

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2018

Bc. Stanislava Serečunová



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

SEGMENTACE CÉVNÍHO ŘEČIŠTĚ VE SNÍMCÍCH SÍTNICE METODAMI HLUBOKÉHO UČENÍ

BLOOD VESSEL SEGMENTATION IN RETINAL IMAGES USING DEEP LEARNING APPROACHES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Stanislava Serečunová

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. Radim Kolář, Ph.D.

BRNO 2018



Diplomová práce

magisterský navazující studijní obor **Biomedicínské a ekologické inženýrství**

Ústav biomedicínského inženýrství

Studentka: Bc. Stanislava Serečunová

ID: 165021

Ročník: 2

Akademický rok: 2017/18

NÁZEV TÉMATU:

Segmentace cévního řečiště ve snímcích sítnice metodami hlubokého učení

POKYNY PRO VYPRACOVÁNÍ:

1) Seznamte se s principem konvolučních neuronových sítí se zaměřením na jejich využití pro segmentaci objektů v obraze. 2) Prostudujte prostředí TensorFlow (s využitím jazyka Python) pro realizaci těchto metod a otestujte vybrané ukázkové úlohy. 3) Zvolte si vhodnou architekturu konvoluční neuronové sítě pro segmentaci cév v retinálních snímcích. Dále navrhnete způsob učení, včetně přípravy dat pro učení a evaluaci. 4) Zvolenou architekturu implementujte a otestujte. 5) Na základě získaných výsledků se pokuste o zlepšení metody (např. změnou architektury nebo změnou způsobu učení sítě). 6) Dosažené výsledky vyhodnoťte, porovnejte s publikovanými výsledky a dostatečně diskutujte.

DOPORUČENÁ LITERATURA:

[1] BROWNLEE, J. Deep Learning with Python, 2017

[2] TensorFlow [online]. Dostupné z: <https://www.tensorflow.org>

Termín zadání: 5.2.2018

Termín odevzdání: 18.5.2018

Vedoucí práce: doc. Ing. Radim Kolář, Ph.D.

Konzultant:

prof. Ing. Ivo Provozník, Ph.D.
předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Táto diplomová práca sa zaoberá využitím hlbokých neurónových sietí so zameraním na segmentáciu obrazu. Teoretická časť obsahuje popis hlbokých neurónových sietí a súhrn využívaných architektúr konvolučných sietí v oblasti segmentácie objektov z obrazu. V praktickej časti bola testovaním existujúcich príkladov používaných architektúr sietí naštudovaná open-source softwarová knižnica pre strojové učenie Tensorflow, implementovaná v programovacom jazyku Python. Obecným problémom použitia konvolučných neurónových sietí je požadované veľké množstvo vstupných dát. Z tohto dôvodu bola vytvorená nová dátová sada skladajúca sa z kombinácie piatich voľne dostupných databáz. Zvolená architektúra siete U-net bola testovaná prvou modifikáciou novo vytvorenej dátovej sady. Na základe výsledkov je zvolená architektúra siete modifikovaná, vďaka čomu bola vytvorená nová sieť, ktorá dosahuje lepšie výsledky než originálna sieť. Modifikovaná architektúra je následne trénovaná na vytvorenej dátovej sade, ktorá obsahuje snímky z rôznych typov fundus kamier. Natrénovaná sieť je vďaka tomuto prístupu vo výsledku robustnejšia a umožňuje segmentáciu cievného riečiska snímkov z rôznymi parametrami. Modifikovaná architektúra bola otestovaná na databázach STARE, CHASE a HRF. Výsledky boli porovnané z publikovanými metódami segmentácie z literatúry, založených na konvolučných neurónových sieťach, ale aj klasickými metódami segmentácie. Vytvorená sieť vykazuje vysokú úspešnosť segmentácie cievného riečiska porovnateľnú so state-of-the-art metódami.

KLÚČOVÉ SLOVÁ

neurón, hlboké neurónové siete, konvolúcia, segmentácia obrazu, Tensorflow, Keras, snímky očného pozadia, dátová sada

ABSTRACT

This diploma thesis deals with the application of deep neural networks with focus on image segmentation. The theoretical part contains a description of deep neural networks and a summary of widely used convolutional architectures for segmentation of objects from the image. Practical part of the work was devoted to testing of an existing network architectures. For this purpose, an open-source software library Tensorflow, implemented in Python programming language, was used. A frequent problem incorporating the use of convolutional neural networks is the requirement on large amount of input data. In order to overcome this obstacle a new data set, consisting of a combination of five freely available databases was created. The selected U-net network architecture was tested by first modification of the newly created data set. Based on the test results, the chosen network architecture has been modified. By these means a new network has been created achieving better performance in comparison to the original network. The modified architecture is then trained on a newly created data set, that contains images of different types taken with various fundus cameras. As a result, the trained network is more robust and allows segmentation of retina blood vessels from images with different parameters. The modified architecture was tested on the STARE, CHASE, and HRF databases. Results were compared with published segmentation methods from literature, which are based on convolutional neural networks, as well as classical segmentation methods. The created network shows a high success rate of retina blood vessels segmentation comparable to state-of-the-art methods.

KEYWORDS

neuron, convolution, deep learning neural network, image segmentation, Tensorflow, Keras, retinal images, image dataset

SEREČUNOVÁ, Stanislava. *Segmentace cévního řečiště pomocí hlubokých neuronových sítí*. Brno, 2018, 96 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedúci práce: doc. Ing. Radim Kolář, PhD.

VYHLÁSENIE

Vyhlasujem, že som svoju diplomovú prácu na tému „Segmentace cévního řečiště pomocí hlubokých neuronových sítí“ vypracovala samostatne pod vedením vedúceho diplomovej práce, využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autorka uvedenej diplomovej práce ďalej vyhlasujem, že v súvislosti s vytvorením tejto diplomovej práce som neporušila autorské práva tretích osôb, najmä som nezasahovala nedovoleným spôsobom do cudzích autorských práv osobnostných alebo majetkových a som si plne vedomá následkov porušenia ustanovenia § 11 a nasledujúcich autorského zákona Českej republiky č. 121/2000 Sb., o práve autorskom, o právach súvisiacich s právom autorským a o zmene niektorých zákonov (autorský zákon), v znení neskorších predpisov, vrátane možných trestnoprávných dôsledkov vyplývajúcich z ustanovenia časti druhej, hlavy VI. diel 4 Trestného zákonníka Českej republiky č. 40/2009 Sb.

Brno

.....

podpis autorky

POĎAKOVANIE

Rada by som poďakovala vedúcemu diplomovej práce pánovi doc. Ing. Radimovi Kolářovi, Ph.D. za odborné vedenie, konzultácie, trpezlivosť, podnetné návrhy a riešenia k práci. Ďakujem Ing. Richardovi Paštěkovi, MSc. za podporu, konzultovanie a trpezlivosť počas celého priebehu vytvárania a písania diplomovej práce. Taktiež ďakujem Ing. Jurajovi Šoltěsovi a Ing. Mikulášovi Serečunovi za korekciu gramatickej časti práce. Zároveň ďakujem svojim rodičom, starým rodičom, súrodencom, priateľom a celej rodine za ich podporu, povzbudenie a pomoc počas celého štúdia, a pri písaní záverečnej diplomovej práce. Prístup k výpočtovým a úložným zdrojom vlastnených skupinami a projektmi podieľajúcimi sa na Národnej Gridovej Infraštruktúre MetaCentrum pod programom "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), je veľmi cenený.

Brno

.....

podpis autorky

OBSAH

Úvod	13
1 Neurónové siete	15
1.1 Neurón	15
1.2 Topológia sietí	16
1.3 Učenie neurónových sietí	17
1.4 Optimalizačné algoritmy	19
2 Hlboké neurónové siete	22
2.1 Konvolučné neurónové siete	22
2.2 Plné prepojené konvolučné siete	25
2.3 Učenie hlbokých neurónových sietí	25
2.4 Techniky minimalizácie syndrómu preučenia	27
3 Metódy segmentácie obrazu	29
3.1 Plne prepojená sieť pre sémantickú segmentáciu	29
3.2 SegNet	31
3.3 ED-CNN	31
3.4 U-net	32
4 Testovanie vybraných úloh	34
4.1 Vyhodnocovacie metriky	35
4.2 Modely klasifikačných konvolučných neurónových sietí	38
4.3 Modely segmentačných konvolučných neurónových sietí	41
5 Návrh a implementácia konvolučnej neurónovej siete	46
5.1 Vytvorenie dátovej sady	46
5.2 Modifikácia dostupných dátových sad	47
5.2.1 Trénovanie na zmenšených obrazoch	48
5.2.2 Trénovanie na výrezoch obrazov	48
5.2.3 Trénovanie na pod-obrazoch	50
5.3 Testovanie zvolenej architektúry siete	51
5.3.1 Modifikácie zvolenej architektúry	52
5.3.2 Prahovanie obrazu	55
6 Výsledky a diskusia	58
6.1 Trénovanie na zmenšených obrazoch	58
6.2 Trénovanie na výrezoch obrazov	60

6.3 Trénovanie na pod-obrazoch	67
7 Záver	82
Literatúra	85
Zoznam symbolov, veličín a skratiek	93
Zoznam príloh	94
A Obsah přiloženého DVD	95

ZOZNAM OBRÁZKOV

1.1	Biologický vs. jednotkový neurón	15
1.2	Viacvrstvá dopredná neurónová sieť	17
2.1	Architektúra konvolučnej neurónovej siete s jednotlivými vrstvami	22
2.2	Princíp konvolúcie	23
2.3	ReLU vrstva	24
2.4	Pooling vrstva	24
2.5	Plne prepojená vrstva	25
2.6	Ukážkový prípad pretrénovania siete.	27
3.1	Plne prepojená sieť pre sémantickú segmentáciu	30
3.2	Výsledky segmentácie plne prepojenej neurónovej siete	30
3.3	Architektúra SegNet	31
3.4	Výsledky segmentácie pľúc architektúrou ED-CNN	32
3.5	Výsledky predikcie pixelov architektúrou U-net	33
3.6	Architektúra siete U-Net	33
4.1	ROC krivka s umiestnením prahových hodnôt	37
4.2	Základná použitá architektúra	39
4.3	Architektúra siete VGG16	40
4.4	Doladenie predtrénovanej siete	40
4.5	Upravená architektúra siete U-net	41
4.6	Ukážka segmentácie nervov upravenej architektúry U-net	42
4.7	Architektúra siete inšpirovaná U-net	43
4.8	Ukážka segmentácie nervovej sústavy v oblasti bruška drozofily	44
4.9	Ukážka segmentácie cievneho riečiska	44
5.1	Rozdelenie obrazu na RGB zložky	47
5.2	Ukážka dátovej sady vytvorenej zmenšením obrazov	48
5.3	Ukážka dátovej sady vytvorenej orezaním obrazov	49
5.4	Ukážka rozdelenia snímku na pod-obrazy	50
5.5	Priebeh chybovej funkcie a úspešnosti tréningu architektúry siete	54
5.6	Modifikácia zvolenej architektúry	55
5.7	Prahovanie Otsu metódou	56
5.8	ROC krivka 20 testovaných snímok databázy STARE na modifikovanej architektúre siete	56
5.9	Zvýšenie prahovej hodnoty na ROC krivke do oblasti optimálneho prahu	57
6.1	Výsledky najlepšej a najhoršej segmentácie snímku z databázy STARE, sieť tréningovaná na zmenšených obrazoch	60

6.2	Výsledky pre dátovú sadu obrazov orezaných na veľkosť 512x512 pixelov po 35 epochách	61
6.3	Výsledky najlepšej a najhoršej segmentácie snímku z databázy STARE, sieť trénovaná na výrezoch obrazov - Modifikácia dátovej sady 2 . . .	62
6.4	Ukážka segmentácie snímok z databázy STARE, ktoré nemajú anotovaný obraz segmentácie cievneho riečiska, sieť trénovaná na výrezoch obrazov - Modifikácia dátovej sady 2	64
6.5	Výsledky najlepšej a najhoršej segmentácie snímku z databázy CHASE, sieť trénovaná na výrezoch - Modifikácia dátovej sady 3	65
6.6	Ukážka výsledku segmentácie pod-obrazu pacienta so zdravým okom	68
6.7	Ukážka výsledku segmentácie pod-obrazu pacienta trpiacim glaukomovou chorobou	68
6.8	Ukážka výsledku segmentácie pod-obrazu pacienta s diabetickou retinopatiou	69
6.9	Ukážka segmentácie rekonštruovaného snímku pacienta so zdravým okom	70
6.10	Ukážka segmentácie rekonštruovaného snímku pacienta s glaukómom	70
6.11	Ukážka segmentácie rekonštruovaného snímku pacienta s DR	71
6.12	Ukážka výsledného snímku po odstránení indexov, nenapojených na hlavné cievne riečisko	72
6.13	Ukážka najlepších a najhorších výsledkov segmentácie cievneho riečiska databázy HRF	74
6.14	Ukážka najlepšej a najhoršej segmentácie snímku databázy CHASE, sieť trénovaná na pod-obrazoch databázy HRF	77
6.15	Ukážka najlepšej a najhoršej segmentácie snímku databázy STARE, sieť trénovaná na vytvorených pod-obrazoch HRF	77

ZOZNAM TABULIEK

1.1	Prehľad aktivačných funkcií	16
4.1	Rozdelenie oblasti pod ROC krivkou	38
4.2	Úspešnosť modelov konvolučných sietí 1	39
4.3	Úspešnosť modelov konvolučných sietí 2	41
5.1	Prehľad dostupných databáz retinálnych snímok	47
5.2	Výsledky pôvodnej architektúry trénovanej na dátovej sade zmenšených obrazov	52
5.3	Výsledky modifikovaných modelov siete	54
6.1	Výsledky modifikovanej siete trénovanej na zmenšených obrazoch - testovaná databáza STARE po prahovaní	59
6.2	Výsledky modifikovanej architektúry siete trénovanej na dátovej sade orezaných obrazov	60
6.3	Výsledky testovania vybraných snímok siete trénovanej na zmenšených obrazoch	61
6.4	Výsledky modifikovanej architektúry siete trénovanej na dátovej sade orezaných obrazov	62
6.5	Výsledky modifikovanej siete trénovanej na orezaných obrazoch - Modifikácia dátovej sady 2, testovaná databáza STARE po prahovaní	63
6.6	Výsledky modifikovanej siete trénovanej na orezaných obrazoch - Modifikácia dátovej sady 3, testovaná databáza CHASE po prahovaní	66
6.7	Výsledky modifikovanej architektúry siete trénovanej na pod-obrazoch databázy HRF	67
6.8	Výsledky úspešnosti siete testovanej na databázach HRF, STARE a CHASE	67
6.9	Porovnanie výsledkov segmentácie cievného riečiska s metódou publikovanou Odstrčilík et al. pre zdravú skupinu jedincov	73
6.10	Porovnanie výsledkov segmentácie cievného riečiska s metódou publikovanou Odstrčilík et al. pre skupinu jedincov s glaukómovou chorobou	73
6.11	Porovnanie výsledkov segmentácie cievného riečiska s metódou publikovanou Odstrčilík et al. pre skupinu jedincov s diabetickou retinopatiou	74
6.12	Porovnanie výsledkov databázy HRF s dostupnými segmentačnými metódami	75
6.13	Výsledky modifikovanej siete trénovanej na pod-obrazoch databázy HRF - testovaná databáza CHASE po prahovaní	76
6.14	Výsledky modifikovanej siete trénovanej na pod-obrazoch databázy HRF - testovaná databáza STARE po prahovaní	78

6.15	Porovnanie dosiahnutých výsledkov testovanej databázy STARE s dostupnými segmentačnými metódami	79
6.16	Porovnanie dosiahnutých výsledkov testovanej databázy CHASE s dostupnými segmentačnými metódami	80
6.17	Prehľad najlepších segmentačných metód cievneho riečiska	81

ÚVOD

Sietnica (retina) je metabolicky aktívna časť nervového tkaniva s vysokou spotrebou kyslíka, vďaka čomu obsahuje veľmi veľké množstvo ciev [37]. Ich štrukturálne usporiadanie je pre každého jedinca unikátne. Pozícia ciev v retine ľudského oka predstavuje dôležité fyziologické a biometrické informácie, ktoré sa využívajú najmä pri diagnostike očných chorôb a získavaní biometrických údajov človeka na jeho identifikáciu.

Diabetická retinopatia, vekom podmienená makulárna degenerácia a glaukómové ochorenie patria v súčasnosti medzi najbežnejšie ochorenia sietnice. Pri neskorej diagnostike, môžu tieto degeneratívne ochorenia sietnice spôsobiť čiastočnú alebo úplnú stratu zraku. Tieto ochorenia sú značne viditeľné na sietnici ľudského oka, preto je segmentácia cievneho riečiska dôležitým faktorom pri ich diagnostike.

V dnešnej dobe existuje množstvo metód ako cievne riečisko z retiny segmentovať. Táto práca sa zaoberá aktuálne jedným z najprogresívnejších odvetí strojového učenia, ktorým sú hlboké neurónové siete so zameraním na segmentáciu. Úlohou tejto diplomovej práce je oboznámiť sa s konvolučnými neurónovými sieťami zameranými na segmentáciu obrazu. Cieľom práce, je vytvoriť architektúru siete, ktorá bude na základe tréningu na snímkach retiny s príslušným ručne segmentovaným cievny riečiskom schopná, s čo najvyššou presnosťou segmentovať hlavný cievny strom na iných snímkach sietnice. Práca je rozčlenená do šiestich kapitol.

V prvej kapitole je stručný úvod do problematiky neurónových sietí. Druhá kapitola je zameraná na podrobnejší rozbor vrstiev hlbokých neurónových sietí, konkrétne konvolučných neurónových sietí a plne prepojených konvolučných sietí.

V tretej kapitole je uvedený prehľad existujúcich architektúr konvolučných neurónových sietí so zameraním na segmentáciu objektu v obraze. Štvrtá kapitola slúži ako stručný prehľad testovania už existujúcich architektúr konvolučných sietí, s cieľom oboznámiť sa s programovacím jazykom Python a open-source softwarovou knižnicou pre strojové učenie Tensorflow. Na základe výsledkov testovania dostupných segmentačných neurónových sietí je zvolená architektúra siete, ktorá je inšpirovaná pôvodnou segmentačnou sieťou U-net určenou pre biomedicínske obrazy.

Piata kapitola sa zaoberá vytváraním dátovej sady. Z dostupných dátových sád sú vytvorené tri návrhy učenia zvolenej architektúry siete. Prvá dátová sada je vytvorená zmenšením pôvodných obrazov. Druhá dátová sada je vytvorená výrezmi pôvodných obrazov. Tretia dátová sada je tvorená pod-obrazmi dostupnej databázy, ktorá má najvyššie rozlíšenie. Zvolená sieť je otestovaná na prvej modifikácii dátovej sady vzniknutej zmenšením pôvodných obrazov. Na základe výsledkov je zvolená sieť upravovaná, s cieľom dosiahnuť lepšie výsledky segmentácie cievneho riečiska.

Posledná šiesta kapitola popisuje výsledky všetkých návrhov učenia siete, ktoré

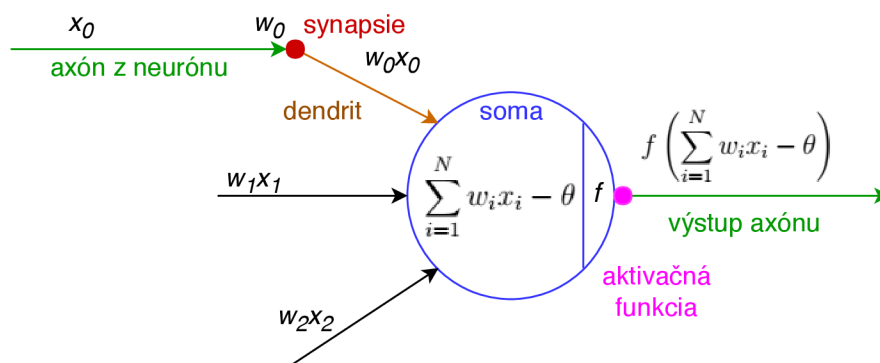
sú otestované na databáze STARE, CHASE a HRF. Dosiiahnuté výsledky sú porovnávané s dostupnými metódami segmentácie cievného riečiska publikované v odbornej literatúre, vrátane segmentačných metód využívajúcich konvolučné neurónové siete.

1 NEURÓNOVÉ SIETE

Umelá neurónová sieť (Artificial Neural Network – ANN) sa skladá z veľkého množstva prepojených procesných prvkov (neurónov), ktoré pracujú spoločne na riešení špecifických problémov. Tak ako sa človek učí príkladom, aj neurónová sieť je nakonfigurovaná pre špeciálnu aplikáciu ako je rozpoznávanie vzorov alebo klasifikácia údajov prostredníctvom procesu učenia. Učenie v biologických systémoch zahŕňa úpravy synaptických spojení, ktoré existujú medzi neurónmi. To platí aj v prípade umelých neurónových sietí.

1.1 Neurón

Matematický model neurónu, nazývaný tiež perceptron, bol v minulosti inšpirovaný anatomickou stavbou biologického nervového systému a spôsobom jeho spracovania informácií. Avšak ustanovením základných technických konceptov prebiehal rozvoj neurónových sietí samostatne [15]. Na Obr. 1.1 sú znázornené súvislosti medzi biologickým a matematickým neurónom.



Obr. 1.1: Biologický vs. jednotkový neurón. Prevzané z [31]

Biologický neurón sa skladá z tela (soma) a niekoľkých výbežkov. Výbežky sú ďalej rozdeľované na dendridy a axón. Dendridy predstavujú z matematického hľadiska vstupné vektory. Axón je dlhý výbežok, ktorým sa vzruch šíri k iným bunkám a ktorý je pomocou synapsíí (váh) napojený na dendridy ďalších neurónov [51].

Matematický neurón je definovaný ako procesný prvok zložený z N vstupov a jedného výstupu, ktorého charakteristická rovnica je [24]:

$$y = f\left(\sum_{i=1}^N w_i x_i - \theta\right), \quad (1.1)$$

kde výstup y je daný vstupným vektorom $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ násobeným vektorom aktuálnych váh $\mathbf{w} = [w_1, w_2, \dots, w_N]^T$. Váhy kladú určitý význam hodnotám na vstupe, t.j. čím vyššia váha, tým sa stáva daný vstup viac významným pre neurón. Od výsledného súčinu vstupu a váhy je odčítavaný aktuálny prah neurónu. Získaná hodnota vnútorného potenciálu neurónu sa stáva argumentom zvyčajne nelineárnej funkcie nazývanej tiež ako charakteristická funkcia neurónu [24]. Typicky používané charakteristiky neurónu sú zobrazené v Tab. 1.1 aj s príslušnými funkciami a oborami hodnôt.

Tab. 1.1: Prehľad aktivačných funkcií

Názov	Funkcia	Obor hodnôt
Skoková funkcia	$f(x) = \text{sign}(x)$	$\{-1, 1\}$
Logistická sigmoida	$f(x) = \frac{1}{1+e^{-x}}$	$(0, 1)$
Hyperbolický tangens	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$(-1, 1)$
ReLU	$f(x) = \max(0, x)$	$\langle 0, \infty \rangle$
Softmax	$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^k e^{z_k}}$	$(0, 1)$
Ostrá nelinearita	$f(x) = \begin{cases} 1 & x > 0,5 \\ 0 & x \leq 0,5 \end{cases}$	$\{0, 1\}$

Z praktických dôvodov je prah modelovaný ako nultá pozícia váh $w_0 = -\theta$, pričom k vstupnému vektoru pribúda tiež nultá hodnota vstupu definovaná ako $x_0 = 1$. V takom prípade je možné prepísať aktiváciu ako [24]:

$$\alpha = \left(\sum_{i=0}^N w_i x_i \right), \quad (1.2)$$

príčom charakteristická rovnica nadobudne tvar [24]:

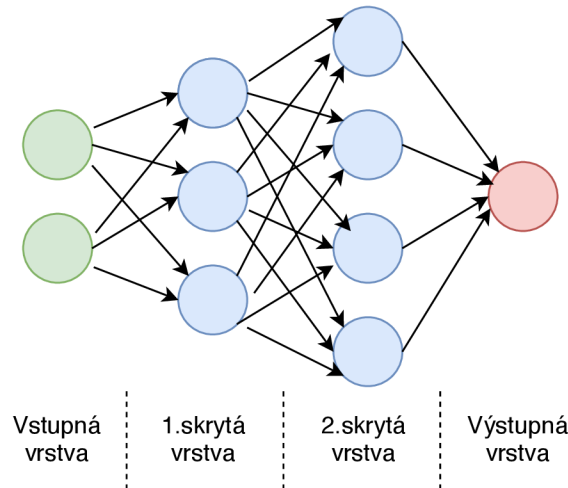
$$y = f(w_i^T x_i), \quad (1.3)$$

Pre N -rozmerný priestor reálnych čísel R^N je definičným oborom neurónu zvyčajne interval reálnych čísel, napr. pre aktivačnú funkciu logistickej sigmoidy $(0, 1)$. Existujú aj iné koncepty neurónov, ktorých výstupné hodnoty sú počítané iným spôsobom, napr. neurón s radiálnou bázou [24], [15].

1.2 Topológia sietí

Neurónová sieť sa skladá z ľubovoľného počtu N vstupov a ľubovoľného počtu M výstupov, ktorá realizuje štruktúralne zobrazenie z priestoru R^N do R^M [24]. Medzi

vstupnou a výstupnou vrstvou siete sa nachádza skrytá vrstva. V súčasnosti existuje veľké množstvo rôznych architektúr neurónových sietí, z ktorých sú najbežnejšie a najjednoduchšie dopredné neurónové siete znázornené na Obr. 1.2, kde zelená znázorňuje vstupnú vrstvu, modrá skryté vrstvy a červená výstupnú vrstvu.



Obr. 1.2: Viacvrstvá dopredná neurónová sieť

Vstupná vrstva veľkosti N je napojená na skrytú vrstvu neurónu. Skrytých vrstiev v topológii viacvrstvých dopredných sietí môže byť viac ako dve (Obr. 1.2). Ich počet závisí na type úlohy. Dopredná neurónová sieť je vyhodnocovaná postupným výpočtom výstupu neurónu v rámci jednotlivých vrstiev. Predchádzajúce výstupy siete sú prepojené s každým neurónom nasledujúcej vrstvy a posledná vrstva je výstupom siete. Medzi dopredné neurónové siete patria aj konvolučné neurónové siete. V prípade plne prepojených vrstiev je počítaný váhový súčet vstupov pre každý neurón a na výstupe je vyhodnocovaný aktivačnou funkciou [5].

1.3 Učenie neurónových sietí

Učenie, resp. tréningovanie neurónových sietí, sa delí do troch hlavných kategórií:

Učenie s učiteľom (supervisor learning) je metóda, pri ktorej je výstup neurónovej siete známy. Neurónová sieť sa učí pomocou vektora vstupných hodnôt, ku ktorému prislúcha vektor výstupných hodnôt. Vektor vstupných hodnôt je počas učenia privedený na vstupnú vrstvu neurónovej siete, kde je spracovávaný neurónovou sieťou a ďalej privedený na výstupnú vrstvu. V prípade, že sa výstupná vrstva líši od požadovaného vektora, je počítaná chybová funkcia, ktorá následne upraví hodnotu váh. Dôležitým problémom týkajúcim sa tohto tréningovania je konvergencia

chýb, t.j. minimalizácia chýb medzi požadovanými hodnotami. Cieľom je určiť súbor vstupných váh, ktoré minimalizujú túto chybu [58].

Učenie bez učiteľa (unsupervised learning) má nastavené váhy len pomocou vstupných vektorov a meniace vektory váh neporovnáva s požadovanými hodnotami, pretože ich nemá definované. Táto metóda tréningu je založená na schopnosti neurónových sietí rozoznať vo vstupných vzoroch rovnaké alebo blízke hodnoty. Na základe týchto informácií triedi prichádzajúce vektory podľa odozvy, ktorá je pre daný vstup najlepšia a podobné vektory spája do zhlukov [58].

Účelové učenie (performance learning) hľadá extrémne hodnoty výstupov siete (minima a maxima) pomocou účelovej funkcie $Y = t(\mathbf{X}, \mathbf{W})$, kde t je transformačná funkcia, \mathbf{X} je vstupný vektor a \mathbf{W} je vektor váh [44].

Algoritmus spätného šírenia chyby

Učenie viacvrstvových dopredných sietí je vykonávané algoritmom spätného šírenia chyby (backpropagation). Nejedná sa o spätné šírenie signálu, ako je to v prípade rekurentných sietí [55], ale ide o propagáciu výpočtu chybovej funkcie z výstupnej vrstvy k vstupnej vrstve neurónovej siete. Po výpočte výstupnej hodnoty neurónu vo výstupnej vrstve je táto hodnota porovnávaná s očakávanou hodnotou výpočtom hodnoty delta. Výpočet delta hodnoty pre novú výstupnú vrstvu je nasledovný [43]:

$$\delta = (d_j - y_j) \cdot y_j(1 - y_j), \quad (1.4)$$

kde d_j je výstupná hodnota požadovaného výstupu a y_j je výstupná hodnota neurónu vo výstupnej vrstve. Na základe hodnoty δ sa prispôbujú váhy a predloží sa nový vstupný tréningový vektor hodnôt s odpovedajúcim výstupom. Aktualizácia váh sa počíta pomocou nasledujúceho vzorca [43]:

$$w_{j,i} = w_{j,i} + \Delta w_{j,i}, \quad (1.5)$$

hodnota $\Delta w_{j,i}$ sa pripočítava k starej váhe a je vypočítaná ako [43]:

$$\Delta w_{j,i} = \mu \delta_j x_i, \quad (1.6)$$

kde μ je koeficient učenia. Vyhodnotením učenia v danej epoche dochádza k ukončeniu procesu učenia alebo zahájeniu novej epochy, pri ktorej sú opäť inicializované nové váhy so vstupným vektorom a požadovaným výstupom. Učenie je ukončené dosiahnutím požadovanej smerodajnej odchýlky alebo vykonaním stanoveného počtu epoch.

1.4 Optimalizačné algoritmy

Optimalizačné algoritmy pomáhajú minimalizovať chybovú funkciu. Chybová funkcia je matematická funkcia závislá na vnútorných parametroch modelu siete. Týmito parametrami sú váhy a prah neurónovej siete, ktoré sa používajú pri výpočte výstupných hodnôt a sú naučené a aktualizované v smere optimálneho riešenia, tj. ich úlohou je minimalizovať chybovú funkciu v procese tréningu modelu neurónovej siete. Vnútorné parametre zohrávajú veľmi dôležitú úlohu pri efektívnom tréningu modelu a vytváraní presných výsledkov. Preto sú používané rôzne optimalizačné stratégie a algoritmy na aktualizáciu a výpočet vhodných a optimálnych parametrov modelu, ktoré ovplyvňujú proces učenia spolu s výstupom modelu siete.

Gradient Descent

Jednou z najdôležitejších techník a základ toho, ako sú tréňované a optimalizované inteligentné systémy, je algoritmus Gradient Descent. Gradient Descent hľadá minimum, kontroluje rozptyl a následne aktualizuje parametre modelu, ktoré vedú k jeho konvergencii. Táto metóda aktualizuje parametre pomocou vzorca [3]:

$$\theta = \theta - \eta \cdot \Delta J(\theta), \quad (1.7)$$

kde η je rýchlosť učenia, $\Delta J(\theta)$ je gradient chybovej funkcie. Najčastejšie sa Gradient Descent používa k aktualizácii váh neurónovej siete, za účelom minimalizácie chybovej funkcie. Uplatňuje sa pri spätnom šírení chyby (backpropagation), v ktorom sa nachádzajú chybové funkcie a aktualizácie hodnôt váh s použitím Gradient Descent, ktoré sa aktualizujú v opačnom smere ako chybová funkcia v doprednom učení siete. V rámci celého súboru dát vykonáva len jednu aktualizáciu váh, preto je pomalý a výpočtovo veľmi náročný.

Stochastic Gradient Descent

Tento optimalizačný algoritmus odstraňuje nedostatky v rýchlosti a výpočtovej náročnosti tradičného algoritmu Gradient Descent. Stochastic Gradient Descent (SGD) iteratívne aktualizuje parametre váh v smere gradientu chybovej funkcie až kým nedosiahne minimum. Na rozdiel od tradičného Gradient Descent nepoužíva celý súbor údajov na výpočet gradientu pri každej iterácii. Používa náhodne vybraný dátový bod z množiny údajov v danej iterácii a presunie sa v smere prechodu k danému bodu. Aktualizácia váh teda prebieha pri výpočte s jedným tréningovým príkladom. Výpočet prebieha pomocou rovnice:

$$\theta = \theta - \eta \cdot \nabla J(\theta; x(i); y(i)), \quad (1.8)$$

kde $x(i); y(i)$ sú príklady tréovania. V prípade, že je chybová funkcia konvexná, je zaručené nájdenie globálneho minima. Dôležitým bodom týkajúceho sa SGD je poradie, v ktorom sú prezentované dáta algoritmu. Ak sú údaje zoradené podľa vzoru, môže to spôsobiť skreslenie, kolísanie gradientu a viesť tak k subjektívnej konvergencii, ktorá spôsobí to, že metóda nebude konvergovať alebo dosiahnutie konvergenzie bude trvať dlhšiu dobu. Vo všeobecnosti je dobrým spôsobom, ako sa tomu vyhnúť, náhodné zoradenie dát pred každou epochou tréovania [39].

Mini Batch Gradient Descent

Problémy s aktualizáciou parametrov s vysokou variabilitou a nestabilnou konvergenciou v prípade SDG je možné odstrániť iným optimalizačným algoritmom nazývaným ako Mini-Batch Gradient Descent. Tento algoritmus aktualizuje váhy pri výpočte s určitou podmnožinou. Tým znižuje rozdiely v aktualizáciách parametrov, čo môže nakoniec viesť k oveľa lepšej a stabilnejšej konvergencii [14].

Momentum

Momentum je optimalizačná metóda, ktorá pomáha urýchliť SDG v príslušnom smere a tlmí kolísanie chybovej funkcie. To robí pridaním zlomku η aktualizáčného vektoru minulého kroku do aktuálneho aktualizáčného vektoru [22]:

$$V(t) = \eta \cdot V(t-1) + \eta \cdot \nabla J(\theta), \quad (1.9)$$

kde hodnota η sa zvyčajne nastavuje na hodnotu 0.99 alebo podobnú hodnotu a parametre váh sú upravované podľa:

$$\theta = \theta - V(t) \quad (1.10)$$

Momentum sa zvyšuje pre rozmery, ktorých gradienty smerujú v rovnakom smere a znižujú sa aktualizácie pre rozmery, ktorých gradienty menia smer. Z tohto dôvodu dosahuje Momentum rýchlejšiu konvergenciu a zníženie oscilácií chybovej funkcie [22].

Existujú ešte ďalšie optimalizačné algoritmy pre hľadanie lokálneho minima ako sú:

- Nesterov [47],
- AdaGrad [7],
- AdaDelta [70],
- RMSProp [62].

V predchádzajúcich optimalizačných algoritmoch boli počítané rôzne hodnoty učenia pre každý parameter. Taktiež bola počítaná dynamika (Momentum) a rýchlosť učenia siete. Existuje modifikovaná technika, Adaptive Moment Estimation,

ktorá počíta jednotlivé zmeny dynamiky pre každý parameter a ukladá ich samostatne.

Adam

Adaptive Moment Estimation (Adam) je ďalšia metóda, ktorá vypočítava adaptívnu rýchlosť učenia pre každý jeden parameter. Okrem ukladania exponenciálne klesajúceho priemeru predchádzajúcich štvorcových gradientov, ako je to v prípade AdaDelta [70], Adam tiež zachováva exponenciálne klesajúci priemer minulých sklonov. Predbežný odhad prvého momentu a druhého momentu odchýlky gradientov je vypočítaný ako:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad (1.11)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \quad (1.12)$$

kde \hat{m}_t je predbežný odhad prvého momentu, m_t je prvý moment, \hat{v}_t je druhý odhad momentu, v_t je druhý moment a β_1^t a β_2^t sú hyperparametre, ktoré riadia exponenciálne rýchlosti poklesu v odhade momentu.

Konečný výpočet pre aktualizáciu parametrov je:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \cdot \hat{m}_t, \quad (1.13)$$

kde θ_t je výsledný parameter, ktorý aktualizuje parametre odhadu prvého a druhého momentu.

Optimalizačný algoritmus Adam dosahuje v praxi veľmi dobré výsledky, v porovnaní s inými adaptačnými algoritmi metódy učenia. Použitím Adam optimalizácie model siete veľmi rýchlo konverguje a rýchlosť učenia je tiež efektívna a rýchla, čím napravnúje každý problém, s ktorým sa stretávajú iné optimalizačné techniky, ako je pomalá konvergencia alebo vysoká odchýlka v aktualizáciách parametrov, ktorá vedie k výkyvom chybovej funkcie [35].

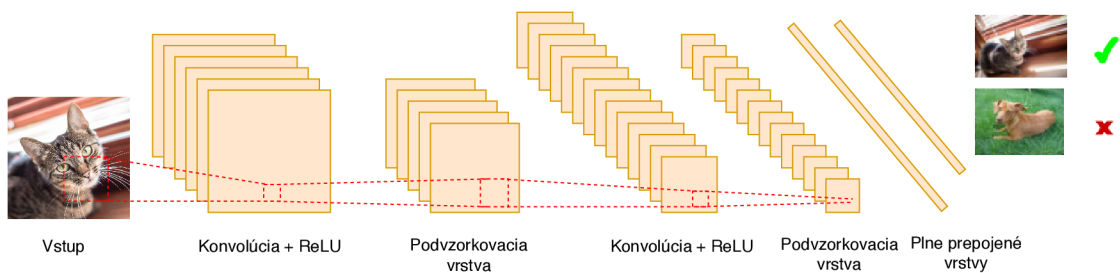
2 HLBOKÉ NEURÓNOVÉ SIETE

Hlboké učenie (deep learning) je metóda strojového učenia, ktorá získava užitočné dáta pomocou mnohovýchvrstvových nelineárnych výpočtových modelov. Modely využívajú na trénovanie veľké množstvo dát, rádovo stovky až tisíce snímok, a ich architektúra sa skladá z hlbokých neurónových sietí. Počet snímok je závislý od toho, či je sieť trénovaná od začiatku alebo je využívaný model predtrénovanej siete. Vďaka súčasnemu technickému stavu sa metódy učenia pomocou hlbokých neurónových sietí z roka na rok dramaticky zlepšujú v oblasti rozpoznávania reči, vizuálnych objektov, detekcii objektov a mnohých ďalších oblastiach. Taktiež priniesli prielom v oblasti spracovania obrazov, videí, reči a zvuku [40].

2.1 Konvolučné neurónové siete

Konvolučné neurónové siete - CNN alebo ConvNet (Convolutional neural networks) sú veľmi podobné klasickým neurónovým sieťam. Sú to klasické dopredné siete, ktoré sa vyznačujú veľkým počtom vrstiev, prípadne modifikáciami v ich prepojení. CNN sú vhodné najmä pre spracovanie 2-D obrazových dát. Ich základom sú vrstvy, ktoré vykonávajú 2-D konvulciu vstupných dát s preddefinovanými filtermi.

Štandardná architektúra konvolučných neurónových sietí (Obr. 2.1) pozostáva zo vstupnej vrstvy, konvolučnej vrstvy (convolutional layers), najčastejšie používanej aktivačnej funkcie ReLU (rectified linearity unit), podvzorkujúcej vrstvy (subsampling layers) a plne prepojenej vrstvy (fully connected layers - FC). Podľa použitia sa štruktúra siete môže od popísanej architektúry líšiť.

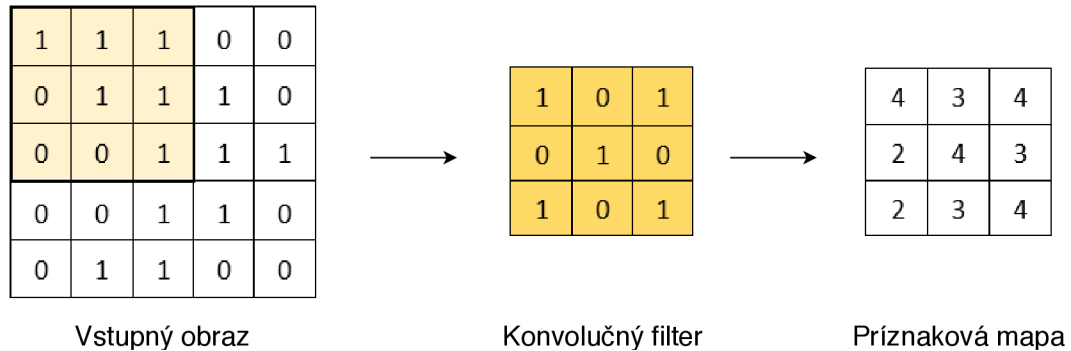


Obr. 2.1: Architektúra konvolučnej neurónovej siete s jednotlivými vrstvami

Do vstupnej vrstvy vstupujú nespracované hodnoty pixelov obrazu, ktoré určujú rozmery vstupu konvolučnej vrstvy, v tomto prípade obraz so šírkou, výškou a hĺbkou vyjadrenou v podobe troch farebných kanálov RGB. V skutočnosti nie je možné vytvoriť viacrozmernú konvolučnú neurónovú sieť, pretože príznakové mapy na konci konvolučnej vrstvy sú jednorozmerné matice. Viacrozmernosť konvolučnej

neurónovej siete je určená hlĺbkou vstupného jednorozmerného obrazu. Príkladom je RGB model pozostávajúci zo zložiek R (červená), G (zelená) a B (modrá), kde sieť tieto zložky spracúva po jednotlivých vrstvách [11].

Konvolučná vrstva a jej princíp je ilustrovaný na Obr. 2.2, kde je definovaný vstupný obraz prostredníctvom matice pixelov s rozmerom 5x5.



Obr. 2.2: Princíp konvolúcie

Konvolučný filter (kernel) je neoddeliteľnou súčasťou vrstevnatej architektúry. Vo všeobecnosti sa aplikuje na celý obraz tak, že transformuje informácie zakódované v pixeloch. Filter (K) je definovaný maskou rozmerov 3x3, ktorá postupne prechádza zľava doprava a zhora nadol celým vstupným obrazom a získava tak príznaky, teda schopnosť extrahovať podstatné informácie. Hodnoty masky filtra sú násobené hodnotami vstupného obrazu nachádzajúcimi sa pod maskou, ktoré pozostávajú z reálnych hodnôt. Súčtom týchto skalárnych súčinov je získaná príznaková mapa. Výsledný počet príznakových máp závisí na počte aplikovaných filtrov. Hodnoty filtra matice sa menia každou iteráciou na tréningovej sade, čo naznačuje, že sieť sa učí identifikovať, ktoré oblasti sú dôležité pre extrahovanie funkcií z dát. Na jeden vstupný obraz sa väčšinou aplikuje viacero filtrov s rôznymi hodnotami vektorových váh [33].

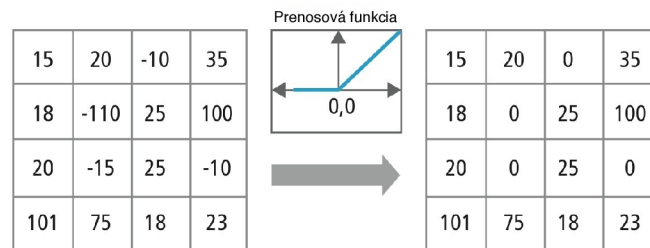
Výstup konvolučnej vrstvy závisí na troch hyperparametroch, ktoré určujú výstupnú veľkosť vrstvy. Prvým hyperparametrom je počet konvolučných filtrov, pričom každý filter sa učí hľadať iné príznaky na vstupe. Druhým hyperparametrom je definovanie kroku (stride), ktorým je filter posúvaný. Ak je krok posunu definovaný na hodnotu 1, filter je posunutý naraz o 1 pixel. Ak je krok definovaný na hodnotu 2, tak filtre pri posuve skočia o 2 pixely naraz. Tým vzniknú priestorovo menšie vstupné objemy. Tretím hyperparametrom je symetrické pridávanie núl do vstupnej matice (zero-padding). Používa sa na presné zachovanie veľkosti vstupného objemu, takže vstupná a výstupná šírka budú rovnaké. Zero-padding (P) je symetrické pridávanie

núl do vstupnej matice. Je to bežne používaná modifikácia, ktorá umožňuje prispôbiť veľkosť vstupu podľa požiadavky. Používa sa hlavne pri navrhovaní CNN, kde je potrebné zachovať rozmery vstupného rozmeru vo výstupe [32].

Vrstva ReLU je najčastejšie používanou aktivačnou funkciou, ktorá nasleduje po konvolučnej vrstve. Táto vrstva používa ako elementárnu aktivačnú funkciu hľadanie prahovej hodnoty:

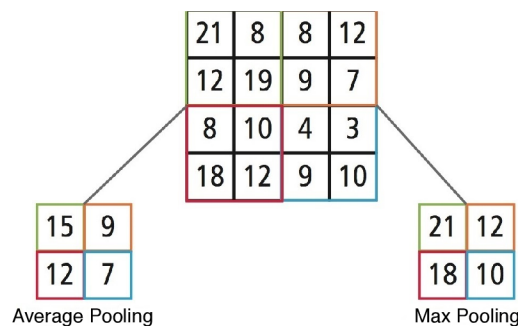
$$f = \max(0, x) \quad (2.1)$$

Touto vrstvou sa zvyšuje vlastnosť rozhodovacej funkcie a celkovej siete bez ovplyvnenia recepčných polí konvolučnej vrstvy. Výhodou ReLU oproti iným nelineárnym funkciám používanými v CNN (napr. hyperbolický tangens, absolútna hyperbolická dotyčnica a sigmoida) je to, že sieť trénuje mnohokrát rýchlejšie. Funkcia ReLU je znázornená na Obr. 2.3, pričom jej prenosová funkcia je vynesená nad šípkou [13].



Obr. 2.3: ReLU vrstva. Prevzané z [13]

Podvzorkovacia vrstva sa konvenčne vkladá medzi dve konvolučné vrstvy. Táto vrstva znižuje v konvolučnej vrstve príznakové pole funkciami Average pooling alebo Max pooling (Obr. 2.4). Najčastejšie sa používa funkcia Max pooling

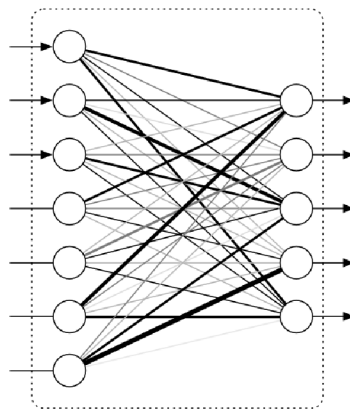


Obr. 2.4: Pooling vrstva. Prevzané z [13]

”hľadanie maxima”, pri ktorej funkcia prechádza príznakovú mapu a vyberá maximálnu hodnotu zvyčajne zo štyroch hodnôt v príznakovej mape, t.j. veľkosti oblasti

2x2 pixelov. Pri funkcii Average pooling sa vyberie priemerná hodnota z daných 4 hodnôt v príznakovej mape. Táto vrstva explicitne neslúži k učeniu neurónovej siete, ale jej úlohou je redukcia priestorovej oblasti a zvýšenie výkonu siete. Taktiež slúži proti pretrénovaniu danej siete [13].

Plne prepojená vrstva sa často používa ako konečná vrstva konvolučných neurónových sietí. Vstup do tejto vrstvy je vektor čísel. Každý zo vstupov je pripojený ku každému z výstupov. Tieto spojenia majú váhy, ktoré určujú dôležitosť daného neurónu. Na Obr. 2.5 sú váhy spojení jednotlivých neurónov ilustrované silnými a tenkými čiarami podľa ich dôležitosti. Plne prepojené vrstvy už nie sú priestorovo umiestnené (je možné ich vizualizovať ako jednorozmerné), takže po plne prepojenej vrstve nemôžu nasledovať žiadne konvolučné vrstvy [11].



Obr. 2.5: Plne prepojená vrstva. Prevzané z [17]

2.2 Plné prepojené konvolučné siete

Plné prepojené konvolučné siete (Fully connected convolutional neural network – FCN) sú také siete, kde všetky vrstvy, ktoré sa učia, sú konvolučné, takže neobsahujú žiadnu plne prepojenú vrstvu. Tieto siete je vhodné použiť pre segmentáciu obrazov, ktoré majú podobné rozloženie. Naopak nevýhodou týchto sietí je vysoký počet váhových parametrov, pri ktorých dochádza k zvýšeniu výpočtovej náročnosti.

2.3 Učenie hlbokých neurónových sietí

Hlboké neurónové siete využívajú najčastejšie k učeniu algoritmus spätného šírenia chyby (backpropagation), ktoré je rozobraté v kapitole 1.3. Medzi chybové funkcie používané pri segmentácii patrí kvadratická chybová funkcia a krížová entropia. Konvolučné vrstvy sa na začiatku siete učia pomalšie ako vrstvy na konci siete, pretože

obsahujú väčšie množstvo filtrov. Napríklad prvá vrstva v sieti sa učí rozpoznávať hrany, ďalšia vrstva rozpoznáva základné tvary, ďalšia vrstva rozpoznáva zložitejšie tvary atď. V závere je hlboká neurónová sieť schopná klasifikovať konkrétne typy úloh ako je rozpoznávanie rôznych objektov v obraze, segmentácia objektov v obraze a iné. Výstupy týchto filtrov sú dočasne ukladané práve kvôli učeniu pomocou spätného šírenia chyby.

V praxi nie je bežné vytvárať celú hlbokú neurónovú sieť od začiatku aj s náhodnou inicializáciou z dôvodu deficitu súboru dát dostatočnej veľkosti. Navyiac takéto obsiahle súbory dát a výpočtové stanice majú len veľké korporáčne spoločnosti na svete. Namiesto toho sa bežne používa učenie pod pojmom *Transfer learning*. Je to technika, ktorá využíva už natrénovanú sieť na veľkom súbore dát a ďalej sa používa na *doladenie* (Fine-tuning), alebo ako pevný extraktor pre konkrétnu úlohu. Existuje niekoľko spôsobov techniky doladovania konvolučnej siete [6]:

- **Mechanizmus extrakcie vlastností** môže v predtrénovanej sieti odstrániť výstupnú vrstvu a potom používať celú sieť ako pevný extraktor pre nový súbor dát.
- **Použitie predtrénovaného modelu**, na ktorom sú inicializované váhy náhodne a tento model je znovu použitý na nový súbor dát.
- **Trénovanie niektorých vrstiev**, pričom iné ostanú ponechané. Podstatou je udržanie váh počiatočných vrstiev architektúry siete, pričom vyššie vrstvy sú menené. Zachovávanie vrstiev je podmienené tým, že sú natrénované na obecné úlohy ako detekcia hrán, či farebných bodov, čo môže byť užitočné pre mnoho úloh.

Pri doladovaní siete je nutné klásť dôraz na použitú trénovanú sadu, ktorá môže byť nasledujúca [6]:

Veľkosť dát je malá, zatiaľ čo podobnosť s dátami v predtrénovanej sieti je veľmi vysoká. V takom prípade sa architektúra siete nemusí pretvárať, nutné je len prispôbiť a upraviť výstupné vrstvy siete podľa požadovanej úlohy.

Veľkosť dát je malá, zatiaľ čo podobnosť s dátami v predtrénovanej sieti je veľmi nízka. Keďže nové vstupné dáta majú nízku podobnosť, je dôležité prispôbiť vyššie vrstvy podľa nového súboru dát. V tomto prípade je malá veľkosť dát kompenzovaná uchovaním počiatočných vrstiev predtrénovanej siete spolu s ich váhami.

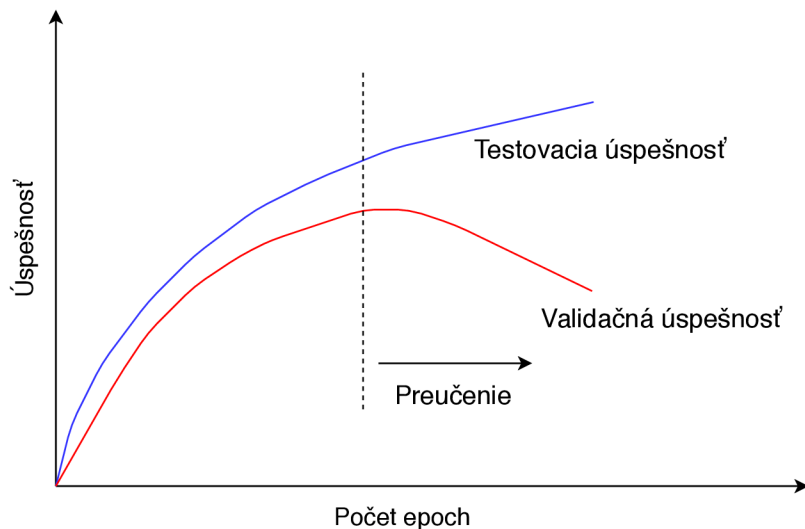
Veľkosť dát je veľká, zatiaľ čo podobnosť s dátami v predtrénovanej sieti je veľmi nízka. Z dôvodu rozsiahlej množiny dát bude tréning neurónovej siete účinný. Avšak keďže sú dáta veľmi odlišné od dát, ktoré boli použité v predtrénovanej sieti, je najlepšie trénovať neurónovú sieť od nuly podľa vlastných dát. Predikcia vykonaná pomocou predtrénovaných dát siete by nebola účinná.

Veľkosť dát je veľká a podobnosť s dátami v predtrénovanej sieti je veľmi vysoká. Jedná sa o ideálnu situáciu, kde by mala byť architektúra predtrénovanej siete najefektívnejšia. Najlepším spôsobom je zachovať architektúru siete a inicializovať počiatočné váhy. Potom je možné tento model trénovať pomocou váh, ktoré boli inicializované v predtrénovanej sieti.

Vzhľadom k výpočtovej náročnosti sietí, ktorých tréning trvá niekedy aj týždne, začali výskumníci zverejňovať ich konvolučné siete, ktoré je možné použiť na dolaďovanie vlastných úloh. Takým príkladom je Model Zoo, kde sú zverejnené váhy, ktoré použili na tréning siete iní výskumníci [33].

2.4 Techniky minimalizácie syndrómu preučenia

Hlboké neurónové siete obsahujú niekoľko nelineárnych skrytých vrstiev, a to z nich robí veľmi expresívne modely, ktoré sa môžu naučiť veľmi zložité vzťahy medzi svojimi vstupmi a výstupmi. Tým môže dôjsť k ich pretrénovaniu. Pretrénovanie siete, nazývané overfitting, nastáva, keď sa tréningová úspešnosť stále zvyšuje, na rozdiel od validačnej úspešnosti, ktorá klesá. Ukážka pretrénovania siete je znázornená na Obr. 2.6.



Obr. 2.6: Ukážkový prípad pretrénovania siete

Niektoré optimalizačné metódy zabráňujú takémuto pretrénovaniu siete, naopak iné sú koncipované na zníženie pamäťovej náročnosti, či zvýšenie rýchlosti tréningu danej neurónovej siete. Najbežnejšie spôsoby, ako zabrániť pretrénovaniu siete, je zníženie kapacity siete, zvýšenie tréningovej sady, tréning na pod-obrazoch,

regulačná metóda dropout, regularizácia pomocou váh a optimalizačné algoritmy, ktoré sú uvedené v 1.4.

Zníženie kapacity siete

Najjednoduchším spôsobom, ako predísť preučeniu siete, je znížiť veľkosť architektúry siete, tj. počet vrstiev. Neexistuje vzorec na určenie správneho počtu vrstiev, preto sa všeobecne začína trénovať sieť na pomerne malom počte vrstiev a zvyšovať sa začne empiricky, kým nebude viditeľný znižujúci sa výnos straty validácie, ktorý signalizuje lepší model.

Umelé rozšírenie trénovanej sady

Navýšenie trénovanej sady (Data Augmentation) zvyšuje schopnosť lepšieho natrénovania neurónovej siete a zároveň znižuje možnosť jej pretrénovania. K rozšíreniu trénovanej sady sa uplatňujú rôzne modifikácie obrazov ako je geometrická rotácia, čiastočné zakrývanie, zašumenie, zväčšenie, zmenšenie, transformácia kontrastu alebo farebnosti a iné. Rozšírenie trénovaných dát musí byť vhodne použité tak, aby nevznikali nereálne prípady, ktoré by negatívne ovplyvnili učenie na trénovanej sade.

Trénovanie na pod-obrazoch

Iným spôsobom, ako je umelé rozširovanie trénovanej sady, je učenie neurónovej siete na pod-obrazoch. Vstupný obraz sa rozdelí na niekoľko menších pod-obrazov, čím dôjde k navýšeniu trénovanej sady a zároveň sa zníži výpočtová náročnosť navrhutej siete, pretože sieť bude pracovať s menším vstupným obrazom ako je originál.

Dropout

Táto regulačná metóda slúži k zníženiu pretrénovania neurónových sietí tým, že bráni komplexnej adaptácii na tréningovú sadu. Dropout by mal byť použitý len v priebehu trénovania neurónovej siete, inak by boli hodnotené výsledky stochastické. Metóda Dropout upravuje počet aktivovaných neurónov v danej iterácii. V najjednoduchšom prípade je počet dočasne deaktivovaných neurónov nastavený koeficientom na hodnotu 0.5, čím sa v jednej iterácii deaktivuje polovica neurónov. Voľba deaktivovaných neurónov je náhodná. Pre vstupné neuróny sa uvádza optimálna deaktivácia bližšie k hodnotám 1 až 0.5 [56].

Regularizátory

Regularizátory sú optimalizačné funkcie, ktoré môžu byť použité pre zníženie pretrénovanosti siete. Tieto regularizátory sa snažia držať váhy modelu dostatočne nízke zavedením dodatočnej penalizácie do chybovej funkcie. Tým bude model čo najjednoduchší a bude mať vlastnosť dobrej generalizácie [48].

3 METÓDY SEGMENTÁCIE OBRAZU

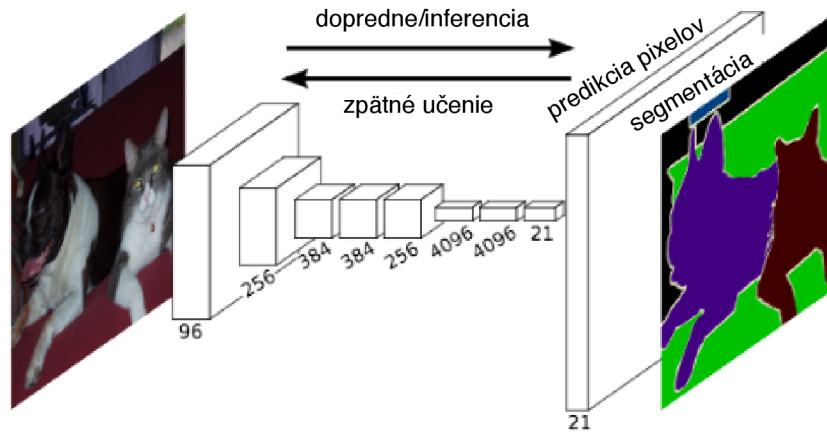
Segmentácia patrí medzi základné metódy analýzy obrazu, kde je potrebné obraz najprv vhodne predspracovať pred jeho ďalšou manipuláciou. Cieľom segmentácie je rozčleniť obraz do častí, ktoré úzko súvisia s oblasťami reálneho sveta alebo majú spoločné príznaky (napr. hranice jednotlivých objektov). Medzi takéto spoločné príznaky patrí napríklad farba, tvar alebo textúra. Hľadaná oblasť v obraze môže byť špecifická, ktorá nadobúda výsledok segmentácie len v binárnych hodnotách, kde hodnota jedna označuje špecifickú oblasť a hodnota nula zodpovedá pozadiu scény v obraze. V opačnom prípade delenia na viacej častí sa jednotlivé segmenty označujú pomocou farebnej škály alebo indexami. Výsledkom segmentácie je súbor vzájomne sa neprekrývajúcich oblastí, ktoré buď jednoznačne korešpondujú s objektami (kompletná segmentácia) alebo nemusia priamo korešpondovať s objektami. Tieto problémy sú ovplyvnené šumom, nerovnomerným osvetlením obrazu, nejednoznačnosťou obrazových dát, zložitnosťou scény, prekrývajúcimi sa objektmi a použitou metódou. Medzi základné metódy segmentácie obrazu patria metódy založené na prahovaní, metódy založené na segmentovaní hrán, metódy orientované na regióny, či hybridné modely, medzi ktoré patria aj neurónové siete [66].

V oblasti segmentácií objektov v obrazoch sa využívajú rôzne typy architektúr sietí založených na hlbokom učení. V tejto časti práce je uvedený prehľad používaných architektúr.

3.1 Plne prepojená sieť pre sémantickú segmentáciu

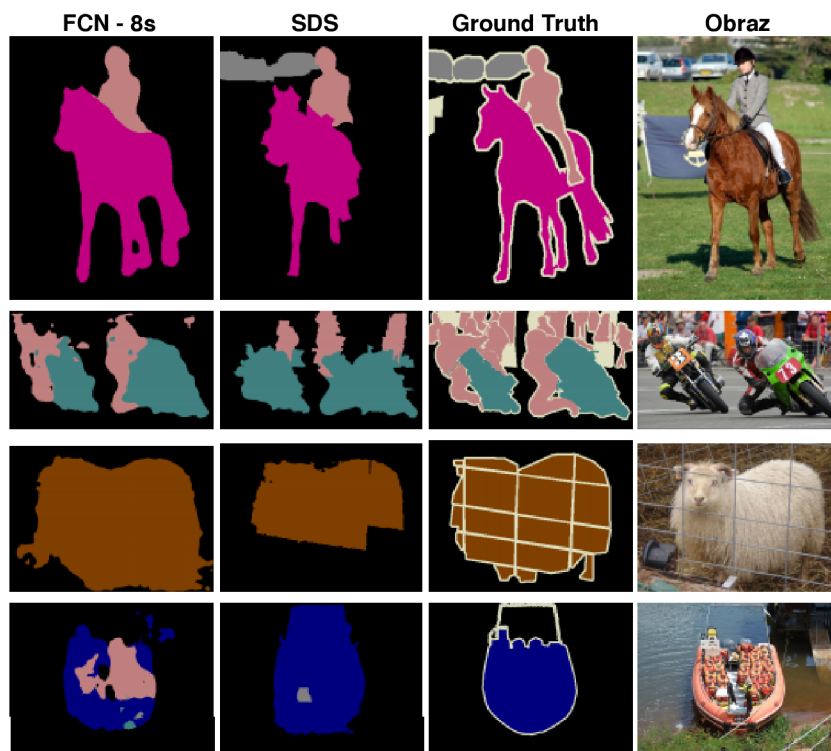
Snahou tejto siete je prispôbiť klasifikačné siete ako sú AlexNet [38] a VGG [54] do plne konvolučnej siete a preniesť ich naučené reprezentácie jemným ladením na úlohu segmentácie. Trénovacia sada obsahuje vstupy ľubovoľnej veľkosti a produkuje výstupný obraz efektívnou interferenciou a učením, ktoré sú vykonávané dopredným a spätným výpočtom siete (viď Obr. 3.1).

Jonathan Long et al. [42] používa v prvých piatich konvolučných vrstvách na doladovanie architektúru siete VGG s predtrénovanými váhami a prahom. Plne prepojené vrstvy sú nahradené konvolučnými vrstvami, aby predišli problému s hustou predikciou ako je segmentácia obrazu. Pri poslednej konvolučnej vrstve je volená veľkosť filtra tak, aby mal výstup z vrstvy rozmer 1x1 pixel a N kanálov. Použitie dekonvolučnej vrstvy vytvorí výsledný obraz o rovnakej veľkosti ako bol vstupný obraz.



Obr. 3.1: Plne prepojená sieť pre sémantickú segmentáciu. Prevzané z [42]

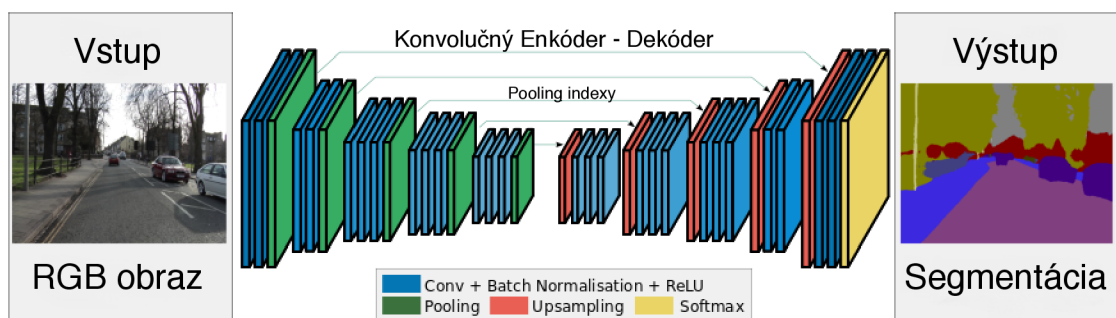
Táto sieť dosahuje na trénovanej sade PASCAL VOC 2011 presnosť 62,7 %, čím J. Long et al. zlepšil výsledky o 20 % ako pri metóde state-of-the-art [42]. Prezentované výsledky sú zobrazené na Obr. 3.2.



Obr. 3.2: Výsledky segmentácie plne prepojenej neurónovej siete. Prevzané z [42]

3.2 SegNet

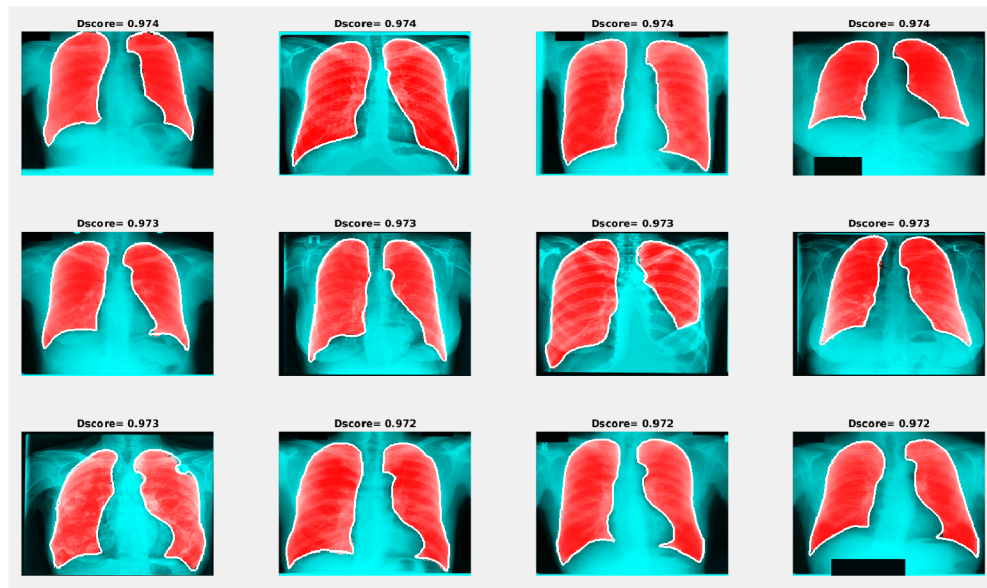
Táto sieť je rozdelená do dvoch častí. V prvej časti nazývanej ako Enkóder je vstupný obraz spracovaný veľkým množstvom konvolučných vrstiev rozdelených do piatich konvolučných blokov, ktoré sú topologické so sieťou VGG16 [54]. Tieto vrstvy slúžia na extrakciu významných príznakov z obrazu, na základe ktorých potom dochádza k učeniu jednotlivých segmentov. Druhá dekódovacia časť siete nazývaná ako Dekóder odpovedá dekonvolúcii, je zrkadlovou štruktúrou prvej časti siete a napája sa do klasifikačnej vrstvy, resp. aktivačnej funkcie Softmax, viď Tab. 1.1, ktorá sa využíva pre klasifikáciu segmentácie obrazu. Sieťová architektúra má rovnaký počet kódovacích a dekódovacích vrstiev. [1] Architektúra siete je zobrazená na Obr. 3.3.



Obr. 3.3: Architektúra SegNet. Prevzané z [1]

3.3 ED-CNN

Táto hybridná konvolučná neurónová kódovacia-dekódovacia sieť (Encoder-Decoder CNN) bola vytvorená A. Krizanovskym et al. [30]. Jej cieľom je segmentácia pľúc v snímkach RTG obrazu hrudníka. Základným prvkom tejto siete je architektúra SegNet. Líši sa počtom použitých konvolučných blokov. V tomto prípade pozostáva architektúra zo štyroch konvolučných blokov v časti enkódera a štyroch blokov v dekódovacej časti. Na vyhodnotenie úspešnosti tejto siete bola využitá metrika Dice koeficient, viď kapitola 4.1, kde priemerná presnosť segmentácie pľúc dosahuje 0,962 pri minimálnom skóre 0,926 a maximálnom skóre 0,974 s odchýlkou 0,008. Na Obr. 3.4 sú znázornené výsledky najlepších segmentácií pľúc z RTG snímk [30].

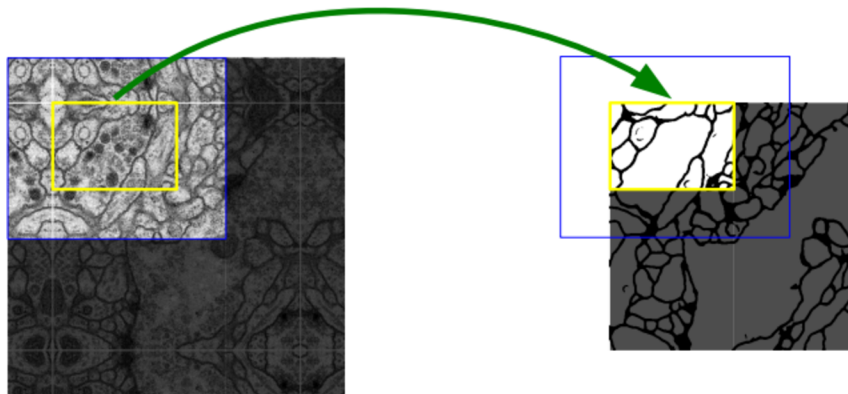


Obr. 3.4: Výsledky segmentácie pľúc architektúrou ED-CNN. Prevzané z [30]

3.4 U-net

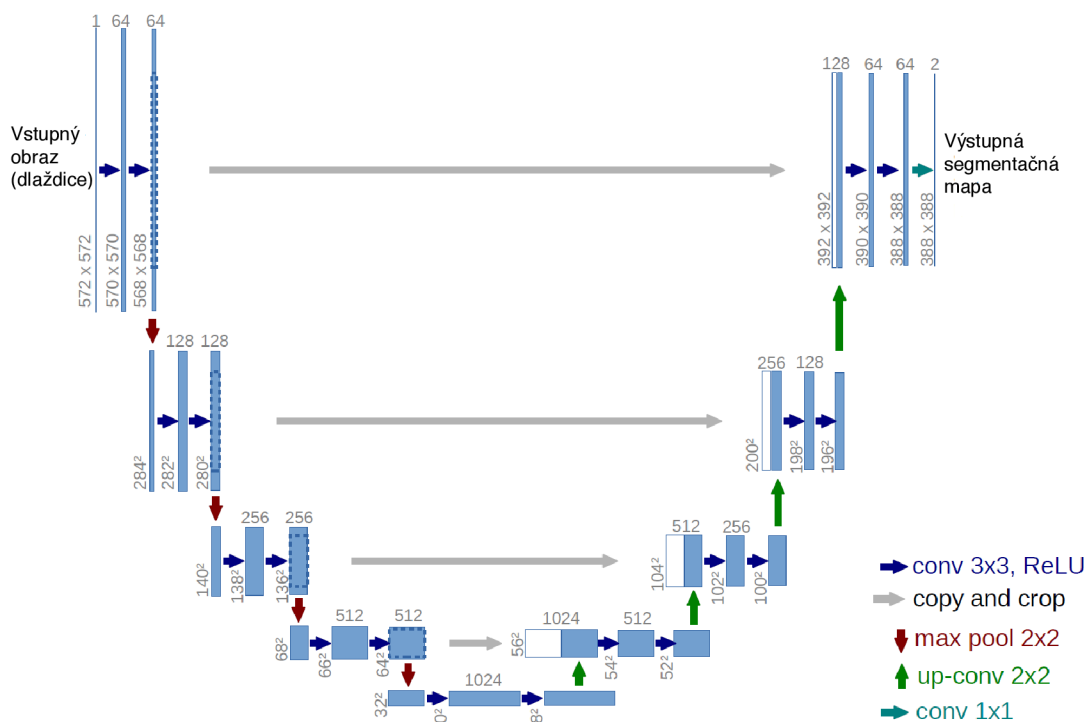
Táto architektúra je postavená na plne konvolučnej sieti, ktorá pracuje s veľmi malým počtom tréovaných obrazov a prináša presnejšie segmentácie. U-net neobsahuje piaty konvolučný blok ako v architektúre VGG, pričom Segnet využíva všetky váhy konvolučných vrstiev z VGG ako predtrénované váhy pre vlastnú architektúru. [1] Sieť sa skladá z kontraktilnej (vľavo) a expanzívnej dráhy (vpravo). Kontraktilná dráha je viac-menej symetrická k expanzívnej dráhe a výsledkom je architektúra v tvare písmena U, viď Obr. 3.6. Sieť neobsahuje žiadne plne prepojené vrstvy a používa iba validnú časť každej konvolučnej vrstvy, t.j. segmentačnú mapu obsahujúcu pixely, pre ktoré je na vstupnom obraze k dispozícii celý kontext. Táto stratégia umožňuje bezproblémovo segmentovať ľubovoľne veľké obrazy pomocou prekrývajúcich sa dlaždíc (Obr. 3.5). Na predikciu pixelov v oblasti hraníc obrazu (žltá oblasť) je chýbajúci kontext (modrá oblasť) extrapolovaný zrkadlom vstupného obrazu. To je dôležité pre aplikáciu siete na veľké obrazy, pretože inak by bolo rozlíšenie obmedzené pamäťou GPU. [52]

U-net používa pre rozšírenie dát elastické deformácie na dostupné tréované obrazy. To umožňuje sieti naučiť sa invariantnosť k takýmto deformáciám bez potreby videnia tejto transformácie v súhrne anotovaných obrazov. Toto je obzvlášť dôležité v biomedicínskej segmentácii, pretože deformácia sa používa na najbežnejšie tkanivá a realistické deformácie môžu byť efektívne simulované. Architektúra U-net dosahuje veľmi dobrú výkonnosť pri veľmi odlišných biomedicínskych segmentačných aplikáciách [52].



Obr. 3.5: Výsledky predikcie pixelov architektúrou U-net

V rámci ISBI cell tracking challenge 2014 and 2015 dosiahla sieť priemernú hodnotu IOU (intersection over union) 92 % pri prvej trénovacej sade obsahujúcej glioblastóm-astrocytóm U373 buniek na polyakrylimidovom substráte zaznamenanom fázovým kontrastným mikroskopom. Pri druhej trénovacej sade buniek DIC-HeLa na plochom skle zaznamenaných difrakčným interferenčným mikroskopom dosiahla sieť priemernú hodnotu IOU 77,5 % [52].



Obr. 3.6: Architektúra siete U-net (vľavo kontraktívna dráha, vpravo expanzívna dráha). Prevzané z [52]

4 TESTOVANIE VYBRANÝCH ÚLOH

Táto kapitola popisuje využitie existujúcich architektúr konvolučných neurónových sietí napísaných v programovacom jazyku Python s využitím prostredia Tensorflow a Keras. Prostredie Keras je momentálne už súčasťou knižnice Tensorflow. V podkapitole 4.1 sú uvedené metriky, ktoré sú využívané k vyhodnocovaniu výsledkov v tejto práci. K naštudovaniu problematiky boli použité architektúry neurónových klasifikačných sietí, popísané v podkapitole 4.2, ktoré slúžili ako opora pre implementáciu konvolučnej neurónovej siete s cieľom segmentácie. Na databáze MNIST bolo testované pridávanie vrstiev siete do základnej architektúry siete. V prípade databázy Dogs vs. Cats competition bola využívaná predtrénovaná sieť na iných dátach, a následne bola doladená pre klasifikáciu mačiek a psov. V podkapitole 4.3 sú popísané dva prístupy segmentácie, ktoré sú inšpirované architektúrou U-net. Prvá architektúra je trébovaná a testovaná na databáze Ultrasound nerve segmentation, za účelom segmentácie nervov Brachial plexus v ultrazvukovom obraze. Na základe nízkych výsledkov segmentácie je testovaná druhá architektúra, ktorá je trébovaná a testovaná na snímkach nervovej sústavy v oblasti bruška drozofili. Úlohy boli testované pomocou výpočetných prostriedkov MetaCentra [46].

Tensorflow

Tensorflow je open source softwarová knižnica, ktorá slúži pre numerické výpočty založené na grafovej štruktúre. Jednotlivé uzly v grafe reprezentujú grafické operácie a hrany reprezentujú multidimenzionálne dátové polia (tenzory), ktoré popisujú vzťahy medzi uzlami a komunikujú medzi sebou. Tensorflow bola pôvodne vyvinutá tímom Google Brain Team spoločnosti Google na účely výskumu strojového učenia a hlbokých neurónových sietí, pričom je verejne prístupná pre vývojárov, ktorí ju využívajú pri tvorbe nových aplikácií. Týmto je vývoj platformy Tensorflow mnohonásobne progresívnejší a prináša tak nové možnosti implementácie strojového učenia. Google uvádza, že systém platformy je natoľko univerzálny, že môže byť použiteľný dostatočne dobre aj v iných oblastiach ako sú neurónové siete. Základná vrstva platformy Tensorflow bola naprogramovaná v jazyku C++, každopádne podporuje tvorbu aplikácií aj v jazyku Python. Veľkou výhodou Tensorflow je, že ponúka možnosť výpočtu nielen na centrálnom procesore (CPU), ale aj na grafickej karte (GPU). Vďaka tomu je možné výpočet mnohonásobne urýchliť a znížiť časovú náročnosť úloh [60].

Keras

Keras je knižnica napísaná v programovacom jazyku Python, ktorá je vyvinutá pre modelovanie konvolučných neurónových sietí. Je zameraná na jednoduché používanie s vysokým výpočtovým výkonom. Podporuje výpočty na CPU a GPU.

Je podporovaná platformou Theano a Tensorflow. Pre jednoduchšie používanie Kerasu v Tensorflow došlo v poslednej aktualizácii k prepojeniu zdrojových kódov, čím sa Keras stal jeho ďalším modulom [34].

Deep learning tool kit

Adaptácia hlbokých neurónových sietí na medicínske obrazy vyžaduje špeciálne operácie a jej implementácie dosahujú pomalé pokroky. Deep learning tool kit (DLTK) je nástrojová sada neurónových sietí napísaná v programovacom jazyku Python s použitím knižnice Tensorflow. Bol vyvinutý, aby umožnil rýchle vytváranie prototypov a zaistil tak reprodukovateľnosť v aplikáciách pre analýzu obrazu s osobitným zameraním na medicínske obrazy. Cieľom DLTK je poskytnúť komunite ľudí najnovšie metódy a modely, ktoré urýchlia výskum v tejto oblasti segmentácie medicínskych obrazov [61].

4.1 Vyhodnocovacie metriky

V tejto časti sú popísané metriky F1 skóre a Dice koeficient, ktoré boli použité na vyhodnocovanie pri testovaní segmentačných modelov konvolučných neurónových sietí. Posledná ROC (Receiver Operating Characteristic) krivka je použitá v podkapitole 5.3.2.

F1-skóre

Skóre F1 [12] je interpretovaný ako vážený priemer hodnôt Precision a Recall, kde skóre F1 dosiahne svoju najlepšiu hodnotu pri 1 a najhoršie skóre pri 0. Pre výpočet Precision a Recall budú zavedené nasledujúce pojmy:

- TP (true positive) - počet pixelov správne detegovaných ako segmentovaný objekt
- TN (true negative) - počet pixelov správne detegovaných ako segmentovaný objekt v ground truth
- FP (false positive) - počet pixelov nesprávne detegovaných ako segmentovaný objekt
- FN (false negative) - počet pixelov nesprávne detegovaných ako segmentovaný objekt v ground truth

Precision (PPV) určuje pravdepodobnosť prítomnosti správne pozitívneho segmentu voči všetkým pozitívnym (správne/nesprávne) segmentom a jeho vzorec je nasledujúci:

$$PPV = \frac{TP}{TP + FP} \quad (4.1)$$

Recall, v binárnej klasifikácii nazývaný ako senzitivita (TPR), určuje pravdepodobnosť pozitívneho segmentu voči všetkým správne pozitívnym a nesprávne negatívnym segmentom a jeho vzorec je nasledujúci:

$$SE = \frac{TP}{TP + FN} = TPR \quad (4.2)$$

Pomocou Precision a Recall je skóre F1 definované ako:

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (4.3)$$

Dice koeficient

Táto štatistická metrika je používaná na porovnanie zhody dvoch súborov údajov, konkrétne pri segmentácii snímky ground truth a výsledné predikované snímky. Pôvodný vzorec pre množiny X a Y je definovaný ako [67]:

$$D(X, Y) = \frac{|X| \cap |Y|}{2 \cdot |X| + |Y|}, \quad (4.4)$$

kde X je predikovaná sada pixlov a Y je ground truth. Dice koeficient je definovaný ako 1, ak sú X a Y prázdne. Výsledné bodové skóre je priemer Dice koeficientov pre každý obrázok v tréningovom, prípadne testovacom súbore.

Tento vzťah môže byť vyjadrený v termínoch TP a FN ako:

$$D = \frac{2 \cdot TP}{(FP + TP) + (TP + FN)} \quad (4.5)$$

Vyššie číslo výsledného Dice koeficientu definuje zhodu týchto dvoch množín. V prípade segmentácie to znamená, že výsledok sa zhoduje s ground truth viac než v prípade výsledkov, pri ktorých by bol Dice coefficient nižší.

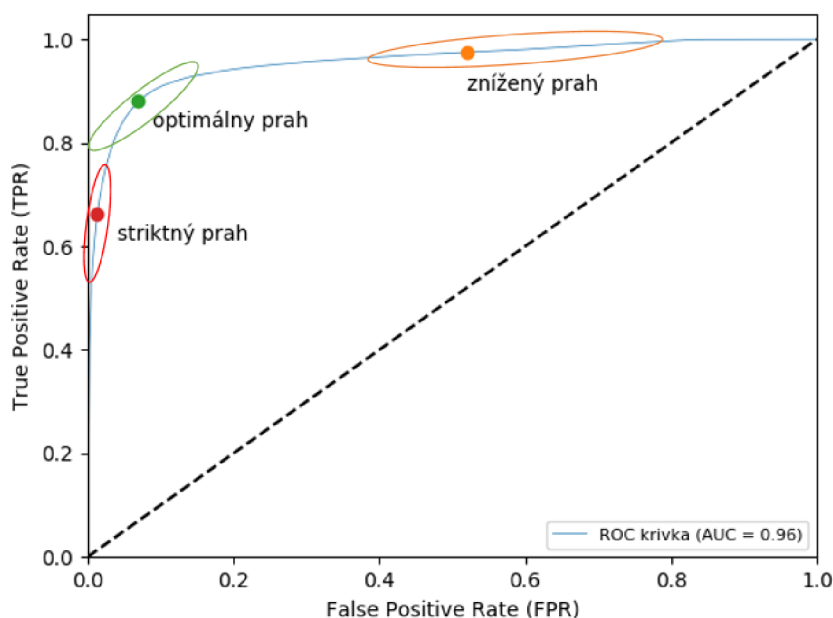
ROC krivka

ROC (Receiver Operating Characteristic) krivka definuje vzťah medzi senzitivitou (TPR) a špecificitou (TNR) pre všetky prípustné hodnoty prahu. Na Obr. 4.1 je znázornená ROC krivka. Na ose Y je znázornená senzitivita, ktorá je definovaná 4.2 a vyjadruje počet správne pozitívnych segmentov (TPR). Na ose X je znázornený počet nesprávne pozitívnych segmentov, ktorý je definovaný pomocou špecificity (TNR) ako $FPR = 1 - TNR$ [63].

$$SP = \frac{TN}{TN + FP} = TNR \quad (4.6)$$

$$FPR = 1 - \frac{TN}{TN + FP} = \frac{FP}{FP + TN} \quad (4.7)$$

Na ROC krivke sú vyznačené oblasti prahu. Prvá oblasť je znázornená červenou farbou a je to približné rozmedzie striktného prahu, tj. pri tomto prahu je síce výsledná senzitivita nižšia, no špecificita je vysoká. Zelenou farbou je označená oblasť optimálneho prahu. Oranžovou farbou je vyznačená oblasť zníženého prahu, pri ktorom je výsledná senzitivita vysoká, no špecificita nízka.



Obr. 4.1: ROC krivka s umiestnením prahových hodnôt

Poslednou dôležitou analýzou ROC krivky je oblasť pod ROC krivkou (the Area under the ROC curve) označovaná ako AUC, ktorá určuje mieru presnosti testu, v tomto prípade segmentácie. Približná klasifikácia presnosti je uvedená v Tab. 4.1 [59]. Ak je plocha pod krivkou rovná 1, znamená to že senzitivita je 100% a segmentovaný obraz odpovedá ground truth. Ak je plocha pod krivkou 0.5, výsledok presnosti je nedostačujúci.

Tab. 4.1: Rozdelenie oblasti pod ROC krivkou

Hodnoty AUC	Význam
0.50 - 0.60	nedostatočný
0.60 - 0.70	slabý
0.70 - 0.80	dobrý
0.80 - 0.90	veľmi dobrý
0.90 - 1	vynikajúci

Výsledky segmentácie cievného riečiska sú v práci porovnávané s dostupnými metódami výpočtom senzitivity a špecificity. Na základe senzitivity 4.2 a špecificity 4.6 je definovaná výsledná presnosť segmentácie cievného riečiska.

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} \quad (4.8)$$

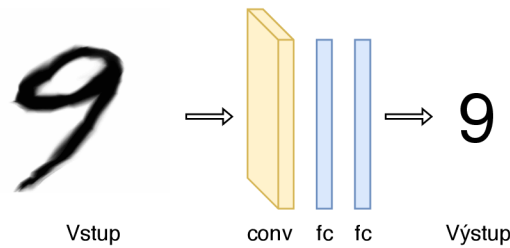
4.2 Modely klasifikačných konvolučných neurónových sietí

Dátové sady pre tréovanie a následné testovanie modelov konvolučných sietí boli získané z platformy Kaggle. [29] Táto platforma ponúka rozsiahle dátové sady, ktoré formou súťaží slúžia na vyriešenie danej problematiky. Snaží sa o vytvorenie najlepšieho prediktívneho modelu a popis dátových súborov, ktoré sú poskytnuté spoločnosťami a užívateľmi.

Pre tréovanie jednoduchej konvulčnej neurónovej siete a hlbšej konvulčnej neurónovej siete bola použitá databáza MNIST (Modified National Institute of Standards and Technology), ktorá bola zverejnená v rámci známych vyučovacích súťaží v oblasti strojového videnia. Databáza MNIST obsahuje ručne písané číslice od 0 po 9 uložené vo formáte .csv, kde každý snímok v dátovej sade má veľkosť 28x28 pixelov. Spolu má 785 stĺpcov, kde prvé číslo v riadku charakterizuje hľadanú číslicu a nasledujúce čísla prislúchajú pixelom definujúce obrázok. Každý pixel je reprezentovaný jedným číslom v rozsahu 0 a 255 vrátane. Úlohou modelov bolo správne identifikovať číslice z dátovej sady [26].

Jednoduchá konvulčná neurónová sieť

Základná architektúra, ktorá bola použitá na klasifikovanie číslic, je znázornená na Obr. 4.2. Obsahuje jednu konvulčnú vrstvu a dve plne prepojené konvulčné vrstvy.



Obr. 4.2: Základná použitá architektúra

Hlbšia konvolučná neurónová sieť

V druhom prípade bol prípad jednoduchej CNN transformovaný pridaním viacerých konvolučných neurónových vrstiev a za nimi nasledujú podvzorkovacie vrstvy, konkrétne Max pooling. V Tab. 4.2 sú zobrazené úspešnosti jednoduchej a hlbšej konvolyčnej neurónovej siete.

Tab. 4.2: Úspešnosť modelov konvolyčných sietí 1

Model	Počet epoch	Úspešnosť		
		Trénovacia	Validačná	Testovacia
Jednoduchá CNN	5	0.945	0.961	0.959
Hlboká CNN	20	0.977	0.977	0.971

Trénovanie siete s nízkym počtom obrazov v dátovej sade je z jedného veľkých nevýhod používania konvolyčných neurónových sietí. V ďalšom príklade je poukávané na niekoľko možností riešenia tohoto problému. Pri experimentovaní s nízkym počtom obrazov v dátovej sade bolo postupované podľa technickej dokumentácie, viď [21]. Dátová sada bola opäť použitá zo súťaže Kaggle, konkrétne z Dogs vs. Cats competition. Úlohou súťaže bolo rozoznať snímok mačky od snímku psa. Dátová sada je pomerne veľká a obsahuje 25 000 snímokov. [27] Pre príklad tréningu siete na menšom počte dát je z tejto dátovej sady vybraných 2000 snímokov (1000 snímokov psov, 1000 snímokov mačiek) pre tréningovú sadu, 800 snímokov pre validačnú sadu a 2000 snímokov pre testovaciu sadu.

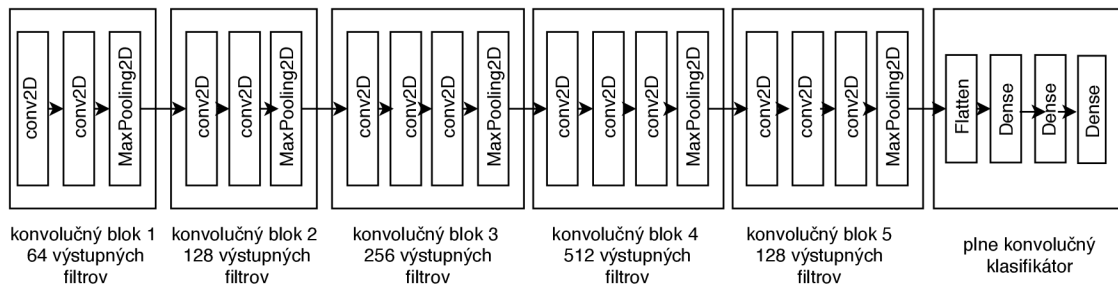
Malá konvolyčná neurónová sieť

Ako základná architektúra (baseline) bola vytvorená malá konvolyčná sieť. Pri takejto štruktúre s nízkym počtom obrazov v dátovej sade dochádza k pretrénovaniu siete, pretože má viac parametrov než dátových bodov pre tréningovanie. Riešením pretrénovania je zavedenie vrstvy Dropout medzi plne prepojené vrstvy. Druhým krokom k zabráneniu pretrénovania je augmentácia dát, ktorou je každý vstupný snímok transformovaný rotovaním a rovinným posúvaním. Augmentácia dát slúži

zároveň na to, aby sa CNN naučila byť invariantná na pozíciu objektu v snímke, tzn. že ju nebude zaujímať, či sa detegovaný objekt (mačka) nachádza v pravom dolnom rohu alebo v ľavom hornom rohu.

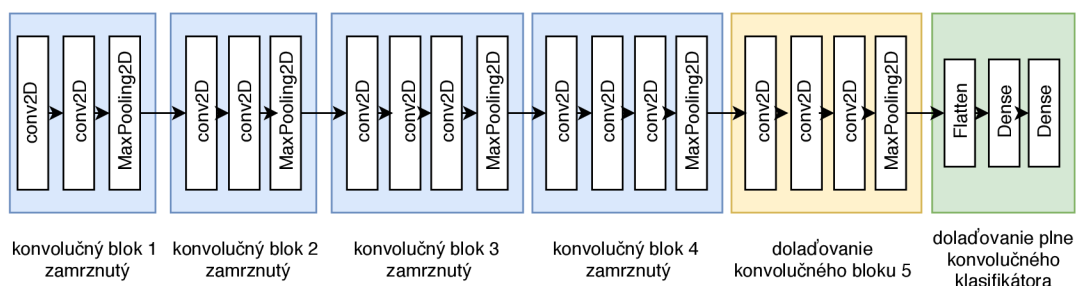
Predtrénovaná sieť

Miesto toho, aby bola trébovaná celá konvolučná sieť, je na trébovanie použitá časť architektúry siete VGG16 (Obr. 4.3), ktorá je natrébovaná na inej dátovej sade, tj. transfer learning. Pri trébovaní budú použité konvulučné vrstvy z tejto siete, aby vyextrahovali dôležité časti snímku užitočné pre klasifikáciu.



Obr. 4.3: Architektúra siete VGG16. Prevzané z [21]

Všetky vrstvy použitej časti architektúry VGG16 okrem poslednej sú zmrazené a nemenia sa. Výnimkou je posledný konvulučný blok modelu VGG16, ktorý je upravený doladením (Fine-tuning) v podobe vlastných plne prepojených vrstiev (Obr.4.4) Tieto vrstvy sú trébované na novej množine dát, v tomto prípade na menšom počte obrazov v dátovej sade mačiek a psov. Týmto sa razantne zníži počet parametrov modelu a taktiež sa zníži aj čas trébovania siete.



Obr. 4.4: Doladenie predtrénovanej siete. Prevzané z [21]

V Tab. 4.3 sú zobrazené výsledky trébovaných sietí. Malá konvulučná neurónová sieť dosahovala pri piatich epochách 70.3 % testovaciu úspešnosť. Zvýšením počtu epoch na dvadsať v malej CNN sa úspešnosť siete zvýšila približne o 6 %. Použitím predtrénovanej siete na menšom datasete bola úspešnosť 89.9 %, čo je skoro o 20 %

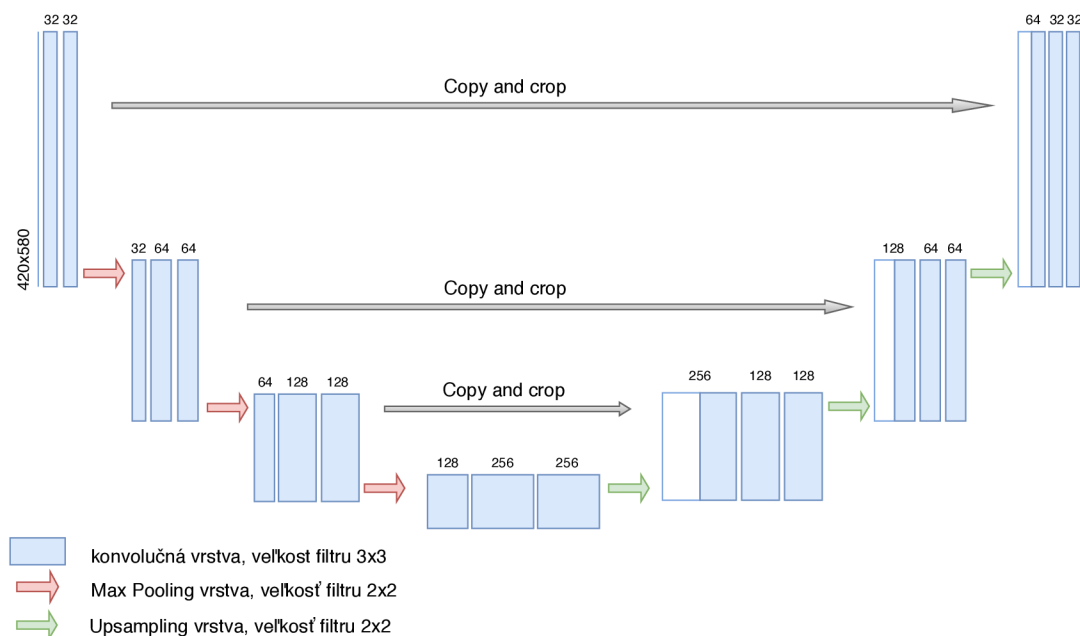
Tab. 4.3: Úspešnosť modelov konvolučných sietí 2

Model	Počet epoch	Úspešnosť		
		Trénovacia	Validačná	Testovacia
Malá CNN	5	0.696	0.712	0.703
Malá CNN	20	0.782	0.756	0.760
Predtrénovaná CNN	20	0.899	0.874	0.899

lepší výsledok ako v pôvodnej malej CNN. Je nutné podotknúť, že využitím len 8 % dát z dátovej sady mačiek a psov sa dokázala sieť natréňovať takmer na 90 % [69].

4.3 Modely segmentačných konvolučných neurónových sietí

Architektúra siete [25] je inšpirovaná U-net, viď Obr. 3.6. Veľkosť vstupných obrázkov a počet konvolučných filtrov sa oproti pôvodnej U-net architektúre líši počtom konvolučných blokov. V tejto architektúre je o jeden konvolučný blok menej v kontraktilnej a expanzívnej dráhe. Rovnako je zmenený aj počet filtrov v jednotlivých vrstvách, ktoré sú na Obr. 4.5 zobrazené nad každou vrstvou.

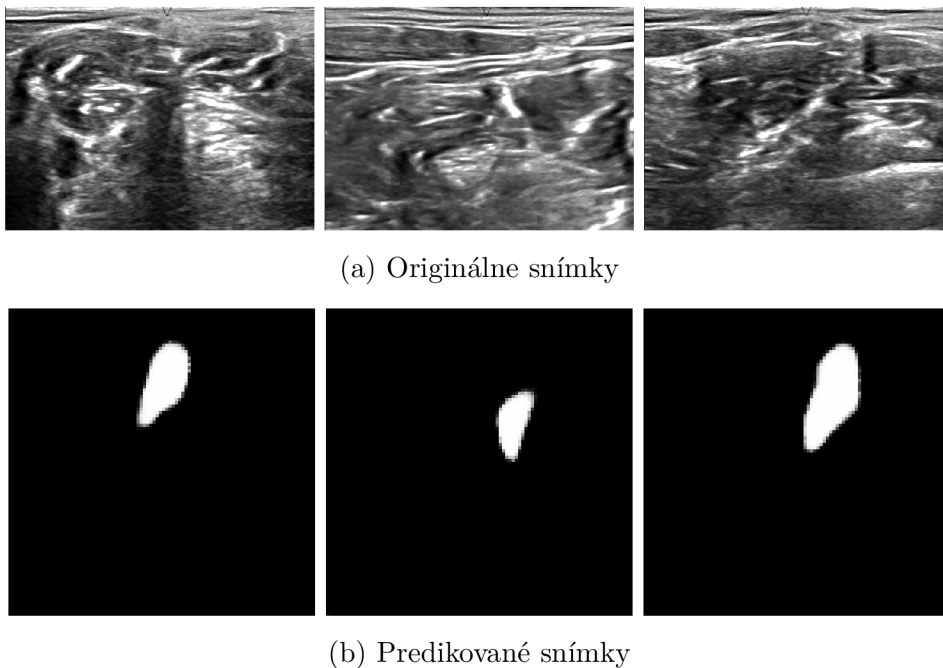


Obr. 4.5: Upravená architektúra siete U-net

Pre tréning a testovanie siete je použitá dátová sada Ultrasound nerve segmen-

tation, za účelom segmentácie nervov Brachial plexus v ultrazvukovom obraze [28]. V tomto prípade je dátová sada rovnako rozdelená do trénovacej, validačnej a testovacej množiny. Trénovacia sada obsahuje 5508 ultrazvukových snímok, ku ktorým patria snímky ručne anotovaných nervov. Súbor trénovacích dát obsahuje aj snímky v ktorých nie sú prítomné nervy Brachial plexus. Z trénovacej sady je odobratých 20 % snímok pre validačnú sadu. Testovacia sada obsahuje ultrazvukové snímky bez anotácie.

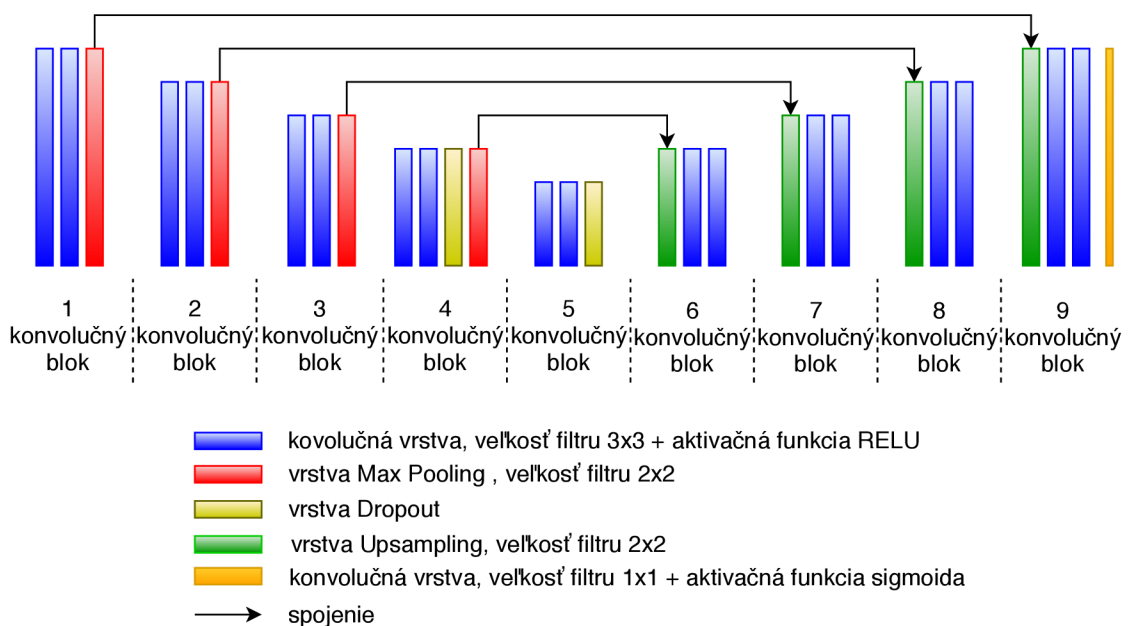
Táto hlboká neurónová sieť je hodnotená pomocou Dice coefficient, viď 4.1, kde výsledná trénovacia úspešnosť po 20 epochách bola 0.68. Pri prepočte na F1-score, viď kapitola 4.1, dosahuje úspešnosť 0.57, čo znamená, že pravdepodobne došlo k jej pretrénovaniu. Váhy sú aktualizované pomocou Adam optimalizácie s rýchlosťou učenia 0.00001. Počas tréningu boli váhy ukladané do súboru hdf5. Dátová sada nebola pred trénovaním nijako upravená. Zmenila sa len veľkosť vstupných snímok na rozmer 96x96 pixelov. Dôkladnejším predspracovaním dát by model siete pravdepodobne dosahoval lepšie výsledky. Výsledok segmentácie nervov je zobrazený na Obr. 4.6, kde je vidieť pôvodný obraz a k nemu prisluchajúci segment nervov.



Obr. 4.6: Ukážka segmentácie nervov upravenej architektúry U-net

Ďalšia testovaná architektúra [71] je taktiež inšpirovaná U-net. Architektúra siete pozostáva z kontraktilnej dráhy (Obr. 3.6), ktorá obsahuje typickú architektúru konvolučnej siete. Skladá sa z opakovaného použitia dvoch konvolučných vrstiev veľkosti 3x3, za ktorými nasleduje vrstva ReLU a podvzorkovacia vrstva Max pooling veľkosti 2x2 s krokom 2. Pri každom kroku sa zdvojnásobí počet príznakov. Každý krok

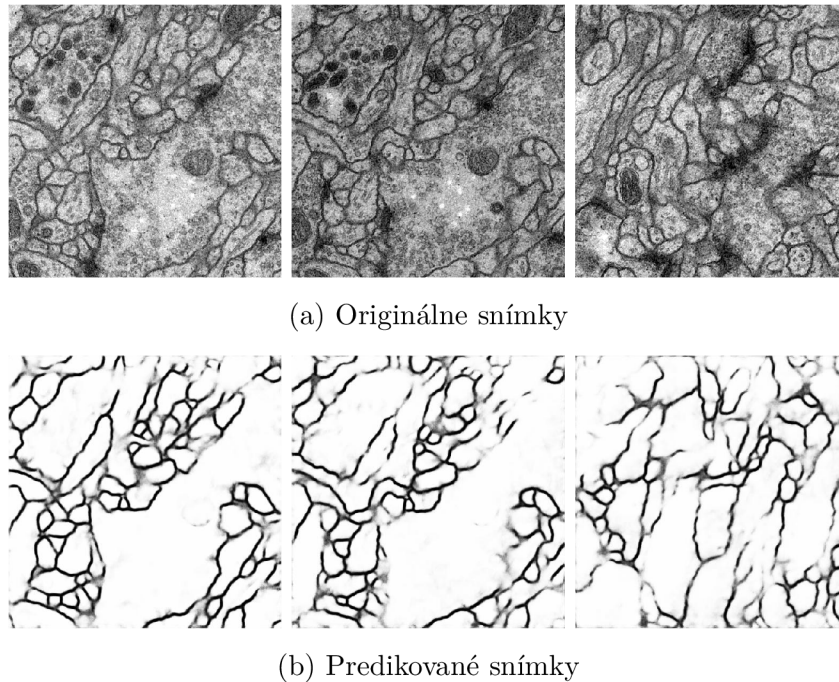
v expanzívnej dráhe pozostáva z vrstvy Upsampling. Tá zredukuje polovicu príznakov a spojí zodpovedajúcu orezanú príznakovú mapu z kontraktilnej dráhy s prislúchajúcou expanzívnu časťou dráhy. Po nej nasledujú opäť dve konvolučné vrstvy veľkosti 3x3, kde po každej jednej nasleduje vrstva ReLU. Celkovo sieť obsahuje 23 konvolučných vrstiev. Počet filtrov jednotlivých vrstiev je identický s pôvodnou architektúrou U-net, viď Obr. 3.6, kde je počet filtrov vyznačený nad každou jednou vrstvou. Počet konvolučných blokov je tiež identický s pôvodnou architektúrou U-net. Rozdiel je v tom, že výsledný predikovaný obraz má na výstupe rovnakú veľkosť ako na vstupe. Taktiež je v štvrtom a piatom konvolučnom bloku pridaná vrstva Dropout, slúžiaca proti preučeniu siete. Architektúra siete je znázornená na Obr. 4.7 a je otestovaná na databáze 30 snímok nervovej sústavy v oblasti bruška drozofily [4].



Obr. 4.7: Architektúra siete inšpirovaná U-net

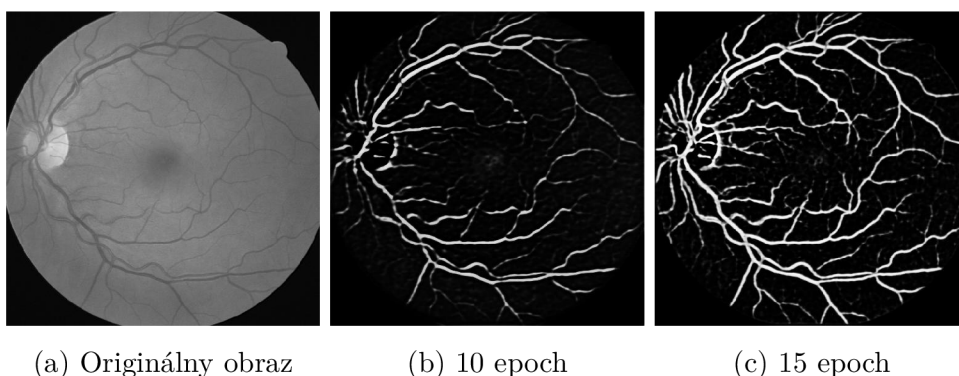
Každý snímok trénovanej sady obsahuje anotáciu segmentovanej časti. Testovacia sada obsahuje 30 snímok bez anotácie. Dátová sada je pred trénovaním augmentovaná tak, že na každý snímok pôvodnej databázy je implementovaná rotácia, priblíženie, horizontálne obrátenie, deformácia šírky a výšky obrazu. Z každého pôvodného snímku vznikne ďalších 30 argumentovaných. Pre výpočet chybovej funkcie je použitá binárna krížová entropia (binary crossentropy), ktorá počíta hodnotu krížovej entropie (cross entropy) pre binárne klasifikačné problémy. Vďaka tejto chybovej funkcii je možné vypočítať presnosť siete metrikou „accuracy“. Váhy sú aktualizované Adam optimalizáciou s rýchlosťou učenia 0.0001. Najlepšie váhy sú po každej epoche ukladané. Po 10 epochách je trénovacia úspešnosť siete 92.28 %

a validačná presnosť 88.9 % Ukážky výsledkov troch predikovaných snímok sú zobrazené na Obr. 4.8 spolu s príslušnými originálnymi snímkami.



Obr. 4.8: Ukážka segmentácie nervovej sústavy v oblasti bruška drozofily

Na tejto architektúre siete boli otestované snímky z databázy DRIVE, ktorej špecifikácia je popísaná v Tab. 5.1. Snímky boli nahrané ako šedo-tónové a tréningová úspešnosť siete bola po 10 epochách 92.63 %, validačná úspešnosť bola 92.24 %. Pridaním 5 epoch sa výsledok tréningovej sady zlepšil na 93.39 % a validačná úspešnosť na 94.08 %. Neurónová sieť síce cievné riečište vysegmentovala, no nežiadúcim faktorom je, že segmentuje aj optický disk a hlbšie štruktúry v sietnici ako je cievnatka, ktorá už nie je súčasťou cievného riečiška, viď Obr. 4.9. Tento problém by



Obr. 4.9: Ukážka segmentácie cievného riečiška, použitá sieť podľa [71]

sa mohol odstrániť predspracovaním dostupných dátových sád o väčšom počte retinálnych snímok, vhodnej augmentácii, ktorá by rozšírila trénovaciu sadu alebo pozmenením architektúry siete, ktorým bude venovaná nasledujúca kapitola práce.

5 NÁVRH A IMPLEMENTÁCIA KONVOLUČNEJ NEURÓNOVEJ SIETE

Pre segmentáciu cievného riečiska bolo nutné najprv pripraviť dáta z dostupných databáz a predspracovať snímky pre tréning siete. Pre účely tréningu boli vytvorené tri typy dátových sád. Prvá dátová sada bola vytvorená zmenšením pôvodných snímkov dostupných databáz. Táto dátová sada tiež slúži na otestovanie zvolenej existujúcej architektúry. Na základe výsledkov je architektúra modifikovaná, s cieľom dosiahnutia lepších výsledkov. Druhá dátová sada bola vytvorená výrezmi pôvodných snímkov dostupných databáz. Táto dátová sada sa delí ešte na tri ďalšie modifikácie, ktoré sú popísané v podkapitole 5.2.2. Tretia dátová sada bola vytvorená pod-obrazmi dostupnej databázy s najvyšším rozlíšením. Vytvorené dátové sady boli testované na modifikovanej verzii architektúry siete. V programovej časti práce je využívaná knižnica Keras, ktorá je súčasťou Tensorflow. Časť experimentov bola počítaná na Národnej Gridovej Infraštruktúre MetaCentrum, konkrétne na clustroch *doom.metacentrum.cz*, ktoré obsahujú 2x grafickú kartu © NVIDIA Tesla K20 5GB. Pre výpočtové preťaženie MetaCentra bola potom využívaná grafická karta © NVIDIA GTX 1060 6GB, ktorá sa nachádza na Ústave biomedicínskej techniky a bioinformatiky. Pre implementáciu bol využívaný programovací jazyk Python, verzia 3.6.5, používaná v integrovanom vývojovom prostredí PyCharm. Pomocné skripty pri rekonštrukcii a vyhodnotení snímkov boli vytvorené v programovacom prostredí MATLAB, verzia R2017a.

5.1 Vytvorenie dátovej sady

Dátová sada bola vytvorená pomocou piatich voľne dostupných databáz, ktoré obsahujú retinálne snímky. Ku každému snímku retiny prislúcha anotovaný obraz cievného riečiska. Snímky v jednotlivých databázach boli vytvorené pomocou odlišných typov fundus kamery s rozličným zorným polom (FOV), čím sa vo výsledku líšia kvalitou a rozmerom. Špecifikácia jednotlivých databáz je uvedená v Tab. 5.1.

Databáza STARE obsahuje ďalších 774 snímkov, ku ktorým neprislúcha žiaden anotovaný obraz. V literatúre sú anotované obrazy označované pod pojmom ground truth. Tieto snímky budú použité na testovanie v rámci tréningu databázy vzniknutej z výrezov pôvodných snímkov všetkých dostupných databáz okrem databázy STARE. Cieľom je zistiť, ako sa bude sieť tréningovaná na obrazoch vzniknutých na inej fundus kamere chovať pri segmentácii snímkov z inej fundus kamery. Snímky tak majú iné vlastnosti ako snímky, na ktorých bola sieť natréningovaná.

Tab. 5.1: Prehľad dostupných databáz retinálnych snímok

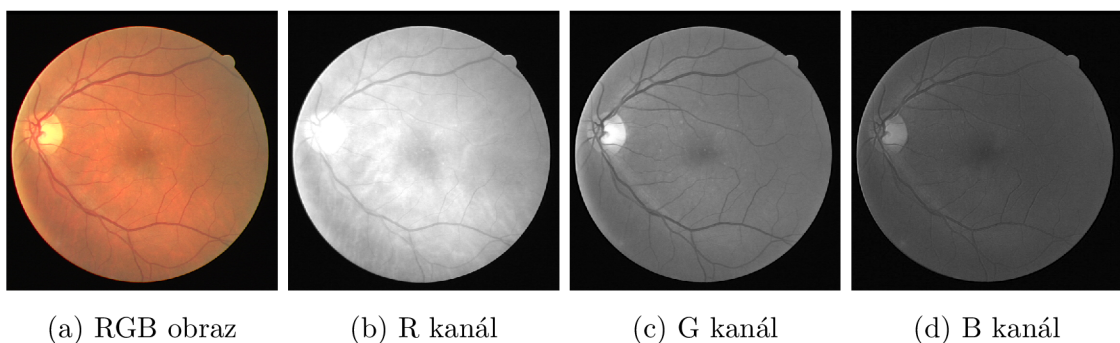
Názov databáze	Počet snímokov	FOV [°]	Veľkosť snímokov	Popis snímok	Typ kamery
DRIVE [57]	40	45°	565 x 584	Trénovanie (20) Testovanie (20)	Canon CR5 non-mydratic 3 CCD
STARE [16]	20	35°	700 x 605	-	TopCon TRV-50
HRF [36]	45	60°	3504 x 2336	Zdravý (15) DR (15) Glaukomatik (15)	Canon CR-1 fundus
CHASE [9]	28	30°	999 x 960	Ľavé oko (14) Pravé oko (14)	Nidek Nm-200-M
ARIA [2]	143	50°	768 x 576	AMD (23) Zdravý (61) DR (59)	Zeus FF450

5.2 Modifikácia dostupných dátových sad

Trénovanie konvolučných sietí prebiehalo na troch modifikáciách z poskytnutých dátových sad:

- trénovanie na celých obrazoch zmenšených na veľkosť 512x512
- trénovanie na obrazoch orezaných na veľkosť 512x512
- trénovanie na pod-obrazoch

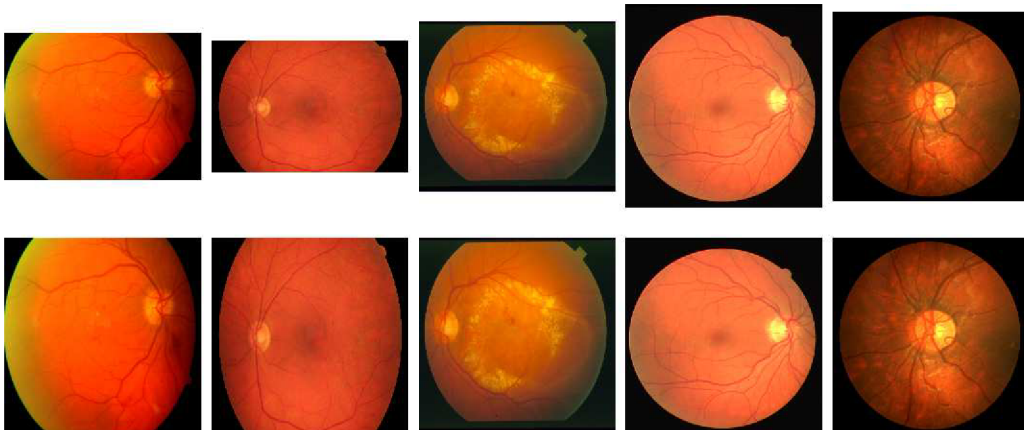
Okrem trénovania konvolučnej neurónovej siete na originálnych RGB snímkoch retiny bolo využité trénovanie aj na jednotlivom zelenom kanáli (G kanál) originálneho snímku. Pri rozdelení RGB originálneho snímku retiny je na Obr. 5.1 vidieť, že najväčší kontrast ciev voči pozadiu obrazu obsahuje zložka G. Zložky R a B neobsahujú významnú informáciu o cievnom riečisku.



Obr. 5.1: Rozdelenie obrazu na RGB zložky

5.2.1 Trénovanie na zmenšených obrazoch

Pri trénovaní na celých obrazoch sú využité dáta získané zo spomínaných dostupných databáz, ktoré sú charakterizované svojou nehomogenitou. Prvá modifikácia dát bola vytvorená úpravou snímok na rovnakú veľkosť. Snímky sú zmenšené na maximálnu veľkosť, pri ktorej siet vracia snímky rovnakej veľkosti. Pri vyššom rozlíšení by boli predikované snímky na konci siete zmenšené. Z tohto hľadiska by ich nebolo možné porovnávať s prislúchajúcim ground truth. Ako je vidieť na Obr. 5.2, zmenšením snímok na veľkosť 512x512 pixelov dôjde k výraznejšej deformácii pôvodných snímok, čím je celá trénovacia sada augmentovaná. Dátová sada je vytvorená z 296 snímok, ktorá je následne rozdelená na trénovaciú, validačnú a testovaciu sadu približne v pomere 2:1:1. Dátová sada určená pre trénovanie siete obsahuje snímky z databáz **DRIVE**, **HRF**, **ARIA** a **CHASE**. Trénovacia sada tak obsahuje 198 snímok a validačná sada 58 snímok. Ku všetkým snímkam prislúchajú snímky ground truth. Testovacia sada obsahuje 20 snímok z databázy STARE. V trénovacej sade sa nenachádza žiaden snímok z databázy STARE, pretože snahou tejto práce je natrénovať zvolenú sieť tak, aby bola vo výsledku adaptabilná na iné dáta, na akých je natrénovaná.

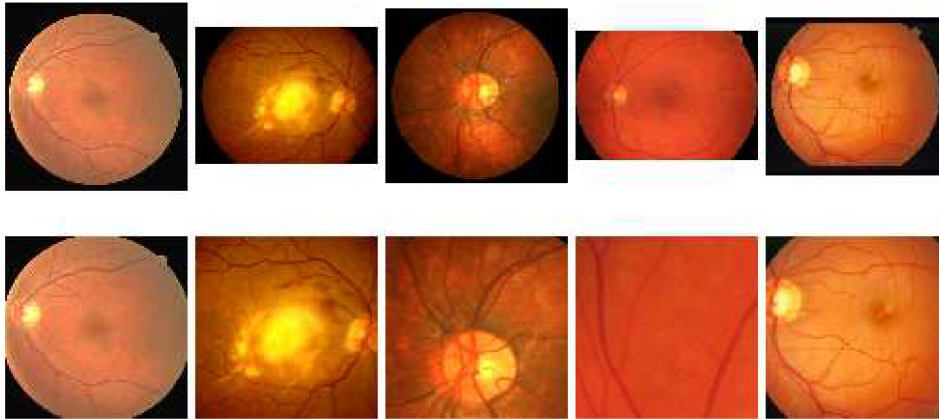


Obr. 5.2: Ukážka dátovej sady vytvorenej zmenšením obrazov na veľkosť 512x512

5.2.2 Trénovanie na výrezoch obrazov

Druhým riešením bolo orezať snímky z jednotlivých databáz na veľkosť 512x512, viď Obr. 5.3. Pri databázach CHASE a HRF, ktoré majú podstatne väčšie rozlíšenie ako ostatné, boli vhodne vybraté časti retinálnej snímky, v ktorej sa nachádza väčší podiel cievneho riečiska. Rozloženie vytvorenej dátovej sady bolo navrhnuté do troch rôznych modifikácií. V tomto prípade nie je trénovacia sada nijako augmentovaná, pretože vytvorením výrezov z pôvodných snímok, ktoré sú

nehomogénne a pôvodne mali rôzne rozlíšenie, sú aktuálne snímky vo vytvorenej databáze rovnakej veľkosti, ale s rôznym priblížením. Týmto spôsobom je dátová sada už augmentovaná metódou priblíženia.



Obr. 5.3: Ukážka dátovej sady vytvorenej orezaním obrazov na veľkosť 512x512

Modifikácia dát 1

Dátová sada obsahuje obrazy orezané na veľkosť 512x512 pixelov všetkých dostupných databáz, tj. **STARE**, **DRIVE**, **ARIA**, **CHASE**, **HRF**. Dátová sada je rozdelená na tréningovú, validačnú a testovaciu sadu v pomere 2:1:1. Tréningová sada obsahuje 455 snímok a validačná sada obsahuje 114 snímok. V tréningovej sade sa nachádza približne rovnomerné rozloženie obrázkov, ktoré obsahujú celé cievne riečisko, detailnejšie priblíženie na oblasť ciev v sietnici a oblasť ciev v okolí optického disku. Tréningová sada obsahuje 114 snímok, ktoré pochádzajú zo spomínaných databáz, no konkrétne snímky neboli použité pre tréningovú a validačnú sadu. Cieľom vytvorenia tejto databázy je zistiť, ako sa sieť vysporiada s tak rôznorodými dátami.

Modifikácia dát 2

Tréningová a validačná sada je rozšírená a obsahuje 1139 snímok. Je tvorená zo snímok databáz **DRIVE**, **ARIA**, **HRF** a **CHASE**. Rozdelenie tréningovej a testovacej sady ostáva v pôvodnom pomere 2:1. Tréningová sada obsahuje 911 snímok a validačná sada obsahuje 228 snímok. Testovanie bude prebiehať len na jednej databázy, ktorá nie je súčasťou tréningovej a validačnej dátovej sady. Testovacia sada je tvorená 794 snímkami databázy **STARE**, z ktorých len k 20 snímkom je dostupný ground truth. Zvyšok snímok bude vyhodnotený subjektívne. Aby bola vo výsledku získaná segmentácia celého snímku, každý obraz je rozdelený na 4 pod-obrazy veľkosti 512x512, ktoré sú po predikcii rekonštruované späť do obrazu pôvodnej veľkosti 605x700 pixelov. Z výsledných predikovaných obrazov sú v prekrývajúcich sa častiach vybrané maximálne hodnoty, ktoré sieť vysegmentovala.

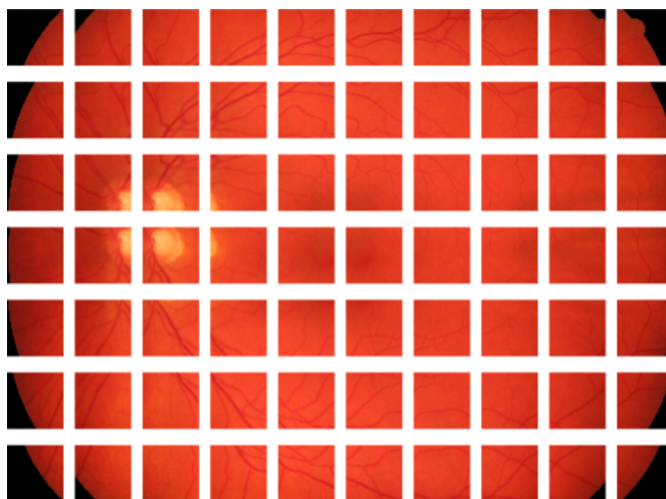
Modifikácia dát 3

Dátová sada obsahuje 1103 snímok. Táto dátová sada bola následne rozdelená v inom pomere. Trénovacia a validačná sada bola tvorená z databáz **DRIVE**, **STARE**, **HRF** a **ARIA**. Trénovacia sada obsahovala 882 snímok a validačná sada 221. Trénovacia sada je tvorená databázou CHASE. Pôvodné snímky boli rozdelené na pod-obrazy, kde z každého pôvodného snímku boli získané opäť 4 snímky. Testovacia sada tak obsahuje 112 snímok, ktoré budú následne zrekonštruované do pôvodnej veľkosti ako v predošlom prípade.

Vytvorením týchto troch modifikácií trénovacích sád je zistiť, ako sa sieť naučí segmentovať snímky s takmer celým cievnym riečiskom a snímky, na ktorých je len časť cievného riečiska s detailnejším záberom. Hlavnou myšlienkou je natrénovať sieť tak, aby dokázala s čo najvyššou úspešnosťou segmentáciu cievného riečiska rôznych typov snímok sietnice, ktoré sa líšia typom fundus kamery, a tým majú odlišné parametre obrazu. V prípade úspešnosti by takto mohol vzniknúť model siete, ktorý bude univerzálny pre segmentáciu cievného riečiska na snímkach akýchkoľvek rozmerov.

5.2.3 Trénovanie na pod-obrazoch

Tretím riešením je vytvorenie pod-obrazov, ktoré je znázornené na Obr. 5.4. Pod-obrazy sú vytvorené na každom jednom snímku databáze HRF, ktorá má z dostupných databáz najväčšie rozlíšenie, a to 2336x3504 pixelov.



Obr. 5.4: Ukážka rozdelenia snímku na pod-obrazy, celkom 10x7 pod-obrazov, veľkosti 512x512 pixelov

Databáza obsahuje 15 snímok zdravých jedincov bez akejkoľvek patológie na sietnici, 15 snímok sietnice jedincov s diabetickou retinopatiou (DR), obsahujúce pato-

logické zmeny, ako sú neovaskulárne siete, hemorágie, jasné lézie, škvrny po ošetrení laserom a iné. Posledná skupina 15 snímok je tvorená jedincami s glaukómom v pokročilom štádiu s príznakmi ohniskovej a difúznej RNFL straty [49]. Pred rozdelením snímok na pod-obrazy sú všetky snímky orezané na veľkosť 2336x3504 pixelov z dôvodu, čo najviac eliminovať čiernu plochu na okrajoch snímok, ktorá by bola pri trénovaní nežiadúca a mohla by znižovať presnosť segmentácie cievného riečiska. Ďalej je prihliadnuté na to, aby sa vytvorené pod-obrazy dostatočne prekrývali a v každom snímku sa nachádzala aspoň jedna cieva, čím je poskytnuté maximálne množstvo informácií pre konvolučnú sieť. Snímky sú rozdelené na pod-obrazy veľkosti 512x512 pixelov, ktoré sa na výšku prekrývajú o 304 pixelov a na šírku 305 pixelov. Prekrývanie pod-obrazov je vypočítané tak, aby obsiahli celý pôvodný obraz. Z jedného pôvodného snímku je získaných 70 nových snímok. Dátová sada pre trénovanie je tvorená 22 pôvodnými snímkami, ktoré obsahujú rovnomerný počet snímok pre zdravé oko, oko glaukomatika a oko s diabetickou retinopatiou. Z týchto snímok vznikne vytvorením pod-obrazov 1540 snímok. Trénovacia sada obsahuje 1232 snímok a validačná sada 308 snímok. Testovacia sada je tvorená zo zvyšných 23 pôvodných snímok, ktorých zastúpenie zdravých a patologických očí je tiež rovnomerné ako pri trénovacej sade. Vytvorením pod-obrazov obsahuje testovacia sada 1610 snímok. Výsledné segmentované pod-obrazy budú opäť rekonštruované do pôvodného rozlíšenia.

Výsledná natrénovaná sieť na pod-obrazoch vzniknutých z pôvodných obrazov s rozlíšením 2336x3506 pixelov bola otestovaná na databázy CHASE s rozlíšením 960x999 pixelov a databázy STARE s rozlíšením 605x700 pixelov.

5.3 Testovanie zvolenej architektúry siete

Zvolená konvolučná neurónová sieť pre segmentáciu je popísaná v kapitole 4.3. Pôvodná architektúra bola otestovaná na prvej modifikácii dát, v ktorej boli retinálne snímky zmenšené na veľkosť 512x512 pixelov. Na základe výsledkov boli na tejto databáze ďalej prevádzané experimenty so zmenou optimalizačného algoritmu, počtom pridaných vrstiev a pod. Modifikácia siete, ktorá dosahovala najlepšie výsledky bola následne použitá na ďalšie modifikácie dátových sad, konkrétne dátová sada vzniknutá výrezom obrazu zo všetkých pôvodných dostupných databáz a dátová sada vzniknutá vytvorením pod-obrazov z databázy HRF.

Trénovanie prebiehalo na RGB snímkoch a samostatne na G kanáli, v ktorom sa nachádza významná informácia o cievnom riečisku. V pôvodnej architektúre bol zmenený len počet epoch z 10 na 20. Taktiež bola pridaná funkcia `model.evaluate`, ktorá počíta testovaciu úspešnosť segmentácie ciev. Výsledky trénovacej, validačnej a testovacej úspešnosti sú znázornené v Tab. 5.2. Po dvadsiatich epochách siet

Tab. 5.2: Výsledky pôvodnej architektúry trénovanej na dátovej sade zmenšených obrazov na veľkosť 512x512 pixelov

RGB obrázky				
Počet epoch	Úspešnosť			Čas tréovania [hh:mm:ss]
	Trénovacia	Validačná	Testovacia	
20	0.9469	0.9295	0.9364	00:22:37
40	0.9479	0.9345	0.9441	00:42:29
60	0.9558	0.9318	0.9423	01:05:13
80	0.9637	0.9306	0.9402	01:26:41
Zelený (G) kanál				
Počet epoch	Úspešnosť			Čas tréovania [hh:mm:ss]
	Trénovacia	Validačná	Testovacia	
20	0.9434	0.9325	0.9430	00:22:37
40	0.9471	0.9340	0.9412	00:42:29
60	0.9549	0.9335	0.9403	01:05:13
80	0.9698	0.9322	0.9380	01:26:40

dosahovala už 94,69% trénovaciu úspešnosť v prípade RGB obrazov a 94,34 % trénovaciu úspešnosť v prípade G kanálu. Počet epoch bol ďalej pridávaný na tridsaťpäť a štyridsať. Sieť po tridsiatich piatich epochách dosahovala trénovaciu úspešnosť až 94,79 % pre RGB obrázky a 94,71 % pre G kanál. Po štyridsiatich epochách sieť dosahovala lepšie výsledky u trénovacej úspešnosti, no validačná úspešnosť oproti 35 epochám klesala, čím došlo k preučeniu siete, nazývanému ako overfitting. Architektúra siete dosahovala prekvapivo nepatrne lepšie výsledky pri RGB obrazoch oproti samotnému G kanálu. Na základe týchto výsledkov prebiehalo tréovanie siete, pre úsporu času, len na RGB obrazoch.

5.3.1 Modifikácie zvolenej architektúry

V tejto časti práce došlo na základe výsledkov k úpravám pôvodnej architektúry siete. V úpravách pôvodnej architektúry siete bolo zahrnuté rôzne pridávanie, resp. odoberanie ďalších vrstiev siete s cieľom zlepšiť jej výsledky. Výsledky experimentov sú uvedené v Tab. 5.3, ktorých zmeny sú uvedené pod každou konkrétnou modifikáciou. Pre prehľadnosť sú najlepšie výsledky tučne zvýraznené. Chovanie siete bolo testované tiež pri zmene optimalizácie chybovej funkcie z Adam [35] na SDG [39] a následne na RMSProp [62]. Výsledky v oboch prípadoch neboli úspešné, pretože sieť dosahovala len približne 89% trénovaciu úspešnosť. Z tohto dôvodu bola ponechaná optimalizácia Adam s rýchlosťou učenia 0.0001, ktorá je podľa dostup-

nej literatúry momentálne najpoužívanejší optimalizačný algoritmus minimalizácie chybovej funkcie.

Modifikácia 1

Keďže úspešnosť siete stagnuje v priebehu prvých epoch tréovania, bola z pôvodnej architektúry odstránená vrstva Dropout, ktorá sa nachádzala v štvrtom konvolučnom bloku. Taktiež bola zmenená vrstva Upsampling, viď pôvodná architektúra na Obr. 4.7, za vrstvu Conv2DTranspose. Dôvodom je, že dekonvolučná vrstva Upsampling zavádza do obrazov artefakty. Pre eliminovanie týchto artefaktov bola táto vrstva zmenená za regulárnu alternatívu nazývanú ako transponovaná konvolučná vrstva (Transposed convolution layer). Vrstva si zachováva spojenie, ktoré je kompatibilné s konvolučnou vrstvou v kontraktilnej časti architektúry. Na rozdiel od dekonvolučnej vrstvy Upsampling by tento prístup k prevzorkovaniu nemal vytvárať artefakty ako štandardné správanie vrstvy Upsampling. Počet epoch bol znížený na 35, pretože pri ďalších epochách začala validačná úspešnosť klesať, čo je nežiadúci faktor. Výsledná sieť dosiahla lepšie výsledky už po 35 epochách v porovnaní s pôvodnou architektúrou. Aby nedochádzalo k preučeniu siete, tak bola do tréovania vložená funkcia, ktorá pri zistení poklesu validačnej úspešnosti včas zastaví tréovanie siete.

Modifikácia 2

V čase kedy bola vyvinutá architektúra U-net neexistovala vrstva Batch Normalization, ktorá urýchľuje konvergenciu učenia a tým by mala sieť dosahovať lepšie výsledky. Táto vrstva bola pridaná po každej konvolučnej vrstve. Výsledky po 35 epochách sú znázornené v Tab. 5.3. Čas tréovania siete je po 35 epochách približne rovnaký ako pri 40 epochách pôvodnej architektúry. Sieť dosahovala o trochu lepšie výsledky v rámci tréovacej úspešnosti, no výsledná validačná a testovacia úspešnosť klesla skoro o 2 % oproti prvej modifikácii.

Modifikácia 3

Pridanie tretej konvolučnej vrstvy do štvrtého konvolučného bloku v kontraktilnej dráhe a v šiestom bloku expanzívnej dráhy siete sa neosvedčilo. Testovacia úspešnosť po 35 epochách učenia dosahovala 94.29 %, čím klesla oproti prvej modifikácii.

Modifikácia 4

Ďalej bolo experimentované s ďalším pridávaním konvolučnej vrstvy do tretieho konvolučného bloku v kontraktilnej dráhe a ôsmeho konvolučného bloku v expanzívnej dráhe. Tretia konvolučná vrstva v štvrtom a šiestom konvolučnom bloku bola ponechaná. Testovacia úspešnosť siete sa znížila na 93.99 %.

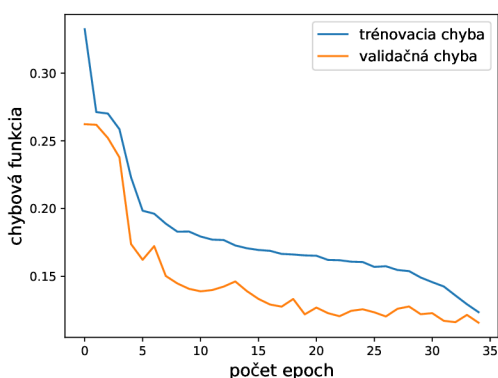
Modifikácia 5

Posledným pokusom bolo pridanie ďalšej konvolučnej vrstvy do druhého a siedmeho konvolučného bloku, pričom konvolučné vrstvy z predošlých pokusov boli tiež ponechané. Sieť dosiahla 94.24 % testovacej úspešnosti a tým nepresiahla úspešnosť siete pri pôvodnom počte dvoch konvolučných vrstiev v každom konvolučnom bloku ako je to v prvej modifikácii.

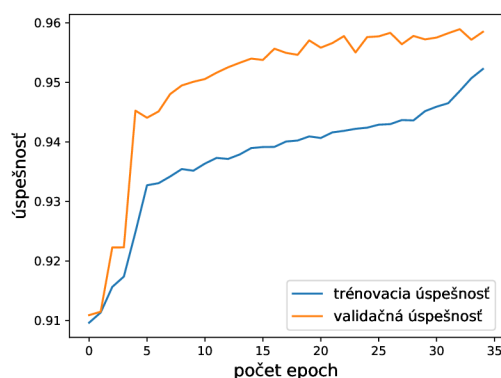
Druhá modifikácia, v ktorej boli pridané vrstvy Batch Normalization dosahuje lepšiu tréningovú úspešnosť ako prvá modifikácia. Na rozdiel od toho prvá modifikácia preyšuje vo validačnej a testovacej úspešnosti, čo je vo výsledku dôležitejší faktor pri vyhodnocovaní segmentácie obrazov. Na základe týchto dosiahnutých výsledkov bude ďalej používaná architektúra siete prvej modifikácie, ktorá je znázornená na Obr. 5.6.

Tab. 5.3: Výsledky modifikovaných modelov siete

Modifikácia	Počet epoch	Úspešnosť			Čas tréningovania [hh:mm:ss]
		Tréningová	Validačná	Testovacia	
1	35	0.9523	0.9585	0.9603	00:58:44
2	35	0.9533	0.9338	0.9425	00:43:12
3	35	0.9492	0.9358	0.9429	00:40:16
4	35	0.9497	0.9347	0.9399	00:42:37
5	35	0.9507	0.9318	0.9424	00:45:31



(a) Minimalizácia chybovej funkcie

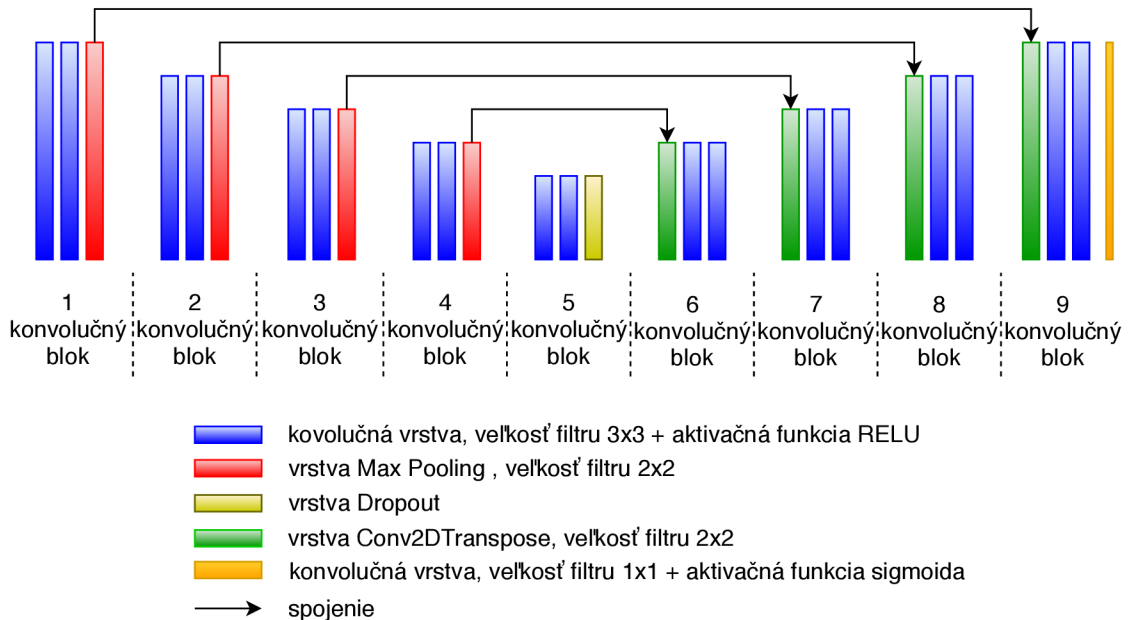


(b) Vývoj tréningovania architektúry siete

Obr. 5.5: Priebeh chybovej funkcie a úspešnosti tréningovania architektúry siete

Priebeh minimalizácie chybovej funkcie v rámci testovacej a validačnej sady je znázornený na Obr. 5.5a. Na základe chybovej funkcie je vypočítaná tréningová a validačná úspešnosť, ktorých priebeh počas 35 epoch je znázornený na Obr. 5.5b.

Výsledná modifikovaná architektúra siete je znázornená na Obr. 5.6. Modifikovaná verzia siete obsahuje vrstvu Dropout len v piatom konvolyčnom bloku, oproti pôvodnej architektúre. Rovnako je vrstva Upsampling nahradená vrstvou Conv2DTranspose. Počet filtrov a veľkosť filtrov v jednotlivých vrstvách ostal nezmenený.



Obr. 5.6: Modifikácia zvolenej architektúry

5.3.2 Prahovanie obrazu

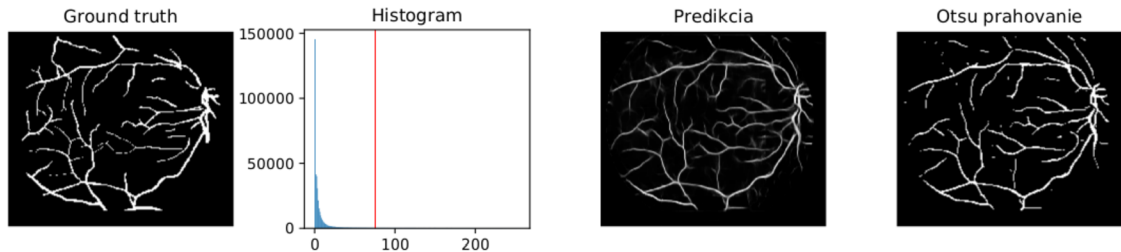
Výsledné snímky zo siete nie sú binárne, ale šedotónové. Práve preto je nutné určiť správny prah, ktorý bude výsledné snímky binarizovať. Pre prahovanie výsledných snímok bola vybraná metóda automatickej voľby prahu, konkrétne metóda Otsu. Metóda pracuje so štatistickými údajmi histogramu, v ktorom skúma intenzitu obrazu. Dochádza pri tom k rozdeľovaniu pixelov histogramu do dvoch tried, ktoré sa rozlišujú na popredie a pozadie obrazu. Optimálny prah je vypočítaný na základe minimalizácie vnútorného rozptylu obidvoch tried:

$$\sigma_2^w = \omega_0(T)\sigma_2^0 + \omega_1(T)\sigma_2^1, \quad (5.1)$$

kde váhy ω_0 a ω_1 sú pravdepodobnosti dvoch tried oddelených zvažovanou prahovou hodnotou T a σ_2^0 , σ_2^1 sú smerodatné odchýlky týchto dvoch tried [50].

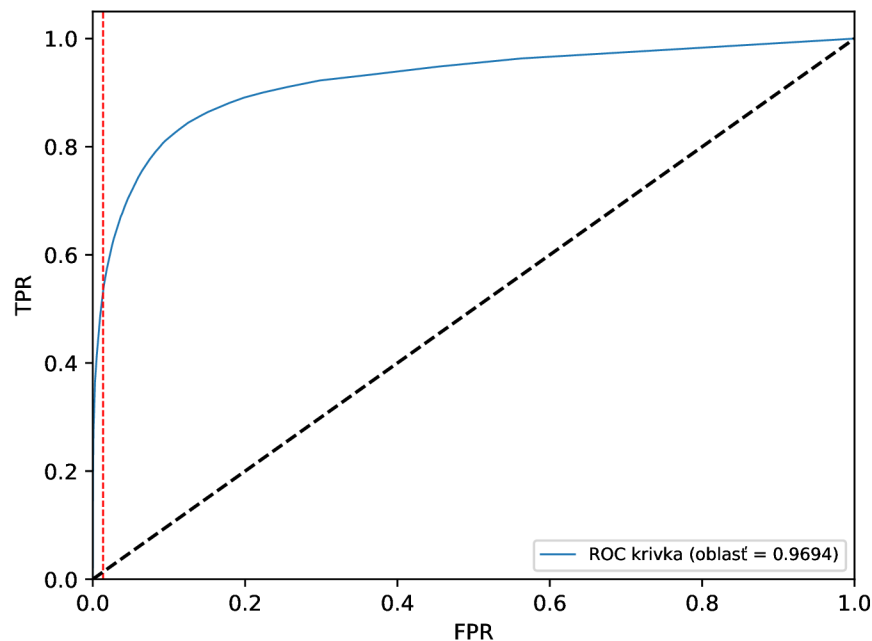
Na Obr. 5.7 je zobrazený výsledok predikovaného obrazu prvej modifikácie architektúry po 35 epochách a histogram tohoto obrazu, kde hodnota 0 predstavuje čiernu farbu a hodnota 1 bielu farbu. Červená čiara v histograme značí prah vypočítaný

pomocou metódy Otsu, ktorá histogram rozdelila na popredie a pozadie. V tomto prípade určila prah na hodnotu 85. Výsledok prahovania je zobrazený za histogramom. Pri tomto prahu je senzitivita výslednej segmentácie len 64.81 % a špecificita 98.29 %. Výsledná presnosť (ACC) je 95.81 %.



Obr. 5.7: Prahovanie Otsu metódou

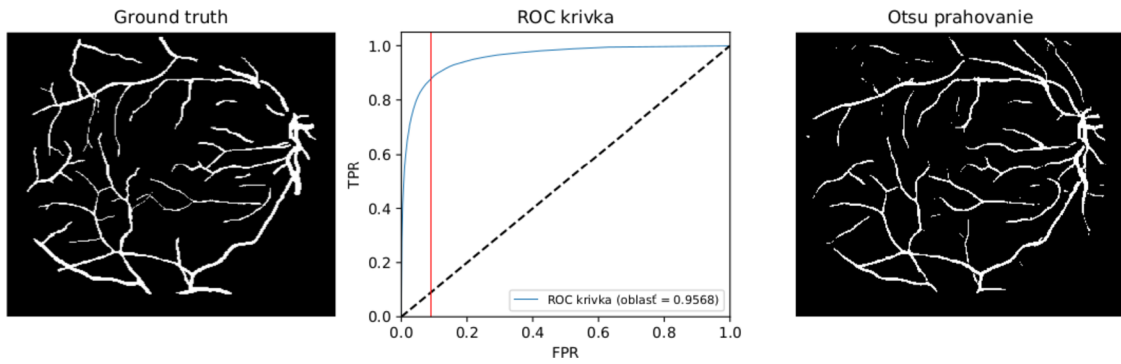
Na Obr. 5.8 je vykreslená priemerná ROC krivka všetkých 20 testovaných snímkov siete trénovanej na zmenšených obrazoch. Červenou prerušovanou čiarou je vyznačená priemerná hodnota Otsu prahovania. Prahovanie sa nachádza v oblasti striktného prahovania, ktoré je definované na Obr. 4.1.



Obr. 5.8: ROC krivka 20 testovaných snímkov databázy STARE na modifikovanej architektúre siete

Pri zvýšení prahovej hodnoty do oblasti optimálneho prahu, ktorý je znázornený na Obr. 5.9 červenou čiarou v ROC krivke, sa síce senzitivita výslednej segmentácie

zvýšila na 86.78 %, no počet nesprávne segmentovaných pixelov sa rapídne zvýšil a špecificita klesla na 91.85 %. Výsledná presnosť klesla na 91.48 %. Taktiež sa hlavné cievne riečisko zvýraznilo hrubšími líniami oproti ground truth. Z dôvodu čo najviac eliminovať počet falošne pozitívnych pixelov vo výslednom obraze bolo ponechané prahovanie Otsu metódou.



Obr. 5.9: Zvýšenie prahovej hodnoty na ROC krivke do oblasti optimálneho prahu

6 VÝSLEDKY A DISKUSIA

Vytvorené dátové sady sú trénované na modifikovanej verzii zvolenej architektúry siete, ktorá je znázornená na Obr. 5.6. Trénovanie na vytvorenej dátovej sade zmenšením pôvodných obrazov je otestované na databázy STARE. Trénovanie na vytvorenej dátovej sade orezaním pôvodných snímok je otestované na databázach STARE a CHASE. Posledné trénovanie siete na vytvorenej dátovej sade pod-obrazov pôvodných snímok databázy HRF je testované na 23 snímkoch databázy HRF a na celej databázy CHASE a STARE.

6.1 Trénovanie na zmenšených obrazoch

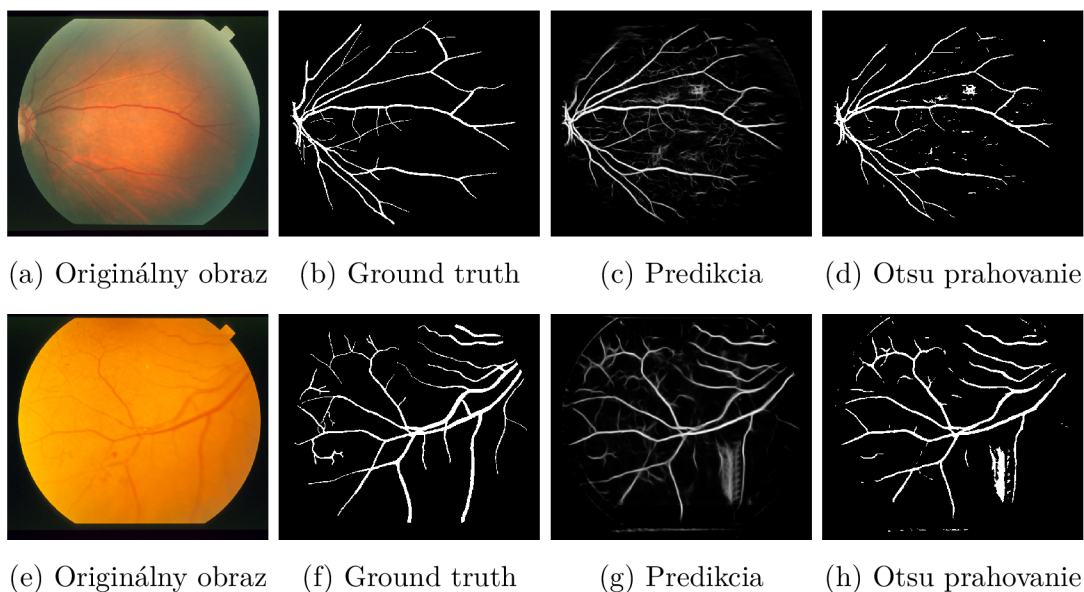
Pri trénovaní na zmenšených obrazoch veľkosti 512x512 pixelov už bola použitá upravená verzia pôvodnej architektúry siete, z ktorej bola odstránená vrstva Dropout nachádzajúca sa vo 4 konvolučnom bloku. Výsledky sú znázornené v Tab. 5.3 pod Modifikáciou 1.

Pre každý prahovaný obraz je následne vypočítaná senzitivita, špecificita a úspešnosť. Oblasť pod ROC krivkou je počítaná na základe predikovaného obrazu a ground truth. Výsledky sú zobrazené v Tab. 6.1, najlepšie a najhoršie výsledky sú zvýraznené tučným fontom. Tieto metriky sú použité pre overenie a porovnanie výsledkov testovacej úspešnosti, ktorá je počítaná v prostredí Keras. Na základe týchto výsledkov je možné výslednú sieť porovnať s inými metódami segmentácie cievneho riečiska pomocou konvolučných neurónových sietí a ďalších dostupných metód.

Najlepšie a najhoršie výsledky segmentácie databázy STARE podľa vypočítanej presnosti ACC sú zobrazené na Obr. 6.1. Najlepší segmentovaný snímok dosahuje úspešnosť 97.56 % a najhoršie segmentovaný snímok dosahuje 95.20 % úspešnosť. Na Obr. 6.1g je vidieť, že sieť vysegmentovala väčšiu oblasť pri cieve, ktorá je po prahovaní na Obr. 6.1h zobrazená už ako biely flak. Tým sa počet falošne pozitívnych pixelov zvýšil. Výhodou je, že tento flak nijako nezasahuje do cievneho riečiska, čím konečný výsledok segmentácie ciev nie je narušený.

Tab. 6.1: Výsledky modifikovanej siete trénovej na zmenšených obrazoch - testovaná databáza STARE po prahovaní

Špecifikácia obrazu	SE	SP	ACC	AUC
im0001	0.6898	0.9824	0.9591	0.9525
im0002	0.7507	0.9834	0.9679	0.9680
im0003	0.8073	0.9851	0.9745	0.9816
im0004	0.6481	0.9829	0.9581	0.9568
im0005	0.8414	0.9712	0.9597	0.9800
im0044	0.7385	0.9786	0.9618	0.9684
im0077	0.9339	0.9712	0.9683	0.9914
im0081	0.9202	0.9740	0.9700	0.9908
im0082	0.8424	0.9787	0.9679	0.9797
im0139	0.8503	0.9667	0.9573	0.9764
im0162	0.8628	0.9745	0.9665	0.9822
im0163	0.8890	0.9753	0.9686	0.9876
im0235	0.8324	0.9828	0.9694	0.9840
im0236	0.8190	0.9843	0.9693	0.9836
im0239	0.7111	0.9865	0.9628	0.9473
im0240	0.7173	0.9816	0.9546	0.9485
im0255	0.8892	0.9786	0.9706	0.9892
im0291	0.7327	0.9885	0.9756	0.9558
im0319	0.6754	0.9772	0.9642	0.9421
im0324	0.6062	0.9768	0.9520	0.9218
Priemer	0.7879	0.9790	0.9649	0.9694
Std	±0.0947	±0.0057	±0.0065	±0.0190



Obr. 6.1: Výsledky najlepšej a najhoršej segmentácie snímku z databázy STARE, sieť trévaná na zmenšených obrazoch

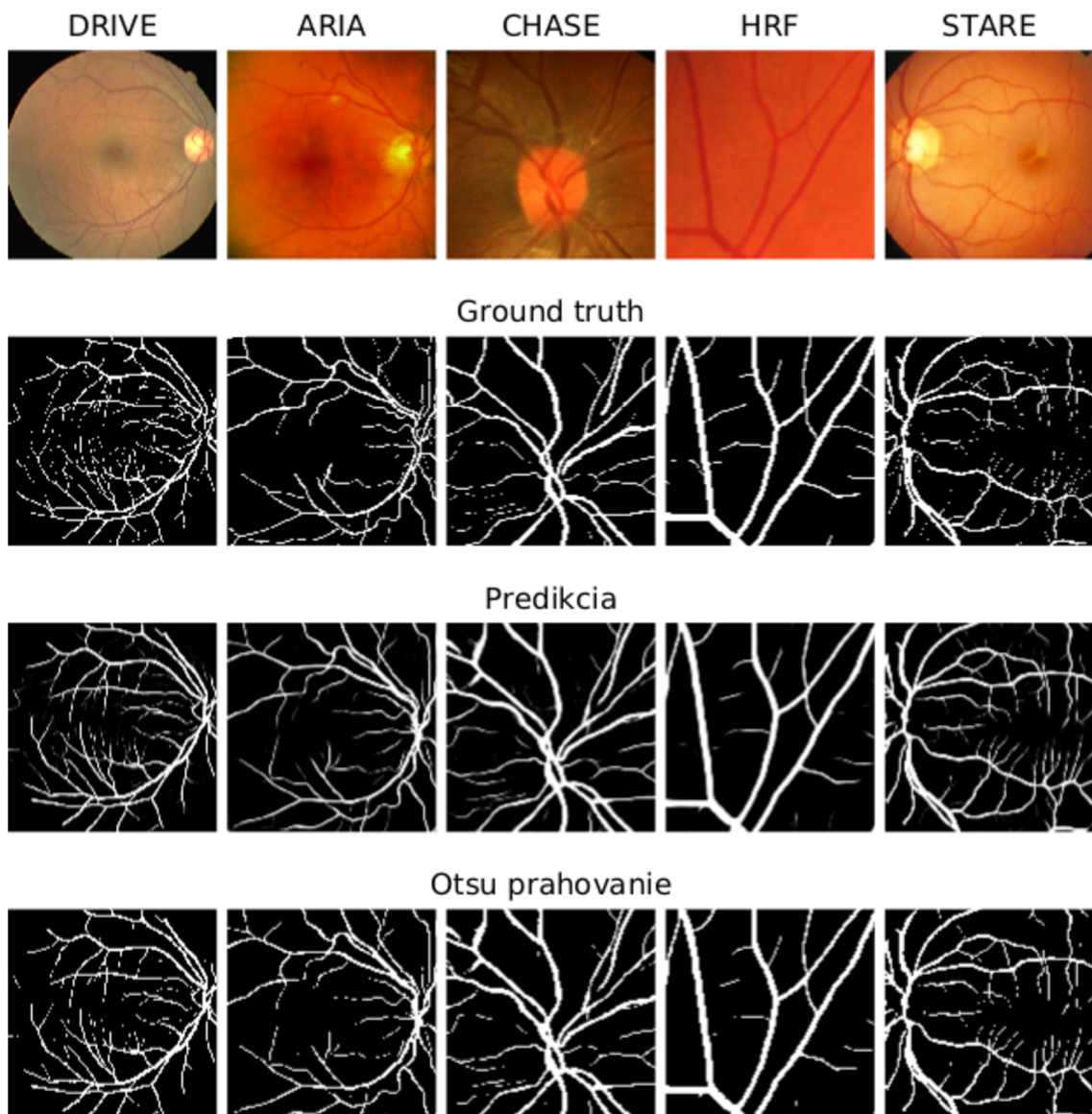
6.2 Tréovanie na výrezoch obrazov

Výsledky úspešnosti siete prvej modifikácie dát sú znázornené v Tab. 6.2. Po 35 epochách sieť dosahovala 98.06 % tréovaciú úspešnosť, 95.80 % validačnú úspešnosť a 95.88 % testovaciu úspešnosť.

Tab. 6.2: Výsledky modifikovanej architektúry siete tréovanej na dátovej sade orezaných obrazov na veľkosť 512x512 pixelov

Počet epoch	Úspešnosť			Čas [hh:mm:ss]
	Tréovacia	Validačná	Testovacia	
20	0.9623	0.9554	0.9575	01:26:17
35	0.9741	0.9578	0.9589	02:52:58

Na Obr. 6.2 sú znázornené jednotlivé ukážky segmentácie snímok z piatich databáz. Pre každý snímok na Obr. 6.2, je vypočítaná senzitivita, špecificita, presnosť a plocha pod ROC krivkou. Najnižšiu senzitivitu dosahuje snímok z databázy ARIA, najvyššiu senzitivitu dosahuje snímok z databázy HRF. Na čím detailnejších snímkach je sieť tréovaná, tým lepšie výsledky dosahuje pre daný testovaný prípad. Taktiež vo všetkých výsledných predikovaných obrazoch sú cievy v oblasti optického disku segmentované bez prítomnosti optického disku vo výslednom segmente. Najlepšie a najhoršie výsledky sú v Tab. 6.3 vyznačené tučným fontom.



Obr. 6.2: Výsledky pre dátovú sadu obrazov orezaných na veľkosť 512x512 pixelov po 35 epochách

Tab. 6.3: Výsledky testovania vybraných snímok siete trénovanej na zmenšených obrazoch

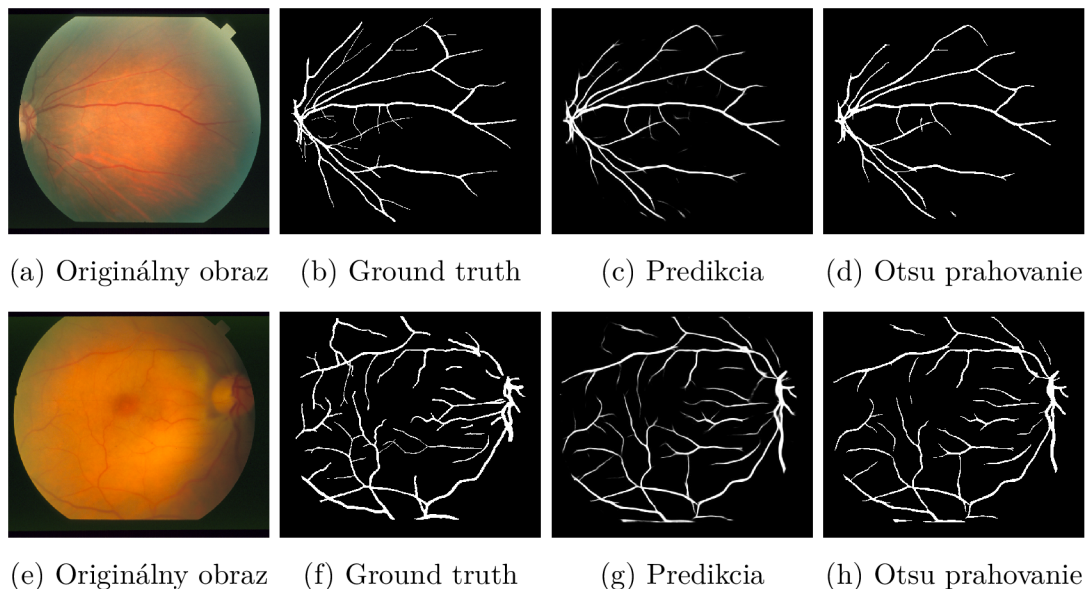
Špecifikácia obrazu obrazu	Senzitivita	Špecificita	ACC	AUC
DRIVE	0.7850	0.9645	0.9453	0.9574
ARIA	0.7197	0.9564	0.9322	0.9423
CHASE	0.8868	0.9535	0.9447	0.9702
HRF	0.8906	0.9838	0.9711	0.9821
STARE	0.8830	0.9606	0.9512	0.9707

Na základe úspešných výsledkov bola architektúra siete ďalej trébovaná na Modifikácii trébovacích dát 2, ktorá je následné testovaná na databázy STARE a Modifikácii trébovacích dát 3, ktorá je testovaná na databázy CHASE. Výsledky dátových sad sú znázornené v Tab. 6.4.

Tab. 6.4: Výsledky modifikovanej architektúry siete trébovanej na dátovej sade orezaných obrazov na veľkosť 512x512 pixelov

Modifikácia	Počet epoch	Úspešnosť			Čas trébovania [hh:mm:ss]
		Trébovacía	Validačná	Testovacia	
1	35	0.9834	0.9601	0.9565	04:17:12
2	35	0.9847	0.9571	0.9535	04:09:18

V obidvoch prípadoch, tj. pre databázu STARE a CHASE, bola vypočítaná senzitivita, špecificita, presnosť a plocha pod ROC krivkou pre každý jeden snímok, ako v predchádzajúcich prípadoch. Výsledky pre databázu STARE sú uvedené v Tab. 6.5. Aj napriek tomu, že je sieť natrébovaná na obecnjšiu dátovú sadu, dosahuje databáza STARE po prahovaní úspešnosť až 95.95 %. Pri trébovaní siete na zmenšených dátach, ktoré sú deformované, ale obsahujú kompletnú informáciu o cievnom riečisku dosahovala databáza STARE po prahovaní úspešnosť 96,49 %, čo je len o 0.54 % viac ako v pri tejto obecnjšej sieti natrébovanej na rôznych priblíženiach cievného riečiska. Najlepší a najhorší výsledok je uvedený na Obr. 6.3, kde výsledok



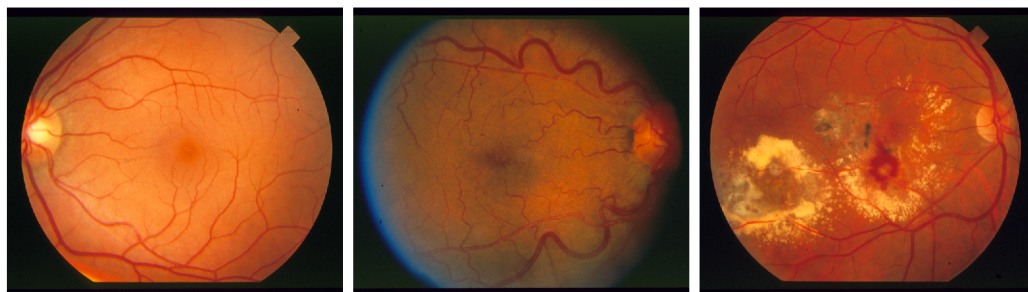
Obr. 6.3: Výsledky najlepšej (hore) a najhoršej (dole) segmentácie snímku z databázy STARE, sieť trébovaná na výrezoch obrazov - Modifikácia dátovej sady 2

snímku s najvyššou úspešnosťou segmentácie cievneho riečiska je rovnaký ako v prípade siete trénovanej na zmenšených obrazoch. Najhoršie segmentovaný snímok dosiahol 94.70 % úspešnosť.

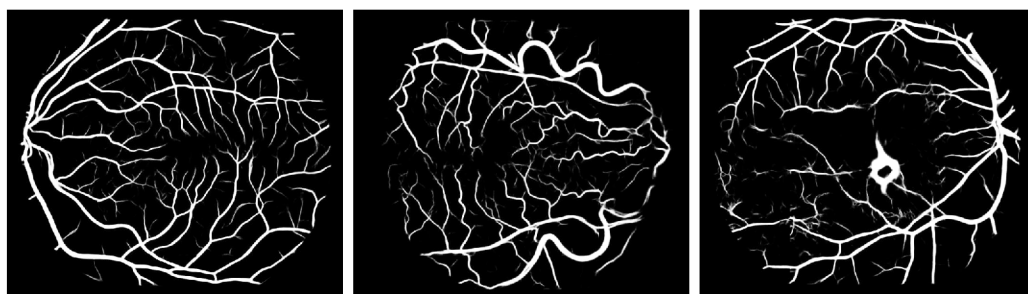
Tab. 6.5: Výsledky modifikovanej siete trénovanej na orezaných obrazoch - Modifikácia dátovej sady 2, testovaná databáza STARE po prahovaní

Špecifikácia obrazu	SE	SP	ACC	AUC
im0001	0.6835	0.9814	0.9576	0.9510
im0002	0.7083	0.9783	0.9603	0.9426
im0003	0.7362	0.9843	0.9694	0.9492
im0004	0.6016	0.9811	0.9529	0.9360
im0005	0.7769	0.9639	0.9470	0.9597
im0044	0.7482	0.9733	0.9576	0.9542
im0077	0.9161	0.9614	0.9578	0.9816
im0081	0.8250	0.9636	0.9593	0.9822
im0082	0.8141	0.9708	0.9594	0.9674
im0139	0.8328	0.9650	0.9529	0.9570
im0162	0.8647	0.9636	0.9543	0.9603
im0163	0.8541	0.9649	0.9572	0.9769
im0235	0.8413	0.9739	0.9632	0.9733
im0236	0.7615	0.9747	0.9627	0.9756
im0239	0.7237	0.9816	0.9626	0.9559
im0240	0.8678	0.9817	0.9554	0.9426
im0255	0.7077	0.9648	0.9561	0.9711
im0291	0.5596	0.9900	0.9757	0.9400
im0319	0.5739	0.9883	0.9698	0.9446
im0324	0.5739	0.9855	0.9580	0.9377
Priemer	0.7651	0.9746	0.9595	0.9580
Std	± 0.1048	± 0.0094	± 0.0066	± 0.0147

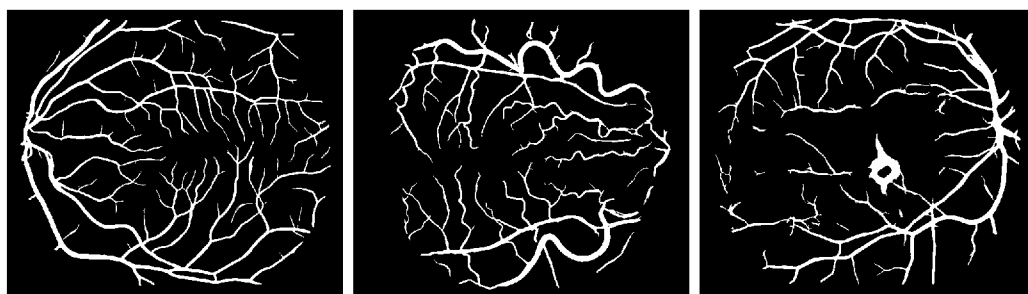
Na Obr. 6.4 sú zobrazené tri ukážky snímok databázy STARE. Výsledky je možné zhodnotiť len subjektívne, pretože neobsahujú ground truth. V prvých dvoch prípadoch sieť zo subjektívneho pohľadu vysegmentovala celý cievny strom správne. V treťom prípade sieť segmentovala okrem cievneho stromu aj červenú škvrnu, ktorá sa nachádza v oblasti makuly.



(a) Originálne snímky



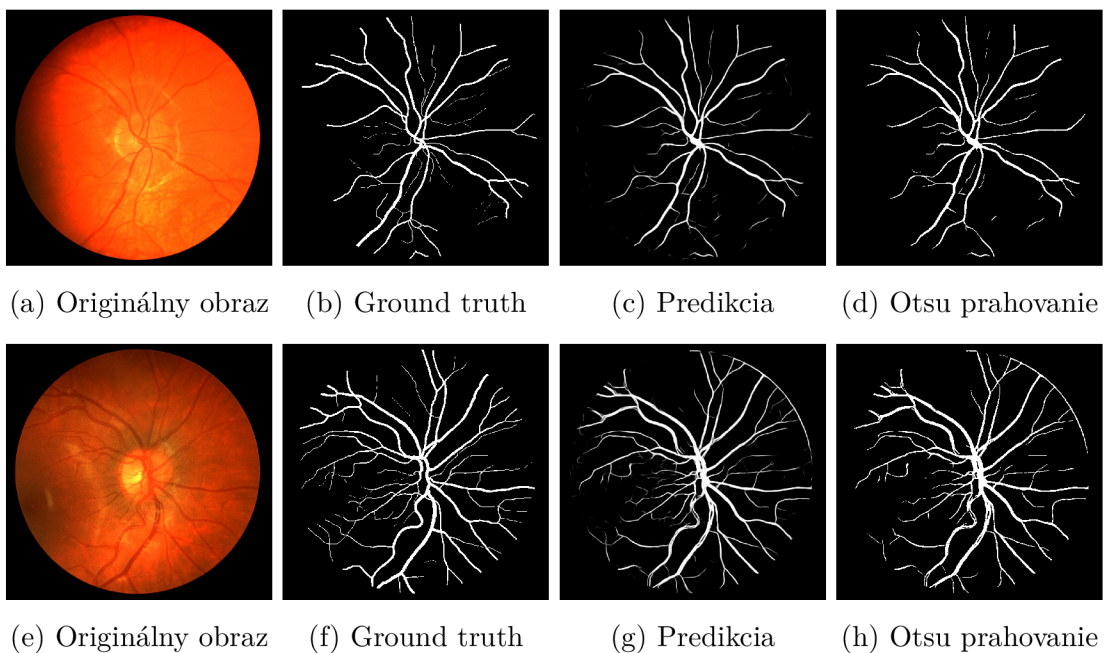
(b) Predikované snímky



(c) Otsu prahovanie

Obr. 6.4: Ukážka segmentácie snímok z databázy STARE, ktoré nemajú anotovaný obraz segmentácie cievneho riečiska, sieť trébovaná na výrezoch obrazov - Modifikácia dátovej sady 2

Na tretej modifikácii dátovej sady orezaných snímok bola otestovaná databáza CHASE, ktorá sa v trénovacej dátovej sade nenachádza. Výsledky sú zobrazené v Tab. 6.6. Sieť dosahuje po prahovaní snímok priemernú úspešnosť 96.56 % so strednou odchýlkou len ± 0.0043 . Tieto výsledky preukazujú, že sa podarilo vytvoriť sieť, ktorá je natrénovaná na rôzne modalitty cievného riečiska. Sieť dosahuje veľmi dobré výsledky na snímkach obsahujúcich celé cieвне riečisko a na snímkach, v ktorých sa nachádza len časť ciev, ktoré sú následne rekonštruované do pôvodného obrazu. Najlepší a najhorší výsledok segmentácie databázy CHASE je znázornený na Obr. 6.5, kde najlepšia úspešnosť segmentácie bola 97.61 % a najhoršia bola 96.00 %. V prípade najhoršie segmentovaného snímku je patrné, viď Obr.6.5g, že bol segmentovaný taktiež okraj zorného poľa, čo znižuje presnosť výslednej segmentácie.



Obr. 6.5: Výsledky najlepšej (hore) a najhoršej (dole) segmentácie snímku z databázy CHASE, sieť trénovaná na výrezoch - Modifikácia dátovej sady 3

Tab. 6.6: Výsledky modifikovanej siete trénuvanej na orezaných obrazoch - Modifikácia dátovej sady 3, testovaná databáza CHASE po prahovaní

Špecifikácia obrazu	SE	SP	ACC	AUC
Image_01L	0.7675	0.9862	0.9710	0.9705
Image_01R	0.6823	0.9879	0.9647	0.9519
Image_02L	0.7338	0.9843	0.9640	0.9580
Image_02R	0.7060	0.9850	0.9632	0.9502
Image_03L	0.6921	0.9856	0.9626	0.9558
Image_03R	0.7220	0.9865	0.9665	0.9622
Image_04L	0.6230	0.9897	0.9610	0.9459
Image_04R	0.6362	0.9878	0.9607	0.9391
Image_05L	0.7022	0.9868	0.9643	0.9541
Image_05R	0.7089	0.9882	0.9644	0.9582
Image_06L	0.6872	0.9879	0.9651	0.9365
Image_06R	0.6743	0.9845	0.9613	0.9497
Image_07L	0.7145	0.9837	0.9627	0.9538
Image_07R	0.6923	0.9846	0.9624	0.9495
Image_08L	0.7379	0.9820	0.9662	0.9549
Image_08R	0.7535	0.9831	0.9686	0.9627
Image_09L	0.7294	0.9874	0.9744	0.9518
Image_09R	0.7222	0.9897	0.9761	0.9572
Image_10L	0.7539	0.9775	0.9635	0.9592
Image_10R	0.7366	0.9762	0.9614	0.9454
Image_11L	0.8283	0.9800	0.9719	0.9502
Image_11R	0.8085	0.9817	0.9724	0.9502
Image_12L	0.6797	0.9824	0.9606	0.9368
Image_12R	0.7273	0.9784	0.9600	0.9573
Image_13L	0.7472	0.9811	0.9667	0.9621
Image_13R	0.6752	0.9818	0.9624	0.9437
Image_14L	0.7123	0.9864	0.9674	0.9576
Image_14R	0.6320	0.9877	0.9668	0.9292
Priemer	0.7138	0.9844	0.9654	0.9528
Priemer	±0.0469	±0.0036	±0.0043	±0.0088

6.3 Trénovanie na pod-obrazoch

Pre trénovanie na pod-obrazoch bola použitá dátová sada databázy HRF. Výsledky po dvadsiatich a tridsiatich piatich epochách sú uvedené v Tab. 6.7. Architektúra siete dosahovala po 35 epochách doposiaľ najlepšie výsledky, 99.40 % pre trénovaciu sadu, 96.85 % úspešnosť pre validačnú sadu a 96.48 % pre testovaciu sadu.

Tab. 6.7: Výsledky modifikovanej architektúry siete trénovanej na pod-obrazoch databázy HRF

Počet epoch	Úspešnosť [%]			Čas [hh:mm:ss]
	Trénovacia	Validačná	Testovacia	
20	0.9907	0.9683	0.9625	03:52:13
35	0.9940	0.9685	0.9648	06:45:27

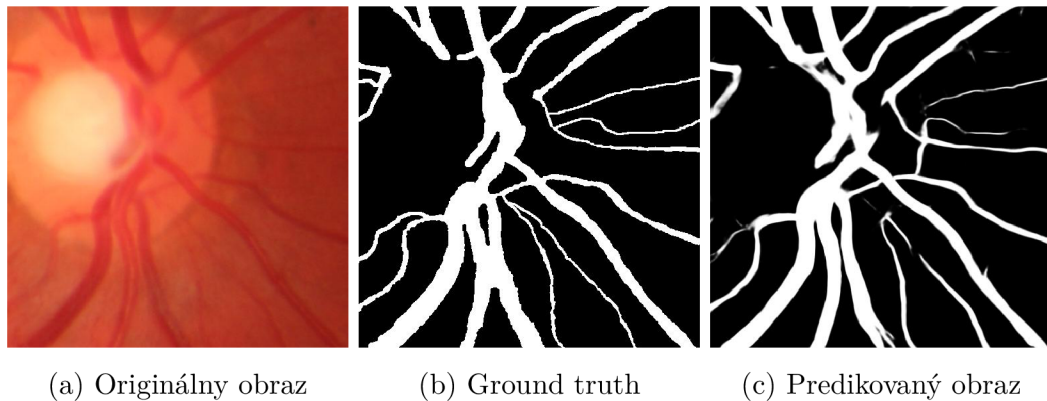
Natrénovaná sieť bola okrem 23 snímok databázy HRF otestovaná aj na databázy CHASE a databázy STARE. Výsledky trénovacej úspešnosti sú uvedené v Tab. 6.8. Sieť dosahuje veľmi dobré výsledky v rámci snímok HRF z dôvodu, že sieť bola natrénovaná za zvyšných 22 snímkach z rovnakej databázy. Pri testovaní databázy CHASE už úspešnosť segmentácie poklesla, no nie tak výrazne, ako to je pri databázy STARE. Je to z dôvodu, že sieť je natrénovaná na detailoch cievneho riečiska. Snímky z databázy STARE majú v jednom testovacom snímku obsiahnuté takmer celé cievne riečište, pričom jeden pod-obraz HRF obsahuje záber na pár detailných ciev.

Tab. 6.8: Výsledky úspešnosti siete testovanej na databázach HRF, STARE a CHASE

Testovacia dátová sada	Testovacia úspešnosť
23 snímok HRF	0.9648
CHASE	0.9485
STARE	0.9293

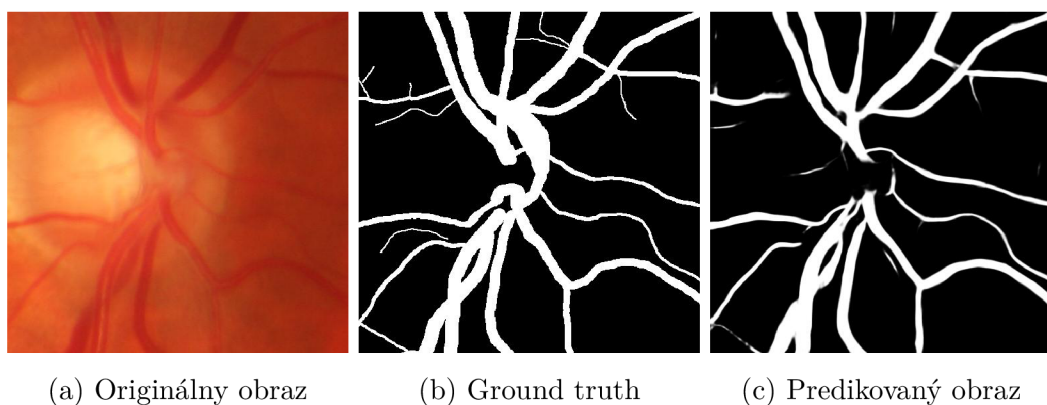
Na Obr. 6.6 je zobrazená ukážka segmentovaného pod-obrazu vybraného zo snímku zdravého pacienta. Pod-obraz je vybraný z miesta, kde sa nachádza optický disk, ktorý je vstupom ciev a optického nervu do sietnice. U zdravého človeka sa na farebných fundus snímkach, viď Obr.6.6a, vyznačuje jasne žltou až bielou farbou. Výsledok segmentácie je zobrazený po 35 epochách. Pri porovnaní predikovaného obrazu (Obr. 6.6c) s originálnym obrazom (Obr.6.6a) a ground truth (Obr. 6.6b)

sa segmentácia hrubších ciev približuje viac k originálnemu snímku ako ku ground truth. Na základe ground truth bola vypočítaná aj senzitivita a špecificita predikovaného pod-obrazu. Senzitivita predikovaného pod-obrazu zdravého oka je 96.82 % a špecificita 87.61 %.



Obr. 6.6: Ukážka výsledku segmentácie pod-obrazu pacienta so zdravým okom

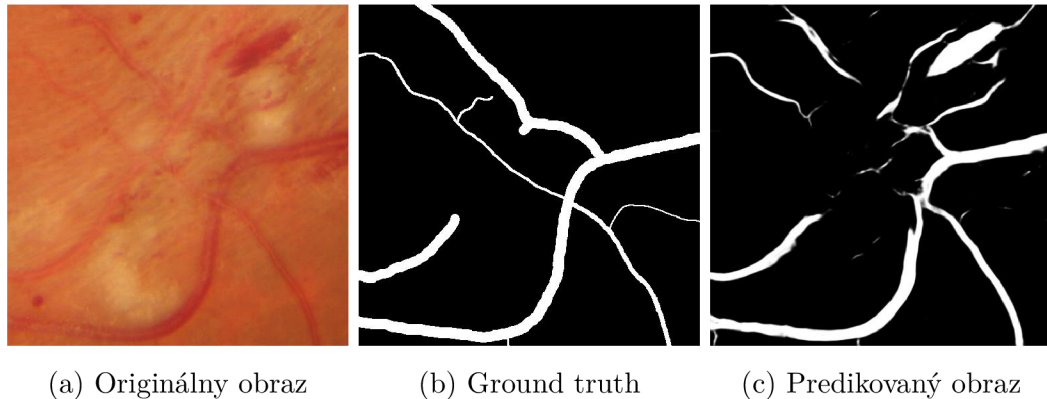
Druhý príklad pod-obrazu je vybraný zo snímku sietnice pacienta s glaukómovou chorobou. Oproti zdravému oku sa vyznačuje hlavne tým, že má v prvom rade iné zafarbenie oblasti optického disku a líši sa aj jeho tvarom. Na Obr. 6.7 je vidieť, že sieť segmentovala len viditeľné časti ciev. Cievky v presvetlenej oblasti optického disku, ktoré v originálnom obraze nie sú viditeľné a v ground truth sú vyznačené, sa vo výslednom predikovanom obraze nenachádzajú. Výsledná senzitivita poklesla oproti pod-obrazu zdravého oka na 88.03 % a špecificita sa zvýšila na 90.98 %.



Obr. 6.7: Ukážka výsledku segmentácie pod-obrazu pacienta trpiacim glaukomovou chorobou

Na Obr. 6.8 je zobrazená tretia ukážka pod-obrazu, ktorý je vybraný zo snímku sietnice pacienta s diabetickou retinopatiou. Oblasť pod-obrazu sa už nenachádza

v oblasti optického disku, ako v predošlých príkladoch. V tomto prípade bol vybraný pod-obraz, v ktorom sa nachádzajú červené škvrny, odborné nazývané ako hemoragie. Tieto škvrny sú jedným z príznakov človeka trpiaceho diabetom. V porovnaní s originálnym obrazom (Obr.6.8a) je vo výslednom predikovanom obraze (Obr. 6.8c) vidieť, že sieť segmentovala aj tieto nežiadúce patológie na sietnici. Senzitivita opäť klesla na 87.42 % a špecificita zvýšila na 90.12 %.

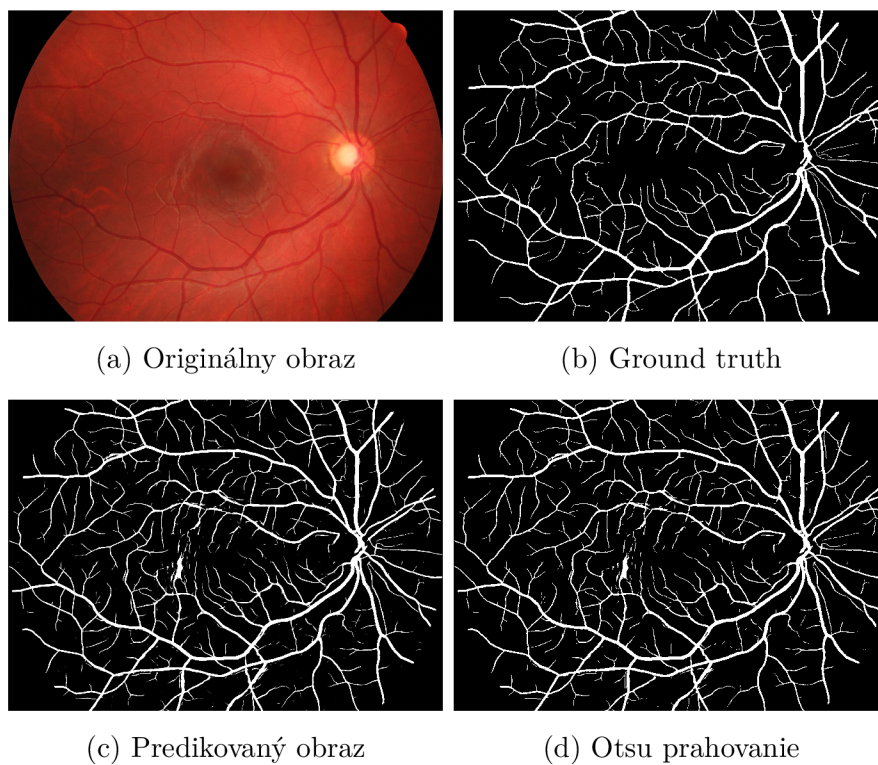


Obr. 6.8: Ukážka výsledku segmentácie pod-obrazu pacienta s diabetickou retinopatiou

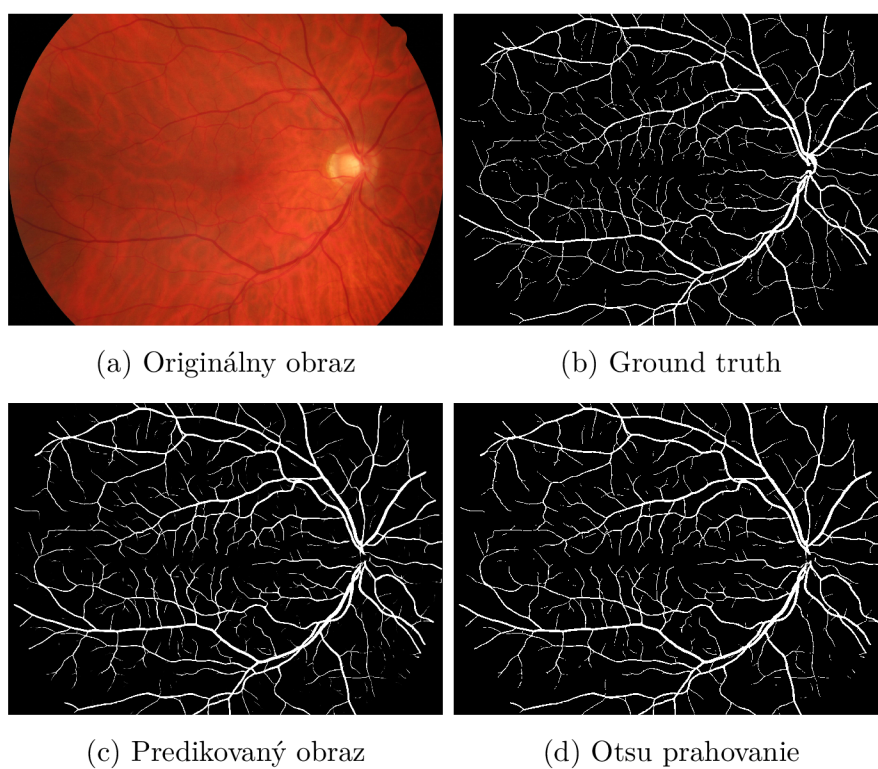
Následne boli výsledné predikované pod-obrazy veľkosti 512x512 pixelov rekonštruované späť do pôvodného obrazu veľkosti 2336x3504 pixelov. Pri spätnej rekonštrukcii obrazu boli v prekrývajúcich sa častiach pod-obrazov vybrané maximálne hodnoty. Výsledky rekonštruovaných obrazov, z ktorých boli vybrané ukážkové pod-obrazy, sú znázornené spolu s prislúchajúcimi originálmi, ground truth a následne prahovaného obrazu pomocou Otsu metódy na Obr. 6.9 - Obr.6.11. V prípade zdravého oka je vidieť v originálnom obraze (Obr. 6.9a) v oblasti okolo makuly biele obrysy okolo ciev. Sieť tieto obrysy vysegmentovala spolu z cievami. Výsledná senzitivita rekonštruovaného predikovaného obrazu bola 97.46 % a špecificita bola 88.10 %. Výsledná plocha pod ROC krivkou (AUC) bola 95.38 %

V prípade sietnice glaukomatika má oblasť optického disku často iné sfarbenie ako je to v prípade zdravej sietnice (Obr. 6.10). Výsledná senzitivita medzi originálnym a rekonštruovaným obrazom je o 3.3 % nižšia ako u zdravého oka, pričom špecificita sa zvýšila o 1.68 %. Výsledná plocha pod krivkou bola 95.70 %.

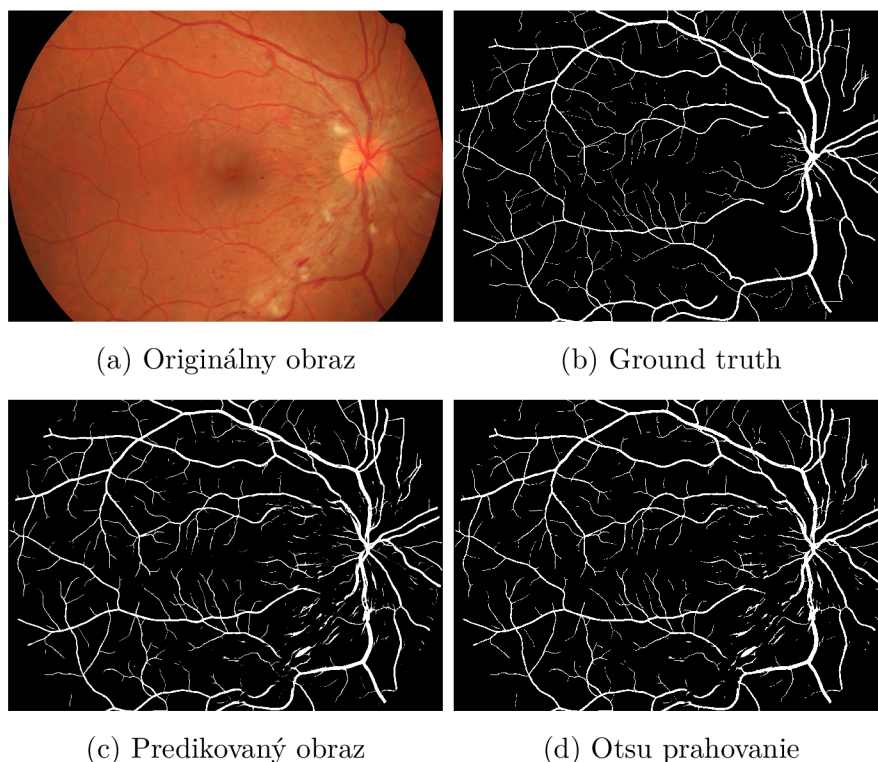
Po rekonštrukcii pod-obrazov sietnice s diabetickou retinopatiou sú na Obr. 6.11c vidieť patologické útvary, nazývané hemoragie, ktoré sú v originálnom obraze (Obr. 6.11a) viditeľné ako červené fličky. Architektúra siete pri týchto snímkach dosahuje najnižšie hodnoty senzitivity. Hodnota senzitivity klesla na 94.36 % a špecificita stúpla na 89.89 %. Výsledná plocha pod ROC krivkou bola 95.38 %. Plocha pod ROC



Obr. 6.9: Ukážka segmentácie rekonštruovaného snímku pacienta so zdravým okom



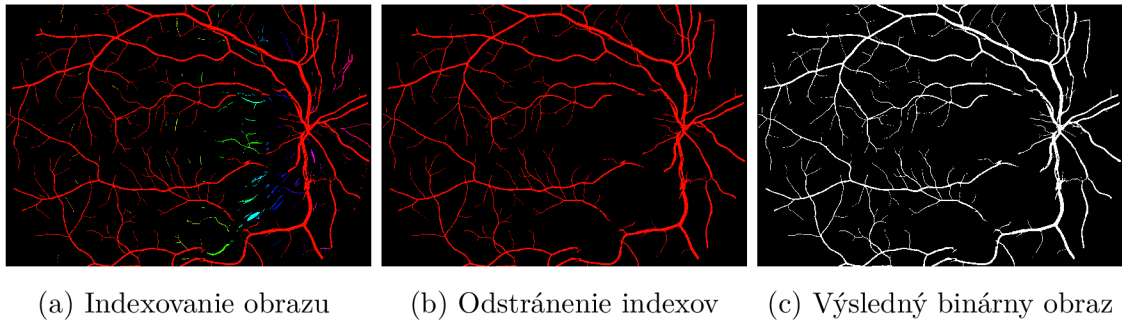
Obr. 6.10: Ukážka segmentácie rekonštruovaného snímku pacienta s glaukómom



Obr. 6.11: Ukážka segmentácie rekonštruovaného snímku pacienta s DR

krivkou sa od predošlých snímok nelíšila a dokonca bola rovnaká ako u zdravého oka.

Z výsledného binárneho obrazu je žiadúce odstrániť patológie, ktoré sú vo výslednom segmente cievného riečiska nežiadúce. Z pohľadu spracovania obrazov sa jedná o oblasti, ktoré nie sú spojené so segmentovaným riečiskom a predstavujú tak malé regióny so značne menšou plochou, než je segmentovaný cievy strom. Použitím funkcie `bwlabel`, ktorá indexuje 2D - binárny obraz na časti, ktoré sú medzi sebou spojené, bol výsledný binárny obraz rozdelený do príslušných indexov. Pozadie snímku bolo zafarbené na čierne a pre zobrazenie ďalších indexov bola použitá farebná mapa `hsv`. Cievné riečište je na Obr.6.12a vyznačené červenou farbou. Ostatné farby znázorňujú cievy, ktoré nie sú pripojené na hlavný cievy strom, resp. v tomto prípade sú farebne vyznačené patológie na sietnici spôsobené ochorením diabetická retinopatia. Účelom tohto indexovania je zbaviť sa týchto nežiadúcich objektov vo výslednom segmente. Preto sú všetky indexy okrem indexu, v ktorom sa nachádza hlavné cievy riečište, odstránené, viď Obr. 6.12b. Výsledné cievné riečište je zobrazené na Obr. 6.12c.



Obr. 6.12: Ukážka výsledného snímku po odstránení indexov, nenapojených na hlavné cieвне riečisko

Výsledky segmentácie cieвнеho riečiska sú porovnávané so segmentačnou metódou, ktorá využíva prispôsobenú filtráciu a minimálnu chybu prahovej techniky na extrakciu binárneho cieвнеho riečiska. Pri segmentácii je používaný iba zelený (G) kanál pôvodného RGB obrazu, ktorý je predspracovaný pomocou nerovnomernej korekcie osvetlenia spolu so zvýraznením kontrastu. Metodológia je podrobne popísaná, viď [49]. Výsledky senzitivity a špecificity prahovaného obrazu a oblast pod ROC krivkou sú pre všetky rekonštruované obrazy uvedené v Tab. 6.9 - 6.11. V tabulkách sú pre porovnanie dosiahnutých výsledkov rovnako uvedené výsledky segmentácie podľa [49]. V Tab. 6.9 je uvedené vyhodnotenie segmentácie cieвнеho riečiska pre zdravú skupinu jedincov snímok 08_h - 15_h databáze HRF. Predošlé snímky nie sú porovnané z dôvodu ich použitia pre tréovanie siete. Priemerné hodnoty senzitivity, špecificity a presnosti dosiahli lepšie výsledky v prípade architektúry siete tréovanej na pod-obrazoch HRF oproti porovnáwanej metóde. Priemerná oblasť pod ROC krivkou je nižšia o 0.7 % oproti porovnáwanej metóde. V Tab. 6.10 sú uvedené výsledky segmentácie cieвнеho riečiska skupiny jedincov s glaukómovou chorobou. V Tab. 6.11 sú znázornené výsledky pre skupinu jedincov s diabetickou retinopatiou. Rovnako ako pri jedincoch so zdravým okom aj v prípade jedincov s glaukómovou chorobou a jedincov s diabetickou retinopatiou boli dosiahnuté lepšie výsledky senzitivity, špecificity a presnosti oproti porovnáwanej metóde. Výsledky plochy pod ROC krivkou boli opäť nižšie ako pri porovnáwanej metóde [49]. Najlepšie a najhoršie výsledky porovnávaných snímok databáze HRF sú vyznačené tučným fontom a najlepšíe a najhoršie výsledky segmentovaných snímok architektúrou siete sú zobrazené na Obr. 6.13.

Tab. 6.9: Porovnanie výsledkov segmentácie cievného riečiska s metódou publikovanou Odstrčilík et al. pre zdravú skupinu jedincov

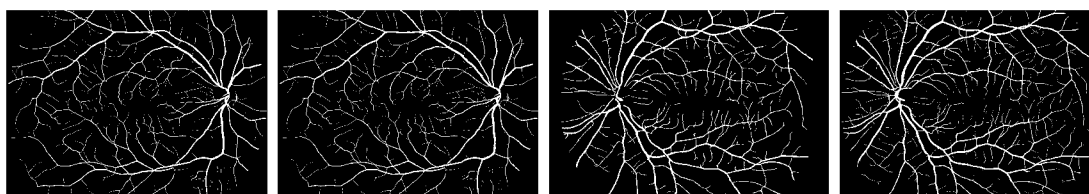
Špecifikácia obrazu	Odstrčilík [49]				Sieť trébovaná na pod-obrazoch			
	SE	SP	ACC	AUC	SE	SP	ACC	AUC
08_h	0.7560	0.9829	0.9560	0.9739	0.8025	0.9920	0.9716	0.9726
09_h	0.7339	0.9805	0.9578	0.9721	0.7473	0.9931	0.9726	0.9626
10_h	0.7393	0.9789	0.9545	0.9687	0.7954	0.9900	0.9720	0.9583
11_h	0.8290	0.9748	0.9584	0.9833	0.8552	0.9817	0.9690	0.9754
12_h	0.7877	0.9804	0.9551	0.9778	0.8467	0.9837	0.9683	0.9753
13_h	0.7949	0.9608	0.9424	0.9701	0.8576	0.9776	0.9664	0.9755
14_h	0.8190	0.9602	0.9456	0.9768	0.8344	0.9797	0.9661	0.9728
15_h	0.8292	0.9739	0.9605	0.9844	0.7476	0.9924	0.9718	0.9584
Priemer	0.7861	0.9741	0.9543	0.9759	0.8108	0.9863	0.9697	0.9689
Std	±0.0391	±0.0089	±0.0067	±0.0058	±0.0452	±0.0063	±0.0026	±0.0072

Tab. 6.10: Porovnanie výsledkov segmentácie cievného riečiska s metódou publikovanou Odstrčilík et al. pre skupinu jedincov s glaukómovou chorobou

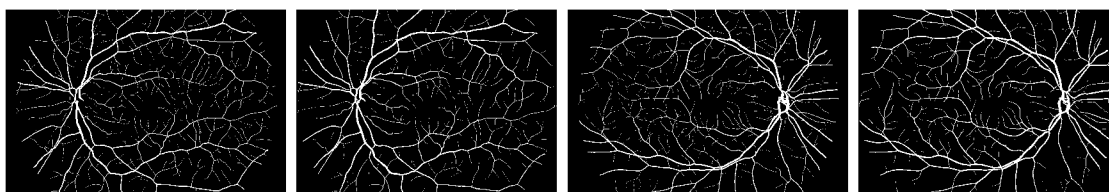
Špecifikácia obrazu	Odstrčilík [49]				Sieť trébovaná na pod-obrazoch			
	SE	SP	ACC	AUC	SE	SP	ACC	AUC
08_g	0.8292	0.9518	0.9419	0.9725	0.8233	0.9823	0.9707	0.9673
09_g	0.7637	0.9705	0.9543	0.9736	0.7799	0.9854	0.9708	0.9588
10_g	0.8061	0.9652	0.9528	0.9732	0.8098	0.9819	0.9698	0.9585
11_g	0.8243	0.9555	0.9437	0.9721	0.7995	0.9830	0.9679	0.9642
12_g	0.8172	0.9533	0.9400	0.9698	0.7960	0.9818	0.9653	0.9588
13_g	0.7684	0.9639	0.9482	0.9670	0.7831	0.9808	0.9664	0.9570
14_g	0.7195	0.9621	0.9418	0.9550	0.7521	0.9810	0.9637	0.9512
15_g	0.7633	0.9677	0.9500	0.9655	0.8112	0.9756	0.9627	0.9586
Priemer	0.7865	0.9613	0.9466	0.9686	0.7944	0.9815	0.9672	0.9593
Std	±0.0386	±0.0069	±0.0055	±0.0062	±0.0224	±0.0028	±0.0031	±0.0045

Tab. 6.11: Porovnanie výsledkov segmentácie cievného riečiska s metódou publikovanou Odstrčilík et al. pre skupinu jedincov s diabetickou retinopatiou

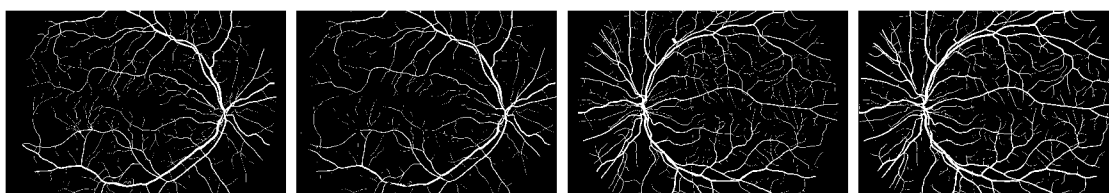
Špecifikácia obrazu	Odstrčilík [49]				Sieť trébovaná na pod-obrazoch			
	SE	SP	ACC	AUC	SE	SP	ACC	AUC
09_dr	0.6728	0.9674	0.9438	0.9487	0.6276	0.9881	0.9619	0.9184
10_dr	0.8006	0.9561	0.9400	0.9655	0.8576	0.9640	0.9540	0.9577
11_dr	0.8075	0.9574	0.9424	0.9690	0.8457	0.9743	0.9626	0.9669
12_dr	0.7446	0.9665	0.9487	0.9646	0.6777	0.9872	0.9646	0.9371
13_dr	0.7597	0.9707	0.9536	0.9701	0.7115	0.9899	0.9691	0.9471
14_dr	0.7266	0.9630	0.9417	0.9561	0.7216	0.9836	0.9621	0.9391
15_dr	0.7730	0.9430	0.9298	0.9558	0.8015	0.9772	0.9648	0.9573
Priemer	0.7550	0.9606	0.9429	0.9614	0.7490	0.9806	0.9627	0.9462
Std	±0.0463	±0.0094	±0.0074	±0.0080	±0.0874	±0.0094	±0.0046	±0.0151



(a) Zdravý jedinci



(b) Jedinci s glaukomatickým ochorením



(c) Jedinci s diabetickou retinopatiou

Obr. 6.13: Ukážka najlepších (vľavo - ground truth a prahovaný obraz) a najhorších (vpravo - ground truth a prahovaný obraz) výsledkov segmentácie cievného riečiska databázy HRF

V Tab. 6.12 sú priemerné výsledky testovania siete 23 snímok databázy HRF natrénovanej na pod-obrazoch databázy HRF porovnané s výsledkami segmentácie cievného riečiska 23 snímok HRF pomocou metódy prahovania podľa Odstrčilík et al. [49]. Ďalej sú výsledky siete porovnané s výsledkami konvolučnej neurónovej siete navrhutej Vengalil et al. [64], ktorá je upravenou verziou architektúry DeepLab [18]. Táto architektúra je tvorená tromi konvolučnými blokmi. Prvý konvolučný blok obsahuje 2 konvolučné vrstvy, po ktorých nasleduje aktivačná funkcia ReLU. Konvolučné vrstvy obsahujú 64 filtrov veľkosti 3x3 s krokom (stride) 1. Druhý konvolučný blok je tvorený rovnako 2 konvolučnými vrstvami, za ktorými nasleduje aktivačná funkcia ReLU. Konvolučné vrstvy obsahujú 128 filtrov veľkosti 3x3 s krokom 1. Tretí konvolučný blok je tvorený 3 konvolučnými vrstvami, za ktorými je aktivačná vrstva ReLU. Konvolučné vrstvy obsahujú 256 filtrov veľkosti 3x3 s krokom 1. Za každým konvolučným blokom nasleduje vrstva Max Pooling. Posledná konvolučná vrstva má jeden filter veľkosti 1x1. Výsledné obrazy sú šedotónové a následne binarizované prahom. Architektúra bola testovaná na 23 náhodne zvolených snímkach a priemerná testovacia úspešnosť dosahuje 93.94 %. Modifikovaná architektúra, trénovaná na pod-obrazoch, dosahuje najlepšie výsledky s vybranými metódami testovaných na databázy HRF okrem plochy pod ROC krivkou, kde metóda publikovaná Odstrčilík et. al. [49] dosahuje o 1.05 % lepšie výsledky.

Tab. 6.12: Porovnanie výsledkov databázy HRF s dostupnými segmentačnými metódami

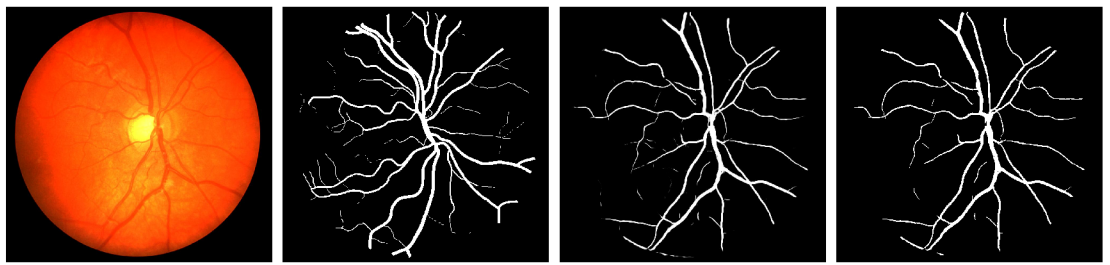
Segmentačné metódy	SE	SP	ACC	AUC
Tréovanie na pod-obrazoch HRF	0.7847	0.9828	0.9665	0.9581
Vengalil et al. [64]	-	-	0.9394	0.8940
Odstrcilik et al. - 23 snímok HRF [49]	0.7759	0.9653	0.9479	0.9686

Sieť bola testovaná tiež na databázy CHASE. Výsledky databáze CHASE po prahovaní sú uvedené v Tab. 6.13. Úspešnosť sa oproti sieti natrénovanej na výrezoch, použitím Modifikácie dátovej sady 3, znížila o 0.54 %. Výraznejšie za znížila priemerná plocha pod ROC krivkou na 89.26 %, čím klesla o 6.02 % oproti sieti tréovanej na výrezoch. Najlepšie a najhoršie výsledky segmentácie databáze CHASE sú znázornené na Obr. 6.14. Pri najhoršie segmentovanom obraze je vidieť v origináli (Obr. 6.14e), v pravom hornom rohu na okraji zorného poľa fialové presvetlenie. Na Obr. 6.14g je vidieť, že sieť v tejto oblasti nesprávne segmentovala cieвне riečisko.

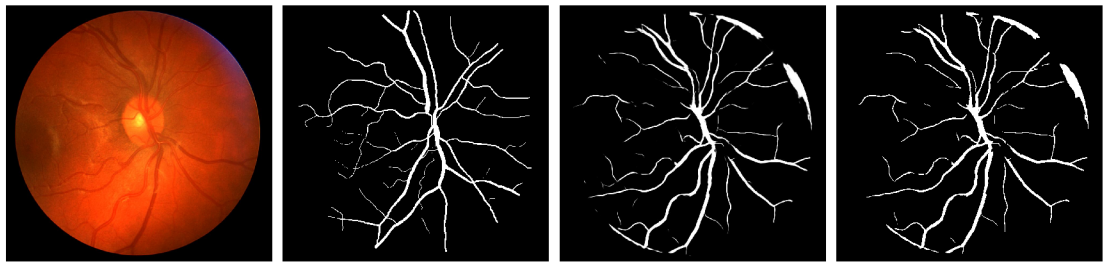
Pod-obrazy databázy HRF obsahujú detailné časti cievného riečiska, pričom pod-obrazy databázy STARE obsahujú takmer celé cieвне riečisko v jednom pod-obraze. Preto je pre porovnanie otestovaná databáza STARE na tréovanej sieti pomocou pod-obrazov databázy HRF.

Tab. 6.13: Výsledky modifikovanej siete trénovej na pod-obrazoch databázy HRF
 - testovaná databáza CHASE po prahovaní

Špecifikácia obrazu	SE	SP	ACC	AUC
Image_01L	0.6389	0.9931	0.9684	0.9097
Image_01R	0.5992	0.9989	0.9593	0.8893
Image_02L	0.5720	0.9863	0.9526	0.8728
Image_02R	0.5892	0.9827	0.9520	0.8829
Image_03L	0.6012	0.9903	0.9598	0.8855
Image_03R	0.6371	0.9878	0.9613	0.9015
Image_04L	0.5297	0.9936	0.9572	0.8727
Image_04R	0.5846	0.9808	0.9503	0.8829
Image_05L	0.6325	0.9863	0.9583	0.9056
Image_05R	0.6495	0.9855	0.9569	0.9050
Image_06L	0.5132	0.9909	0.9569	0.8671
Image_06R	0.6216	0.9887	0.9612	0.8942
Image_07L	0.6145	0.9799	0.9514	0.8870
Image_07R	0.6232	0.9806	0.9534	0.8973
Image_08L	0.6121	0.9903	0.9658	0.9020
Image_08R	0.6807	0.9900	0.9703	0.9214
Image_09L	0.6801	0.9798	0.9646	0.9077
Image_09R	0.6505	0.9910	0.9737	0.9064
Image_10L	0.6420	0.9823	0.9609	0.8995
Image_10R	0.6514	0.9791	0.9588	0.8892
Image_11L	0.7512	0.9750	0.9631	0.8829
Image_11R	0.7666	0.9871	0.9753	0.8829
Image_12L	0.6372	0.9762	0.9518	0.8899
Image_12R	0.6706	0.9822	0.9593	0.9088
Image_13L	0.5751	0.9905	0.9650	0.8914
Image_13R	0.5257	0.9898	0.9604	0.8643
Image_14L	0.6421	0.9872	0.9634	0.8997
Image_14R	0.5515	0.9897	0.9641	0.8716
Priemer	0.6240	0.9859	0.9606	0.8926
Std	±0.0571	±0.0052	±0.0065	±0.0171

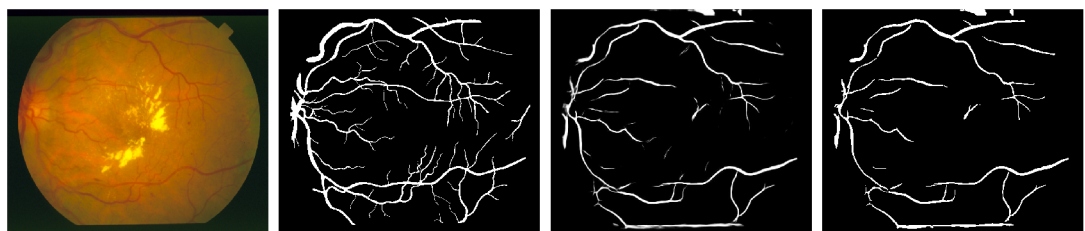


(a) Originálny obraz (b) Ground truth (c) Predikcia (d) Otsu prahovanie

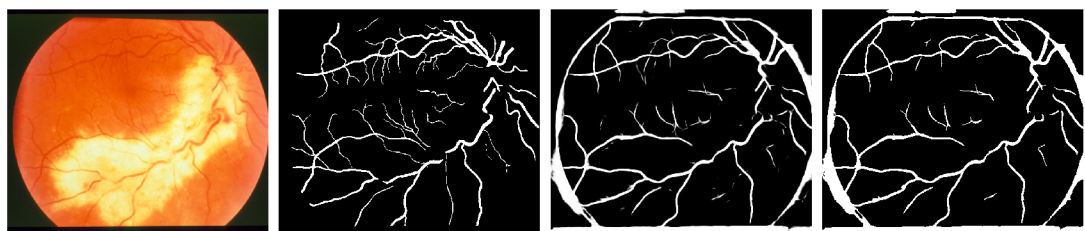


(e) Originálny obraz (f) Ground truth (g) Predikcia (h) Otsu prahovanie

Obr. 6.14: Ukážka najlepšej (hore) a najhoršej (dole) segmentácie snímku databázy CHASE, sieť trébovaná na pod-obrazoch databázy HRF



(a) Originálny obraz (b) Ground truth (c) Predikcia (d) Otsu prahovanie



(e) Originálny obraz (f) Ground truth (g) Predikcia (h) Otsu prahovanie

Obr. 6.15: Výsledky najlepšej (hore) a najhoršej (dole) segmentácie snímku databázy STARE, sieť trébovaná na vytvorených pod-obrazoch HRF

Výsledky databázy STARE po rekonštrukcii a prahovaní sú uvedené v Tab. 6.14. Napriek tomu, že úspešnosť segmentácie klesla na 92.47 %, to stále nie je zlý výsledok z pohľadu, že je sieť trénovaná na odlišný typ obrazov. Najlepší a najhorší výsledok segmentácie je zobrazený na Obr. 6.15. Vo výsledku sieť segmentuje len hlavné cievy a má problém s oblasťou okraja zorného poľa (FOV), ktorú segmentuje tiež, viď Obr. 6.15g. Najlepšia úspešnosť výslednej segmentácie obrazu je 94.31 % a najhoršia len 87.57 %.

Tab. 6.14: Výsledky modifikovanej siete trénovanej na pod-obrazoch databázy HRF - testovaná databáza STARE po prahovaní

Špecifikácia obrazu	SE	SP	ACC	AUC
im0001	0.4628	0.9848	0.9431	0.8412
im0002	0.4753	0.9663	0.9336	0.8183
im0003	0.5602	0.9512	0.9278	0.8343
im0004	0.4387	0.9536	0.9154	0.8063
im0005	0.6327	0.9344	0.9072	0.8461
im0044	0.6546	0.8923	0.8757	0.8411
im0077	0.8376	0.9264	0.9192	0.9281
im0081	0.8437	0.9436	0.9361	0.9385
im0082	0.7241	0.9544	0.9363	0.8928
im0139	0.6941	0.9434	0.9233	0.8771
im0162	0.6839	0.9469	0.9282	0.8608
im0163	0.7712	0.9399	0.9268	0.9109
im0235	0.7741	0.9374	0.9228	0.9068
im0236	0.7486	0.9447	0.9270	0.9000
im0239	0.5904	0.9563	0.9247	0.8469
im0240	0.6279	0.9527	0.9195	0.8768
im0255	0.7400	0.9412	0.9232	0.8733
im0291	0.7400	0.9551	0.9381	0.8516
im0319	0.3319	0.9557	0.9289	0.7491
im0324	0.4671	0.9698	0.9363	0.8511
Priemer	0.6340	0.9475	0.9247	0.8626
Std	±0.1418	±0.0185	±0.0143	±0.0434

V Tab. 6.15 sú uvedené priemerné hodnoty všetkých dosiahnutých výsledkov segmentácie cievneho riečiska na databázy STARE. Pod nimi sú taktiež uvedené tri dostupné metódy segmentácie, ktoré používajú tiež konvolučné neurónové siete. Kon-

volučná neurónová sieť podľa Fu et al. [10] je obsahuje päť vrstiev, ktoré majú plné prepojenie s neurónmi v predchádzajúcej vrstve. Sieť je doladovaná (fine-tuning) algoritmom spätného šírenia chyby (backpropagation). Navrhnutá sieť je trébovaná na databázy ARIA a testovaná na databázy STARE s úspešnosťou 95.45 %.

Tab. 6.15: Porovnanie dosiahnutých výsledkov testovanej databázy STARE s dostupnými segmentačnými metódami

Segmentačné metódy	Trébovacía sada	SE	SP	ACC	AUC
Sieť trébovaná na zmenšených obrazoch	DRIVE, ARIA, HRF, CHASE	0.7879±0.0947	0.9790±0.0057	0.9649±0.0065	0.9694±0.0190
Sieť trébovaná na výrezoch	DRIVE, ARIA, HRF, CHASE	0.7651±0.1048	0.9746±0.0094	0.9595±0.0066	0.9580±0.0147
Sieť trébovaná na pod-obrazoch	HRF	0.6340±0.1418	0.9475±0.0185	0.9247±0.0143	0.8626 ±0.0434
Fu et al. [10]	ARIA	0.7140	-	0.9545	-
Liskowski et al.(NO-POOL) [41]	DRIVE	0.7763	0.9768	0.9505 ±0.0093	0.9595±0.0222
Liskowski et al. (NO-POOL) [41]	STARE	0.7867±0.0689	0.9754±0.0099	0.9566 ±0.0082	0.9785±0.0066
Liskowski et al. (NO-POOL) [41]	DRIVE, STARE	-	-	0.9566 ±0.0067	0.9776±0.0068
Liskowski et al. (NO-POOL) [41]	CHASE	0.7091	0.9791	0.9525	0.9543
Liskowski et al. (NO-POOL-SP) [41]	CHASE	0.8145	0.9866	0.9696	0.9880
Fan et al. [8]	STARE	0.7834	0.9799	0.9654	-

Ďalšia konvolučná neurónová sieť NO-POOL podľa Liskowski et al. [41] je tvorená štyrmi konvolučnými vrstvami a tromi plne prepojenými vrstvami. Prvé dve konvolučné vrstvy majú každá počet filtrov 64, zvyšné dve majú 128 filtrov. Sieť používa konvolučné filtre veľkosti 3x3 s krokom (stride) 1. V tejto práci tiež experimentovali s rôznymi databázami, kde trébovanie a testovanie prebiehalo na rovnakých a rozličných databázach. Ak bola sieť trébovaná a testovaná na rovnakej databázy, samozrejme dosahovala lepšie výsledky, ako keď sa trébovacía a testovacía sada líšila. Táto konvolučná sieť dosiahla 95.66 % úspešnosť testovania na databázy STARE, pričom bola sieť trébovaná na databázy DRIVE. Rovnakú testovaciu úspešnosť dosiahla pri trébovaní na databázy STARE a DRIVE. Posledná konvolučná sieť NO-POOL-SP podľa Liskowski et al. [41] dosiahla najlepšie výsledky testovacej úspešnosti, a to 96.96 %, čím prevýšila sieť trébovanú na zmenšených obrazoch o 0.47 %. Sieť bola trébovaná na databázy CHASE a testovaná na databázy STARE.

Posledná konvolučná sieť podľa Fan et al. [8] je tvorená štyrmi konvolučnými blokmi. Prvý konvolučný blok obsahuje 2 konvolučné vrstvy so 64 filtrami. Druhý konvolučný blok obsahuje 3 konvolučné vrstvy so 128 filtrami. Tretí konvolučný blok obsahuje 3 konvolučné vrstvy s 256 filtrami. Posledný štvrtý konvolučný blok obsahuje 3 konvolučné vrstvy s 512 filtrami. Za každou konvolučnou vrstvou nasleduje aktivačná funkcia ReLU a posledná vrstva siete je plne prepojená vrstva. Táto sieť dosahuje po sieti NO-POOL-SP podľa Liskowski et al. [41] druhé najlepšie výsledky testovacej úspešnosti, a to 96,54 %. Je nutné podotknúť, že sieť bola trébovaná na rovnakej databáze ako bola testovaná. Na rozdiel od siete NO-POOL-SP podľa

Liskowski et al. [41] a od vytvorenej modifikovanej siete trénovanej na zmenšených obrazoch, predstavenej v tejto práci, ktorá bola trénovaná na 4 rôznych databázach, pričom testovaná bola na 5 databáze, ktorá nie je zahrnutá v trénovacej sade.

V Tab. 6.16 sú uvedené výsledky segmentácie snímkov databázy CHASE dosiahnuté na sieti trénovanej na výrezoch, pomocou Modifikácie dátovej sady 3, a sieti trénovanej na pod-obrazoch databázy HRF. Pod nimi sú uvedené výsledky konvolučnej neuronovej siete podľa Fan et al. [8], ktorých štruktúra neuronovej siete je uvedená pri hodnotení výsledkov databázy STARE. Táto metóda dosiahla testovaciu úspešnosť 95.73 %. Architektúra trénovaná na pod-obrazoch dosiahla úspešnosť 96.06 % a architektúra siete trénovaná na výrezoch dosiahla najlepšiu testovaciu úspešnosť až 96.54

Tab. 6.16: Porovnanie dosiahnutých výsledkov testovanej databázy CHASE s dostupnými segmentačnými metódami

Segmentačné metódy	Trénovacia sada	SE	SP	ACC	AUC
Sieť trénovaná na výrezoch	DRIVE,ARIA, HRF, STARE	0.7138±0.0469	0.9844±0.0036	0.9654±0.0043	0.9528±0.0088
Sieť trénovaná na pod-obrazoch	HRF	0.6240±0.0571	0.9859±0.0052	0.9606±0.0065	0.8926 ±0.0171
Fan et al. [8]	CHASE	0.7656	0.9704	0.9573	-

V Tab. 6.17 sú uvedené tri najlepšie prístupy segmentácie cievného riečiska pre každú jednu testovanú databázu, tj. STARE, CHASE a HRF. Tieto metódy už nie sú založené na princípe konvolučných neurónových sietí. Žiaden z výsledkov trénovania databázy STARE pomocou konvolučných neurónových sietí uvedených v Tab. 6.15 nepresiahol úspešnosť metódy podľa Wang et al. [68], ktorá dosiahla až 98.10 % úspešnosť segmentácie cievného riečiska databázy STARE. Rovnako metóda podľa Moghimirad et al. dosiahla úspešnosť 97.60 %, čo je o 0.64 lepší výsledkok, ako v prípade konvolučnej neuronovej siete podľa Liskowski et al. [41], ktorá dosiahla najlepšie výsledky v rámci segmentácie cievného riečiska databázy STARE pomocou konvolučných neurónových sietí. V prípade testovania databázy CHASE a databázy HRF dosiahla modifikovaná architektúra siete v oboch prípadoch lepšie výsledky ako dostupné metódy uvedené v Tab. 6.17.

Tab. 6.17: Prehľad najlepších segmentačných metód cievného riečiska

Metóda segmentácie	STARE			
	SE	SP	ACC	AUC
Wang et al. [68]	0.8100	0.9790	0.9810	0.9750
Moghimirad et al. [45]	0.8130	0.9910	0.9760	0.9680
Imani et al. [23]	0.7500	0.9750	0.9590	-
Metóda segmentácie	CHASE			
	SE	SP	ACC	AUC
Roychowdhury et al. [53]	0.7200	0.9820	0.9530	0.9530
Fraz et al. [9]	0.7220	0.9710	0.9470	0.9710
Vostatek et al. (Soares method) [65]	0.6900	0.9770	0.9470	0.9640
Metóda segmentácie	HRF			
	SE	SP	ACC	AUC
Cheng et al. [19]	0.7040	0.9860	0.9610	-
Vostatek et al. (Soares method) [65]	0.7340	0.9800	0.9580	0,9700
Christodoulidis et al. [20]	0.8510	0.9580	0.9480	-

7 ZÁVER

Témou diplomovej práce bola segmentácia cievného riečiska v snímkach retiny pomocou hlbokých neurónových sietí. Úlohou práce bolo oboznámenie sa s princípom konvolučných neurónových sietí so zameraním na segmentáciu objektov z obrazu. Vybrané existujúce úlohy konvolučných neurónových sietí boli otestované s cieľom oboznámiť sa s programovacím jazykom Python a využitím prostredia Tensorflow, v ktorom je implementované riešenie práce. Na základe rešerše a testovania vybraných úloh bola pre segmentáciu cievného riečiska zvolená architektúra siete U-net, ktorá je koncipovaná na použitie v medicínskych obrazoch.

Z piatich voľne dostupných databáz retinálnych snímok boli vytvorené tri dátové sady pre tréovanie zvolenej architektúry. Prvá dátová sada pre tréovanie bola vytvorená zmenšením pôvodných snímok databáz DRIVE, CHASE, HRF a ARIA. Vytvorená dátová sada zmenšených obrazov bola použitá na tréovanie pôvodnej zvolenej architektúry. Tréovanie bolo uskutočnené na RGB snímkach a taktiež na zelenom kanáli RGB snímok. Sieť bola tréovaná pri počte epoch dvadsať, štyridsať, šesťdesiat a osemdesiat. Po šesťdesiatich epochách začala validačná úspešnosť siete klesať, čím došlo k jej pretréovaniu. Z tohto dôvodu bola do tréovania siete vložená funkcia, ktorá pri poklese validačnej úspešnosti tréovanie siete včas zastaví. Po štyridsiatich epochách dosahovala sieť tréovaná na RGB snímkoch 94.79 % úspešnosť. V prípade tréovania siete na G kanáli sieť dosiahla 94.71 % úspešnosť. Keďže úspešnosť tréovania na RGB snímkoch bola vyššia, ako na G kanáli, pri ďalších tréovaniach boli používané len dátové sady obsahujúce RGB snímky.

Na základe výsledkov bola pôvodná architektúra zmenená, s cieľom dosiahnuť lepšie výsledky. Modifikovaná architektúra obsahuje len jednu vrstvu Dropout v piatom konvolučnom bloku. Zo štvrtého konvolučného bloku bola táto vrstva odstránená, čím sa zlepšila konvergencia siete. Rovnako bola z pôvodnej architektúry odstránená dekonvolučná vrstva Upsampling, ktorá v tréovaní siete zavádza do obrazov artefakty, čím znižuje jej úspešnosť. Táto vrstva bola nahradená transponovanou konvolučnou vrstvou Conv2Dtranspose. Pri znížení počtu epoch z štyridsať na tridsať päť sieť dosiahla vyššiu tréovaciu, validačnú a testovaciu úspešnosť ako pri pôvodnej architektúre. Sieť tréovaná na tejto dátovej sade bola testovaná na databáze STARE. Úspešnosť siete bola vypočítaná na základe senzitivity a špecificity segmentovaných snímok po prahovaní Otsu metódou. Priemerná úspešnosť segmentácie bola 96.48 %.

Druhá dátová sada na tréovanie bola vytvorená orezaním pôvodných snímok. Táto dátová sada sa delí ešte na tri ďalšie modifikácie. Prvá modifikácia dátovej sady bola tréovaná a testovaná na výrezoch všetkých dostupných databáz, pričom snímky v tréovacej a testovacej sade boli odlišné. Cieľom bolo zistiť, ako sa sieť

vysporiada so segmentovaním cievneho riečiska s tak rôznymi dátami, tj. snímkami, ktoré obsahovali takmer celé cievne riečisko, a snímkami, na ktorých bol len detail cievy. Sieť dosiahla 95.75 % testovaciu úspešnosť. Čím je sieť trébovaná na detailnejších snímkach cievneho riečiska, tým dosahuje pre daný testovaný prípad lepšie výsledky. Druhá modifikácia dátovej sady pre trébovanie bola vytvorená z databáz DRIVE, ARIA, HRF a CHASE. Sieť bola testovaná na databáze STARE. Priemerná testovacia úspešnosť dosiahla 95.95 %. Tretia modifikácia dátovej sady bola vytvorená z databáz DRIVE, ARIA, HRF a STARE. Testovaná je na databáze CHASE, pričom výsledná priemerná úspešnosť segmentácie bola 96.56 %.

Tretia posledná dátová sada bola vytvorená z pod-obrazov databázy HRF. Na trébovanie bolo použitých 22 pôvodných snímkov a na testovanie 23 snímkov. Priemerná testovacia úspešnosť segmentácie na týchto snímkach bola 96.65 %. Sieť trébovaná na pod-obrazoch bola tiež otestovaná na databázy CHASE a databázy STARE. Databáza CHASE dosiahla 96.06 % úspešnosť a databáza STARE 92.47 % úspešnosť segmentácie cievneho riečiska. Na prvý pohľad sú tieto výsledky nižšie ako v predošlých prípadoch. Vzhľadom na to, že sieť bola trébovaná na detailných záberoch ciev a testovaná bola na snímkach databázy STARE, ktoré obsahujú takmer celý cievny strom, je tento výsledok stále uchádzajúci. Sieť cievne riečisko vysegmentovala. Horšie výsledky boli spôsobené segmentáciou v oblasti zorného poľa, ktoré prispieva ku zvýšeniu počtu falošne segmentovaných pixelov.

Najlepší výsledok segmentácie databázy STARE bol 96.54 %, dosiahnutý trébovaním siete na zmenšených obrazoch. Pri porovnaní dosiahnutých výsledkov segmentácie databázy STARE úspešnosť neprevýšila segmentáciu pomocou konvolučnej neurónovej siete podľa Liskowski et al [41]. Táto konvolučná sieť dosahuje 96.96 % úspešnosť segmentácie cievneho riečiska. Doposiaľ najlepšie výsledky, tj. 98.10 %, dosahuje metóda podľa Wang et al. [68], ktorá nie je založená na segmentácii konvolučnými neurónovými sieťami. V prípade databázy CHASE bol najlepší výsledok segmentácie dosiahnutý trébovaním siete na orezaných obrazoch. Segmentácia dosiahla 96.54 %, čím dosiahla najlepšie výsledky v rámci porovnávaní konvolučných neurónových sietí a ostatných metód, ktoré nie sú založené na hlbokom učení. Rovnako dosiahnutá úspešnosť segmentácie databázy HRF, ktorá bola 96.65 %, dosiahla najlepšie výsledky medzi porovnávanými metódami. Druhý najlepší výsledok metódou hlbokého učenia dosahuje Fan et al. [8] s úspešnosťou segmentácie 93.94 %. V prípade ostatných metód je najlepší výsledok segmentácie databázy HRF 96.10 % metódou podľa Cheng et al. [19].

Vytvorená sieť trébovaná na orezaných snímkach a sieť trébovaná na pod-obrazoch, prekonáva úspešnosť segmentácie databáz CHASE a HRF, oproti ostatným metódam segmentácie pomocou hlbokého učenia, ale aj metódam, ktoré nie sú koncipované na segmentácii cievneho riečiska pomocou konvolučných neurónových sietí.

V prípade segmentácie databázy STARE úspešnosť segmentácie neprevýšila segmentáciu konvolučnou neurónovou sieťou podľa Liskowski et al. [41]. V rámci segmentácie tejto databázy, je stále najúspešnejšia metóda podľa Wang et al. [68]. Práca by mohla byť v budúcnosti rozšírená porovnaním výsledkov vytvorenej siete s druhou najpoužívanejšiu architektúrou siete SegNet, ktorá v originálnych snímkach veľmi dobre segmentuje tenké štruktúry, ako sú pouličné lampy. Z tohto dôvodu je možné predpokladať jej vhodnosť pre segmentáciu cievneho riečiska. Ďalším zaujímavým prístupom by bolo vytvoriť architektúru siete, ktorá by neobsahovala dekonvolučné vrstvy, ako architektúry U-net a SegNet.

LITERATÚRA

- [1] BADRINARAYANAN, Vijay, Alex KENDALL, Roberto CIPOLLA. *SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation*. IEEE transactions on pattern analysis and machine intelligence [online]. 2017, 39(12), 2481 [cit. 2018-05-15]. DOI: 10.1109/TPAMI.2016.2644615.
- [2] BANKHEAD, Peter, C. Norman SCHOLFIELD, J. Graham MCGEOWN, Tim M CURTIS, Teresa SERRANO-GOTARREDONA. *Fast Retinal Vessel Detection and Measurement Using Wavelets and Edge Location Refinement (Retinal Vessel Detection and Measurement)*. PLoS ONE [online]. San Francisco, USA: Public Library of Science, 2012, 7(3), e32435 [cit. 2017-12-15]. DOI: 10.1371/journal.pone.0032435.
- [3] BOTTOU, Léon. *Large-scale machine learning with stochastic gradient descent*. Proceedings of COMPSTAT'2010. Physica-Verlag HD, 2010. p. 177-186.
- [4] CARDONA, Albert, Stephan SAALFELD, Stephan PREIBISCH, et al. *An Integrated Micro- and Macroarchitectural Analysis of the Drosophila Brain by Computer-Assisted Serial Section Electron Microscopy (Serial TEM Analysis of Drosophila Brain Circuitry)*. PLoS Biology [online]. San Francisco, USA: Public Library of Science, 2010, 8(10), e1000502 [cit. 2018-05-15]. DOI: 10.1371/journal.pbio.1000502. ISSN 1544-9173.
- [5] CELEBI, Ömer Cengiz. *Neural Networks and Pattern Recognition Using MATLAB: Chapter 7 Neural Network* [online]. [cit. 2017-11-22]. Dostupné z: https://www.byclb.com/TR/Tutorials/neural_networks/ch7_1.htm
- [6] DISHASHREE, Gupta. *Transfer learning and The art of using Pre-trained Models in Deep Learning*. [online]. [cit. 2017-11-23]. Dostupné z: <https://www.analyticsvidhya.com/blog/2017/06/transfer-learning-the-art-of-fine-tuning-a-pre-trained-model/>
- [7] DUCHI, John, Elad HAZAN, Yoram SINGER. *Adaptive subgradient methods for online learning and stochastic optimization*. Journal of Machine Learning Research, 2011, 12.Jul: 2121-2159.
- [8] FAN, Zhun, Jia-Jie MO. *Automated blood vessel segmentation based on denoising auto-encoder and neural network*. In: Machine Learning and Cybernetics (ICMLC), 2016 International Conference on [online]. IEEE, 2016, 2, s. 849-856 [cit. 2018-05-15]. DOI: 10.1109/ICMLC.2016.7872998.

- [9] FRAZ, Muhammad Moazam, Paolo REMAGNINO, Andreas HOPPE, Buny-
arit UYYANONVARA, Alicja R. RUDNICKA, Christopher G. OWEN, Sa-
rah A. BARMAN. *An Ensemble Classification-Based Approach Applied to Re-
tinal Blood Vessel Segmentation*. In: Biomedical Engineering, IEEE Transac-
tions on [online]. USA: IEEE, 2012, 59(9), 2538-2548 [cit. 2017-12-15]. DOI:
10.1109/TBME.2012.2205687. ISSN 0018-9294.
- [10] FU, Huazhu, Yanwu XU, Damon Wing Kee WONG, Jiang LIU. *Retinal ves-
sel segmentation via deep learning network and fully-connected conditional
random fields*. In: Biomedical Imaging (ISBI), 2016 IEEE 13th Internati-
onal Symposium on [online]. IEEE, 2016, s. 698-701 [cit. 2018-05-15]. DOI:
10.1109/ISBI.2016.7493362.
- [11] GIBIANSKY, Andrew. *Convolutional Neural Networks*. [online]. [cit. 2017-10-
18]. Dostupné z: [http://andrew.gibiansky.com/blog/machine-learning/
convolutionalneural-networks/](http://andrew.gibiansky.com/blog/machine-learning/convolutionalneural-networks/)
- [12] HAO HUANG, Wushour, Wushour HAIHUA XU, Wushour XIANHUI WANG,
Wushour SILAMU. *Maximum F1-Score Discriminative Training Criterion for
Automatic Mispronunciation Detection*. Audio, Speech, and Language Proces-
sing, IEEE/ACM Transactions on [online]. USA: IEEE, 2015, 23(4), 787-797
[cit. 2018-03-12]. DOI: 10.1109/TASLP.2015.2409733. ISSN 2329-9290.
- [13] HIJAZI, Samer, Rishi KUMAR, Chris ROWEN. *Using convolutional neural
networks for image recognition*. Tech. Rep., 2015.[online]. [cit. 2017-10-18]. Do-
stupné z: https://ip.cadence.com/uploads/901/cnn_wp-pdf
- [14] HINTON, Geoffrey. et al. *Neural networks for machine learning lecture 6a
overview of mini-batch gradient descent*. 2012. [online]. [cit. 2017-03-18]. Do-
stupné z: [https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_
slides_lec6.pdf](https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf)
- [15] HOLČÍK, Jiří, Martin KOMENDA. et al. *Matematická biologie: e-learningová
učebnice* [online]. [cit. 2017-10-18]. 1. vydání. Brno: Masarykova univerzita,
2015. ISBN 978-80-210-8095-9.
- [16] HOOVER, Adam, Valentina KOUZNETSOVA a Michael GOLDBAUM. *Loca-
ting blood vessels in retinal images by piecewise threshold probing of a matched
filter response*. Medical Imaging, IEEE Transactions on [online]. USA: IEEE,
2000, 19(3), 203-210 [cit. 2017-12-15]. DOI: 10.1109/42.845178. ISSN 0278-0062.

- [17] HOLLEMANS, Matthijs. *Convolutional neural networks on the iPhone with VGGNet*. 2016. [Online]. [cit. 2017-10-19]. Dostupné z: <http://machinethink.net/blog/convolutional-neural-networks-on-the-iphone-with-vggnet/>
- [18] CHEN, Liang-Chieh, George PAPANDREOU, Iasonas KOKKINOS, Kevin MURPHY, Alan L. YUILLE. *DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs*. Pattern Analysis and Machine Intelligence, IEEE Transactions on [online]. USA: IEEE, 2018, 40(4), 834-848 [cit. 2018-05-15]. DOI: 10.1109/T-PAMI.2017.2699184. ISSN 0162-8828.
- [19] CHENG, Erkang, Liang DU, Yi WU, Ying ZHU, Vasileios MEGALOOIKONOMOU, Haibin LING. *Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features*. Machine Vision and Applications [online]. Berlin/Heidelberg: Springer Berlin Heidelberg, 2014, 25(7), 1779-1792 [cit. 2018-05-13]. DOI: 10.1007/s00138-014-0638-x. ISSN 0932-8092.
- [20] CHRISTODOULIDIS, Argyrios, Thomas HURTUT, Housseem Ben TAHAR a Farida CHERIET. *A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images*. Computerized Medical Imaging and Graphics [online]. Elsevier, 2016, 52, 28-43 [cit. 2018-05-13]. DOI: 10.1016/j.compmedimag.2016.06.001. ISSN 0895-6111.
- [21] CHOLLET, Francois. *Building powerful image classification models using very little data* 2016. [online]. [cit. 2017-09-23]. Dostupné z: <https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html>
- [22] QIAN, Ning. *On the momentum term in gradient descent learning algorithms*. Neural networks,[online] 1999, 12.1: 145-151 [cit. 2018-05-13].
- [23] IMANI, Elaheh, Malihe JAVIDI a Hamid-Reza POURREZA. *Improvement of retinal blood vessel detection using morphological component analysis*. Computer Methods and Programs in Biomedicine [online]. Elsevier Ireland, 2015, 118(3), 263-279 [cit. 2018-05-13]. DOI: 10.1016/j.cmpb.2015.01.004. ISSN 0169-2607.
- [24] JAN, Jíří. *Číslíková filtrace, analýza a restaurace signálů*. Vyd. 2. Brno: VUTIUUM, 2002, 427 s. ISBN 80-214-1558-4.

- [25] JOCIĆ, Marko. *Deep Learning Tutorial for Kaggle Ultrasound Nerve Segmentation competition, using Keras* [online]. [cit. 2017-12-19]. Dostupné z: <https://github.com/jocicmarko/ultrasound-nerve-segmentation>
- [26] Kaggle datasets. *Digit Recognizer*[online]. [cit. 2017-12-12]. Dostupné z: <https://www.kaggle.com/c/digit-recognizer>.
- [27] Kaggle datasets. *Dogs vs. Cats*[online]. [cit. 2017-12-12]. Dostupné z: <https://www.kaggle.com/c/dogs-vs-cats>.
- [28] Kaggle datasets. *Ultrasound Nerve Segmentation*[online]. [cit. 2017-12-12]. Dostupné z: <https://www.kaggle.com/c/ultrasound-nerve-segmentation>.
- [29] Kaggle Inc. *Data Science and Machine Learning*. [online]. [cit. 2017-12-12]. Dostupné z: <https://www.kaggle.com/>.
- [30] KALINOVSKY, Alexander, Vassili KOVALEV. *Lung Image Segmentation Using Deep Learning Methods and Convolutional Neural Networks*. 2016. [online]. [cit. 2017-12-19].
- [31] KARPATY, Andrej. *Neural Networks Part 1: Setting Up the Architecture*. Notes for CS231n Convolutional Neural Networks for Visual Recognition, Stanford University. 2015. [online]. [cit. 2017-10-12]. Dostupné z: <http://cs231n.github.io/neural-networks-1/>
- [32] KARPATY, Andrej. *Module 2: Convolutional Neural Networks*. Notes for CS231n Convolutional Neural Networks for Visual Recognition, Stanford University. 2015. [online]. [cit. 2017-10-11]. Dostupné z: <http://cs231n.github.io/convolutional-networks/>
- [33] KARPATY, Andrej. *CS231n: Convolutional Neural Networks for Visual Recognition*. [online]. [cit. 2017-11-23]. Dostupné z: <http://cs231n.github.io/transfer-learning/>
- [34] Keras Documentation. [online]. [cit. 2017-12-12]. Dostupné z: <https://keras.io/>
- [35] KINGMA, Diederik P., Jimmy BA. *Adam: A Method for Stochastic Optimization*. [online]. 2014 [cit. 2018-05-10]. arXiv preprint arXiv:1412.6980, 2014.
- [36] KOHLER, Thomas, Attila BUDAI, Martin F. KRAUS, Jan ODSTRČILIK, Georg MICHELSON a Joachim HORNEGGER. *Automatic no-reference quality assessment for retinal fundus images using vessel segmentation*.

- In: Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on [online]. IEEE, 2013, s. 95-100 [cit. 2017-12-15]. DOI: 10.1109/CBMS.2013.6627771.
- [37] KOLÁŘ, Petr. *Věkem podmíněná makulární degenerace*. 1. vyd. Praha: Grada, 2008, 148 s. ISBN 978-802-4726-052.
- [38] KRIZHEVSKY, Alex, Ilya SUTSKEVER, Geoffrey HINTON. *ImageNet classification with deep convolutional neural networks*. Communications of the ACM [online]. ACM, 2017, 60(6), 84-90 [cit. 2018-05-15]. DOI: 10.1145/3065386. ISSN 0001-0782.
- [39] LE, Quoc V., et al. *On optimization methods for deep learning*. Proceedings of the 28th International Conference on International Conference on Machine Learning [online]. Omnipress, 2011. p. 265-272 [cit. 2018-04-05]. Dostupné z: <http://ai.stanford.edu/~quocle/LeNgiCoaLahProNg11.pdf>
- [40] LECUN, Yann, Yoshua BENGIO, Geoffrey HINTON. *Deep learning*. Nature [online]. Nature Publishing Group, 2015, 521(7553), 436 [cit. 2018-05-15]. DOI: 10.1038/nature14539. ISSN 0028-0836.
- [41] LISKOWSKI, Pawel, Krzysztof KRAWIEC. *Segmenting Retinal Blood Vessels With Deep Neural Networks*. Medical Imaging, IEEE Transactions on [online]. USA: IEEE, 2016, 35(11), 2369-2380 [cit. 2018-05-15]. DOI: 10.1109/TMI.2016.2546227. ISSN 0278-0062.
- [42] LONG, Jonathan, Evan SHELHAMER, Trevor DARRELL. *Fully Convolutional Networks for Semantic Segmentation*. Pattern Analysis and Machine Intelligence, IEEE Transactions on [online]. USA: IEEE, 2017, 39(4), 640-651 [cit. 2018-05-15]. DOI: 10.1109/TPAMI.2016.2572683. ISSN 0162-8828.
- [43] MEHROTRA, Kishan, Chulukuri K. MOHAN, Sanjay RANKA. *Elements of artificial neural networks*. The MIT Press. 1997. ISBN 0-262-13328-8.
- [44] MILLER, W. Thomas, Paul J. WERBOS, Richard SUTTON et al. *Neural networks for control*. MIT press, 1995.
- [45] MOGHIMIRAD, Elahe, Seyed HAMID REZATOFIGHI, Hamid SOLTANIAN-ZADEH. *Retinal vessel segmentation using a multi-scale medialness function*. Computers in Biology and Medicine [online]. 2012, 42(1), 50-60 [cit. 2018-05-13]. DOI: 10.1016/j.compbiomed.2011.10.008. ISSN 0010-4825.
- [46] *Národní Gridové Infrastruktury* MetaCentrum NGI. [online]. [cit. 2017-11-19] Dostupné z: <https://www.metacentrum.cz/cs/>

- [47] NESTEROV, Yurii. *A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$* . In: Doklady AN USSR. 1983. p. 543-547.
- [48] NOWLAN, Steven J., Geoffrey E. HINTON. *Simplifying neural networks by soft weight-sharing*. Neural computation, 1992, 4.4: 473-493.
- [49] ODSTRČILÍK, Jan. *Analýza obrazových dat sítnice pro podporu diagnostiky glaukomu*. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2014.
- [50] OTSU, Nobuyuki. *A Threshold Selection Method from Gray-Level Histograms*. Systems, Man and Cybernetics, IEEE Transactions on [online]. USA: IEEE, 1979, 9(1), 62-66 [cit. 2018-05-15]. DOI: 10.1109/TSMC.1979.4310076. ISSN 0018-9472.
- [51] ROKYTA, Richard, et al. *Fyziologie: pro bakalářská studia v medicíně, přírodních a tělovýchovných oborech*. 1. vyd. Praha: ISV nakladatelství, 2000, 359 s. ISBN 80-858-6645-5.
- [52] RONNEBERGER, Olaf, Philipp FISCHER, Thomas BROX. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. [online]. MICCAI. Springer, Cham, 2015. p. 234-241 [cit. 2018-05-15].
- [53] ROYCHOWDHURY, Sohini, Dara D. KOOZEKANANI, Keshab K. PARHI. *Blood Vessel Segmentation of Fundus Images by Major Vessel Extraction and Subimage Classification*. Biomedical and Health Informatics, IEEE Journal of [online]. USA: IEEE, 2015, 19(3), 1118-1128 [cit. 2018-05-13]. DOI: 10.1109/JBHI.2014.2335617. ISSN 2168-2194.
- [54] SIMONYAN, Karen, Andrew ZISSERMAN. *Very deep convolutional networks for large-scale image recognition*. [online]. arXiv preprint arXiv:1409.1556, 2014 [cit. 2018-05-15].
- [55] SOCHER, Richard, et al. *Parsing natural scenes and natural language with recursive neural networks*. Proceedings of the 28th international conference on machine learning (ICML-11). [online]. 2011. p. 129-136 [cit. 2018-03-25]. Dostupné z: <http://ai.stanford.edu/~ang/papers/icml11-ParsingWithRecursiveNeuralNetworks.pdf>
- [56] SRIVASTAVA, Nitish, et al. *Dropout: a simple way to prevent neural networks from overfitting*. Journal of machine learning research [online]. 2014, 15.1: 1929-1958 [cit. 2017-11-24]. Dostupné z: <http://jmlr.org/papers/volume15/srivastava14a.old/srivastava14a.pdf>

- [57] STAAL, Joes, Michael D. ABRAMOFF, Meindert NIEMEIJER, Max A. VIERGEVER, Bram van GINNEKEN. *Ridge-based vessel segmentation in color images of the retina*. Medical Imaging, IEEE Transactions on [online]. USA: IEEE, 2004, 23(4), 501-509 [cit. 2017-12-15]. DOI: 10.1109/TMI.2004.825627. ISSN 0278-0062.
- [58] STERGIOU, Christos, Dimitrios SIGANOS. *NEURAL NETWORKS* [online]. [cit. 2017-11-22]. Dostupné z: https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#The%20Learning%20Process
- [59] TAPE Thomas G. *Interpreting Diagnostic Tests* [online]. University of Nebraska Medical Center. [cit. 2018-3-13]. Dostupné z: <http://gim.unmc.edu/dxtests/ROC3.htm>
- [60] TensorFlow. *An open-source software library for Machine Intelligence*. [online]. [cit. 2017-12-12]. Dostupné z: <https://www.tensorflow.org>
- [61] The Python Deep Learning library. *Deep Learning Toolkit for Medical Imaging*. [online]. [cit. 2017-12-12]. Dostupné z: <https://dltk.github.io/>
- [62] TIELEMAN, Tijmen; HINTON, Geoffrey. *Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude*. COURSERA: Neural networks for machine learning, 2012, 4.2: 26-31.
- [63] VAN ERKEL, Arian. R., Peter M. Th. PATTYNAMA. *Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology*. European Journal of Radiology [online]. Elsevier Ireland, 1998, 27(2), 88-94 [cit. 2018-03-13]. DOI: 10.1016/S0720-048X(97)00157-5. ISSN 0720-048X.
- [64] VENGALIL, Sunil, Kumar, Neelam SINHA, Srinivas S. S. KRUTHIVENTI, R. Venkatesh BABU. *Customizing CNNs for blood vessel segmentation from fundus images*. In: Signal Processing and Communications (SPCOM), 2016 International Conference on [online]. IEEE, 2016, s. 1-4 [cit. 2018-05-13]. DOI: 10.1109/SPCOM.2016.7746702.
- [65] VOSTATEK, Pavel, et al. *Blood vessel segmentation in the analysis of retinal and diaphragm images*. Acta Universitatis Lappeenrantaensis, 2018.
- [66] WALEK, Petr, Martin LAMOŠ, Jiří JAN. *Analýza biomedicínských obrazů*. Brno, 2013, 1-138.

- [67] WANG, Bei, Xuefeng YAN, Qingchao JIANG, Zhaomin LV. *Generalized Dice's coefficient-based multi-block principal component analysis with Bayesian inference for plant-wide process monitoring*. Journal of Chemometrics [online]. 2015, 29(3), 165-178 [cit. 2018-03-12]. DOI: 10.1002/cem.2687. ISSN 0886-9383.
- [68] WANG, Shuangling, Yilong YIN, Guibao CAO, Benzhen WEI, Yuanjie ZHENG, Gongping YANG. *Hierarchical retinal blood vessel segmentation based on feature and ensemble learning*. Neurocomputing [online]. Elsevier B.V, 2015, 149, 708-717 [cit. 2018-05-13]. DOI: 10.1016/j.neucom.2014.07.059. ISSN 0925-2312.
- [69] WINDMAIER, Eric, P., Hershel RAFF, Kevin T. STRANG. *Vander's human physiology: The Mechanisms of Body Function*. 13th edition. New York: McGraw-Hill, 2014. ISBN 978-0-07-337830-5.
- [70] ZEILER, Matthew D. *ADADELTA: an adaptive learning rate method*. arXiv preprint arXiv:1212.5701, [online]. 2012 [cit. 2018-04-02].
- [71] ZHI XU, Hao. *Implementation of deep learning framework – Unet, using Keras*. [online]. [cit. 2017-02-27]. Dostupné z: <https://github.com/zhixuhao/unet>

ZOZNAM SYMBOLOV, VELIČÍN A SKRATIEK

Adam	Adaptive Moment Estimation
AMD	degenerácia buniek v žltej škvrne – Age Related Macular Degeneration
ANN	umelá neurónová sieť – Artificial Neural Network
AUC	oblasť pod ROC krivkou – Area Under the ROC Curve
CNN	konvolučná neurónová sieť – Convolutional Neural Network
ConvNet	konvolučná neurónová sieť – Convolutional Neural Network
DR	diabetická retinopatia
ED-CNN	encoder-decoder convolutional neural networks
FOV	zorné pole – field of view
FC	plne prepojené vrstvy – fully connected layers
FCN	plne prepojená konvolučná neurónová sieť – Fully Connected Convolutional Neural Network
ReLU	rectified linearity unit
RGB	aditívny farebný model, ktorý vytvára všetky farby z kombinácie červenej (R), zelenej (G) a modrej (B) farby
RNFL	vrstva nervových vlákien sietnice – Retinal Nerve Fiber Layer
ROC	Receiver Operating Characteristic
RTG	rentgenový (obraz)
SGD	Stochastic Gradient Descent
Std	stredná kvadratická odchýlka – Standard deviation

ZOZNAM PRÍLOH

A Obsah priloženého DVD

95

A OBSAH PŘILOŽENÉHO DVD

Súčasťou diplomovej práce je DVD, ktoré obsahuje nasledujúce položky:

```
/. . . . . koreňový adresár priloženého DVD
├ Stanislava_Serečunová_DP.pdf . . . . . elektronická verzia diplomovej práce
├ zdrojové kódy . . . . . zdrojové kódy diplomovej práce
│   ├── README.txt
│   ├── data.py
│   ├── model.py
│   ├── train.py
│   ├── test.py
│   ├── roc_curve.py
│   ├── patch.m
│   ├── preprocessing.py
│   ├── reconstruction.m
│   ├── calculation_results.m
│   └── remove_indexes.m
├ Dátové sady . . . . . vytvorené dátové sady
│   ├── Dátové sady na tréningovanie
│   │   ├── Databáza zmenšených obrazov
│   │   │   ├── image
│   │   │   └── label
│   │   ├── Databáza orezaných obrazov
│   │   │   ├── Modifikácia dátovej sady 2
│   │   │   │   ├── image
│   │   │   │   └── label
│   │   │   └── Modifikácia dátovej sady 3
│   │   │       ├── image
│   │   │       └── label
│   │   └── Databáza pod-obrazov
│   │       ├── image
│   │       └── label
│   └── Dátové sady na testovanie
│       ├── Databáza STARE
│       │   ├── image
│       │   ├── label
│       │   └── image bez ground truth
│       ├── Databáza CHASE
│       │   ├── image
│       │   └── label
│       └── Databáza HRF
│           ├── image
│           └── label
└ Výsledky . . . . . výsledky modelov architektúry siete
    ├── Sieť tréningovaná na zmenšených obrazoch
    └── výsledky predikcie
```