BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

USING ADVANCED SEGMENTATION METHODS FOR IMAGES FROM TEM MICROSCOPES

VYUŽITIE POKROČILÝCH SEGMENTAČNÝCH METÓD PRE OBRAZY Z TEM MIKROSKOPOV

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR — **Bc. Štefan Mocko**
AUTOR PRÁCE

SUPERVISOR — **Ing. Tomáš Potočňák**
VEDOUCÍ PRÁCE

BRNO 2018

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Master's Thesis

Master's study field **Biomedical Engineering and Bioinformatics**
Department of Biomedical Engineering

Student: Bc. Štefan Mocko ***ID:*** 155593

Year of study: 2 ***Academic year:*** 2017/18

TITLE OF THESIS:

Using advanced segmentation methods for images from TEM microscopes

INSTRUCTION:

1) Elaborate a literary research in the field of adaptive segmentation methods. Focus mainly on methods using artificial neural networks and their applicability on transmural electron microscope (TEM) images. 2) Choose theoretically the most appropriate approach to segmenting the characteristic patterns of TEM microscope calibration samples. 3) Define a way of evaluating the results according to the acquisition parameters of the electron microscope. 4) Create a test set of images in collaboration with FEI Czech Republic (part of Thermo Fisher Scientific). 5) Test the selected methods on model and real images and compare the results with the commercially available methods. Process the selected algorithms in Matlab or Python. 6) Discuss the results obtained and evaluate the efficiency and usability of the methods. The diploma thesis is created in cooperation with the FEI of the Czech Republic (part of Thermo Fisher Scientific).

RECOMMENDED LITERATURE:

[1] KARLIK M., Úvod do transmisní elektronové mirkoskopie. Praha ČVUT, 2011. ISBN 978-80-01-04729-3.

[2] JAN, J. Digital Signal Filtering, Analysis and Restoration. volume 44. volume 44. London: The Institution of Electrical Engineers, 2000. 407 s. ISBN: 0-85296-760- 8.

Date of project specification: 5.2.2018 ***Deadline for submission:*** 18.5.2018

Leader: Ing. Tomáš Potočňák
Consultant:

prof. Ing. Ivo Provazník, Ph.D.
Subject Council chairman

ABSTRACT

This master's thesis deals with the use of a convolutional neural networks for the segmentation task on images from transmission electron microscope. It also describes chosen neural network topology - U-NET, used augmentation techniques and programming environment. ThermoFisher Scientific (formerly FEI Czech Republic s.r.o.) provided data for this thesis. Obtained segmentation results are presented in the form of curves (ROC, PRC) and numerical values (ARI, DSC, Confusion matrices). Chosen U-NET topology achieved excellent results in the field of pixel-wise segmentation, and hopefully, these results will serve as a starting point for internal company research.

KEYWORDS

Neural network; U-net; Convolution neural network; TEM; segmentation

ABSTRAKT

Tato magisterská práce se zabývá využitím konvolučních neuronových sítí pro segmentační účely v oblasti transmisní elektronové mikroskopie. Také popisuje zvolenou topologii neuronové sítě - U-NET, použíté augmentační techniky a programové prostředí. Firma Thermo Fisher Scientific (dříve FEI Czech Republic s.r.o) poskytla obrazová data pro účely této práce. Získané segmentační výsledky jsou prezentovány ve formě křivek (ROC, PRC) a ve formě numerických hodnot (ARI, DSC, Chybová matice). Zvolená U-NET topologie dosáhla excelentních výsledků v oblasti pixelové segmentace. S největší pravděpodobností, budou tyto výsledky sloužit jako odrazový můstek pro interní firemní výzkum.

KLÍČOVÁ SLOVA

Neuronová síť; U-net; Konvoluční neuronová síť; TEM; segmentace

MOCKO, Štefan. *Using advanced segmentation methods for images from TEM microscopes*. Brno, 2018, 80 p. Master's Thesis. Brno University of Technology, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Advised by Ing. Tomáš Potočňák

DECLARATION

I declare that I have written the Master's Thesis titled "Using advanced segmentation methods for images from TEM microscopes" independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the thesis and listed in the comprehensive bibliography at the end of the thesis.

As the author I furthermore declare that, with respect to the creation of this Master's Thesis, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll., Section 2, Head VI, Part 4.

Brno ....

author's signature

ACKNOWLEDGEMENT

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Tomáši Potočňákovi, za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno ....

author's signature

CONTENTS

Introduction **11**

1 TEM **12**
1.1 Components . . . 12
1.1.1 Electron source . . . 13
1.1.2 Vacuum system . . . 14
1.1.3 Lenses . . . 15
1.1.4 Apertures . . . 16
1.1.5 Detectors . . . 17
1.2 Image acquisition . . . 17
1.3 Aberrations . . . 18
1.3.1 Spherical aberration . . . 18
1.3.2 Chromatic aberration . . . 18
1.3.3 Astigmatism . . . 19
1.4 Contrast Formation . . . 20

2 Segmentation **21**
2.1 Definition of segmentation . . . 21
2.1.1 Segmentation terminology . . . 22
2.2 Segmentation methods . . . 22
2.2.1 Edge-based detection . . . 23
2.2.2 Region-based detection . . . 24
2.2.3 Statistic-based detection . . . 25
2.2.4 Knowledge detection . . . 26
2.2.5 Hybrid-based detection . . . 26

3 Convolution neural networks **28**
3.1 Convolution . . . 28
3.1.1 Convolution layers . . . 29
3.1.2 Convolution kernels . . . 29
3.2 Architecture . . . 30
3.2.1 AlexNet . . . 30
3.2.2 VGG Net . . . 33
3.2.3 GoogLeNet . . . 33
3.2.4 Microsoft ResNet . . . 34
3.2.5 U-NET . . . 35

4 Testing concept **37**
4.1 Deep Learning frameworks . . . 37
4.2 Input data . . . 38
4.3 Augmentation . . . 39
4.3.1 Rotate . . . 40
4.3.2 Zoom . . . 40
4.3.3 Distort . . . 40
4.4 CNN selection . . . 41
4.5 Hardware . . . 42
4.6 Learning . . . 43
4.6.1 Score and Loss . . . 43
4.6.2 Optimizer . . . 43
4.6.3 Activation . . . 44
4.6.4 Number of iterations . . . 46
4.7 Evaluation . . . 46
4.7.1 Dice similarity coefficient . . . 46
4.7.2 Rand index . . . 47
4.7.3 ROC curves . . . 47
4.7.4 Precision recall-curves . . . 48
4.7.5 Confusion matrices . . . 49

5 Results **50**
5.1 Training size influence . . . 50
5.2 Number of iterations influence . . . 53
5.3 Edge and Region comparison . . . 56
5.4 Neural network comparison . . . 57

6 Conclusion **60**

Bibliography **62**

List of appendices **73**

A Graphical results **74**

B Numerical results **79**

LIST OF FIGURES

1.1 Transmission electron microscope – simplified schema [6]. 12
1.2 Secondary signals from electron interaction [10]. 13
1.3 Gun tips [11]. 14
1.4 Electrons path through electromagnetic lens [14]. 16
1.5 Astigmatism example[24]. 19
2.1 Edge types (a) Step edge, (b) Ramp edge, (c) Line edge, (d) Roof edge [40]. 23
2.2 Kohonen map topology [47]. 25
3.1 Example of kernel slide [72]. 29
3.2 Chart of different network topologies. [61] 31
3.3 AlexNet layout [62]. 32
3.4 Topology of GoogLeNet [68]. 33
3.5 Inception module of GoogLeNet [67]. 34
3.6 Residual learning: a building block [70]. 35
3.7 U-Net architecture. [73] 36
4.1 Workflow diagram. 38
4.2 Test data variability example. 39
4.3 Graphical representation of rotation operation on model image [74]. . 40
4.4 Graphical representation of random zoom operation on model image [74]. 41
4.5 Graphical representation of random distortion operation on model image [74]. 41
4.6 Relu activation function [60]. 45
4.7 Sigmoid activation function [60]. 46
4.8 ROC example. 48
5.1 Test image - spinel sample. 51
5.2 Test image - graphitized carbon sample 2. 51
5.3 Test image - crossgrating sample. 52
5.4 Dice results. 53
5.5 ARI results. 53
5.6 ROC comparison across training datasets. 54
5.7 PRC comparison across training datasets. 55
5.8 Test image - graphitized carbon sample 1. 55
5.9 Test image - graphitized carbon sample 3. 56
5.10 Test image - spinel sample 1. 56
5.11 Confusion matrix results - best models. 58
5.12 a - original images, b - edge based detection, c - region based detection 59

A.1 ROC. Training Dataset RDZ1000, 10 iterations. Model 1. 74
A.2 PRC. Training Dataset RDZ1000, 10 iterations. Model 1. 74
A.3 ROC. Training Dataset RDZ5000, 10 iterations. Model 8. 75
A.4 PRC. Training Dataset RDZ5000, 10 iterations. Model 8. 75
A.5 ROC. Training Dataset RDZ10000, 10 iterations. Model 5. 76
A.6 PRC. Training Dataset RDZ10000, 10 iterations. Model 5. 76
A.7 ROC. Training Dataset RDZ10000, 15 iterations. Model 4. 77
A.8 PRC. Training Dataset RDZ10000, 15 iterations. Model 4. 77
A.9 ROC. Training Dataset RDZ10000, 20 iterations. Model 1. 78
A.10 PRC. Training Dataset RDZ10000, 20 iterations. Model 1. 78

LIST OF TABLES

1.1 Conversion - pressure units. 15
1.2 Vacuum levels. 15
4.1 Example of confusion matrix. 49
5.1 Final DSC and ARI results with 2σ 57
B.1 DSC - score and Adjusted Rand Index results for RDZ1K 79
B.2 DSC - score and Adjusted Rand Index results for RDZ5K 79
B.3 DSC - score and Adjusted Rand Index results for RDZ10K 80
B.4 DSC - score and Adjusted Rand Index results for RDZ10KITER15 and RDZ10KITER20 80

INTRODUCTION

Importance and necessity of digital image processing originate from two principal application areas: the first is an improvement of the acquired information for human interpretation. The second is processing the obtained data for autonomous machine perception [1]. The area of image processing has a broad range of applications, for example, image storage, medical imaging, acoustic imaging, forensic science, industrial automatization, remote sensing or segmentation. Authors [1] consider as the most challenging step in image analysis the autonomous segmentation. Mostly because the resulting quality of the segmentation then determines the success of the subsequent image analysis. Therefore, be able to provide excellent results is a must.

In 2010 The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [3] has been introduced. Since then is this challenge running annually and serves as an object recognition benchmark. In 2015 was for the first time that neural network solution outperformed human accuracy [4]. This, together with hardware and data centres capacity improvements, leads to availability to create more deep and successful neural architectures. The area of segmentation in biomedical area is widespread. Non-adaptive methods are usually highly specific for particular segmentation task. Therefore neuron networks seem to be an effective tool for competition against other approaches. Because of that, this thesis is focusing on segmentation methods for images from TEM microscopes with the primary focus on the application of neural networks.

The first chapter provides an introductory tour through the area of transmission electron microscopy. Next chapter´s focus is on available segmentation techniques. These techniques are divided and described for user‘s understanding.

The most important delivery of this thesis starts from the third chapter and continues to the end of the thesis. At first, the concept and famous architectures of convolution neural networks are introduced. Then the most significant part follows – description of the experimental section. The chapter provides information about chosen neural network topology – U-NET, test data, augmentation principles, neural network training and evaluation parameters. Segmentation results are then presented with a thorough description. Results are compared with available edge and region-based techniques.

Simultaneously with this thesis was also written a conference paper [2] that presented achieved segmentation results on Conference Student EEICT 2018.

1 TEM

Transmission electron microscopy (TEM) is in principle very similar to a regular optical microscope [5]. Both microscopes serve for magnification of an examine specimen. Optical lenses are not used in TEM. Instead, TEM systems use electromagnetic ones. Another difference is in particles that form us output image. Optical microscope uses photons while TEM uses accelerated electrons. The specimen for electron microscope has to be placed in the vacuum chamber and electrons, that pass through the sample are detected with suitable detectors.

1.1 Components

The transmission electron microscope can be represented by three main parts, see figure 1.1:

1. top part - a source of the particles - an electron gun;
2. middle part - electron lenses and a specimen area;
3. bottom part - detectors.

This simple division is sure not sufficient. Therefore the followed up paragraphs will provide further explanation of mentioned parts.

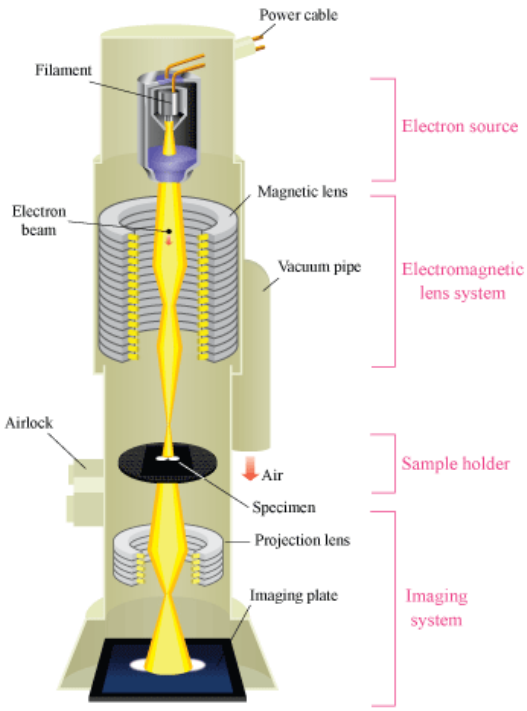

Fig. 1.1: Transmission electron microscope – simplified schema [6].

1.1.1 Electron source

Electron particles are and will be mentioned multiple times. But what is the reason that electrons are used so many times? Why the protons or other particles are not used? Technically there is a possibility to use protons [7], even the alpha particles [8] for microscopy, but the technology and cost/usability ratio is given a substantial advantage to electrons. Another reason why electrons are used is due to their possibility to carry sufficient amount of kinetic energy. Electrons are considered as an ionisation source [8]. This means that they can remove tightly bound inner-shell electrons from atoms or molecules - ionising them. The advantage here is that the ionising radiation produces a wide range of secondary signals from the specimen, see Figure 1.2. With appropriate detecting devices, one can get a better picture of specimen properties. In order to obtain the best signal for detection device, there must also be the best signal on an entering side. This leads to high requirements for quality of electron source [9].

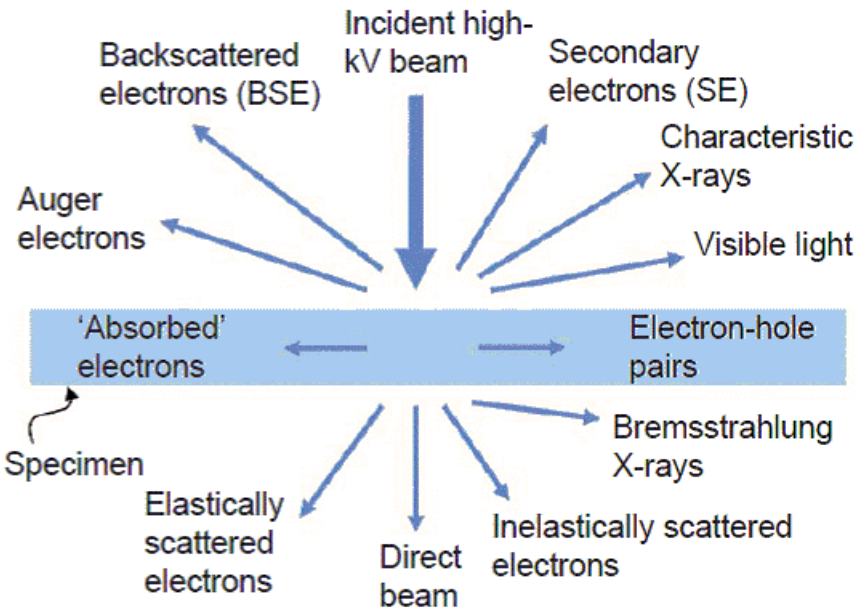

Fig. 1.2: Secondary signals from electron interaction [10].

On current market, two types of electron sources are in use. TEM systems can either use thermionic source or field-emission one [8]. Unfortunately, these two sources are not interchangeable. Both technologies have their specific technical challenges that prevent their replaceability.

Thermionic emission

Thermionic sources were tungsten (W) filaments. Now they are usually substituted with crystals of the lanthanum hexaboride (LaB_6) [13]. This method requires high temperatures that give electrons sufficient energy to overcome the barrier that prevents them from leaking out from a material. The temperature is so high that most materials with such energy donation melt or vaporise. Therefore the two mentioned materials are most commonly used [9].

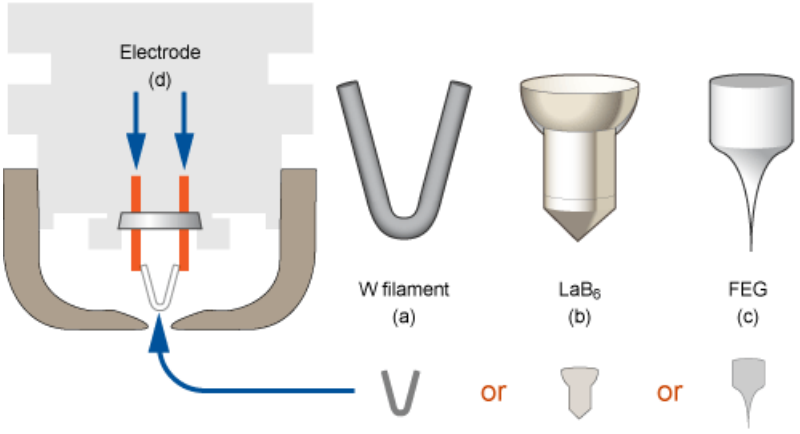

Fig. 1.3: Gun tips [11].

Field emission

Field emission (FE) sources (commonly known as "guns") looks like fine-tuned tungsten needles. The fundamental operating principle is not based on high temperature but the high electric field. Due to this high electrical potential difference, electrons are drawn from the tip of very fine needle wire (tungsten). For FE to occur, a surface has to be entirely clear and pristine – no oxides or other contaminants. The same is applicable for surroundings – to achieve such cleanest, an ultra-high vacuum is demanded (less than $10^{-8}Pa$). When the conditions of vacuum are weaker, the tip can be slightly heated up – but again in a long-term scenario, the source´s life will be shortened [12].

1.1.2 Vacuum system

The quality of vacuum inside the system significantly influences final image quality. This statement is based on an idea, that better vacuum scatters less electron beam particles. When the TEM system is operational, the pumps are maintaining proper vacuum level. The system usually needs to be vented when something goes wrong,

Tab. 1.1: Conversion - pressure units.

Torr	Bar	Pascal
760	1	10^5

Tab. 1.2: Vacuum levels.

Vacuum level	Vacuum Pressure [Pa]
Rough	$100 - 0.1$
Low	$0.1 - 10^{-4}$
High	$10^{-4} - 10^{-7}$
Ultra high	10^{-7}

and service engineer needs to act. Otherwise, there is no reason to vent the system. The process of acquiring appropriate vacuum level is time-consuming. Therefore even specimen replacement occurs when a system has a high vacuum inside. Specially designed sample holders enable specimen replacement. Thanks to these holders, samples can be manipulated on different axes. Various axis movements depend on the manufacturer and technical requirements for a specific holder [9].

Vacuum levels

In TEM articles around the world are still used in non-SI units. Therefore in table 1.1 are included conversions among Pascals, Torrs and Bars.

Vacuum pressures are very low, but in electron microscopy, a naming convention is exactly opposite - lower pressure means a higher vacuum. In this area, vacuum levels are split into four groups – Rough (RV), Low (LV), High (HV), Ultra high (UHV) vacuums [9]. For better understanding, see table 1.2.

1.1.3 Lenses

Electron beam needs to be somehow focused. For the focusation, either electric or magnetic field is used. Commonly electromagnetic lenses are in use. An electromagnetic lens is a coil of wire through which current flows [14]. During the time that current flows, a magnetic field is created around the coil. This field pushes inwards into the hole in the centre of the lens. Therefore the resulting path of passing electrons has a spiral shape. Purpose of these electromagnetic lenses is same as the optical ones. They change the path of the electrons in the desired direction.

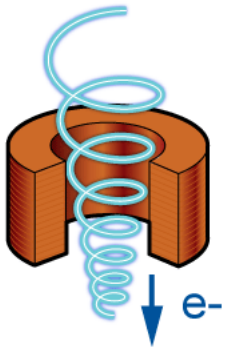

Fig. 1.4: Electrons path through electromagnetic lens [14].

Resolution

Louis de Broglie was first, who theorised in 1925 that electrons could have wave-like character, where the wavelength should be substantially shorter than visible light – the only "particle" source at that time microscopy.

$$\delta = \frac{0.61\lambda}{\mu \sin \beta} \tag{1.1}$$

Resolution can be presented as the smallest distance between two points that can be recognised. Light microscopes can achieve resolution of approx. 300 nm (When the green light 550 nm is used, and the numerical aperture is approximated to[1], equation 1.2.

$$\lambda = \frac{1.22}{E^{\frac{1}{2}}} \tag{1.2}$$

When electrons are accelerated up to 100 keV, then accordingly to equation 1.2 the result $\lambda = 0.004$ nm. Putting this value back into equation 1.1 and final δ will be approx. 2 pm. Of course, this is a very simplified way how to highlight the finer resolution of TEM. There are technological drawbacks that cannot allow achieving such a high-resolution [9].

1.1.4 Apertures

Function principle of an aperture in TEM system is very similar to aperture behaviour in photography or optical microscopy. A smaller opening leads to greater depth of field and the other way round [14]. In TEM there are apertures placed in two locations. The first aperture is located near the top of the column. Here it maintains the coherence of the electron beam. The second aperture is located

[1] $\mu sin\beta = 1$

below the sample, just after the objective lens [15]. Purpose of this second aperture is to control the contrast of the image [9]. From the technological point of view are apertures supplied as interchangeable strips that can be swapped when different conditions are needed.

1.1.5 Detectors

Historically, electrons were observed on a viewing screen that was coated with a material able to emit light. Nowadays, the technological development led to getting rid of this old design habit and provide the output only on a computer display. Several alternatives to fluorescent screens were introduced, but they can be assigned to three main categories:

1. semiconductor detectors,
2. scintillator-photomultiplier detectors.
3. CCD.

Generally, semiconductor detectors are easy to fabricate, cheap to replace and can be cut into any shape [9]. Drawbacks are in large dark current (current is registered even when no signal is incident on the detector) and insensitivity to low-energy electrons. The Second group of detectors has lower noise level and very high gain of the system. Unfortunately, there are also disadvantages. The energy-conversion efficiency is lower [9] compared to semiconductor detector and is substantially more expensive to manufacture. The CCDs group advantages are in very low noise and excellent detective quantum efficiency (DQE)[16], when they are cooled. They have a high dynamic range and linear response to input signal changes. With further development, the high price will eventually drop to tolerable level. Currently, the most severe drawback is in "blooming" [17], when too much signal fills up the pixel, and another incoming signal overflows into rounding pixels.

1.2 Image acquisition

Starting path of image acquisition is in an electron source, where electrons are leaving the tungsten or LaB_6 tip. This beam of electrons is accelerated towards to anode plate and travel through a hole in anode plate. The path continuous when they enter the columns section. First "contact" is with the condenser lens, where the beam is concentrated and brought to the point of focus, which lies above the specimen plane [9]. The condenser system is composed of small physical aperture holes in a shape of a disc. This first aperture controls the intensity and the convergence of electron beam. This chose of aperture size with a level of focus forms the final image

"brightness". When the electrons encounter the specimen, three possible scenarios can occur [18]:

1. they pass unimpeded,
2. they are elastically scattered – no energy is lost,
3. they are in-elastically scattered – exchange of energy happen.

The image contrast varies on an ability to prevent in-elastically scattered electrons contribute to the final image. This is achieved by second small aperture hole just below the specimen [9]. The specimen itself is surrounded by the objective lens which magnifies the image of the sample. The image is further magnified with a use of two different lenses below the sample – intermediate and projector lenses. Final path of the electrons ends when they hit the screen, coated with a fluorescent material or they hit the detector [19].

1.3 Aberrations

The similarity of electromagnetic lenses with the optical ones is also in aberrations. Optical lenses can have manufacturing flaws. Electromagnetic lenses have them also, but they are also other factors to consider - temperature, voltage supply stability, quality of the vacuum. If all lens aberrations have been eliminated, then there would be a possibility to achieve maximal theoretical resolution.

1.3.1 Spherical aberration

The structure of electromagnetic lens is made from coils of wire that are surrounded by a soft iron core. This technological design is creating an electromagnetic field that affects the electron beam. The effect on electron beam is similar to the behaviour of converging glass lens to photons. The effect can be introduced as a point of focus that displays as a disk of finite size, with a high-intensity region surrounded by a halo of decreasing intensity [9]. The problem here is that that part of the beam that enters the lens near the perimeter is brought to focus at a slightly different spot than the beam near centre [20]. This phenomenon is known as a spherical aberration. The easiest way how to deal with this problem is to cut off the parts of a beam that is on the edge sections of a lens with an aperture hole.

1.3.2 Chromatic aberration

Term "chromatic" can indicate something related to the colour of electrons. Due to wave-particle duality, chromatic aberration is related to the energy of electrons. The assumption so far was that electrons are monochromatic, although this is not

exactly correct. With very good high voltage supply the energy variation among the electrons is in 100 keV beam less than 0.1 eV (0.3 eV cold FEG) [9]. Because of such a small range, the chromatic aberration can be ignored unless the specimen is inserted into the beam path. Electrons passing the specimen lost a various amount of energies, and objective lens then bends lower energy electrons more strongly [21] resulting in similar blurred disc object as in previous spherical aberration. Ways how to solve this problem is several, either we can put another aperture below the specimen, or do something cheaper - create thinner sample [9]. Thinner the sample, smaller the chromatic aberration.

1.3.3 Astigmatism

Astigmatism refers to the geometry/shape of the beam. A common way how to see it is with sphere samples, where the spheres are in some axes deformed into an elliptical shape. It occurs when a non-uniform magnetic field influences the electrons on their path through the column. Main reasons are due to inability to manufacture perfectly cylindrical soft iron parts and possible microstructural inhomogeneities that can cause local irregularity in the magnetic field [9]. Another reason could be due to contamination that can be charged up and cause local deflections in parts of the beam. Correction of astigmatism is done with stigmators (electromagnets) – small octupoles that create compensation field to balance the inhomogeneities [22].

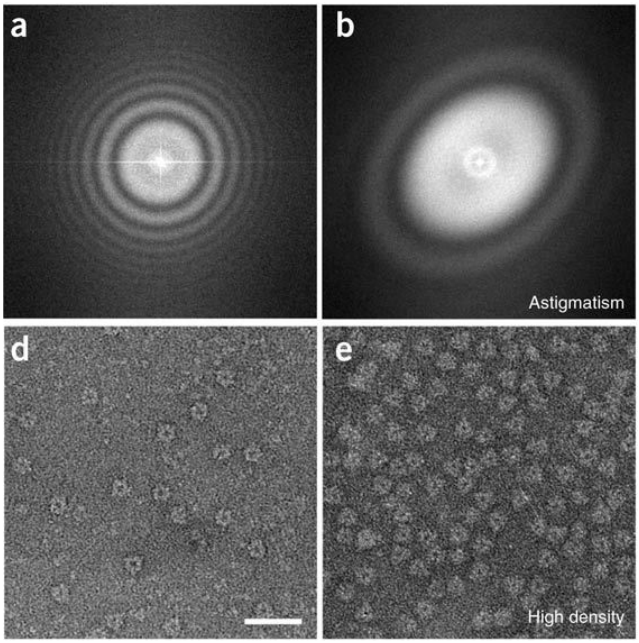

Fig. 1.5: Astigmatism example[24].

1.4 Contrast Formation

TEM systems can through various modes provide desired contrast images. Different imaging modes modify electron waves that are exiting from samples in order to discern useful information.

- **Bright field** - The most common mode of TEM systems is bright field imaging mode. Contrast formation is made from different absorption of electrons in a specimen. Thicker regions or segments made from higher atomic numbers will absorb more, and as a result, they will form darker areas in the output image [9].
- **Diffraction contrast** - This particular mode is based on the ability of a specimen to variously disperse electrons in the back of the focal plane. Various placement of the objective aperture can filter dispersed electron, which will result in the image composed of specific diffraction conditions. Electrons scattered in other areas are not contributing to the final image [9].
- **Phase contrast** - Contrast can also be made due to differences in the phase of electrons that are scattered through a specimen. The appearance of the final image is very sensitive to many factors, not only in phase contrast. Factors that are contributing to the resulting quality are a thickness of the sample, orientation or astigmatism of the objective lens [9]. Specific of this method is that final image is not created from one electron beam, but with thin-film plates (or defocus can be used) [25] there are several individual beams created that contribute into the final high contrast image.

2 SEGMENTATION

Segmentation is considered as a method of image processing [28]. It can mark areas that correspond to objects in particular scene. This "marking" is performed based on similar properties of each pixel (e.g., colour, brightness, material, ...) that represent objects. Results, obtained from segmentation, are later processed by higher-order methods [26] to fulfil tasks like a presence of a specific object or finding and classifying objects in the scene.

2.1 Definition of segmentation

Image segmentation is the process of partitioning original digital image $f(x, y)$ into sub-images $R_1, R_2, ..., R_n$ (segments), without any attempt to understand what these segments represent [31], in a way where these conditions are fulfilled [27]:

1. If all segments are merged - original image is obtained.
2. Two different segments are mutually disjoint.
3. Every segment fulfils some other statement (e.g. All pixels in segment R_j have the same value).

Two situations can occur. Any segments correspond to objects in the scene or not, then a total or partial segmentation is happening. The last condition makes defining objects easier. The segmentation can occur on various presumptions:

1. Object has the same color,
2. Object has same texture,
3. Object has a square shape, etc.

Due to the possibility of various object definitions, segmentation is considered as a challenging image processing technique [27]. No single computational algorithm could be implemented in every application [29]. Various approaches can be used [27], [30]:

- edge-based methods,
- region-based methods,
- statistical methods,
- hybrid methods,
- knowledge-based.

One another reason, why the segmentation algorithms are considered as a challenging, is also a different quality of acquired images [38]. Noise or inhomogeneous illumination usually corrupt obtained data. Objects can have complicated structures, and they can overlap etc. Sometimes even trained human eye can have problems to distinguish individual objects.

2.1.1 Segmentation terminology

In the area of segmentation it is possible to encounter several segmentation terms, namely: Semantic segmentation [32, 33, 34], scene labeling [35], pixel-level segmentation [36] pixel-wise segmentation [37]. This paragraphs intention is to clarify these terms.

Semantic segmentation attempts to partition the image into semantically meaningful parts and then classify each segment into one of the pre-determined classes. There is also a possibility to classify each pixel (rather than the entire segment) to achieve the same goal. In this case, the pixel-wise classification is performed - both approaches lead to the same result but with a slightly different path. We can say that semantic segmentation, scene labelling and pixel-wise/pixel-level segmentation are trying to achieve the same thing - semantically understand the role of each pixel in the image. This understanding is accomplished via slightly different paths that causes slight nuances in terminology.

2.2 Segmentation methods

There is a lot of confusion among segmentation algorithms and segmentation methods [39]. For simplification, a method is an idea how to solve a problem, and an algorithm is solution implementation. Implementation can be performed multiple times, every time the actual implementation can be different. Adaptive segmentation methods should share at least one characteristic property, that should be their ability to adapt. Either they should adjust locally or globally. Local adaptation can be pictured as an image that is partially corrupted, e.g. overexposure. A picture still carries the same information, but only on the part of the image. Something inside the algorithm has to be slightly changed to obtain the best result. Global adaptation can one imagines as a pallet of different images of the same thing, a picture of a tree. Every time it has different bark, colour, leaf shapes but the algorithm has to adapt to these different tree variations. One possible approach how to separate segmentation methods is according to the list below [27]:

- **Edge-based** – This first group detects significant edges inside the image. Several edge detection algorithms can provide sufficient outputs.
- **Region-based** – These methods are similar to the first group. If edges are detected, then algorithms should be able to mark regions in the image, although object contours may be broken.
- **Statistic-based** – the Core principle of segmentation is here statistical analysis performed on pixel values. Structure information is most commonly neglected.

- **Hybrid-based** – Some methods are hard to put only into one category. Usually, all three mentioned methods are somehow incorporated.
- **Knowledge-based** – If information about colour, shape or structure is passed into the system, final segmentation can benefit from it. These methods usually require a database of already segmented examples. During the segmentation, algorithms are trying to fit object from test image into an object from the database of segmented objects.

2.2.1 Edge-based detection

One of the essential parts of low-level image processing is edge detection. Even though that in real images, edges can sometimes be hard to identify. As an edge is considered location in an image, where intensity value of pixels is changing rapidly [41]. From the mathematical point of view, an edge can be displayed as a function, see figure 2.1. Ideal edge looks like a step function [42], but in real images is nearly impossible to obtain such an edge. Real case scenario looks like ramp edges. Special cases can also occur, line or roof - again roof is more probable to find.

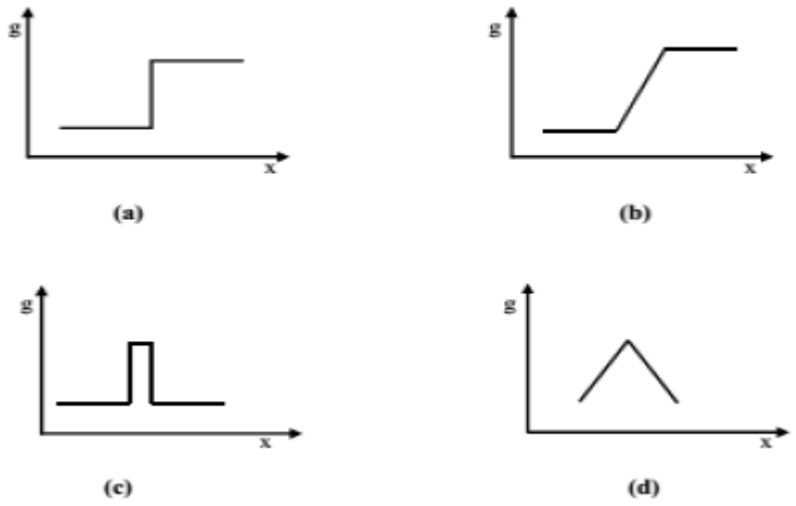

Fig. 2.1: Edge types (a) Step edge, (b) Ramp edge, (c) Line edge, (d) Roof edge [40].

Typical things that occur in images are edge changes or splits. This can be helpful in marking the object's position [41]. To be able to perform best edge detection, generally, three steps need to be followed [42]: filtering, differentiating and detection. Noise caused by sampling should be filtered. Then differentiation can

enhance image areas where an intensity of pixels is changing. Lastly, detection step can find the most significant intensity changes. Edge detection concepts can be divided into two groups. One that uses the first derivative and one that uses second derivative [42]. When the first derivative concept is used, output gradient is compared to a threshold value that determines if an edge is detected or not. The second derivative concept is based on spatial change polarity of the second derivative, if a spatial change is significant enough - edge is detected. Gradient computation can be displayed as convolution filtration. There are plenty of convolution edge detector´s core designs [44]. Gradient image is obtained due to convolution filter(detector core) with the image itself. When the second derivative concept is used, detection is focused on zero value – value when the function has null on the y-axis. Laplacian [42] is one of the examples of the operators [43], properties are in all directions same. Therefore it is rotationally invariant.

2.2.2 Region-based detection

Region growing method is the simplest concept that can create segmented area [27]. Neighbour pixels, which have similar amplitude are joined together. The actual rules that lead to the joining process can be more complicated [45]. When segmentation is applied, elementary pixel areas are created. As a next step, weak edges are removed. So-called first joining occurred accordingly defined rules. Available methods differ in rules for initial elementary pixel areas creation or in joining rules. As an example rule can "strength of the edge" be used. If edge between two areas is weak, areas are merged. For edge testing, "super grid" is used. The principle that also carries information about the difference between neighbour pixels ($s_{i,j}$). The strength of edge is computed by [27]:

$$v_{i,j} = \begin{cases} 1, & s_{i,j} < T_1 \\ 0, & otherwise \end{cases} \tag{2.1}$$

where T_1 is a threshold that determines if a boundary is strong or weak and $v_{i,j}$ is a strength of the edge. According to another rule, the circumference of individual regions (l_i, l_j) and their shared edge (W- the amount of weak common edges) can be tested [27]:

$$\frac{W}{min(l_i, l_j)} < T_2 \tag{2.2}$$

The last criterion is computed by [27]:

$$\frac{W}{l} < T_3 \tag{2.3}$$

where l is length of shared edge.

2.2.3 Statistic-based detection

Due to working principle, Kohonen maps are considered as a statistical-based method [47]. Kohonen maps are also a member of neuron network group - type SOM (Self-organizing Map). They use competitive learning strategy. The principle itself is based on that output neurons – those in the last layer, compete against each other, in whose going to end up active. In one moment there is only one active neuron [48]. As an important attribute of this network is considered aggregation. Inputs of this network are sorted into groups by the winning neuron. Every neuron in the output layer is connected to neurons from input layer [49]. Each of those connections has their weight factor. Winning neuron index corresponds to the position of the segment in an image. As an input, it can be taken amplitude of the pixel or other characteristic sign extracted from an image.

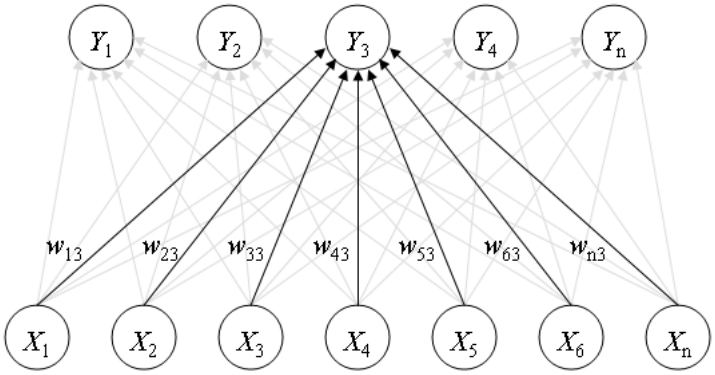

Fig. 2.2: Kohonen map topology [47].

An actual contest among neurons is done by computation of weighted distance vector between each neuron and input neuron [48]. A neuron that has the lowest output is considered as a winner, and his index (c) is noted. A distance between both vectors can be computed as a Euclidean distance or as a simple subtraction [50].

$$c = argminD(x - w_i) \tag{2.4}$$

Learning process

Training dataset needs to be imported into the network. When training pattern is introduced, competition can start [47]. When a winning neuron is discovered, weight factors are adjusted – winning neuron weight and neighbour neurons weights. The

point here is to enhance neurons with new training pattern. At the beginning of training process is the weight value equal to one. Neurons are more willing to change, at the end the value is shifted towards zero, which means they are not likely to change. This is also known as unsupervised learning [49]. One of the problems of clustering algorithms is many regions (clusters) that equal to output neurons [47]. The success rate is dependent on the complexity of input and number of output clusters. More neurons mean better differentiation, but if the input has fewer objects/ groups than the network was designed for, then one object can be segmented into multiple neurons, and solution for correct joining needs to be found. And the problem is same the other way around – a few neurons means a lot of objects that need to be clustered. There is a way how actually to handle this [49] – adaptively increase/decrease the number of output neurons.

2.2.4 Knowledge detection

In this category, methods use already obtained knowledge about objects found in an image. Knowledge can be represented either by templates or models of objects [27]. These templates are then compared with unknown image and search for matches can begin. The main disadvantage of these methods is significant variability of unidentified objects. Exact segmentation of complicated structures is then very difficult. On the other hand, if a structure of interest objects is very similar – methods are then very successful. As an example, Active appearance models method can be introduced [50]. The method uses PCA analysis for model creation. The model itself is composed of manually segmented data. As an example from medicine can serve specific bone in the arm. On test fleet, bone is manually segmented and labelled. This batch of labelled images is then via PCA transferred into the model for actual segmentation on unknown data.

2.2.5 Hybrid-based detection

Majority of segmentation methods is based on knowledge and experiences how the things should proceed. The exact opposite approach is used in segmentation with neural networks (NN). Training of the NN is based on learning by examples principle [27]. There are two approaches how to train the network. The first approach is looking for characteristic signs in the input data and then classify them without any extra interpretation. This is known as an unsupervised learning [51]. The second approach is called supervised learning [52]. In this approach, data has to be segmented beforehand. With this training dataset is network trained and prepared for further segmentation of unknown data samples.

Convolution neural networks

The big bang in neural network area happened is dated to the year 2012. A revolution was made in ImageNet competition [53]. Several reasons lead to massive usage of the neural networks, mainly due to efficient usage of graphics processing units, a discovery of new rectified linear units, new dropout regularisation and effective data augmentation [53]. CNNs have become once again popular learning machine in a various application filed. The main power of CNNs is held in its deep architecture, which allows extracting discriminant features at multiple levels of abstraction [54].

3 CONVOLUTION NEURAL NETWORKS

To gather a complete list of network architectures is nearly impossible because new network designs are invented almost on a daily basis. For the illustration purposes see figure 3.2 where are some of the known architectures. As the first architecture is basic perceptron but let's focus on Convolution neural networks (DCN/CNN). These architectures are partially different than others. Image processing tasks use them frequently, but lately, they are in favour of audio inputs tasks [55]. A typical function of these networks is classification, they obtain images as input and spits out image class. The starting layer can be imagined as a scanner that does not parse all the data at once. For an image of dimensions 512×512 px, there would be 262144 nodes required in the first layer. Therefore there is this scanning layer that takes 16×16 px starting from the top left corner and slowly moves and scans all image and continuously feeds the network with data. The data are not going through usual but convolution layers. These layers are not necessarily fully connected. There is also a tendency to shrink as the layer become deeper. Powers of two are commonly used in dividing the layer to obtain clean and complete division. Besides the convolution layers, these architectures also have pooling layers. These specific layers tend to filter out details, commonly used is max pooling, where the maximal pixel value continues further computation.

3.1 Convolution

Mathematical operation of convolution represents a signal $s(t)$ passing through a system with impulse response $h(t)$. The resulting output is the convolution of $s(t)$ with $h(t)$. Operation of convolution can be also visualised as a integral of the product of those two functions, equation 3.1, where one is reversed [56].

$$(s * h)(t) = \int_0^\infty s(\tau)h(t-\tau)d\tau \tag{3.1}$$

Convolution is not limited to one dimension. Extremely useful is 2D convolution in image processing operations. Image is 2D signal $s(x, y)$ and can be the input to a 2D filter $h(x, y)$, (x, y) is the position of a pixel [56]. Then the operation of convolution is:

$$(s * h)(x, y) = \int_x \int_y s(\alpha, \beta)h(x-\alpha, y-\beta)\partial\beta\partial\alpha \tag{3.2}$$

3.1.1 Convolution layers

A simple convolution network is made from a sequence of layers (pooling, fully-connected, convolution, ...). Every layer transforms the volume of activations to another through a differentiable function. The most common form of convolution network architectures stack a few Convolution-ReLu layers, followed by Pooling layers [62, 66]. This structure repeats several times until the image achieves desired spatial size. Possible last Fully-connected layer holds the output - e.g. class scores. Another possibility is to use Deconvolution layers to attain original image size [73]. With such pattern, every pixel will belong to some of the available classes.

3.1.2 Convolution kernels

Operation of convolution allows to do many things, like calculate derivatives, apply distortions, detect edges/features, etc. All of this is done via convolution kernel. The kernel is a small matrix that has numbers in each cell and has an anchor point. Purpose of the anchor point is to determine the position of the kernel in respect to the image, in which the kernel slides and do its operation.

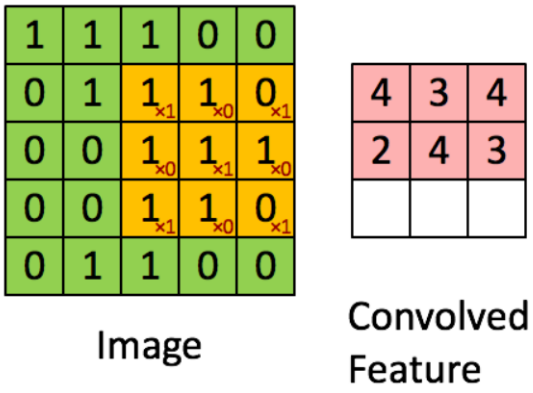

Fig. 3.1: Example of kernel slide [72].

A standard rule about the kernel size is that has the same size in (x,y) axis and that the size of the kernel belongs to $2\mathbb{Z}+1$ [58, 62, 66]. The distance the kernel moves each time is called stride, depending on the stride value is the input matrix cut on dimensions more or less aggressively. Purpose of this kernel application is filtering that search for patterns in that section of an image. During the image training, weights are changing. Later when the image needs to be evaluated, these weights return high values if they think that they see the pattern that has seen before. Combination of filters enables the network to predict the content of an image[64].

3.2 Architecture

Neural networks can be visualised as collections of neurons in an acyclic graph. This type of graph means that outputs from some neurons can serve as an input to other neurons. From the word, acyclic is evident that neurons connection does not create an infinite loop. The most common layer type of conventional neural network is the fully-connected layer where neurons from neighbour layers are fully connected, but inside the same layer are not connected with each other. N-layer neural network means that the input layer is not counting. Single-layer neural networks indicate that there are no hidden layers. The uniqueness of the output layer is not having the activation function. Most commonly this layer represents the class scores which are real-valued numbers. The typical metrics that can measure the size of the neural networks are the number of neurons and number of parameters. Nowadays in modern CNNs, one hundred million parameters can be found, and the networks are usually composed of 10 - 20 layers [57]. The layer design structure made very easy to use matrix-vector operations. All connection strengths for a layer can be stored in a single matrix. The full feed-forward computation is then simple several matrix multiplications where activation function is incorporated [59].

3.2.1 AlexNet

One of the most influential publications in the field of deep learning is the paper, titled "ImageNet Classification with Deep Convolution Networks" [62]. Authors were able to create large, deep convolution neural network to attend 2012 ILSVRC (ImageNet Large-Scale Visual Recognition Challenge). In this competition, teams compete across the world who has the best computer vision model for the desired category: classification, localisation, detection, object recognition, etc. AlexNet was the first CNN that achieve top-5 test error rate of 15.4 % (Top-5 error is the rate at which, given an image, the model does not output the correct label with its top-5 predictions [62]), the next best entry achieved an error of 26.2 %. This improvement shocked the computer vision community.

AlexNet´s layout, figure 3.3, is relatively simple - compare to modern architectures. Network design consists of five convolution layers, max-pooling, dropout and three fully connected layers. The initial purpose was to be able to classify one thousand possible categories. One interesting remark about the layout - there are two different streams, due to computationally expensive training, that had to be split into two GPUs.

The specific contributions from this paper [62] are the introduction of new and unusual features which improved performance, reduce training time and several

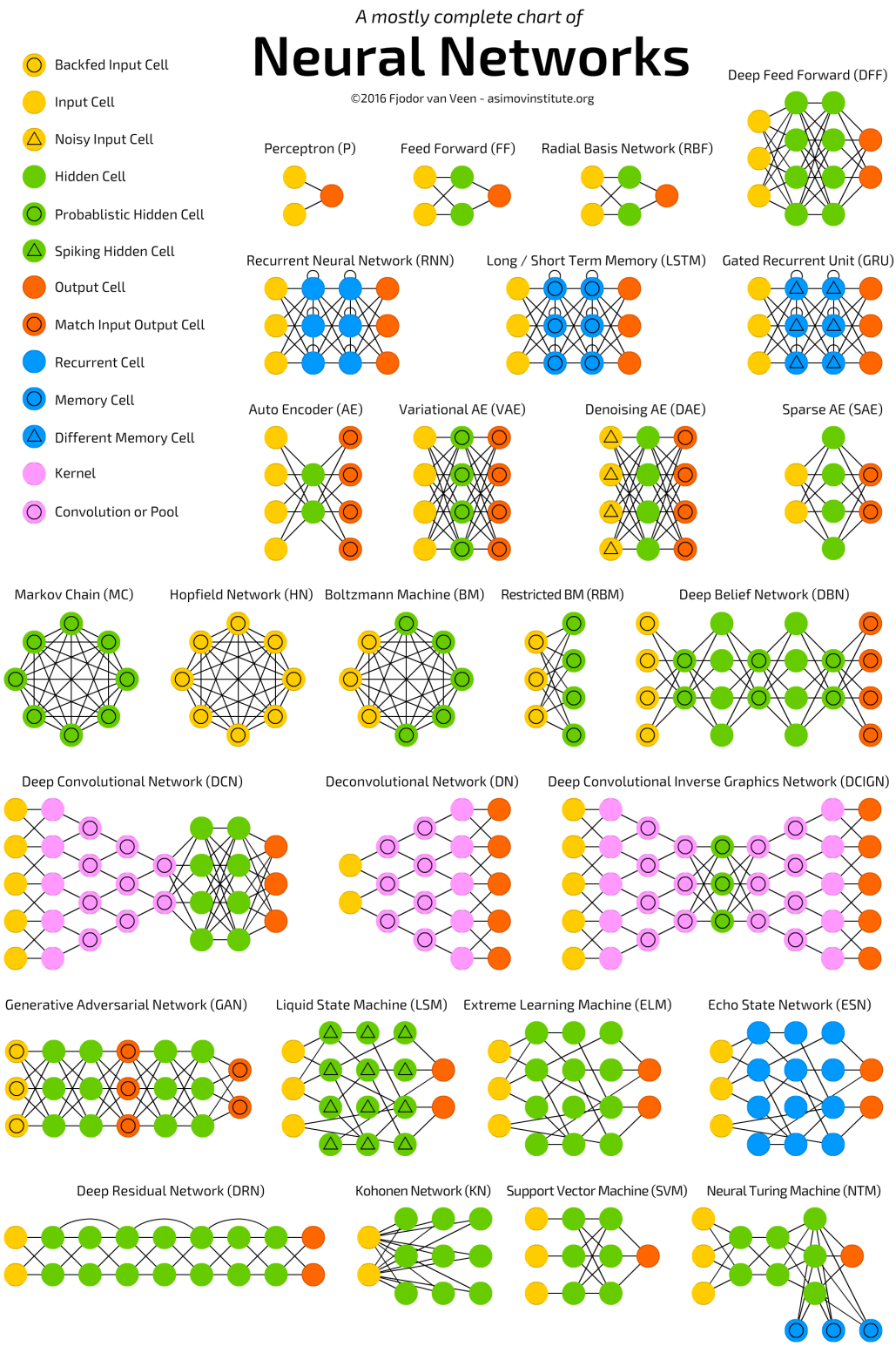

Fig. 3.2: Chart of different network topologies. [61]

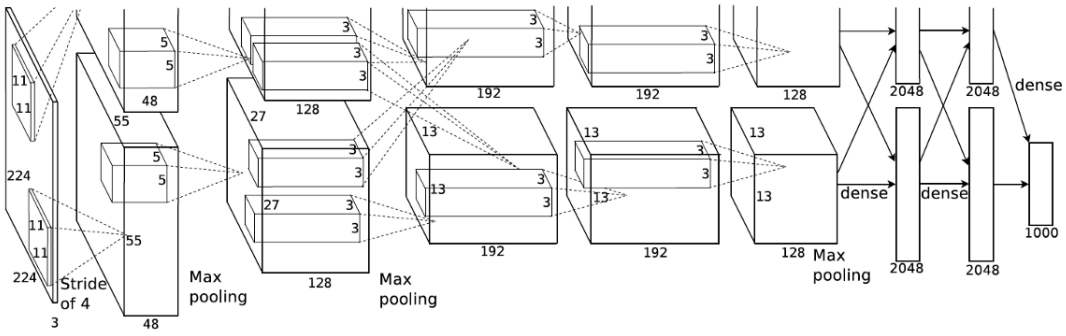

Fig. 3.3: AlexNet layout [62].

techniques how to prevent overfitting.

$$f(x) = max(0, x) \tag{3.3}$$

The main benefits can be summarized in five key points:

- **ReLu** - Authors [62] decided to use ReLu activation function. This activation function act as a zero thresholder, even though that the output values are not zero centred there are some arguments for its usage. It can significantly accelerate the convergence of gradient descent compared to Sigmoid or tanh functions. This property is argued due to linear non-saturating form. Another enhancement is in an implementation form. No mathematically complicated operations are needed, simple thresholding at zero is applied. Unfortunately, there is also a disadvantage. Due to large gradient flow, the weights can be updated in a way that they have never activated again. That will result in zero gradients from that point. This is known as "dead" neuron (neuron that never activates across the entire training dataset). Controlling the number of dead neurons with proper setting of learning rate is, therefore, a must [59].
- **GPUs** - Precisely two GPUs were tuned especially for computation purposes. In the paper, [62] is described a specific pattern of connectivity to correctly optimise the amount of communication between the two GPUs to allow them to communicate only on particular layers.
- **Overlapping Pooling** - Traditional pooling layers downsample the output of neighbouring groups of the neurons without overlapping [64]. Here in [62] came up with an overlapping setup that resulted in the reduction in top-1 and top-5 error rates by 0.4 % and 0.3 %
- **Data Augmentation** - Two types of data augmentation were used. First consists of generating image translations and horizontal reflections, second

consists of altering the intensities of the RGB channels.

- **Dropout** - In [63] new technique to improve results of the neural network was introduced - dropout. Neurons from each hidden layer become zero valued with 0.5 probability. These "dropped" neurons do not continue in contributing to the forward pass and do not participate in backpropagation. This technique reduces complex relations amongst the neurons in near area. As a result, the network is forced to learn more robust features.

3.2.2 VGG Net

Compare to AlexNet, VGG NET is deeper and simple. This CNN has nineteen layers and strictly uses 3×3 filters and 2×2 max-pooling. Authors [66] decide to use benefits of smaller sized filters (decrease number of parameters) to create from two 3×3 filters effective receptive field of 5×5. Usage of two convolution layers allows using two ReLU layers instead of one that leads to faster learning. The primary outcome of this topology is to keep the design deep and simple.

3.2.3 GoogLeNet

Simple is the only word that does not characterise this convolution neural topology, see figure 3.4. Authors [67] claim that the topology consists of twenty-two layers with parameters. If you count the independent building blocks, the number rises to one hundred. The exact number of layers is dependent on how the machine learning infrastructure counts the layers.

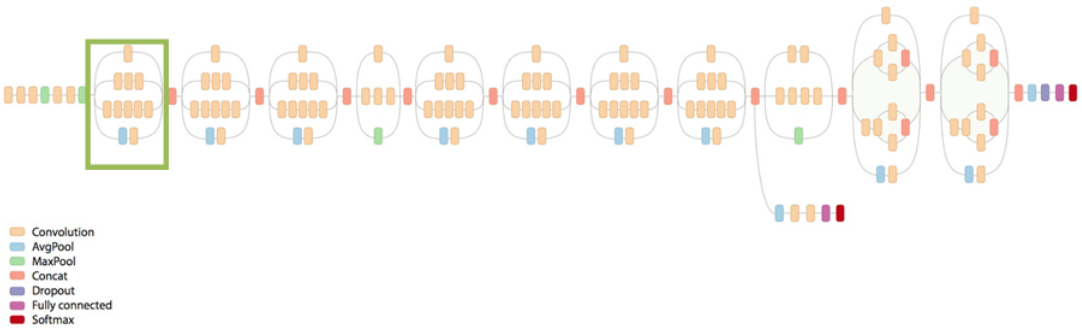

Fig. 3.4: Topology of GoogLeNet [68].

The primary outcome from this topology is the introduction of an "Inception module", figure 3.5 – neural networks does not need to be designed for sequential computation, but there is a possibility to use pieces with parallel computation. On traditional layer designs [62, 66], the user has to decide whether is better to

use pooling or convolution. The inception module allows performing all of these operations in parallel. But how exactly is this architecture helpful? The module has medium and large sized filter convolution with a pooling operation. The network in network convolution extracts information about the very fine grain details in volume, while the average convolution filter covers large spatial sizes of the input. This leads to better extraction of information. For reduction of the spatial size and overfitting is used pooling operation. ReLu layers deal with the nonlinearities. Even though that this looks like a lot of computations to perform, from the computational point of view is this still considerable.

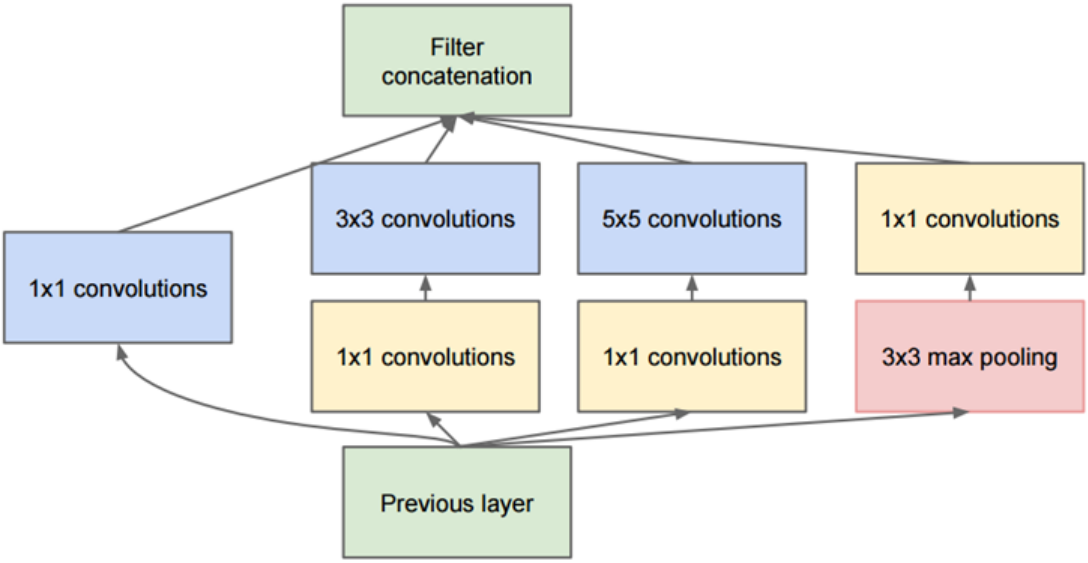

Fig. 3.5: Inception module of GoogLeNet [67].

3.2.4 Microsoft ResNet

In 2015 incredible error rate of 3.6 % was overcome in ILSVRC challenge. Microsoft Research Asia came up with deep CNN architecture. ResNet consists of 152 layers. This was for the first time when general human error rate (5 %) [69] was beaten. ResNet introduce another new concept - Residual block, see figure 3.6. The idea is that the input pass through convolution-ReLu-convolution series. This results in function $F(x)$. Then input x is added to this function $H(x) = F(x) + x$. The module is computing small delta that serves for slightly altered input representation[1]. According to [70] is "...easier to optimise the residual mapping than to optimise the original, unreferenced mapping".

[1]Traditional CNNs goes from x to F(x) and results in completely new representation.

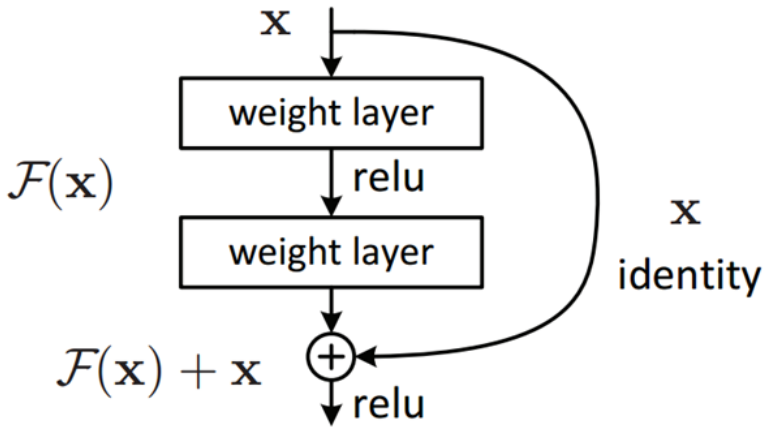

Fig. 3.6: Residual learning: a building block [70].

3.2.5 U-NET

A significant amount of neural networks designs is strictly focusing on general classification problem (ImageNet, GoogLeNet, etc.), but in the more areas localisation of the particular label is desired. Each pixel should obtain a label. The problem that is frequently appearing is not sufficient amount of labelled training data in specific areas. In [71] developed a sliding principle. Authors created large training dataset with use of image patch – differently sized masks were applied to an image in order to obtain only a section of the image.

With sliding principle, a large number of training samples were created from one image. In [73] got inspired by this sliding principle, but they saw drawbacks of this technique, the network was run separately for each patch (mini-batch size = 1), and colossal data redundancy was present, patches were overlapping in large areas. Therefore improved architecture was presented, see figure 4.3. They changed the typical design, see DCN in figure 4.2, where pooling layers successively contracted a network into a layout wherein certain point the contracting is stopped, and they upsampled the data to obtain an increase in outputs resolution. Another improvement was in combining upsampling layers with data obtained from the pooling layers, that led to better localisation. The network itself does not have any fully-connected layers and uses only the valid part of the convolution. Authors claim that they applied augmentation to data to enhance the training dataset. Their architecture can be used in various biomedical tasks. It can segment and classify objects from an image and was successfully tested on images from light and electron microscope [73].

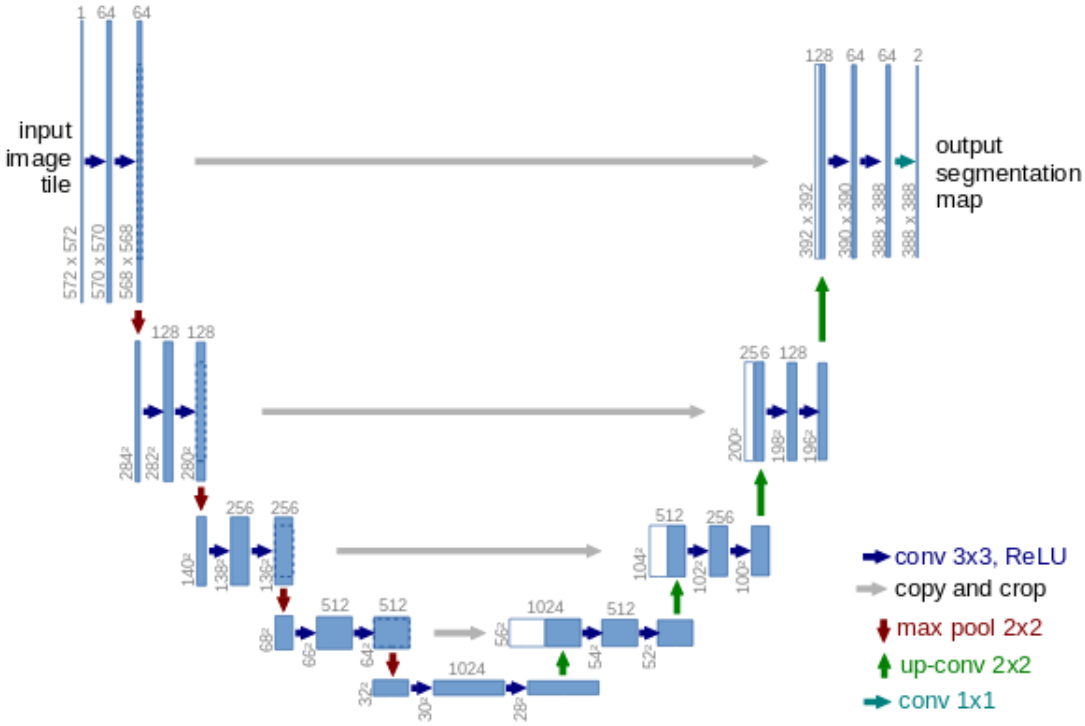

Fig. 3.7: U-Net architecture. [73]

4 TESTING CONCEPT

In recent years, massive improvement in parallelisation tasks was achieved. Large companies like Intel, AMD or NVIDIA presented a trend in the field of multicore development. This multicore development is most significantly evident in the area of graphics card development. Last years summer boom in a field of mining cryptocurrencies [89, 90] with a graphics card is one piece of the evidence, GPUs became affordable and able to compute a large number of parallel tasks. This parallelisation is particularly useful when it comes to neural networks. Because of the simple operation of multiplication and addition that can be computed on many parallel cores, more network topologies were introduced. Possibility to combine several cards in parallel [88] also improved the growth in the neural network area.

4.1 Deep Learning frameworks

From the Cambridge dictionary definition, a framework is "a supporting structure around which something can be built" [91]. Then deep learning framework can be imagined as a software supporting structure to built neural networks. Among the developers, several frameworks are in favour due to different language capabilities and ways how to write in them. There is a possibility to rewrite network implementations from framework to framework. This is typical practice when some of the frameworks have particular layer implemented in a faster or more efficient way. Because of my knowledge of MATLAB and Python, the decision had to be made between Caffe and TensorFlow frameworks. After several consultations and discovery that TensorFlow is popular amongst the deep learning scientists (Google‘s Machine Intelligence research organisation developed TF. Caffe was developed by Berkeley Vision and Learning Center. Caffee community users can also contribute to it.), the decision use TensorFlow as a low-level framework was done. Mind the low-level, because even people from TF does not claim that this is a deep-learning framework, instead "TensorFlow is an open source software library for numerical computation using data flow graphs." [88] It is written in such a low-level code, that even people from TF have to introduce an additional layer framework on top of that. To be able to use the benefits of TF, this paper uses Keras as a communication tool between user and TF. Keras is a high-level library that works on top of Theano or Tensorflow. It is configurable and also works with the Python. One of the main advantages is that Keras has well-written documentation and reinforces minimalism. There is no need to write large segments of code. Therefore the code and modifications in it can be done very quickly [85, 88].

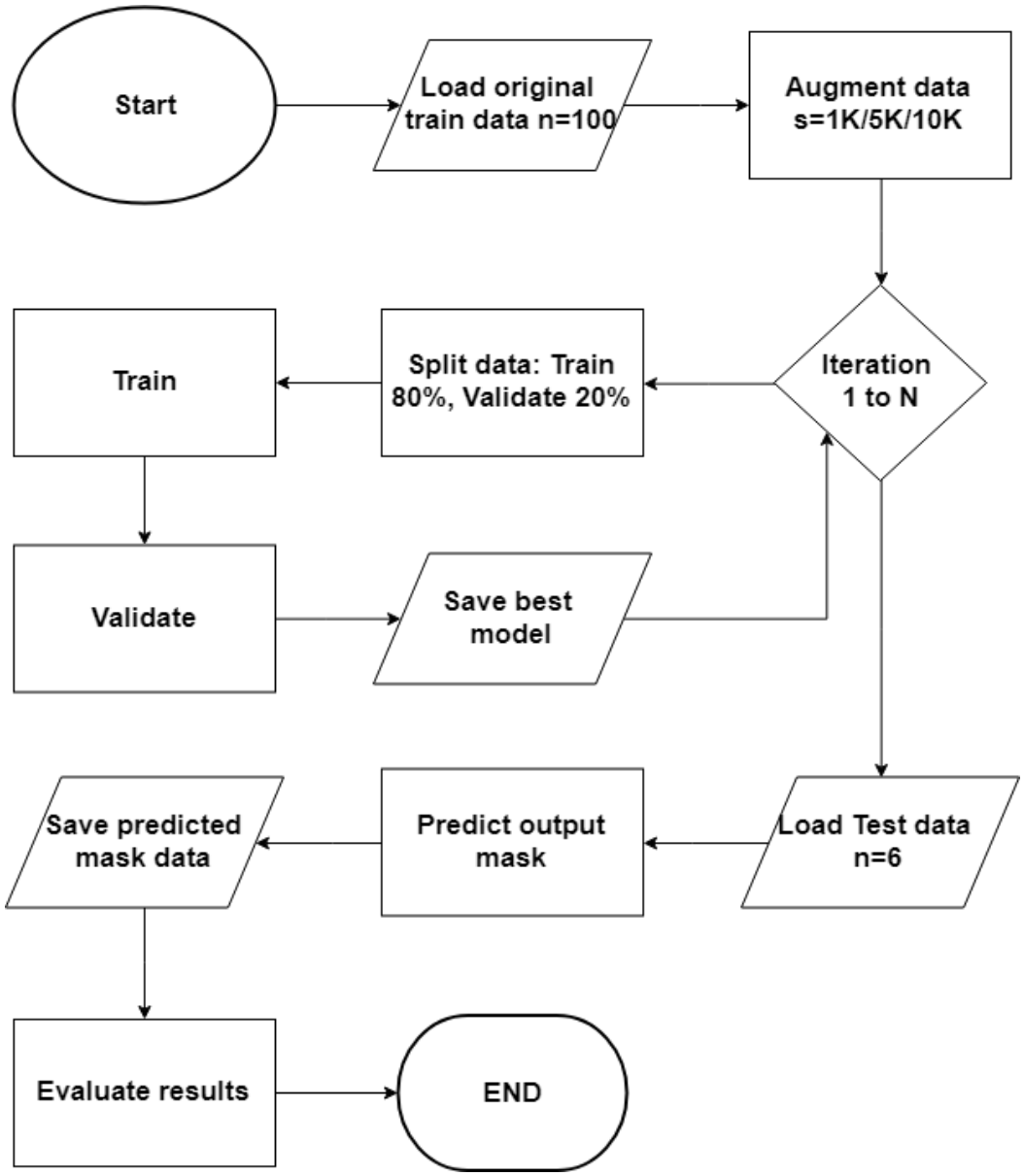

Fig. 4.1: Workflow diagram.

4.2 Input data

Thermo Fisher Scientific (formerly FEI Czech Republic s.r.o.) provided image data for this thesis. This provided image dataset was chosen based on the internal company requirements. Original image data were acquired on Thermo scientific TEM systems, 200 kV, imaging modes: bright/dark field, STEM and magnification range 2k-5000kx. Original images had different resolutions, 4K and 2K images were randomly cropped to fit dimensions 512×512 px. Images that did not require dimension changes kept their original size. The initial training dataset consists of

100 images and the appropriate number of binary masks. Images are made up of various elements and samples (Au, C, Zn, Bronz, KCl, Spinel, Az, Grid). Due to such variability in samples, see figure 4.2 (Starting from top left corner: Spinel, Zinc, Crossgrating, Gold, Graphitized Carbon, Zinc, Spinel, Crossgrating.), was decided that for segmentation purposes will be used a convolutional neural network.

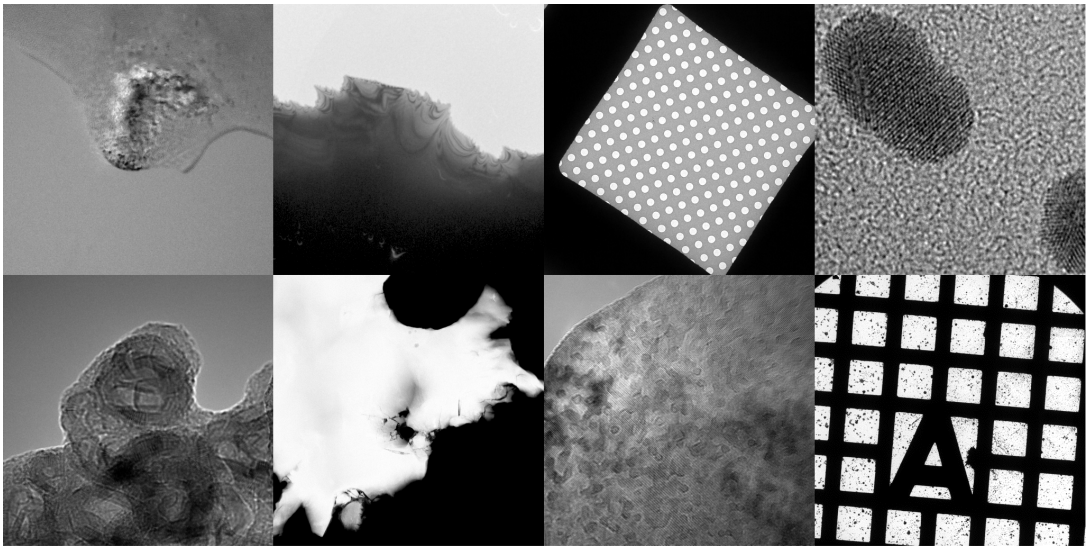

Fig. 4.2: Test data variability example.

4.3 Augmentation

The training dataset consists only of 100 images. That is surely not enough for reliable training purposes, some data enlargement is required. For this purposes serves operation of data augmentation. We decided to use Augmentor tool [75] for well-documented GitHub repository and availability on official pip install base. This tool consists of various functions. Some of them are usable for our purposes, and some of them don't. In this thesis is working with the three available functions: Zoom, Rotate and Distort. With the use of these three functions, we are trying to simulate the real case scenarios. Delivered samples could be easily in different magnifications. We would like to have topology resistant to this variety of magnifications. Then during zooming process, the image is rotated in the system - again we want resistance to this rotation phenomenon and lastly distortion. Some sort of astigmatism is always present due to manufacturing flaws and electric fluctuations in the components. Therefore small random distortion is applied. Augmentor tool is beneficial for us because it can simultaneously augment the corresponding layer mask, and then we do not have to segment and label these areas manually. This help

leads to tremendous time-saving. Variously sized datasets will be tested to obtain the best model for segmentation purposes.

4.3.1 Rotate

The rotation of image could be considered as the most straightforward way how to augment original dataset. Used function from Augmentor tool [75] turned every point from the image by an angle about a centre of rotation. In this case, the centre of rotation was the centre of the image. When the rotation angle was not by modulo 90, the largest possible crop from the centre of the newly rotated image was taken and stretched to obtain original size. Bicubic interpolation served during the stretching process. The input parameters of the rotate function were: probability = 0.7, max left rotate = 25, max right rotate = 25.

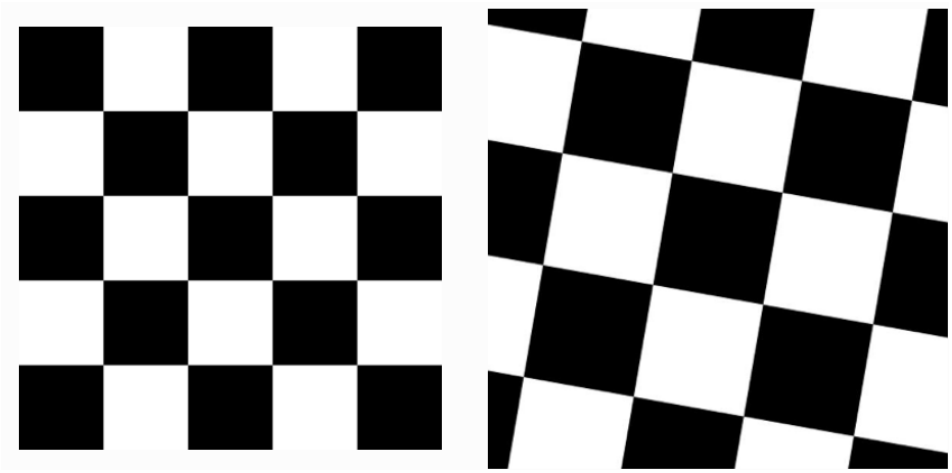

Fig. 4.3: Graphical representation of rotation operation on model image [74].

4.3.2 Zoom

During zooming operation is image zoomed (enlarged) in the defined interval, while the image size is maintained (cropped to the original size). The input parameters of the zoom function were: probability = 0.3, min factor = 1.1, max factor = 1.6.

4.3.3 Distort

Last used operation from Augmentor tool [75] was elastic distortion. This operation creates distortion and maintains the image‘s aspect ratio, see figure 4.5. These variations are useful because they can simulate close enough real-world scenarios.

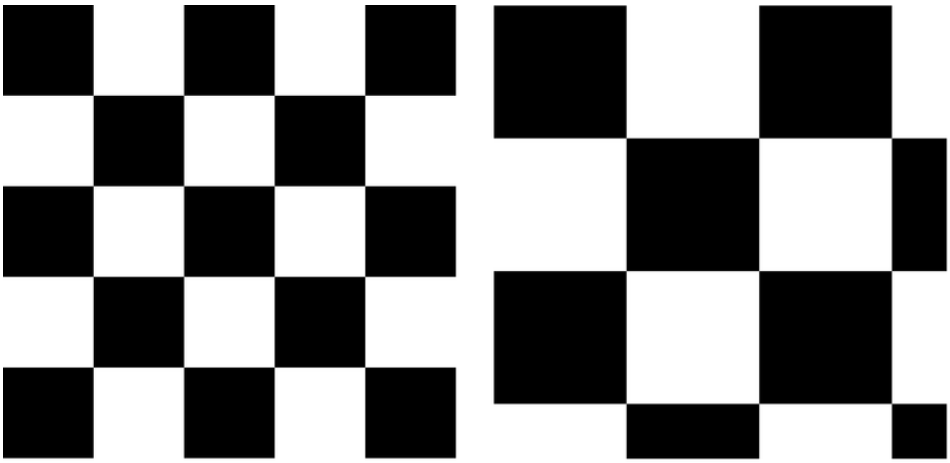

Fig. 4.4: Graphical representation of random zoom operation on model image [74].

Realistically looking new samples can be quickly generated. The input parameters of the distort function were: probability = 0.9, grid width = 8, grid height = 8, magnitude = 2. Grid width represents number of rectangles in the grid´s horizontal/vertical axis [75], magnitude stands for magnitude of the distortions.

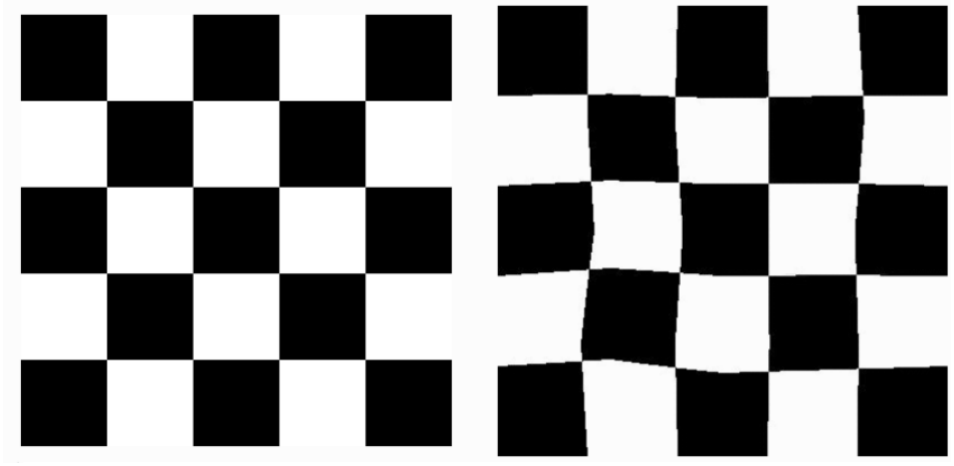

Fig. 4.5: Graphical representation of random distortion operation on model image [74].

4.4 CNN selection

Based on the general idea in CNN area - It is better to tune already finished neural network topology than try to invent a new one for the specific task [64]. With this approach, we decided to define parameters that we require from the topology.

1. It has to be open source topology with available and documented architecture.
2. Available topology in one of the listed neural network frameworks - TensorFlow, Keras, Caffee.
3. Designed for segmentation task.
4. Designed for work on grayscale images.
5. Successfully tested neural network topology.
6. Computationally undemanding. Keep it deep, keep it simple.

When these six requirements were set up, search for the fitting topology could begin. In [73] we discovered the topology for our testing purposes. The topology is well documented and presented in the corresponding article. Then the neural topology was written initially for Matlab in Caffee framework, after a bit of searching we also found Keras implementation for Python. The main reason for this option was due to significant python and Keras community if some problem would occur. Another reason why to use this network is connected to the 6th requirement. My test environment consists of one Nvidia Titan GTX graphics card, and I wanted to benefit from CUDA neural enhancement. Corresponding thorough documentation is written for TensorFlow/Keras enviroment[1], and possible problems again would be solved with higher probability than on any other edge framework. The third requirement is also fulfilled, the neural network was tuned for segmentation task on neuronal structures in electron microscope [73] and more - It also won the ISBI cell tracking challenge in 2015 by a large margin. Authors tried to design the topology, which for training would not require massive dataset [62] and one high-performance graphics card should handle this. This last requirement is especially interesting nowadays when on the market are graphics cards hard to obtain or the prices are set to high. From this point on, everything suggests that the U-NET should be possible to achieve excellent results on available TEM samples.

4.5 Hardware

Implementation: All experiments were conducted on a workstation equipped with one NVIDIA GeForce graphics card: NVIDIA GeForce GTX 1080, and two 3.20 GHz Intel Xeon X5482 processors with a 64-bit Windows 7 and 32 GB RAM.

[1]Reason why Keras and TensorFlow are bound together is that of the Keras is a framework that uses on lower levels the TensorFlow. Amongst the scientists, the Keras is more favourable due to more readable and user-friendly syntax.

4.6 Learning

One of the basic building blocks of ANN is a classification task, assigning a single label to an image from a list of known categories. The recent and more powerful approach how to classify images that can be extended to entire ANNs have two major components [65]:

- Score function that maps the raw data to class scores.
- Loss function that quantifies the similarity between the predicted score and the ground truth labels.

Using these two parameters classification can be transformed into the optimisation problem in which minimising the loss function to the parameters of the score function is desired.

4.6.1 Score and Loss

Purpose of Score function is to map the raw data to class scores. This function can be formulated as:

$$f : R^D \mapsto R^K \tag{4.1}$$

where R^D is training dataset with a dimensionality D and R^K training dataset mapped to class scores K. Implemented U-Net topology uses TensorFlow accuracy function [88] as a score value. This function calculates how often predictions match labels.

A loss function is a tool that serves as a measurement of unhappiness of the score function. The loss function will be high when the score will not correspond with the ground truth label and will be low in the opposite case. With the help of the Loss and Score functions, we can create an optimisation problem, in which the minimisation of the loss for score function parameters. For this thesis the KerasTensorFlow binary cross entropy function [85] will be used, just because of only two object labels need to be distinguished.

4.6.2 Optimizer

The third and last key component of setting the correct weight to minimise the loss function is an optimisation. As a simple idea, random search approach can be used. Several random weights can be generated, and one combination can work better than others, but this approach can also lead to the dead end. Therefore optimisation algorithms are in use. U-Net topology [73] has as a default optimiser the Adam optimisation algorithm. Choosing the right algorithm for deep learning

model can have a considerable effect how good or bad the results will be and when they will be - minutes, hours, days [86].

Adam updates network weights iteratively based on the training data. According to authors [86] is:

- computationally efficient;
- require little memory;
- is suited for large problems in term of data/parameters;
- appropriete for noisy gradients;
- hyperparameters does require very little tuning.

Adam combines advantages of two other algorithms: Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp). From the AdaGrad took advantage of the per-parameter learning that helps with the improvement of the performance with sparse gradients [87]. From the RMSProp maintained the per-parameter learning rate that is adapted on the average of the recent magnitudes of the gradients (how quickly is changing), which resulted in better handling the noisy data [86]. The algorithm calculates an exponential moving average of the gradient, and also the squared gradient - β_1, β_2 parameters control decay rates of these moving averages.

4.6.3 Activation

The goal of activation function is to take a single number as an input and perform a specific mathematical operation. Based on that resulting number the neuron will be activated or not. For deeper layers was discovered that ReLu activation function provides significantly better results than other activation functions [62]. Manipulating with deeper layer activation functions is generally due to complicated relationships and dependencies not recommended. To be able to set the probability to the correct label, activation function for this purpose needs to be used. This task is handled via Sigmoid activation function.

ReLU

$$f(x) = max(0, x) \tag{4.2}$$

This activation function act as a zero thresholder, even though that the output values are not zero centred there are some arguments for its usage. It can significantly accelerate the convergence of gradient descent compared to previous two functions. This property is argued due to linear non-saturating form. Another enhancement is in an implementation form. No mathematically complicated operations are needed, simple thresholding at zero is applied. Unfortunately, there is

also a disadvantage. Due to significant gradient flow, the weights update in a way that they would not be activated again. That will result in zero gradients from that point. This is known as dead neuron (neuron that never activates across the entire training dataset). Controlling the number of dead neurons with proper setting of learning rate is, therefore, a must [57, 59].

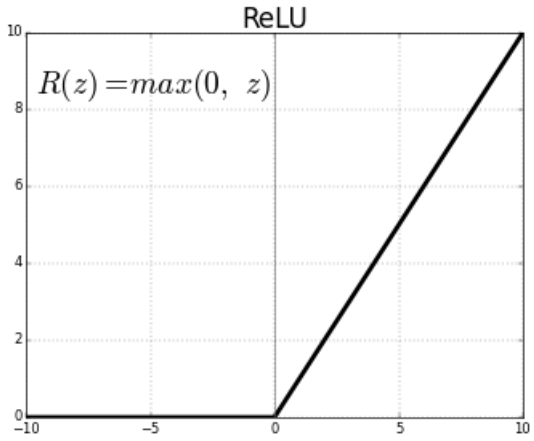

Fig. 4.6: Relu activation function [60].

Sigmoid

The mathematical form of the sigmoid function is:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{4.3}$$

It can transpose input number into the range of (0,1). There can be seen that substantial negative values would appear as zero and large positive numbers as one. The sigmoid function used to be very popular due to a relatively straightforward interpretation of the activation rate: not activate (0), active (1). The disadvantage of this activation function is in minimal gradient in tail values. The signal will almost not flow, and the learning process can be stopped. This is related to a situation during the initial set of weights, if the weights are set too high, most neurons saturate very quickly, and the network will barely learn. Another problematic part of this function is that the outputs are not zero centred. This centring will become more important in later layers of processing where the layers are receiving either positive or zero value input. It affects updating the gradient during backpropagation. Weights would become more positive or more negative, due to zero value that acts as no update situation. In large batches, a zig-zag effect can be seen, but the resulting

differences in variable signs can somehow compensate this drawback. The main inconvenience persists from the saturation effect [57, 59].

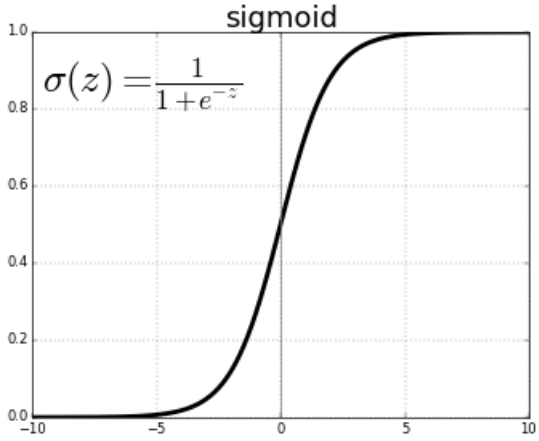

Fig. 4.7: Sigmoid activation function [60].

4.6.4 Number of iterations

Authors [73] does not mention how many iterations, nor how many images needs to be used for the training. My computational resources allowed me approx. 12-hour training with the use of 10 000 images. During this time, I was able to run approx. 20 iterations. I believe that 20 might be too much and after a specific number, there will be no improvement, nor the neural network might overfit the data.

4.7 Evaluation

For evaluation of the segmentation quality of the neural network design, this thesis uses these four metrics: Dice similarity coefficient (DSC), Rand index (RI), ROC and Precision recall-curve. These parameters are not evaluating every single neuron or layer. Instead, they assess the final trained model.

4.7.1 Dice similarity coefficient

Original name of this Dice similarity coefficient (DSC) is Sørensen-Dice index, but currently, many other names are used for this evaluation metric (similarity coefficient, F1 score, ...). Original intention [80] was to apply this coefficient to binary data, given two sets X, Y is defined as:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \tag{4.4}$$

where $|X|$ and $|Y|$ are the cardinalities of the two sets. In the particular case when the coefficient is applied on boolean data, using the definitions of true positive (TP), false positive (FP) and false negative (FN), it can be rewritten as:

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{4.5}$$

Because the DSC does not satisfy the triangle inequality, it is considered as the semi-metric version of the Jaccard index. Benefits of comparison the Sørensen distance to Euclidean distance is in retainment sensitivity in more heterogeneous data sets, and it gives fewer weights to outliers [81]. It is also commonly used in image segmentation [82, 83, 84], when algorithm output masks are compared to the reference.

4.7.2 Rand index

The Rand Index (RI) serves for measuring the similarity between two data clusters in statistics and data clustering. From a mathematical standpoint is RI related to the accuracy, but in segmentation, evaluation is applicable even when class labels are not used.

Given an n object set $S = O_1, ..., O_n$ and two different partitions of S - $U - u_1, ..., u_R$ and $V - v_1, ..., v_C$ then RI measures correspondence between those two partitions U, V on how object pairs are classified in contingency table [79]. Based on this two partitions principle, four different pair types could be found:

- objects in the pair are placed in the same class as in U and V;
- objects in the pair are placed in different classes in U and in different in V;
- objects in the pair are placed in different classes in U and the same class in V;
- objects in the pair are placed in the same class in U and different classes in V.

Unfortunately, RI has known problems such as the fact that the expected value of the RI of the two random parts does not result in a constant value. Another problem is when a number of clusters rise, then RI approaches its unity upper limit. Therefore researchers created alternative measures. In this thesis, the decision was made to use Adjusted Rand Index (ARI) [79] for its improvement of RI. Originally RI can yield a value between 0 and 1. The ARI has interval -1 to 1. Values close to negative one means that labels do not correspond; values close to zero are the result of a random labelling and values close to one indicates that labels are identical.

4.7.3 ROC curves

The receiver operating characteristics curve is a plot consist of false positive rate against true positive rate when some other parameter is varied. This curve was

taken up in the 1970s in a medical profession [76], where they were used for graphical representation of sensitivity against specificity in medical trials. The interpretation of ROC curves follows [77]:

- The accuracy of the test can be easily determined by the direction of the curve. When the curve approaches 45° diagonal, the test is worse. On the other side, when curve approaches the top left corner, the more accurate the test.
- The area under the curve also measure the accuracy of the test.
- From the plot can be seen a trade-off between TP and FP - an increase in FP follows an increase in TP.

Fig. 4.8 shows ROC curves for three different quality test (perfect, good and imperfect). When area under the curve is 0.5 (poor test) test acts randomly and represents failed test - useless test. The opposite situation is when the area gets to unity - that can theoretically result in the perfect test[78].

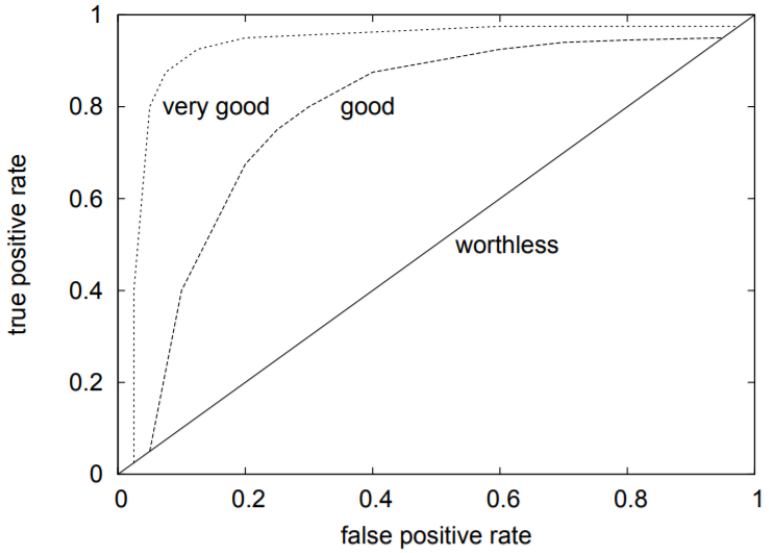

Fig. 4.8: ROC example.

4.7.4 Precision recall-curves

From TP, FP, FN, TN values, other numbers of quantities can be derived.

$$Sensitivity = Recall = \frac{TP}{TP + FN} \tag{4.6}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4.7}$$

$$Precision = \frac{TP}{TP + FP} \tag{4.8}$$

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4.9}$$

Unlike the ROC curve, precision-recall curves have precision on the y-axis, and they run from upper left to lower right corner. In vision related scientific papers these curves are increasingly replacing ROC curves [92]. The main reason is that in the field of machine learning was introduced new challenge - classification of imbalanced datasets[2]. Such imbalanced data can easily trick the ROC representation. Balanced or imbalanced data [93] does not influence Precision-Recall curve. It will provide the desired output. This leads to an idea of Precision-recall curves can be useful in assessing how particular parameter affect the performance of an algorithm [77].

4.7.5 Confusion matrices

A confusion matrix stores information from actual and predicted classifications performed by an algorithm. This is demonstrated in table 4.1, where digit recognition task was taken. The left column contains digits that were loaded into the system and rest of the matrix shows how these entry digits were classified. In the ideal system, every entry value would be classified as it should be and numbers would be only in diagonal. Here the misclassifications occurred and numbers are also off-diagonal. Quickly can also be discovered, which entry digit was classified correctly and which poorly.

Tab. 4.1: Example of confusion matrix.

actual	predicted			
	0	1	2	3
0	20	0	0	5
1	0	25	0	0
2	0	0	23	2
3	0	0	3	22

[2]Picture it as image were a region of interest covers 25 % of the full area instead of 50 %.

5 RESULTS

This part covers test results of differently sized training datasets and a different number of training iterations to confirm a theory, which with increasing number of iterations and training data the neural network model would achieve better prediction score. Dice Similarity coefficient (DSC) and Adjusted Rand Index (ARI) values will serve for presenting this improvement. Graphically these results will be displayed in the form of Receiver operating characteristics (ROC), Precision recall-curves (PRC) and Confusion matrices along with the segmentation masks next to the original ones.

All obtained data were processed in Python 3.5. All input data, codes and related documentation is provided on the attached disk.

5.1 Training size influence

The initial size and variability of the training dataset come particularly evident when the trained model is trying to process unknown data. During the training period is neural network training and validating on data that is available for it. Therefore the model can be extremely precise when is processing data that is familiar with it. It is almost impossible to train a model on every possible scenario. Therefore appropriate variability in training dataset comes handy. Larger the dataset – larger the variability and better output model. This idea was tested with the increasing augment datasets. Three datasets were created: RDZ1K, RDZ5K and RDZ10K[1]. Initial ROC analysis, see figures A.1, A.3, A.5 suggested that the trained models should be able to perform segmentation task. For this idea confirmation is good to look at corresponding PRC curves, see figures A.2, A.4, A.6 and results from the confusion matrices, see 5.11. Information obtained from appropriate matrices (1, 2 and 3) shows that even the smallest data model was able to select correct label on 96 % of the pixels.

From ROC and PRC curves is obvious that the most problematic test image for these particular models was the image 5.1. This image is composed of several smaller output masks. Each is representing best output model within the category, where **1a** is original test image, **1b** original segmentation mask, **1c** original image overlayed with the model´s probability output mask, transparency 60%, **2** predicted mask by model RDZ1000 with 10 train iterations, **3** predicted mask by model RDZ5000 with 10 train iterations, **4** predicted mask by model RDZ10K with 10 train iterations, **5** predicted mask by model RDZ10K with 15 train iterations and **6** predicted mask by

[1] R stand for rotate function, D for distort, Z for zoom and 1, 5, 10 are multiple numeric prefixes for thousands

model RDZ10K with 20 train iterations. This identical layout is used for comparison purposes also on other test images.

The most probable reason of these problems is that this image was taken on high resolution and large magnification. Models are trying to discover burned holes in the atomic structure, with various grey level areas and small edges of the atoms.

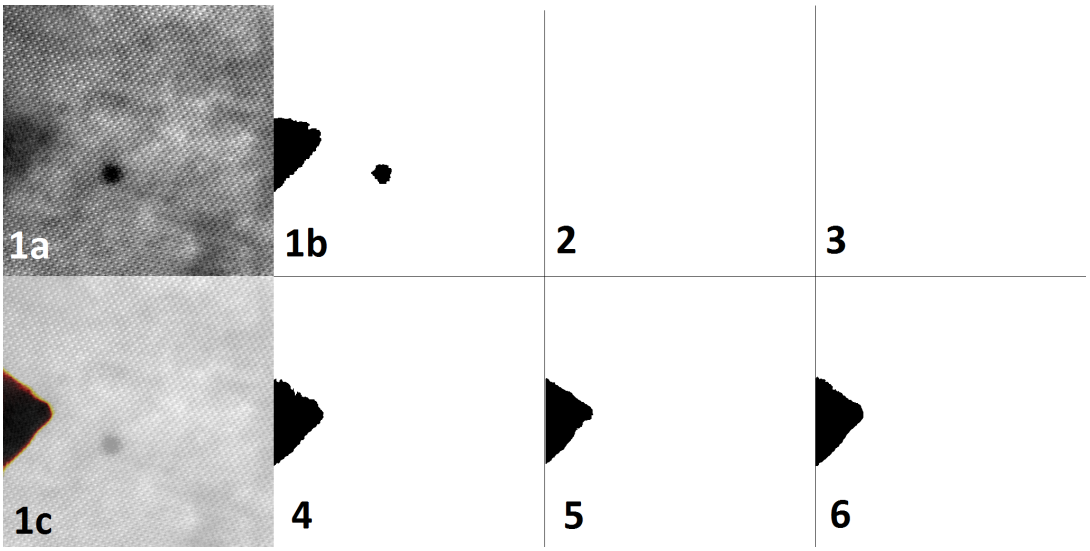

Fig. 5.1: Test image - spinel sample.

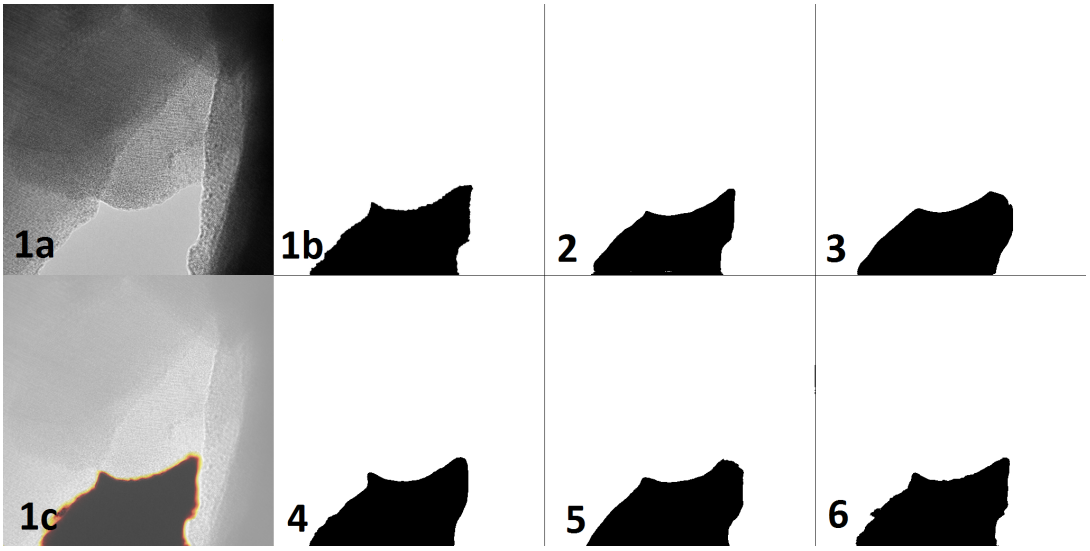

Fig. 5.2: Test image - graphitized carbon sample 2.

The least problematic test images are on figures 5.2 and 5.3 where first test image is graphitised carbon and second the cross grating sample. The first mentioned sample is particularly interesting for image processing because there is evidence that sample is composed of several carbon layers. Presented models had no problem to segment this sample. Particularly interesting it that the RDZ1K achieved better results[2] in the form of maintaining the sharp edge of the area. Apparently, with bigger variability, some smoothing the edge effect is occurring, at least from my training data. The second mentioned sample of the cross grating was interesting due to the rectangular windows. Such artificial pattern and regularity should be easy to learn for the model. Presented segmentation masks 5.3 confirm this, even though that small false positive fragments are present. Again, the presented smallest dataset provided better segmentation results that the two bigger ones.

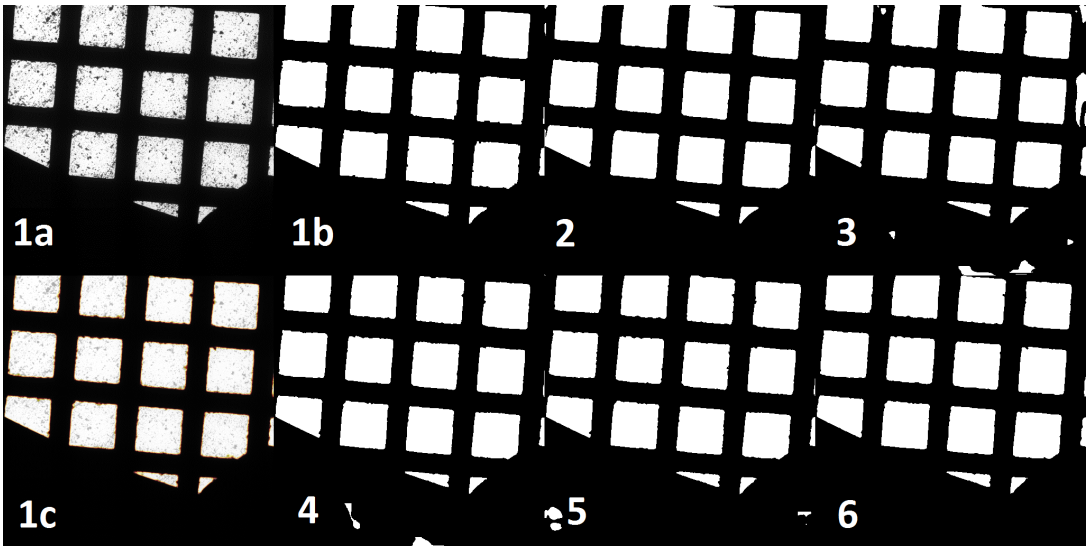

Fig. 5.3: Test image - crossgrating sample.

To numerically evaluate these results, see tables B.1, B.2 and B.3. Values of DSC and ARI are graphically presented in boxplots 5.4 and 5.5. As you can see, the models based on the smallest dataset have interesting segmentation potential, but large deviation does not provide the look of certainty. The ARI boxplots are even more drastic. The smallest training dataset has the lowest ARI mean value that highly suggests that the segmentation results are mostly random acts. As you can see from these two charts, the larger training dataset, with sufficient variability can provide better and more reliable segmentation results.

[2]Listed segmentation masks are actual masks created from the best segmentation model in each categorry

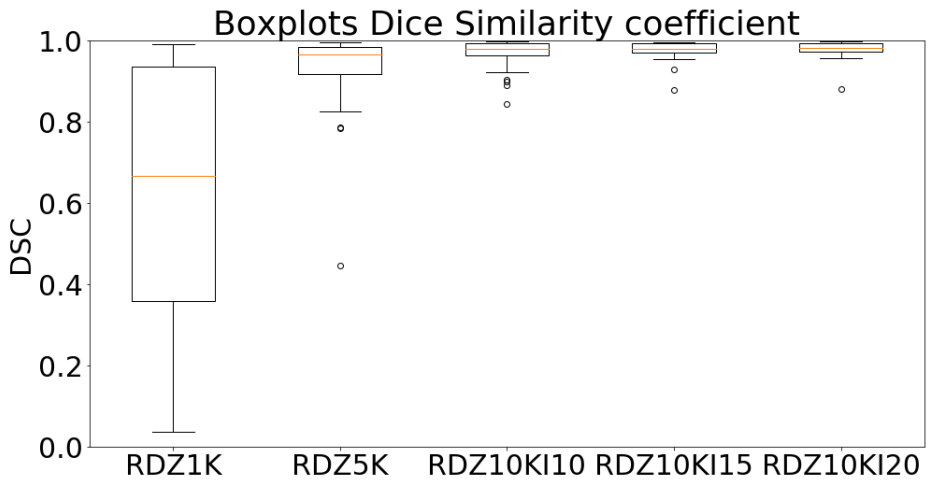

Fig. 5.4: Dice results.

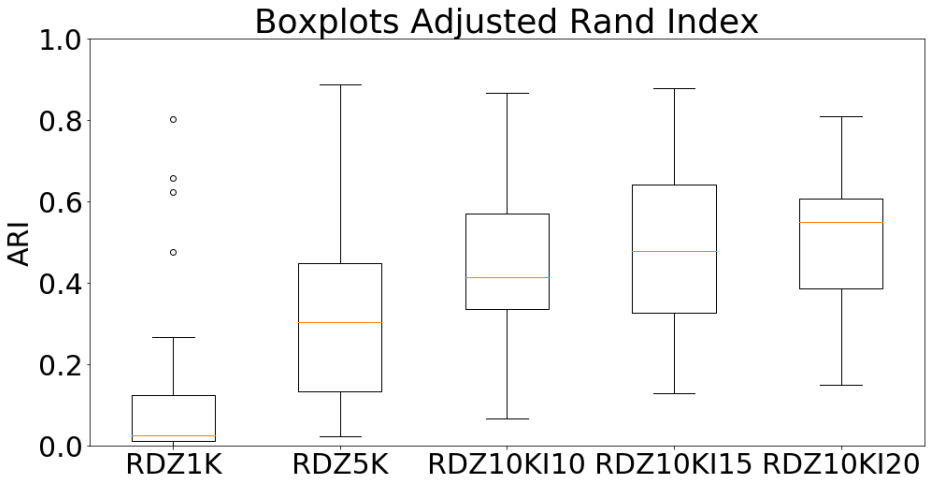

Fig. 5.5: ARI results.

5.2 Number of iterations influence

Another possibility how to improve model‘s behaviour is to let the model train for more iterations. There is a chance for improvement with every iteration. But there is also a chance to get stuck with the model training and with further iteration not improve the previous result. This could be overcome with appropriate setting of the training parameters mostly with the learning rate parameter. My models did not suffer from this situation, and I decided to stuck with the default and recom-

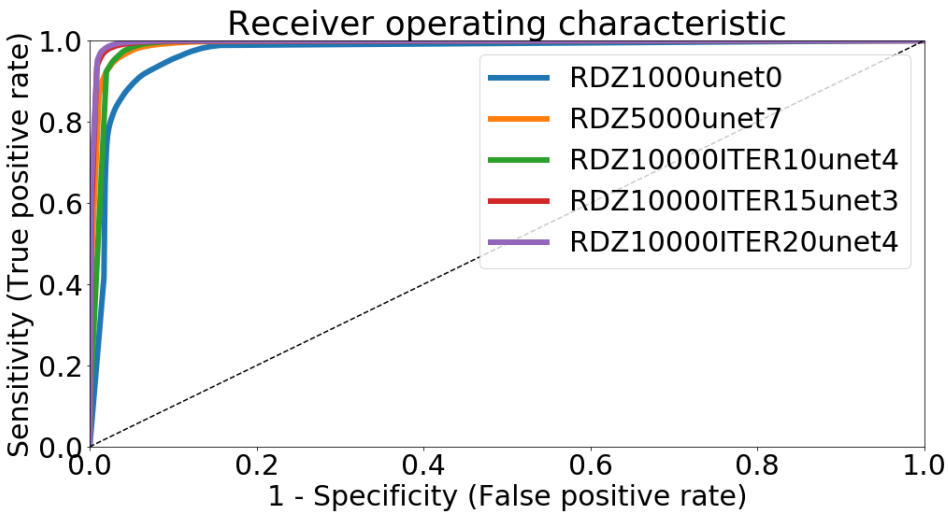

Fig. 5.6: ROC comparison across training datasets.

mended parameters set up, I was playing only with the training iteration parameter. From the figure, 5.6 is obvious that the increasing number of training iterations led to improvement in the behaviour of the receiver operating characteristics. As a confirmation of these results will serve figure 5.7 with Precision-Recall results was generated. These curves backed up the previous statement about the segmentation improvement with the larger number of iterations. To be able to visualise this improvement, another confusion matrices were generated, see fig. 5.11, which shows continuous elimination of false labeling. The amount of background to sample and sample to background pixel label dropped by 2 % in total.

The generated segmentation masks suffer from image artefacts despite previous curves results. Segmentation task on another sample of graphitised carbon, see 5.8 was not that successful. It is true that the most iteration model was the most successful one, but during the testing square shapes artefacts occurred. It is possible that during the randomised selection of training and validating set, it resulted that this shape was preferred and weights were set on it. With the continuous increase in the size of the dataset and number of training iterations, this artefact was almost eliminated. Notice, that the small area was detected only partially with the final model. The model was mostly trained in situations when the background formed only one consistent area and was not divided. This could be overcome with an improvement of the original training dataset.

With another graphitised carbon sample, where monolayers are particularly noticeable, see 5.9. The fifteen and twenty iterations models did not have the problem. Edges of the area are not smoothed out. This is happening mostly on smaller train-

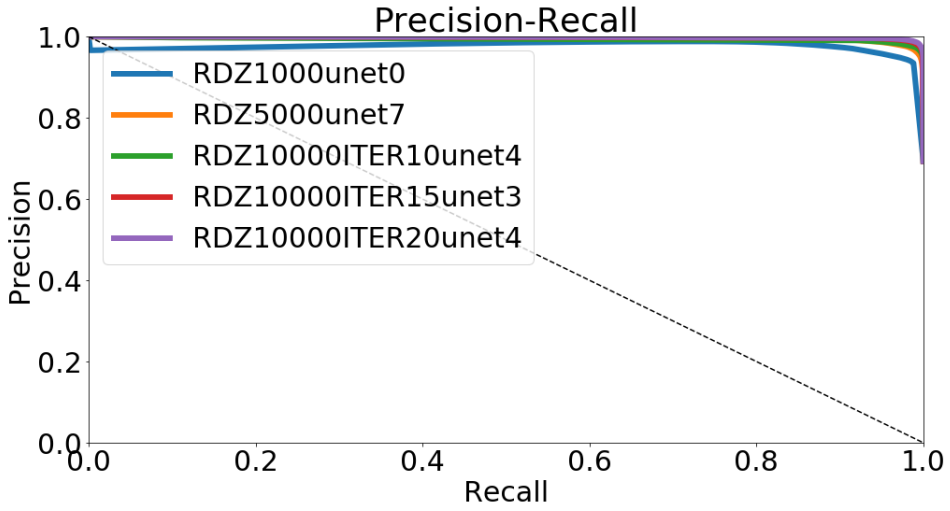

Fig. 5.7: PRC comparison across training datasets.

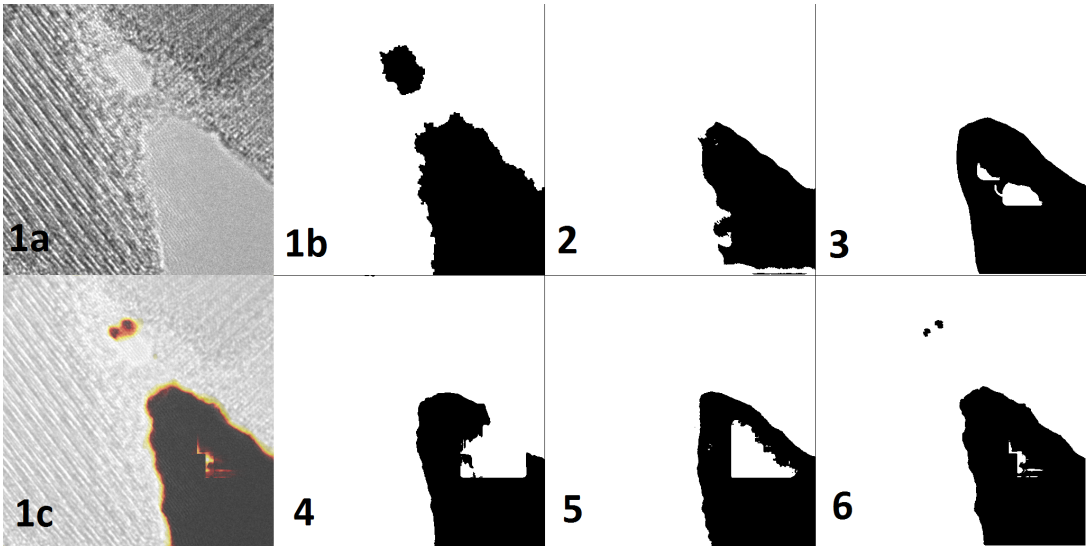

Fig. 5.8: Test image - graphitized carbon sample 1.

ing datasets RZD1K, and RZD5K suffers from this smoothing.

The last test image contained spinel. Notice that the image has very vague and blurry edge on the right side and distinctive J shape area in the centre. Here, the best segmentation results were provided by the RDZ10K with fifteen iterations. From the boxplots of DSC 5.4 and ARI 5.5 can be seen, that the fifteen and twenty iterations models should behave very similar, in this case, the smaller iteration number outperformed, the bigger one. All DSC and ARI results are in tables B.4.

Fig. 5.9: Test image - graphitized carbon sample 3.

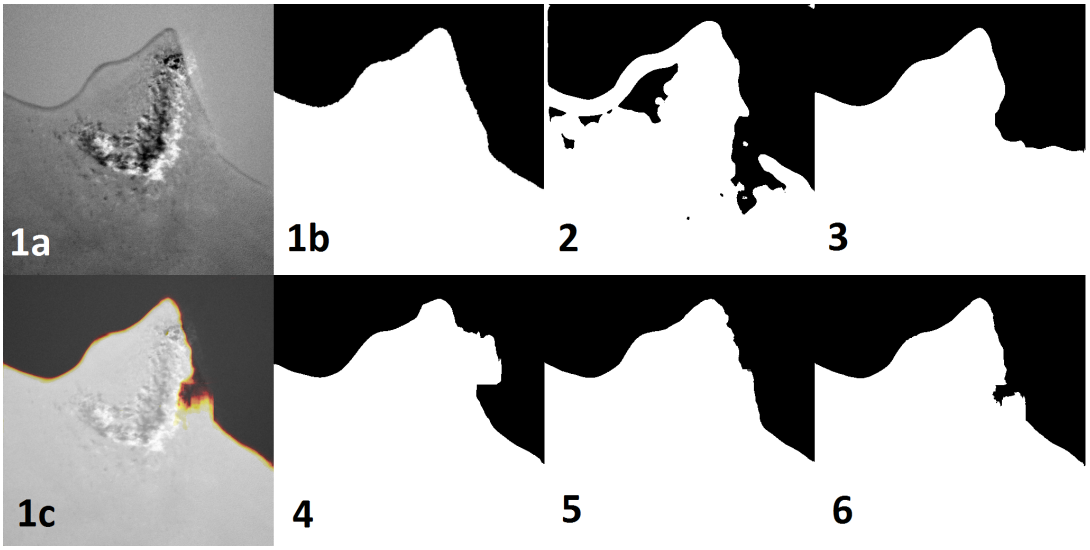

Fig. 5.10: Test image - spinel sample 1.

5.3 Edge and Region comparison

Computed neural network models were compared with two detection algorithms. Scikit-image processing library [94] has implemented edge-based and region-based segmentation algorithms. For the edge detection Canny detector was used, for region-based detection Sobel gradient method and watershed algorithm was used. Both approaches theoretically could compete with the neural network. Test images have decent edges and large consistent regions. But the numerical results in table

5.1 provides precise information about the lack of universal application. Obtained DSC values for both approaches suggest that some achievement was reached. But these results are not consistent and according to ARI score are more random than predictable. Especially edge-based algorithm acts randomly, which is also supported by its confusion matrix in 5.11. Resulting segmentation masks in figure 5.12 shows where these algorithms succeed and where they failed. The clean and consistent regions as the cross grating sample 1a showed where both of these algorithms showed some success. Especially take a look at 5c, where region-based algorithm outperformed even the best neural network model when it found the small burned hole in the sample. If you take a closer look at the edge based samples 1-6b there are apparent efforts to find the leading edge of the object, but the noise is keeping the output score low. Intentionally I did not perform any tuning on these implementations, because no adjustment was made on neural network topology.

Tab. 5.1: Final DSC and ARI results with 2σ

Model	$\overline{DSC}\pm 2\sigma$	$\overline{ARI}\pm 2\sigma$
RDZ1K	0.63±0.59	0.10±0.34
RDZ5K	0.94±0.17	0.32±0.40
RDZ10KI10	0.97±0.06	0.43±0.34
RDZ10KI15	0.98±0.05	0.50±0.41
RDZ10KI20	0.98±0.04	0.52±0.30
Edge	0.72±0.33	0.07±0.14
Region	0.57±0.80	0.28±0.87

5.4 Neural network comparison

There is a limited amount of possibilities to evaluate and compare different neural topologies correctly. Most of the neural topologies are very complicated [62, 66, 67] and code implementation is not friendly to use. They usually require specialised hardware, training is time-consuming, and they are generally tuned for the particular task - the majority of them is set up for classification problem. In this thesis was an effort to find a comparable neural network that would perform segmentation of grey-tone images and would be well documented for testing implementation. Unfortunately, such comparison wasn't accomplished.

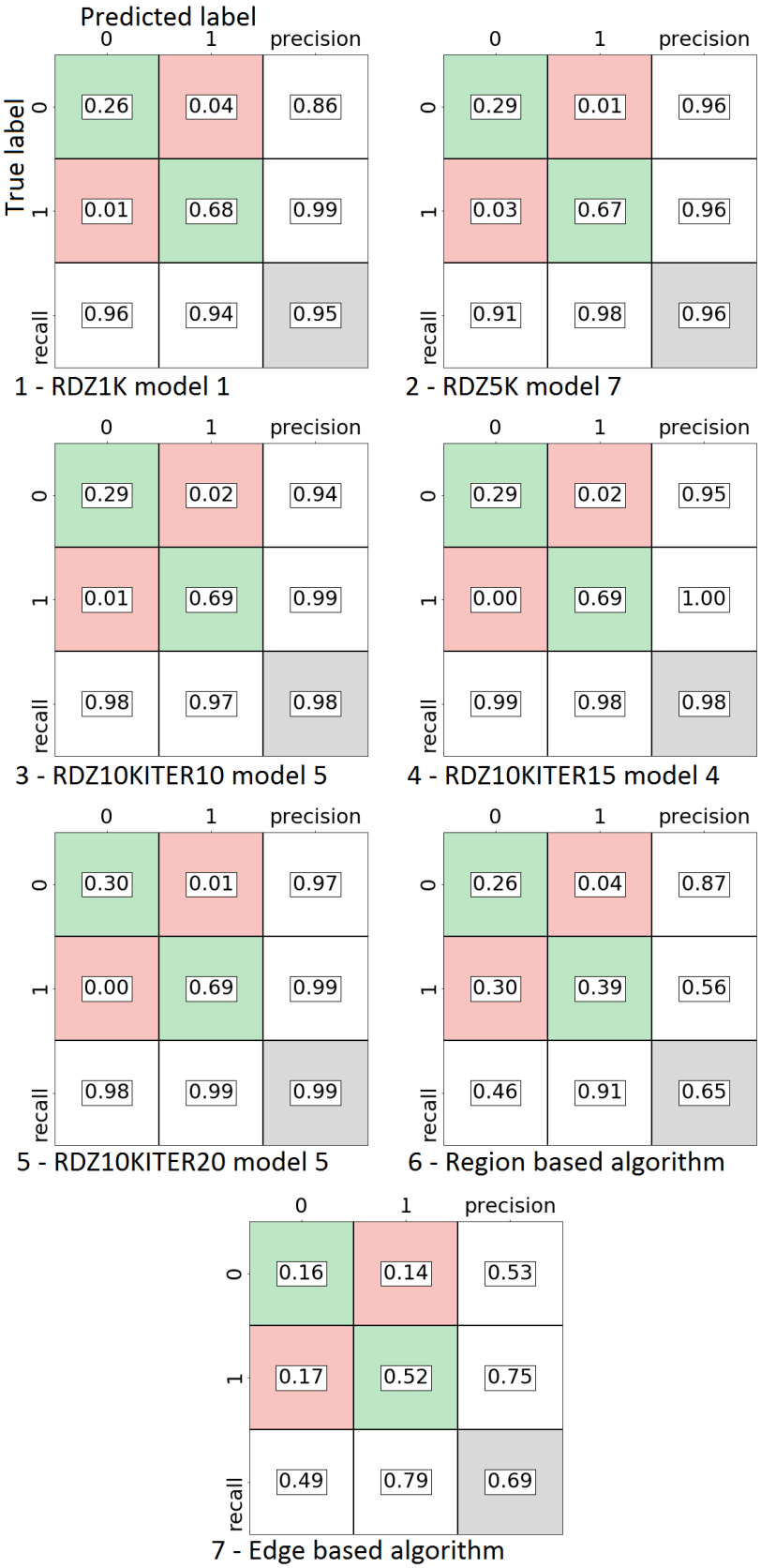

Fig. 5.11: Confusion matrix results - best models.

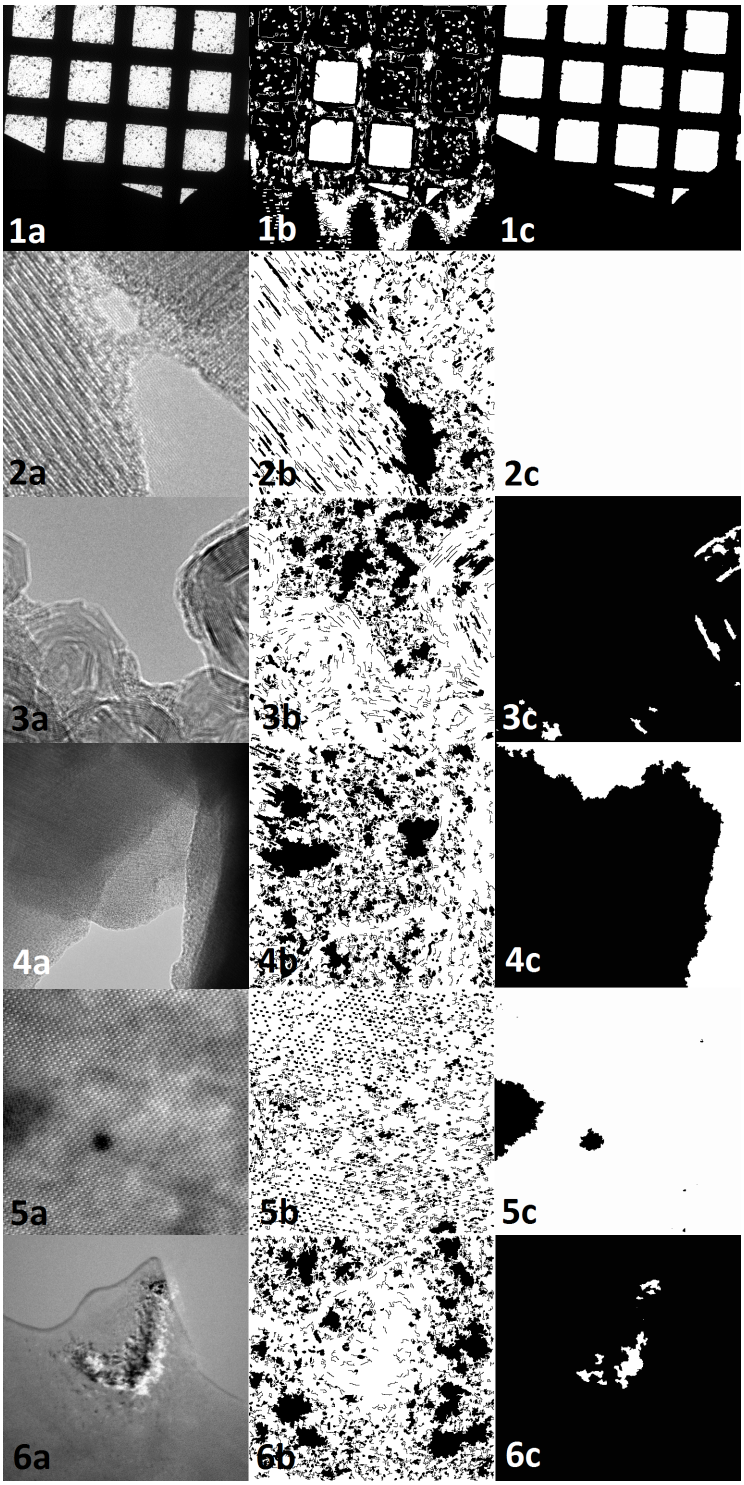

Fig. 5.12: a - original images, b - edge based detection, c - region based detection

6 CONCLUSION

This diploma thesis focuses on the problematics of adaptive segmentation methods with the primary focus on segmentation technique using neural networks. Thesis familiarises the reader with the area of transmission electron microscopy, key components, image acquisition and three types of aberration. In follow up part adaptive segmentation methods are described. This description serves for understanding the critical difference in operating principle of neural networks.

The central part of the thesis focuses on the area of convolution neural networks. A thorough description of their architecture development, how they work and their ability to serve as segmentation tool is provided. This information then helped for the decision of the convolution neural network topology U-NET for segmentation task. With the cooperation of the Thermo Fisher Scientific (formerly FEI Czech Republic s.r.o) was created training and testing dataset. For training, purposes are this dataset appropriately enlargened with the use of Augmentation library - Augmentor. Then a plan for testing the segmentation task was designed. The plan is based on the variability and number of training iteration of the training dataset. For numerical evaluation of achieved results are presented two different metrics: first Dice similarity coefficient for evaluating the similarity between original and output segmentation mask; second the Adjusted Rand index for assessing the credibility of the particular segmentation model.

Achieved results are graphically presented in the form of segmentation masks, receiver-operating characteristics and precision-recall curves.

Thesis confirms two hypothesis: first about the influence of the training dataset on the quality of the resulting segmentation; second about the number of training iterations on the resulting segmentation. Trained models are also compared with the available implementation of the Edge and Region based detection algorithms. The thesis also confirms the versatility of convolution neural networks. When the topology is designed for segmentation purposes, with small or none modifications could be applied to segmentation task on different data that it was initially intended for. This statement is based on the original design of the U-Net that was designed for segmentation the structures of the Drosophila first instar larva ventral nerve cord (VNC) [73] and for this thesis TEM inorganic samples were used.

During the testing period, I encountered the usual implementation problems - incompatible libraries, missing documentation, GPU problems, RAM limitation etc. These problems were overcome, and achieved results are astonishing. With such high variability of the input data: different magnification, different imaging modes, various types of samples; the best model produced the mean value of DSC = 0.99 ± 0.01 and ARI = 0.56 ± 0.24. This result confirms the quality of convolution

neural networks when it comes to the segmentation problematics with high data variability.

Increasing size of the training dataset and number of iterations displayed that the trained model - RDZ10KITER20 eliminated false positive and false negative mask artefacts. But there are still possibilities how to improve the trained models. At first, the initial segmentation masks were created by two company experts. A chance to consult these segmentation masks with more experts from the particular field could enhance the segmentation results. Another possible improvement would be to add more variability to the original training dataset. The third solution would be to modify the original network design. Author of this thesis performed modification in the form of changing the activation function, adjusting the kernel sizes, changing the optimisation algorithms and also adding a layer. Such changes led either to computing errors or unpublishable results. Therefore recommendation is to stick with the original design and modify the controllable part - training data.

BIBLIOGRAPHY

[1] KUMAR, Amit and Fahimuddin SHAIK. 2016. Importance of Image Processing. Image Processing in Diabetic Related Causes [online]. Singapore: Springer Singapore, 5-7. SpringerBriefs in Applied Sciences and Technology. Available at: http://link.springer.com/10.1007/978-981-287-624-9_2

[2] MOCKO, Štefan (ed.). 2018. U-net convolutional neural network for TEM image segmentation. In: Proceedings of the 24 th Conference STUDENT EEICT 2018. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, p. 270-272.

[3] About ImageNet. 2010. ImageNet [online]. Available at: http://imagenet.org/about-overview

[4] HE, Kaiming, Xiangyu ZHANG, Shaoqing REN and Jian SUN. 2015. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. 2015 IEEE International Conference on Computer Vision (ICCV) [online]. IEEE, 1026-1034. Available at: http://ieeexplore.ieee.org/document/7410480/

[5] KUBÍNEK, Roman, Klára ŠAFÁŘOVÁ and Milan VŮJTEK. 2011. Elektronová mikroskopie [online]. Olomouc: Univerzita Palackého v Olomouci. Available at: https://fyzika.upol.cz/cs/system/files/download/vujtek/granty/elmikro.pdf

[6] The schematic outline of a TEM. In: Atomic World - Transmission electron microscope(TEM): Principle of TEM [online]. Available at: http://www.hk-phy.org/atomic_world/tem/tem02_e.html

[7] PINHEIRO, T., M. DOLORES YNSA and L. ALVES. 2007. Imaging biological structures with proton microprobe [online]. Available at: http://www.formatex.org/microscopy3/pdf/pp237-244.pdf

[8] ZSCHORNACKA, G., M. SCHMIDT and A. THORN. 2014. Electron Beam Ion Sources. CAS-CERN Accelerator School: Ion Sources - Proceedings [online]. , 1-37. Available at: http://cds.cern.ch/record/1965922

[9] WILLIAMS, David B. and C. Barry. CARTER. 2008. Transmission electron microscopy: a textbook for materials science. 2nd ed. New York: Springer.

[10] MACARTHUR, By Katherine E. 2016. The Use of Annular Dark-Field Scanning Transmission Electron Microscopy for Quantitative Characterisation. Johnson Matthey Technology Review [online]. 60(2), 117-131. Available at: http://www.ingentaconnect.com/content/10.1595/205651316X691186

[11] Gun parts. 2014. In: Australian Microscopy & Microanalysis Research Facility [online]. Australia. Available at: http://www.ammrf.org.au/myscope/images/sem/gun_parts.png

[12] CREWE, A. V., D. N. EGGENBERGER, J. WALL and L. M. WELTER. 1968. Electron Gun Using a Field Emission Source. Review of Scientific Instruments [online]. 39(4), 576-583. Available at: http://aip.scitation.org/doi/10.1063/1.1683435

[13] Denka M-3 LaB_6 EM cathodes: High brightness, direct replacement electron sources. 1999. Micro to Nano: Innovative Microscopy Supplies [online]. Haarlem NL: Micro to Nano. Available at: https://www.microtonano.com/Denka-M-3-LaB6-EM-cathodes.php

[14] JEOL, A Guide to Scanning Microscope Observation. 2002. Akishima JPN, 36 p. Available at: http://www.geology.wisc.edu/j̃ohnf/g777/JEOLguide.pdf

[15] KESHRI, Anup Kr., Srikanth KORLA, Sushma AMRUTHALURI, Venkata PASUMARTHI and Xudong CHEN. 2014. Lenses and Apertures of A TEM. Miami, Florida. Available at: https://tinyurl.com/ybgzbzxp

[16] RANGER, Nicole T., Ehsan SAMEI, James T. DOBBINS and Carl E. RAVIN. 2007. Assessment of Detective Quantum Efficiency: Intercomparison of a Recently Introduced International Standard with Prior Methods 1. Radiology [online]. 243(3), 785-795. Available at: http://pubs.rsna.org/doi/abs/10.1148/radiol.2433060485

[17] ISHIHARA, Y., E. ODA, H. TANIGAWA, et al. 1982. Interline CCD image sensor with an anti blooming structure. 1982 IEEE International Solid-State Circuits Conference. Digest of Technical Papers [online]. IEEE, , 168-169. Available at: http://ieeexplore.ieee.org/document/1156370/

[18] PERKES, Paul. 1987. Electron Microscopy: Image Formation in the TEM. Arizona State University [online]. Tempe AZ: Arizona Board of Regents. Available at: https://www.asu.edu/courses/phs208/patternsbb/PiN/rdg/elmicr/elmicr-tem.shtml

[19] NEBESÁŘOVÁ, Jana. 2002. Elektronová mikroskopie pro biology: Tvroba obrazu. Laboratory of Electron Microscopy: Biology Centre of ASCR - Institute of Parasitology [online]. České Budějovice. Available at: http://triton.paru.cas.cz/old-lem/book/Podkap/3.4.html

[20] HETHERINGTON, Crispin. 2004. Aberration correction for TEM. Materials Today [online]. 7(12), 50-55. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1369702104005711

[21] LINCK, Martin, Peter HARTEL, Stephan UHLEMANN, et al. 2016. Chromatic Aberration Correction for Atomic Resolution TEM Imaging from 20 to 80 kV. Physical Review Letters [online]. 117(7), -. Available at: https://link.aps.org/doi/10.1103/PhysRevLett.117.076101

[22] AHN, Jae Hyung, Tai-Wook KIM and Heui Jae PAHK. 2015. Fast focus and astigmatism correction algorithm for critical dimension measurement using electron microscopy. International Journal of Precision Engineering and Manufacturing [online]. 16(9), 1941-1947. Available at: http://link.springer.com/10.1007/s12541-015-0252-5

[23] ONG, K. H., J. C. H. PHANG and J. T. L. THONG. 1997. A Robust Focusing and Astigmatism Correction Method for the Scanning Electron Microscope. Scanning [online]. 19(8), 553-563. Available at: http://onlinelibrary.wiley.com/doi/10.1002/sca.4950190805/pdf

[24] SCHERES, Sjors H W, Rafael NÚÑEZ-RAMÍREZ, Carlos O S SORZANO, José María CARAZO and Roberto MARABINI. 2008. Image processing for electron microscopy single-particle analysis using XMIPP. Nature Protocols [online]. 3(6), 977-990. Available at: http://www.nature.com/articles/nprot.2008.62

[25] NAGAYAMA, K. and R. DANEV. 2008. Phase contrast electron microscopy: development of thin-film phase plates and biological applications. Philosophical Transactions of the Royal Society B: Biological Sciences [online]. 363(1500), 2153-2162. Available at: http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2008.2268

[26] LÉZORAY, Olivier and Leo GRADY. 2012. Image processing and analysis with graphs theory and practice [online]. Online-Ausg. Boca Raton, FL: CRC Press. Available at: https://tinyurl.com/yc9fbrn2

[27] ŠPANĚL, Michal and Vítězslav BERAN. 2005. Obrazové segmentační techniky: Přehled existujících metod [online]. Brno. Available at: http://www.fit.vutbr.cz/španel/segmentace/

[28] YUHENG, Song and Yan HAO. 2017. Image Segmentation Algorithms Overview. CoRR [online]. , 6. Available at: https://arxiv.org/abs/1707.02051

[29] LU MING. 2010. Image segmentation algorithm research and improvement. 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE) [online]. IEEE, V5-211-V5-214. Available at: http://ieeexplore.ieee.org/document/5579114/

[30] PAL, Nikhil R and Sankar K PAL. 1993. A review on image segmentation techniques. Pattern Recognition [online]. 26(9), 1277-1294. Available at: http://linkinghub.elsevier.com/retrieve/pii/003132039390135J

[31] JIANBO SHI and J. MALIK. 2000. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence [online]. 22(8), 888-905. Available at: http://ieeexplore.ieee.org/document/868688/

[32] LONG, Jonathan, Evan SHELHAMER and Trevor DARRELL. 2015. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [online]. IEEE, , 3431-3440. Available at: http://ieeexplore.ieee.org/document/7298965/

[33] SEUNGHOON, Hong and Han BOHYUNG. 2015. Decoupled Deep Neural Network for Semi-supervised Semantic Segmentation. CoRR[online]. 1-9. Available at: https://arxiv.org/abs/1506.04924

[34] LEMPITSKY, Victor, Andrea VEDALDI and Andrew ZISSERMAN. 2011. A Pylon Model for Semantic Segmentation. In: Proceedings of the 24th International Conference on Neural Information Processing Systems [online]. Granada, Spain: Curran Associates, p. 9. Available at: https://dl.acm.org/citation.cfm?id=2986625

[35] FARABET, Clement, Camille COUPRIE, Laurent NAJMAN and Yann LECUN. 2013. Learning Hierarchical Features for Scene Labeling. IEEE Transactions on Pattern Analysis and Machine Intelligence [online]. 35(8), 1915-1929. Available at: http://ieeexplore.ieee.org/document/6338939/

[36] PINHEIRO, Pedro O. and Ronan COLLOBERT. 2014. From Image-level to Pixel-level Labeling with Convolutional Networks. CoRR [online]. 1-9. Available at: https://arxiv.org/abs/1411.6228

[37] HONGSHENG, Li, Rui ZHAO and Wang XIAOGANG. 2014. Highly Efficient Forward and Backward Propagation of Convolutional Neural Networks for Pixelwise Classification. CoRR [online]. , 1-10. Available at: http://arxiv.org/abs/1412.4526

[38] ZHANG, Yu Jin. 1995. Influence of segmentation over feature measurement. Pattern Recognition Letters [online]. 16(2), 201-206. Available at: http://linkinghub.elsevier.com/retrieve/pii/016786559400083F

[39] HARALICK, Robert M. and Linda G. SHAPIRO. 1985. Image segmentation techniques. Computer Vision, Graphics, and Image Processing[online]. 29(1), 100-132. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0734189X85901537

[40] DUVAL, Laurent. 2018. Edge Types. In: Stackexchange [online]. Available at: https://dsp.stackexchange.com/questions/47556/why-edge-sharpening-produces-high-frequency

[41] DAVIS, E.R. 2005. Machine vision: Chapter 5. edge detection. 3rd ed. Boston: Elsevier. 140-185. Available at: https://tinyurl.com/ybsqdwyk

[42] VIKRAM MUTNEJA, Dharampal. 2015. Methods of Image Edge Detection: A Review. Journal of Electrical & Electronic Systems [online]. 04(02), -. Available at: https://www.omicsonline.org/open-access/methods-of-image-edge-detection-a-review-2332-0796-1000136.pdf

[43] ÖZTÜRK, Şaban and Bayram AKDEMIR. 2015. Comparison of Edge Detection Algorithms for Texture Analysis on Glass Production. Procedia - Social and Behavioral Sciences [online]. 195, 2675-2682. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1877042815039567

[44] MUTHUKRISHNAN, R and M RADHA. 2011. Edge Detection Techniques For Image Segmentation. International Journal of Computer Science and Information Technology [online]. 3(6), 259-267. Available at: http://www.airccse.org/journal/jcsit/1211csit20.pdf

[45] KAGANAMI, Hassana Grema and Zou BEIJI. 2009. Region-Based Segmentation versus Edge Detection. 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing [online]. IEEE, 1217-1221. Available at: http://ieeexplore.ieee.org/document/5337327/

[46] MANCAS, Matei, Edward R. DOUGHERTY, Jaakko T. ASTOLA, Bernard GOSSELIN, Benoit MACQ and Karen O. EGIAZARIAN. 2005. Segmentation using a region-growing thresholding [online]. 388-. Available at: https://tinyurl.com/ycbzaoae

[47] KOHONEN, Teuvo. Self-Organizing Maps. In: Mnemosyne_Studio [online]. Available at: http://mnemstudio.org/neural-networks-kohonen-self-organizing-maps.htm

[48] KOHONEN, Teuvo. 2013. Essentials of the self-organizing map. Neural Networks [online]. 37, 52-65. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0893608012002596

[49] KOHONEN, Teuvo and Timo HONKELA. 2007. Kohonen network. Scholarpedia [online]. 2(1), 1568-. Available at: http://www.scholarpedia.org/article/Kohonen_network

[50] MITCHELL, S.C., J.G. BOSCH, B.P.F. LELIEVELDT, R.J. VAN DER GEEST, J.H.C. REIBER and M. SONKA. 2002. 3-D active appearance models: segmentation of cardiac MR and ultrasound images. IEEE Transactions on Medical Imaging [online]. 21(9), 1167-1178. Available at: http://ieeexplore.ieee.org/document/1166645/

[51] ATIYA, Amir F. 1990. An unsupervised learning technique for artificial neural networks. Neural Networks [online]. 3(6), 707-711. Available at: http://linkinghub.elsevier.com/retrieve/pii/089360809090058S

[52] XIE, Xiurui, Hong QU, Guisong LIU, Malu ZHANG, Jürgen KURTHS and Gennady CYMBALYUK. 2016. An Efficient Supervised Training Algorithm for Multilayer Spiking Neural Networks. PLOS ONE [online]. 11(4), e0150329-. Available at: http://dx.plos.org/10.1371/journal.pone.0150329

[53] TAJBAKHSH, Nima, Jae Y. SHIN, Suryakanth R. GURUDU, R. Todd HURST, Christopher B. KENDALL, Michael B. GOTWAY and Jianming LIANG. 2016. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? IEEE Transactions on Medical Imaging [online]. 35(5), 1299-1312. Available at: http://ieeexplore.ieee.org/document/7426826/

[54] SHELHAMER, Evan, Jonathan LONG and Trevor DARRELL. 2017. Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence [online]. 39(4), 640-651. Available at: http://ieeexplore.ieee.org/document/7478072/

[55] DORFLER, Monika, Roswitha BAMMER and Thomas GRILL. 2017. Inside the spectrogram: Convolutional Neural Networks in audio processing. 2017 International Conference on Sampling Theory and Applications (SampTA) [online]. IEEE, 152-155. Available at: http://ieeexplore.ieee.org/document/8024472/

[56] 2000. BRACEWELL, Ronald N. The Fourier transform and its applications [online]. 3rd ed. Boston: McGraw Hill, p. 24-27, 331-335.

Available at: http://www.zuj.edu.jo/download/the-fourier-transform-and-its-applications-bracewell-pdf/

[57] LI, Fei-Fei, Justin JOHNSON and Serena YEUNG. 2017. Neural Network architectures. CS231n: Convolutional Neural Networks for Visual Recognition: Spring 2017 [online]. Stanford, CA: Stanford University. Available at: http://cs231n.github.io/neural-networks-1/

[58] GIRSHICK, Ross, Jeff DONAHUE, Trevor DARRELL and Jitendra MALIK. 2016. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence [online]. 38(1), 142-158. Available at: http://ieeexplore.ieee.org/document/7112511/

[59] SHARMA, Avinash. 2017. Understanding Activation Functions in Neural Networks. In: The Theory Of Everything: Musings of an AI , Deep Learning, Mathematics addict [online]. US: A Medium Corporation. Available at: https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0

[60] SHARMA, Sagar. 2017. Activation Functions. In: Towards Data Science [online]. Available at: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

[61] VAN VEEN, Fjodor. 2016. The Neural network ZOO. In: The Asimov Institute [online]. Utrecht, NL: The Asimov Institute. Available at: http://www.asimovinstitute.org/neural-network-zoo/

[62] KRIZHEVSKY, Alex, Ilya SUTSKEVER and Geoffrey E. HINTON. 2017. ImageNet classification with deep convolutional neural networks. Communications of the ACM [online]. 60(6), 84-90. Available at: http://dl.acm.org/citation.cfm?doid=3098997.3065386

[63] HINTON, Geoffrey E., Nitish SRIVASTAVA, Alex KRIZHEVSKY, Ilya SUTSKEVER and Ruslan R. SALAKHUTDINOV. 2012. Improving neural networks by preventing co-adaptation of feature detectors. Corr [online]. 18. Available at: https://arxiv.org/abs/1207.0580

[64] LI, Fei-Fei, Justin JOHNSON and Serena YEUNG. 2017. Neural Network architectures. CS231n: Convolutional Neural Networks for Visual Recognition: Spring 2017 [online]. Stanford, CA: Stanford University. Available at: http://cs231n.github.io/convolutional-networks/

[65] LI, Fei-Fei, Justin JOHNSON and Serena YEUNG. 2017. Linear Classification. CS231n: Convolutional Neural Networks for Visual Recognition: Spring 2017 [online]. Stanford, CA: Stanford University. Available at: http://cs231n.github.io/linear-classify/

[66] SIMONYAN, Karen and Andrew ZISSERMAN. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR [online]. 14. Available at: https://arxiv.org/abs/1409.1556

[67] SZEGEDY, Christian, WEI LIU, YANGQING JIA, et al. 2015. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [online]. IEEE, 1-9. Available at: http://ieeexplore.ieee.org/document/7298594/

[68] GoogLeNet architecture [online]. In: . Available at: https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/single-shot-detectors.html

[69] DODGE, Samuel and Lina KARAM. 2017. A Study and Comparison of Human and Deep Learning Recognition Performance Under Visual Distortions. CoRR [online]. 7. Available at: http://arxiv.org/abs/1705.02498

[70] HE, Kaiming, Xiangyu ZHANG, Shaoqing REN and Jian SUN. 2016. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [online]. IEEE, 770-778. Available at: http://ieeexplore.ieee.org/document/7780459/

[71] CIRESAN, Dan, Alessandro GIUSTI, Luca M. GAMBARDELLA and Jürgen SCHMIDHUBER. 2012. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. Advances in Neural Information Processing Systems [online]. (25). Available at: https://papers.nips.cc/paper/4741-deep-neural-networks-segment-neuronal-membranes-in-electron-microscopyimages.pdf

[72] Convolution sliding window. In: NVIDIA Developer Blog [online]. Available at: https://devblogs.nvidia.com/deep-learning-nutshell-core-concepts/convolution_schematic/

[73] RONNEBERGER, Olaf, Philipp FISCHER and Thomas BROX. 2015. UNet: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 [online]. Cham: Springer International Publishing, 234-241. Lecture Notes in Computer Science. Available at: http://link.springer.com/10.1007/978-3-319-24574-4_28

[74] Checkerboard pattern. 2007. In: Wikimedia commons [online]. Available at: https://commons.wikimedia.org/wiki/File:Checkerboard_pattern.svg

[75] BLOICE, Marcus D., Christof STOCKER, and Andreas HOLZINGER. 2017. Augmentor: An Image Augmentation Library for Machine Learning, arXiv preprint arXiv:1708.04680, Available at: https://arxiv.org/abs/1708.04680

[76] PESCE, Lorenzo L., Charles E. METZ and Kevin S. BERBAUM. 2010. On the Convexity of ROC Curves Estimated from Radiological Test Results. Academic Radiology [online]. 17(8), 960-968.e4. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1076633210001789

[77] CLARK, Adrian F. and Christine CLARK. 2017. Performance Characterization in Computer Vision. In: University of Essex [online]. Colchester, UK: VASE Laboratory, Computer Science and Electronic Engineering University of Essex. Available at: http://peipa.sx.ac.uk/vase/talks/performance-evaluation.pdf

[78] BROWNLEE, Jason. 2016. Metrics To Evaluate Machine Learning Algorithms in Python. In: Machine learning mastery [online]. Vermont Victoria, Australia: Machine Learning Mastery. Available at: https://machinelearningmastery.com/metrics-evaluate-machine-learningalgorithms-python/

[79] HUBERT, Lawrence and Phipps ARABIE. 1985. Comparing partitions. Journal of Classification [online]. 2(1), 193-218. Available at: http://link.springer.com/10.1007/BF01908075

[80] Sørensen, T. (1948). "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". Kongelige Danske Videnskabernes Selskab. 5 (4): 1–34.

[81] MCCUNE, Bruce and James GRACE. 2002. Analysis of Ecological Communities. Mjm Software Design; ISBN 0-9721290-0-6.

[82] ZIJDENBOS, A.P., B.M. DAWANT, R.A. MARGOLIN and A.C. PALMER. 1994. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Transactions on Medical Imaging [online]. 13(4), 716-724. Available at: http://ieeexplore.ieee.org/document/363096/

[83] ZOU, Kelly H., Simon K. WARFIELD, Aditya BHARATHA, et al. 2004. Statistical validation of image segmentation quality based on a spatial

overlap index1. Academic Radiology [online]. 11(2), 178-189. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1076633203006718

[84] EL-BAZ, Ayman, Xiaoyi JIANG and Jasjit S. SURI (eds.). 2017. Biomedical image segmentation: advances and trends [online]. Boca Raton: CRC Press, p. 278-279. Available at: https://tinyurl.com/yb89yy6v

[85] CHOLLET, François. 2015. Keras. GitHub [online]. Available at: https://github.com/fchollet/keras

[86] KINGMA, Diederik P. and Jimmy BA. 2014. Adam: A Method for Stochastic Optimization. CoRR [online]. abs/1412.6980, 1-15. Available at: https://arxiv.org/abs/1412.6980

[87] RUDER, Sebastian. 2016. An overview of gradient descent optimization algorithms. CoRR [online]. abs/1609.04747, 1-14. Available at: https://arxiv.org/abs/1609.04747

[88] ABADI, Martín, Ashish AGARWAL, Paul BARHAM, et al. 2015. TensorFlow: A System for Large-Scale Machine Learning. In: Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16) [online]. Savannah, GA, USA, p. 265-283. Available at: https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf

[89] ROBERTS, Jeff John. 2017. Why Tech Investors Love ICOs-and Lawyers Don't. In: Fortune [online]. California, US: Meredith. Available at: http://fortune.com/2017/06/26/ico-initial-coin-offering-investing/

[90] DANNEN, Chris. 2017. Crypto-News Recap: June-July 2017. In: Medium [online]. US: A Medium Corporation. Available at: https://tinyurl.com/y86v4bjk

[91] Cambridge Dictionary. 2017. Cambridge Advanced Learner's Dictionary & Thesaurus [online]. Cambridge, UK: Cambridge University Press. Available at: https://dictionary.cambridge.org/dictionary/english/framework

[92] SAITO, Takaya, Marc REHMSMEIER and Guy BROCK. 2015. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLOS ONE [online]. 10(3), e0118432-. Available at: http://dx.plos.org/10.1371/journal.pone.0118432

[93] DAVIS, Jesse and Mark GOADRICH. 2006. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd international conference on Machine learning - ICML '06 [online].

New York, New York, USA: ACM Press, , 233-240. Available at: http://portal.acm.org/citation.cfm?doid=1143844.1143874

[94] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu and the scikit-image contributors. scikit-image: Image processing in Python. PeerJ 2:e453 (2014) http://dx.doi.org/10.7717/peerj.453

LIST OF APPENDICES

A Graphical results **74**

B Numerical results **79**

A GRAPHICAL RESULTS

In this part of the apendix are ROC and Precision-Recall curves that represent the behaviour of the best neural network model from each training category.

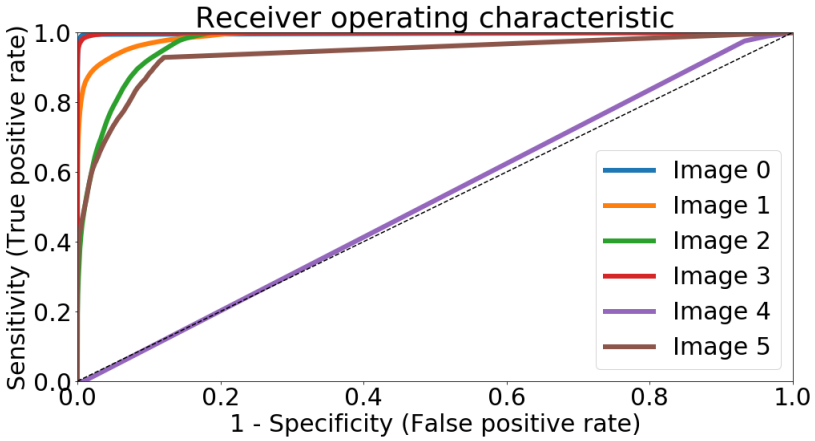

Fig. A.1: ROC. Training Dataset RDZ1000, 10 iterations. Model 1.

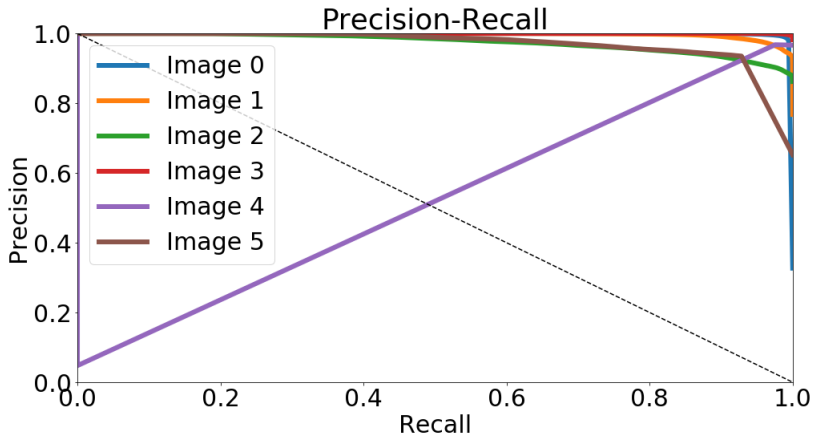

Fig. A.2: PRC. Training Dataset RDZ1000, 10 iterations. Model 1.

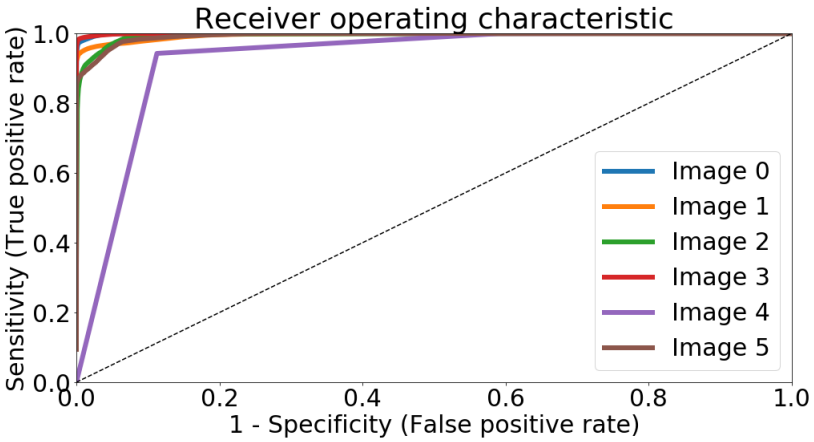

Fig. A.3: ROC. Training Dataset RDZ5000, 10 iterations. Model 8.

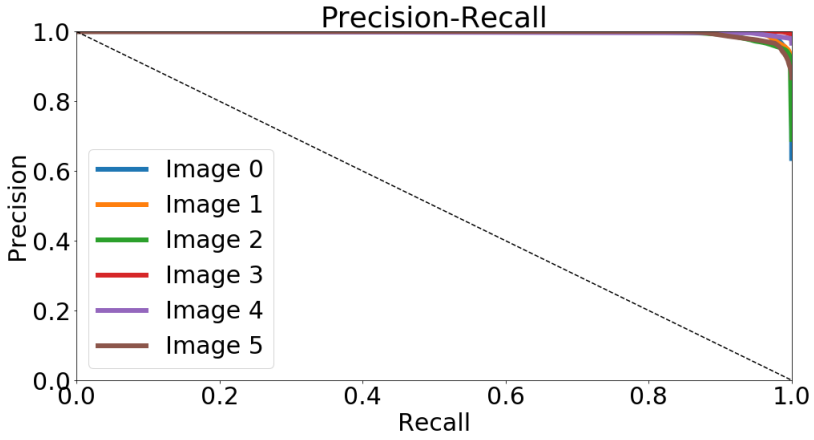

Fig. A.4: PRC. Training Dataset RDZ5000, 10 iterations. Model 8.

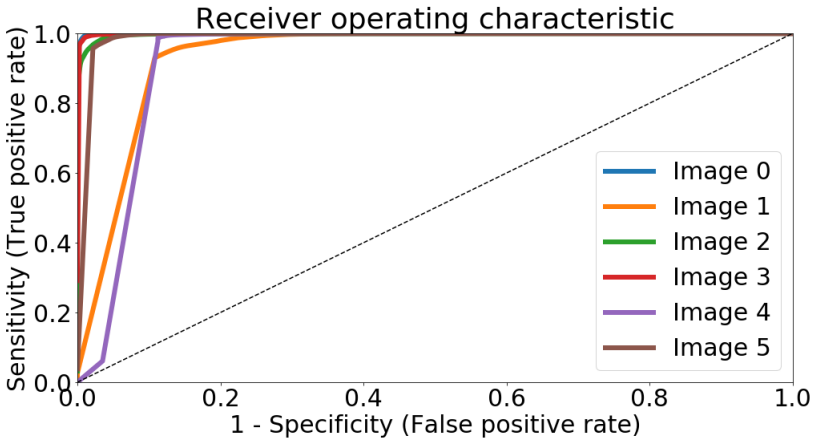

Fig. A.5: ROC. Training Dataset RDZ10000, 10 iterations. Model 5.

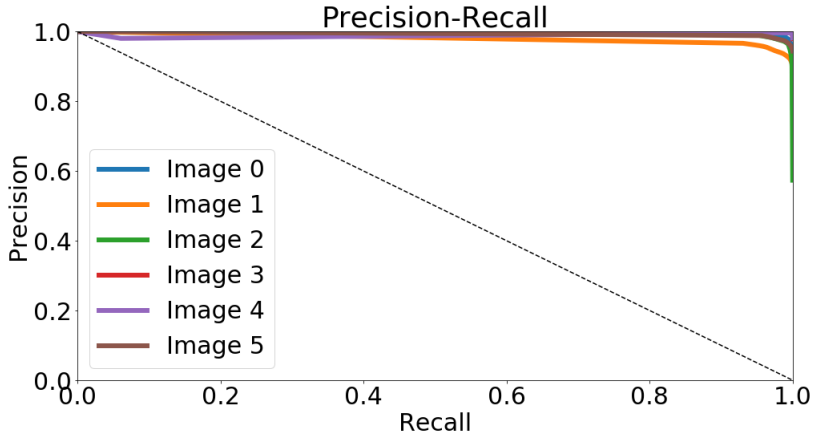

Fig. A.6: PRC. Training Dataset RDZ10000, 10 iterations. Model 5.

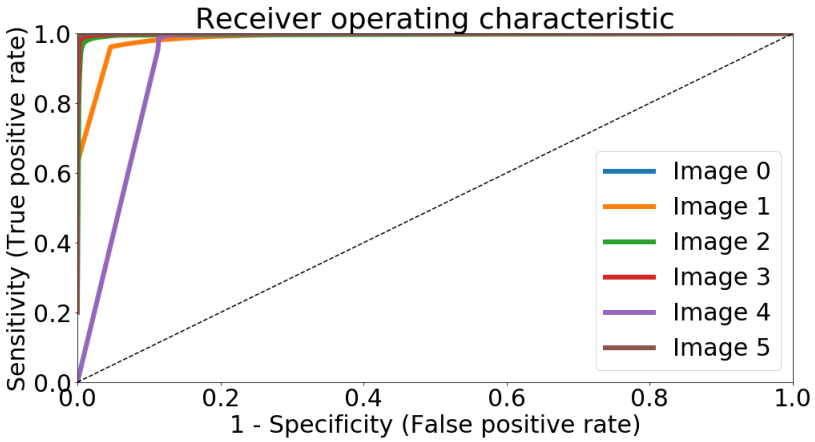

Fig. A.7: ROC. Training Dataset RDZ10000, 15 iterations. Model 4.

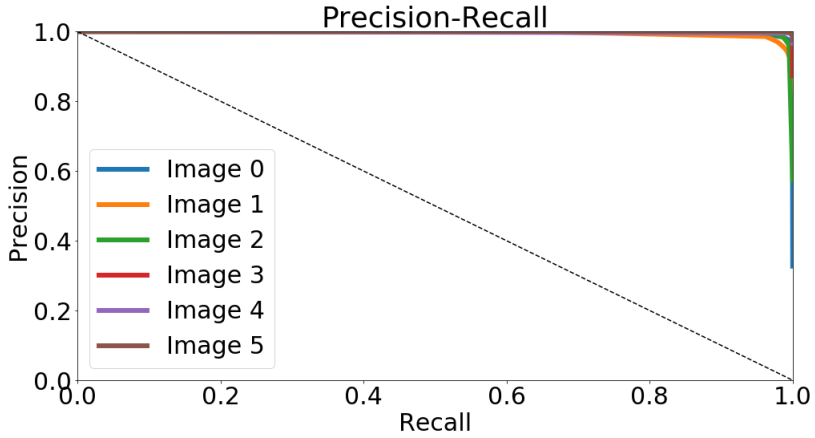

Fig. A.8: PRC. Training Dataset RDZ10000, 15 iterations. Model 4.

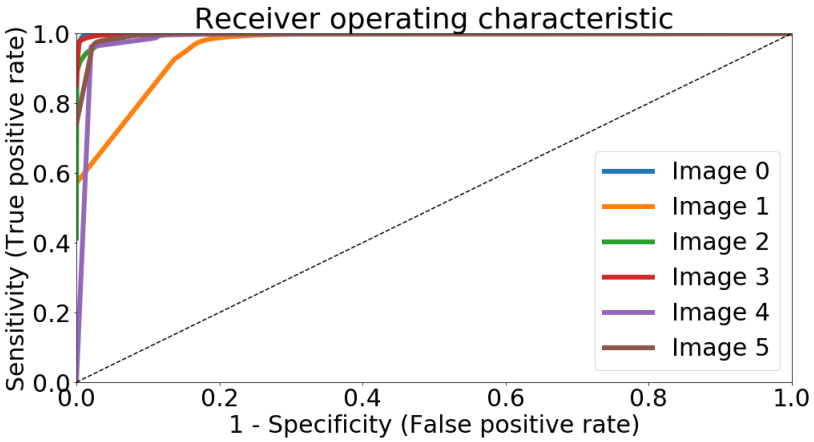

Fig. A.9: ROC. Training Dataset RDZ10000, 20 iterations. Model 1.

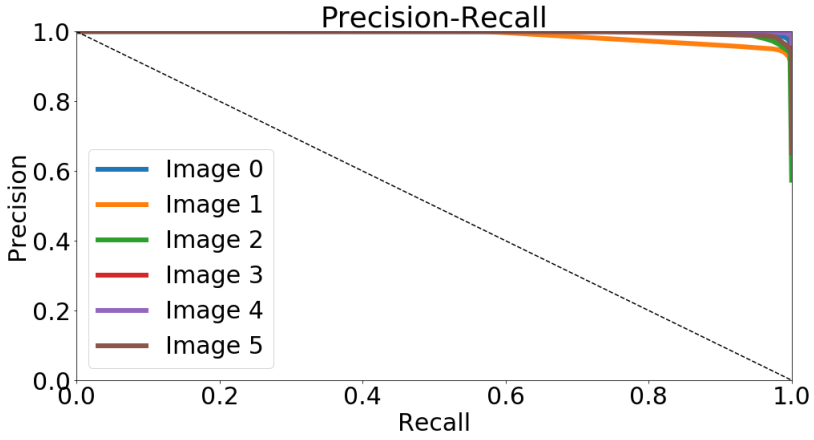

Fig. A.10: PRC. Training Dataset RDZ10000, 20 iterations. Model 1.

B NUMERICAL RESULTS

In this part of the apendix are DSC and ARI tables that represent the behaviour of the all computed models from each training category.

Tab. B.1: DSC - score and Adjusted Rand Index results for RDZ1K

Dataset	RDZ1K	
	DSC	ARI
1	0.94±0.14	0.25±0.51
2	0.53±0.68	0.04±0.11
3	0.72±0.44	0.01±0.04
4	0.56±0.57	0.05±0.16
5	0.76±0.43	0.18±0.46
6	0.47±0.55	0.05±0.14
7	0.63±0.58	0.02±0.16
8	0.72±0.50	0.06±0.17
9	0.41±0.36	0.14±0.35
10	0.55±0.62	0.17±0.45
$\overline{x} \pm 2\sigma$	0.63±0.59	0.10±0.34

Tab. B.2: DSC - score and Adjusted Rand Index results for RDZ5K

Dataset	RDZ5K	
	DSC	ARI
1	0.95±0.09	0.24±0.42
2	0.96±0.06	0.39±0.25
3	0.83±0.38	0.23±0.44
4	0.95±0.07	0.23±0.35
5	0.92±0.15	0.19±0.26
6	0.95±0.07	0.45±0.40
7	0.95±0.09	0.40±0.51
8	0.97±0.03	0.31±0.15
9	0.94±0.10	0.30±0.22
10	0.94±0.12	0.44±0.43
$\overline{x} \pm 2\sigma$	0.94±0.17	0.32±0.40

Tab. B.3: DSC - score and Adjusted Rand Index results for RDZ10K

Dataset	RDZ10KITER10	
	DSC	ARI
1	0.98±0.07	0.51± 0.21
2	0.98±0.03	0.42±0.18
3	0.97±0.03	0.42±0.21
4	0.96±0.10	0.43±0.25
5	0.98±0.03	0.54±0.29
6	0.98±0.04	0.50±0.29
7	0.97±0.04	0.50±0.35
8	0.98±0.03	0.44±0.44
9	0.95±0.08	0.21±0.34
10	0.97±0.05	0.37±0.22
$\overline{x} \pm 2\sigma$	0.97±0.06	0.43±0.34

Tab. B.4: DSC - score and Adjusted Rand Index results for RDZ10KITER15 and RDZ10KITER20

Dataset	RDZ10KITER15		RDZ10KITER20	
	DSC	ARI	DSC	ARI
1	0.97±0.08	0.45±0.31	0.98±0.03	0.56±0.23
2	0.98±0.02	0.65±0.44	0.98±0.03	0.46±0.29
3	0.97±0.04	0.38±0.38	0.98±0.02	0.53±0.46
4	0.99±0.02	0.54±0.27	0.97±0.08	0.48±0.15
5	0.98±0.02	0.48±0.42	0.99±0.01	0.56±0.24
$\overline{x} \pm 2\sigma$	0.98±0.05	0.50±0.41	0.98±0.04	0.52±0.30