



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**

FACULTY OF INFORMATION TECHNOLOGY

**ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ**

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

**HLBOKÉ NEURÓNOVÉ SIETE PRE ANALÝZU  
MEDICÍNSKÝCH OBRAZOVÝCH DÁT**

DEEP NEURAL NETWORKS FOR MEDICAL DATA ANALYSIS

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**ALBERT SZÖLLÖSI**

**VEDOUcí PRÁCE**

SUPERVISOR

**Ing. MICHAL ŠPANĚL, Ph.D.**

BRNO 2020

## Zadání bakalářské práce



Student: **Szöllösi Albert**  
Program: Informační technologie  
Název: **Hluboké neuronové sítě pro analýzu medicínských obrazových dat**  
**Deep Learning for Medical Image Analysis**  
Kategorie: Zpracování obrazu

### Zadání:

1. Seznamte se s problematikou hlubokých neuronových sítí a jejich učení.
2. Zorientujte se v metodách segmentace medicínských obrazových dat s využitím hlubokých neuronových sítí (medicínská CT data, RTG snímky, apod.).
3. Vyberte vhodnou metodu použitelnou pro řešení zvoleného problému analýzy medicínského obrazu.
4. Implementujte navrženou metodu s využitím existujících nástrojů pro trénování hlubokých neuronových sítí.
5. Proveďte experimenty nad připravenou datovou sadou.
6. Porovnejte dosažené výsledky a diskutujte možnosti budoucího vývoje.
7. Vytvořte stručný plakát nebo video prezentující vaši práci, její cíle a výsledky.

### Literatura:

- Krizhevsky, A., Sutskever, I. and Hinton, G. E.: ImageNet Classification with Deep Convolutional Neural Networks, NIPS 2012

Pro udělení zápočtu za první semestr je požadováno:

- Body 1 až 3.

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Španěl Michal, Ing., Ph.D.**

Vedoucí ústavu: Černocký Jan, doc. Dr. Ing.

Datum zadání: 1. listopadu 2019

Datum odevzdání: 28. května 2020

Datum schválení: 1. listopadu 2019

## Abstrakt

Táto práca ponúka možné riešenie automatickej lokalizácie anatomických landmarkov v 3D dentálnych skenoch. Tieto skeny slúžia k uľahčeniu modelovania zubných korún a ortodontických aparátov pomocou špecializovaného softvéru. Predtým však musia byť anotované, aby softvér vedel určiť, kde sa jednotlivé zuby nachádzajú. Anotácia prebieha ručne, čo síce zaručuje presnosť, ale zaberá veľa času. Výsledok tejto práce by mohol spomínaný postup výrazne zjednodušiť aplikovaním hlbokého učenia. Lokalizácia landmarkov bola riešená pomocou konvolučnej neurónovej siete.

## Abstract

This thesis offers possible solution to automatic 3D dental scan landmark localization. These scans are used in dental crown design and digital orthodontics to make the design process easier using specialized software. Before that, though, the scan has to be annotated for the software to know the positions of the teeth. The annotation process is done manually, which guarantees precision, but takes a lot of time. The result of this work could make said process much simpler by applying deep learning. Landmark localization was implemented using a convolutional neural network.

## Klíčové slová

hlboké učenie, analýza medicínskych obrazových dát, lokalizácia landmarkov, dentálny sken

## Keywords

deep learning, medical image data analysis, landmark localization, dental scan

## Citácia

SZÖLLÖSI, Albert. *Hlboké neurónové siete pre analýzu medicínskych obrazových dát*. Brno, 2020. Bakalárska práca. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Michal Španěl, Ph.D.

# Hlboké neurónové siete pre analýzu medicínskych obrazových dát

## Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána Michala Španěla. Uviedol som všetky literárne pramene, publikácie a ďalšie zdroje, z ktorých som čerpal.

.....  
Albert Szöllösi  
28.5.2020

## Podakovanie

Chcel by som sa poďakovať pánovi Michalovi Španělovi, ktorého rady mi veľmi pomohli pri vypracovaní tejto práce.

# Obsah

<b>1</b>	<b>Úvod</b>	<b>2</b>
<b>2</b>	<b>Lokalizácia anatomických landmarkov</b>	<b>3</b>
2.1	Lokalizácia landmarkov v Röntgenových snímkach kolena pomocou hourglass siete . . . . .	3
2.1.1	Architektúra siete . . . . .	3
2.1.2	Loss funkcia . . . . .	5
2.1.3	Výsledky . . . . .	5
2.2	Lokalizácia landmarkov v medicínskych snímkach pomocou umelo vytvorených dát . . . . .	6
2.2.1	Vytváranie dát . . . . .	6
2.2.2	Návrh úlohy . . . . .	7
2.2.3	Výsledky . . . . .	8
2.3	Lokalizácia cephalometrických landmarkov regresiou heatmáp . . . . .	9
2.3.1	Návrh úlohy . . . . .	9
2.3.2	Experimenty a výsledky . . . . .	11
<b>3</b>	<b>Návrh úlohy</b>	<b>13</b>
3.1	Dataset . . . . .	13
3.2	Metóda lokalizácie landmarkov . . . . .	14
3.3	Trénované modely . . . . .	14
<b>4</b>	<b>Implementácia</b>	<b>16</b>
4.1	Vytvorenie datasetu . . . . .	16
4.2	Augmentácia dát . . . . .	17
4.3	Loss funkcia . . . . .	18
4.4	Trénovanie . . . . .	18
<b>5</b>	<b>Experimenty a výsledky</b>	<b>19</b>
5.1	Trénované modely . . . . .	19
5.2	Modifikácia modelov . . . . .	21
5.3	Testovací dataset . . . . .	23
5.4	Vyhodnotenie modelov . . . . .	24
5.5	Zhodnotenie riešenia . . . . .	26
<b>6</b>	<b>Záver</b>	<b>29</b>
	<b>Literatúra</b>	<b>30</b>

# Kapitola 1

## Úvod

Rýchly rozvoj techniky a strojového učenia v posledných rokoch viedol k tomu, že algoritmy strojového učenia sú čoraz častejšie integrované v rôznych pracovných oblastiach. Jednou z oblastí, v ktorej má využitie strojového učenia veľký potenciál je medicína. Dôvody pre zapojenie strojového učenia do pracovných postupov v medicíne sú napríklad diagnostika a predpovedanie vývoja choroby, eliminovanie chýb spôsobených nedokonalosťou ľudských zmyslov a šetrenie času a tým pádom aj peňanzí. Táto práca je zameraná na posledné dva z uvedených dôvodov.

Cielom mojej práce je lokalizácia landmarkov v 3D dentálnych skenoch pomocou konvolučnej neurónovej siete. Spomínané skeny sú používané pri jednej z metód plánovania ortodontickej liečby. Ortodontickou liečbou sa rozumie úprava chrupu pomocou ortodontického aparátu, t. j. zubného strojčeka. Táto metóda prebieha nasledovne. V špecializovanom softvéri sa zobrazí 3D model chrupu. Následne odborník na modeli ručne označí základné anatomické body. Na základe týchto bodov vie softvér určiť natočenie modelu a pozície jednotlivých zubov. Tieto informácie sú použité pri následnej segmentácii modelu. Segmentáciou sa v tomto prípade rozumie separovanie individuálnych zubov od čeluste tak, aby bolo možné nimi pohybovať a natáčať ich. Odborník potom nesprávne umiestnené zuby popresúva do požadovaných pozícií a na základe porovnania pôvodných pozícií zubov s tými požadovanými dokáže softvér simulovať priebeh ortodontickej liečby, teda postupný pohyb jednotlivých zubov počas nej. Tieto informácie sú potom použité k naplánovaniu samotnej liečby. Ako bolo uvedené vyššie, spomínané anatomické body na skenoch chrupu musia byť označované ručne, čo môže byť zbytočne zdĺhavé, a ak nie sú označené presne, je treba proces opakovať. Cielom tejto práce je lokalizácia práve týchto bodov.

Na riešenie tohoto problému bola natrénovaná neurónová sieť inšpirovaná architektúrou U-Net. Táto architektúra bola pôvodne navrhnutá na segmentáciu obrázkov, kde dosahuje výborné výsledky. S malými úpravami je však možné použiť ju aj na lokalizáciu landmarkov, a to regresiou heatmáp. Sieť bola natrénovaná na 2D obrázkoch, ktoré boli vytvorené ako hĺbkové mapy dentálnych skenov. Tieto obrázky boli následne ručne anotované súradnicami anatomických bodov. Na základe jednotlivých súradníc boli vytvorené obrázky s rovnakým rozlíšením ako hĺbkové mapy vytvorené z dentálnych skenov, pričom tieto obrázky obsahujú 2D Gaussovú krivku so stredom v mieste, kde sa konkrétny landmark nachádza. Toto sú práve vyššie spomínané heatmapy, ktoré sa natrénovaná sieť snaží regresiou vytvoriť. Z výstupov siete boli následne vypočítané súradnice jednotlivých landmarkov.

Kapitola 2 poskytuje zhrnutie niektorých zo súčasných riešení lokalizácie anatomických landmarkov. V kapitolách 3 a 4 je opísaná formulácia úlohy a jej implementácia. Na záver, v kapitole 5, sú vyhodnotené a porovnané natrénované modely.

## Kapitola 2

# Lokalizácia anatomických landmarkov

Cieľom lokalizácie anatomických landmarkov je čo najpresnejšie určiť polohu anatomicky významných bodov v snímkach obsahujúcich časti ľudského tela. Pri riešení tohoto typu úloh môže nastať množstvo problémov.

Najčastejšie sa lokalizácia anatomických landmarkov rieši na fotografiách tváří alebo medicínskych snímkach, či už sú to Röntgenove snímky, CT alebo MRI. Keďže po anatomickej stránke sú si ľudia navzájom často veľmi odlišní, vzhľad landmarkov sa tiež často líši medzi jednotlivými vzorkami. Toto ovplyvňuje presnosť lokalizácie a môže spôsobiť to, že model, ktorý je dobre natrénovaný na jednom datasete bude na inom datasete dosahovať oveľa horšie výsledky. Na výkon modelov na lokalizáciu landmarkov môže negatívne vplyvať aj výber požadovaných landmarkov, keďže niektoré landmarky môžu byť ťažko rozpoznateľné, a tiež požadovaná presnosť modelu.

Táto kapitola obsahuje prehľad súčasných riešení lokalizácie anatomických landmarkov.

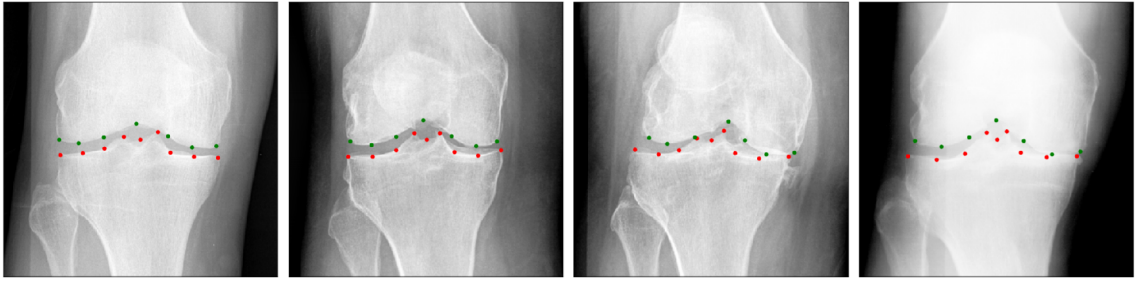
### 2.1 Lokalizácia landmarkov v Röntgenových snímkach kolena pomocou hourglass siete

Tiulpin et al. [18] vo svojej práci ponúka riešenie lokalizácie anatomických landmarkov v Röntgenových snímkach kolena v rôznych štádiách osteoartritídy. Ručná anotácia landmarkov v snímkach kolena nie je jednoduchá úloha, ktorá môže odborníkom zabráť veľmi veľa času. Tento proces sa stáva ešte náročnejším so zvyšujúcou sa závažnosťou osteoartritídy, čo je možné vidieť aj na obrázku 2.1. Na tomto obrázku je možné vidieť ako deformácia kolena so zvyšujúcim sa stupňom osteoartritídy ovplyvňuje celkový vzhľad snímky.

V tejto práci je lokalizácia landmarkov riešená ako regresný problém, kde sú súradnice landmarkov predikované priamo zo snímok kolien.

#### 2.1.1 Architektúra siete

Architektúra modelu použitého v tejto práci je znázornená na obrázku 2.3. Použitý model sa skladá z troch častí: vstupného bloku (entry block), hourglass bloku a výstupného bloku (output block). Namiesto tradičného reziduálneho bottleneck bloku opísaného v [10] je použitý HMP (hierarchical multi-scale parallel) reziduálny blok [3], ktorý zlepšuje priebeh gradientu.

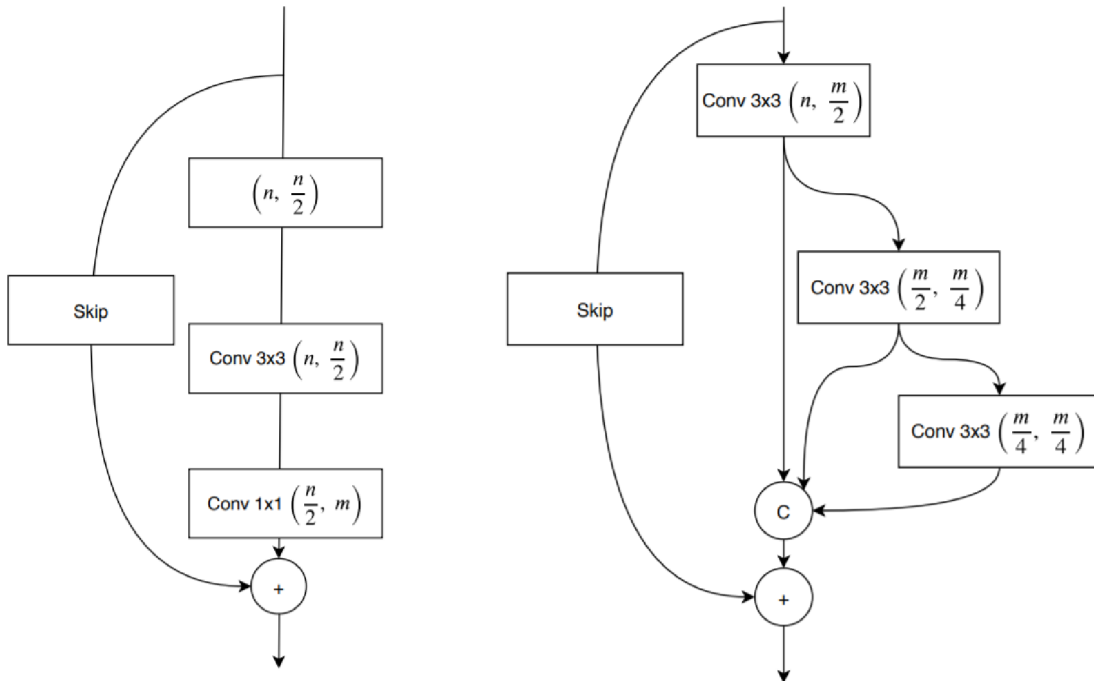


Obr. 2.1: Typické príklady snímok kolena s označenými landmarkami. Na 2. až 4. obrázku zľava sú zobrazené snímky kolien s osteoartritídou so zvyšujúcou sa závažnosťou zľava doprava [18].

### Vstupný blok

Tento blok aplikuje  $7 \times 7$  konvolúciu s krokom 2 a výsledok privedie na vstup reziduálneho bloku. Ďalej je v tomto bloku aplikovaný  $2 \times 2$  max-pooling, ktorého výstup následne prechádza cez ďalšie tri reziduálne bloky pred hourglass blokom.

Tento blok teda umožňuje štvornásobné podvzorkovanie vstupného obrázku a zároveň získanie embeddingov vhodných pre následné spracovanie HMP blokmi.



Obr. 2.2: Grafické znázornenie rozdielu medzi reziduálnym bottleneck blokom [10] a HMP reziduálnym blokom [3, 18], kde  $n$  je počet vstupných a  $m$  počet výstupných príznakových vektorov. Prechod *skip*, ktorý reprezentuje  $1 \times 1$  konvolúciu, je aplikovaný, ak  $n \neq m$ .

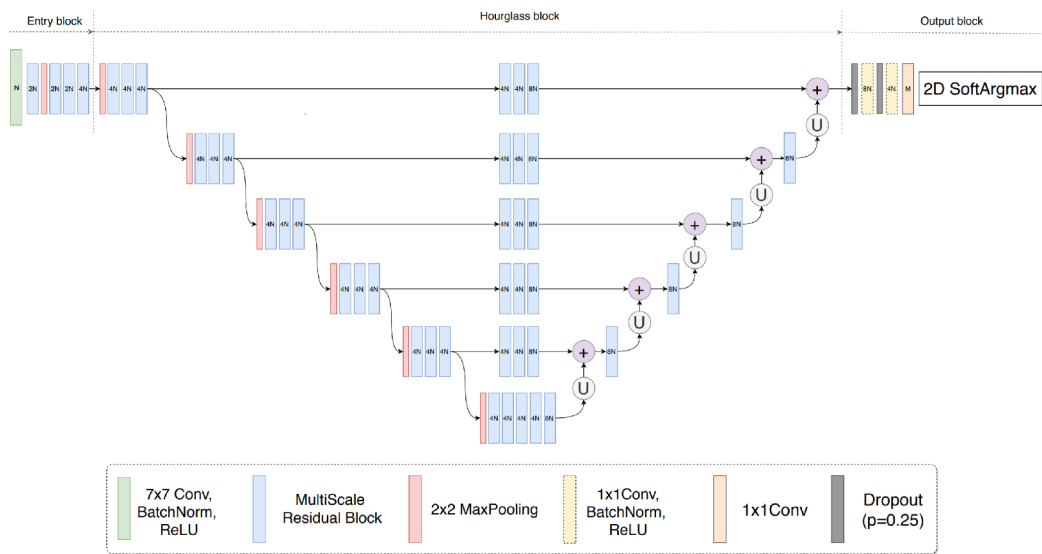


## Hourglass blok

Každá úroveň tohoto bloku začína max-pool operáciou, ktorú nasledujú 3 HMP reziduálne bloky. Ďalej je výstup odovzdávaný do nižšej vrstvy a následne je tiež sčítaný s nadvzorkovaným výstupom nižšej vrstvy. Na najnižšej úrovni je na výstup napojený len ďalší HMP blok, za ktorým nasleduje vyššie spomínané nadvzorkovanie a sčítanie s výstupom podvzorkovacej časti vyššej vrstvy.

## Výstupný blok

Posledný blok má na vstupe výstup hourglass bloku, na ktorý postupne aplikuje dvakrát dropout a  $1 \times 1$  konvolúciu s batch normalizáciou a ReLU aktivačnou funkciou. Nakoniec je na regresiu súradníc použitá  $1 \times 1$  konvolúcia s aktivačnou funkciou soft-argmax [4].



Obr. 2.3: **Architektúra navrhnutej siete [18]**.  $N$  je šírka siete a  $M$  je počet výstupných landmarkov.

### 2.1.2 Loss funkcia

Táto práca využíva wing loss funkciu [9], ktorá je veľmi podobná  $L_1$  loss funkcii. Rozdiel je ten, že chyby v blízkosti 0 –  $(-w, w)$  sú vďaka logaritmickému priebehu funkcie lepšie amplifikované:

$$L(y, \hat{y}) = \begin{cases} w \log(1 + \frac{1}{e}|y - \hat{y}|) & |y - \hat{y}| < w \\ |y - \hat{y}| & |y - \hat{y}| \geq w \end{cases},$$

kde  $y$  je cieľový výstup,  $\hat{y}$  je predikcia,  $(-w, w)$  je rozsah nelineárnej časti funkcie a  $C$  je konštanta, ktorá prepája lineárnu a nelineárnu časť funkcie.

### 2.1.3 Výsledky

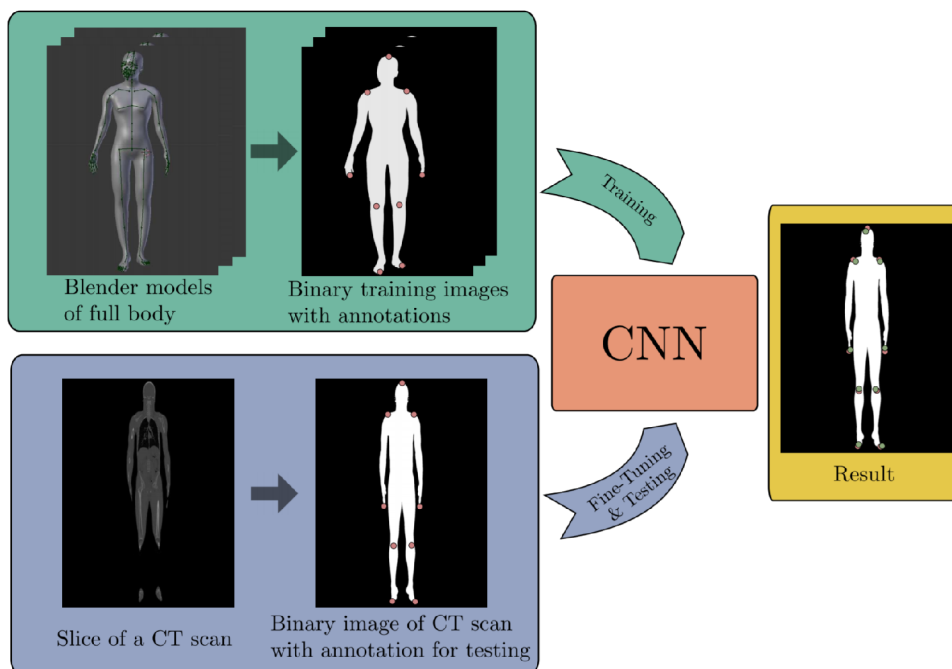
Práca opisuje mnohé variácie modelu a experimenty s nimi. Navyššia dosiahnutá presnosť bola  $90,91 \pm 3.21\%$  predikcií s chybou menšou ako 2,5 mm.

## 2.2 Lokalizácia landmarkov v medicínskych snímkach pomocou umelo vytvorených dát

Práca z roku 2015 od Rieglera et al. [15] sa snaží riešiť pomerne často vyskytujúci sa problém nedostatku obrazových dát pre strojové učenie v oblasti medicíny. Autori tejto práce prišli na spôsob, ktorým je možné umelo vytvárať tréningové dáta na detekciu anatomických landmarkov na snímkach ľudského tela a ruky. To umožňuje prispôbiť počet tréningových dát, ako aj ich vlastnosti, tak, aby boli dosiahnuté požadované kritériá pre riešenie úlohu. Detekcia landmarkov je tu riešená regresiou súradníc priamo zo vstupných obrázkov, ale na rozdiel od práce spomínanej v 2.1, táto kladie väčší dôraz na samotné vytváranie a spracovanie tréningových dát ako na architektúru siete.

### 2.2.1 Vytváranie dát

Boli vytvorené dva datasey. Jeden zo snímkov celého tela a druhý zo snímkov dlane. Oba datasey boli rozdelené na tréningové a ladiace dáta. Na tréning boli použité umelo vytvorené dáta, ktorých bolo vytvorených veľké množstvo aby bola sieť schopná čo najväčšej generalizácie. Na ladenie boli použité dáta vytvorené zo skutočných MR/CT snímkov. Na vytváranie umelých dát bol použitý 3D model MakeHuman [2], z ktorého boli vytvorené jednotlivé obrázky na tréning. Pre každú vzorku boli vytvorené dva obrázky. Prvý obrázok bol binárny s hodnotou 1 v miestach, kde sa telo resp. dľaň nachádza a hodnotou 0 všade inde. Druhý obrázok bol vytvorený ako hĺbková mapa vypočítaná z rovnakého pohľadu ako prvý obrázok. Presná pozícia landmarkov bola definovaná v 3D, čo vďaka mapovaniu z 3D do 2D zaručilo takmer okamžitú a bezchybnú anotáciu k 2D obrázkom.



Obr. 2.4: Prispôsobenie dát pre odhad polohy tela na účely lokalizácie anatomických landmarkov [15].

## 2.2.2 Návrh úlohy

Riešená úloha bola formulovaná ako regresná. Ku každému tréningovému obrázku  $s_n$  bol priradený anotačný vektor  $y_n = (x_1, y_1, \dots, x_k, y_k)$ , ktorý obsahuje  $k$  skutočných pozícií landmarkov. Anotačné vektory boli následne transformované tak, aby súradnice boli v intervale  $[-1, 1]$ :

$$y^p = \left( \frac{x_1 - \frac{p_w}{2}}{p_w}, \frac{y_1 - \frac{p_h}{2}}{p_h}, \dots \right)^T,$$

kde  $p_w$  je šírka a  $p_h$  výška obrázku.

Boli natréňované tri modely na lokalizáciu landmarkov v 2D s architektúrou opísanou nižšie, pričom prvý bol tréňovaný na syntetických dátach a ladený pomocou dát vytvorených z CT/MR skenov, druhý bol tréňovaný len na syntetických dátach a tretí len na dátach z CT/MR skenov. Tieto modely boli tréňované len na binárnych obrázkoch. Ďalej bol natréňovaný model na odhad hĺbky jednotlivých landmarkov s architektúrou zhodnou s vyššie spomenutými modelmi, ktorý bol tréňovaný na umelo vytvorených hĺbkových mapách a ladenej na hĺbkových mapách z CT/MR skenov.

### Dataset

Na tréňovanie bolo vytvorených 2,2 milióna umelých obrázkov dlane a 600 000 obrázkov celého tela. Na účely ladenia bolo vytvorených 132 MR skenov dlane a 20 CT skenov tela. Obrázky dlane boli anotované 5 súradnicami reprezentujúcimi polohy špičiek prstov. V obrázkoch celého tela bolo anotovaných 9 landmarkov. Jeden na vrchu hlavy a dva na ramenách, prostredných prstoch, kolenách a chodidlách. Ako bolo spomenuté v časti 2.2.1 pre každú vzorku bol vytvorený binárny obrázok a hĺbková mapa vytvorená z rovnakého pohľadu.

### Architektúra siete

Natréňovaná bola sieť  $\phi$  s váhami  $\mathbf{w}$  a pri tréňovaní bola použitá  $L_2$  loss funkcia:

$$\mathbf{w} = \arg \min_{\mathbf{w}} \frac{1}{2} \sum_{n=1}^N \|\phi(s_n; \mathbf{w}) - \mathbf{y}_n^P\|^2.$$

Architektúra siete je opísaná v tabuľke 2.1, kde  $\text{conv}w\text{-}c$  reprezentuje konvolučnú vrstvu s veľkosťou filtra  $w \times w$  a  $c$  výstupnými kanálmi,  $\text{maxpool}p$  je max-pooling so šírkou poolu  $p$  a  $\text{fco}$  je plne prepojená vrstva s  $o$  výstupmi. Počet výstupov poslednej vrstvy  $k$  závisí na počte landmarkov, ktoré chceme lokalizovať.

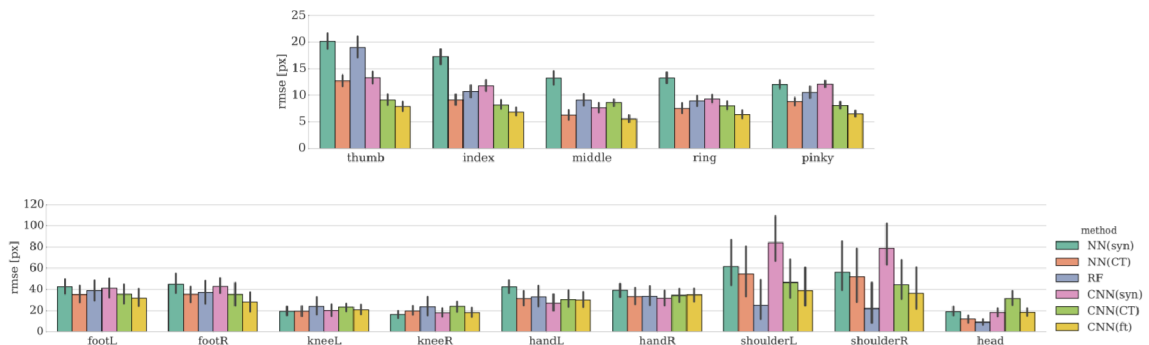
conv5-32
maxpool2
conv5-32
maxpool2
conv3-32
fc2048
fc2048
fc $k$

Tabuľka 2.1: Architektúra siete.

Za každou konvolučnou a plne prepojenou vrstvou bola ako aktivačná funkcia použitá funkcia ReLU [14]. Za každou plne prepojenou vrstvou bol použitý Dropout [17] ako forma regularizácie.

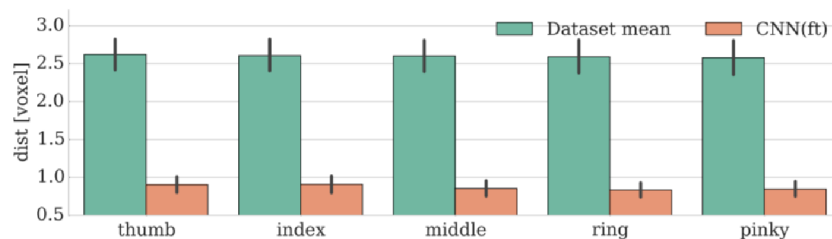
### 2.2.3 Výsledky

Výsledky lokalizácie landmarkov sú znázornené na obrázku 2.7. Pre každý landmark bola vypočítaná stredná odmocnená kvadratická chyba (RMSE) a štandardná odchylka. Na grafe je zobrazená RMSE a štandardná odchylka každého landmarku pre každú použitú metódu. Porovnané sú výsledky siete trénovanej na umelo vytvorených dátach a ladenej na dátach vytvorených z MR/CT snímkov (CNN(ft)) s výsledkami siete trénovanej len na syntetických dátach (CNN(syn)) a siete trénovanej len na binarizovaných obrázkoch z MR/CT (CNN(CT)). Ďalej sú výsledky porovnané s metódou  $k$ -najbližších susedov s parametrom  $k = 3$  (NN(CT) a NN(syn)), kde sú výsledné odhady vypočítané ako stredná hodnota 3 najbližších susedov predikcie.



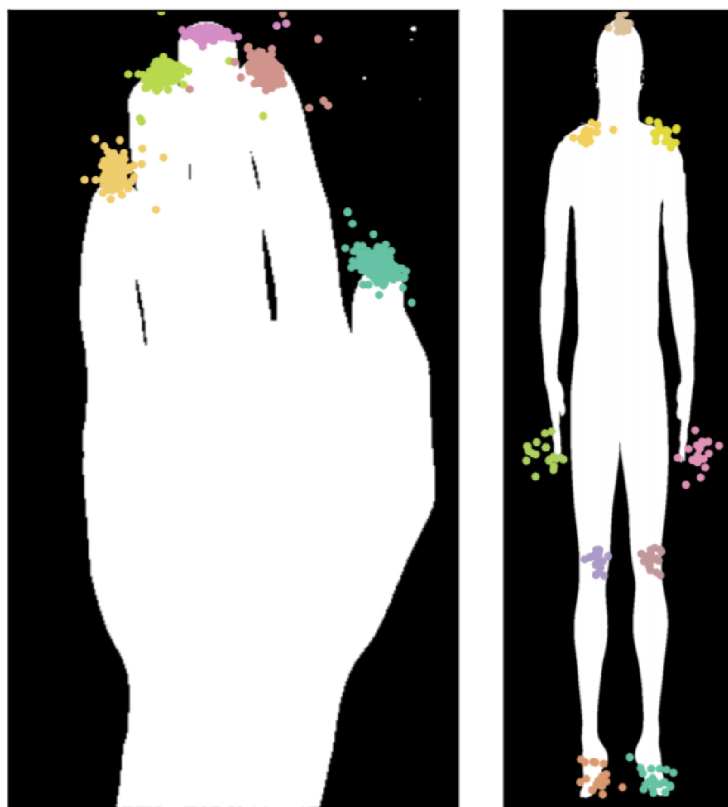
Obr. 2.5: Presnosť lokalizácie landmarkov v MR skenoch dlane (hore) a CT skenoch tela (dole) [15].

Ako prvé je možné si povšimnúť, že NN(syn) dosahuje oveľa horšie výsledky ako NN(CT), čo môže byť pripísané nepresnému modelovaniu syntetických dát. To isté platí aj o CNN(syn) a CNN(ft), ale výsledky z CNN(ft) sú stále lepšie ako výsledky z CNN(CT). Nakoniec sú použité metódy porovnané so state-of-the-art metódou random regression forest (RF) [7].



Obr. 2.6: Výsledky odhadu hĺbky landmarkov v MR skenoch dlane [15].

Výsledky odhadu hĺbky landmarkov dlane sú znázornené na obrázku 2.6, kde sú pre každý landmark porovnané dve hodnoty. Prvá bola vypočítaná ako chyba strednej hĺbky vypočítanej na trénovacích dátach a druhá bola získaná zo siete natrénovanej na odhad hĺbky landmarkov.



Obr. 2.7: Ukážka odhadovaných pozícií landmarkov zobrazená na spoločnom obrázku [15].

## 2.3 Lokalizácia cephalometrických landmarkov regresiou heatmáp

Cephalometrická analýza je proces určovania polohy kľúčových bodov (landmarkov) v cephalogramoch, t .j. v Röntgenových snímkach hlavy. Poskytuje odborníkovi informácie o štruktúre pacientových kostí, zubov a mäkkého tkaniva. Lokalizácia landmarkov je vykonávaná ručne, a to tak, že odborník anotuje cephalogram súradnicami jednotlivých landmarkov. Anotované snímky sú následne použité pri diagnóze, plánovaní liečby a chirurgických zákrokoch. Ručná anotácia snímok zaberie odborníkovi veľa času a môže byť subjektívna. Tieto problémy môžu byť potenciálne odstránené automatickou lokalizáciou landmarkov [19].

Drevický [8] vo svojej diplomovej práci predstavuje riešenie automatickej lokalizácie cephalometrických landmarkov pomocou konvolučnej neurónovej siete.

### 2.3.1 Návrh úlohy

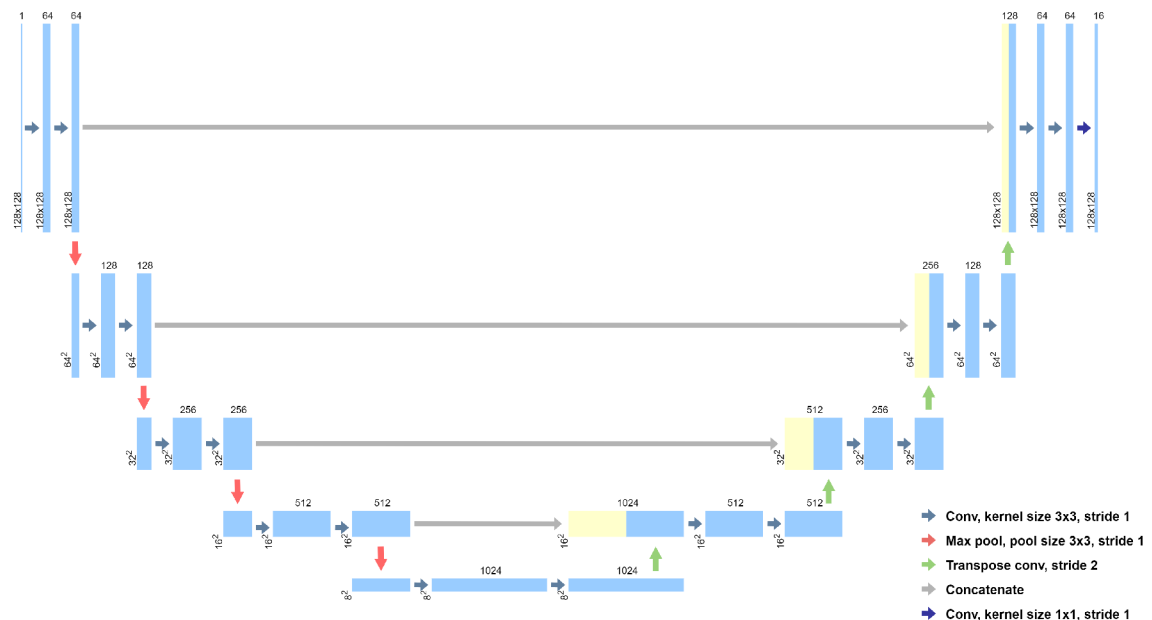
Úloha bola formulovaná ako regresná, kde sa však pozície landmarkov neodhadujú priamo ako súradnice. Na miesto toho bola sieť trénovaná na regresiu viacerých heatmáp, pričom každá heatmapa reprezentuje polohu jedného landmarku.

## Dataset

V tejto práci bol použitý verejne dostupný dataset, ktorý bol zverejnený pre účely Grand Challenge in Dental X-ray Image Analysis [19]. Obsahuje 400 cephalogramov a ku každému cephalogramu anotácie súradníc 19 cephalometrických landmarkov. Pri tréovaní sieť dostane na vstupe cephalogram o veľkosti  $d \times d$  a k nemu zodpovedajúci požadovaný výstup obsahujúci 19 heatmáp vytvorených na základe anotácií daného cephalogramu. Výstup siete má teda rozmery  $19 \times d \times d$ . Každá heatmapa zodpovedá jednému landmarku a obsahuje Gaussovú krivku s fixnou varianciou a amplitúdou a so stredom v mieste, kde sa daný landmark nachádza.

## Architektúra siete

Návrh siete je odvodený od architektúry U-Net [16]. Táto architektúra bola pôvodne navrhnutá na segmentáciu obrazových dát, ale po zmene počtu výstupných kanálov a za použitia vhodnej loss funkcie dosahuje dobré výsledky aj pri lokalizácii landmarkov. Skladá sa z podvzorkovacej časti, ktorá je nasledovaná symetrickou nadvzorkovacou časťou.



Obr. 2.8: Architektúra U-Net.

Podvzorkovanie na veľmi nízke rozlíšenie umožňuje sieti získať informácie o globálnom kontexte obrázku, na základe ktorého je možné zistiť relatívne pozície landmarkov. Sieť obsahuje 4 podvzorkovacie bloky a k nim zodpovedajúce nadvzorkovacie bloky. Podvzorkovací blok sa skladá z dvoch konvolučných vrstiev a jednej max pool vrstvy. Najprv je konvolúciou zvýšená dimenzionalita kanálov vstupu, čo umožňuje modelovať bohatšie príznaky. Nasleduje ďalšia konvolúcia s rovnakým počtom kanálov a max pool vrstva, ktorá zníži rozlíšenie príznakov na polovicu.

Nadvzorkovací blok aplikuje transponovanú konvolúciu na príznakovú mapu z nižšej úrovne, čím zdvojnásobí jej rozlíšenie, a navyše zníži dimenzionalitu kanálov na polovicu. Výsledok tejto operácie je konkatenovaný s mapou príznakov z korešpondujúcej podvzorkovacej vrstvy a následne prechádza cez dve konvolučné vrstvy.

Na výstup posledného nadvzorkovacieho bloku je pripojená ešte jedna konvolučná vrstva, ktorá má jadro s rozmermi  $1 \times 1$  a krok o veľkosti 1. Táto vrstva je zároveň vrstvou výstupnou a jej jediným účelom je znížiť počet kanálov výstupu tak, aby zodpovedal počtu landmarkov ktoré chceme lokalizovať.

Za každou konvolučnou vrstvou bola vložená vrstva na normalizáciu batchu za účelom urýchlenia konvergencie modelu [11].

Pri tréovaní bola ako loss funkcia použitá stredná kvadratická chyba (MSE):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

### 2.3.2 Experimenty a výsledky

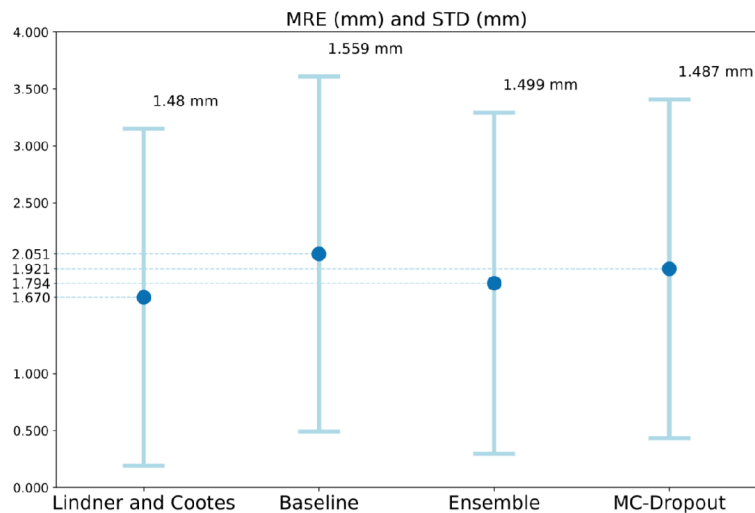
Boli natréované tri modely, ktoré boli založené na architektúre opísanej v 2.3.1. Model **Baseline**, ktorý neobsahoval žiadne modifikácie, model **Ensemble**, ktorý s skladal z 15 paralelne a nezávisle natréovaných instancií modelu Baseline, kde boli výsledné predikcie vypočítané spriemerovaním predikcií všetkých 15 instancií modelu a nakoniec model **Dropout**, ktorý obsahoval Dropout vrstvy [17] na konci každého podvzorkovacieho bloku a na začiatku každého nadvzorkovacieho bloku.

#### Vyhodnotenie modelov

Na vyhodnotenie modelov bola ako metrika použitá priemerná radiálna chyba MRE a jej štandardná odchyľka STD:

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N \sqrt{\Delta x_i^2 + \Delta y_i^2}$$

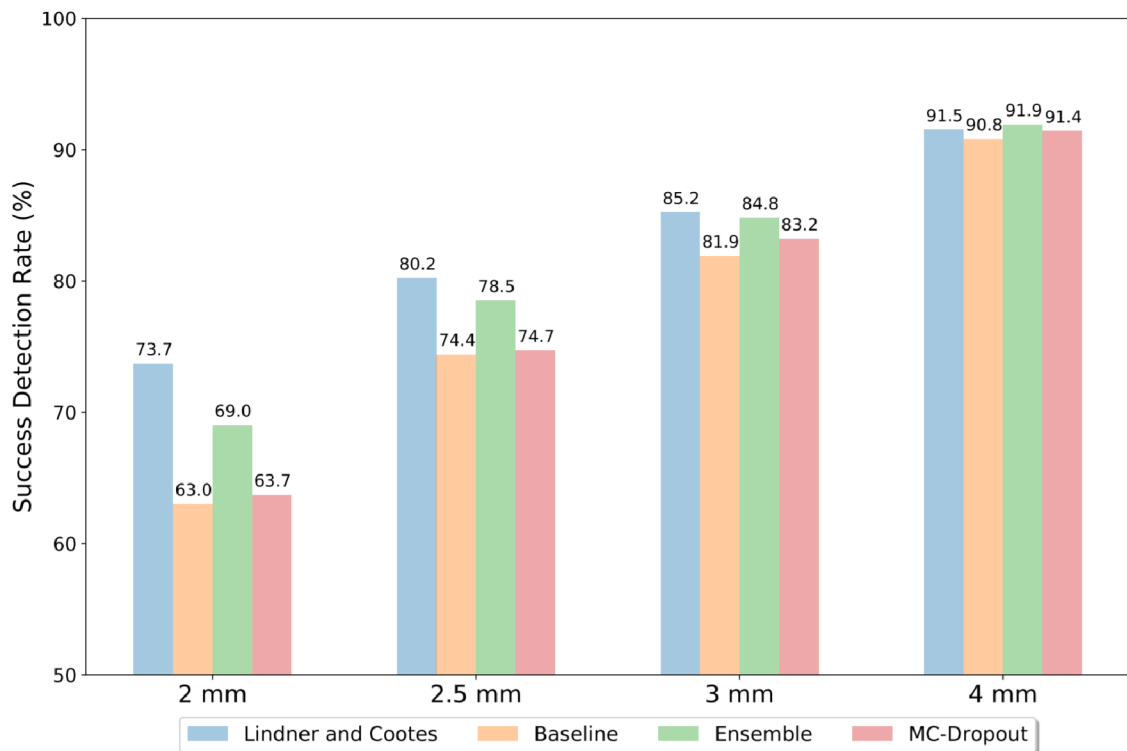
$$\text{STD} = \sqrt{\frac{\sum_{i=1}^N (\sqrt{\Delta x_i^2 + \Delta y_i^2} - \text{MRE})^2}{N}},$$



Obr. 2.9: Radiálna chyba tréovaných modelov porovnaná s najlepšou metódou z Grand Challenge in Dental X-ray Image Analysis (Linder, Cootes) [8, 19].

kde  $\sqrt{\Delta x_i^2 + \Delta y_i^2}$  je euklidovská vzdialenosť medzi skutočnou pozíciou landmarku a predikciou a  $N$  je počet všetkých predikovaných landmarkov.

Ďalej bol ako metrika presnosti modelov vypočítaný podiel úspešných predikcií (angl. Success Detection Rate - SDR), ktorý bol definovaný ako podiel predikcií s radiálnou chybou menšou ako  $z$  mm, kde  $z \in \{2, 2.5, 3, 4\}$ . Najlepšiu presnosť z hľadiska podielu úspešných predikcií dosiahol model Baseline. Najmenšiu priemernú radiálnu chybu dosiahol model Ensemble.



Obr. 2.10: Podiel úspešných predikcií(SDR) tréovaných modelov porovnaný s najlepšou metódou z Grand Challenge in Dental X-ray Image Analysis od dvojice Linder a Cootes [8, 19].



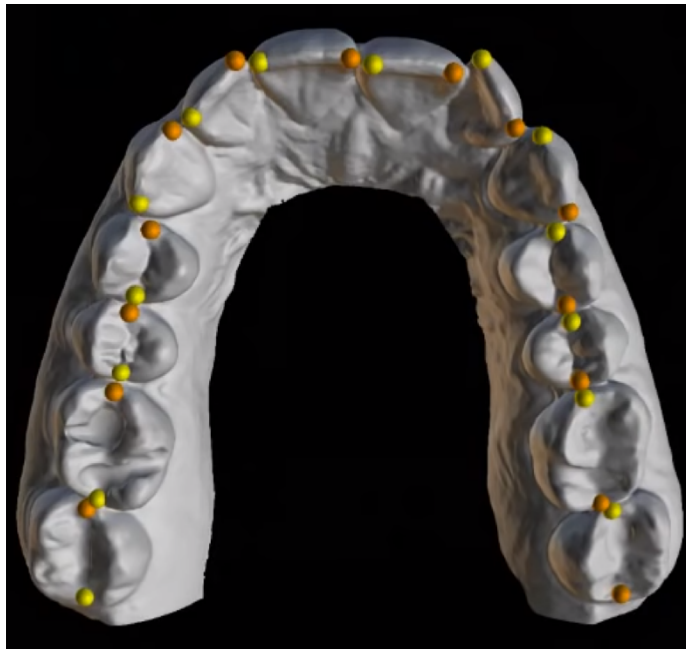
## Kapitola 3

# Návrh úlohy

Cieľom mojej práce bola automatická lokalizácia anatomických landmarkov v 3D skenoch chrupu za použitia konvolučnej neurónovej siete. Motivácia k tejto úlohe je bližšie opísaná v kapitole 1. Táto kapitola opisuje metódy a postupy použité pri riešení mojej úlohy.

### 3.1 Dataset

Dataset použitý na riešenie úlohy mi bol poskytnutý mojim vedúcim bakalárskej práce a skladal sa z 3D polygonálnych modelov chrupu rôznych pacientov. Za účelom zjednodušenia tréningového procesu som sa rozhodol, že lokalizáciu landmarkov nebudem tréňovať priamo na 3D modeloch, ale na 2D hĺbkových mapách vytvorených z daných modelov. Aby bol proces tréningu ešte jednoduchší, snímky (hĺbkové mapy) boli vytvárané z takého pohľadu, aby boli všetky modely približne rovnako natočené. Dataset bol následne rozdelený do dvoch kategórií podľa toho, či sa jednalo o snímku dolnej alebo hornej čeluste.



Obr. 3.1: Dentálny sken s landmarkami označenými odborníkom.

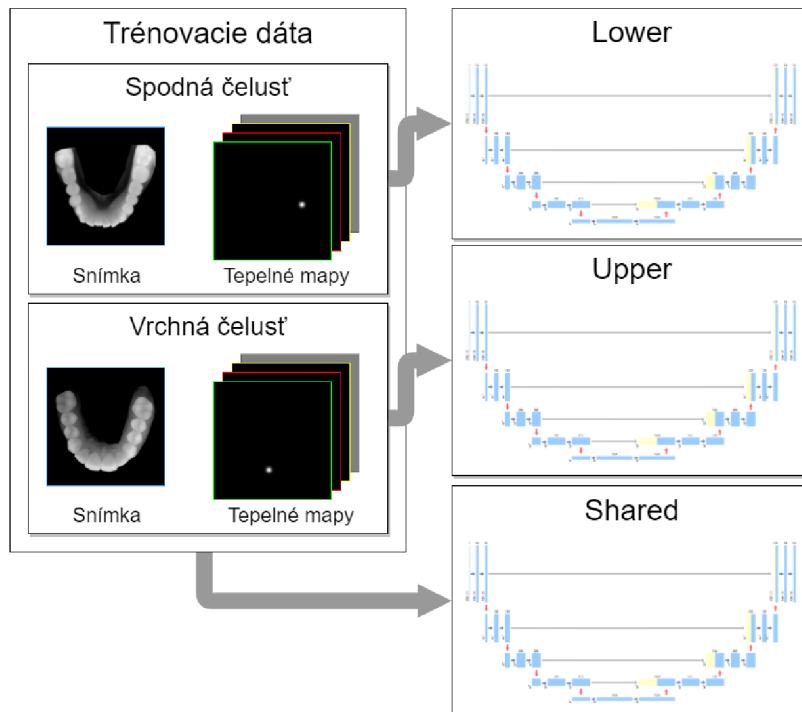
Vytvorené snímky som následne anotoval súradnicami udávajúcimi polohy jednotlivých landmarkov v snímke. V praxi sa na modeloch označuje 32 landmarkov, kde na každý zub pripadajú dva: jeden na pravej a druhý na ľavej strane zubu. Pre jednoduchosť som sa však rozhodol pracovať len zo 16 landmarkami, ktorých pozície boli v strede jednotlivých zubov. V prípade, že niektorý zub chýbal, bola anotácia landmarku zodpovedajúceho danému zubu vynechaná. Jedna vzorka vytvoreného datasetu sa teda skladala z 2D hĺbkovej mapy dentálneho skenu a 16 párov súradníc udávajúcich pozície jednotlivých zubov(landmarkov, ktoré majú byť lokalizované).

### 3.2 Metóda lokalizácie landmarkov

Na lokalizáciu landmarkov som zvolil metódu regresie heatmáp ako v [8]. Vstupom siete je hĺbková mapa vytvorená z 3D modelu chrupu v stupňoch sivej a s veľkosťou  $d \times d$ . Zodpovedajúci cieľový výstup je súbor 16 heatmáp o veľkosti  $d \times d$ . Každá heatmapa zodpovedá jednému landmarku a obsahuje Gaussovú krivku s fixnou varianciou a amplitúdou, ktorá má stred v mieste, kde sa landmark nachádza. Výstupom siete je teda matica  $16 \times d \times d$  skladajúca sa z heatmáp, z ktorých každá obsahuje informáciu o pozícii jedného landmarku.

### 3.3 Trénované modely

Natrénované boli tri modely nasledujúce architektúru U-Net [16], bližšie opísanú v 2.3.1. Model **Shared**, ktorý bol natrénovaný na všetkých tréningových dátach, model **Lower** natrénovaný na snímkach dolných čelustí a model **Upper** natrénovaný na snímkach horných čelustí.



Obr. 3.2: Rozdelenie tréningových dát.

Takéto rozdelenie som zvolil, pretože dolná a horná čelusť vyzerajú odlišne, čo by potenciálne mohlo viesť k nižšej účinnosti tréningu pri modeli tréningu na snímkach oboch čelustí. Výsledky porovnania výkonu tréningových modelov sú opísané v kapitole 5.

Z predpovedí vytvorených natréningovanými sieťami boli následne odhadnuté súradnice landmarkov. Súradnice vypočítané z heatmáp boli porovnané so skutočnými súradnicami landmarkov, na základe čoho bola vyhodnotená presnosť jednotlivých modelov. Detaily o vstupných dátach a architektúre siete sú uvedené nižšie v tejto kapitole.

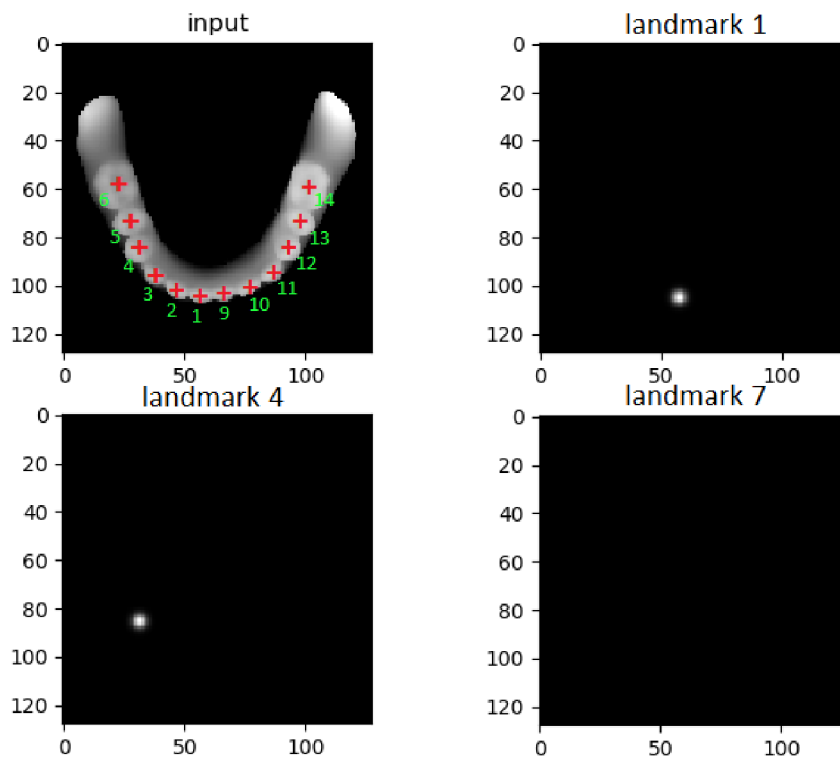
# Kapitola 4

## Implementácia

V tejto kapitole sú uvedené implementačné detaily týkajúce sa predspracovania dát a trénovacieho procesu.

### 4.1 Vytvorenie datasetu

Ako prvé som vytvoril dataset hĺbkových máp 3D modelov chrupu. Ten som vytváral v softvéri MeshLab [6], ktorý obsahuje nástroj na vytváranie hĺbkových máp. Výsledný dataset obsahuje 282 snímok, pričom 118 z nich bolo vytvorených z modelov dolnej čeluste a zvyšných 164 z modelov hornej čeluste. Dataset bol rozdelený do troch podmnožín tak, ako je opísané v 3.3. Takto vytvorené datasety som ďalej rozdelil na trénovacie a validačné dáta v pomere  $r \approx 3 : 1$ .



Obr. 4.1: Ukážka trénovacieho vstupu a jeho cieľových výstupov.

Následne bolo potrebné snímky anotovať súradnicami landmarkov. Na tento účel som vytvoril skript `anotator.py`, ktorý po spustení otvorí okno, v ktorom sa zobrazí neanotovaná hĺbková mapa z vytvoreného datasetu. Anotácia sa následne vykoná klikaním na miesta v obrázku, kde chceme aby sa jednotlivé landmarky nachádzali. Súradnice miesta kliknutia sa následne zapisujú do textového súboru označeného rovnakým identifikátorom ako otvorený obrázok. Po naklikaní všetkých 16 landmarkov sa automaticky otvorí ďalší neanotovaný obrázok a postup sa opakuje kým nie je anotovaný celý dataset.

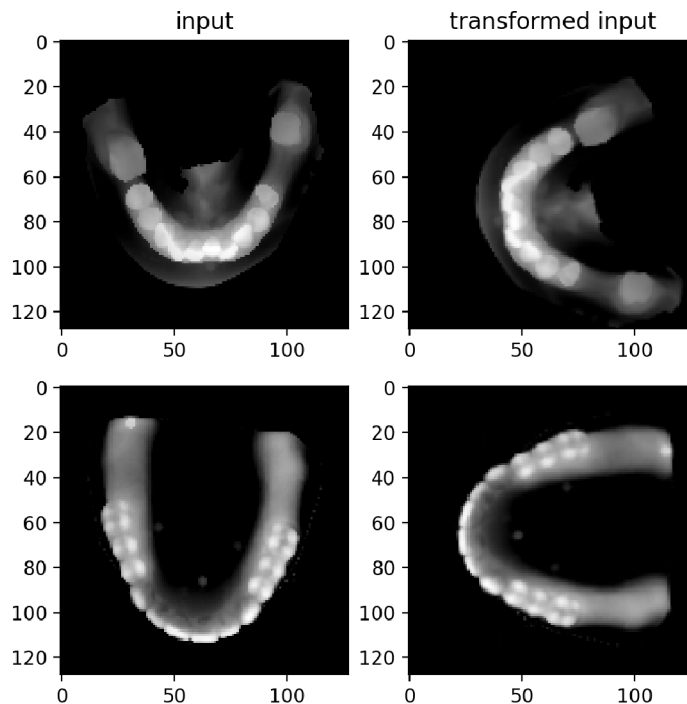
Keďže som úlohu riešil metódou regresie heatmap, bolo ešte potrebné transformovať anotácie snímok na heatmápy. Na to som vytvoril funkciu `generate_heatmaps`, ktorá sa nachádza v skripte `data_generator.py`. Táto funkcia postupne prechádza súbory obsahujúce anotácie snímok a pre každú snímku vytvorí 16 obrázkov, ktoré obsahujú informáciu o pozícii jednotlivých landmarkov tak, ako je opísané v 3.2.

## 4.2 Augmentácia dát

Keďže použitý dataset bol relatívne malý, bola naň za účelom zväčšenia počtu trénovacích vzoriek a zvýšenia variability aplikovaná augmentácia. Z každej snímky bola vytvorená jedna ďalšia, na ktorú boli aplikované nasledujúce transformácie:

- **Translácia:** Vertikálna aj horizontálna v intervale  $\langle -3, 3 \rangle$  pixely
- **Škálovanie:** Vertikálne aj horizontálne v intervale  $\langle -0.95, 1.05 \rangle$
- **Rotácia:** v intervale  $\langle -5, 5 \rangle$  stupňov, následne bola na každú snímku aplikovaná rotácia o 90, 180 alebo 270 stupňov, aby sa sieť dokázala naučiť rozpoznávať príznaky na rôzne natočených modeloch

Augmentácia bola aplikovaná na vstupné snímky aj na cieľové heatmápy.



Obr. 4.2: Príklad augmentácie vstupných obrázkov.

### 4.3 Loss funkcia

Keďže ide o regresnú úlohu, rozhodol som sa ako loss funkciu použiť strednú kvadratickú chybu MSE:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

kde  $y_i$  sú cieľové heatmapy,  $\hat{y}_i$  sú predikcie modelu a  $n$  je veľkosť batchu. Keďže heatmapy obsahujú nenulové hodnoty iba v malej oblasti okolo pozície landmarku, keď bola amplitúda Gaussovej krivky nastavená na 1, sieť sa sústredila na predikciu oveľa väčšej nulovej oblasti a Gaussovú krivku v mieste landmarku ignorovala. To viedlo k veľmi pomalej konvergencii tréningového procesu. Aby sa tomuto predišlo, amplitúda Gaussovej krivky bola nastavená na 1000 ako je navrhované v [8].

### 4.4 Trénovanie

Na zostavenie modelov neurónovej siete, ako aj na ich vyhodnotenie, som vytvoril funkcie, ktoré sa nachádzajú v skripte `model.py`. Samotné tréningovanie a následné vyhodnotenie modelov je implementované v skriptoch `train.py` a `landmark_detection.py`. Na vytváranie modelov, ich tréningovanie a testovanie som použil Keras API [5] s Tensorflow backendom [1].

Všetky tri modely som trénoval po dobu 100 epôch za použitia optimalizátora Adam [12]. Najrýchlejšia konvergencia bola dosiahnutá pri koeficiente učenia  $lr = 10^{-3}$  a veľkosti batcha 8.

## Kapitola 5

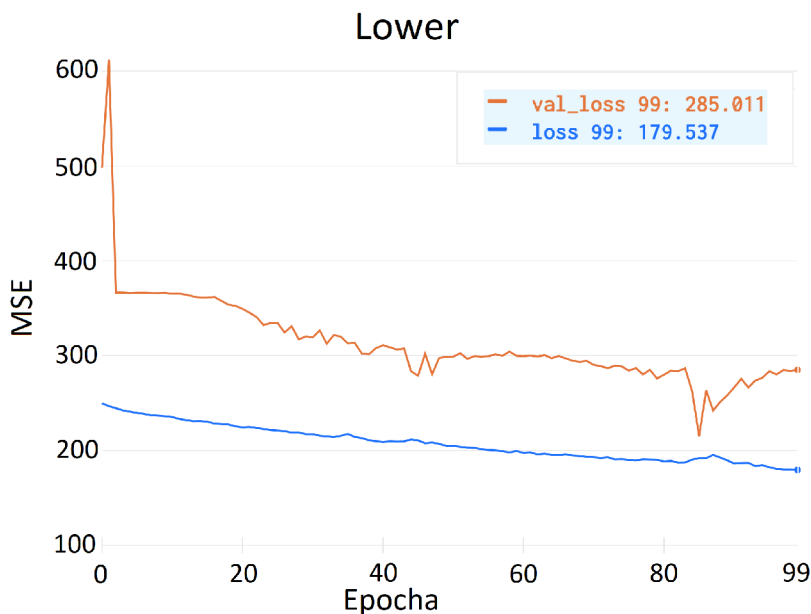
# Experimenty a výsledky

Táto kapitola obsahuje detaily o natrénovaných modeloch, ich vyhodnotenie a porovnanie presností. V závere tejto kapitoly sa zaoberám faktormi, ktoré mohli negatívne ovplyvniť presnosť modelov a predkladám návrhy na riešenie.

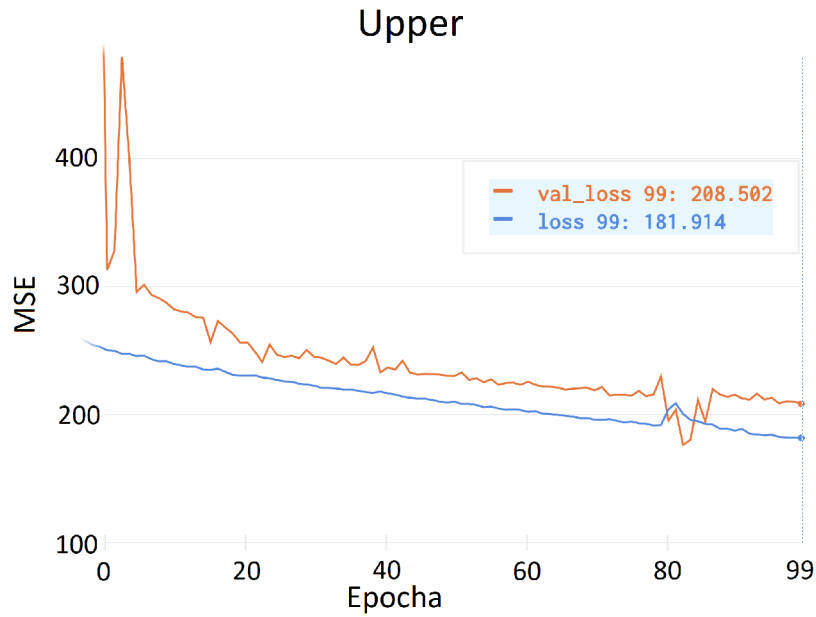
### 5.1 Trénované modely

Táto a nasledujúca sekcia sa zaoberajú krížovou validáciou natrénovaných modelov s ich modifikovanými variantmi. Na základe krížovej validácie boli vybrané najlepšie modely, ktoré boli neskôr použité pre ďalšie experimenty opísané nižšie.

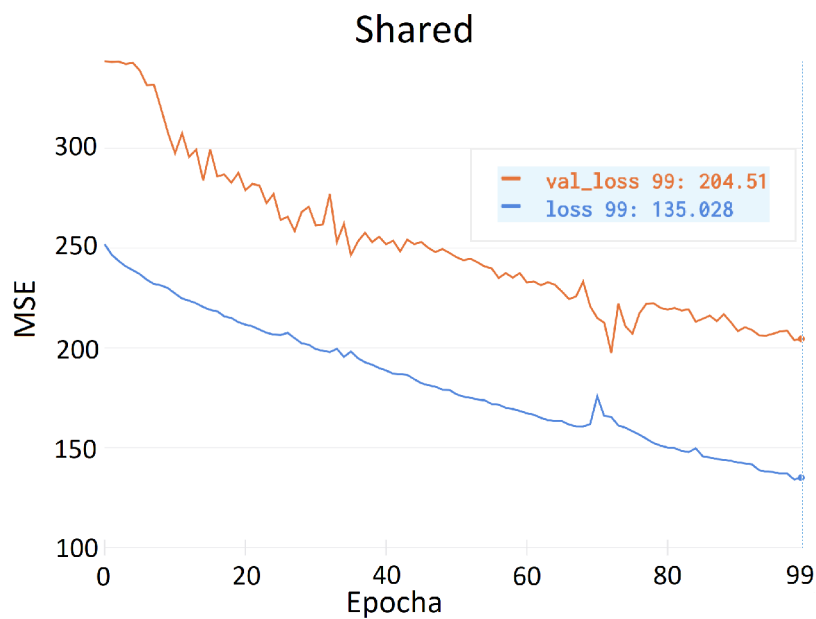
Najprv boli natrénované tri modely, ktoré presne sledovali architektúru opísanú v časti 2.3.1. Ako je uvedené v kapitole 3, model Lower bol natrénovaný na snímkach dolných čelustí, model Upper bol natrénovaný na snímkach horných čelustí a model Shared na snímkach oboch čelustí. Trénovanie prebiehalo tak, ako je opísané v kapitole 4.



Obr. 5.1: Priebek loss funkcie modelu Lower počas trénovania.



Obr. 5.2: Priebek loss funkcie modelu Upper počas tréovania.



Obr. 5.3: Priebek loss funkcie modelu Shared počas tréovania.

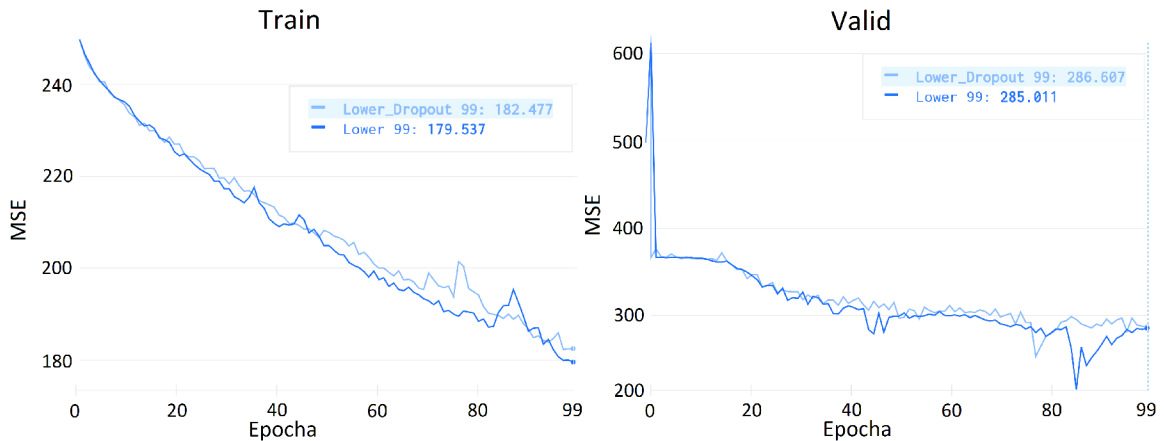
Pri analýze priebehu loss funkcie počas tréovania bolo zistené, že pri všetkých troch modeloch došlo k overfittingu. To znamená, že sa modely príliš konkretizovali na tréovacie dáta, následkom čoho dosahuje validačná loss funkcia výrazne vyššie hodnoty ako tá, ktorou sa modely optimalizujú. Najväčšia miera overfittingu bola zaznamenaná na modeli Shared, aj napriek tomu, že tento model bol tréovaný na najväčšom počte dát. Za ním nasleduje model Lower a najmenšia, takmer zanedbateľná miera overfittingu bola zaznamenaná pri modeli Upper.



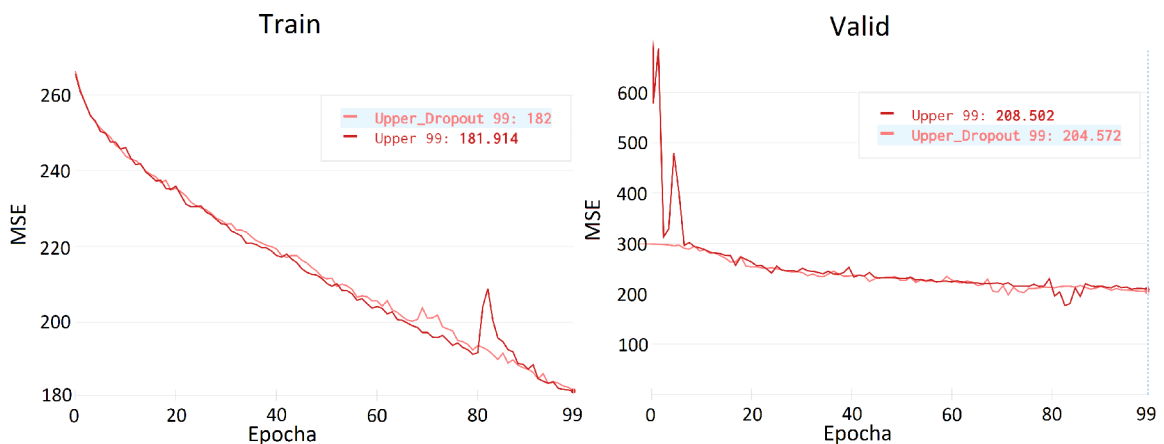
## 5.2 Modifikácia modelov

Za účelom obmedzenia overfittingu boli následne modely modifikované, a to dvomi rôznymi metódami.

Prvou metódou bolo vloženie dropout vrstiev do architektúry siete, ako bolo navrhnuté v [17]. Dropout vrstvy boli vložené pred každou podvzorkovacou operáciou a za každou nadvzorkovacou operáciou. Podiel vynechaných neurónov bol nastavený na 0.4.

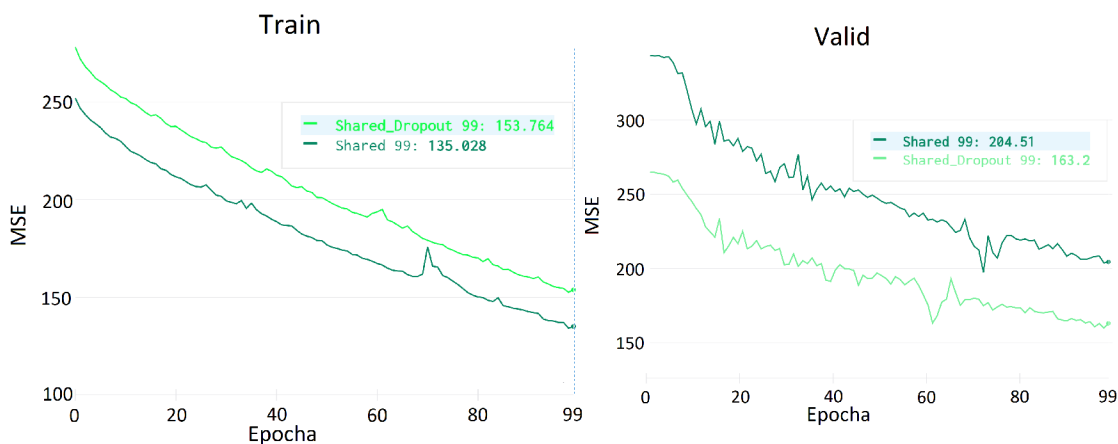


Obr. 5.4: Porovnanie priebehov loss funkcií počas tréningu modelu Lower s a bez dropout vrstiev.



Obr. 5.5: Porovnanie priebehov loss funkcií počas tréningu modelu Upper s a bez dropout vrstiev.

Po natrénovaní takto modifikovaných modelov na pôvodných dátach bolo analýzou zistené, že Dropout vrstvy nemali na výsledky modelov Lower a Upper takmer žiadny vplyv. V prípade modelu Shared sa hodnoty loss funkcie síce o niečo zvýšili, ale zároveň sa dosiahlo zníženie validačnej loss funkcie, čiže overfitting bol čiastočne odstránený.

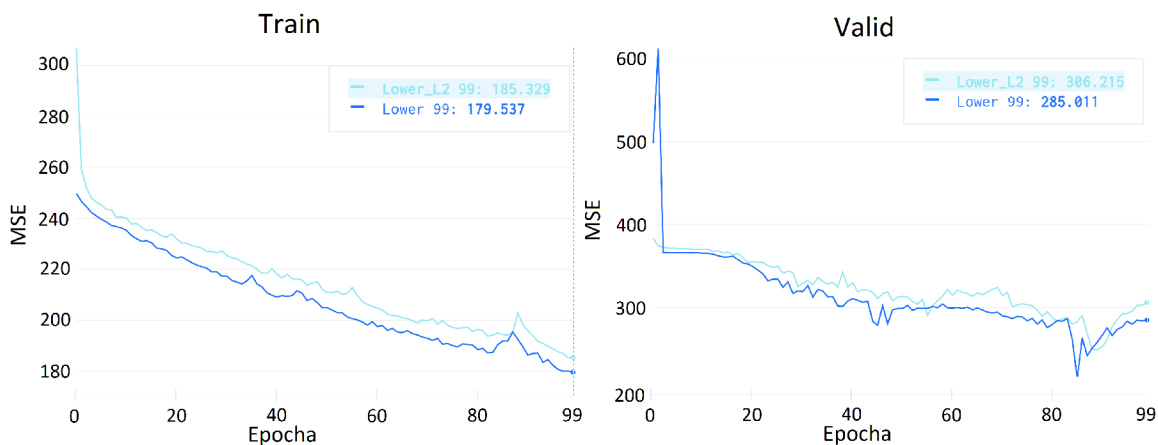


Obr. 5.6: Porovnanie priebehov loss funkcií počas tréovania modelu Shared s a bez dropout vrstiev.

Druhou metódou bola L2 regularizácia váh [13] inak nazývaná aj hrebeňová regresia. Táto metóda upravuje hodnotu loss funkcie pri jej výpočte. Úprava spočíva v tom, že sa k loss funkcii pri jej výpočte pripočíta súčet druhých mocnín súčasných váh, vynásobený koeficientom  $\lambda$ :

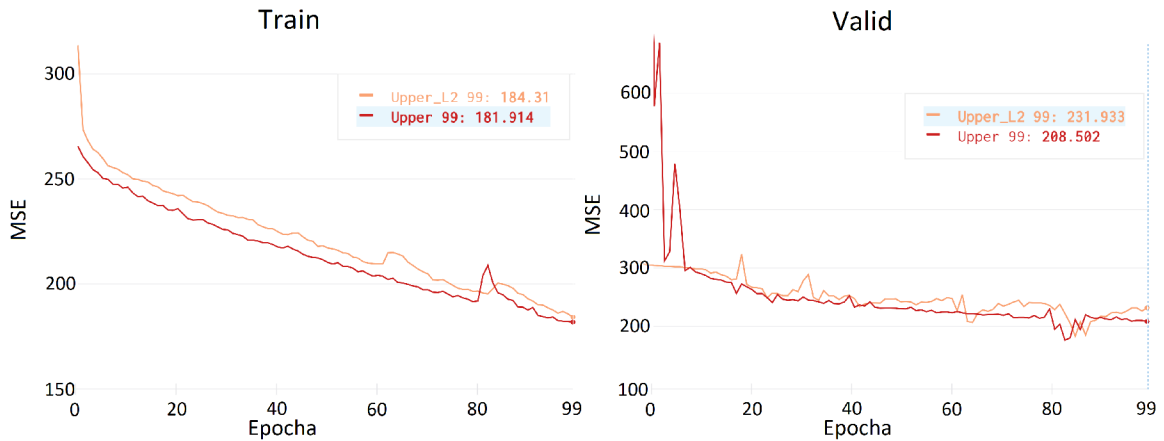
$$Loss = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \frac{\lambda}{n} \sum_{i=1}^n w_i^2$$

Pre experimenty bola zvolená hodnota koeficientu  $\lambda = 2 * 10^{-2}$ . Regularizácia bola aplikovaná na všetky konvolučné vrstvy.

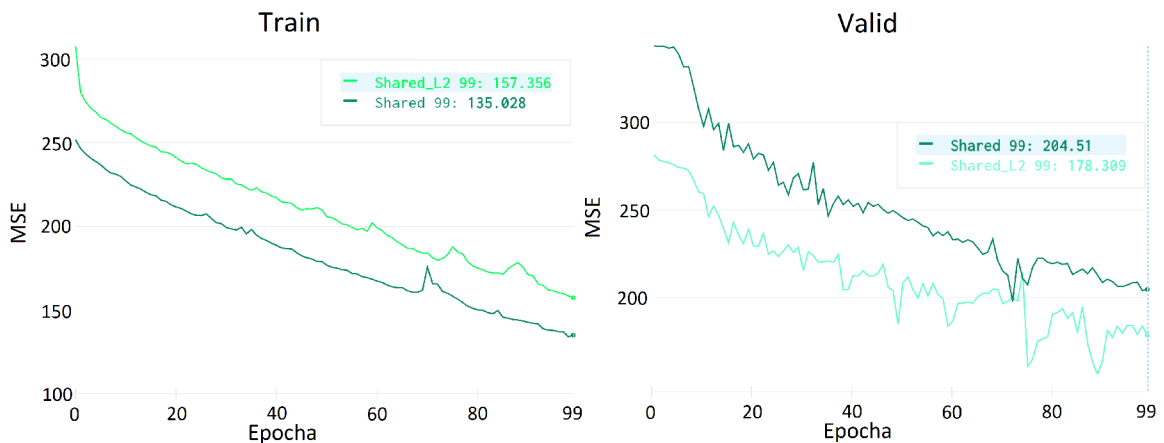


Obr. 5.7: Porovnanie priebehu loss funkcií počas tréovania modelu Lower s a bez L2 regularizácie.

Po ukončení tréovania bolo zistené, že hrebeňová regresia mala na presnosť modelov Lower a Upper len negatívny dopad. Tréovacia aj validačná loss funkcia oboch modelov dosahovali vyššie hodnoty ako modely bez regularizácie váh. Na model Shared mala L2 regularizácia podobný vplyv ako pridanie dropout vrstiev, ale pri L2 regularizácii dosahovala loss funkcia vyšších hodnôt.



Obr. 5.8: Porovnanie priebehu loss funkcií počas tréovania modelu Upper s a bez L2 regularizácie.



Obr. 5.9: Porovnanie priebehu loss funkcií počas tréovania modelu Shared s a bez L2 regularizácie.

Na základe krížovej validácie boli pre použitie v nasledujúcich experimentoch vybrané najlepšie modely. Keďže pridanie dropout vrstiev do modelov Lower a Upper nijako nezlepšilo výkon týchto modelov a L2 regularizácia ich výkon zhoršila, boli pre nasledujúce experimenty vybrané varianty týchto modelov nasledujúce pôvodný návrh siete. V prípade modelu Shared aplikácia oboch modifikácií spôsobila o niečo vyššiu hodnotu tréovacej loss funkcie, ale pomohla znížiť validačnú loss funkciu. Rozhodujúce teda boli samotné hodnoty loss funkcií. Na základe nich bol pre nasledujúce experimenty zvolený variant modelu Shared s Dropout vrstvami, ktorý dosahoval nižšie hodnoty oboch loss funkcií.

### 5.3 Testovací dataset

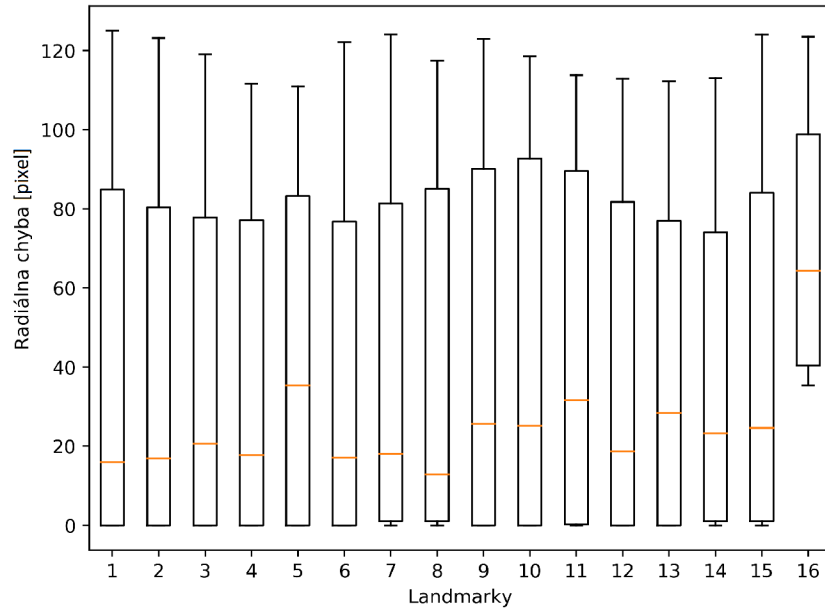
Keďže z poskytnutého datasetu bolo možné zhotoviť len relatívne málo snímok, na testovanie boli použité tie isté dáta, ako na tréovanie a validáciu. Aby však bol zaistený určitý stupeň variability medzi tréovacím a testovacím datasetom, testovacie dáta boli tiež augmentované podobným spôsobom ako dáta tréovacie.

## 5.4 Vyhodnotenie modelov

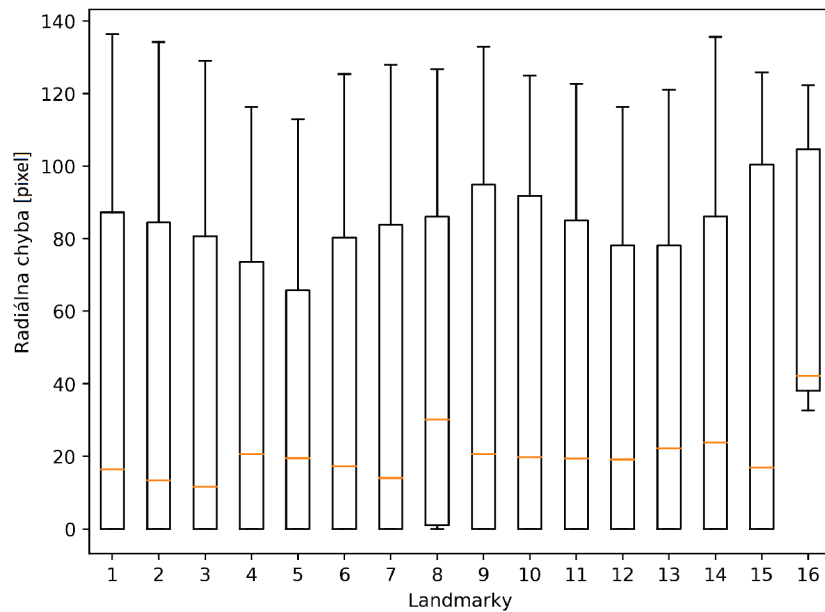
Na vyhodnotenie jednotlivých modelov boli použité dve metriky. Radiálna chyba  $R$ , ktorá počíta euklidovskú vzdialenosť medzi skutočnou pozíciou landmarku a tou predikovanou:

$$R = \sqrt{\Delta x^2 + \Delta y^2},$$

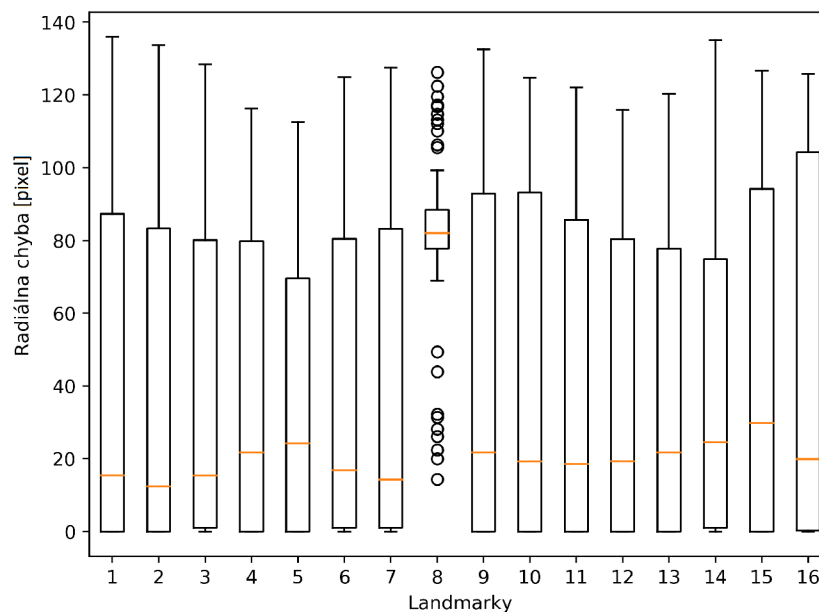
kde  $\Delta x$  a  $\Delta y$  sú vzdialenosti predpovedanej a skutočnej pozície landmarku na osiach  $x$  a  $y$ . Táto funkcia bola vypočítaná pre každý landmark každej predikcie zvlášť. Na základe výsledkov bola získaná presnosť predpovedí modelov pre jednotlivé landmarky.



Obr. 5.10: Radiálna chyba modelu Lower na jednotlivých landmarkoch.



Obr. 5.11: Radiálna chyba modelu Shared na jednotlivých landmarkoch.

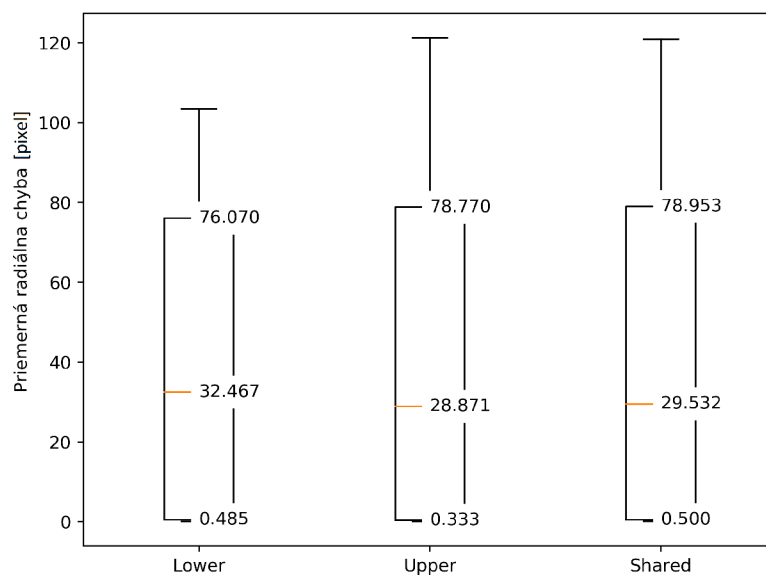


Obr. 5.12: Radiálna chyba modelu Shared na jednotlivých landmarkoch.

Druhou použitou metrikou bola priemerná radiálna chyba  $AR$ , ktorá na rozdiel od vyššie spomínanej funkcie vracia priemer radiálnych chýb cez všetky landmarky jednej predikcie:

$$AR = \frac{1}{n} \sum_{i=1}^n (\sqrt{\Delta x_i^2 + \Delta y_i^2}),$$

kde  $n$  je dimenzionalita výstupu siete, teda počet landmarkov na jednej snímke. Táto funkcia bola vypočítaná pre každú predikciu zvlášť. Na základe výsledkov bola získaná celková presnosť predpovedí jednotlivých modelov.



Obr. 5.13: Priemerná radiálna chyba jednotlivých modelov.

Pre obidve metriky boli odhadované pozície landmarkov vypočítané ako súradnice bodu na heatmape s najvyššou amplitúdou:

$$\hat{x}, \hat{y} = \underset{\hat{x}, \hat{y} \in \{0, \dots, d-1\}}{\operatorname{arg\,max}} f(x, y),$$

kde  $d$  je veľkosť heatmapy a  $f(x, y)$  je funkcia, ktorá vracia amplitúdu predikovanej heatmapy v súradniciach  $x$  a  $y$ .

Na záver bola vyhodnotená úspešnosť modelov. Tá bola vypočítaná ako podiel úspešných predikcií zo všetkých predikcií modelu. Ako kritérium úspešnosti predikcie bolo stanovené, že radiálna chyba predikcie nesmie presiahnuť hodnotu 2 pixely. V praxi to znamená, že úspešná predikcia môže byť od skutočnej polohy landmarku vzdialená najviac o jeden pixel vertikálne aj horizontálne.

Model	Presnosť
Lower	0.4618
Upper	0.4768
Shared	0.4694

Tabuľka 5.1: Presnosť modelov

## 5.5 Zhodnotenie riešenia

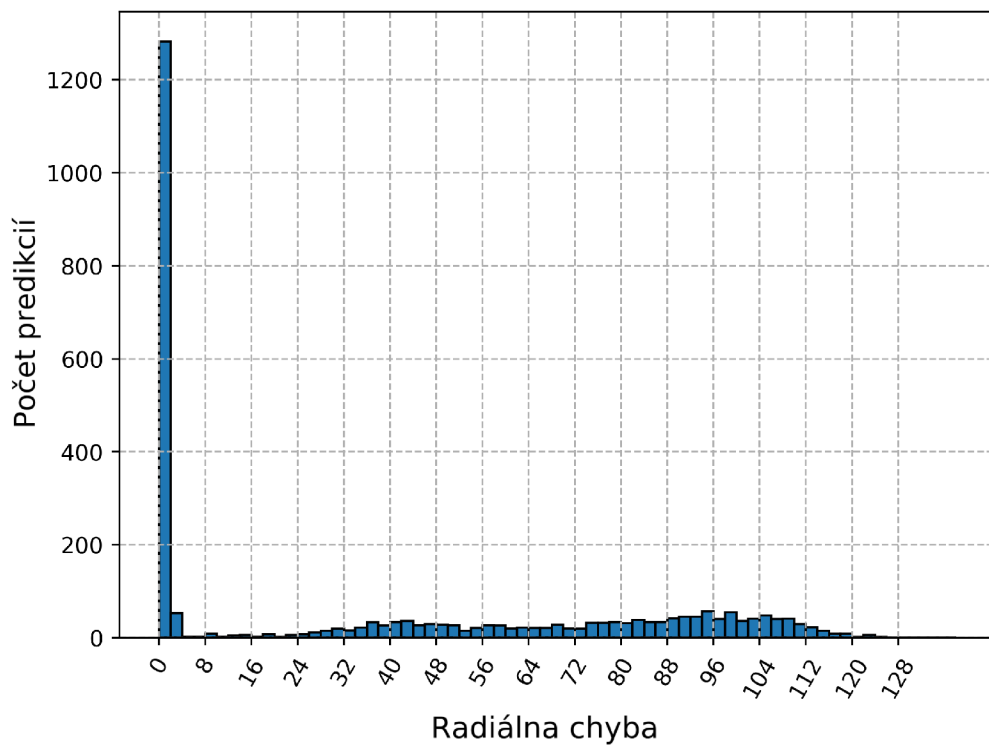
Keďže všetky natrénované modely dosiahli úspešnosť menšiu ako 50%, riešenie navrhnuté v tejto práci nie je použiteľné v praxi. Navyše, v tejto práci nebolo riešené spätné mapovanie pozícií landmarkov z 2D obrázkov do 3D modelov. Na presnosť modelov mohli mať vplyv viaceré faktory. Ako vidno v tabuľke 5.2, dataset obsahuje veľké množstvo snímok, v ktorých niektoré zuby chýbajú.

Landmark	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Dolná čelusť	10	9	3	19	32	80	70	149	7	9	2	20	31	83	76	165
Horná čelusť	22	21	10	33	45	63	102	214	25	19	17	40	40	68	118	230

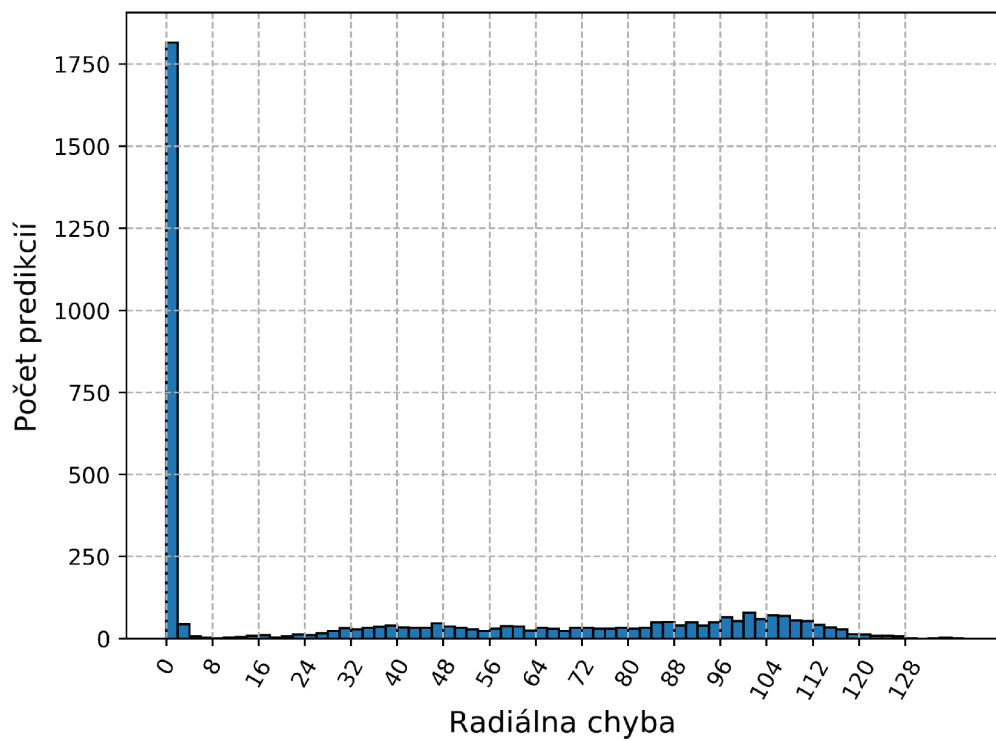
Tabuľka 5.2: Počty tréningových dát (po augmentácií), na ktorých chýbajú jednotlivé zuby

Keď tieto počty porovnáme s radiálnou chybou na daných zuboch, zistíme, že tieto hodnoty sú v niektorých prípadoch korelované, čo najviac vidieť pri landmarku č. 16. Za toto však môže aj malá veľkosť tréningového datasetu. V prípade väčšieho datasetu by bolo možné dosiahnuť väčšiu variabilitu dát a viac obrázkov, na ktorých by sa jednotlivé zuby nachádzali, čo by pomohlo aj pri predikcií často chýbajúcich zubov.

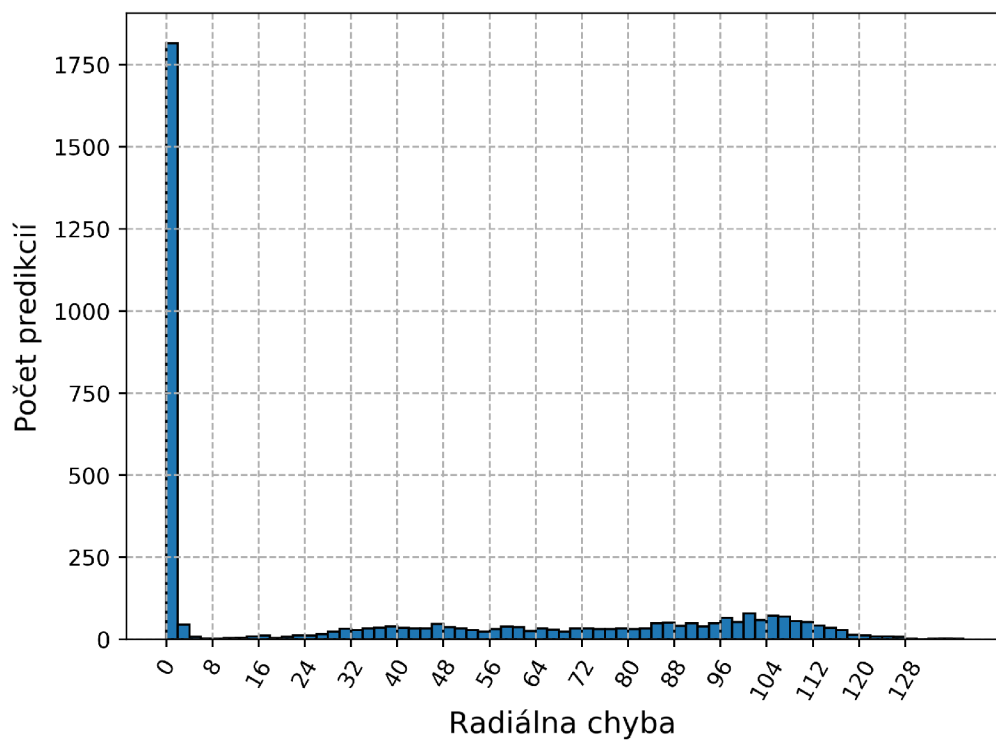
Zvláštnosťou predikcií všetkých modelov je rozloženie veľkostí ich chyby. Aj napriek tomu, že sa mediány radiálnej chyby jednotlivých modelov pohybujú až okolo hodnoty 30, ako je možné vidieť na obrázku 5.13, viac ako 46% predikcií dosahuje chybu  $R < 2$ . To platí pre všetky tri modely, ako je možné vidieť v tabuľke 5.1.



Obr. 5.14: Rozloženie veľkosti chyby modelu Lower.



Obr. 5.15: Rozloženie veľkosti chyby modelu Upper.



Obr. 5.16: Rozloženie veľkosti chyby modelu Shared.



## Kapitola 6

# Záver

V tejto práci som riešil problém detekcie landmarkov v hĺbkových mapách 3D skenov chrupu, ktoré sa používajú pri digitálnej ortodoncii alebo digitálnom návrhu zubných koruniek.

Problém bol riešený regresiou heatmáp, kde na každý landmark pripadá jedna heatmapa, ktorá obsahuje 2D Gaussovú krivku so stredom v mieste landmarku, ktorému prislúcha. Boli natrénované tri modely nasledujúce architektúru U-Net, každá natrénovaná na inom datase. Následne boli na modely aplikované dve rôzne optimalizácie za účelom zvýšenia presnosti modelov. Ako prvé boli do modelov pridané dropout vrstvy. Tie mali pozitívny vplyv iba na model Shared, ktorý bol trénovaný na všetkých tréningových dátach. Druhou aplikovanou optimalizáciou bola L2 regularizácia. Tá mala opäť pozitívny vplyv iba na model Shared, nie však až taký výrazný ako pridanie dropout vrstiev.

Ako najlepšie modely boli teda vybraté model Lower bez vyššie spomínaných optimalizácií, ktorý bol trénovaný na snímkach dolnej čeluste, model Upper tiež bez modifikácií, ktorý bol trénovaný na snímkach horných čelustí a model Shared s dropout vrstvami. Po vytvorení predikcií z testovacieho datasetu bola spočítaná radiálna chyba jednotlivých predikcií, t. j. euklidovská vzdialenosť odhadovaných súradníc od skutočnej polohy landmarku. Na základe hodnôt tejto chyby bola vypočítaná presnosť natrénovaných modelov. Na záver bola vykonaná krátka analýza výsledkov.

Výsledok tejto práce by mohol byť vylepšený natrénovaním modelov na rozsiahlejšom datase. Tým by sa zvýšila presnosť predikcií. Aby toto riešenie mohlo byť použité v praxi, bolo by potrebné použiť spoľahlivý spôsob mapovania súradníc landmarkov z 2D snímok naspäť do 3D modelov.

# Literatúra

- [1] ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z. et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Software available from tensorflow.org. Dostupné z: <https://www.tensorflow.org/>.
- [2] BASTIONI, M., RE, S. a MISRA, S. Ideas and methods for modeling 3D human figures: the principal algorithms used by MakeHuman and their implementation in a new approach to parametric modeling. In: Január 2008, s. 10. DOI: 10.1145/1341771.1341782.
- [3] BULAT, A. a TZIMIROPOULOS, G. Binarized Convolutional Landmark Localizers for Human Pose Estimation and Face Alignment with Limited Resources. *CoRR*. 2017, abs/1703.00862. Dostupné z: <http://arxiv.org/abs/1703.00862>.
- [4] CHAPELLE, O. a WU, M. Gradient descent optimization of smoothed information retrieval metrics. *Information Retrieval*. Jun 2010, zv. 13, č. 3, s. 216–235. DOI: 10.1007/s10791-009-9110-3. ISSN 1573-7659. Dostupné z: <https://doi.org/10.1007/s10791-009-9110-3>.
- [5] CHOLLET, F. et al. *Keras* [<https://github.com/fchollet/keras>]. GitHub, 2015.
- [6] CIGNONI, P., CALLIERI, M., CORSINI, M., DELLEPIANE, M., GANOVELLI, F. et al. MeshLab: an Open-Source Mesh Processing Tool. In: *Eurographics Italian Chapter Conference*. 2008.
- [7] CRIMINISI, A., SHOTTON, J., ROBERTSON, D. a KONUKOGLU, E. Regression Forests for Efficient Anatomy Detection and Localization in CT Studies. In: MENZE, B., LANGS, G., TU, Z. a CRIMINISI, A., ed. *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, s. 106–117. ISBN 978-3-642-18421-5.
- [8] DREVICKÝ, D. *Deep Learning Model Uncertainty in Medical Image Analysis*. Brno, CZ, 2019. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: <https://www.fit.vut.cz/study/thesis/22094/>.
- [9] FENG, Z., KITTLER, J., AWAIS, M., HUBER, P. a WU, X.-J. Wing Loss for Robust Facial Landmark Localisation with Convolutional Neural Networks. In: *Proceedings - 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2018*. United States: IEEE Computer Society, December 2018, s. 2235–2245. DOI: 10.1109/CVPR.2018.00238. © 2018, The Author(s). ; 31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2018 ; Conference date: 18-06-2018 Through 22-06-2018.

- [10] HE, K., ZHANG, X., REN, S. a SUN, J. Deep Residual Learning for Image Recognition. *CoRR*. 2015, abs/1512.03385. Dostupné z: <http://arxiv.org/abs/1512.03385>.
- [11] IOFFE, S. a SZEGEDY, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *CoRR*. 2015, abs/1502.03167. Dostupné z: <http://arxiv.org/abs/1502.03167>.
- [12] KINGMA, D. P. a BA, J. *Adam: A Method for Stochastic Optimization*. 2014.
- [13] NAGPAL, A. L1 and L2 Regularization Methods. *Towards Data Science*, 13. oktobra 2017 [cit. 2020-05-20]. Dostupné z: <https://towardsdatascience.com/l1-and-l2-regularization-methods-ce25e7fc831c>.
- [14] NAIR, V. a HINTON, G. Rectified Linear Units Improve Restricted Boltzmann Machines Vinod Nair. In: Jún 2010, sv. 27, s. 807–814.
- [15] RIEGLER, G., URSCHLER, M., RÜTHER, M., BISCHOF, H. a STERN, D. Anatomical Landmark Detection in Medical Applications Driven by Synthetic Data. In: *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*. 2015, s. 85–89.
- [16] RONNEBERGER, O., FISCHER, P. a BROX, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015.
- [17] SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. a SALAKHUTDINOV, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. 2014, zv. 15, č. 56, s. 1929–1958. Dostupné z: <http://jmlr.org/papers/v15/srivastava14a.html>.
- [18] TIULPIN, A., MELEKHOV, I. a SAARAKKALA, S. KNEEL: Knee Anatomical Landmark Localization Using Hourglass Networks. *CoRR*. 2019, abs/1907.12237. Dostupné z: <http://arxiv.org/abs/1907.12237>.
- [19] WANG, C.-W., HUANG, C.-T., LEE, J.-H., LI, C.-H., CHANG, S.-W. et al. A benchmark for comparison of dental radiography analysis algorithms. *Medical Image Analysis*. 2016, zv. 31, s. 63 – 76. DOI: <https://doi.org/10.1016/j.media.2016.02.004>. ISSN 1361-8415. Dostupné z: <http://www.sciencedirect.com/science/article/pii/S1361841516000190>.