

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIZERTAČNÍ PRÁCE

Brno, 2020

Ing. JIŘÍ CHMELÍK



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 DETEKCE, SEGMENTACE A KLASIFIKACE
OBTÍŽNĚ DEFINOVATELNÝCH KOSTNÍCH
NÁDOROVÝCH LÉZÍ VE 3D CT DATECH**

METHODS OF DETECTION, SEGMENTATION AND CLASSIFICATION OF DIFFICULT TO DEFINE
BONE TUMOR LESIONS IN 3D CT DATA

DIZERTAČNÍ PRÁCE

DOCTORAL THESIS

AUTOR PRÁCE

AUTHOR

Ing. Jiří Chmelík

ŠKOLITEL

ACADEMIC ADVISOR

prof. Ing. Jiří Jan, CSc.

BRNO 2020

ABSTRAKT

Cílem této práce byl vývoj algoritmů pro detekci, segmentaci a klasifikaci obtížně definovatelných kostních metastatických nádorových lézí v oblasti páteře z obrazových CT dat. Pro tyto účely byla vytvořena patientská databáze anotovaná lékařskými experty. Postupně byly navrženy tři metody, z nichž první je založena na přepracování a kombinaci metod dostupných z předchozího řešení projektu, druhá je velmi rychlá varianta založená na fuzzy k-means shlukové analýze, třetí metoda využívá moderních metod strojového učení, konkrétně hluboké učení konvolučních neuronových sítí. Dále byla navržena modifikace upřesňující výsledky následnou meta-analýzou nalezených kandidátů lézí pomocí náhodného lesa. Dosažené výsledky byly objektivně vyhodnoceny a porovnány s výsledky algoritmů publikovaných jinými autory. Pro hodnocení byly použity dva přístupy: technický, voxelově založený a klinický, objektově založený. Dosažené výsledky byly následně zhodnoceny a diskutovány.

KLÍČOVÁ SLOVA

segmentace, klasifikace, detekce, nádor, léze, metastáza, obratel, páteř, CT, strojové učení, CNN, hluboké učení

ABSTRACT

The aim of this work was the development of algorithms for detection segmentation and classification of difficult to define bone metastatic cancerous lesions from spinal CT image data. For this purpose, the patient database was created and annotated by medical experts. Successively, three methods were proposed and developed; the first of them is based on the reworking and combination of methods developed during the preceding project phase, the second method is a fast variant based on the fuzzy k-means cluster analysis, the third method uses modern machine learning algorithms, specifically deep learning of convolutional neural networks. Further, an approach that elaborates the results by a subsequent random forest based meta-analysis of detected lesion candidates was proposed. The achieved results were objectively evaluated and compared with results achieved by algorithms published by other authors. The evaluation was done by two objective methodologies, technical voxel-based and clinical object-based ones. The achieved results were subsequently evaluated and discussed.

KEYWORDS

segmentation, classification, detection, tumour, lesion, metastasis, vertebra, spine, CT, machine learning, CNN, deep learning

CHMELÍK, Jiří. *Metody detekce, segmentace a klasifikace obtížně definovatelných kostních nádorových lézí ve 3D CT datech*. Brno, 2020, 124 s. Dizertační práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Školitel: prof. Ing. Jiří Jan, CSc.

PROHLÁŠENÍ

Prohlašuji, že svou dizertační práci na téma „Metody detekce, segmentace a klasifikace obtížně definovatelných kostních nádorových lézí ve 3D CT datech“ jsem vypracoval samostatně pod vedením školitele dizertační 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 autor uvedené dizertační práce dále prohlašuji, že v souvislosti s vytvořením této dizertační práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom 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 autora

PODĚKOVÁNÍ

Rád bych poděkoval školiteli prof. Ing. Jiřímu Janovi, CSc. za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci. Dále také děkuji odborným konzultantům MUDr. Petru Ouředníčkovi, PhD. a doc. MUDr. Ing. et Ing. Lukáši Lambertovi, PhD. za přínosné rady a nápady, jakož i za pomoc při získávání dat a jejich anotaci.

Obsah

Úvod	10
1 Výchozí stav řešení	14
1.1 Detekce, segmentace a klasifikace lézí	14
1.2 Přístupy k hodnocení úspěšnosti metod	18
2 Cíle práce	23
3 Přehled používaných metod	24
3.1 Extrakce obrazových příznaků	24
3.1.1 Statistické příznaky	24
3.1.2 Strukturální příznaky	26
3.1.3 Filtračně odvozené příznaky	27
3.1.4 Modelově založené příznaky	28
3.1.5 Tvarové příznaky	30
3.2 Klasifikační přístupy	31
3.2.1 Shluková analýza	32
3.2.2 Grafové metody	33
3.2.3 Náhodný les	34
3.2.4 AdaBoost algoritmus	35
3.2.5 Umělé neuronové sítě	36
3.2.6 Konvoluční neuronové sítě	36
4 Vlastní řešení	46
4.1 Pacientská databáze a expertní značení dat	46
4.2 Přepřacování metod předcházejícího projektu	50
4.2.1 Úprava statistických modelů	52
4.2.2 Snížení výpočetní náročnosti	53
4.2.3 Kombinace předchozích metod	55
4.2.4 Dosažené výsledky – přepracované verze původních metod	56
4.2.5 Závěr k výsledkům přepracovaných původních metod	59
4.3 Fuzzy k-means rychlá klasifikace lézí	60
4.3.1 Dosažené výsledky – fuzzy k-means	61
4.3.2 Závěr k metodě fuzzy k-means	62
4.4 Klasifikace lézí konvoluční neuronovou sítí	63
4.4.1 Předzpracování obrazových dat	64
4.4.2 Návrh původní architektury specializované sítě	65
4.4.3 Učení sítě a optimalizace parametrů	71

4.4.4	Interpretace neuronové sítě	73
4.4.5	Dosažené výsledky – konvoluční neuronová síť	74
4.4.6	Meta-analýza detekovaných lézí	76
4.4.7	Závěr k metodě využívající konvoluční neuronové sítě	78
4.5	Finální realizovaný algoritmus	79
4.6	Použité metody hodnocení úspěšnosti	80
4.6.1	Voxelově založená metodika	81
4.6.2	Objektově založená metodika	83
5	Výsledky a diskuze	86
5.1	Voxelové hodnocení	86
5.2	Objektové hodnocení	90
5.3	Porovnání s dalšími autory	94
5.4	Subjektivní hodnocení	99
5.4.1	Hodnocení na neanotovaných datech	99
5.5	Shrnutí dosažených výsledků	106
6	Závěr	110
	Literatura	111

Seznam obrázků

1	Ukázka kostních metastáz v obratli v CT obrazech	12
1.1	Vysvětlení 2D „ROI metodiky“ hodnocení	19
1.2	Vysvětlení 2D „bodové metodiky“ hodnocení	20
1.3	Vysvětlení 2D „bounding-box metodiky“ hodnocení	21
1.4	Vysvětlení 2D „objektové metodiky“ hodnocení	22
1.5	Vysvětlení 2D „voxelové metodiky“ hodnocení	22
4.1	Informace o patientských datech: dostupná CT data a jejich rozdělení podle počtu a velikostí anotovaných lézí	47
4.2	Informace o patientských datech: CT data s různými rekonstrukčními parametry	48
4.3	Ukázka softwaru „Marker“ pro expertní voxelové značení nádorových tkání v páteři	49
4.4	Karabicové grafy denzity jednotlivých tkání pro jednotlivé typy obratlů určené z DATABÁZE A	51
4.5	Ukázka segmentace a klasifikace nádorových tkání přepracovanými metodami	58
4.6	Ukázka segmentace a klasifikace nádorových tkání rychlou fuzzy k-means metodou	62
4.7	Blokové schéma navržené metody využívající CNN klasifikátor a RF meta-analýzu lézí	64
4.8	Blokové schéma předzpracování dat před vstupem do CNN	65
4.9	Rámcové schéma CNN klasifikátoru pro jeden klasifikovaný voxel	67
4.10	Detailní schéma navržené konvoluční neuronové sítě pro klasifikaci voxelů	67
4.11	Grafy průběhu učení	72
4.12	Ukázka učební množiny	73
4.13	Vizualizace naučených filtrů CNN	74
4.14	Vizualizace příznaků naučené CNN	75
4.15	Voxelové hodnoty ROC křivky pro jednotlivé klasifikované tkáně metodou CNN	77
4.16	Ukázka segmentace a klasifikace nádorových tkání	78
4.17	Ukázka hlavního grafického rozhraní	80
4.18	Ukázka grafického rozhraní pro zobrazení výsledků	81
4.19	Princip objektového hodnocení ve 2D pro práh překryvu	85
5.1	Porovnání dosažených průměrných výsledků podle voxelově založené metodiky	89
5.2	Porovnání dosažených výsledků podle objektově založené metodiky	91

5.3	Porovnání dosažených výsledků podle objektově založené metodiky: nízkodenzní léze	92
5.4	Porovnání dosažených výsledků podle objektově založené metodiky: vysokodenzní léze	93
5.5	Příklady výsledků segmentace a klasifikace tkání 1	101
5.6	Příklady výsledků segmentace a klasifikace tkání 2	102
5.7	Příklady výsledků segmentace a klasifikace tkání 3	103
5.8	Příklady výsledků segmentace a klasifikace tkání 4	104
5.9	Příklad výsledku snížení počtu falešně detekovaných lézí	105

Seznam tabulek

4.1	Porovnání výsledků původních a přepracovaných verzí dostupných metod (voxelově)	56
4.2	Dosažené výsledky kombinací metod (voxelově)	59
4.3	Tabulka parametrů rychlé klasifikační metody	61
4.4	Dosažené výsledky fuzzy k-means (voxelově)	63
4.5	Tabulka hyperparametrů vrstev konvoluční neuronové sítě	70
4.6	Tabulka globálních hyperparametrů konvoluční neuronové sítě	71
4.7	Dosažené výsledky CNN (voxelově)	76
5.1	Porovnání výsledků navržených metod s dalšími publikovanými algoritmy – nízkodenzní léze	97
5.2	Porovnání výsledků navržených metod s dalšími publikovanými algoritmy – vysokodenzní léze	98
5.3	Přehled výsledků nemodifikovaných metod	106
5.4	Přehled výsledků modifikovaných metod	106
5.5	Výsledek kombinace modifikovaných metod	107
5.6	Výsledek rychlé metody s využitím fuzzy k-means	107
5.7	Výsledek metody využívající konvoluční neuronové sítě	108
5.8	Výsledek metody využívající konvoluční neuronové sítě s meta-analýzou	108

Úvod

Výsledky dizertační práce jsou součástí komplexního projektu řešeného ve spolupráci s firmou Philips® divize Philips® Healthcare Eindhoven (Nizozemí), jehož cílem je vývoj pokročilých algoritmů a metod pro počítačové zpracování obrazových dat se zaměřením na analýzu skeletálních částí lidského těla s pomocí počítačové tomografie (CT¹). Momentálním zaměřením projektu je automatická analýza kostních metastatických lézí v oblastech páteře a sledování jejich odezvy na aplikovanou léčbu. Z tohoto hlediska bylo možné projekt rozdělit na dílčí části: první část řeší detekci, identifikaci a segmentaci obratlů z celotělových CT dat, a byla řešena v rámci dizertační práce a publikací Jakubicek (2020); Jakubicek a kol. (2020, 2019*a,b*); Jakubicek a kol. (2019*c*) Romanem Jakubíčkem, druhá část se zabývá detekcí, segmentací a klasifikací kostních nádorových lézí v obratlích, a byla řešena v této dizertační práci a publikována autory Chmelik a kol. (2018*b*, 2019). Tato část je doplněna algoritmem pro zjednodušování tvarů binárních objektů vytvořené Petrem Walkem a byla publikována také v Chmelik a kol. (2018*b*). Další částí jsou moduly pro automatickou analýzu minerální hustoty kostí (BMD²) z CT dat, které byly řešeny v rámci diplomových prací Jakuba Šalplachty a Oldřicha Kodyma pro automatickou segmentaci specifických tkání (tuk a paraspinální sval) (Salplachta, 2017) a automatický výpočet BMD spolu s analýzou tloušťky kortikální části obratlů (Kodym, 2017; Jakubicek a kol., 2020). Poslední část se zaměřuje na analýzu časových řad obrazových dat, sledováním vývoje lézí a odezvy na aplikovanou léčbu, která bude řešena v návaznosti na obě dizertační práce.

Specialisty, zabývajícími se expertní anotací dat a odbornými konzultacemi, jsou prof. Giampaolo Gavelli³, doc. MUDr. Ing. et Ing. Lukáš Lambert, Ph.D.⁴ a MUDr. Petr Ouředníček, Ph.D.⁵.

U onkologicky nemocných pacientů se kromě primárního nádorového ložiska (oblast prvotního vzniku nádorových buněk) mohou objevit také sekundární nádory (metastáze, dceřiná ložiska). Tato ložiska vznikají šířením nádorových buněk z primárního nádoru krevním řečištěm nebo lymfatickými cestami do dalších oblastí v těle pacienta, kde se mohou usazovat a začít formovat nádor (Gasbarrini a kol.,

¹z anglického Computed Tomography

²z anglického Bone Mineral Density

³Radiologia, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) Srl a Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Meldola, Itálie

⁴Radiodiagnostická klinika 1. lékařské fakulty Univerzity Karlovy a Všeobecné fakultní nemocnice v Praze (VFN)

⁵Klinika zobrazovacích metod Lékařské fakulty Masarykovy univerzity v Brně a Fakultní nemocnice u sv. Anny v Brně (FNUSA); International Clinical Research Center (ICRC) FNUSA; Philips, Philips Healthcare, Eindhoven, Nizozemí

2004; Kakhki a kol., 2013). Z uvedených statistických studií (Gasbarrini a kol., 2004; Kakhki a kol., 2013) vyplývá, že 30 až 70 % metastáz vzniká v oblasti kostí (kostní metastáze) v závislosti na anatomickém umístění primárního nádoru. Mezi nejčastěji zasažené kosti patří zejména páteř (přibližně 20 % všech kostních metastáz), žebra (15 %), pánev (10 %), sternum (7 %), femur (5 %), lopatky a lebeční kost (4 %). Prvotním podezřením na přítomnost metastatických nádorů mohou být projevující se symptomy spojené jak s primárním nádorem, tak se samotnými metastázami, jako například bolesti, fraktury kostí, poškození míchy, vysoký obsah vápníku v krvi, apod. Pro potvrzení diagnózy se provádějí laboratorní a zobrazovací testy a biopsie. Mezi laboratorní vyšetření patří různé krevní a močové testy, případně podání tzv. nádorových markerů (The American Cancer Society, 2016).

Nejjednoduššími a nejdostupnějšími diagnostickými technikami jsou konvenční planární rentgen, planární gamagrafie a jejich kombinace, přičemž tyto techniky nesou nejméně diagnostické informace pro odhalení metastáz v kostech a případné sledování odezvy na aplikovanou léčbu (Rybak a Rosenthal, 2001; Costelloe a kol., 2009; Tokuhashi a kol., 2014; Evangelista a kol., 2016). Nejvhodnější zobrazovací technikou z hlediska schopnosti odhalení metastáz včetně sledování jejich progresu a nulové radiační zátížení pacienta i personálu je celotělový sken magnetickou rezonancí (MRI⁶). Tato však není příliš rozšířena z důvodů vysoké ceny přístroje i vyšetření, horší dostupnosti na pracovištích a dlouhé době vyšetření (Rybak a Rosenthal, 2001; Costelloe a kol., 2009; Tokuhashi a kol., 2014; Evangelista a kol., 2016). Nejčastěji používané jsou techniky kombinace CT a jednofotonové emisní tomografie (SPECT⁷) nebo pozitronové emisní tomografie (PET⁸), které nesou dostatek diagnostické informace, za cenu vysoké dávky radiace samotného CT, ještě navýšené o dávku z aplikovaného radionuklidu. Cena je velmi závislá na použitém radionuklidu (Rybak a Rosenthal, 2001; Costelloe a kol., 2009; Tokuhashi a kol., 2014; Evangelista a kol., 2016).

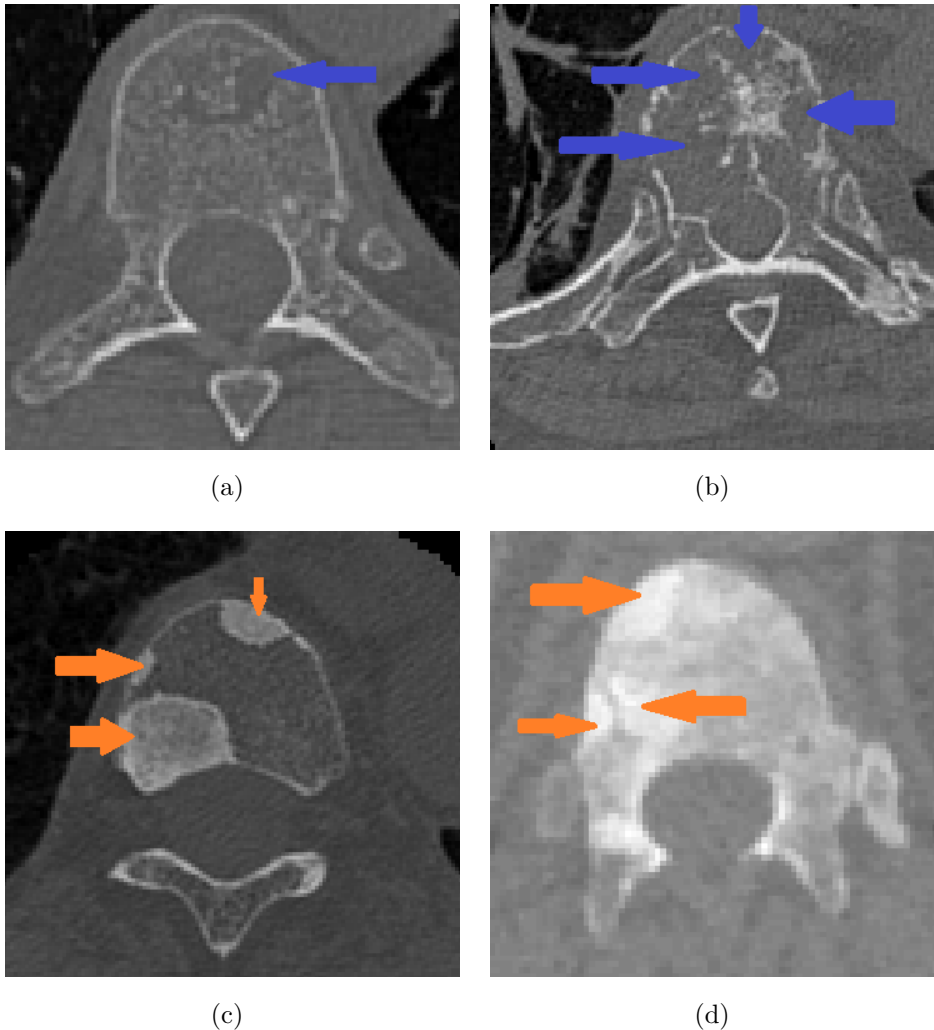
Kosti obsahují dva hlavní typy buněk; osteoblasty, které provádějí syntézu nové kostní tkáně (tzv. kostní matrice) vylučováním kolagenu, a osteoklasty plnící funkci degradace staré kostní matrice rozkládáním kolagenu. Za normálních podmínek tyto buňky zajišťují regeneraci a dorůstání kostí. V případě ovlivnění kostních buněk metastatickými buňkami, dochází k narušení jejich správné funkce, což vede na vznik kostního nádoru. V zásadě se rozlišují dva typy těchto lézí, podle druhu zasažených kostních buněk. První typ (osteolytické/lytické léze) vzniká při postižení osteoklastů (jejich zvýšená aktivita a růst), kdy dochází k degradaci kostní matrice bez vzniku nové. Oblasti lytických lézí se vyznačují sníženou denzitou, a tím pádem oslabením

⁶z anglického Magnetic Resonance Imaging

⁷z anglického Single-Photon Emission Computed Tomography

⁸z anglického Positron Emission Tomography

kostní tkáň, kde hrozí nebezpečí zhroucení nebo zlomení kosti. V CT obrazech se projevují jako oblasti s nižší jasovou hodnotou (Obr. 1(a) a Obr. 1(b)). Druhým typem jsou léze (osteoblastické/blastické/sklerotické) vzniklé zvýšenou aktivitou osteoblastů, kde dochází k výrazné tvorbě kostní matrice bez odpovídající degradace. Tyto oblasti se vyznačují zvýšenou denzitou a tvrdostí (tzv. skleróza), nicméně jsou také mnohem křehčí než normální kostní tkáň, a z tohoto důvodu jsou náchylnější ke zlomení nebo odštípnutí (The American Cancer Society, 2016). V CT obrazech se projevují jako oblasti s vyšší hodnotou jasu (Obr. 1(c) a Obr. 1(d)).



Obr. 1: Ukázka kostních metastáz v obratli v CT obrazech: osteolytické léze – lépe (a) a obtížně (b) definovatelné; osteoblastické léze – lépe (c) a obtížně definovatelné (d).

Vzhledem k relativně častému výskytu metastatických lézí v páteři a jejich vlivu na kvalitu a dobu života takto postižených pacientů, byl vyvinut hodnotící systém k zjištění závažnosti onemocnění a k vyhodnocení dalšího postupu péče o pacienta

(Tokuhashi a kol., 2005; Zou a kol., 2006; Zeng a kol., 2007). Tento systém využívá k výpočtu velkého množství apriorních informací o stavu pacienta, přičemž informace ze zobrazovacích modalit je jejich podstatnou částí. Nevýhodou těchto informací je však jejich velké množství (2D/3D data) a jejich časově náročné hodnocení lékařským personálem (radiologický technik, klinický radiolog, ošetřující lékař, případně chirurg). Z tohoto důvodu je snaha o automatizaci některých rutinních kroků a extrakci co největšího množství diagnosticky užitečné informace, při co největším zjednodušení její interpretace. Segmentace a analýza samotných metastatických lézí z obrazových dat a sledování jejich časového vývoje se jeví jako vhodný nástroj.

V této dizertační práci je navrženo a testováno několik možností řešení segmentace a klasifikace metastatických lézí v páteři ve 3D CT datech. Základním problémem je vysoký rozptyl mezi experty anotovanými daty, kdy se často anotované léze neshodují. Rozhodnutí nejen o tvaru, ale dokonce i o samotné existenci nebo klasifikaci léze je i pro zkušené radiology nejednoznačné. Během řešení byly s ohledem na tuto problematiku navrženy, modifikovány, přepracovávány a testovány klasické přístupy klasifikace založené na statistických modelech nebo na extrakci expertně navržených texturních a tvarových příznaků s jejich následnou redukcí a klasifikací umělou neuronovou sítí. Z důvodu příliš složitého a neurčitého návrhu použitelných příznaků byl jako alternativní navržen přístup založený na moderních metodách hlubokého učení, konkrétně konvoluční neuronové sítě s původní strukturou. Toto se ukázalo jako vhodnější řešení, protože nejrelevantnější příznaky jsou automaticky získávány samotným procesem učení. Dosažené výsledky všech testovaných metod jsou v práci diskutovány a porovnávány s publikacemi v impaktovaných časopisech; navržené řešení s využitím konvoluční neuronové sítě dosahuje slibných výsledků (výrazně lepších než klasické přístupy) a i v porovnání s meziexpertními rozdíly se jeví jako potenciálně použitelné v klinické praxi.

1 Výchozí stav řešení

V této kapitole jsou stručně popsány dostupné publikované metody a přístupy zabývající se problematikou detekce, segmentace a klasifikace kostních lézí se zaměřením na obrazová data páteře snímaná pomocí CT. Dále jsou uvedeny nejčastěji používané metodiky vyhodnocení výsledků dosažených automatickými metodami v porovnání s experty anotovanými daty. Obsah této kapitoly byl v rozšířené podobě publikován jako přehledový článek v impaktovaném časopise (Chmelik a kol., 2018a).

1.1 Detekce, segmentace a klasifikace lézí

Yao a kol. (2006, 2007) a O'Connor a kol. (2007) představili metodu segmentace a klasifikace lytických lézí ze 2D CT obrazových řezů. V prvním kroku aplikovali metodu rozvodí na gradientní parametrickou mapu, čímž získali přesegmentovaný obraz, kde byla dále jednotlivá sousedící povodí na základě průměrné intenzity slučována. Pro jednotlivé kandidáty bylo poté vypočteno 26 kvantitativních příznaků vybraných na základě pozorování a expertních znalostí; využito bylo 13 tvarových příznaků (prostorové momenty binárních objektů), 9 jasových příznaků (statistické momenty z jasových histogramů) a 4 poziční příznaky. Následně byli kandidáti filtrováni pomocí příznakových filtrů (tvarový, velikostní a denzitní) s pevně nastavenými prahovými hodnotami, čímž bylo odstraněno velké množství falešných detekcí. Zbylí kandidáti poté vstupovali do souboru sedmi binárních SVM¹ klasifikátorů s radiální bází, kde každý klasifikátor obsahoval podmnožinu třech příznaků. Soubor SVM a příslušné příznakové vektory byly určeny na základě dvoustranné ANOVA² analýzy. Výsledná klasifikace byla dána většinovým hlasováním. Za správně pozitivní autoři považovali jakoukoliv detekci, která měla alespoň jeden pixel společný s expertní anotací.

Whyne a kol. (2007) testovali hypotézu o možnosti segmentace a odlišení lytických a blastických lézí a následného sledování jejich vývoje na základě analýzy histogramů zdravých a nemocných obratlů. Nejprve provedli analýzu trabekulárních částí zdravých obratlů, kde zjistili Gaussovské rozložení Hounsfieldových jednotek. Následně testovali různé hodnoty prahů, kde experimentálně byly určeny prahy jako $\mu - \sigma$ pro lytické léze a $\mu + 2\sigma$ pro blastické léze, přičemž μ byla střední hodnota jasových hodnot zdravé tkáně daného pacienta a σ jejich rozptyl.

Wiese a kol. (2011) upravili metodu od autorů Yao a kol. (2007) pro detekci sklerotických lézí. Změněna byla metodika slučování 2D oblastí rozšířením o další

¹z anglického Support Vector Machine

²z anglického ANalysis Of VAriance (analýza rozptylu)

podmínky (rozdíl směrodatných odchylek, pozičních souřadnicích a velikosti kandidátů; všechny hodnoty byly stanoveny experimentálně). Dále byl algoritmus rozšířen o slučování 2D kandidátů do 3D „kapek“ na základě vzájemného překryvu a rozdílu intenzit 2D kandidátů v sousedních CT řezech; hodnoty byly opět stanoveny experimentálně. Dále bylo vypočteno 30 kvantitativních příznaků zvolených na základě pozorování a expertních znalostí, které byly dále filtrovány pomocí příznakových filtrů s experimentálně nastavenými prahy a následně klasifikovány souborem SVM. Za správně pozitivní autoři považovali jakoukoliv detekci, která měla alespoň jeden pixel společný s expertní anotací.

Huang a Chian (2012) publikovali metodu klasifikace trabekulárních částí obratlů do dvou tříd – zdravá nebo obsahující metastáze (bez rozlišení jejich typu). Metoda byla založena na analýze texturních příznaků z 2D šedotónových matic vzájemného výskytu (GLCM³), konkrétně 9 Haralickových příznaků a 2 lokálních statistických příznaků (průměr a směrodatná odchylka). Jejich použitelnost byla ověřena nepárovým Studentovým t-testem. Ke snížení počtu falešně negativních výsledků byly zahrnuty také sousední CT řezy tak, že pro každý řez byly vypočteny rozdíly mezi texturními příznaky. Celkově je tedy dostupný 33prvkový příznakový vektor pro každou 2D oblast zájmu (ROI⁴). Samotná klasifikace je provedena pomocí obecné vícevrstvé perceptronové neuronové sítě se zpětným šířením chyby. Byla použita 3vrstvá síť, kde v první vstupní vrstvě bylo 33 neuronů, ve druhé skryté vrstvě 18 neuronů a výstupní vrstva obsahovala 2 neurony. Učení probíhalo 500 epoch s učebním krokem nastaveným na 0,01 a hybností 0,01.

Wels a kol. (2012) představili metodu detekce lytických lézí pomocí víceúrovňového detektoru s vnitřní kontrolou senzitivity se třemi diskriminativními modely. Všechny modely byly tvořeny klasifikátory náhodného lesa (RF⁵), vytvořené na základě podmnožiny komponent z příznakového souboru, který byl tvořen pouze příznaky vykazujícími vysoký Pearsonův korelační koeficient s danou klasifikační třídou. Vstupní data byla nejprve prostorově normována. Víceúrovňový detektor využíval pyramidový přístup detekce, kde první model využíval pouze výpočetně méně náročné 3D Haarovy příznaky pro hrubou detekci středů prvotních kandidátů. Následující dva modely již využívaly ke klasifikaci jednotlivých kandidátů celou relevantní databázi příznaků. Takto byly získány pouze středy jednotlivých kandidátů, proto byl poté implementován detektor škálování založený na strojovém učení, kterým bylo detekováno prostorové rozšíření kandidátů na základě škálovací hypotézy (*scale hypothesis*) získaných na základě učebních dat. Tím byly získány hranice detekovaných lézí. Poté bylo provedeno aglomerativní shlukování, jehož pomocí byly

³z anglického Gray-Level Co-occurrence Matrix

⁴z anglického Region Of Interest

⁵z anglického Random forest

eliminovány redundantní detekce sloučením blízkých detekcí. Posledním krokem bylo odstranění kandidátů, jejichž středové voxely neměli dostatečně nízké jasové hodnoty oproti okolním voxelům. Za správně pozitivní byly považovány detekce, jejichž střed se vyskytoval uvnitř expertem označeného hraničního boxu.

Wiese a kol. (2012) upravili metody autorů Yao a kol. (2006) a Wiese a kol. (2011) pro detekci a klasifikaci sklerotických lézí v páteři. Segmentaci provedli 2D metodou rozvodí s následným slučováním oblastí pomocí grafových metod. Pro každého 2D kandidáta bylo vypočteno 30 charakteristických příznaků. Finální klasifikace kandidátů je poté provedena pomocí SVM klasifikátoru. Za správně pozitivní autoři považovali jakoukoliv detekci, která měla alespoň jeden pixel společný s expertní anotací.

Hammon a kol. (2013) použili dříve publikovanou metodu autory Wels a kol. (2012) a rozšířili ji o detekci sklerotických lézí. Rozšíření spočívalo v použití dvou identických, na sobě nezávislých přístupů dle autorů Wels a kol. (2012), avšak každý byl naučen na odlišný typ lézí. Za správně pozitivní byly považovány detekce, jejichž střed se vyskytoval uvnitř expertem označeného hraničního boxu.

Burns a kol. (2013) a Yao a kol. (2015) využili a upravili metody autorů Yao a kol. (2006); Wiese a kol. (2011) a Wiese a kol. (2012) pro detekci a klasifikaci sklerotických lézí v páteři. Segmentaci provedli 2D metodou rozvodí s následným slučováním oblastí pomocí grafových metod k získání 2D kandidátů. Následně bylo provedeno slučování do 3D kandidátů s využitím analýzy spojených komponent (*connected component analysis*). Dále bylo detekováno těžiště každého 3D objektu a bylo použito jako inicializační bod (*seed point*) pro segmentaci metodou geodetických kontur *level set*. Pro každého 3D kandidáta bylo vypočteno 25 charakteristických příznaků (jasová hodnota, morfologické příznaky, poziční příznaky), které byly dále filtrovány s pevně danými experimentálně zjištěnými prahy. Finální klasifikace kandidátů byla provedena pomocí souboru SVM klasifikátorů. Za správně pozitivní autoři považovali jakoukoliv detekci, která měla alespoň jednovoxelový společný překryv s expertní anotací, stejně tak pokud na jednu expertem označenou lézi připadlo více automaticky detekovaných lézí.

Peter a kol. (2013) představili metodu detekce a segmentace sklerotických a lytických lézí na základě klasifikačního modelu založeného na Bayesově větě kombinované s přístupem Markovských náhodných polí (MRF⁶). K vytvoření pravděpodobnostního modelu využili znalosti o rozložení intenzity zdravé tkáně aproximované Gaussovou funkcí. Pro každý individuální typ obratle vypočítali parametry aproximované Gaussovy funkce (průměrná intenzita a průměrný rozptyl), které tvořily model zdravé tkáně pro daný typ obratle. Optimalizace kritériální funkce založená

⁶z anglického Markov Random Field

na modelované distribuci kombinované s MRF byla řešena metodou *Modified Metropolis dynamic*. Jednotlivé voxely byly klasifikovány do 3 tříd: sklerotická, zdravá a lytická tkáň. Hodnocení senzitivit a specifit bylo počítáno voxelově.

Muñoz a kol. (2013) publikovali metodu pro detekci degenerativního onemocnění meziobratlových disků (DDD⁷) na kortikální části těl obratlů z thorakolumbální páteře. Prvním krokem autoři segmentovali kortikální část obratle pomocí deformovatelných duálních ploch (*deformable dual surfaces*) (první plocha je uvnitř těla obratle, druhá z vnějšku) za použití vnitřních sil (*internal forces*), obrazových potenciálových sil (*image potential forces*) a omezení odvozených z vlastností mezi vnitřní a vnější plochou. Poté byla kortikální část rozvinuta do 2D mapy podle cylindrických souřadnic, přičemž byly brány průměrné jasové hodnoty mezi odpovídajícími si body duálních ploch. Poté byla vypočtena gradientní mapa, která byla prahována, čímž byla získána binární maska představující hyperdenzní oblasti na kortikální části tvořící kandidáty na DDD. Z takto získaných oblastí byly vypočteny 2D charakteristické příznaky (tvarové, poziční, jasové). Vzhledem k nepřesnostem segmentace kortikální části, byla přesnost detekce ve 2D prostoru limitována, proto byly tyto oblasti zpětně mapovány do 3D prostoru za použití párování odpovídajících si hodnot mezi získanou parametrickou mapou a duálními plochami. Zde bylo využito geodezických kontur *level set*, které oblast následně rozšířily do celé oblasti DDD. Poté byly vypočteny další příznaky ve 3D prostoru. Celkem bylo získáno 54 příznaků, které vstupovaly do souboru SVM klasifikátorů pro provedení finální klasifikace.

Roth a kol. (2015, 2016) rozšířili metodu publikovanou autory Burns a kol. (2013) o krok využívající hlubokého učení konvolučních neuronových sítí (CNN⁸) pro snížení falešných detekcí sklerotických lézí. Výstupy po SVM klasifikaci ze stávající metody zde byly brány jako vstupní 2D ROI. K umělému rozšíření učební množiny z důvodu omezení přeučení neuronové sítě autoři využili přístupy rozšiřování (*augmentation*) databáze, kde každé ROI bylo posunuto podél náhodného vektoru v axiální rovině, rotováno kolem středu o náhodný úhel a škálováno s náhodným škálovacím faktorem. Z každého takto upraveného ROI byl poté vyříznut 2D vzorek o velikost 32×32 pixelů. Pro klasifikaci byla použita CNN se dvěma konvolučními vrstvami dvěma sdružovacími (*max-pool*) vrstvami, dvěma plně propojenými vrstvami, *DropConnect* vrstvou a výstupní *softmax* vrstvou. Použity byly ReLU⁹ aktivační funkce neuronů. Za správně pozitivní autoři považovali jakoukoliv detekci, která měla alespoň jednovoxelový společný překryv s expertní anotací, stejně tak pokud na jednu expertem označenou lézi připadlo více automaticky detekovaných

⁷z anglického Disc Degenerative Disease

⁸z anglického Convolutional Neural Network

⁹z anglického Rectified Linear Unit

lézí.

Jan a kol. (2015) prezentovali dva nové přístupy segmentace a klasifikace metastatických lézí v tělech obratlů. Anotace dat i hodnocení účinnosti metod bylo provedeno voxelově pro lytické i sklerotické tkáně.

První metoda byla založena na texturní a tvarové analýze, kde autoři otestovali celkem 27 236 charakteristických 2D a 3D příznaků. Texturní příznaky byly získány řadou metod texturní analýzy, zatímco tvarové příznaky byly obdrženy za prvotní segmentace metodou rozvodů, následným slučováním oblastí a tvarové analýzy takto získaných objektů. Příznakový prostor byl následně redukován pomocí Kruskal-Wallisova testu za použití expertně označených dat na 118 texturních a 30 tvarových statisticky relevantních příznaků. Dále na základě genetických algoritmů byly příznaky dále redukovány na 43 relevantních a neredundantních příznaků. Jako klasifikátor byla využita umělá neuronová síť, která byla experimentálně opakovaně učena na redukováném příznakovém vektoru po dobu 100 učebních epoch. Pro každou naučenou síť byla poté určena hodnota efektivity experimentálně stanoveným kritériem na základě senzitivity a specifity každé klasifikované tkáně. Z tohoto hlediska nejefektivnější neuronová síť byla použita jako finální klasifikátor.

Druhý přístup byl založen na metodě grafových řezů využívajících hledání minimálního řezu grafem pomocí maximálního toku (*min-cut max-flow theorem*). Jako apriorní znalosti byly použity průměrné hodnoty jasů a směrodatných odchylek jednotlivých typů tkání (lytická, zdravá, sklerotická) pro každý typ obratle. Tyto hodnoty byly získány z expertně označených dat. Následně byla provedena dvoufázová segmentace obrazových dat, kdy v první fázi byl hledán minimální řez grafem zkonstruovaným ze 3D obrazových dat těla obratle, přičemž byly odděleny oblasti sklerotických lézí od ostatních tkání. Poté byl zkonstruován nový graf ze zbývajících obrazových dat (bez sklerotické tkáně) a opět byl hledán minimální řez grafem, který oddělil oblasti lytických lézí od zdravé tkáně.

1.2 Přístupy k hodnocení úspěšnosti metod

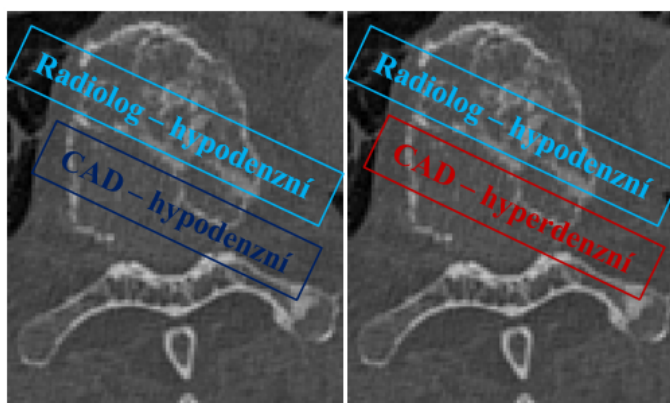
Ověření správnosti a přesnosti detekce, segmentace a klasifikace lézí prováděné systémy pro počítačem podporovanou detekci/diagnózu (CAD¹⁰) je nejčastěji prováděno metodami, které vyžadují referenční, lékařským expertem anotovaná data. Tato data jsou poté různými způsoby porovnávána s výsledky automatické metody a jsou vypočítány různé hodnoty statistického měření (například senzitivita, specifita, FROC¹¹ analýza, apod.). Následuje přehled způsobů porovnávání výsledků s expertními daty, které jsou často používané v publikovaných metodách popsaných

¹⁰z anglického Computer Aided Detection/Diagnosis

¹¹z anglického Free-response Receiver Operating Characteristic

v Kap. 1.1.

„ROI metodika“ – nejjednodušší metodou použitou autory Wels a kol. (2012) je hodnocení celého 3D snímku obratle, jednoho 2D řezu, případně objem zájmu (VOI¹²) nebo ROI. Lékařský expert pouze určí, zda hodnocený objem nebo řez je zdravý či obsahuje nějaký typ lézí. Automatická metoda také hodnotí daný vzorek jako celek, a za správné detekce se považují ty, které se shodují s expertem. Ukázka tohoto přístupu je ilustrována na Obr. 1.1.



Obr. 1.1: Vysvětlení 2D „ROI metodiky“ hodnocení: vlevo – CAD generuje jeden výsledek správné klasifikace (stejné rozhodnutí jako radiolog pro celý snímek), vpravo – generován je jeden falešně pozitivní a falešně negativní výsledek (CAD určí jiné rozhodnutí než radiolog). Další informace o lézích zde není dostupná.

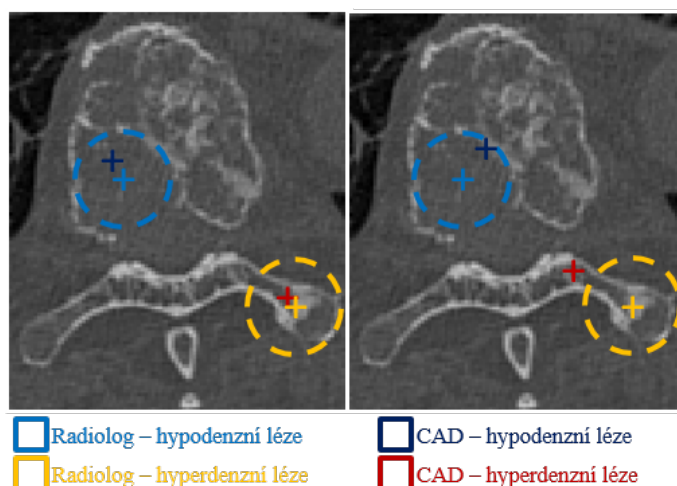
„Bodová metodika“ – další jednoduchou metodou hodnocení využitou autory Muñoz a kol. (2013) je určení jednoho bodu přibližně ve středu léze (*foci*¹³) lékařským expertem. Výstupem automatické metody mohou být také jednotlivé „foci“, přičemž za správnou detekci je považována taková léze, kde např. Eukleidovská vzdálenost mezi jednotlivými *foci* leží v určité toleranci, případně jsou data ověřena pouze subjektivně. Pokud je výstupem automatické metody 2D/3D *bounding-box*¹⁴ nebo pixelově/voxelově segmentovaný objekt, jsou za správnou detekci považovány ty, kde expertem určené *foci* leží uvnitř *bounding-boxů* nebo objektů. Ukázka tohoto přístupu je ilustrována na Obr. 1.2.

„Bounding-box metodika“ – metoda hodnocení detekce a klasifikace lézí, kdy lékařský expert označí 3D *bounding-box* kolem dané léze je použita autory Wiese a kol. (2012) a Burns a kol. (2013). Expert také určí typ takto označené léze. Výsledkem automatické metody je také 3D *bounding-box*. Za správně detekované léze jsou poté považovány takové detekce, kdy střed automatického *bounding-boxu* leží

¹²z anglického Volume of Interest

¹³plurál z anglického Focus (ohisko)

¹⁴ohraničující rámeček

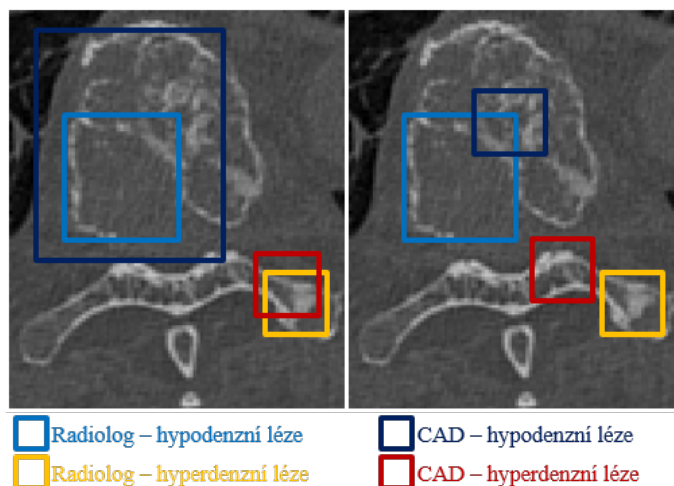


Obr. 1.2: Vysvětlení 2D „bodové metodiky“ hodnocení: vlevo – CAD generuje dva výsledky správných detekcí (bod určený CAD metodou je uvnitř toleranční zóny kolem bodu určeného radiologem), vpravo – hypodenzní detekce není na správné pozici, nicméně je stále počítána jako správná detekce kvůli toleranční zóně, hyperdenzní léze generuje jeden falešně pozitivní a jeden falešně negativní výsledek. Zde je dostupná alespoň přibližná pozice lézí.

kdekoliv uvnitř *bounding-boxu* určeného expertem. Pokud na jednu expertní oblast připadá větší počet automatických detekcí, jsou počítány jako jedna správná detekce. Ukázka tohoto přístupu je ilustrována na Obr. 1.3.

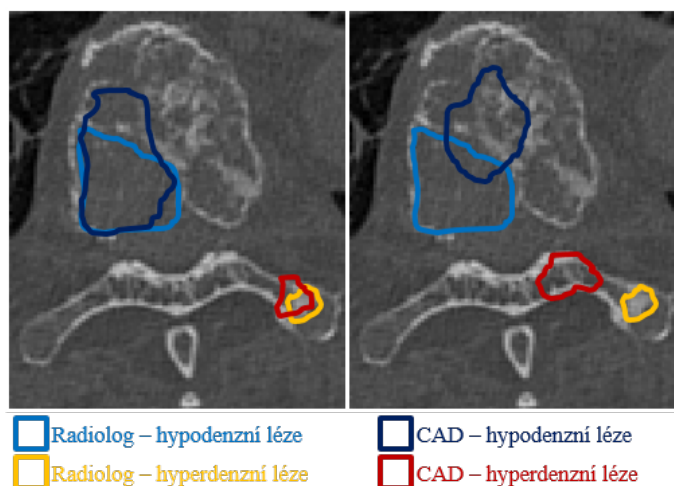
„Objektová metodika“ – obdobná metoda s přesnějším určením lézí segmentací byla použita autory Yao a kol. (2006, 2007, 2015); O’Connor a kol. (2007); Wiese a kol. (2011); Huang a Chian (2012); Hammon a kol. (2013); Muñoz a kol. (2013) a Roth a kol. (2015, 2016), přičemž místo 3D *bounding-boxů* využívá k hodnocení konkrétní objekty. Lékařským expertem jsou označeny voxely každé léze, tak aby vznikly jednotlivé kompaktní objekty, a je určen typ léze. Automatická metoda zde neprovádí pouze detekci a klasifikaci, ale přesnější segmentaci, čímž vzniknou jednotlivé objekty. Stejně jako v předchozím případě jsou poté tyto objekty porovnány s objekty označenými expertem, přičemž jejich překryv jediným voxelem značí správně určenou lézi. Pokud na jeden expertem určený objekt připadá větší počet automaticky nalezených objektů, jsou počítány jako jedna správná detekce. Ukázka tohoto přístupu je ilustrována na Obr. 1.4.

„Voxelová metodika“ – odlišná metoda použitá autory Peter a kol. (2013) a Jan a kol. (2015) využívá k hodnocení voxelově označená data lékařským expertem. Expert tedy označí každý voxel, který náleží dané lézi; zároveň výstupem automatické metody jsou přímo klasifikované jednotlivé voxely. Za správně určené se považují voxely překrývající se s expertním značením. Rozdílem této metody od předchozí

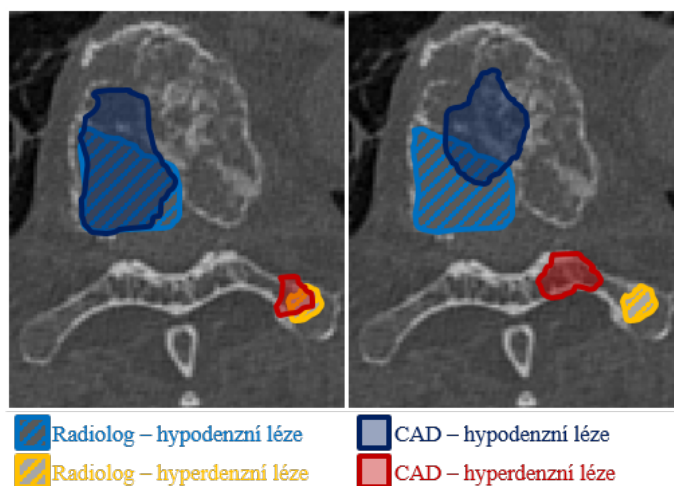


Obr. 1.3: Vysvětlení 2D „bounding-box metodiky“ hodnocení: vlevo – oba výsledky generují správné detekce (i když je hypodenzní léze detekována větší a hyperdenzní je mírně mimo správnou pozici), vpravo – hypodenzní detekce není na správné pozici, nicméně je stále počítána jako správná detekce, protože její střed leží uvnitř anotovaného „bounding-boxu“, hyperdenzní léze generuje jeden falešně pozitivní a jeden falešně negativní výsledek. Dostupná je přibližná pozice lézí a jejich velikost.

„objektové metodiky“ je to, že na výstupy CAD systému i experta je pohlíženo z hlediska jednotlivých voxelů a nikoliv jako celých objektů. Princip značení expertem i CAD systémem však může být v obou metodách stejný. Ukázka tohoto přístupu je ilustrována na Obr. 1.5.



Obr. 1.4: Vysvětlení 2D „objektové metodiky“ hodnocení: vlevo – oba výsledky generují správné detekce (hypodenzní léze je detekována větší a hyperdenzní je mírně mimo správnou pozici), vpravo – hypodenzní detekce není na správné pozici, nicméně je stále počítána jako správná detekce, protože je překryta s anotovaným objektem, hyperdenzní léze generuje jeden falešně pozitivní a jeden falešně negativní výsledek. Přesná pozice a tvar lézí jsou dostupné, velikost lézí je možné odhadnout.



Obr. 1.5: Vysvětlení 2D „voxelové metodiky“ hodnocení: Tento přístup umožňuje nejpreciznější možné určení pozic, tvarů a velikostí lézí oproti dalším zmíněným metodikám, a navíc umožňuje přesný výpočet senzitivity, specificity, apod. pro každý hodnocený voxel.

2 Cíle práce

Cílem dizertační práce byl návrh vhodných analytických metod a realizace automatického algoritmu pro detekci, segmentaci a klasifikaci těžko definovatelných kostních metastatických lézí v obratlích, který má sloužit jako součást komplexního automatického klinicky využitelného počítačem podporovaného detekčního/diagnostického (CAD¹) systému pro analýzu kostních patologií a sledování jejich vývoje a odezvy na aplikovanou léčbu. Cíle práce a stručnou charakteristiku přístupů je možné rozdělit do následujících úkolů:

1. Tvorba patientské databáze

Součástí dizertační práce bylo prvotní vytvoření kompletní databáze patientských dat včetně anotace metastatických lézí (DATABÁZE A). Tato databáze vznikala ve spolupráci s firmou Philips[®] divize Philips Healthcare, Eindhoven (Nizozemí) a nemocnicí s výzkumným onkologickým pracovištěm (IRST) v Meldole (Itálie). Data pro naše účely hodnotili tito specialisté: prof. Giampaolo Gavelli (IRST), doc. MUDr. Ing. et Ing. Lukáš Lambert, Ph.D. (radiodiagnostická klinika 1. LF UK a VFN v Praze) a MUDr. Petr Ouředníček, Ph.D. (klinika zobrazovacích metod LF MU v Brně a FNUSA v Brně; Philips Healthcare Eindhoven, Nizozemí). Příprava patientských databází byla v rovnocenné spolupráci řešena s Ing. Romanem Jakubíčkem.

2. Detekce, segmentace a klasifikace metastatických lézí

Stěžejní částí dizertační práce byl návrh a realizace vhodné metodologie a vytvoření automatického algoritmu pro detekci, segmentaci a klasifikaci těžko definovatelných metastatických lézí v obratlích, který byl připraven jako nástroj komplexního CAD systému pro klinické použití. Vývoj tohoto algoritmu spočíval v aplikaci a testování známých jednoduchých i komplexních metod z oblastí zpracování a analýzy obrazů, které byly v dizertační práci s ohledem na charakter metastatických lézí, jejich obtížnou definovatelnost a rozlišitelnost od jiných kostních struktur dále modifikovány a přepracovávány. Vstupem do algoritmu byly detailně segmentované obratle obdržené z první části projektu, řešené Ing. Romanem Jakubíčkem v rámci jeho dizertační práce (Jakubicek, 2020).

3. Vyhodnocení výsledků pro klinickou využitelnost

Dosažené výsledky navržených algoritmů byly porovnány z pohledu úspěšnosti na dostupných patientských databázích, včetně porovnání se současným stavem řešení daného problému jinými autory. Vyhodnocení bylo diskutováno s lékařskými specialisty a byla zhodnocena využitelnost algoritmů v klinické praxi.

¹z anglického Computer Aided Detection/Diagnosis

3 Přehled používaných metod

Tato kapitola se zabývá stručnou rešerší metod v práci používaných pro klasifikaci a segmentaci obrazových dat a jejich publikovaných rozšíření či modifikací. První část je věnována extrakci obrazových příznaků, včetně výčtu nejběžněji používaných příznaků, jejich jednoduchého dělení a použití. Druhá část shrnuje běžně používané klasifikátory či sofistikovanější segmentační metody. Segmentační přístupy mohou být děleny do dvou základních kategorií – regionově a hranově orientované. Velká část moderních segmentačních přístupů však tyto dvě oblasti kombinuje různými způsoby. Dizertační práce se zabývá zejména segmentací a klasifikací využívající příznakového prostoru, proto je tomuto tématu věnována většinová část teoretického rozboru. Příznakový prostor využitý k následné segmentaci a klasifikaci je obecně N -dimenzionální, kde velikost dimenze závisí na počtu použitých příznaků.

3.1 Extrakce obrazových příznaků

Základními obrazovými příznaky mohou být informace o globálních nebo lokálních hodnotách intenzity, barvy a jejich distribuci, dále různé texturní vlastnosti a informace o tvaru objektů v obraze. Obecně je textura velmi obtížně definovatelná, avšak lidským okem rozpoznatelná. Textura může být chápána jako oblast obrazu, ve které mají změny intenzity jisté charakteristické vlastnosti, jevící se pozorovateli jako uniformní. Základními prvky každé textury jsou tzv. primitiva, což jsou vzájemně podobné útvary, které svým prostorovým uspořádáním tvoří strukturu textury. Na základě vlastností primitiv je možné textury rozdělit do několika kategorií, a to podle velikosti primitiv (hrubé a jemné), podle směrovosti (izotropní a směrové) a podle prostorového uspořádání (silné a slabé) (Jan, 2019; Theodoridis a Koutroumbas, 2008; Sonka a kol., 2015). Účelem texturních příznaků je popsat analyzovanou texturu pomocí objektivních parametrů. Tvarové příznaky jsou většinou počítány z binarizovaných obrazů, například z binárních objektů, kontur objektů nebo jejich hranové reprezentace (Costa a Cesar Jr, 2001; Neal a Russ, 2012), a slouží k objektivnímu popisu tvaru objektů.

Texturní příznaky mohou být děleny na statistické, strukturální, filtračně odvozené, modelově založené a tvarové (Costa a Cesar Jr, 2001; Jan, 2019; Sonka a kol., 2015; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Neal a Russ, 2012).

3.1.1 Statistické příznaky

Statistické příznaky vypovídají obecně o prostorovém rozložení (distribuci) intenzit nebo příznaků voxelů. Obecně rozlišujeme statistické příznaky podle řádu.

Statistiky prvního řádu (statistické momenty) lze získat z analýzy lokálních histogramů obrazu podle

$$m_n = \frac{1}{K} \sum_{k=0}^{K-1} (k - \mu)^n h(k), \quad (3.1)$$

kde K je dynamický rozsah intenzit, k je aktuální hodnota intenzity, μ je střední hodnota, $h(k)$ je histogram a n udává řád momentu. Takto lze získat lokální rozptyl (2. řádu), šikmost (3. řádu) a špičatost (4. řádu). Dalšími příznaky jsou například medián, rozsah, modus, směrodatná odchylka, střední hodnoty, případně energie a entropie. V případě znalosti více rozložení je možné zjistit také korelační koeficient, Euklidovskou vzdálenost, kosinové kritérium, Chí kvadrát nebo vzájemnou informaci (Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Sonka a kol., 2015).

Statistiky druhého řádu vypovídají o relativní pozici páru voxelů. Do této kategorie jsou zařazeny tzv. matice vzájemného výskytu šedotónových hodnot (GLCM¹), které udávají četnosti výskytu dvou voxelů se stejnou hodnotou intenzity v přesně definované vzájemné relativní pozici. Pozice je obecně kódována dvěma parametry: orientací θ , která udává směr v prostoru a relativní vzdáleností d . Pro každou kombinaci uvažovaných prostorových parametrů je vypočítána matice četností $H(d, \theta)$ s prvky $h_{i,j}$, kde i (j) je jasová hodnota prvního (druhého) porovnávaného voxelu. Velikost těchto matic je dána počtem úrovní šedi K v obraze; každá matice H tedy bude mít rozměr $K \times K$. Následně jsou vypočteny normované matice relativních četností $C(d, \theta)$, přičemž její prvky $c_{i,j}$ jsou dány jako

$$c_{i,j} = \frac{h_{i,j}}{\sum_{ij} h_{i,j}}. \quad (3.2)$$

Z těchto matic je možné vypočítat řadu texturních příznaků (Haralickovi příznaky): *Angular Second Moment, Contrast, Inverse Difference Moment, Entropy, Correlation, Sum Average, Difference Average, Sum Entropy, Difference entropy, Variance, Sum Variance, Difference Variance, Information Measures of Correlation, Maximal Correlation Coefficient* (Haralick a kol., 1973; Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Sonka a kol., 2015).

Mezi statistiky vyšších řádů, které vyšetřují vztahy mezi větším počtem voxelů, je možné zařadit matice délky běhu šedotónových hodnot (GRLM²). Tato metoda hodnotí počty sousedících voxelů v řadě o stejné intenzitě v určitém směru (tzv. délka běhu). Výstupem je tedy množina matic $H_\theta(m, k)$, kde θ je směr v prostoru a elementy matice $h_{k,m}$ jsou počty běhů dané délky m při intenzitě k . Každá matice

¹z anglického Gray-Level Co-Occurrence Matrices

²z anglického Gray-Level Run-Length Matrices

má tedy rozměr $K \times M$, kde K je počet úrovní šedi a M je maximální možná délka běhu (největší rozměr obrazu). Z těchto matic je opět možné získat řadu texturních příznaků: *Short Run Emphasis*, *Long Run Emphasis*, *Gray-Level Nonuniformity*, *Run Length Nonuniformity*, *Run Percentage*, *Low Gray-Level Run Emphasis*, *High Gray-Level Emphasis*, *Short Run Low Gray-Level Emphasis*, *Long Run Low Gray-Level Emphasis*, *Short Run High Gray-Level Emphasis*, *Long Run High Gray-Level Emphasis* (Jan, 2019; Sonka a kol., 2015; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008).

Další metodou řazenou mezi statistiky vyšších řádů jsou lokální binární vzory (LBP³). Základním principem této metody je binární kódování definovaného lokálního okolí voxelu. Nejdříve jsou podle centrálního prvku ohodnoceny všechny uvažované okolní voxely, a to tak, že voxely s nižší intenzitou než má centrální voxel jsou nahrazeny 0 a voxely s vyšší nebo rovnou intenzitou pak 1. Následně je takto vzniklá matice vynásobena s maskou binárního kódu (např. pro 26okolí bude maska tvořena prvky $2^0 - 2^{25}$). Nakonec jsou všechny hodnoty sečteny a vloženy na místo centrálního prvku. Existují různé modifikace této metody jako NI-LBP⁴, CI-LBP⁵, CS-LBP⁶, RD-LBP⁷ AD-LBP⁸, VAR⁹ a C¹⁰. Rozdíly těchto metod spočívají v odlišném hodnocení okolních voxelů podle různých kritérií. Výhodou LBP metod je jejich nízká výpočetní náročnost a robustnost vůči lineárním jasovým změnám (Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Brahmam a kol., 2013; Sonka a kol., 2015).

3.1.2 Strukturální příznaky

Strukturálními příznaky je možné získat informace o texturních primitivech a jejich prostorovém rozložení.

Nejběžnější metodou je mikrostrukturní analýza, která je založena na filtraci obrazu souborem 2D filtrů (konvolučních masek) s konečnou impulsní charakteristikou (FIR¹¹). Pro každý voxel je tak získán příznakový vektor, jehož délka je závislá na počtu použitých masek. Základem je soubor 1D filtrů o délce impulsní charakteristiky 3 nebo 5. Jmenovitě jsou to filtry $L3$ a $L5$ pro výpočet lokálního průměru, $E3$ a $E5$ pro zvýraznění hran, $S3$ a $S5$ zajišťujících detekci bodů, $W5$ detekujících vlny

³z anglického Local Binary Pattern

⁴z anglického Neighboring Intensities (sousedící intenzity)

⁵z anglického Central Intensity (centrální intenzita)

⁶z anglického Circular Symmetric (kruhově symetrické)

⁷z anglického Radial Differences (radiálně diferenční)

⁸z anglického Angular Differences (úhlově diferenční)

⁹z anglického Variance (rozptylové)

¹⁰z anglického Contrast (kontrastní)

¹¹z anglického Finite Impulse Response

a $R5$ pro detekci vlnění. Vzájemnou konvolucí impulsních charakteristik jednotlivých 1D filtrů je možné získat konvoluční masky pro vícerozměrná obrazy (9 a 25 masek pro 2D, 27 a 125 masek pro 3D). Každá z těchto masek popisuje odlišné charakteristiky textury. Vzhledem k velikosti masek je tato metoda vhodná zejména pro jemné textury s malými primitivy (Suzuki a kol., 2006; Jan, 2019; Theodoridis a Koutroumbas, 2008; Sonka a kol., 2015).

Další metodou řazenou do odvětví strukturálních příznaků je například analýza texturního spektra, kde je možné ve spektru pozorovat různé charakteristické změny závislé na charakteru textury, případně pozorovat změny lokálních spekter obrazu (Jan, 2019; Theodoridis a Koutroumbas, 2008). Dříve zmíněná LPB se také dá považovat za metodu popisující strukturální příznaky. Dále zde může být zmíněna také metoda TEXEMS¹², která využívá předpokladu, že každý obraz je tvořen superpozicí různého množství různě velkých sub-obrazů („texemy“) s určitým rozptylem na každém voxelu. Tyto sub-obrazy jsou vytvořeny s využitím víceúrovňové analýzy, mixačních modelů apod. Následně je měřena podobnost „texemů“ s částmi obrazu, případně je využito některých klasifikačních přístupů. Tato metoda se využívá zejména v detekci defektů materiálů apod. (Xie a Mirmehdi, 2007).

3.1.3 Filtračně odvozené příznaky

Tyto příznaky jsou získány na základě aplikace filtrů různého typu nebo bank filtrů na obraz a následné analýzy. Podle původu těchto příznaků je možné je rozdělit na příznaky získané z prostorové oblasti, z frekvenční oblasti nebo z prostorově frekvenční oblasti (Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Sonka a kol., 2015).

Příkladem příznaků z prostorové oblasti může být aplikace gradientních filtrů, hranových nebo rohových detektorů, průměrujících filtrů, přizpůsobených filtrů, apod. Dalšími používanými metodami je využití Gáborové analýzy nebo výpočet autokorelační funkce (Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Sonka a kol., 2015). Zařazeny zde mohou být i výše zmíněné strukturální příznaky.

Příznaky z frekvenční oblasti jsou získány analýzou spektra případně výkonového spektra. Využívá se různých tvarů a směrů frekvenčních charakteristik filtrů a porovnává se výkon složek v různých pásmech výkonového spektra. Příkladem mohou být kulovité (radiální) filtry, kde výkony korespondují zejména s velikostí primitiv (větší primitiva mají větší výkon blíže středu výkonového spektra a naopak). Pomocí klínových (úhlových) filtrů lze získat informace o směrovosti textury

¹²z anglického Textural Exemplars (texturní vzory)

(Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Sonka a kol., 2015).

Prostorově frekvenční vlastnosti lze z obrazu získat za použití Fourierovy transformace s proměnnou velikostí okna nebo vlnkové transformace s proměnnou dilatací vlnky. Nejznámějšími příznaky jsou Haarovy příznaky, kdy jsme schopni získat informace o lokálních frekvencích (Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Sonka a kol., 2015).

3.1.4 Modelově založené příznaky

Využití modelově založených metod vede ke dvěma praktickým výsledkům: za prvé je možné vytvořit popisný model textury nebo tvaru, a za druhé je možné texturu či tvar syntetizovat. Každý model je popsán určitým množstvím parametrů, které mohou sloužit také jako příznaky obrazové analýzy (Jan, 2019; Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Sonka a kol., 2015).

Mezi modelově založené metody patří například fraktální geometrie, matematicky popisující složité tvary, které nemohou být jednoduše popsány. Fraktální objekty, pokud jsou zobrazeny v jiném měřítku, mohou být invariantní, soběpodobné, statisticky soběpodobné nebo soběpříbuzné. Těchto vlastností objektů je využito při odhadu fraktální dimenze, jejíž hodnota udává složitost objektu. Základní metodou odhadu fraktální dimenze je mřížková (krabicová) metoda (*box-counting*), kde odhad fraktální dimenze D_B je dán podílem logaritmu počtu objektů potřebných k pokrytí původního objektu N a logaritmem měřítkového faktoru ϵ (zmenšení objektu) dle

$$D_B = \frac{\log N}{\log \epsilon}. \quad (3.3)$$

Při použití této metody je možné uvažovat n topologických dimenzí, kde čtvrtým rozměrem může být jasová hodnota voxelů ve 3D prostoru. Existuje mnoho dalších metod odhadu fraktální dimenze, například *Range/Standard Deviation Analysis*, *Variance Method*, *2D Variation Procedure*, *Triangular Prism Surface Area Procedure*, *Cumulative – Intersection Method*, *Parallel – Lines Method*, *Mass – Radius Method*, a další (Theodoridis a Koutroumbas, 2008; Mirmehdi a kol., 2008; Bedin a kol., 2010; Sonka a kol., 2015).

Další skupinou metod založených na využití modelů jsou například Gibbsova náhodná pole (GRF¹³) a Markovská náhodná pole (MRF¹⁴). Tyto metody využívají k analýze prostorových závislostí jednotlivých voxelů pravděpodobnostní (statistické) teorie. MRF umožňují popsat a modelovat vzájemné vztahy mezi voxely a jejich charakter na základě podmíněných pravděpodobností. Praktické použití MRF ve velké

¹³z anglického Gibbs Random Fields

¹⁴z anglického Markov Random Fields

míře vychází z věty o ekvivalenci mezi MRF a Gibbsovou distribucí. Je to dáno tím, že ve většině praktických případů je nutné znát rozložení pravděpodobností vzájemného výskytu jevů, přičemž její získání z podmíněné pravděpodobnosti MRF je mnohdy velmi obtížné. Větou o MRF–GRF ekvivalenci je dokázáno, že distribuce pravděpodobností vzájemného výskytu jevů v MRF je Gibbsovou distribucí (Berthod a kol., 1996; Li, 2000, 2009).

Máme tedy originální obraz, ve kterém \mathcal{S} je množina jeho voxelů. Vztahy mezi voxely jsou dány systémem sousedů, který je definován jako

$$\mathcal{N} = \{\mathcal{N}_i | \forall i \in \mathcal{S}\}, \quad (3.4)$$

kde \mathcal{N}_i je množina voxelů sousedících s voxelem i . Zároveň pro každou tuto množinu platí, že $i \notin \mathcal{N}_i$ a $i \in \mathcal{N}_{i'} \Leftrightarrow i' \in \mathcal{N}_i$. Počet prvků v každé množině \mathcal{N}_i je dán velikostí uvažovaného okolí a určuje tzv. řád sousedního systému (např. 6okolí = 6 prvků = 1. řád, 26okolí = 26 prvků = 2. řád, atd.). Pár $(\mathcal{S}, \mathcal{N}) \triangleq \mathcal{G}$ představuje graf, kde \mathcal{S} obsahuje uzly a \mathcal{N} určuje spoje mezi jednotlivými uzly v závislosti na definovaném okolí. Následně může být pro $(\mathcal{S}, \mathcal{N})$ vytvořen tzv. *clicque* („klik“) c , který je definován jako podmnožina voxelů z \mathcal{S} . Teoreticky může existovat jednoprvkový „klik“ $c = \{i\}$ až n prvkový „klik“ $c = \{i, i', i'', \dots, i^n\}$. Soubory těchto „kliků“ jsou pak značeny \mathcal{C}_1 pro jednoprvkové až \mathcal{C}_n pro n prvkové „kliky“. Podmínkou je, že všechny prvky v libovolném „kliku“ musejí být vzájemně sousedící. Dále jsou všechny prvky v každém „kliku“ pevně uspořádané; je tedy rozdíl mezi „klikem“ s prvky $\{i, i'\}$ a $\{i', i\}$. Soubor všech „kliků“ pro $(\mathcal{S}, \mathcal{N})$ je poté dán sjednocením všech souborů „kliků“ $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \dots \cup \mathcal{C}_n$. (Berthod a kol., 1996; Li, 2000, 2009)

Mějme náhodné pole F , které má pravděpodobnost konfigurace $P(F_i = f_i)$, kde F_i je index voxelu a f_i je funkční hodnota (intenzita voxelu). Pro zjednodušení zápisu uvažujme, že $P(F_i = f_i) = P(f)$. Aby toto náhodné pole mohlo být současně MRF, musí být splněny dvě podmínky (Berthod a kol., 1996). První je podmínka positivity

$$P(f) > 0, \forall f \in \mathbb{F}, \quad (3.5)$$

kde \mathbb{F} je množina všech možných stavů, kterých může náhodné pole nabývat (Berthod a kol., 1996). Druhou podmínkou je Markovská podmínka

$$P(f_i | f_{\mathcal{S} - \{i\}}) = P(f_i | f_{N_i}), \quad (3.6)$$

kde $\mathcal{S} - \{i\}$ je rozdíl od indexu \mathcal{S} , f_{N_i} jsou hodnoty v sousedních voxelích (Berthod a kol., 1996). MRF může mít i další vlastnosti jako homogenitu a izotropii. Pokud je MRF homogenní znamená to, že pravděpodobnost $P(f_i | f_{N_i})$ je nezávislá na pozici v \mathcal{S} . Izotropní MRF má pravděpodobnost nezávislou na směru (relativní pozici) kliku v \mathcal{S} (Berthod a kol., 1996).

Pokud bude konfigurace náhodného pole F řízena Gibbsovou (Rov. 3.7) distribucí, bude označováno jako GRF.

$$P(f) = Z^{-1} \cdot e^{-\frac{1}{T}U(f)}, \quad (3.7)$$

kde

$$Z = \sum_{f \in \mathbb{F}} e^{-\frac{1}{T}U(f)} \quad (3.8)$$

je normalizační konstanta (*partition function*), T je konstanta nazývaná jako teplota a $U(f)$ je energetická funkce (Berthod a kol., 1996)

$$U(f) = \sum_{c \in \mathcal{C}} V_c(f). \quad (3.9)$$

Energie je dána součtem „klikových“ potenciálů $V_c(f)$ přes všechny možné „kliky“ \mathcal{C} . Hodnota těchto potenciálů je závislá na konfiguraci v každém „kliku“ c (Berthod a kol., 1996). Gaussovské rozložení je speciálním případem Gibbsova rozložení. Důležitou částí výpočtu těchto modelů je výpočet normalizační konstanty Z . V praxi je její přesné vyčíslení nemožné z důvodu sumy přes všechny možné konformace náhodného pole, proto se tato konstanta většinou pouze odhaduje. Další obtížností při modelování textur a získávání parametrů texturních příznaků je také to, že v těchto případech nemusí být pole homogenní ani izotropní. Tyto parametry dále ovlivní jak normalizační konstantu Z , tak i samotnou energetickou funkci $U(f)$ a jejich distribuce pravděpodobností. Tento problém vede na využití řady optimalizačních algoritmů, pomocí kterých jsou postupně všechny parametry odhadovány. Výsledkem je tedy sada odhadnutých parametrů, popisujících model textury. K tomuto účelu jsou využívány algoritmy *Maximum Likelihood*, *Maximum Entropy*, *Maximum Description Length*, *Maximum A Posteriori Probability Estimate*, *Monte Carlo simulation*, *Simulated Annealing*, *Metropolis Dynamic*, apod. (Berthod a kol., 1996; Li, 2000, 2009).

3.1.5 Tvarové příznaky

Tvarové příznaky (deskriptory) jsou skupinou, která popisuje parametry tvaru objektu, jež mohou být vyjádřeny například na základě kontury objektu a jeho povrchu. Mezi obecné tvarové deskriptory patří například obvod, povrch, obsah, těžiště, poloměr (průměr), hlavní a vedlejší osy, tloušťka, apod. Další speciální deskriptory popisují topologii (počet děr, počet spojených prvků v objektu), symetrii, úhlové funkce (graf vzdálenosti bodů kontury od těžiště objektu), počty rohů, hran a vrcholů, konvexnost, distanční vlastnosti, kulatost, hranatost, a řadu dalších. Speciální skupinou jsou příznaky vycházející z fraktální geometrie (viz Kap. 3.1.4), analýzy zakřivenosti, Fourierovské deskriptory, analýza založená na škálovacích transformacích

(*multiscale shape characterization*), Zernikovy momenty, analýza skeletů, analýza středové osy a další morfologicky založené metody, vlnkové deskriptory, apod. (Costa a Cesar Jr, 2001; Neal a Russ, 2012).

3.2 Klasifikační přístupy

Výsledná segmentace obrazu může být, mimo jiné přístupy, dána také zařazením jednotlivých voxelů do tříd s využitím klasifikačních algoritmů. Vstupem do těchto algoritmů může být pouze jasová hodnota voxelů (jako nejjednodušší příznak) nebo celé vektory různých příznaků. Klasifikaci objektů provádí tzv. klasifikátor. Jeho úkolem je přiřadit každému objektu danou třídu na základě vstupních dat a parametrů samotného klasifikátoru. Tyto parametry jsou určovány trénováním (učením), přičemž mohou být rozlišeny dva základní přístupy učení: učení s učitelem, kdy je klasifikátoru předkládána učební množina a odpovídající známý výstup příslušnosti k dané třídě, nebo učení bez učitele, kdy je předkládán pouze vstupní vektor bez znalosti příslušnosti. Podmínkou téměř každého klasifikátoru je, že vstupní příznaky by neměly být korelované. Zároveň je vhodné, aby jednotlivé příznaky přinášely co nejvíce informace (maximální vzdálenost mezi třídami a minimální rozptyl uvnitř tříd). Před učením klasifikátoru je tedy vhodné provést test korelace mezi příznaky, statistické testy relevantnosti příznaků a případně provést analýzu hlavních komponent (PCA¹⁵) s vyčerpáním maximální variability nebo diskriminační analýzu (Holčík, 2012; Haruštiaková, 2012).

Nezákladnějšími klasifikačními metodami jsou například diskriminační funkce, kdy je vektor příznaků přiřazován do třídy pro kterou je diskriminační funkce největší, nebo metoda nejmenší vzdálenosti, kdy je vektor přiřazen do třídy s nejmenší nepodobností. Obě tyto metody jsou v principu podobné pouze se liší v typu hledaného extrému funkce (maximum/minimum). Dalším klasifikátorem je Bayesův (Bayesovský) klasifikátor, který pracuje se znalostí apriorní pravděpodobnosti příslušnosti objektu do dané třídy, a hledá takové řešení, kdy aposteriorní pravděpodobnost bude maximální a chyba minimální.

Speciálními klasifikátory jsou takové, které oddělují příznakový prostor určitými hranicemi (nadrovinami). Do této skupiny patří například umělé neuronové sítě, rozhodovací stromy, náhodné lesy a SVM¹⁶ klasifikátory. Dále existují klasifikátory, které dokáží vhodné příznaky extrahovat automaticky. Příkladem mohou být autoenkodéry, které obsahují dvě hlavní části: první část převádí vstupní data na zakódovaná data (příznaky) a druhá část se na základě těchto příznaků snaží zrekonstruovat vstupní data. Učení probíhá na základě minimalizace chyby rekonstrukce. Využít se

¹⁵z anglického Principal Component Analysis

¹⁶z anglického Support Vector Machine

poté dá autoenkodér jako celek pro klasifikaci, nebo pouze první část pro extrakci příznaků v kombinaci s jiným klasifikátorem. Dalším příkladem klasifikátorů s automatickou extrakcí příznaků mohou být konvoluční neuronové sítě (CNN¹⁷).

3.2.1 Shluková analýza

Možnou klasifikační metodou je shluková analýza. Principem metody je zařazování objektů do určitých shluků na základě jejich vlastností. Existuje několik přístupů shlukové analýzy, které lze rozdělit na hierarchické a nehierarchické.

Hierarchické shlukování je tvořeno systémem podmnožin, přičemž průnikem dvou množin je buď prázdná množina nebo jedna z těchto množin. Každá podmnožina může obsahovat další podmnožiny. Podle směru přístupu jsou definovány dvě možnosti hierarchického shlukování: aglomerativní, kde shlukování začíná od jednotlivých objektů (spodní konec hierarchie) a postupně dochází k jejich slučování, dokud nejsou všechny objekty zařazeny do jednoho shluku. Druhou možností je divizní přístup, kde shlukování začíná od celé množiny (vrchol hierarchie) a postupuje pomocí dělení shluků směrem k jednotlivým objektům. Proces dělení nebo slučování objektů a podmnožin je dán různými metrikami; příkladem může být *Eukleidovská*, *Manhattanovská*, *maximální*, *Mahalanobisova*, *Hammingova* nebo *Levenstheinova vzdálenost*. Výsledkem hierarchického shlukování je tzv. dendrogram, který mapuje hierarchii objektů. Výsledného rozdělení do disjunktních shluků pokrývajících celou množinu je dosaženo metodami řezu dendrogramu. Hierarchické shlukování je však výpočetně náročné a není proto vhodné pro velké objemy dat, kde výpočetní náročnost exponenciálně stoupá (Haruštiaková, 2012).

Do skupiny nehierarchických metod se řadí *k-means*¹⁸ shlukování, shlukování založené na rozložení objektů, shlukování založené na hustotě objektů a další. Výsledkem nehierarchického shlukování je skupina vzájemně disjunktních množin, které obsahují všechny původní objekty (Haruštiaková, 2012). *K-means* algoritmus shlukování probíhá iterativně. Inicializačním krokem je náhodné vygenerování *k* průměrů z pozorovaných dat. Na základě těchto průměrů je prostor rozdělen na *k* shluků pomocí Voronoiova diagramu. Poté je vypočítáno těžiště každé oblasti z diagramu. Následně je prostor opět rozdělen a takto se pokračuje dokud bude docházet k významné změně pozice průměrů/těžišť. Existují různé obměny tohoto algoritmu, např. *k-medians*, *k-medoids*, *fuzzy k-means*, apod. (Bezdek a kol., 1984; Ahmed a kol., 2002; Haruštiaková, 2012) Fuzzy *k-means* algoritmus (Bezdek a kol., 1984)

¹⁷z anglického Convolutional Neural Network

¹⁸shlukování *k*-průměrů (shlukování založené na hledání *k*-průměrů)

hledá minimum kriteriální funkce:

$$J_m = \sum_{k=1}^c \sum_{i=1}^N u_{k,i}^p \|y_i - v_k\|^2, \quad (3.10)$$

kde y_i je pozorovaná hodnota na pozici i , v_k je těžiště (průměr) k shluku z maximálního počtu c , a p je exponent váhování ($1 \leq p \leq \infty$) funkcí příslušnosti $u_{k,i}$, přičemž musí dle Bezdek a kol. (1984) platit, že

$$u_{k,i} \in \langle 0, 1 \rangle \left| \sum_{k=1}^c u_{k,i} = 1 \forall k \quad \wedge \quad 0 < \sum_{i=1}^N u_{k,i} < N \forall i \right. . \quad (3.11)$$

3.2.2 Grafové metody

Obraz může být popsán pomocí grafu tak, že jednotlivé voxely obrazu představují jeho uzly (*nodes*) a vazby se sousedními voxely definují tzv. hrany (*edges*). Tyto hrany se označují také jako n hrany (*n-links*). Dále jsou všechny vrcholy obrazového grafu spojeny se speciálními terminálními vrcholy, zdrojem (*source*) a stokem (*sink*). Tyto spoje jsou označovány jako t hrany (*t-links*) (Sonka a kol., 2015; Lézoray a Grady, 2012). Každá hrana a každý vrchol má své váhy, které mohou být určeny vztahy vycházejícími z různých obrazových příznaků, tvarovými odchylkami, pozíčními odchylkami, apod. V nejjednodušší variantě jsou váhy vrcholů uvažovány jako rovny jedné, existuje jeden zdroj a jeden stok a váhy hran jsou definovány dvěma vztahy:

- regionální vztah (*region term*) – složka energie vyjadřující celkovou penalizaci R zařazení voxelů do oblasti objektu nebo pozadí na základě jasové hodnoty daného voxelu a apriorní znalosti obou histogramů (objektu a pozadí) dle

$$R(A) = \sum_k R_i(A_i), \quad (3.12)$$

kde A je binární vektor všech voxelů v obraze a i je index daného voxelu. Hodnoty vektoru mohou nabývat hodnot 1 pro objekt a 0 pro pozadí, přičemž penalizace R_i pro daný voxel i je dána rovnicemi

$$R_i(1) = -\ln p(I_i|1), \quad (3.13)$$

$$R_i(0) = -\ln p(I_i|0), \quad (3.14)$$

kde I_i je intenzita daného voxelu, p je pravděpodobnost příslušnosti daného voxelu do dané klasifikační skupiny podle apriorního histogramu. Váhy t hran mezi voxely a zdrojem jsou dány $R_i(1)$ a mezi voxely a stokem $R_i(0)$ (Boykov a Jolly, 2001; Sonka a kol., 2015; Lézoray a Grady, 2012).

- hraniční vztah (*boundary term*) – složka $B(A)$ energie penalizující vzájemné propojení N sousedních voxelů k ležících na hranicích mezi objektem a pozadím je definována jako

$$B(A) = \sum_{\{i,k\} \in N} B_{i,k} \cdot \delta_{A_i \neq A_k}, \quad (3.15)$$

$$\delta_{A_i \neq A_k} = \begin{cases} 1 & \text{pokud } A_i \neq A_k \\ 0 & \text{pokud } A_i = A_k \end{cases}, \quad (3.16)$$

kde δ je funkce určující, zda dvojice voxelů $\{i, k\}$ leží na rozhraní mezi objektem a pozadím. Míra penalizace $B_{i,k}$ je pak dána vztahem

$$B_{i,k} = \frac{1}{\text{dist}(i, k)} \cdot e^{-\frac{(I_i - I_k)^2}{2\sigma^2}}, \quad (3.17)$$

vyjadřujícím podobnost jasových hodnot dvou voxelů ležících na rozhraní (Boykov a Jolly, 2001; Sonka a kol., 2015; Lézoray a Grady, 2012).

Celková energie slučuje obě penalizační složky dle vztahu

$$E(A) = \lambda \cdot R(A) + B(A), \quad (3.18)$$

kde λ je parametr řídící poměr obou energetických složek přispívajících do celkové energie. Optimální segmentace takto reprezentovaného obrazu může být provedena například některou z metod hledání minimálního řezu grafem pomocí maximálního toku tak, aby byl úplně oddělen zdroj od stoku (Boykov a Jolly, 2001; Boykov a Kolmogorov, 2004; Sonka a kol., 2015; Lézoray a Grady, 2012).

3.2.3 Náhodný les

Algoritmus náhodného lesa (RF¹⁹) (Breiman, 2001) využívá výhod vycházejících z kombinace množiny slabých klasifikátorů, např. rozhodovacích stromů (DT²⁰) (Breiman, 1984), a tvoří tzv. silný klasifikátor. Nejpoužívanější metodou učení RF je *brute-force* optimalizace v kombinaci se vzorkováním učebních dat tzv. „*bagging*“²¹ a s náhodným výběrem vstupních příznaků pro každý uzel každého DT. Tímto způsobem je naučena množina odlišných (s různou učební podmnožinou a jen některými náhodně vybranými příznaky) DT, které v klasifikační fázi poskytují odlišné výsledky z okolí hledané hodnoty. Celkový výsledek RF klasifikátoru je poté dán např. většinovým pravidlem nebo průměrem výstupních pravděpodobností množiny naučených DT.

¹⁹z anglického Random Forest

²⁰z anglického Decision Trees

²¹z anglického Bootstrap AGGregation

3.2.4 AdaBoost algoritmus

AdaBoostM1²² (Freund a Schapire, 1995) je algoritmus učení s učitelem, na jehož základě je vytvořen binární klasifikátor s využitím učební množiny. Výsledný klasifikátor $H(x)$ (Rov. 3.19) je lineární kombinací tzv. slabých klasifikátorů $h_t(x)$.

$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t \cdot h_t(x) \right), \quad (3.19)$$

přičemž T je počet použitých slabých klasifikátorů, α_t je koeficient daného slabého klasifikátoru a x je vstupní hodnota. V každém kroku učení je do této lineární kombinace přidán jeden slabý klasifikátor z množiny klasifikátorů \mathcal{H} . Výběr v každém kroku učení je realizován hladovým (*greedy*) způsobem tak, aby byl minimalizován horní odhad chyby klasifikátoru. Učební množina se skládá ze dvojic (x_n, y_n) , kde x_n je jeden bod vstupních dat a y_n výstupní (známá) hodnota příslušnosti k jedné ze dvou tříd $\{-1, 1\}$. AdaBoost využívá při učení váhování indikátorové funkce $I(\cdot)$ vahami $D^{(t)}$ tak, že každý prvek vstupních dat je na počátku inicializován stejnou hodnotou ($D_n^{(t)} = 1/N$, kde N je délka celého učebního vektoru). Učení probíhá iterativně s počtem T kroků, přičemž v každém cyklu jsou provedeny následující kroky:

1. nejdříve je nalezen nejlepší slabý klasifikátor h_t při daných vahách $D^{(t)}$ z učebních dat podle

$$h_t = \arg \min_{h_j \in \mathcal{H}} (\epsilon_t), \quad \text{kde} \quad \epsilon_t = \sum_{n=1}^N D_n^{(t)} \cdot I(y_n \neq h_t(x_n)), \quad (3.20)$$

kde ϵ_t je váhovaná klasifikační chyba a $I(\cdot)$ je indikátorová funkce nabývající hodnot $\{0, 1\}$ podle pravdivosti výroku,

2. dále je ověřena velikost chyby nalezeného klasifikátoru jako zastavovacího kritéria zajišťující konvergenci algoritmu podle pravidla

$$\text{pokud} \quad \epsilon_t \geq \frac{1}{2} \quad \text{stop}, \quad (3.21)$$

3. poté je spočítán koeficient slabého klasifikátoru jako

$$\alpha_t = \frac{1}{2} \cdot \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right), \quad (3.22)$$

4. a jsou aktualizovány váhy dle

$$D_n^{(t+1)} = \frac{D_n^{(t)} \cdot e^{-\alpha_t \cdot y_n \cdot h_t(x_n)}}{\sum_{n=1}^N D_n^{(t)} \cdot e^{-\alpha_t \cdot y_n \cdot h_t(x_n)}}. \quad (3.23)$$

²²z anglického Adaptive Boosting

Aktualizace vah v kroku 4. způsobí, že váha špatně klasifikovaných vstupů je zvětšena a váha správně klasifikovaných zmenšena. V následujícím kroku bude tedy hledán slabý klasifikátor, který bude muset lépe klasifikovat doposud nesprávně klasifikované vstupy. Pro klasifikaci do více tříd byl tento algoritmus rozšířen a je označován jako AdaBoostM2. (Freund a Schapire, 1995)

3.2.5 Umělé neuronové sítě

Neuronové sítě se skládají z neuronů, což jsou procesní prvky s několika vstupy a jedním výstupem. Neuron je obecně popsán rovnicí:

$$y = f(\alpha) = f \left(\sum_{i=1}^N w_i \cdot x_i - \vartheta \right), \quad (3.24)$$

kde y je výstup neuronu, x jsou prvky vstupního vektoru, w jsou prvky vah neuronu, ϑ je práh neuronu, N je počet vstupů a $f(\alpha)$ je aktivační funkce (charakteristika) neuronu, přičemž α je označena jako aktivace neuronu. Aktivační funkce neuronu popisuje jeho chování na vstupní hodnoty a jeho možnosti výstupu. Může být definována lineární, po částech lineární nebo nelineární funkcí. Každý neuron dokáže prostor příznaků rozdělit na dvě oblasti. Učením neuronových sítí dochází ke změnám vah a prahů všech neuronů na základě definovaných pravidel/algoritmů. Příkladem mohou být Hebbovo pravidlo, δ pravidlo, gradientní optimalizace, evoluční metody, *Expectation-Maximization*, simulované žhání, *Particle Swarm Optimization*, a další. Kromě klasických dopředných perceptronových sítí (jednovrstvé a vícevrstvé sítě složené z neuronů) existují také odlišné sítě: sítě s radiální bází (RBF²³), Kohonenovy samoorganizující se mapy, rekurentní neuronové sítě (RNN²⁴) (např. Hopfieldova síť), konvoluční neuronové sítě (CNN²⁵), apod. (Jan, 2002; Holčík, 2012; LeCun a kol., 2015; Jan, 2019).

3.2.6 Konvoluční neuronové sítě

Konvoluční neuronové sítě jsou založené na velmi zjednodušené představě funkce primárního vizuálního kortexu (V1²⁶) mozku, která byla zjištěna měřením aktivity neuronů (Hubel a Wiesel, 1959, 1962, 1968), kde bylo zjištěno, že každý jednotlivý neuron dává odlišnou míru odezvy na odlišně prostorově orientovaný stimul. Dále bylo ukázáno, že V1 je uspořádán do jakési 2D prostorové struktury, která odráží obraz ze sítnice oka (analogií je myšlenka 2D konvoluce na 2D obrazech) (LeCun a kol.,

²³z anglického Radial Basis Function

²⁴z anglického Recurrent Neural Network

²⁵z anglického Convolutional Neural Network

²⁶primary visual cortex

2015). V1 obsahuje tzv. „jednoduché buňky“, jejichž aktivita může být zjednodušeně charakterizována lineární funkcí obrazu v malém prostorově umístěném receptivním poli (analogicky je použita malá konvoluční maska, která „pozoruje“ pouze definovanou část obrazu; výsledek konvoluce je poté převeden přes aktivační funkci jednoho neuronu). Dále V1 obsahuje tzv. „komplexní buňky“, které jsou svou funkcí podobné „jednoduchým buňkám“, ale reagují na příznaky získané „jednoduchými buňkami“ a jsou současně invariantní k malým pozičním posunům vstupních příznaků (analogií jsou tzv. slučovací vrstvy – *pooling*). Obecná myšlenka konvolučních sítí je pak tvořena představou, že tento postup je opakovaně aplikován několikrát po sobě, jako bychom procházeli přes několik anatomických vrstev hlouběji do mozku. Na konci této cesty jsou buňky nazývané „*grandmother cells*“, které odpovídají aktivaci na nějaký specifický globální stimul (např. konkrétní osoba) a jsou invariantní na různé geometrické transformace vstupního obrazu. V CNN je jejich ekvivalentem poslední klasifikační vrstva, kde každý neuron odpovídá dané klasifikační třídě (LeCun a kol., 2015).

V posledních letech je vývoj strojového učení s využitím CNN a hlubokého učení velmi znatelný, a to nejen v oblastech zpracování obrazů ale také signálů, bioinformatických a sociálních dat. Ve zpracování obrazů o tom svědčí velké množství vydaných publikací na toto téma v různých oblastech výzkumu a vývoje (Arel a kol., 2010; LeCun a kol., 2015; Litjens a kol., 2017). První použití CNN bylo představeno autory LeCun a kol. (1989), kde autoři porovnávají různé architektury sítí a porovnávají její efektivitu se standardní neuronovou sítí v problematice klasifikace písmen. Nejznámějšími navrženými sítěmi jsou například *LeNet* (LeCun a kol., 1998) pro rozpoznávání číslic v dokumentech z databáze MNIST, dále vítězná síť AlexNet (Krizhevsky a kol., 2012) pro klasifikaci fotografických obrazů do tisíce různých tříd v rámci výzvy ILSVRC-2010²⁷, vítězná síť GoogLeNet (Szegedy a kol., 2013, 2015, 2017) na výzvě ILSVRC-2014, kde byly představeny tzv. *inception* moduly, VGG síť (Simonyan a Zisserman, 2014) vítězná síť ResNet Microsoft (He a kol., 2016) na výzvě ILSVRC-2015 (Russakovsky a kol., 2015) se zavedením reziduálního učení řešícím problematiku zpětného šíření gradientu u velmi hlubokých sítí, a mnoho dalších. Zvláštním typem CNN mohou být tzv. *image-to-image* architektury, které vstupní obraz převádějí přímo na výstupní obraz o stejné velikosti. Výstupní obraz může nabývat nějakých reálných hodnot nebo i binárních či indexových hodnot. Tyto architektury jsou dnes velmi populární pro řešení různých regresních nebo segmentačních problémů. Nejznámější jsou CNN typu U-Net (Ronneberger a kol., 2015) nebo SegNet (Badrinarayanan a kol., 2017) a jejich nejrůznější modifikace.

Učební proces CNN je velmi podobný jako u standardních neuronových sítí

²⁷z anglického ImageNet Large Scale Visual Recognition Challenge

(je možné použít klasický přístup zpětného šíření chyby), avšak zásadním rozdílem je samotná architektura sítě a její uspořádání. Standardní neuronové sítě mají každý neuron propojený se všemi neurony v předchozí vrstvě (označováno jako *fully-connected*), a navíc mají pro každý tento spoj unikátní parametr (tzv. váhu), což vede na velké množství parametrů sítě. Počet parametrů rychle narůstá s narůstající velikostí vstupního vektoru (což může být velký problém právě u obrazových dat) a se zvětšujícím se počtem skrytých vrstev a skrytých neuronů. Mimo vysoké paměťové (uložení velkého množství parametrů) a výpočetní nároky (úprava velkého množství parametrů během každého učebního cyklu) je hlavním problémem vysoká pravděpodobnost přeučení sítě (*over-fitting*). CNN sítě jsou oproti tomu uspořádány do 3D objemů neuronů, každá vrstva má tedy kromě počtu neuronů také dány rozměry šířky (*width*), výšky (*height*) a hloubky (*depth*) vrstvy. Snížení počtu parametrů je provedeno tak, že jeden neuron je spojen pouze s omezeným počtem lokálních neuronů v předchozí vrstvě (označováno jako *locally-connected*). Další výrazné snížení počtu parametrů je dáno jejich sdílením mezi neurony ve stejném hloubkovém řezu (*depth slice*) dané vrstvy sítě.

Základními typy vrstev CNN jsou konvoluční vrstva (*convolutional layer*), aktivační vrstva (*activation layer*), slučovací vrstva (*pooling layer*) a standardní plně propojenou (FC²⁸) vrstvou (Jan, 2019).

Konvoluční vrstva provádí konvoluci vstupních dat se sadou masek filtrů (někdy označováno jako *kernels*), přičemž impulsní charakteristiky těchto filtrů jsou dány jednotlivými vahami neuronů, které se během učebního procesu upravují. Tato vrstva má čtyři základní hyper-parametry:

- receptivní pole neuronu (*receptive field*) – odpovídá velikosti konvolučního filtru (jeho šířce a výšce); vyjadřuje, jak velká oblast předchozí vrstvy je spojena s neuronem; velikost ve směru hloubkové osy je vždy stejně velká jako hloubka předchozí vrstvy,
- výstupní hloubka (*depth*) – definuje počet použitých filtrů v dané vrstvě sítě a tedy hloubku výstupu z této vrstvy; množina neuronů, která sleduje stejnou vstupní oblast je označována jako hloubkový řádek (*depth column*) nebo vlákno (*fibre*),
- krok (*stride*) – určuje o kolik prvků se daný filtr posouvá po vstupních datech,
- rozšíření nulami (*zero-padding*) – určuje, o kolik nulových řádků/sloupců bude vstupní obraz rozšířen; umožňuje tak řídit velikost výstupu.

Aktivační vrstva počítá výstup každého neuronu na základě hodnot vycházejících z konvoluční vrstvy a definované aktivační funkce. Nejznámější jsou aktivační funkce s průběhem:

²⁸z anglického Fully-Connected

- sigmoidy, které jsou oblíbeny pro jejich obor funkčních hodnot, který je v intervalu $\langle 0, 1 \rangle$; v dnešní době se tento typ aktivační funkce opouští, protože trpí dvěma nevýhodami. Při vysokých kladných nebo záporných vstupech dochází k jejich saturaci, a během učebního procesu při zpětném šíření chyby je poté gradient velmi malý nebo nulový a dochází k velmi malým nebo žádným úpravám vah neuronů. S tím je spojeno také nebezpečí saturace při inicializaci vah na začátku učení. Druhou nevýhodou je, že funkce není centrována na nule; to může způsobit problém během učení, pokud hodnoty vstupních dat také nebudou centrovány.
- hyperbolické tangenty, stále trpí nevýhodou saturace, ale bývá používána častěji než sigmoidální funkce; je centrována na nule a nabývá hodnot z intervalu $\langle -1, 1 \rangle$,
- ReLU²⁹, je definována jako kladná stoupající lineární funkce pro kladné vstupní hodnoty a nulová pro záporné; v posledních letech bývá často používána pro své výhody: nedochází k saturaci, jednoduchá programová implementace a rychlý výpočet gradientu i aktivace, velká rychlost konvergence při použití metody stochastického gradientního sestupu (SGD³⁰) (Krizhevsky a kol., 2012). Nevýhodou je, že během trénování může dojít ke „smrti“ těchto jednotek pokud dojde k velkému toku gradientu a následně nevhodné úpravě odpovídajících vah, které budou generovat vždy záporný výstup. Během celého učení pak nedojde k aktivaci na žádný vstupní prvek, přičemž gradient procházející tímto neuronem bude vždy nevratně nulový. Většinou je tento jev způsoben nevhodně zvoleným (příliš vysokým) parametrem rychlosti učení (LR³¹).
- *Leaky ReLU*, je definován podobně jako ReLU, s tím rozdílem, že záporná oblast není nulová, ale nabývá záporných hodnot s mírným lineárním poklesem; tím je omezena nevýhoda „smrti“ ReLU. Nevýhodou je přítomnost učitelého parametru navíc, který ovlivňuje strmost záporného poklesu. Tento přístup byl dále upraven tak, že parametr strmosti je adaptivně učen spolu s dalšími parametry neuronů (He a kol., 2015).

Slučovací vrstva slouží ke snížení prostorové velikosti dat v osách šířky a výšky ve vnitřních vrstvách sítě, čímž jsou redukovány počty parametrů v dalších vrstvách. Hloubkový rozměr zůstává nezměněn. Tato vrstva také zvyšuje robustnost sítě na lokální posun objektů v obraze. V praxi v podstatě dochází ke specifickému podvzorkování dat (většinou aktivačních map), které je ovlivněno třemi hyper-parametry:

- použitá operace (*pooling operation*) – matematická operace, která je prováděna při podvzorkování; nejčastěji se používá maximum, jedná se tedy o nelineární

²⁹z anglického Rectified Linear Unit

³⁰z anglického Stochastic Gradient Descent

³¹z anglického Learning Rate

maximový filtr; obecně mohou být použity také operace průměrování, L2 vzdálenost, apod.,

- receptivní pole (*receptive field*) – odpovídá velikosti masky filtru (jeho šířce a výšce); vyjadřuje, jak velká oblast předchozí vrstvy je brána v potaz,
- krok (*stride*) – určuje o kolik prvků se daný filtr posouvá po vstupních datech.

Kromě základních typů vrstev se také běžně používají tzv. *Dropout* regularizační vrstvy ((Hinton a kol., 2012; Krizhevsky a kol., 2012; Srivastava a kol., 2014)), která během učebního procesu v každé epoše náhodně „vypíná“ určité procento neuronů (v praxi kolem 50 %), čímž pro každou epochu učení je uměle využita pouze část dané vrstvy sítě. Použitím těchto vrstev se zamezuje přeučení sítě na trénovací množině, přičemž pro vybavování sítě je již použita celá síť, která dosahuje lepšího zobecnění problému. V podstatě je naučeno větší množství různých menších sítí, které poté provádí klasifikaci společně. Tato myšlenka byla dále zobecněna do *DropConnect* vrstvy (Wan a kol., 2013), která místo celých neuronů „vypíná“ náhodně jednotlivé spoje mezi vrstvami.

Dalším speciálním typem vrstvy je normalizační vrstva, která normalizuje data před vstupem do další vrstvy sítě. Při použití tzv. *Batch normalization*, je během učení každý vstupní prvek z trénovacího *mini-batch* (omezený počet prvků z trénovací množiny – řádově desítky až stovky) v každé vrstvě normalizován odečtením průměrné hodnoty z daného *mini-batch* z dané vrstvy a dané dimenze (příznaku) a podělen jeho směrodatnou odchylkou (Ioffe a Szegedy, 2015). Takto normalizované hodnoty jsou poté lineárně transformovány podle vztahu

$$\tilde{h}_d^{(i)} = \gamma_d \tilde{x}_d^{(i)} + \beta_d, \quad (3.25)$$

kde $\tilde{x}_d^{(i)}$ je i prvek d příznaku z normalizovaného *mini-batch*, $\tilde{h}_d^{(i)}$ je lineárně transformovaný výstupní prvek, γ_d a β_d jsou transformační parametry, které se během učení sítě upravují obdobně jako váhy neuronů v jiných vrstvách (Ioffe a Szegedy, 2015). Tímto krokem je eliminován problém, kdy se během učení mění distribuce vstupních hodnot v každé vrstvě sítě, podle toho, jak se mění parametry vah vrstvy předchozí. Normalizace také umožňuje použití vyšších hodnot učebního kroku a přispívá k rychlejší konvergenci algoritmu (dosažení stejné přesnosti sítě za menší počet trénovacích epoch); současně také slouží jako regularizace a zamezuje tak přeučení sítě a eliminuje problematiku počátečního nastavení hyperparametrů (Ioffe a Szegedy, 2015).

Během učebního procesu dochází k hledání optima kriteriální funkce (*loss function*, *cost function*, *objective function*), která odráží úspěšnost klasifikace dat. Pro zjednodušení definujme vstupní vektory jako \mathbf{x}_i , \mathbf{y}_i jsou *one-hot encoded* (vektor nul s jedinou jedničkou na pozici správné třídy c) vektory správné třídy z C možných tříd a \mathbf{W} je matice všech vah v neuronové síti.

Nyní můžeme definovat jednu z nejpoužívanějších kriteriálních funkcí, tzv. *cross-entropy loss*:

$$L_i = - \sum_{c=1}^C y_{ic} \cdot \ln(p(y_{ic})) + (1 - y_{ic}) \cdot \ln(1 - p(y_{ic})), \quad (3.26)$$

Celková hodnota kriteriální funkce je pak dána průměrem přes všechna i , a je snaha nalézt takové hodnoty matice \mathbf{W} , které budou odpovídat globálnímu minimu této funkce (ideálně hodnota 0). Problémem je, že pokud existují takové hodnoty \mathbf{W} , které správně klasifikují všechny třídy, nemusí tyto hodnoty být nutně unikátní (příkladem může být například vynásobení celé matice libovolnou skalární hodnotou). Aby se zamezilo nejednoznačnosti řešení, přidává se ke kriteriální funkci regularizační vztah. Nejběžněji používaná je $L2$ regularizace, která upravuje celkovou hodnotu kriteriální funkce takto:

$$L = \frac{1}{N} \sum_i^N L_i + \lambda R(\mathbf{W}), \quad (3.27)$$

přičemž N je počet vstupních prvků, λ je hyper-parametr míry regularizace a $R(\mathbf{W})$ je regularizační vztah definovaný jako

$$R(\mathbf{W}) = \sum_k W_k^2, \quad (3.28)$$

kde k je počet prvků matice \mathbf{W} . Vztah tak k hodnotě kriteriální funkce přičítá hodnotu danou součtem kvadrátů všech prvků matice \mathbf{W} , která není závislá na vstupních datech a správnosti jejich klasifikace; účelem je penalizovat vysoké hodnoty vah. Je nutné si také uvědomit, že při použití regularizačního vztahu již nelze dosáhnout nulové hodnoty kriteriální funkce.

Existuje celá řada kriteriálních funkcí a možností jejich regularizace, např. *hinge/SVM loss*, *Dice loss*, *weighted cross-entropy loss*, *generalized Dice loss* (Sudre a kol., 2017), *mean squared error/L2 loss*, apod.

Učení sítě, a tím také hledání optima kriteriální funkce probíhá standardně zpětným šířením chyby celou sítí na základě gradientu kriteriální funkce, obdobně jako u klasických neuronových sítí. Úprava parametrů vah tedy probíhá v závislosti na velikosti gradientu chyby, tak aby bylo dosaženo její minimalizace. Váhy mohou být iterativně měněny například pomocí některého z následujících používaných přístupů:

- metoda gradientního sestupu (GD³²) (Robbins a Monro, 1951; Amari, 1998) – nejprve je vypočítána celková hodnota kriteriální funkce L_{epoch} pro všechny prvky učební množiny při daných vahách v dané epoše učení W_{epoch} a její gradient ∇L_{epoch} , a následně je provedena úprava vah podle

$$W_{epoch+1} = W_{epoch} - \alpha \nabla L_{epoch}, \quad (3.29)$$

³²z anglického Gradient Descent

kde α je koeficient rychlosti učení (učební krok). V praxi se tento přístup většinou nevyužívá, protože v každé epoše učení je nutné neuronové síti předložit celou učební množinu, která pro učení CNN bývá velmi rozsáhlá, aby došlo k jednomu kroku úpravy vah.

- metoda stochastického gradientního sestupu (SGD³³) (Robbins a Monro, 1951; Amari, 1998) – gradient je aproximován pomocí pouze jednoho náhodně vybraného prvku z učební množiny a na jeho základě je provedena úprava vah podle

$$W_{it+1} = W_{it} - \alpha \nabla L_i, \quad (3.30)$$

což vede na efektivnější učení z hlediska časové a výpočetní náročnosti, než při použití GD algoritmu. V praxi se většinou využívá tzv. *Mini-Batch Gradient Descent*, který je z aplikačního hlediska totožný s SGD, ale pro výpočet aproximace gradientu využívá místo jednoho prvku sadu několika (desítky až stovky) náhodně vybraných vzorků. Výhodou je možnost využití paralelizace a urychlení výpočtu.

- SGD s hybností (*SGD with Momentum*) (Yu, 1993) – přidává do rovnice úpravy vah koeficient hybnosti μ a rychlost v podle

$$v_{it+1} = \mu v_{it} + \alpha \nabla L_i, \quad (3.31)$$

$$W_{it+1} = W_{it} - v_{it+1}. \quad (3.32)$$

Jak je patrné, pohyb po kriteriální funkci ve směru sestupu gradientu je nyní upravován v závislosti na velikosti rychlosti (tedy v podstatě gradientu) z předchozí iterace upravený koeficientem hybnosti μ (většinou v intervalu $\langle 0, 5, 1 \rangle$). To vede k možnosti pohybovat se částečně ve směru, ve kterém byl gradient v předchozí iteraci; výhodu je určité zbrzdění působící ve směru předchozího gradientu, což umožňuje překonávat určité lokální extrémy a poskytuje určitou odolnost proti oscilacím na lokálních výchylkách gradientu. Pro problémy hloubkového učení (*deep learning*) vykazuje větší rychlost konvergence než SGD metoda.

- Nesterovův adaptivní gradient (NAG³⁴) (Nesterov, 1983) – obdobný princip jako u SGD s hybností, avšak s tím rozdílem, že nová hodnota rychlosti je vypočítána na základě gradientu v místě kroku hybnosti

$$v_{it+1} = \mu v_{it} + \alpha \nabla L_i(W_{it} - v_{it}), \quad (3.33)$$

$$W_{it+1} = W_{it} - v_{it+1}, \quad (3.34)$$

V současné době je tento přístup často používán a vykazuje mírně lepší výsledky než standardní SGD s hybností.

³³z anglického Stochastic Gradient Descent

³⁴z anglického Nesterov Adaptive Gradient

- *Adagrad* (Duchi a kol., 2011) – metoda využívající sledování gradientu během celého učení k adaptivnímu určování učebního kroku pro každý parametr vah W individuálně. Pro váhy s vysokým gradientem je učební krok snížen a naopak, pokud je gradient nízký, učební krok je zvýšen. Úprava probíhá podle

$$C_{it+1} = C_{it} + \nabla L_i^2, \quad (3.35)$$

$$W_{it+1} = W_{it} - \frac{\alpha}{\sqrt{C_{it+1}} + eps} \odot \nabla L_i, \quad (3.36)$$

kde C je matice součtů umocněných gradientů za celý průběh učení, který je v každém kroku použit k normalizaci (s narůstajícím C bude míra úpravy parametrů vah snižována), eps je malé číslo zamezující dělení nulou, a \odot je násobení prvek po prvku. Nevýhodou je, že v případě hlubokého učení může dojít k předčasnému zastavení učení z důvodu rychlého poklesu učebního kroku způsobeného neustálým narůstáním hodnot matice C .

- *AdaDelta* (Zeiler, 2012) – řeší problematiku rychlého poklesu učebního kroku a předčasného zastavení metody *Adagrad* tak, že exponenciálně snižuje hodnoty matice C podle rovnice

$$C_{it+1} = \rho C_{it} + (1 - \rho) \nabla L_i^2, \quad (3.37)$$

$$W_{it+1} = W_{it} - \frac{\alpha}{\sqrt{C_{it+1}} + eps} \odot \nabla L_i, \quad (3.38)$$

kde ρ je koeficient exponenciálního poklesu.

- *Adam* (Kingma a Ba, 2014) – jedna z nejpoužívanějších metod v poslední době využívající automaticky se adaptující učební krok, jejíž výpočet probíhá podle rovnic

$$m_{it+1} = \beta_1 m_{it} + (1 - \beta_1) \nabla L_i, \quad (3.39)$$

$$v_{it+1} = \beta_2 v_{it} + (1 - \beta_2) \nabla L_i^2, \quad (3.40)$$

$$\widehat{m}_{it+1} = \frac{m_{it+1}}{(1 - \beta_1^{it+1})}, \quad (3.41)$$

$$\widehat{v}_{it+1} = \frac{v_{it+1}}{(1 - \beta_2^{it+1})}, \quad (3.42)$$

$$W_{it+1} = W_{it} - \frac{\alpha}{\sqrt{\widehat{v}_{it+1}} + eps} \odot \widehat{m}_{it+1}, \quad (3.43)$$

kde m a v jsou exponenciálně plovoucí průměry gradientu ∇L_i a umocněného gradientu ∇L_i^2 , β_1 a β_2 jsou hyper-parametry řídicí exponenciální pokles m a v , a \widehat{m} a \widehat{v} jsou upravené hodnoty z důvodu nestability na začátku učení, kdy jsou m a v inicializovány jako nulové.

- Metody druhého řádu – mezi tyto metody lze zařadit například ty, jež jsou založené na Newtonově metodě, přičemž je nutné počítat inverzní Hessovu matici. Pro učení velkých sítí s velkým počtem parametrů nebo velkými učebními

množinami nejsou ani *quazi-Newtonovské* (Likas a Stafylopatis, 2000) metody téměř používány z důvodu velké paměťové a výpočetní náročnosti. V poslední době zejména z důvodu rozvoje grafických jader s vysokou paralelizovatelností se již tyto metody začínají využívat a vznikají různé modifikace metod druhého řádu (Martens, 2010).

- Negradientní metody – nevyužívají k úpravám vah zpětné šíření chyby pomocí gradientu, ale jiné přístupy. Příkladem mohou být metody evolučních algoritmů (Angeline a kol., 1994; D’Ambrosio a Stanley, 2007; James a kol., 2011; Sinha a kol., 2017; Chhachhiya a kol., 2017) apod. Tyto metody se s výhodou používají pro učení rekurentních neuronových sítí, například v oblastech počítačového vidění pro řízení autonomních robotů. Jejich výhodou je rychlost optimalizace parametrů.

Moderní aplikace hlubokého učení v medicíně

Metody hlubokého učení jsou momentálně ve fázi rozvoje ve všech oblastech zpracování medicínských (i nemedicínských) obrazových dat v různých výzkumných a klinických odvětvích (Litjens a kol., 2017). Příkladem mohou být nejnovější aplikace CNN pro segmentaci (Dou a kol., 2017) a detekci mozkových nádorů z MRI obrazů s využitím dvoucestné CNN s dvoufázovým učením (Havaei a kol., 2017) nebo přístup s 11vrstvou pyramidální 3D CNN (Kamnitsas a kol., 2017). Další metoda segmentující nádory prostaty z MRI dat byla navržena autory Yang a kol. (2017b), kteří využili metod slabého učení (*weakly supervised learning*). Segmentace plicních nodulů založená na CNN využívající víceúrovňové 2D + 3D příznaky a centrální slučovací vrstvu byla implementována autory Wang a kol. (2017). Autoři Charbonnier a kol. (2017) navrhli metodu segmentace dýchacích cest, kterou rozšířili o detekci a opravu neuzavřených ploch pomocí CNN. Ngo a kol. (2017) navrhli metodu pro detekci a segmentaci levé srdeční komory pomocí kombinace *deep belief network* a geometrických kontur (*level-sets*). Odlišný přístup segmentace srdeční komory popsali autoři Tan a kol. (2017), který byl založen na dvou oddělených CNN; první pro lokalizaci těžiště a druhou pro vymezení okrajů komory. Metoda segmentace objektů z histologických obrazů využívající víceúrovňovou *Fully Convolutional Network* byl prezentován autory Chen a kol. (2017). CAD metody využívající CNN byly také navrženy pro segmentaci novotvarů prsu (Dhungel a kol., 2017; Kooi a kol., 2017). Metoda segmentace tepenných kalcifikací a jejich následného hodnocení (Wolterink a kol., 2016) byla navržena za pomoci dvojice CNN. Metody hlubokého učení jsou využívány v mnoha dalších oblastech, např. pro detekci patologií sítnice (Quellec a kol., 2017), známkování patologií deformace páteře (Jamaludin a kol., 2017), diagnostiku onemocnění mozku (Suk a kol., 2017), potlačení šumu (Benou a kol.,

2017), nebo pro potlačení nežádoucích anatomických struktur v analyzovaných obrazech (Yang a kol., 2017a), a řada dalších. V budoucnu se dá předpokládat další rozvoj, nové aplikace a využitelnost těchto metod.

4 Vlastní řešení

Tato kapitola pojednává o navržených a realizovaných přístupech řešení dané problematiky, včetně přípravy patientské databáze, anotace dostupných dat experty a návrhu metodik hodnocení dosažených výsledků. Během řešení bylo nejprve provedeno přepracování dostupných algoritmů vytvořených během dřívějšího řešení projektu pro aktuální dostupnou patientskou databázi. Následoval návrh a realizace několika původních přístupů a jejich vzájemné porovnání a zhodnocení dosažených výsledků.

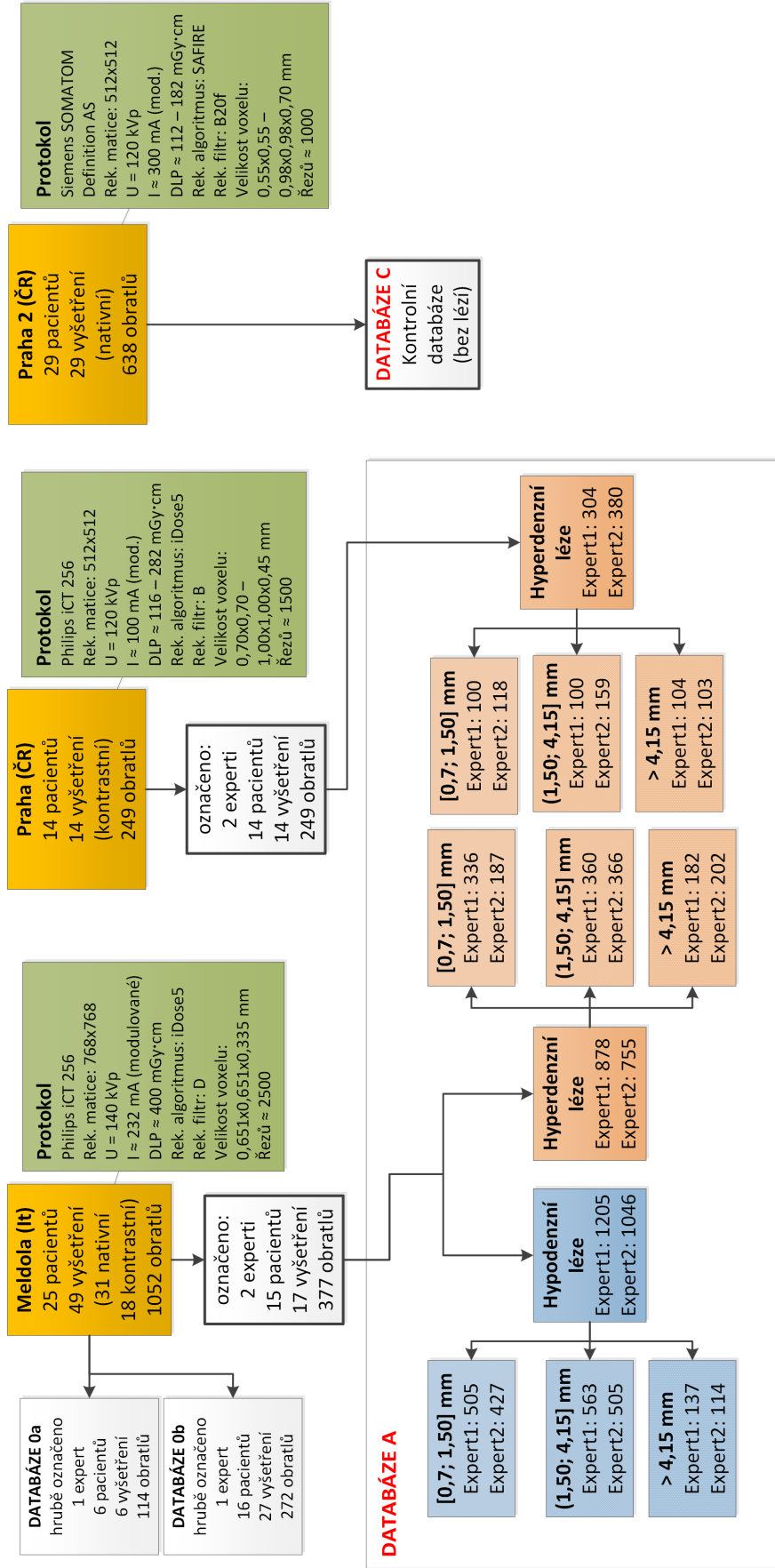
Veškeré algoritmy byly realizovány na osobním počítači se 4jádrovým procesorem Intel® Core™ i5–3330 s frekvencí jader 3 GHz, 32 GB operační paměti, standardním magnetickým diskem, grafickým procesorem NVIDIA® GeForce® GTX970 se 4 GB GDDR5 paměti, 1664 CUDA¹ jádery a taktem 1050 MHz a s Microsoft® Windows® 10 operačním systémem. Dále bylo použito programovací prostředí MathWorks® MATLAB® 2017b, Microsoft® Visual Studio® 2015, volně dostupný toolbox MatConvNet v.1.0–beta25 (Vedaldi a Lenc, 2015), NVIDIA® CUDA toolkit 9.0 a knihovny NVIDIA® cuDNN 7.

4.1 Patientská databáze a expertní značení dat

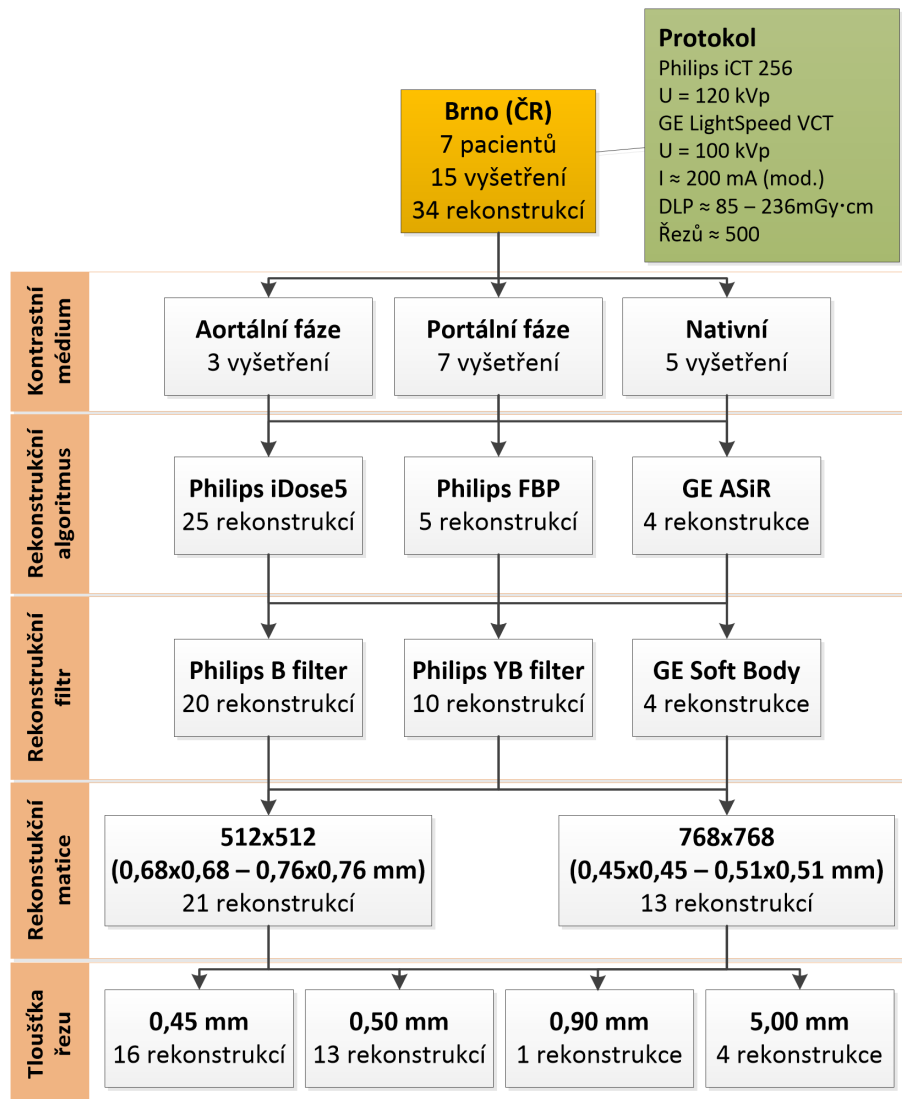
Součástí dizertační práce bylo vytvoření patientské databáze s expertním značením metastatických lézí. Během řešení byly získány databáze 25 onkologicky nemocných pacientů s celkově 49 vyšetřeními (časový vývoj) s experimentálním CT protokolem (Obr. 4.1), dále databáze 14 onkologicky nemocných pacientů se standardním protokolem pro snímání těla (*standard body protocol*) (Obr. 4.1), databáze 29 pacientů bez onkologických patologií se standardním protokolem, a databáze 7 pacientů s různými parametry rekonstrukce (Obr. 4.2). Všechny databáze byly anonymizovány a obsahují pro každé vyšetření informace o skenovacím protokolu, věku, pohlaví, případně druh a množství podané kontrastní látky, rekonstrukční parametry, případně informace o diagnóze primárního nádoru a aplikované léčbě. Data byla získána pomocí Philips Healthcare Brilliance iCT skeneru s 256vrstevným detektorem (případně GE LightSpeed VCT a Siemens SOMATOM Definition AS) z IRST Meldola, Itálie, dále z Fakultní nemocnice u svaté Anny v Brně, a Všeobecné fakultní nemocnice v Praze. Všechna data byla získána v souladu s Helsinskou deklarací Světová lékařská asociace (WMA²) aktualizované roku 2008 a s informovaným souhlasem pacientů. Experimentální data byla získána se souhlasem etické komise onkologického centra IRST a IRCCS Meldola, Prot. N. 343/V.3 z 10. ledna 2012.

¹z anglického Compute Unified Device Architecture

²z anglického World Medical Association



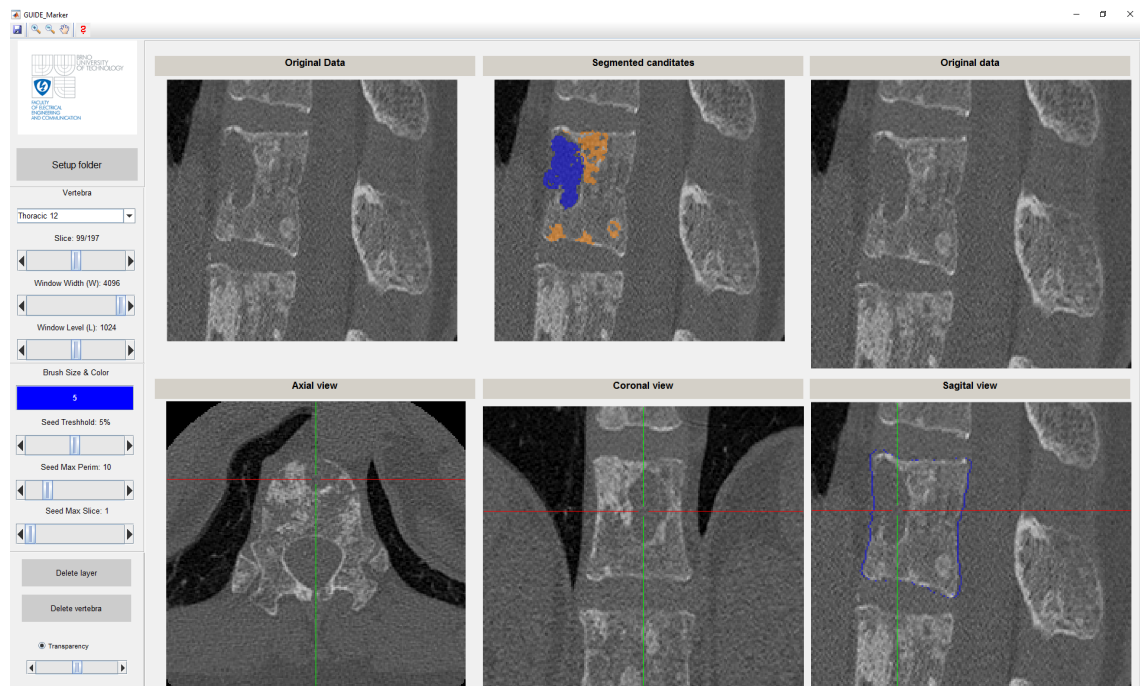
Obr. 4.1: Informace o patientských datech.



Obr. 4.2: Informace o patientských datech: CT data s různými rekonstrukčními parametry.

K realizaci značení obrazových dat lékařskými experty byl používán námi (ve spolupráci s Ing. Romanem Jakubíčkem) v rámci projektu navržený a realizovaný software „Marker“, který obsahuje prvky usnadňující expertní značení a byl vyvíjen dle požadavků lékařského personálu (Obr. 4.3). Součástí programu je možnost volby radiologického okna, možnost pohledu i značení v axiálním, koronálním i sagitálním řezu, možnosti změny velikosti značícího nástroje a možnost značení metodou narůstání oblastí. Cílem této části bylo vytvoření kompletní databáze (DATABÁZE A) obsahující voxelově označená data radiologickými experty. Tato databáze byla nezbytná pro analýzu efektivity segmentačních a klasifikačních algoritmů, jako učební podklad pro učení klasifikátorů s učitelem, jako nástroj pro získání mezi-expertních rozdílů a pro další statistické hodnocení z pohledu technického i lékařského.

DATABÁZE A byla dále rozdělena na trénovací (11 vyšetření, 220 obratlů) a testovací část (20 vyšetření, 406 obratlů) – všechny dosažené objektivní výsledky jsou vždy prezentovány na testovací části databáze.



Obr. 4.3: Ukázka softwaru „Marker“ pro expertní voxelové značení nádorových tkání v páteři.

Dalším krokem byl návrh a realizace softwaru „Validator“ umožňujícího dodatečnou validaci dosažených výsledků segmentace a klasifikace nádorových tkání. Koncepce programu je podobná „Markeru“, s tím rozdílem, že program pracuje s klasifikovanými tkáněmi jako s objekty (kandidáty) namísto jednotlivých voxelů a umožňuje expertům určité objekty přidat či odstranit.

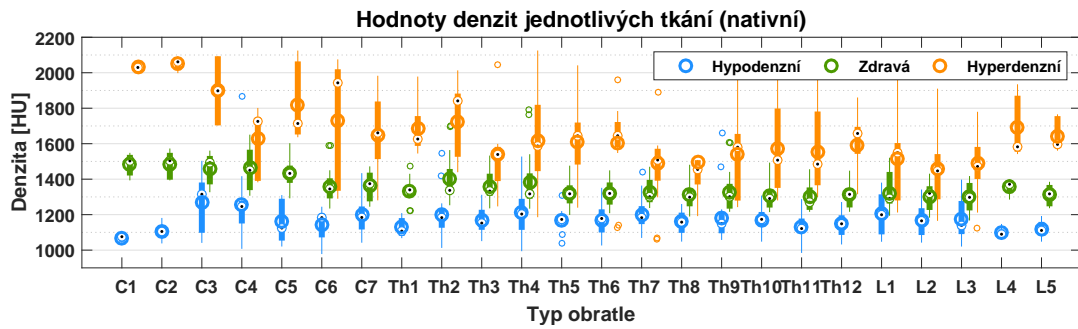
4.2 Přepřerování metod předcházejícího projektu

Během řešení projektu předchozími skupinami, na které tato práce navazuje, byly vyvinuty algoritmy pro segmentaci a klasifikaci lytických, sklerotických a zdravých kostních tkání pomocí třech metodik, z nichž první (Malinsky, 2012; Peter a kol., 2013; Jan a kol., 2015) byla založena na optimalizaci energie Markovských náhodných polí (Kap. 3.1.4) metodou *Modified Metropolis Dynamic* (Berthod a kol., 1996) s apriorní znalostí distribuce jasových hodnot pro každou z řešených tkání získaných z medicínsky označených dat (dále značeno jako **MRF** metoda). Druhá použitá metoda (Demel, 2014; Jan a kol., 2015) pracovala na základě dvoukrokového hledání optimálního řezu grafem (první krok odděluje sklerotickou (hyperdenzní) tkáň od ostatních, a druhý odlišuje zdravou od lytické (hypodenzní)), přičemž opět využíval apriorní statistické znalosti o distribuci jasových hodnot (dále značeno jako **GRAF** metoda). Poslední dříve navržená metoda (Novosadova, 2014; Jan a kol., 2015) řešila problém pomocí extrakce a ověření relevantnosti řady texturních a tvarových příznaků, které byly klasifikovány standardní třívrstvou neuronovou sítí, učenou na voxelově medicínsky označených datech (dále značeno jako **TEX** metoda).

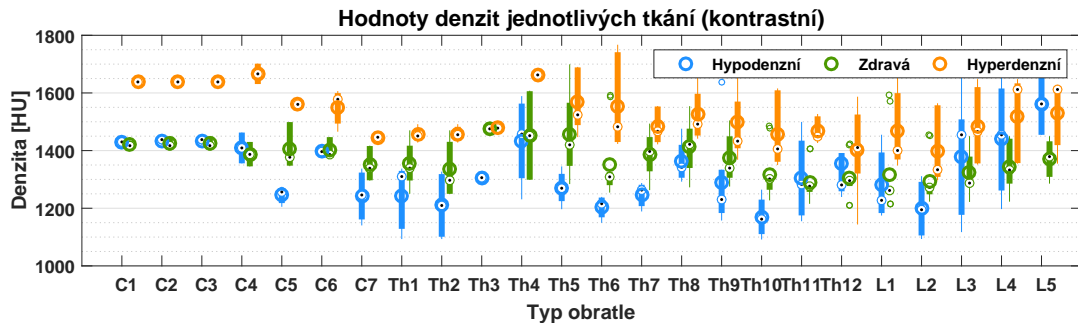
Z hlediska použitelnosti metod se všechny jevily jako efektivní, avšak nebyly ověřeny na větší patientské databázi. V době vzniku těchto tří metod byla dostupná pouze omezená patientské databáze a několik lékařsky označených dat (DATABÁZE 0a pro **MRF** a DATABÁZE 0b pro **GRAF** a **TEX** (viz Obr. 4.1)). Z tohoto důvodu bylo přistoupeno k rozšíření patientské databáze a doznačení dat více lékařskými experty (DATABÁZE A), aby bylo dostupné objektivní porovnání těchto metod a ověřena jejich využitelnost. Po otestování se ukázalo, že všechny tři metody dosahovaly na rozšířené databázi určitých výsledků (z nichž nejlépe se jevila metoda texturní a tvarové analýzy), které ale nebyly pro praktické využití dostatečné, a zároveň trpěly velkou výpočetní a časovou náročností, která byla z pohledu klinického použití neakceptovatelná (viz Tab. 4.1). Hlavním nedostatkem byl odhad apriorních znalostí a některých parametrů či postupů výpočtu pouze na základě omezené patientské databáze. Proto byly metody přepřerovány tak, aby jejich potenciál mohl být maximálně využit. Součástí změn byla úprava statistických parametrů na nově získané z DATABÁZE A (Obr. 4.4(a),(b)), přičemž byly zvlášť hodnoceni pacienti kterým nebyla během snímání aplikována kontrastní látka (CM³) (nativní), a zvlášť pacienti s CM (kontrastní).

Hodnoty denzit pro jednotlivé zájmové tkáně byly získány z expertně označených dat dvěma klinickými radiology celkem ze 31 vyšetření (26 bez CM a 5 s CM). Jak je vidět z Obr. 4.4, jsou zde patrné tři zásadní vlastnosti: za prvé, hodnoty mají poměrně velký rozptyl v rámci jednoho typu obratle (to může být způsobeno

³z anglického Contrast Media



(a)



(b)

Obr. 4.4: Karabicové grafy denzity jednotlivých tkání pro jednotlivé typy obratlů určené z DATABÁZE A: bez aplikace (a) a s aplikací (b) kontrastního média. Spodní/horní hrana boxu značí 25/75procentní percentil, spodní/horní vousy značí hodnoty pod/nad 25/75procentním percentilem, konce vousů minimální/maximální hodnotu z 1,5 násobku mezikvartilového rozpětí (mimo odlehlé hodnoty), terč uvnitř boxu značí medián, kružnice průměr a malé kružnice odlehlé hodnoty.

mezi-pacientskou, případně mezi-expertní variabilitou); za druhé, průměrné hodnoty i rozptyly se velmi liší v závislosti na typu obratlů (může být způsobeno, rozdílným jevem utvrzování rentgenového svazku v důsledku přítomnosti různých struktur – ramenní klouby, tuk, apod., dále skutečně rozdílnou densitou spongiózní části obratlů); za třetí, dochází k výraznému překryvu hodnot denzity mezi jednotlivými tkáněmi (zejména při pohledu na graf pro data s CM je patrné, že hodnoty lytických tkání jsou v důsledku CM výrazně zvýšené).

V závislosti na rozšířené databázi byly také upraveny další parametry dostupných metod a současně byly hledány možnosti optimalizace výpočetní náročnosti dostupných algoritmů.

Dalším problémem byla dostupnost třech různých přístupů, které v různých případech poskytovaly velmi odlišné výsledky, přičemž v některých z nich byly metody **MRF** a **GRAF** účinnější než klasifikátor **TEX**. Zároveň pro využití metody v klinické praxi je vhodné předložit pouze jeden výsledek místo třech odlišných, aby byla

zjednodušena jeho interpretace. Z těchto důvodů bylo přistoupeno k návrhu vhodné kombinace všech tří dostupných výsledků v jeden celek.

4.2.1 Úprava statistických modelů

Navržené metody **MRF** (Malinsky, 2012; Peter a kol., 2013; Jan a kol., 2015) a **GRAF** (Demel, 2014; Jan a kol., 2015) využívají ve svých algoritmech modely vytvořené na základě analýzy distribuce jasových hodnot v jednotlivých klasifikovaných tkáních.

MRF metoda zde využívá apriorní model v podobě funkcí s normálním rozložením, pomocí kterých jsou generovány pravděpodobnostní mapy p zařazení voxelu s danou jasovou hodnotou do dané klasifikační třídy L , které mohou být definovány jako

$$p(I_i|L) = \frac{1}{\sqrt{2\pi}\sigma_L} \cdot e^{-\frac{(I_i-\mu_L)^2}{2\sigma_L^2}}, \quad (4.1)$$

kde I_i je jasová hodnota i voxelu, μ_L je střední hodnota jasu dané třídy a σ_L je směrodatná odchylka jasových hodnot pro danou třídu. Pro vytvoření pravděpodobnostních map p je tedy zapotřebí apriorní znalost všech parametrů μ_L a σ_L . V navrženém algoritmu podle autora Malinsky (2012) byly tyto parametry pro zdravou trabekulární tkáň experimentálně odvozeny z databáze 114 dostupných obrázků z CT snímků (DATABÁZE 0a – Obr. 4.1). Zároveň byly tyto parametry experimentálně odhadnuty pro každý typ obratle ($C3 - L4$) zvlášť dle doporučení autorů Whyne a kol. (2007). Parametry byly odhadnuty pro oba typy nádorových tkání zvlášť. Úprava metody zde spočívala v aktualizaci těchto parametrů podle nově získané DATABÁZE A dle Obr. 4.4.

Metoda **GRAF** využívá stejného apriorního modelu daného funkcí s normálním rozložením s parametry μ_L a σ_L jako metoda **MRF**. Tento model je zde využit pro výpočet regionálního vztahu $R(A)$. Výsledná segmentace je dána dvoukrokově, přičemž v prvním kroku je segmentována sklerotická tkáň od ostatních tkání. Regionální vztahy jsou tedy v prvním kroku dány jako

$$R_i(1) = -\ln p(I_i|L_{hyper}), \quad (4.2)$$

$$R_i(0) = -\ln p(I_i|L_{hypo}) - \ln p(I_i|L_{healthy}). \quad (4.3)$$

Ve druhém kroku je segmentována tkáň lytická od zdravé, přičemž vstupním obrazem je již pouze oblast ostatních tkání získaná v předchozím kroku. Obdobně jsou tedy regionální vztahy ve druhém kroku dány jako

$$R_i(1) = -\ln p(I_i|L_{hypo}), \quad (4.4)$$

$$R_i(0) = -\ln p(I_i|L_{healthy}). \quad (4.5)$$

Úprava metody zde spočívala opět v aktualizaci parametrů podle nově získané DATABÁZE A dle Obr. 4.4.

4.2.2 Snížení výpočetní náročnosti

Za účelem klinického využití dostupných metod bylo přistoupeno k úpravám, směřujícím ke snížení výpočetní náročnosti, a tím také ke zkrácení doby potřebné pro získání výsledků aplikovaných metod. Výrazného urychlení bylo dosaženo u metod MRF a TEX přepracováním následujících algoritmů:

Metoda využívající Markovských náhodných polí (MRF)

V navrženém algoritmu (Malinsky, 2012; Peter a kol., 2013; Jan a kol., 2015) bylo využito Rov. 3.9 z Kap. 3.1.4 pro výpočet globální energie Markovského náhodného pole, přičemž pokud globální energii rozdělíme na sumy lokálních energií ve vztahu k pozičním souřadnicím počítaného voxelu ve 3D datech s omezením na systém okolí druhého řádu (26okolí voxelu) a s uvážením pouze jednovoxelových a párových „kliků“, potom získáme vztah

$$U(f) = \sum_{i \in \mathcal{S}} V_1(f_i) + \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{N}_i} V_2(f_i, f_{i'}), \quad (4.6)$$

kde U je energie MRF, V_1 je potenciál jednovoxelového „kliku“, V_2 potenciál párového „kliku“, i je index voxelu z množiny všech voxelů v obraze \mathcal{S} , i' je index voxelu z 26okolí \mathcal{N}_i kolem voxelu i , přičemž $i \notin \mathcal{N}_i$. Pokud bude uvažován víceúrovňový logistický model (MLL⁴) (Elliott a kol., 1984; Derin a Cole, 1986; Derin a Elliott, 1987; Li, 2009), potom potenciál V_1 bude definován jako

$$V_1(f_i) = \alpha_l \quad \text{pokud} \quad f_i = l \in \mathcal{L}_d, \quad (4.7)$$

kde α_l je potenciál pro daný voxel i , zařazený do třídy nabývající hodnoty l , s indexem třídy d z množiny tříd \mathcal{L} . V našem případě je index třídy d shodný s hodnotou l pro daný index ($d = \mathcal{L}_d$). V praxi α_l vyjadřuje penalizaci/zvýhodnění zařazení voxelu do dané třídy. Dále bude potenciál V_2 definován jako

$$V_2(f_i, f_{i'}) = \begin{cases} \beta_c & \text{pokud} \quad f_i = f_{i'} \\ -\beta_c & \text{pokud} \quad f_i \neq f_{i'} \end{cases}, \quad (4.8)$$

kde β_c je parametr penalizující rozdílné hodnoty indexu třídy mezi počítaným voxellem i a voxellem i' v jeho okolí v daném kliku c v určitém směru, a naopak zvýhodňující hodnoty stejné. V praxi je tento parametr označován jako parametr homogenity.

⁴z anglického Multi-Level Logistic model

Obecně je možné matici energií $U(f)$ získat postupným procházením všech voxelů a jejich lokálních okolí, což bylo provedeno v původní verzi algoritmu. Značného urychlení bylo dosaženo úpravou části výpočtu V_2 klikového potenciálu s využitím maticových operací a 3D konvoluce se speciálně navrženou konvoluční maskou. Tento postup může být formalizován takto:

$$\sum_{i' \in \mathcal{N}_i} V_2(f_i, f_{i'}) = \sum_{d \in \mathcal{L}_d}^D (C_d * h_\beta) \odot C_d, \quad (4.9)$$

kde C_d je binární matice o velikost f , obsahující nenulové hodnoty na pozicích, kde $f_i = d$, D je celkový počet tříd a h_β je $3 \times 3 \times 3$ konvoluční maska v jednotlivých řezech z definovaná jako

$$h_\beta^{z_1} = 2 \cdot \begin{pmatrix} \beta_{12} & \beta_8 & \beta_{10} \\ \beta_6 & \beta_5 & \beta_7 \\ \beta_{11} & \beta_9 & \beta_{13} \end{pmatrix} h_\beta^{z_2} = 2 \cdot \begin{pmatrix} \beta_4 & \beta_2 & \beta_3 \\ \beta_1 & -\sum \beta_c & \beta_1 \\ \beta_3 & \beta_2 & \beta_4 \end{pmatrix} h_\beta^{z_3} = 2 \cdot \begin{pmatrix} \beta_{13} & \beta_9 & \beta_{11} \\ \beta_7 & \beta_5 & \beta_6 \\ \beta_{10} & \beta_8 & \beta_{12} \end{pmatrix}, \quad (4.10)$$

kteřá pro 26okolí obsahuje 13 parametrů homogenity pro různé směry. Jak je vidět za Rov. 4.9, tento postup výpočtu je vhodný zejména pro aplikace s malým počtem uvažovaných tříd (v této aplikaci 3 třídy), protože s jejich narůstajícím počtem bude úměrně stoupat výpočetní náročnost v závislosti na zvyšujícím se počtu aplikovaných operací 3D konvoluce. Obecně lze také tento postup využít i při aplikacích s vyšším řádem systému okolí, případně s víceprvkovými kliky, kdy by byla přidána vždy obdobná konvoluční maska větších rozměrů (např. $5 \times 5 \times 5$) obsahující další parametry. Současně s narůstající velikostí masky nebo aplikací na obrazová data větších rozměrů by bylo možné využít výpočtu 3D konvoluce ve frekvenční oblasti. Další části výpočtu a optimalizace zůstali oproti metodám autorů Malinsky (2012); Peter a kol. (2013) a Jan a kol. (2015) nezměněny. Konkrétní hodnoty zrychlení jsou uvedeny v Kap. 4.2.4.

Metoda využívající texturní a tvarové analýzy (TEX)

Navržená metoda (Novosadova, 2014; Jan a kol., 2015) využívá ke klasifikaci voxelů neuronovou síť, jejímž vstupem je vektor lokálně počítaných texturních a tvarových příznaků (viz Kap. 3.1). Kromě paralelizace jednotlivých výpočtů bylo zásadního urychlení dosaženo v části algoritmu pro extrakci tvarových příznaků. Původní postup vytváří oddělené objekty z originálních 3D obrazových šedotónových dat na základě iterativní segmentace 3D metodou rozvodí (*watershed*) a následného slučování jednotlivých segmentovaných oblastí podle jejich jasové podobnosti. Tvarové příznaky jsou poté odvozeny pro každý takto vzniklý objekt a jejich hodnoty přiřazeny každému voxelu náležícímu do daného objektu.

Vzhledem k časové náročnosti několikrát opakované 3D metody rozvodí a slučování oblastí, bylo přistoupeno k vytvoření aproximovaných objektů s využitím filtrace obrazových dat dolnopropustním Gaussovským 3D filtrem ve frekvenční oblasti, podkvantování obrazových dat, a následnou analýzou spojených komponent (CCA⁵) (He a kol., 2009). Konkrétní hodnoty zrychlení jsou uvedeny v Kap. 4.2.4.

Současně s urychlením výpočtu byly také upraveny některé použité texturní příznaky, kde v původní verzi byl použit pouze jeden směr výpočtu GLRM (Kap. 3.1), v upravené verzi tři směry – axiální, koronální a sagitální. Dále byla změněna velikost masky použité pro výpočet lokálních spektrálních vlastností z velikosti $3 \times 3 \times 3$ na $9 \times 9 \times 9$, čímž bylo zlepšeno frekvenční rozlišení těchto příznaků za cenu horšího prostorového rozlišení.

S provedenými úpravami použitých příznaků byly současně naučeny nové neuronové sítě pro klasifikaci voxelů do tříd odpovídajícím zájmovým patologiím (osteolitická, osteoblastická, zdravá tkáň).

4.2.3 Kombinace předchozích metod

Z důvodů klinické využitelnosti a jednoznačnosti interpretace výsledků získaných všemi třemi upravenými metodami bylo přistoupeno ke kombinaci jejich výsledků (metoda **CMB**). Pro tyto účely byly navrženy dvě slučovací metody; první metoda využívá algoritmu AdaBoostM2 (viz Kap. 3.2.4), a druhá logického většinového pravidla.

Využití algoritmu AdaBoostM2

Pro kombinaci dostupných výsledků bylo využito algoritmu AdaBoostM2 pro klasifikaci do více tříd, přičemž vstupem do algoritmu byl vektor třech výstupů předchozích klasifikátorů spolu s expertně určenou klasifikací pro každý prvek učební množiny. Jako slabé klasifikátory byly zvoleny rozhodovací stromy (*decision trees*). Učební množina obsahovala 100 000 náhodně vybraných voxelů z patientské databáze s rovnoměrným zastoupením klasifikovaných tkání. Naučeno bylo celkem 1 500 klasifikátorů se změnou permutace učební množiny a změnami v počtu slabých klasifikátorů v rozmezí 1 – 150, a každý byl testován na rovnoměrně zastoupené testovací množině o velikosti 1 000 000 náhodně vybraných prvků mimo učební množinu. Vybrán byl klasifikátor s největší úspěšností.

⁵connected component analysis

Využití většinového pravidla – modus

Druhou možností kombinace metod bylo jednoduché logické rozhodnutí na základě většinového pravidla:

1. pokud 2 ze 3 nebo všechny klasifikátory mají na výstupu stejnou třídu, je tato třída finálním výsledkem,
2. pokud má každý z klasifikátorů na výstupu jinou třídu, je za finální výsledek považována volba z AdaBoostM2 algoritmu.

4.2.4 Dosažené výsledky – přepracované verze původních metod

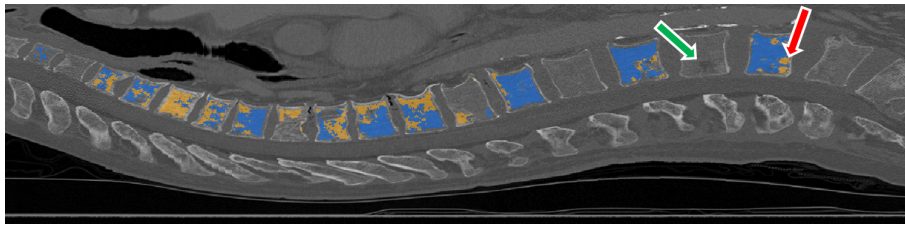
Výše popsané metody dosahovaly bez jakýchkoliv změn po otestování na DATA-BÁZI A výsledků, které se jeví jako nevhodné pro použití v klinické praxi, jak podle dosažených hodnot senzitivity a specificity dle voxelově založené metodiky (viz Kap. 4.6.1), tak i z hlediska časové náročnosti výpočtů, což je patrné z Tab. 4.1.

Tab. 4.1: Porovnání výsledků původních a přepracovaných verzí dostupných metod (voxelově) – průměrné hodnoty. Jednotlivé sloupcové bloky rozdělují tři různé přístupy, sloupce poté určují typ tkáně a průměrnou hodnotu přes všechny tkáně. Na řádcích jsou postupně průměrné hodnoty senzitivity (TPR), specificity (TNR), Youdenova indexu (YI) a času výpočtu na pacienta. Řádky jsou rozděleny na původní metody (pův.) a metody přepracované v rámci dizertační práce (pře.).

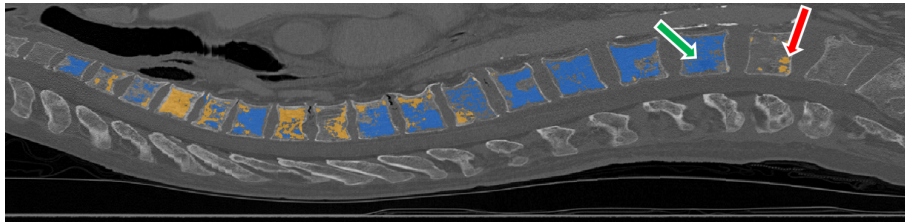
Metoda	MRF				GRAF				TEX			
	Tkáň	<i>Hypo</i>	<i>Hyper</i>	<i>Zdravá</i>	<i>Průměr</i>	<i>Hypo</i>	<i>Hyper</i>	<i>Zdravá</i>	<i>Průměr</i>	<i>Hypo</i>	<i>Hyper</i>	<i>Zdravá</i>
TPR pův. [%]	10,5	24,0	94,3	42,9	20,1	19,8	44,2	28,0	30,9	16,2	62,7	36,6
TPR pře. [%]	29,4	41,6	76,2	49,1	32,7	39,8	79,3	50,6	32,2	35,0	81,0	49,4
TNR pův. [%]	99,0	95,7	20,9	71,9	74,2	73,5	70,7	72,8	80,9	83,4	57,2	73,8
TNR pře. [%]	83,8	93,7	38,8	72,2	86,6	94,0	39,1	73,2	87,5	94,3	38,0	73,2
YI pův. [-]	9,5	19,7	15,2	14,8	-5,7	-6,7	14,9	0,8	11,8	-0,4	19,9	10,4
YI pře. [-]	13,2	35,3	15,0	21,2	19,3	33,8	18,4	23,8	19,7	29,3	19,0	22,7
Čas/pac. pův.	42,1 min				6,0 min				13,1 hod			
Čas/pac. pře.	1,5 min				6,0 min				6,6 hod			

Podle dosažených hodnot je patrné, že nejrychlejším původním algoritmem byla metoda **GRAF**, v průměru nejúspěšnější metodou byla metoda **MRF**. Z tabulky je dále patrné, že po přepracování algoritmů došlo ke značnému snížení výpočetní náročnosti metod **MRF** a **TEX**, za současného zvýšení úspěšnosti všech tří metod. Vzhledem k vzájemně velmi podobným úspěšnostem přepracovaných metod byla nejvhodnější samostatná metoda **MRF** pro její nízkou náročnost. Jednotlivé metody však přes podobné úspěšnosti vykazovaly subjektivně odlišné výsledky segmentace a klasifikace, což je patrné z ukázkových sagitálních výsledků na Obr. 4.5(a),(b),(c). Z tohoto důvodu byla provedena kombinace všech metod (**CMB**), čímž došlo k nepatrnému zvýšení úspěšnosti (viz Tab. 4.2) oproti jednotlivým samostatným metodám.

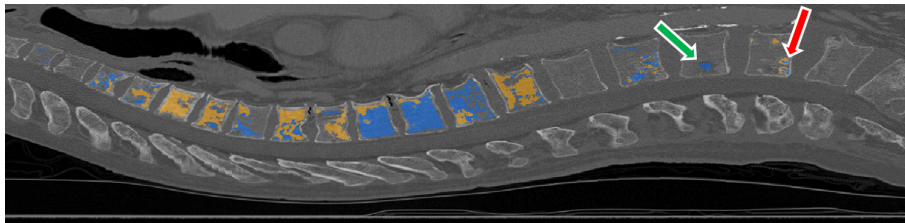
Ukázka tkání, které byly klasifikovány metodou **CMB** lépe je patrné na Obr. 4.5(d) v porovnání s expertně označenými daty (Obr. 4.5(e)). Zelená šipka zde ukazuje lytickou lézi, která byla správně segmentována a klasifikována pouze metodou **TEX** (metoda **MRF** tuto lézi nesegmentuje vůbec; metoda **GRAF** segmentuje a klasifikuje jako lytickou tkáň chybně téměř celý obratel), na druhou stranu červená šipka ukazuje chybnou segmentaci lytické léze uvnitř lézí blastických opět metodou **TEX** (metoda **GRAF** tyto léze segmentuje bez chybných lytických lézí; metoda **MRF** provede jejich segmentaci správně, avšak s chybnou lytickou tkání v jejich okolí).



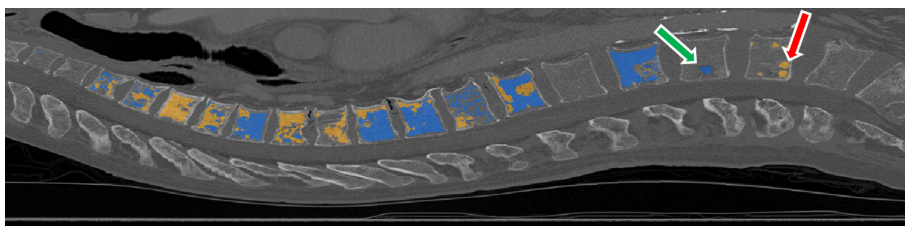
(a)



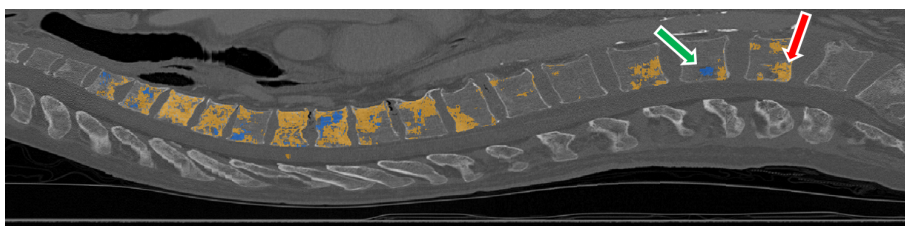
(b)



(c)



(d)



(e)

Obr. 4.5: Ukázka segmentace a klasifikace nádorových tkání přepracovanými metodami (DATABÁZE A – pacient 4): **MRF** (a), **GRAF** (b), **TEX** (c), jejich kombinace pomocí AdaBoostM2 algoritmu s většinovým rozhodováním **CMB** (d) a ukázka expertně označených dat expertem 1 (e) (modrá – lytické tkáně, oranžová – blastické tkáně, šipky – příklady rozdílů v detekcích mezi metodami).

V případě použití metody **CMB** jsou oba tyto případy vyhodnoceny správně. Cenou je však velká výpočetní náročnost, která je nyní složena z časů všech tří metod navýšeným o čas potřebný pro jejich kombinaci (viz Tab. 4.2).

Tab. 4.2: Dosažené výsledky kombinací metod (voxelově) – průměrné hodnoty. Sloupce určují typ tkáně a průměrnou hodnotu přes všechny tkáně. Na řádcích jsou postupně průměrné hodnoty senzitivity (TPR), specifity (TNR), Youndenova indexu (YI) a času výpočtu na pacienta.

	CMB			
Tkáň	<i>Hypo</i>	<i>Hyper</i>	<i>Zdravá</i>	<i>Průměr</i>
TPR [%]	33,2	41,1	78,5	50,9
TNR [%]	86,0	93,7	40,5	73,4
YI [-]	19,2	34,8	19,0	24,3
Čas/pac.	1, 5 min + 6, 0 min + 6, 6 hod + 0, 6 min \approx 6, 7 hod			

4.2.5 Závěr k výsledkům přepracovaných původních metod

V první fázi práce byla vytvořena patientská DATABÁZE A (Kap. 4.1) obsahující 31 CT vyšetření, která byla dále označena dvěma nezávislými lékařskými experty se zaměřením na radiologii a onkologii pomocí navrženého softwaru „Marker“ (Kap. 4.1). Dále byly dříve vyvinuté algoritmy **MRF** (využívající Markovských náhodných polí), **GRAF** (využívající optimálního řezu grafem) a **TEX** (využívající texturní a tvarovou analýzu) otestovány na této databázi, přičemž bylo dokázáno, že nejsou pro přímé klinické použití vhodné z důvodu vysoké výpočetní náročnosti a poměrně nízké úspěšnosti segmentace a klasifikace hledaných lézí.

Pro zlepšení jejich využitelnosti bylo přistoupeno k úpravě parametrů těchto algoritmů na základě DATABÁZE A, vedoucí na zlepšení jejich úspěšnosti a výraznému snížení výpočetní náročnosti (viz Tab. 4.1). Vzhledem k dosaženým výsledkům jednotlivých přepracovaných metod, a jejich značné subjektivní rozdílnosti (metody dosahují srovnatelných hodnot úspěšnosti, avšak při pohledu na výsledné výstupy segmentace a klasifikace dosahují rozdílných výsledků – každá nalezne jinou lézi nebo má falešně pozitivní nálezy na různých pozicích), byl navržen postup pro kombinaci jejich výstupů v jediný výsledek s využitím AdaBoostM2 algoritmu a logické kombinace většinou volbou (**CMB**). Touto kombinací bylo dosaženo mírného zlepšení úspěšnosti, avšak za cenu nutnosti výpočtu výsledků pomocí všech tří přepracovaných přístupů (viz Tab. 4.2).

I přes provedené úpravy nebylo dosaženo dostatečné úspěšnosti segmentace a klasifikace (průměr (voxelově): senzitivita 50, 9 %, specifita 73, 4 %, Youndenův index

24, 3) v době vhodné pro rutinní klinické využití (průměrně 6, 7 *hod/pac.*). Tato fáze však byla nutná ke zlepšení náhledu na danou problematiku. Ukázalo se, že jak samotná jasová hodnota v obrazových datech, tak různé texturní a tvarové příznaky přinášejí důležitou informaci potřebnou pro úspěšnou segmentaci a klasifikaci kostních nádorových lézí. Dalším důležitým poznatkem byl přístup, jakým lékaři hodnotí hledané léze – spíše globálnějším pohledem, než po jednotlivých voxelech. Také bylo zjištěno, že není možné využít exaktního výběru obrazových příznaků na základě expertních znalostí; daná problematika je příliš neurčitá, aby bylo možné vytvořit automatický expertní systém na základě určitých přesně definovaných pravidel.

4.3 Fuzzy k-means rychlá klasifikace lézí

Vzhledem k požadavkům spolupracujících klinických lékařů na nízkou časovou náročnost algoritmu pro segmentaci a klasifikaci nádorových tkání byla navržena metoda pracující pouze s jasovými hodnotami obrazu. Podle poznatků získaných z předchozích metod z Kap. 4.2 lze očekávat, že při zahrnutí pouze jasových hodnot nebude dosaženo takové úspěšnosti jako s metodami využívající další tvarové a texturní příznaky, a proto byla tato metoda navržena pouze pro předběžné vyhodnocení tkání. Podle výsledku rychlé metody může lékař zhodnotit, zda má smysl aplikovat některou náročnější a přesnější metodu, nebo zda je tento méně přesný výsledek pro jeho potřeby dostatečný.

Pro tyto účely byl navržen algoritmus využívající fuzzy k-means (**FCM**) podle Rov. 3.10, který byl dále částečně upraven podle autorů Ahmed a kol. (2002). Výsledná kritériální funkce pro klasifikaci, která je tvořena navíc jasovou informací z okolí daného voxelu, je dána jako:

$$J_m = \sum_{k=1}^c \sum_{i=1}^N u_{k,i}^p \|y_i - v_k\|^2 + \frac{\alpha}{N_R} \cdot \sum_{k=1}^c \sum_{i=1}^N u_{k,i}^p \left(\sum_{y_r \in \mathcal{N}_i} \|y_r - v_k\|^2 \right), \quad (4.11)$$

kde \mathcal{N}_i je množina prvků z definovaného okolí prvku y_i , y_r je prvek z množiny \mathcal{N}_i , N_R je počet prvků v množině \mathcal{N}_i a α je parametr řídicí vliv okolí. Optimum kritériální funkce je nalezeno iterativně, kde odhady v jednotlivých krocích jsou odvozeny pomocí Lagrangeových multiplikátorů. Postup výpočtu je následující:

1. volba inicializačních hodnot průměrů $\mathbf{v}_{it=0}$,
2. výpočet odhadu funkcí příslušnosti \mathbf{u}_{it} ,
3. výpočet nových pozic průměrů \mathbf{v}_{it} , z odhadnutých funkcí příslušnosti \mathbf{u}_{it} ,
4. kontrola zastavovacích podmínek $\|\mathbf{v}_{it} - \mathbf{v}_{it-1}\| < \epsilon$ a $it < M$.

Kroky 2 až 4 se opakují, dokud nedojde ke splnění alespoň jedné ze zastavovacích podmínek. Hodnota ϵ je malé kladné číslo odpovídající Eukleidovské vzdálenosti

mezi pozicemi průměrů v_k v aktuálním a předchozím iteračním kroku it , a M je kladné celé číslo vyjadřující maximální počet iteračních kroků.

Pro účely řešení problému segmentace a klasifikace nádorových tkání byl nastaven parametr $c = 4$, tedy rozdělení do čtyř klasifikačních tříd (pozadí, lytická, zdravá a sklerotická tkáň). Vstupem do algoritmu jsou objemová CT data (Obr. 4.6(a)), která jsou nejprve převedena z Hounsfieldových jednotek do intervalu $\langle 0, 1 \rangle$ podělením bitovou hloubkou CT dat a filtrována (Obr. 4.6(b)) pomocí 3D filtru typu dolní propust se symetrickou Gaussovskou impulsní charakteristikou (se směrodatnou odchylkou σ a velikostí impulsní charakteristiky K). Dále jsou předzpracovaná data maskována binární maskou těla obratle (v okolí každého obratle je hodnota nula); masky obratlů jsou získány algoritmem podle autorů Jakubicek (2020); Jakubicek a kol. (2019b, 2020, 2019a). Vymaskovaná data odpovídají vstupu \mathbf{y} (Obr. 4.6(c)). Inicializační pozice průměrů $\mathbf{v}_{it=0}$ byly nastaveny jako průměrné hodnoty z vytvořeného apriorního jasového modelu s normálním rozložením z Kap. 4.2 (Obr. 4.4(a) a Obr. 4.4(b)). Do výpočtu bylo zahrnuta 26okolí každého voxelu. Další parametry metody byly nastaveny experimentálně podle dosažených výsledků a jsou dostupné v Tab. 4.3. Finální klasifikační výsledek (Obr. 4.6(h)) je získán tak, že na odpovídající pozici je uložena hodnota třídy s maximální hodnotou příslušnosti k dané třídě (Obr. 4.6(d)–(g)). Vzhledem k možnosti záměny pozic jednotlivých průměrů během iteračního cyklu je na konci výpočtu provedeno ověření správnosti pozic průměrů, podle podmínky:

$$v_1 < v_2 < v_3 < v_4. \quad (4.12)$$

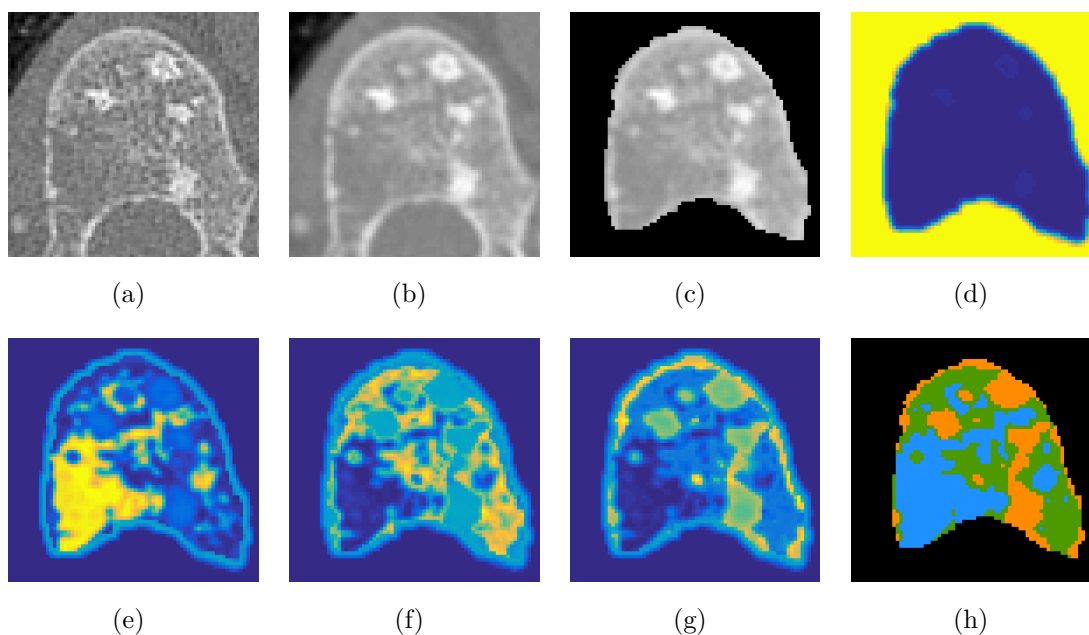
Pokud tato podmínka není splněna, jsou následně výsledné třídy vzájemně přeskládány. Senzitivita a specificita metody může být řízena pomocí prahování matic příslušností namísto výběru třídy s maximální příslušností.

Tab. 4.3: Tabulka parametrů rychlé klasifikační metody

Parametr	c	p	N_R	α	ϵ	M	σ	K
Hodnota	4	2	26	2	0,03	4	0,85	5

4.3.1 Dosažené výsledky – fuzzy k-means

Vzhledem k požadavkům na velkou rychlost výpočtu byla navržena a testována metoda založená na fuzzy k-means nehierarchickém shlukování, která využívá stejné apriorní statistické modely jako metody **MRF** a **GRAF** z Kap. 4.2 (Obr. 4.4(a) a Obr. 4.4(b)). Navíc byla přidána znalost o jasových hodnotách okolních voxelů. Výsledky dosažené navrženou metodou **FCM** jsou patrné v Tab. 4.4, kde oproti



Obr. 4.6: Ukázka segmentace a klasifikace nádorových tkání rychlou fuzzy k-means metodou: vstupní data (a), data po předzpracování (b), maskovaná data (c), matice příslušnosti třídy pozadí (modrá – 0, žlutá – 1) (d), lytické tkáně (e), zdravé tkáně (f), sklerotické tkáně (g) a výsledný výstup (černá – pozadí, modrá – hypodenzní, zelená – zdravá, oranžová – hyperdenzní) (h).

kombinované **CMB** metodě stoupla průměrná senzitivita přibližně o 9 % a specifická o 1 %.

Tento nárůst samotný není nijak výrazný, avšak důležitý je zejména nárůst senzitivity pro lytickou (ze 33,2 % na 69,7 %) a sklerotickou (ze 41,1 % na 55,0 %) tkáň, i za cenu snížení senzitivity pro zdravou (ze 78,5 % na 54,8 %) tkáň, což zvyšuje šanci na odhalení nádorových tkání během rychlé segmentace. Nejdůležitějším přínosem této metody je značná rychlost (48 s na pacienta), umožňující využití metody v klinické praxi.

4.3.2 Závěr k metodě fuzzy k-means

V další fázi práce byl navržen a realizován algoritmus pro rychlou segmentaci a klasifikaci nádorových tkání (**FCM**), založený na fuzzy k-means algoritmu využívající znalosti o jasových hodnotách okolních voxelů a statistického modelu rozložení denzity jednotlivých tkání v obrátcích. Testování bylo provedeno na DATABÁZI A (Kap. 4.1), kde bylo dosaženo průměrných hodnot (voxelové) senzitivity 59,8 %, specifity 74,8 % a Youdenova indexu 34,7. Spolu s dosaženou průměrnou dobou výpočtu 48 s/pac. je tento algoritmus vhodný jako nástroj pro předběžnou přehled-

Tab. 4.4: Dosažené výsledky fuzzy k-means (voxelově) – průměrné hodnoty. Sloupce určují typ tkáně a průměrnou hodnotu přes všechny tkáně. Na řádcích jsou postupně průměrné hodnoty senzitivity (TPR), specificity (TNR), Youdenova indexu (YI) a času výpočtu na pacienta.

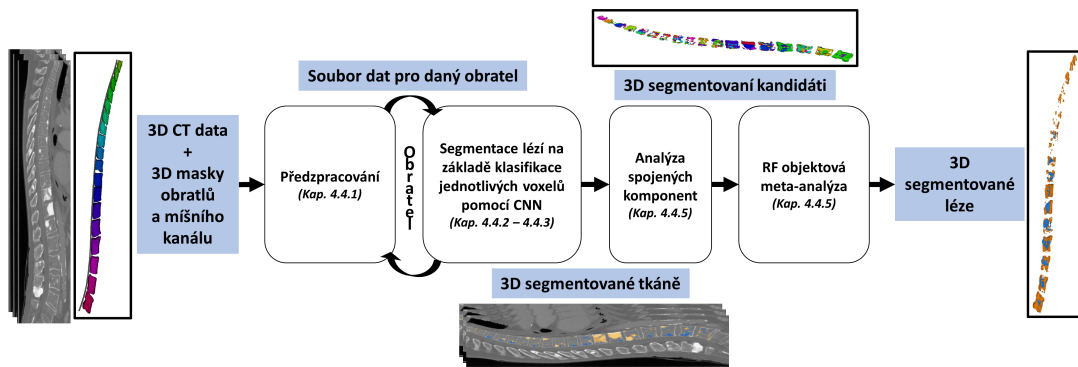
	FCM			
Tkáň	<i>Hypo</i>	<i>Hyper</i>	<i>Zdravá</i>	<i>Průměr</i>
TPR [%]	69,7	55,0	54,8	59,8
TNR [%]	69,8	87,8	66,9	74,8
YI [-]	39,5	42,8	21,7	34,7
Čas/pac.	48 s			

vou segmentaci a klasifikaci zájmových tkání. Z hlediska přesné segmentace ovšem stále nedosahoval požadovaných výsledků úspěšnosti.

4.4 Klasifikace lézí konvoluční neuronovou sítí

Zjištění získaná během řešení projektu vedla k návrhu metody (CNN) využívající hlubokého učení (*Deep Learning*), konkrétně hluboké konvoluční neuronové sítě (CNN⁶). Prvním důvodem byla nedostatečná informace pro klasifikaci voxelu (případně jeho blízkého okolí) pouze na základě jeho jasové hodnoty, a současně předpokládaná výhoda použití metod využívajících texturních a tvarových příznaků (viz Kap. 4.2.5 a Kap. 4.3.2). Druhým důvodem byl expertně komplexní a obtížně definovatelný problém klasifikace. Po konzultaci s lékařskými experty bylo zjištěno, že na daný problém nahlíží z globálnějšího hlediska než je klasifikace jednotlivých voxelů, a zároveň hodnotí, kromě jasové hodnoty, různá kritéria (pozice léze, celkový stav obratle, apod.), která však nemohou být všechna konkrétně explicitně popsána (například pro vytvoření expertního systému nebo návrhu konkrétních příznaků). Třetím důvodem byla dostupnost expertně označených dat, která lze s výhodou využít pro učení a ověření neuronové sítě s učitelem. Předpokladem tedy bylo, že CNN si na základě dostupných učebních dat během učebního procesu automaticky vytvoří sady příznaků potřebné pro vhodnou klasifikaci dat. Obecné blokové schéma navržené metodiky je znázorněno na Obr. 4.7, přičemž jednotlivé bloky jsou popsány detailně v následujících podkapitolách.

⁶z anglického Convolutional Neural Network



Obr. 4.7: Blokové schéma navržené metody využívající CNN klasifikátor a meta-analýzu lézí. Vstupními daty jsou původní 3D CT data a současně 3D masky obratlů a míšního kanálu poskytnuté algoritmem Romana Jakubíčka (Jakubicek, 2020).

4.4.1 Předzpracování obrazových dat

Před dalším použitím jakýchkoliv CT dat bylo provedeno předzpracování obrazových dat (Obr. 4.8) složené z kroků:

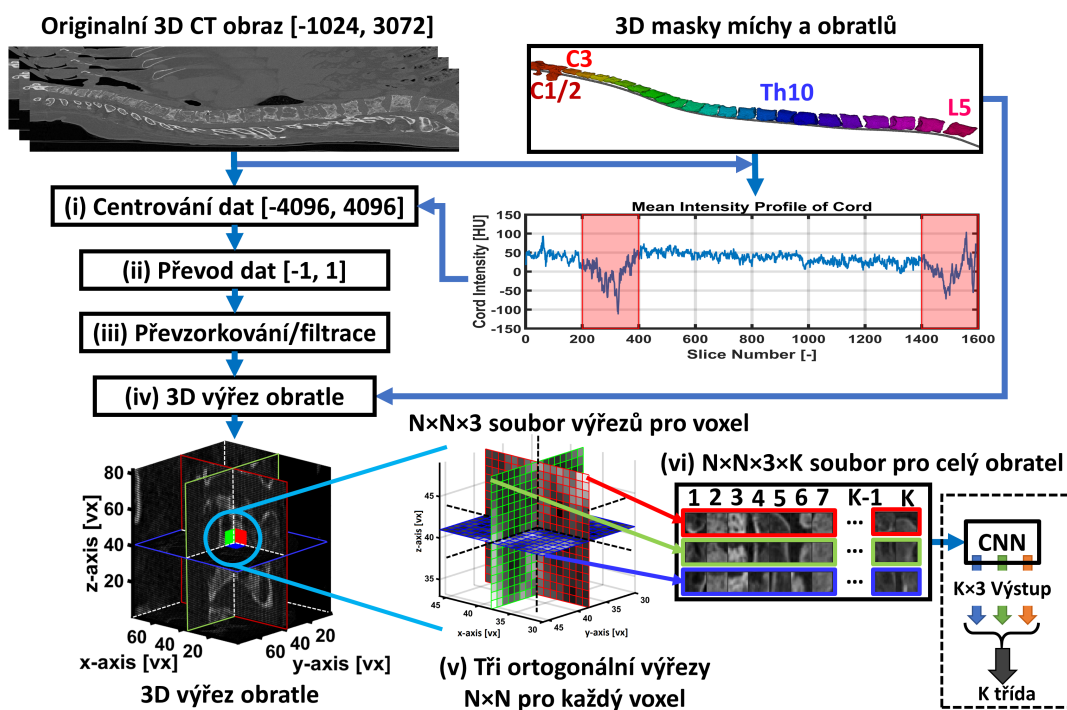
- (i) centrování dat na nulovou hodnotu v oblasti míchy odečtením její průměrné hodnoty – toto je prováděno individuálně pro každý řez CT dat. Mícha byla zvolena pro její nejstabilnější hodnoty Hounsfieldových jednotek (HU^7) z dostupných tkání a současně je prostorově nejbližší tkání k analyzovaným obratlům (nejvhodnější by byla tkáň zdravé kostní tkáně v obratli, ale její určení není z důvodu patologických změn v této fázi možné). Tím mohou být částečně potlačeny vlivy jevu utvrzování svazku a globální rozdíly mezi pacienty,
- (ii) normalizace dat do intervalu $\langle -1, 1 \rangle$ dělením hodnot bitovou hloubkou snímaných dat (není upravován tvar histogramu),
- (iii) filtrace dat dolní propustí typu 3D izotropický Gaussovský filtr (velikost $5 \times 5 \times 5$ se směrodatnou odchylkou $\sigma = 0,75$) – voleno experimentálně jako kompromis dostatečného potlačení šumu v CT datech při zachování detailů textury potřebných pro extrakci příznaků (v případě odlišného rozlišení analyzovaných dat od učebních není tento filtr využit – dochází zde k převzorkování dat na stejný fyzikální rozměr voxelu v závislosti na rozlišení, přičemž dochází k interpolaci kubickým splajnem),
- (iv) kvádrový výřez daného obratle (každý obratel je zpracováván individuálně),
- (v) vzhledem k velké výpočetní náročnosti následného učení CNN, i s ohledem na počet vah sítě, která je také značně závislá na počtu vstupních hodnot, byly místo objemových dat použity pouze tři vzájemně ortogonální 2D řezy,

⁷z anglického Hounsfield Unit

příčměž klasifikovaný voxel leží ve středu těchto řezů (viz Obr. 4.9). Pro každý klasifikovaný voxel je tedy vstupem $N \times N \times 3$ voxelů místo $N \times N \times N$, kde N je velikost okolí klasifikovaného voxelu. Toto okolí bylo heuristicky zvoleno ($N = 16$) kompromisem mezi dostatečně globálním pohledem s dostatkem informace, výpočetní náročností, velikostí obrazových dat a faktem, že účelem výsledné sítě je segmentace obrazu pomocí klasifikace jednotlivých voxelů,

- (vi) a následné seřazení těchto trojic ortogonálních výřezů do *mini-batch* s velikostí K (*mini-batch* obsahuje K náhodných vzorků učební množiny – řádově desítky až stovky).

Z důvodu obecné využitelnosti naučené CNN během vybavovací fáze jsou data vždy převzorkována dle aktuálního rozlišení CT dat tak, aby fyzikální rozměry odpovídaly datům použitým k učení sítě (pokud je rozlišení původních dat horší než učebních, jsou data nadvzorkována, a naopak – toto je prováděno pouze pro funkčnost vytvořené sítě a přirozeně nemůže dojít k doplnění neexistujících informací; např. léze menší než nativní rozlišení dat nemohou být detekovány).



Obr. 4.8: Blokové schéma předzpracování dat před vstupem do CNN. Římské číslice odpovídají bodům v Kap. 4.4.1. Blok CNN je dále popsán detailněji v Kap. 4.4.2.

4.4.2 Návrh původní architektury specializované sítě

Při návrhu vhodného řešení byla uvažována řada existujících standardních architektur CNN a jejich předučených verzí, jako např. AlexNet (Krizhevsky a kol., 2012),

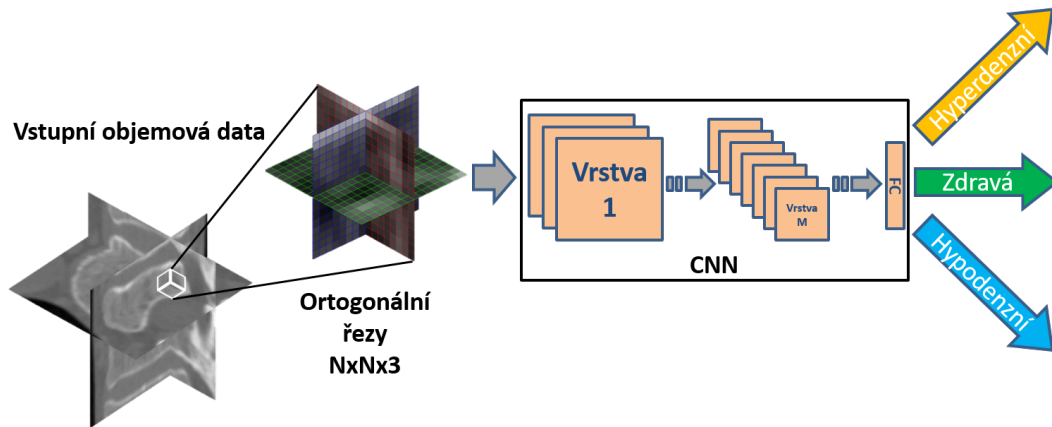
VGG (Simonyan a Zisserman, 2014), GoogLeNet (Szegedy a kol., 2015), ResNet (He a kol., 2016), Inception-v4 (Szegedy a kol., 2017) a podobné. Tyto architektury byly převedeny na databázích 2D RGB obrazů formátu $N \times N \times 3$ pro účely detekce objektů v obrazech. Těchto architektury by se teoreticky dalo využít i pro naše účely, a to tak, že naše vzájemně ortogonální řezy ze 3D CT dat by byly vkládány na místo RGB kanálů. Problémem však je, že vyžadovaná vstupní velikost výřezů je řádově ve stovkách pixelů, což je příliš mnoho v poměru k dostupným CT datům, kde celá obratlová těla mají rozměr řádově v desítkách voxelů v každém směru (v závislosti na rozlišení CT dat a typu obratle). Tento problém by bylo možné obejít například převzorkováním CT dat na odpovídající velikost nebo doplněním okolí výřezu nulami, nicméně to by vedlo k mnohonásobně vyšším výpočetním nárokům. Navíc, zmíněné architektury obsahují velké množství vah a vyžadují mnoho výpočetních operací během jednoho dopředného průchodu dat sítí, např. síť VGG-19 (Simonyan a Zisserman, 2014) obsahuje $\sim 125 \cdot 10^6$ parametrů (vah) a provede $\sim 39 \cdot 10^9$ operací násobení a sčítání (detailní rozbor nároků jednotlivých architektury byl publikován autory Bianco a kol. (2018)).

Další možností by bylo využití architektury přímo navržených pro segmentaci/regresi obrazů (tzv. *image-to-image*), jako je například U-Net (Ronneberger a kol., 2015), avšak tyto sítě umožňují využít pouze 2D vstupní data. Z našich experimentů vycházela senzitivita testované sítě typu U-Net v rozmezí od 16% do 66% a specifická od 6% do 31% v závislosti na typu lézí, učební databázi, a dalším nastavení architektury (počet úrovní, počty filtrů, učební algoritmus, augmentace dat, atd.). Během testování se tedy ukázalo, že objemová informace je pro správnou segmentaci nepostradatelná, a proto tyto architektury sítě nebyly dále využity. Existují také 3D varianty těchto architektury (např. Čižek a kol. (2016)), kdy vstupem jsou přímo objemová data, avšak paměťové a výpočetní nároky těchto architektury jsou enormní a pro následné využití v klinické praxi jsou prozatím z těchto důvodů nepoužitelné. Dalším problémem je zde dostupnost vhodných učebních dat – v případě, že by vstupem byl celý objem jednoho obratle, tak by bylo k dispozici pouze 1.264 obratlů a z toho méně než 626 s metastatickými lézemi. Navíc při rozdělení na trénovací a testovací data by bylo k učení sítě reálně použito například pouze 400 obratlů.

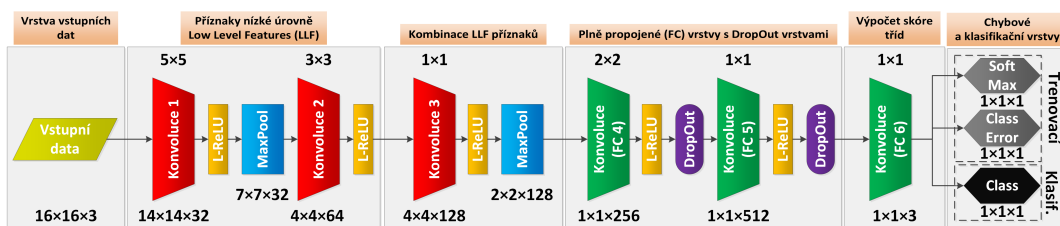
Z těchto důvodů bylo přistoupeno k návrhu, učení a testování vlastní původní specializované architektury CNN, které je popsána dále. Výsledná síť má pouze $\sim 850 \cdot 10^3$ parametrů a provede $\sim 5 \cdot 10^6$ operací během jednoho dopředného průchodu, což je v obou případech řádově méně než u známých převedených sítí.

Podle daných předpokladů byla navržena CNN pro klasifikaci jednotlivých voxelů předzpracovaných dat do třech tříd (hypodenzní, zdravá, hyperdenzní) podle rámcového schématu na Obr. 4.9. Síť byla navržena v programovém prostředí MATLAB®

s využitím volně dostupných knihoven MatConvNet (Vedaldi a Lenc, 2015). Během řešení práce vznikla a byla testována řada různých architektur, přičemž výsledná navržená síť je zobrazena na Obr. 4.10.



Obr. 4.9: Rámcové schéma CNN klasifikátoru pro jeden klasifikovaný voxel: zleva – vstupní objemová data po základním předzpracování, tři vzájemně ortogonální řezy $N \times N$ (klasifikovaný voxel leží uprostřed), M vrstvá konvoluční neuronová síť, výstupy klasifikátoru.



Obr. 4.10: Detailní schéma navržené konvoluční neuronové sítě pro klasifikaci voxelů.

Navržená CNN je tvořena 17 vrstvami pro fázi učení (14 pro fázi vybavování). Vstupní obrazová data obsahují pro každý klasifikovaný voxel tři ortogonální řezy o velikosti 16×16 , jež se ukázaly dostatečné k charakterizaci struktury v okolí hodnoceného voxelu. Tato data dále prochází prvními pěti vrstvami, kde dochází k extrakci tzv. nízkourovňových obrazových příznaků (LLF⁸). Zde je opakovaně využit standardní sled operací konvoluce, ReLU⁹ aktivační funkce a operace slučování (*pooling*) pomocí lokálního maximového filtru. Vzhledem k relativně malým vstupním datům je operace slučování použita pouze po prvních vrstvách, aby bylo možno využít výhod více vrstev. V této části CNN jsou využity konvoluční masky s velikostmi 5×5 pro první konvoluční vrstvu a 3×3 pro druhou, jejíž aplikací dochází

⁸z anglického Low Level Features

⁹z anglického Rectified Linear Unit

ke zmenšení dat v osách x a y . Současně je v první konvoluční vrstvě použito 64 různých masek, 128 ve druhé; tímto je zajištěna tvorba různých nízkourovňových příznaků a narůstání dat ve směru hloubky.

Druhou částí CNN je další kombinace konvoluční, ReLU aktivační a slučovací vrstvy, avšak s tím rozdílem, že konvoluční vrstva je složena z 256 masek o velikosti 1×1 . Toto uspořádání má za úkol vytvořit příznaky, které vzniknou kombinací předchozích nízkourovňových příznaků mezi sebou. Po samotné konvoluci zde tedy nedojde ke zmenšení dat v osách x a y , avšak dojde k jejich nárůstu do hloubky. Výstupem z této části jsou tedy složitější obrazové příznaky (HLF¹⁰) s vyšší úrovní abstrakce.

Další část sítě je složena ze dvou sledů konvoluce, aktivace a *DropOut* vrstvy. Zvláštností těchto konvolučních vrstev je stejná velikost jejich konvolučních masek jako je velikost vstupních dat. Jedná se tedy v podstatě o plně propojené (FC¹¹) vrstvy, stejně jako u klasických neuronových sítí. Jejich výstupem je vždy příznakový vektor hodnot, velikostí odpovídající počtu použitých masek. *DropOut* vrstvy mají za úkol „vypínat“ určité procento vstupů do následující vrstvy sítě; které vstupy mají být vypnuty je náhodně měněno v každé epoše učení. Výhodou je, že během učebního procesu je síť využívána pouze část vytvářených příznaků, což pomáhá regulovat jev přeučení sítě a zvyšuje její výslednou robustnost. Jinými slovy síť je učena na náhodně zvolených „neúplných“ informacích, což vede ke snížené efektivitě sítě během jejího učení. V globálním měřítku je výsledná naučená síť ekvivalentem velkého množství menších sítí. Přínos se poté projeví v lepším zobecnění sítě, jež mohla být učena po více epoch bez přeučení, kde pro vybavení sítě jsou poté již využity všechny dostupné informace (obě *DropOut* vrstvy jsou vyřazeny).

V další vrstvě sítě je použita konvoluční vrstva opět jako FC vrstva, přičemž je navržena jako standardní klasifikátor. Vstupem do sítě je celý výstupní příznakový vektor z předchozí části a výstupem 3 hodnoty (skóre) získané konvolucí se třemi maskami (v podstatě vrstva se třemi neurony). Nejvyšší z jejich výstupních hodnot odpovídá výsledné klasifikační třídě.

Poslední dvě ze tří vrstev jsou využity pouze v procesu učení, kde *SoftMax* vrstva obsahuje funkci pro výpočet normalizované pravděpodobnosti příslušnosti vstupu k dané klasifikační třídě. Dále je tato hodnota logaritmována přirozeným logaritmem a násobena hodnotou -1 , čímž je získána hodnota vzájemné entropie (*cross-entropy*) mezi vstupem a danou klasifikační třídou. Tato funkce je použita jako kritériální během učebního procesu a je hledáno její minimum. *ClassErr* vrstva má pouze informační charakter a slouží pro kontrolu učebního procesu; je zde vypočtena relativní chyba klasifikace sítě během každé epochy učení. Třetí vrstva *Class*

¹⁰z anglického High Level Features

¹¹z anglického Fully Connected

je naopak použita pouze ve fázi vybavování sítě. Zde je možné místo maxima určit výstup sítě pomocí tří prahů a preferovat tak určitou klasifikační třídu/y nad ostatními – tímto je možné se pohybovat po výsledné ROC¹² křivce klasifikátoru – a zvyšovat tak senzitivitu klasifikátoru jedné jedné třídy na úkor specificity ke stejné třídě a senzitivity k třídám ostatním.

Pro všechny ReLU aktivační funkce byla použita úprava na tzv. *Leaky* ReLU, kde pro záporné vstupní hodnoty není výstupní hodnota nulová, ale závislost má mírný záporný sklon daný konstantou. Tato úprava omezuje množství „zemřelých“ jednotek v důsledku násobení zpětně šířené chyby nulou, a tím vymizení gradientu. Detailní popis hyperparametrů jednotlivých vrstev sítě, včetně velikostí dat v každé vrstvě, je uveden v Tab. 4.5.

¹²z anglického Receiver Operating Characteristic

Tab. 4.5: Tabulka hyperparametrů vrstev konvoluční neuronové sítě

Vrstva	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Typ	Konv.	ReLU	Pool	Konv.	ReLU	Konv.	ReLU	Pool	Konv.	ReLU	DropOut	Konv.	ReLU	DropOut	Konv.	SoftMax	ClassErr
Vstup	16x16x3	14x14x32	14x14x32	7x7x32	4x4x64	4x4x64	4x4x128	4x4x128	2x2x128	1x1x256	1x1x256	1x1x256	1x1x512	1x1x512	1x1x512	1x1x3	1x1x3
Výstup	14x14x64	14x14x32	7x7x32	4x4x64	4x4x64	4x4x128	4x4x128	2x2x128	1x1x256	1x1x256	1x1x256	1x1x512	1x1x512	1x1x512	1x1x3	1x1x1	1x1x1
Velikost masky	5x5	-	2x2	3x3	-	1x1	-	2x2	2x2	-	-	1x1	-	-	1x1	-	-
Počet masek	32	-	-	64	-	128	-	-	256	-	-	512	-	-	3	-	-
Krok	1	-	2	2	-	1	-	2	1	-	-	1	-	-	1	-	-
Rozšíření	1	-	0	1	-	0	-	0	0	-	-	0	-	-	0	-	-
Modifikátor učebního kroku vah	1	-	-	1	-	1	-	-	0,50	-	-	0,25	-	-	0,10	-	-
Modifikátor učebního kroku prahu	2	-	-	2	-	2	-	-	1	-	-	0,50	-	-	0,20	-	-
Další	-	Leak: 0,005	max	-	Leak: 0,005	-	Leak: 0,005	max	-	Leak: 0,005	Rate: 0,6	-	Leak: 0,005	Rate: 0,6	-	-	-

4.4.3 Učení sítě a optimalizace parametrů

Pro učení CNN byl zvolen algoritmus *Mini-Batch Gradient Descent with Momentum* (viz Kap. 3.2.6) pro jeho jednoduchost a snadnou využitelnost v kombinaci s grafickou výpočetní jednotkou (GPU¹³) – paralelizace výpočtu a říditelné paměťové nároky. Učení bylo realizováno v programovém prostředí MATLAB[®] s využitím volně dostupných knihoven MatConvNet (Vedaldi a Lenc, 2015). Dále byla použita kriteriální funkce vzájemné entropie s $L2$ regularizací vah pro zabránění přeučení sítě. Všechny váhy byly inicializovány náhodně s normálním rozložením s nulovou střední hodnotou a rozptylem (dle doporučení při použití ReLU aktivačních funkcí podle He a kol. (2015)), daným vztahem

$$\sigma_l^2 = \sqrt{2/n_l}, \quad (4.13)$$

kde n_l je celková velikost vstupního vektoru z předchozí vrstvy a l je index dané vrstvy sítě. Toto nastavení zrychluje konvergenci učebního procesu a zajišťuje homogenní rozložení aktivací ve všech vrstvách sítě, zejména u velmi hlubokých architektur. Všechny hodnoty prahů byly na počátku učení nastaveny na nulu. Hyperparametry hybnosti μ , $L2$ regularizace λ a počátečního učebního kroku α_0 byly optimalizovány na základě náhodného prohledávání parametrického prostoru během učebního procesu: učení probíhalo po dobu 100 epoch, byly náhodně vygenerovány množiny těchto hyperparametrů z omezených intervalů, a následně hodnocena úspěšnost naučené sítě po 100 epochách a celkové chování sítě během učení dle průběhu její klasifikační chyby. Tento proces byl poté opakován s menším rozptylem parametrů kolem nalezeného optimálního nastavení z předchozího kroku. Velikost sady výřezů použité pro jednu úpravu vah (*mini-batch*) β byla zvolena s ohledem na paměťové nároky. Optimální nalezené hyperparametry jsou uvedeny v Tab. 4.6.

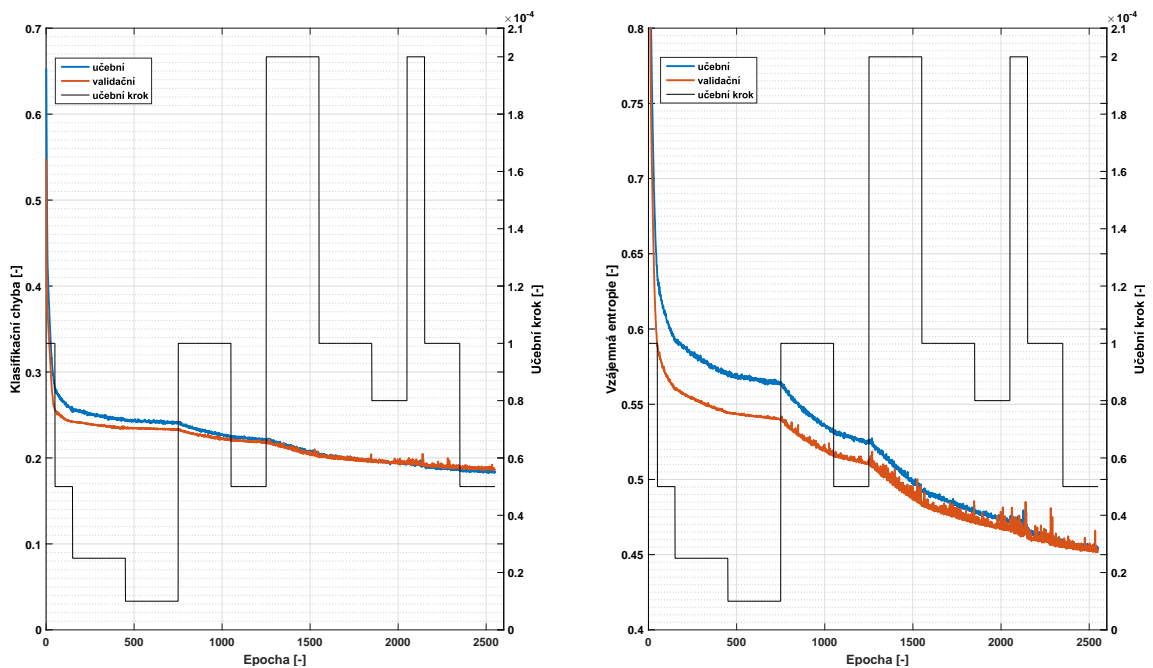
Tab. 4.6: Tabulka globálních hyperparametrů konvoluční neuronové sítě

Parametr	λ	μ	α_0	β
Hodnota	0,002	0,97	10^{-4}	512

Učební krok pro úpravu prahů neuronů byl nastaven na dvojnásobek standardního kroku α pro úpravu vah. Pro FC vrstvy sítě byl učební krok poměrově snížen oproti standardnímu kroku, a to z důvodu průběhu úpravy vah v jednotlivých vrstvách během učení – vzhledem k potřebě vytvoření masek filtrů pro extrakci příznaků prvními konvolučními vrstvami je vhodné zpočátku upravovat zpětným šířením zejména tyto váhy bez výrazných změn vah klasifikačních FC vrstev. Jakmile jsou váhy v prvních vrstvách téměř neměnné, je možné doučit váhy v posledních

¹³z anglického Graphical Processor Unit

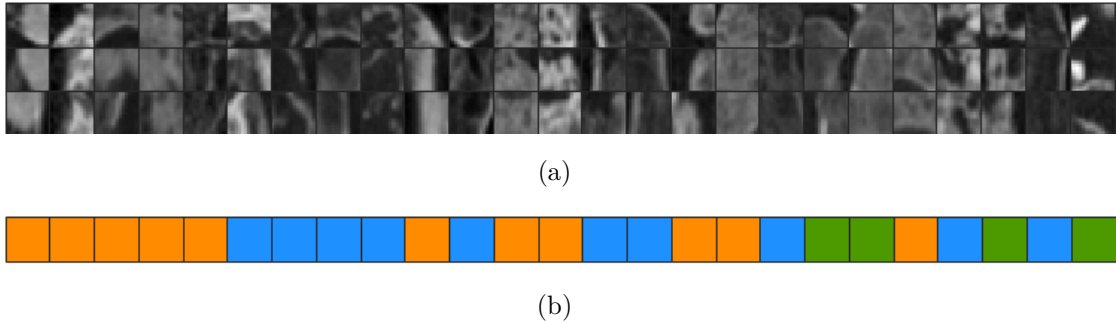
vrstvách na základě téměř stabilních výstupních příznaků z předchozích vrstev. Toto by bylo možné provést také dvoukrokově – tedy nejprve naučit vrstvy extrahující příznaky a poté na tyto příznaky naučit klasifikační vrstvy – nicméně postup se sníženým učebním krokem ve výstupních vrstvách je méně časově náročný. Hyperparametry specifické pro jednotlivé vrstvy sítě jsou uvedeny v Tab. 4.5. Učební krok α byl v průběhu učení vždy po několika epochách experimentálně snižován, případně zvyšován podle chování průběhu hodnot kritériální funkce. Snižování učebního kroku umožňuje dosáhnout přesnější pozice hledaného extrému, zatímco zvýšení učebního kroku vede k rychlejší konvergenci a umožňuje opustit lokální extrémy. Průběhy chybové a kritériální funkce během epoch učení včetně aktuálního učebního kroku jsou znázorněny na Obr. 4.11. Z obrázku je patrné vždy postupné snižování kroku α po několika epochách; jakmile dojde ke stagnaci hodnot kritériální funkce, je krok skokově zvýšen. Dále je patrný vliv *DropOut* vrstev v jejichž důsledku je chyba na validační množině téměř po celou dobu učení netradičně o něco menší než na množině učební (pro učební množinu je použita pouze část příznaků pro klasifikaci). Učení bylo zastaveno jakmile hodnota kritériální funkce učební množiny začala dosahovat stejné hodnoty jako na validační množině.



Obr. 4.11: Grafy průběhu učení: průběh klasifikační chyby během jednotlivých epoch a velikost učebního kroku v dané epoše (vlevo), průběh hodnot kritériální funkce během jednotlivých epoch a velikost učebního kroku v dané epoše (vpravo).

Učební množina použitá ve všech fázích učení byla tvořena 120 000 náhodně vybranými vzorky z trénovací části DATABÁZE A. Množina obsahovala 1/3 vzorků

od každé klasifikované tkáně; byly vybírány pouze vzorky, kde byla shoda obou lékařských expertů. Data byla dále náhodně rozdělena na trénovací (60%), validační (20%) a testovací (20%) podmnožiny. Trénovací a validační data byla použita během učebního procesu, testovací po naučení sítě k prvotnímu ověření její funkčnosti. Ukázka části učební podmnožiny je znázorněna na Obr. 4.12.



Obr. 4.12: Ukázka učební množiny: 25 vstupních vzorků předzpracovaných ortogonálních řezů (slupce jsou jednotlivé vzorky, řádky odpovídají ortogonálním řezům) (a) a odpovídajících očekávaných výstupů (modrá – lytická, zelená – zdravá a oranžová – sklerotická tkáň) (b).

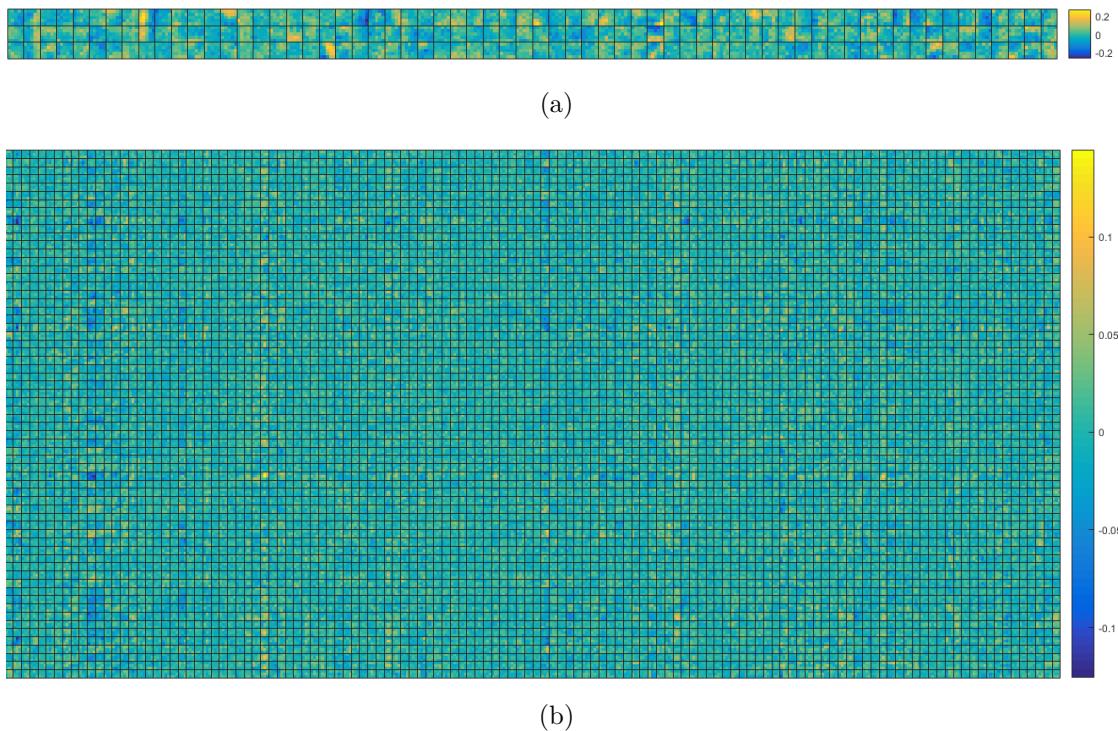
Pro snížení výpočetního času pro učení a vybavování sítě byl algoritmus optimalizován pro využití grafických jader s využitím NVIDIA® CUDA toolkitu, NVIDIA® cuDNN knihoven pro hluboké učení a MatConvNet toolboxu.

4.4.4 Interpretace neuronové sítě

Subjektivně byla posuzována naučená CNN z hlediska vizualizace filtrů. Na ukázce (Obr. 4.13(a)) 64 filtrů $5 \times 5 \times 3$ v první konvoluční vrstvě naučené CNN je patrný charakter použitých učebních dat; interpretace filtrů není v tomto případě jednoznačná (jako v případě běžných fotografií), avšak můžeme pozorovat jisté „tvarové“ vlastnosti klasifikovaných dat. Filtry v dalších vrstvách jsou méně interpretovatelné, protože jejich účelem je filtrovat a propojovat extrahované parametrické mapy z předchozích vrstev. Ukázka naučených filtrů z druhé konvoluční vrstvy je na Obr. 4.13(b), kde je 128 filtrů $3 \times 3 \times 64$.

Další ukázka (Obr. 4.14) znázorňuje obrazy získané při maximální aktivaci daného neuronu v dané vrstvě sítě. K získání těchto obrazů bylo využito řízeného zpětného šíření chyby (GBP¹⁴), s využitím operací dekonvoluce a nadzvorkování (*unpooling*). Na Obr. 4.14(a) jsou ukázány odhady nejjednodušších příznaků odpovídající obrazům, které by teoreticky maximálně aktivovaly jednotlivé neurony v první konvoluční vrstvě (jejich počet odpovídá počtu filtrů v dané vrstvě, velikost obrazů

¹⁴z anglického Guided Back-Propagation

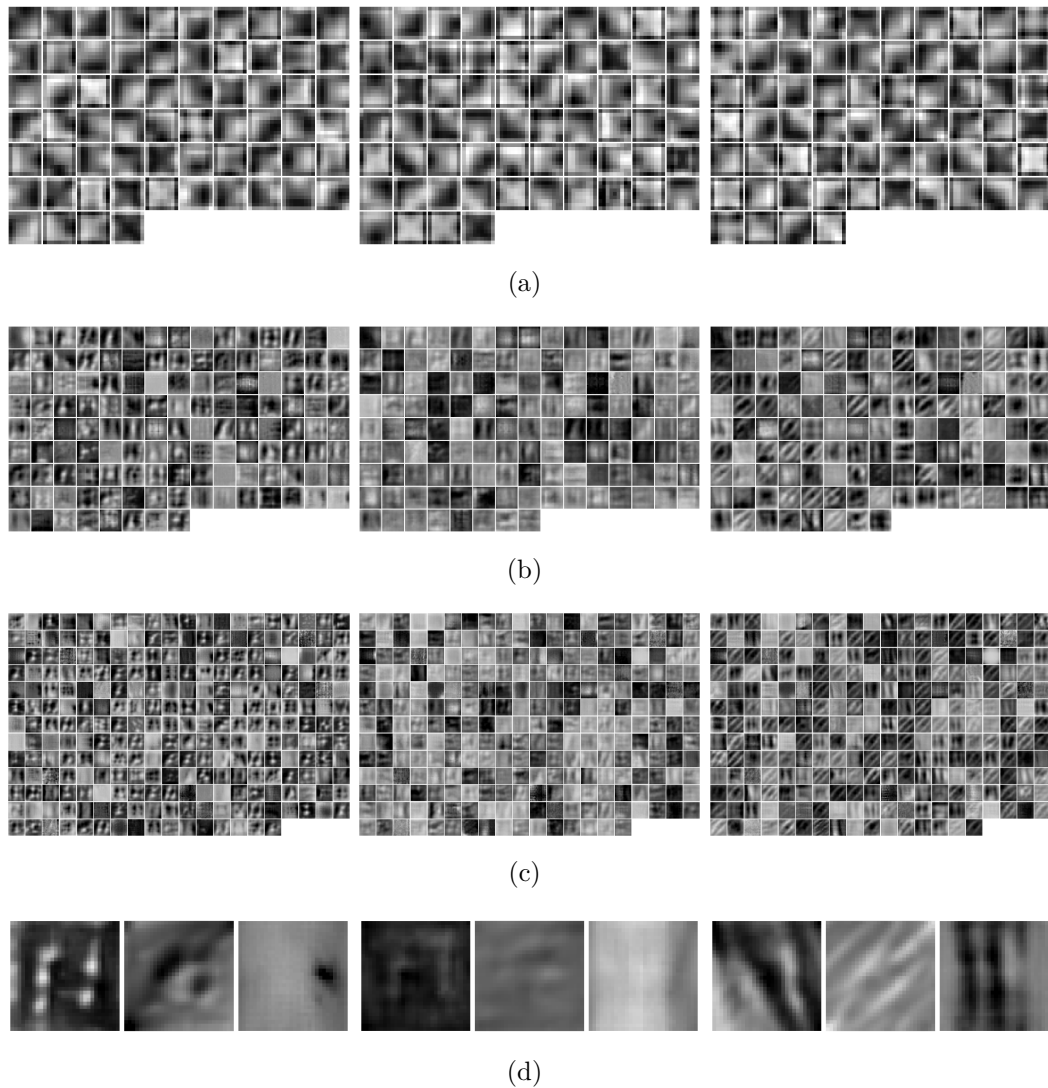


Obr. 4.13: Vizualizace naučených filtrů CNN: první konvoluční vrstva (a) a druhá konvoluční vrstva (b) (sloupce odpovídají jednotlivým maskám filtrů, řádky jednotlivým ortogonálním vstupům (a) / vstupním vrstvám z předchozí vrstvy (b)).

poté velikosti receptivního pole aktivovaného neuronu). V každé vrstvě jsou zobrazeny 3 sady příznaků, vzhledem k ortogonálnímu uspořádání vstupních dat (pro RGB (barevné) obrazy se s výhodou tyto příznaky zobrazují v barevném provedení znázorňujícím také barevné vlastnosti). Jak je patrné, v první vrstvě největší odezva neuronů odpovídá jednoduchým směrovým přechodům s různou orientací – tedy zejména různě orientovaným hranám, rohům, apod. V dalších konvolučních vrstvách (Obr. 4.14(b) a Obr. 4.14(c)) jsou poté extrahovány složitější příznaky vzniklé kombinacemi předchozích jednodušších – tedy například různé představy kortikální části obratle, téměř homogenních struktur, textur s trabekulárním charakterem, nízkodenzní homogenní struktury, kapkovité vysokodenzní struktury, apod. Obr. 4.14(d) zobrazuje struktury, které nejvíce aktivují neurony v poslední FC vrstvě.

4.4.5 Dosažené výsledky – konvoluční neuronová síť

Pro konečnou segmentaci a klasifikaci nádorových tkání byla navržena metoda využívající konvoluční neuronové sítě. Učení probíhalo na množině 120 000 náhodně vybraných voxelových vzorků z trénovací části DATABÁZE A s rovnoměrným zastoupením všech tří klasifikovaných tkání. Objemově toto množství odpovídá přibližně



Obr. 4.14: Vizualizace příznaků naučené CNN: aktivace neuronů v první konvoluční vrstvě (a), druhé konvoluční vrstvě (b), třetí konvoluční vrstvě (c), a poslední klasifikační vrstvě (d) (pro každou vrstvu jsou ukázány 3 sady příznaků; vlevo axiální, uprostřed sagitální, vpravo koronární ortogonální vrstva). Každý příznak odpovídá obrazu získaného při maximální aktivaci daného neuronu v dané vrstvě sítě.

pouze jednomu průměrnému hrudnímu obratli. Bylo dosaženo průměrné voxelové senzitivity 82,2 % na trénovací části (60 % z celé množiny), 81,3 % na validační části (20 % z celé množiny) a 81,2 % na testovací části (20 % z celé množiny). Na testovací části množiny byla senzitivita 87,8 % pro lytickou, 87,9 % pro sklerotickou a 67,8 % pro zdravou tkáň. Tyto výsledky sloužily pouze pro výběr nejvhodnějších naučených sítí, které byly dále testovány na testovací části DATABÁZE A, a následně byla vybrána síť s nejlepší úspěšností. Dosažené výsledky na reálné validační databázi jsou shrnuty v Tab. 4.7, kde dle předpokladu došlo ke snížení senzitivity

oproti testování na učební množině přibližně o 10 až 15 %. Zároveň je patrné, že v porovnání se všemi metodami z předchozích kapitol bylo dosaženo výrazně lepší senzitivity i specifity u všech zájmových tkání. Odpovídající ROC křivka je na Obr. 4.15. Současně je časová náročnost výpočtu při využití grafických procesorů dostatečně nízká pro klinické použití (hodnota v závorce platí při využití pouze jader počítačového procesoru).

Tab. 4.7: Dosažené výsledky CNN (voxelově) – průměrné hodnoty. Sloupce určují typ tkáně a průměrnou hodnotu přes všechny tkáně. Na řádcích jsou postupně průměrné hodnoty senzitivity (TPR), specifity (TNR), Youdenova indexu (YI) a času výpočtu na pacienta. Hodnota času v závorce byla získána bez využití paralelizace na grafickém procesoru.

	CNN			
Tkáň	<i>Hypo</i>	<i>Hyper</i>	<i>Zdravá</i>	<i>Průměr</i>
TPR [%]	71,0	74,2	61,5	68,9
TNR [%]	88,0	82,2	78,5	82,9
YI [-]	59,0	56,4	40,0	51,8
Čas/pac.	4 min (19 min)			

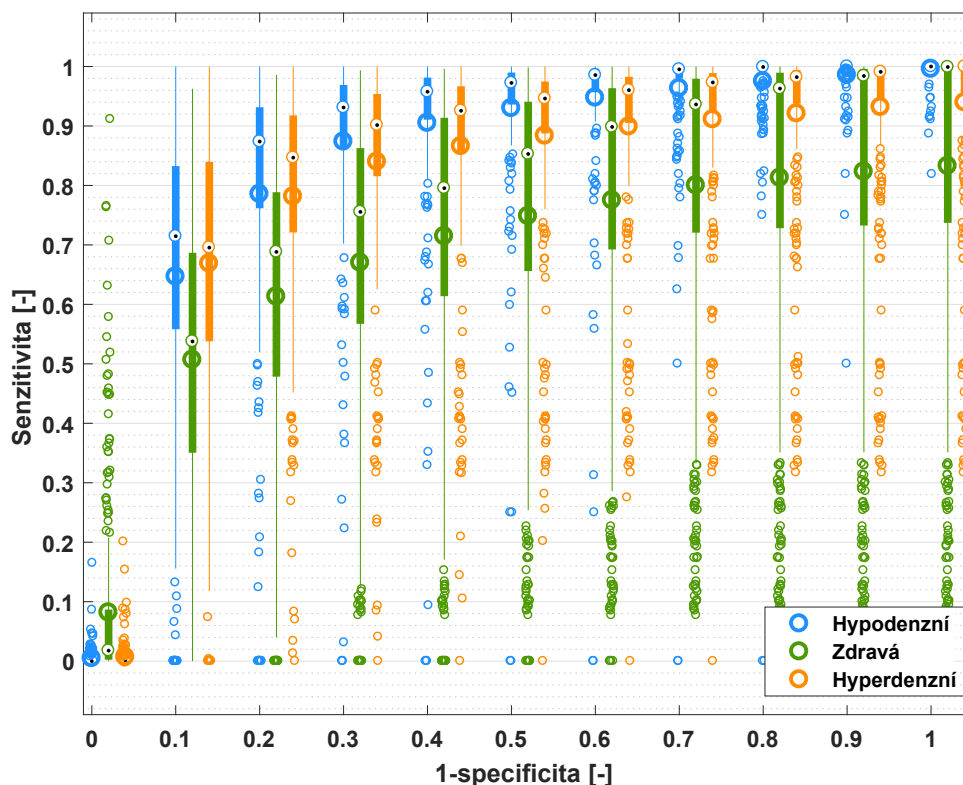
Ukázka výsledné segmentace a klasifikace a odpovídajících expertních značení je zobrazena na Obr. 4.16; současně jsou patrné rozdílnosti mezi dvěma experty (viz Obr. 4.16(b) a (c)). S přihlédnutím k patrným mezi-expertním rozdílům je vidět velmi dobrá subjektivní shoda navržené CNN metody s oběma lékařskými experty.

4.4.6 Meta-analýza detekovaných lézí

Po předchozí segmentaci lézí postavené na základě klasifikace jednotlivých voxelů byly výsledné voxely spojeny do objektů pomocí 3D analýzy spojených komponent (CCA¹⁵) (He a kol., 2009) s použitím 26okolím. Za účelem omezení počtu falešně detekovaných lézí byla pro každý takto vzniklý objekt vypočtena řada tvarových, jasových a texturních příznaků, které byly zpracovány binárním klasifikačním modelem typu náhodný les (RF¹⁶) realizovaným v programu (CNN+RF). Příznaky byly voleny heuristicky z hlediska jejich nízké výpočetní náročnosti a relevantnosti dle spolupracujících lékařských expertů. Použito bylo celkem 17 příznaků, z toho:

¹⁵z anglického Connected Component Analysis

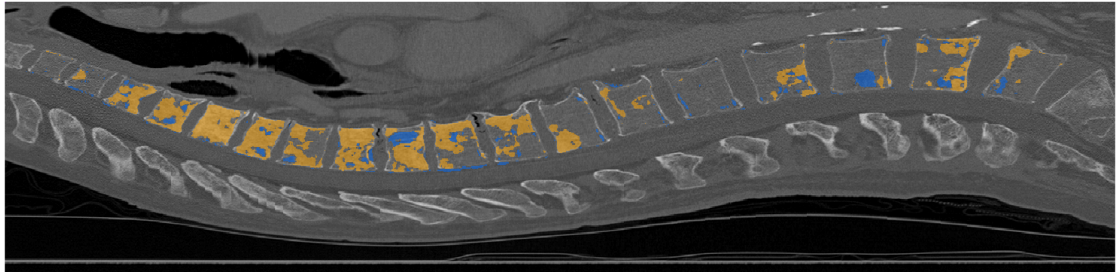
¹⁶z anglického Random Forest



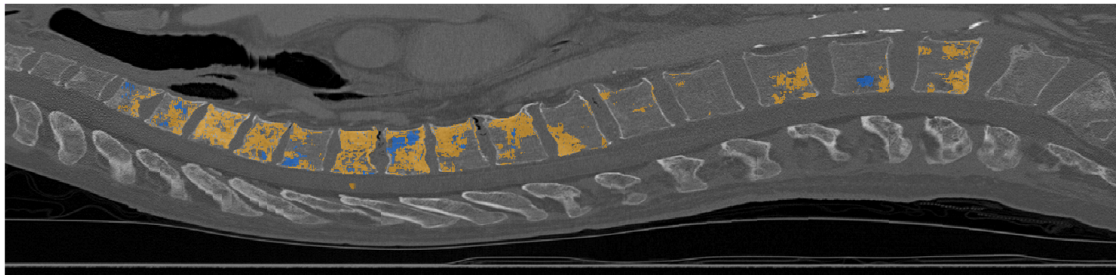
Obr. 4.15: Voxelové hodnoty ROC křivky pro jednotlivé klasifikované tkáně metodou CNN: pro každý operační bod specificity je vytvořen krabicový graf senzitivity pro danou tkáň; testovací množina z DATABÁZE A. Spodní/horní hrana boxu značí 25/75procentní percentil, konce vousů minimální/maximální hodnotu z 1,5 násobku mezikvartilového rozpětí (mimo odlehlé hodnoty), teč uvnitř boxu značí medián, kružnice průměr a malé kružnice odlehlé hodnoty.

- 4 získané z originálních vstupních dat – průměrná, maximální a minimální jasová hodnota a směrodatná odchylka jasových hodnot daného analyzovaného objektu,
- 8 získaných z výstupů použité CNN – průměrná, maximální a minimální SoftMax hodnota pro lytickou a sklerotickou tkáň a směrodatné odchylky jasových hodnot těchto tkání pro každý objekt,
- 5 získaných z výsledků segmentace a binární masky obratlů – Eukleidovská vzdálenost mezi těžištěm objektu a hranicí obratle, tato vzdálenost normovaná velikostí obratle, Eukleidovská vzdálenost mezi těžištěm objektu a těžištěm obratle, a dále dva úhly definující pozici objektu ve sférických souřadnicích se středem souřadného systému v těžišti obratle.

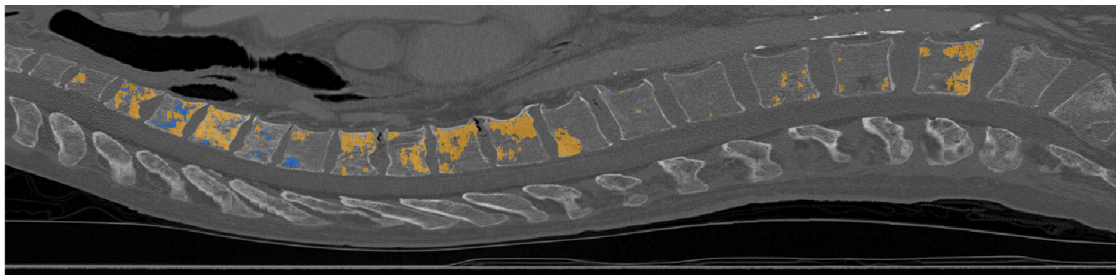
Učení RF klasifikátoru bylo provedeno na stejné učební databázi jako byla použita pro učení CNN.



(a)



(b)



(c)

Obr. 4.16: Ukázka segmentace a klasifikace nádorových tkání (DATABÁZE A – pacient 4): výsledek CNN metody (a), ukázka expertně označených dat expertem 1 (b) a expertem 2 (c) (modrá – lytické tkáně, oranžová – blastické tkáně).

4.4.7 Závěr k metodě využívající konvoluční neuronové sítě

Jako finální metoda pro automatickou segmentaci byla navržena hluboká konvoluční neuronová síť, která byla naučena metodou *mini-batch* gradientního sestupu s hybridností na učební množině vytvořené z DATABÁZE A (Kap. 4.1). Výsledný přístup (CNN) byl testován na zbytku databáze. Pro upřesnění: učební množina byla tvořena voxelovými vzorky s odpovídajícími ortogonálními 2D řezy a odpovídajícími hodnotami správné klasifikace získanými ze shody mezi expertně označenými daty. Výsledné testování proběhlo na 64,86 % DATABÁZE A odpovídající její testovací části. Na testovací části databáze bylo dosaženo průměrných (voxelových) hodnot senzitivity 70,5 %, specificity 81,9 % a Youdenova indexu 52,5. Tato metoda dosahovala z výše navržených a testovaných metod nejlepší úspěšnosti (oproti metodě

FCM přibližně o 10 % vyšší senzitivita a o 6 % vyšší specificita, a oproti **CMB** o 20 % vyšší senzitivita a o 7 % vyšší specificita).

Časová náročnost navržené **CNN** metody byla přibližně 19 *min* na celotělový sken pacienta, což nebylo pro klinické využití vhodné, a proto byl algoritmus realizován s využitím paralelizace na grafických procesorech. Po této úpravě bylo dosaženo podstatně nižšího času 4 *min* na pacienta, která již vyhovuje zamýšlenému využití. Obdobné časové úspory bylo dosaženo také v učební fázi, kdy původní doba učení sítě (pro 1 550 epoch) byla přibližně 10×24 *hod*, a doba po provedení paralelizace cca 12 *hod*.

Dále byla navržena modifikace této metody založená na odstraňování segmentovaných falešných objektů na základě jejich meta-analýzy (**CNN+RF**). Přínos této modifikace je dále diskutován až v Kap. 5.3 a Kap. 5.4, protože vzhledem k charakteru klasifikace založené na objektech nemá hodnocení této modifikace po jednotlivých voxelech význam.

4.5 Finální realizovaný algoritmus

Výsledný algoritmus je realizován v rámci komplexního týmově vytvořeného experimentálního softwaru, který umožňuje ověřování na reálných klinických datech. Popsané navržené metody byly po ověření implementovány jako součást komplexního softwaru složeného z dílčích kroků:

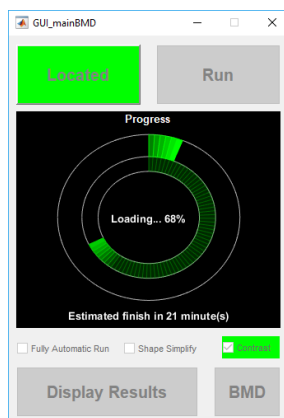
1. načtení a předzpracování obrazových dat ve formátech *DICOM* nebo *MHD/RAW*,
2. detekce páteře a páteřní osy, detekce obratlů, identifikace obratlů a segmentace obratlů (Jan a kol., 2018; Jakubicek, 2020; Jakubicek a kol., 2020, 2019*b,a*),
3. **detekce, segmentace a klasifikace lytických a blastických lézí a extrakce pravděpodobnostní mapy výskytu zdravé tkáně** (metody **FCM** nebo **CNN** (Jan a kol., 2018; Chmelik a kol., 2018*b*, 2019)),
4. **meta-analýza (CNN+RF)** a post-processing lézí (Chmelik a kol., 2018*b*),
5. analýzy tloušťky kortikální části obratlů (Kodym, 2017; Jakubicek a kol., 2020),
6. a segmentace kalibračních tkání (paraspinální sval a tuk) s následnou analýzou minerální hustoty kostí (BMD¹⁷) (Kodym, 2017; Salplachta, 2017).

Tučně jsou uvedeny kroky, jejichž řešení je předmětem této dizertace.

Software je ovládán jednoduchým uživatelským rozhraním (Obr. 4.17), které umožňuje volbu patientské složky a zda má být využito zjednodušování tvarů segmentovaných lézí. Postup běhu softwaru je zobrazován v hlavním okně rozhraní

¹⁷z anglického Bone Mineral Density

spolu s odhadovaným zbývajícím časem. Přes toto rozhraní je také přístup k zobrazení výsledků a analýze BMD.



Obr. 4.17: Ukázka hlavního grafického rozhraní pro ovládání realizovaného softwaru.

Grafické rozhraní pro zobrazení výsledků (Obr. 4.18) umožňuje volit mezi výsledky detekce a identifikace meziobratlových disků, segmentace obratlů a segmentace a klasifikace lézí. Rozhraní obsahuje běžné prvky aplikací pro vizualizaci medicínských obrazů, jako zobrazení originálních dat pomocí volby radiologického okna a zobrazení ve třech základních rovinách spolu se dvěma přehledovými snímky a dvěma informačními tabulkami (informace o každé lézi a míře postižení páteře lézemi). Dále umožňuje lékařskému personálu odstranit detekované léze na základě velikosti absolutního rozdílu z-skóre mezi zdravou a patologickou tkání, velikosti lézí, měnit operační bod na FROC¹⁸ křivce RF klasifikátoru nebo nastavit šířku tzv. ochranného pásma v okolí kortikálních částí obratlů. Rozhraní dále slouží k validaci aktuálních výsledku, kde je možné přijmout/zamítnout automaticky získané výsledky pro celého pacienta, každý obratel nebo individuální lézi. Program umožňuje výpočet, zobrazení a export hodnot míry zasažení jednotlivých obratlů i celého pacienta lézemi, a tím napomoci k diagnostice a sledování vývoje nemoci, či reakce lézí na aplikovanou léčbu.

4.6 Použité metody hodnocení úspěšnosti

Dosažené úspěšnosti navržených počítačem podporovaných detekčních/diagnostických (CAD¹⁹) algoritmů z Kap. 4 z pohledu detekce, segmentace a klasifikace nádorových tkání byly objektivně hodnoceny na základě standardně používaných přístupů založených na porovnání výstupů algoritmů s expertně označenými daty. Pro

¹⁸z anglického Free-Response Receiver Operating Characteristic

¹⁹Computer Aided Detection/Diagnosis



Obr. 4.18: Ukázka grafického rozhraní umožňujícího zobrazit, validovat a exportovat získané výsledky.

hodnocení byly využity dva přístupy: první (spíše technický) založený na vyhodnocení správnosti klasifikace jednotlivých voxelů, a druhý (spíše klinický) řešený porovnáváním celých detekovaných objektů. K vyhodnocení obou přístupů byly využity běžně používané metriky (např.: senzitivita, specificita, ROC/FROC analýza, apod.).

Pro vyhodnocení navržených metod byl zlatý standard tvořen daty anotovanými každým expertem a úspěšnost metod byla hodnocena v porovnání s každým expertem zvlášť – finální hodnoty byly pak dány jako průměr mezi oběma výsledky.

Pro zjištění mezi-expertních rozdílů bylo využito obdobného přístupu jako pro vyhodnocení navržených metod, přičemž byl vždy jeden expert považován za zlatý standard a druhý expert jako porovnávaná metoda. Poté byly role zlatého standardu a experta vyměněny a jako výsledné hodnoty byly vzaty průměry z těchto měření.

4.6.1 Voxelově založená metodika

Z technického pohledu je hodnocení všech jednotlivých klasifikovaných nejmenších prvků (voxelů) nejpřesnějším měřítkem míry úspěšnosti segmentačních algoritmů. Jejich výhodami jsou možnosti jednoduchého porovnání různých algoritmů navzájem, jednoznačná definice hodnocených prvků a jednoduchý výpočet. Nevýhodou je problematika a časová náročnost expertního značení dat a klinicky ne příliš interpretovatelné výsledky (lékař v klinické praxi většinou nesleduje a nehodnotí jednotlivé

voxely). Tato metodika tedy vypovídá o kvalitě a přesnosti segmentace a klasifikace hodnocených tkání.

Vzhledem k dostupnosti voxelově označené DATABÁZE A (Kap. 4.1) nezávisle dvěma lékařskými experty, bylo možné vypočítat shodu mezi experty a současně shodu testovaných metod s experty. Pro vyhodnocení byly využity některé z běžně používaných metrik využívající počtů správně klasifikovaných pozitivních vzorků (TP^{20}), správně klasifikovaných negativních vzorků (TN^{21}), vzorků určených falešně jako pozitivní (FP^{22}) a vzorků určených jako falešně negativní (FN^{23}). Jmenovitě byly využity následující metriky:

- senzitivita (TPR^{24})

$$TPR = \frac{TP}{(TP + FN)}, \quad (4.14)$$

- specificita (TNR^{25})

$$TNR = \frac{TN}{(TN + FP)}, \quad (4.15)$$

- pozitivní prediktivní hodnota (PPV^{26})

$$PPV = \frac{TP}{(TP + FP)}, \quad (4.16)$$

- negativní prediktivní hodnota (NPV^{27})

$$NPV = \frac{TN}{(TN + FN)}, \quad (4.17)$$

- míra pozitivní věrohodnosti (LR^{+28})

$$LR^+ = \frac{TPR}{(1 - TNR)}, \quad (4.18)$$

- míra negativní věrohodnosti (LR^{-29})

$$LR^- = \frac{(1 - TPR)}{TNR}, \quad (4.19)$$

- správnost (ACC^{30})

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)}, \quad (4.20)$$

²⁰z anglického True Positive

²¹z anglického True Negative

²²z anglického False Positive

²³z anglického False Negative

²⁴z anglického True Positive Rate

²⁵z anglického True Negative Rate

²⁶z anglického Positive Predictive Value

²⁷z anglického Negative Predictive Value

²⁸z anglického Positive Likelihood Ratio

²⁹z anglického Negative Likelihood Ratio

³⁰z anglického Accuracy

- F1-score³¹

$$F1 = \frac{2TP}{(2TP + FP + FN)}, \quad (4.21)$$

- ROC³² analýza,
- plocha pod ROC křivkou (AUC³³),
- vyvážená správnost (AUCb³⁴) (Brodersen a kol., 2010)

$$AUCb = \frac{(TPR + TNR)}{2}, \quad (4.22)$$

- Youndenův index (YI³⁵)

$$YI = TPR + TNR - 1, \quad (4.23)$$

- diskriminační síla (DP³⁶) (Sokolova a kol., 2006; Taha a Hanbury, 2015)

$$DP = \frac{\sqrt{3}}{\pi} \log(LR + LR-), \quad (4.24)$$

- a koeficient diagnostické šance (DOR³⁷) (Glas a kol., 2003).

$$DOR = LR + /LR-, \quad (4.25)$$

Tyto voxelově založené metriky byly použity pro prvotní (technické) vyhodnocení úspěšnosti navržených metod.

4.6.2 Objektově založená metodika

Vyhodnocení úspěšnosti CAD algoritmu vhodnější pro klinickou praxi je zřejmě založeno na hodnocení jednotlivých detekovaných a klasifikovaných kandidátů jako celku namísto jednotlivých voxelů. Výsledky tohoto hodnocení jsou lépe interpretovatelné z hlediska počtu, typu a velikosti detekovaných lézí. Problematická je však definice objektů v případě tvarově velmi komplikovaných lézí (nemají kompaktní tvar). Tato metodika vyhodnocuje spíše kvalitu detekce hledaných lézí.

Aby bylo možno využít voxelově označené DATABÁZE A, bylo nutné převést expertně označená data i výsledky testovaných metod na objektově označená data. K tomuto účelu bylo využito algoritmu 3D CCA³⁸ (Ronse a Devijver, 1984) ve 26okolí

³¹označováno také jako Dice koeficient

³²z anglického Receiver Operating Characteristic

³³z anglického Area Under Curve

³⁴z anglického Balanced Accuracy

³⁵z anglického Younden's index

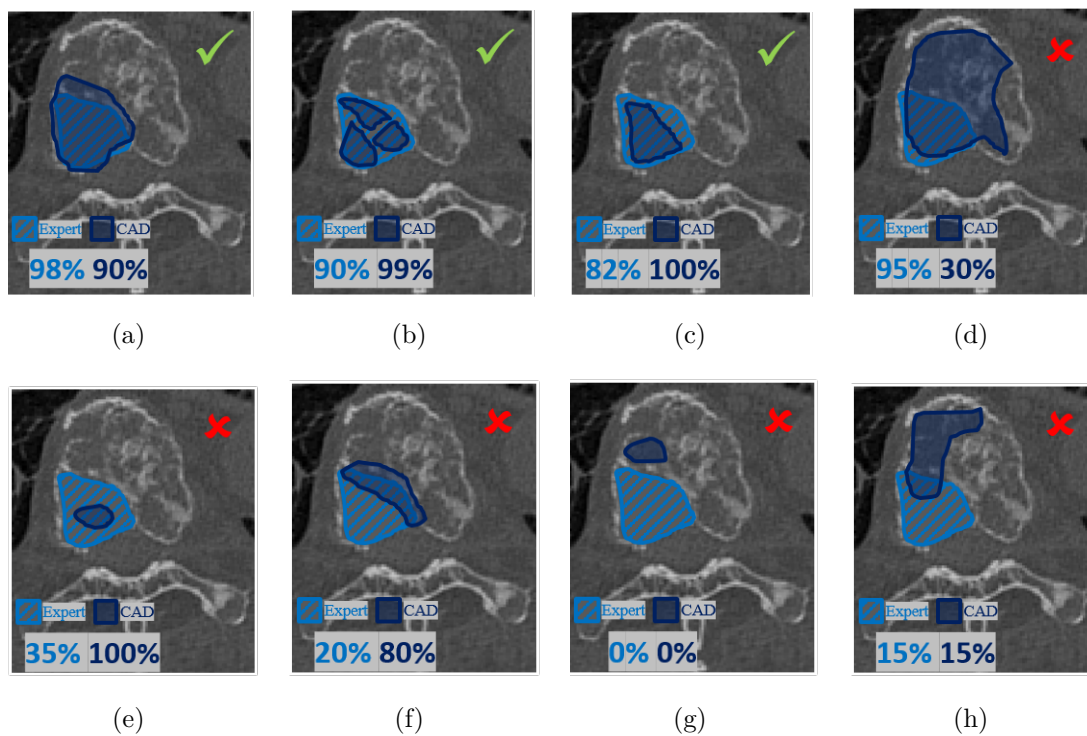
³⁶z anglického Discriminant Power

³⁷z anglického Diagnostic Odds Ratio

³⁸z anglického Connected Components Analysis (analýzy spojených komponent)

voxelu pro každou klasifikovanou patologickou tkáň, který byl aplikován na anoto-
vaná i CAD metodami označená data. Dále byly ze všech dat odstraněny objekty
menší než 10 voxelů (objem $\sim 1,40 \text{ mm}^3$, poloměr koule ekvivalentního objemu
 $\sim 0,70 \text{ mm}$), které byly často generovány přítomným šumem v datech. Následné vy-
hodnocení úspěšnosti poté probíhalo v závislosti na míře překryvu objektů získaných
CAD metodou s objekty určenými expertem, přičemž vyhodnocení bylo provedeno
pro různé hodnoty prahu překryvu. Hodnocení bylo rozděleno do třech kategorií
velikosti lézí (malé: $V < 14,20 \text{ mm}^3$ ($r < 1,50 \text{ mm}$), střední: $V < 300,00 \text{ mm}^3$
($r < 4,15 \text{ mm}$) a velké: $V > 300,00 \text{ mm}^3$ ($r > 4,15 \text{ mm}$)) a také dle typu léze (níz-
kodenzní, vysokodenzní). Pro zjednodušení je princip přístupu hodnocení vysvětlen
pouze na 2D příkladu na Obr. 4.19. Výsledkem tohoto přístupu jsou počty správně
a falešně detekovaných a klasifikovaných objektů v závislosti na jejich velikosti, typu
a nastavené hodnotě prahu překryvu. Dále byly opět vypočteny metriky vhodné pro
detekci podobně jako v Kap. 4.6.1.

Objektově založené metodiky byly využity pro hodnocení klinické aplikovatel-
nosti navržených metod a pro objektivní porovnání s dalšími dostupnými publika-
cemi, protože tento přístup hodnocení je využíván ve většině publikovaných článků
(viz Kap. 1.2).



Obr. 4.19: Princip objektového hodnocení ve 2D pro práh překryvu = 80 %: světle modře je znázorněn zlatý standard, tmavě modře výsledek hodnocené CAD metody, procenta v obrázcích vyjadřují překryv každého objektu. Správné detekce (a), (b), (c); správná a současně falešná detekce (d); falešně negativní detekce (e), (f); falešně negativní a současně falešně pozitivní detekce (g), (h).

5 Výsledky a diskuze

V této kapitole jsou shrnuty a diskutovány detailní výsledky všech navržených metod detekce, segmentace a klasifikace lézí podle obou navržených metodik hodnocení (4.6). Výsledky hodnocení jsou poté v další části využity pro porovnání navržených metod s výsledky algoritmů publikovaných jinými autory.

5.1 Voxelové hodnocení

Z technického hlediska bylo provedeno objektivní porovnání navržených metod pomocí běžně používaných metrik (viz Kap. 4.6.1). Dosažené výsledky jsou vyjádřeny průměrnými hodnotami vypočtenými přes všechny dostupné voxely z testovací části DATABÁZE A (přibližně $406 \text{ obratlů} \times 120\,000 \text{ voxelů}$ na průměrný obratel, tedy $\sim 48,9$ milionů voxelů), včetně mezi-expertních hodnot získaných stejnou metodikou, na Obr. 5.1. Prevalence (procentuální zastoupení) lytické tkáně v celé dostupné DATABÁZI A je 2,6 % a sklerotické 8,0 %.

Při bližším pohledu na jednotlivé hodnoty je na Obr. 5.1(a) patrná vyšší senzitivita metody CNN pro patologické tkáně než u ostatních automatických metod, včetně senzitivity vypočtené mezi experty (viz Kap. 4.6). Lze tedy předpokládat, že CNN metoda dokáže s vyšší mírou odhalit zájmové patologické tkáně než ostatní testované metody; současně dosahuje vyšší senzitivity vypočtené mezi metodou a experty než experti mezi sebou. Naopak pro zdravou tkáň je senzitivita CNN nižší, což je způsobeno zvýšenou mírou falešně označené zdravé tkáně jako patologické, kde je tento jev patrný také v grafu specificity na Obr. 5.1(b). Snížená specificita metody CNN pro patologické tkáně ukazuje na vyšší míru falešně pozitivních případů než u metody CMB a mezi experty. Vyšší hodnota specificity CNN pro zdravou tkáň je provázána právě s vyšší senzitivitou pro patologické tkáně.

Dalšími hodnocenými metrikami byla správnost (*accuracy*) na Obr. 5.1(c), která vyjadřuje celkovou úspěšnost algoritmu (všechny správně zařazené voxely ku všem hodnoceným), a *F1*-skóre (*Dice* koeficient) na Obr. 5.1(d) – zde jsou testované automatické metody na podobné úrovni, a ve všech případech o něco horší než mezi-expertní hodnoty. Důvodem může být metodika použitá při expertním značení, kdy oba experti používali totožný značící nástroj „Marker“ s implementovanou segmentační metodou narůstání semínka (Kap. 4.1), čímž mohlo docházet k podobné výsledné segmentaci stejné léze, kterou experti anotovali. Naproti tomu, automatické metody využívají každá odlišný přístup segmentace, a dochází například k rozdílu ve tvaru a perforaci detekovaných lézí oproti expertně označeným tkáním.

Dále jsou ukázány hodnoty, shrnující senzitivitu a specificitu, které vyjadřují celkovou efektivitu metod pomocí vyvážené správnosti (*balanced accuracy*) (Obr. 5.1(f))

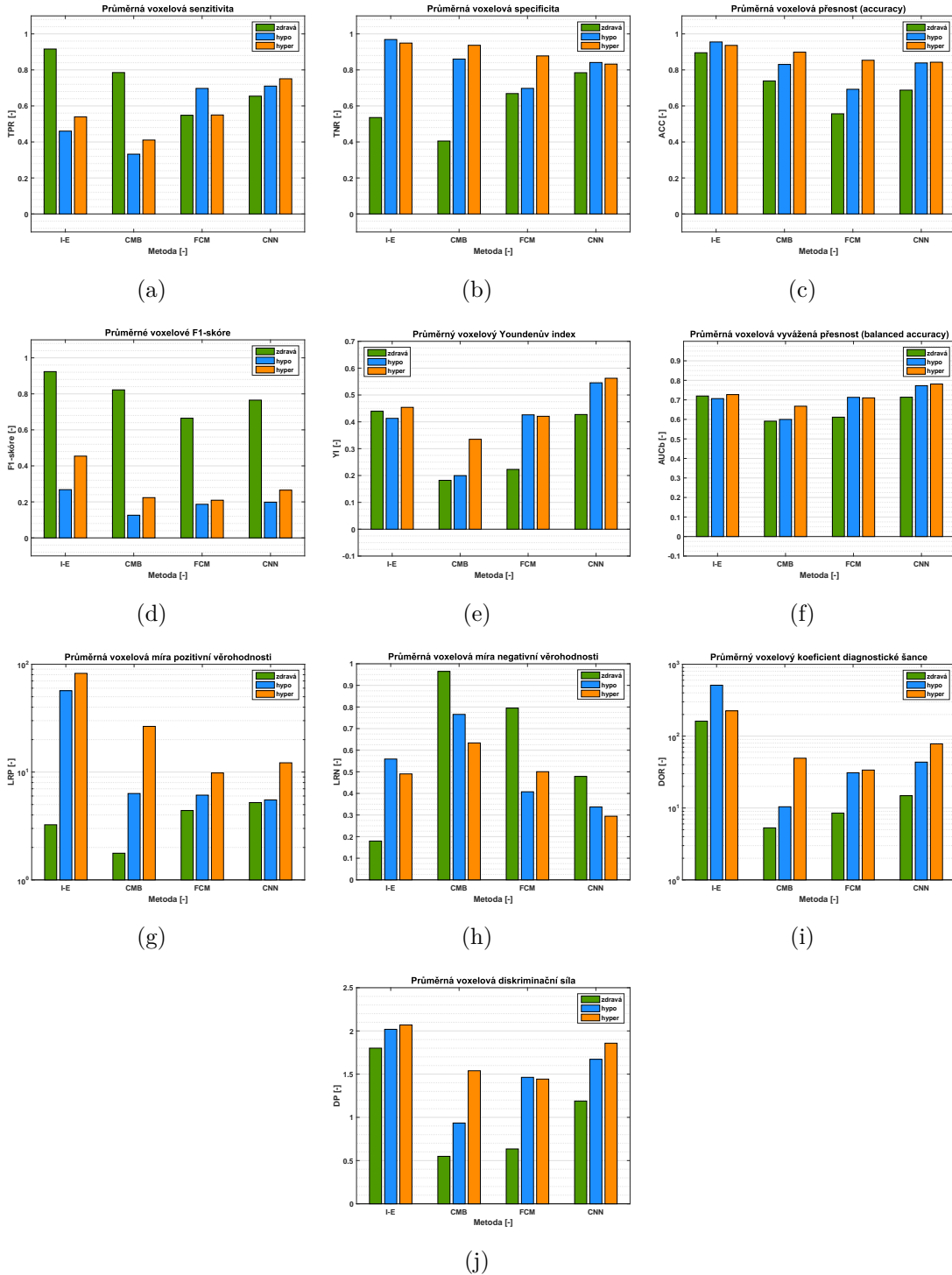
a schopnosti algoritmu vyhnout se chybám podle Youndenova indexu (Obr. 5.1(e)). Mezi oběma těmito parametry je lineární převodní vztah; vyjadřují tedy jiným číslem podobnou vlastnost. Vyšší hodnota Youndenova indexu (v intervalu $\langle -1, 1 \rangle$) značí lepší schopnost metody vyhnout se chybám, obdobně vyšší hodnota vyvážené přesnosti (v intervalu $\langle 0, 1 \rangle$) vyjadřuje větší počet správně zařazených voxelů. Výhodou těchto parametrů je nezávislost na prevalenci tkání. Pro oba parametry vychází metoda **CNN** nejlépe, včetně porovnání s mezi-expertními hodnotami.

Hodnotami vyjadřujícími schopnost testované metody rozlišovat mezi klasifikovanými tkáněmi jsou míra pozitivní a negativní věrohodnosti, koeficient diagnostické šance a diskriminační síla. Míra pozitivní věrohodnosti (Obr. 5.1(g)) vyjadřuje, jak důvěryhodný je pozitivní výsledek klasifikace voxelu tkáně, přičemž rozsah hodnot je v intervalu $\langle 1, +\infty \rangle$. Vyšší hodnota znamená vyšší důvěryhodnost pozitivního výsledku. Zde je patrné, že experti mají pro obě patologické tkáně řádově větší důvěryhodnost než všechny metody (osa y je v logaritmickém měřítku), což je dáno také jejich vysokou specificitou pro patologické tkáně (minimálně označují zdravou tkáň jako patologickou). Oproti tomu, míra negativní věrohodnosti (Obr. 5.1(h)) vyjadřuje důvěryhodnost negativních výsledků klasifikace, kde interval hodnot je $\langle 0, 1 \rangle$ a nižší hodnota značí vyšší důvěryhodnost. Pokud metoda často určuje některé tkáně falešně jako patologické, je její míra negativní důvěryhodnosti vysoká. Je patrné, že **CNN** metoda má nižší míru negativní důvěryhodnosti pro všechny typy tkáně než ostatní automatické metody, což značí, že pokud metoda **CNN** podá negativní výsledek o dané tkáni je důvěryhodnost tohoto výsledku vyšší než u ostatních metod. Vyšší hodnoty míry negativních věrohodností pro patologické tkáně u mezi-expertně hodnocených dat jsou pravděpodobně dány vysokou mírou neurčitosti patologických lézí, případně možností, že jeden z expertů lézí ve velkém množství dat přehlédne. Naopak nízká míra negativní důvěryhodnosti pro zdravou tkáň je u expertů nejvyšší, což znamená, že lékaři se často shodují v rozhodnutí v oblastech zdravé tkáně.

Koeficient diagnostické šance (Obr. 5.1(i)) je přenesen a použit pro míru „klasifikační“ šance. Vyjadřuje kolikrát je větší šance na danou pozitivní klasifikaci voxelu v případě, že je voxel skutečně pozitivní oproti šanci na pozitivní klasifikaci negativního voxelu. Vyšší hodnota značí vyšší šanci správné klasifikace (např. hodnota diagnostické šance **CNN** metody je cca 80, což říká, že šance na klasifikaci sklerotického voxelu jako sklerotický je 80 krát větší než šance na klasifikaci voxelu jiné tkáně jako sklerotický). Z testovaných automatických metod je **CNN** z tohoto hlediska nejlepší pro všechny typy tkáně. Z důvodu výrazně vyšší míry pozitivní věrohodnosti (a specificity) patologických tkání a nižší míry negativní věrohodnosti (a senzitivity) zdravé tkáně je koeficient diagnostické šance řádově vyšší mezi experty než u automatických metod (osa y je zde v logaritmickém měřítku). Šance správně klasifikovat daný voxel je tedy pro experty vyšší než u všech automatických

metod.

Diskriminační síla (Obr. 5.1(j)) vyjadřuje schopnost odlišovat klasifikované tkáně. Vyšší hodnota značí vyšší schopnost diskriminovat danou tkáň. Zde metoda **CNN** opět převyšuje ostatní automatické metody, a současně jsou mezi-expertní hodnoty vyšší než všechny porovnávané automatické metody. Pro tento parametr byl určen rozsah hodnot, kdy hodnota < 1 je špatná, $(1, 2)$ limitovaná $(2, 3)$ ucházející a ≥ 3 dobrá schopnost diskriminace (Sokolova a kol., 2006). Dle tohoto parametru ani mezi-expertní hodnocení nedosahuje dostatečné diskriminační schopnosti, což může být opět zapříčiněno neurčitostí patologických tkání, případně přehlédnutím lézí jedním z expertů. Pro naše účely jsou tedy hodnoty vypočtené mezi experty považovány za maximálně dosažitelné na daných datech a dané zobrazovací modalitě. Metoda **CNN** pro patologické léze dosahuje v porovnání s experty vysokých hodnot a jednoznačně je z testovaných automatických metod pro daný problém nejvhodnější.



Obr. 5.1: Porovnání dosažených průměrných výsledků podle voxelově založené metody (Kap. 4.6.1): senzitivita (a), specifita (b), správnost (c), F1-skóre (d), Youdenův index (e), vyvážená správnost (f), míra pozitivní věrohodnosti (g), míra negativní věrohodnosti (h), koeficient diagnostické šance (i) a diskriminační síla (j) (na ose x jsou navržené metody **CMB**, **FCM** a **CNN**, **I-E** vyjadřuje mezi-expertní hodnoty, jednotlivé sloupce vyjadřují typ klasifikované tkáně).

5.2 Objektové hodnocení

Z klinického hlediska není voxelově založené hodnocení příliš vypovídající, proto byly testované metody dále objektivně zhodnoceny také objektově (viz Kap. 4.6.2 a princip metodiky na Obr. 4.19), což je více podobné lékařské praxi, kdy lékař hodnotí, zda je léze vůbec přítomna, jak je velká a jakého je typu. Rozhodnout o tom, zda je daný objekt léze či nikoliv je pro lékaře lépe definovaný problém, než hodnocení, zda daný voxel je ještě patologická tkáň či nikoliv. Hodnocena byla testovací část DATABÁZE A s celkem 406 obratli. Vzhledem k nemožnosti určení počtu správně negativních objektů touto metodikou, byly pro vyhodnocení uvažovány jen parametry senzitivity a počtu falešně detekovaných lézí. Dosažené výsledky jsou zobrazeny na Obr. 5.2, Obr. 5.3 a Obr. 5.4, kde v každém krabicovém grafu je na ose x míra překryvu automaticky segmentované a klasifikované léze s expertně anotovanou, při kterém je léze brána jako správná detekce. Alternativní interpretací grafů lze zjistit, že hodnota objektově vypočtené senzitivity při daném prahu odpovídá relativnímu počtu lézí z celé hodnocené databáze, které dosahují voxelově zjištěné senzitivity rovné alespoň zvolenému prahu (pokud je tedy průměrná objektová senzitivita např. 0,6 při prahu překryvu 75 %, lze výsledek interpretovat tak, že průměrně 60 % hodnocených lézí dosahuje voxelové senzitivity alespoň 0,75). Jednotlivé sloupce (boxy) vyjadřují jednotlivé testované metody a mezi-expertní shodu.

Při detailním pohledu na Obr. 5.2(a), je patrná nejvyšší senzitivita metody **CNN** oproti ostatním metodám při všech hodnotách prahového překryvu. Dále je vidět také vyšší hodnota senzitivity v porovnání s mezi-expertní shodou. Výrazně nižší je také rozptyl hodnot senzitivity přes všechny testované obratle a minimální hodnoty senzitivity. Dle očekávání je ovšem nejvyšší počet falešných detekcí také u metody **CNN** a nejnižší v případě mezi-expertního hodnocení, což je vidět na Obr. 5.2(b). Při zpřísnění kritéria správné detekce (tedy při zvýšení minimální hodnoty požadovaného překryvu lézí) je patrný pokles senzitivity u všech metod včetně expertního značení a nárůst počtu falešných detekcí.

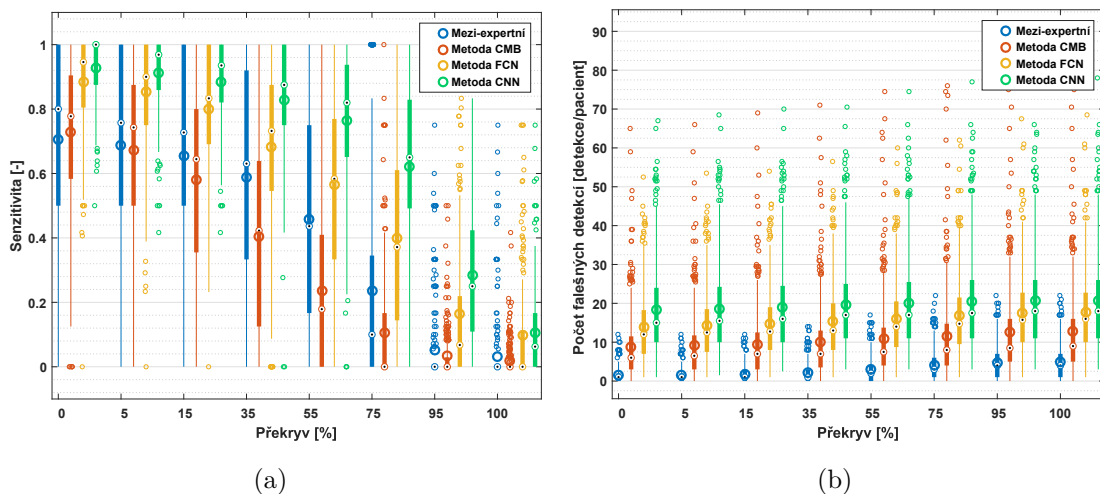
V dalších grafech (Obr. 5.3(a), (b) a Obr. 5.4(a), (b)) je možné porovnat účinnost metod samostatně pro lytické a sklerotické léze, přičemž je patrné, že jak navržená metoda **CNN**, tak i mezi-expertní shoda je lepší pro sklerotické léze podle senzitivity i počtu falešných detekcí. Metoda **CMB** dosahuje ve všech případech nižšího počtu falešných detekcí, avšak za cenu výrazně nižší senzitivity, což je z hlediska CAD systému nežádoucí.

Chování testovaných metod v závislosti na velikosti i typu lézí je shrnuto v Tab. 5.1 a Tab. 5.2, kde dle očekávání pro velmi malé léze o velikosti od 10 vx do 100 vx (objem od $\sim 1,40 \text{ mm}^3$ do $\sim 14,20 \text{ mm}^3$, poloměr koule o ekvivalentním objemu od $\sim 0,70 \text{ mm}$ do $\sim 1,50 \text{ mm}$) je senzitivita nižší a současně je větší

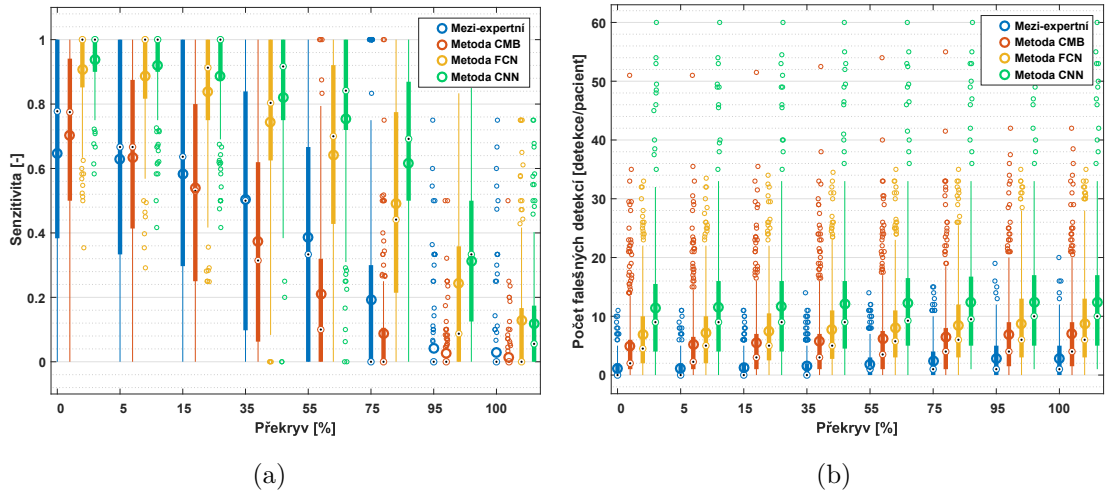
počet falešných detekcí než u středně velkých lézí od 100 vx do 2133 vx (objem $\sim 14,20 \div 300,00 \text{ mm}^3$, poloměr $\sim 1,50 \div 4,15 \text{ mm}$) a velkých lézí nad 2.133 vx (objem nad $\sim 300,00 \text{ mm}^3$, poloměr nad $\sim 4,15 \text{ mm}$).

Z těchto výsledků vyplývá, že jak navržená metoda **CNN**, tak **FCM** dosahují uspokojivých výsledků z pohledu objektové senzitivity pro oba typy patologických tkání, a to zejména pro střední a velké léze. Senzitivita obou navržených metod také převyšuje senzitivitu získanou při vzájemném mezi-expertním hodnocení, což ukazuje že automatické metody detekují velkou část lézí, které byly označeny oběma experty. Cenou je však vyšší počet falešných detekcí automatických metod oproti expertům, nicméně pro využití jako CAD algoritmu je vyšší senzitivita preferována před nízkým počtem falešných detekcí, protože je vhodnější upozornit na více možných kandidátů s možností jejich případného zamítnutí, než některé kandidáty nedetekovat vůbec.

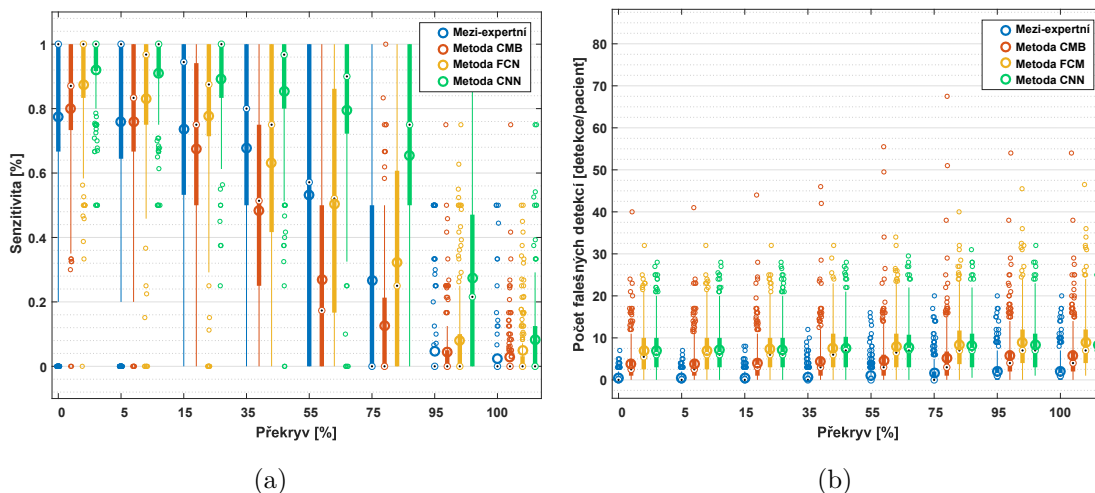
Snížení počtu falešných detekcí lézí bylo dále provedeno s využitím meta-analýzy **CNN+RF** aplikované na metodu **CNN**. Dosažené výsledky jsou dále komentovány v následující kapitole 5.3.



Obr. 5.2: Porovnání dosažených výsledků navržených metod podle objektově založené metodiky (Kap. 4.6.2) bez uvážení typu lézí: senzitivita (a) a počet falešných detekcí (b); testovací množina DATABÁZE A. Na ose x je minimální hraniční překryv CAD metody s expertním značením, kdy je léze počítána jako správně detekovaná jednotlivé sloupce vyjadřují testovanou metodu **CMB** (Kap. 4.2), **FCM** (Kap. 4.3) a **CNN** (Kap. 4.4), přičemž **Inter-Expert** jsou meziexpertní hodnoty. Každý sloupec obsahuje 406 hodnot (obratlů) z DATABÁZE A. Spodní/horní hrana boxu značí 25/75procentní percentil, konce vousů minimální/maximální hodnotu z 1,5 násobku mezikvartilového rozpětí (mimo odlehlé hodnoty), terč uvnitř boxu značí medián, kružnice průměr a malé kružnice odlehlé hodnoty.



Obr. 5.3: Porovnání dosažených výsledků navržených metod podle objektově založené metodiky (Kap. 4.6.2) pro nízkodenní léze: senzitivita (a) a počet falešných detekcí (b); testovací množina DATABÁZE A. Na ose x je minimální hraniční překryv CAD metody s expertním značením, kdy je léze počítána jako správně detekovaná jednotlivé sloupce vyjadřují testovanou metodu **CMB** (Kap. 4.2), **FCM** (Kap. 4.3) a **CNN** (Kap. 4.4), přičemž **Inter-Expert** jsou meziexpertní hodnoty. Každý sloupec obsahuje 406 hodnot (obratlů) z DATABÁZE A. Spodní/horní hrana boxu značí 25/75procentní percentil, konce vousů minimální/maximální hodnotu z 1,5 násobku mezikvartilového rozpětí (mimo odlehlé hodnoty), teč uvnitř boxu značí medián, kružnice průměr a malé kružnice odlehlé hodnoty.



Obr. 5.4: Porovnání dosažených výsledků navržených metod podle objektově založené metodiky (Kap. 4.6.2) pro vysokodenzní léze: senzitivita (a) a počet falešných detekcí (b); testovací množina DATABÁZE A. Na ose x je minimální hraniční překryv CAD metody s expertním značením, kdy je léze počítána jako správně detekovaná jednotlivé sloupce vyjadřují testovanou metodu **CMB** (Kap. 4.2), **FCM** (Kap. 4.3) a **CNN** (Kap. 4.4), přičemž **Inter-Expert** jsou meziexpertní hodnoty. Každý sloupec obsahuje 406 hodnot (obratlů) z DATABÁZE A. Spodní/horní hrana boxu značí 25/75procentní percentil, konce vousů minimální/maximální hodnotu z 1,5 násobku mezikvartilového rozpětí (mimo odlehlé hodnoty), terč uvnitř boxu značí medián, kružnice průměr a malé kružnice odlehlé hodnoty.

5.3 Porovnání s dalšími autory

V další fázi objektivního hodnocení byly navržené metody porovnány s publikacemi zabývajícími se obdobnou problematikou v impaktovaných časopisech (jejich rešerše je provedena v Kap. 1.1). Vzhledem k nedostupnosti veřejné databáze s expertně anotovaným zlatým standardem klasifikovaných a segmentovaných lézí je přímé objektivní srovnání nemožné, avšak s přihlédnutím ke všem aspektům autory používaných dat (počty pacientů, lézí, parametry CT a metodikou hodnocení) je možné alespoň částečné porovnání a diskuze.

Srovnání je provedeno vždy metodikou, která byla využita autory dané publikace – tedy voxelově vypočtené výsledky byly porovnány s publikacemi s voxelovým hodnocením a objektové s objektovým („bounding-box“ metodika hodnocení byla zařazena do objektového značení pro její podobnost, avšak je nutné si uvědomit, že přesnost tohoto hodnocení je potenciálně nižší než u objektového). Protože autoři publikací uvádějí výsledky hodnocené pouze pro práh překryvu detekované léze se skutečnou $\sim 1\%$ (překryv jediným voxelem je brán jako správná detekce), jsou v této práci z důvodu možnosti objektivního porovnání uváděny také výsledky dosažené s tímto prahem. Tyto hodnoty mohou být pro klinické hodnocení velmi zavádějící (vhodnější by byly např. hodnoty dosažené při překryvu alespoň 75% , jako viz Kap. 5.2), nicméně autoři zmíněných publikací jiné dosažené hodnoty neuvádějí.

Tabulky Tab. 5.1 a Tab. 5.2 shrnují autory uváděné výsledky senzitivity a falešných detekcí spolu s minimálními velikostmi hodnocených lézí a počty anotovaných lézí, obratlů a případů, pokud jsou tyto informace dostupné, pro nízkodenzní a vysokodenzní léze. Tučně jsou v obou tabulkách uvedeny navržené metody a jejich dosažené výsledky. Pro navržené metody byly uvedeny výsledné hodnoty pro různé velikosti lézí.

Jak je patrné z obou tabulek (Tab. 5.1 a Tab. 5.2), velikost velkých lézí ($> 300\text{ mm}^3$) odpovídá velikosti hodnocených lézí použité ve všech zmíněných publikacích využívajících objektové hodnocení. Při porovnání dosažených výsledků pro nízkodenzní léze (Tab. 5.1) je zřejmé, že při hodnocení pouze velkých lézí dosahují všechny tři navržené metody senzitivity blízké 100% a do 10 falešně detekovaných lézí na pacienta, a dosahují tedy vyšší nebo srovnatelné úspěšnosti s jinými metodami. Je nutné si uvědomit, že většina autorů využila data s mnohem horším rozlišením v ose z (přibližně 15krát, tloušťka řezu až $5,000\text{ mm}$ oproti $0,335\text{ mm}$ v DATABÁZI A), proto je vhodnější porovnávat výsledky dosažené navrženými metodami pro střední léze ($> 14,20\text{ mm}^3$). Zde je znatelný mírný pokles senzitivity (CNN na $97,9\%$) a značný nárůst počtu falešných detekcí na pacienta (CNN na $73,03$). Zejména počet falešných detekcí je v porovnání s jinými autory o řád vyšší, avšak je vhodné zahrnout objektivnější hodnocení počtu falešných detekcí,

který vezme v úvahu také počty anotovaných případů a lézí. Při bližším pohledu je patrné, že dle údajů o počtu anotovaných lézí jsou námi navržené metody hodnoceny na řádově větší databázi než používají jiní autoři (619 anotovaných lézí střední velikosti oproti např. 16 velkým lézím). Z těchto údajů je patrné, že dostupná DATABÁZE A obsahuje celkově mnohem více patologií než databáze použité dalšími autory.

Oproti jiným metodám jsou navržené metody **FCM** a **CNN** schopné uspokojivě detekovat, segmentovat a klasifikovat také velmi malé léze ($> 1,40 \text{ mm}^3$) se senzitivitou $\sim 90,0 \%$ ovšem s výrazně vyšším počtem falešných detekcí.

Vzhledem k tomu, že ve většině autory uváděných publikacích chybí veškeré údaje o počtech hodnocených dat, prevalenci patologických voxelů, celkovém počtu hodnocených voxelů, apod., tak přímé voxelové hodnocení není možné. Nicméně, voxelové hodnocení bylo provedeno alespoň na publikované metody dalších autorů, jejichž algoritmy byly v rámci projektu dostupné a byly tedy přímo testovány a přepracovány v Kap. 4.2. Objektivní výsledky jak původních, tak přepracovaných metod jsou dostupné v Kap. 4.2.4 v Tab. 4.1. Tab. 5.1 poté uvádí hodnoty prezentované přímo v dostupných publikacích na autory použitých datech.

Obdobné srovnání metod bylo provedeno v Tab. 5.2 pro vysokodenzní léze. Metoda **CNN** zde opět dosahuje nejvyšších hodnot senzitivity ($\sim 99,4 \%$) s řádově obdobným počtem falešných detekcí. Velmi dobré výsledky z hlediska nízkého počtu falešných detekcí na srovnatelně velké databázi anotovaných lézí dosahuje algoritmus publikovaný autory Roth a kol. (2015, 2016) s hodnotou 3 falešných detekcí na pacienta, což je řádově lepší než navržená metoda **CNN**; cenou je ovšem znatelně nižší senzitivita ($70,0 \%$). Pro voxelové hodnocení vysokodenzních lézí platí totéž, co pro hodnocení nízkodenzních lézí v předchozím odstavci.

Porovnáním dosažených výsledků navržených metod s výsledky metod publikovanými dalšími autory bylo ukázáno, že navržené metody dosahují srovnatelných a často lepších výsledků senzitivity a počtu falešných detekcí pro řešenou problematiku velkých lézí. Navržené metody jsou, oproti ostatním, zároveň schopny detekovat a klasifikovat léze menší než $300,00 \text{ mm}^3$ téměř se zachováním efektivity metod, a také detekovat a klasifikovat léze menší než $14,20 \text{ mm}^3$ s odpovídajícím poklesem senzitivity a nárůstem počtu falešných detekcí. Dosažené výsledky navržených metod byly získány na základě porovnání se dvěma nezávislými lékařskými experty, což u řady jiných publikovaných algoritmů neplatí; další autoři často disponují pouze anotacemi od jednoho experta, případně více expertů na anotacích spolupracuje a své značení mezi sebou sdílí nebo navzájem korigují.

V případě navržené metody kombinující **CNN** segmentaci a **RF** meta-analýzu (**CNN+RF**) bylo opět provedeno hodnocení na DATABÁZI A. Jak je patrné v Tab. 5.1 a Tab. 5.2, ve všech případech došlo k výraznému poklesu falešně de-

tekovaných lézí, přičemž v případě hůře definovatelných nízkodenzních lézí došlo ke snížení senzitivity až o 10 %. Obdobných výsledků počtu falešných detekcí bylo dosaženo také na kontrolní DATABÁZI C.

Tab. 5.1: Porovnání výsledků navržených metod s dalšími publikovanými algoritmy – nízkodenní léze (pro voxelové hodnocení je použito FPR (False Positive Rate) v procentech, pro objektové je FP (False Positive) v počtu detekcí na pacienta a na obratel). TPR (True Positive Rate) vyjadřuje senzitivitu. Počty případů, obratlů a lézí odpovídají celým uvedeným databázím; dosažené výsledky jsou počítány pouze na jejich testovacích částech. Databáze „0“ obsahovala pouze omezené množství obrazových dat, databáze „A“ je námi vytvořená databáze a databáze „C“ obsahuje kontrolní skupinu pacientů bez metastatických změn.

Autoři	Hodnocení	Databáze	Případů	Obratlů	Lézi	Velikost [mm ³]	TPR [%]	FPR[%]/FP [na pac.]	FP [na obratel]
Peter a kol. (2013) (MRF)	voxelové	0a	6	114	–	>1,4	55,7	8,6	–
Jan a kol. (2015) (GRAF)	voxelové	0b	27	227	–	>0,0	80,5	42,5	–
Jan a kol. (2015) (TEX)	voxelové	0b	27	227	–	>0,0	57,5	39,0	–
Mezi-expertní	voxelové	A	31	626	1046	>1,4	46,0	3,1	–
CMB	voxelové	A	31	626	1046	>1,4	33,2	14,0	–
FCM	voxelové	A	31	626	1046	>1,4	69,7	30,2	–
CNN	voxelové	A	31	626	1046	>1,4	71,0	12,0	–
Yao a kol. (2007)	objektové	–	21	–	16	>400,0	93,8	5,7	–
O'Connor a kol. (2007)	objektové	–	21	–	16	>400,0	94,0	4,5	–
Wels a kol. (2012)	bounding-box	–	34	–	105	>500,0	75,0	3,0	–
Hammon a kol. (2013)	bounding-box	–	20	–	42	>500,0	88,0	3,7	–
Mezi-expertní	objektové	A	31	626	1046/619/114	>1,4/14,2/300,0	64,8/75,4/95,7	21,65/8,12/0,24	1,06/0,40/0,01
CMB	objektové	A	31	626	1046/619/114	>1,4/14,2/300,0	70,2/85,5/99,4	63,73/22,56/6,23	3,64/1,49/0,45
FCM	objektové	A	31	626	1046/619/114	>1,4/14,2/300,0	90,7/97,8/99,7	124,81/34,72/8,61	2,05/1,00/1,33
CNN	objektové	A	31	626	1046/619/114	>1,4/14,2/300,0	94,2/97,9/99,8	271,29/73,03/7,71	15,79/4,68/0,65
CNN+RF	objektové	A	31	626	1046/619/114	>1,4/14,2/300,0	80,0/88,1/89,2	31,02/14,44/2,87	1,59/0,74/0,15
CNN	objektové	C	29	638	0	>1,4/14,2/300,0	–	182,90/80,38/8,79	9,19/4,04/0,44
CNN+RF	objektové	C	29	638	0	>1,4/14,2/300,0	–	21,76/19,62/5,45	1,09/0,99/0,27

Tab. 5.2: Porovnání výsledků navržených metod s dalšími publikovanými algoritmy – vysokoznížní léze (pro voxelové hodnocení je použito FPR (False Positive Rate) v procentech, pro objektové je FP (False Positive) v počtu detekcí na pacienta a na obratel). TPR (True Positive Rate) vyjadřuje senzitivitu. Počty případů, obratlů a lézí odpovídají celým uvedeným databázím; dosažené výsledky jsou počítány pouze na jejich testovacích částech. Databáze „0“ obsahovala pouze omezené množství obrazových dat, databáze „A“ je námi vytvořená databáze a databáze „C“ obsahuje kontrolní skupinu pacientů bez metastatických změn.

Autoři	Hodnocení	Databáze	Případů	Obratlů	Lézí	Velikost [mm ³]	TPR [%]	FPR[%]/FP [na pac.]	FP [na obratel]
Peter a kol. (2013) (MRF)	voxelové	0a	6	114	–	>1,4	63,5	4,9	–
Jan a kol. (2015) (GRAF)	voxelové	0b	27	272	–	>0,0	64,5	28,0	–
Jan a kol. (2015) (TEX)	voxelové	0b	27	272	–	>0,0	84,0	35,5	–
Mezi-expertní	voxelové	A	31	626	1135	>1,4	54,9	5,1	–
CMB	voxelové	A	31	626	1135	>1,4	41,1	6,3	–
FCM	voxelové	A	31	626	1135	>1,4	55,0	12,2	–
CNN	voxelové	A	31	626	1135	>1,4	74,2	17,8	–
Wiese a kol. (2011)	objektové	–	10	–	212	>300,0	77,4	9,4	–
Wiese a kol. (2012)	objektové	–	10	–	83	>300,0	71,2	8,8	–
Hammon a kol. (2013)	bounding-box	–	30	–	172	>300,0	83,0	3,5	–
Burns a kol. (2013)	objektové	–	42	–	352	>300,0	79,0	10,9	–
Yao a kol. (2015)	objektové	–	43	–	372	>300,0	84,0	11,3	–
Roth a kol. (2015, 2016)	objektové	–	59	–	532	>300,0	70,0	3,0	–
Mezi-expertní	objektové	A	31	626	1135/830/305	>1,4/14,2/300,0	76,8/82,8/94,9	8,88/4,65/0,47	0,43/0,23/0,02
CMB	objektové	A	31	626	1135/830/305	>1,4/14,2/300,0	72,9/83,0/96,2	47,94/19,16/2,52	2,73/1,21/0,16
FCM	objektové	A	31	626	1135/830/305	>1,4/14,2/300,0	87,5/95,1/99,6	124,22/37,38/7,18	2,83/1,11/0,64
CNN	objektové	A	31	626	1135/830/305	>1,4/14,2/300,0	95,1/99,4/99,7	130,31/40,40/7,65	7,83/2,58/0,55
CNN+RF	objektové	A	31	626	1135/830/305	>1,4/14,2/300,0	92,0/97,4/99,5	45,63/20,47/5,92	3,40/1,56/0,46
CNN	objektové	C	29	638	0	>1,4/14,2/300,0	–	131,83/53,14/21,79	6,63/2,67/1,10
CNN+RF	objektové	C	29	638	0	>1,4/14,2/300,0	–	21,31/20,79/17,14	1,07/1,05/0,86

5.4 Subjektivní hodnocení

Vzhledem k dostupnosti expertně anotovaných dat pouze z naší DATABÁZE A bylo pro zbylé neoznačené pacienty z dalších databází provedeno subjektivní hodnocení na základě experty stanovených pravidel – takto byly vyhodnoceny nejčastější chyby navržené automatické metody. Dále byla subjektivně zhodnocena naučená CNN vizualizací filtrů a obrazů maximálně aktivujících určité neurony získané s využitím GBP¹ (Simonyan a kol., 2013; Zeiler a Fergus, 2014; Yosinski a kol., 2015) techniky.

5.4.1 Hodnocení na neanotovaných datech

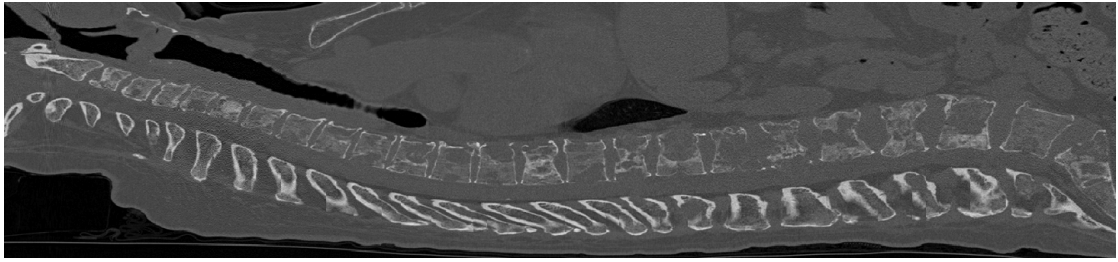
Během testování byla vyzkoušena data získaná s jinými rozlišeními, akvizičními parametry, a s různými použitými rekonstrukčními filtry. Dle navrženého subjektivního hodnocení experty jsou dosažené výsledky CNN srovnatelné (kromě YB rekonstrukčního filtru – viz dále) s výsledky získanými na datech z DATABÁZE A. Nejčastějšími chybami navržených metod je označování kortikální části obratlů jako sklerotické léze (Obr. 5.6(b), Obr. 5.6(d) a Obr. 5.7(b), Obr. 5.7(d), Obr. 5.7(b) – A), obdobně se vyskytují problémy s falešnou detekcí sklerotických lézí v oblastech kalcifikovaných tkání, ve zhuštěných oblastech v důsledku patologické komprese obratlů nebo v oblastech, kde dochází k pozitivní odezvě na léčbu lytické tkáně (dochází zde k přestavbě kostní matrice provázené zvýšenou denzitou (Obr. 5.5(b) – B)).

Falešně detekované lytické léze se nejčastěji vyskytují v oblastech těla obratle směrem k obratlovému otvoru, kde jsou napojeny cévy, které mají výrazně nižší denzitu než spongiózní část obratle (Obr. 5.6(b), Obr. 5.6(d) a Obr. 5.7(b), Obr. 5.7(d), Obr. 5.7(b) – C). Dále jsou falešné detekce způsobovány poklesem denzity v důsledku osteopenie nebo osteoporózy. Dalšími důvody mohou být nedokonalosti v segmentaci obratlů, kde se falešné detekce mohou vyskytovat v okolí kortikální části obratle nebo v intervertebrálních discích (Obr. 5.6(d) a Obr. 5.7(b) – D). Další výskyt falešně detekovaných lytických lézí vzniká v důsledku použití YB rekonstrukčního filtru u nízkodávkových snímacích protokolů (Obr. 5.8(d) – E). Na Obr. 5.8 je vidět příklad segmentace aplikované na nízkodávková CT data totožného pacienta s totožnými akvizičními parametry, přičemž pro Obr. 5.8(a) byl použit Philips B rekonstrukční filtr, zatímco pro Obr. 5.8(c) Philips YB rekonstrukční filtr; rozdíl v poměru signál/šum je značný, stejně tak i výsledek segmentačního algoritmu. Pro data rekonstruovaná s YB filtrem tedy navržená metoda CNN není vhodná.

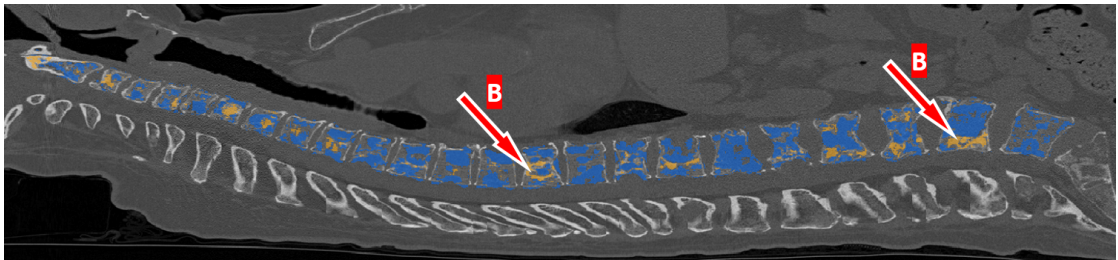
Vzhledem k charakteru výše popsaných falešných detekcí bylo možné jejich počet dále snížit jejich dodatečným odstraňováním pomocí klasifikátoru, který je založen

¹z anglického Guided Backpropagation (řízené zpětné šíření)

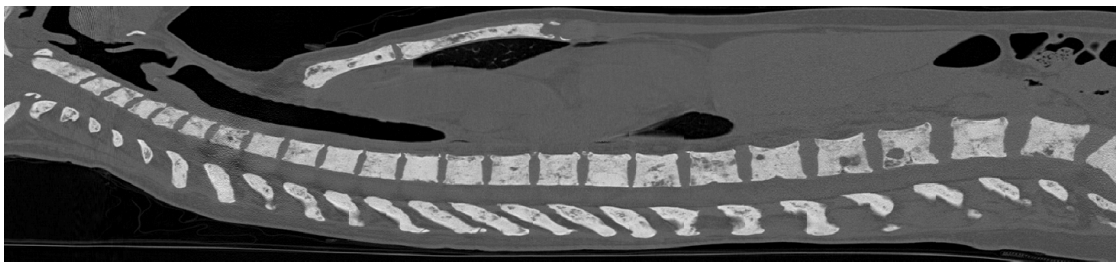
na formě meta-analýzy jednotlivých objektů (**CNN+RF**). Ukázka výsledků je patrná na obrázku Obr. 5.9



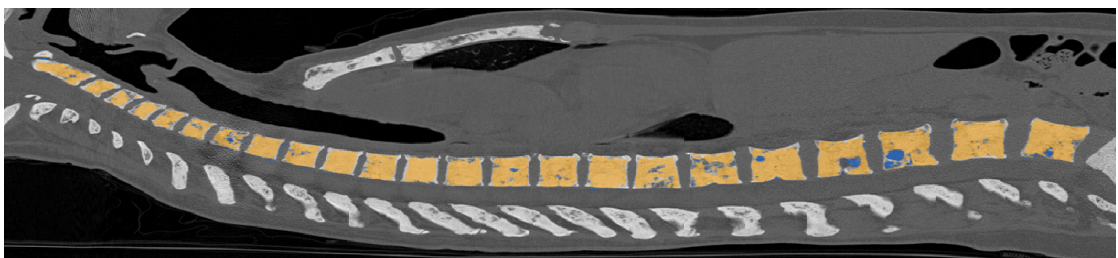
(a)



(b)



(c)

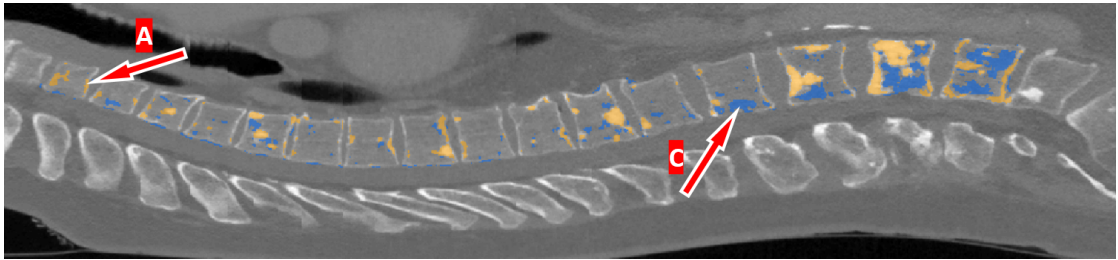


(d)

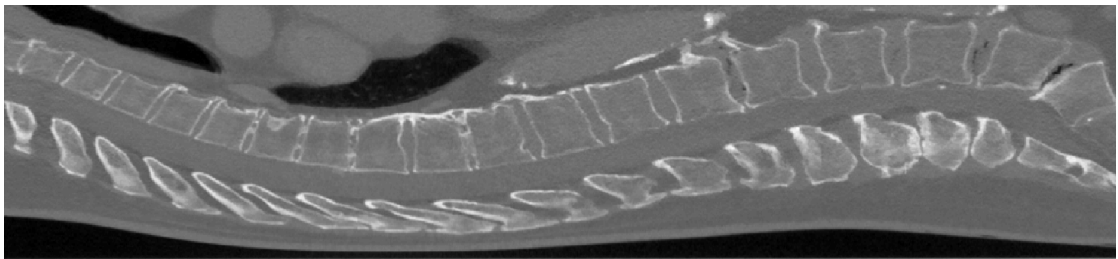
Obr. 5.5: Příklady segmentace a klasifikace nádorových tkání na datech s experimentálním protokolem (Meldola – Obr. 4.1) – sagitální: pacient 16 s vysokou mírou osteolytických změn, patrné jsou komprese a deformace řady obratlů – originální obraz (a) a výsledek CNN metody (b); pacient 25 s vysokou mírou smíšených změn s následnou přestavbou tkáně v oblasti osteolytických lézí v reakci na aplikovanou léčbu – originální obraz (c) a výsledek CNN metody (d) (modrá – lytická tkáň, oranžová – blastická tkáň, červené šipky označují příklady nejčastějších případů falešných detekcí, písmena značí druh chyby).



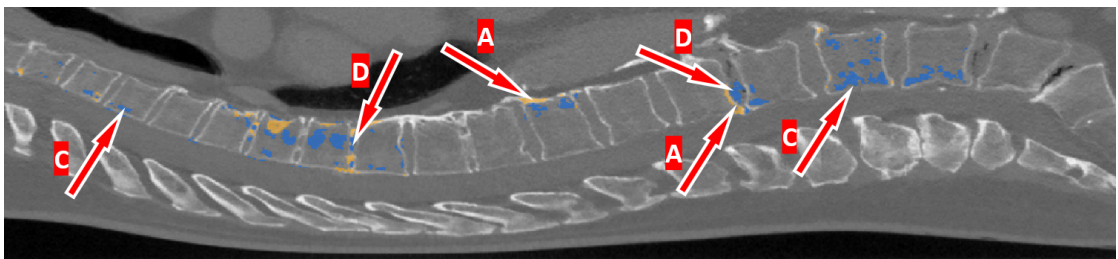
(a)



(b)

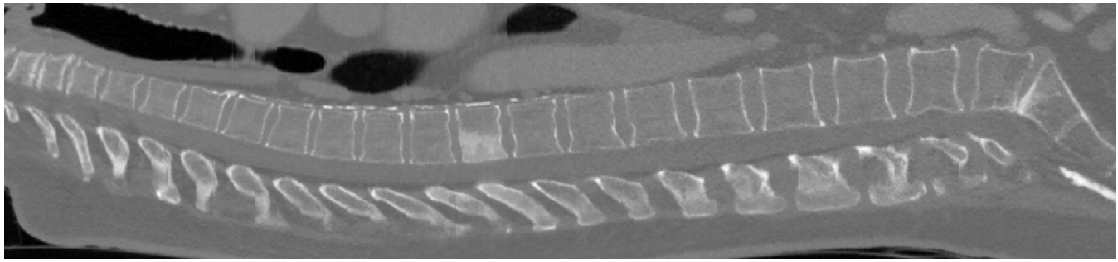


(c)

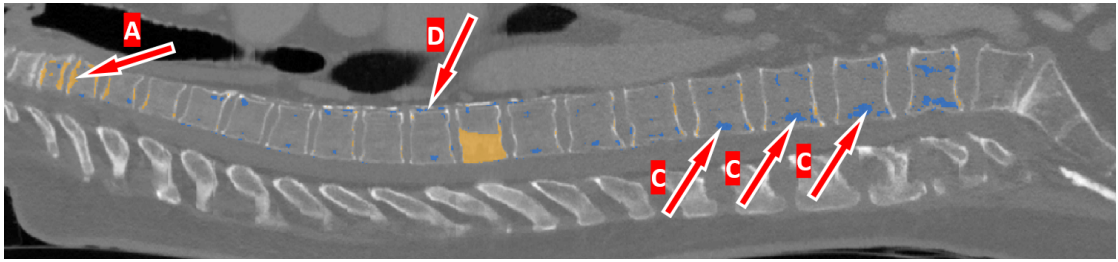


(d)

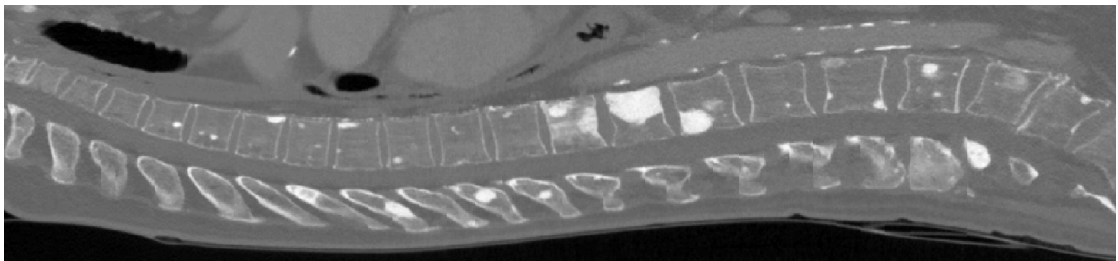
Obr. 5.6: Příklady segmentace a klasifikace nádorových tkání na datech se standardním protokolem (Praha – Obr. 4.1) – sagitální: pacient *S33850* s vícenásobnými osteoblastickými lézemi – originální obraz (a) a výsledek CNN metody (b); pacient *S16080* s počátečními metastázami obou typů – originální obraz (c) a výsledek CNN metody (d) (modrá – lytická tkáň, oranžová – blastická tkáň, červené šipky označují příklady nejčastějších případů falešných detekcí, písmena značí druh chyby).



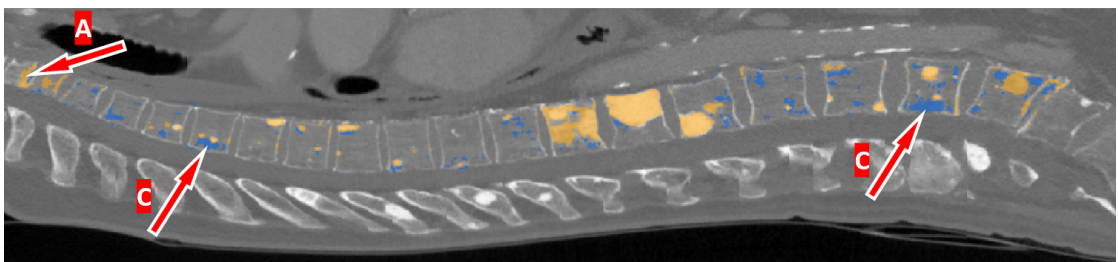
(a)



(b)

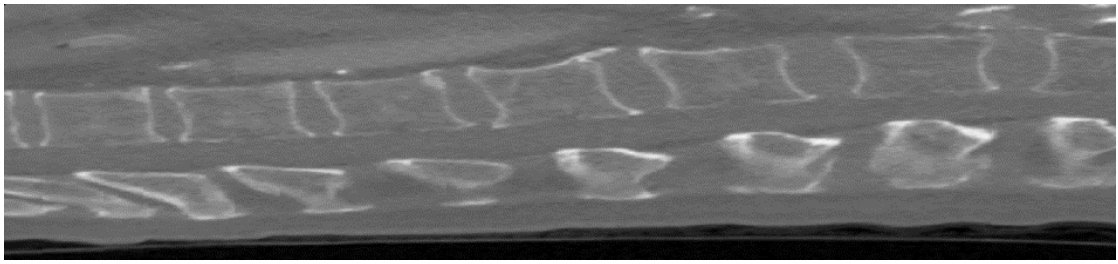


(c)

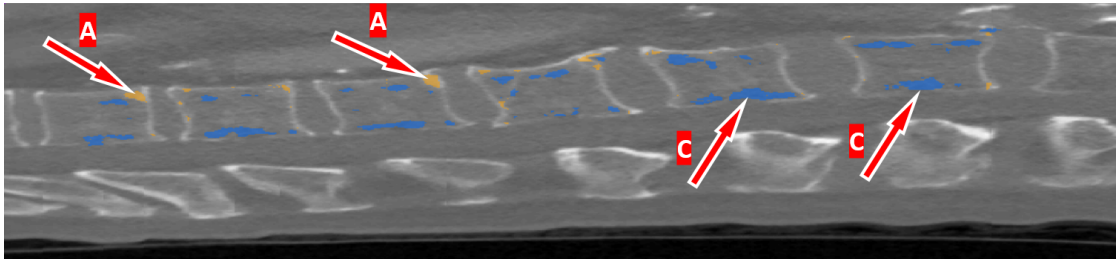


(d)

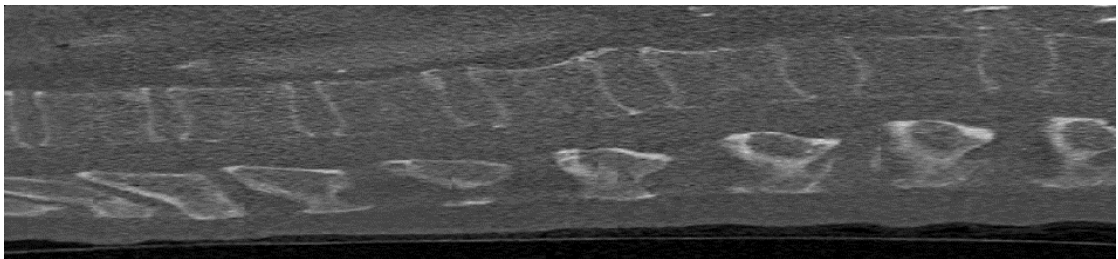
Obr. 5.7: Příklady segmentace a klasifikace nádorových tkání na datech se standardním protokolem (Praha – Obr. 4.1) – sagitální: pacient *S33880* s osteoblastickou lézí v jednom obratli – originální obraz (a) a výsledek CNN metody (b); pacient *S32350* s vícenásobnými osteoblastickými lézemi – originální obraz (c) a výsledek CNN metody (d) (modrá – lytická tkáň, oranžová – blastická tkáň, červené šipky označují příklady nejčastějších případů falešných detekcí, písmena značí druh chyby).



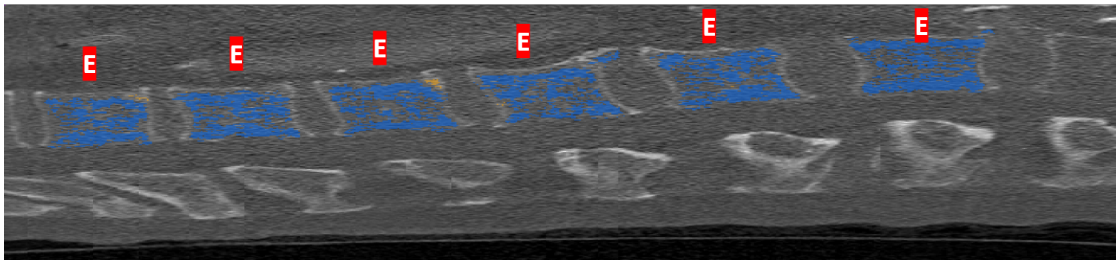
(a)



(b)

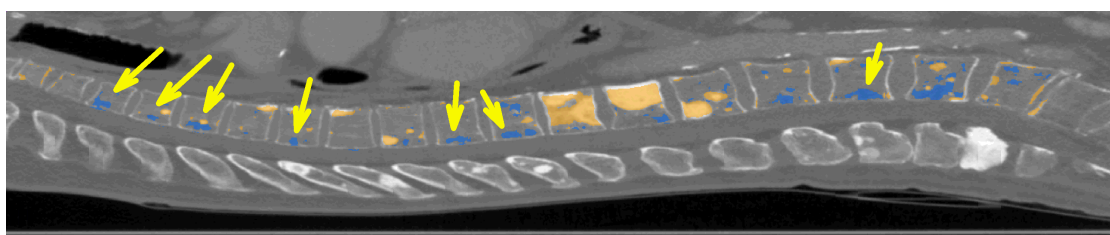


(c)

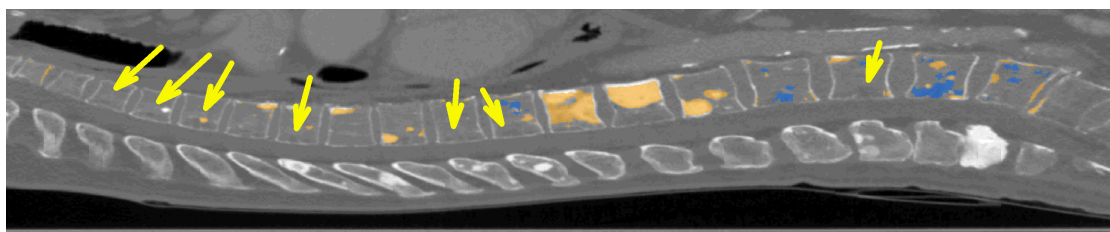


(d)

Obr. 5.8: Příklady segmentace a klasifikace nádorových tkání na datech s nízkodávkovým protokolem (Brno – Obr. 4.2) – sagitální: pacient *S143970* bez diagnostikovaných lézí rekonstruovaný s B filtrem – originální obraz (a) a výsledek CNN metody (b); totožný snímek pacienta *S143970*, pouze rekonstruovaný s YB filtrem – originální obraz (c) a výsledek CNN metody (d) (modrá – lytická tkáň, oranžová – blastická tkáň, červené šipky označují příklady nejčastějších případů falešných detekcí, písmena značí druh chyby).



(a)



(b)

Obr. 5.9: Příklad výsledku snížení počtu falešně detekovaných lézí (Praha – Obr. 4.1) – sagitální: pacient *S32350* s vícenásobnými osteoblastickými lézemi – výsledek **CNN** metody (a) a výsledek **CNN+RF** metody (b) (modrá – lytická tkáň, oranžová – blastická tkáň, žluté šipky označují příklady správně odstraněných falešných detekcí).

5.5 Shrnutí dosažených výsledků

Hlavní částí dizertační práce byl návrh a realizace výpočetních metod pro automatickou detekci, segmentaci a klasifikaci kostních metastatických lézí v obratlích. Pro tyto účely byly nejprve na vytvořené DATABÁZI A otestovány algoritmy dostupné z předchozího řešení projektu (metody využívající k segmentaci lézí Markovských náhodných polí – **MRF**, algoritmu řezu obrazovým grafem – **GRAF**, a klasifikaci voxelů umělou neuronovou sítí se vstupními daty v podobě texturních a tvarových příznaků – **TEX**), abychom zjistili jejich spolehlivost. Průměrné výsledky dosažené těmito metodami jsou shrnuty v Tab. 5.3.

Tab. 5.3: Přehled průměrných výsledků nemodifikovaných metod (voxelově).

Metoda	Senzitivita	Specifická	Youndenův index	Výpočetní čas/pac.
MRF	42,9	71,9	14,8	42,1 min
GRAF	28,0	72,8	0,8	6,0 min
TEX	36,6	73,8	10,4	13,1 hod

Vzhledem k relativně nízkým hodnotám úspěšnosti a vysoké časové náročnosti těchto metod bylo provedeno jejich přepracování a optimalizace algoritmů z hlediska úpravy použitých statistických modelů a apriorních znalostí, paralelizace výpočtů a zjednodušení zdrojového kódu. Následně proběhlo opětovné testování, které ukázalo zlepšení všech tří dostupných algoritmů z hlediska úspěšnosti i výrazného snížení časové náročnosti. Dosažené průměrné výsledky jsou shrnuty v Tab. 5.4.

Tab. 5.4: Přehled průměrných výsledků modifikovaných metod (voxelově).

Metoda	Senzitivita	Specifická	Youndenův index	Výpočetní čas/pac.
MRF	49,1	72,2	21,2	1,5 min
GRAF	50,6	73,2	23,8	6,0 min
TEX	49,4	73,2	22,7	6,6 hod

Vzhledem k tomu, že každý z přepracovaných algoritmů dosahoval obdobné objektivní úspěšnosti a současně vzájemně rozdílných výsledků (odlišně segmentované léze a falešné detekce na odlišných místech), bylo přistoupeno k jejich kombinaci do jediného algoritmu (**CMB**) a jeho realizaci v podobě experimentálního softwaru. Tento krok se ukázal jako vhodný, což dokazuje mírné zlepšení úspěšnosti (Tab. 5.5); tímto přístupem došlo k navýšení časové náročnosti v důsledku nutnosti výpočtu všech 3 metod a jejich kombinace (Tab. 5.5).

Tab. 5.5: Výsledek kombinace modifikovaných metod (voxelově).

Metoda	Senzitivita	Specifická	Youndenův index	Výpočetní čas/pac.
CMB	50,9	73,4	24,3	6,7 hod

Z důvodu požadavku projektu na algoritmus pro rychlou předběžnou segmentaci a klasifikaci byla navržena metoda založená na fuzzy k-means shlukové analýze s využitím znalosti intenzit okolních voxelů a statistického modelu distribucí intenzit jednotlivých klasifikovaných tkání (**FCM**). Touto metodou bylo dosaženo největší efektivity dle poměru časové náročnosti k úspěšnosti segmentace (Tab. 5.6). Úspěšnost však stále nebyla dostatečná pro splnění požadavků lékařů na přesnou segmentaci a klasifikaci nádorových tkání.

Tab. 5.6: Výsledek rychlé metody s využitím fuzzy k-means (voxelově).

Metoda	Senzitivita	Specifická	Youndenův index	Výpočetní čas/pac.
FCM	59,8	74,8	34,7	48,0 s

Z předchozího řešení projektu (zejména **TEX** metody) vyplynulo, že tvarové a texturní příznaky přinášejí podstatnou informaci potřebnou pro správnou klasifikaci lézí. Problémem však bylo obtížné definování relevantních příznaků, jejichž určení nebylo na základě konzultace s lékaři možné vzhledem k velké míře nejistoty. Bylo tedy nutné extrahovat velké množství různých příznaků a provést rozsáhlou statistickou analýzu jejich relevantnosti a diskriminačních vlastností. Tento postup byl využit právě v případě **TEX** metody v rámci diplomové práce Novosadova (2014) a dále v rámci jejího přepracování v této dizertační práci. Výsledkem však vždy byl algoritmus s vysokou výpočetní náročností způsobenou výpočtem složitých příznaků, které však stále nevedly na požadovanou vyšší úspěšnost segmentace a klasifikace lézí.

Z výše zmíněných důvodů bylo navrženo využití moderních metod strojového učení s učitelem, konkrétně konvolučních neuronových sítí, které dokáží relevantní příznaky samy extrahovat během procesu učení a následně tyto příznaky využít pro klasifikaci. Výsledkem je původní, realizovaná a naučená 17 vrstvá konvoluční neuronová síť (metoda **CNN**), která dosahuje nejlepších výsledků (Tab. 5.7) při čase 19 *min* na analýzu celé páteře. Dále byla tato síť optimalizována pro paralelní výpočet na grafické kartě, čímž bylo dosaženo času 4 *min* na pacienta. Učení sítě trvalo 10 dní bez použití grafického procesoru, s jejím využitím pouze 12 *hod*.

Navržená metoda **CNN** byla dále doplněna následnou meta-analýzou lézí užívající klasifikátoru typu náhodného lesa (metoda **CNN+RF**). Hodnocení klinické využitelnosti tohoto algoritmu bylo s ohledem na objektový charakter meta-analýzy

Tab. 5.7: Výsledek metody využívající konvoluční neuronové sítě (voxelově).

Metoda	Senzitivita	Specifická	Youdenův index	Výpočetní čas/pac.
CNN	68,9	82,9	51,8	19 (CPU) 4 (GPU) min

provedeno na jednotlivých segmentovaných a klasifikovaných lézích rovněž objektově; nebyly tedy hodnoceny jednotlivé voxely zvlášť. Automaticky detekované objekty byly hodnoceny na základě míry překryvu s expertně anotovanými objekty. Odděleně byly hodnoceny různé velikosti objektů a různé typy nádorové tkáně. Dosažené výsledky navržených metod byly objektivně porovnány s mezi-expertními rozdíly (rozdíly v nezávislé anotaci dat různými lékaři), a také s algoritmy publikovanými dalšími autory a dále diskutovány se spolupracujícími lékařskými specialisty. Dle objektového hodnocení pro jednovoxelový (z důvodu zachování stejného kritéria hodnocení jako v jiných publikacích) překryv dosahuje nejlepší navržená **CNN+RF** metoda pro lytické (sklerotické) léze pro velikosti lézí $> 1,40 \text{ mm}^3$ a $> 300,00 \text{ mm}^3$ následujících průměrných výsledků (Tab. 5.8). Detailnější výsledky s analýzou dopadu míry překryvu na výslednou úspěšnost metody jsou popsány v Kap. 5.2.

Tab. 5.8: Výsledek metody využívající konvoluční neuronové sítě v kombinaci s meta-analýzou lézí pomocí klasifikátoru náhodného lesa (objektově s rozdělením výsledků dle typu lézí).

Typ lézí	Senzitivita	Počet falešně pozitivních detekcí/pac.
Lytické – všechny	80,0	31,02
Lytické – jen velké	89,2	2,87
Sklerotické – všechny	92,0	45,63
Sklerotické – jen velké	99,5	5,92

Dosažené výsledky jsou lepší nebo srovnatelné s výsledky algoritmů publikovaných dalšími autory a také dosahují vyšších hodnot senzitivity než byla vypočtena mezi experty (kdy byl vždy jeden z expertů určen jako zlatý standard a druhý byl k němu porovnáván – výsledné hodnoty jsou poté dány průměrem z těchto dvou výpočtů). Počet falešně pozitivních detekcí je u všech automatických metod v porovnání s mezi-expertními hodnotami výrazně vyšší. Za účelem využití metody jako systému počítačem podporované detekce/diagnostiky je však preferována vyšší senzitivita a následný vyšší počet falešně pozitivních detekcí je tolerován (pro lékařského experta je snazší odstranit nebo ignorovat přesně dané falešně pozitivní detekce než dohledávat chybějící léze v celých datech).

V porovnání s metodami jiných autorů je výhodou námi navržené metody schopnost analyzovat také krční (cervikální) obratle, které v jiných publikacích nebyly

testovány z důvodů nízkého rozlišení v ose z a malých rozměrů těchto obratlů. Při dostupnosti segmentovaných masek je z principu možná aplikace navržené metody také na obratlové výběžky či žebra, avšak úspěšnost metody v těchto oblastech nebyla testována.

Dále byly výsledky zhodnoceny také subjektivně na neanotovaných databázích a konzultovány se spolupracujícími specialisty; dle expertního subjektivního pohledu nebyly mezi výsledky získanými na experimentálních (onkologických) i rutinních datech výrazné rozdíly. Výhodou je tedy využitelnost metody také na data získaná při jiných vyšetřeních (příkladem může být analýza hrudních obratlů při vyšetření primárního nádoru plic), bez navýšení patientské dávky. Pro další klinické testování byl realizován komplexní experimentální software, zahrnující detekci, identifikaci a segmentaci obratlů (Jakubicek, 2020), detekci, segmentaci a klasifikaci lytických a sklerotických lézí v obratlích (tato dizertační práce), analýzu tloušťky kortikální části obratlů (Kodym, 2017), a segmentaci kalibračních tkání pro účely následného odhadu minerální hustoty trabekulární části obratlů (Kodym, 2017; Salplachta, 2017).

6 Závěr

V rámci řešení dizertační práce byla získána řada patientských obrazových CT dat. Dostupná 3D CT obrazová data byla anotována nezávisle dvěma radiology. Dále byl vytvořen software pro pozdější ověřování funkčnosti automatických počítačem podporovaných detekčních/diagnostických metod (CAD). Vytvořená databáze anotovaných obrazových dat byla po oddělených částech použita během návrhu automatických metod a jako zlatý standard pro jejich vyhodnocení.

Hlavní částí dizertační práce byl návrh a realizace metod pro automatickou detekci, segmentaci a klasifikaci kostních metastatických lézí v obratlích. Pro tyto účely byly nejprve na získané databázi otestovány algoritmy dostupné z předchozího řešení projektu, abychom zjistili jejich spolehlivost. Některé části těchto algoritmů byly přepracovány a optimalizovány ke zlepšení jejich efektivity. Pro rychlou předběžnou segmentaci a klasifikaci byla dále navržena metoda založená na fuzzy k-means shlukové analýze s využitím znalosti intenzit okolních voxelů a statistického modelu distribucí intenzit jednotlivých klasifikovaných tkání. Touto metodou bylo dosaženo největší efektivity dle poměru časové náročnosti k úspěšnosti segmentace.

Z předchozího řešení projektu vyplynulo, že tvarové a texturní příznaky přinášejí podstatnou informaci potřebnou pro správnou klasifikaci lézí. Problémem však bylo obtížné definování relevantních příznaků, jejichž určení nebylo na základě konzultace s lékaři možné vzhledem k velké míře nejistoty. Z tohoto důvodu bylo navrženo využití moderních metod strojového učení s učitelem, konkrétně konvolučních neuronových sítí, které dokáží relevantní příznaky samy extrahovat během procesu učení a dále tyto příznaky využít pro klasifikaci. Výsledkem je navržená, realizovaná a naučená původní 17 vrstvá konvoluční neuronová síť. Navržená metoda byla dále doplněna meta-analýzou kandidátů lézí klasifikátorem ve formě náhodného lesa za účelem redukce falešně pozitivních detekcí lézí.

Dosažené výsledky navržených a realizovaných metod byly objektivně porovnány s algoritmy publikovanými jinými autory, a dále diskutovány se spolupracujícími specialisty. Dosažené výsledky jsou lepší nebo porovnatelné s algoritmy publikovanými jinými autory. Hlavní výsledky dizertace byly publikovány (Chmelik a kol., 2018b) v recenzovaném impaktovaném časopise *Medical Image Analysis* (JIF 2018: 8,880 – Q1, D1 dle WoS). Předběžné výsledky původní konvoluční neuronové sítě byly publikovány v konferenčním příspěvku (Chmelik a kol., 2019) ve sborníku mezinárodní konference *IUPESM 2018*. Přehled výchozího stavu řešení problému ve světě byl rovněž publikován jako přehledový článek Chmelik a kol. (2018a).

Z výše uvedeného přehledu vyplývá, že všechny tři definované cíle dizertační práce lze považovat za dosažené.

Literatura

- AHMED, M. N., YAMANY, S. M., MOHAMED, N., FARAG, A. A., A MORIARTY, T., 2002. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*. **21**(3), 193 – 199. ISSN 0278-0062.
- AMARI, S.-I., 1998. Natural gradient works efficiently in learning. *Neural Computation*. **10**(2), 251 – 276. ISSN 0899-7667.
- ANGELINE, P. J., SAUNDERS, G. M., A POLLACK, J. B., 1994. An evolutionary algorithm that constructs recurrent neural networks. *IEEE Transactions on Neural Networks*. **5**(1), 54 – 65. ISSN 1045-9227.
- AREL, I., ROSE, D. C., A KARNOWSKI, T. P., 2010. Deep machine learning – a new frontier in artificial intelligence research [research frontier]. *IEEE Computational Intelligence Magazine*. **5**(4), 13 – 18. ISSN 1556-603X.
- BADRINARAYANAN, V., KENDALL, A., A CIPOLLA, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **39**(12), 2481–2495. ISSN 0162-8828.
- BEDIN, V., ADAM, R. L., DE SÁ, B. C., LANDMAN, G., A METZE, K., 2010. Fractal dimension of chromatin is an independent prognostic factor for survival in melanoma. *BMC cancer*. **10**(260), 1 – 6. ISSN 1471-2407.
- BENOU, A., VEKSLER, R., FRIEDMAN, A., A RAVIV, T. R., 2017. Ensemble of expert deep neural networks for spatio-temporal denoising of contrast-enhanced MRI sequences. *Medical Image Analysis*. **42**, 145 – 159. ISSN 1361-8415.
- BERTHOD, M., KATO, Z., YU, S., A ZERUBIA, J., 1996. Bayesian image classification using markov random fields. *Image and Vision Computing*. **14**(4), 285 – 295. ISSN: 0262-8856.
- BEZDEK, J. C., EHRLICH, R., A FULL, W., 1984. Fcm: The fuzzy c-means clustering algorithm. *Computers and Geosciences*. **10**(2), 191 – 203. ISSN 0098-3004.
- BIANCO, S., CADENE, R., CELONA, L., A NAPOLETANO, P., 2018. Benchmark analysis of representative deep neural network architectures. *IEEE Access*. **6**(1), 64270–64277. ISSN 2169-3536.

- BOYKOV, Y., A KOLMOGOROV, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **26**(9), 1124 – 1137. ISSN 0162-8828.
- BOYKOV, Y. Y., A JOLLY, M.-P., 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. **In: Proceedings of the Eighth IEEE International Conference on Computer Vision**. 1. Vancouver: IEEE, 105 – 112. ISBN 0-7695-1143-0.
- BRAHNAM, S., JAIN, L. C., NANNI, L., LUMINI, A., A KOL., 2013. *Local binary patterns: new variants and applications*. 1. vyd. Heidelberg: Springer. ISBN 978-3-64239-288-7.
- BREIMAN, L., 1984. *Classification and regression trees*: Routledge. ISSN 978-0-412-04841-8.
- BREIMAN, L., 2001. Random forests. *Machine Learning*. **45**(1), 5 – 32. ISSN 1573-0565.
- BRODERSEN, K. H., ONG, C. S., STEPHAN, K. E., A BUHMANN, J. M., 2010. The balanced accuracy and its posterior distribution. **In: Proceedings of the 20th International Conference on Pattern Recognition**. Istanbul: IEEE, 3121 – 3124. ISBN 978-1-42447-542-1.
- BURNS, J. E., YAO, J., WIESE, T. S., MUÑOZ, H. E., JONES, E. C., A SUMMERS, R. M., 2013. Automated detection of sclerotic metastases in the thoracolumbar spine at CT. *Radiology*. **268**(1), 69 – 78. ISSN 0033-8419.
- CHARBONNIER, J.-P., VAN RIKXOORT, E. M., SETIO, A. A., SCHAEFER-PROKOP, C. M., VAN GINNEKEN, B., A CIOMPI, F., 2017. Improving airway segmentation in computed tomography using leak detection with convolutional networks. *Medical Image Analysis*. **36**, 52 – 60. ISSN 1361-8415.
- CHEN, H., QI, X., YU, L., DOU, Q., QIN, J., A HENG, P.-A., 2017. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis*. **36**, 135 – 146. ISSN 1361-8415.
- CHHACHHIYA, D., SHARMA, A., A GUPTA, M., 2017. Designing optimal architecture of neural network with particle swarm optimization techniques specifically for educational dataset. **In: Proceedings of the 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence**. Noida: IEEE, 52 – 57. ISBN 978-1-5090-3519-9.

- CHMELIK, J., JAKUBICEK, R., A JAN, J., 2018a. Tumorous spinal lesions: Computer aided diagnosis and evaluation based on CT data – A review. *Current Medical Imaging Reviews*. **14**(5), 1 – 10. ISSN 1573-4056.
- CHMELIK, J., JAKUBICEK, R., JAN, J., OUREDNICEK, P., LAMBERT, L., AMADORI, E., A GAVELLI, G., 2018b. Deep convolutional neural network-based segmentation and classification of difficult to define spinal lesions in 3D CT data. *Medical Image Analysis*. **49**(C), 76 – 88. ISSN 1361-8415.
- CHMELIK, J., JAKUBICEK, R., JAN, J., OUREDNICEK, P., LAMBERT, L., AMADORI, E., A GAVELLI, G., 2019. Fully automatic CAD system for segmentation and classification of spinal metastatic lesions in CT data. **In:** *World Congress on Medical Physics and Biomedical Engineering 2018*. 68 of *IFMBE Proceedings*. Singapore: Springer, 155 – 158. ISBN 978-981-10-9034-9.
- ÇIÇEK, Ö., ABDULKADIR, A., LIENKAMP, S. S., BROX, T., A RONNEBERGER, O., 2016. 3d u-net: Learning dense volumetric segmentation from sparse annotation. **In:** Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G., a Wells, W. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Cham: Springer International Publishing, 424–432. ISBN 978-3-319-46723-8.
- COSTA, L. D. F. D., A CESAR JR, R. M., 2001. *Shape analysis and classification: theory and practice*. 1. vyd. Boca Raton: CRC Press. ISBN 978-1-4398-5598-0.
- COSTELLOE, C. M., ROHREN, E. M., MADEWELL, J. E., HAMAOKA, T., THE-RIAULT, R. L., YU, T.-K., LEWIS, V. O., MA, J., STAFFORD, R. J., TARI, A. M., HORTOBAGYI, G. N., A UENO, N. T., 2009. Imaging bone metastases in breast cancer: techniques and recommendations for diagnosis. *The Lancet Oncology*. **10**(6), 606 – 614. ISSN 1470-2045.
- D’AMBROSIO, D. B., A STANLEY, K. O., 2007. A novel generative encoding for exploiting neural network sensor and output geometry. **In:** *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*. London: ACM Press, 974 – 981. ISBN 978-1-59593-697-4.
- DEMEL, J., 2014. *Segmentace 3D obrazových dat s využitím grafové reprezentace*. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Jiří Jan.
- DERIN, H., A COLE, W. S., 1986. Segmentation of textured images using gibbs random fields. *Computer Vision, Graphics, and Image Processing*. **35**(1), 72 – 98. ISSN 0734-189X.

- DERIN, H., A ELLIOTT, H., 1987. Modeling and segmentation of noisy and textured images using gibbs random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **9**(1), 39 – 55. ISSN 0162-8828.
- DHUNGEL, N., CARNEIRO, G., A BRADLEY, A. P., 2017. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. *Medical Image Analysis*. **37**, 114 – 128. ISSN 1361-8415.
- DOU, Q., YU, L., CHEN, H., JIN, Y., YANG, X., QIN, J., A HENG, P.-A., 2017. 3D deeply supervised network for automated segmentation of volumetric medical images. *Medical Image Analysis*. **41**, 40 – 54. Special Issue on the 2016 Conference on Medical Image Computing and Computer Assisted Intervention (Analog to MICCAI 2015). ISSN 1361-8415.
- DUCHI, J., HAZAN, E., A SINGER, Y., 2011. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*. **12**(Jul), 2121 – 2159. ISSN 1532-4435.
- ELLIOTT, H., DERIN, H., CRISTI, R., A GEMAN, D., 1984. Application of the gibbs distribution to image segmentation. **In:** *Proceedings of the '84 IEEE International Conference on Acoustics, Speech, and Signal Processing*. 9. San Diego: IEEE, 678 – 681.
- EVANGELISTA, L., BERTOLDO, F., BOCCARDO, F., CONTI, G., MENCHI, I., MUNGAI, F., RICARDI, U., A BOMBARDIERI, E., 2016. Diagnostic imaging to detect and evaluate response to therapy in bone metastases from prostate cancer: current modalities and new horizons. *European Journal of Nuclear Medicine and Molecular Imaging*. **43**(8), 1546 – 1562. ISSN 1619-7070.
- FREUND, Y., A SCHAPIRE, R. E., 1995. A decision-theoretic generalization of on-line learning and an application to boosting. **In:** *Proceedings of the Second European Conference on Computational Learning Theory*. Barcelona: Springer, 23 – 37. ISBN 978-3-540-59119-1.
- GASBARRINI, A., CAPPuccio, M., MIRABILE, L., BANDIERA, S., TERZI, S., BARBANTI BRÒDANO, G., A BORIANI, S., 2004. Spinal metastases: treatment evaluation algorithm. *European Review for Medical and Pharmacological Sciences*. **8**(6), 265 – 274. ISSN 2284-0729.
- GLAS, A. S., LIJMER, J. G., PRINS, M. H., BONSEL, G. J., A BOSSUYT, P. M., 2003. The diagnostic odds ratio: a single indicator of test performance. *Journal of Clinical Epidemiology*. **56**(11), 1129 – 1135. ISSN 0895-4356.

- HAMMON, M., DANKERL, P., TSYMBAL, A., WELS, M., KELM, M., MAY, M., SUEHLING, M., UDER, M., A CAVALLARO, A., 2013. Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography. *European Radiology*. **23**(7), 1862 – 1870. ISSN 0938-7994.
- HARALICK, R. M., SHANMUGAM, K., A KOL., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*. **3**(6), 610 – 621. ISSN 0018-9472.
- HARUŠTIAKOVÁ, D., 2012. *Vícerozměrné statistické metody v biologii*. 1. vyd. Brno: Akademické nakladatelství CERM. ISBN 978-80-7204-791-8.
- HAVAEI, M., DAVY, A., WARDE-FARLEY, D., BIARD, A., COURVILLE, A., BENGIO, Y., PAL, C., JODOIN, P.-M., A LAROCHELLE, H., 2017. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*. **35**, 18 – 31. ISSN 1361-8415.
- HE, K., ZHANG, X., REN, S., A SUN, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. **In: Proceedings of the IEEE International Conference on Computer Vision 2015**. Santiago: IEEE, 1026 – 1034. ISBN 978-1-4673-8390-5.
- HE, K., ZHANG, X., REN, S., A SUN, J., 2016. Deep residual learning for image recognition. **In: Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition**. Las Vegas: IEEE, 770 – 778. ISSN 1063-6919.
- HE, L., CHAO, Y., SUZUKI, K., A WU, K., 2009. Fast connected-component labeling. *Pattern Recognition*. **42**(9), 1977 – 1987. ISSN 0031-3203.
- HINTON, G. E., SRIVASTAVA, N., KRIZHEVSKY, A., SUTSKEVER, I., A SALAKHUTDINOV, R. R., 2012. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint*. **1207.0580**(v1), 1 – 18.
- HOLČÍK, J., 2012. *Analýza a klasifikace dat*. 1. vyd. Brno: Akademické nakladatelství CERM. ISBN 978-807-2047-932.
- HUANG, S.-F., A CHIAN, K.-H., 2012. Automatic detection of bone metastasis in vertebrae by using CT images. **In: Proceedings of the World Congress on Engineering 2012**. 2. London: IEEE, 1166 – 1171. ISBN 978-988-19252-1-3.
- HUBEL, D. H., A WIESEL, T. N., 1959. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*. **148**(3), 574 – 591. ISSN 1469-7793.

HUBEL, D. H., A WIESEL, T. N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*. **160**(1), 106 – 154. ISSN 1469-7793.

HUBEL, D. H., A WIESEL, T. N., 1968. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*. **195**(1), 215 – 243. ISSN 1469-7793.

IOFFE, S., A SZEGEDY, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint*. **1502.03167**(v3), 1 – 11.

JAKUBICEK, R., 2020. *Metody segmentace a identifikace deformovaných obratlů ve 3D CT datech onkologických pacientů*. Brno. Dizertační práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Jiří Jan.

JAKUBICEK, R., CHMELIK, J., JAN, J., OUREDNICEK, P., LAMBERT, L., A GAVELLI, G., 2019a. Fully automatic CAD system for spine localisation and vertebra segmentation in CT data. **In: World Congress on Medical Physics and Biomedical Engineering 2018**. 68 of *IFMBE Proceedings*. Singapore: Springer, 223 – 226. ISBN 978-981-10-9034-9.

JAKUBICEK, R., CHMELIK, J., JAN, J., OUREDNICEK, P., LAMBERT, L., A GAVELLI, G., 2019b. Learning-based vertebra localisation and labelling in 3D CT data of possibly incomplete and pathological spines. *Computer Methods and Programs in Biomedicine*. **183**(105081), 1 – 9. ISSN 0169-2607.

JAKUBICEK, R., CHMELIK, J., KODYM, O., JAN, J., OUREDNICEK, P., LAMBERT, L., A GAVELLI, G., 2020. An online learned approach for 3D sub-voxel cortical bone thickness estimation with a robust segmentation of pathological vertebrae in CT data. . **1**(1), 1 – 1. [unpublished].

JAKUBICEK, R., CHMELIK, J., OUREDNICEK, P., A JAN, J., 2019c. Deep-learning-based fully automatic spine centerline detection in ct data. **In: 2019 41th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. 19. IEEE: IEEE, 2407–2410. ISBN 978-1-5386-1312-2.

JAMALUDIN, A., KADIR, T., A ZISSERMAN, A., 2017. SpineNet: Automated classification and evidence visualization in spinal MRIs. *Medical Image Analysis*. **41**, 63 – 73. ISSN 1361-8415.

- JAMES, J., LAM, A. Y., A LI, V. O., 2011. Evolutionary artificial neural network based on chemical reaction optimization. **In:** *Proceedings of the IEEE Congress on Evolutionary Computation 2011*. New Orleans: IEEE, 2083 – 2090. ISBN 978-1-4244-7834-7.
- JAN, J., 2002. *Číslicová filtrace, analýza a restaurace signálů*. 2. vyd. Brno: VU-TIUM. ISBN 80-214-2911-9.
- JAN, J., 2019. *Medical Image Processing, Reconstruction and Analysis: Concepts and Methods*. 2. vyd. Boca Raton: CRC Press Taylor & Francis Group. ISBN 978-113-8310-285.
- JAN, J., CHMELÍK, J., JAKUBÍČEK, R., OUŘEDNÍČEK, P., A GAVELLI, G., 2018. Spine lesion analysis in 3D CT data – reporting on research progress. **In:** *Proceedings of the 6th International Conference on Biotechnology and Bioengineering*. 1956 of *AIP Conference Proceedings*. New York: AIP Publishing, 11 – 14. ISBN 978-0-7354-1657-4.
- JAN, J., NOVOSADOVÁ, M., DEMEL, J., OUŘEDNÍČEK, P., CHMELÍK, J., A JAKUBÍČEK, R., 2015. Combined bone lesion analysis in 3D CT data of vertebrae. **In:** *Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Milano: IEEE, 6374 – 6377. ISBN 978-1-42449-270-1.
- KAKHKI, V. R. D., ANVARI, K., SADEGHI, R., MAHMOUDIAN, A.-S., A TORABIAN-KAKHKI, M., 2013. Pattern and distribution of bone metastases in common malignant tumors. *Nuclear Medicine Review*. **16**(2), 66 – 69. ISSN 1644-4345.
- KAMNITSAS, K., LEDIG, C., NEWCOMBE, V. F., SIMPSON, J. P., KANE, A. D., MENON, D. K., RUECKERT, D., A GLOCKER, B., 2017. Efficient multi-scale 3D CNN with fully connected crf for accurate brain lesion segmentation. *Medical Image Analysis*. **36**, 61 – 78. ISSN 1361-8415.
- KINGMA, D., A BA, J., 2014. Adam: A method for stochastic optimization. *arXiv preprint*. **1412.6980**(v9), 1 – 15.
- KODYM, O., 2017. *Analýza 3D CT obrazových dat se zaměřením na stanovení hustoty kostních elementů*. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Jiří Jan.

- KOOI, T., LITJENS, G., VAN GINNEKEN, B., GUBERN-MÉRIDA, A., SÁNCHEZ, C. I., MANN, R., DEN HEETEN, A., A KARSSEMEIJER, N., 2017. Large scale deep learning for computer aided detection of mammographic lesions. *Medical Image Analysis*. **35**, 303 – 312. ISSN 1361-8415.
- KRIZHEVSKY, A., SUTSKEVER, I., A HINTON, G. E., 2012. Imagenet classification with deep convolutional neural networks. **In: *Proceedings of the Advances in Neural Information Processing Systems 25***. Lake Tahoe: MIT Press, 1097 – 1105. ISBN 978-1-62748-003-1.
- LECUN, Y., BENGIO, Y., A GEOFFREY, H., 2015. Deep learning. *Nature*. **521**(7553), 436 – 444. ISSN 0028-0836.
- LECUN, Y., BOTTOU, L., BENGIO, Y., A HAFFNER, P., 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. **86**(11), 2278 – 2324. ISSN 0018-9219.
- LECUN, Y., A KOL., 1989. Generalization and network design strategies. **In: *Connectionism in Perspective***. 1. vyd. Zurich: Elsevier, 143 – 155. ISBN 978-0-44459-876-9.
- LÉZORAY, O., A GRADY, L., 2012. *Image processing and analysis with graphs: theory and practice*. 1. vyd. Boca Raton: Taylor. ISBN 978-143-9855-072.
- LIKAS, A., A STAFYLOPATIS, A., 2000. Training the random neural network using quasi-Newton methods. *European Journal of Operational Research*. **126**(2), 331 – 339. ISSN 0377-2217.
- LI, S., 2009. *Markov Random Field Modeling in Image Analysis*. 3. vyd. London: Springer. ISBN 978-1-84800-278-4.
- LI, S. Z., 2000. *Stochastic Processes: Modeling and Simulation*. 1. vyd. 20. ISBN 978-0-44450-013-7.
- LITJENS, G., KOOI, T., BEJNORDI, B. E., SETIO, A. A. A., CIOMPI, F., GHAFORIAN, M., VAN DER LAAK, J. A., VAN GINNEKEN, B., A SÁNCHEZ, C. I., 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis*. **42**(1), 60 – 88. ISSN 1361-8415.
- MALINSKY, M., 2012. *Advanced Algorithms for 3D Medical Image Data Fusion in Specific Medical Problems*. Dizertační práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Jiří Jan.

- MARTENS, J., 2010. Deep learning via Hessian-free optimization. **In:** *Proceedings of the 27th International Conference on Machine Learning*. Haifa: International Machine Learning Society, 735 – 742. ISBN 978-1-60558-907-7.
- MIRMEHDI, M., XIE, X., A SURI, J., 2008. *Handbook of Texture Analysis*. 1. vyd. Hackensack: Imperial College Press. ISBN 978-184-8161-153.
- MUÑOZ, H. E., YAO, J., BURNS, J. E., A SUMMERS, R. M., 2013. Detection of vertebral degenerative disc disease based on cortical shell unwrapping. **In:** *Proceedings of SPIE, Medical Imaging 2013: Computer-Aided Diagnosis*. International Society for Optics and Photonics, 86700C. ISBN 978-0-81949-444-3.
- NEAL, F. B., A RUSS, J. C., 2012. *Measuring shape*. 1. vyd. Boca Raton: CRC Press. ISBN 978-143-9855-980.
- NESTEROV, Y., 1983. A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Soviet Mathematics Doklady*. **27**(2), 372 – 376. ISSN 0197-6788.
- NGO, T. A., LU, Z., A CARNEIRO, G., 2017. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Medical Image Analysis*. **35**, 159 – 171. ISSN 1361-8415.
- NOVOSADOVA, M., 2014. *Segmentace 3D obrazových dat s využitím pokročilých texturních a tvarových příznaků*. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Jiří Jan.
- O'CONNOR, S. D., YAO, J., A SUMMERS, R. M., 2007. Lytic metastases in thoracolumbar spine: Computer-aided detection at CT – preliminary study. *Radiology*. **242**(3), 811 – 816. ISSN 0033-8419.
- PETER, R., MALINSKY, M., OUREDNICEK, P., A JAN, J., 2013. 3D CT spine data segmentation and analysis of vertebrae bone lesions. **In:** *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Osaka: IEEE, 2376 – 2379. ISBN 978-1-45770-215-0.
- QUELLEC, G., CHARRIÈRE, K., BOUDI, Y., COCHENER, B., A LAMARD, M., 2017. Deep image mining for diabetic retinopathy screening. *Medical Image Analysis*. **39**, 178 – 193. ISSN 1361-8415.
- ROBBINS, H., A MONRO, S., 1951. A stochastic approximation method. *The Annals of Mathematical Statistics*. **22**(3), 400 – 407. ISSN 0003-4851.

- RONNEBERGER, O., FISCHER, P., A BROX, T., 2015. U-net: Convolutional networks for biomedical image segmentation. **In:** Navab, N., Hornegger, J., Wells, W. M., a Frangi, A. F. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing, 234–241. ISBN 978-3-319-24574-4.
- RONSE, C., A DEVIJVER, P. A., 1984. *Connected Components in Binary Images: The Detection Problem*. 1. vyd. New York: Wiley. ISBN 04-719-0456-2.
- ROTH, H. R., LU, L., LIU, J., YAO, J., SEFF, A., CHERRY, K., KIM, L., A SUMMERS, R. M., 2016. Improving computer-aided detection using convolutional neural networks and random view aggregation. *IEEE Transactions on Medical Imaging*. **35**(5), 1170 – 1181. ISSN 0278-0062.
- ROTH, H. R., YAO, J., LU, L., STIEGER, J., BURNS, J. E., A SUMMERS, R. M., 2015. Detection of sclerotic spine metastases via random aggregation of deep convolutional neural network classifications. **In:** Yao, J., Glocker, B., Klinder, T., a Li, S. *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. 1. vyd. Cham: Springer, 3 – 12. ISBN 978-3-319-14148-0.
- RUSSAKOVSKY, O., DENG, J., SU, H., KRAUSE, J., SATHEESH, S., MA, S., HUANG, Z., KARPATY, A., KHOSLA, A., BERNSTEIN, M., A KOL., 2015. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*. **115**(3), 211 – 252. ISSN 0920-5691.
- RYBAK, L. D., A ROSENTHAL, D. I., 2001. Radiological imaging for the diagnosis of bone metastases. *Quarterly Journal of Nuclear Medicine*. **45**(1), 53 – 64. ISSN 0392-0208.
- SALPLACHTA, J., 2017. *Analýza 3D CT obrazových dat se zaměřením na detekci a klasifikaci specifických struktur tkání*. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Jiří Jan.
- SIMONYAN, K., VEDALDI, A., A ZISSERMAN, A., 2013. Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv preprint*. **1312.6034**(v2), 1 – 8.
- SIMONYAN, K., A ZISSERMAN, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint*. **1409.1556**(v6), 1 – 14.
- SINHA, A., SARKAR, M., MUKHERJEE, A., A KRISHNAMURTHY, B., 2017. Introspection: Accelerating neural network training by learning weight evolution. *arXiv preprint*. **1704.04959**(v1), 1 – 15.

- SOKOLOVA, M., JAPKOWICZ, N., A SZPAKOWICZ, S., 2006. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. **In:** *Proceedings of the 19th Australian Joint Conference on Artificial Intelligence*. 4304. Hobart: Springer, 1015 – 1021. ISBN 978-3-540-49787-5.
- SONKA, M., HLAVAC, V., A BOYLE, R., 2015. *Image Processing, Analysis, and Machine Vision*. 4. vyd. Stamford: Cengage Learning. ISBN 978-1-133-59360-7.
- SRIVASTAVA, N., HINTON, G. E., KRIZHEVSKY, A., SUTSKEVER, I., A SALAKHUTDINOV, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*. **15**(1), 1929 – 1958. ISSN 1532-4435.
- SUDRE, C. H., LI, W., VERCAUTEREN, T., OURSELIN, S., A JORGE CARDOSO, M., 2017. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. **In:** Cardoso, M. J., Arbel, T., Carneiro, G., Syeda-Mahmood, T., Tavares, J. M. R., Moradi, M., Bradley, A., Greenspan, H., Papa, J. P., Madabhushi, A., Nascimento, J. C., Cardoso, J. S., Belagiannis, V., a Lu, Z. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer International Publishing, 240–248. ISBN 978-3-319-67558-9.
- SUK, H.-I., LEE, S.-W., A SHEN, D., 2017. Deep ensemble learning of sparse regression models for brain disease diagnosis. *Medical Image Analysis*. **37**, 101 – 113. ISSN 1361-8415.
- SUZUKI, M. T., YAGINUMA, Y., YAMADA, T., A SHIMIZU, Y., 2006. A shape feature extraction method based on 3D convolution masks. **In:** *Proceedings of the Eighth IEEE International Symposium on Multimedia*. San Diego: IEEE, 837 – 844. ISBN 0-7695-2746-9.
- SZEGEDY, C., IOFFE, S., VANHOUCKE, V., A ALEMI, A. A., 2017. Inception-v4, Inception-ResNet and the impact of residual connections on learning. **In:** *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. San Francisco: AAAI, 4278 – 4284. ISSN 2159-5399.
- SZEGEDY, C., LIU, W., JIA, Y., SERMANET, P., REED, S., ANGUELOV, D., ERHAN, D., VANHOUCKE, V., A RABINOVICH, A., 2015. Going deeper with convolutions. **In:** *Proceedings of the 28th IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 1 – 9. ISBN 978-1-46736-965-7.

- SZEGEDY, C., ZAREMBA, W., SUTSKEVER, I., BRUNA, J., ERHAN, D., GOODFELLOW, I., A FERGUS, R., 2013. Intriguing properties of neural networks. *arXiv preprint*. **1312.6199**(v4), 1 – 10.
- TAHA, A. A., A HANBURY, A., 2015. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*. **15**(1), 29. ISSN 1471-2342.
- TAN, L. K., LIEW, Y. M., LIM, E., A MCCLAUGHLIN, R. A., 2017. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine mr sequences. *Medical Image Analysis*. **39**, 78 – 86. ISSN 1361-8415.
- THE AMERICAN CANCER SOCIETY, 2016. *Advanced cancer, metastatic cancer, and bone metastasis*. [online]. New Yourk: The American Cancer Society. [cit. 31. 8. 2017]. Dostupné z: <http://www.cancer.org>
- THEODORIDIS, S., A KOUTROUMBAS, K., 2008. *Pattern Recognition*. 4. vyd. San Diego: Academic Press. ISBN 978-1-59749-272-0.
- TOKUHASHI, Y., MATSUZAKI, H., ODA, H., OSHIMA, M., A RYU, J., 2005. A revised scoring system for preoperative evaluation of metastatic spine tumor prognosis. *Spine*. **30**(19), 2186 – 2191. ISSN 0362-2436.
- TOKUHASHI, Y., UEI, H., OSHIMA, M., A AJIRO, Y., 2014. Scoring system for prediction of metastatic spine tumor prognosis. *World Journal of Orthopedics*. **5**(3), 262 – 271. ISSN 2218-5836.
- VEDALDI, A., A LENC, K., 2015. MatConvNet – convolutional neural networks for MATLAB. **In:** *Proceeding of the 23rd ACM International Conference on Multimedia*. Brisbane: ACM. ISBN 978-1-4503-3459-4.
- WANG, S., ZHOU, M., LIU, Z., LIU, Z., GU, D., ZANG, Y., DONG, D., GEVAERT, O., A TIAN, J., 2017. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. *Medical Image Analysis*. **40**, 172 – 183. ISSN 1361-8415.
- WAN, L., ZEILER, M., ZHANG, S., CUN, Y. L., A FERGUS, R., 2013. Regularization of neural networks using dropconnect. **In:** *Proceedings of the 30th International Conference on Machine Learning*. 3. Atlanta: PMLR, 1058 – 1066. ISBN 978-1-62993-306-1.
- WELS, M., KELM, B., TSYMBAL, A., HAMMON, M., SOZA, G., SÜHLING, M., CAVALLARO, A., A COMANICIU, D., 2012. Multi-stage osteolytic spinal bone

lesion detection from CT data with internal sensitivity control. **In:** *Proceedings of SPIE, Medical Imaging 2012: Computer-Aided Diagnosis*. International Society for Optics and Photonics, 831513. ISBN 978-0-81948-964-7.

WHYNE, C., HARDISTY, M., WU, F., SKRINSKAS, T., CLEMONS, M., GORDON, L., A BASRAN, P. S., 2007. Quantitative characterization of metastatic disease in the spine. part ii. histogram-based analyses. *Medical Physics*. **34**(8), 3279 – 3285. ISSN 0094-2405.

WIESE, T., BURNS, J., YAO, J., A SUMMERS, R. M., 2011. Computer-aided detection of sclerotic bone metastases in the spine using watershed algorithm and support vector machines. **In:** *Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro 2011*. Chicago: IEEE, 152 – 155. ISBN 978-1-4244-4128-0.

WIESE, T., YAO, J., BURNS, J. E., A SUMMERS, R. M., 2012. Detection of sclerotic bone metastases in the spine using watershed algorithm and graph cut. **In:** *Proceedings of SPIE, Medical Imaging 2012: Computer-Aided Diagnosis*. International Society for Optics and Photonics, 831512. ISBN 978-0-81948-964-7.

WOLTERINK, J. M., LEINER, T., DE VOS, B. D., VAN HAMERSVELT, R. W., VIERGEVER, M. A., A IŠGUM, I., 2016. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Medical Image Analysis*. **34**, 123 – 136. Special Issue on the 2015 Conference on Medical Image Computing and Computer Assisted Intervention. ISSN 1361-8415.

XIE, X., A MIRMEHDI, M., 2007. Texems: Texture exemplars for defect detection on random textured surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **29**(8), 1454 – 1464. ISSN 0162-8828.

YANG, W., CHEN, Y., LIU, Y., ZHONG, L., QIN, G., LU, Z., FENG, Q., A CHEN, W., 2017a. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. *Medical Image Analysis*. **35**, 421 – 433. ISSN 1361-8415.

YANG, X., LIU, C., WANG, Z., YANG, J., MIN, H. L., WANG, L., A CHENG, K.-T. T., 2017b. Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric {MRI}. *Medical Image Analysis*. **42**, 212 – 227. ISSN 1361-8415.

YAO, J., BURNS, J. E., A SUMMERS, R. M., 2015. Computer aided detection of bone metastases in the thoracolumbar spine. **In:** Li, S., a Yao, J. *Spinal Imaging and Image Analysis*. 1. vyd. Cham: Springer, 97 – 130. ISBN 978-3-319-12508-4.

YAO, J., O'CONNOR, S. D., A SUMMERS, R., 2006. Computer aided lytic bone metastasis detection using regular CT images. **In:** *Proceedings of SPIE, Medical Imaging 2006: Image Processing*. International Society for Optics and Photonics, 614459. ISBN 978-0-81946-23-1.

YAO, J., O'CONNOR, S. D., A SUMMERS, R. M., 2007. Computer aided detection of lytic bone metastases in the spine using routine CT images. **In:** *Proceedings of the 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro 2007*. Arlington: IEEE, 512 – 515. ISBN 978-1-4244-0672-2.

YOSINSKI, J., CLUNE, J., NGUYEN, A., FUCHS, T., A LIPSON, H., 2015. Understanding neural networks through deep visualization. *arXiv preprint*. **1506.06579**(v1), 1 – 12.

YU, X.-H., 1993. Acceleration of backpropagation learning using optimised learning rate and momentum. *Electronics Letters*. **29**(14), 1288 – 1290. ISSN 0013-5194.

ZEILER, M. D., 2012. ADADELTA: an adaptive learning rate method. *arXiv preprint*. **1212.5701**(v1), 1 – 6.

ZEILER, M. D., A FERGUS, R., 2014. Visualizing and understanding convolutional networks. **In:** *European conference on computer vision*. Cham: Springer, 818 – 833. ISBN 978-3-319-10590-1.

ZENG, J., SONG, Y., LIU, H., GONG, Q., LI, T., LIU, L., HU, Y., PEI, F., A RAO, S., 2007. The predictive value of the tokuhashi revised scoring system for the survival time of patients with spinal metastases. *Journal of Sichuan University (Medical Science Edition)*. **38**(3), 488 – 491. ISSN 1672-173X.

ZOU, X., GREJS, A., LI, H., HØY, K., HANSEN, E., A BÜNGER, C., 2006. Estimation of life expectancy for selecting surgical procedure and predicting prognosis of extradural spinal metastases. *Chinese Journal of Cancer*. **25**(11), 1406 – 1410. ISSN 1000-467X.