

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE

Brno, 2018

Matúš Bafnec



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

## FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

### ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

## AUTOMATICKÁ 3D SEGMENTACE OBRAZU MOZKU

AUTOMATIC 3D SEGMENTATION OF BRAIN IMAGES

### BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

### AUTOR PRÁCE

AUTHOR

**Matúš Bafrnec**

### VEDOUCÍ PRÁCE

SUPERVISOR

**Ing. Martin Kolařík**

**BRNO 2018**



# Bakalářská práce

bakalářský studijní obor **Teleinformatika**  
Ústav telekomunikací

**Student:** Matuš Bafnec

**ID:** 186025

**Ročník:** 3

**Akademický rok:** 2017/18

**NÁZEV TÉMATU:**

## Automatická 3D segmentace obrazu mozku

### POKYNY PRO VYPRACOVÁNÍ:

Nastudujte současné metody segmentace obrazu, které používají tzv. hluboké učení a popište jejich vlastnosti. Ze zkoumaných postupů vyberte vhodnou metodu na problém segmentace 3D snímků mozku snímaných pomocí magnetické rezonance. Výběr zdůvodněte a metodu otestujte pro použití na daný problém. Výsledky testování vhodně prezentujte.

### DOPORUČENÁ LITERATURA:

[1]LONG, Jonathan, Evan SHELHAMER a Trevor DARRELL, Fully convolutional networks for semantic segmentation [online]. [cit. 2017-09-13]. DOI: 10.1109/CVPR.2015.7298965. ISBN 10.1109/CVPR.2015.7298965. Dostupné z: <http://ieeexplore.ieee.org/document/7298965/>

[2]SCHMIDHUBER, Jürgen, Deep learning in neural networks: An overview [online]. [cit. 2017-09-13]. DOI: 10.1016/j.neunet.2014.09.003. ISBN 10.1016/j.neunet.2014.09.003. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0893608014002135>

**Termín zadání:** 5.2.2018

**Termín odevzdání:** 29.5.2018

**Vedoucí práce:** Ing. Martin Kolařík

**Konzultant:**

**prof. Ing. Jiří Mišurec, CSc.**  
*předseda oborové rady*

### UPOZORNĚNÍ:

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

## ABSTRAKT

Táto bakalárska práca popisuje návrh a implementáciu systému na automatickú 3D segmentáciu mozgu založeného na konvolučných neurónových sieťach. Prvá časť práce je venovaná krátkej histórii neurónových sietí a teoretickému popisu funkcionality konvolučných neurónových sietí. Predstavuje rýchly úvod do problematiky a poskytuje teoretický základ pre pochopenie a vytvorenie samostatného systému. Popísané sú jednotlivé vrstvy neurónovej siete, princípy ich fungovania a vzájomnej náväznosti. Druhá časť práce sa zaoberá analýzou problému, návrhom konkrétneho riešenia a porovnaním neurónových sietí s inými riešeniami. Výsledkom magnetickej rezonancie hlavy je séria čierno-bielych snímok, ktoré reprezentujú 3D sken. Úlohou je označiť mozog a tým odstrániť nepotrebné informácie vo forme okolitých tkanív. Výsledný obraz mozgu môže byť využitý pri volumetrii alebo diagnostike neurodegeneratívnych chorôb. Výhodou neurónových sietí v porovnaní s deterministickými systémami je ich flexibilita. Umožňujú prispôbienie na iné typy segmentačných problémom len zmenou množiny tréningových dát, bez nutnosti zásahu do architektúry. Jedným zo systémov vykonávajúcich plne automatickú 3D segmentáciu sa nazýva U-Net – jeho názov vychádza z podobnosti architektúry s písmenom U. Prezentované boli tri reálne riešenia, prvotný návrh U-Netu, rozšírený U-Net a rekurentný U-Net. Prvá verzia U-Netu bola pamäťovo veľmi náročná, vyžadovala tréningovanie na procesore miesto grafickej karty a neumožňovala spracovanie dát v plnom rozlíšení. Rozšírený U-Net tieto problémy vyriešil načítavaním dát v prekrývajúcich sa skupinách troch snímok. Okrem možnosti tréningovania siete na grafickej karte s následným skrátením doby učenia, bola zvýšená aj presnosť pridaním ďalších interkonekcií do vnútornej architektúry siete. Posledná verzia, rekurentný U-Net, sa zameriava na optimalizáciu rozšíreného U-Netu založenej na princípe znovupoužitia existujúcich úrovní siete. To pri-náša zníženie časovej a hardwarovej náročnosti. Bol znížený počet parametrov siete na menej ako 20%, bez akéhokoľvek zvýšenia tohto počtu v prípade pridania ďalších úrovní. Takáto sieť patrí k jednej z prvých rekurentných sietí použitých na problém 3D segmentácie a poskytuje tak základ pre ďalší výskum. Posledná časť práce sa zameriava na vyhodnotenie výsledkov a porovnanie presnosti, rýchlosti a náročnosti jednotlivých sietí. Porovnávaná je aj presnosť ľudskej a strojovej segmentácie. Rozšírený a rekurentný U-Net dokázali poraziť svojho ľudského oponenta, čím by v prípade reálneho nasadenia mohli ušetriť množstvo lekárskeho času a predísť ľudským chybám. Výsledkom práce je teoretický základ poskytujúci úvod do problematiky konvolučných neurónových sietí a segmentácie, plne funkčné systémy na automatickú 3D segmentáciu a základ pre ďalší výskum v oblasti rekurentných sietí.

## KLÚČOVÉ SLOVÁ

3D, CNN, hlboké učenie, konvolučná neurónová sieť, mozog, MRI, obraz, rekurentný, segmentácia, U-Net



## ABSTRACT

This bachelor thesis describes the design and implementation of the system for automatic 3D segmentation of a brain based on convolutional neural networks. The first part is dedicated to a brief history of neural networks and a theoretical description of the functionality of convolutional neural networks. It represents a fast introduction to the problematics and provides theoretical basics needed for the understanding and creation of the system. Individual layers of the neural network and principles of their functionality and mutual relations are also described in this part. The second part of the thesis is about problem analysis, designing of a solution and a comparison between neural networks and other solutions. The result of a magnetic resonance imaging of the head is a series of black-and-white images representing a 3D scan. The task is to tag a brain and to remove unnecessary information in the form of surrounding tissues. The final image of the brain can be utilized in a volumetry or during a diagnostic of neurodegenerative diseases. The advantage of neural networks in comparison with deterministic systems is their flexibility. They allow an adaptation to other segmentation problems just by changing the training dataset, without a need of changes in the architecture. One of the systems performing fully automatic 3D segmentation is called U-Net – its name comes from the similarity of the architecture with the letter U. Three real solutions, the first implementation of U-Net, extended U-Net and recurrent U-Net were presented. The first version of U-Net has been very memory-demanding, it required a training on a processor instead of a graphic card and has not allowed data processing in full resolution. The extended U-Net has resolved these problems by loading data in overlaying series of three images. In addition to the possibility of a training on a graphic card with related decrease in learning time, the accuracy was increased by adding interconnections to the internal architecture of the network. The last version, recurrent U-Net, aims for the optimization of extended U-Net based on the reuse of existing levels. This brings a decrease in a time and resource difficulty. The number of parameters of the network was lowered to less than 20%, without any increase in case of further level addition. This network is one of first recurrent networks used on the problem of 3D segmentation and provides a foundation to further research. The last part focuses on the evaluation of results and the comparison of accuracy, speed and requirements between particular networks. The accuracy of human and machine segmentation is also compared. The extended and recurrent U-Net have surpassed their human opponent, which in real case could save a lot of doctors time and prevent human mistakes. The result of this work is a theoretical basis providing an introduction to the problematics of convolutional neural networks and segmentation, fully working systems for automatic 3D segmentation and the foundation for further research in the field of recurrent networks.

## KEYWORDS

3D, CNN, deep learning, convolutional neural network, brain, MRI, image, recurrent, segmentation, U-Net

BAFRNEC, Matúš. *Automatická 3D segmentace obrazu mozku*. Brno, 2018, 54 s. Bakalárska práca. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedúci práce: Ing. Martin Kolařík

## VYHLÁSENIE

Vyhlasujem, že som svoju bakalársku prácu na tému „Automatická 3D segmentace obrazu mozku“ vypracoval(a) samostatne pod vedením vedúceho bakalárskej práce, využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

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

Brno .....

.....

podpis autora(-ky)

## POĎAKOVANIE

Rád by som poďakoval vedúcemu bakalárskej práce pánovi Ing. Martinovi Kolaříkovi, za odborné vedenie, konzultácie, trpezlivosť a podnetné návrhy k práci.

Brno .....

.....

podpis autora(-ky)



Faculty of Electrical Engineering  
and Communication  
Brno University of Technology  
Purkynova 118, CZ-61200 Brno  
Czech Republic  
<http://www.six.feec.vutbr.cz>

## POĎAKOVANIE

Výzkum popsaný v tejto bakalárskej práci bol realizovaný v laboratóriách podporených projektom SIX; registračné číslo CZ.1.05/2.1.00/03.0072, operačný program Výzkum a vývoj pro inovace.

Brno .....

.....  
podpis autora(-ky)



EVROPSKÁ UNIE  
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ  
INVESTICE DO VAŠÍ BUDOUCNOSTI



# OBSAH

Úvod	12
<b>1 Konvolučné neurónové siete</b>	<b>14</b>
1.1 Úvod do všeobecných neurónových sietí . . . . .	14
1.1.1 Biologická podobnosť . . . . .	14
1.1.2 História . . . . .	15
1.1.3 Princíp fungovania . . . . .	15
1.1.4 Moderné aplikácie neurónových sietí . . . . .	17
1.2 Architektúra siete . . . . .	18
1.2.1 Konvolučná vrstva . . . . .	20
1.2.2 Združovacia vrstva . . . . .	21
1.2.3 Plne prepojená vrstva . . . . .	22
1.2.4 Aktivačné funkcie . . . . .	24
1.3 Hyperparametre neurónovej siete . . . . .	25
1.4 Segmentácia obrazu pomocou CNN . . . . .	26
<b>2 Analýza problému a implementácia riešenia</b>	<b>28</b>
2.1 Analýza problému . . . . .	28
2.1.1 Magnetická rezonancia . . . . .	28
2.1.2 Popis dát . . . . .	28
2.2 Súčasnú metódy segmentácie obrazu . . . . .	30
2.2.1 Metódy založené na znalosti obrazu . . . . .	30
2.2.2 Metódy založené na neurónových sietach . . . . .	31
2.3 U-Net . . . . .	32
2.3.1 Architektúra a implementácia . . . . .	32
2.3.2 Behové prostredie a požiadavky . . . . .	34
2.4 Rozšírený U-Net . . . . .	35
2.4.1 Predspracovanie a koncept . . . . .	35
2.4.2 Architektúra . . . . .	35
2.5 Rekurentný U-Net . . . . .	36
<b>3 Vyhodnotenie výsledkov</b>	<b>38</b>
3.1 Presnosť výsledkov . . . . .	38
3.1.1 U-Net . . . . .	40
3.1.2 Rozšírený U-Net . . . . .	40
3.1.3 Rekurentný U-Net . . . . .	41
3.2 Náročnosť spracovania údajov . . . . .	42

3.3 Uplatnenie v praxi . . . . .	44
<b>4 Záver</b>	<b>45</b>
<b>Literatúra</b>	<b>47</b>
<b>Zoznam symbolov, veličín a skratiek</b>	<b>51</b>
<b>Zoznam príloh</b>	<b>52</b>
<b>A Obsah priloženého disku</b>	<b>53</b>

# ZOZNAM OBRÁZKOV

1.1	Biologický model neurónu. . . . .	14
1.2	Model umelého neurónu. . . . .	16
1.3	Dopredná neurónová sieť. . . . .	16
1.4	Rozpoznanie symbolov v obrázkoch pomocou konvolučnej neurónovej siete. . . . .	18
1.5	Obrázok ako dvojrozmerné pole hodnôt. . . . .	19
1.6	Príznyaky symbolu X. . . . .	19
1.7	Princíp 2D konvolúcie. . . . .	20
1.8	Výsledok 2D konvolúcie vstupného obrázka a príznaku. . . . .	21
1.9	Výsledok konvolúcie so všetkými príznakmi. . . . .	21
1.10	Výsledok združovania. . . . .	22
1.11	Výsledok hlasovania. . . . .	23
1.12	Nastavovanie váh. . . . .	24
1.13	Obrázok po aplikácii ReLU. . . . .	24
1.14	Segmentácia obrazu. . . . .	26
2.1	Snímka z magnetickej rezonancie mozgu. . . . .	29
2.2	Referenčná maska mozgu snímky z obr. 2.1. . . . .	29
2.3	Zapojenie U-Netu. . . . .	33
2.4	Spracovanie vstupných dát. . . . .	36
2.5	Architektúra rozšíreného U-Netu. . . . .	37
2.6	Architektúra rekurentného U-Netu. . . . .	37
3.1	Výstupná maska pred a po upravení rozhodovacej úrovne. . . . .	38

## ZOZNAM TABULIEK

3.1	Presnosť ľudskej segmentácie. . . . .	40
3.2	Porovnanie presnosti človeka a U-Netu. . . . .	40
3.3	Porovnanie presnosti človeka a rozšíreného U-Netu. . . . .	41
3.4	Porovnanie presnosti rozšíreného a pôvodného U-Netu. . . . .	41
3.5	Porovnanie presnosti človeka a rekurentného U-Netu. . . . .	41
3.6	Relatívne zlepšenie presnosti rekurentného U-Netu oproti pôvodnému a rozšírenému U-Netu. . . . .	42
3.7	Porovnanie všetkých sietí. . . . .	42
3.8	Porovnanie presnosti segmentácie všetkých sietí s človekom podľa Jac- card indexu. . . . .	42
3.9	Počet parametrov v závislosti od počtu úrovní rozšíreného U-Netu. .	43
3.10	Počet parametrov v závislosti od počtu úrovní rekurentného U-Netu.	44



# ÚVOD

V súčasnej dobe sa s umelou inteligenciou a neurónovými sieťami stretávame čoraz viac. Prudký nárast výkonu počítačov za posledné roky umožňuje navrhovať programy, ktoré majú schopnosť učiť sa, rovnako ako ľudia. To prináša nové možnosti v oblastiach dátovej analýzy, navigácie, riadenia a spracovávania signálu a obrazu.

Práve spracovanie obrazu je téma, ktorou sa táto práca zaoberá. Magnetická rezonancia zohráva v medicíne veľmi dôležitú úlohu. Umožňuje lekárom nahliadnúť do ľudského tela bez nutnosti invazívneho chirurgického zásahu. Výsledný obraz pozostáva z veľkého množstva tkaniva, ktoré je pre vyhodnotenie výsledkov často len rušivým faktorom. Konkrétnym prípadom môže byť 3D magnetická rezonancia hlavy, pri ktorej tvorí dôležité dáta len jedna časť – mozog.

Práca sa zaoberá segmentáciou (vyčlenenie určitej časti) obrazu pomocou neurónových sietí. Cieľom je navrhnúť systém, ktorý na snímkach magnetickej rezonancie označí a oddelí mozog. Tento proces umožňuje následnú vizualizáciu a volumetriu (počítanie objemu) mozgu, ktorá je využívaná pri diagnostike neurodegeneratívnych chorôb.

K spracovaniu obrazu neurónovými sieťami sa využívajú prevažne konvolučné neurónové siete. Tie pozostávajú z niekoľkých vrstiev – konvolučná a združovacia – zabezpečujúce spracovanie obrazu a plne prepojená vrstva zabezpečujúca logiku a rozhodovanie. Počas celej práce boli skúmané tri systémy, pričom každý ďalší priniesol určité zlepšenia. Prvý navrhnutý systém, tzv. U-Net nedokázal spracovať dáta v plnom rozlíšení, preto bol navrhnutý rozšírený U-Net, ktorý tento problém odstránil a vďaka pridaniu interkonekcií do vnútra siete zvýšil aj presnosť segmentácie. Posledný systém je rekurentný U-Net, ktorý znižuje pamäťovú náročnosť rozšíreného U-Netu a tým znižuje hardwarovú aj časovú náročnosť.

Hlavným prínosom práce je popis princípu činnosti neurónových sietí určených na spracovanie obrazu – konvolučných neurónových sietí, navrhnutý systém a vytvorenie základu pre ďalší výskum rekurentných neurónových sietí, ktoré výrazne redukujú pamäťové nároky. Práca takisto porovnáva neurónové siete s tradičnými spôsobmi spracovania obrazu a s ľudským spracovaním. Vyzdvihuje výhody konvolučných neurónových sietí, medzi ktoré patrí vysoká adaptabilita na zadaný problém, možnosť prispôsobenia na iný typ zadania bez nutnosti zmeny architektúry a hlavne schopnosť učiť sa. Nezabúda sa ani na nevýhody a obmedzenia tejto technológie, medzi ktoré môžeme zaradiť vysokú hardwarovú náročnosť a nutnosť dostatočného množstva tréningových dát. Výsledkom je plne automatický systém na 3D segmentáciu mozgu zo snímkov magnetickej rezonancie hlavy. V prípade reálneho nasadenia by systém ušetril čas lekárov a zamedzil by ľudským chybám, ktoré sú bohužiaľ bežnou súčasťou lekárskej praxe.

Práca je rozdelená do troch častí. Prvá z nich zahŕňa úvod do problematiky neurónových sietí a jednoduchou formou popisuje princíp funkčnosti konvolučných neurónových sietí. V druhej časti je analyzovaný problém segmentácie mozgu spolu s popisom a návrhom implementácie niekoľkých riešení, z ktorých každé využíva iný prístup k problému. Posledná kapitola je venovaná prezentácii a vyhodnoteniu dosiahnutých výsledkov.

# 1 KONVOLUČNÉ NEURÓNOVÉ SIETE

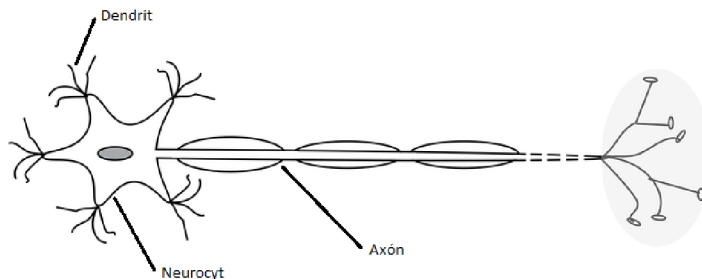
Konvolučná neurónová sieť – *Convolutional Neural Network* (CNN) je trieda hlbkej doprednej umelej neurónovej siete, ktorá nachádza široké uplatnenie v strojovom spracovaní obrazu pomocou hlbokého učenia, tzv. *deep learningu*. Prvé použitie konvolučných neurónových sietí sa datuje do roku 1980, kedy Francúz Yann LeCun ukázal svetu LeNet-5, ktorá bola použitá na rozpoznanie ručne písaných číslíc. Od tej doby prešli CNN dlhým vývojom, ktorý zaznamenal úspech v roku 2012 na súťaži *ImageNet Large Scale Visual Recognition Challenge* (súťaž zameraná na počítačové rozpoznávanie obrazu), kedy sa naplno ukázala výkonnosť a potenciál týchto sietí. Táto udalosť spôsobila enormný nárast záujmu o CNN, ktoré sú v dnešnej dobe lídrom na poli spracovania obrazu pomocou neurónových sietí [1, 2].

## 1.1 Úvod do všeobecných neurónových sietí

K chápaniu konvolučných neurónových sietí je nutné dôkladné pochopenie princípov fungovania základných stavebných jednotiek týchto sietí – umelých neurónov, ich biologickej podobnosti s mozgom a výhod, ktoré prinášajú v porovnaní s tradičnými počítačovými riešeniami.

### 1.1.1 Biologická podobnosť

Vývoj umelých neurónových sietí je založený na snahe strojovo napodobniť ľudský mozog, ktorého základnou stavebnou jednotkou je **neurón** (obr. 1.1). Centrálna nervová sústava – časť mozgu zodpovedná za spracovanie vnemov, pamäť, učenie, ale aj celé vedomie človeka – je tvorená viac ako piatimi miliardami neurónov. Najdôležitejšie časti biologického neurónu sú **neurocyt** (samotné telo neurónu), **axón** (alebo **neurit**, dlhý výbežok) a **dendrity** (krátke výbežky) [3].



Obr. 1.1: Biologický model neurónu.

Funkčnou jednotkou mozgu sú *synapsie* (nervové spojenia), ktoré umožňujú presun signálu medzi neurónmi. Denrity, ktoré okolo neurónu vytvárajú stromovú štruktúru slúžia ako prijímače nervových vzruchov, ktoré neurón po spracovaní pošle ďalej pomocou axónu. Tieto spojenia vytvárajú v mozgu obrovskú štruktúru, ktorá je vysoko modulárna vďaka možnosti vytvárania a zanikania synapsí. Tento proces sa nazýva učenie [3].

### 1.1.2 História

Snaha vytvoriť stroje, ktoré by boli schopné učiť sa a myslieť siaha do ďalekej histórie [4]. Medzi dôležité mílniky patrí rok 1943, kedy sa neurobiológ Warren S. McCulloch a štatistik Walter Pitts vytvorili model umelého neurónu, čím položili základy novej vednej disciplíny – *umelým neurónovým sieťam*. Postupný rozvoj v tejto oblasti zaznamenal prelom v roku 1958, kedy Franck Rosenblatt publikoval prvý algoritmus učenia pre neurónové siete s dopredným šírením signálu – túto sieť nazval **Perceptron** [5].

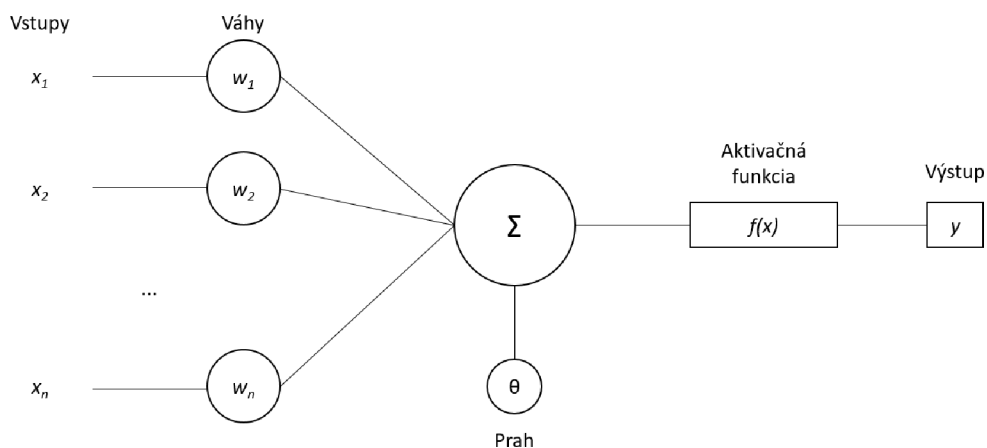
O 11 rokov neskôr však dvaja známi a vplyvní matematici, Marvin L. Minski a Seymour S., zverejnili kvalitne spracovanú publikáciu, ktorá významne kritizovala niektoré vlastnosti dobových neurónových sietí, čo malo za následok dlhodobý útlm rozvoja v tejto oblasti. Zvrat nastal až po takmer 20-tich rokoch, kedy sa niekoľkým vedcom podarilo znovu vzbudiť záujem o neurónové siete prácou, ktorá elegantne vyriešila problémy predošlých sietí. Tým odštartovala éra učiacich sa strojov, ktoré vďaka neustálemu nárastu výkonu a rozvoju v tejto oblasti napredujú míľovými krokmi [5].

### 1.1.3 Princíp fungovania

Základnou funkciou neurónu (či už biologického alebo umelého) je na základe signálov ( $x_1, x_2, \dots$ ) a ich jednotlivých váh ( $w_1, w_2, \dots$ ) – ich dôležitosť – odoslať výstup  $y$  vo forme aktivačnej funkcie  $f(x)$ . Umelý neurón (obr. 1.2) má však oproti jeho biologickej verzii (obr. 1.1) definovaný prah  $\theta$ , ktorého úlohou je aktivovať neurón až po prekonaní tejto hodnoty. Fungovanie umelého neurónu je možné charakterizovať vzťahom 1.1:

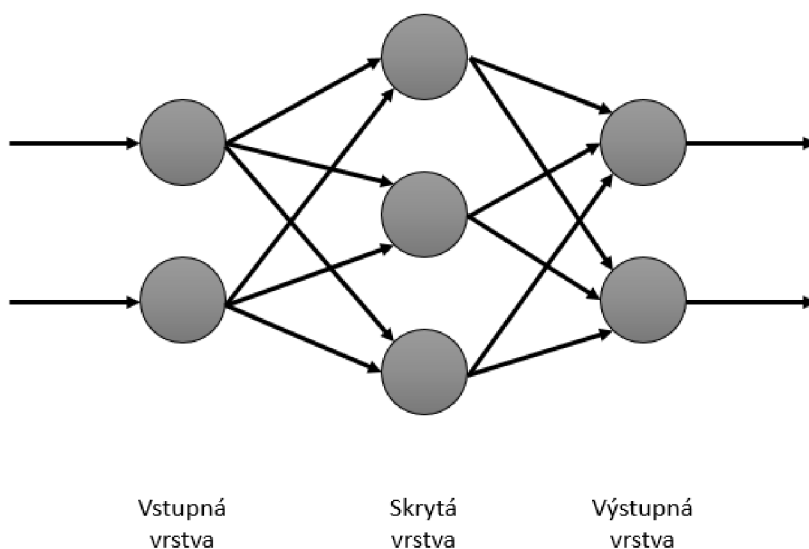
$$y = f\left(\sum_{i=1}^n x_i w_i + \theta\right) \quad (1.1)$$

Samotný neurón však neumožňuje vykonávať žiadnu zložitejšiu funkciu. Ako aj v mozgu, je len základnou stavebnou jednotkou. Funkčnou jednotkou sú synapsie – spojenia s ďalšími neurónmi, ktoré v závislosti od zložitosti a počtu vrstiev (skupina



Obr. 1.2: Model umelého neurónu.

neurónov na rovnakej úrovni signálového toku) vedia vykonávať rôzne funkcie. Najjednoduchším príkladom sú dopredné neurónové siete (obr. 1.3), v ktorých signály prechádzajú zo vstupnej, postupne cez všetky skryté, až do výstupnej vrstvy [4].



Obr. 1.3: Dopredná neurónová sieť.

Výnimočnosť týchto signál-spracúvajúcich spojení tkvie v tom, že sa v priebehu času menia – **učia sa**. Predstavme si dieťa, ktoré sa učí rozprávať. Zvyčajne to funguje tak, že sa dieťaťu ukáže písmeno (vstup) a vysloví sa (požadovaný výstup). Dieťa sa to následne pokúša zopakovať, pričom priebežne mení spôsob, akým dané písmeno vyslovuje, až kým ho nedokáže dostatočne kvalitne zopakovať. Rovnakým

spôsobom fungujú aj neurónové siete, je zadaný vstup, požadovaný výstup a sieť upravuje jednotlivé váhy až do doby, kým nie je jej výstup zhodný s požadovaným výstupom – sieť sa učí. Tento spôsob učenia je hojne využívaný vo väčšine neurónových sietí a nazýva sa algoritmus spätného šírenia chyby (angl. BPE – *Back-Propagation of Error*). Po fáze učenia je sieť pripravená riešiť problémy, pre ktoré bola navrhnutá [4, 5].

Pri použití programovania ako prostriedku pre počítačové spracovanie informácií je nutné, aby najskôr programátor daný problém dôkladne analyzoval a následne navrhol konkrétne riešenie. V prípade, že sa zadanie zmení, je nutné absolvovať celý tento proces odznovu, čo je v mnohých prípadoch nereálne. Tu sa nám ukazuje jedna z hlavných predností neurónových sietí – v prípade zmeny zadania nie je nutné prerábať celú sieť, stačí jej len poskytnúť nové dáta, na základe ktorých sa sama uspokojí k riešeniu nového problému. Táto vlastnosť neurónových sietí môže vyvolávať obavy z bezpečnosti tejto technológie, ktorá už v mnohých prípadoch prekonala najlepších ľudí vo svojom obore.

#### 1.1.4 Moderné aplikácie neurónových sietí

Od rozpoznávania číslíc z minulého storočia výskum a technika výrazne pokročili a umelé neurónové siete našli uplatnenie takmer v každom odvetví technologického priemyslu – autonómne vozidlá, webové vyhľadávače, optimalizačné úlohy, predikcie trhu na základe historických dát, návrhy a konštrukcie súčiastok alebo len bežné hranie hier. Spoločným menovateľom všetkých aplikácií je fakt, že umelé neurónové siete dosiahli výsledky ako svoje biologické ekvivalenty – ľudia – za výrazne kratší čas, dokonca sa im podarilo vyriešiť úlohy, ktoré boli pre ľudí prakticky nemožné [6, 7, 8].

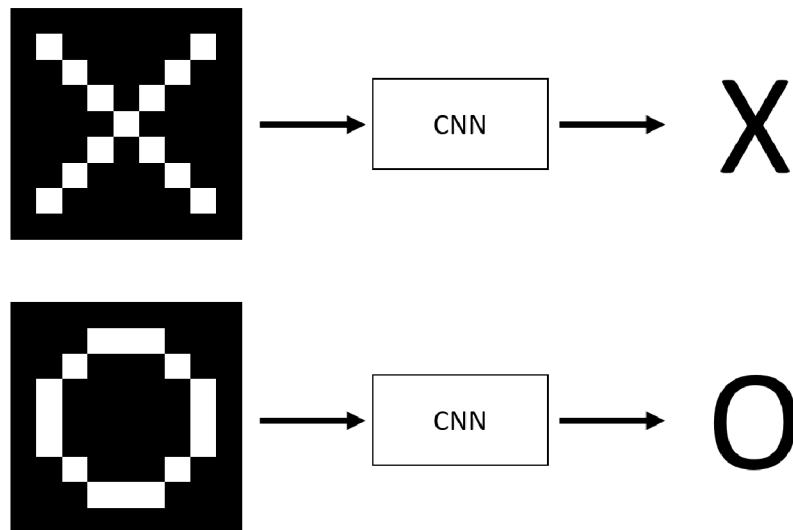
Úžasným príkladom pokroku je projekt AlphaGo od Googlu, ktorého cieľom je využiť strojové učenie na natrénovanie hry Go. Bližšia znalosť tejto hry nie je nutná, avšak je potrebné si uvedomiť, že počet všetkých možných variant tejto hry je viac ako samotných atómov vo vesmíre, tým pádom nie je možné naučiť počítač všetky výherné postupy. Prvá verzia s menom AlphaGo Lee bola trénovaná formou všetkých dostupných záznamov hier, na základe ktorých sa sieť naučila vyhrávať a už v roku 2016 porazila najlepšieho ľudského hráča Lee Sedola. Pri druhej verzii AlphaGo Zero zvolili autori iný postup, pri ktorom boli jedinými vstupnými dátami pravidlá hry. Následne algoritmus trénoval s použitím princípov spomenutých v predošlej kapitole. Už po 21 dňoch od spustenia tréningu bol porazený šampión z Číny, Ke Jie. Po 41 dňoch porazil AlphaGo Zero všetky ostatné umelé inteligencie, ktoré boli pre hru Go vytvorené [9].

Rovnako fascinujúce sú aj výsledky ďalšej neurónovej siete od Googlu, ktorej

úlohou bolo zdokonaliť existujúcu neurónovú sieť na predikciu textu. Výsledkom bola architektúra neurónovej siete, ktorej optimalizácia a efektivita bola na výrazne vyššej úrovni ako pri verzii vytvorenej človekom [10].

## 1.2 Architektúra siete

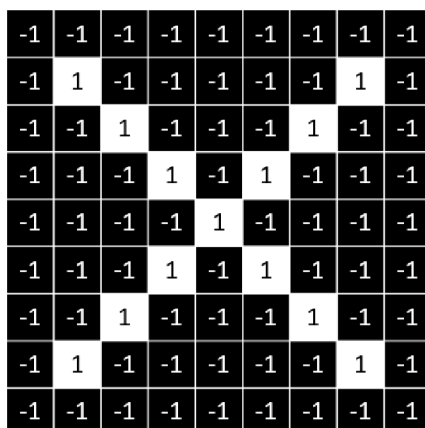
Pod pojmom architektúra neurónovej siete si môžeme predstaviť prepojenie jednotlivých vrstiev (funkčných blokov) siete, ktoré sa výrazne podieľajú na jej vlastnostiach a definujú jej presnosť a výkonnosť. Konvolučné neurónové siete pozostávajú z troch základných vrstiev, ktoré si popíšeme na nasledujúcich stranách. Počas vysvetľovania sa budeme držať jednoduchého príkladu, ktorým je rozoznávanie X a O z obrázka (obr. 1.4). Tento príklad je dostatočne obsiahly na to, aby na ňom bolo možné vysvetliť princípy fungovania a popísať jednotlivé vrstvy, avšak stále dostatočne jednoduchý na pochopenie širším obecnstvom.



Obr. 1.4: Rozpoznanie symbolov v obrázkoch pomocou konvolučnej neurónovej siete.

Prvoplánovým riešením by bolo vziať vstupný obrázok a porovnať ho s obrázkom X alebo O a rozhodnúť, s ktorým dosahuje vyššiu zhodu. Pre počítač je obrázok iba dvojrozmernou maticou hodnôt  $-1$  pre bielu a  $1$  pre čiernu (obr. 1.5). Ak sa pri porovnávaní obrázkov nezhodujú hodnoty jednotlivých pixelov, tak sa nezhoduje ani samotný obrázok. V ideálnom prípade by mal počítač rozpoznať aj rôzne posunuté, zdeformované, otočené a rozmerovo zmenené obrázky [11]. Tu prichádzajú na rad CNN.

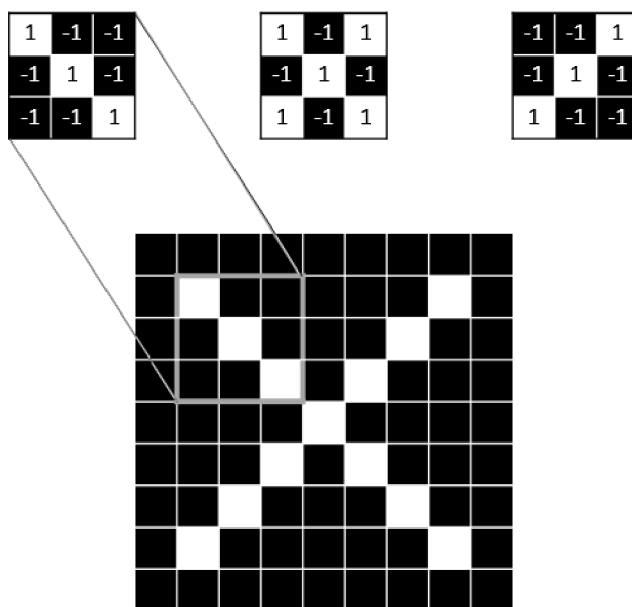
Základným rozdielom medzi konvolučnými neurónovými sieťami a tradičnými metódami spracovania obrazu je ten, že na rozdiel porovnávanie celého obrázku



Obr. 1.5: Obrázok ako dvojrozmerné pole hodnôt.

s určitým vzorom, CNN porovnávajú obrázok kúsok po kúsku s časťami nazývanými **príznaky** (angl. *features*). Hľadaním zhody po častiach je možné dosiahnuť oveľa vyššej presnosti ako pri celoobrazovom spracovaní a zároveň zabezpečiť rozpoznanie rôzne modifikovaných obrázkov [12].

Každý príznak sa dá popísať ako mini-obrázok – malá dvojrozmerná matica (obr. 1.6). Príznaky reprezentujú jednotlivé aspekty obrázkov, napr. príznaky symbolu X sú dve diagonálne čiary a jeden kríž.



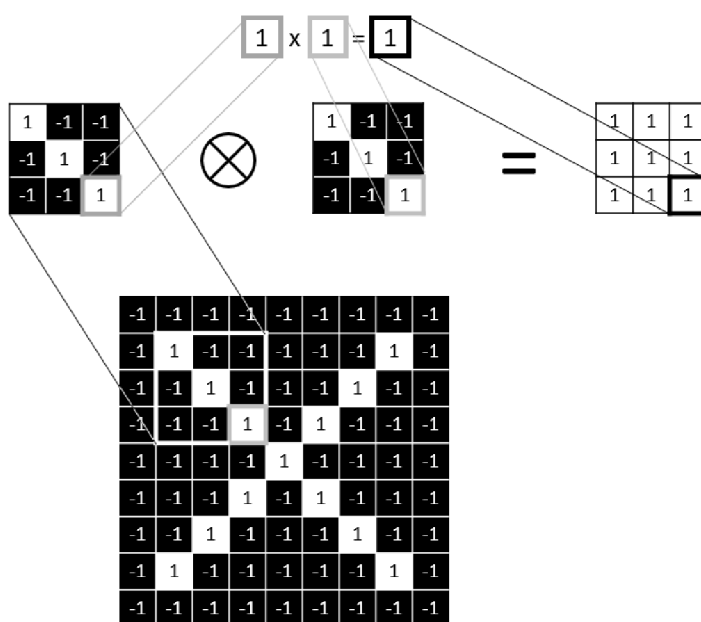
Obr. 1.6: Príznaky symbolu X.



### 1.2.1 Konvolučná vrstva

Pri vložení nového obrázku CNN nevie, kde konkrétne sa príznaky nachádzajú, preto ich začne hľadať na každej možnej pozícii – z príznaku spravíme filter. Pre výpočet miery zhody príznaku a práve zvolenej časti obrázka sa používa operácia, ktorej základom je 2D konvolúcia [13]. Operácia konvolúcie je fundamentálnou operáciou konvolučných neurónových sietí, po ktorej dostali aj svoje meno.

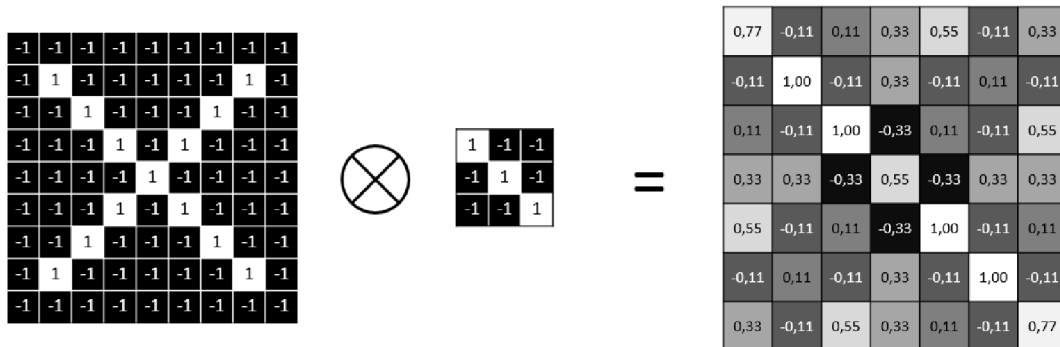
Na vypočítanie zhody príznaku s aktuálne zvolenou časťou obrázka jednoducho vynásobíme všetky pixely príznaku s príslušnými pixelmi na obrázku (obr. 1.7). Tieto súčiny následne sčítame a vydělíme počtom pixelov príznaku. Ako výsledok dostaneme mieru zhody, ktorá sa pohybuje v intervale od  $-1$  po  $1$ .



Obr. 1.7: Princíp 2D konvolúcie.

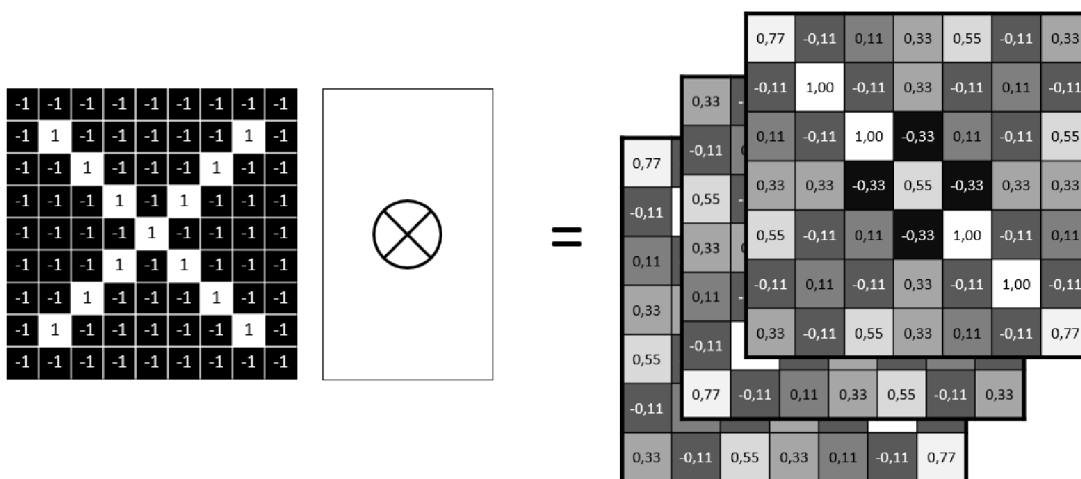
Pre dokončenie konvolúcie túto operáciu opakujeme s každým kúskom obrázka. Po zobrazení všetkých výsledkov na ich príslušných pozíciách v obrázku dostávame novú dvojrozmernú maticu (obr. 1.8). Táto matica zhôd je zároveň filtrovanou verziou nášho pôvodného obrázka, ktorá nám pomocou miery zhody určuje časti obrázka, v ktorých sa nachádza daný príznak. Hodnoty blízke 1 reprezentujú silnú zhodu, hodnoty blízke  $-1$  zobrazujú zhodu s fotografickým negatívom nášho príznaku (tzn. invertované hodnoty každého prvku matice) a hodnoty blízke 0 znamenajú nulovú zhodu akéhokoľvek typu.

Ďalším krokom je opakovanie konvolúcie so všetkými príznakmi. Výsledkom je sada filtrovaných obrázkov (obr. 1.9), jeden na každý filter (príznak). Spravidla má



Obr. 1.8: Výsledok 2D konvolúcie vstupného obrázka a príznaku.

výsledný obrázok menšie rozmery ako pôvodný, čo je dané charakterom samotnej konvolúcie. Táto vlastnosť bude dôležitá v nasledujúcich vrstvách. Súhrn všetkých operácií konvolúcie môžeme považovať za jeden procesný krok, ktorý reprezentuje práve konvolučná vrstva [14].

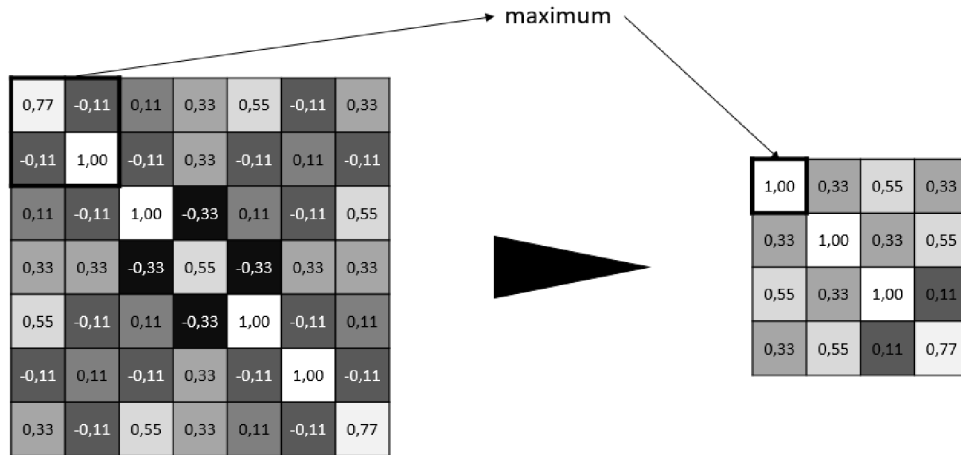


Obr. 1.9: Výsledok konvolúcie so všetkými príznakmi.

### 1.2.2 Zdužovacia vrstva

Ďalším z rady mocných nástrojov konvolučných neurónových sietí je zdužovanie alebo *pooling*. Zdužovanie je spôsob, akým je možné zmenšiť veľké obrázky a zároveň zachovať podstatné informácie v nich. Podstata zdužovania spočíva v posúvaní malého okna (zvyčajne 2 x 2 alebo 3 x 3 pixely) o určitý krok a vyberaní najvyššej hodnoty z okna pri každom kroku. Po dokončení zdužovania má výsledný obrázok približne štvrtinu (resp. devätinu) pôvodných pixelov (obr. 1.10). Pretože táto

funkcia vyberá iba najvyššie hodnoty z okna, informácie o najpravdepodobnejšom výskyte príznakov zostávajú nedotknuté [15].



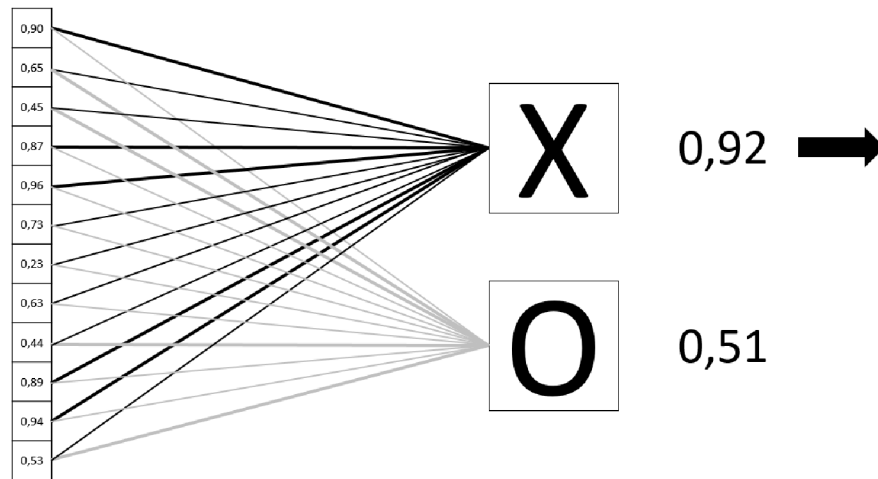
Obr. 1.10: Výsledok združovania.

To znamená, že nezáleží na tom, kde presne sa príznaky nachádzajú, pokiaľ sa v danom obrázku vyskytujú. Vďaka tejto vlastnosti dosahujú CNN úspech aj s obrázkami, ktoré sú rôzne deformované, posunuté alebo otočené, čo je pri deterministických metódach spracovania jedným z najväčších problémov.

### 1.2.3 Plne prepojená vrstva

V tejto vrstve prichádza na rad učenie a samotné rozhodovanie. Plne prepojená vrstva zoberie vysoko filtrované obrázky spracované predošlými vrstvami a spraví z nich hlasy. V našom prípade je možné hlasovať len o dvoch možnostiach, X alebo O. Plne prepojené vrstvy sú hlavnými stavebnými blokmi tradičných neurónových sietí. Namiesto spracovania dvojrozmerných obrázkov (matic) sú všetky vstupné hodnoty spracované ako jednorozmerný vektor. Každá hodnota reprezentuje jeden hlas, avšak jednotlivé hlasy nie sú spracované rovnocenne. Každý jeden hlas má totiž svoju váhu (istotu), s akou vie o danom obrázku povedať, do ktorej kategórie patrí. Hlasy s vyššou váhou majú pri rozhodovaní väčší vplyv a po sčítaní hlasov je výstupom jedna hodnota, ktorá získala najviac hlasov a definitívne rozhoduje o príslušnosti daného objektu k skupine (obr. 1.11).

V praxi sa pre rozhodovanie využíva väčší počet plne prepojených vrstiev, ktoré sa nazývajú vstupná (prvá v rade), výstupná (posledná v rade) a skrytá (všetky ostatné). Každá ďalšia plne prepojená vrstva umožňuje neurónovým sieťam naučiť sa oveľa komplexnejšie vzťahy medzi príznakmi a tým vykonávať presnejšie rozhodnutia [16]. Dôležitým prvkom tejto vrstvy je učenie, ku ktorému sme sa ešte nedostali.



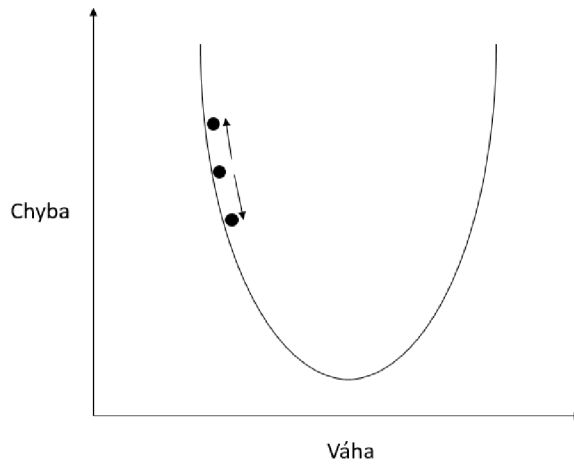
Obr. 1.11: Výsledok hlasovania.

Odkiaľ pochádzajú príznaky, ako nastaviť váhy v plne prepojených vrstvách? Pokiaľ by boli tieto parametre nastavované ručne, stratilo by sa všetko čaro neurónových sietí a ich učenia. Tu sa dostáva do popredia malý trik strojového učenia, BPE – algoritmus spätného šírenia chyby [17].

Aby bolo možné BPE použiť, je nutné CNN predložiť sériu obrázkov, ktorých príslušnosť k určitej skupine vieme. To znamená trpezlivého človeka, ktorý stovkám až tisíckam obrázkov priradí popis X alebo O. Na začiatku máme netrénovanú konvolučnú neurónovú sieť, ktorej príznaky v konvolučnej vrstve a váhy v plne prepojenej vrstve sú nastavené náhodne. Následne sieti predložíme obrázky, jeden po druhom.

Po spracovaní každého obrázka dostávame z rôznych vrstiev rôzne výstupy. Po porovnaní so správnou hodnotou dostávame mieru chyby, ktorá určuje, ako dobre sme na začiatku nastavili príznaky a váhy. Následne sú príznaky a váhy upravené, o malý krok hore alebo dole (obr. 1.12), obrázok je znovu spracovaný a výsledkom je nový hlas, ktorý je taktiež znovu porovnaný s očakávaným hlasom.

V prípade, že je chyba väčšia ako pôvodná sú zmeny zahodené, v opačnom prípade sa zachovávajú a pokračuje sa ďalej. Po upravení každého pixelu v príznakoch a každej váhy v plne prepojených vrstvách dostávame výsledok, ktorý je o kúsok presnejší ako predošlý. Proces opakujeme až do spracovania všetkých tréningových obrázkov. Náhodné odchýlky vyskytujúce sa v niektorých obrázkoch zapríčiňujú zvyšovanie chyby, avšak vďaka ich malému počtu je tento „krok späť“ rýchlo zabudnutý. Po spracovaní veľkého množstva obrázkov sa príznaky a váhy ustália na určitých hodnotách, ktoré budú o ďalších prípadoch rozhodovať s vysokou presnosťou.

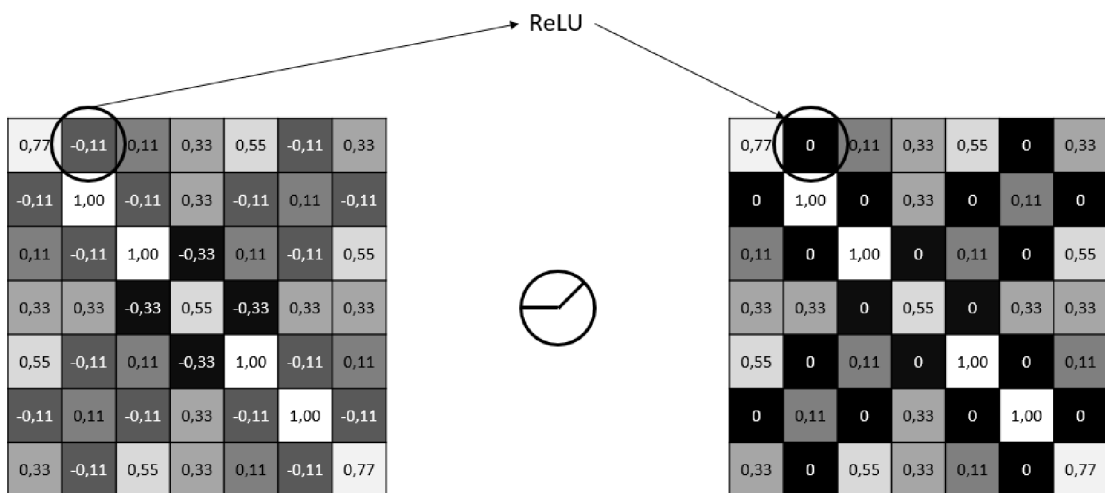


Obr. 1.12: Nastavovanie váh.

### 1.2.4 Aktivačné funkcie

Aktivačná funkcia sa aplikuje pre aktiváciu neurónu na jeho výstup. Pod aktivačnou funkciou môžeme chápať ľubovoľnú matematickú funkciu, avšak podmienkou je jej nelinearita. V prípade zvolenia lineárnej aktivačnej funkcie by každá sekvencia neurónov vykonávala iba lineárne transformácie, v dôsledku čoho by sme mohli celú neurónovú sieť reprezentovať iba jediným neurónom [18].

V posledných rokoch si veľkú popularitu získala aktivačná funkcia *Rectified Linear Near Unit* (ReLU) [19] (obr. 1.13), ktorá plní jednoduchú, avšak pre CNN veľmi dôležitú úlohu.



Obr. 1.13: Obrázok po aplikácii ReLU.

Jej aplikácia spôsobí to, že všetky záporné hodnoty sa nahradia 0 a kladné hodnoty sa bez zmeny prevedú na výstup, podobne ako to vyjadruje vzťah 1.2:

$$f(x) = \max(0, x), \quad (1.2)$$

kde  $x$  je vstupom neurónu. Prínos tejto funkcie spočíva vo výraznom zlepšení rýchlosti učenia CNN.

Ďalšou funkciou je funkcia **Softmax** [20]. Je to výstupná logická funkcia, ktorá sa používa pri klasifikačných úlohách, kedy je cieľom určiť objektu (obrázku) príslušnosť k jednej z mnohých skupín. Funkcia priraduje pravdepodobnostné hodnoty všetkým výstupom siete tak, aby boli splnené všetky definície a vety o pravdepodobnosti (každá hodnota je z intervalu  $(0, 1)$  a súčet všetkých pravdepodobností je rovný 1). Funkciu je možné definovať nasledovne:

$$y_i = \frac{e^z}{\sum_{k=1}^K e^z} \quad (1.3)$$

Jej jednoduchším ekvivalentom je **sigmoidná** funkcia, ktorá je vyžívaná pri klasifikácií do dvoch skupín.

### 1.3 Hyperparametre neurónovej siete

Nanešťastie, príznaky a váhy nie sú jediné parametre, ktoré sú pre fungovanie konvolučných neurónových sietí dôležité. Musíme si položiť aj nasledujúce otázky:

- Pre každú konvolučnú vrstvu – koľko príznakov a akých rozmerov?
- Pre každú združovaciu vrstvu – aká má byť veľkosť okna a krok?
- Pre každú plne prepojenú vrstvu – koľko skrytých neurónov a v akom zapojení?
- Ktorú aktivačnú a optimalizačnú funkciu zvoliť?
- Ako nastaviť rýchlosť učenia a zmien?

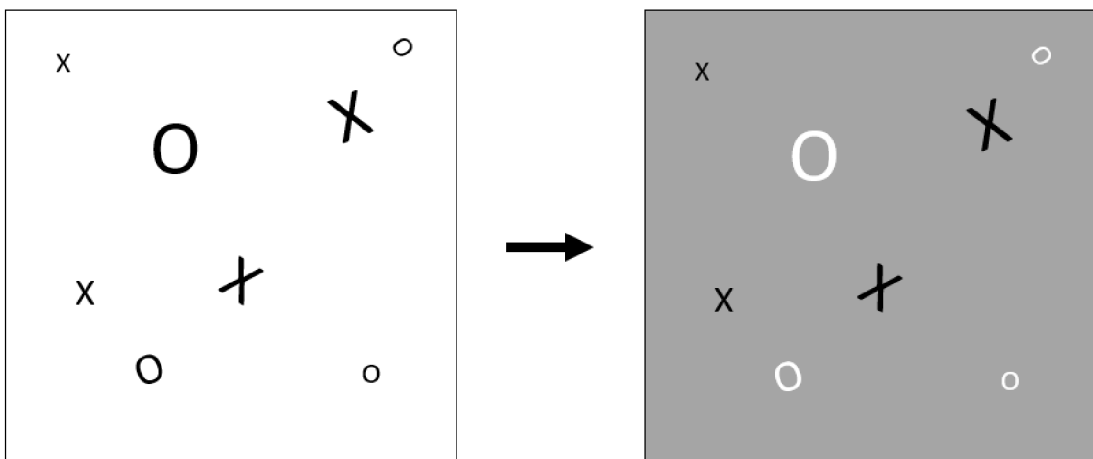
Všetky tieto parametre môžeme zaradiť medzi hyperparametre neurónovej siete. Nesmieme zabudnúť ani na to, že veľké neurónové siete môžu pozostávať až z tisícok vrstiev. Ktoré typy vrstiev? V akom poradí? Všetky predošlé otázky nám poskytujú nespočetné množstvo možností, ktorými vieme zostaviť konvolučnú neurónovú sieť. V praxi bol otestovaný len zlomok kombinácií a veľká časť výskumu neurónových sietí sa venuje práve tejto problematike. Aj pre človeka nepodstatné zmeny môžu mať obrovský dopad na presnosť a výkonnosť výslednej siete. V súčasnosti existujú pokusy navrhnúť nové architektúry neurónových sietí pomocou iných neurónových sietí [10].

## 1.4 Segmentácia obrazu pomocou CNN

Rozhodovanie o tom, či je na obrázku X alebo O nie je naším hlavným cieľom. Našou úlohou je nie len rozhodnúť, ale aj určiť, v ktorej časti obrázku sa daný symbol nachádza. Doteraz sme sa stretli len s vrstvami, ktoré svojou činnosťou kvalitu obrázka znížili, resp. zachovali. Na výstupe poslednej plne prepojenej vrstvy máme vektor pravdepodobnosti, ktorý percentuálne vyjadruje istotu, s akou je daný útvar X alebo O.

Pokiaľ sa chceme dostať k pôvodnému obrázku spolu s mapou výskytu segmentovaného objektu, je nutné vykonať operáciu nadvzorkovania alebo inak povedané **dekonvolúcie** s rovnakou veľkosťou okna a kroku ako pri konvolúcii. Proces dekonvolúcie patrí k jednej z najzložitejších operácií v procese segmentácie obrazu a jej konkrétne prevedenie má výrazný vplyv na výslednú kvalitu segmentácie. Napríklad, konvolučná neurónová sieť s názvom U-Net používa k dekonvolúcii výstupy z vrstiev na rovnakej úrovni, čo umožňuje presnejšiu rekonštrukciu obrazu a výrazné zlepšenie dosiahnutých výsledkov [21].

Segmentácia obrazu je proces rozdeľovania obrázku na viacero segmentov (častí). Cieľom segmentácie je zjednodušenie reprezentácie obrázku na niečo zmyslupnnejšie a jednoduchšie reprezentovateľné. Zvyčajne je používaná na lokalizáciu objektov a ich hraníc v obrázkoch. Pokiaľ by sme chceli našu definíciu spresniť, jedná sa o proces priradovania popisu jednotlivým pixelom, ktoré zdieľajú spoločné charakteristiky. Výsledkom segmentácie je sada segmentov, ktoré pokrývajú celý obrázok alebo skupina kontúr reprezentujúcich hranice objektov. V našom prípade môžeme vidieť príklad segmentácie, kde sú symboly X označené čiernou farbou, symboly O bielou a pozadie šedou (obr. 1.14) [22].



Obr. 1.14: Segmentácia obrazu.

Samotnú segmentáciu je možné rozdeliť na dva prípady. Prvý, kedy sú jednou značkou (farbou) označené všetky objekty spadajúce do segmentovanej skupiny (všetky X sú čiernou). Tento typ segmentácie sa nazýva sémantická a využíva sa pri detekcii výskytu istých prvkov v obrázku, ako napríklad detekcia ľudí vo vyhradených oblastiach. Druhým prípadom je tzv. *instance-aware* segmentácia, ktorá nielenže deteguje výskyt prvkov v obrázku, ale zároveň rozlišuje ich jednotlivé výskyty (jednotlivé osoby) [23].



## 2 ANALÝZA PROBLÉMU A IMPLEMENTÁCIA RIEŠENIA

Nasledujúca kapitola bude zameraná na analýzu zadaného problému a návrh konkrétneho riešenia. Bude vysvetlený princíp vzniku medicínskych dát a predstavené v súčasnosti používané metódy spracovania týchto dát. Následný návrh bude zahŕňať neurónové siete a jeho predpokladom budú poznatky z predošlej kapitoly.

### 2.1 Analýza problému

V rámci bakalárskej práce bolo úlohou naštudovať súčasné metódy segmentácie obrazu, ktoré využívajú hlboké učenie. Zo skúmaných metód bola následne vybratá jedna, ktorá bola najvhodnejšia na automatickú 3D segmentáciu snímok mozgu snímaných pomocou MRI – *Magnetic Resonance Imaging*, magnetickej rezonancie.

Vstupné dáta predstavuje množina 22 3D snímok lebky nasnímaných pomocou magnetickej rezonancie. Každý 3D snímok je zložený z 257-mich rezov o rozlíšení 400 x 400 pixelov. Úlohou je automatická segmentácia mozgu z poskytnutých snímok a následné porovnanie presnosti segmentácie s ľudským spracovaním.

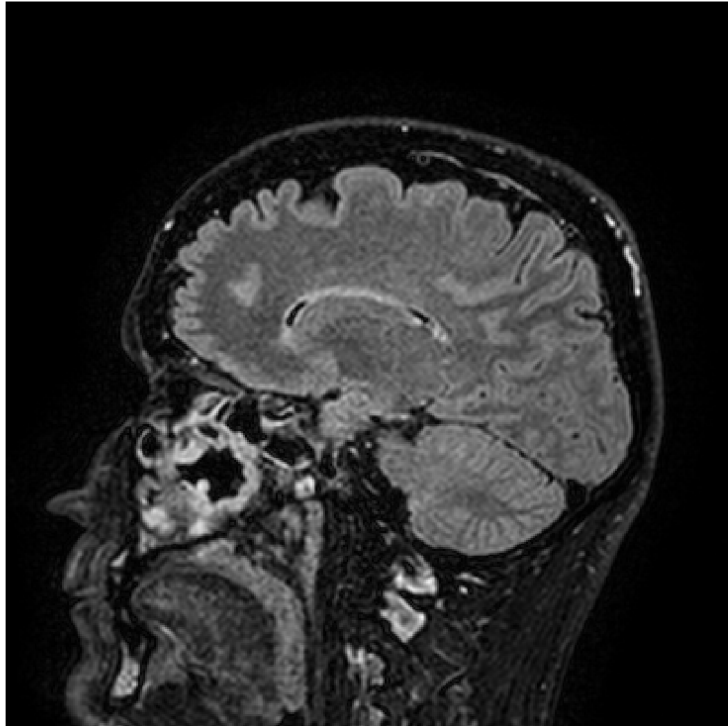
#### 2.1.1 Magnetická rezonancia

Zobrazovanie magneticou rezonanciou patrí medzi moderné rádiologické metódy, pomocou ktorej je možné diagnostikovať patologické zmeny v ľudskom tele bez nutnosti invazívneho zásahu a poškodenia tkaniva. Skúmaná časť tkaniva (pacientovho tela) je vložená do prístroju generujúceho veľmi silné homogénne magnetické pole. Následne je do tela pacienta vyslaný krátky rádiový impulz vyvolávajúci nepatrné zmeny v atónoch vodíka, ktoré prístroj následne deteguje a spracováva do grafickej podoby [24].

Výsledkom takéhoto snímania je séria obrázkov v 8-bitovej farebnej palete reprezentujúca kompletnú biologickú stavbu skúmanej časti tela. V súčasnej dobe sú výsledky spracovávané a reprezentované ľuďmi, čo vytvára priestor na chybovosť a nesprávnu diagnostiku pacienta.

#### 2.1.2 Popis dát

Ako už bolo spomenuté, na bakalársku prácu bolo poskytnutých 22 3D snímok lebky (obr. 2.1). K snímkam sú dostupné aj 1-bitové referenčné masky (obr. 2.2) reprezentujúce vysegmentovaný mozog a takisto masky vysegmentované ľuďmi, ktoré slúžia k porovnaniu presnosti strojového a ľudského spracovania.



Obr. 2.1: Snímka z magnetickej rezonancie mozgu.



Obr. 2.2: Referenčná maska mozgu snímky z obr. 2.1.

Úlohou bude naučiť neurónovú sieť rozpoznať a označiť mozog v snímkach lebky. Druhý krok spočíva v následnom použití neurónovej siete na nespracované dáta a kontrole dosiahnutých výsledkov. K segmentácii sa v dnešnej dobe používa množstvo metód, z ktorých niektoré si popíšeme v nasledujúcej časti.

## 2.2 Súčasné metódy segmentácie obrazu

Drvivá väčšina doteraz používaných metód počítačového spracovania obrazu, konkrétne segmentácie, bola založená na matematických operáciách určujúcich hranice objektu na základe istých charakteristík ako sú jas, textúra a iné. Tento prístup nevyžaduje bližšiu znalosť segmentovanej časti obrazu, postačuje len jej popis voči okoliu v ktorom sa nachádza.

S rozvojom výpočtovej techniky a umelej inteligencie vznikol istý trend používania neurónových sietí na spracovanie obrazu. Oproti spomenutému postupu umožňujú neurónové siete segmentáciu bez bližšej znalosti charakteristík obrazu.

### 2.2.1 Metódy založené na znalosti obrazu

Metódy založené na znalosti obrazu je možné rozdeliť do troch hlavných skupín. Prvou sú metódy využívajúce globálne znalosti obrazu, zvyčajne reprezentované histogramom určitých vlastností. Segmentáciu je ďalej možné rozdeliť na postupy vychádzajúce z hraníc medzi oblasťami obrazu a na postupy vytvárajúce priamo tieto oblasti. Nie je pritom podstatné z akých charakteristík (jas, textúra, rýchlostné pole, ...) pri hľadaní či tvorbe hraníc vychádzame.

Kombináciou nasledujúcich postupov by bolo možné vytvoriť systém, ktorý by plnil zadanú úlohu, avšak jeho implementácia by bola nesmierne zložitá a neflexibilná. Napriek tomu sú takéto riešenie stále hojne používané.

#### Segmentácia prahovaním

Najjednoduchším segmentačným postupom je prahovanie. Vychádza zo skutočnosti, že množstvo objektov či oblastí obrazu je charakterizovaných konštantným jasom. Túto skutočnosť je následne možné využiť na stanovenie jasových konštánt – prahu, pomocou ktorého sú objekty oddelené od pozadia. Prahovanie patrí k najstarším segmentačným metódam a v istých prípadoch je aj naďalej využívané [25].

#### Segmentácia na základe detekcie hrán

Veľkú časť segmentačných metód tvoria práve metódy založené na detekcii hrán. Tento prístup patril k historicky prvým a zároveň najdôležitejším prístupom v seg-

mentácii obrazu.

Využíva skutočnosti, že hranice oblastí obrazu pozostávajú z hrán, ktoré je možné nájsť aplikáciou niektorého z hranových operátorov. Takto sú označené miesta, na ktorých dochádza v obraze k istej nespojitosti, zvyčajne v hodnote jasnosti alebo tiež vo farbe alebo textúre.

Výsledok aplikácie samotného hranového operátora je vo svojej podobe prakticky nepoužiteľný. Preto je nutné ďalšie spracovanie, ktoré spája jednotlivé hrany do hraníc objektu. V závislosti od použitého hranového operátora a algoritmu na spájanie hrán je možné dosiahnuť rôznu úspešnosť výsledkov na rôznych problémoch [25].

### **Segmentácia narastaním oblasti**

Určiť oblasti obrazu pri dokonalej znalosti ich hraníc nie je zložité, rovnako ako nie je problém následne zostrojiť ich hranice. V praxi sa často stretávame so zašumeným obrazom, v ktorom sa hranice určujú veľmi ťažko. Spojením predošlých dvoch metód je možné tento problém vyriešiť.

Z hľadiska segmentácie je významnou vlastnosťou oblastí ich homogenita. Základnou myšlienkou segmentácie s narastaním oblastí je rozčlenenie obrazu do maximálne súvislých oblastí tak, aby boli z hľadiska zvoleného popisu homogénne. Kritérium homogenity sa môže opierať o rôzne vlastnosti obrazu, od jasnosti až po zložitejšiu štruktúru. Kritérium sa volí v konkrétnom použití algoritmu a výrazne ovplyvňuje dosahované výsledky [25].

### **Segmentácia porovnávaním so vzorom**

Princíp segmentácie porovnávaním so vzorom (angl. *matching*) je očividný – obrázok sa porovnáva so zadaným vzorom a určuje sa ich miera súhlasu – vzájomná korelácia. Okrem detekcie oblastí umožňuje tento prístup aj sledovanie pohybu v sérii obrázkov. Pre nájdenie najlepšej zhody je používané optimalizačné kritérium, ktoré môže byť od jednoduchej korelácie až po dômyselné porovnávanie grafov tvorených podľa nájdených vlastností porovnávaných objektov [25].

## **2.2.2 Metódy založené na neurónových sieťach**

Spracovaniu obrazu pomocou neurónových sietí sa venovala celá prvá kapitola – jedná sa o typ neurónovej siete, ktorý sa výhradne používa k spracovaniu obrazu. Každá takáto sieť je postavená na troch vrstvách – konvolučná, združovacia a plne prepojená – vykonávajúci operácie vedúce k segmentácii obrazu.

Na začiatku sú zvolené príznaky, ktoré sú v konvolučnej vrstve porovnávané s práve spracovávaným obrázkom a je počítaná miera ich zhody s obrázkom. Následne sú z obrázka vyčlenené len významné hodnoty, čím dochádza k jeho výraznému zmenšeniu, ktoré zabezpečuje združovacia vrstva. Po opakovanej aplikácii týchto dvoch vrstiev prichádza na rad plne prepojená vrstva, ktorá zabezpečuje rozhodovanie a tvorí „inteligenciu“ neurónovej siete. Výstupom plne prepojenej vrstvy je rozhodnutie o príslušnosti objektu k danej skupine – v prípade segmentácie mozgu z lebky sa jedná o rozhodnutia je to mozog a nie je to mozog. V prípade segmentácie viacerých objektov nie je výstupom len príslušnosť k istej skupine, ale aj skupina, do ktorej daný objekt patrí. Po rozhodnutí je obraz zrekonštruovaný a výsledkom je segmentovaný obraz.

Oproti metódam založeným na znalosti obrazu umožňujú neurónové siete obrovskú flexibilitu v spracovaní konkrétnych obrázkov. Systém založený na znalosti obrazu je úzko špecializovaný softvér navrhnutý na konkrétnu úlohu – napríklad segmentácia mozgu. V prípade zmeny tejto úlohy je prispôsobenie systému veľmi obtiažne, niekedy dokonca až nemožné.

Riešenia založené na neurónových sieťach sa nezameriavajú na žiadnu konkrétnu segmentáciu obrazu, pretože majú možnosť učiť sa. Jedinými predpokladmi sú vstupné dáta a vzorové výsledky, pomocou ktorých vie neurónová sieť upraviť svoje rozhodovanie k dosiahnutiu želaných výsledkov. V prípade zmeny problému stačí sieti poskytnúť nové dáta a razom je možné upraviť systém na segmentáciu mozgu na systém na segmentáciu srdca zo snímky hrudníka, ľudí z fotky, áut z kamerových záznamov. Na riešenie zadaného problému bola zvolená architektúra U-net, ktorá je všeobecne používaná na segmentáciu obrazu neurónovými sieťami s veľmi vysokou presnosťou [21].

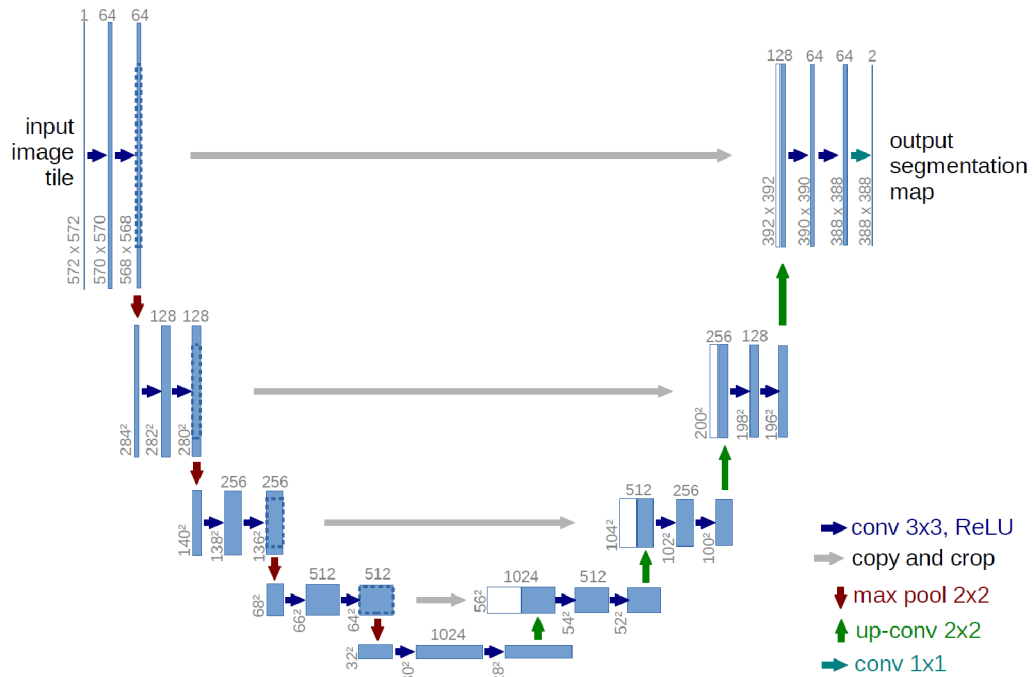
## 2.3 U-Net

*U-Net* je konkrétnym príkladom konvolučnej neurónovej siete určenej na precízne a rýchle spracovanie obrazu. Svoj názov dostala na základe výraznej podoby jej usporiadania s písmenom U. Výhodou takto implementovanej CNN je jej prehľadnosť, rekonštrukcia obrazu pomocou dát z predošlých vrstiev a malé množstvo dát potrebné na jej natréovanie [21].

### 2.3.1 Architektúra a implementácia

Na obr. 2.3 [26] je zobrazené konkrétne zapojenie U-Netu.

Ako je možné vidieť, celá architektúra tvorí 5 pomyselných úrovní spracovania. Každá úroveň s výnimkou poslednej pozostáva z dvoch konvolučných a jednej



Obr. 2.3: Zapojenie U-Netu.

združovacej vrstvy, resp. dvoch konvolučných a jednej nadvzorkovacej vrstvy. Medzi vstupnou a výstupnou časťou rovnakej úrovne je možné vidieť signálny tok obsahujúci výstupy daných vrstiev, vďaka ktorým je možná precíznejšia rekonštrukcia výsledného obrázku.

Najspodnejšia úroveň U-Netu predstavuje pomyselný most medzi vstupnou a výstupnou časťou. Na poslednej vrstve je výstup, na ktorý je aplikovaná aktivačná funkcia, ktorá „jednoduchý“ obrázok prevedie na výstupnú masku tak, že každému voxelu priradí istú triedu (v našom prípade sú len dve možnosti – je (1) alebo nie je (0) mozog).

Implementácia celého systému je rozdelená do troch samostatných častí. Prvou z nich je načítanie a spracovanie vstupných dát. Vzhľadom na pamäťovú náročnosť riešenia je nutné zmenšiť rozlíšenie vstupných snímok. Spracované snímky sú uložené na pevný disk do rýchlo dostupného formátu *.npy*.

Po spracovaní dát nasleduje samotné tréningovanie siete. Po natréňovaní sú všetky potrebné parametre neurónovej siete (príznačky, váhy, ...) uložené do jediného binárneho súboru, ktorý umožňuje ďalšie použitie bez predošlého tréningovania. V praxi to predstavuje spôsob, akým je možné distribuovať parametre siete dosahujúce lepšie výsledky bez nutnosti výpočtov u všetkých klientov.

Tretiu časť tvorí systém na samotnú segmentáciu, ktorý vyžaduje binárny súbor z predošlej časti a súbor so snímkami určenými na segmentáciu. Po vykonaní

segmentácie sú výstupné binárne masky prevedené na formu čierno-bielých obrázkov a uložené do priečinkov, ktorých adresárová štruktúra zodpovedá adresárovej štruktúre snímok určených na segmentáciu.

### 2.3.2 Behové prostredie a požiadavky

Táto konkrétna architektúra konvolučnej neurónovej siete prináša aj svoje nevýhody. Prvou z nich je kritérium pre rozmery vstupných obrázkov, ktoré musia byť násobkom čísla 16. Tento fakt je spôsobený tým, že sú využité 4 združovacie vrstvy s veľkosťou okna 2, ktoré vstupný obrázok pred spracovaním mostom zmenšia 16-násobne. Samotný U-Net je postavený na Pythone vo verzii 3.6.3 a niekoľkých v ňom naprogramovaných knižniciach, ktoré si popíšeme na nasledujúcich riadkoch.

#### TensorFlow

TensorFlow je otvorená knižnica určená k programovaniu dátovo prepojených úloh a tvorí základ zvoleného riešenia. Je to symbolická matematická knižnica hojne používaná pri strojovom učení a vývoji neurónových sietí. Jej flexibilná architektúra umožňuje rozdelenie výpočtov medzi jedno alebo viac jadier procesoru alebo grafickej karty. TensorFlow bol vyvinutý výskumníkmi a inžiniermi z *Google Brain* tímu v rámci organizácie zaoberajúcej sa výskumom hlbokého učenia a neurónových sietí, avšak je rovnako vhodný aj pre široké spektrum iných aplikácií [27].

#### Keras

Keras je otvorená knižnica určená pre priamy vývoj neurónových sietí. Je navrhnutá s dôrazom na umožnenie rýchlych experimentov na poli neurónových sietí a zameriava sa hlavne na jednoduchosť, modularitu a rozšíriteľnosť. K svojej funkčnosti vyžaduje niektorú z knižníc určenú pre spracovanie dát, v tomto prípade TensorFlow. Rovnako umožňuje paralelizovať výpočty na grafickej karte a procesore, čo má za následok výrazné zrýchlenie vykonávaných činností [28].

#### Scikit-image

Scikit-image je balíček v Pythone zahrňajúci sadu algoritmov pre prácu s obrázkami. Prostredníctvom tohto balíčku sú načítavané, predspracované a uložené všetky vstupné dáta spolu s maskami. K svojmu fungovaniu vyžaduje ďalšie balíky pre prácu so súborami, poliami a pre organizáciu dát. Bližšie informácie je možné nájsť v oficiálnej dokumentácii [29].

## **h5py**

Balík h5py prestravuje pre Python rozhranie pre prácu s binárnymi dátami. Umožňuje ukladať a spravovať obrovské množstvá numerických dát na pevný disk a je nevyhnutný pre uloženie jednotlivých parametrov neurónovej siete [30].

## **2.4 Rozšírený U-Net**

Prvotný návrh U-Netu mal jednu veľkú nevýhodu, ktorou bola pamäťová náročnosť, kvôli ktorej bolo nutné vstupné dáta zmenšiť, čím došlo k nenávratnej strate dôležitých údajov. Zároveň bolo potrebné trénovať na CPU, čo výrazne zvyšovalo čas tréovania. Upravená verzia U-Netu všetky tieto problémy rieši vďaka pozmenenému konceptu načítavania a predspracovania dát. Ten umožňuje plne automatické 3D spracovanie ľubovoľného počtu snímok na grafickej karte bez akéhokoľvek zásahu do vstupných dát.

### **2.4.1 Predspracovanie a koncept**

Vstupné dáta sú k dispozícii s veľkosťou 400 x 400 x 257 voxelov. Pôvodne sa dáta zmenšovali tak, aby bolo možné načítať celú 3D snímku mozgu. Nový prístup spočíva v tom, že celý mozog sa rozdelí na skupiny snímok o veľkosti 3, ktoré sa budú vzájomne prekrývať. Tým pádom je možné na vstup dodať ľubovoľný počet snímok bez nutnosti redukcie ich počtu. Prvá skupina bude obsahovať snímky 1 až 3, druhá 2 až 4, tretia 3 až 5, ... ako je možné vidieť na obrázku 2.4.

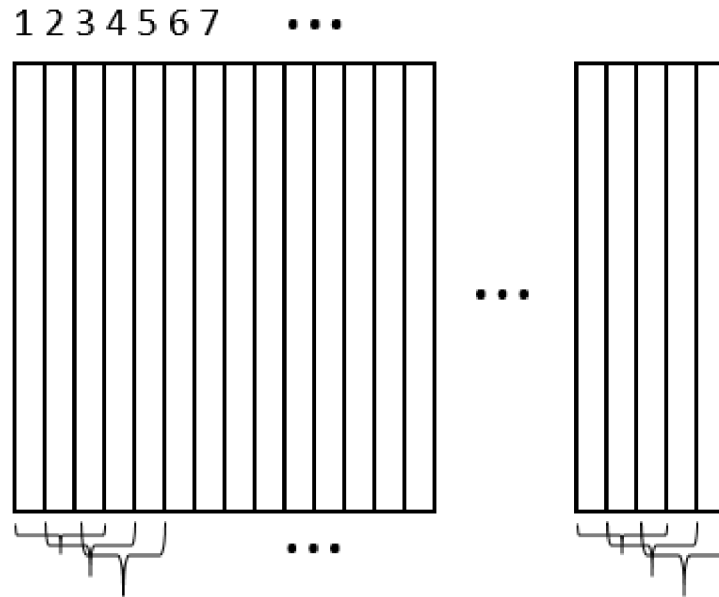
Takýmto prepojením jednotlivých skupín zostane zachovaná 3D informácia o celom mozgu (tzn., pri segmentácii strednej snímky môže sieť nahliadnuť do susedných snímok a podľa toho sa rozhodnúť). Praktickým prínosom je zníženie pamätevej náročnosti a novovzniknutá možnosť trénovať na grafickej karte, dokonca bez zmenšovania rozlíšenia vstupných dát.

K vedľajším produktom takéhoto prístupu patrí aj zväčšenie trénovacej množiny, ktoré priaznivo ovplyvňuje výslednú presnosť a znižuje počet epôch potrebných na učenie. Po natrénovaní siete nasleduje predikcia, ktorá používa rovnaké predspracovanie dát, v závere však vyberie vždy iba stredný snímok z každej skupiny, čím dochádza k strate krajných dvoch snímok, ktoré neobsahujú žiadne časti mozgu.

### **2.4.2 Architektúra**

Okrem zmeny spôsobu spracovávania dát bola vykonaná aj zmena v architektúre – boli pridané nové interkonekcie v jednotlivých úrovniach, ktoré zlepšujú presnosť





Obr. 2.4: Spracovanie vstupných dát.

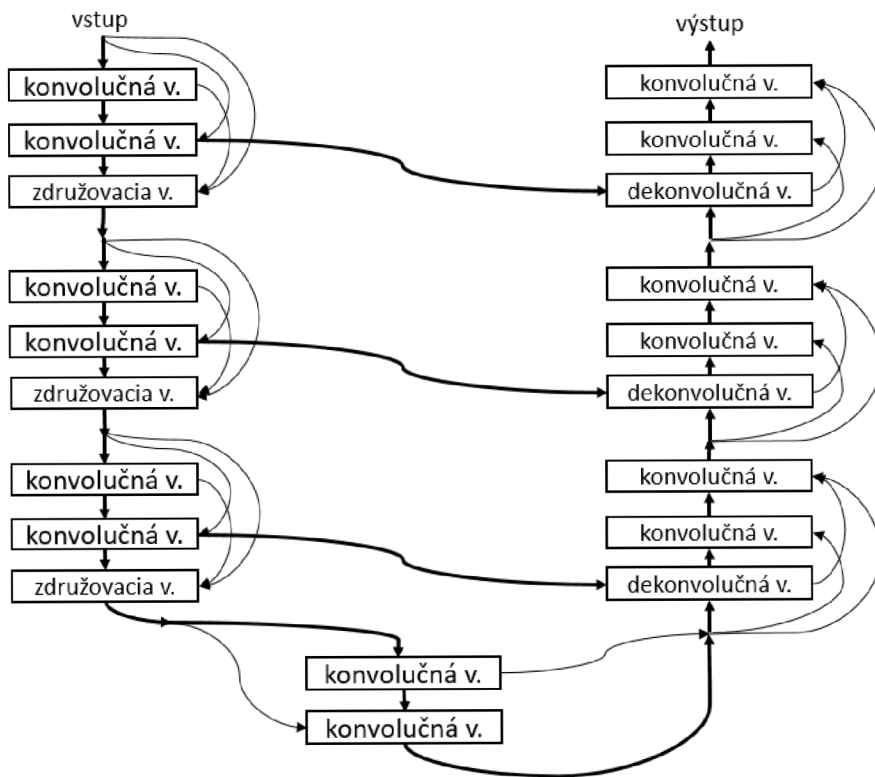
siete vďaka možnosti spracovávanía dát so znalosťami z iných vrstiev. Tým vznikol husto prepojený U-Net, tzv. *Dense U-Net*.

Hrubou čiaru v obrázku 2.5 sú znázornené prepojenia existujúce aj v pôvodnom U-Nete, tenké čiary reprezentujú nové interkonekcie. Súčasné riešenie má iba tri úrovne a tomu zodpovedajúce množstvo parametrov. Každou ďalšou úrovňou sa tento počet môže zvýšiť až do momentu, kedy je vyčerpaná celá pamäť grafickej karty. Otázka znie, akým spôsobom je možné pridať ďalšie úrovne bez zväčšovania pamäťovej náročnosti? Odpoveďou je **rekurentná sieť**.

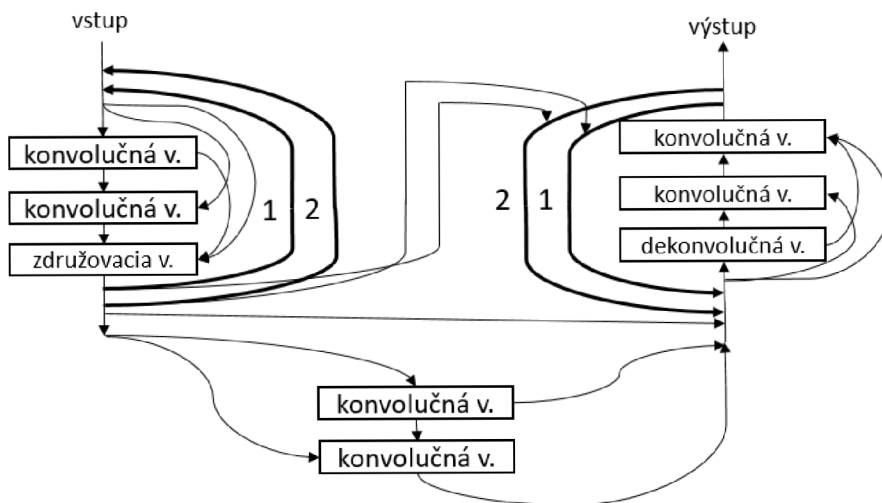
## 2.5 Rekurentný U-Net

Myšlienkou rekurentného U-Netu je znovupoužitie jednotlivých úrovní. Tým je možné virtuálne vytvoriť niekoľkoúrovňovú štruktúru, ktorá má rovnaké pamäťové nároky ako jednoúrovňový U-Net.

Hrubou čiarou na obrázku 2.6 sú zvýraznené znovupoužitia jednotlivých úrovní, ktoré je možné takmer neobmedzene zvyšovať bez zvyšovania zložitosti a počtu parametrov siete. V tomto konkrétnom prípade sa jedná o zníženie počtu parametrov neurónovej siete o viac ako 80% oproti nerekurentnej verzii s tromi úrovňami.



Obr. 2.5: Architektúra rozšíreného U-Netu.



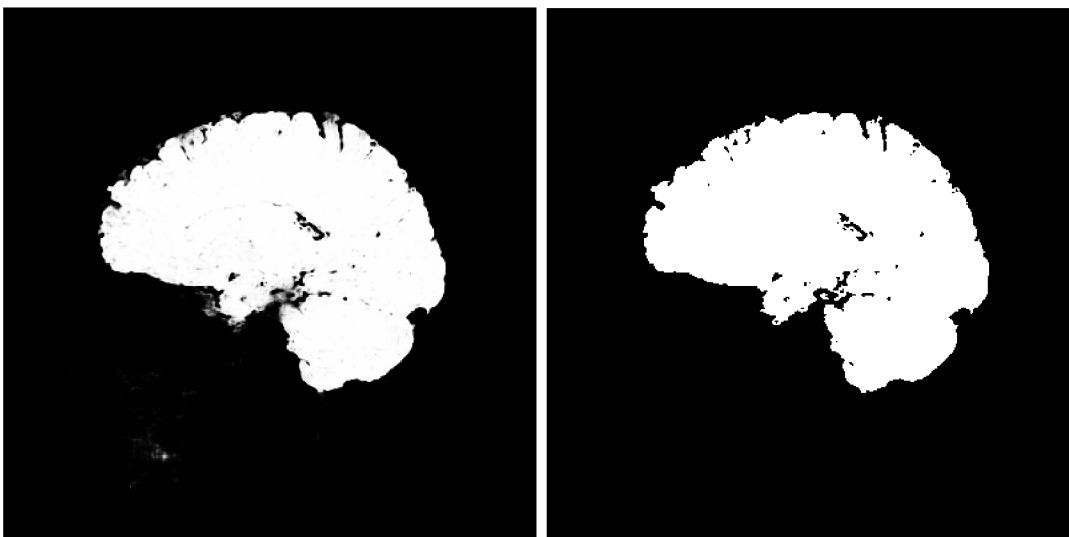
Obr. 2.6: Architektúra rekurentného U-Netu.

### 3 VYHODNOTENIE VÝSLEDKOV

Posledná časť bakalárskej práce bude venovaná prezentácii dosiahnutých výsledkov a ich vyhodnoteniu. Po načítaní boli dáta predložené jednotlivým neurónovým sieťam a bolo spustené učenie. V prípade prvotného U-Netu bolo nutné zredukovať vstupné dáta na rozmery 256 x 256 x 128 voxelov. Učením vznikol binárny súbor s príponou .h5 s uloženými parametrami siete, ktorý je načítaný a použitý pri každej predikcii (pozn. každá sieť má vlastný binárny súbor). Finálnym výstupom je séria masiek predikovaných zo zadaných dát.

#### 3.1 Presnosť výsledkov

Oproti teoretickému popisu, kedy bolo rozhodnuté o príslušnosti pixelu k určitej skupine na plne prepojenej vrstve bol zvolený iný postup. Ako výsledok boli uložené istoty vyjadrujúce pravdepodobnosť s akou je daný pixel mozog alebo nie. To nám umožňuje experimentovať s rozhodovacím prahom a optimalizovať tak rozhodovanie pre čo najlepšie výsledky. Z dôvodu grafickej reprezentácie boli hodnoty 0 až 1 prevedené do 8-bitovej farebnej palety, ktorá nadobúda rozsah od 0 (čierna) až po 255 (biela). Na obrázku č. 3.1 naľavo môžeme vidieť reprezentáciu masky z výstupu neurónovej siete. Zjavné sú aj artefakty a pixely mimo mozgu, ktoré vznikli ako chyba pri redukcii rozmerov dát alebo pri predikcii sieťou. Zvolením vhodnej rozhodovacej úrovne je možné tieto chybné časti odstrániť a tým získať presnejší výsledok (obr. č. 3.1 napravo).



Obr. 3.1: Výstupná maska pred a po upravení rozhodovacej úrovne.

Pri vyhodnocovaní presnosti jednotlivých sietí budú použité nasledujúce metriky získané nástrojom *Visceral – EvaluateSegmentation* [31] a rozhodovací prah 50%, ktorý vyjadruje predvolené správanie siete (tj., zaokrúhľovanie).

### ***Per voxel accuracy* – voxelová presnosť**

Voxelová presnosť vyjadruje absolútnu mieru zhody dvoch 3D snímok  $\mathbf{X}$  a  $\mathbf{Y}$ . pričom sa berú do úvahy:

- $N_{TP}$  – počet TP – *true positive* – skutočne pozitívnych voxelov (tj., správne označených ako mozog (1)),
  - $N_{TN}$  – počet TN – *true negative* – skutočne negatívnych voxelov (tj., správne označených ako pozadie (0)),
  - $N_{FP}$  – počet FP – *false positive* – falošne pozitívnych voxelov (tj., nesprávne označených ako mozog (1)),
  - $N_{FN}$  – počet FN – *false negative* – falošne negatívnych voxelov (tj., nesprávne označených ako pozadie (0)),
- nasledujúcim spôsobom [31]:

$$A_{PV}(X, Y) = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}. \quad (3.1)$$

### **Dice koeficient**

Dice koeficient vyjadruje mieru podobnosti dvoch datasetov. Oproti predošlej metrike neberie do úvahy *true negative* voxely, čím je dosiahnuté objektívnejšie hodnotenie datasetov s veľkým množstvom pozadia (voxelov označených ako 0). Vzťah pre výpočet je nasledujúci [31]:

$$D(X, Y) = \frac{2 * N_{TP}}{2 * N_{TP} + N_{FP} + N_{FN}}. \quad (3.2)$$

### **Jaccard index**

Jaccard index oproti Dice koeficientu berie *true positive* voxely do úvahy len raz, v čitateli aj menovateli, čím je ešte prísnejší pri hodnotení kvality segmentácie. Jaccard index je reprezentovaný vzorcom 3.3 [31]:

$$J(X, Y) = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}. \quad (3.3)$$

### 3.1.1 U-Net

Pri všetkých troch sieťach bola ako trérovací dataset použitá séria 21 kompletných 3D skenov mozgu, z ktorých 10% snímok bolo využitých na validáciu. Následne bol sieť predložený 3D sken, ktorý nikdy nevidela a bola vykonaná predikcia. Presnosť ľudskej segmentácie porovnaná podľa vyššie spomínaných metrík voči referenčným maskám reprezentuje tabuľka 3.1.

Tab. 3.1: Presnosť ľudskej segmentácie.

Per voxel accuracy	Dice koeficient	Jaccard index
0.995210	0.980368	0.961493

Veľkou nevýhodou tohto U-Netu je nutnosť zmenšovania vstupných dát, ktorá sa odrazila na výslednej presnosti. Problémom je takisto doba trérovania – sieť sa učila len 20 epôch, čo je v porovnaní s bežnou praxou, rádovo stovky epôch, výrazne menej. Dôvodom takto krátkej učiacej doby bolo zdieľanie výpočtového serveru s ostatnými študentmi – učenie počas adekvátneho počtu epôch by zabralo desiatky dní. Presnosť U-Netu v porovnaní s človekom je možné vidieť v tabuľke 3.2.

Tab. 3.2: Porovnanie presnosti človeka a U-Netu.

Zdroj	Per voxel accuracy	Dice koeficient	Jaccard index
Človek	0.995210	0.980368	0.961493
U-Net	0.992613	0.964964	0.932299
<b>Porovnanie</b>	<b>-0,002597</b>	<b>-0,015404</b>	<b>-0,029194</b>
<b>Relatívne zlepšenie</b>	<b>-0,260950%</b>	<b>-1,571247%</b>	<b>-3,036320%</b>

### 3.1.2 Rozšírený U-Net

Rozšírený U-Net zmenil prístup k spracovaniu údajov a umožnil spracovanie v plnom rozlíšení, preto bola eliminovaná strata presnosti spôsobená komprimáciou údajov. Ako je možné vidieť v tabuľke č. 3.3, presnosť rozšíreného U-Netu dokázala prekonať ľudskú segmentáciu. Sieť sa učila počas 30-tich epôch a z praxe je isté, že v prípade dostatočne dlhého učenia by sieť svoje výsledky ešte zlepšila.

Oproti prvej verzii U-Netu dosiahla segmentácia rozšíreným U-Netom učným v plnom rozlíšení na grafickej karte presnosť o viac ako 4,12% vyššiu (Jaccard index). Viac detailov je možné vidieť v tabuľke 3.4.

Tab. 3.3: Porovnanie presnosti človeka a rozšíreného U-Netu.

Zdroj	Per voxel accuracy	Dice koeficient	Jaccard index
Človek	0.995210	0.980368	0.961493
Rozšírený U-Net	0.996758	0.985147	0.970730
<b>Porovnanie</b>	<b>+0,001548</b>	<b>+0,004779</b>	<b>+0,009237</b>
<b>Relatívne zlepšenie</b>	<b>0,155545%</b>	<b>0,487470%</b>	<b>0,960693%</b>

Tab. 3.4: Porovnanie presnosti rozšíreného a pôvodného U-Netu.

Zdroj	Per voxel accuracy	Dice koeficient	Jaccard index
U-Net	0.992613	0.964964	0.932299
Rozšírený U-Net	0.996758	0.985147	0.970730
<b>Porovnanie</b>	<b>+0,004145</b>	<b>+0,020183</b>	<b>+0,038431</b>
<b>Relatívne zlepšenie</b>	<b>0,417585%</b>	<b>2,091581%</b>	<b>4,122175%</b>

### 3.1.3 Rekurentný U-Net

Rekurentný U-Net pozostával na rozdiel od rozšíreného U-Netu len z 3 úrovní – úroveň smerom dolu, ktorá sa opakovane využívala, prepojovacia úroveň, tzv. most a úroveň, ktorá sa takisto opakovane používala smerom hore. To zaistilo výrazne nižšiu pamäťovú náročnosť a čas tréovania, dokonca aj vyššiu presnosť ako v predošlých verziách. Ako je možné vidieť v tabuľke 3.5, rozdiel medzi človekom a rekurentným U-Netom je ešte väčší (v prospech neurónovej siete) a v prípade dostatočne dlhého tréovania by rekurentný U-Net svoju presnosť ešte zvýšil.

Tab. 3.5: Porovnanie presnosti človeka a rekurentného U-Netu.

Zdroj	Per voxel accuracy	Dice koeficient	Jaccard index
Človek	0.995210	0.980368	0.961493
Rekurentný U-Net	0.997210	0.987221	0.974764
<b>Porovnanie</b>	<b>+0,002000</b>	<b>+0,006853</b>	<b>+0,013271</b>
<b>Relatívne zlepšenie</b>	<b>0,200963%</b>	<b>0,699023%</b>	<b>1,380249%</b>

V tabuľke 3.6 je vidieť porovnanie doteraz najpresnejšieho (rozšíreného U-Netu) s rekurentným U-Netom. Ten oproti pôvodnému U-Netu zaznamenal zlepšenie pres-

nosti o 4,55% a oproti rozšírenému U-Netu viac ako 0,41% pri porovnávaní najprísnejšieho sledovaného kritéria – Jaccard indexu.

Tab. 3.6: Relatívne zlepšenie presnosti rekurentného U-Netu oproti pôvodnému a rozšírenému U-Netu.

Porovnané s	Per voxel accuracy	Dice koeficient	Jaccard index
U-Net	0,463121%	2,306511%	4,554869%
Rozšírený U-Net	0.045347%	0.210527%	0.415564%

Na záver je v tabuľke 3.8 prehľad všetkých metrík každej zo skúmaných sietí, v tabuľke 3.9 je možné vidieť porovnanie presnosti segmentácie neurónových sietí s človekom podľa Jaccard indexu.

Tab. 3.7: Porovnanie všetkých sietí.

Zdroj	Per voxel accuracy	Dice koeficient	Jaccard index
Človek	0.995210	0.980368	0.961493
U-Net	0.992613	0.964964	0.932299
Rozšírený U-Net	0.996758	0.985147	0.970730
Rekurentný U-Net	0.997210	0.987221	0.974764

Tab. 3.8: Porovnanie presnosti segmentácie všetkých sietí s človekom podľa Jaccard indexu.

Sieť	Absolútne zlepšenie	Relatívne zlepšenie
U-Net	-0,029194	-3,036320%
Rozšírený U-Net	+0,009237	+0,960693%
Rekurentný U-Net	+0,013271	+1,380249%

## 3.2 Náročnosť spracovania údajov

Prvým krokom každého U-Netu bolo predspracovanie dát. Jedná sa o proces, kedy sú tréningové a testovacie dáta spojené do jedného súboru pre následné využitie neurónovou sieťou. Doba tohto spracovania závisí od množstva vstupných dát a rýchlosti disku a pohybuje sa na úrovni niekoľkých minút. Pre každú sieť je tento krok nutné

vykonať iba raz, vzniknuté súbory je možné použiť opakovane aj pri iných sieťach. Pôvodný U-Net v rámci predspracovania dát znižoval veľkosť všetkých vstupných dát na rozmery 256 x 256 x 128 voxelov, čím sa trvanie tohto kroku predĺžilo zhruba o 30%. Výhodou neurónových sietí je ukladanie všetkých váh a konfigurácie do jediného binárneho súboru, ktorý je jednoducho prenositeľný a odstraňuje potrebu učenia pri každom použití siete.

### Pôvodný U-Net

Pôvodný U-Net spracovával kompletne 3D snímky mozgu, čím bol pamäťovo veľmi náročný a bolo nutné ho spustiť na procesore, ktorý mal dostupný dostatočné množstvo pamäte. Počas celej doby spracovania zaberol proces približne 50 GB operačnej pamäte pri dobe trvania jednej epochy 2 hodiny. Pri 20-tich epochách sa jedná o 40 hodín čistého výpočtového času, ku ktorému treba pripočítať dobu potrebnú na priebežné ukladanie aktuálneho stavu siete (rádovo minúty). Sieť celkovo pozostávala z 22 575 329 trénovateľných parametrov. Samotná predikcia masky kompletnej 3D snímky trvala približne 1 minútu (jedná sa o rozmerovo redukované snímky).

### Rozšírený U-Net

Zmena konceptu načítavania dát pri zachovaní 3D spracovania umožnila spustiť sieť na grafickej karte, ktorá v porovnaní s procesorom výrazne zrýchliła proces učenia. Rozšírený U-Net pozostával z 2 531 650 trénovateľných parametrov, pričom jedna epocha trvala približne 2,5 hodiny a celková doba učenia viac ako 3 dni. Je to viac ako pri pôvodnom U-Nete, avšak treba brať do úvahy predspracovanie, ktoré virtuálne zväčšilo počet vstupných dát na trojnásobok. Po natrénovaní trvala predikcia jednej 3D snímky v plnom rozlíšení 2 minúty. Pre zaujímavosť, v tabuľke 3.9 je možné vidieť nárast počtu parametrov siete v závislosti od počtu použitých úrovní.

Tab. 3.9: Počet parametrov v závislosti od počtu úrovní rozšíreného U-Netu.

Počet úrovní	Počet parametrov
3 úrovne	2 531 650
4 úrovne	4 139 842
5 úrovní	6 135 106
6 úrovní	8 517 442



## Rekurentný U-Net

Poslednou zo skúmaných sietí bol rekurentný U-Net. Jedná sa o experimentálnu sieť, ktorá patrí k jednej z prvých rekurentných sietí použitých za účelom 3D segmentácie obrazu. Znovupoužitím úrovní sa podarilo znížiť počet trénovateľných parametrov na 443 553 a ako je možné vidieť v tabuľke 3.10, tento počet zostával konštantný bez ohľadu na počet použitých úrovní. Sieť bežala ako v predošlom prípade na grafickej karte s dĺžkou jednej epochy jedna hodina a 20 minút. Predikcia masiek následne trvala len o niečo menej ako jeden a pol minúty.

Tab. 3.10: Počet parametrov v závislosti od počtu úrovní rekurentného U-Netu.

Počet úrovní	Počet parametrov
3 úrovne	443 553
4 úrovne	443 553
5 úrovní	443 553
6 úrovní	443 553

Všetky výpočty prebiehali na školskom serveri *Gravy* a na grafickej karte *NVIDIA GeForce GTX 1080*.

## 3.3 Uplatnenie v praxi

Zvolený systém vie nájsť široké uplatnenie v praxi. Aktuálne riešenie poskytuje lekárom možnosť plne automatizovaného procesu **3D volumetrie** ľudského mozgu, ktorá nachádza uplatnenie pri diagnostike neurodegeneratívnych chorôb, pri ktorých dochádza k úbytku mozgovej hmoty. V súčasnosti je volumetria vykonávaná lekármi v rámci ich pracovného času. Zvolený systém dokáže nielenže ušetriť tento drahocenný čas, ale dosahuje vyššej presnosti, čím zamedzuje ľudským chybám.

Výhodou neurónových sietí je ich možnosť rýchleho prispôsobenia na úplne iný problém behom krátkeho času, len zmenou trénovacích dát. Navrhnuté riešenie je všeobecne použiteľné na problém 3D segmentácie obrazu, jedinou podmienkou je dodržanie potrebnej adresárovej štruktúry.

Netreba zabúdať ani na prínos samotnej práce, ktorá predstavuje úvod do konvolučných neurónových sietí a problematiky 3D segmentácie obrazu. Rekurentné siete patria jednej k z mála oblastí, o ktorých neexistujú rozsiahle štúdie. V čase písania nebol dostupný žiadny vedecký článok skúmajúci rýchlosť, presnosť a obmedzenia rekurentných sietí v oblasti 3D segmentácie, čím táto práca okrem vyžadovaného prínosu pokladá aj základy ďalšieho výskumu.

## 4 ZÁVER

Táto bakalárska práca sa venovala problému 3D segmentácie obrazu pomocou konvolučných neurónových sietí. Cieľom bolo vytvorenie systému pre automatickú 3D segmentáciu mozgu zo snímkov magnetickej rezonancie. Po teoretickom úvode do problematiky a následnom popise a analýze problému bola zvolená konvolučná neurónová sieť U-Net. V rámci riešenia bolo vytvorených niekoľko verzií U-Netu, z ktorých každá využívala rozdielny prístup k danému problému. Siete boli natrénované na konkrétny problém a použité na samotnú segmentáciu. Výsledky ukázali, že po vhodnej úprave predspracovania dát je možné znížiť pamäťové nároky natolko, že je možné spustiť sieť na grafickej karte bez znižovania rozlíšenia vstupných dát a zároveň zvýšiť jej presnosť. Práca takisto ukázala, že rekurentné neurónové siete dokážu konkurovať svojim nerekurentným verziám a tým položila základ pre ďalší výskum. Navrhnuté siete dokázali segmentovať celý mozog behom niekoľkých desiatok sekúnd, a to dokonca presnejšie ako človek.

Zvolený systém umožňuje veľmi rýchlu adaptáciu na iný typ spracovania obrazu a takisto veľmi jednoduchý spôsob distribúcie naučených poznatkov. Parametre naučenej siete sú uložené v jednom binárnom súbore, ktorý je možné nahradiť súborom zo siete, ktorá bola trébovaná na väčšej vstupnej množine a po dlhšiu dobu.

Prvá verzia U-Netu musela kvôli svojej pamäťovej náročnosti znižovať rozlíšenie spracovávaných dát takmer na polovicu v každom rozmere. Pamäťová náročnosť takisto neumožňovala trébovanie siete na grafickej karte, ale na procesore, čo ju zaradilo k jedným z pomalších riešení. Takto vznikla potreba optimalizácie, ktorej výsledkom bol rozšírený U-Net. Prvá vec, ktorá prešla zmenou, bol spôsob načítavania vstupných dát. Pôvodný U-Net spracovával kompletný 3D sken mozgu. Rozšírený U-Net načítaval prekrývajúce sa skupiny troch snímkov, čím zostala zachovaná 3D informácia a znížila sa pamäťová náročnosť natolko, že dáta nebolo nutné zmenšovať a sieť bolo možné trébovať na grafickej karte. Okrem spracovania dát sa zmeny dočkala aj architektúra, v ktorej pribudli interkonekcie zvyšujúce kvalitu a presnosť vykonávaných operácií. Výsledkom bol systém, ktorý bol niekoľkokrát rýchlejší a aj výrazne presnejší ako pôvodné riešenie.

Po dokončení a overení rozšíreného U-Netu bola pozornosť venovaná ďalšej optimalizácii, z ktorej vznikol rekurentný U-Net. Jeho myšlienkou je znovupoužitie už existujúcich úrovní nachádzajúcich sa v sieti, čím dochádza k zníženiu počtu parametrov neurónovej siete a zrýchleniu celého procesu učenia a následnej predikcie. Navrhnuté riešenie takisto umožňuje zvyšovanie počtu úrovní bez vplyvu na počet parametrov siete. Táto sieť patrí k prvým rekurentným sieťam použitým na problém 3D segmentácie a vytvára veľké množstvo otázok spolu s priestorom na ďalší výskum.

Medzi prínosy tejto práce môžeme určite zaradiť teoretický obsah, ktorý poskytuje stručný a prehľadný úvod do neurónových sietí. Poskytuje ucelený prehľad o spôsobe segmentácie obrazu konvolučnými neurónovými sieťami a popisuje rozdiely medzi týmto riešením a tradičnými metódami spracovania obrazu. Popísané boli 3 riešenia určené k automatickej 3D segmentácii mozgu, z ktorých menovite posledné, rekurentný U-Net, predstavuje základy pre ďalší výskum. Výstupom sú takisto aj programy umožňujúce plne automatickú 3D segmentáciu mozgu.

# LITERATÚRA

- [1] LECUN, Y. *LeNet-5, convolutional neural networks* [online]. 2015, [cit. 1. 10. 2017]. Dostupné z URL: <http://yann.lecun.com/exdb/lenet>.
- [2] ZARAPINA, N. *Anotace obrazu s použitím hlubokého učení* [online]. 43 strán. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2017 [cit. 7. 10. 2017]. Dostupné z URL: [https://www.vutbr.cz/studium/zaverecne-prace?zp\\_id=101896](https://www.vutbr.cz/studium/zaverecne-prace?zp_id=101896).
- [3] DAVALO, E., NAIM. P. *Neural networks*. Houndmills: Macmillan Education, 1991. ISBN 0-333-54996-1.
- [4] ZELINKA, I. *Umělá inteligence: hrozba nebo naděje?*. Praha: BEN - technická literatura, 2003. ISBN 80-7300-068-7.
- [5] TUČKOVÁ, J. *Vybrané aplikace umělých neuronových sítí při zpracování signálů*. Praha: České vysoké učení technické v Praze, 2009. ISBN 978-80-01-04229-8.
- [6] CHIA-YU SU, E., LI, Y., IQBAL, U. *Deep learning revolutionizes healthcare and precision medicine: the next wave of artificial intelligence applications in biomedicine*. Computer Methods and Programs in Biomedicine, 2017, poslední aktualizácia 1. 1. 2017 [cit. 13. 10. 2017]. Dostupné z URL: [https://www.researchgate.net/publication/312016309\\_Deep\\_learning\\_revolutionizes\\_healthcare\\_and\\_precision\\_medicine\\_the\\_next\\_wave\\_of\\_artificial\\_intelligence\\_applications\\_in\\_biomedicine](https://www.researchgate.net/publication/312016309_Deep_learning_revolutionizes_healthcare_and_precision_medicine_the_next_wave_of_artificial_intelligence_applications_in_biomedicine).
- [7] KASPAROV, G. K., GREENGARD M. *Deep thinking: where machine intelligence ends and human creativity begins*. New York: PublicAffairs, an imprint of Perseus Books, 2017. ISBN 9781610397865.
- [8] YOUSSEF, A., EL-TELBANY, M., ZEKRY, A. The role of artificial intelligence in photo-voltaic systems design and control: A review. In *Renewable and Sustainable Energy Reviews* [online]. 2017, [cit. 13. 10. 2017]. DOI: 10.1016/j.rser.2017.04.046. ISSN 13640321. Dostupné z URL: <http://linkinghub.elsevier.com/retrieve/pii/S13640321173055550>.
- [9] SILVER, D., SCHRITTWIESER, J., SIMONYAN, K. Mastering the game of Go without human knowledge. In *Nature* [online]. 2017, [cit. 13. 10. 2017].

DOI: 10.1038/nature24270. ISSN 0028-0836. Dostupné z URL:  
<<http://www.nature.com/doifinder/10.1038/nature24270>>.

- [10] LE, Q., ZOPH, B. *Using Machine Learning to Explore Neural Network Architecture* [online]. 2017, posledná aktualizácia 17.5.2017 [cit. 13.10.2017]. Dostupné z URL:  
<<https://research.googleblog.com/2017/05/using-machine-learning-to-explore.html>>.
- [11] LAWRENCE, S., GILES, C. L., TSOI, A. CH., BACK, A. D. Face recognition: a convolutional neural-network approach. In *IEEE Transactions on Neural Networks* [online]. 1997, [cit. 19.10.2017]. DOI: 10.1109/72.554195. ISSN 10459227. Dostupné z URL:  
<<http://ieeexplore.ieee.org/document/554195/>>.
- [12] KRIZHEVSKY, A., SUTSKEVER, I., HINTON, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* [online]. 2012, [cit. 19.10.2017]. Dostupné z URL:  
<<http://papers.nips.cc/paper/4824>>.
- [13] SMĚKAL, Z. *Číslíkové zpracování signálů*. Brno: Vysoké učení technické v Brně, 2012. ISBN 978-80-214-4639-7.
- [14] LIU, L., SHEN, CH., VAN DEN HENGEL, A. The treasure beneath convolutional layers: Cross-convolutional-layer pooling for image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* [online]. 2015, [cit. 22.10.2017]. Dostupné z URL:  
<[https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2015/html/Liu\\_The\\_Treasure\\_Beneath\\_2015\\_CVPR\\_paper.html](https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Liu_The_Treasure_Beneath_2015_CVPR_paper.html)>.
- [15] SCHERER, D., MÜLLER, A., BEHNKE, S. Evaluation of pooling operations in convolutional architectures for object recognition. In *Artificial Neural Networks-ICANN* [online]. 2010, [cit. 23.10.2017]. Dostupné z URL:  
<[https://link.springer.com/chapter/10.1007%2F978-3-642-15825-4\\_10?LI=true](https://link.springer.com/chapter/10.1007%2F978-3-642-15825-4_10?LI=true)>.
- [16] SIMARD, P. Y. Best practices for convolutional neural networks applied to visual document analysis. In *ICDAR* [online]. 2003, [cit. 23.10.2017]. Dostupné z URL:  
<<http://www.cs.cmu.edu/~bhiksha/courses/deeplearning/Fall.2016/pdfs/Simard.pdf>>.

- [17] HECHT-NIELSEN Theory of the backpropagation neural network. In *International Joint Conference on Neural Networks* [online]. 1989, [cit. 26. 10. 2017]. DOI: 10.1109/IJCNN.1989. Dostupné z URL: [<http://ieeexplore.ieee.org/document/118638/>](http://ieeexplore.ieee.org/document/118638/).
- [18] LESHNO, M., LIN, V. Y., PINKUS, A., SCHOCKEN, S. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. In *Neural Networks* [online]. 1993, [cit. 29. 10. 2017]. DOI: 10.1016/S0893-6080(05)80131-5. ISSN 08936080. Dostupné z URL: <http://linkinghub.elsevier.com/retrieve/pii/S0893608005801315>.
- [19] DAHL, G., SAINATH, T. N., HINTON, G. E. Improving deep neural networks for LVCSR using rectified linear units and dropout. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* [online]. 2013, [cit. 4. 11. 2017]. DOI: 10.1109/ICASSP.2013.6639346. ISBN 978-1-4799-0356-6. Dostupné z URL: <http://ieeexplore.ieee.org/document/6639346/>.
- [20] BISHOP, CH. M. *Pattern recognition and machine learning*. Springer, 2006. Dostupné z URL: [http://cds.cern.ch/record/998831/files/9780387310732\\_T0C.pdf](http://cds.cern.ch/record/998831/files/9780387310732_T0C.pdf).
- [21] RONNEBERGER, O., FISCHER, P., BROX, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* [online]. Springer, Cham, 2015, [cit. 10. 11. 2017]. Dostupné z URL: <https://arxiv.org/abs/1505.04597>.
- [22] LONG, J., SHELHAMER, E., DARRELL, T. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. 2015, [cit. 10. 11. 2017]. DOI: 10.1109/CVPR.2015.7298965. ISBN 978-1-4673-6964-0. Dostupné z URL: <http://ieeexplore.ieee.org/document/7298965/>.
- [23] DAI, J., HE, K., SUN, J. Instance-aware semantic segmentation via multi-task network cascades. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* [online]. 2016, [cit. 11. 11. 2017]. Dostupné z URL: <https://arxiv.org/abs/1512.04412>.
- [24] MAYHEW, T. M., OLSEN, D. R. Magnetic resonance imaging (MRI) and model-free estimates of brain volume determined using the Cavalieri principle.

- In *Journal of anatomy*. 1991, [cit. 27. 11. 2017]. Dostupné z URL:  
<<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1260542/>>.
- [25] HLAVÁČ, V., ŠONKA, M. *Počítačové vidění*. Praha: Grada, 1992. ISBN isbn80-85424-67-3.
- [26] RONNEBERGER, O. Image. In *U-net: Convolutional networks for biomedical image segmentation*. [online]. 2015, [cit. 1. 12. 2017]. Dostupné z URL:  
<<https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/>>.
- [27] ABADI, M., ASHISH. TensorFlow. *Large-Scale Machine Learning on Heterogeneous Systems* [online]. 2015, [cit. 4. 12. 2017]. Dostupné z URL:  
<<https://www.tensorflow.org/>>.
- [28] CHOLLET, F. Keras. 2015, [cit. 4. 12. 2017]. Dostupné z URL:  
<<https://github.com/fchollet/keras>>.
- [29] VAN DER WALT, S., SCHNONBERGER, J. L. scikit-image. In *scikit-image: Image processing in Python* [online]. 2014, [cit. 4. 12. 2017]. Dostupné z URL:  
<<https://peerj.com/articles/453/>>.
- [30] COLLETTE, A. h5py. *Python and HDF5* [online]. 2013, [cit. 4. 12. 2017]. Dostupné z URL:  
<<http://www.h5py.org/>>.
- [31] TAHA, A. A., HANBURY, A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. In *BMC Medical Imaging* [online]. 2015, [cit. 16. 5. 2018]. DOI: 10.1186/s12880-015-0068-x. ISSN 1471-2342. Dostupné z URL:  
<<http://bmcmimedimaging.biomedcentral.com/articles/10.1186/s12880-015-0068-x>>.

## ZOZNAM SYMBOLOV, VELIČÍN A SKRATIEK

BPE	Spätne šírenie chyby – <i>Back-Propagation of Error</i>
CNN	Konvolučná neurónová sieť – <i>Convolutional Neural Network</i>
GB	Gigabajt – <i>GigaByte</i>
MRI	Zobrazovanie magnetickou rezonanciou – <i>Magnetic Resonance Imaging</i>
ReLU	<i>Rectified Linear Unit</i>
SSD	Mechanika s nepohyblivým médiom – <i>Solid State Drive</i>
TP	Skutočne pozitívne – <i>True positive</i>
TN	Skutočne negatívne – <i>True negative</i>
FP	Falošne pozitívne – <i>False positive</i>
FN	Falošne negatívne – <i>False negative</i>



# ZOZNAM PRÍLOH

A Obsah priloženého disku

53

## A OBSAH PRILOŽENÉHO DISKU

Nasledujúci zoznam obsahuje zoznam súborov na priloženom disku spolu s ich stručným popisom. Priečinky a súbory označené symbolom \* sa nenachádzajú na disku a vzniknú až činnosťou programu. Priečinky a súbory s názvom písaným veľkými písmenami sú príkladom vyžadovanej adresárovej a súborovej štruktúry.

```
/.....koreňový adresár priloženého disku
├── 186025.pdf ..... digitálna verzia bakalárskej práce
├── software/..... priečinok programu
│   ├── data.py.....program na prípravu dát
│   ├── data_unet.py.....program na prípravu dát pre pôvodný U-Net
│   ├── *imgs_train.npy.....spracované tréningové dáta
│   ├── *imgs_mask_train.npy ..... spracované masky
│   ├── *imgs_test.npy ..... spracované dáta k predikcii
│   ├── *imgs_train_unet.npy..... spracované tréningové dáta pre pôvodný U-Net
│   ├── *imgs_mask_train_unet.npy.....spracované masky pre pôvodný U-Net
│   ├── *imgs_test_unet.npy.....spracované dáta k predikcii pre pôvodný U-Net
│   ├── unet.py ..... U-Net
│   ├── 3_layer.py.....rozšírený U-Net
│   ├── 3_layer_recurrent.py.....rekurentný U-Net
│   ├── *logs/.....logy z tréningovania
│   │   ├── *UNET.txt
│   │   ├── *3_layer.txt
│   │   └── *3_layer_recurrent.txt
│   └── masks/.....masky k tréningovaniu
│       ├── TRAIN_01_MASK
│       │   ├── IMG_001.png
│       │   ├── IMG_002.png
│       │   └── IMG_XXX.png
│       ├── TRAIN_02_MASK
│       ├── TRAIN_03_MASK
│       └── TRAIN_XX_MASK
```

