wide-field Smartphone Fundus Video Camera Based On Miniaturized Indirect Ophthalmoscopy. *Retina (Philadelphia, Pa.)* 2018, vol. 38, no. 2, pp. 438–441. ISSN 1539-2864. Available from DOI: 10.1097/IAE.0000000000001888.

21. RAJU, Biju; RAJU, N. S. D.; AKKARA, John Davis; PATHENGAY, Avinash. Do it yourself smartphone fundus camera - DIYretCAM. *Indian journal of ophthalmology*. 2016, vol. 64, no. 9, pp. 663–667. ISSN 1998-3689. Available from DOI: 10.4103/0301-4738.194325.

22. JAMES, Gareth; WITTEN, Daniela; HASTIE, Trevor; TIBSHIRANI, Robert. *An Introduction to Statistical Learning*. Springer New York, 2013. ISSN 2197-4136. Available from DOI: 10.1007/978-1-4614-7138-7.

23. RAY, Sunil. *Commonly used Machine Learning Algorithms (with Python and R Codes)*. 2017. Available also from: https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms.

24. VALERII, Filipets. *4 Types of Machine Learning Algorithms*. Available also from: https://theappsolutions.com/blog/development/machine-learning-algorithm-types.

25. PATEL, Nitin. *Data Mining*. 2003. Available also from: https://ocw.mit.edu/courses/sloan-school-of-management/15-062-data-mining-spring-2003/index.htm.

26. BURKE, Laura Ignizio. Introduction to artificial neural systems for pattern recognition. *Computers & Operations Research*. 1991, vol. 18, no. 2, pp. 211–220. ISSN 0305-0548. Available from DOI: https://doi.org/10.1016/0305-0548(91)90091-5.

27. VIEIRA, Sandra; PINAYA, Walter; MECHELLI, Andrea. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neuroscience & Biobehavioral Reviews*. 2017, vol. 74. Available from DOI: 10.1016/j.neubiorev.2017.01.002.

28. GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. *Deep Learning*. MIT Press, 2016. Available also from: http://www.deeplearningbook.org.

29. BISHOP, Christopher. *Pattern recognition and machine learning*. New York: Springer, 2006. ISBN 978-0-387-31073-2.

30. SEXTON, Randall S.; DORSEY, Robert E.; JOHNSON, John D. Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing. *European Journal of Operational Research*. 1999, vol. 114, no. 3, pp. 589–601. ISSN 0377-2217. Available from DOI: https://doi.org/10.1016/S0377-2217(98)00114-3.

31. LI, Fei-Fei; KRISHNA, Ranjay; XU, Danfei. *CS231n: Convolutional Neural Networks for Visual Recognition*. 2020. Available also from: http://cs231n.stanford.edu/.

32. ZHAO, Yang; WONG, Zoie Shui-Yee; TSUI, Kwok Leung. A Framework of Rebalancing Imbalanced Healthcare Data for Rare Events' Classification: A Case of Look-Alike Sound-Alike Mix-Up Incident Detection. *Journal of Healthcare Engineering*. 2018, vol. 2018, pp. 1–11. Available from DOI: 10.1155/2018/6275435.

33. MISHRA, Aditya. *Metrics to Evaluate your Machine Learning Algorithm*. 2018. Available also from: https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234.

34. KOMOROWSKI, M.; CELI, L. A.; BADAWI, O.; GORDON, A. C.; FAISAL, A. A. The Artificial Intelligence Clinician Learns Optimal Treatment Strategies for Sepsis in Intensive Care. *Nat. Med.* 2018, vol. 24, no. 11, pp. 1716–1720. Available from DOI: 10.1038/s41591-018-0213-5.

35. OPENAI et al. Learning Dexterous In-Hand Manipulation. *CoRR*. 2018, vol. abs/1808.00177. Available from arXiv: 1808.00177.

36. LU, Lu; SHIN, Yeonjong; SU, Yanhui; KARNIADAKIS, George Em. *Dying ReLU and Initialization: Theory and Numerical Examples*. 2019. Available from arXiv: 1903.06733.

37. KAUR, Prabhjot; GOSAIN, Anjana. Issues and Challenges of Class Imbalance Problem in Classification. *International Journal of Information Technology*. 2018. ISSN 2511-2112. Available from DOI: 10.1007/s41870-018-0251-8.

38. CHAWLA, N. V.; BOWYER, K. W.; HALL, L. O.; KEGELMEYER, W. P. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*. 2002, vol. 16, pp. 321–357. ISSN 1076-9757. Available from DOI: 10.1613/jair.953.

39. GARDNER, G G; KEATING, D; WILLIAMSON, T H; ELLIOTT, A T. Automatic Detection of Diabetic Retinopathy Using an Artificial Neural Network: a Screening Tool. *British Journal of Ophthalmology*. 1996, vol. 80, no. 11, pp. 940–944. ISSN 0007-1161. Available from DOI: 10.1136/bjo.80.11.940.

40. CALLEJA, Jorge de la; TECUAPETLA, Lourdes; AUXILIO MEDINA, Ma.; BÁRCENAS, Everardo; URBINA NÁJERA, Argelia B. LBP and Machine Learning for Diabetic Retinopathy Detection. In: CORCHADO, Emilio; LOZANO, José A.; QUINTIÁN, Héctor; YIN, Hujun (eds.). *Intelligent Data Engineering and Automated Learning – IDEAL 2014*. Cham: Springer International Publishing, 2014, pp. 110–117. ISBN 978-3-319-10840-7.

41. CARRERA, Enrique; GONZÁLEZ, Andrés; CARRERA, Ricardo. Automated detection of diabetic retinopathy using SVM. In: 2017. Available from DOI: 10.1109/INTERCON.2017.8079692.

42. LI, Feng; LIU, Zheng; CHEN, Hua; JIANG, Minshan; ZHANG, Xuedian; WU, Zhizheng. Automatic Detection of Diabetic Retinopathy in Retinal Fundus Photographs Based on Deep Learning Algorithm. *Translational vision science & technology*. 2019, vol. 8, no. 6, pp. 4. ISSN 2164-2591. Available from DOI: 10.1167/tvst.8.6.4.

43. SZEGEDY, Christian; VANHOUCKE, Vincent; IOFFE, Sergey; SHLENS, Jonathon; WOJNA, Zbigniew. *Rethinking the Inception Architecture for Computer Vision*. 2015. Available from arXiv: 1512.00567 [cs.CV].

44. PIZER, Stephen M.; AMBURN, E. Philip; AUSTIN, John D.; CROMARTIE, Robert; GESELOWITZ, Ari; GREER, Trey; ROMENY], Bart [ter Haar; ZIMMERMAN, John B.; ZUIDERVELD, Karel. Adaptive Histogram Equalization and Its Variations. *Computer Vision, Graphics, and Image Processing*. 1987, vol. 39, no. 3, pp. 355–368. ISSN 0734-189X. Available also from: http://www.sciencedirect.com/science/article/pii/S0734189X8780186X.

45. BUADES, A.; COLL, B.; MOREL, J. -. A Non-local Algorithm for Image Denoising. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. 2005, vol. 2, pp. 60–65.

46. ALBAN, Marco; GILLIGAN, Tanner. Automated Detection of Diabetic Retinopathy using Fluorescein Angiography Photographs. In: Stanford University, 2016.

47. SZEGEDY, Christian; LIU, Wei; JIA, Yangqing; SERMANET, Pierre; REED, Scott; ANGUELOV, Dragomir; ERHAN, Dumitru; VANHOUCKE, Vincent; RABINOVICH, Andrew. *Going Deeper with Convolutions*. 2014. Available from arXiv: 1409.4842 [cs.CV].

48. KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. ImageNet Classification with Deep Convolutional Neural Networks. In: PEREIRA, F.; BURGES, C. J. C.; BOTTOU, L.; WEINBERGER, K. Q. (eds.). *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc., 2012, pp. 1097–1105. Available also from:

http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.

49. RUSSAKOVSKY, Olga et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*. 2015, vol. 115, no. 3, pp. 211–252. Available from DOI: 10.1007/s11263-015-0816-y.

50. GULSHAN, Varun et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*. 2016, vol. 316, no. 22, pp. 2402–2410. ISSN 0098-7484. Available from DOI: 10.1001/jama.2016.17216.

51. PORWAL, Prasanna et al. IDRiD: Diabetic Retinopathy – Segmentation and Grading Challenge. *Medical Image Analysis*. 2020, vol. 59, pp. 101561. ISSN 1361-8415. Available also from: http://www.sciencedirect.com/science/article/pii/S1361841519301033.

52. VOETS, Mike; MØLLERSEN, Kajsa; BONGO, Lars Ailo. Reproduction Study Using Public Data of: Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *PLOS ONE*. 2019, vol. 14, no. 6, pp. 1–11. Available from DOI: 10.1371/journal.pone.0217541.

53. TAYLOR, Luke; NITSCHKE, Geoff. *Improving Deep Learning using Generic Data Augmentation*. 2017. Available from arXiv: 1708.06020.

54. SIMONYAN, Karen; ZISSERMAN, Andrew. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2014. Available from arXiv: 1409.1556.

55. SMITH, Leslie N. *Cyclical Learning Rates for Training Neural Networks*. 2015. Available from arXiv: 1506.01186.

56. CHALLA, Uday Kiran; YELLAMRAJU, Pavankumar; BHATT, Jignesh S. A Multi-class Deep All-CNN for Detection of Diabetic Retinopathy Using Retinal Fundus Images. In: DEKA, Bhabesh; MAJI, Pradipta; MITRA, Sushmita; BHATTACHARYYA, Dhruba Kumar; BORA, Prabin Kumar; PAL, Sankar Kumar (eds.). *Pattern Recognition and Machine Intelligence*. Cham: Springer International Publishing, 2019, pp. 191–199. ISBN 978-3-030-34869-4. Available from DOI: 10.1007/978-3-030-34869-4_21.

57. MASOOD, S.; LUTHRA, T.; SUNDRIYAL, H.; AHMED, M. Identification of Diabetic Retinopathy in Eye Images Using Transfer Learning. In: *2017 International Conference on Computing, Communication and Automation (ICCCA)*. 2017, pp. 1183–1187.

58. XU, Jun; DUNAVENT, John; KAINKARYAM, Raghu. *Summary of our Solution to the Kaggle Diabetic Retinopathy Detection Competition*. 2015. Available also from: `https://www.kaggle.com/c/diabetic-retinopathy-detection/discussion/15845`.

59. ANTONY, Mathis; BRÜGGEMANN, Stephan. *Kaggle Diabetic Retinopathy Detection Team o_O solution*. 2015. Available also from: `https://github.com/sveitser/kaggle_diabetic`.

60. GRAHAM, Ben. *Kaggle Diabetic Retinopathy Detection competition report*. 2015. Available also from: `https://www.kaggle.com/c/diabetic-retinopathy-detection/discussion/15801`.

Appendix A

Acronyms

AI Artificial intelligence

ANN Artificial neural network

AUC-ROC Area under the ROC Curve

AUPR Area under precision-recall curve

CLAHE Contrast-limited adaptive histogram equalization

CNN Convolutional neural network

CT Computed tomography

DR Diabetic retinopathy

FN False negatives

FOV Field of view

FP False positives

MAE Mean absolute error

MRI Magnetic resonance imaging

MSE Mean squared error

NPDR Nonproliferative diabetic retinopathy

PDR Proliferative diabetic retinopathy

ROC Receiver operating characteristic curve

ReLU Rectified linear unit

SGD Stochastic gradient descent

SMOTE Synthetic minority over-sampling technique

TN True negatives

TP True positives

Appendix B

Contents of enclosed SD card

- `src` ..... the directory of source codes
  - `models` ..... the directory with models implementations
  - `modules` ..... the directory with helping codes
  - `preprocessing` ..... the directory with preprocessing codes
- `text` ..... the thesis text directory
  - `src` ..... the directory of LaTeX source codes of the thesis