Bachelor Project

Czech Technical University in Prague

F3

Faculty of Electrical Engineering
Department of Cybernetics

Prediction of Atherosclerotic Plaque Parameters from In-Vivo Ultrasound Carotid Artery Images

Artem Moroz

Supervisor: Prof. Dr. Ing. Jan Kybic
Field of study: Cybernetics and robotics
May 2022

BACHELOR‘S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Moroz Artem** Personal ID number: **492340**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Cybernetics**

Study program: **Cybernetics and Robotics**

II. Bachelor’s thesis details

Bachelor’s thesis title in English:

Prediction of Atherosclerotic Plaque Parameters from In-Vivo Ultrasound Carotid Artery Images

Bachelor’s thesis title in Czech:

Predikce parametrů aterosklerotického plátu z in-vivo ultrazvukových obrázků karotidy

Guidelines:

The goal of this thesis is to develop an automatic procedure of determining ultrasonographic parameters of a carotid plaque from in-vivo ultrasound carotid artery images. Such parameters include the plaque width, degree of stenosis, smoothness, homogeneity, echogenicity and others. Expert annotations are available for training on a per patient-level.
1. Get familiar with the available data and existing methods for carotid artery localization and segmentation.
2. Test existing as well as new analytical methods to evaluate the desired parameters from both automatic and manual segmentations. Attempt to identify the key images for each patient.
3. Design, implement and test regression CNN methods to estimate the desired parameters from both automatic and manual segmentations. Suggest a method to handle per patient annotations by predicting the relevance of each image.
4. Extend the CNN architecture to accept the input images as additional inputs. Test several variants and compare their performance.
5. Attempt to predict the desired parameters directly from images, without using the segmentations.
6. [optional] Predict the long term plaque stability based on derived parameters and other known clininical parameters.

Bibliography / sources:

[1] Kostelanský M., Manzano-Rodríguez A., Kybic J., Hekrdla M., Dvorský O., Kozel J., Baurová P., Pakizer D. and Školoudík D.. "Differentiating between stable and progressive carotid atherosclerotic plaques from in-vivo ultrasound images using texture descriptors." SIPAIM: International Symposium on Medical Information Processing and Analysis, 2021.
[2] Kostelanský M.. "Localization and segmentation of in-vivo ultrasound carotid artery images.", master thesis, CTU 2021.
[3] Lekadir-Galimzianova-etal_2017, "A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound", IEEE journal of biomedical and health informatics, 2017.
[4] Sampa Misra , Seungwan Jeon , Ravi Managuli , Seiyon Lee , Gyuwon Kim , Seungchul Lee , Richard G Barr , and Chulhong Kim. "Ensemble Transfer Learning of Elastography and B-mode Breast Ultrasound Images", 2021.
[5] Arijit Das, Srinibas Rana. "Exploring Residual Networks for Breast Cancer Detection from Ultrasound Images", 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021
[6] Philipp Fischer, Alexey Dosovitskiy , Eddy Ilg, Philip Hausser, Caner Hazırbaş¸, Vladimir Golkov. "FlowNet: Learning Optical Flow with Convolutional Networks", IEEE International Conference on Computer Vision, 2015.
[7] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, Quoc V. Le. "AutoAugment: Learning Augmentation Strategies from Data", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

Name and workplace of bachelor's thesis supervisor:

prof. Dr. Ing. Jan Kybic Biomedical imaging algorithms FEE

Name and workplace of second bachelor's thesis supervisor or consultant:

Date of bachelor's thesis assignment: **10.02.2022** Deadline for bachelor thesis submission: **20.05.2022**

Assignment valid until: **30.09.2023**

prof. Dr. Ing. Jan Kybic	prof. Ing. Tomáš Svoboda, Ph.D.	prof. Mgr. Petr Páta, Ph.D.
Supervisor's signature	Head of department's signature	Dean's signature

III. Assignment receipt

The student acknowledges that the bachelor's thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the bachelor's thesis, the author must state the names of consultants and include a list of references.

.	
Date of assignment receipt	Student's signature

Acknowledgements

Declaration

I declare that the presented work was developed independently and that I have listed all sources of information used within it in accordance with the methodical instructions for observing the ethical principles in the preparation of university theses. Prague, 20 May 2022

Abstract

The main goal of my bachelor thesis is to predict atherosclerotic plaque parameters such as echogenicity, homogeneity and degree of stenosis from in-vivo ultrasound carotid artery images using deep-learning methods. ResNet 34 and DenseNet 161 models were proposed as deep learning architectures. A novel context restoration strategy was successfully implemented in order to obtain good initialization parameters for target task models. Three architectures able to conduct fusion of ultrasound image and corresponding segmentation were designed. The best performance for echogenicity and homogeneity prediction has been achieved using ultrasound and segmentation images fusion and are equal to 75 % and 87 % correlation respectively.

Keywords: Deep learning, dataset, ultrasound image, atherosclerosis, atherosclerotic plaque, degree of stenosis, echogenicity, homogeneity

Supervisor: Prof. Dr. Ing. Jan Kybic

Abstrakt

Hlavním cílem moje bakalářské práce je predikce parametrů aterosklerotického plátu jako jsou echogenita, homogenita a procento stenózy z in-vivo ultrazvukových obrázků karotidy pomocí metod hlubokého učení. ResNet 34 a DenseNet 161 architektury byly navrženy. Inovační strategie Context restoration byla implementována za účelem získání vhodných parametrů pro inicializaci modelů. Byly implementovány tři architektury schopné provádět fúzi ultrazvukového obrázku a odpovídající segmentace. Nejlepší výsledky predikce echogenity a homogenity byly dosaženy metodou fúze ultrazvukového obrázku a segmentace a jsou rovny 75% a 87% korelace.

Klíčová slova: Hluboké učení, dataset, ultrazvukový obrázek, ateroskleróza, aterosklerotický plát, procento stenózy, echogenita, homogenita

Překlad názvu: Predikce parametrů aterosklerotického plátu z in-vivo ultrazvukových obrázků karotidy

Contents

Figures

Tables

BACHELOR‘S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Moroz Artem** Personal ID number: **492340**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Cybernetics**

Study program: **Cybernetics and Robotics**

II. Bachelor’s thesis details

Bachelor’s thesis title in English:

Prediction of Atherosclerotic Plaque Parameters from In-Vivo Ultrasound Carotid Artery Images

Bachelor’s thesis title in Czech:

Predikce parametrů aterosklerotického plátu z in-vivo ultrazvukových obrázků karotidy

Guidelines:

The goal of this thesis is to develop an automatic procedure of determining ultrasonographic parameters of a carotid plaque from in-vivo ultrasound carotid artery images. Such parameters include the plaque width, degree of stenosis, smoothness, homogeneity, echogenicity and others. Expert annotations are available for training on a per patient-level.
1. Get familiar with the available data and existing methods for carotid artery localization and segmentation.
2. Test existing as well as new analytical methods to evaluate the desired parameters from both automatic and manual segmentations. Attempt to identify the key images for each patient.
3. Design, implement and test regression CNN methods to estimate the desired parameters from both automatic and manual segmentations. Suggest a method to handle per patient annotations by predicting the relevance of each image.
4. Extend the CNN architecture to accept the input images as additional inputs. Test several variants and compare their performance.
5. Attempt to predict the desired parameters directly from images, without using the segmentations.
6. [optional] Predict the long term plaque stability based on derived parameters and other known clininical parameters.

Bibliography / sources:

[1] Kostelanský M., Manzano-Rodríguez A., Kybic J., Hekrdla M., Dvorský O., Kozel J., Baurová P., Pakizer D. and Školoudík D.. "Differentiating between stable and progressive carotid atherosclerotic plaques from in-vivo ultrasound images using texture descriptors." SIPAIM: International Symposium on Medical Information Processing and Analysis, 2021.
[2] Kostelanský M.. "Localization and segmentation of in-vivo ultrasound carotid artery images.", master thesis, CTU 2021.
[3] Lekadir-Galimzianova-etal_2017, "A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound", IEEE journal of biomedical and health informatics, 2017.
[4] Sampa Misra , Seungwan Jeon , Ravi Managuli , Seiyon Lee , Gyuwon Kim , Seungchul Lee , Richard G Barr , and Chulhong Kim. "Ensemble Transfer Learning of Elastography and B-mode Breast Ultrasound Images", 2021.
[5] Arijit Das, Srinibas Rana. "Exploring Residual Networks for Breast Cancer Detection from Ultrasound Images", 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021
[6] Philipp Fischer, Alexey Dosovitskiy , Eddy Ilg, Philip Hausser, Caner Hazırbaş¸, Vladimir Golkov. "FlowNet: Learning Optical Flow with Convolutional Networks", IEEE International Conference on Computer Vision, 2015.
[7] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, Quoc V. Le. "AutoAugment: Learning Augmentation Strategies from Data", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

Name and workplace of bachelor's thesis supervisor:

prof. Dr. Ing. Jan Kybic Biomedical imaging algorithms FEE

Name and workplace of second bachelor's thesis supervisor or consultant:

Date of bachelor's thesis assignment: **10.02.2022** Deadline for bachelor thesis submission: **20.05.2022**

Assignment valid until: **30.09.2023**

prof. Dr. Ing. Jan Kybic	prof. Ing. Tomáš Svoboda, Ph.D.	prof. Mgr. Petr Páta, Ph.D.
Supervisor's signature	Head of department's signature	Dean's signature

III. Assignment receipt

The student acknowledges that the bachelor's thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the bachelor's thesis, the author must state the names of consultants and include a list of references.

.	
Date of assignment receipt	Student's signature

Chapter 1

Introduction

Atherosclerosis is a very common disease among people over the age of 60 or among people who do not lead a healthy lifestyle. [35] This illness occurs when cholesterol and other fatty products begin accumulating on inner walls of arteries consequently reducing the lumen area and hindering the blood flow. Severe narrowing of an artery lumen blocks the oxygen supply to organs and tissues, increases the probability of cerebrovascular and cardiovascular events and may even lead to lethal outcome. Signs of atherosclerosis develop slowly and gradually so patient may not observe any noticeable symptoms till the artery is heavily narrowed and blood flow is restricted. This fact insists on the importance of early detection, diagnosing and treatment of this disease.

Deep learning is a subfield of Artificial intelligence. Deep learning works on designing, developing and improving models inspired by the neural networks in human brain and are able to imitate human decision process. Such models have become widely used thanks to a large amount of data and rapid development of computational power. Application of deep learning techniques in medicine is a novel approach intending to automatize disease development prediction, grading of illness severeness and characterization of tissues composition.

This project aims to apply deep learning methods in order to predict atherosclerotic plaque parameters using in-vivo ultrasound carotid artery images and their semantic segmentations prepared by Martin Kostelanský in his Master degree thesis "Localization and segmentation of in-vivo ultrasound carotid artery images". Experts annotation including atherosclerotic plaque parameters grading is available for such parameters as echogenicity index, homogeneity of an atherosclerotic plaque, percentage of calcification, atherosclerotic plaque width, type of plaque surface and degree of stenosis.

1.1 Goals of the project

The intention of this thesis work is to propose a solution for atherosclerotic plaque parameters prediction from in-vivo ultrasound carotid artery images. In line with the proposed implementation the whole stated problem divides into following intermediate tasks:

- Firstly, we will get familiar with the available dataset and the corresponding experts annotation, describe atherosclerotic plaque parameters and find mathematical interpretation for each of them;
- Secondly, we will propose some methods for images processing;
- Thirdly, we will implement various deep learning techniques including novel pre-training method and models for image fusion;
- Lastly, we will conduct experiments and make quantitative evaluation of the achieved performance.

1.2 Structure of the document

My bachelor thesis has in total 6 chapters and is structured in the next way: The first chapter makes a concise overview of studied problematics, specifies the main goals of the project and outlines the documents structure. The second chapter provides the necessary understanding of carotid artery anatomy, closely related disease called atherosclerosis, its diagnosing and treatment. In chapter number three the target dataset and several atherosclerotic plaque parameters are presented. The chapter number four is going to describe methods applied for data processing and explains deep learning techniques implemented for problem resolving. In chapter number five results of the experiments are presented and discussed. In the last chapter I am going to draw a conclusion about work done.

Chapter 2

Background

This chapter briefly explains and outlines the general medical background of my thesis project. The section 2.1 explains the structure and significance of the carotid artery. The section 2.2 offers insight into the most widespread disease that is connected with carotid artery. The 2.3 sections main purpose is to make us familiar with the most common diagnostic procedures and investigations. The 2.4 section intends to describe the main treatment techniques and preventative recommendations.

2.1 The carotid artery structure

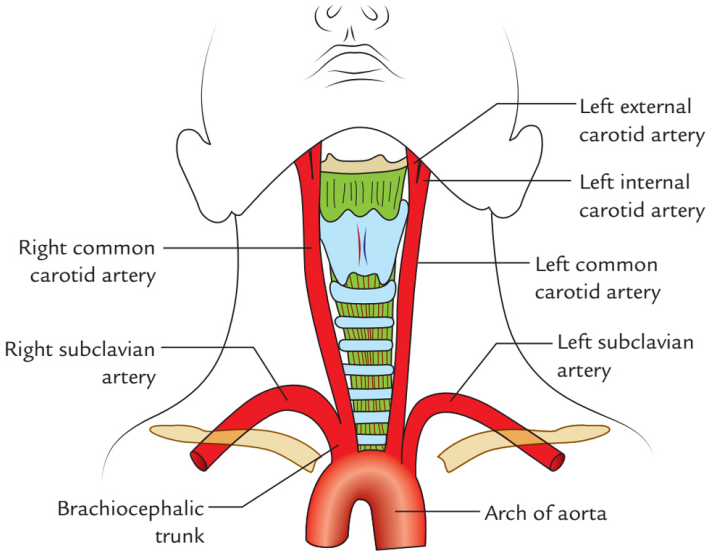


Figure 2.1: Carotid artery anatomy

The carotid arteries are the main blood vessels that carry blood and oxygen to the head. The are two Common Carotid Arteries (CCA) located in the neck - symmetrically one on either side. The right CCA arises from the brachiocephalic artery while the left CCA takes its origin directly off the

aorta. Both carotid arteries splits into two parts:

- Internal Carotid Artery (ICA) is the terminal branch of the Common Carotid Artery (CCA) that carries blood and oxygen to the brain, eyes, scalp, scull and the mennings;
- External Carotid Artery (ECA) divides into seven branches that carry blood and oxygen to the face and neck.

Carotid artery walls, similarly to all other arteries in our body, are composed of three layered structures:

- Intima, a smooth inner layer that allows blood to flow easily;
- Media, a muscular middle layer that controls the diameter of the artery;
- Adventitia, the protective outer layer of the artery.

2.2 Atherosclerosis

Atherosclerosis is a slow-progressive disease which can be explained as narrowing, stiffening and hardening of the arteries that is caused by a buildup of plaques on the inner walls of the arteries. Plaques consist of deposits of cholesterol, calcium, fibrin, cellular waste products and fatty products. Atherosclerosis causes restrictions of the flow of oxygen-rich blood to different parts of the body [36].

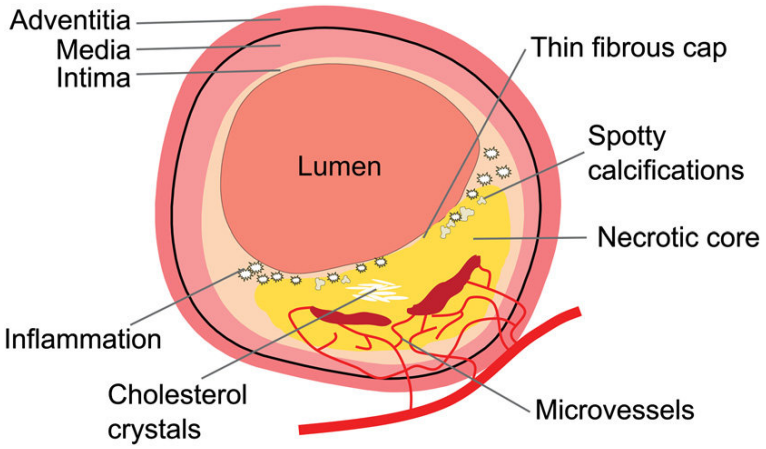


Figure 2.2: Atherosclerotic plaque composition

The exact cause of atherosclerosis is not well explored but it is assumed that any damage or injury of intima layer of an artery may instigate this disease. In case when the inner wall of an artery is injured blood cells and other constituents may gather at the injury place over time fatty deposits build up and harden so this process results in lumen area restriction. The initial damage of the inner layer of an artery may become an outcome of high

blood pressure, smoking, obesity or diabetes, high concentration of cholesterol or other lipid constituents in blood [34].

Signs of atherosclerosis usually progress gradually and patient would not feel any symptoms till blood flow isn't highly restricted because of narrowing or clogging of artery what consequently leads to situation when artery can not supply enough oxygen to tissues and organs. However, in case when blood flow through major artery is restricted symptoms may become severe. So the signs and symptoms depend on affected artery. Atherosclerotic plaques in heart arteries may lead to chest pressure and pain. Restricted blood flow through carotid arteries may make patient feel weakness or numbness in limbs, difficulties in speaking, temporary loss of vision or it can even progress to a stroke. Atherosclerosis in kidney arteries causes high blood pressure or even kidney failure. Affected arteries in limbs may develop decreased blood pressure or pain in affected leg or arm. Also this disease leads to emerging of aneurysms anywhere in body.

There are several major and most widespread risk factors that can cause atherosclerosis disease or support its progress with very high probability: high blood pressure, smoking, stress, physical inactivity, obesity, unhealthy and fat saturated diet, diabetes, high cholesterol and triglyceride level.

2.3 Diagnosis

The diagnosis of atherosclerosis disease usually starts with a physical test and continue with one or several next investigations:

- Blood test is a simple test which can reveal the risk of having atherosclerosis, like increased level of lipids like cholesterol or blood sugar;
- Ultrasound (including Doppler sonography) is the procedure when doctor uses a special probe to direct sound waves into a blood vessel to evaluate blood flow. Faintness of sound may mean there is narrowing or blockage;
- Computerized tomography angiography (CTA) is a type of X-ray investigation that can reveal calcifications;
- Magnetic resonance angiography (MRA) is the test that is quite similar to CTA which does not use ionizing radiation;
- Blood pressure comparison (Ankle-brachial index) is the test when specialist compares blood pressure in ankles and in arms. Significant difference may acknowledge existence of narrowings because of atherosclerosis;
- Cerebral angiography is the procedure in which doctor puts dye into arteries and they become visible on an X-ray. It helps to spot and localize arterial narrowings;
- Stress test is the survey when patient is subjected to exercising while doctors monitor patients heart rate, blood pressure, breathing and other health parameters.

2.4 Prevention and treatment

Atherosclerosis disease is prevented or delayed by reducing influence of the most common risk factors. Lifestyle changes like health diet, quit smoking, physical activity lead to reducing the risk. Patients that are at risk because of high cholesterol or high blood pressure should take medicines which help to slow down or even halt atherosclerosis progress. [34]

More serious cases of this disease require invasive techniques to get rid of blockages:

- Stenting and angioplasty is the surgery when catheter with a balloon tip is injected into an artery and balloon tip is inflated to clear the artery, insert a stent and hold it open
- Bypass surgery is the procedure when doctor uses a healthy blood vessel from another part of patients body to go around blocked segment;
- Endarterectomy is the surgery when doctor removes plaque with any injured part of artery to ease blood flow;
- Fibrinolytic therapy - some medicines are used to dissolve blockages in arteries [33].

Chapter 3

Dataset and atherosclerotic plaque parameters

Chapter number 3 intends to present the target datasets and the atherosclerotic plaque parameters which are going to be predicted in further experiments. In section 3.1 we are going to get familiar with the available datasets that consist of in-vivo ultrasound images of the carotid artery and prepared automatic segmentation of ultrasound images based on manual segmentation. Section number 3.2 presents experts annotation for target dataset. Further in section 3.3 we will specify and make a brief description of several atherosclerotic plaque parameters in accordance with its medical definitions.

Machine learning is a fast emerging field of Artificial intelligence. The main reason of machine learnings rapid development is a large amount of available annotated datasets, which become an essential resource to build smart systems, required for an improvement of machine learning models performance and are suitable for variety of tasks. However the situation is contrary when we take into consideration the application of machine learning techniques on medical problems and tasks. Collecting a qualitative large medical dataset is challenging because of 2 main reasons: firstly, patient health data are mostly not publicly available and are well protected by patient data laws, secondly, annotation and labelling of such datasets require expert knowledge compare to annotating natural images, when just basic human experience is enough.

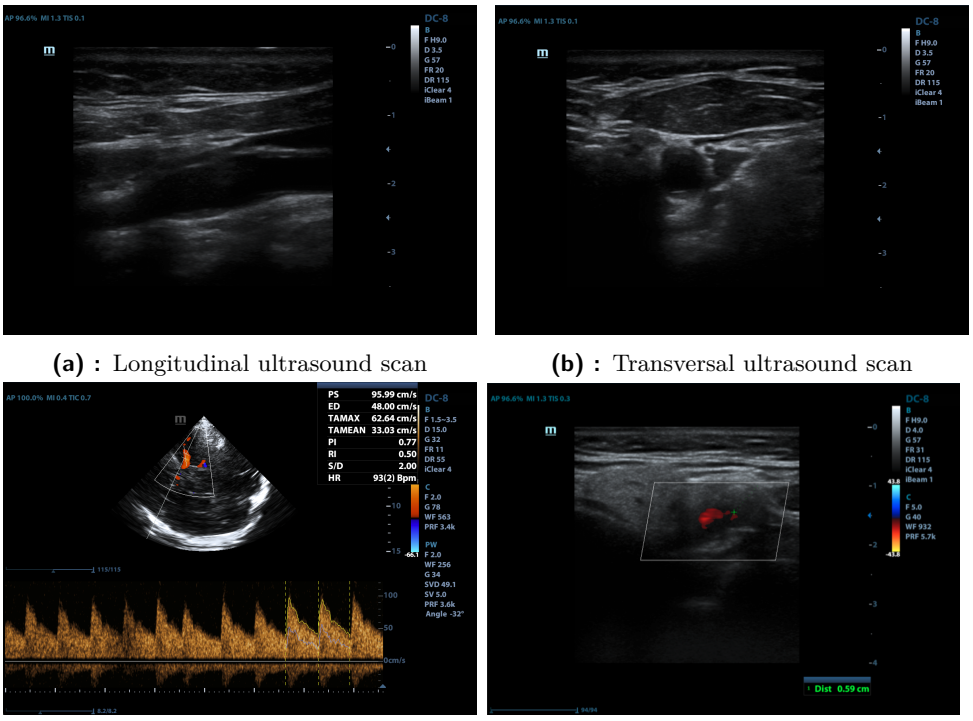
The concept "labelled dataset" can be explained as a collection of pairs of observations X_i and tag Y_i. Each such pair observation-label is considered as a single unit by machine learning model.

3.1 Dataset description

The ANTIQUE dataset was collected during the research project called "Atherosclerotic Plaque Characterisation Associated With a Progression Rate of the Plaque in Carotids and a Risk of Stroke" [22]. The main goal of this study was to examine dependence between risk factors such as greater plaque thickness, coronary heart disease, smoking etc. and progression of carotid

plaque thickness of patients after the preventive treatment was introduced [22]. Consequently key factors were specified.

The project development was taken 54 month in period between 2015 and 2020 [22][2]. The investigation has been conducted in sonographic laboratories of the Military University Hospital in Prague and University Hospital Ostrava. In total 1591 patients were screened and 1391 were included to the further analysis. Only patients without serious diseases with a high probability of fatal outcome and with diagnosed degree of stenosis over 30% have been included into the research. Overall 8 duplex ultrasound measurements were performed for each patient within 36 months and in total 28,178 ultrasound scans were collected. [22] Ultrasound scans are divided into four main types: Longitudinal, Transversal, Conical and Doppler. Examples of each type can be seen in Figure 3.1. From obtained scans the next plaque characteristics were evaluated: echogenicity index of the plaque, percentage of calcification, type of plaque surface and homogeneity [22]. Dataset has no supplementary annotation of the ultrasound probe orientation, artery location etc. Also raw ultrasound scans figure not only an artery location but redundant surrounding tissues and some information from scanner [2].



(a) : Longitudinal ultrasound scan

(b) : Transversal ultrasound scan

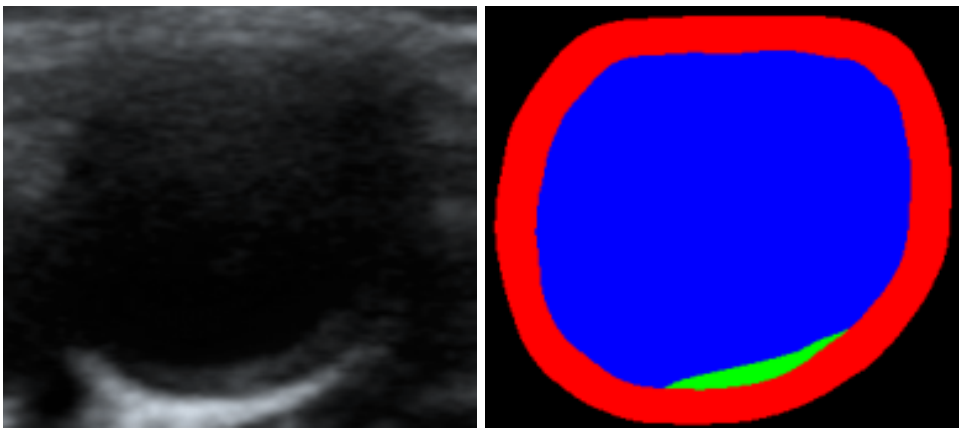
(c) : Conical ultrasound scan

(d) : Doppler ultrasound scan

Figure 3.1: Demonstration of ANTIQUE dataset examples

In my bachelor thesis I am working with two datasets each of them was derived from the ANTIQUE dataset. The first dataset consists of ultrasound images where only a carotid artery location is represented and other irrelevant content is cropped. The second dataset, prepared by M.Kostelanský in his

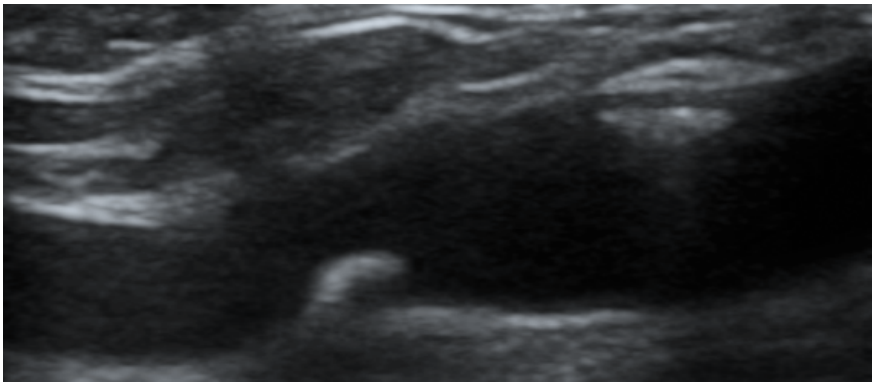
master thesis work, collects the semantic segmentation images of a localized artery. Semantic segmentation is a technique which assigns a label to each pixel of the image, so that each pixel belongs to the class which encloses the same or similar objects and regions. This technique brings significant benefits in our task that requires distinguishing sections of an artery. Having a prepared semantic segmentation image any person without medical experience and knowledge is possible to observe an atherosclerotic plaque position and size relatively to a carotid artery. In addition images in both datasets are divided into Longitudinal and Transversal accordingly to ultrasound probe orientation. Ultrasound images with two types of orientation and their segmentations we can see in the Figures 3.2 and 3.3. Red pixels represent an artery walls, blue pixels represent an artery lumen and green pixels represent an area which is blocked by atherosclerotic plaque. Cropped ultrasound scans and segmentation both types of orientation will be used in the next part of this work for prediction of atherosclerotic plaque parameters.



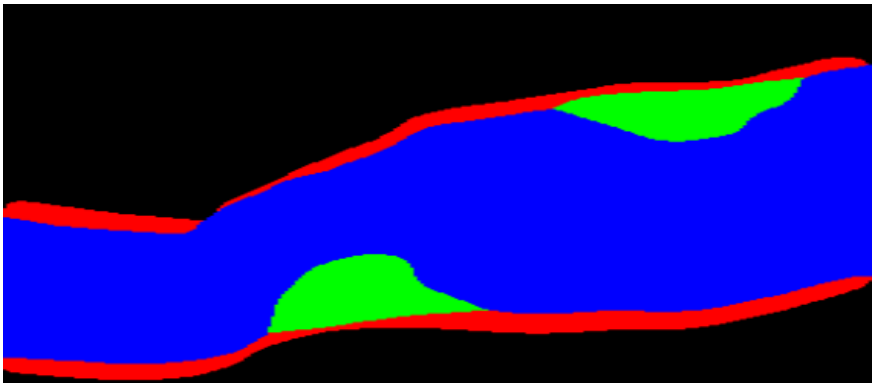
(a) : Ultrasound image with transversal orientation

(b) : Segmentation image with transversal orientation

Figure 3.2: Transversal ultrasound images and corresponding segmentation



(a) : Ultrasound image with longitudinal orientation



(b) : Segmentation image with longitudinal orientation

Figure 3.3: Ultrasound images and corresponding segmentations

3.2 Experts annotation

The document *data_with_attributes.csv* is available in folder */datagrid/Medical/ArteryPlaque/in_vivo_platy_progrese_*20191015*_anonym/* *data_description* provides us with all essential information about data which will be used for our research. The table contains such valuable content as grouping all images which belonging to one particular patient, longitudinal or transversal orientation tag, left or right label describing a carotid position, path to original ultrasound scans taken for the purpose of ANTIQUE research, path to cropped ultrasound images and grading of atherosclerotic plaque parameters such as echogenicity index, homogeneity, surface type, and calcification percentage provided by experts. Information is available in total for 25765 images which belong to 1321 patients. Unfortunately carotid artery localization was not prepared for group of images. Also experts annotations are not available for subset of patients.

An important notice for our further work with data is that the experts annotation varies for different images belonging to each individual patient depending on the left or right side of a carotid position. It means that scans taken from left and right side must be dealt separately with distinct labels although they still belong to one patient.

3.3 Atherosclerotic plaque parameters

Atherosclerotic plaque parameters are in general medical biomarkers which characterize a plaque composition. In the next three paragraphs we are going to describe three such plaque parameters, which can be predicted based on our datasets and annotations using deep learning methods.

3.3.1 Percentage of stenosis (degree of stenosis)

Stenosis percentage is a biomarker, which numerically characterizes a vessel narrowing. The formulas for degree of stenosis are defined separately for transversal and longitudinal images because the visibility of plaque differs depending on orientation.

For the transversal orientation this parameter can be calculated as follows:

$$S = \frac{P}{L + P} \tag{3.1}$$

where S is the degree of stenosis, L is the lumen area and P is the plaque area [3]. An example is shown in the Figure 3.4, where the degree of stenosis is computed per pixel and equals to 12.3 %.

For the longitudinal orientation percentage of stenosis is estimated according to the following formula:

$$S = \frac{H_2^2}{H_1^2} \tag{3.2}$$

where H_2 is the maximum width of atherosclerotic plaque, H_1 is the maximum width of a vessel as shown in the Figure 3.5.

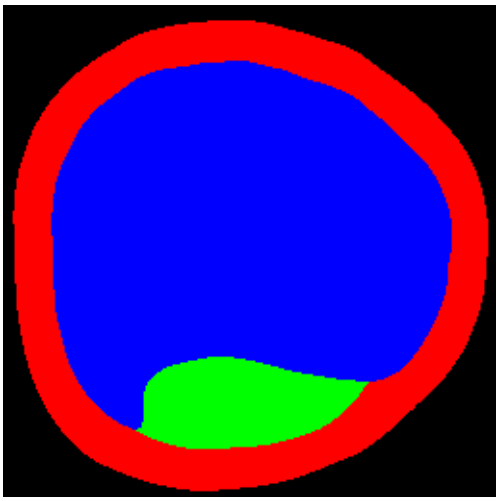


Figure 3.4: Degree of stenosis demonstration on transversal segmentation

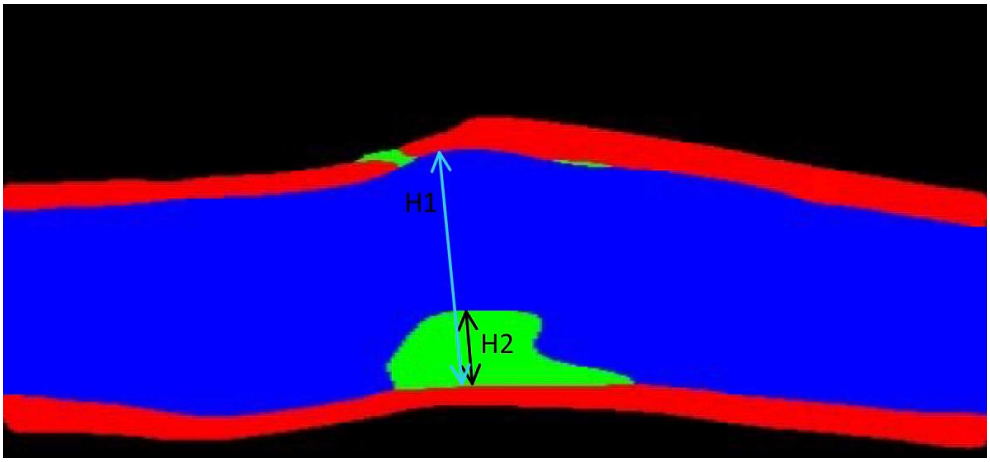
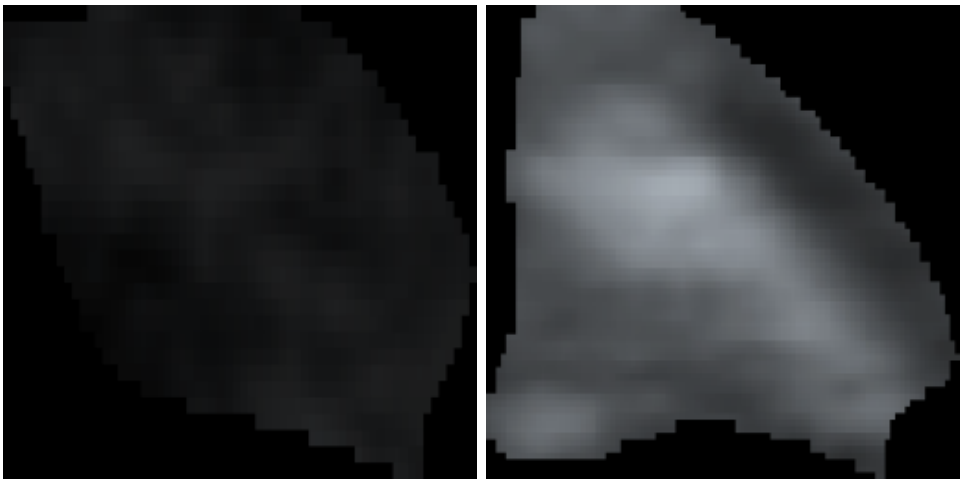


Figure 3.5: Degree of stenosis demonstration on longitudinal segmentation

3.3.2 Echogenicity index

Echogenicity is a measure that expresses the ability of different tissues to reflect ultrasound waves. This parameter is intimately related with acoustic resistance of various types of tissues e.g. bones have higher acoustic resistance and so their echogenicity is high accordingly. Medical terminology exploits four terms to grade an echogenicity of tissue: Anechoic (the lowest rate of tissue echogenicity), Hypoechoic, Isoechoic and Hyperechoic (the highest rate of tissue echogenicity). For simplicity in the remaining part of our work only number grading is used, where 1 - Anechoic, 2 - Hypoechoic, 3 - Isoechoic and 4 - Hyperechoic [46][47]. Two examples of atherosclerotic plaque segments with different echogenicity grading can be seen in Figure 3.6.



(a) : Atherosclerotic plaque segment with echogenicity equals to 1

(b) : Atherosclerotic plaque segment with echogenicity equals to 4

Figure 3.6: Atherosclerotic plaque segments with different grade of echogenicity

Mean pixel intensity over the entire atherosclerotic plaque was proposed

as statistic in order to obtain mathematical interpretation and consequently quantitive grading of echogenicity :

$$E = \frac{1}{N}\sum_{j=1}^{N} I_j \tag{3.3}$$

where

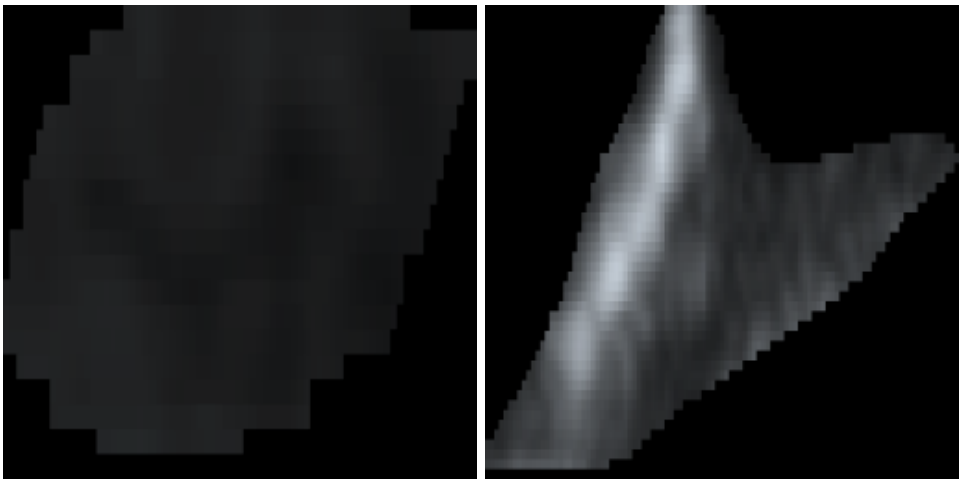
$$I_j = \frac{R_j + G_j + B_j}{3} \tag{3.4}$$

is the intensity of pixel and R_j, G_j, B_j are red, green and blue primary colors of pixel indexed with j, N is the total number of pixels in atherosclerotic plaque.

Mean pixel intensity is expected to be relatively high for a plaque with the index of echogenicity equals to 4 and relatively low for a plaque with the index of echogenicity equals to 1.

3.3.3 Homogeneity

Homogeneity is a parameter which characterizes uniformity and consistency of tissue among its different parts. In experts annotation a simple grading is used: 1 means that plaque is considered as homogenous and 2 means that plaque is considered as unhomogenous. Two images representing different homogeneity are demonstrated in Figure 3.7.



(a) : Atherosclerotic plaque segment with homogeneity equals to 1

(b) : Atherosclerotic plaque segment with homogeneity equals to 2

Figure 3.7: Atherosclerotic plaque segments with different grade of homogeneity

Sample standard deviation of pixel intensity for all pixels in atherosclerotic plaque is proposed as mathematical interpretation of homogeneity and the following formula is used for its computation:

$$S = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}(I_j - E)^2} \tag{3.5}$$

where I_j is pixel intensity computed as 3.4, E is mean pixel intensity for all pixels in atherosclerotic plaque computed as 3.3, N is the total number of pixels in atherosclerotic plaque.

The case when a value of the standard deviation is low indicates that samples tend to be closer to their average hence the distribution of samples is more homogenous. Vise versa high standard deviation corresponds to a wider range of values and unhomogenous distribution.

Chapter 4

Methodology

In this chapter we will describe the main data processing and deep learning techniques that were implemented and used in our work. Dataset preparation is explained in 4.1 section. In section 4.2 we will get familiar with transfer learning principles and weights initialization techniques. Section 4.3 explains the importance of data augmentation and describes the strategy that was chosen in purpose to expand and diversify available dataset. In section 4.4 we will discuss deep learning models that were implemented in order to predict desired atherosclerotic plaque parameters. Finally section 4.5 explains conception of correlation and provides us with several correlation metrics.

4.1 Data preprocessing

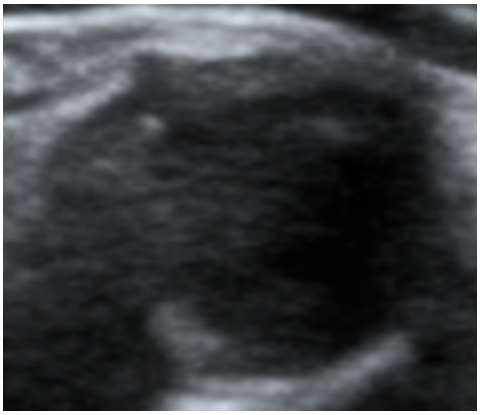
Data preprocessing is the technique that includes some steps which allow us to get data in more suitable form with respect to the particular application, extract additional helpful information and get rid of redundant, ineffective content.

4.1.1 Plaque localization in ultrasound image using corresponding segmentation

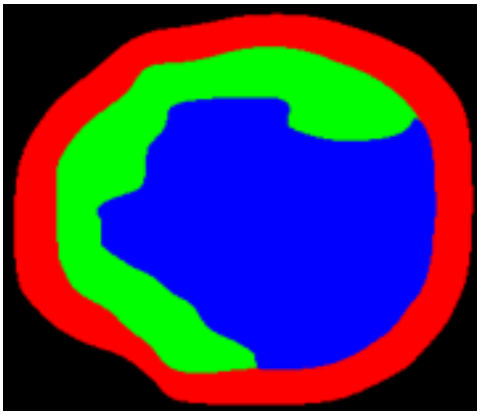
Based on the fact that relevant scope of information about atherosclerotic plaque echogenicity and homogeneity is concentrated only in plaques location, remaining context of ultrasound image can be considered as unnecessary for such parameters prediction and exceedingly confusing for deep learning architectures. As was mentioned in 3.1, segmentation divides image content into 3 groups, so that plaque positioning becomes obvious. Thereby segmentation can be leveraged with the purpose of atheroselerotic plaque location extraction from ultrasound image. The following procedure for plaque content separation was proposed:

- Take an ultrasound image and the corresponding segmentation;
- From ultrasound image extract only those pixels which accord with green area in segmentation;
- Get rid of the rest ultrasound image content.

After this procedure the dataset collecting images where only plaque segments are present was created. For better understanding and orienting in further experiments in the remaining part of this work an abbreviation PSD (Plaque Segments Dataset) will be used as reference to this dataset. Example of image from PSD can be seen in the Figure 4.1. Simultaneously with PSD generation also mean and standard deviation of pixel intensity for the entire plaque were calculated according to formulas 3.3 and 3.5.



(a) : Ultrasound image with transversal orientation



(b) : Segmentation image with transversal orientation

(c) : Atherosclerotic plaque location extracted from ultrasound image

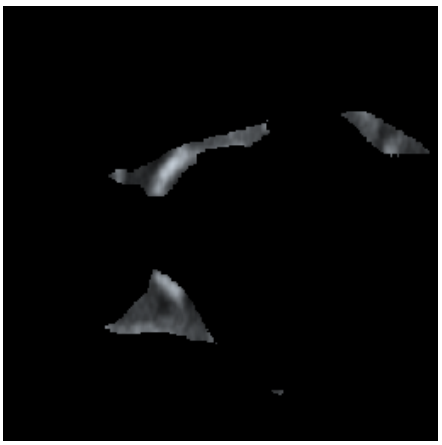
Figure 4.1: Atherosclerotic plaque localization

During the exploration of PSD was exposed a subset of images containing several plaque segments, so another one proposition for improvement of deep learning model performance was to use images where one and only one plaque segment is present. Clarification of this approach assumes that focusing on only one plaque segment may be more beneficial for convolutional neural network than taking into consideration several disjunct segments. Subsequently dataset collecting only images with single plaque segment was derived from PSD by selecting proper images, removing redundant black area

and resizing to the shape 256x256 pixels through the use of opencv-python library. We can see the detailed implementation in the following code:

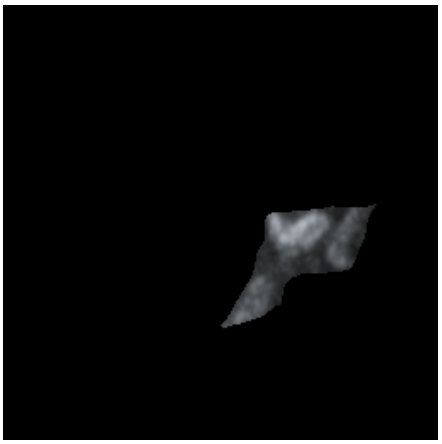
```
  gray    = cv2.cvtColor(plaque_img, cv2.COLOR_BGR2GRAY)
edged = cv2.Canny(gray, 30, 200)

contours, hierarchy = cv2.findContours(edged,
cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_NONE)

if len(contours) == 1:
        x, y, w, h = cv2.boundingRect(contours[0])
        single_plaque = plaque_img[y + 1:y + h, x + 1: x + w]
        reshaped = cv2.resize(single_plaque, 256, 256,
        interpolation=cv2.INTER_AREA)
```

Listing 4.1: Single plaque segment detection Python 3.7 implementation

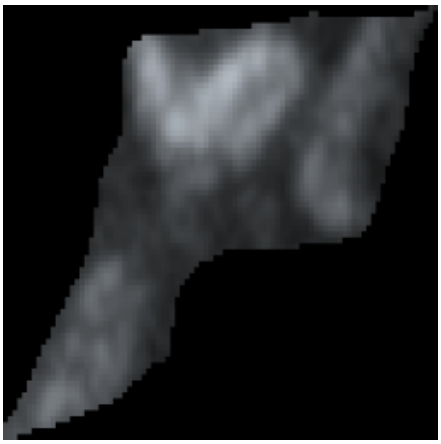


(a) : Multiple plaque segments example

(b) : Multiple plaque segments example



(c) : PSD example



(d) : Example of image from SPSD

Figure 4.2: Examples of images with multiple and single plaque segments

The abbreviation SPSD (Single Plaque Segment Dataset) was introduced as the name for the proposed dataset. PSD counts in total 11774 images. SPSD has totally 4673 images. Several examples of images from PSD dataset

that were not included into SPSD and as well an example of image from SPSD can be seen in the Figure 4.2.

4.1.2 Estimation of reliability of the experts annotation

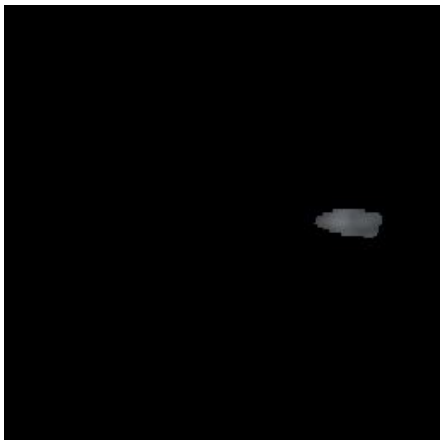
As was mentioned in 4.1 paragraph, for each pair ultrasound scan-segmentation mean value and sample standard deviation of pixel intensity over plaque area were computed and stored in file. Afterwards the detailed investigation was done.

It was revealed that a substantial part of images has identical experts annotation, although values of statistics 3.3 and 3.5 differs significantly. Several examples demonstrating inconsistency can be seen in the Figure 4.3 for echogenicity grading and in the Figure 4.4 for homogeneity grading.

(a) : Plaque with echogenicity annotation equals 4 and mean plaque intensity equals to 78.1

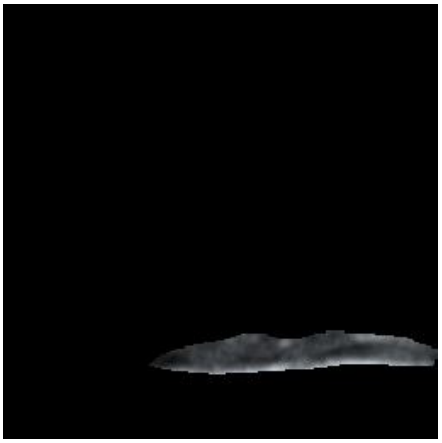
(b) : Plaque with echogenicity annotation equals 4 and mean plaque intensity equals to 32.3

(c) : Plaque with echogenicity annotation equals 1 and mean plaque intensity equals to 29.2



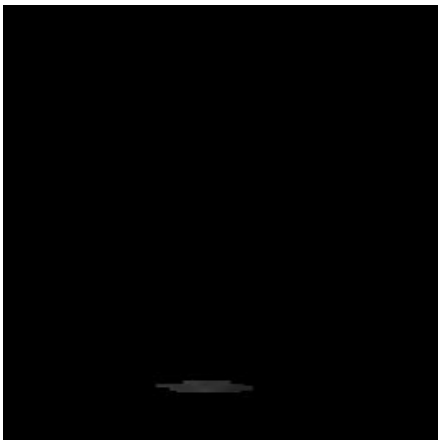
(d) : Plaque with echogenicity annotation equals 1 and mean plaque intensity equals to 73.8

Figure 4.3: Examples of inconsistency in echogenicity grading



(a) : Plaque with homogeneity annotation equals 2 and plaque intensity standard deviation equals to 38.0

(b) : Plaque with homogeneity annotation equals 2 and plaque intensity standard deviation equals to 7.9



(c) : Plaque with homogeneity annotation equals 1 and plaque intensity standard deviation equals to 6.4

(d) : Plaque with homogeneity annotation equals 1 and plaque intensity standard deviation equals to 39.4

Figure 4.4: Examples of inconsistency in homogeneity grading

Conducted analysis has shown that not all images can be selected for further training and testing of deep learning models. Consequently only a subset of them must be selected in order not to confuse the proposed algorithms by giving them inconsistent data samples. According to notice about different annotation for left and right side of carotid position, we will made an assumption that **at least one ultrasound scan from both left and right sides was reviewed and graded by an expert**.

Further two sections propose methods which at least partially help to tackle inconsistency in dataset for both echogenicity and homogeneity grading.

Reliability estimation for echogenicity grading

Algorithm for reliability testing for echogenicity annotation is based on the confrontation between computed mean plaque intensity per-pixel and

label assigned by expert. The whole method consists of several simple steps. Firstly, after applying the boxplot (Figure 4.5) was examined that the set of mean intensity values contains a big number of outliers, which can negatively affect further normalization. Deleting of outliers was chosen as the most simple method to handle them [39]. After outliers have been deleted, mean value computed through mean pixel intensity set has reduced from 56.435 to 51.173.

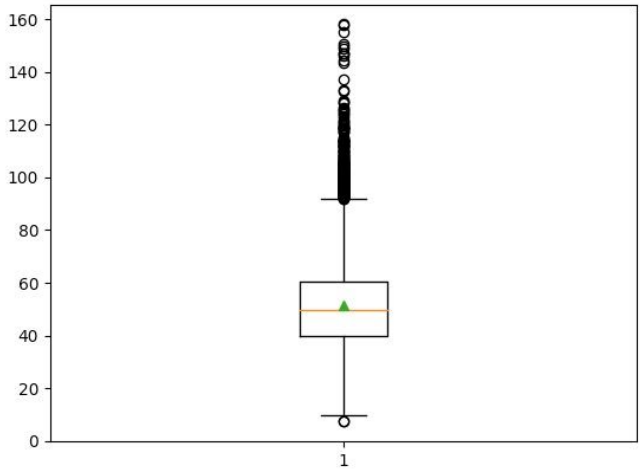


Figure 4.5: Boxplot for mean value of mean pixel intensity

Secondly, mini-max normalization was applied (Equation 4.1) in order to map a set of mean intensity values into [0, 1] range [40]. Thirdly, the [0, 1] interval was equidistantly spaced to four disjunct intervals [0, 0.25), [0.25, 0.5), [0.5, 0.75) and [0.75, 1) which correspond to echogenicity labels 1, 2, 3, 4.

$$\overline{x_i} = \frac{x_i - \min_{x_1,...,x_N}}{\max_{x_1,...,x_N} - \min_{x_1,...,x_N}} \tag{4.1}$$

where $x_1, ..., x_N$ are all data samples, $\overline{x_i}$ is the normalized data sample with index i, N is the total number of samples. Afterwards, for each patient and both left and right sides reliable images were chosen according to the simple rule: if normalized mean plaque intensity, computed for an image, falls into correspondent interval, then conforming image is considered as reliable, otherwise is excluded from further analysis. As we said, for each patient and for each side of carotid artery position at least one image must be selected. So, in case when no one image falls into corresponding interval, the closest distance between normalized mean intensity among all images and interval boundaries is computed and respective image is then taken as reliable. The next equation represents the rule according to which the only one proper image is selected.

$$reliable\ image = arg(\min_{\substack{i\in\{1,...,M\} \\ j\in[\text{"}left\text{"},\text{"}right\text{"}]}} (dist(E_{i,j}) - R_l)) \tag{4.2}$$

where j is "left" or "right" tag, M is the number of images either left or right side for one patient, $E_{i,j}$ is mean plaque intensity, computed as 3.3, R_l is an interval corresponding to echogenicity label $l = \{1, 2, 3, 4\}$.

Also code with implementation can be seen in the Listing 4.2.

```
rngLabel = {"1.0": [0, 0.25], "2.0": [0.25, 0.5], "3.0": [0.5, 0.75], "4.0": [0.75, 1]}
for pat in all_patients:
        patImgIntens = patientsImgIntensities[pat]
        for carotid_side in ["left", "right"]:
                reliabilityIdxEcho = list()
                echogenicityLabel = patSideEchoLabel[pat][carotid_side]:
                rng = rngLabel[echogenicityLabel]
                for img_intens in patImgIntens:
                        if img_intens >= rng[0] and img_intens < rng[1]:
                                reliabilityIdxEcho.append(True)
                        else:
                                reliabilityIdxEcho.append(False)
                if True not in reliabilityIdxEcho:
                        minDist = 1
                        index_true = None
                        for idx in range(len(patImgIntens)):
                                if abs(patImgIntens[idx] - rng[0]) < minDist:
                                        minDist = abs(patImgIntens[idx] - rng[0])
                                        index_true = idx
                                if abs(patImgIntens[idx] - rng[1]) < minDist:
                                        minDist = abs(patImgIntens[idx] - rng[1])
                                        index_true = idx
                        reliabilityIdxEcho[index_true] = True
```

Listing 4.2: Reliability estimation for echogenicity Python 3.7 implementation

Reliability estimation for homogeneity grading

The algorithm for reliability estimation for homogeneity grading is similar to that one is used in the previous paragraph with two changes: firstly, no outliers in data were reported, secondly, the standardization presented in the 4.3 equation, is used instead of min-max normalization.

$$\overline{x_i} = \frac{x_i - \mu_x}{\sigma_x} \tag{4.3}$$

where $x_1, ..., x_N$ are all obtained values of statistic 3.5, $\overline{x_i}$ is the normalized value of respective statistic with index i, μ_x is the mean value computed for $x_1, ..., x_N$, σ_x is the mean standard deviation, computed for $x_1, ..., x_N$.

Afterwards, we divide normalized values $\overline{x_i}$ into negative and positive. If $\overline{x_i}$ is negative and homogeneity label is equal to 1, then data sample is considered as reliable (for positive value of $\overline{x_i}$ and homogeneity grading 2 it works in the same way). Similarly as for echogenicity, the case when no one value $\overline{x_i}$ lies in the range corresponding to homogeneity label, the only one data sample is selected by taking the minimal absolute value of $\overline{x_i}$. The minimal absolute value of $\overline{x_i}$ implies that this data sample is the closest to interval boundary.

Code with implementation can be seen in the Listing 4.3.

```
for pat in all_patients:
        patImgSTD = patientsImgSTD[pat]
        for carotid_side in ["left", "right"]:
                reliabilityIdxHomo = list()
                homogeneityLabel = patSideHomoLabel[pat][carotid_side]:
                for img_std in patImgSTD:
                        if homogeneityLabel == "1":
                                if img_std < 0:
                                        reliabilityIdxHomo.append(True)
                                else:
                                        reliabilityIdxHomo.append(False)
                        elif homogeneityLabel == "2":
                                if img_std > 0:
                                        reliabilityIdxHomo.append(True)
```

```
                else:
                    reliabilityIdxHomo.append(False)
        if True not in reliabilityIdxHomo:
            minDist = float('inf')
            index_true = None
            for idx in range(len(patImgSTD)):
                if abs(patImgSTD[idx]) < minDist:
                    minDist = abs(patImgSTD[idx])
                    index_true = idx
            reliabilityIdxHomo[index_true] = True
```

Listing 4.3: Reliability estimation for homogeneity Python 3.7 implementation

4.1.3 Data preparation for the degree of stenosis prediction

Experts annotation for percentage of stenosis parameter was no available in the *data_with_attributes.csv* document, so the information has been taken from the */datagrid/Medical/ArteryPlaque/AS_platy − histologie − UZ.xlsx* table. This table provides percentage of stenosis grading for only 208 patients. Also only one ultrasound image is specified for each patient in contrast to *data_with_attributes.csv* document. Therefore a method to handle such data representation was proposed:

- Find a patient to whom the specified image belongs;
- Take all ultrasound scans belonging to each particular patient;
- Assign the same degree of stenosis label taken from table to all ultrasound images belonging to patient;
- Divide ultrasound scans into transversal and longitudinal datasets;
- Select corresponding segmentations for each ultrasound scan.

After this procedure 515 transversal and 359 longitudinal scans were obtained.

4.1.4 Normalization

Deep learning models in the vast majority cases work better if inputs are normalized into smaller range than [0, 255], because of the exploding gradient problem or oscillation which may appear during the training process. So the data normalization was implemented in order to prevent an emerging of big values on the input or output of a model. Values of red, green and blue channels for all images were linearly normalized into the range [0, 1].

$$\overline{x} = \frac{x}{255} \tag{4.4}$$

Taking into consideration the fact that for different atherosclerotic plaque parameters corresponding labels belong to distinct ranges it is necessary to propose label normalization suitable for each task.

Values of stenosis percentage belonging to the range [0, 100] for degree of stenosis regression were normalized into the range [0, 1]:

$$\overline{y} = \frac{y}{100} \tag{4.5}$$

An index of echogenicity was mapped from the values {1, 2, 3, 4} to the values {0.25, 0.5, 0.75, 1} by dividing each label by 4:

$$\overline{y} = \frac{y}{4} \tag{4.6}$$

The following transformation was used for each homogeneity label: $\{1, 2\} \mapsto \{0.5, 1\}$:

$$\overline{y} = \frac{y}{2} \tag{4.7}$$

where x is an tensor representing input image into a model, y represents a label and $\overline{x}, \overline{y}$ represent normalized input tensor and normalized corresponding label.

4.2 CNN pretraining and weights initialization

Transfer learning method is a machine learning approach when some useful knowledge obtained by model while solving one task are transferred and applied to another task. So this principle assumes that we can develop machine learning model for one particular task, solve this task, store model and then use stored model parameters like initial point for another task. The source task may include a subset of domain tasks in order to acquire maximum effect from transfer learning, so that models are often pre-trained on large and challenging datasets which comprise a lot of useful semantic features. Target model therefore can be only fine-tuned on source task dataset or adapted to a new data units from the task of interest [41], [14].

4.2.1 Self-supervised learning and context restoration method

As was mentioned in section 3.1 any machine learning method requires sufficient amount of annotated data samples and also it was said that collecting any labelled medical dataset is challenging task [6], [42]. Therefore self-supervised learning can be considered as a possible solution. Self-supervised learning methods do not require labelled datasets in order to extract useful semantic features from images. Context restoration was presented as novel self-supervised learning strategy. This method has several major advantages that make it impactful: firstly, the model will acquire necessary relevant knowledge that can be further used for variety of subsequent tasks and secondly, context restoration implementation is very simple and straightforward. For our application, where we are working with medical images, context restoration may become highly beneficial. Deep learning models used for regression tasks can be initialized directly with parameters that were pretrained in context restoration process. Obviously models pretrained from natural images may not be so useful in medical imaging domain due to the significant difference in intensity distribution [6]. The entire algorithm can be described as follows:

- Take original image, disorder and corrupt its context by swapping pairs of disjunct patches N times;

- Consider corrupted image as input for the model and original image as reference;

- Model will recognize disordered parts of the image and reconstruct it by comparing the original and corrupted image;

- During this process model is learning relevant semantic features from images.

4.2.2 Implementation of context restoration method

In my implementation swapped patches are squares of size 10x10 pixels. N is equal to 50, so in total 50 pairs of disjunct patches were randomly selected and swapped to corrupt image context.

The model which is used for context restoration consists of two main parts: Analysis part is actually the model which will be used for subsequent task purposes and is going to be pretrained during context restoration process, Reconstruction part consists of several upsampling (also transposed convolution) layers that restores context information of an image [6], [42] It was decided to use 4 transposed convolution layers with proper input and output padding and stride, so that image context is restored preserving the initial size. The model with all hyperparameters for upsampling part can be seen in the Figure 4.6. The L2 loss is chosen for training model, although it causes some blurriness in restored image [6]. Nevertheless perfect restoration of image is not our main purpose. But for applications which require better quality and sharpness of restored image other loss functions may lead to a better result.

All models were trained for 1500 epochs with learning rate 1.5×10^{-4} and batch size equals 1. It is important to mention, that only training subset was used in context restoration pretraining. Validation and testing subsets remained unseen for the model in order to obtain unbiased result on these data. Several illustrative examples and comparisons of original, disordered and restored images can be seen in the Figures 4.7 for segmentation and 4.8 for original ultrasound image.

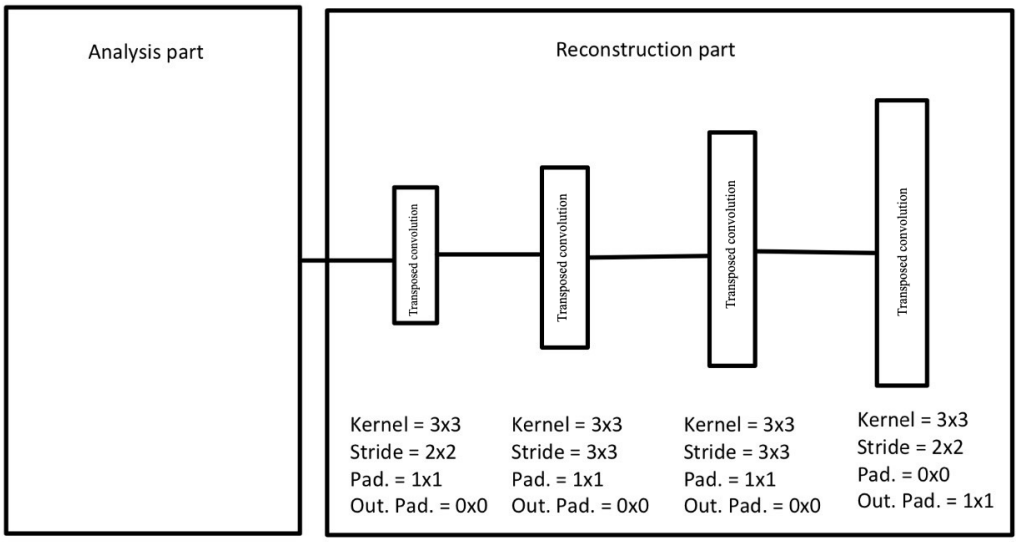
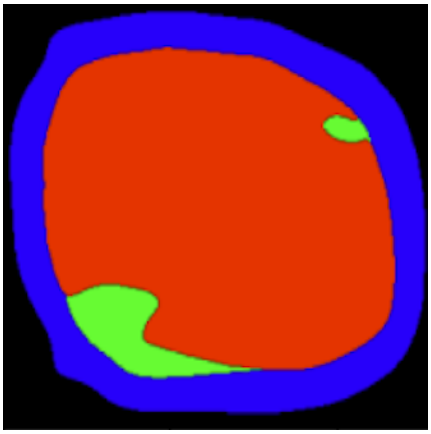
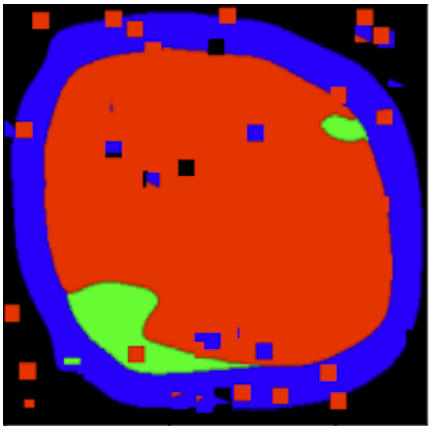


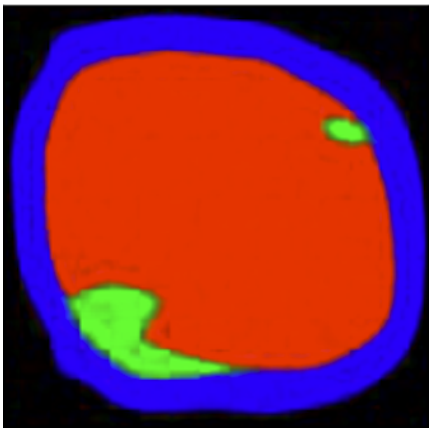
Figure 4.6: Context restoration model



(a) : Original image segmentation

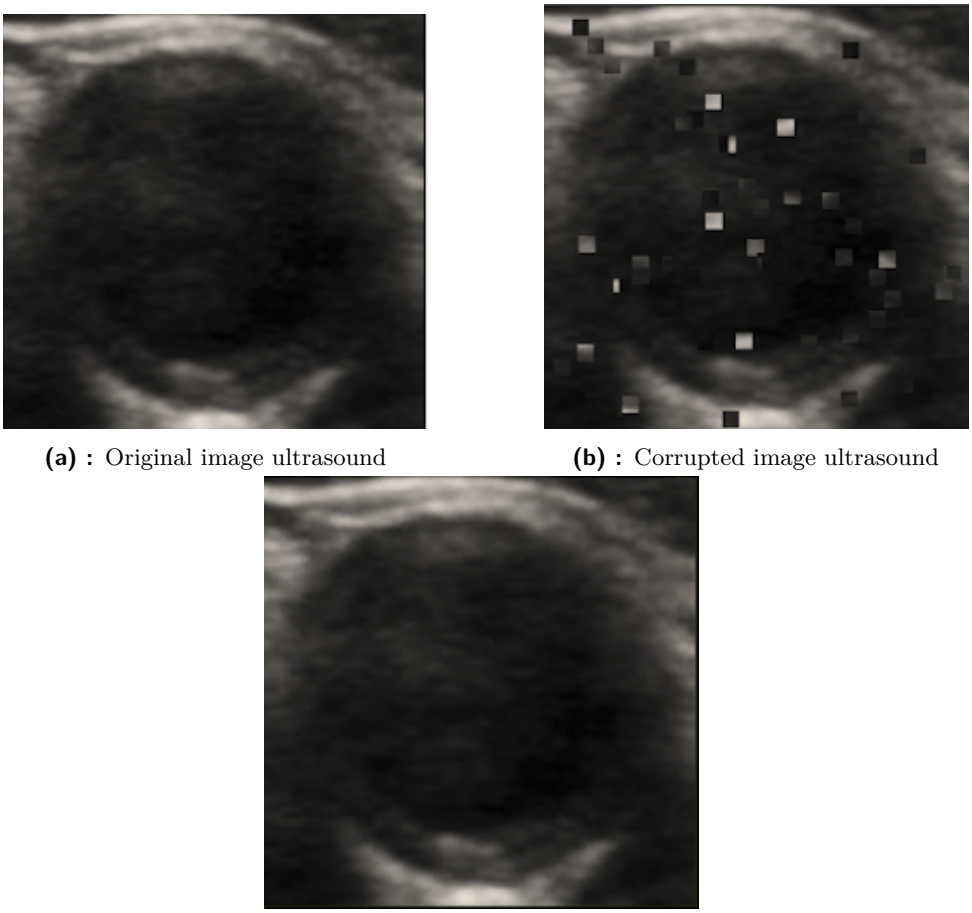


(b) : Corrupted image segmentation



(c) : Restored image segmentation

Figure 4.7: Performance of context restoration method on segmentation image



(a) : Original image ultrasound

(b) : Corrupted image ultrasound

(c) : Restored image ultrasound

Figure 4.8: Performance of context restoration method on ultrasound image

4.2.3 Weights initialization

For the regression of the atherosclerotic plaque parameters only pretrained models are going to be used. Weights obtained after context restoration would serve as initialization for models.

Xavier initializion is used for the regression part (fully connected layers) of the deep learning architecture for all models.

4.3 Data augmentation

The generalization ability and performance of machine learning model are highly dependent on the size of dataset which is used to train model. Deep learning models generalize better on large datasets with high diversity of data points, but for some applications big amount of data is not available due to several reasons. Data augmentation in machine learning is a set of techniques which helps to increase the number and the diversification of annotated training samples by preparing new data samples from already existing [16].

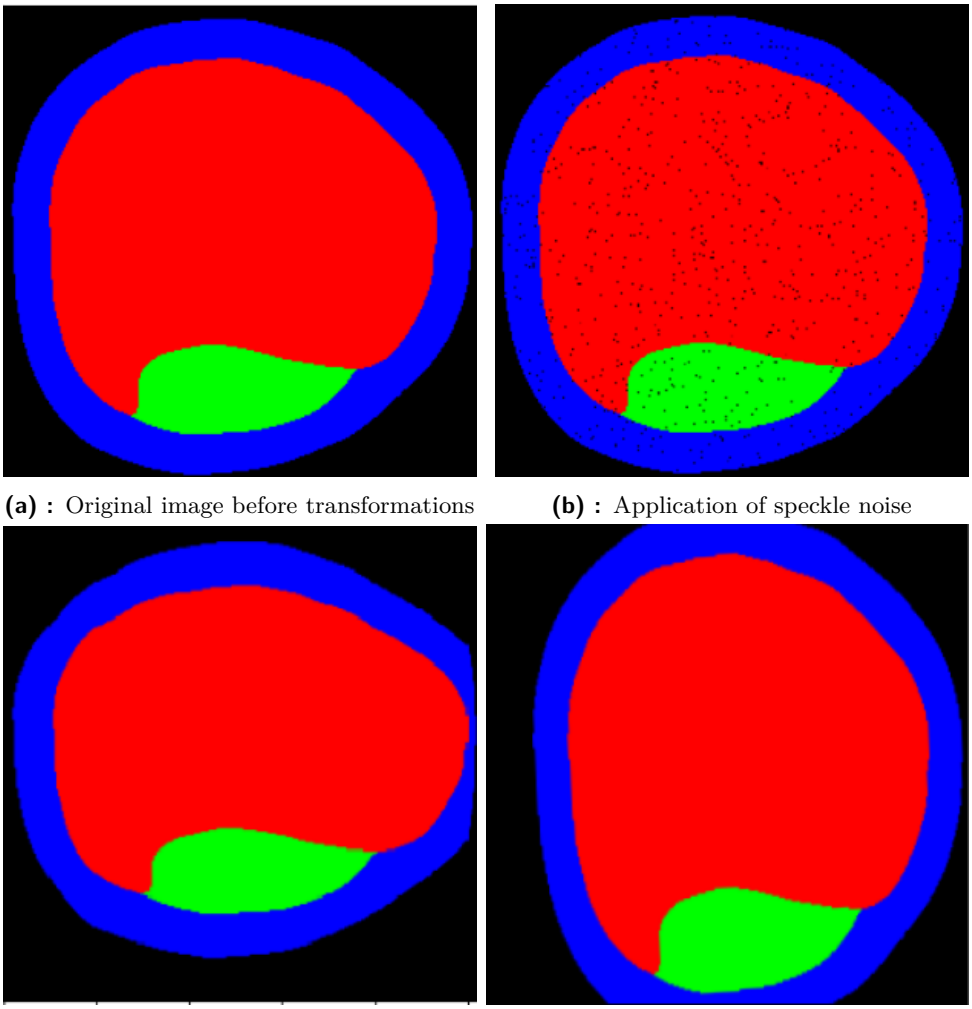
This method takes data samples and apply some set of geometric or color transformations preserving the corresponding labels. Data augmentation is also considered as a type of regularization strategy that contributes to increased model robustness, prevents overfitting and improves generalization of the model on new unseen data points. Data augmentation can be applied in two different ways: offline augmentation (generation of dataset) means that new data points are created before training process [43] and online (real-time) augmentation is more common choice in deep learning that supposes the transformations are applied on random data sample at random time. Combination of both is also possible.

The following set of geometric transformations will be used for online augmentation during a training process:

- Horizontal flip;
- Vertical flip
- Horizontal shift in range [-4, 4] pixels;
- Vertical shift in range [-4, 4] pixels;
- Rotation in range [0, 360) degrees;
- Cropping in range [0, 5] pixels from image borders in x and y axis;
- Speckle noise; [16]
- Grid distortion; [16]
- Elastic transformation; [16]
- Identity transformation.

The subset of proposed geometric transformations will be modified based on the target dataset, because some transformations may become destructive for important semantic features present in data. Also it was decided not to apply any color transformations because changing color distribution may create a significant gap between training and testing data subsets distributions and consequently affect model performance in negative way.

Several illustrative examples of transformations can be seen in the Figure 4.9.



(a) : Original image before transformations

(b) : Application of speckle noise

(c) : Application of grid transformation

(d) : Application of elastic transformation

Figure 4.9: Effect of some geometric transformations from data augmentation strategy

4.4 CNN architectures

From the mathematical point of view any neural network architecture is considered as approximation of some mathematical function that maps an observation to an output. In agreement with universal approximation theorem, even a network consisting of a single sufficiently wide layer, is enough to provide high quality approximation [44], [45]. However shallow and wide architectures are more prone to memorize training data, so that such models fail in further generalization to new data. Deep models with multiple layers are able to train itself to recognize transitional features at different levels of abstraction. Nevertheless it turns to be impossible to achieve deep model performance improvement by only straightforward stacking layers due

to the vanishing gradients, which approach infinitesimally small values if architecture becomes deeper [31],[4],[5]. Therefore two widely-known deep learning architectures with techniques preventing vanishing gradients problem are used in my work. In addition three models able to accept and process two input images simultaneously were implemented in order to obtain relevant semantic information contained in both pictures.

4.4.1 ResNet

Leveraging of skip connections (also called identity shortcut connections) is the main distinctive feature of the ResNet models family. Such engineering proposal solves degradation problem and enables to create networks with deep architectures. Moreover, shortcut technique does not bring additional learning parameters, thus computational complexity remains the same. A feature map x is passed and added to an output feature map F(x) of deeper layer according to tensor addition rules. [4] Skip connection block used in ResNet models can be seen in the Figure 4.10.

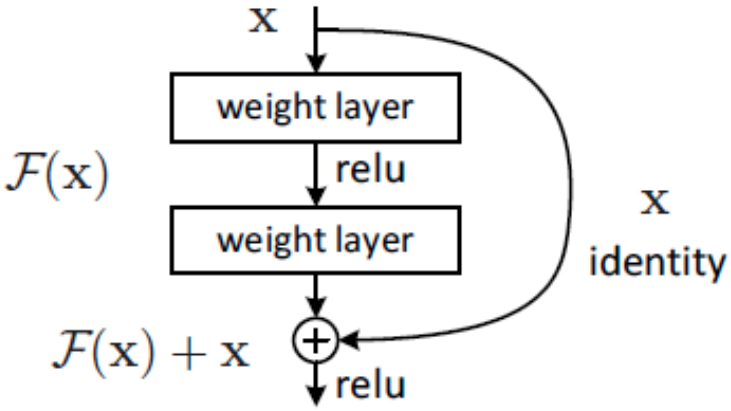


Figure 4.10: Skip connection block in ResNet model

For my work purposes model ResNet 34 was chosen. Exactly this ResNet model variant is considered by me as a good tradeoff between performance and time (also memory and computational) complexity. Model was implemented based on *torchvision.models.resnet*34(). The 2D model variation involves in total 21.5 million trainable parameters. 3×3 kernels are applied in all convolutional layers except for the initial one, where 7×7 filters are implemented. Maxpooling layer is used after image is processed by the first convolutional layer. Rectifier linear units are utilized as activation functions. Auxiliary batch normalisation layers are introduced after all convolutional layers. Fully connected layer and adaptive average pooling layer at the end of the architecture were exclueded from the model. Model description including size of filters, number of feature maps and stride can be viewed in the Figure 4.11.

Layer type	[Kernel size	№ Input feature maps	№ Output feature maps	Stride]
Conv2d	$1 \times [7 \times 7$	3	64	$2 \times 2]$
MaxPool2d	$1 \times [3 \times 3$	–	–	$2 \times 2]$
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 64 & 64 & 1 \times 1 \\ 3 \times 3 & 64 & 64 & 1 \times 1 \end{bmatrix}$			
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 64 & 128 & 2 \times 2 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$			
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 128 & 128 & 1 \times 1 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$			
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 128 & 256 & 2 \times 2 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$			
Conv2d	$5 \times \begin{bmatrix} 3 \times 3 & 256 & 256 & 1 \times 1 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$			
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 256 & 512 & 2 \times 2 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$			
Conv2d	$2 \times \begin{bmatrix} 3 \times 3 & 512 & 512 & 1 \times 1 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$			

Figure 4.11: ResNet 34 model implementation with layers description

4.4.2 DenseNet

DenseNet models utilize densely-connected layers which enhance and extend the shortcut connection technique implemented in ResNet architecture. Feature maps, produced in all previous layers, are passed to all deeper layers, form connectivity and information flow between each pair of convolutional layers within one denseblock. This technique combines feature maps by using tensor concatenation in distinction from skip connection, where feature maps are simply added up [5][30]. An example of densely connected block is demonstrated in the Figure 4.12.

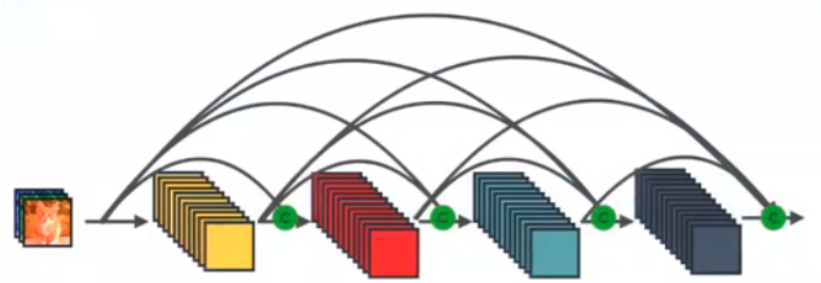


Figure 4.12: Denseblock in DenseNet model

DenseNet 161 architecture was selected and implemented by using *torchvision.models.densenet*161(). Totally this deep learning architecture has 29 million learnable parameters. Model consists of four dense blocks interconnected by implementing transition layers. Input image is processed by 7×7 kernels with stride 2. Following max pooling layers with 22 filter size and stride 2 reduces the dimensionality by half. Each dense block utilizes convolutional filters with 1×1 and 3×3 sizes. Rectifier linear units and batch normalization layers are inserted into every densely connected block. Average pooling layers with receptor field 2×2 and stride 2 are added to

transition layers besides convolutions. Model description including size of filters, number of feature maps and stride can be viewed in the Figure 4.13.

Layer type	[Kernel size Stride]	№ Input feature maps	№ Output feature maps
Conv2d	$1 \times [7 \times 7 \quad 2 \times 2]$	3	96
MaxPool2d	$1 \times [3 \times 3 \quad 2 \times 2]$	-	-
denseblock1	$6 \times \begin{bmatrix} 1 \times 1 & 1 \times 1 \\ 3 \times 3 & 1 \times 1 \end{bmatrix}$	96	384
transition1	$1 \times \begin{bmatrix} 1 \times 1 & 1 \times 1 \\ 2 \times 2 & 2 \times 2 \end{bmatrix}$	384	192
denseblock2	$12 \times \begin{bmatrix} 1 \times 1 & 1 \times 1 \\ 3 \times 3 & 1 \times 1 \end{bmatrix}$	192	768
transition2	$1 \times \begin{bmatrix} 1 \times 1 & 1 \times 1 \\ 2 \times 2 & 2 \times 2 \end{bmatrix}$	768	384
denseblock3	$36 \times \begin{bmatrix} 1 \times 1 & 1 \times 1 \\ 3 \times 3 & 1 \times 1 \end{bmatrix}$	384	2112
transition3	$1 \times \begin{bmatrix} 1 \times 1 & 1 \times 1 \\ 2 \times 2 & 2 \times 2 \end{bmatrix}$	2112	1056
denseblock4	$24 \times \begin{bmatrix} 1 \times 1 & 1 \times 1 \\ 3 \times 3 & 1 \times 1 \end{bmatrix}$	1056	2208

Figure 4.13: DenseNet 161 model implementation with layers description

4.4.3 Multi-input models

From our target datasets examination in the Chapter 3 we know that ultrasound images contain necessary content for echogenicity and homogeneity plaque parameters prediction. Also we mentioned that segmentation corresponding to ultrasound image provides us with the information about atherosclerotic plaque location. Multiple-input models accepting both an ultrasound image and corresponding segmentation are introduced in order to combine images during the training process and extract valuable content. In total three deep learning architectures, based on ResNet 34 model, were implemented. The stage at which features extracted from two input images are combined is the only one significant distinction among three proposed convolutional neural networks. In order to avoid long confusing names in the next parts of my work models will be called as Multi-input model 1, Multi-input model 2 and Multi-input model 3. The main inspiration for models with multiple inputs comes from [7].

Multi-input model 1

The first model that is able to process two input images simultaneously is based on straightforward image stacking before they are fed to the convolutional neural network. Let us assume that both images have three color channels R, G, B and identical size, so that their representation corresponds to a tensor with shape [H, W, 3], where H, W are height and width of an image respectively. Stacking both images does simple tensor concatenation hence the tensor of shape [H, W, 6] is given as input for the proposed architecture. The initial convolutional layer of the former ResNet 34 model was changed to a layer having the input dimension equals to 6. Kernel size, stride and output dimension were preserved in accordance with the replaced layer. Graphic

model representation and layers description are available in the Figures 4.14 and 4.15 respectively.

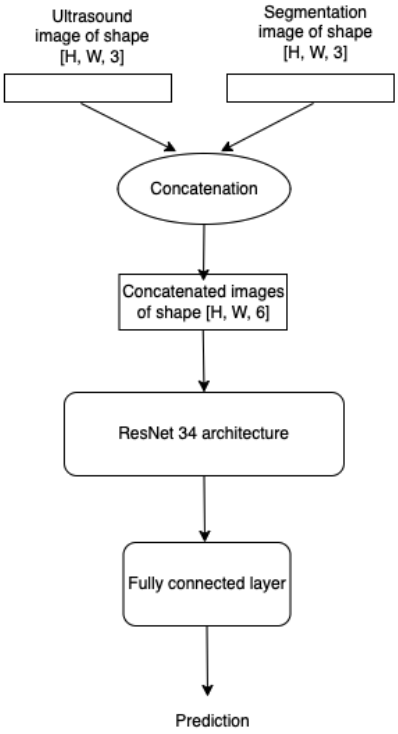


Figure 4.14: Model 1 graphic representation

Layer type	[K. In f. Out f. S.]
Conv2d	$1 \times [7 \times 7 \quad 6 \quad 64 \quad 2 \times 2]$
MaxPool2d	$1 \times [3 \times 3 \quad - \quad - \quad 2 \times 2]$
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 64 & 64 & 1 \times 1 \\ 3 \times 3 & 64 & 64 & 1 \times 1 \end{bmatrix}$
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 64 & 128 & 2 \times 2 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 128 & 128 & 1 \times 1 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 128 & 256 & 2 \times 2 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$
Conv2d	$5 \times \begin{bmatrix} 3 \times 3 & 256 & 256 & 1 \times 1 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 256 & 512 & 2 \times 2 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$
Conv2d	$2 \times \begin{bmatrix} 3 \times 3 & 512 & 512 & 1 \times 1 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$

Figure 4.15: Model 1 architecture description

Multi-input model 2

The second possible method of image combination was to create model with two separate pipes that will be processing each image independently and afterwards will be merged in one stream. First 5 layers of ResNet 34 were taken as separate pipes each of them has output dimension equals to 128. The layer with input dimensionality 128 and output dimenstionality 256 was removed and the remaining deep part of the ResNet 34 was taken as combined stream for further processing. After both images were processed in separate pipes, obtained feature maps of equal shape [H', W', 128] are concatenated into a tensor of shape [H', W', 256]. H', W' mean height and width of obtained feature maps respectively. The resulting tensor is subsequently taken as input for the combined stream part. Graphic model representation and layer description are available in the Figure 4.16 and 4.17 respectively.

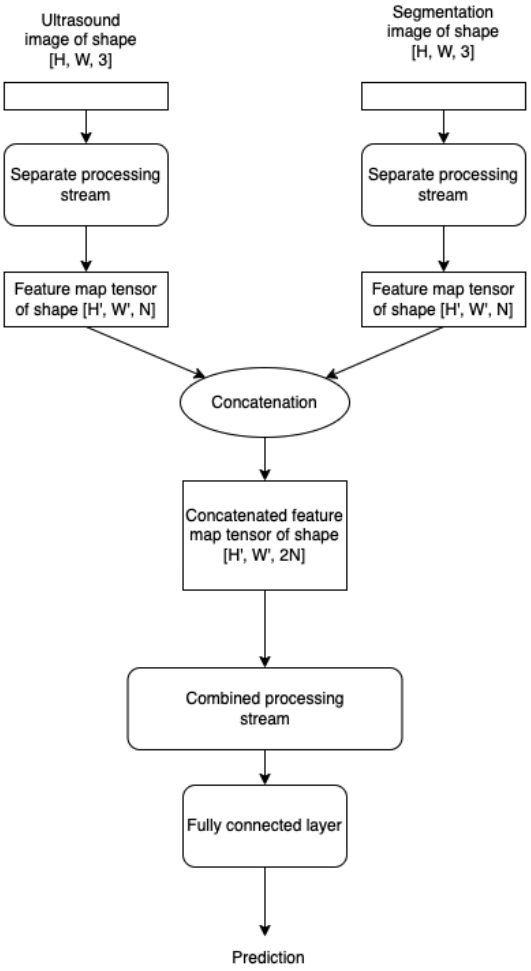


Figure 4.16: Model 2 graphic representation

Layer type	[K. In f. Out f. S.]	
Conv2d	$1 \times [7 \times 7 \quad 3 \quad 64 \quad 2 \times 2]$	$1 \times [7 \times 7 \quad 3 \quad 64 \quad 2 \times 2]$
MaxPool2d	$1 \times [3 \times 3 \quad - \quad - \quad 2 \times 2]$	$1 \times [3 \times 3 \quad - \quad - \quad 2 \times 2]$
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 64 & 64 & 1 \times 1 \\ 3 \times 3 & 64 & 64 & 1 \times 1 \end{bmatrix}$	$3 \times \begin{bmatrix} 3 \times 3 & 64 & 64 & 1 \times 1 \\ 3 \times 3 & 64 & 64 & 1 \times 1 \end{bmatrix}$
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 64 & 128 & 2 \times 2 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$	$1 \times \begin{bmatrix} 3 \times 3 & 64 & 128 & 2 \times 2 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 128 & 128 & 1 \times 1 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$	$3 \times \begin{bmatrix} 3 \times 3 & 128 & 128 & 1 \times 1 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$
Concatenation		
Conv2d	$5 \times \begin{bmatrix} 3 \times 3 & 256 & 256 & 1 \times 1 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$	
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 256 & 512 & 2 \times 2 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$	
Conv2d	$2 \times \begin{bmatrix} 3 \times 3 & 512 & 512 & 1 \times 1 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$	

Figure 4.17: Model 2 architecture description

Multi-input model 3

The final implemented model uses the same idea as the second architecture with two separate processing streams that are going to be merged in later processing phase. In distinction from previous model, separate pipes are combined right before the fully connected layer. The whole ResNet 34 model, apart from the final average pooling layer, was taken as independent processing stream for both input images. Graphic model representation and layer description are available in the Figures 4.18 and 4.19 respectively.

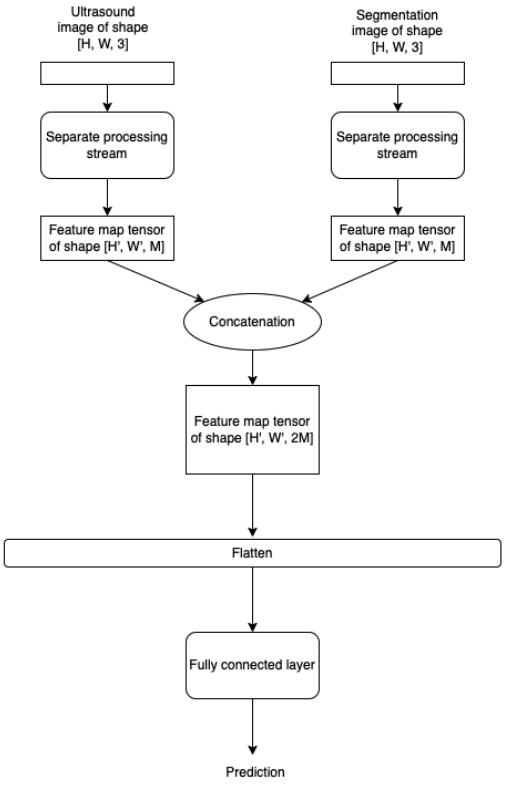


Figure 4.18: Model 3 graphic representation

Layer type	[K. In f. Out f. S.]	
Conv2d	$1 \times [7 \times 7 \quad 3 \quad 64 \quad 2 \times 2]$	$1 \times [7 \times 7 \quad 3 \quad 64 \quad 2 \times 2]$
MaxPool2d	$1 \times [3 \times 3 \quad - \quad - \quad 2 \times 2]$	$1 \times [3 \times 3 \quad - \quad - \quad 2 \times 2]$
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 64 & 64 & 1 \times 1 \\ 3 \times 3 & 64 & 64 & 1 \times 1 \end{bmatrix}$	$3 \times \begin{bmatrix} 3 \times 3 & 64 & 64 & 1 \times 1 \\ 3 \times 3 & 64 & 64 & 1 \times 1 \end{bmatrix}$
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 64 & 128 & 2 \times 2 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$	$1 \times \begin{bmatrix} 3 \times 3 & 64 & 128 & 2 \times 2 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$
Conv2d	$3 \times \begin{bmatrix} 3 \times 3 & 128 & 128 & 1 \times 1 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$	$3 \times \begin{bmatrix} 3 \times 3 & 128 & 128 & 1 \times 1 \\ 3 \times 3 & 128 & 128 & 1 \times 1 \end{bmatrix}$
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 128 & 256 & 2 \times 2 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$	$1 \times \begin{bmatrix} 3 \times 3 & 128 & 256 & 2 \times 2 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$
Conv2d	$5 \times \begin{bmatrix} 3 \times 3 & 256 & 256 & 1 \times 1 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$	$5 \times \begin{bmatrix} 3 \times 3 & 256 & 256 & 1 \times 1 \\ 3 \times 3 & 256 & 256 & 1 \times 1 \end{bmatrix}$
Conv2d	$1 \times \begin{bmatrix} 3 \times 3 & 256 & 512 & 2 \times 2 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$	$1 \times \begin{bmatrix} 3 \times 3 & 256 & 512 & 2 \times 2 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$
Conv2d	$2 \times \begin{bmatrix} 3 \times 3 & 512 & 512 & 1 \times 1 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$	$2 \times \begin{bmatrix} 3 \times 3 & 512 & 512 & 1 \times 1 \\ 3 \times 3 & 512 & 512 & 1 \times 1 \end{bmatrix}$
Concatenation		
Flatten		

Figure 4.19: Model 3 architecture description

4.4.4 Fully connected layer for regression

Atherosclerotic plaque parameters prediction task leads to building regression models. Consequently fully connected layer proper for regression must be implemented for all above said deep learning models. The following three-layer architecture was proposed:

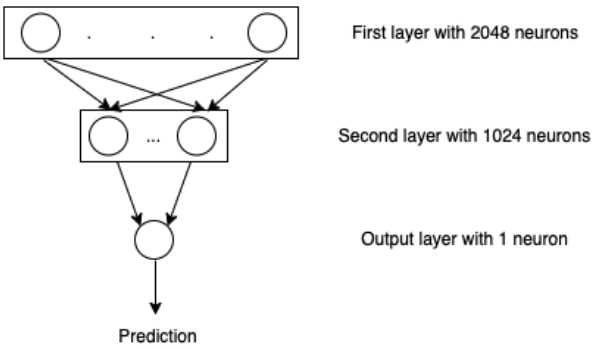


Figure 4.20: Fully connected layer for regression

The first layer consists of 2048 neurons, the second one has a half number of neurons and the last layer is made using only one neuron. Rectified linear unit is used as activation function in each layer except for the output one. Bias term was also set for each layer. Fully connected layer follows the flattening layer in every deep learning model.

4.5 Performance evaluation and correlation

Correlation is a measure that estimates the strength of the relationship between two random variables. The value of correlation belongs to the range [-1, 1], where 1 means perfect positive correlation and -1 means perfect

negative correlation. Correlation coefficient near or equals to zero means negligeable relationship or full absence of relationship between two random variables. For estimation of our regression task performance two different correlation coefficients were used.

4.5.1 Pearson correlation coefficient

Pearson correlation coefficient measures the strength of the linear relationship and is formally defined as covariance coefficient of two random variables normalized by the product of their standard deviations [28].

$$r = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{\sqrt{\mathbb{E}(X^2) - (\mathbb{E}(X))^2}\sqrt{\mathbb{E}(Y^2) - (\mathbb{E}(Y))^2}} \tag{4.8}$$

where X and Y means random variables, $\mathbb{E}$ means an expected value of random variable [38].

Another one definition of the Pearsons correlation coefficient is used when some pairs of observations $\{(x_1, y_1), ..., (x_n, y_n)\}$ are given:

$$\begin{aligned} r &= \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} = \\ &= \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}} \end{aligned} \tag{4.9}$$

where n is the total number of samples, x_i y_i is the pair of observations indexed with i, $\overline{x}$ $\overline{y}$ are mean values for all x and y observations [38].

4.5.2 Spearman rank correlation coefficient

In general Spearman rank correlation coefficient is defined as Pearson correlation among the ranks of random variables [27].

$$r_s = \frac{cov(R(X), R(Y))}{\sigma_{R(X)}\sigma_{R(Y)}} \tag{4.10}$$

where R(X) and R(Y) are random variables representing ranks of X and Y random variables [37].

Another one formula is used when some pairs of observations $\{(x_1, y_1), ..., (x_n, y_n)\}$ are given:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{4.11}$$

where $d = R(x_i) - R(y_i)$ the difference between ranks for the two corresponding samples, n is the total number of samples [37].

Chapter 5

Experiments and results

In this chapter we are going to present all quantitative results and achievements in atherosclerotic plaque parameters prediction. Section 5.1 provides us with the results of the percentage of stenosis estimation and explains several additional experiments that were done. In section 5.2 results of the echogenicity and homogeneity regression are presented. It was decided to combine the results discussion for echogenicity and homogeneity into one section, because identical datasets and methods were used for both atherosclerotic plaque parameters prediction.

5.1 Degree of stenosis regression

I have used the following settings for all ecperiments done in this part: Each dataset was divided into three parts with proportion 85 % for testing, 7.5 % for validation and 7.5 % for testing. ResNet 34 and DenseNet 161 models were trained for 50 epochs with learning rates $\alpha = 10^{-6}$ and $\alpha = 10^{-5}$ respectively. Weight decay $\mu = 0.00005$ has been set.

5.1.1 Degree of stenosis regression based on automatic segmentation and experts annotation

As was mentioned in 3.1 section, semantic segmentation provides information about atherosclerotic plaque size and location, so that degree of stenosis parameter can be predicted with relatively high by utilizing only segmentations. The datasets prepared after the procedure explained in 4.1.3 section were used for degree of stenosis prediction. The model was evaluated three times. Mean and standard deviation are computed based on three experiments. Models were initialized with two methods: parameters obtained after context restoration pretraining on segmentations and parameters after training on ImageNet dataset. Also models were trained separately on Transversal and Longitudinal datasets. The table 5.1 presents the results of the prediction of the degree of stenosis parameter

Model	Pretraining	Orient.	Pearson	Spearman
ResNet34	ImageNet	Trans	0.22 ± 0.03	0.19 ± 0.05
ResNet34	ImageNet	Long	0.20 ± 0.13	0.24 ± 0.13
ResNet34	Context rest.	Trans	0.10 ± 0.03	0.07 ± 0.03
ResNet34	Context rest.	Long	0.05 ± 0.04	0.10 ± 0.05
DenseNet161	ImageNet	Trans	0.14 ± 0.01	0.05 ± 0.02
DenseNet161	ImageNet	Long	0.30 ± 0.76	0.20 ± 0.12
DenseNet161	Context rest.	Trans	0.14 ± 0.07	0.13 ± 0.04
DenseNet161	Context rest.	Long	0.26 ± 0.07	0.13 ± 0.01

Table 5.1: Results of the degree of stenosis prediction from segmentation using experts annotation

5.1.2 Model performance tests

In order to prove that prepared deep learning models are able to learn such simple feature as proportion of plaque area and lumen area several approaches based on per pixel degree of stenosis evaluation were proposed. The main benefit of such methods is that we exactly know the degree of stenosis for each image in synthetic dataset, so that we are able to evaluate model performance without taking into consideration any mistakes in annotation and information distortion during the segmentation process preparing.

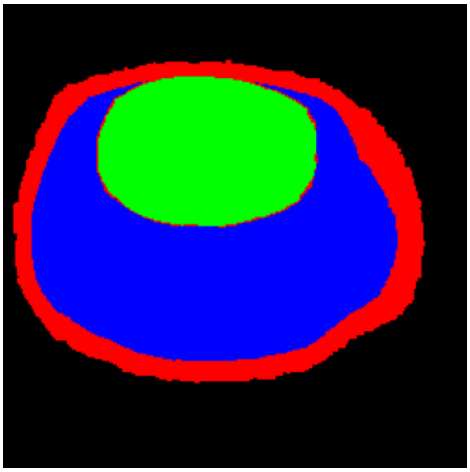
Experiment 1: Synthetic dataset

Synthetic dataset which imitates the segmentation was generated. The whole dataset consists of 5000 images of size 256x256 pixels with different green to blue area ratio in range [0, 0.45]. In total 4250 images were used for training and 375 for both validation and testing.

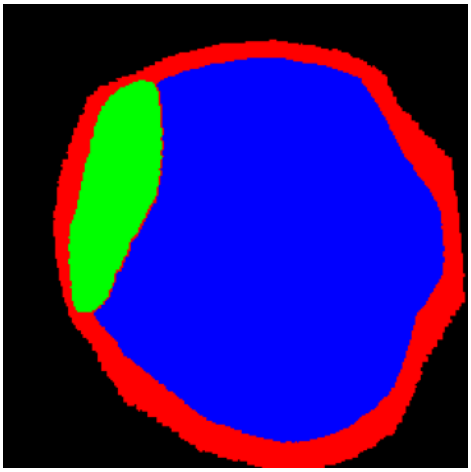
Implementation of this datatset is straightforward and can be described in only four steps:

- Draw 3 ellipses of red, blue and green colors with randomly chosen diameters;
- Red, blue and green colors represent artery wall, lumen of the artery and plaque according to segmentation;
- Apply grid distortion and elastic transformation in order to make images more unsmooth, irregular and consequently more similar to segmentations;
- Apply rotations.

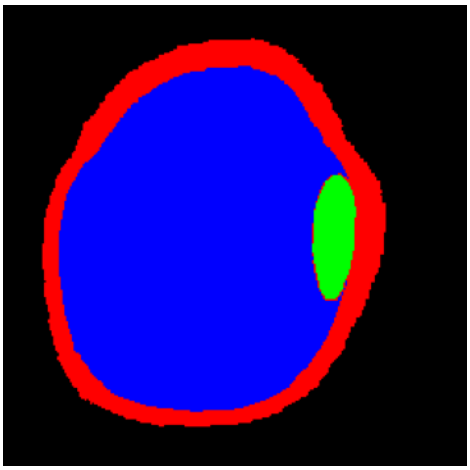
Examples of images from generated synthetic dataset can be seen in the Figure 5.1:



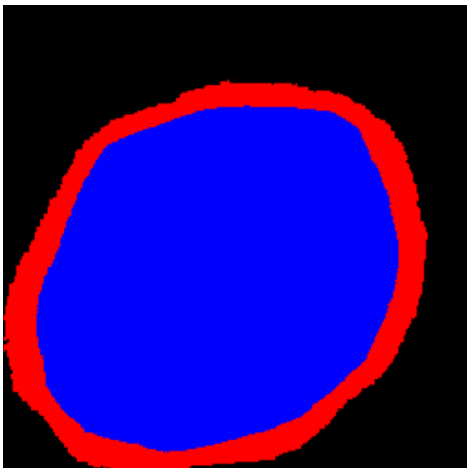
(a) : Synthetic image with green to blue ratio equals 0.324



(b) : Synthetic image with green to blue ratio equals 0.125



(c) : Synthetic image with green to blue ratio equals 0.051



(d) : Synthetic image with green to blue ratio equals 0

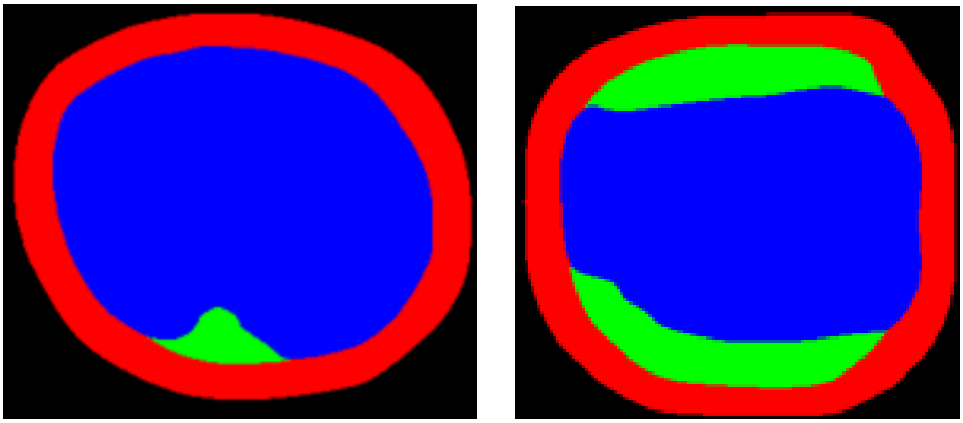
Figure 5.1: Examples of images from the synthetic dataset

Final model evaluation on testing subset is stated in the Table 5.2.

Experiment 2: Degree of stenosis approximation from segmentation

Another one test was proposed in order to verify ability of the model to extract necessary features for degree of stenosis prediction from segmentation. Test assumes that percentage of stenosis is estimated directly from segmentation image and is consequently used as label. For transversal segmentation image degree of stenosis was computed per pixel as area ratio from formula 3.1 where green pixels represent a plaque area P and blue pixels represent area of an artery lumen L. Degree of stenosis approximation was applied to all available segmentation images with transversal orientation. In total 9700 images were labelled in such way. 8245 images were used for training process, 727 images for model validation and 728 images for final testing evaluation.

Images used in context restoration pretraining were added only into training subset. Several examples can be seen in the Figure 5.2.



(a) : Segmentation image with 0.034 green to blue pixels ratio

(b) : Segmentation image with 0.236 green to blue pixels ratio

Figure 5.2: Examples of per pixel approximation

Final model evaluation on testing subset is stated in the Table 5.2.

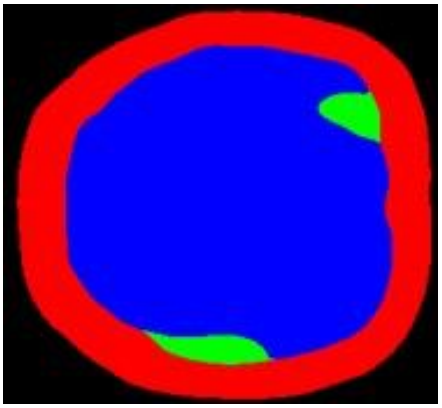
Experiment 3: Application of model trained on synthetic data to segmentation dataset

The final one test investigates if model trained on synthetic data can be applied to the segmentation dataset. This experiment assumes direct model testing on segmentation images without any additional fine-tuning.

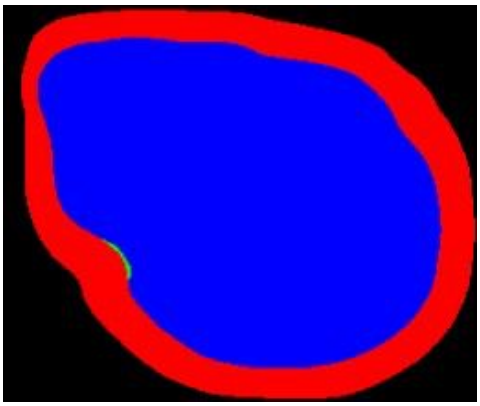
Name of the test	Pearson corr. coeff.	Spearman corr. coeff
Experiment 1	0.9632	0.9723
Experiment 2	0.9897	0.9742
Experiment 3	0.9908	0.9704

Table 5.2: Correlation between labels and predictions on testing data for all three experiments

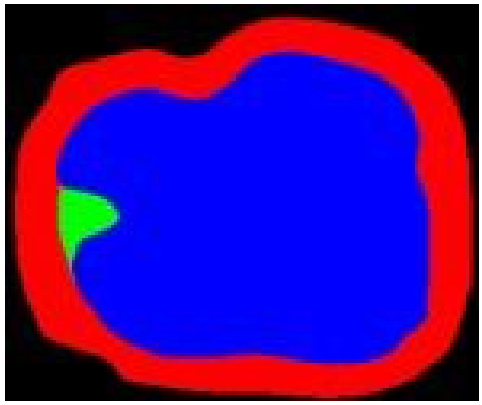
High values of correlation obtained in each experiment proves that proposed deep learning model is able to extract all necessary information to predict stenosis percentage from segmentation. On the other hand significant gap between results obtained based on experts annotation and from per - pixel evaluation can not be definitely explained. Based on visual exploration and confrontation of experts grading and segmentation my reasoning is that annotations in many cases do not correspond to illustrative evaluation. We can see three examples of such inconsistency in the Figure 5.3



(a) : Segmentation corresponding to 90 % degree of stenosis grading



(b) : Segmentation corresponding to 50 % degree of stenosis grading



(c) : Segmentation corresponding to 35 % degree of stenosis grading

Figure 5.3: Examples of data inconsistency in degree of stenosis grading

5.2 Echogenicity and homogeneity regression results

Totally 12 experiments were conducted for both echogenicity and homogeneity prediction using various combinations of proposed models and available as well as implemented datatsets. Each dataset was divided into three subsets: training subset contains 85 % of dataset, remaining 15 % of data were equally divided between validation and testing subsets. Also several data augmentation strategies were implemented depending on a target dataset, because some of geometric transformations, mentioned in 4.3 Section, can be destructive consequently leading to models performance degradation. For PSD and SPSD the following data augmentation strategy was implemented: with probability 60 % from one to four geometric transformations from a subset {Rotation, Vertical flip, Horizontal flip, Identity transformation} are applied. Randomness of speckle noise, grid distortion and elastic transformations almost always leads to the situation when two input images are transformed variously. Therefore is was impossible to use stated transforma-

tions for multiple input models training. From the remaining subset from one to four transformations are applied with probability 25 %.

ResNet 34 and DenseNet 161 were both initialized with context restoration weights. Because of qualitative pre-training strategy models were trained only for 50 epochs. Adam optimizer was used as optimization algorithm for both models. Learning rate was set as $\alpha = 5 \times 10^{-6}$ and $\alpha = 10^{-5}$ for ResNet 34 and DenseNet 161 architectures respectively. Weight decay $\mu = 0.0005$ was used for ResNet 34 model training. Multiple input models were all trained for 250 epochs. Adam optimizer with learning rate $\alpha = 5 \times 10^{-6}$ was used. MSE loss was chosen as criterion for regression tasks.

Model name	Dataset	Pearson	Spearman
ResNet34	PSD	0.547	0.430
DenseNet161	PSD	0.598	0.476
ResNet34	SPSD	0.669	0.563
DenseNet161	SPSD	0.694	0.587
Model 1	Ultrasound + Segmentation Transversal	0.726	0.548
Model 2	Ultrasound + Segmentation Transversal	**0.751**	**0.612**
Model 3	Ultrasound + Segmentation Transversal	0.587	0.408
Model 1	Ultrasound + Segmentation Longitudinal	0.347	0.294
Model 2	Ultrasound + Segmentation Longitudinal	0.313	0.275
Model 3	Ultrasound + Segmentation Longitudinal	0.256	0.213
ResNet34	Ultrasound Tranversal	0.537	0.419
DenseNet161	Ultrasound Transversal	0.524	0.389

Table 5.3: Echogenicity prediction results

Model name	Dataset	Pearson	Spearman
ResNet34	PSD	0.745	0.160
DenseNet161	PSD	0.752	0.183
ResNet34	SPSD	0.781	**0.451**
DenseNet161	SPSD	0.815	0.412
Model 1	Ultrasound + Segmentation Transversal	**0.870**	0.422
Model 2	Ultrasound + Segmentation Transversal	0.793	0.409
Model 3	Ultrasound + Segmentation Transversal	0.778	0.403
Model 1	Ultrasound + Segmentation Longitudinal	0.659	-0.064
Model 2	Ultrasound + Segmentation Longitudinal	0.679	-0.0766
Model 3	Ultrasound + Segmentation Longitudinal	0.586	-0.143
ResNet34	Ultrasound Transversal	0.691	0.342
DenseNet161	Ultrasound Transversal	0.663	0.329

Table 5.4: Homogeneity prediction results

The complete evaluation for echogeneicty and homogeneity regression is stated in the Tables 5.3 and 5.4 respectively. The best results obtained are highlighted. The best correlation for echogenicity prediction equals to 0.751

was obtained from the second variation of model accepting both the ultrasound image and segmentation with transversal orientation. The first multiple input model trained on transversal images has shown the highest correlation (0.870) for homogeneity regression task. Also we were able to achieve moderate values of correlation on raw ultrasound images with transversal orientation (0.537 for echogenicity and 0.691 for homogeneity). However achieved results can not be considered as satisfactory in medical image processing, where higher precision is required.

Chapter 6

Conclusion

This projects objective was to propose deep leaning methods for atherosclerotic plaque parameters estimation. Echogenicity, homogeneity and degree of stenosis were chosen to predict. In total 12142 ultrasound images and corresponding segmentations were used for parameters prediction. Experts annotations were also available and used as reference labels for models training and testing. For each chosen atherosclerotic plaque parameter we were able to find a simple mathematical representation. Echogenicity represents mean plaque intensity computed per pixel. Homogeneity is modelled as standard deviation of pixels intensity over the whole plaque. Degree of stenosis is represented as proportion between plaque area and artery cross-section. Prediction of all three parameters was formulated as regression problem. During the project two additional datasets were created by extracting and combining semantic information from both ultrasound image and correspondent segmentation. The first created dataset is called PSD (Plaque Segment Dataset) and collects images where only an atherosclerotic plaque segments are represented while another unwanted content is excluded, simultaneously proposed statistics representing plaque echogenicity and homogeneity were calculated. The second dataset was derived from PSD. It collects only images where only one plaque segment is present and is called SPSD (Single Plaque Segment Dataset). Leveraging computed mean and standard deviation of plaque intensity we were able to handle inconsistency in data and select proper images. Two conventional deep learning architectures such as ResNet 34 and DenseNet 161 were chosen and implemented using PyTorch library. The novel self-supervised learning strategy called context restoration was implemented for pre-training of above-said models. Also three deep learning architectures providing an opportunity to simultaneously process both the segmentation and ultrasound image were designed. For both echogenicity and homogeneity regression totally 12 experiments were conducted with different variation of target datasets and implemented convolutional neural networks. We were able to achieve 75 % correlation for the task of echogenicity regression by using the second model combining transversal ultrasound scan and segmentation. The best performance for homogeneity estimation of 87 % correlation was achieved with the first multiple input model on transversal orientation type data.

Implemented codes and models are stored and can be found in the faculty Gitlab in the following link: https://gitlab.fel.cvut.cz/morozart/bachelor-thesis

Bibliography

[1] Kostelanský M., Manzano-Rodríguez A., Kybic J., Hekrdla M., Dvorský O., Kozel J., Baurová P., Pakizer D. and Školoudík D.. *"Differentiating between stable and progressive carotid atherosclerotic plaques from in-vivo ultrasound images using texture descriptors."* SIPAIM: International Symposium on Medical Information Processing and Analysis, 2021

[2] Kostelanský M.. *"Localization and segmentation of in-vivo ultrasound carotid artery images."*, Master thesis, Czech Technical University Faculty of Electrical Engineering, 2021

[3] Ana Manzano Rodriguez. *"Processing in vivo ultrasound images of the carotid artery"*, Degree Thesis, Faculty of the Escola Tecnica d'Enginyeria de Telecomunicacio de Barcelona Universitat Politecnica de Catalunya 2021

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. *"Deep Residual Learning for Image Recognition"*, Conference: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015

[5] G. Huang, Z. Liu and L. van der Maaten. *"Densely Connected Convolutional Networks"*, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), August 2016.

[6] Liang Chen, Paul Bentley, Kensaku Mori, Kazunari Misawa, Michitaka Fujiwara, Daniel Rueckert. *"Self-supervised learning for medical image analysis using image context restoration"* , Medical Image Analysis, July 2019

[7] Philipp Fischer, Alexey Dosovitskiy , Eddy Ilg, Philip Hausser, Caner Hazırbas, Vladimir Golkov. *"FlowNet: Learning Optical Flow with Convolutional Networks"*, IEEE International Conference on Computer Vision, Project: Deep Learning for Inverse Problems, April 2015.

[8] Lekadir-Galimzianova-etal 2017. *"A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound"*, IEEE journal of biomedical and health informatics, November 2016.

[9] Elham Yousef Kalaf, Ata Jodeiri, Seyed Kamaledin Setarehdan, Ng Wei Lin, Kartini Binti Rahman, Nur Aishah Taib, Sarinder Kaur Dhillon. *"Classification of Breast Cancer Lesions in Ultrasound Images by using Attention Layer and loss Ensembles in Deep Convolutional Neural Networks"*, February 2021

[10] Germán González, George R. Washko, Raúl San José Estépar. *"Deep learning for biomarker regression: application to osteoporosis and emphysema on chest CT scans"*, Proceedings of SPIE - The International Society for Optical Engineering, Image Processing, March 2018

[11] Ali Nawaz, Syed Muhammad Anwar, Rehan Liaqat, Javid Iqbal, Ulas Bagci, Muhammad Majid. *"Deep Convolutional Neural Network based Classification of Alzheimer's Disease using MRI data"*, January 2021

[12] Miloš Anić, Branko Arsić, Smiljana Đorović, Nenad Filipović. *"Deep learning methods for detection of carotid artery wall"*, International Scientific Student Conference RI-STEM-2021, 10.06.2021.

[13] Walid Al-Dhabyani, Mohammed Gomaa, Mohammed Gomaa, Hussien Khaled, Aly Fahmy. *"Deep Learning Approaches for Data Augmentation and Classification of Breast Masses using Ultrasound Images"*, International Journal of Advanced Computer Science and Applications, January 2019

[14] Sampa Misra , Seungwan Jeon , Ravi Managuli , Seiyon Lee , Gyuwon Kim , Seungchul Lee , Richard G Barr , and Chulhong Kim. *"Ensemble Transfer Learning of Elastography and B-mode Breast Ultrasound Images"*, February 2021.

[15] Lok Hin Lee, Yuan Gao, J. Alison Noble. *"Principled Ultrasound Data Augmentation for Classification of Standard Planes"*, Information Processing in Medical Imaging, June 2021

[16] Alexander Buslaev, Alex Parinov, Eugene Khvedchenya, Vladimir I. Iglovikov, Alexandr A. Kalinin. *"Albumentations: fast and flexible image augmentations"*, Information (Switzerland), February 2018

[17] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, Quoc V. Le. *"AutoAugment: Learning Augmentation Policies from Data"*, 24 May 2018

[18] Ekin D. Cubuk, Barret Zoph, Jonathon Shlens, Quoc V. Le, *"RandAugment: Practical automated data augmentation with a reduced search space"*, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 30 September 2019

[19] Bashir Zeimarani, Marly Guimaraes Fernandes Costa, Nilufar Zeimarani Nurani, Sabrina Ramos Bianco, Wagner Coehlo De Albuquerque Pereira and Cicero Ferreira Fernandes Costa Filho. *"Breast Lesion Classification*

in Ultrasound Images Using Deep Convolutional Neural Network", IEEE Access, 2020

[20] Hyungsuk Kim, Juyoung Park, Hakjoon Lee, Geuntae Im, Jongsoo Lee, Ki-Baek Lee, Heung Jae Lee. *"Classification for Breast Ultrasound Using Convolutional Neural Network with Multiple Time-Domain Feature Maps"*, 31 October 2021

[21] Florian Dubost, Hieab Adams, Gerda Bortsova, M. Arfan Ikram, Wiro Niessen, Meike Vernooij , Marleen de Bruijnea. *"3D Regression Neural Network for the Quantification of Enlarged Perivascular Spaces in Brain MRI"*, 2018

[22] David Školoudík, Petra Kešnerová, Tomáš Hrbáč, David Netuka, Jaroslav Vomáčka, Kateřina Langová, Roman Herzig, Tomáš Belšan. *"Risk factors for carotid plaque progression after optimising the risk factor treatment: substudy results of the Atherosclerotic Plaque Characteristics Associated with a Progression Rate of the Plaque and a Risk of Stroke in Patients with the carotid Bifurcation Plaque Study (ANTIQUE)"*, 2021

[23] Adrian Rosebrock. "Regression with Keras", 2019

[24] Nathan Inkawhich. "FINETUNING TORCHVISION MODELS", PyTorch tutorials

[25] abhik jha. "Fastai — Image Regression — Age Prediction based on Image", Published in Analytics Vidhya, 2019

[26] Alex P.. "How to Train an Ensemble of Convolutional Neural Networks for Image Classification"

[27] "Spearman's correlation"

[28] QuestionPro. "Pearson correlation coefficient: Introduction, formula, calculation, and examples"

[29] Pritha Bhandari. "Correlation Coefficient Types, Formulas and Examples", August 2, 2021

[30] Lerd Statistics. "Spearman's Rank-Order Correlation"

[31] Pablo Ruiz. "Understanding and visualizing DenseNets", published in Towards Data Science, Oct 10, 2018

[32] Burak Aybar. "Self-supervised Learning for Medical Image Analysis Using Image Context Restoration", published in Towards Data Science, Mar 2, 2020

[33] n0obcoder. "Everything You Need To Know About Saving Weights In PyTorch", published in Towards Data Science, Aug 13, 2019

[34] Mayo Clinic, Patient Care and Health Information, Diseases and Conditions. "Arteriosclerosis / atherosclerosis"

[35] Johns Hopkins Medicine, Conditions and Diseases. "Atherosclerosis"

[36] Cleveland Clinic, Disease and Conditions. "Atherosclerosis: Arterial Disease"

[37] Wikipedia. Spearman's rank correlation coefficient

[38] Wikipedia. Pearson correlation coefficient

[39] Harika Bonthu. Detecting and Treating Outliers | Treating the odd one out!, May 21, 2021

[40] Codecademy Team. Normalization

[41] Dipanjan (DJ) Sarkar. A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning, Published in Towards Data Science, Nov 14, 2018.

[42] Burak Aybar. Self-supervised Learning for Medical Image Analysis Using Image Context Restoration, published in Towards Data Science, Mar 2, 2020

[43] Simi Sanya. Offline Data Augmentation for multiple images, published as a part of the Data Science Blogathon, June 14, 2021

[44] Milind Sahay. Neural Networks and the Universal Approximation Theorem, Published in Towards Data Science, Jun 6, 2020.

[45] Andre Ye. You Don't Understand Neural Networks Until You Understand the Universal Approximation Theorem, Published in Analytics Vidhya, Jul 1, 2020

[46] Barys Ihnatsenka and André Pierre Boezaart. "Ultrasound: Basic understanding and learning the language", 2010 Jul-Sep.

[47] Wikipedia. Echogenicity