
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2019

Bc. Andrea Rozhoňová



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

**METODY HLUBOKÉHO UČENÍ PRO SEGMENTACI CÉV A
OPTICKÉHO DISKU V OFTALMOLOGICKÝCH
SEKVENCÍCH**

DEEP LEARNING METHODS FOR VESSEL AND OPTIC DISC SEGMENTATION IN OPHTHALMOLOGIC
SEQUENCES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Andrea Rozhoňová

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Branislav Hesko

BRNO 2019

Diplomová práce

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

Ústav biomedicínského inženýrství

Studentka: Bc. Andrea Rozhoňová

ID: 174526

Ročník: 2

Akademický rok: 2018/19

NÁZEV TÉMATU:

Metody hlubokého učení pro segmentaci cév a optického disku v oftalmologických sekvencích

POKYNY PRO VYPRACOVÁNÍ:

1) Nastudujte si problematiku oftalmoskopie, detekce optického disku a cév. Provedte literární rešerši používaných metod pro segmentaci. 2) Navrhněte vhodnou metodu pro automatickou segmentaci optického disku a cév v oftalmologických sekvencích. 3) Vytvořte dostatečnou databázi manuálně segmentovaných obrazů pro trénování. 4) Vámi navrženou metodu implementujte, proveďte natrénování modelu a ověřte jeho funkčnost. 5) Vhodným způsobem vyhodnoťte dosažené výsledky a proveďte diskuzi.

DOPORUČENÁ LITERATURA:

[1] ALMAZROA, A. et al., Optic disc and optic cup segmentation methodologies for glaucoma image detection: A survey, Journal of ophthalmology, 2015, 28p.

[2] LISKOWSKI, P. and KRAWIEC, K. Segmenting retinal blood vessels with deep neural networks, IEEE Transactions on medical imaging, 2016, 11p.

Termín zadání: 4.2.2019

Termín odevzdání: 17.5.2019

Vedoucí práce: Ing. Branislav Hesko

Konzultant:

prof. Ing. Ivo Provazník, 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

Cílem následující diplomové práce bylo studium problematiky segmentace optického disku a cév sítnice v oftalmologických sekvencích. Teoretická část práce shrnuje principy různých přístupů z oblasti hlubokého učení, které se v souvislosti s danou problematikou využívají. Na základě teoretické části jsou navrženy metody pro segmentaci optického disku a pro segmentaci cév sítnice založené na konvolučních neuronových sítích Linknet, PSPNet, Unet a MaskRCNN. Popisem jejich implementace a následným zhodnocením se zabývá praktická část diplomové práce.

KLÍČOVÁ SLOVA

cévy sítnice, hluboké učení, Keras, konvoluční neuronové sítě, MaskRCNN, oftalmologie, optický disk, segmentace obrazu, Unet

ABSTRACT

The aim of the following thesis was to study the issue of optical disc and retinal vessels segmentation in ophthalmologic sequences. The theoretical part of the thesis summarizes the principles of different approaches in the field of deep learning, which are used in connection with the given issue. Based on the theoretical part, methods for optical disk segmentation and retinal vessel segmentation based on the convolutional neural networks Linknet, PSPNet, Unet and MaskRCNN are proposed. The practical part of the thesis deals with the description of their implementation and subsequent evaluation.

KEYWORDS

Retina vessels, Deep learning, Keras, Convolutional neural networks, MaskRCNN, Ophthalmology, Optic disc, Image segmentation, Unet

ROZHOŇOVÁ, Andrea. *Metody hlubokého učení pro segmentaci cév a optického disku v oftalmologických sekvencích*. Brno, Rok, 77 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce: Ing. Branislav Hesko

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Metody hlubokého učení pro segmentaci cév a optického disku v oftalmologických sekvencích“ jsem vypracovala samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autorka uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědoma následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autorky

PODĚKOVÁNÍ

Ráda bych poděkovala svému vedoucímu diplomové práce Ing. Branislavu Heskovi za trpělivost, ochotu a mnoho cenných rad při psaní diplomové práce.

Brno

.....

podpis autorky

Obsah

Úvod	11
1 Teoretická část studentské práce	12
1.1 Oftalmoskopie	12
1.1.1 Přímá oftalmoskopie	12
1.1.2 Nepřímá oftalmoskopie	13
1.1.3 Nemoci detekovatelné oftalmoskopií	14
1.2 Konvoluční neuronové sítě - základní prvky	16
1.2.1 Konvoluční vrstva	17
1.2.2 Sdružovací vrstva	18
1.2.3 Plně propojená vrstva	19
1.3 Konvoluční neuronové sítě - princip učení	20
1.3.1 Chybové funkce	20
1.3.2 Optimalizační algoritmy	22
1.3.3 Transfer learning	25
1.4 Konvoluční neuronové sítě - segmentace obrazu	27
1.4.1 První konvoluční neuronové sítě a segmentace	27
1.4.2 Unet	29
1.4.3 SegNet	30
1.4.4 LinkNet	31
1.4.5 PSPNet	32
1.4.6 Mask RCNN	33
1.5 Metody segmentace optického disku	35
1.5.1 Fuzzy konvergence krevních cév	35
1.5.2 Fuzzy C-means a iterativní prahování	37
1.5.3 Modifikace architektury Unet	38
1.6 Metody segmentace cév sítnice	41
1.6.1 Level-set metoda	41
1.6.2 Metoda zesílení lokálních znaků	42
1.6.3 Metoda založená na 3-fázovém modelu hlubokého učení	44
2 Výsledky studentské práce	47
2.1 Použité nástroje	47
2.2 Příprava dat	47
2.2.1 Tvorba trénovací databáze	47
2.2.2 Rozšíření učební množiny	48
2.3 Segmentace optického disku	50

2.3.1	Předzpracování obrazů pro segmentaci optického disku	50
2.3.2	MaskRCNN	52
2.3.3	Metoda Unet	54
2.3.4	Zhodnocení použitých metod	55
2.4	Segmentace cév sítnice	58
2.4.1	Předzpracování obrazů pro segmentaci cév	59
2.4.2	Metoda Unet	60
2.4.3	Metoda Linknet	61
2.4.4	Zhodnocení použitých metod	61
2.5	Oftalmologické sekvence	64
2.5.1	Zhodnocení metod na oftalmologických sekvencích	65
2.6	Zhodnocení dosažených výsledků	67
3	Závěr	70
	Literatura	71
	Seznam symbolů, veličin a zkratk	76
A	Obsah přiloženého CD	77

Seznam obrázků

1.1	Přímý oftalmoskop	13
1.2	Nepřímý oftalmoskop	14
1.3	Schéma nepřímého oftalmoskopu	14
1.4	Glaukom	15
1.5	Další nemoci	15
1.6	Konvoluční filtry	17
1.7	Posouvání konvoluční masky	17
1.8	Aktivační funkce	18
1.9	Podvzorkování obrazu	18
1.10	Schéma CNN	19
1.11	Klesání ve směru gradientu	23
1.12	Klesání ve směru gradientu	25
1.13	Hluboká neuronová síť	27
1.14	Plně konvoluční neuronová síť	28
1.15	Architektura Unet	29
1.16	Architektura SegNet	30
1.17	Architektura LinkNet	31
1.18	Architektura PSPNet	32
1.19	Metoda Faster RCNN - detekované objekty	33
1.20	Architektura Mask RCNN	34
1.21	Fuzzy konvergence	36
1.22	Ekvalizace jasových hodnot	37
1.23	Architektura modifikované metody Unet	39
1.24	Modifikovaná metoda Unet - příklad	40
1.25	Segmentované cévy sítnice	41
1.26	Třídy pro rozřazování pixelů cév.	43
1.27	SVM - dodatečné zpracování a pouhá klasifikace	44
1.28	Anotované cévy - třífázový model.	45
1.29	Architektura třífázového modelu pro segmentaci cév.	46
2.1	Využité nástroje	47
2.2	Anotované snímky cév a optického disku	48
2.3	Rozšíření trénovací množiny	49
2.4	Blokové schéma segmentace optického disku	50
2.5	Aplikované kroky přezpracování obrazu	51
2.6	Detekované optické disky metodou MaskRCNN	53
2.7	Detekované optické disky metodou Unet	54
2.8	Graf úspěšnosti optimalizačních algoritmů	57

2.9	Graf úspěšnosti vnořených architektur	58
2.10	Blokové schéma segmentace cév sítnice	59
2.11	Rozdělení obrazu na podoblasti	60
2.12	Detekované cévy	61
2.13	Graf úspěšnosti různého předzpracování obrazu	63
2.14	Graf úspěšnosti vnořených architektur sestupné části pro cévy	64
2.15	Segmentace disku a cév v sekvencích	65

Seznam tabulek

2.1	Trénovací databáze	49
2.2	Velikost rozšířené trénovací databáze	50
2.3	Segmentační architektury pro optický disk - srovnání	56
2.4	Optimalizační algoritmy - srovnání	56
2.5	Architektury sestupné části pro optický disk - srovnání	57
2.6	Segmentační architektury pro cévy - srovnání	62
2.7	Úspěšnost na základě předzpracování obrazu	62
2.8	Architektury sestupné části pro cévy sítnice - srovnání	63
2.9	Databáze pro zhodnocení sekvencí	66
2.10	Segmentace optického disku - sekvence	66
2.11	Segmentace cév sítnice - sekvence	67

Úvod

Jelikož v současnosti stále narůstá využití medicínských zobrazovacích technik jako je magnetická rezonance, výpočetní tomografie nebo v případě této práce oftalmoskopie, narůstá i potřeba využití počítačů se spolehlivými algoritmy za účelem pomoci při diagnostice či plánování léčby. Mezi hlavní cíle, na které je při počítačovém zpracování medicínských obrazů potřeba myslet, patří automatizace zpracování velkého objemu obrazů se stále stejnou přesností, dosažení co nejvyšší rychlosti zpracování a podpora rychlejší komunikace mezi pacientem a zpracovávajícím expertem.

Následující práce je věnována segmentaci dat z experimentálního oftalmoskopu. Hlavním úkolem při segmentaci medicínských obrazů je rozdělit obraz na základě studovaných anatomických struktur a následně identifikovat oblasti zájmu. K těm mohou patřit léze, tumory i jiné abnormality, které podstupují další měření a zhodnocování.

V případě oftalmoskopických obrazů dochází k segmentaci obnažených cév a optického disku v sítnici, což může být velice nápomocné pro diagnostiku řady onemocnění. Mezi nejčastější z nich můžeme řadit například glaukom, hypertenzi či cukrovku. Často je však důležité pro zajištění včasné diagnostiky provádět kontrolu sítnice pravidelně. Metoda automatické extrakce cév a optického disku a jejich následná analýza by v tomto případě mohla ušetřit velké množství úsilí a času.

Práce je rozdělená na dvě části - teoretický úvod do problematiky a popis implementovaných metod segmentace optického disku a cév sítnice.

V teoretické části je nejprve popsána samotná metoda oftalmoskopie a její využití. Podrobněji je popsán princip fungování konvolučních neuronových sítí s několika vzorovými architekturami. Následuje shrnutí segmentačních metod, které jsou pro segmentaci optického disku a cév dosud využívány.

Praktická část zahrnuje konkrétnější popis navržených a implementovaných metod založených na architekturách Linknet, PSPNet, Unet a MaskRCNN pro segmentaci cév a optického disku ze snímků a sekvencí pořízených experimentálním oftalmoskopem. Nejlepších výsledků je dosaženo s využitím architektury Unet. Kódy s implementovanými metodami se nachází v repozitáři na adrese: https://github.com/AndreaR23/vessel_segmentation.

1 Teoretická část studentské práce

1.1 Oftalmoskopie

Oftalmoskopie je medicínský obor využívající přístroj zvaný oftalmoskop pro vyšetření očního pozadí, kam řadíme sítnici, optický disk a cévy. Vzhledem k obnaženým cévám v oblasti sítnice je oftalmoskopie jednou z nejvýznamnějších metod pro diagnostiku chorob jako například glaukom, cukrovka, či hypertenze. [25]

Cévy a optický disk patří mezi nejvýznamnější struktury sítnice a jsou nápomocné pro diagnostiku onemocnění sítnice. Pro včasnou diagnostiku některých onemocnění je nutné provádět vyšetření a zhodnocování opakovaně v pravidelných intervalech. To může být náročné jak ze stránky finanční, tak z důvodu časových nároků na manuální lékařské zhodnocování velkého množství snímků. Jelikož kvalitní extrakcí cév a optického disku a následnou analýzou je možné identifikovat některé choroby již v jejich zárodku, je automatizace tohoto úkolu jednou z častých a stále zdokonalovaných oblastí výzkumu. [11]

První vyšetření očního pozadí pomocí přímé oftalmoskopie bylo poprvé provedeno již v roce 1851 díky německému vědci Hermannu von Helmholtzovi. Rakouský vědec je dnes považován za vynálezce prvního oftalmoskopu, jehož vynález odstartoval první popisy cév a optického disku z oftalmologického zobrazení. [30]

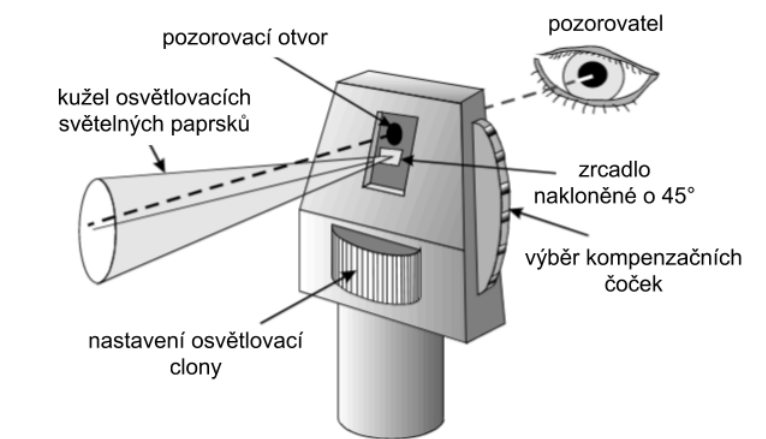
Základem přístroje bylo Helmholtzovo zrcátko, které odráželo paprsky vycházející ze zdroje světla přímo do oka vyšetřované osoby. Oči vyšetřujícího a pacienta byly uspořádány protisměrně, díky čemuž mohl vyšetřující pozorovat osvětlenou plochu pacientovy sítnice, ze které se odrážely paprsky. [25]

Na základě tohoto vynálezu jsou dnes využívány nejčastěji dvě oftalmoskopické metody vyšetření - oftalmoskopie přímá a nepřímá.

1.1.1 Přímá oftalmoskopie

Vzhled moderního přímého oftalmoskopu je zobrazen na obrázku 1.1. Osvětlení oftalmoskopu je zde umístěnou velmi blízko pozorovacím osám, v těsné blízkosti pod pozorovacím otvorem se nachází v průměru asi 3 mm velké, nakloněné zrcadlo.

S uspořádáním takového oftalmoskopu jaký je zobrazený na obrázku je neodmyslitelně spjatý rovněž vynález elektrické žárovky, která umožňuje znázorněné uspořádání zrcadla a pozorovacího otvoru. K dalším součástem oftalmoskopu patří posuvná kolečka, kterými se nastavuje v prvním případě průměr, tvar a barva osvětlovacího paprsku - horizontální kolečko a v druhém případě jsou pomocí kolečka zasouvány kompenzační čočky za pozorovací clonu - vertikální kolečko. [30]



Obr. 1.1: Schéma moderního přímého oftalmoskopu.

Přímý oftalmoskop funguje za pomoci vyzáření divergentního paprsku světla, kterým je osvětlena pacientova sítnice a odražené světlo se vrací zpět k vyšetřujícímu. Je pozorován přímý obraz sítnice, který je závislý na refrakčním stavu jak vyšetřujícího, tak pacienta. Může být proto využíváno zaměření za pomoci zakřivování nebo použití čoček, které jsou vestavěné v samotném oftalmoskopu. Mimo vzdálenost oftalmoskopu od pacienta je do velké míry ovlivňováno zorné pole a zvětšení oftalmoskopu také refrakčním stavem oka. V případě krátkozrakosti - myopie, je zvětšení vyšší, zmenšuje se však zorné pole. Naopak při dalekozrakosti - hypermetropii, není tak velké zvětšení, nicméně se zvětšuje zorné pole. Stejně jako u dalekozrakosti je tomu v případě rozšířených zornic. Velké nepřesnosti mohou být způsobeny především při pohybu pacienta, v tomto ohledu je dosahováno lepších výsledků při použití nepřímé oftalmoskopie. [5]

1.1.2 Nepřímá oftalmoskopie

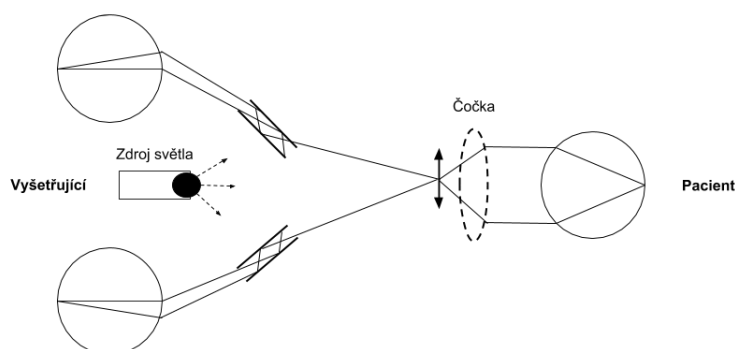
Metoda nepřímé oftalmoskopie byla objevena až přibližně rok po uvedení oftalmoskopie přímé, kdy největší motivací bylo zvětšení zorného pole. Nepřímý oftalmoskop se nasazuje přímo na hlavu vyšetřujícího, disponuje vlastním zdrojem světla a binokulární optikou - viz obrázek 1.2. [25]

Existuje mnoho typů nepřímých oftalmoskopů odlišujících se například na základě jejich stáří či designu, princip fungování je nicméně u všech stejný. Před oko je umístěna konvexní čočka, která je prosvětlována světelným paprskem. Toto světlo je odraženo od sítnice a následně soustředěno za pomoci kondenzační čočky na bod nacházející se mezi vyšetřujícím a objektivem. Tímto způsobem vzniká reálný obraz



Obr. 1.2: Nepřímý oftalmoskop.Vantage Plus

sítnice. Kondenzační čočky v okulárech mají obvykle $+2\text{ D}$, což umožňuje relaxaci při vyklenutí a obraz je proto možné i při presbyopii - stařecké vetchozrakosti, sledovat na délku paže. I zde se obraz liší v závislosti na refrakční chybě, ke které dochází v pacientově oku. [5] Schéma nepřímého oftalmoskopu je na obrázku 1.3.



Obr. 1.3: Schéma nepřímého oftalmoskopu. [8]

K základním výhodám nepřímých oftalmoskopů můžeme zařadit širší zorné pole než u přímých oftalmoskopů či jasnější zdroj světla. K hygienickým výhodám lze zařadit zejména větší vzdálenost mezi vyšetřujícím a pacientem, nicméně na úkor této skutečnosti není získaný obraz tak detailní jako u oftalmologie přímé. [25, 5]

1.1.3 Nemoci detekovatelné oftalmoskopií

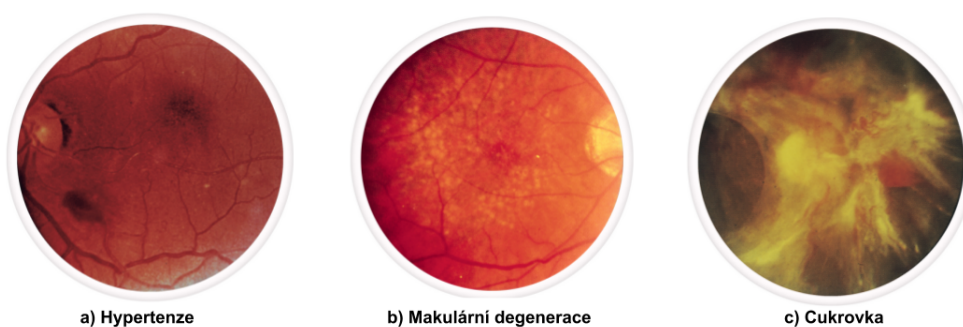
Za pomoci oftalmoskopie lze detekovat řadu onemocnění. Nejčastěji je hodnoceno, zda není poškozený optický nerv, což může vést až ke vzniku zeleného zákalu - **glau-**

komu (obrázek 1.2). Glaukom je druhou nejčastější příčinou slepoty po celém světě. Je způsobený poškozením očních nervů v důsledku degenerace gangliových buněk, které jsou důležité pro správné vidění. Ve většině případů je toto poškození způsobeno zvýšeným nitroočním tlakem zapříčiněným nahromaděním tekutiny proudící skrze oko. [43]



Obr. 1.4: Glaukom. [43]

Dalším detekovatelným onemocněním je **hypertenze** - vysoký krevní tlak, který je dnes jednou z nejčastějších příčin úmrtí. Diagnostikovat lze rovněž **makulární degeneraci**, kdy dochází ke ztrátě vidění ze středu zorného pole, nebo další velice běžné onemocnění, kterým je **cukrovka**. Na obrázku 1.2 můžeme vidět snímky sítnice u zmiňovaných nemocí



Obr. 1.5: Další nemoci detekovatelné oftalmoskopií. [45]

1.2 Konvoluční neuronové sítě - základní prvky

Obraz je nositelem mnoha užitečných informací, pro jejichž získání je nutné jej hlouběji pochopit, což nám umožňují technologie digitálního zpracování obrazu. [34] Některé části obrazu bývají pro pozorovatele často nezajímavé a nedůležité, a proto je za jeden z nejdůležitějších a nejnáročnějších procesů v oblasti digitálního zpracování obrazu považována segmentace obrazu. Jejím hlavním úkolem je zjednodušení obrazu prostřednictvím smysluplného rozdělení na jeho části, které se vyznačují podobnými vlastnostmi a charakteristikami. Tím dojde k odlišení pozadí a oblasti zájmu, ze které můžeme zjistit více užitečných informací. Tento krok bývá často nezbytný pro požadovanou následující analýzu obrazu. [4, 34]

Zejména v posledních letech se začínají stále více rozšiřovat metody založené na hlubokém učení. Za rozmach v oblasti počítačového vidění a zpracování obrazu jsou zodpovědné především konvoluční neuronové sítě, díky kterým byl zaznamenán výrazný pokrok mimo jiné v rozpoznání obrazů jako celků, v detekci jednotlivých objektů nacházejících se v analyzovaných obrazech, či v predikci dílčích částí a bodů v obrazech. Rozvoj konvolučních neuronových sítí je umožněn zejména díky stále vyšší dostupnosti obrázků pro jejich trénování a učení. [27]

V oblasti zpracování biomedicínských obrazů se vyskytuje hned několik problémů, se kterými je potřeba se při implementaci konvolučních neuronových sítí vypořádat. V první řadě může nastat problém právě s dostatečným množstvím trénovacích obrazů, které často nejsou dostupné. V klasifikačních úlohách je rovněž v mnoha případech za účelem vysegmentování požadované struktury nutno klasifikovat postupně každý pixel v obraze, ne pouze obraz jako celek. Pro vyřešení vyskytujících se komplikací bylo postupně navrženo mnoho metod, popsanych v následujícím textu. [22]

Základ konvolučních neuronových sítí je stejný jako u většiny ostatních. Sítě jsou tvořeny pracovními jednotkami - neurony. Každý neuron je definován příslušným matematickým popisem, na základě kterého prochází jednotlivé vstupy neuronovou sítí. Pro konvoluční neuronové sítě jsou charakteristické dva až tři druhy vrstev - vrstvy konvoluční, sdružovací a v některých případech vrstvy plně propojené. Vstupní obraz postupně prochází neuronovou sítí o navržené architektuře vrstvu po vrstvě a je postupně transformován na výslednou hodnotu, která udává jeho zařazení do příslušné třídy. [32]

V architektuře běžné konvoluční sítě dochází k pravidelnému opakování funkčních bloků, které jsou tvořeny různým uspořádáním výše zmiňovaných vrstev. Průchod obrazu definovanou sítí za účelem získání výstupních hodnot je nazýván dopředným průchodem. [32]

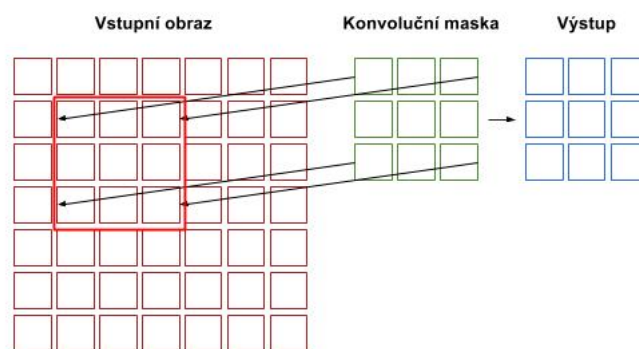
1.2.1 Konvoluční vrstva

Základním stavebním blokem konvoluční neuronové sítě je konvoluční vrstva sloužící k extrahování typických příznaků z obrazu. Vrstva obsahuje mnoho filtrů - masek (obrázek 1.6), kdy každý filtr je přizpůsoben na vyhledání jiných důležitých znaků v obraze, například výrazných hran, barevných skvrn i komplexnějších struktur. Mezi nejčastější rozměry filtrů patří 3×3 nebo 5×5 pixelů (pro obrazy se 3 barev-



Obr. 1.6: Příklad filtrů v konvoluční vrstvě. [18]

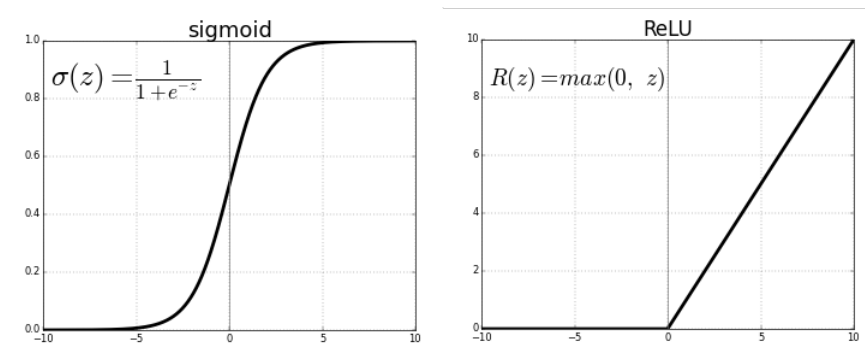
nými kanály po tom $3 \times 3 \times 3$ nebo $5 \times 5 \times 3$ pixelů). Filtry jsou postupně posouvány po vstupním obraze a probíhá matematická operace konvoluce filtru s oblastí obrazu ohraničenou aktuální polohou masky - viz obrázek 1.7. Výstupem je odpoví-



Obr. 1.7: Postupné posouvání konvoluční masky po obraze. [40]

vídající aktivační mapa zaznamenávající výsledek ve všech prostorových pozicích, přičemž všechny hodnoty výstupu v jedné mapě mají sdílené váhy. Každému filtru náleží na výstupu jedna aktivační mapa s příznaky, pro jejichž extrakci je použitý filtr navržen. [40]

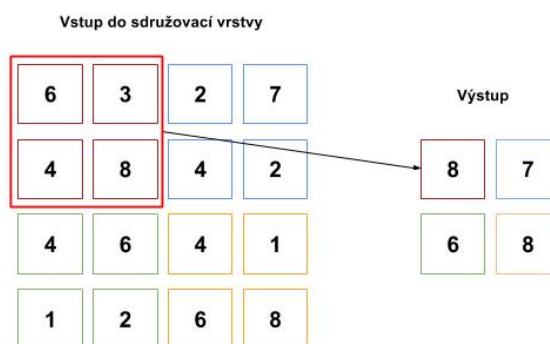
Každá konvoluční vrstva bývá následována aktivační funkcí udávající přenos každého neuronu. V konvolučních sítích bývá touto funkcí nejčastěji funkce ReLU, případně sigmoida. Průběh přenosových funkcí je zobrazen na obrázku 1.8 [22].



Obr. 1.8: Průběh aktivačních funkcí - sigmoida a ReLU. [32]

1.2.2 Sdružovací vrstva

Ve vrstvě sdružovací, která bývá zařazena za vrstvou konvoluční, dochází k potřebnému podvzorkování obrazu pro snížení rozměrů vstupujících příznakových map vzniklých v konvoluční vrstvě. [32] Díky snížení rozměrů dochází k úbytku parametrů a redukci výpočetní náročnosti, s čímž souvisí i menší pravděpodobnost přetrénování sítě. Nejčastější operací používanou při podvzorkování vstupu je výběr maxima z okna o stanoveném rozměru, nejčastěji 2×2 pixely - obrázek 1.9. Další variantou je namísto maximální hodnoty vložit na výstup průměrnou hodnotu z okna o stanovené. [40] velikosti.

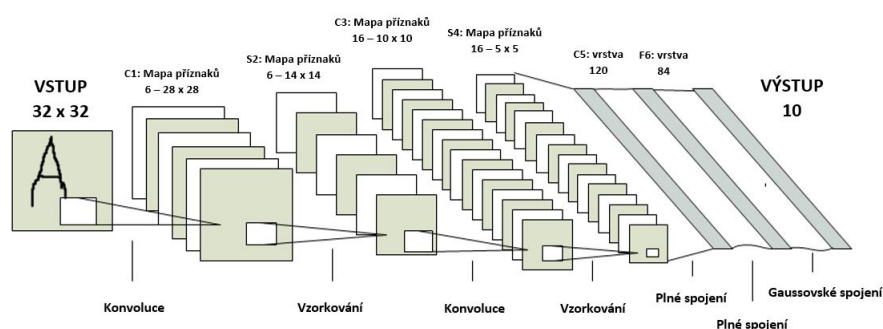


Obr. 1.9: Příklad podvzorkování obrazu - výběr maximální hodnoty. [40]

1.2.3 Plně propojená vrstva

Výstup z poslední sdružovací vrstvy je nejprve převeden do jednorozměrného pole čísel, které postupuje do jedné či více vrstev plně propojených. Charakteristickou vlastností plně propojených vrstev je, že každý vstup do této vrstvy je propojený s každým výstupem s definovanou váhou. Plně propojené vrstvy bývají umístěny na závěr architektury jako klasifikátory zařazující výstup z poslední sdružovací vrstvy do konkrétní kategorie. Množství neuronů v poslední plně propojené vrstvě tedy odpovídá počtu kategorií, do kterého jsou výstupy rozřazovány.

Na obrázku 1.10 se nachází obecné schéma architektury konvoluční neuronové sítě. Konkrétní architektura je vždy volena s ohledem na požadované využití sítě.



Obr. 1.10: Schéma základní konvoluční neuronové sítě. [40]

V dalších kapitolách budou popsány architektury navržené za účelem segmentace obrazu, které jsou často navrženy bez bloku plně propojených vrstev.

1.3 Konvoluční neuronové sítě - princip učení

Jak již bylo zmíněno prvním a zároveň jedním z hlavních úkolů konvoluční neuronové sítě je zařadit konkrétní obraz nebo jeho část do některé z definovaných kategorií - fungují tedy jako klasifikátory. Velmi důležitým je proto proces učení sítě, při kterém se snažíme optimalizovat hodnoty jednotlivých vah na základě předkládání obrazů, u kterých známe požadovaný výstup sítě. [17]

Běžně používanou metodou pro učení neuronové sítě je zpětné šíření stanovené chyby s využitím výpočtu chybové funkce, kterou se snažíme minimalizovat za pomoci některé z níže zmíněných optimalizačních algoritmů.

1.3.1 Chybové funkce

Při učení konvoluční neuronové sítě je počítána tzv. chybová funkce. Chybová funkce vyjadřuje odchylku hodnoty predikované sítí od požadované hodnoty výstupu pro jeden konkrétní obraz. Typ chybové funkce je jedním z volených parametrů, kterým lze ovlivnit výsledek učení sítě.

Na základě vypočtených hodnot chybové funkce dochází k úpravě vah v jednotlivých vrstvách. Korekce hodnot vah probíhá postupně od poslední vrstvy, kde známe reálnou i predikovanou hodnotu výskytu, směrem k vrstvě první. Postup, při kterém dochází k šíření chyby z poslední, klasifikační vrstvy směrem k první vrstvě, se nazývá **algoritmus zpětného šíření chyby**. [17]

Váhy v jednotlivých vrstvách jsou upravovány na základě tzv. δ – pravidla, kdy vztah pro korekci je definován jako:

$$\vec{w}(t+1) = \vec{w}(t) - \mu_t \nabla \varepsilon. \quad (1.1)$$

Zde $\vec{w}(t+1)$ značí vektor vah v následující iteraci. Ten je počítán na základě vektoru vah v současné iteraci $\vec{w}(t)$ a zjištěné vypočtené chybové funkce ε váhované nastaveným parametrem učení μ pro danou iteraci.

Chybová funkce udává odchylku pro jeden konkrétní vstup sítě. Při optimalizaci parametrů bývá proto využívána tzv. hodnotící funkce, která průměruje pro daný průchod sítí chybové funkce vypočítané na základě všech předložených obrazů.

$$f(w) = \frac{1}{N} \sum_{i=1}^N \varepsilon, \quad (1.2)$$

kde N je počet průměrovaných vstupů pro výpočet zpětného šíření chyby - tzv. batch. [44]

Střední kvadratická odchylka

Jedním z nejběžněji využívaných chybových funkcí je výpočet střední kvadratické odchylky:

$$\varepsilon = \vec{e}^T \vec{e} = \sum_{j=1}^M (p_j - y_j)^2, \quad (1.3)$$

kde p jsou požadované hodnoty výstupu, které odpovídají M vstupním obrazům. Hodnoty y jsou reálné hodnoty, získané po průchodu obrazu sítí. [32, 17]

Tato chybová funkce je nejčastěji využívána při lineární regresi, kdy dochází k proložení prostoru křivkou, na základě které lze stanovit závislost proměnných. Ke snižování střední kvadratické odchylky dochází minimalizací součtu vzdáleností všech bodů od této regresní křivky. [48]

Vzájemná entropie

V oblastech klasifikace bývá často využívána chybová funkce zvaná vzájemná entropie. Vzájemná entropie stanovuje rozdíl dvou pravděpodobnostních rozložení. Vysoká hodnota této entropie značí velký rozdíl porovnávaných rozložení. Při snižování hodnoty se podobnost mezi rozloženími zvyšuje, vzájemná entropie tedy měří výkonnost klasifikačního modelu. Jsou porovnávány predikované pravděpodobnosti příslušnosti k jednotlivým třídám. V případě vysoké pravděpodobnosti příslušnosti do nesprávné třídy dochází k penalizaci a hodnota vzájemné entropie se zvyšuje. [48]

Vzájemná entropie může být počítána pro binární klasifikaci, kdy dochází k rozdělení do dvou tříd, na základě vztahu:

$$\varepsilon = -(y \log(p) + (1 - y) \log(1 - p)), \quad (1.4)$$

kde p je predikovaná hodnota na výstupu a y je hodnota požadovaná. Funkce penalizuje odchylku každého pixelu od hodnoty 1. [46, 48]

V případě klasifikace mezi více tříd je počítána a následně sečtena odchylka pro všechny třídy:

$$\varepsilon = - \sum_{j=1}^M y_{i,j} \log(p_{i,j}). \quad (1.5)$$

Zde M je počet tříd, do kterého je prováděna klasifikace, hodnota $p_{i,j}$ je predikovaná hodnota pravděpodobné příslušnosti každého pixelu k třídě j porovnávaná s požadovanou odpovídající hodnotou $y_{i,j}$. [46]

IoU - průnik sjednocení

Pro segmentaci obrazu, která je předmětem této práce, bývá v mnoha studiích využívána chybová funkce, která započítává hodnotu průniku sjednocení. Původní vztah

pro její vypočtení je dán: [7]

$$\varepsilon = \frac{|y \cap p|}{|y \cup p|}. \quad (1.6)$$

Výsledek je dán průměrem přes predikované hodnoty pro všechny pixely. [7]

Pro správný výpočet IoU je však nutné porovnávat pouze hodnoty 0 a 1. To je splněno pro hodnoty, y , nikoli však pro predikované pravděpodobnosti p . Zde je nutné využít výstup po aplikaci aktivační funkce jako například sigmoidy. Pro korektní výpočet však lze využít aproximaci, kdy dostaneme odvozený vztah: [7]

$$\varepsilon' = \frac{|y \times p|}{|y + p - (y \times p)|} = \frac{I}{U}. \quad (1.7)$$

Definovaným vztahem získáme v čitateli touto aproximací pravděpodobnost hodnoty p_x v případě, že se odpovídající hodnota y_x rovná 1. V okamžiku, kdy je odpovídající hodnota y_x rovna 0, je vlivem násobení roven 0 i celý čítec. Hodnota IoU je tedy nejvyšší v okamžiku, kdy p_x spolu s y_x jsou rovny 1. Jmenovatel slouží pouze pro korekci, kdy dochází ke snížení vlivu dvojnásobného výpočtu průniku v dané oblasti. [7]

Diceův koeficient

Na podobném principu jako IoU je založen i výpočet chybové funkce pomocí Diceova koeficientu podobnosti. Ten rovněž hodnotí prostorový překryv detekované vysegmentované oblasti a oblasti z vzorové segmentační mapy, podle které je chybová funkce počítána: [14]

$$\varepsilon = 2 \frac{|y \cap p|}{|y + p|}. \quad (1.8)$$

V některých případech může být provedeno zlogaritmování výše uvedeného vztahu, jelikož výsledná hodnota Diceova koeficientu se nachází v intervalu $\langle 0, 1 \rangle$, přičemž se často blíží právě hodnotě 1. Výsledný vztah je znázorněn: [14]

$$\varepsilon' = \log\left(\frac{\varepsilon}{1 - \varepsilon}\right). \quad (1.9)$$

1.3.2 Optimalizační algoritmy

V konvolučních neuronových sítích se za účelem snižování odchylky od požadovaného výstupu využívá několik druhů optimalizačních algoritmů. Cílem těchto algoritmů je dosáhnout postupným snižováním odchylky globálního minima a tedy nejvíce přiblížit predikovaný výstup výstupu reálnému.

SGD - pokles ve směru gradientu

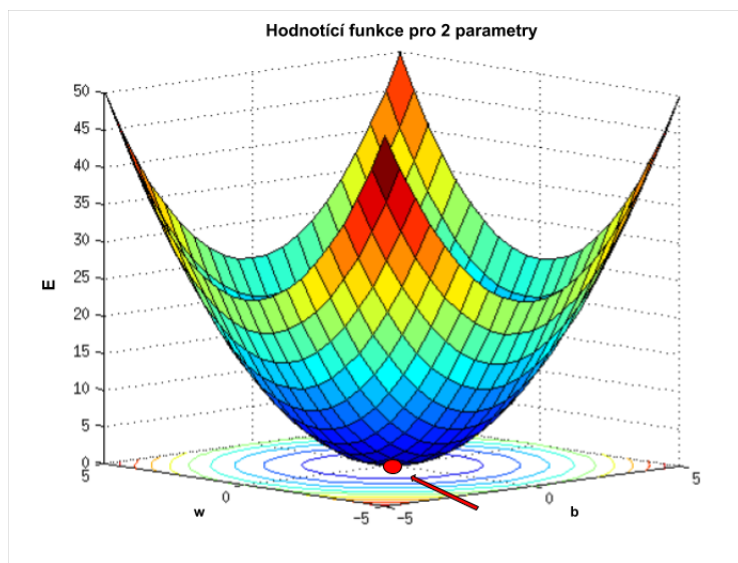
SGD je dnes vzhledem ke své jednoduchosti a efektivitě jedním z nejběžněji využívaných optimalizačních algoritmů.

Základem je výpočet derivace hodnotící funkce podle optimalizovaného parametru, který poskytuje informaci o směru nárůstu funkce. Následně je postupováno proti směru tohoto nalezeného nárůstu - ve směru nejstrmějšího sestupu. Důležitá je definice délky učicího kroku při postupování. Optimalizační metoda klesání ve směru gradientu pak vychází z předpokladu, že postupováním podle nejstrmějšího sestupu je postupně dosaženo hledaného minima hodnotící funkce. [44]

Jelikož v konvoluční neuronové síti je možné mít na výstupu více parametrů, které je nutné zpětným šířením chyby upravit, bývá pomocí klesání ve směru poklesu gradientu optimalizována funkce o více proměnných. Funkce zhodnocující výsledný průchod bývá nejčastěji ovlivňována již zmiňovanými vahami v jednotlivých vrstvách, druhým optimalizovaným parametrem je ve většině případů hodnota zkreslení - tzv. bias.

$$f(w, b) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M (d_{ij} - (my_{ij} + b))^2. \quad (1.10)$$

Zde b značí zmiňovanou hodnotu zkreslení - bias.



Obr. 1.11: Závislost chybové funkce na optimalizovaných parametrech - hledané minimum hodnotící funkce. [47]

Pro nalezení gradientu funkce o více proměnných je nutné využít výpočet parciálních derivací, kdy hodnotící funkci derivujeme podle upravovaných parametrů: [44]

$$f'(w, b) = \begin{bmatrix} \frac{df}{dw} \\ \frac{df}{db} \end{bmatrix} = \begin{bmatrix} \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M -2y_{ij}(d_{ij} - (wy_{ij} + b)) \\ \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M -2(d_{ij} - (wy_{ij} + b)) \end{bmatrix}. \quad (1.11)$$

Směr nejstrmějšího sestupu je tedy hledán v závislosti na dvou proměnných. Závislost hodnotící funkce na jednotlivých parametrech je znázorněna na obrázku 1.11.

Adam - přizpůsoblivý odhad momentů

Existuje řada algoritmů fungujících na podobných principech jako je metoda SGD. Z kombinace jejich výhod je odvozená další pokročilá a běžně užívaná optimalizační metoda označovaná zkratkou Adam. Algoritmus Adam je přizpůsoblivý algoritmus, který počítá jednotlivé rychlosti učení různých parametrů. Jejich výpočet je založen na stanovení momentů gradientů, kdy momentem proměnné rozumíme očekávanou hodnotu její energie umocněnou na n . [37] Momenty pro váhy jednotlivých proměnných jsou počítány na základě vztahu:

$$m_w^{t+1} = \beta_1 m_w^t + (1 - \beta_1) \varepsilon \quad (1.12)$$

pro výpočet průměru jako prvního momentu a na základě vztahu:

$$v_w^{t+1} = \beta_2 v_w^t + (1 - \beta_2) \varepsilon^2 \quad (1.13)$$

pro výpočet rozptylu jako druhého momentu. V uvedených vztazích hodnota t udává pořadí trénovací iterace. Konstanty β_1 a β_2 slouží jako faktory zapomínání a ovlivňují váhu těchto hodnot z předchozí iterace. Parametr ε udává hodnotu gradientu chybové funkce. [15]

Důležitým krokem pro správné upravení koeficientů vah je upravit u definovaných vztahů korekci zkreslení. Tím jsou vztahy pro výpočet momentů upraveny do následující podoby:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad (1.14)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}. \quad (1.15)$$

Hodnotu vah pro daný parametr už lze následně vypočítat dosazením do vztahu:

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}, \quad (1.16)$$

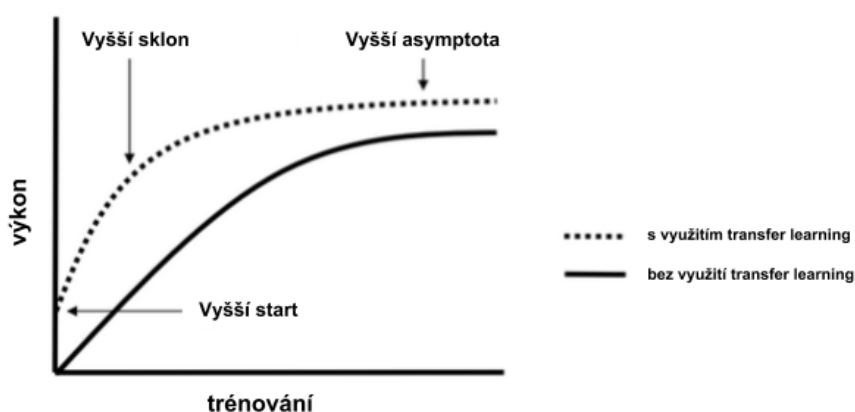
kde η je velikost kroku, často závislá na aktuální iteraci a konstanta ϵ je číslo o zanedbatelné hodnotě pro ochranu před dělením 0. [37, 15]

1.3.3 Transfer learning

Jen ve velmi malém množství studií je využíváno učení konvoluční neuronové sítě od počátku pouze na vlastních trénovacích datech. Běžně využívané jsou například předtrénované sítě z webu <http://www.image-net.org/>, na kterém lze nalézt přes 1 milion obrázků rozdělených do 1000 kategorií. [41]

Případ, kdy je trénovaný model naučen na dostupné obsáhlé databázi, je označován jako tzv. transfer learning. Ze sítě natrénované na této databázi přeneseme získané funkce do cílové sítě, která už bude pro učení využívat cílovou množinu dat pro požadovanou úlohu. [38]

Ačkoli skutečnost, zda bude v praxi transfer learning přínosem lze zhodnotit až po provedení experimentu, ve většině případů tomu tak je. Tento očekávaný přínos může být různého charakteru jako je znázorněno na obrázku 1.12. V prvním případě



Obr. 1.12: Výhody, které přináší transfer learning. [38]

dojde k vyššímu startu, model má tedy od začátku lepší výkon. Dalším možným přínosem je vyšší sklon, který je způsoben rychlejším učením modelu. Poslední výhoda může spočívat ve vyšší asymptotě, která je zapříčiněna lepší konvergencí při trénování modelu. [38]

V praxi jsou využívány tři hlavní scénáře pro transfer learning.

Konvoluční síť jako pevný extraktor příznaků

Nejjednodušším případem je situace, kdy se převezme předtrénovaná konvoluční síť. Dále je odstraněna poslední plně propojená vrstva, která zprostředkovává rozdělení vstupů do původních kategorií. Přidá se vrstva s požadovaným množstvím výstupních kategorií podle zadaného úkolu a předtrénovaná síť je využívána pro predikci obrazů z nové databáze. [41]

Ladění konvoluční neuronové sítě

V tomto případě už dochází kromě nahrazení plně propojené vrstvy zároveň k doučení sítě na vlastních datech. Posíláním obrazů z vlastní databáze jsou algoritmem zpětného šíření chyby dále upravovány přednastavené koeficienty vah. Z důvodu zabránění přetrénování jsou často některé vrstvy původní vrstvy tzv. zafixovány, jejich váhy se již dalším učením nemění a upravují se pouze váhy ve vyšších vrstvách. Vychází se z předpokladu, že první zafixované vrstvy extrahují obecnější příznaky, ke kterým patří například hrany, upravované vyšší vrstvy potom vypovídají o konkrétnějších příznacích vlastních dat. [41]

Předtrénované modely

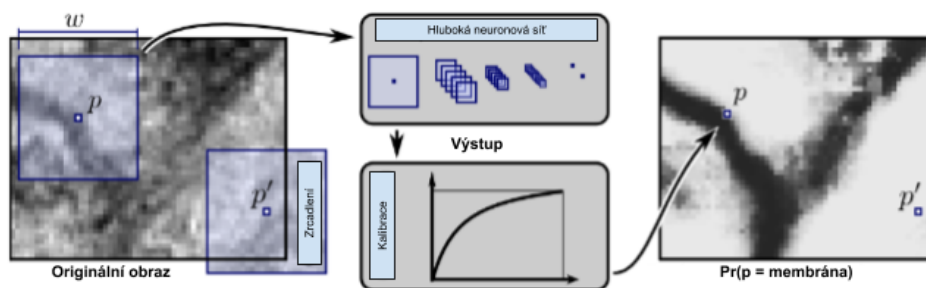
I s využitím GPU je časová náročnost předtrénování modelu na velkém množství obrazů i v řádech týdnů. Mnoho lidí proto zveřejňuje jejich výsledky a průběh trénování, což usnadní mnoho času a práce při výběru modelu pro ladění konvoluční sítě. [41]

1.4 Konvoluční neuronové sítě - segmentace obrazu

1.4.1 První konvoluční neuronové sítě a segmentace

Jedním z prvních přístupů, který využívá dříve zavedené konvoluční neuronové sítě umožňující hlubší analýzu obrazů za účelem jejich segmentování, je implementace hlubokých neuronových sítí jako klasifikátorů jednotlivých pixelů. Přístup je popsán ve studii [9], kterou provedl v roce 2012 Ciresan et al.. Studie vychází z výpočtu pravděpodobností, se kterými náleží jednotlivé pixely do výsledných segmentovaných oblastí. Pro výpočet pravděpodobností je počítáno s plovoucím oknem o definované délce v původním obraze, jehož střed se nachází v počítaném pixelu. Délka strany okna je obvykle stanovena na lichý počet pixelů z důvodu symetrie. V okamžiku, kdy je vyšetřovaný pixel na okraji obrazu a je okrajem okna o definované délce vystoupeno mimo hranici obrazu, dochází k zrcadlení podle hranice obrazu. Výsledný obraz vzniká následným vyhlazením a zavedením prahování.

Oproti dřívějším přístupům nastává rovněž změna ve vrstvách zodpovědných za podvzorkování obrazu. V konvolučních neuronových sítích, ze kterých studie vychází bylo podvzorkování zprostředkováno průměrováním pixelů v okně o definované délce. Autoři přístupu nicméně zavádí snížení rozměrů výběrem maxima z takto definovaného okna.



Obr. 1.13: Popsaný algoritmus hluboké neuronové sítě. [9]

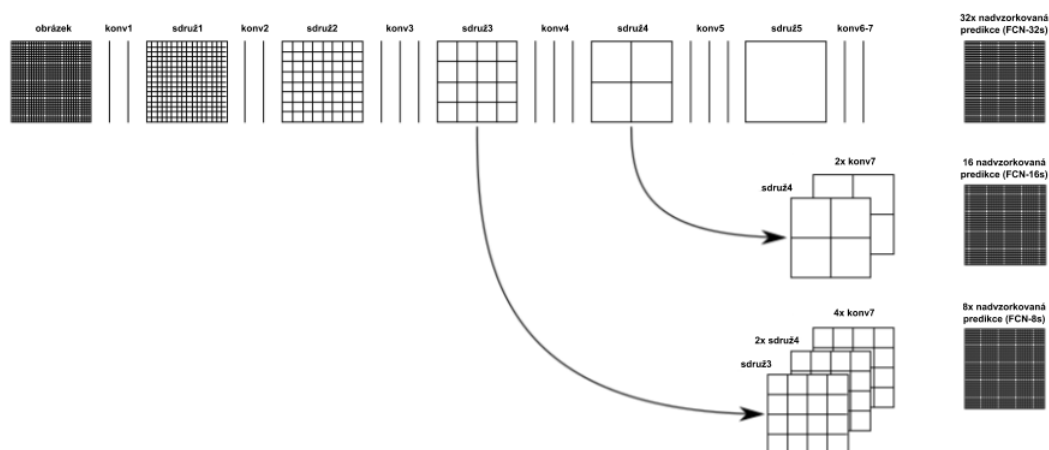
Ke klasifikaci pixelů dochází na úrovni plně propojených vrstev, kdy po průchodu obrazu konvolučními a sdružovacími vrstvami je obraz převeden na jednorozměrný vektor a jednotlivé pixely jsou následně na základě zmiňovaných výpočtů pravděpodobností rozřazovány do tříd. Průběh popsaného algoritmu spolu s ukázkou zrcadlení je na obrázku 1.13. [9]

Mezi hlavní omezení této metody můžeme zařadit její výpočetní náročnost, způsobenou postupným posouváním plovoucího okna po obraze s vysokým překryvem.

V případě zvětšení tohoto okna a urychlení algoritmu dochází ke zhoršení přesnosti lokalizace klasifikovaných pixelů. Naopak zmenšením okna je mimo vyšší čas průchodu sítí pro rozhodování v každém okně využit pro klasifikaci pouze malý kontext. [22]

Shelhamer et al. ve své studii z roku 2015 navazují na dosavadní metody a popisují síť plně konvoluční, kdy navržená architektura neobsahuje plně propojené vrstvy na svém konci. Hlavním rozdílem je skutečnost, že běžná hluboká konvoluční neuronová síť má na výstupu pouze nelineární funkci, rozměry výstupu u plně konvolučních sítí však v běžné praxi odpovídají rozměrům jejich vstupu.

Velkou roli tedy hraje převedení klasifikační funkce sítě do formy, kdy výstupem je hrubá segmentovaná mapa obrazu. Za tímto účelem je síť tvořena dvěma hlavními bloky - blokem pro podvzorkování, kdy dochází k extrahování informací důležitých pro konkrétní obraz a blokem pro zpětné převzorkování na původní velikost, kdy dochází k navrácení prostorové informace a tvorbě segmentovaného obrazu. Nejjednodušší metoda opětovného zvyšování rozměrů je založena interpolací mezi okolními hodnotami pixelů. Plně konvoluční síť zachycuje obrázek 1.14. [27]



Obr. 1.14: Popsaný algoritmus plně konvoluční neuronové sítě. [27]

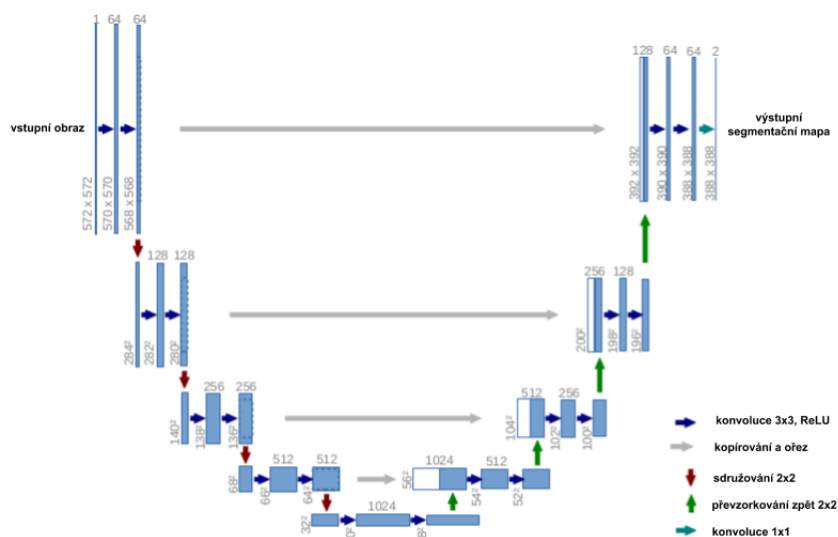
Existuje celá řada konkrétních architektur založených na plně konvolučních neuronových sítích, které jsou využívány pro segmentaci zájmových oblastí z obrazu. Některé z nich jsou podrobněji popsány v následujících kapitolách.

1.4.2 Unet

Architektura Unet byla navržena konkrétně pro biomedicínské využití. Jejím cílem je mimo zefektivnění procesu segmentace také její zpřesnění i pro situace, kdy není k dispozici příliš vysoké množství trénovacích dat. Vychází ze základního principu plně propojené konvoluční sítě, která je upravena a rozšířena. [22]

Princip

Jelikož základem architektury je plně konvoluční síť, vychází architektura Unet z části, kdy dochází k podvzorkování obrazu, kterou následuje část v níž je obraz opětovně skládán pro zajištění správného prostorového namapování segmentovaného obrazu na obraz původní. Samotná síť je složena jako v případě plně konvolučních sítí z vrstev konvolučních a z vrstev sdružovacích, které zprostředkovávají podvzorkování sítě. Pro extrakci charakteristických znaků v obraze v sestupné části je využita



Obr. 1.15: Architektura Unet. [22]

architektura Resnet, navržena pro klasifikaci obrazu, která má výhodu především ve využití skrytých spojení v případě dostatečného natrénování, čímž je zabráněno přetrénování konvoluční sítě. Rozdílem oproti předcházejícím strukturám je, že následné zvýšení rozměrů obrazu na původní je prováděno postupně, po stejných krocích, po kterých probíhalo podvzorkování. Další výhodou je skutečnost, že při tomto procesu nedochází k pouhé interpolaci mezi okolními pixely, ale následné navyšování rozměrů je prováděno opět pomocí definovaných filtrů a pomocí skrytých

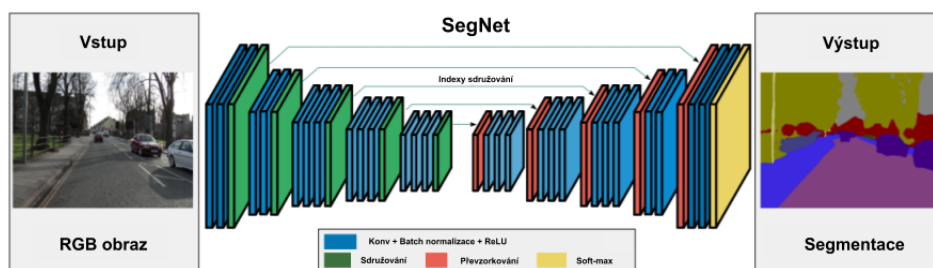
spojení mezi odpovídajícími vrstvami v části pro podvzorkování a v části pro převzorkování na vyšší rozměry pro zajištění kvalitnějšího výstupu především z pohledu přesnosti lokalizace segmentovaných oblastí. [22] Architektura Unet je znázorněna na obrázku 1.15.

1.4.3 SegNet

Další variantou plně konvoluční sítě je architektura nazvaná SegNet. Stejně jako v případě architektury Unet je síť složena ze sestupné části, kdy jsou rozměry sítě snižovány a z části vzestupné, kdy jsou rozměry opět navyšovány. V případě architektury SegNet je sestupná část nazvána kodér a část vzestupná dekodér. Hlavní rozdíl spočívá v procesu převzorkování při navyšování rozměrů v dekodéru. Návrh architektury SegNet je popsán ve studii [6] od Badrinarayanan et al., 2017.

Princip

Část kódující v architektuře SegNet je tvořena 13 konvolučními bloky, což odpovídá běžně využívané architektuře nazvané VGG. Část dekódující se tvořena vrstvami v přesně opačném pořadí. Jako v předchozím případě se jedná o síť plně konvoluční, neobsahuje tedy žádné plně propojené vrstvy na závěr a umožňuje vstup i obrazům s vyšším rozlišením. [6]



Obr. 1.16: Architektura SegNet. [6]

Důležitým znakem architektury SegNet je uložení informací o jednotlivých příznacích v příznakových mapách zařazených za jednotlivými konvolučními vrstvami kodéru pro jejich následné využití při opětovném převzorkování v dekodéru. Pro ušetření místa v paměti je však ukládání zprostředkováno 2 bity pro každé podvzorkované okno velikosti 2×2 pixely namísto paměťově náročného ukládání všech příznakových map. [6]

Dekodér provádí převzorkování na základě uložených indexů příznakové mapy v odpovídající vrstvě kódující části. Po převzorkování vzniká řídká příznaková mapa,

na kterou je opět aplikována řada konvolučních filtrů. Výsledkem je hustá příznaková mapa postupující po normalizaci do další vrstvy dekodéru dokud nevzniká výsledná segmentační mapa obrazu. Popsaný princip převzorkování na základě indexů je znázorněn na obrázku 1.16. [6]

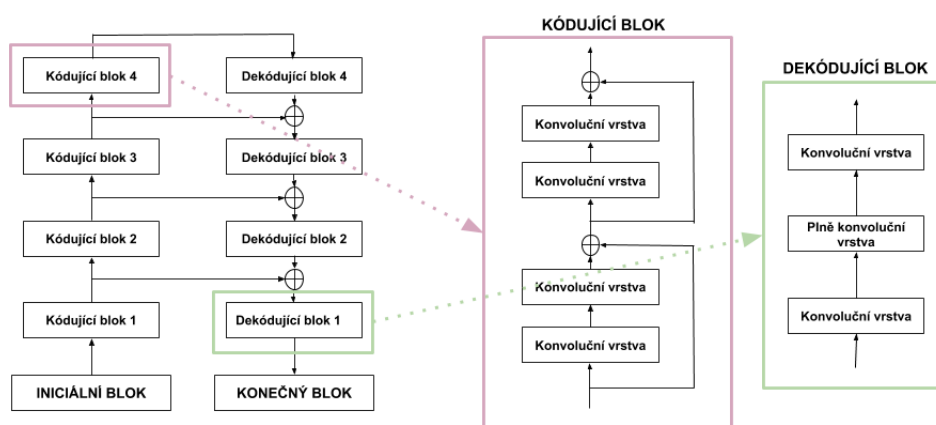
1.4.4 LinkNet

I další využívaná architektura, skládající se rovněž z kódující a dekódující části, se snaží především optimalizovat poměr maximálního zachování informace o prostorovém rozvržení pixelů během průchodu původního obrazu sdružovacími vrstvami vůči co nejnižší výpočetní náročnosti při celém průchodu obrazu sítí. Stejně jako předchozí architektury ke svému cíli využívá skrytých spojení mezi kódující a dekódující částí. [10]

Princip

Vrstvy v kódující části jsou opět nejčastěji uskupeny podle architektury zvané ResNet. V dekódující části dochází opět k převzorkování na původní rozměr obrazu. [10]

Hlavním rozdílem konvoluční sítě s architekturou LinkNet je propojení mezi kódující částí a odpovídajícími místy v části dekódující. Spojení, přítomná i v některých z předcházejících metod (např. Unet či SegNet), však v tomto případě posílají do dekódující části celou získanou příznakovou mapu, pro následné přesnější převzorkování na původní rozměr. Propojení mezi kódující a dekódující částí sítě LinkNet je znázorněno na obrázku 1.17. [10]



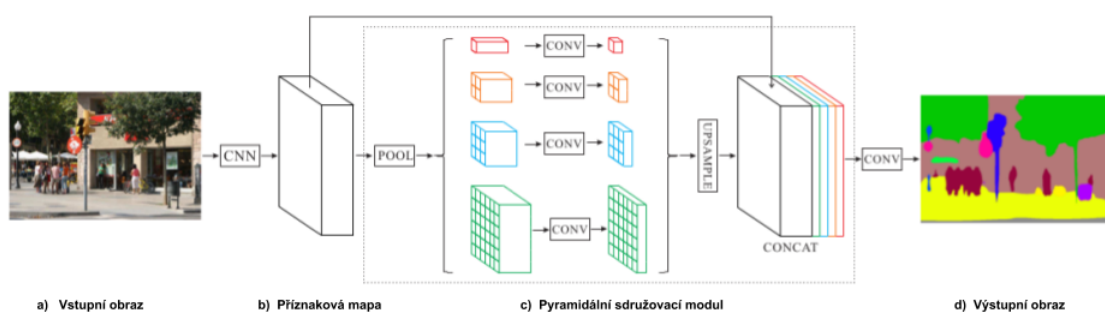
Obr. 1.17: Propojení kódující a dekódující části architektury LinkNet. [10]

1.4.5 PSPNet

S poněkud odlišným přístupem přicházejí autoři ve studii *Pyramid Scene Parsing Network* z roku 2017, kde navrhuji zařazení tzv. pyramidového sdružovacího modulu do architektury konvoluční sítě. Ten zachovává další kontextuální informace, čímž zajišťuje komplexnější porozumění obrazu a tudíž i jeho kvalitnější sémantickou segmentaci. Jako ukázka je ve studii uvedena například situace, kdy automobil lze rozpoznat nejen na základě vlastních příznaků, ale lze rovněž zhodnotit jeho pozici v obraze na základě okolních informací. V okolí automobilu se velice často může vyskytovat silnice, naopak vodní plocha jej s největší pravděpodobností ze všech stran obklopovat nebude. [39]

Princip

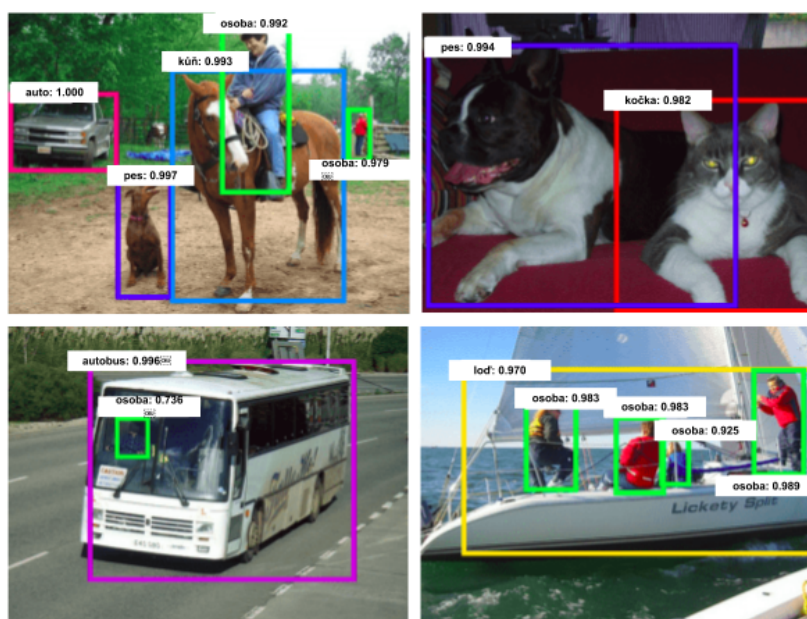
Základní přínos navržené architektury spočívá v zařazení pyramidálního sdružovacího modulu. Příznaky ze vstupního obrazu jsou nejprve extrahovány zařazením některé z dosud dostupných architektur konvolučních neuronových sítí. Autoři studie navrhuji využití architektury ResNet. Výstupní mapy příznaků jsou posílány na výstup sítě a zároveň vstupují do pyramidálního sdružovacího modulu. Zde je použito několikaúrovňové sdružování podle velikosti kontextu, který má být na dané úrovni zachován. Pro zachování vah je zařazena konvoluce s velikostí masky 1×1 pixel. Následuje převzorkování na původní velikost obrazu a spojení s původní příznakovou mapou, která vznikla po průchodu první částí algoritmu. Pro získání výsledné segmentační mapy je zařazena poslední konvoluční vrstva. [39] Celá popsaná architektura je vykreslena na obrázku 1.18.



Obr. 1.18: Architektura PSPNet. [39]

1.4.6 Mask RCNN

Konvoluční síť pro segmentaci obrazu spolu s konvoluční sítí pro detekci objektu ve své studii kombinují He et al. v roce 2017. Jejich přístup vychází z konvoluční sítě s názvem *Faster RCNN*, která slouží k detekci jednotlivých objektů v obraze. Ve vstupním obrazu je nalezeno velké množství návrhů na zařazení částí obrazu k hledaným objektům. Každý takový návrh ohraničuje obdélníkový rámeček a je ohodnocený skórem, které udává příslušnost ke každému z hledaných objektů či k pozadí. Na základě dosaženého skóre jsou následně vykresleny rámečky ohraničující hledané objekty v obraze, viz obrázek 1.19. [21]

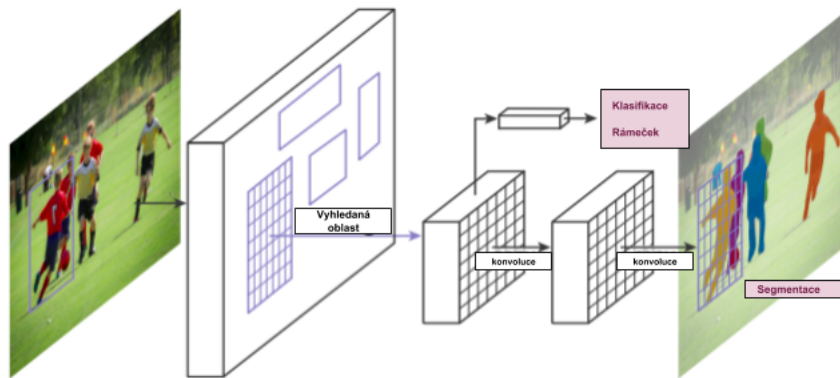


Obr. 1.19: Objekty detekované pomocí metody Faster RCNN. [21]

Metoda Mask RCNN navíc v detekovaných rámečcích segmentuje samotné označené objekty.

Princip

Metoda Mask RCNN vychází ze zmiňované Faster (rychlejší) RCNN, která je rozdělena do dvou základních fází. V první jsou v obraze vyhledány obdélníkové rámečky, v druhé fázi dochází v těchto oblastech k extrakci charakteristických příznaků, za pomoci kterých je zprostředkována klasifikace do hledaných tříd a vyhledává se výsledné ohrazení objektu rámečkem. Metoda se označuje za rychlejší, jelikož příznaky v těchto stádiích jsou sdílené, což výrazně snižuje výpočetní náročnost.



Obr. 1.20: Architektura Mask RCNN.HE

Samotná Mask RCNN přidává ke zmiňovaným dvěma fázím ještě fázi třetí, která probíhá současně s klasifikací objektu ve vyhledané oblasti. Dochází při ní k průchodu nalezené oblasti zájmu ohraničené obdélníkem některou z plně konvolučních sítí. I v tomto případě se nejčastěji jedná o architekturu Resnet. Na výstupu celé metody je tedy ohraničená oblast objektu, který je klasifikován do vyhledávané třídy, a segmentační mapa hledaného objektu.

Při trénování jsou v každé kandidátní oblasti zájmu počítány chybové funkce jako součet chybových funkcí jednotlivých úkonů probíhajících v síti, tedy:

$$L = L_{klasifikace} + L_{rámeček} + L_{maska}. \quad (1.17)$$

Postup průchodu obrazu sítí je znázorněný na obrázku 1.20.

1.5 Metody segmentace optického disku

Jedna z největších motivací pro segmentaci optického disku se nazývá glaukom. Jedná se o chronické onemocnění oka, kdy dochází k postupnému poškozování optického nervu. Toto oční onemocnění může vyústit až ve slepotu, přičemž v současné době je glaukom druhou nejčastější příčinou slepoty na světě. Jeden ze způsobů diagnostiky glaukomu vychází z hodnocení optického disku v sítnici. [3]

Optický disk je hlavou optického nervu, který do oka přichází z mozku. Uvnitř oblasti optického disku se nachází prohloubený útvar, nazývaný oční pohárek. Poměr velikosti pohárku a optického disku je jedním ze zásadních kritérií při diagnostice glaukomu. [23]

Díky nutnosti pravidelného opakování vyšetření pro včasnou diagnostiku se do popředí zájmu stále více dostává snaha o automatickou detekci a segmentaci optického disku. [3] Optický disk je ve většině případů místo s vyšší intenzitou jasu, což jeho detekci výrazně usnadňuje. Na druhou stranu je vždy do určité míry zakryt procházejícími cévami, proto je častým krokem předzpracování v podobě odstranění těchto cév. [23]

Existuje řada přístupů využívaných pro segmentaci optického disku, od jednodušších, založených na prahování obrazu, přes metody založené na detekci hran či oblastí až po složitější metody využívající hluboké neuronové sítě. Základem všech je lokalizace optického disku a jeho segmentace, kdy úspěšnost metody je zhodnocována na ručně anotované databázi s binárními maskami označujícími reálnou pozici hledaných optických disků. [1]

1.5.1 Fuzzy konvergence krevních cév

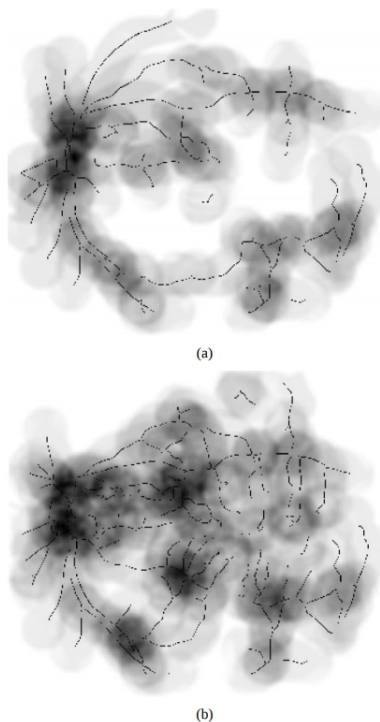
Jeden z prvních přístupů k detekci zračkového nervu, který vystupuje z optického disku, ve své práci *Locating the Optic Nerve in Retinal Image Using the Fuzzy Convergence of the Blood Vessels* popisují Hoover et Goldbaum v roce 2003.

Princip metody

Jak plyne z názvu studie, metoda je založena na tzv. fuzzy konvergenci, kdy optický disk je lokalizován jako oblast, ve které dochází k největšímu soustředění krevních cév. Oblasti krevních cév jsou modelovány za pomoci křivek a hledaná konvergence lze následně převést na problém vyhledání nejhustšího průniku těchto cév. Pro modelování křivek představujících krevní cévy je využita Houghova transformace a metoda nejmenších čtverců. [12]

Samotná fuzzy konvergence je metoda hlasovací, kdy každá oblast podobná hledané linii je modelována jako fuzzy oblast, v níž dochází k ohodnocení jednotlivých

pixelů. Výsledným hlasováním dochází ke zvolení oblasti označené jako oblast největší konvergence cév. Ohodnocené kandidátní oblasti fuzzy hodnotami jsou znázorněné na obrázku 1.21. [12]



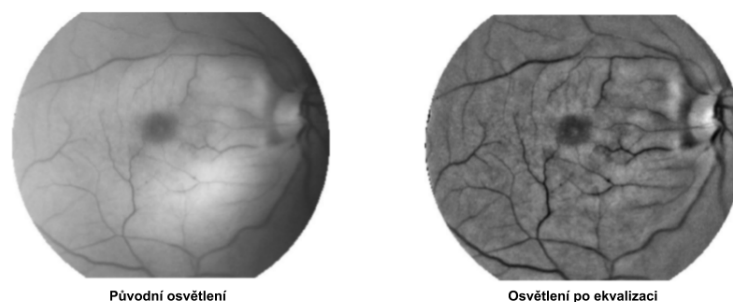
Obr. 1.21: Dva příklady obrazů sítnice a fuzzy konvergence. [12]

Jako druhý znak, který je charakteristický pro optický disk, využívají autoři studie úroveň jasu, kdy optický disk je v obrazu nejjasnější oblastí. Pro jeho extrakci by proto stačilo využít metodu prahování. Pravidlo nejsvětlejší oblasti však nemusí platit vzhledem k nepravidelnému osvětlení oka, které je způsobené jeho zakřivením. Pro odstranění nepravidelného osvětlení je proto využito ekvalizace obrazu podle vztahu:

$$I_{ekv}(x, y) = I(x, y) + m - A(x, y), \quad (1.18)$$

kde je počítána ekvalizovaná hodnota každého pixelu $I(x, y)$, m je žádoucí průměrná intenzita přičtená ke každému pixelu a $A(x, y)$ je naopak odečítaná lokální intenzita pixelu. Ekvalizace probíhá v okně o definované velikosti posouváním po obraze. Příklad ekvalizovaného obrazu je na obrázku 1.22.

Následně může být využito zmiňované metody prahování a nalezení pravděpodobné oblasti výskytu optického disku. Na základě dalších definovaných pravidel byla následně porovnávána oblast extrahovaná pomocí fuzzy konvergence spolu s nejjasnější oblastí v ekvalizovaném obraze. [12]



Obr. 1.22: Ukázka ekvalizace jasových hodnot. [12]

Zhodnocení

Kombinací fuzzy konvergence a extrakce nejsvětlejší oblasti bylo dosaženo kvalitních výsledků, zejména však u fotografií zdravých osob.

1.5.2 Fuzzy C-means a iterativní prahování

Metoda je založená na převedení obrazu na šedotónový, odstranění cév sítnice pomocí morfologických operací a na zkombinování dvou základních metod segmentace - metodě fuzzy C-means a metodě prahování.

Předzpracování dat

Prvním krokem je převedení obrazu na šedotónový zavedením následujícího vztahu:

$$Gr = 0,3 \times R + 0,59 \times G + 0,11 \times B, \quad (1.19)$$

kde R představuje červenou, G zelenou a B modrou složku obrazu.

Pro redukci šumu, který se může v obraze vyskytovat, je následně aplikován mediánový filtr.

Za účelem odstranění cév, které překrývají optický disk, je aplikována morfologická operace dilatace s oknem o velikosti 10×10 , a poté operace uzavření s oknem stejné velikosti. Výsledkem je vyhlazený obraz bez cév sítnice. [23]

Princip metody

Pro segmentaci optického disku je použita metoda Fuzzy C-means. Počet shluků, do kterých jsou pixely obrazu přiřazovány je nastaven na 3. Nejjasnější pixely jsou charakteristické pro oblast hledaného optického disku. [23]

Metodu C-means následuje iterativní prahování s prahovou hodnotou definovanou vztahem:

$$T = \mu \times (Gr_{min} + Gr_{max}), \quad (1.20)$$

kde hodnoty Gr_{min} a Gr_{max} jsou odvozeny z šedotónového obrazu. Gr_{min} odpovídá minimální hodnotě v tomto obraze a Gr_{max} hodnotě maximální. Podle proměnné μ nese metoda název iterativní prahování. V prvním kroku je $\mu = 0,5$, postupně však dochází k jeho nárůstu. K nárůstu dochází na základě poměru množství pixelů, které jsou vyšší než definovaný práh a pixelů ostatních. Pokud je poměr vyšší než $1/25$, je faktor μ navýšen. V případě nižšího poměru je dosaženo požadované nalezené prahové hodnoty, na základě které dochází k prahování obrazu.. [23]

Zhodnocení

Po aplikaci na několik testovacích databází dosahují ve své studii [23] Rust et al., 2017 kvalitních výsledků, což dokazuje její robustnost, zejména z důvodu vhodné kombinace různých metod aplikovaných na jednotlivé dílčí úkoly.

1.5.3 Modifikace architektury Unet

Za pomoci modifikace architektury Unet lze dosáhnout segmentace optického disku i uvnitř se nacházejícího optického pohárku. [24]

Předzpracování dat

Předzpracování obrazů je prováděno prostřednictvím adaptivní ekvalizace histogramu kontrastu šedotónového obrazu. V jednotlivých oblastech obrazu je postupně měněna barva obrazu a kontrast je ekvalizován interpolací mezi těmito oblastmi.

Vhodné je také následné rozšíření trénovací množiny pomocí různých rotací a posunů v obrazech, které jsou pro učení k dispozici. [24]

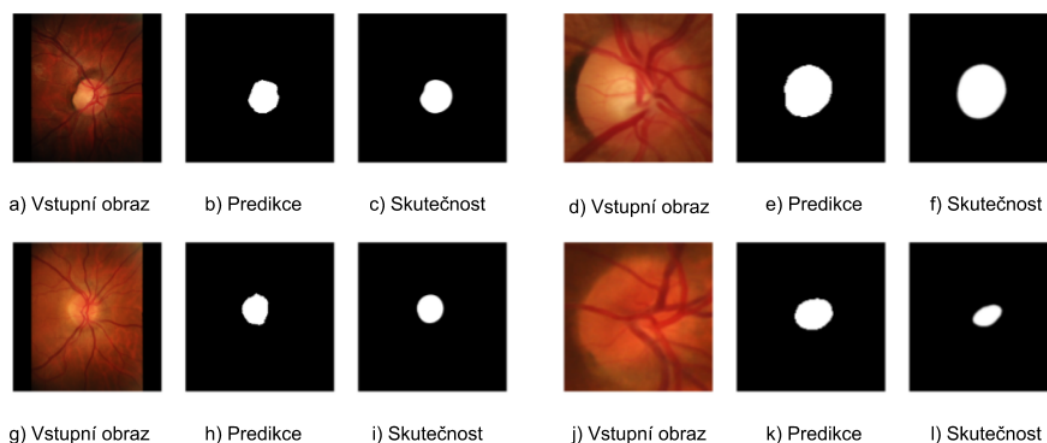
Princip metody

Metoda je založena na architektuře konvoluční neuronové sítě Unet popisované v jedné z předchozích kapitol. Je složena z části klesající, kde dochází ke snižování rozměrů obrazu a z části stoupající, kde jsou rozměry opět navyšovány.

Sít má v každé z konvolučních vrstev méně filtrů než klasická architektura Unet. Tato úprava je zavedena za účelem snížení množství parametrů a času pro učení a pro průchod sítí, zároveň ovšem v případě segmentace optického disku a pohárku není pozorováno snížení kvality výstupu. [23] Schéma použité architektury je na obrázku 1.23.

Zhodnocení

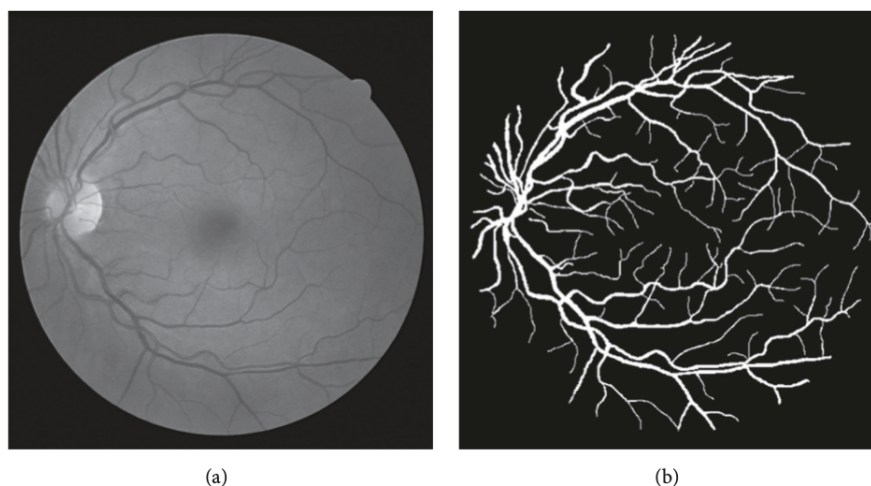
Navržená metoda dosahuje velice kvalitních výsledků, mezi její největší klady však patří krátký čas učení a čas průchodu obrazu sítí, což je způsobeno upravením architektury. Vlivem upravení architektury nese síť i méně parametrů, díky čemuž zabírá uložený model v paměti jen několik MB. Výsledky metody jsou ukázány na obrázku 1.24.



Obr. 1.24: Výsledky po implementaci modifikované Unet. Obrázky a) - c) zobrazují segmentaci optického disku s nejlepším dosaženým výsledkem, obrázky g) - i) pak nejhorší segmentaci optického disku. Stejně tak obrázky d) - f) ukazují nejlepší dosažený výsledek při následné segmentaci očního pohárku a obrázky j) - l) výsledek nejhorší. [24]

1.6 Metody segmentace cév sítnice

Stejně jako segmentace optického disku i segmentace cév sítnice může poskytnout informaci o řadě důležitých onemocnění mezi které se řadí například hypertenze, cukrovka nebo i degenerace samotné sítnice. Příklad segmentovaných cév sítnice je na obrázku 1.25.



Obr. 1.25: Manuálně segmentované cévy sítnice. Na obrázku a) je původní obraz a na obrázku b) jsou z tohoto obrazu manuálně vysegmentovány cévy sítnice.⁹

1.6.1 Level-set metoda

Tato běžně využívaná metoda je založena na nalezení optimálního obrysu segmentované oblasti. [13]

Předzpracování dat

U vstupního obrazu nejprve dochází k odstranění nežádoucího pozadí v obraze prostřednictvím metody K-means, která rozděluje data do shluků na základě společné informace. V tomto případě jsou rozdělovány pixely na základě jejich barevné informace. [13]

Provedením tohoto kroku můžou v obraze vzniknout nevyplněné oblasti, které je nutno opravit. Za účelem odstranění nežádoucího šumu v obraze je dále aplikován některý z potlačujících filtrů, například filtr mediánový nebo průměrovací. [13]

Po aplikaci filtru pro redukci šumu je provedena morfologická operace zvaná Bottom-hats, kdy je původní obraz odečtený od obrazu vzniklého operací uzavření. [35] Protože jsou cévy v obraze v obvyklém případě nejtmavší strukturou, napomáhá operace Bottom-hats jejich zvýraznění.

Posledním krokem předzpracování je extrakce pouze zeleného barevného kanálu, který obsahuje nejvíce informací o cévách sítnice. [13]

Princip metody

Metoda level-set je velice podobná metodě aktivních kontur a užitečná zejména při snaze vyhledat obrys a povrch charakterizující určitou strukturu v obraze. Křivka je definována svojí nulovou hladinou - řezem v dané rovině určité vícerozměrné funkce. Tato funkce nazývaná level-set každému bodu přiřadí hodnotu, která udává jeho výšku vůči definované hladině. Vzhledem ke gradientu obrazu a dalším metrikám je funkce postupně adaptována na požadovaný obrys hledané struktury. [13]

Odchylka nulové hladiny level-set funkce od hledaného obrysu struktury je charakterizována energetickou funkcí, která je v průběhu algoritmu minimalizována. Obecně ji lze definovat: [36]

$$F^{MS} = \int_{\Omega} (u - I(X))^2 dX + \mu \int_{\Omega/C} |\nabla u|^2 dX + v |C|, \quad (1.23)$$

V tomto vztahu I je originální obraz, Ω představuje prostor obrazu, a u je obraz, který originální obraz aproximuje. $|C|$ je délka definované kontury. První člen ovlivňuje přiblížení aproximovaného obrazu u k reálnému I , druhý člen vyhlazuje oblast definovanou konturou C a třetí reguluje délku kontury $|C|$ za pomoci parametru v . [36]

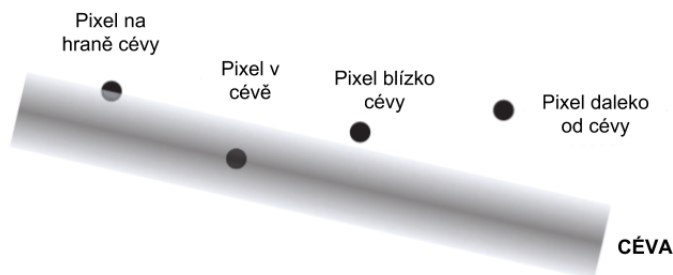
1.6.2 Metoda zesílení lokálních znaků

S odlišným přístupem přicházejí autoři metody založené na zdůraznění lokálních znaků v obraze. Algoritmus pracuje se zelenou barevnou složkou v obraze, která obsahuje nejvíce kontrastní informaci. V té jsou nalezeny definované znaky na jejichž základě je natrénován SVM klasifikátor (metoda podpůrných vektorů), který slouží k následné segmentaci cév. Důležité je také dodatečné zpracování obrazu po průchodu klasifikátorem. [19]

Extrakce lokálních znaků

Na základě obdélníkového a liniového charakteru cév v sítnici jsou definovány znaky, podle kterých jsou rozřazovány jednotlivé pixely do tříd. Pixely jsou prohledávány

ve všech směrech a v případě splnění podmínek je postupně prodlužována linie představující detekovanou cévu. Jedním z pravidel pro rozřazení pixelu do tříd je situace, kdy rozdíl intenzity mezi vyšetřovaným pixelem a pixelem sousedním je menší než stanovená prahová hodnota. Podle jasové intenzity jsou pixely v obraze rozřazovány do 4 skupin, ukázaných na obrázku 1.26. [19]



Obr. 1.26: Třídy, do kterých jsou rozřazovány pixely cév. [19]

Dalším extrahovaným znakem, který definuje přítomnost cévy je její tmavší jasová hodnota v porovnání s hodnotou pozadí. Vyšetřovaný pixel je ohraničen oknem o definované velikosti. Ze stanoveného okna je počítána průměrná intenzita, což je robustnější než pokud by byla uvažovaná jasová hodnota samotného pixelu. [19]

Jako poslední znak je extrahován morfologický gradient, který zajišťuje přesnější segmentaci tenčích cév. Je definován jako rozdíl mezi morfologickými operacemi dilatací a erozí a je vypočten: [19]

$$F_{grad} = \frac{1}{6} \sum_{i=1}^6 (I \oplus d_i - I \otimes d_i). \quad (1.24)$$

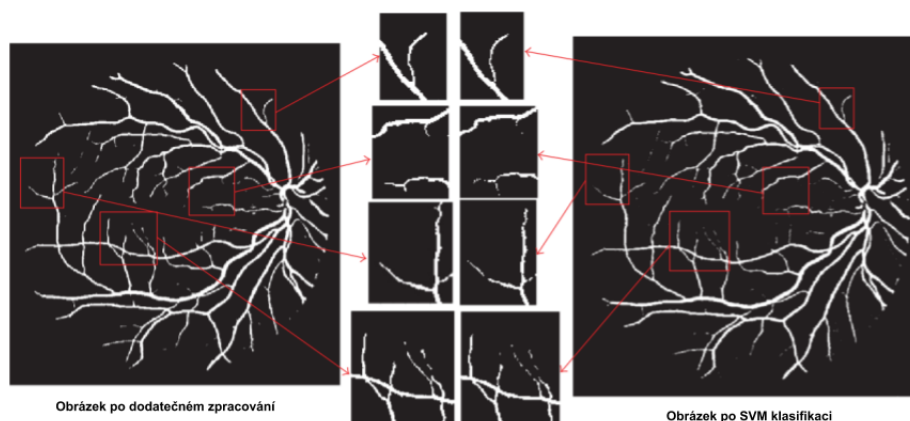
Ve vztahu $I \oplus d_i$ představuje dilataci a $I \otimes d_i$ erozi obrazu I s diskovitým strukturním elementem d o poloměru i .

Princip metody

Pixely v obraze jsou na základě extrahovaných charakteristických znaků děleny do skupin pomocí metody podpůrných vektorů - SVM klasifikátoru. Ten prokládá prostor lineárně do dvou tříd, kdy hodnota minima vzdáleností od oddělovací roviny je co nejvyšší. [19]

Následné zpracování obrazu

Dodatečné zpracování je nutné zejména z důvodu nespojitostí, které se mohou po klasifikaci vyskytnout, jelikož výstupem ze SVM klasifikátoru je pravděpodobnostní mapa. Pro jejich propojení byla použita morfologická rekonstrukce založená na stanovené prahové hodnotě. Použitím prahování je výrazně zvýšena přesnost implementované metody. Rozdíl mezi výsledkem dosaženým bez dodatečného zpracování a s ním je na obrázku 1.27. [19]



Obr. 1.27: Rozdíl mezi výsledkem dosaženým bez dodatečného zpracování a s ním. [19]

1.6.3 Metoda založená na 3-fázovém modelu hlubokého učení

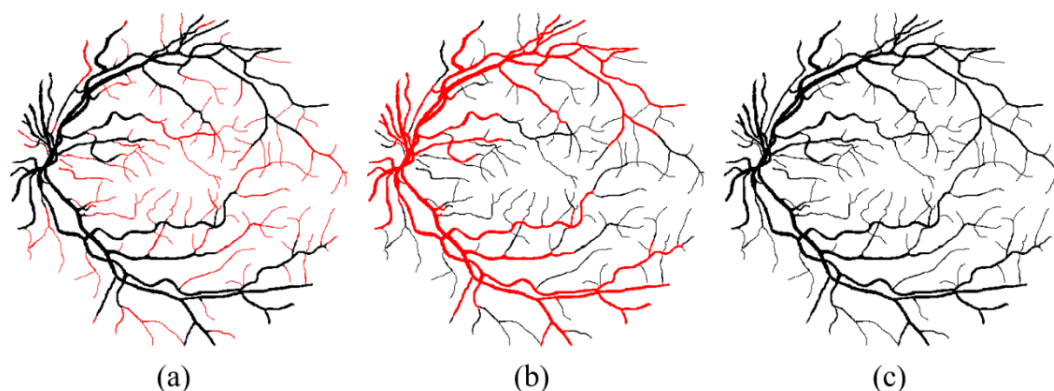
Sofistikovanější přístup představuje následující metoda založená na hlubokém učení. Obsahuje tři bloky, kdy první je zodpovědný za segmentaci silných cév, druhý za segmentaci tenkých cév a třetí zprostředkovává spojení segmentovaných obrazů. [33]

Problémy segmentace cév sítnice

Při segmentaci očních cév pomocí hlubokého učení se může vyskytovat několik problémů. Pro diagnostiku je důležitá segmentace cév silných i slabších, většina pixelů však náleží cévám silnějším. Všechny metody založené na hlubokých neuronových sítích mají tendenci minimalizovat celkovou ztrátu, což způsobuje přesnější segmentaci silných cév a nepřesnou segmentaci neméně důležitých cév tenkých. Souvisejícím problémem je fakt, že silné cévy mají zpravidla vyšší kontrast a poměr signálu vůči šumu než cévy tenké. To způsobuje, že příznaky, které se osvědčí pro segmentaci silnějších cév, nemusí nutně zajistit správnou segmentaci cév tenkých. [33]

Předzpracování dat

Pro urychlení průběhu jsou snímky opět nejprve převedeny na šedotónové výběrem zelené složky z původního obrazu s největším množstvím potřebné informace. I v tomto případě jsou pro umělé navýšení množství obrazů v databázi prováděny různé transformace jako rotace, úprava kontrastu či změny velikosti. Následně je pro učení neuronové sítě provedena manuální anotace cév. Na základě anotace jsou odděleny tenké a silné cévy. Příklad anotace můžeme vidět na obrázku 1.28, kde černou barvou jsou anotované cévy pro segmentaci a červenou cévy, ze kterých není počítána chybová funkce. [33]



Obr. 1.28: Manuálně anotované cévy sítnice. Na obrázku a) jsou černě označeny silné cévy, na obrázku b) jsou naopak černě anotovány tenké cévy sítnice a obrázek c) je anotace pro blok spojující předchozí segmentace. [33]

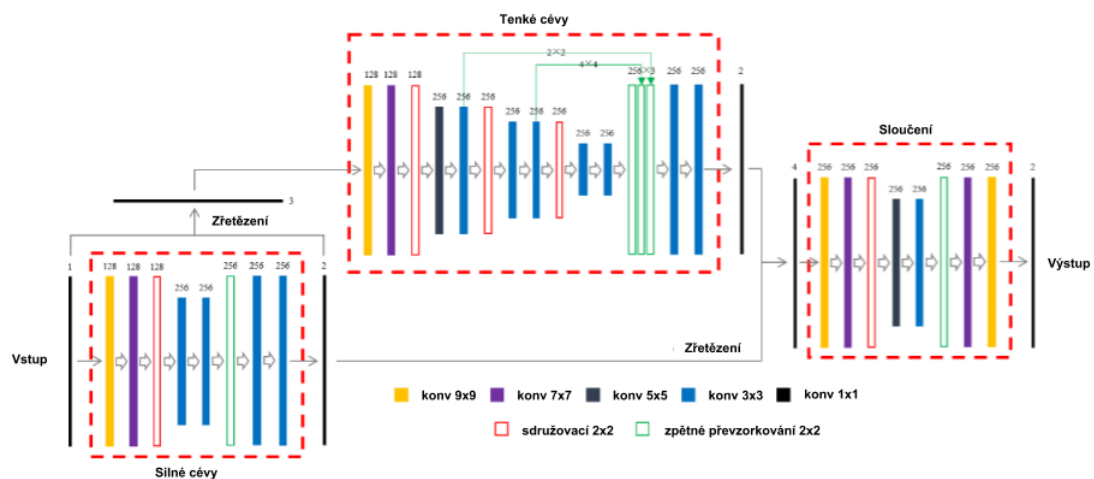
Princip metody

Metoda je založena na plně konvolučních architekturách tří bloků o odlišném množství konvolučních vrstev. Architektura sítě je na obrázku 1.29.

První blok je pro segmentaci **silných cév**. Jako pozitivní jsou označeny pouze pixely silných cév a při výpočtu chybové funkce je ztráta způsobená tenkými cévami zanedbávána.

Druhý blok je založený na segmentaci **tenkých cév**, kdy je naopak počítáno pouze s pixely tenkých cév. Vstupem do obou těchto obrazů je zelená složka vstupního obrazu a výstupem jsou odpovídající segmentované obrazy.

Poslední blok má na starosti **spojení obrazů z předchozích bloků** spojení obrazů z předchozích bloků dohromady, které jsou zároveň na jeho vstupu. Jako pozitivní pixely jsou označeny pixely tenkých i silných cév. Výstupem je hledaný segmentovaný obraz s tenkými i silnými cévami.



Obr. 1.29: Architektura 3 bloků plně konvolučních sítí. [33]

Chybová funkce je opět vyjádřena výpočtem vzájemné entropie. [33]

Zhodnocení

Snahou navržené metody je potlačení dominance silnějších cév. Za pomoci zavedení oddělených bloků je dosaženo velice dobrých výsledků napříč různými databázemi dat, což svědčí o robustnosti metody. [33]

2 Výsledky studentské práce

2.1 Použité nástroje

Navržená metoda pro segmentaci optického disku a cév v sítnici byla implementována v programovacím jazyce **Python**. Jeho využití je vhodné zejména díky jednoduchosti, srozumitelnosti a hlavně stručnosti. Většina knihoven potřebných pro práci se nachází na webových stránkách <https://pypi.org/>. Pro účely této práce je navržena práce s knihovnou pro zpracování obrazu **OpenCV** a knihovnou pro hluboké učení **Keras**.

Knihovna Keras pro hluboké učení je navržena přímo v programovacím jazyce Python za účelem umožnění provádění rychlých experimentů. Pro svoji funkci potřebuje na pozadí běh nástroje s názvem Tensorflow nebo Theano pro numerické výpočty. V případě této práce je pro běh knihovny Keras doporučený nástroj **Tensorflow**.



Obr. 2.1: Využité nástroje

Obrázek 2.1 ukazuje ikony využitých nástrojů.

2.2 Příprava dat

2.2.1 Tvorba trénovací databáze

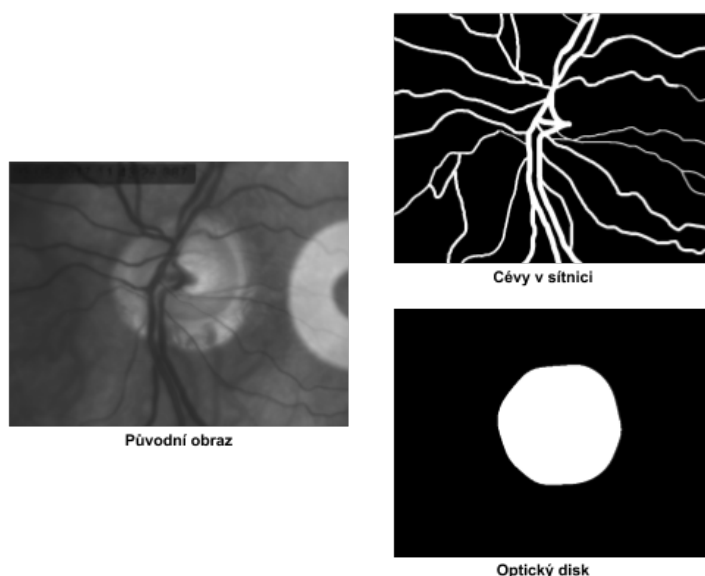
Zvolená metoda je založena na učení neuronové sítě tzv. s učitelem. To znamená, že je síť učena na základě množiny anotovaných dat. Tvorba učební množiny je jedním z nejzásadnějších úkolů při implementaci navržené metody.

Obrazy pro implementaci metody vhodné k segmentaci optického disku a cév sítnice byly získány ze sekvencí pořízených experimentálním oftalmoskopem, kdy každý obraz je získán průměrem několika snímků sekvence.

Obrazy bylo pro účely diplomové práce nutné anotovat. Anotace dat spočívá v postupném procházení připravených obrazů a co nejpřesnějším ručním vykreslení viditelných cév a optického disku. Zejména anotace cév sítnice je vzhledem k jejich rozvětvenosti a často nezřetelným obrysům časově velmi náročná.

Na základě takto vzniklého ručního označení jsou z obrazů dále vytvořeny jejich masky představující ideální výsledek, kterého by mělo být po provedené segmentaci dosaženo.

Pomocí množiny manuálně označených dat se navržená síť učí postupně zařazovat jednotlivé pixely daného obrazu. Příklad obrazu a masek, které vznikly na základě manuální anotace cév a optického disku v sítnici je na obrázku 2.2.



Obr. 2.2: Obraz s odpovídající maskou po manuální anotaci pro cévy sítnice a pro optický disk.

Popsaným způsobem byla vytvořena trénovací databáze zvláště pro segmentaci optického disku a zvláště pro segmentaci cév v sítnici. Celá dostupná databáze snímků pro segmentaci tvoří 483 šedotónových snímků, přičemž rozlišení obrazů je 1000×770 pixelů. Manuálně bylo anotováno 100 snímků pro segmentaci cév v oku, pro segmentaci optického disku je označeno obrazů 200. Hodnoty popisující databázi jsou shrnuty v následující tabulce 2.1.

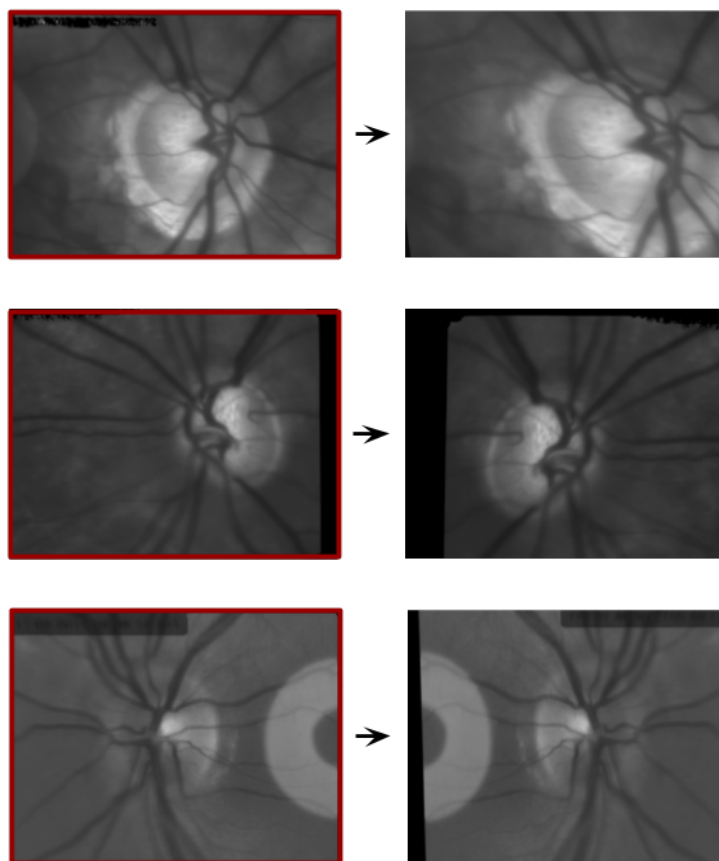
2.2.2 Rozšíření učební množiny

Jelikož ruční anotace dostatečného množství dat je především v případě segmentace cév, ale i optického disku nereálná, ideálním řešením je po anotování menší mno-

Tab. 2.1: Vytvořená trénovací databáze.

	Počet snímků
Celá databáze	483
Anotované snímky sítnice	100
Anotované snímky optického disku	100

žiny velikost vytvořené trénovací databáze uměle navýšit. Toho lze docílit aplikací vybraných elastických operací, mezi které řadíme například různé posuny v obraze, změny měřítka, přidání šumu, či rotace obrazů na základě Gaussovy distribuce a maximální odchylky s následnou interpolací pixelů na jejich nová umístění. [22] Příklady, jak mohou obrazy po aplikaci zmiňovaných operací vypadat, jsou znázorněny na obrázku 2.3, kde červeně jsou ohrazeny obrazy původní, na které byly náhodně aplikovány některé z transformací.



Obr. 2.3: Příklad rozšíření trénovací množiny obrazů.

Navržená metoda pro segmentaci optického disku byla po provedení umělého rozšíření sady dostupných obrazů trénovaná na 450 obrazech a validována na 150

validačních obrazech a metoda pro segmentaci cév sítnice byla trénována na 528 obrazech a validována na 192 - viz tabulka 2.3.4.

Tab. 2.2: Velikost rozšířené trénovací databáze.

	Optický disk	Cévy sítnice
Počet trénovacích obrazů	450	528
Počet validačních obrazů	150	192

2.3 Segmentace optického disku

Cílem praktické části bylo implementovat vhodnou metodu pro segmentaci optického disku a cév, které se nacházejí v sítnici. Zvolené metody jsou založené na konvolučních neuronových sítích, pro optický disk se konkrétně jedná o architektury Unet a MaskRCNN. Metoda Unet byla vybrána především proto, že důvodem jejího vzniku byla práce s biomedicínskými obrazy. Metoda MaskRCNN zejména pro porovnání, jelikož se jedná o několik let novější přístup než jakým je metoda Unet. Postup navržené metody je znázorněný v blokovém schématu na obrázku 2.4.

Obě metody jsou vhodné také díky svému specifickému zpracování učebních dat, kterých nemusí být příliš velké množství, a zároveň z důvodu jednoduché implementace za předpokladu kvalitních dosažených výsledků.

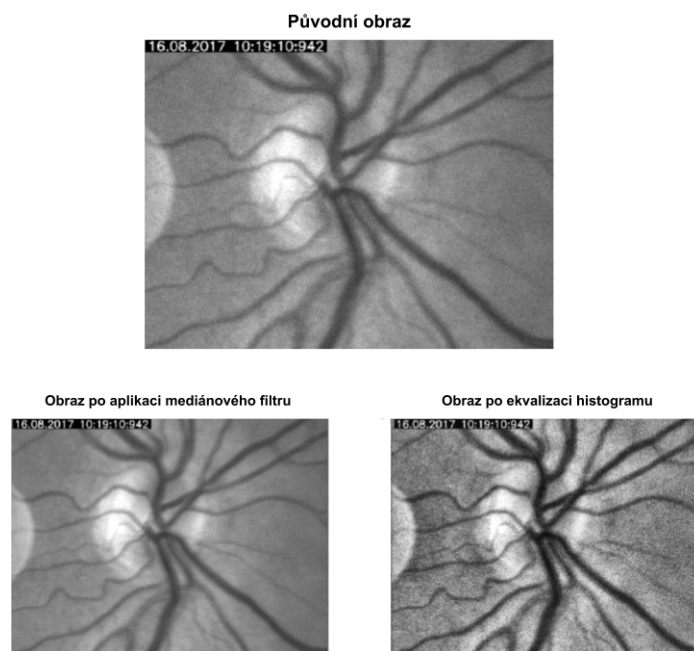


Obr. 2.4: Blokové schéma segmentace optického disku.

2.3.1 Předzpracování obrazů pro segmentaci optického disku

Obrazy využívané pro segmentaci optického disku je možné zejména pro zjednodušení převést na šedotónové. Nejvhodnější metodou je převedení na základě všech barevných složek. Navrhovaná metoda je však přizpůsobena i na vstup barevných obrazů.

Pro odstranění impulzního šumu v obraze, je dále navrženo použití některého z vhodných filtrů, například filtru mediánového. Pro upravení kvality obrazu a zlepšení jeho kontrastu lze rovněž využít ekvalizaci histogramu.



Obr. 2.5: Předzpracování obrazu pro segmentaci optického disku.

V této práci bylo nejlepších výsledků pro segmentování optického disku dosaženo v případě použití všech barevných kanálů obrazu. Porovnávány byly rovněž dosažené výsledky v případě, že byla provedena ekvalizace histogramu využívaná ve velkém množství výzkumů. Ta má za následek zvýšení kontrastu mezi oblastí optického disku a pozadím. Ta však vzhledem k neuspokojivým výsledkům nebyla ve výsledné práci využita. Docházelo k zesvětlení některých oblastí a následně k falešně pozitivním detekcím optického disku.

Jelikož zvolená metoda pro nalezení optického disku Unet využívá pro segmentaci architekturu konvoluční neuronové sítě zvanou Resnet, je důležitým krokem upravení vstupního obrazu podle požadavků této sítě. Zde dochází za pomoci funkce *preprocess_input* již implementované v používané knihovně Keras k odečtení průměrné jasové hodnoty od každého pixelu z obrazu.

Následuje upravení obrazu na požadovanou velikost vstupu v závislosti na zvolené architektuře neuronové sítě. Vstupní velikost je závislá na použitých konvolučních maskách uvnitř architektury. Vzhledem k původní velikosti trénovacích obrazů a vzhledem k využívaným konvolučním maskám byl vstupní obraz upraven na rozměr 768×768 . Za pomoci každého takto upraveného obrazu dochází následně k rozšíření

trénovací databáze postupem popsaným v předcházející kapitole. Poslední úpravou je normalizace jasových hodnot obrazu na rozmezí 0 - 1.

Předzpracované obrazy jsou na vstupu rozděleny na trénovací a validační množinu, kdy trénovací množina slouží k naučení sítě na charakteristických znacích ve vstupních obrazech a validační množina poskytuje informaci o úspěšnosti segmentace na neznámých obrazech a zajišťuje úpravu koeficientů vah v jednotlivých vrstvách sítě v jednotlivých epochách.

2.3.2 MaskRCNN

Pro segmentaci optického disku byla vyzkoušena metoda z roku 2017 s názvem MaskRCNN. Ta funguje na principu vyhledání a ohraničení oblastí zájmu. Objekty v ohraničených oblastech jsou klasifikovány do tříd. V případě, že je objekt klasifikován jako optický disk, dochází k jeho segmentaci. Metoda, blíže popsána v teoretické části práce, byla implementována s využitím repozitáře:

https://github.com/matterport/Mask_RCNN.git.

V části, kde dochází k vyhledávání charakteristických znaků pro klasifikaci a segmentaci objektů v zájmových oblastech je i v tomto případě zařazena konvoluční síť s architekturou Resnet. V tomto případě bylo dosaženo nejkvalitnějších výsledků s hlubší sítí - *Resnet101* se 101 váženými vrstvami.

Upravení obrazu na vstup sítě je zprostředkováno za pomoci konfigurační třídy *InitialConfig*. Zde dochází k nastavení parametrů od použité architektury Resnet101, velikosti vstupního obrazu - 640×640 pixelů, přes velikosti konvolučních masek, počet tříd, do kterých je klasifikováno, počet epoch, až k parametrům týkajících se počtu využívaných jader při trénování - v našem případě 8. Na základě vytvořené konfigurační instance dochází rovněž k předzpracování vstupních obrazů v podobě odečtení průměrných hodnot od jednotlivých pixelů. Příklad inicializace konfigurační instance pro síť MaskRCNN je na výpisu 2.1.

Výpis 2.1: Konfigurační parametry

```
class InitialConfig(Config):
    """
    Konfigurace pro segmentaci optického disku.
    """

    NAME = "disks" # Jmeno tridy detekovanych objektu
    BACKBONE = "resnet101" # Architektura pro ekstrakci
    GPU_COUNT = 1 # Pocet vyuzivanych zdroju CPU
    IMAGES_PER_GPU = 8 # Pocet obrazu na jedno jadro GPU
    NUM_CLASSES = 1 + 1 # Pocet trid = pozadi + disks
```

```

IMAGE_SHAPE = [640, 640, 3] # Definovane rozmery
IMAGE_RESIZE_MODE = "square" # Zachovany pomer stran
TRAIN_ROIS_PER_IMAGE = 32 # Mnozstvi oblasti zajmu
STEPS_PER_EPOCH = 250 # Pocet kroku v jedne epose
VALIDATION_STEPS = 20 # Pocet kroku pro validaci
BATCH_SIZE = 2 # Mnozstvi obrazy na vstupu

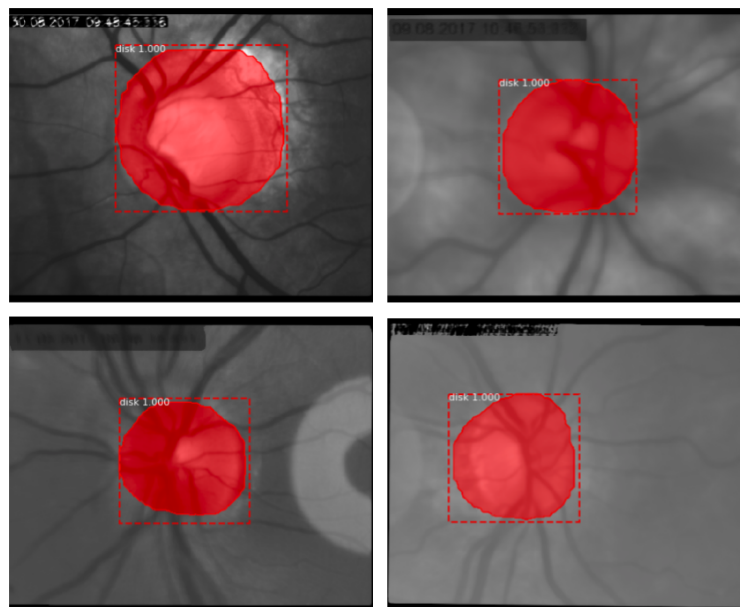
config = InitialConfig()
config.display()

```

Načítání masek je zprostředkováno upravením zděděné třídy *Dataset*, pomocí jejíž instance dochází následně k naplnění trénovacími a validačními daty.

I v případě metody MaskRCNN je architektura předtrénována na obrazové databázi *ImageNet*, na základě čehož dochází k počáteční inicializaci vah.

V metodě je několik chybových funkcí, které hodnotí především správnost zařazení objektů v ohraničených oblastech zájmu do konkrétní třídy. V případě segmentace je nejdůležitější metrika počítána metodou *mrcnn_mask_loss_graph*. Ta počítá vzájemnou entropii pro masku a obraz konkrétní oblasti označené jako oblasti, ve které se nachází optický disk. Pomocí vzájemné entropie je stejně tak zhodnocována správnost nalezení ohraničení požadované oblasti zájmu. Výsledné segmentované optické disky jsou na obrázku 2.6.



Obr. 2.6: Detekované optické disky za pomoci metody MaskRCNN.

2.3.3 Metoda Unet

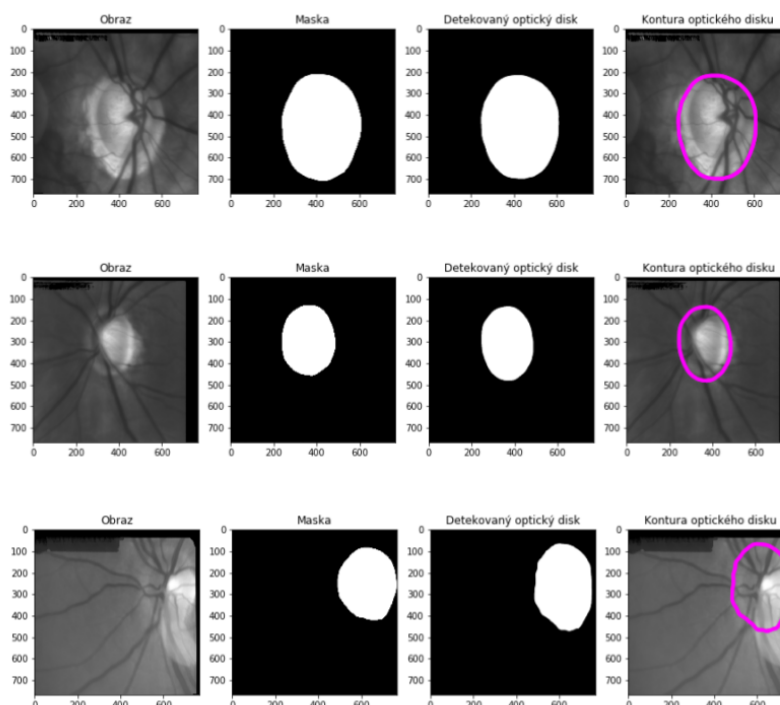
Další z metod využitých pro segmentaci optického disku je konvoluční neuronová síť Unet popsaná v teoretické části.

Sestupná část této sítě sloužící k extrakci charakteristických znaků a oblastí v obraze je tvořena architekturou *Resnet18* s 18 váženými vrstvami. Architektura Resnet je charakteristická zejména svými skrytými propojeními, které zabraňují přetrénování sítě.

Za pomoci této metody byly jednotlivé pixely obrazu rozděleny na pixely ležící v oblasti optického disku a na pixely pozadí. Aktivační funkcí na konci sestupné části architektury byla tedy *sigmoidea*, rozdělující pixely na hodnoty 0 a 1.

Pro zkvalitnění výsledků byl využit transfer learning a hodnoty vah v sestupné části byly inicializovány podle natrénované sítě Resnet18 na databázi *ImageNet*. K jejich další úpravě dochází průchodem vytvořené trénovací databáze architekturou.

Odchylka predikované hodnoty od reálné hodnoty byla stanovena na základě *Diceova koeficientu* a jako optimalizační algoritmus byla zvolena adaptivní metoda - *Adam*.



Obr. 2.7: Detekované optické disky za pomoci metody Unet.

Množství epoch pro učení implementované sítě bylo nastaveno na 150, kdy jedna epocha představuje ukázání sítě všech trénovacích obrazů spolu s jejich upravenými

variantami pro navýšení trénovací množiny.

Výsledky využití implementované metody Unet, implementované za pomoci repositáře: https://github.com/qubvel/segmentation_models jsou ukázány na obrázku 2.7.

2.3.4 Zhodnocení použitých metod

Pro segmentaci optického disku byla vyzkoušena řada přístupů s různými kombinacemi parametrů, které vedly k dosažení různě kvalitních výsledků. Pro porovnání úspěšnosti jednotlivých algoritmů byly využity 3 metriky - senzitivita (recall), specifita a Diceův koeficient podobnosti.

Vztah pro výpočet **senzitivity**, která udává míru správně zařazených pixelů optického disku [29]:

$$\text{Senzitivita} = \frac{TP}{TP + FN}, \quad (2.1)$$

kde TP značí pixely, které jsou správně zařazeny jako pixely optického disku. FN naopak představuje množství pixelů chybně označených jako pixelů pozadí přesto, že ve skutečnosti patří do oblasti optického disku.

Specifita oproti tomu vyjadřuje míru správně zařazených pixelů pozadí a lze ji spočítat jako [29]:

$$\text{Specifita} = \frac{TN}{TN + FP}. \quad (2.2)$$

Zde TN označuje pixely správně zařazené mezi pixely pozadí a FP naopak pixely nesprávně přiřazené mezi pixely optického disku.

Poslední porovnávanou metrikou je **Diceův koeficient podobnosti** mezi ručně anotovanou maskou a maskou predikovanou po průchodu implementovanou metodou. Je definován vztahem [29]:

$$\text{Dice} = \frac{2TP}{2TP + FP + FN}. \quad (2.3)$$

Právě Diceův koeficient podobnosti slouží jako hlavní metrika pro porovnání úspěšnosti implementovaných segmentačních technik.

Segmentační architektura

Asi nejzásadnější byla volba hlavní použité architektury konvoluční neuronové sítě zodpovědné za segmentaci. Při segmentaci optického disku bylo porovnáno několik architektur, jmenovitě architektury Linknet, PSPNet, Unet a MaskRCNN.

Tab. 2.3: Úspěšnost segmentačních architektur.

	Senzitivita	Specifita	Diceův koeficient
PSPNet	75,33 %	98,01 %	85,68 %
MaskRCNN	86,24 %	89,20 %	85,72 %
Linknet	90,67 %	98,99 %	92,34 %
Unet	92,75 %	98,78 %	92,86 %

V tabulce 2.3.4 jsou znázorněny nejlepší dosažené výsledky pro každou z použitých architektur při různých kombinacích parametrů jako je počet epoch, optimalizační algoritmus, předzpracování obrazu v podobě ekvalizace histogramu či architektura v sestupné části segmentační sítě. Jak je možné v tabulce vidět, nejvyšší úspěšnosti bylo dosaženo v případě použití metody Unet. Zatímco metoda s využitím architektury Linknet dosahuje jen nepatrně horších výsledků, novější metoda PSPNet je v porovnání s ostatními testovanými přístupy výrazně horší. Přístup s využitím architektury MaskRCNN spolehlivě ohraničí oblast zájmu obsahující optický disk. Ohraničení oblasti zájmu s optickým diskem dosahuje úspěšnosti až 98 %. Samotná jeho segmentace však při implementaci s využitím zmiňovaného repozitáře v porovnání s metodami Unet i Linknet poskytuje nepatrně horší výsledky.

Veškeré další uvedené analýzy jsou vzhledem k nejvyšší úspěšnosti prováděné se segmentační architekturou Unet.

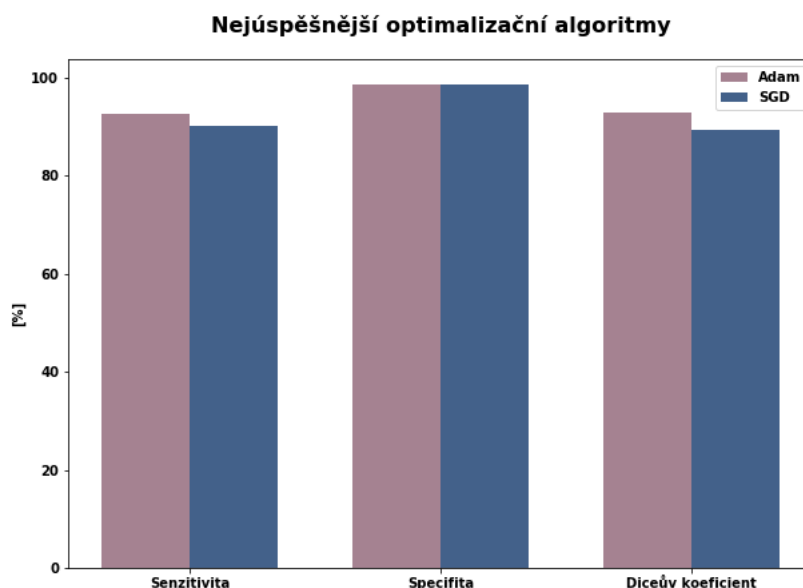
Optimalizační algoritmus

Další z významných parametrů se byla při trénování optimalizační metoda použita pro snižování chybové funkce při učení sítě. Jako nejúspěšnější se ukázaly algoritmy SGD a Adam. Zhodnocení úspěšnosti u každého z nich ukazuje graf 2.8 spolu s tabulkou 2.3.4.

Tab. 2.4: Úspěšnost optimalizačních algortimů.

	Senzitivita	Specifita	Diceův koeficient
SGD	90,23 %	98,77 %	89,44 %
Adam	92,75 %	98,78 %	92,86 %

Nejlépeších výsledků bylo dosaženo s využitím optimalizačního algoritmu Adam. Ten adaptivně upravuje rychlost učení jednotlivých parametrů a lze jej považovat za následovníka druhého zhodnocovaného přístupu s algoritmem SGD.



Obr. 2.8: Porovnání nejúspěšnějších optimalizačních algoritmů.

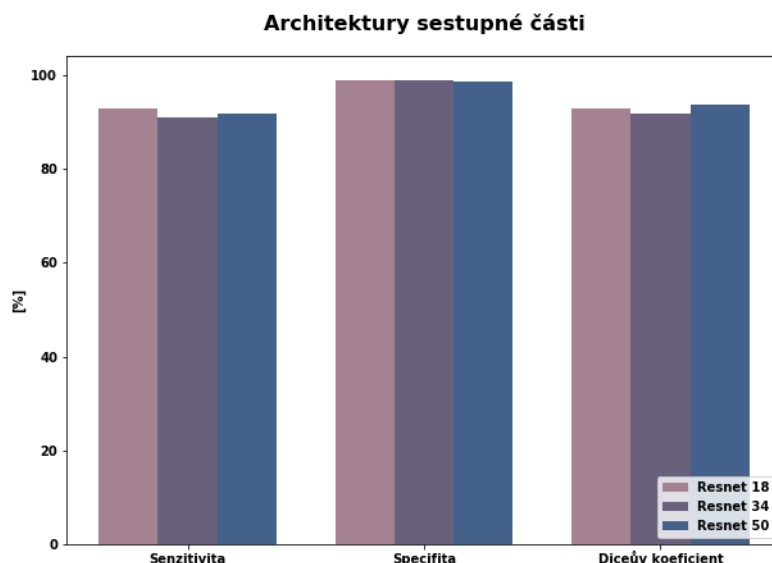
Vnořená architektura sestupné části

Vnořená architektura v sestupné části segmentační sítě je zodpovědná za extrakci charakteristických příznaků v obraze a za následné rozpoznání jednotlivých struktur v obraze. Jako nejúspěšnější se ukázala architektura Resnet. Výsledné porovnání bylo provedeno pro různý počet vážených vrstev v této architektuře - 18, 34 a 50. Porovnání je znázorněno v grafu 2.9 a hodnoty shrnuty v tabulce 2.3.4.

Tab. 2.5: Úspěšnost různě hlubokých architektur sestupné části Resnet.

	Senzitivita	Specifita	Diceův koeficient
Resnet 18	92,75 %	98,78 %	92,86 %
Resnet 34	91,05 %	98,98 %	91,90 %
Resnet 50	91,77 %	98,67 %	93,68 %

Jelikož dosažené hodnoty nejrelevantnější z měřených metrik - Diceova koeficientu podobnosti se u všech architektur pohybují kolem hodnoty 92 %, byla jako nejoptimálnější řešení vybrána architektura Resnet 18 s 18 váženými vrstvami. Množství parametrů modelu a výpočetní náročnost jsou vlivem menšího počtu vážených vrstev výrazně nižší.



Obr. 2.9: Porovnání nejúspěšnějších vnořených architektur.

Výsledné zhodnocení

Jako nejúspěšnější byla zhodnocena metoda navržená pro biomedicínské aplikace - Unet, jejíž architektura a princip byly podrobněji shrnuty v teoretické části práce. Pro tuto architekturu bylo hledáno nejoptimálnější předzpracování obrazu spolu s ideálními parametry pro učení sítě. V nejlepším případě se při zhodnocení jednotlivých snímků Diceův koeficient podobnosti pohyboval přibližně kolem hodnoty 93 %, což svědčí o dosažení vysoce kvalitních výsledků. Robustnost metody je zajištěna především rozšířením učební databáze v podobě navržených transformací.

Metoda MaskRCNN, do které byly vkládány vzhledem k jejímu nedávnému navržení velké naděje dosáhla výsledků o poznání horších. Přesto, že byly s jistotou nalezeny oblasti výskytu optického disku v obraze, samotný Diceův koeficient podobnosti ručně vytvořené masky a predikované segmentační mapy dosahoval hodnoty 86,72 %.

2.4 Segmentace cév sítnice

Pro segmentaci cév sítnice bylo opět testováno několik různých architektur, přičemž nejlepších výsledků bylo dosaženo při použití metody s architekturou Linknet a stejně jako u optického disku také s architekturou Unet. Postup navržené metody je znázorněn v blokovém schématu na obrázku 2.10.



Obr. 2.10: Blokové schéma segmentace cév sítnice.

2.4.1 Předzpracování obrazů pro segmentaci cév

Při segmentaci cév sítnice se na základě dostupné literatury jako nejlepší volba pro převod do šedotónového obrazu vybírá v mnoha studiích z každého obrazu jeho zelená složka. Ta nese o hledaných cévách největší množství informací.

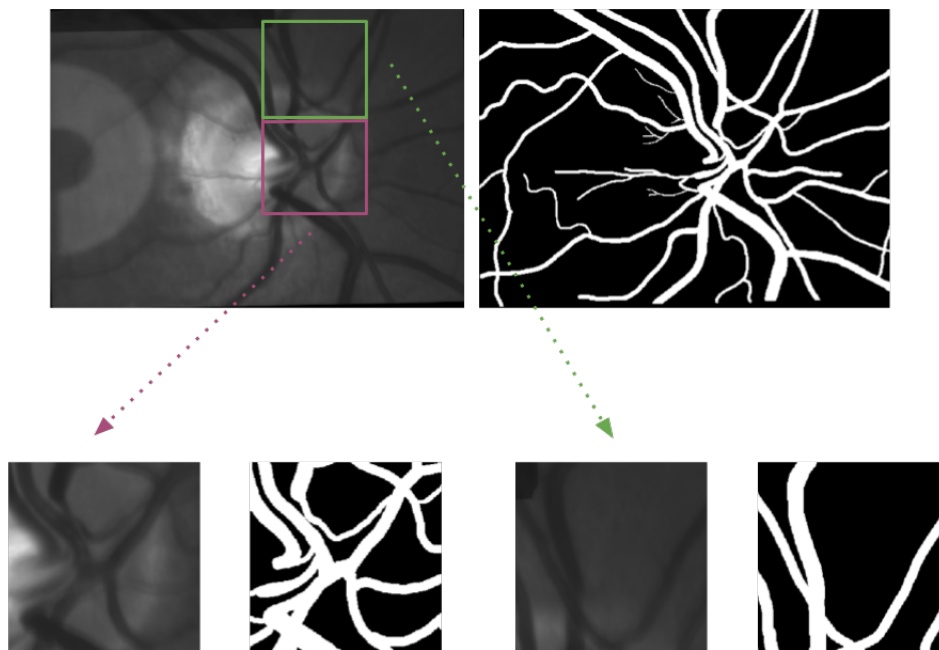
Stejně jako v případě segmentace optického disku i zde může být v některých případech vhodně obraz dále zbavit nadbytečného šumu využitím například mediánového filtru, případně zvýšit jeho kontrast úpravou histogramu.

V případě segmentace cév se v prováděném výzkumu nicméně rovněž osvědčilo využití všech barevných kanálů. Naopak ekvalizace histogramu oproti předzpracování při segmentaci optického disku způsobuje v tomto případě zvýraznění cév a jejich hran, které lze následně detekovat přesněji.

Ekvalizace je zprostředkována za pomoci knihovny *OpenCV* a její metody *cv2.createCLAHE*. Obraz je nejprve převeden do jiného barevného prostoru, konkrétně prostoru *LAB*. Zde informaci o nerovnoměrném osvětlení v obraze nese složka *L* (lightness). Využitá metoda provede adaptivní ekvalizaci histogramu ve složce *L*, následuje opětovné složení barevných kanálů dohromady a převedení ekvalizovaného obrazu do známého barevného prostoru *RGB*.

Jelikož pro segmentaci cév sítnice je využita také architektura konvoluční neuronové sítě Unet, následné zpracování je stejné jako v případě předzpracování optického disku a na každý obraz je aplikována metoda *preprocess_input*. Následuje opět rozšíření trénovací množiny a normalizace hodnot jasu do rozmezí hodnot 0 - 1.

Rozdíl oproti segmentaci optického disku může nastat v úpravě obrazu na požadovaný vstupní rozměr sítě. V případě segmentace cév sítnice je vhodné zachovat co nejlepší rozlišení obrazu pro detekci velmi jemných cév. V této práci byly proto vyzkoušeny dva přístupy. První přístup zachová původní rozlišení a obraz spolu s odpovídající maskou je rozdělen na podoblasti o velikosti 256×256 pixelů, které jsou přivedeny na vstup sítě. Příklad podoblastí je na obrázku 2.11. Druhý přístup ořízne obraz na nejvyšší možné rozlišení, tedy stejně jako u optického disku na rozměry



Obr. 2.11: Rozdělení trénovacího obrazu na podoblasti.

768 × 768 pixelů. Jelikož v případě rozdělení obrazu na jeho podoblasti docházelo k výrazným přechodovým jevům na okrajích jednotlivých podoblastí, bylo i v tomto případě kvalitnějších výsledků dosaženo s využitím druhého přístupu.

2.4.2 Metoda Unet

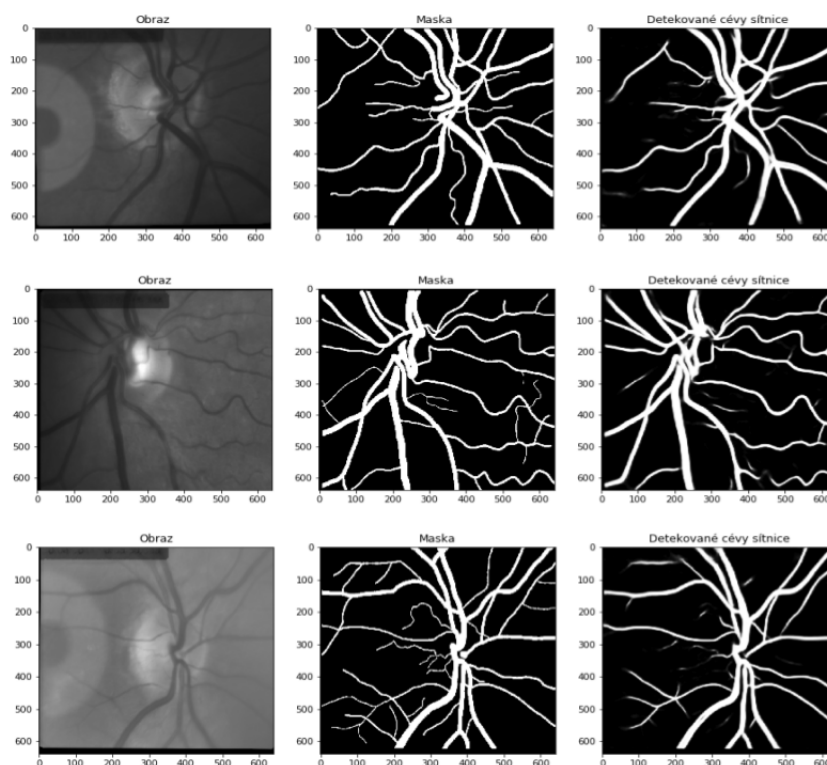
Pro samotnou segmentaci cév sítnice byla jako první využita stejná architektura jako pro segmentaci optického disku, architektura Unet.

Sestupnou část rovněž tvoří konvoluční neuronová síť *Resnet18* předučená na obrazové databázi *ImageNet*, s aktivací funkcí *sigmoidou* na jejím konci.

Chybová funkce zodpovědná za správné učení sítě je počítána jako *Diceův koeficient podobnosti* a nejlepší optimalizační algoritmus je i v tomto případě adaptivní optimalizace - *Adam*.

Množství epoch s nejlepším výsledkem je 150.

Příklady výsledné segmentace cév sítnice s využitím architektury Unet jsou ukázány na obrázku 2.12.



Obr. 2.12: Detekované cévy za pomoci metody Unet.

2.4.3 Metoda Linknet

Další metoda poskytující kvalitní výsledky byla založena na konvoluční neuronové síti Linknet.

Nejlepších výsledků bylo dosaženo se stejně nastavenými parametry jako u sítě Unet - sestupná část s konvoluční neuronovou sítí *Resnet18* předtrénovaná na databázi *ImageNet* se *sigmoidou* jako aktivační funkcí na závěr. Snižování chybové funkce v podobě *Diceova koeficientu podobnosti* je zajištěno optimalizačním algoritmem *Adam*. Rozdíl nastává pouze v počtu epoch, při níž je dosaženo nejlepších výsledků. Protože architektura je založena na sníženém počtu parametrů, trvá její konvergence k optimu delší dobu a nejlepšího výsledku bylo dosaženo s počtem epoch 200.

2.4.4 Zhodnocení použitých metod

Pro zhodnocení kombinace různých parametrů byly i v případě segmentace cév sítnice využity metriky senzitivita, specifita a Diceův koeficient podobnosti. Rozdíl je pouze v zařazování pixelů, kdy jednu třídu tvoří pixely v oblastech výskytu cév a druhou opět pixely pozadí.

Segmentační architektura

Opět bylo hlavním cílem nalézt vhodnou segmentační architekturu dosahující nejlepších výsledků. Bylo proto provedeno několik experimentů pro architektury PSPNet, Linknet a Unet. Výsledné hodnoty jsou v tabulce 2.4.4.

Tab. 2.6: Úspěšnost segmentačních architektur.

	Senzitivita	Specifita	Diceův koeficient
PSPNet	81,37 %	85,82 %	61,08%
Unet	78,09 %	94,26 %	81,93 %
Linknet	77,32 %	94,68 %	82,68 %

Z tabulky je patrné, že v případě segmentace cév sítnice dosahují vzhledem k náročnosti úlohy metody výrazně nižší úspěšnosti. Nejnižší úspěšnosti je opět dosaženo při využití konvoluční sítě PSPNet, kde Diceův koeficient dosahuje pouze 61,08 %.

V porovnání s výsledky segmentace optických disků je Diceův koeficient u architektur Unet a Linknets nejlepšími výsledky nižší dokonce o více než 10 %. Protože metody s využitím architektur Unet a Linknet dosahují velice podobných výsledků, byla dále upravována metoda použitá i pro segmentaci optického disku - Unet.

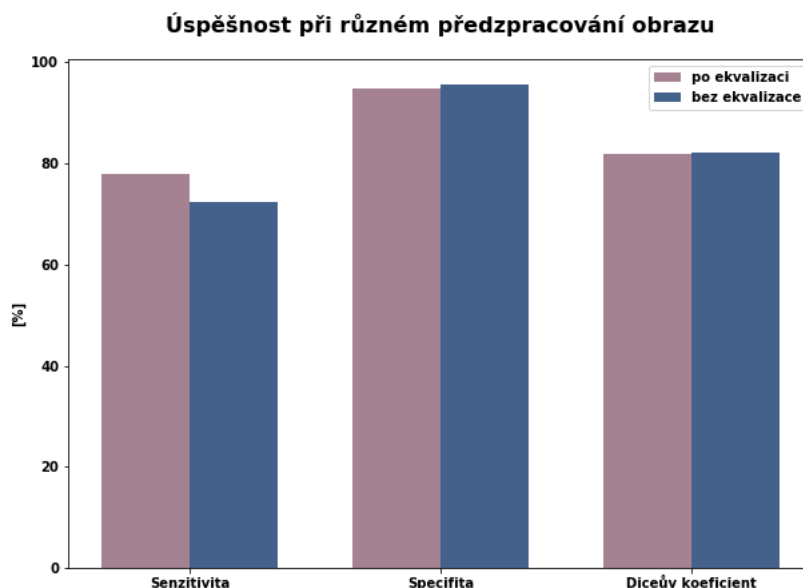
Ekvalizace histogramu

Oproti segmentaci optického disku se jako jeden z důležitých kroků jevílo předzpracování obrazu v podobě ekvalizace histogramu, kdy jsou zvýrazněny cévy v obraze. Porovnání úspěšnosti v případě předzpracování a bez něj je znázorněno v grafu 2.13 a hodnoty jsou shrnuty v tabulce 2.4.4.

Tab. 2.7: Úspěšnost na základě předzpracování obrazu.

	Senzitivita	Specifita	Diceův koeficient
Bez ekvalizace	72,32 %	95,70 %	82,20 %
S ekvalizací	78,01 %	94,68 %	81,93 %

Z výsledků je patrné, že ekvalizace se nejvíce projevila při hodnocení specifity, Diceův koeficient, naopak poklesl o 0,73 %. Vzhledem ke všem porovnávaným netrikám je výsledná metoda přizpůsobena na vstup obrazů po ekvalizaci histogramu.



Obr. 2.13: Graf úspěšnosti pro různé předzpracování obrazu.

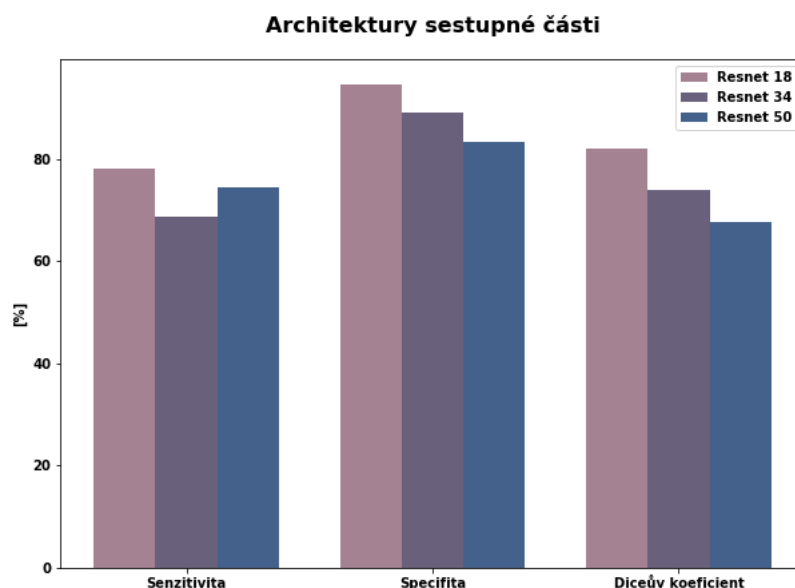
Vnořená architektura sestupné části

Důležitým parametrem byla při zhodocení algoritmu pro segmentaci cév vnořená architektura pro extrakci charakteristických znaků obrazů. Výrazný rozdíl se ukázal při vyzkoušení hlubších architektur sestupné části. Protože nejlepších výsledků bylo i v tomto případě dosaženo s architekturou Resnet, byla proto vyzkoušena ve variantách s 18, 34 a 50 vrstvami. Výsledky popisuje graf 2.14 a tabulka 2.4.4.

Tab. 2.8: Úspěšnost různě hlubokých architektur sestupné části Resnet.

	Senzitivita	Specifita	Diceův koeficient
Resnet 50	74,47 %	83,44 %	67,75 %
Resnet 34	68,80 %	89,14 %	73,94 %
Resnet 18	78,01 %	94,68 %	81,93 %

Z dosažených výsledků je patrné, že zvyšováním hloubky sítě dochází v případě segmentace cév k výraznému zhoršení úspěšnosti, u Diceova koeficientu podobnosti jde rozdíl mezi architekturou s 18 vrstvami a s 50 vrstvami až 14,18 %.



Obr. 2.14: Graf úspěšnosti vnořených architektur sestupné části

Výsledné zhodnocení

Pro segmentaci cév sítnice byl na základě provedených experimentů rovněž vybrán přístup s konvoluční neuronovou sítí Unet. Téměř identických výsledků bylo nicméně dosaženo i s využitím segmentační architektury sítě Linknet.

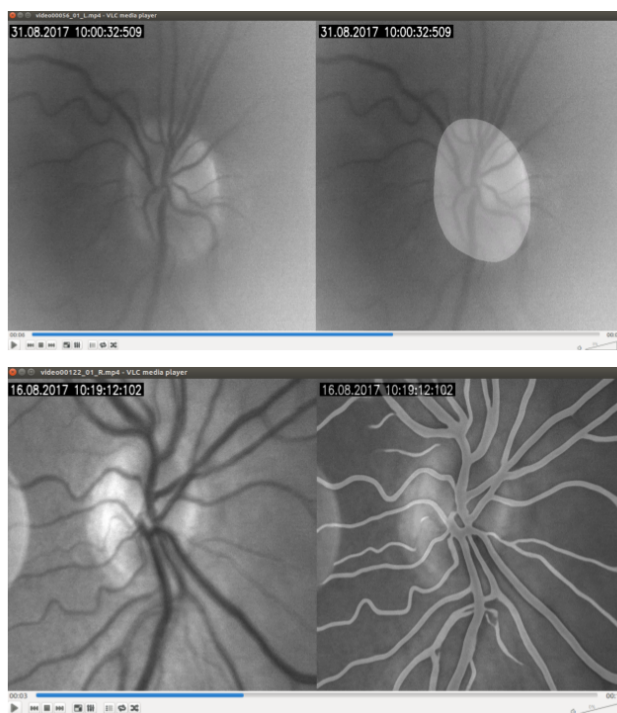
Největším rozdílem v porovnání s metodou implementovanou pro segmentaci optického disku je předzpracování obrazu v podobě ekvalizace histogramu, které napomáhá úspěšnosti metody.

Nejllepší dosažená úspěšnost v podobě Diceova koeficientu podobnosti po implementaci navržené metody dosahuje hodnot přibližně kolem 82 %, je tedy o více než 10 % méně úspěšná než metoda segmentující optický disk. Jelikož segmentace cév představuje díky jejich rozvětvenosti a často nevýraznému kontrastu především jemnějších cév v porovnání s pozadím velice náročný úkol, je tento výsledek očekávaný. Velkou roli může při provádění zhodnocení hrát také ruční tvorba trénovací databáze, kdy nepřesné označení nekонтastních cév vede k tvorbě nepřesné masky a horším výsledkům.

2.5 Oftalmologické sekvence

Cílem diplomové práce byla segmentace optického disku a sév sítnice v oftalmologických sekvencích. Implementovaná metoda byla naučena na snímcích vytvořených na

základě těchto sekvencí. Posledním krokem je tedy aplikovat navržený přístup přímo na oftalmologické sekvence, kdy výsledkem jsou videa s nalezenými optickými disky a cévami sítnice - 2.15.



Obr. 2.15: Segmentovaný optický disk a cévy v oftalmologické sekvenci.

Příklady segmentovaných optických disků v sekvenci se nacházejí v repozitáři: https://github.com/AndreaR23/vessel_segmentation/tree/master/result_code/UNET/disks/disk_video a sekvencí se segmentovanými cévami sítnice: https://github.com/AndreaR23/vessel_segmentation/tree/master/result_code/UNET/vessels/vessels_video.

2.5.1 Zhodnocení metod na oftalmologických sekvencích

Stejné zhodnocení, které bylo provedeno pro jednotlivé zprůměrované obrazy získané z oftalmologických sekvencí bylo následně provedeno i na samotných sekvencích. Aby nebylo nutné anotovat ručně každý snímek, byly sekvence, na kterých zhodnocení probíhalo, registrovány. Registrace obrazu má za následek, že odpovídající struktury se ve všech snímcích sekvence vyskytují na stejné pozici. To je nutné zejména kvůli artefaktům způsobeným pohybem očí. Registrace byla prováděna ve dvou krocích na základě studie *Registration of retinal sequences from new video-ophthalmoscopic camera*. Základem je vhodný výběr referenčního snímku v sekvenci, následovaný vhodným přezpracováním za účelem získání rovnoměrnějšího kontrastu a osvětlení

v obraze. Referenční snímky jsou vybrány na základě míry entropie s ohledem na šum způsobený pohybem očí a odraz světla. V dalším přezpracování obrazů dochází ke snaze zvýšit poměr signálu vůči šumu pomocí mediánové filtrace a ekvalizace histogramu. Pro posun obrazu v rámci registrace byla dále využita fázová korelace. [31, 16]

Druhý krok registrace slouží pro zpřesnění výsledků kroku prvního. Na základě Hessovy matice a prahování Otsu je segmentován obraz, ve kterém jsou vyznačeny kontrolní body cévách. Po dalších úpravách byla použita binární skeletonizace pro nalezení sledovaných kontrolních bodů pro registraci. Pro nalezení transformační matice byl využit odhad euklidovské vzdálenosti. [31, 16]

Zhodnocení navržených metod bylo provedeno pro optické disky na 6 sekvencích, z nichž každá byla složena z 253 snímků. Celková validační databáze tedy činí 1518 obrazů. Pro segmentované cévy bylo zhodnocení provedeno na 5 sekvencích, rovněž s 253 snímky v jedné, validační databáze tedy obsahuje 1256 snímků.

Tab. 2.9: Velikost databáze pro zhodnocení sekvencí.

	Optický disk	Cévy sítnice
Počet sekvencí	6	5
Počet snímků na sekvenci	253	253
Celkový počet snímků	1518	1265

Segmentace optického disku

Zhodnocení segmentace optického disku bylo provedeno s využitím nejlépe pracující metody s konvoluční neuronovou sítí Unet. Dosažené výsledky pro jednotlivé registrované sekvence spolu s průměrnou celkovou hodnotou jsou zobrazeny v tabulce 2.5.1.

Tab. 2.10: Úspěšnost segmentace optického disku v sekvencích.

	Senzitivita	Specifita	Diceův koeficient
Sekvence č. 1	94.35 %	98,61 %	95,82 %
Sekvence č. 2	91.31 %	98,91 %	93,93 %
Sekvence č. 3	97.13 %	98,72 %	98,02 %
Sekvence č. 4	93.77 %	98,76 %	96,21 %
Sekvence č. 5	93.25 %	98,26 %	93,19 %
Sekvence č. 6	88.11 %	98,93 %	94,80 %
Průměr	92,99 %	98,70 %	95,33 %

Diceův koeficient podobnosti, který se ve velké míře využívá ke zhodnocování výsledků segmentace dosahuje hodnoty 95,33 %. Kvalitní výsledky potvrzují i vysoké hodnoty senzitivity - 92,99 % a specifity - 98,7 %.

Segmentace cév sítnice

I pro segmentaci cév sítnice byla pro zhodnocování vybrána metoda s konvoluční neuronovou sítí Unet. Dosažené výsledky pro jednotlivé registrované sekvence spolu s průměrnou celkovou hodnotou jsou zobrazeny v tabulce 2.5.1.

Tab. 2.11: Úspěšnost segmentace cév sítnice v sekvencích.

	Senzitivita	Specifita	Diceův koeficient
Sekvence č. 1	60,84 %	91,52 %	67,32 %
Sekvence č. 2	65,66 %	91,01 %	69,89 %
Sekvence č. 3	56,03 %	91,91 %	66,38 %
Sekvence č. 4	64,98 %	90,58 %	69,45 %
Sekvence č. 5	61,82 %	91,49 %	69,76 %
Průměr	61,87 %	91,30 %	68,56 %

Při testování implementované metody na segmentaci cév v oftalmologických sekvencích byly dosaženy výsledky horší než v případě zhodnocení pro jednotlivé snímky. Hodnota Diceova koeficientu podobnosti například dosáhla pouze hodnoty 68,56 %, což je přibližně o 13,4 % horší výsledek než při segmentaci ze samotných snímků. V tomto případě může být výsledek do jisté míry ovlivněný menší kvalitou jednotlivých snímků v sekvenci oproti trénovacím snímkům, kde každý vznikl na základě několika ostatních.

2.6 Zhodnocení dosažených výsledků

V diplomové práci bylo porovnáno několik přístupů založených na konvolučních neuronových sítích pro segmentaci obrazu.

Pro **segmentaci optického disku** byly vyzkoušeny metody založené na architekturách PSPNet, Linknet, Unet a MaskRCNN. Nejslabších výsledků bylo dosaženo s architekturou PSPNet, kdy i po kombinování a obměňování parametrů sítě spolu s různými variantami předzpracování obrazu bylo dosaženo velice nízkých hodnot - senzitivita 75,33 %, specifita 98,01 % a Diceův koeficient podobnosti 85,68 %.

Lepší výsledky poskytovala metoda založená na konvoluční síti MaskRCNN, která nejprve vyhledá oblasti zájmu v obraze a v okamžiku, kdy je obsah této oblasti klasifikován jako optický disk, dochází k segmentaci této podoblasti. Zde byly

spolehlivě nalezeny oblasti s optickým diskem, samotná segmentace však dosahovala nižších hodnot než následující architektury - senzitivita 86,24 %, specifita 89,20 % a Diceův koeficient podobnosti 85,72 %.

Velice podobných výsledků bylo dosaženo s využitím metod s architekturami Linknet a Unet. O něco slabší architektura Linknet dosahovala hodnot senzitivity 90,67 %, specifity 98,99 % a Diceova koeficientu podobnosti 92,34 %. Pomocí zmíněné segmentační architektury Unet, která byla původně navržena přímo pro segmentaci biomedicínských obrazů, bylo v případě segmentace optického disku dosaženo senzitivity 92,75 %, specifity 98,78 % a Diceova koeficientu podobnosti 92,86 %, což svědčí o vysoké úspěšnosti a robustnosti tohoto přístupu.

Výsledků bylo dosaženo za pomoci extrakce pomocí vnořené sítě Resnet 18, pro kterou bylo také provedeno požadované předzpracování obrazů. Výpočet chybové funkce byl zprostředkován pomocí Diceova koeficientu podobnosti a její optimalizace algoritmem Adam. Celá síť byla učena na 150 epochách, kdy doba učení se s využitím GPU pohybovala v rozmezí 3,5 až 4,5 hodiny.

Síť natrénovaná s těmito parametry byla následně využita pro segmentaci v oftalmologických sekvencích. Při zhodnocení, které probíhalo na registrovaných sekvencích, bylo dosaženo dokonce vyšší úspěšnosti než u samotných zprůměrovaných obrazů - senzitivita 92,99 %, specifita 98,7 % a Diceův koeficient podobnosti 95,33 %. Důvodem dosažení lepších hodnot mohla být menší variabilita při zhodnocování obrazů pouze z 6 sekvencí.

Při **segmentaci cév sítnice** byly srovnávány stejné architektury, pouze architektura MaskRCNN vycházející z vyhledávání oblastí zájmu vzhledem k její nevhodnosti pro tento úkol uvažována nebyla.

I zde dosahovala nejhorších výsledků metoda využívající segmentační architekturu PSPNet. Její výsledné hodnoty dosahovaly hodnot - senzitivita 81,37 %, specifita 85,82 % a Diceův koeficient podobnosti 61,08 %.

Výsledky metod Linknet a Unet byly téměř srovnatelné, metoda Linknet dosáhla výsledků - senzitivita 77,32 %, specifita 94,68 % a Diceův koeficient podobnosti 82,68 % a metoda se sítí Unet potom dosáhla hodnot senzitivity 78,09 %, specifity 94,26 % a Diceova koeficientu podobnosti 81,93 %. Vzhledem k využití hodnoty Unet pro segmentaci optického disku byla i zde zvolena metoda Unet pro další analýzu.

Nejlepší výsledky byly získány při učení na předzpracovaných obrazech pro využitou vnořenou architekturu Resnet18 zodpovědnou za extrakci důležitých charakteristik v obrazech. U obrazů dále proběhla v rámci předzpracování ekvalizace histogramu. Chybová funkce byla zprostředkována stejně jako v případě segmentace optického disku výpočtem Diceova koeficientu podobnosti a úprava za pomoci optimalizačního algoritmu Adam. Výsledný počet epoch pro trénování sítě byl i v tomto

případě 150. Doba učení pravděpodobně vlivem předzpracování obrazu při načítání obrazů generátorem stoupla na rozmezí 5,5 až 6,5 hodiny při využití GPU.

Za pomoci připravené natrénované sítě s architekturou Unet se zmíněnými parametry bylo provedeno zhodnocení i na registrovaných oftalmologických sekvencích. Výsledky, kterých bylo dosaženo, byly v porovnání s úspěšností segmentace cév u samotných obrazů výrazně horší - senzitivita 61,87 %, specifita 91,30 % a Diceův koeficient podobnosti 68,56 %. Jako jeden z hlavních důvodů této nízké úspěšnosti může být skutečnost, že jednotlivé snímky v sekvenci mají nižší kvalitu oproti zprůměrovaným snímkům v učební a validační databázi. Při náročném úkolu jakým je segmentace cév sítnice může hrát tento rozdíl výraznou roli. Velký vliv může mít stejně jako u segmentace ve zprůměrovaných obrazech manuální anotace obrazů. Tamůže být především v oblastech tenkých cév nepřesná a zhoršit tak dosažené výsledky.

3 Závěr

Cílem diplomové práce byla implementace vhodných metod pro automatickou segmentaci optického disku a cév v oftalmologických sekvencích, protože právě na základě zhodnocení těchto struktur lze včasné diagnostikovat velkou řadu více či méně závažných onemocněních.

Práce je rozčleněna na teoretickou a praktickou část. V teoretické části práce jsou popsány využívané metody snímání očního pozadí za účelem jeho vyšetření. Následuje shrnutí metod využívaných v problematice segmentace obrazu. Na závěr úvodní teoretické části je zařazeno spojení segmentačních technik a problému s diagnostikou očních chorob a je popsáno několik přístupů historicky i současně využívaných pro automatickou detekci požadovaných struktur, kterými jsou optický disk a cévy sítnice.

V praktické části jsou popsány a zhodnoceny metody navržené a implementované za účelem segmentace v rámci diplomové práce. Pro segmentaci optického disku bylo nejlepších výsledků dosaženo s využitím metody Unet. O tři roky novější metoda MaskRCNN dosahovala o poznání horších výsledků pro segmentaci disků, spolehlivě však našla oblast, ve které se optický disk nacházel. Horší výsledky mohly být způsobeny například jiným způsobem načítání dat pomocí generátorů, které metoda využívá, v porovnání s generátory vlastními vytvořenými konkrétně za účelem segmentace optického disku a cév.

I výsledná segmentace cév byla implementována s využitím konvoluční neuronové sítě Unet, i když velice podobných výsledků bylo dosaženo i s architekturou Linknet. U segmentace cév se jako stěžejní ukázala volba architektury zodpovědné za extrakci příznaků v sestupné části algoritmu a předzpracování vstupních obrazů.

V poslední části je prováděno zhodnocení v oftalmologických sekvencích, ze kterých byly obrazy použité pro učení sítí získány. V případě segmentace optického disku dosahují hodnoty pozorovaných metrik při zhodnocení na oftalmologických sekvencích lepších výsledků, což může být způsobeno nižší variabilitou testované databáze. Naopak v případě zhodnocení segmentace cév sítnice v sekvencích jsou výsledky výrazně horší. Zde může hrát roli horší kvalita jednotlivých snímků v sekvenci oproti zprůměrovaným obrazům využívaným při implementaci segmentační metody.

Kódy s implementovanými metodami v praktické části se nachází v repozitáři: https://github.com/AndreaR23/vessel_segmentation.

Literatura

- [1] ABDULLAH, Muhammad, Muhammad Moazam FRAZ a Sarah A. BARMAN. Localization and segmentation of optic disc in retinal images using Circular Hough transform and Grow Cut algorithm. *PeerJ*. 2016, **4**. ISSN 2167-8359. Dostupné také z: <https://peerj.com/articles/2003>
- [2] AJALA, Funmilola A., O. A. OKE, T. O. ADEDEJI, O. M. ALADE a E. A. ADEWUSI. Fuzzy k-c-means Clustering Algorithm for Medical Image Segmentation. *Journal of Information Engineering and Applications*. 2012, **2**(6), 21-32.
- [3] ALMAZROA, Weiwei SUN, Sami ALODHAYB, Kaamran RAAHEMIFAR a Vasudevan LAKSHMINARAYANAN. Optic disc segmentation for glaucoma screening system using fundus images. *Clinical Ophthalmology*. Dovepress, 2017, (11), 2017-2029.
- [4] ANJNA, Er. a Er.Rajandeep KAUR. Review of Image Segmentation Technique. *International Journal of Advanced Research in Computer Science*. 2017, **8**(4), 36-39.
- [5] ATSBURY, Nick. *Clinical Techniques in Ophthalmology*. China: Churchill Livingstone, 2006.
- [6] BADRINARAYANAN, Vijay, Alex KENDALL a Roberto CIPOLLA. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017, **39**(12), 2481-2495.
- [7] BEERS, Floris van, Arvis LINDSTRÖM, Emmanuel OKAFOR a Marco A. WIERING. Deep Neural Networks with Intersection over Union Loss for Binary Image Segmentation. *International Conference on Pattern Recognition Applications and Methods*. Bernoulli Institute, Department of Artificial Intelligence: University of Groningen, 2019.
- [8] CERMAK, Marek. *Detekce optického disku v sériích snímků z video oftalmoskopu*. Brno, 2017. Diplomová práce. Vysoké učení technické v Brně. Vedoucí práce Jan Odstrčilík.
- [9] CIREŞAN, Dan C., Alessandro GIUSTI, Jürgen SCHMIDHUBER a Luca M. GAMBARDELLA. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. *Advances in Neural Information Processing Systems*. Curran Associates, 2012, (25), 2843-2851.

- [10] CHAURASIA, Abhishek a Eugenio CULURCIELLO. LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation. *IEEE Visual Communications and Image Processing*. 2017, 1-4.
- [11] DEVI, S. Saranya, K. I. RAMACHANDRAN a Ashish SHARMA. Retinal Vasculature Segmentation in Smartphone Ophthalmoscope Images. *7th WACBE World Congress on Bioengineering 2015*. Cham: Springer International Publishing, 2015, 2015, , 64-67. IFMBE Proceedings. ISBN 978-3-319-19451-6. Dostupné také z: http://link.springer.com/10.1007/978-3-319-19452-3_18
- [12] HOOVER, Adam a Michael GOLDBAUM. Locating the Optic Nerve in a Retinal Image Using the Fuzzy Convergence of the Blood Vessels. *Transactions of Medical Imaging*. 2003, **22**(8), 951-958.
- [13] CHEN, Guannan, Meizhu CHEN, Jichun LI a Encai ZHANG. Retina Image Vessel Segmentation Using a Hybrid CGLI Level Set Method. *BioMed Research International*. Hindawi, 2017.
- [14] KELLY, Zou H., Simon K. WARFIELD, Aditya BHARATHA, et al. Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *National Institutes of Health*. 2004, **11**(2), 178-189.
- [15] KINGMA, Diederik a Jimmy BA. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*. 2014, (12).
- [16] KOLAR, Radim, Ralf. P. TORNOW, Jan ODSTRČILIK a Ivana LIBERDOVA. Registration of retinal sequences from new video-ophthalmoscopic camera. *BioMedical Engineering OnLine*. 2016, **15**(1).
- [17] KOZUMPLÍK, Jiří a Ivo PROVAZNÍK. *Umělá inteligence v medicíně*. Brno, 2007. Skripta. Fakulta elektrotechniky a komunikačních technologií - Vysoké učení technické v Brně.
- [18] KRIZHEVSKY, Alex, Ilya SUTSKEVER a Geoffrey E. HINTON. ImageNet Classification with Deep Convolutional Neural Networks. *NIPS'12 Proceedings of the 25th International Conference on Neural Information*. 2012, **1**, 1097-1105.
- [19] LI, Meng, Zhenshen MA, Chao LIU, Guang ZHANG a Zhe HAN. Robust Retinal Blood Vessel Segmentation Based on Reinforcement Local Descriptions. *BioMed Research International*. Hindawi, 2017, 1-9.
- [20] MOHAMMED, Hind Rustum, Husein Hadi ALNOAMANI a Ali AbdulZahraa JALIL. Improved Fuzzy C-Mean Algorithm for Image Segmentation. *International Journal of Advanced Research in Artificial Intelligence*. 2016, **5**(6), 7-10.

- [21] REN, Shaoqing, Kaiming HE, Ross GIRSHICK a Jian SUN. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Microsoft Research*. 2017, (39), 1137-1149.
- [22] RONNEBERGER, Olaf, Philipp FISCHER a Thomas BROX. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2015, (9351).
- [23] RUST, Caterina, Stephanie HÄGER, Nadine TRAULSEN a Jan MODERSITZKI. A robust algorithm for optic disc segmentation and fovea detection in retinal fundus images. *Current Directions in Biomedical Engineering*. De Gruyter, 2017, **3**(2), 533-537.
- [24] SEVASTOPOLSKY, Artem. *Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network*. Moscow, 2017. Department of Mathematical Methods of Forecasting, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University.
- [25] SEVERA, David. *Oftalmologická zařízení pro vyšetřování sítnice*. Brno, 2009. Bakalářská práce. Masarykova univerzita - Lékařská fakulta. Vedoucí práce Jan Richter.
- [26] SHARMA, Neeraj a Lalit M. AGGARWAL. Automated medical image segmentation techniques. *Journal of Medical Physics*. 2009, **35**(1), 3-14.
- [27] SHELHAMER, Evan, Jonathan LONG a Trevor DARRELL. Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2015, 2015, , 3431-3440. ISBN 978-1-4673-6964-0. Dostupné také z: <http://ieeexplore.ieee.org/document/7298965/>
- [28] SUGANYA a MENAKA. Various Segmentation Techniques in Image Processing: A Survey. *International Journal of Innovative Research in Computer and Communication Engineering*. 2014, **2**(1), 1048-1052.
- [29] TAHA, Abdel, Aziz a Allan HANBURY. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*. 2015, , 1-28.
- [30] TIMBERLAKE, George T. a Michael KENNEDY. The Direct Ophthalmoscope How it Works and How to Use It. *University of Kansas Medical Center*. University of Kansas, 2005, , 1-39.

- [31] TORNOW, Ralf-Peter, Jan ODSTRČILÍK a Radim KOLÁŘ. Time-resolved quantitative inter-eye comparison of cardiac cycle-induced blood volume changes in the human retina. *BIOMEDICAL OPTICS EXPRESS*. 2018, **9**(12).
- [32] YAMASHITA, Rikiya, Mizuho NISHIO, Richard KINH GIAN DO a Kaori TOGASHI. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*. Springer, 2018, (9), 611-629.
- [33] YAN, Zengqiang, Xin YANG a Kwang-Ting Tim CHENG. A Three-stage Deep Learning Model for Accurate Retinal Vessel Segmentation. *IEEE Journal of Biomedical and Health Informatics*. 2018, , 1-1. ISSN 2168-2194. Dostupné také z: <https://ieeexplore.ieee.org/document/8476171/>
- [34] YUHENG, Song a Yan HAO. Image Segmentation Algorithms Overview. *ArXiv*. 2017, (abs/1701.02051).
- [35] WALEK, Petr, Martin LAMOŠ a Jiří JAN. *Analýza biomedicínských obrazů počítačová cvičení FEKT VUT v Brně*. Vysoké učení technické v Brně, 2015.
- [36] WANG, Chuang. *Level Set Segmentation of Retinal Structures*. London, 2016. Department of Computer Science Brunel University London.
- [37] Adam - latest trends in deep learning optimization. *Towards Data Science* [online]. [cit. 2019-04-18]. Dostupné z: <https://towardsdatascience.com/adam-latest-trends-in-deep-learning-optimization-6be9a291375c>
- [38] A Gentle Introduction to Transfer Learning for Deep Learning. *Machine Learning Mastery* [online]. [cit. 2019-01-01]. Dostupné z: <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>
- [39] ZHAO, Hengshuang, Jianping SHI, Xiaojuan QI, Xiaogang WANG a Jiaya JIA. Pyramid Scene Parsing Network. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, **1**, 6230-6239.
- [40] CS231n Convolutional Neural Networks for Visual Recognition *CS231n Convolutional Neural Networks for Visual Recognition* [online]. [cit. 2018-12-26]. Dostupné z: <http://cs231n.github.io/convolutional-networks/#layers>
- [41] CS231n Convolutional Neural Networks for Visual Recognition *CS231n Convolutional Neural Networks for Visual Recognition* [online]. [cit. 2019-01-01]. Dostupné z: <http://cs231n.github.io/transfer-learning/>

- [42] ENEE 631 Web Journal - Sriram Vasudeam. *Laplacian of Gaussian operator, Directional smoothing filter and Morphological* [online]. [cit. 2018-11-18]. Dostupné z: <https://sriramvasu.wordpress.com/2017/03/01/laplacian-of-gaussian-operator-directional-smoothing-filter-and-morphological/>
- [43] Glaucoma Tests and Diagnosis - Glaucoma Associates of Texas. *Glaucoma Associates of Texas - Glaucoma Eye Doctor, Speciality Ophtalmologists* [online]. [cit. 2018-12-26]. Dostupné z: <http://glaucomaassociates.com/glaucoma-tests-and-diagnosis/>
- [44] Gradient Descent - ML Cheatsheet documentation. *Machine Learning Cheatsheet - ML Cheatsheet documentation* [online]. [cit. 2018-12-26]. Dostupné z: https://ml-cheatsheet.readthedocs.io/en/latest/gradient_descent.html
- [45] Ophtalmoscopy Pathologies. *Welch Allyn - Medical Diagnostic Device Manufacturer* [online]. [cit. 2018-12-26]. Dostupné z: <https://www.welchallyn.com/en/students/ophtalmoscopy-pathologies.html>
- [46] Loss Functions - ML Cheatsheet documentation. *Machine Learning Cheatsheet - ML Cheatsheet documentation* [online]. [cit. 2019-04-19]. Dostupné z: https://ml-cheatsheet.readthedocs.io/en/latest/loss_functions.html
- [47] Loss Functions and Optimization Algorithms - XpertUp. *XpertUp - Let's Expert Up* [online]. [cit. 2018-12-26]. Dostupné z: <http://www.xpertup.com/2018/05/11/loss-functions-and-optimization-algorithms/>
- [48] Loss and Loss Functions for Training Deep Learning Neural Networks. *Machine Learning Mastery* [online]. [cit. 2019-04-19]. Dostupné z: <https://machinelearningmastery.com/loss-and-loss-functions-for-training-deep-learning-neural-networks/>
- [49] *Vantage Plus* [online]. In: . [cit. 2018-11-18]. Dostupné z: <http://www.spiritmedical.cz/cs/ocni-ambulance/oftalmoskopy/vantage-plus.html>

Seznam symbolů, veličin a zkratek

GPU	Graphics Processing Unit
ReLU	Rectified Linear Unit
IoU	Intersection Over Union
SGD	Stochastic Gradient Descent
Adam	Adaptive Moment Estimation
SVM	Support Vector Machines
PSPNet	Pyramid Scene Parsing Network
RCNN	Region Convolutional Neural Network

A Obsah přiloženého CD

MaskRCNN	Repozitář s MaskRCNN
classes	Vlastní třídy pro architekturu MaskRCNN
inference_config.py	Inicializační konfigurační třída
initial_config.py	Zhodnocující konfigurační třída
mask_dataset.py	Třída pro načítání masek
metrics.py	Metriky pro zhodnocení
models	Modely využité pro zhodnocení
mrcnn	Složka s kódy původního repozitáře
notebooks	Implementace metody Mask RCNN
MaskRCNN_disk_train.ipynb	Postup trénování sítě.
MaskRCNN_display_result.ipynb	Výsledky a zhodnocení
UNET	Architektura UNET
classes	Vlastní třídy pro architekturu UNET
create_patches.py	Rozdělení obrazu na podobrazy
data_generator.py	Generátor pro načítání dat
image_utils.py	Metody pro práci s obrazy
metrics.py	Metriky pro zhodnocení
model_utils.py	Metody pro práci s modely
process_results.py	Metody pro zpracování výsledků
video_parser.py	Metody pro zpracování sekvencí
visualize.py	Metody pro zobrazení výsledků
disks	Segmentace optických disků
disk_video	Výsledné sekvence
Disk_from_video.ipynb	Segmentace v sekvencích
Evaluate_registred_video_disk.ipynb	Sekvence
Evaluation_disk_models.ipynb	Porovnání modelů
Postprocess_disk_results.ipynb	Zpracování výsledků
UNET_disks_train.ipynb	Trénování
models	Výsledné modely
vessels	Segmentace cév sítnice
vessels_video	Videa s cévami sítnice
Evaluate_registred_video_vessels.ipynb	Sekvence
Evaluation_vessels_models.ipynb	Porovnání modelů
Postprocess_vessels_results.ipynb	Zpracování výsledků
UNET_vessels_train.ipynb	Trénování
data	Data
reg_disks	Registrované sekvence optických disků
reg_vessels	Registrované sekvence cév sítnice
train_disks	Trénovací obrazy optických disků
train_vessels	Trénovací obrazy cév sítnice
diplomova_prace.pdf	PDF s textem práce
requirements.txt	Závislé knihovny
README.txt	Soubor s popisem obsahu CD