

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE

Brno, 2022

Kryštof Kopečný



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

LOKALIZACE INTRAKRANIÁLNÍCH HEMORAGIÍ V AXIÁLNÍCH ŘEZECH CT SNÍMKŮ HLAVY.

INTRACRANIAL HEMORRHAGE LOCALIZATION IN AXIAL SLICES OF HEAD CT IMAGES

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Kryštof Kopečný

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Jakub Nemček

BRNO 2022

Abstrakt

Tématem této práce je lokalizace intrakraniálních krvácení na snímcích počítačové tomografie pomocí jednofázového a dvoufázového detektoru založených na konvolučních neuronových sítích. Práce mimo základní poznatky z umělé inteligence klíčové pro pochopení procesu detekce objektů podává medicínské informace o patologii a zobrazování krvácení. Znalosti ze zmíněných okruhů se využijí pro vlastní návrh algoritmu pracujícím s veřejně dostupnými databázemi CT snímků hlavy a jejich anotací CQ500 a BHX. Součástí práce je vyhodnocení výsledků a jejich diskuse.

Klíčová slova

Neuronové sítě, konvoluční neuronové sítě, detekce objektů, YOLO, Faster-RCNN, intrakraniální krvácení

Abstract

This thesis is focused on detection of intracranial hemorrhage in CT images using both one-stage and two-stage object detectors based on convolutional neural networks. The fundamentals of intracranial hemorrhage pathology and CT imaging as well as essential insight into computer vision and object detection are listed in this work. The knowledge of these fields of studies is a starting point for the implementation of hemorrhage detector. The use of open-source CT image datasets is also discussed. The final part of this thesis is a model evaluation on a test dataset and results examination.

Keywords

Convolutional neural networks, YOLO, Faster-RCNN, bounding box, object detection, intracranial hemorrhage

Prohlášení autora o původnosti díla

| | |
|-----------------------------------|-------------------------------------------------------------------------------------|
| Jméno a příjmení studenta: | <i>Kryštof Kopečný</i> |
| VUT ID studenta: | <i>212737</i> |
| Typ práce: | <i>Bakalářská práce</i> |
| Akademický rok: | <i>2021/2022</i> |
| Téma závěrečné práce: | <i>Lokalizace intrakraniálních hemoragií v axiálních řezech CT snímků hlavy</i> |

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné 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 závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne: 6. ledna 2022

podpis autora

Bibliografická citace

KOPEČNÝ, Kryštof. *Lokalizace intrakraniálních hemoragií v axiálních řezech CT snímků hlavy*. Brno, 2022. Dostupné také z: <https://www.vutbr.cz/studenti/zav-prace/detail/142067>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Jakub Nemček.

Poděkování

Děkuji vedoucímu bakalářské práce Ing. J. Nemčekovi za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé práce.

V Brně dne: 27. května 2022

podpis autora

Obsah

| | |
|-----------------------------------------------------------------------|-----------|
| SEZNAM OBRÁZKŮ | 8 |
| SEZNAM TABULEK | 9 |
| ÚVOD | 10 |
| 1. ANATOMIE HLAVY | 11 |
| 1.1 LEBKA | 11 |
| 1.2 MOZEK..... | 12 |
| 1.3 KOMOROVÝ SYSTÉM A LIKVOR | 14 |
| 1.4 OBALY MOZKU..... | 15 |
| 1.5 CÉVNÍ ZÁSOBENÍ MOZKU | 15 |
| 2. KARDIOVASKULÁRNÍ ONEMOCNĚNÍ | 16 |
| 2.1 CEREBROVASKULÁRNÍ ONEMOCNĚNÍ | 16 |
| 3. DIAGNOSTIKA HEMORRHAGÍ | 18 |
| 4. NEURONOVÉ SÍTĚ | 20 |
| 4.1 MODEL NEURONU | 20 |
| 4.2 AKTIVAČNÍ FUNKCE..... | 21 |
| 4.3 UČENÍ NEURONOVÉ SÍTĚ | 23 |
| 5. KONVOLUČNÍ NEURONOVÉ SÍTĚ | 26 |
| 5.1 ARCHITEKTURA | 26 |
| 5.2 DETEKCE OBJEKTŮ POMOCÍ CNN..... | 30 |
| 5.3 YOLO („YOU ONLY LOOK ONCE“) | 30 |
| 5.4 FASTER-RCNN..... | 34 |
| 6. PUBLIKOVANÉ METODY PRO DETEKCI A KLASIFIKACI HEMORAGÍ | 37 |
| 7. DATABÁZE CQ500 | 39 |
| 7.1 BRAIN HEMORRHAGE EXTENDED (BHX)..... | 39 |
| 8. PŘEDZPRACOVÁNÍ DAT A TRÉNOVÁNÍ DETEKTORŮ | 40 |
| 8.1 OBECNÝ NÁVRH..... | 40 |
| 8.2 DATABÁZE A PŘEDZPRACOVÁNÍ DAT | 40 |
| 8.3 TRÉNOVÁNÍ DETEKTORŮ..... | 44 |
| 9. DETEKCE A VYHODNOCENÍ ÚSPĚŠNOSTI | 51 |
| 9.1 METRIKY PRO VYHODNOCENÍ VÝSLEDKŮ..... | 51 |
| 9.2 REALIZACE DETEKCE A DOSAŽENÉ VÝSLEDKY | 53 |
| 9.3 DISKUSE DOSAŽENÝCH VÝSLEDKŮ | 56 |
| 9.4 MOŽNÉ VYUŽITÍ DETEKTORU V LÉKAŘSKÉ DIAGNOSTICE..... | 61 |
| ZÁVĚR | 62 |
| LITERATURA | 63 |
| PŘÍLOHA A - OBSAH PŘILOŽENÝCH SOUBORŮ | 71 |

SEZNAM OBRÁZKŮ

| | |
|----------------------------------------------------------------------------------------------|----|
| Obrázek 3.1: CT snímky jednotlivých typů krvácení | 19 |
| Obrázek 4.1: Model umělého neuronu | 20 |
| Obrázek 4.2: Grafy aktivačních funkcí. | 22 |
| Obrázek 4.3: CT snímky jednotlivých typů krvácení. | 22 |
| Obrázek 5.1: Konvoluce vstupní matice | 27 |
| Obrázek 5.2: „Max-pooling“ | 28 |
| Obrázek 5.3: Detektor YOLO | 31 |
| Obrázek 5.4: Schéma architektury YOLOv5 | 33 |
| Obrázek 5.5: Klouzání okna po mapě příznaků. | 35 |
| Obrázek 5.6: Architektura Faster-RCNN..... | 36 |
| Obrázek 8.1: Předzpracování CT snímku. | 42 |
| Obrázek 8.2: Snímky po předzpracování | 43 |
| Obrázek 8.3: Bodový graf velikostí ohraničujících oblastí krvácení v trénovací databázi..... | 46 |
| Obrázek 8.4: Transformační funkce..... | 47 |
| Obrázek 8.5: Vývoj chyby kritériální funkce..... | 48 |
| Obrázek 8.6: Transformace mozaikováním | 49 |
| Obrázek 8.7: Hodnoty chybových funkcí během učení. | 50 |
| Obrázek 9.1: IOU | 51 |
| Obrázek 9.2: Snímky se skutečnými ohraničujícími oblastmi a predikcemi | 53 |
| Obrázek 9.3: Detekce po postprocessingu. | 54 |
| Obrázek 9.4: Snímky se subdurálním krvácením..... | 57 |
| Obrázek 9.5: Bodový graf středových souřadnic | 57 |
| Obrázek 9.6: Porovnání predikcí obou detektorů..... | 59 |
| Obrázek 9.7: Porovnání predikcí obou detektorů 2..... | 60 |

SEZNAM TABULEK

| | |
|---------------------------------------------------------------------------------------|----|
| Tabulka 8.1: Počty ohraničujících oblastí a CT skenů jednotlivých typů krvácení | 40 |
| Tabulka 8.2: Počty ohraničujících oblastí napříč částmi databáze..... | 41 |
| Tabulka 9.1: Tabulka výsledků detektoru Faster-RCNN..... | 54 |
| Tabulka 9.2: Tabulka výsledků detektoru YOLOv5 | 55 |

ÚVOD

Intrakraniální hemoragie je závažný stav, při kterém dochází ke krvácení uvnitř lebky. Důsledkem mohou být celoživotní komplikace nebo v krajních případech smrt. Krvácení se může objevit spontánně či jeho příčinou mohou být traumata hlavy. Celkově se rozlišuje pět typů: epidurální (EDH), subdurální (SDH), subarachnoidální (SAH), intraparenchymální (IPH) a nitrokomorové (IVH). Všechny typy vyžadují neodkladnou zdravotní péči, a především rychlou diagnózu, na které je závislá rychle se měnící prognóza pacienta [1].

Modalitou první volby pro diagnózu intrakraniálního krvácení je počítačová tomografie (CT). Ačkoli pořízení snímku touto metodou je rychlé a poměrně snadné, velké nároky se zde kladou na zdravotnický personál, který se při hledání krvácení potýká s jeho vysokou variabilitou, a tudíž náročnou identifikací. Jisté východisko představují podpůrné diagnostické systémy („Computer aided diagnosis“). Jedná se o technologie, které kombinují poznatky z umělé inteligence a počítačového vidění se zpracováním medicínského obrazu. Součástí počítačového vidění je zejména detekce objektů, kde se v posledních asi 10 letech uplatňují hlavně techniky založené na konvolučních neuronových sítích. Ty dosahují dobrých výsledků a umožňují rozpoznávat objekty prakticky v reálném čase. Svě uplatnění postupně nachází na poli medicíny, zatím však pouze jako pomocný prvek pro profesionály, jejichž úsudek je stále definitivní [2].

Práce se zabývá využitím strojového učení, konkrétně konvolučních neuronových sítí, k detekci intrakraniálního krvácení na axiálních snímcích CT. Nachází se zde základní anatomické, patologické a radiologické poznatky. Popsané jsou zde principy neuronových sítí společně s publikovanými metodami pro detekci objektů. Je zde vypracovaná rešerše prací zabývajících se detekcí intrakraniálního krvácení s využitím konvolučních neuronových sítí. Popsané jsou veřejně dostupné databáze CT snímků hlavy společně s jejich anotacemi.

Praktická část práce vychází z provedené rešerše a jejím cílem je navrhnout detektor schopný lokalizovat a klasifikovat jednotlivé typy krvácení na snímcích CT. Celkem byli na databázích CT snímků natrénovány dva detektory, u kterých byla následně podle vhodných metrik vyhodnocena úspěšnost na testovací databázi. Součástí praktické části je rozprava o předzpracování dat a jejich postprocessingu. Na závěr práce se diskutuje využití detektorů na poli medicíny.

1. ANATOMIE HLAVY

Hlava jakožto část lidského těla je místem průběhu významných lidských funkcí a procesů, což z ní dělá co do struktury poměrně komplikovaný objekt. Je tak třeba se zaměřit na anatomická fakta, která společně s patologií a radiologií (popsaných v následujících kapitolách) tvoří základ pro budoucí práci.

1.1 Lebka

Z funkčního hlediska má lebka úlohu pevné a zároveň pružné schránky pro mozek a smyslové orgány. Jednotlivé kosti jsou mezi sebou pevně spojeny švy, dají se od sebe oddělit jen těžko. Výjimku tvoří pouze dolní čelist, která je se sousedními kostmi spojena kloubně, a jazyka, které je zavěšena pomocí systému šlach a vaziv. Kosti skládající lebku jsou velmi variabilní a rozmanité, vyskytují se zde jak ploché a pneumatické, tak i nepravidelného tvaru. Ploché kosti pak zejména tvoří objemné dutiny a jámy pro mozek se svými obaly a smyslové orgány [3][4][5],

Lebku celkově tvoří 22 kostí, které se řadí do dvou oddílů. Prvním oddílem je obličejová část lebky (*splanchnocranium*), která vytváří kostěný podklad pro obličej. Druhým oddílem je pak mozková část lebky (*neurocranium*), ta vytváří pouzdro kolem mozku a smyslových orgánů.

Ke *splanchnocraniu* se řadí horní čelist (*maxilla*), kost patrová (*os palatinum*), kost lící (*os zygomaticum*), dolní čelist (*mandibula*), jazyka (*os hyoideum*), a kosti sluchové: kladívko (*malleus*), kovádlínka (*incus*), třmínek (*stapes*). K *neurocraniu* pak patří kost týlní (*os occipitale*), kost klínová (*os sphenoidale*), kost čichová (*os ethmoidale*), kost radličná (*vomer*), kost čelní (*os frontale*), kost spánková (*os temporale*), kost temenní (*os parietale*), kost nosní (*os nasale*), kost slzní (*os lacrimale*), a dolní skořepa nosní (*concha nasalis inferior*) [4].

Na lebku se mimo jiné upíná velké množství příčně pruhovaných svalů, což umožňuje pohyb hlavy jako celku [3].

1.2 Mozek

Mozek je nenahraditelná součást nervového systému, který se společně s endokrinním a imunitním systémem stará o výměnu informací mezi vnitřním a vnějším prostředím mnohobuněčných živočichů. Jeho hlavní funkcí je příjem, analýza a integrace informací, které přišly z vnějšího nebo vnitřního prostředí, a vytvoření odpovídající reakce na změny v tomto prostředí. Mozek je utvořen v kaudokraniálním směru z mozkového kmene, na který z dorzálního směru nasedá mozeček, mezimozku a koncového mozku [6][7].

1.2.1 Mozkový kmen

Mozkový kmen (truncus encephalius) představuje pokračování hřbetní míchy. V lebce leží v zadní jámě lební a je překrytý mozečkem. Tvoří ho tři oddíly. Nejkaudálněji sousedí s hřbetní míchou prodloužená mícha (medulla oblongata). Nad prodlouženou míchou se nachází Varlův most (pons Varoli) a nejkraniálnější úsek mozkového kmene představuje střední mozek (mesencephalon). Jsou zde lokalizovaná centra nutná pro základní vitální funkce. Pro mozkový kmen je typické uspořádání šedé a bílé hmoty mozkové. Šedá hmota se zde vyskytuje ve formě jader, kdyžto bílá odpovídá spojům [6][7].

1.2.2 Mozeček

Mozeček (cerebellum) je umístěn v zadní jámě lební a kryje dorzální stranu prodloužené míchy a Varlova mostu, tvoří tak strop 4. mozkové komory. Od ostatních částí mozku je kryt tvrdou plenou (tentorium cerebelli). Tvoří ho střední část (vermis cerebelli), na kterou jsou z obou stran připojeny dvě hemisféry (hemisphaeria cerebelli). S mozkovým kmenem je mozeček spojen pomocí tří páru svazků bílé hmoty (pedunculli cerebellares inferiores, medii et superiores). Povrch mozečku kryje šedá hmota, mozečková kůra, která je rozčleněná do velkého množství folií. Bílá hmota se nachází hlouběji pod kůrou.

Hlavní úlohou mozečku je zpracování informace z vestibulárního aparátu, držení rovnováhy a vzpřímené polohy, regulace svalového tonu [6][7].

1.2.3 Mezimozek

Mezi koncovým mozkem a mozkovým kmenem se nachází mezimozek (diencephalon). Jedná se o dutý útvar uložený mezi hemisférami koncového mozku obklopující III. mozkovou komoru. Člení se na následující struktury: thalamus a metathalamus, epithalamus, subthalamus a hypothalamus.

Thalamus obsahuje množství jader, jejichž funkcí je přivádět a integrovat signály z míchy, mozkového kmene, mozečku a bazálních ganglií. Tyto signály pak následně pošlou do mozkového kmene a do striata. Epithalamus představuje dorzokaudální část mezimozku. Je tvořen hubenulárními jádry, která jsou součástí kmenových okruhů limbického systému, a šišinkou (corpus pineale, epifýza). Ta má díky melatoninu, hormonu, který se zde metabolizuje, vliv na sexuální chování a spánek. Další částí mezimozku je hypothalamus. Malá, ale funkčně nenahraditelná část mezimozku se nachází na pod thalamem a představuje soubor jader, která skrz tzv. hypothalamo-hypofyzární systém podílejí na řízení celé řady funkcí (např. tělesný růst, činnost štítné žlázy, produkce mléka v období laktace). Mimo jiné je hypothalamus také považován za pacemaker cirkadiálních rytmů. Poslední částí mezimozku je subthalamus. Jedná se o několik svazků vláken spojující bazální ganglia s thalamem [6][7].

1.2.4 Koncový mozek

Koncový mozek představuje nejrostrálnější část mozku. Tvoří ho dvě hemisféry, které jsou navzájem odděleny zářezem, do kterého vybíhá tvrdá plena mozková. Hluběji jsou obě hemisféry spojeny mohutným svazkem vláken (corpus callosum). Po povrchu obou hemisfér se táhnou hluboké rýhy (sulci), které ohraničují jednotlivé laloky (lobi). Celkově rozeznáváme 5 mozkových laloků: čelní, temenní, týlní, spánkový a insulární. Menší brázdy pak na každém z laloků oddělují samostatné mozkové závity (gyri).

Povrch koncového mozku kryje šedá mozková kůra. Jedná se o složitou tkáň o několika vrstvách mezi sebou propojených neuronů. Většina kůry je ukryta uvnitř brázd mezi jednotlivými závity, zhruba jedna třetina připadá na volný povrch gyrů. Vymezuje se zde tzv. funkční korové oblasti. Ty jsou buď motorické nebo senzorické, citlivé. Motorické řídí motorickou aktivitu organismu, senzorické a citlivé přijímají dostředivou informaci z těla. Pod povrchem hemisfér se nachází bílá hmota (corpus medullare). Je složena s myelinizovaných a nemyelinizovaných vláken, glií, cévní a kapilární sítě. Vlákná propojují mozkovou kůru s prakticky každým útvarem mozku, spojují korové okruhy téže hemisféry, nebo spojují kůru pravé a levé hemisféry, a tím zajišťují jejich součinnost.

Hluboko uvnitř bílé hmoty se nachází bazální ganglia. Tyto vývojově staré útvary šedé hmoty se uplatňují při koordinaci reflexní a úmyslné pohybové aktivity [6][7].

1.3 Komorový systém a likvor

Komorový systém centrálního nervového systému tvoří 4 dutiny vyplněné likvorem (vnitřní likvorový systém). Tato tekutina se mimo komory nachází též v subarachnoidálním prostoru, který plášťovitě obklopuje CNS (zevní likvorový systém). Oba tyto systémy mezi sebou komunikují skrz IV. mozkovou komoru.

Čtvrtá komora mozková představuje rozšíření úzkého kanálku, který prostupuje šedou hmotou míchy po celé její délce. Nachází se mezi mozkovým kmenem a mozečkem. Ve stropu komory je vytvořena tenká lamela (tela choroidea), na kterou se zvenku přikládají piální cévy, které se zde bohatě větví a vytváří hustou síť (plexus choroideus), jejich funkcí je produkce mozkomíšního moku. Komoru opouští mozkomíšní mok do subarachnoidálního prostoru třemi otvory, jsou to párové laterální foramina Luschkea a nepárový mediální foramen Magendi. Kaudálně se komora zužuje a přechází do míšního kanálku, kraniálně pak přes *aquaeductus mesencephali* komunikuje s III. mozkovou komorou. Třetí komora mozková je úzká štěrbiná zasazená mezi pravý a levý thalamus. Na stěně se podobně jako u IV. komory vyskytuje síť cév pro produkci likvoru. Směrem ventrokranálně se nachází dva otvory (foramina ventricularia), z nichž každý vede do jedné z postranních komor. I. a II. mozková komora (postranní komory, *ventriculi laterales*) jsou zrcadlově identické dutiny zasazené v hemisférách koncového mozku. Mají podkovitý tvar a ze všech již uvedených komor jsou to ty nejobjemnější. Na vnitřních stěnách v částech komor, které se nachází ve spánkovém laloku, se nachází plexus choroideus.

Likvor je průzračná žlutá tekutina, ve které se nachází malé množství glukózy, proteinů, solí a lymfocytů. Objem likvoru je asi 150 ml, z toho asi pětina se nachází v mozkových komorách, zbytek připadá na subarachnoidální prostor. Denně plexus choroideus vyprodukuje likvoru asi 500 ml, takže se každý den celkový objem cca třikrát obmění. Jak již bylo zmíněno, všechen likvor vytvořený v mozkových komorách odtéká přes IV. komoru do subarachnoidálního prostoru, který obklopuje mozkou a míchu. Z oblasti báze mozku se likvor dostává směrem nahoru a obtéká stěny hemisfér. Vstřebává se do žilních splavů v řasách tvrdé pleny mozkové. Likvor má funkci mechanické ochrany měkké pulsující mozkové tkáně. Komunikuje také s extracelulárním prostředím mozkové tkáně, odvádí odtud metabolity, které by mohly ohrozit správnou činnost mozkových buněk [3][6][7].

1.4 Obaly mozku

Celkově rozlišujeme 3 vazivové obaly. Nejbližší povrchu se nachází tvrdá plena (dura mater). Je to tuhá vazivová membrána pevně srůstající se švy kalvy, bází lebni a okraji otvoru pro míchu. Skládá se ze dvou vrstev, které se na několika místech od sebe oddělují a vytváří tak prostory pro žilní splavy. Tvrdá plena je vyživována meningeálními tepnami, žilní odtok je zajištěn stejnojmennými žilami, které ústí do žilních splavů.

Hluběji pod tvrdou plenu se nachází pavučnice (arachnoidea). Tenká, měkká a na rozdíl od tvrdé pleny bezcévná blána přiléhá těsně na dura mater. Pavučnice obaluje mozek volně a nevniká do hlubokých brázd. Společně s omozečnicí (nejhlubší z mozkových obalů) vytváří subarachnoidální prostor, který je vyplněný mozkomíšním mokem. V oblasti žilních splavů se vyklenují výběžky subarachnoidálního prostoru skrz jejich stěnu a dochází zde k vstřebávání likvoru zpět do žilní krve.

Nejhlubší blána pia mater (omozčnice) proniká do nejhlubších zářezů mozku a přesně sleduje jeho povrch, ke kterému těsně přiléhá. Obaluje a doprovází všechny cévy, které vstupují a vyživují mozkovou tkáň. S pavučnicí je spojena vazivovými trámečky [3][6][7].

1.5 Cévní zásobení mozku

Pro správnou činnost mozku, optimální řízení organismu a udržení samotného vědomí je průtok krve mozkem nezbytný. Lidský mozek představuje sice 2 % hmotnosti celého těla, ale spotřebuje až 20 % objemu kyslíku určeného pro celý organismus. Příjem látek z krve je korigován hematoencefalickou bariérou, která zajišťuje, že prochází pouze kyslík, oxid uhličitý a voda, pro aminokyseliny a glukózu existují přenašeče. Při stavech, kdy se sníží přívod arteriální krve, dochází k ischemickým změnám nebo dokonce k odúmrťi dané oblasti.

Hlavními tepnami, které přivádí krev do mozku jsou a. vertebralis a a. carotis interna. Těmito tepnami dohromady každou minutu přiteče asi 1 litr krve, z toho asi 700 ml připadá na karotidy. Společně ještě s dalšími arteriemi vytváří na bázi mozku circulus arteriosus cerebri (Willisi), jeho úkolem je vyrovnat tlakové rozdíly a průtok krve v obou řečištích. Z tohoto okruhu vystupují tři typy tepen. Tepny korové jsou tři, arteria cerebri anterior, media et posterior. Míří směrem k povrchu mozku a v pia mater vyživují mozkovou kůru a bílou hmotu. Tepny centrální odstupují z počátečních úseků korových tepen a zanořují se do tkáň bezprostředně u Willisova okruhu. Poslední skupinou jsou tepny zásobující plexus choroideus v mozkových komorách.

Mozkové žíly odvádějící odkysličenou krev rozdělujeme na povrchové a hluboké. Povrchové vycházejí z mozkové kůry a bílé hmoty jako malé žilní větvičky, vytváří piální pleteně, ze kterých krev mozkovými žilami putuje pryč. Hluboké žíly jsou situované kolem horní stěny thalamu, kde sbírají krev z okolních struktur jako například thalamu a plexus choroideus postranních komor. Žilní splavy (sinus durae matris) jsou prostory mezi řasami tvrdé pleny, které sbírají krev z hlubokých a povrchových žil [6][7].

2. KARDIOVASKULÁRNÍ ONEMOCNĚNÍ

Adekvátní přísun krve je nezbytný pro každý orgán a tkáň v lidském těle. Hlavní rolí kardiovaskulárního systému je kontinuálně přivádět okysličenou a odvádět odkysličenou krev do, respektive z tkáně. Nemoci kardiovaskulárního systému jsou jednou z nejběžnějších příčin úmrtí v západním světě a většina se zakládá na jednom ze dvou základních problémů.

Lokální selhání cirkulace je způsobeno zúžením (stenóza) nebo kompletním ucpaním (okluze) krevní cévy. Stav, kdy buňky nejsou zásobovány dostatečným množstvím krve nebo krev není odváděná, je známý jako ischemie. Šok vlivem ztráty nadměrného množství krve či selhání srdce jako pumpy je druhou kategorií pro většinu kardiovaskulárních onemocnění, která vede k poškození orgánů způsobeným nedostačující perfuzí.

Mezi nejčastější pojmy kardiovaskulárních onemocnění patří trombóza a infarkt. Trombóza je častou příčinou okluze krevní cévy. Za normálních okolností homeostatický systém tvoří krevní sraženiny pouze v místech poranění cévy. Může však nastat situace, kdy nastane nerovnováha mezi hemokoagulací a fibrinolýzou. Masa krevních destiček, fibrinu a zachycených červených krvinek formují trombus, který může kompletně ucpat cévu. Trombus se však může od cévní stěny urvat a putovat krevním řečištěm jako embolus, který může nyní blokovat jiné krevní cévy po celém těle. Pokud je ischemie dlouhodobá, dochází k nekróze buněk, tento proces se nazývá infarkt [8].

2.1 Cerebrovaskulární onemocnění

Cerebrovaskulární onemocnění je název pro skupinu onemocnění, při kterých jsou okrsky mozku dočasně či trvale postihnuty ischemií či krvácením jedné nebo více mozkových cév. Základní rozdělení těchto onemocnění je na ischemické příhody a krvácení, pro obě skupiny pak existuje celá škála příčin vzniku. U ischemických příhod jsou to aterosklerózy, embolie, zvýšená srážlivost nebo viskozita krve či disekce karotidy. Nejčastějšími příčinami krvácení pak je zvýšený tlak, aneurysma, malformace stěn cév nebo komplikace při užívání antikoagulačních léků. Ischemické příhody jsou pak dále členěny na podskupiny podle místa, kde se symptomy objevily, a době jejich trvání. V případě krvácení mluvíme o nitrolebním krvácení, které se dále dělí na epidurální, subdurální, subarachnoidální, intracerebrální a nitrokomorové [9][10].

Epidurální krvácení bývá zpravidla traumatického původu, prochází do prostoru mezi lebkou a tvrdou plenu. Hematom roste pomalu a v průběhu několika dnů začíná utlačovat mozek. Subdurální krvácení vniká mezi tvrdou plenu a arachnoideu, také bývá nejčastěji traumatického původu. Nejčastěji se lokalizuje na konvexitách hemisfér ve frontoparietální oblasti. Subarachnoidální krvácení směřuje do subarachnoidálního prostoru. Nejčastější příčinou bývá ruptura aneurysmatu či cévní malformace. Dochází k rozvoji intrakraniální hypertenze a dráždění bazálních nervových struktur a cévním spasmům. Symptomy jsou bolest hlavy, zvracení, zmatenost a ztráta vědomí. Můžou se také vyskytovat srdeční arytmie. Intracerebrální krvácení je masivní, vzniká náhle a následky jsou vážné, směřuje přímo do mozkové tkáně. Příčinou číslo jedna je ruptura cévy při arteriální hypertenzi. Za krvácením můžou ale také stát aneurysmata, tumory, traumata, abscesy či hemangiomy. Ložisko krvácení se skládá z krve a zbytků tkáně, bývá nepravidelného tvaru. Okolní tkáň je často stlačená, dislokovaná a prokrvácená. Často bývá intracerebrální krvácení doprovázeno nitrolební hypertenzí a edémem. Krvácení se může také dostat do mozkových komor, pak se jedná o krvácení nitrokomorové. Tato komplikace nastává asi u 25–45 % případů. Asi polovina postižených přežívá [11].

3. DIAGNOSTIKA HEMORRHAGIÍ

Nitrolební krvácení může být život ohrožující stav. Zdraví pacienta je pak velmi poutané na rychlou a přesnou diagnózu, která je ale kvůli vysoké vzhledové variabilitě krve na snímcích stále náročným úkolem. Počítačová tomografie je v této oblasti díky své rychlosti vyšetření a dostupnosti diagnostickou metodou číslo jedna [12][13].

Objevují se však statistiky, ze kterých vyplývá, že magnetická rezonance je minimálně stejně senzitivní na odhalení intrakraniálního krvácení v akutních případech, a dokonce úspěšnější v subakutních a chronických případech ve stejné činnosti. Kromě toho nové metody jako difuzní tenzorové zobrazování nebo magnetická rezonanční spektroskopie mohou přinést v budoucnosti nový pohled na možnosti zobrazení tohoto závažného onemocnění [12][13].

3.1.1 CT zobrazování mozku

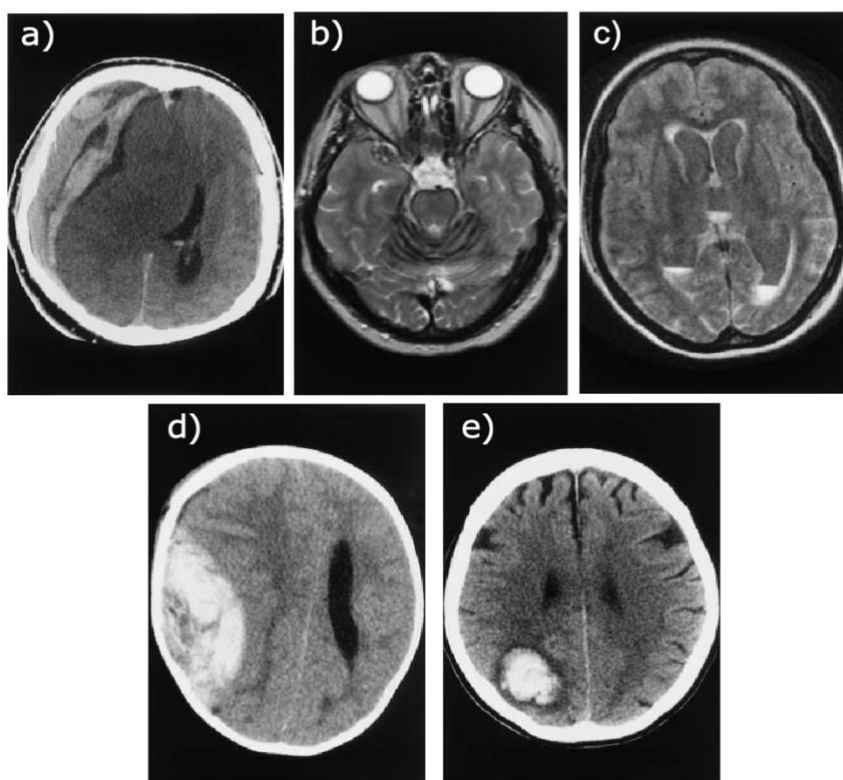
Při vyšetření se pracuje s tzv. radiologickými okny. Jedná se o transformace kontrastu ke zvýraznění určitých struktur k lepší interpretaci snímku. Každé radiologické okno popisujeme dvěma parametry, jsou to šířka okna a střed [14]. Používá se okno pro mozkovou tkáň o šířce 80 HU (Hounsfieldova jednotka) a středu 40 HU ke kontrole tvarů mozkových závitů a kontrole útlumu záření. Zvláštní pozornost si zaslouží subarachnoidální prostor a mozkové komory. Pozoruje se nejen tvar a velikost, ale i zda se v těchto prostorech nenachází krev. Dalším oknem je subdurální okno o šířce v rozmezí 130–300 HU a středem mezi 50 a 100 HU. Používá se ke kontrole subdurálních nebo extradurálních oblastí. Ke zvýraznění kostní tkáně se využije okno o šířce 2800 HU se středem v 600 HU. Sleduje se, zda nedošlo k traumatu, který by mohl být příčinou krvácení [15].

Zobrazit patologie můžeme buď s kontrastní látkou nebo bez ní. Ve většině případů se provádí bezkontrastní CT, ale existují i indikace pro případy, kdy je použití kontrastní látky výhodné [16].

3.1.2 CT zobrazování intrakraniálních krvácení

Při dopadu záření na tkáň, dochází k jeho útlumu vlivem interakcí rentgenových fotonů s elektrony v obalu atomu tkáně. Útlum je tím větší, čím je hustota elektronů v obalu atomu větší. V kontextu intrakraniálního krvácení jsou významnými strukturami, na kterých dochází k útlumu, agregované molekuly globinu v hematomu. Existuje tedy vztah mezi útlumem rentgenového záření, proteinovou stavbou dané tkáně (především hemoglobinu) a hematokritem [17][18].

Bezprostředně po krvácení se v oblasti okolo zasažené cévy formuje masa obsahující červené a bílé krvinky spolu s krevními destičkami obklopená krevním sérem. Tuto formaci lze na CT snímcích pozorovat jako ložisko o denzitě mezi 40–60 Hounsfieldovými jednotkami, což má za následek poměrně složité rozeznání akutního krvácení od okolní šedé (30–40 HU) a bílé (20–30) hmoty mozkové. V následujících 12 hodinách se denzita hematomu rapidně zvýší až na 80 HU. To je zapříčiněno vznikem krevního trombu, který vycytává denzní globin, a jeho následnou retrakcí. Důsledkem tohoto procesu se hematom první týden po krvácení na snímcích CT jeví jako hyperdenzní léze. V následujících týdnech dochází k proteolýze, proteiny hematomu jsou degradovány a postupně odklizeny, krevní trombus ubývá na denzitě. Tento proces probíhá směrem od periferie do centra hematomu. Po týdnech nebo dokonce měsících dochází k úplnému vstřebání krevního trombu. Jediný důkaz, co zůstane po už zhojeném krvácení je jizva s mírně hyperdenzními okraji. Pro zobrazení epidurálního a subdurálního krvácení se využívá subdurální okno, naopak intraparenchymální krvácení je detekováno nejlépe pomocí okna pro mozkovou tkáň [17][18].



Obrázek 3.1: CT snímky jednotlivých typů krvácení. a) SDH; b) SAH (vermis mozečku); c) IVH (temporální výběžek 3. a 4. mozkové komory); d) EDH; e) IPH [18].

4. NEURONOVÉ SÍTĚ

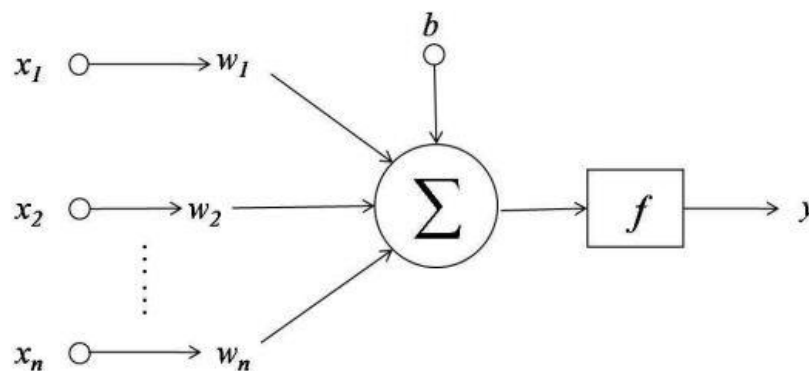
Neuronové sítě jsou výpočetní systémy oblasti umělé inteligence, které se do jisté míry inspirojí biologickými neuronovými sítěmi v mozku živých organismů. Nejdůležitější vlastností neuronových sítí je schopnost se učit. Činí tak adaptací své struktury takovým způsobem, aby vyhovovala zadanému problému. Aplikace neuronových sítí je rozsáhlá, dají se použít při řešení klasifikačních a predikčních úkolů [19][20].

4.1 Model neuronu

Neurony představují základní výpočetní a stavební jednotku neuronových sítí. Každý neuron je definován vektorem vah $\mathbf{w} = [w_1, w_2, \dots, w_n]$ a prahem b (bias). Na neuron je přiveden vektor vstupních hodnot $\mathbf{x} = [x_1, x_2, \dots, x_n]$ z předešlé vrstvy sítě. Následně se všechny součiny vstupních hodnot a korespondujících vah sumují a jsou předány aktivační funkci. Výstupem každého neuronu je pak pouze jedna hodnota, která směřuje do následující vrstvy neuronové sítě. Pokud se jedná o neuron poslední vrstvy, odpovídá výstup neuronu výstupu celé sítě. Neuron lze popsat následující rovnicí:

$$y = f\left(\sum_{i=1}^N w_i x_i - \theta\right), \quad (4.1)$$

kde člen $\sum_{i=1}^N w_i x_i - \theta$ je tzv. aktivace a a $f(a)$ je aktivační funkce [20].



Obrázek 4.1: Model umělého neuronu se vstupním vektorem hodnot, kde jsou jednotlivé vstupy vynásobeny korespondujícími váhami. K sumě těchto hodnot je přičten bias b a tato hodnota je dále předána aktivační funkci f . Výstupem neuronu je jediná hodnota y .

4.2 Aktivační funkce

Aktivační funkce jsou v neuronových sítích využívány za účelem transformace vstupního signálu na signál výstupní. Uplatňují se v případech, kdy je potřeba mezi vstup a výstup zanechat nelineární vztah. Existuje několik druhů aktivačních funkcí.

Prahová funkce zavádí limitní hodnotu pro aktivaci daného neuronu. Je to nejjednodušší aktivační funkce. Svě uplatnění nachází při binární klasifikaci. Je definována jako:

$$f(x) = \begin{cases} x & \text{pro } x \geq t \\ 0 & \text{pro } x < t \end{cases} \quad (4.2)$$

kde x je vstupní hodnota prahové funkce, a t je prahová hodnota pro aktivaci („threshold“). Dalším druhem aktivační funkce je sigmoidní. Je to nejvíce se využívající nelineární funkce, která transformuje vstupní hodnoty na výstupní v intervalu (0, 1). Její rovnici lze zapsat jako:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4.3)$$

kde x značí vstupní hodnotu prahové funkce a e je Eulerovo číslo. Podobný sigmoidní funkci je hyperbolický tangens. Ten je na rozdíl od sigmoidu symetrický kolem počátku a jeho výstup je v intervalu (-1, 1):

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (4.4)$$

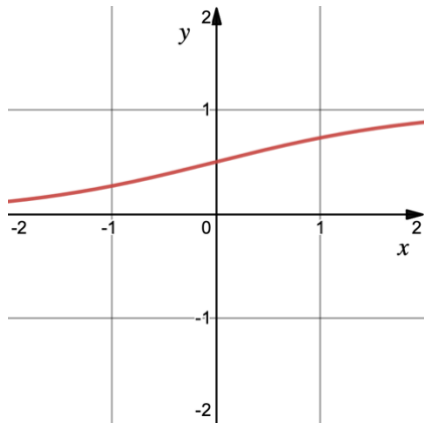
Problémem při používání posledních dvou zmíněných aktivačních funkcí může být jev zvaný „vanishing gradient“, který zabraňuje se neuronové síti efektivně učit. V poslední době je možné se setkat s ReLU („Rectified linear unit“). Tato aktivační funkce nepodléhá saturaci jako sigmoid a hyperbolický tangens a zároveň zmenšuje výpočetní náročnost sítě. ReLu je definováno následujícím vztahem:

$$f(x) = \begin{cases} 1 & \text{pro } x \geq 0 \\ 0 & \text{pro } x < 0 \end{cases} \quad (4.5)$$

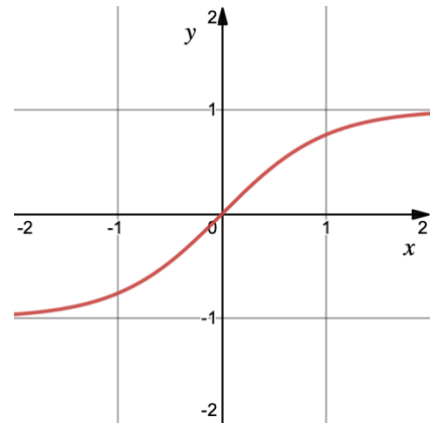
Další aktivační funkce, která se používá hlavně pro řešení klasifikačních problémů do více tříd je „Softmax“. Umísťuje se hlavně do posledních vrstev sítí. Jejím výstupem je vektor pravděpodobností, se kterými vstup patří do daných tříd. Všechny pravděpodobnosti ve vektoru mají součet roven 1. „Softmax“ lze zapsat následující rovnicí:

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad (4.6)$$

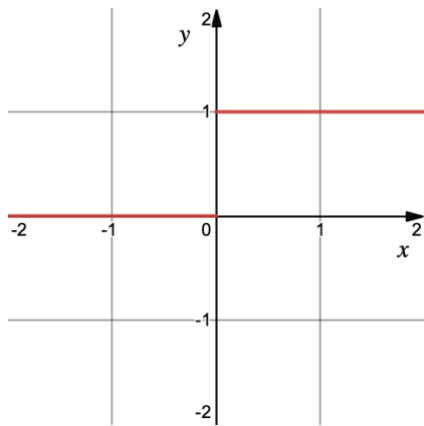
pro $j = 1, \dots, K$, kde „Softmax“ normalizuje K -rozměrný vstupní vektor \mathbf{z} do K -rozměrného výstupního vektoru $\sigma(\mathbf{z})$ [21][22].



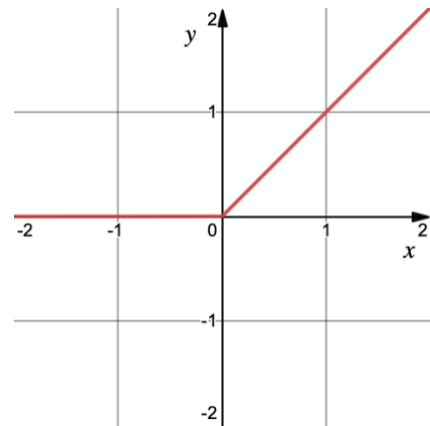
a)



b)



c)



d)

Obrázek 4.2: Grafy aktivačních funkcí. a) Sigmoidní; b) Hyperbolický tangens; c) Prahová funkce; d) ReLU.

4.3 Učení neuronové sítě

Jakmile je navržena architektura neuronové sítě, nastává proces učení, aby síť poskytovala požadované výsledky. Na počátku učení můžou být parametry neuronů nastaveny náhodně, nebo mohou být využity jiné inicializační metody (viz kapitola 2.3.2). Existují dva přístupy učení neuronových sítí: s učitelem a bez učitele, přičemž drtivá většina je prvního typu.

Při učení s učitelem poskytujeme síti vedle vstupů i požadované výstupy a hodnotíme úspěšnost sítě. Je zde používán algoritmus, který pak systematicky modifikuje parametry jednotlivých neuronů takovým způsobem, aby se výstupy sítě a požadované výstupy lišily co nejméně.

Při učení bez učitele neposkytujeme neuronové síti požadované výstupy. Síť musí sama rozpoznat vzory v datech. Typickým úkolem je zde shlukování [23].

4.3.1 Inicializace vah

Před samotným učením neuronové sítě je potřeba nějakým způsobem nastavit parametry neuronů a zavést tak výchozí pozici jejich optimalizace. Od této startovací pozice se pak odvíjí finální stav parametrů neuronové sítě po ukončení učení. Pokud jsou parametry dobře přednastaveny, může to pozitivně ovlivnit rychlost učení a konečnou přesnost sítě. Jedním z možných postupů je tzv. „fine-tuning“. Jako výchozí pozice parametrů před samotným učením slouží hodnoty již z natrénované sítě. Tato metoda vychází z pozorování, že první vrstvy neuronových sítí většinou rozpoznávají obecné prvky jako jsou například hrany a barevné shluky, zatímco konečné vrstvy stojí za detekcí specifických detailů pro daný dataset. Běžně se dá setkat s postupem, kdy se váhy neuronů ve finální/finálních vrstvách/vrstvách vynulují a optimalizují se pak až na cílovém datasetu. Jako předtrénované sítě se dnes používají publikované architektury především konvolučních neuronových sítích naučených na rozsáhlých datasetech se statistickými obrazy jako jsou například COCO dataset či ImageNet. Lze se také setkat s postupem, že při „fine-tuningu“ se vyberou z již natrénované architektury pouze některé vrstvy a zbylé se na cílovém datasetu dotrénují z počátečních náhodných hodnot [24].

4.3.2 Chybové funkce

Chybová funkce („loss function“) se během trénování využívá ke zjištění, jak se liší výstup neuronové sítě od skutečné očekávané hodnoty. Požadujeme tedy, aby se během učení hodnota chybové funkce blížila co nejbližší nulové hodnotě.

Jednou z chybových funkcí je střední kvadratická chyba (MSE, „Mean squared error“). MSE je definována jako:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2, \quad (4.7)$$

kde \hat{Y}_i je vektor predikcí, Y_i je vektor skutečných požadovaných výsledků a n je velikost těchto vektorů.

V dnešní době se dá také při řešení problémů setkat s vzájemnou entropií (CE, „Cross-entropy“). Ta se používá v případech klasifikace, kdy se třídy vzájemně nevylučují. CE se zapisuje vzorcem:

$$CE = - \sum_{i=1}^N y_i \cdot \log(\hat{y}_i), \quad (4.8)$$

kde y_i je požadované rozdělení pravděpodobnosti, \hat{y}_i je predikované rozdělení pravděpodobnosti a N je počet klasifikačních tříd [25][26].

4.3.3 Zpětné šíření chyby

Algoritmus zpětného šíření chyby („Backpropagation“) hraje klíčovou roli při samotném učení sítě. Jeho pomocí lze vypočítat, jak se jednotlivé neurony podílejí na celkovém chybování neuronové sítě. Jinými optimalizačními algoritmy (viz kapitola 2.3.4) pak upravíme parametry těchto neuronů tak, aby neuronová síť poskytovala co nejpřesnější predikce.

Samotný algoritmus se skládá ze tří etap. V první dopředné („feedforward“) etapě se šíří signál sítě na výstup. V druhé etapě je na výstupu pomocí chybové funkce vypočtena chyba mezi hodnotou „ground-truth“ a predikcí sítě. Tato chyba je pak následně šířena v opačném směru a v poslední fázi jsou aktualizovány parametry neuronů.

Rovnici pro velikost chyby ve výstupní vrstvě lze zapsat ve tvaru:

$$\delta^L = \nabla_{y_L} \mathbf{E} \odot f'(\mathbf{a}^L) \quad (4.9)$$

kde člen $\nabla_{y_L} \mathbf{E}$ je gradient chybové funkce podle výstupů neuronů y_L ve výstupní vrstvě L a člen $f'(\mathbf{a}^L)$ je derivace aktivační funkce f podle aktivace a . Symbol \odot značí násobení prvek po prvku. Chyba se následně zpětně šíří do další l -té vrstvy. Můžeme ji zapsat následující rovnicí:

$$\delta^l = ((\mathbf{w}^{l+1})^T \delta^{l+1}) \odot f'(\mathbf{a}^l), \quad (4.10)$$

kde \mathbf{w} je vektor vah neuronů ve vrstvě $l + 1$. Z následující rovnice pak je možné získat gradient chybové funkce ve vrstvě l podle vektoru vah \mathbf{w} :

$$\nabla_{\mathbf{w}_l} \mathbf{E} = y^{l-1} \delta^l \quad (4.11)$$

[27][28].

4.3.4 Optimalizační algoritmy

Optimalizační algoritmy stojí za úpravou parametrů jednotlivých neuronů během učení. Jedním z nejpoužívanějších algoritmů je tzv. „Gradient descend“. Ten využívá k úpravě parametrů neuronu informaci o gradientu chybové funkce. Algoritmus iterativně optimalizuje parametry proti směru gradientu chybové funkce. Změna parametrů se vyjádří následující rovnicí:

$$\mathbf{w}^{t+1} = \mathbf{w}^t - \eta \nabla_{\mathbf{w}^t} E(\mathbf{w}^t), \quad (4.12)$$

kde t značí číslo epochy, \mathbf{w} je vektor vah, člen $\nabla E(\mathbf{w}^t)$ představuje gradient chybové funkce a η je krok učení („learning rate“). Krok učení udává, s jakou rychlostí budou upravovány parametry neuronů. Při zvolení příliš velké hodnoty „learning rate“, může model konvergovat příliš rychle k neoptimálnímu výsledku. Zatímco když je tato hodnota příliš malá, může být proces učení zaseknutý v lokálním minimu. Častěji se dá setkat s variantou „Stochastic gradient descend“, kdy dochází k úpravě vah po každé iteraci napříč vzorkem dat z trénovací databáze („minibatch“). Výhodou je snadnější únik z lokálního minima a rychlejší konvergence [29][30].

5. KONVOLUČNÍ NEURONOVÉ SÍTĚ

Konvoluční neuronové sítě (CNN) jsou speciálním typem neuronových sítí popsaných v předešlé kapitole. Stejně jako plně propojené neuronové sítě se skládají z neuronů, které provádějí skalární součin vstupních hodnot s váhami a aplikují aktivační funkci. Sítě dostaly svůj název po matematické operaci – konvoluci, která zde probíhá v tzv. konvolučních vrstvách.

CNN v posledním desetiletí dosáhly výborných výsledků v oblastech počítačového vidění a zpracování přirozeného jazyka. Nejprínosnějším aspektem je oproti klasickým neuronovým sítím redukce parametrů, které popisují jednotlivé vrstvy. To umožnilo navrhnout komplexní architektury k řešení složitějších problémů, které by v dnešní době pomocí plně propojených sítí z výpočetně kapacitních důvodů nebylo možné překonat [31][32][33].

5.1 Architektura

Konvoluční neuronové sítě jsou zpravidla tvořeny třemi druhy vrstev. Jsou to konvoluční, podvzorkovací („pooling“) a plně propojené vrstvy. Z hlediska funkcionality může být architektura CNN rozdělena do čtyř klíčových procesů. Vstupem je na rozdíl od klasických neuronových sítí obraz (nebo jiný signál) reprezentován N -dimenzionální maticí. V druhé fázi je v neuronech konvolučních vrstev vypočítán skalární součin vah se vstupními hodnotami pixelů. Tento výstup je dále upraven aktivační funkcí (u CNN typicky ReLU). Následuje „pooling“ vrstva, která provede podvzorkování. Nakonec plně propojená vrstva v závislosti na řešeném problému buď určí skóre tříd pro klasifikaci, nebo na výstup sítě přivede pouze jedinou hodnotu v případě, že je jedná o regresi. [31][32][33].

5.1.1 Konvoluční vrstva

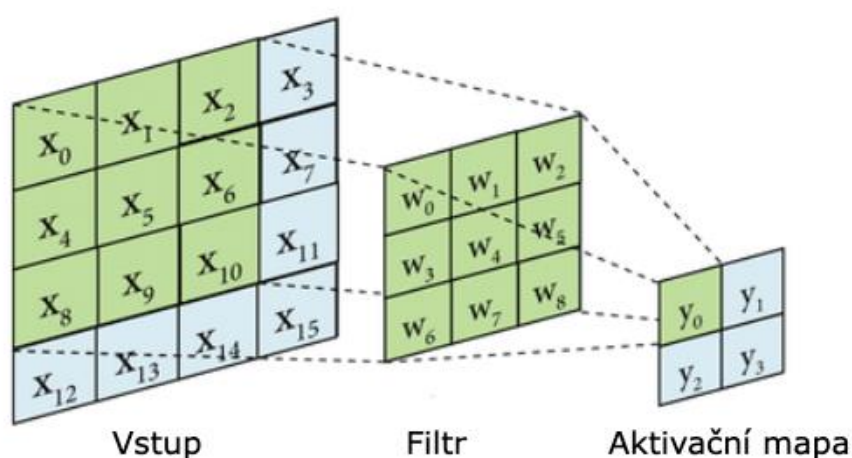
Konvoluční vrstva je základním stavebním kamenem konvolučních sítí. Parametry neuronů jsou zde soustředěny do filtru. Filtr („kernel“, jádro) má obvykle oproti vstupnímu obrazu menší horizontální a vertikální rozměr, avšak prostupuje jeho celou hloubkou (všemi kanály). Každý neuron má tedy $h \times d \times w$ vah, kde h je výška filtru, w je šířka a d je hloubka. Když vstupní data přijdou na konvoluční vrstvu, nastane konvoluce mezi filtrem a lokální oblastí vstupního obrazu o stejné velikosti, se kterou je propojen. Výstupem jednoho neuronu je tedy pouze jedna hodnota, tzv. aktivace.

U filtrů se mimo parametry definuje krok („stride“), ohraničení („padding“) a výstupní hloubka obrazu. Krok udává, o kolik pixelů je posunutý vedlejší filtr propojený s další lokální oblastí. Obvykle nabývá hodnot jedna nebo dva, ale je možné se setkat i s většími hodnotami. Ohraničení znamená doplnění vstupního obrazu nulami. Výstupní hloubka určuje počet k filtrů, které budou sledovat jednu lokální oblast obrazu. Tento tzv. hyperparametr se zavádí z důvodu, že je žádoucí sledovat v dané oblasti několik atributů (např. barva, tvar, hrany apod.) nejdnou. Současně filtry v jedné hloubce sdílejí svoje váhy, což vychází z potřeby sledovat tu samou vlastnost ve všech oblastech obrazu.

Krok a ohraničení mají dopad na výslednou šířku a výšku aktivační mapy, která je výstupem filtrace obrazu neurony se stejnými váhami. Její rozměry se spočítají pomocí následujícího vztahu:

$$x = \frac{(W - F + 2P)}{S + 1}, \quad (5.1)$$

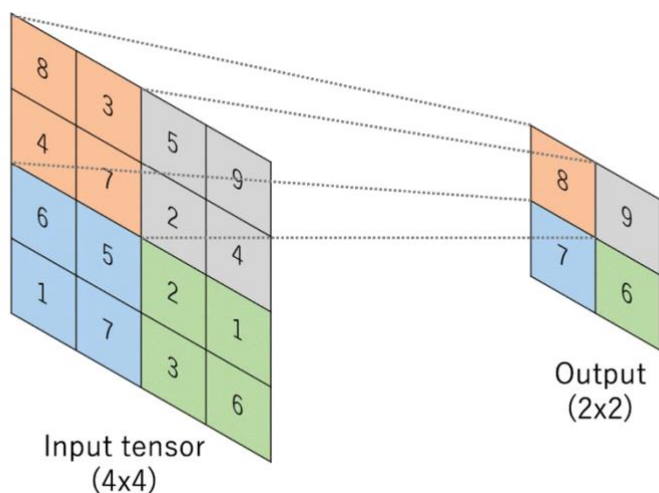
kde W je rozměr vstupního obrazu, F je velikost filtru, P je ohraničení a S je krok. Výstupem jedné konvoluční vrstvy je tedy aktivační mapa o rozměrech popsány vztahem výše a o dané hloubce. Aktivační mapy jsou dále transformovány pomocí aktivačních funkcí [31][32][33].



Obrázek 5.1: Konvoluce vstupní matice s krokem jedna bez ohraničení s filtrem produkuje aktivační mapu [33].

5.1.2 „Pooling“ vrstva

U architektury konvolučních sítí se běžně mezi dvě po sobě následující konvoluční vrstvy zavádí „pooling“ vrstva. Slouží k podvzorkování aktivačních map, tím redukuje počet parametrů filtrů v následujících vrstvách a snižuje výpočetní náročnost celé sítě. Vstupní aktivační mapa se rozdělí na menší oblasti (obvykle o velikosti 2×2 pixely), ze kterých do výstupní matice přijde pouze jedna hodnota. Jedním z nejčastěji používaným typem poolingů je „max-pooling“, kde se z oblasti o definované velikosti vybere maximální hodnota aktivace. V praxi je ale možné se setkat i s „average-pooling“, zde se vybírá průměrná hodnota. Poolingové vrstvy obecně zmenšují rozměry vstupních map a současně zavádějí prostorovou invarianci. Není tedy důležitá poloha aktivace, ale její existence sama o sobě [31][32][33].



Obrázek 5.2: „Max-pooling“ aktivační mapy [32].

5.1.3 Plně propojená vrstva

V plně propojené vrstvě jsou neurony uspořádány stejně jako ve vrstvách standardních neuronových sítí. Každý neuron je tedy propojený se všemi neurony předešlé a následující vrstvy. Před samotným vstupem jsou trojrozměrné aktivační mapy přetransformovány („flattening“) do jednorozměrného vektoru, aby mohly být plně propojenou vrstvou zpracovány.

Je nutné podotknout, že se zavedení plně propojených vrstev do architektury CNN způsobuje jisté komplikace. Obrovské navýšení počtu parametrů neuronů může vést k „overfittingu“, tedy neschopnosti sítě se učit. K tomuto problému se v praxi přistupuje aplikací tzv. „dropout“ vrstvy [31][32][33].

5.1.4 „Dropout“ vrstva

„Dropout“ vrstva se používá při trénování sítě jako prevence proti jejímu „overfittingu“ na cílovém datasetu. Každý neuron skryté vrstvy je s pravděpodobností p vypnut a jeho výstupní hodnota je uměle nastavena na hodnotu nula. Tyto neurony se nepodílejí na finálním výstupu celé neuronové sítě a během procesu zpětného šíření chyby se jejich váhy neoptimalizují. Pokaždé, když je tedy na vstup sítě přiveden vzorek, použije se k získání výstupu jiná architektura, protože jsou deaktivovány odlišné neurony. Tato technika zamezuje komplexní adaptaci skupin neuronů na vstupní data, jelikož se neurony nemohou spolehnout na přítomnost ostatních neuronů v rámci dané skupiny. Všechny neurony jsou tak přinuceny se naučit reagovat na různorodá data. Výsledek každého nevypnutého neuronu během učení kvůli zachování energie je nutné vynásobit hodnotou p . Během testování nebo samotném nasazení pak vypnuté neurony nejsou žádné [34][35]

5.1.5 „BatchNormalization“ vrstva („BatchNorm“)

„Batch“ normalizace je vrstva, která umožňuje rychlejší a efektivnější trénování hlubokých neuronových sítí. Hlavní motivací pro vývoj tohoto mechanismu byla snaha o snížení vlivu „internal covariate shift“ (ICS). Jedná se o změnu vstupních hodnot neuronové vrstvy vlivem optimalizace parametrů vrstvy předešlé. Neurony se tedy musí stále přizpůsobovat na různá rozdělení. To zpomaluje a komplikuje proces učení.

„BatchNorm“ normalizuje hodnoty výstupů, respektive aktivací předešlých vrstev napříč tzv. „mini-batch“. Aktivace mají tedy průměr 0 a standardní odchylku rovné 1. Během procesu učení dochází k optimalizaci parametrů posunu a škálování, které dále upravují normalizovanou hodnotu. „BatchNorm“ vrstva se u konvolučních neuronových sítí umísťuje před vrstvu aktivační. „BatchNorm“ má také vlastnost regularizace. Nevýhodou vrstvy je větší výpočetní náročnost [36][37].

5.2 Detekce objektů pomocí CNN

Obecně se detekcí objektů zabývá vědní disciplína nazvaná počítačové vidění. Vedle detekce je oblastí zájmu v tomto oboru i klasifikace. Tyto dvě navzájem si blízké úlohy jsou podmnožinou rozpoznávání objektů. U všech zmíněných pojmů bylo v posledních několika letech zaznamenáno prudké zdokonalení, a to zejména díky postupu v metodách strojového učení a konvolučních neuronových sítí [36].

Cílem klasifikace je na základě vstupních dat předpovědět, do jaké třídy objekt na snímku spadá. Jedná se o klasifikaci obrazu jako celku, výstupem tedy je daná třída objektu, nebo informace, že žádný objekt nebyl rozpoznán. Při detekci se na snímek nedívá jako na celek, cílem je vyhledat jednotlivé objekty, přiřadit k nim třídu a označit, kde na snímku se nacházejí. Detekce je tedy predikce souřadnic všech rozpoznávaných objektů a jejich klasifikace do tříd.

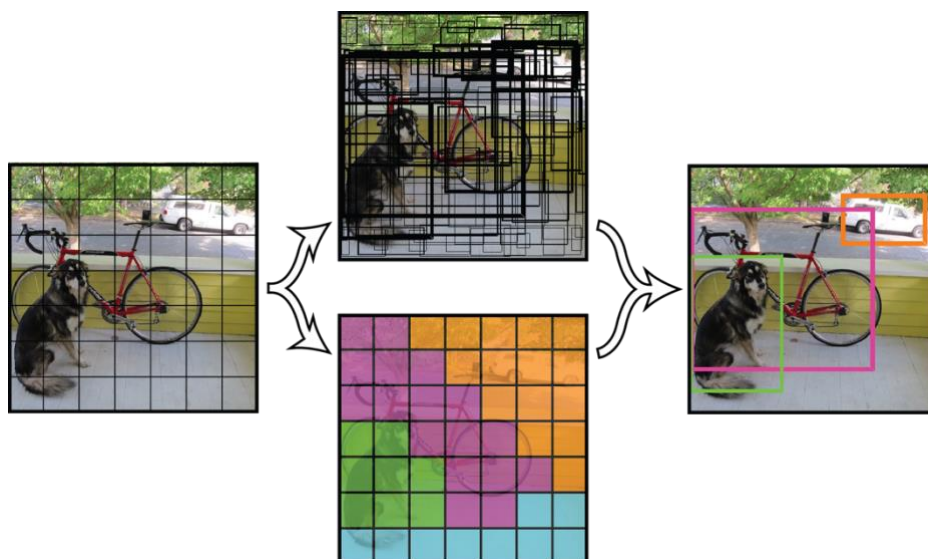
Publikované algoritmy pro detekci objektů se dělí do dvou hlavních skupin. Rozlišují se dvoustupňové detektory jako například Faster R-CNN (Region-based Convolutional Neural Networks) a Mask R-CNN, které využívají komponentu Region Proposal Network (RPN) ke generování oblastí zájmu v první fázi, a následně tyto oblasti zájmu v druhé fázi předávají ke klasifikaci tříd a regresi ohraničujících oblastí („bounding box“). Druhou skupinou jsou jednostupňové detektory YOLO („You only look once“) a SSD („Single Shot Multibox Detector“), které přistupují k úloze detekce jako k obyčejné klasifikaci a regresi. Vyznačují se vysokou rychlostí, která umožňuje detekovat objekty v reálném čase, v porovnání s dvoustupňovými detektory jsou srovnatelně přesné [39][40][41].

5.3 YOLO („You Only Look Once“)

Většina publikovaných algoritmů přistupuje k problému detekce ve dvou fázích. Jsou nalezeny regiony zájmů, které se následně pomocí konvoluční sítě určené pro řešení klasifikačních problémů rozpoznají. Model YOLO využívá jediné konvoluční síť, která současně předpovídá pozici ohraničujících oblastí v obraze a pravděpodobnosti příslušnosti k dané třídě. Jelikož je celý proces detekce prováděn jednou sítí, klesá výpočetní náročnost a je možné detekovat objekty i na videu v reálném čase. Dalším benefitem, který tento model přináší, je pohled na obraz jako na celek. Na rozdíl od metod založených na posuvném okně nebo na RPN (Region Proposal Network) se YOLO dívá na celý obraz a dokáže tak kromě obecného vzhledu objektů zpracovat i jejich kontextuální informace. Tento přístup dodává robustnosti modelu a algoritmus tak méně chybí v situacích, kdy se setkal s úplně novým typem dat.

Samotná detekce objektů začíná rozdělením vstupního obrazu na mřížku o rozměrech $S \times S$, jestliže střed objektu spadá do jedné z buněk mřížky, je tato buňka zodpovědná za detekci objektu. Každá buňka dále predikuje B ohraničujících oblastí, pro které zároveň určí skóre. Tyto skóre reflektují jistotu, s jakou model určil ohraničující oblasti a jak přesné toto ohraničení kolem objektu je. Jestliže se tedy žádný objekt uvnitř buněk nenachází, odpovídá skóre nule.

V opačném případě se skóre odvíjí od poměru průniku a sjednocení (IOU–Intersection over Union) predikované ohraničující oblasti a anotace. Každá ohraničující oblast je definována souřadnicemi x a y , které udávají její relativní střed vzhledem k hranicím buňky. Šířka a výška ohraničující oblasti jsou predikovány relativně k celému obrazu. Nakonec je vypočteno skóre jistoty vycházející z IOU predikované ohraničující oblasti a anotace. Dále každá buňka generuje C predikcí tříd. Nehledě na počet ohraničujících oblastí B je pro každou buňku generován pouze jeden vektor predikcí [40].



Obrázek 5.3: Detektor YOLO rozdělí vstupní obraz na mřížku o velikosti buněk. Pro každou buňku vytvoří B ohraničujících oblastí, jejich jistoty a vektor predikcí. Výsledný tenzor má pak velikost $S \times S \times (5 \times B + C)$. 5 zde představuje čtyři souřadnice ohraničující oblasti a skóre jistoty [40].

5.3.1 YOLOv2 (dřívější název YOLO9000)

Největším nedostatkem původní verze je v porovnání s jinými detektory jako jsou například Faster-RCNN nebo Fast-RCNN velká chybovost při lokalizaci objektů a poměrně nízká senzitivita. Autoři tedy publikují optimalizovanou verzi původního algoritmu nesoucí název YOLO9000, která odstraňuje nedostatky původního modelu při zachování jeho jednoduchosti a rychlosti.

Autoři do své sítě přidali „BatchNorm“ vrstvu, která je podrobněji popsána v kapitole 3.1.5. Uvádějí, že dosáhli 2% zlepšení v mAP. Zároveň mohli odstranit všechny „dropout“ vrstvy bez toho, aby došlo k přetrénování modelu.

U původní verze YOLO se k trénování klasifikační sítě používají vstupní obrazy o velikosti 256×256 pixelů. Samotná detekce se uskutečňuje na velikosti 448×448 pixelů. YOLOv2 se už při samotném tréninku klasifikační sítě po několika prvních epoch učí na velikostech 448×488 pixelů. Parametry filtrů mají čas se upravit tak, aby byla síť schopna lépe reagovat na vstupní obrazy o větších rozměrech.

Asi největší změnou oproti předešlé verzi přineslo použití tzv. „anchor boxes“. Jedná se o oblasti zájmu, kde se může vyskytovat hledaný objekt, a které se automaticky generují pro každou buňku aktivační mapy (podrobnější popis v kapitole [5.4](#)). Pro každý „anchor box“ se predikuje klasifikační třída a skóre, které je úměrné průniku nad sjednocením (IOU) „anchor box“ a skutečnou ohraničující oblastí. Tato změna si vyžádala úplné odstranění posledních plně propojených vrstev detektoru. Zároveň se odstranila „maxpool“ vrstva páteřní sítě k získání aktivační mapy o větším rozlišení a celá páteřní síť se přizpůsobila na vstupní snímek o velikosti 416×416 pixelů. Aktivační mapa má nyní rozměr 13×13 a vzniká tak jediná buňka uprostřed mapy. Autoři uvádějí, že tato buňka zodpovídá za detekci velkých objektů, které obvykle okupují právě střed obrazu. Uvedená úprava podle publikace mírně snižuje hodnotu mAP, avšak výrazně vylepšuje sensitivitu, a tudíž vzniká prostor pro další vylepšování detektoru [\[42\]](#).

5.3.2 YOLOv3

Nová verze YOLOv3 přináší další optimalizaci původní architektury YOLOv2. Pro klasifikaci do třídy se pro každé krabicové ohraničení již nepoužívá funkce „Softmax“, jako to bylo u předešlých verzí. Nahradila ji funkce binární křížové entropie. Autoři uvádějí, že využívání funkce „Softmax“ již nebylo nutné pro dosažení dobrých výsledků, a zároveň jim tato optimalizace umožňuje lépe pracovat s komplexnějšími daty, kde se jednotlivé klasifikační třídy mohou u objektů překrývat. Funkce „Softmax“ totiž každému krabicovému ohraničení přiřazuje pouze jednu klasifikační třídu, což vždy neodpovídá skutečnosti.

Změnila se také páteřní síť pro extrakci příznaků. Nově se používá konvoluční neuronová síť Darknet-53, kterou tvoří 53 konvolučních vrstev. Je inspirovaná reziduálními neuronovými sítěmi [\[43\]](#).

5.3.3 YOLOv4

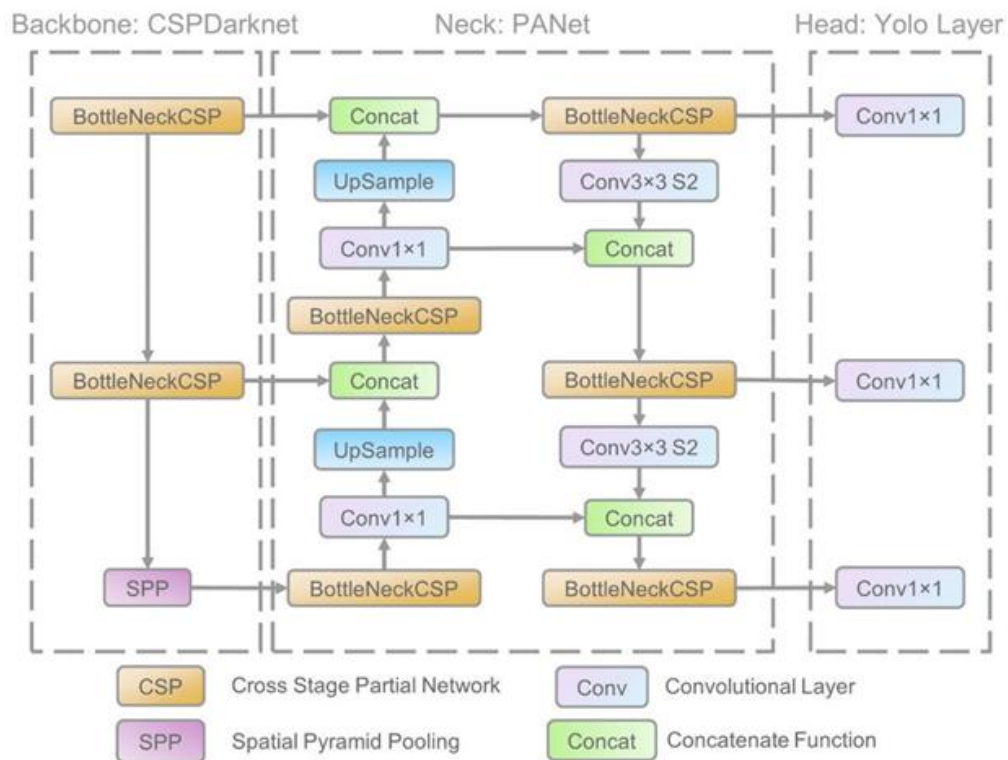
Hlavním cílem autorů u čtvrté verze detektoru rodiny YOLO bylo nadále zdokonalovat úspěšnost a rychlost. Zároveň usilují o to, aby bylo možné síť natrénovat a testovat na běžném GPU. Na začátku své publikace autoři shrnují tzv. „Bag of freebies“. Patří sem metody augmentace dat a různé varianty chybové funkce pro dosažení lepšího trénování sítě. Dále jsou zde uvedeny i tzv. „Bag of specials“, coby plugin moduly, které mají zlepšit přesnost detektoru. Patří sem komponenty pro zvětšení recepčního pole sítě, komponenty pro dokonalejší zpracování příznaků a postprocessingové metody pro sledování predikcí sítě.

Následně autoři provedli rozsáhlé testování a sledovali vliv „Bag of specials“ a „Bag of freebies“ na trénování a výsledky. Nakonec vycházejíce z výsledků představili finální architekturu YOLOv4 [\[44\]](#).

5.3.4 YOLOv5

Ačkoli je architektura YOLOv5 volně dostupná jako open source, v době psaní této práce ještě nevyšla žádná oficiální publikace s prokazatelnými výsledky. Společnost Ultralytics, která YOLOv5 implementovala, na svém Github repozitáři uvádí, že tento model je v porovnání s YOLOv4 zabere asi o 90 % méně paměti, přičemž dosahuje podobných výsledků. Mimo jiné je podle autorů asi třikrát rychlejší než jeho předchůdce. Existuje tedy jistá kontroverze ohledně dosažených výsledků a názvu samotného [45]

Architekturu detektoru je možné rozdělit na tři části: páteřní síť, krk a hlavu. Páteřní síť CSP Darknet, kterou detektor používá, je inspirovaná tzv. „Cross Stage Partial Network“ [46]. Ta využívá postupného řetězení výstupů jednotlivých konvolučních vrstev, které při optimalizaci umožňuje sdílení gradientu chybové funkce, a tak výrazně snižuje výpočetní náročnost. Jako krk autoři použili PANet, kde dochází k vhodné kombinaci příznakových vrstev [47]. Hlava je pak stejná jako u předchozí verze YOLOv3.



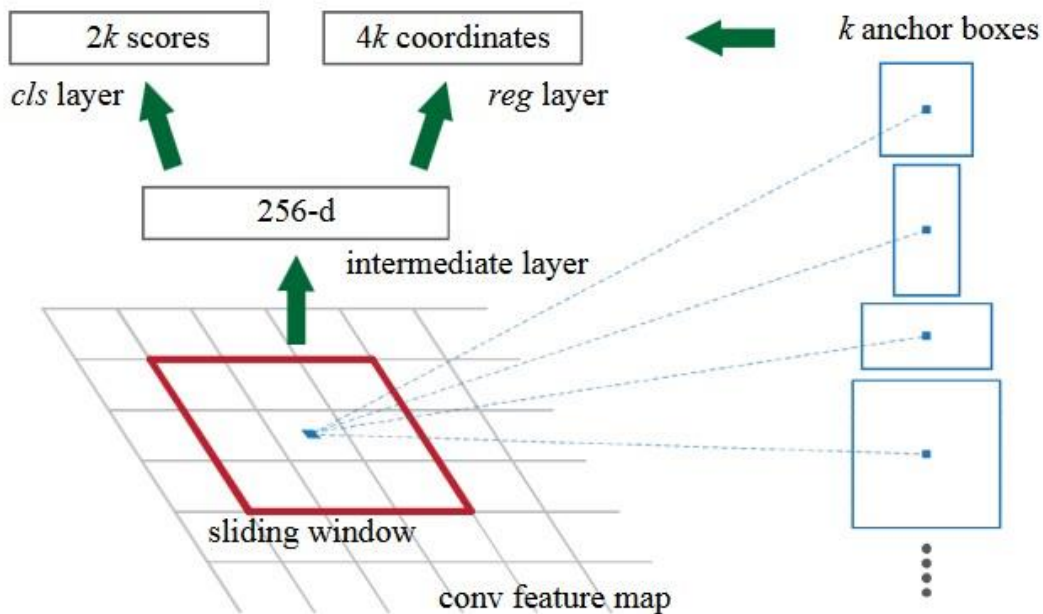
Obrázek 5.4: Schéma architektury YOLOv5. Neuronové vrstvy páteřní sítě („Backbone“) s menším počtem neuronů („Bottleneck“) předávají svůj výstup do krku („Neck“), kde dochází k jejich zřetězení („Concat“) a převzorkování („UpSample“). Blok „Conv“ představuje operaci konvoluce a blok „SPP“ tzv. „Spacial Pyramid Pooling“, který se umísťuje na konec páteřní sítě a umožňuje, aby na vstup detektoru mohl být přiveden obraz různé velikosti. Výstup z krku je dále předán hlavě („Head“), kde probíhá samotná klasifikace a regrese ohraničujících oblastí [45].

5.4 Faster-RCNN

Dvoustupňový „State of art“ detektor z rodiny RCNN představuje další alternativu k řešení klasifikačních a lokalizačních problémů počítačového vidění. Na rozdíl od jeho předchůdců používá nově během detekování objektů tzv. „Region Proposal Networks“ (RPNs), klíčový prvek architektury zkracující dobu detekce pro snímek z 2 sekund (jak tomu bylo u předchozích architektur z této rodiny detektorů) na 10 ms a výrazným způsobem regulující výpočetní náročnost.

Detektor Faster-RCNN se skládá z dvou modulů. Prvním je hluboká plně konvoluční neuronová síť, která generuje oblasti zájmu („region proposals“). Následuje druhý modul, kterým je předchůdce ze stejné rodiny detektorů: Fast-RCNN. Celý systém představuje jedinou sjednocenou síť schopnou detekce objektů.

RPN jako vstup přijímá mapu příznaků a na výstup přivádí stovky tzv. „object proposals“ (obdélníkové ohraničující oblasti zájmu, ve kterých se může vyskytovat hledaný objekt). Každý „object proposal“ je navíc ohodnocen skórem („objectness score“), který popisuje jistotu detektoru, se kterou rozlišuje pozadí od skutečného objektu. RPN přijímá příznakovou mapu z páteřní hluboké konvoluční sítě („backbone“), která za účelem snížení výpočetní náročnosti sdílí vrstvy s hlubokou konvoluční sítí Fast-RCNN. RPN se skládá z menší konvoluční sítě, která klouže po aktivační mapě a při každém posunutí provádí konvoluci s daným okřskem aktivační mapy (v publikaci uveden rozměr $n \times n \times d$, kde $n = 3$ a d je hloubka výsledné mapy). Výsledek konvoluce následně putuje do dvou plně propojených neuronových vrstev: „box-regression layer“ („*reg*“, vrstva pro regresi ohraničujících oblastí) a „box-classification layer“ („*cls*“, vrstva pro klasifikaci ohraničujících oblastí). Při posunu okna přes aktivační mapu, se pro každý její pixel vygeneruje k oblastí zájmů. *Reg* vrstva přivádí na výstup $4k$ souřadnic oblastí zájmů, *cls* vrstva potom na výstup dává $2k$ skóre, které popisují pravděpodobnosti, že oblast zájmu je pozadí nebo hledaný objekt. k oblastí zájmu je odvozeno od referenčních obdélníkových ohraničujících oblastí („anchor boxes“, „anchors“). Ty jsou vycentrovány pro každé posunutí přes aktivační mapu. Autoři uvádějí, že výchozí parametry pro „anchors“ jsou tři možné poměry stran a tři možné rozměry, celkově se tedy získá 9 „anchor boxes“. Pro konvoluční mapu o rozměrech $H \times W$ je celkem vytvořeno HWk „anchors“ [41][48].

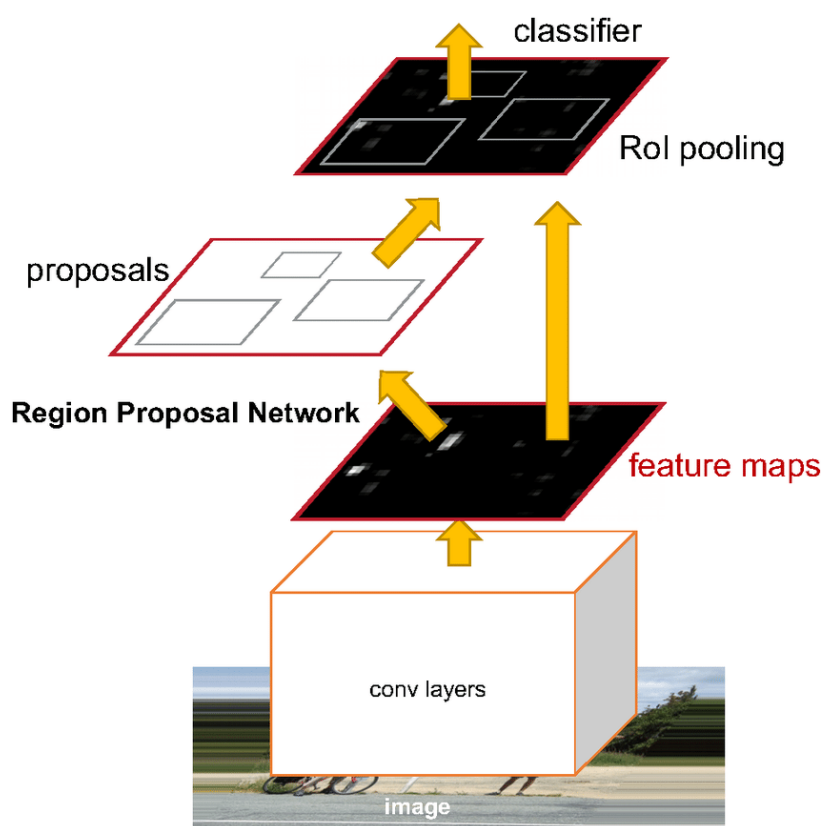


Obrázek 5.5: „Region Proposal Network“. Menší konvoluční síť („intermediate layer“) o hloubce 256 klouzá („sliding window“) po aktivací mapě („conv feature map“) generované páteřní sítí. Pro každou pozici je vytvořeno k obdélníkových referenčních ohraničujících oblastí („anchor boxes“). Oblasti zájmu jsou předány vrstvám pro regresi krabicových ohraničení („reg layer“) a jejich klasifikaci („cls layer“) [41].

RPN lze samostatně trénovat pomocí algoritmu zpětného šíření chyby a stochastického gradientního sestupu. Nastává však komplikace v případě, kdy se za RPN dosadí neuronová síť, která dále zpracovává všechny oblasti zájmu, které RPN generuje. Jak bylo zmíněno na začátku kapitoly, druhým modulem, který pracuje právě s výstupem RPN je Fast-RCNN. Pokud se RPN a Fast-RCNN trénují samostatně, váhy filtrů se modifikují odlišně. Je tedy nutný algoritmus, který umožní vzájemné sdílení konvolučních vrstev mezi dvěma sítěmi. Autoři představují pragmatický čtyř-krokový algoritmus, který toto sdílení vrstev umožňuje. Jeho popis je dostupný v [41][48].

Oblasti zájmu vygenerované RPN společně s aktivační mapou získanou pomocí páteřní sítě vstupují společně do tzv. „Region of interest pooling layer“ („ROI pooling“). Jedná se o speciální podvzorkovací vrstvu, která každou oblast zájmu na aktivační mapě rozdělí na pevný počet oken. Na každém okně následně provede „max-pooling“ a získá tak výstup o požadované velikosti. Výstupem „ROI“ vrstvy je pak vektor o velikosti $N \times W \times H \times C$, kde N je počet oblastí zájmu, W a H jsou počty oken v horizontálním a vertikálním rozměru, C je hloubka aktivační mapy.

Tento vektor dále vstupuje do dvou plně propojených sítí a následně do sítě pro finální regresi ohraničujících oblastí a do sítě pro finální klasifikaci tohoto ohraničení do třídy [41][48].



Obrázek 5.6: Obrázek znázorňující architekturu Faster-RCNN. Páteřní síť („conv layers“) ze vstupního obrazu („image“) vytvoří aktivační mapu („feature map“). „Region Proposal Network“ vygeneruje oblasti zájmu („proposals“), které společně s aktivační mapou vstupují do „ROI pooling“ vrstvy. Ta svůj výstup předá klasifikátoru a regresoru ohraničujících oblastí [41].

6. PUBLIKOVANÉ METODY PRO DETEKCI A KLASIFIKACI HEMORAGIÍ

V posledních letech bylo publikováno mnoho prací zabývajících se detekcí intrakraniálních krvácení zejména pomocí strojového učení. V kapitole se nachází soupis několika z nich, který bude sloužit jako východisko k vlastními návrhu řešení.

Chilamkurthy a kol. (2018) publikují rozsáhlou práci, jejíž součástí jsou dva datasey CT snímků hlavy a algoritmus provádějící detekci intrakraniálních krvácení, fraktur lebky, posunu středové linie a přesunu mozkové tkáně. Pro detekci krvácení použili autoři konvoluční neuronovou síť o architektuře ResNet18. Tutu síť autoři modifikovali tak, aby u snímku předpověděla pravděpodobnost výskytu každého z typů krvácení. Dále byla provedena segmentace krvácení pomocí architektury UNet. Při předzpracování dat byly vybrány pouze bezkontrastní snímky s rekonstrukčním filtrem pro měkké tkáně, následně se upravili na velikost 244×244 pixelů. Posledním krokem předzpracování je transformace kontrastu podle mozkového, subdurálního a kostního okna. Všechny tři transformace snímku byly následně přivedeny na vstup konvoluční neuronové sítě. Autoři uvádějí, že nasbírali přes 300 tisíc CT skenů hlavy, ze kterých následně sestavily datasey Qure25k a CQ500 (viz kapitola 5). Na nich pak testovali úspěšnost navrženého algoritmu. Na datasetu CQ500 AUC (plocha pod ROC křivkou) zde pro nitrolební krvácení nabývala hodnoty 0,9419 [49].

Ye a kol. (2019) ve své publikaci k detekci a klasifikaci intrakraniálního krvácení použili kaskádový CNN-RNN model založený na hlubokém učení. Jedná se o konvoluční neuronovou síť následovanou rekurentní neuronovou sítí. Autoři uvádějí, že se algoritmem chtěli přiblížit postupu, jakým radiologové interpretují CT snímky. První neuronová síť provádí extrakci příznaků a klasifikaci do dvou skupin podle přítomnosti, respektive nepřítomnosti krvácení. Druhá síť na základě získaných příznaků určí, o jaký typ krvácení se jedná. Použila se databáze o 2836 CT skenech, kde prevalence krvácení je 65 %. Celá databáze byla rozdělena na tři nestejně velké celky k učení, validaci a testování v poměrech 8:1:1. Algoritmus dosahuje při určování přítomnosti krvácení 98% sensitivity a 99% specificity. Při klasifikaci krvácení dosáhla specifita nad 80 % napříč všemi typy krvácení [50].

Pro detekci intrakraniálního krvácení a jeho klasifikaci navrhli Cho. a kol. (2019) diagnostický systém skládající se z dvou neuronových konvolučních sítí (CNN) provádějících binární klasifikaci podle (ne)přítomnosti krvácení, za kterými v kaskádě následují duální plně konvoluční sítě (FCN). Ty vedle klasifikace krvácení do konkrétního typu provádějí i segmentaci obrazu. Autoři se inspirovali běžným postupem radiologů, kteří používají radiologická okna pro lepší diagnózu, a na vstup konvolučních neuronových sítí přivedli CT snímky transformované jiným radiologickým oknem. Byla použita databáze 135 974 bezkontrastních snímků od 5702 pacientů. U binární klasifikace dosahoval algoritmus 98,28% přesnosti („accuracy“). Při klasifikaci krvácení do jednoho z pěti typů a segmentaci se preciznost („precision“) pohybovala mezi 70 a 90 % v závislosti na daném typu krvácení [51].

Ertugrul a kol. (2021) využili veřejně dostupného datasetu BHX coby rozšíření CQ500 o ohraničující oblasti anotovaných krvácení k jejich detekci za pomoci architektury YOLOv4. Provedli „fine-tuning“ předtrénované architektury na cílovém datasetu, při kterém sledovali vliv optimalizačních algoritmů, „batch-size“ a učícího kroku na výsledky. Ke zhodnocení úspěšnosti použili metriky sensitivitu, pozitivní prediktivní hodnotu a F1 skóre. Nejlepších výsledků dosáhli, pokud před samotným trénováním celý dataset rozdělili na části podle jednotlivých typů krvácení a trénovali detektor samostatně pro každý typ. V tomto případě při detekci dosáhli u každého typu krvácení sensitivity nad 85 % a prediktivní pozitivní hodnoty nad 89 %. F1 skóre nekleslo pod 87 % [52].

7. DATABÁZE CQ500

Jedná se o volně dostupnou databázi 491 CT skenů hlavy s 193 317 2D řezy. Všechny data byla při sběru anonymizována. Sběr dat proběhl ve dvou vlnách. První vlna obsahuje všechny CT skeny, které byly provedeny v konkrétních lékařských centrech po dobu 30 dnů počínaje 20. listopadu 2017. Druhá část dat byla získána pomocí algoritmu na zpracování přirozeného jazyka z radiologických zpráv tak, aby byl podtypy intrakraniálního krvácení (IPH, SDH, EDH a SAH) zastoupeny 80 skeny. Odebrány byly pak všechny skeny, které porušují zavedená kritéria. Pacient by neměl mít žádné postoperační defekty a neměl by být mladší 7 let. Měl by se zde nacházet taky jeden bezkontrastní sken zahrnující celý mozek s rekonstrukčním algoritmem pro měkkou tkáň.

Součástí samotných CT skenů jsou i anotace, které byly vytvořeny třemi nezávislými radiology. Ti nebyli součástí personálu, který se podílel na léčbě daných pacientů, a neměli přístup k jejich klinické historii. Každý z třech radiologů hodnotil, zda se na konkrétních snímcích nevyskytuje, nebo vyskytuje intrakraniální krvácení a popřípadě o jaký podtyp krvácení se jedná (epidurální, subdurální, subarachnoidální, intracerebrální, nitrokomorové). Dále je v anotacích uvedeno, jestli se jedná o chronické krvácení či nikoli. Hodnotí se, jaká hemisféra byla postihnuta. Nakonec radiologové uvedli, zda se na snímku vyskytuje či nevyskytuje posun středové linie, přesun mozkové tkáně a zlomenina. O posun se středové linie se jedná v případě, že je větší než 5 mm. Přesun mozkové tkáně je definován jako lokální přesuny, útlak mozkových komor nebo herniace. Pokud došlo k neshodě mezi třemi radiology, finální anotace je pak většinový názor [49].

7.1 Brain hemorrhage extended (BHX)

Brain hemorrhage extended představuje rozšíření původní anotace databáze CQ500. Autoři zveřejnili 39 668 ohraničujících oblastí (v angl. „Bounding Box“) hemoragií ve 23 409 snímcích CT pro pět podtypů intrakraniálního krvácení včetně dodatečného šestého označení pro chronické subdurální krvácení. Na vytvoření anotace se podíleli tři neuroradiologové. Ti ručně označili hranice hematomů na 3 mm tlustých CT řezech. Tato práce přinesla 6238 krabicových ohraničení na 3558 různých CT snímcích. Pomocí extrapolace krabicového ohraničení na tenčí 1 mm široké řezy došlo k rozšíření databáze na finální velikost.

Dostupné jsou celkem tři datasety. První obsahuje pouze ručně vytvořené anotace pro 3 mm tlusté řezy. V druhém datasetu jsou navíc zahrnuty extrapolace na 1 mm tlusté řezy nehledě na to, jestli se jedná o kontrastní CT nebo jestli byl použit rekonstrukční filtr pro kosti. Poslední dataset na rozdíl od druhého neobsahuje post-kontrastní CT a nachází se zde pouze snímky získané rekonstrukčním filtrem pro měkké tkáně [53].

8. PŘEDZPRACOVÁNÍ DAT A TRÉNOVÁNÍ DETEKTORŮ

Tato kapitola se věnuje formulaci zadaného problému a návrhu samotného řešení rozděleného do dílčích nezávislých kroků.

8.1 Obecný návrh

Cíl praktické části představuje přesně klasifikovat a lokalizovat intrakraniální krvácení na axiálních CT snímcích hlavy. Pro řešení tohoto problému budou jako zdroje dat využity veřejně dostupné databáze CQ500 a BHX podrobněji popsané v kapitole 7. Samotná klasifikace a lokalizace bude prováděna detektory objektů založenými na konvolučních neuronových sítích popsanými v kapitole [5.2](#).

Celý postup je vhodné rozdělit na několik dílčích kroků. Prvním krokem při řešení zadaného problému bude surová data z databází podrobit vhodnému předzpracování. Následně vybrat architekturu detektoru a implementovat odlišné postupy pro efektivní doučení na cílové databázi („fine-tuning“, viz kapitola [4.2.1](#)).

8.2 Databáze a předzpracování dat

Jako surová data posloužila veřejně dostupná databáze CT skenů hlavy CQ500 a databáze ohraničujících oblastí intrakraniálních krvácení Brain Hemorrhage Extended (BHX). Důležitým krokem přípravy dat je vyseparování všech CT snímků z databáze CQ500, pro které existuje v databázi BHX anotace ve formě ohraničujících oblastí. Pouze tyto snímky bude možné použít pro trénování, validaci a testování detektoru. Zastoupení jednotlivých typů krvácení v obou databázích je uvedeno v následující tabulce.

Tabulka 8.1: Počty ohraničujících oblastí a CT skenů jednotlivých typů krvácení

| Typ krvácení | Počet ohraničujících oblastí (BHX) | Počet CT skenů (CQ500) |
|-----------------------------|------------------------------------|------------------------|
| Intraparenchymální krvácení | 5264 | 134 |
| Epidurální krvácení | 548 | 13 |
| Subdurální krvácení | 7906 | 53 |
| Intraventriculární krvácení | 2219 | 28 |
| Subarachnoidální krvácení | 7596 | 60 |
| Chronické | 3672 | NA |

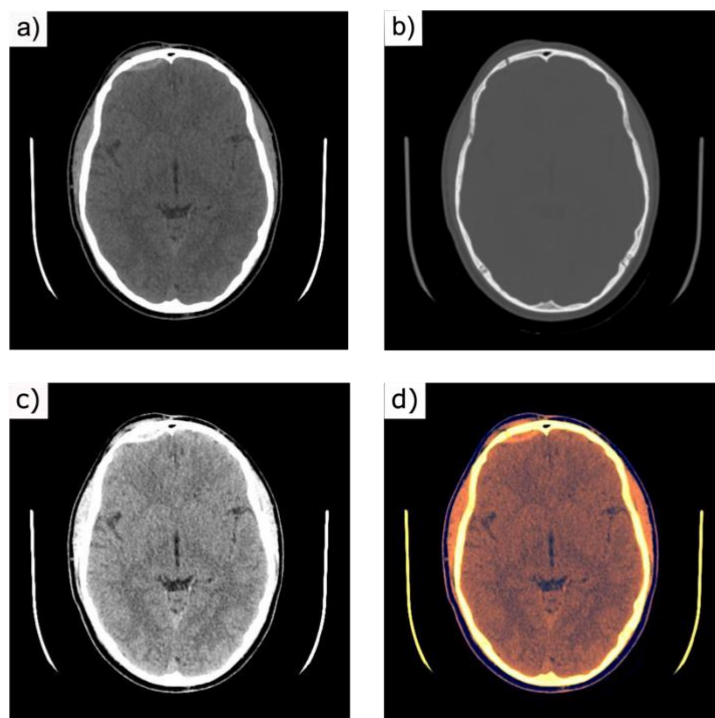
Po extrahování všech potřebných dat se tedy velikost databáze dostala na celkových 27203 ohraničujících oblastí na 15979 obrazech. Někteří pacienti mají anotovaný ohraničujícími oblastmi více než jeden CT sken. Při pilotních měřeních bylo však zjištěno, že je výhodnější pro jednoho pacienta mít pouze snímky z jednoho skenu. To udržuje variabilitu dat a brání rychlému přeučení sítě, což byl problém prvotních měření. Velikost databáze se tak zmenšila oproti hodnotám uvedených v tabulce výše asi o polovinu.

Před trénováním detektoru byl dataset rozdělen na trénovací, testovací a validační část v poměru 5: 1: 1. Rozdělení proběhlo na úrovni pacientů, aby nedošlo k situaci, že snímky jednoho CT skenu pacienta budou v trénovací i testovací části. Trénovací část byla použita k trénování detektoru, validační část na sledování chyby kritériální funkce během učení a testovací část na vyhodnocení detekční schopnosti modelu.

Tabulka 8.2: Počty ohraničujících oblastí napříč částmi databáze

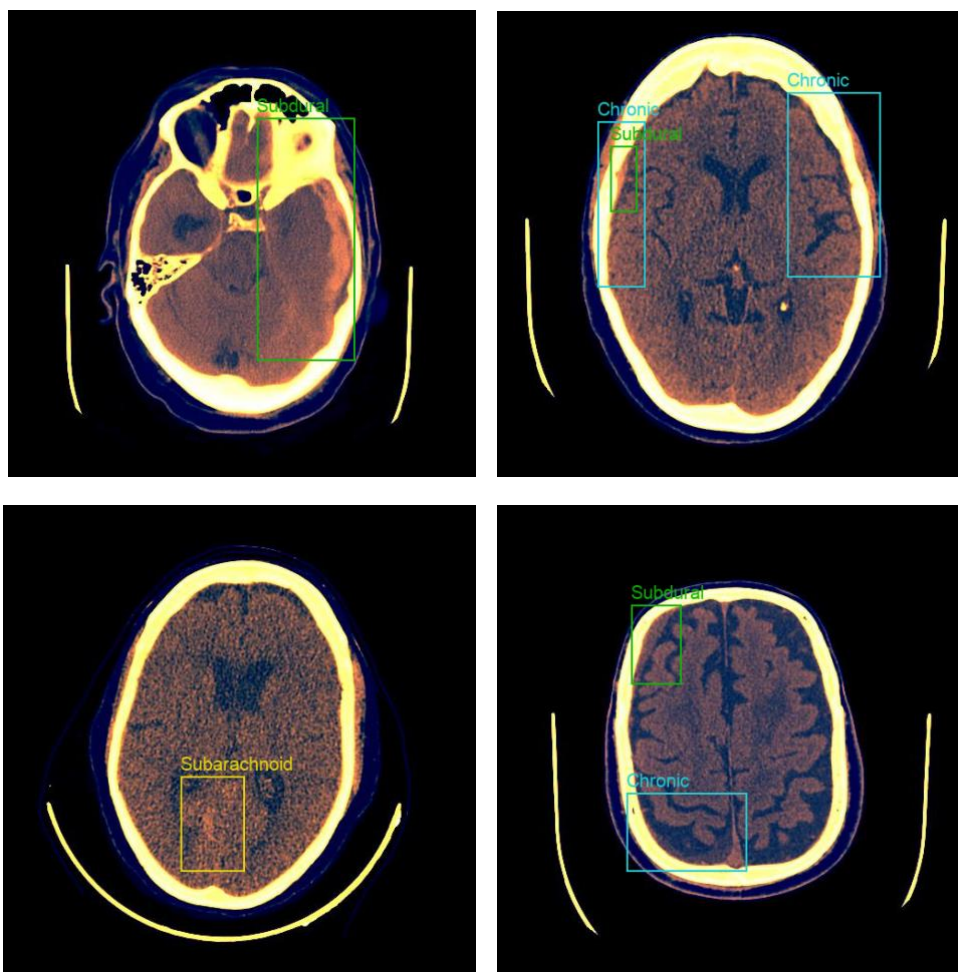
| Typ krvácení | Počet ohraničujících oblastí | | |
|--------------------|------------------------------|----------------|----------------|
| | Trénovací část | Testovací část | Validační část |
| Intraparenchymální | 1689 | 420 | 357 |
| Subdurální | 2509 | 1031 | 602 |
| Epidurální | 108 | 29 | 54 |
| Intraventriculární | 585 | 210 | 236 |
| Subarachnoidální | 2573 | 707 | 694 |
| Chronický | 1554 | 261 | 427 |

Na základě rešerše provedené v kapitole 6 byl každý CT snímek před vstupem do detektoru podroben následujícímu předzpracování. Šedotónový 12bitový dvourozměrný CT snímek o velikosti 514×514 pixelů se transformoval pomocí subdurálního, kostního a mozkového radiologického okna. Každá ze tří transformací se následně použila jako jeden kanál RGB obrazu, který je pak samotným vstupem do detektoru. Tento krok je vhodné provést z toho důvodu, že každý z typů nitrolebečního krvácení je snadnější rozpoznat za použití jiného radiologického okna. Obraz během fáze předzpracování nebylo nutné žádným způsobem podvzorkovat či ořezat.



Obrázek 8.1: Předzpracování CT snímku. Transformace pomocí a) subdurálního, b) kostního, c) mozkového radiologického okna. Obrázek d) představuje snímek přivádění na vstup detektoru.

V neposlední řadě v závislosti na zvoleném detektoru bylo nutné vstupní snímky a anotace převést do takového formátu, který detektor na vstupu požaduje. V případě detektoru YOLOv5 musí být pro každý vstupní snímek zvlášť vytvořen textový dokument obsahující čtyři souřadnice a klasifikační třídu. Dvě souřadnice ukazují na střed ohraničující oblasti, zbylé dvě určují její výšku a šířku. Všechny souřadnice jsou navíc normalizované od nuly do jedné relativně k velikosti celého snímku. Detektor Faster-RCNN během fáze učení požaduje kromě vstupního snímku, čtyř souřadnic (souřadnice zde ukazují na levý horní a pravý spodní roh ohraničující oblasti) a klasifikační třídy také obsah jednotlivých ohraničujících oblastí.



Obrázek 8.2: Snímky po předzpracování včetně ohraničujících oblastí a klasifikačních tříd.

8.3 Trénování detektorů

Během trénování detektoru probíhá obrovské množství poměrně jednoduchých matematických výpočtů jako například násobení a sčítání matic. Tyto jednotlivé výpočty jsou na sobě nezávislé a vzniká tak možnost je provádět paralelně. Grafické karty, které v dnešní době obsahují stovky až tisíce jader, představují ideální řešení pro problém paralelizace výpočtů a mohou až několikanásobně urychlit trénování neuronových sítí [54].

Pytorch, open-source framework podporující akceleraci výpočtů pomocí grafických karet s podporou architektury CUDA, byl použit pro řešení praktické části bakalářské práce. Pytorch představuje velmi robustní ekosystém nástrojů a knihoven, které rozšiřují uplatnění frameworku také na úlohy spojené se zpracováním přirozeného jazyka nebo video záznamu [55].

Pro účely hlubokého učení nad architekturou CUDA stojí ještě knihovna cuDNN („CUDA Deep Neural Network Library), která poskytuje perfektně vyladěné a optimalizované implementace pro standardní úlohy jako je konvoluce, podvzorkování, normalizace a aktivace vrstev. Tyto API pak vývojáři knihoven hlubokého učení (např. právě Pytorch) přímo využívají, aby se tak oprostili od ladění nízkourovňového výkonu GPU [54].

Trénování detektoru probíhalo na stolním počítači s grafickou kartou nVIDIA GeForce GTX 1060 s 1280 CUDA jádry. V případě detektoru YOLOv5 byly vyžadovány knihovny Pytorch ve verzi 1.7.0 a Torchvision ve verzi 0.8.1 či novější. Pro jejich správnou funkci je nutné mít verzi CUDA alespoň 9.2. CUDA 9.2 požaduje ovladač grafické karty verze 397, nebo novější. Vše tedy muselo být podřízeno těmto požadavkům. Trénování bylo realizováno na instalaci Windows 10 s ovladačem GPU 511.23 a CUDA 11.6.

8.3.1 Trénování detektoru Faster-RCNN

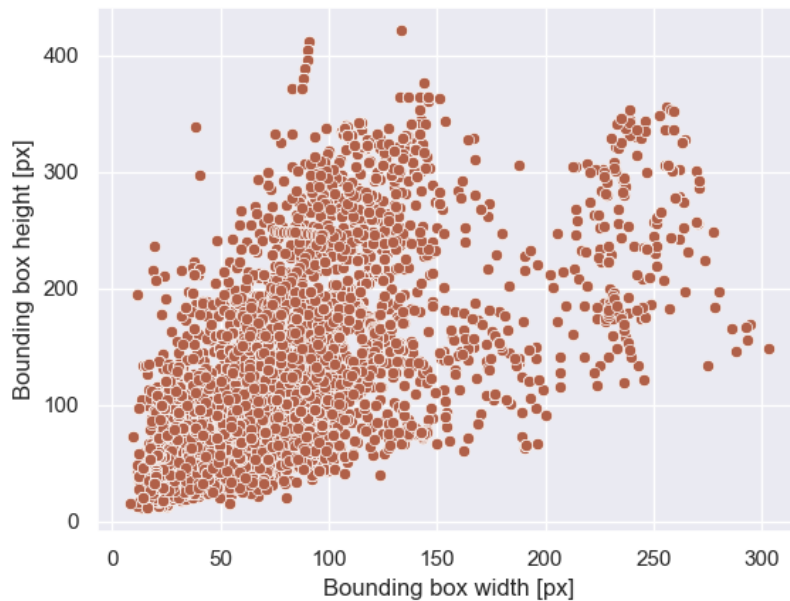
Během procesu trénování detektoru Faster-RCNN byla použita trénovací a validační část celého datasetu. Trénovací dataset přímo sloužil k zisku chyb kritériální funkcí při dopředném průchodu trénovacích vzorků detektorem. Tyto chyby se následně zpětně šířili neuronovou sítí a v posledním kroku došlo k optimalizaci parametrů detektoru.

Během trénování „Region Proposal Network“ může být každý „anchor box“ (predikovaná oblast ve které se může nacházet objekt) zařadit do jedné ze dvou tříd v závislosti na tom, jak velký tvoří podíl průniku nad sjednocením se skutečnou ohraničující oblastí (IOU, viz kapitola 9.1). Do pozitivní skupiny patří všechny „anchor boxes“ s největšími IOU nebo takové s IOU nad 0.7. Do negativní skupiny jsou zařazeny všechny „anchor boxes“ s IOU menším než 0.3. Pokud nebyl „anchor box“ přiřazen ani do jedné ze skupin, nepodílí se na optimalizaci parametrů. Chybová funkce má následující tvar:

$$L(\{\mathbf{p}_i\}, \{\mathbf{t}_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(\mathbf{p}_i, \mathbf{p}_i^*) + \lambda \frac{1}{N_{reg}} \sum_i \mathbf{p}_i^* L_{reg}(\mathbf{t}_i, \mathbf{t}_i^*), \quad (8.1)$$

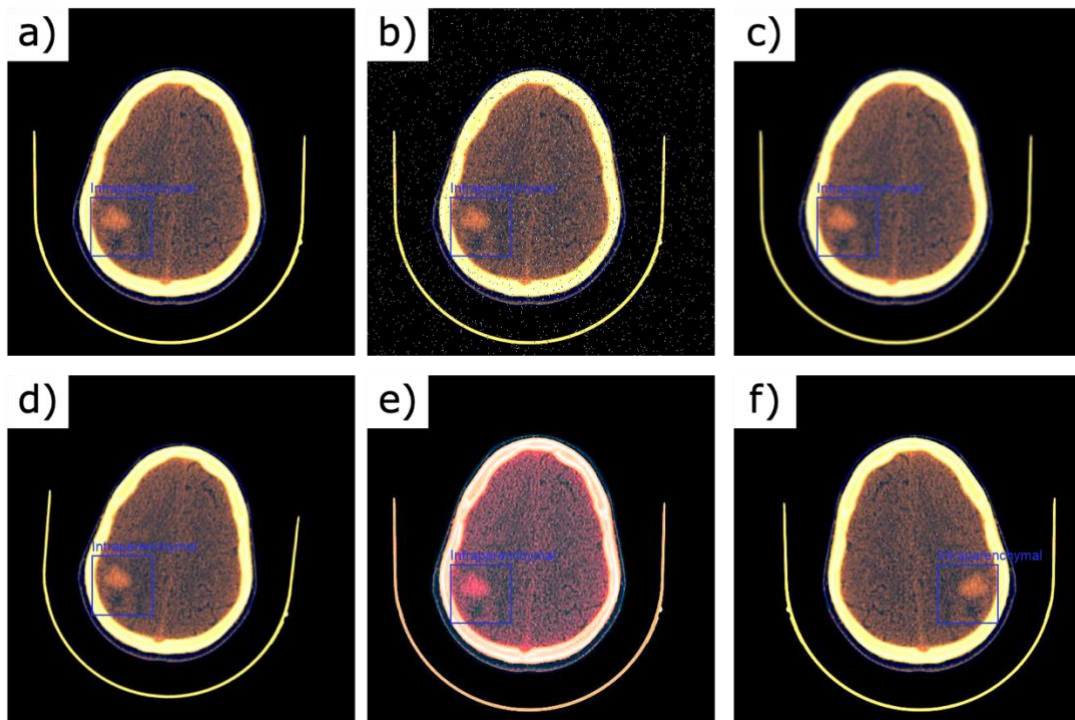
kde i je index daného „anchor box“ v rámci „mini-batch“ a \mathbf{p}_i je pravděpodobnost, že se v něm opravdu nachází objekt. Pravděpodobnost existence objektu ve skutečné ohraničující oblasti \mathbf{p}_i^* nabývá hodnoty 1, pokud byl „anchor box“ zařazen do pozitivní skupiny, či 0 v případě, že patří do negativní skupiny. \mathbf{t}_i představuje vektor souřadnic „anchor boxu“ a \mathbf{t}_i^* vektor souřadnic skutečné ohraničující oblasti, podle které byl „anchor box“ zařazen buďto do pozitivní nebo negativní třídy. Chybovou funkcí pro klasifikaci L_{cls} je logaritická ztráta („log loss“). Pro regresi krabicových ohraničení autoři použili „robust loss function“ popsanou podrobněji v [48]. Člen $\mathbf{p}_i^* L_{reg}$ znamená, že chyba regrese bude deaktivována pro „anchor boxes“, které byli dříve přiřazeny do negativní skupiny (\mathbf{p}_i^* zde nabývá hodnoty 0). N_{cls} je velikost „mini-batch“ ($N_{cls} = 256$) a N_{reg} je počet předpovězených „anchor boxes“ ($N_{reg} \sim 2400$). λ je balanční parametr, který zaručuje, že chyba klasifikace a chyba regrese budou zhruba stejně váhovány [41].

Pytorch umožňuje uživatelům si detektor Faster-RCNN upravovat podle vlastních požadavků. Je zde na výběr mezi několika páteřními sítěmi, přizpůsobit lze i generátor „anchor boxes“ či Fast-RCNN. Pro potřeby řešení této práce byl jako detektor použit Faster-RCNN naučený na COCO datasetu [56]. Regresor a klasifikátor Fast-RCNN byl přizpůsoben na počet klasifikačních tříd v cílovém datasetu a jeho váhy byly inicializovány podle výchozího nastavení Pytorch. Zbylé nastavení Fast-RCNN a RPN bylo ponecháno na výchozích hodnotách. Jako páteřní síť se zvolila ResNet-50 [57] s FPN („Feature Pyramid Network“) [58].



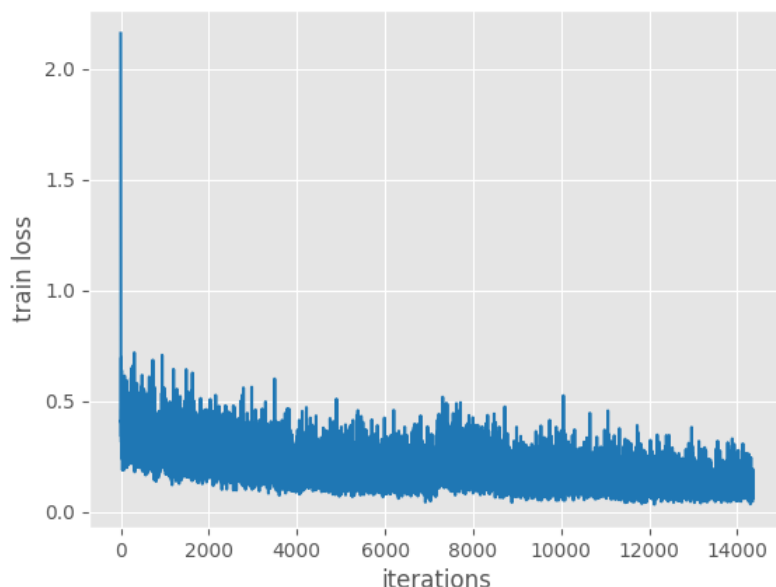
Obrázek 8.3: Bodový graf velikostí ohraničujících oblastí krvácení v trénovací databázi. Generátor „anchor boxes“ byl ponechán na výchozím nastavení. „Anchor boxes“ mají velikost 32, 62, 256 a 512 pixelů při poměrech stran 0.5, 1 a 2.

Bylo také implementováno 5 transformačních metod, které každý snímek před vstupem do detektoru upraví, a tak zvýší variabilitu dat. Transformace byly zároveň zakomponovány do cyklu učení takovým způsobem, aby se každou epochu uplatňovala pouze jedna a navzájem se nepřekrývaly. Každou šestou epochu však na vstup přicházely netransformované originální snímky. Výhodou tohoto přístupu je, že během jedné epochy „vidí“ detektor každý snímek z trénovací databáze pouze jednou.



Obrázek 8.4: Transformační funkce. a) originální snímek; b) impulsní šum; c) Gaussovské rozostření; d) rotace v rozmezí (-30, 30) stupňů; e) úprava jasu a odstínu barvy; f) převrácení podle horizontální osy

Jako optimalizační algoritmus byl zvolen stochastický gradientní sestup s krokem učení 0.0005. „Mini-batch“, pakl snímků přiváděných na vstup detektoru v jeden moment, se nastavil na velikost 4. Celkově byl detektor trénován 12 epoch. Validační dataset sloužil k průběžnému sledování chyby kriteriální funkce během trénování detektoru. V paměti jsou kontinuálně uloženy parametry z poslední epochy a takové epochy, při kterých dosahoval model nejmenší chyby na validačním datasetu. Ty byly pak použity při samotném testování úspěšnosti detektoru.



Obrázek 8.5: Vývoj chyby („train loss“) kriteriální funkce během iterace přes „mini-batch“.

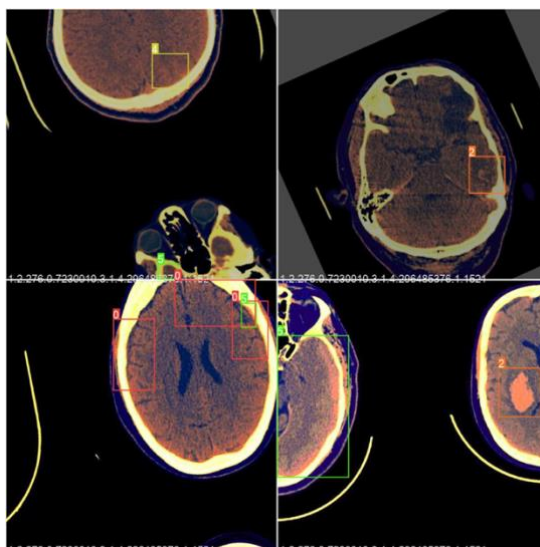
Trénování detektoru YOLOv5

Jako další detektor pro klasifikaci a lokalizaci intrakraniálního krvácení byla zvolena architektura YOLOv5 vyvinutá společností Ultralytics. Jedná se o open-source projekt poprvé zveřejněný 25. června 2020. Kromě samotné architektury detektoru se zde nachází všechny potřebné skripty pro trénování, testování a modifikaci parametrů modelu. Mezi největší výhody oproti Faster-RCNN patří krátká doba tréninku a již integrované prostředí pro rychlé doučení modelu na cílové databázi.

Celková chyba kriteriální funkce je rozdělena mezi chybu regrese ohraničujících oblastí, chybu jejich klasifikace a chybu jistoty existence objektu („objectness loss“). Při výchozím nastavení YOLOv5 využívá pro určení chyby regrese ohraničujících oblastí „GIoU loss“, pro chybu klasifikace a jistoty pak „Binary cross entropy with logits loss“. Nastavení pro chybové funkce pak bylo ponecháno na výchozích hodnotách.

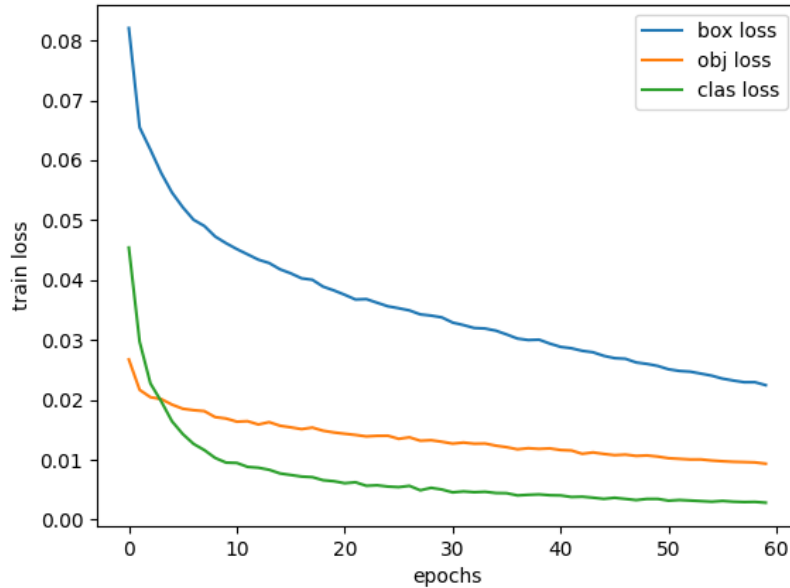
Kvůli malé velikosti cílového datasetu se podobně jako u předchozího detektoru YOLOv5 netrénovalo od náhodně inicializovaných vah, nýbrž se použily váhy již z natrénovaného detektoru na COCO datasetu. Před trénováním je ještě potřeba definovat velikost „anchor boxes“, která může mít zásadní vliv na efektivitu učení a finálních výsledcích. Součástí celého projektu s architekturou je však algoritmus, který pomocí k-means shlukování určí z výšky a šířky všech ohraničujících oblastí v trénovacím datasetu konečné přizpůsobené rozměry „anchor boxes“.

Při trénování detektoru je možné použít některé z implementovaných transformací snímků. Nevýhodou však oproti postupu popsaném u detektoru Faster-RCNN je menší kontrola nad tím, kdy se jaká transformace bude uplatňovat. Uživatel si zde může pouze navolit pravděpodobnosti, se kterou transformace proběhne. Při trénování na cílovém datasetu intrakraniálních krvácení se uplatňovali následující transformace: rotace v rozmezí (-30, 30) stupňů, převrácení podle horizontální osy, úprava jasu, odstínu a sytosti barvy. Poslední transformací snímků je tzv. mozaikování, kdy se 4 snímky nerovnoměrně spojí do jednoho finálního. Tato augmentace dat se poprvé objevila u předchozí verze YOLOv3 a podle autorů zvyšuje celkovou přesnost detektoru [43]. Tato transformace probíhala s pravděpodobností 1.



Obrázek 8.6: Transformace mozaikováním

Detektor byl trénován po 60 epoch, jako optimalizační algoritmus byl zvolen stochastický gradientní sestup s krokem učení 0,001. Celý průběh trénování lze sledovat pomocí knihovny Weights and Biases [59] v reálném čase na internetovém prohlížeči. Pokud chyba kritériální funkce stagnuje či dokonce stoupá, je na místě trénování ukončit a zabránit tak přeučení detektoru.



Obrázek 8.7: Hodnoty chybových funkcí během učení.

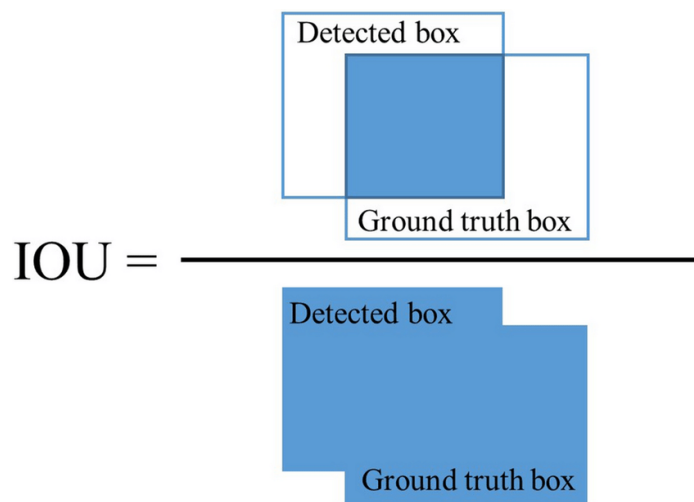
9. DETEKCE A VYHODNOCENÍ ÚSPĚŠNOSTI

V této kapitole jsou popsány metriky, kterými se posuzovala detekční schopnost architektur. Dále jsou zde uvedeny samotné dosažené výsledky pro každý z obou detektorů a v neposlední řadě také jejich diskuse.

9.1 Metriky pro vyhodnocení výsledků

V komunitě počítačového vidění se pravidelně pořádají soutěže, kde je úkolem natrénovat detektor objektů s cílem dosažení nejlepších výsledků na předloženém datasetu. Každá soutěž má zavedené vlastní metriky, které používá k ohodnocení daného detektoru. Tyto metriky jsou vesměs založené na podobném principu, ale nachází se zde malé odlišnosti v jejich interpretaci mezi jednotlivými soutěžemi. Mezi nejznámější soutěže, které v posledním desetiletí proběhly, patří COCO Detection Challenge (2019) [56], ImageNet Challenge (2017) [60] a Pacal VOC Challenge (2012) [61]. Asi nejaktuálnější proběhlou největší soutěží v době psaní této práce je LVIS 2021 Challenge [62].

Pro ohodnocení natrénovaných detektorů v této bakalářské práci jsou použity metriky inspirované těmi, jež se používají v soutěži COCO Detection Challenge. Před jejich samotným uvedením je nutné nejdříve definovat pojem „Intersection over Union“ (IOU), neboli podíl průniku a sjednocení. Jedná se o metriku popisující shodu mezi „ground truth“ ohraničující oblastí a predikcí modelu. Nejčastěji se používá jako prahová hodnota pro určení, zdali detekce byla úspěšná nebo nikoli [63].



Obrázek 9.1: IOU neboli průnik nad sjednocením predikované („detected box“) a skutečné („ground truth“) ohraničující oblasti [64].

Použité metriky jsou sensitivita („recall“), pozitivní prediktivní hodnota (PPV, „precision“) a tzv. „mean average precision“ (mAP). Pro výpočet uvedených metrik je potřeba zaznamenat počet správně pozitivních (TP, „true positive“), falešně pozitivních (FP, „false positive“) a falešně negativních (FN, „false negative“) detekcí. Sensitivita je vyjádřena následujícím vztahem:

$$sensitivita = \frac{TP}{TP + FN} \quad (9.1)$$

Pozitivní prediktivní hodnotu můžeme pak zapsat následující rovnicí:

$$PPV = \frac{TP}{TP + FP} \quad (9.2)$$

Jako skutečně pozitivní detekce jsou označeny takové, kde IOU nabývá hodnoty větší než stanovený práh. Obvykle se jako práh zavádí hodnota 0,5, ale je možné se setkat s menšími i vyššími hodnotami v závislosti na nárocích na detektor [40][41][48]. Falešně pozitivní jsou pak takové detekce, kde je IOU pod definovaným prahem. Jako falešně negativní se považují případy, kdy objekt nebyl detekován vůbec. V případě, že se skutečná a predikovaná klasifikační třída liší, je výsledek falešně pozitivní neohledně na hodnotu IOU. Pokud detektor predikuje více ohraničujících oblastí (všechny s IOU nad prahovou hodnotou) na pouze jeden objekt, jedná se pouze o 1 správně pozitivní detekci, zbylé detekce jsou falešně pozitivní. Jako práh IOU na základě rešerše pro tuto práci byla zvolena hodnota 0,5, která představuje střední cestu mezi striktností a přílišnou volností [40][41][48].

Často se dá setkat i s křivkou závislosti PPV na sensitivitě („Precision X Recall Curve“). Výhodou tohoto způsobu ohodnocení je informace o tom, jak se mění detekční schopnost sítě při postupném snižování skóre jistoty, se kterou jsou predikovány ohraničující oblasti objektů. Dobrý detektor si zachovává vysokou prediktivní pozitivní hodnotu, mezitím co sensitivita roste. Se snižováním skóre jistoty tedy sensitivita a PPV zůstávají pořád na vysokých hodnotách. Naopak špatný detektor musí predikovat více ohraničujících oblastí (nárůst FP, pokles PPV), aby nakonec hledaný objekt správně označil (nárůst TP, nárůst sensitivity). Z toho důvodu křivka většinou začne ve vysokých hodnotách PPV, které dále klesají mezitím co sensitivita roste.

„Mean average precision“ (mAP) představuje plochu pod křivkou závislosti pozitivní prediktivní hodnoty na sensitivitě interpolovanou 101 body [63][65].

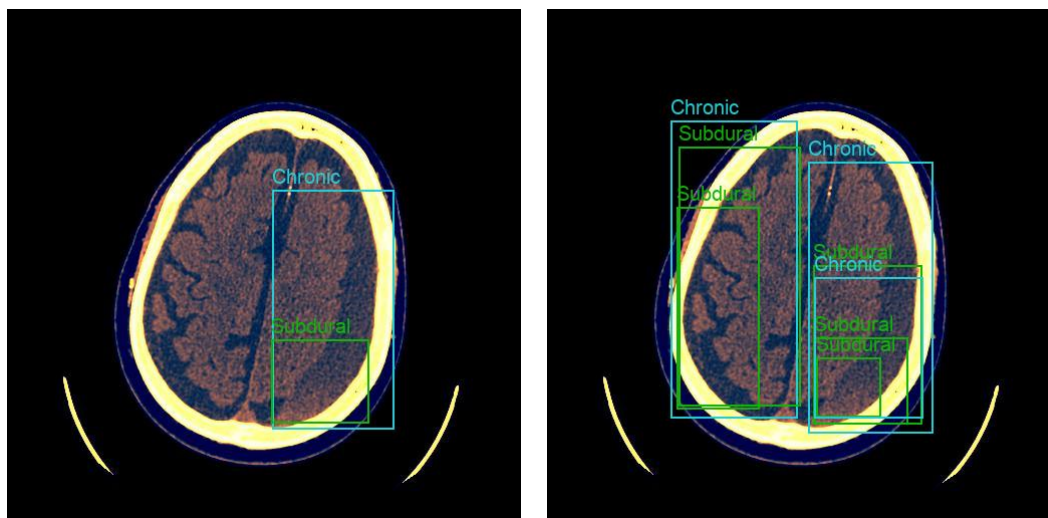
9.2 Realizace detekce a dosažené výsledky

Vyhodnocení úspěšnosti proběhlo na testovací části datasetu. Pro každý z detektorů byla vytvořena tabulka dosažených výsledků podle metrik uvedených v předcházející kapitole. Nacházejí se zde též příklady snímků s predikovanými krabicovými ohraničeními.

9.2.1 Faster-RCNN

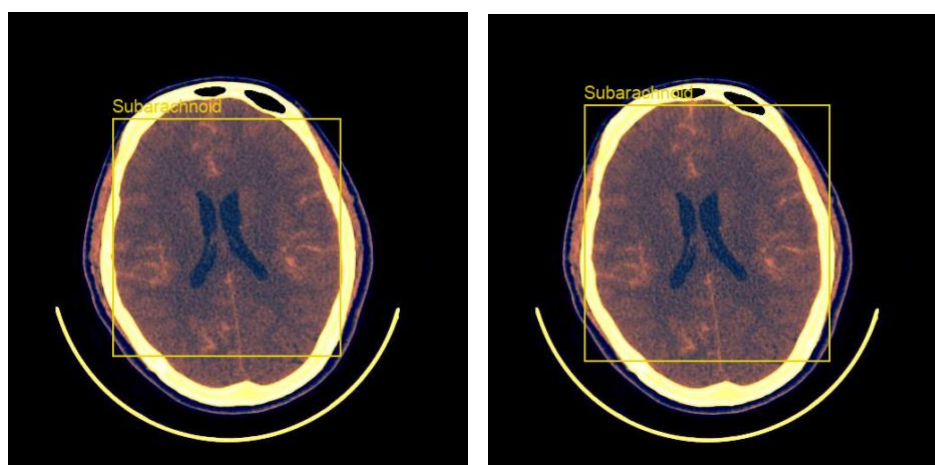
Pokud se detektor přepne z tréninkové fáze do fáze predikční, přivádí na svůj výstup pro každý vstupní snímek N vektorů o velikosti šest, kde N je počet predikovaných objektů a vektor jsou čtyři souřadnice definující levý horní a pravý spodní roh, predikovaná třída a jistota, se kterou danou ohraničující oblast určil.

Při ověřování detekční schopnosti sítě bylo zpozorováno, že přivádí na svůj výstup mnoho nadbytečných překrývajících se ohraničujících oblastí. Je tedy nutné její výstup podrobit nějaké formě postprocessingu. Nabízí se několik možností. Zprv je možné zavedení prahu pro jistotu, s jakou detektor ohraničující oblast určil, a následné odebrání všech detekcí se skórem jistoty menším, než je stanovený práh. Druhou variantou je použití techniky „Non-maximum suppression“. Poslední možností je zkombinovat tyto dvě metody.



Obrázek 9.2: Snímek se skutečnými ohraničujícími oblastmi (vlevo) a snímek s predikcemi (vpravo) Faster-RCNN přivádí na svůj výstup mnoho nadbytečných detekcí, které je nutné podrobit postprocessingu

Pro postprocessing byla nakonec vybrána kombinace dvou zmíněných přístupů, kdy se nejprve odstranily predikce s nízkým skórem a na zbylé ohraničující oblasti se aplikovala technika „Non-maximum suppression“. Poměrně dobrých výsledků se dosahovalo v případě, kdy se práh pro skóre nastavil na hodnotu 0,45 a práh IOU pro NMS na hodnotu 0,4. Příklad predikce s postprocessingem je zachycen na následujícím obrázku.



Obrázek 9.3: Detekce po postprocessingu. Skutečné ohraničující oblasti (vlevo) a predikce (vpravo).

Predikce detektoru Faster-RCNN po postprocessingu byly následně pomocí metrik popsaných v předchozí kapitole evaluovány. Úspěšnost detekce napříč všemi typy intrakraniálních krvácení je zaznamenána v tabulce níže.

Tabulka 9.1: Tabulka výsledků detektoru Faster-RCNN

| Typ krvácení | Počet ohraničujících oblastí | Sensitivita | PPV | mAP |
|----------------|------------------------------|--------------|--------------|--------------|
| IPH | 419 | 0,619 | 0,447 | 0,558 |
| EDH | 29 | 0,333 | 0,0345 | 0,178 |
| SDH | 1031 | 0,193 | 0,0686 | 0,11 |
| CHRONIC | 261 | 0,391 | 0,0307 | 0,201 |
| IVH | 210 | 0,629 | 0,414 | 0,557 |
| SAH | 705 | 0,147 | 0,131 | 0,12 |
| všechny | 2650 | 0,386 | 0,188 | 0,281 |

Z tabulky 9.1 vyplývá, že přesnost detekce je závislá na tom, o jaký typ krvácení se přesně jedná. Nejlépe si detektor z pohledu mAP vedl u intraparenchymálního a intraventrikulárního krvácení. Nejhůře Faster-RCNN detekovalo subdurální krvácení, přestože se u tohoto typu nacházel největší počet ohraničujících oblastí. Celkové mAP dosahovalo pro všechny objekty hodnoty 0,311.

9.2.2 Detektor YOLOv5

Podobně jako předchozí detektor YOLOv5 po přepnutí do predikční fáze přivádí na svůj výstup souřadnice predikovaných ohraničujících oblastí společně s klasifikační třídou a skórem jistoty. Mnoho z predikovaných oblastí má malé skóre nebo se překrývají, proto je stejně jako u předchozího detektoru potřeba protřídění. Nejprve se z detekcí odeberou všechny, které mají skóre jistoty menší než stanovený práh (0,45). V druhém kroku se provede technika „Non-maximum suppression“, která ponechá pouze ohraničující oblasti s největším skórem jistoty a odebere všechny ostatní s IOU hodnotou větší než stanovený práh (0,4). Tento algoritmus už je součástí samotné architektury (na rozdíl od Faster-RCNN), uživatel může pouze definovat požadované prahy. Výsledky detekce jsou uvedeny v následující tabulce.

Tabulka 9.2: Tabulka výsledků detektoru YOLOv5

| Typ krvácení | Počet ohraničujících oblastí | Sensitivita | PPV | mAP |
|----------------|------------------------------|--------------|--------------|--------------|
| IPH | 419 | 0,495 | 0,495 | 0,448 |
| EDH | 29 | 0 | 0 | 0 |
| SDH | 1031 | 0,339 | 0,216 | 0,181 |
| CHRONIC | 261 | 0,215 | 0,314 | 0,128 |
| IVH | 210 | 0,35 | 0,433 | 0,332 |
| SAH | 705 | 0,492 | 0,213 | 0,356 |
| všechny | 2650 | 0,378 | 0,273 | 0,310 |

Nejlepších výsledků dosáhnul detektor u intraparenchymálního a intraventrikulárního typu krvácení. U epidurálního typu se nepodařilo detekovat ani jedno krvácení. Sensitivita, pozitivní prediktivní hodnota a mAP mají nulovou hodnotu.

9.3 Diskuse dosažených výsledků

Jak můžeme vidět v příložených tabulkách [9.1](#) a [9.2](#), detektory mají při predikci některých typů krvácení podobné tendence, jinde však dosahují rozdílných výsledků. Sensitivita u Faster-RCNN nabývá větší hodnoty, ale nižší hodnota pozitivní prediktivní hodnoty způsobuje, že celkové mAP je v porovnání s YOLOv5 menší. Pro oba detektory však platí, že se jednotlivé metriky napříč všemi typy krvácení velmi lišily.

Tabulka [9.1](#) přináší hned několik zajímavostí. U epidurálního, subdurálního a chronického typu krvácení se vyskytují velmi malé hodnoty PPV. Síť tedy měla tendenci tyto typy krvácení často predikovat, avšak ohraničující oblasti měly buď nepřesnou pozici (IOU menší než práh) nebo špatnou klasifikační třídu, a byly tak zařazeny mezi falešně pozitivní detekce. V případě subdurálního krvácení jich bylo samo o sobě málo detekováno, což vzhledem k faktu, že se jedná o nejvíce zastoupený typ krvácení v celé databázi, a detektor se měl tak šanci dobře natrénovat, je zajímavá informace. Lepších výsledků Faster-RCNN dosahovalo u intraparenchymálního typu, detektor zde zachytil přes polovinu ze všech hemoragií, přičemž si zachoval i téměř poloviční PPV. Podobné výsledky je možné pozorovat u intraventrikulárního krvácení, ale rozdíl mezi sensitivitou a PPV hodnotou je zde oproti předešlému typu větší.

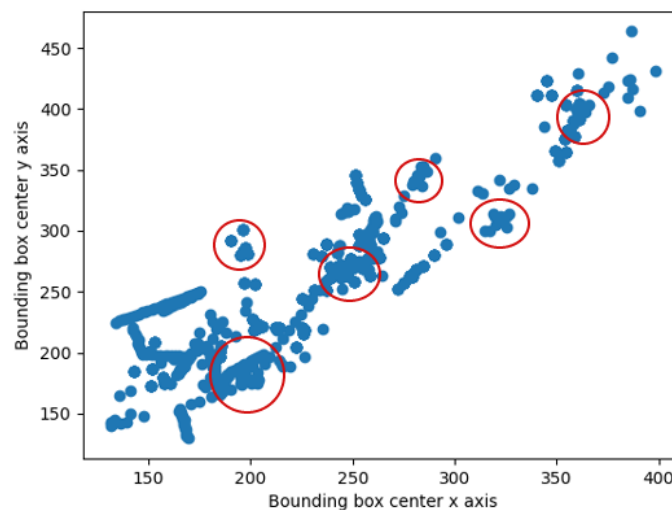
Detektor YOLOv5 podle výsledků v tabulce [9.2](#) si v porovnání s předešlým detektorem vedl o trochu lépe. Největším rozdílem zde je vyšší hodnota PPV napříč všemi typy krvácení vyjma epidurálního, které nebylo detekováno vůbec. Otázkou je, kterou z uvedených architektur by v kontextu detekce intrakraniálních krvácení dávalo větší smysl použít. Pravděpodobně Faster-RCNN za cenu velkého množství falešně pozitivních detekcí, které by sice radiologického pracovníka či doktora zbytečně zahltily, ale byla by zde menší pravděpodobnost, že krvácení unikne zpozorování úplně.

Jak už bylo nastíněno dříve, pozoruhodná je situace u subdurálního krvácení. V trénovací databázi se vyskytuje 2509 ohraničujících oblastí tohoto typu, což ho dělá druhým nejzastoupenějším krvácením. Přesto jeho detekce u obou architektur v porovnání s jinými typy krvácení nedosahuje zvláště dobrých výsledků. Za tímto problémem by mohla stát malá variabilita datasetu, kdy krvácení na několika po sobě jdoucích snímcích vypadají velmi podobně. Situace je znázorněná na následujícím obrázku.



Obrázek 9.4: Snímky se subdurálním krvácením jednoho pacienta. Tři velmi podobné hemoragie snižují variabilitu datasetu.

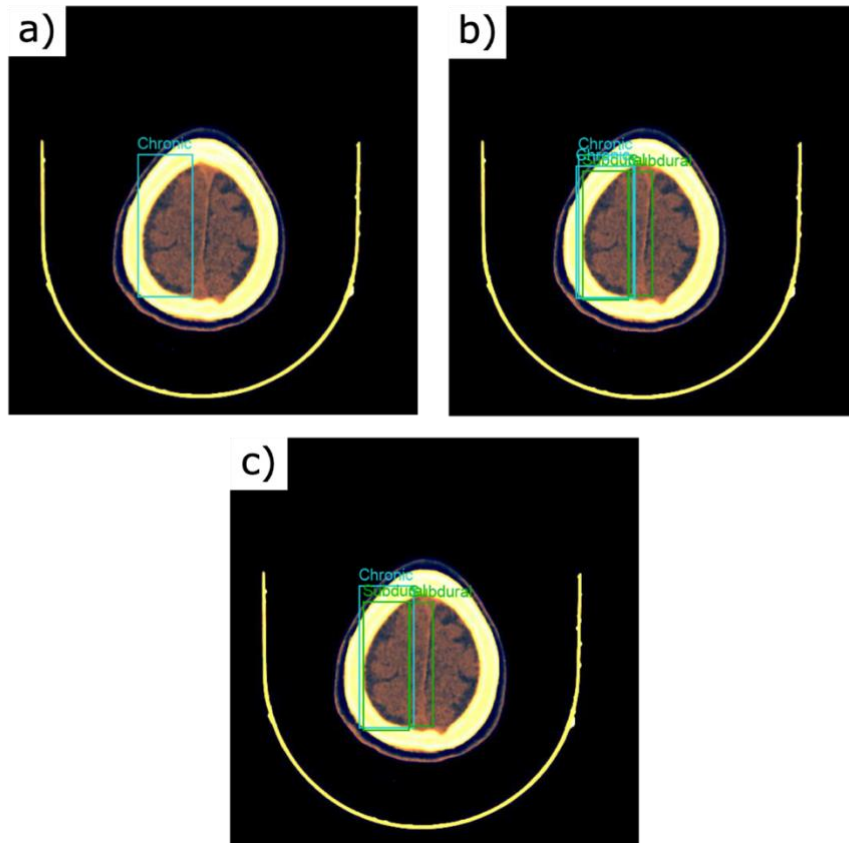
Tento jev je možné taky dobře znázornit pomocí bodového grafu, kde na osách jsou souřadnice středů ohraničujících oblastí hemoragií. Jednotlivé shluky jsou tvořeny krvácením zachyceným na několika po sobě jdoucích snímcích. Graf je na následujícím obrázku.



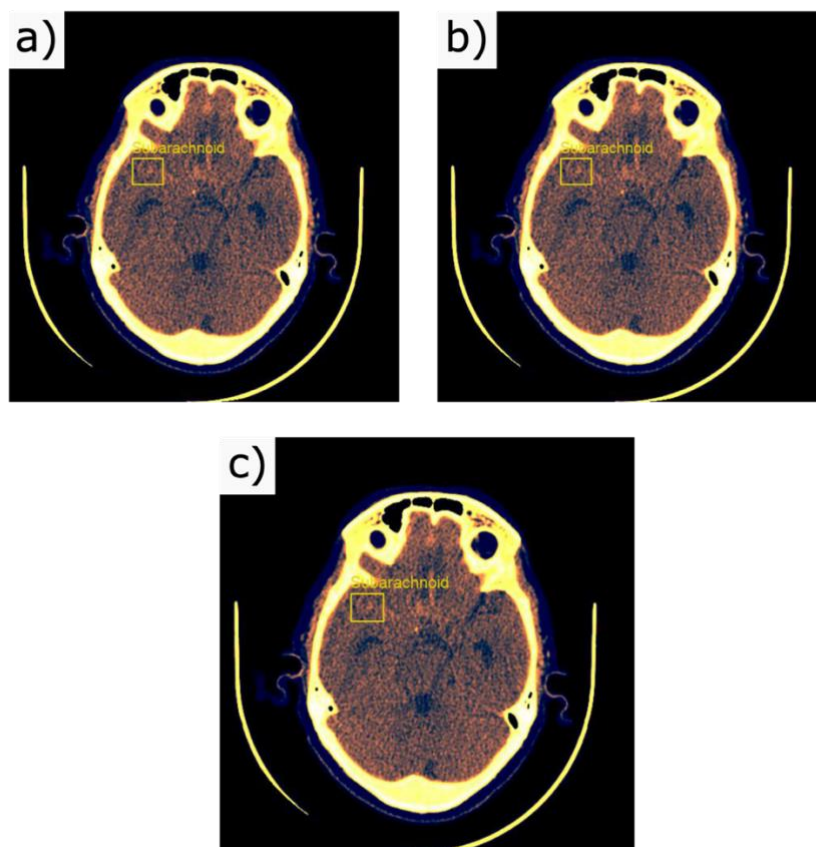
Obrázek 9.5: Bodový graf středových souřadnic ohraničujících oblastí subdurálního krvácení s ohraničenými shluky.

Popsaná situace by se dala v podstatě aplikovat na každý typ krvácení, což může mít vliv na nepříliš dobré výsledky i přes poměrně velkou databázi. U subdurálního krvácení může navíc detekci komplikovat poloha v blízkosti lebky nebo podél výchlípků tvrdé pleny. Tato skutečnost do budoucna komplikuje sběr dat s intrakraniálními hemoragiemi pro podobné účely. Databáze, na kterých by bylo možné efektivně natrénovat detektor, by teoreticky musela obsahovat protříděné snímky od tisíců ne-li desetitisíců pacientů. Samotné snímky by pak navíc musely být ručně anotovány kvalifikovaným personálem.

Na následujících obrázcích jsou ukázky detekcí obou architektur společně s krátkými komentáři.



Obrázek 9.6: Porovnání predikcí obou detektorů. a) originální snímek; b) Faster-RCNN; c) YOLOv5. YOLOv5 správně predikovalo chronické krvácení, ale navíc hned dvakrát falešně označilo pozadí jako subdurální krvácení. Faster-RCNN předpovědělo chronické krvácení hned dvakrát, které se v postprocessingu pravděpodobně kvůli malé hodnotě IOU a velkému skóre neodstranilo. Stejně jako předchozí detektor následně falešně označil pozadí jako subdurální krvácení.



Obrázek 9.7: Porovnání predikcí obou detektorů. a) originální snímek; b) Faster-RCNN; c) YOLOv5. Příklad povedené detekce. Oba detektory přesně lokalizovaly a správně klasifikovaly subarachnoidální krvácení

Pokud se vezmou v potaz všechny skutečnosti zmíněné v této kapitole, dosáhlo se poměrně úctyhodných výsledků. Vyzkoušelo se několik přístupů při řešení zadání a byly naučeny dva detektory, které nakonec dosahovaly srovnatelných výsledků. V oblasti této bakalářské práce je mnoho prostoru pro další testování a optimalizaci, které by samo sobě mohlo být předmětem několika dalších bakalářských prací. Kdybych v době odevzdávání této práce měl začít od začátku, pravděpodobně bych se zaměřil pouze na jeden detektor (spíše Faster-RCNN) a provedl detailnější analýzu parametrů na konečné výsledky. Vyzkoušel bych několik páteřních sítí, implementoval další transformační funkce, experimentoval s nastavením „Region Proposal Network“ či Fast-RCNN. Detektory objektů určitě mají na poli biomedicíny své uplatnění a byl zajímavý experiment je použít na data s intrakraniálním krvácením.

9.4 Možné využití detektoru v lékařské diagnostice

Podpůrné diagnostické systémy („Computer-aided diagnosis“, CAD) představují jednu z možných oblastí pro uplatnění algoritmů hlubokého učení. Jedná se o systémy, které mohou lékařskému personálu výrazně napomoci při vytváření konečného úsudku a ovlivnit tak celkovou diagnózu.

Nitrolebeční krvácení je závažným onemocněním, které vyžaduje rychlou a hlavně přesnou diagnózu. Vzniká tak velký tlak na doktory a radiologické pracovníky, kteří v této situaci musí postupovat neomylně. Algoritmus popsáný v této práci má poměrně velký potenciál stát se coby pomocníkem lékařského personálu při vyhodnocování snímků počítačové tomografie s podezřením na nitrolební krvácení. Největší překážkou však zůstává poměrně malá přesnost detektoru, která se odvíjí zejména od neexistence velké databáze snímků CT dat s jejich anotacemi, na kterých by se detektor dal efektivně naučit. V blízké budoucnosti je spíše pravděpodobnější použití klasifikačních hlubokých konvolučních sítí, které sice neoznačí místo s teoretickým výskytem krvácení ohraničujících oblastí, ale s vysokou přesností (viz kapitola 6) dokážou rozhodnout o tom, zda se na snímku krvácení vyskytuje, či nevyskytuje. Velkou výhodou dnešních „State-of-the-art“ detektorů je také rychlost, se kterou dokážou lokalizovat hledané objekty. Zakomponování detekčního algoritmu do nemocničního softwaru by tedy nijak znatelně neprodloužilo celkovou dobu diagnostiky.

ZÁVĚR

Tato práce se zabývá lokalizací intrakraniálního krvácení na CT snímcích hlavy a zkoumá možnosti uplatnění detektorů objektů na poli medicíny. V teoretické části práce je vypracována základní rešerše anatomie, patologie, konvolučních neuronových sítí a architektur samotných detektorů objektů. Tyto informace sloužily jako podklad pro samostatnou praktickou práci, která tvoří druhou část bakalářské práce.

Praktická část zahrnovala přípravu dat z dvou volně dostupných databází snímků hlav s intrakraniálními hemoragiemi a jejich anotacemi ve formě souřadnic ohraničujících oblastí. Tyto databáze sloužily k naučení jednofázového detektoru YOLOv5 a dvoufázového detektoru Faster-RCNN za účelem predikce hemoragií na nových dosud nespátných snímcích. V práci se také nachází rešerše již publikovaných prací zabývajících se problematikou klasifikace a lokalizace intrakraniálních hemoragií na CT snímcích, které sloužily jako zdroj inspirace pro návrh řešení a samotný postup implementace.

Přestože použité databáze obsahovaly více než 27 tisíc ohraničujících oblastí na zhruba 16 tisíc obrazech, kvůli malé variabilitě dat měly detektory tendenci se rychle přeučit. Bylo proto implementováno několik transformačních funkcí, které snímek před vstupem do detektoru při fázi učení upravili. Úspěšnost takto natrénovaných detektorů byla poté ohodnocena zvolenými metrikami.

Přesnost detekce se u obou detektorů napříč typy krvácení poměrně lišila, avšak bylo zde možné spatřit společné tendence, které detektory vykazovaly. Pro oba detektory platí, že s největší přesností detekovaly intraparenchymální a intraventrikulární typ. Nejhůře si pak vedly u subdurálního a epidurálního typu, kde detekce navíc mohla být zkomplikována výskytem krvácení okolo tvrdé pleny mozkové. Celkově YOLOv5 dosahovalo větší PPV, predikovalo tedy méně falešně pozitivních detekcí. mAP pro všechny typy krvácení mělo pro Faster-RCNN a YOLOv5 hodnotu 0,281 a 0,310 respektive.

Tato bakalářská práce dává skromný pohled na možnosti využití detektorů na poli biomedicíny, které mají do budoucna určitě velký potenciál. V případě detekce hemoragií ovšem představuje problém sběr a anotace dat, která musí být provedena odborníky v oboru. Každým rokem se publikují architektury detektorů, které dosahují stále lepších výsledků. Zveřejnění nové databáze CT snímků hlav s hemoragiemi by pak dalo nový impulz k bádání v tomto směru.

LITERATURA

- [1] MAJUMDAR, Arjun, Laura BRATTAIN, Brian TELFER, Chad FARRIS a Jonathan SCALERA. Detecting Intracranial Hemorrhage with Deep Learning. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* [online]. IEEE, 2018, 583-587 [cit. 2021-12-05]. ISBN 978-1-5386-3646-6. Dostupné z: doi:10.1109/EMBC.2018.8512336
- [2] HALALLI, Bhagirathi a Aziz MAKANDAR. Computer Aided Diagnosis - Medical Image Analysis Techniques. *Breast Imaging* [online]. InTech, 2018, 2018-01-17 [cit. 2021-12-05]. ISBN 978-953-51-3732-0. Dostupné z: doi:10.5772/intechopen.69792
- [3] NAŇKA, Ondřej a Miroslava ELIŠKOVÁ. *Přehled anatomie*. Třetí, doplněné a přepracované vydání. Praha: Galén, 2009, s. 46. ISBN 978-80-7492-206-0.
- [4] GRIM, Miloš a Rastislav DRUGA. *Základy anatomie: Obecná anatomie a pohybový systém*. Praha: Galén, 2001. ISBN 80-7262-111-2.
- [5] PÁČ, Libor a Ladislava HORÁČKOVÁ. *Anatomie pohybového systému člověka*. Brno: Masarykova Univerzita, 2011, s. 31-32. ISBN 978-80-87192-14-6.
- [6] DRUGA, Rastislav, Petr DUBOVÝ a Miloš GRIM. *Anatomie centrálního nervového systému*. Praha: Galén, 2011. ISBN 978-80-7262-706-6.
- [7] GRIM, Miloš a Rastislav DRUGA. *Základy anatomie: Centrální nervový systém*. Druhé, přepracované vydání. Praha: Galén, 2014. ISBN 978-80-7262-938-1.
- [8] CARTON, James, Richard DALY a Pramila RAMANI. *Clinical Pathology*. Oxford University Press, 2007. ISBN 978-0-19-856946-6.
- [9] KENNETH WALKER, H., W. DALLAS HALL a J. WILLIS HURST. *Clinical Methods: The History, Physical, and Laboratory Examinations*. Boston: Butterworths, 1990, Kapitola 55. ISBN 0-409-90077-X. Dostupné také z: <https://www.ncbi.nlm.nih.gov/books/NBK378/>

- [10] Cerebrovascular Disease. *American Association of Neurological Surgeons* [online]. [cit. 2021-11-05]. Dostupné z: <https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Cerebrovascular-Disease>
- [11] POVÝŠIL, Ctibor, Ivo ŠTEINER a Pavel DUŠEK. *Speciální patologie: III. díl*. Praha: Karolinum, 2003, s. 13-14. ISBN 80-7184-526-4.
- [12] KIDWELL, Chelsea S. a Max WINTERMARK. Imaging of intracranial haemorrhage. In: *The Lancet Neurology*. 2008, s. 256-267. ISSN 1474-4422. Dostupné z: doi:10.1016/S1474-4422(08)70041-3
- [13] CACERES, J. Alfredo a Joshua N. GOLDSTEIN. Intracranial Hemorrhage. *Emergency Medicine Clinics of North America* [online]. 2012, **30**(3), 771-794 [cit. 2022-01-03]. ISSN 07338627. Dostupné z: doi:10.1016/j.emc.2012.06.003
- [14] BABA, Yahya. Windowing (CT). *Radiopaedia.org* [online]. 2017 [cit. 2021-11-06]. Dostupné z: <https://doi.org/10.53347/rID-52108>
- [15] MURPHY, Andrew. CT head (an approach). *Radiopaedia.org* [online]. 2010 [cit. 2021-11-06]. Dostupné z: <https://doi.org/10.53347/rID-8998>
- [16] LUONG, David. CT head. *Radiopaedia.org* [online]. 2010 [cit. 2021-11-06]. Dostupné z: <https://doi.org/10.53347/rID-8996>
- [17] NAIDICH, Thomas P., Mauricio CASTILLO a Soonmee CHA. *Imaging of the Brain: Expert Radiology Series*. Elsevier Health Sciences, 2012, s. 389-390. ISBN 9781416050094.
- [18] PARIZEL, P., S. MAKKAT, E. VAN MIERT, J. VAN GOETHEM, L. VAN DEN HAUWE a A. DE SCHEPPER. Intracranial hemorrhage: principles of CT and MRI interpretation. *European Radiology* [online]. 2001, **11**(9), 1770-1783 [cit. 2022-01-03]. ISSN 0938-7994. Dostupné z: doi:10.1007/s003300000800
- [19] MEHROTRA, Kishan a Chilukuri K. MOHAN. *Elements of artificial neural network*. 2. vydání. USA: Massachusetts Institute of Technology, 1996. ISBN 0-262-13328-8.

- [20] JAIN, A.K., JIANCHANG MAO a K.M. MOHIUDDIN. Artificial neural networks: a tutorial. *Computer* [online]. **29**(3), 31-44 [cit. 2021-12-30]. ISSN 00189162. Dostupné z: doi:10.1109/2.485891
- [21] SHARMA, Siddharth, Simone SHARMA a Anidhya ATHAIYA. Activation functions in neural networks. *International Journal of Engineering Applied Sciences and Technology* [online]. 2020, 310-316 [cit. 2021-12-30]. ISSN 2455-2143. Dostupné z: <https://www.ijeast.com/papers/310-316,Tesma412,IJEAST.pdf>
- [22] XU, Bing a Naiyan WANG. *Empirical evaluation of rectified activations in convolutional network* [online]. [cit. 2021-12-30]. Dostupné z: <https://arxiv.org/pdf/1505.00853.pdf>
- [23] MAIND, Sonali B. a Priyanka WANKAR. Research Paper on Basic of Artificial Neural Network. *International Journal on Recent and Innovation Trends in Computing and Communication* [online]. 2014, 96-100 [cit. 2021-12-30]. ISSN 2321-8169. Dostupné z: <http://www.ijritcc.org>
- [24] RENDA, Alex, Jonathan FRANKLE a Michael CARBIN. *Comparing Rewinding and Fine-tuning in Neural Network Pruning* [online]. 2020 [cit. 2021-12-30]. Dostupné z: <https://arxiv.org/pdf/2003.02389.pdf>
- [25] JANOCHA, Katarzyna a Wojciech Marian CZARNECKI. *On Loss Functions for Deep Neural Networks in Classification* [online]. 2017 [cit. 2021-12-30]. Dostupné z: <https://arxiv.org/pdf/1702.05659.pdf>
- [26] QI, Jun, Jun DU, Sabato Marco SINISCALCHI, Xiaoli MA a Chin-Hui LEE. On Mean Absolute Error for Deep Neural Network Based Vector-to-Vector Regression. *IEEE Signal Processing Letters* [online]. 2020, **27**, 1485-1489 [cit. 2021-12-30]. ISSN 1070-9908. Dostupné z: doi:10.1109/LSP.2020.3016837
- [27] LEUNG, H. a S. HAYKIN. The complex backpropagation algorithm. *IEEE Transactions on Signal Processing* [online]. **39**(9), 2101-2104 [cit. 2021-12-30]. ISSN 1053587X. Dostupné z: doi:10.1109/78.134446
- [28] NIELSEN, M. A. Neural Networks and Deep Learning. *Determination Press* [online]. 2018 [cit. 2021-12-30]. Dostupné z: <http://neuralnetworksanddeeplearning.com/index.html>

- [29] MEHLIG, Bernhard. *Machine Learning with Neural Networks* [online]. Cambridge University Press, 2021-10-14 [cit. 2021-12-30]. Dostupné z: doi:10.1017/9781108860604
- [30] BOTTOU, Léon. Stochastic Gradient Descent Tricks. *Neural Networks: Tricks of the Trade* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, 2012, 421-436 [cit. 2021-12-30]. Lecture Notes in Computer Science. ISBN 978-3-642-35288-1. Dostupné z: doi:10.1007/978-3-642-35289-8_25
- [31] ALBAWI, Saad, Tareq Abed MOHAMMED a Saad AL-ZAWI. Understanding of a convolutional neural network. *2017 International Conference on Engineering and Technology (ICET)* [online]. IEEE, 2017, 2017, 1-6 [cit. 2021-12-22]. ISBN 978-1-5386-1949-0. Dostupné z: doi:10.1109/ICEngTechnol.2017.8308186
- [32] O'SHEA, Keiron a Ryan NASH. *An Introduction to Convolutional Neural Networks* [online]. 2015 [cit. 2021-12-22]. Dostupné z: <https://arxiv.org/abs/1511.08458>
- [33] YAMASHITA, Rikiya, Mizuho NISHIO, Richard Kinh Gian DO a Kaori TOGASHI. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging* [online]. 2018, **9**(4), 611-629 [cit. 2021-12-22]. ISSN 1869-4101. Dostupné z: doi:10.1007/s13244-018-0639-9
- [34] HINTON, Geoffrey E., Nitish SRIVASTAVA a Alex KRIZHEVSKY. *Improving neural networks by preventing co-adaptation of feature detectors* [online]. 2012 [cit. 2021-12-22]. Dostupné z: <https://arxiv.org/abs/1207.0580>
- [35] KRIZHEVSKY, Alex, Ilya SUTSKEVER a Geoffrey E. HINTON. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* [online]. 2012, 1097-1105 [cit. 2021-12-22]. Dostupné z: <https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>
- [36] SANTURKAR, Shibani, et al. How does batch normalization help optimization?. In: *Proceedings of the 32nd international conference on neural information processing systems*. 2018, 2488-2498 [cit. 2021-12-22]. Dostupné z:

<https://papers.nips.cc/paper/7515-how-doesbatch-normalization-help-optimization.pdf>

- [37] IOFFE, Sergey; SZEGEDY, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: *International conference on machine learning*. PMLR, 2015, 448-456 [cit. 2021-12-22]. Dostupné z: <https://arxiv.org/pdf/1502.03167v3.pdf>
- [38] SANTIAGO TELES DE MENEZES, Richardson, Rafael MARROCOS MAGALHAES a Helton MAIA. Object Recognition Using Convolutional Neural Networks. *Recent Trends in Artificial Neural Networks - from Training to Prediction* [online]. IntechOpen, 2020, 2020-3-4 [cit. 2021-12-10]. ISBN 978-1-78985-419-0. Dostupné z: doi:10.5772/intechopen.89726
- [39] SOVIANY, Petru a Radu Tudor IONESCU. Optimizing the Trade-off between Single-Stage and Two-Stage Object Detectors using Image Difficulty Prediction [online]. 2018 [cit. 2021-12-10]. Dostupné z: arXiv:1803.08707
- [40] REDMON, Joseph, Santosh DIVVALA, Ross GIRSHICK a Ali FARHADI. You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. IEEE, 2016, 2016, 779-788 [cit. 2021-12-10]. ISBN 978-1-4673-8851-1. Dostupné z: doi:10.1109/CVPR.2016.91
- [41] REN, Shaoqing, Kaiming HE, Ross GIRSHICK a Jian SUN. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online]. 2017, **39**(6), 1137-1149 [cit. 2021-12-10]. ISSN 0162-8828. Dostupné z: doi:10.1109/TPAMI.2016.2577031
- [42] REDMON, Joseph a Ali FARHADI. *YOLO9000: Better, Faster, Stronger* [online]. 2016 [cit. 2021-12-31]. Dostupné z: <https://arxiv.org/pdf/1612.08242.pdf>
- [43] REDMON, Joseph a Ali FARHADI. *YOLOv3: An Incremental Improvement* [online]. 2018 [cit. 2021-12-31]. Dostupné z: <https://arxiv.org/pdf/1804.02767.pdf>

- [44] BOCHKOVSKIY, Alexey, Chien-Yao WANG a Hong-Yuan Mark LIAO. *YOLOv4: Optimal Speed and Accuracy of Object Detection* [online]. 2020 [cit. 2021-12-31]. Dostupné z: <https://arxiv.org/pdf/2004.10934.pdf>
- [45] ULTRALYTICS. *Yolov5*. *GitHub* [online]. [cit. 2022-05-02]. Dostupné z: <https://github.com/ultralytics/yolov5>
- [46] WANG, Chien-Yao, Hong-Yuan MARK LIAO, Yueh-Hua WU, Ping-Yang CHEN, Jun-Wei HSIEH a I-Hau YEH. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* [online]. IEEE, 2020, 2020, 1571-1580 [cit. 2022-05-02]. ISBN 978-1-7281-9360-1. Dostupné z: doi:10.1109/CVPRW50498.2020.00203
- [47] LIU, Shu, Lu QI, Haifang QIN, Jianping SHI a Jiaya JIA. Path Aggregation Network for Instance Segmentation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* [online]. IEEE, 2018, 2018, 8759-8768 [cit. 2022-05-02]. ISBN 978-1-5386-6420-9. Dostupné z: doi:10.1109/CVPR.2018.00913
- [48] GIRSHICK, Ross. Fast R-CNN. *2015 IEEE International Conference on Computer Vision (ICCV)* [online]. IEEE, 2015, 2015, 1440-1448 [cit. 2022-04-22]. ISBN 978-1-4673-8391-2. Dostupné z: doi:10.1109/ICCV.2015.169
- [49] CHILAMKURTHY, Sasank, Rohit GHOSH, Swetha TANAMALA a Mustafa BIVIJI. *Development and Validation of Deep Learning Algorithms for Detection of Critical Findings in Head CT Scans*. 2018. Dostupné také z: <https://arxiv.org/pdf/1803.05854.pdf>
- [50] YE, H., F. GAO, Y. YIN a spol. Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. *European Radiology* [online]. 2019, 6191–6201 [cit. 2021-11-07]. ISSN 0938-7994. Dostupné z: doi:10.1007/s00330-019-06163-2
- [51] CHO, Junghwan, Ki-Su PARK, Manohar KARKI a spol. Improving Sensitivity on Identification and Delineation of Intracranial Hemorrhage Lesion Using Cascaded Deep Learning Models. *Journal of Digital Imaging* [online]. 2019, 450-461 [cit. 2021-11-07]. ISSN 1618-727X. Dostupné z: <https://link.springer.com/article/10.1007/s10278-018-00172-1>

- [52] ERTUĞRUL, Ömer Faruk a Muhammed Fatih AKIL. Detecting hemorrhage types and bounding box of hemorrhage by deep learning. *Biomedical Signal Processing and Control* [online]. 2022, **71** [cit. 2022-01-05]. ISSN 17468094. Dostupné z: doi:10.1016/j.bspc.2021.103085
- [53] PONTES REIS, Eduardo, Felipe NASCIMENTO a Mateus ARANHA. *Brain Hemorrhage Extended (BHX): Bounding box extrapolation from thick to thin slice CT images*. 2020. Dostupné také z: <https://doi.org/10.13026/9cft-hg92>
- [54] *Nvidia Corporation: CUDA Toolkit Documentation* [online]. [cit. 2022-05-25]. Dostupné z: <https://docs.nvidia.com/cuda/>
- [55] *PyTorch* [online]. [cit. 2022-05-25]. Dostupné z: <https://pytorch.org/>
- [56] LIN, Tsung-Yi, Michael MAIRE, Serge BELONGIE, James HAYS, Pietro PERONA, Deva RAMANAN, Piotr DOLLÁR a C. Lawrence ZITNICK. Microsoft COCO: Common Objects in Context. *Computer Vision – ECCV 2014* [online]. Cham: Springer International Publishing, 2014, 2014, 740-755 [cit. 2022-04-22]. Lecture Notes in Computer Science. ISBN 978-3-319-10601-4. Dostupné z: doi:10.1007/978-3-319-10602-1_48
- [57] HE, Kaiming, Xiangyu ZHANG, Shaoqing REN a Jian SUN. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. IEEE, 2016, 2016, 770-778 [cit. 2022-05-26]. ISBN 978-1-4673-8851-1. Dostupné z: doi:10.1109/CVPR.2016.90
- [58] LIN, Tsung-Yi, Piotr DOLLAR, Ross GIRSHICK, Kaiming HE, Bharath HARIHARAN a Serge BELONGIE. Feature Pyramid Networks for Object Detection. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. IEEE, 2017, 2017, 936-944 [cit. 2022-05-26]. ISBN 978-1-5386-0457-1. Dostupné z: doi:10.1109/CVPR.2017.106
- [59] *Weights & Biases* [online]. [cit. 2022-05-26]. Dostupné z: <https://wandb.ai/site>
- [60] DENG, Jia, Wei DONG, Richard SOCHER, Li-Jia LI, KAI LI a LI FEI-FEI. ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition* [online]. IEEE, 2009, 2009, 248-255 [cit. 2022-04-22]. ISBN 978-1-4244-3992-8. Dostupné z: doi:10.1109/CVPR.2009.5206848

- [61] EVERINGHAM, Mark, S. M. Ali ESLAMI, Luc VAN GOOL, Christopher K. I. WILLIAMS, John WINN a Andrew ZISSERMAN. The Pascal Visual Object Classes Challenge: A Retrospective. *International Journal of Computer Vision* [online]. 2015, **111**(1), 98-136 [cit. 2022-04-22]. ISSN 0920-5691. Dostupné z: doi:10.1007/s11263-014-0733-5
- [62] GUPTA, Agrim, Piotr DOLLÁR a Ross GIRSHICK. LVIS: A Dataset for Large Vocabulary Instance Segmentation. *Proceedings of the {IEEE} Conference on Computer Vision and Pattern Recognition* [online]. 2019 [cit. 2022-04-22]. Dostupné z: <https://arxiv.org/pdf/1908.03195.pdf>
- [63] PADILLA, Rafael, Wesley L. PASSOS, Thadeu L. B. DIAS, Sergio L. NETTO a Eduardo A. B. DA SILVA. A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. *Electronics* [online]. 2021, **10**(3) [cit. 2022-05-26]. ISSN 2079-9292. Dostupné z: doi:10.3390/electronics10030279
- [64] MAHDI, Fahad Parvez, Kota MOTOKI a Syoji KOBASHI. Optimization technique combined with deep learning method for teeth recognition in dental panoramic radiographs. *Scientific Reports* [online]. 2020, **10**(1) [cit. 2022-04-22]. ISSN 2045-2322. Dostupné z: doi:10.1038/s41598-020-75887-9
- [65] CARVALHO, Diogo V., Eduardo M. PEREIRA a Jaime S. CARDOSO. Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics* [online]. 2019, **8**(8) [cit. 2022-01-06]. ISSN 2079-9292. Dostupné z: doi:10.3390/electronics8080832

Příloha A - Obsah přiložených souborů

Kopecny_BP/

| | |
|------------------------|---------------------------------------------|
| — data/..... | složka s daty |
| — demo.csv | |
| — demo_file1.jpg | |
| — demo_file2.jpg | |
| — demo_file3.jpg | |
| — demo_file4.jpg | |
| — dcm2jpg/..... | složka se skripty pro přípravu surových dat |
| — config.py | |
| — data_split.py | |
| — pydicom_to_jpeg.py | |
| — requirements.txt | |
| — FasterRCNN/..... | implementace detektoru Faster-RCNN |
| — transforms/..... | transformační funkce |
| — my_transforms.py | |
| — transforms_utils.py | |
| — __init__.py | |
| — config.py | |
| — custom_utils.py | |
| — dataset.py | s |
| — model.py | |
| — model_evaluation.py | |
| — requirements.txt | |
| — train.py | |
| — YOLOv5/..... | implementace detektoru YOLOv5 |
| — cq500.yaml | |
| — params.yaml | |
| — yolo.ipynb | |
| — inference.ipynb..... | notebook pro porovnání predikcí detektorů |

