



**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 MEDI-  
CÍNSKYCH 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 landmarkov v 3D dentálnych skenoch. Tieto skeny slúžia k uľahčeniu modelovania zubných koruniek 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.

Na lokalizáciu landmarkov bola použitá regresia tepelných máp. Z výsledkov regresie boli následne vypočítané odhadované súradnice jednotlivých landmarkov.

## Abstract

This thesis offers possible solution to automatic 3D dental scan landmark localization. These scans are used in dental crown design 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 as heat map regression. Results of the regression were then used to compute the estimated landmark coordinates.

## 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

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana X... Další informace mi poskytli... Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....

Albert Szöllösi

28.5.2020

## Podakovanie

V této sekci je možno uvést poděkování vedoucímu práce a těm, kteří poskytli odbornou pomoc (externí zadavatel, konzultant apod.).

# Obsah

<b>1</b>	<b>Úvod</b>	<b>2</b>
<b>2</b>	<b>Lokalizácia landmarkov</b>	<b>3</b>
2.1	Problémy pri lokalizácii landmarkov . . . . .	3
2.2	Regresia súradníc landmarkov . . . . .	3
2.3	Lokalizácia oblastí záujmu . . . . .	3
2.4	Lokalizácia landmarkov regresiou tepelných máp . . . . .	4
<b>3</b>	<b>Návrh riešenia</b>	<b>5</b>
3.1	Dataset . . . . .	6
3.2	Architektúra siete . . . . .	6
<b>4</b>	<b>Implementácia</b>	<b>8</b>
4.1	Trénovanie . . . . .	8
4.2	Trénovacie dáta . . . . .	8
4.3	Augmentácia dát . . . . .	8
4.4	Stratová funkcia . . . . .	9
4.5	Trénovacie parametre . . . . .	9
<b>5</b>	<b>Experimenty a výsledky</b>	<b>10</b>
5.1	Trénované modely . . . . .	10
5.2	Testovací dataset . . . . .	14
5.3	Vyhodnotenie modelov . . . . .	14
5.4	Zhodnotenie riešenia . . . . .	16
<b>6</b>	<b>Záver</b>	<b>20</b>
	<b>Literatúra</b>	<b>21</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 predpovedanie vývoja choroby, eliminovanie chýb spôsobených nedokonalosťou ľudských zmyslov a šetrenie času a tým pádom aj peňanží. Táto práca je zameraná na posledné dva z uvedených dôvodov.

Cielom mojej práce je lokalizácia landmarkov v 3D skenoch chrupu pomocou konvolučnej neurónovej siete. Spomínané skeny sú používané pri jednej z metód plánovania ortodontickej liečby. Ortodontickou liečbou sa rozumie úprava chrupu pomocou ortodonticého 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 potom softvér vie určiť natočenie modelu a pozície jednotlivých zubov. Tieto informácie sú použité pri následnej segmentácii modelu. Segmentáciou sa 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 ju však možné použiť aj na lokalizáciu landmarkov, a to regresiou tepelných máp. Sieť bola natrénovaná na 2D obrázkoch, ktoré boli vytvorené ako hĺbkové mapy 3D skenov chrupu. Tieto obrázky boli následne ručne anotované súradnicami anatomických bodov. Na základe jednotlivých súradníc boli vytvorené obrázky s rovnakým rozlíšením ako hĺbkové mapy vytvorené zo skenov chrupu, pričom tieto obrázky obsahujú 2D Gaussovu krivku so stredom v mieste, kde sa konkrétny landmark nachádza. Toto sú práve vyššie spomínané tepelné mapy, ktoré sa natrénovaná sieť snaží regresiou vytvoriť. Z výstupov siete boli následne vypočítané súradnice jednotlivých landmarkov. Na výpočet boli použité rôzne metódy, aby bolo možné vybrať tú s najvyššou presnosťou.

V kapitole 2 sú bližšie opísané najčastejšie používané metódy lokalizácie landmarkov.

**Dopísať obsah**

## Kapitola 2

# Lokalizácia landmarkov

V dnešnej dobe sa lokalizácia landmarkov, ako väčšina problémov spojených s analýzou obrazových dát, rieši pomocou konvolučných neurónových sietí. Existuje už mnoho overených, ale aj nových spôsobov ako pristupovať k problému lokalizácie landmarkov. Väčšina z nich formuluje problém ako regresnú úlohu, sú však aj prístupy využívajúce klasifikáciu [1]. Záleží na tom aké informácie je potrebné o landmarkoch dostať. V tejto kapitole sú vysvetlené niektoré z možných prístupov k lokalizácii landmarkov.

### 2.1 Problémy pri lokalizácii landmarkov

Pri lokalizácii landmarkov môže nastať množstvo problémov. Najčastejšie sa lokalizácia landmarkov rieši na tvárach 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 landmarkov a mnohokrát spôsobuje to, že model, ktorý je dobre natrénovaný na jednom datasete bude na inom datasete dosahovať oveľa horšie výsledky. Výkon modelov na lokalizáciu landmarkov môže negatívne ovplyvniť aj výber požadovaných landmarkov, keďže niektoré landmarky môžu byť ťažko rozpoznateľné, a tiež požadovaná presnosť modelu.

### 2.2 Regresia súradníc landmarkov

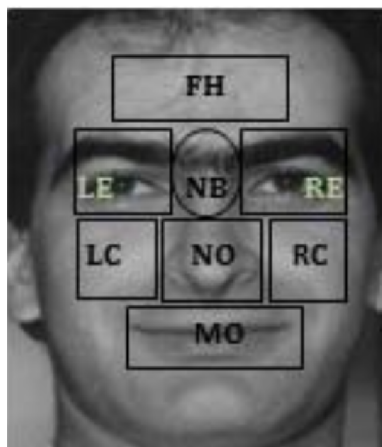
Priama regresia súradníc je v praxi asi najbežnejšou metódou lokalizácie landmarkov. Princíp je jednoduchý. Na vstupe je snímka, v ktorej chceme landmarky lokalizovať. Požadovaným výstupom je matica obsahujúca súradnice jednotlivých landmarkov:

$$y = \begin{bmatrix} x_1 & \cdots & x_n \\ y_1 & \cdots & y_n \end{bmatrix}$$

### 2.3 Lokalizácia oblastí záujmu

Lokalizácia oblastí záujmu (angl. regions of interest, skrátene ROI) spočíva v tom, že pozície landmarkov sa odhadujú ako súradnice štvoruholníkov, ktoré ohraničujú jednotlivé landmarky (angl. bounding boxes). Vstupom je teda obrázok, na ktorom má byť lokalizácia vykonaná a cieľovým výstupom sú súradnice spomínaných bounding boxov. Takéto riešenie

však nie je vhodné pre účely tejto práce, pretože jej cieľom je lokalizovať konkrétne body na jednotlivých zuboch, nie väčšie oblasti.



Obr. 2.1: Príklad ROI na tvári [1].

Aj napriek tomu je táto metóda používaná aj pri lokalizácii landmarkov definovaných len ako body. V takýchto prípadoch je model natrénovaný na lokalizáciu ROI použitý na extrakciu ROI ako určitú formu predspracovania dát [2]. Regresia pozícií landmarkov je potom aplikovaná len na jednotlivé ROI, čo má za následok zvýšenie rýchlosti výpočtov, keďže je dimenzionalita obrázkov znížená na veľkosť jednotlivých ROI. Keďže sa týmto zároveň zo vstupnej matice odstránia nepotrebné príznaky, ktoré by mohli byť pri predikcii rušivé, zvyšuje sa tým aj presnosť predikcií.

## 2.4 Lokalizácia landmarkov regresiou tepelných máp

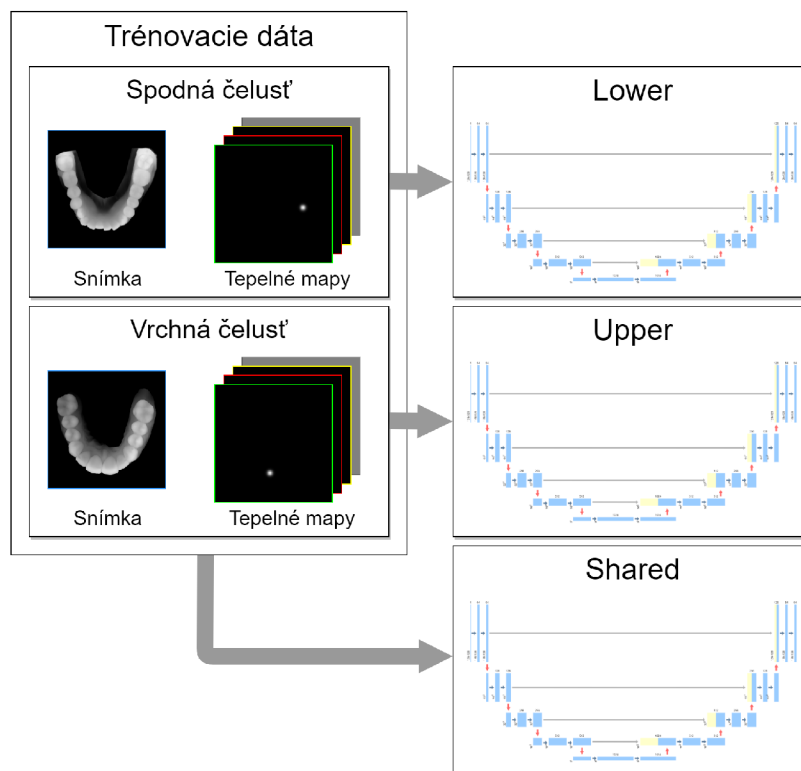
Táto metóda sa snaží regresiou predpovedať tepelné mapy, a to priamo zo vstupných snímok. Tepelné mapy reprezentujú pozície landmarkov. Obsahujú len jednu Gaussovú krivku s malou varianciou a so stredom v mieste landmarku. Odhadované súradnice sa počítajú ako súradnice, na ktorých sa nachádza maximálna hodnota v predikovanej tepelnej mape.



## Kapitola 3

# Návrh riešenia

Na lokalizáciu landmarkov bola použitá metóda regresie tepelných máp. Natrénované boli tri modely nasledujúce architektúru U-Net [6]. Prvý model bol natrénovaný na všetkých trénovacích dátach, druhý bol natrénovaný na obrázkoch spodných čelustí a tretí na obrázkoch vrchných čelustí. Následne boli všetky modely porovnané, na základe čoho bolo vyhodnotené, či by v praxi bolo vhodné skeny chrupu pred lokalizáciou landmarkov klasifikovať na spodnú a vrchnú čelusť.



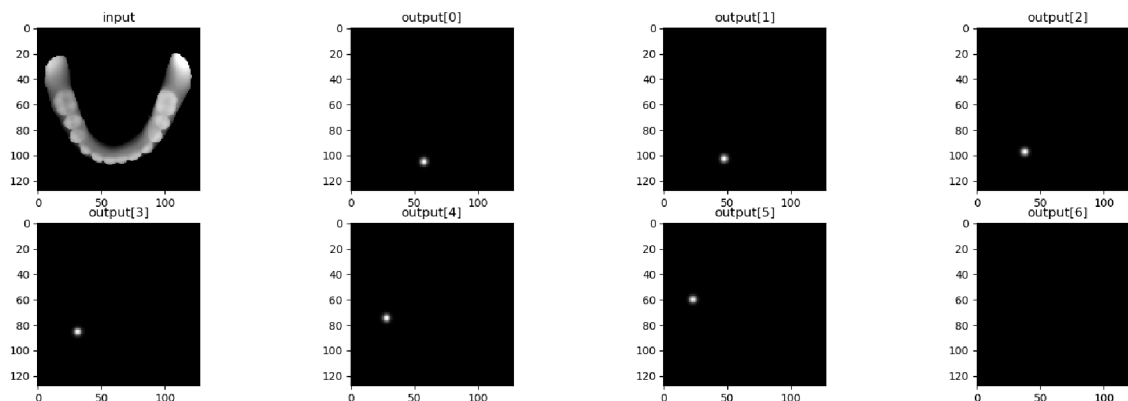
Obr. 3.1: Rozdelenie trénovacích dát.

Z predpovedí vytvorených natrénovanými sieťami boli následne odhadnuté súradnice landmarkov. Súradnice vypočítané z tepelných máp 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.

### 3.1 Dataset

Dataset použitý na riešenie úlohy bol vytvorený z 3D dentálnych skenov poskytnutých mojím pánom vedúcim. Z jednotlivých skenov boli vytvorené hĺbkové mapy. Aby bol proces tréningu jednoduchší, boli snímky vytvárané tak, aby boli všetky modely približne rovnako natočené. Dataset bol následne rozdelený podľa toho, či sa jednalo o dolnú alebo hornú čelusť. Jednotlivé snímky boli následne ručne anotované súradnicami udávajúcimi polohy jednotlivých zubov v snímke. V prípade, že niektorý zub chýbal, bola anotácia daného landmarku vynechaná. Jedna vzorka datasetu teda obsahuje hĺbkovú mapu 3D dentálneho skenu a 16 pozícií landmarkov.

Ako bolo už spomenuté vyššie, lokalizácia landmarkov bola riešená metódou regresie tepelných máp. To znamená, že neurónová sieť dostane na vstupe snímku chrupu v stupňoch sivej o veľkosti  $d \times d$ . Zodpovedajúci cieľový výstup je súbor 16 tepelných máp o veľkosti  $d \times d$ . Každá tepelná mapa 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  $16 \times d \times d$  obsahujúca odhadované tepelné mapy pre jednotlivé landmarky.

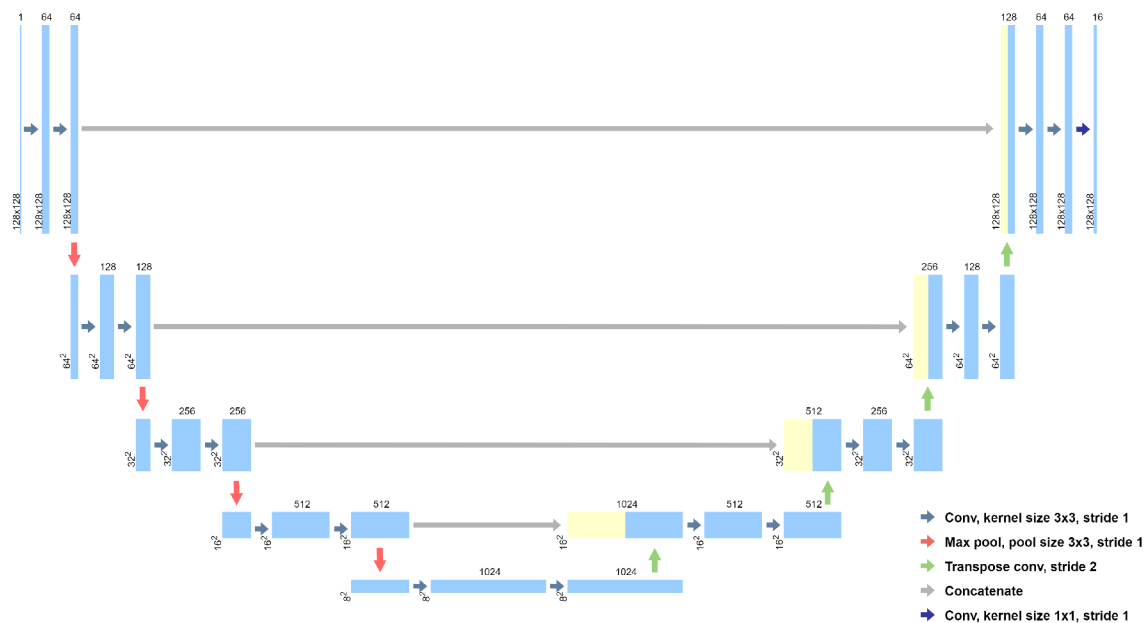


Obr. 3.2: Ukážka tréningového vstupu a jeho cieľových výstupov.

### 3.2 Architektúra siete

Návrh siete je odvodený architektúry U-Net. Táto architektúra bola pôvodne navrhnutá na segmentáciu obrazových dát [6], ale po zmene počtu výstupných kanálov a za použitia vhodnej stratovej funkcie dosahuje dobré výsledky aj pri lokalizácii landmarkov. Skladá sa z podvzorkovacej časti, ktorá je nasledovaná symetrickou nadvzorkovacou časťou.

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 konvolúčných 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 príznakovej mapy na polovicu.



Obr. 3.3: Architektúra U-Net.

Nadvzorkovací blok aplikuje trasponovanú 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, ktorý 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 konvergence modelu[3].

## Kapitola 4

# Implementácia

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

### 4.1 Trénovanie

Vstupom siete je snímka v stupňoch sivej. Všetky modely boli trénované na obrázkoch o rozmeroch  $128 \times 128$ , čo urýchlilo trénovanie, ale viedlo to k strate presnosti v porovnaní s inými modelmi pracujúcimi s obrázkami väčších rozmerov.

Modely, ako aj ich trénovanie boli implementované v Keras API s Tensorflow backendom. Trénovanie prebiehalo na grafickej karte NVIDIA GeForce GTX960M.

### 4.2 Trénovacie dáta

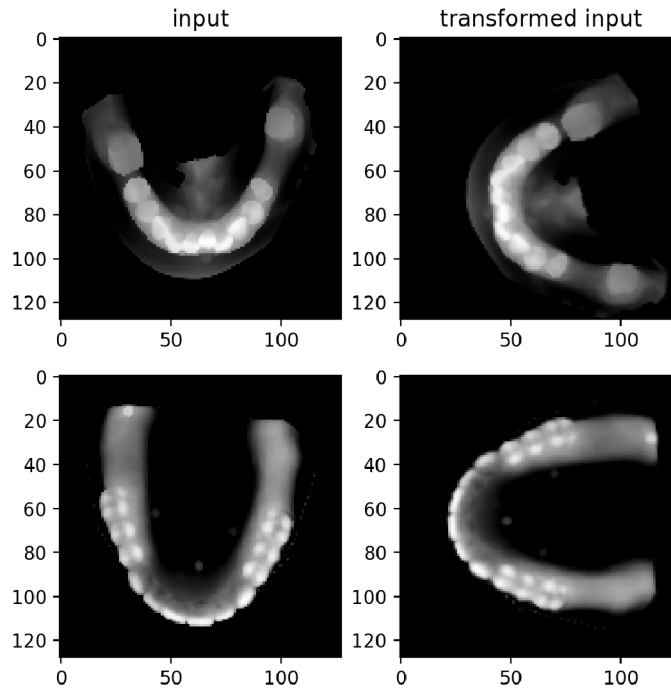
Vytvorený dataset obsahuje 282 snímok, pričom 118 z nich obsahuje skeny spodnej čeluste a ostatných 164 obsahuje skeny vrchnej čeluste. Z tohoto datasetu boli vytvorené dva ďalšie datasety. Jeden obsahuje snímky len vrchnej čeluste a druhý obsahuje snímky len spodnej čeluste. Všetky tri datasety boli rozdelené na trénovacie a validačné dáta v pomere  $r \approx 3 : 1$ .

### 4.3 Augmentácia dát

Augmentácia dát bola použitá za účelom vyššej variability a zväčšenia datasetu, ktorý je relatívne malý. Z každej snímky bola vytvorená jedna ďalšia, na ktorú boli aplikované nasledujúce transformácie:

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

Augmentácia bola aplikovaná na snímky aj na cieľové tepelné mapy.



Obr. 4.1: Príklad augmentácie vstupných obrázkov.

#### 4.4 Stratová funkcia

Keďže ide o regresnú úlohu, ako stratová funkcia bola použitá stredná kvadratická chyba (MSE):

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

kde  $y_i$  sú cieľové tepelné mapy,  $\hat{y}_i$  sú predikcie modelu a  $n$  je veľkosť batchu. Keďže tepelné mapy 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 zabránilo, amplitúda Gaussovej krivky bola nastavená na 1000 ako je navrhované v [?].

#### 4.5 Tréningové parametre

Pri tréningu bol použitý optimalizátor Adam s koeficientom učenia  $10^{-3}$  a tréning prebiehalo po batchoch o veľkosti 8.

# Kapitola 5

## Experimenty a výsledky

Táto kapitola obsahuje detaily o natrénovaných modeloch a vyhodnotenie a porovnanie ich presností. Na záver tiež faktormi, ktoré mohli negatívne ovplyvniť presnosť modelov a predkladám návrhy na ich riešenie.

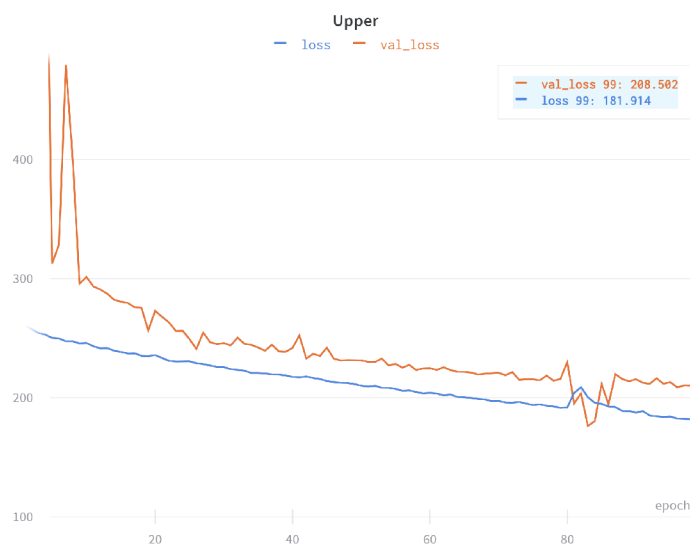
### 5.1 Trénované modely

V tejto kapitole sú vyhodnotené tri modely so spoločnou architektúrou a príbehom tréovania, ktorý je opísaný v sekcii 4.1. Táto sekcia sa zaoberá krížovou validáciou natrénovaných modelov a ich modifikovaných variantov. Na základe krížovej validácie boli vybraté najlepšie modely, ktoré boli neskôr použité pre ďalšie experimenty opísané v nasledujúcich sekciiach.

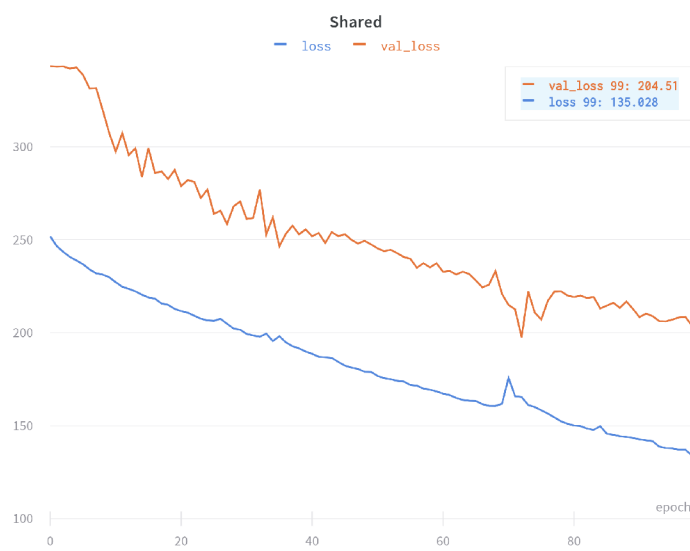
Najprv boli natrénované modely presne nasledujúce architektúru opísanú v sekcii 3.2. Ako je uvedené v kapitole 3, jeden model bol natrénovaný na snímkach spodných čelustí (ďalej len Lower), druhý model bol natrénovaný na snímkach vrchných čelustí (ďalej len Upper) a tretí na snímkach oboch čelustí (ďalej len Shared).



Obr. 5.1: Stratová funkcia modelu Lower počas tréovania



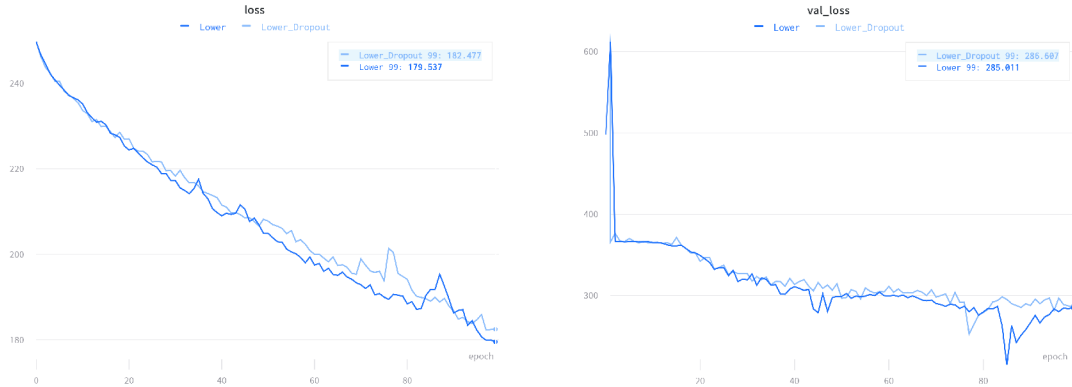
Obr. 5.2: Stratová funkcia modelu Upper počas tréningovania



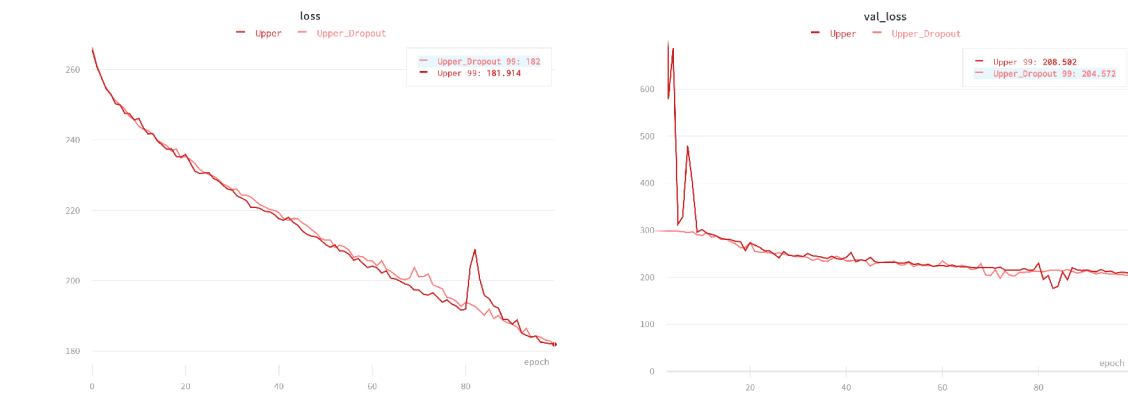
Obr. 5.3: Stratová funkcia modelu Shared počas tréningovania

Pri analýze vývoja stratovej funkcie počas tréningovania bolo zistené, že pri všetkých troch modeloch došlo k overfittingu. To znamená, že sa modely príliš konkretizovali na tréningové dáta a následkom toho dosahuje stratová funkcia slúžiaca na validáciu modelov výrazne vyššie hodnoty ako tá, ktorou sa modely optimalizujú. Najväčšia miera overfittingu bola zaznamenaná na modeli Shared, aj napriek tomu, že tento model bol tréningovaný na najväčšom počte dát. Za ním nasleduje model Lower a najmenšia, takmer zanedbateľná miera overfittingu bola zaznamenaná pri modeli Upper. 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 [7]. 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 hodnôt stratových funkcií počas tréningu modelu Lower s a bez dropout vrstiev.



Obr. 5.5: Porovnanie hodnôt stratových funkcií počas tréningu modelu Upper s a bez dropout vrstiev.

Po natréningu takto modifikovaných modelov na pôvodných dátach bolo analýzou zistené, že Dropout vrstvy nemali na výsledky modelov Lower a Upper žiaden pozitívny vplyv. V prípade modelu Shared sa hodnoty tréningovej stratovej funkcie síce o iečo zvýšili, ale zároveň bolo dosiahnuté zníženie validačnej stratovej funkcie.

Druhou metódou bola L2 regularizácia váh [4] inak nazývaná aj hrebeňová regresia. Táto metóda upravuje hodnotu stratovej funkcie pri jej výpočte. Úprava spočíva v tom, že sa k stratovej funkcii pri jej výpočte pripočíta súčet druhých mocnín súčasných váh vynásobený koeficientom  $\lambda$  k stratovej funkcii:

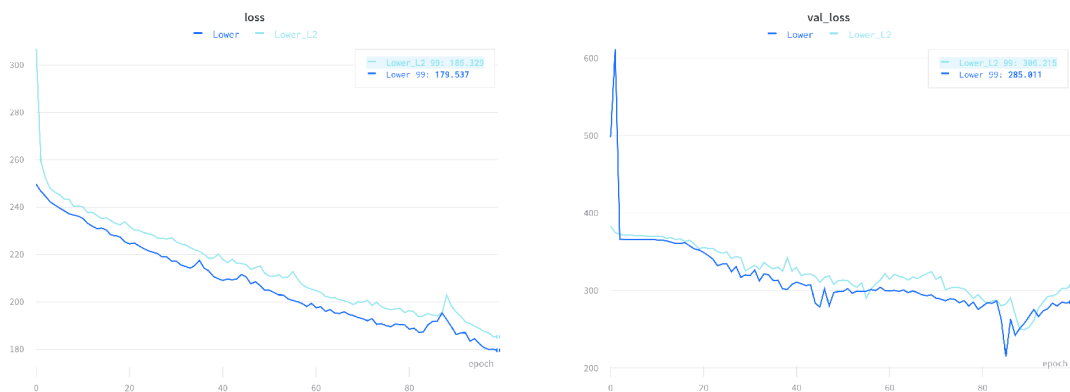
$$Loss = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \frac{\lambda}{n} \sum_{i=1}^n w_i^2$$

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





Obr. 5.6: Porovnanie hodnôt stratových funkcií počas tréningovania modelu Shared s a bez dropout vrstiev.



Obr. 5.7: Porovnanie hodnôt stratových funkcií počas tréningovania modelu Lower s a bez L2 regularizácie.



Obr. 5.8: Porovnanie hodnôt stratových funkcií počas tréningovania modelu Upper s a bez L2 regularizácie.

Po ukončení tréningovania bolo zistené, že hrebeňová regresia mala na presnosť modelov Lower a Upper len negatívny dopad. Obe stratové funkcie oboch modelov dosahovali vyššie



Obr. 5.9: Porovnanie hodnôt stratových funkcií počas tréovania modelu Shared s a bez L2 regularizácie.

hodnoty ako modely bez regularizácie váh. Na model Shared mala L2 regularizácia podobný vplyv ako prídanie dropout vrstiev.

Na základe krížovej validácie boli pre použitie na nasledujúce experimenty vybrané najlepšie modely. Keďže prídanie dropout vrstiev do modelov Lower a Upper nijak nezlepšilo výkon týchto modelov a prídanie L2 regularizácia ich výkon zhoršila, boli pre nasledujúce experimenty vybrané varianty týchto modelov nasledujúce pôvodný návrh siete. V prípade modelu Upper došlo pri aplikácii oboch metód k underfittingu. Rozdiel medzi tréovacou funkciou modelu Upper bez modifikácií a jeho variantu s L2 regularizáciou bol však relatívne malý a validačná stratová funkcia modelu s L2 regularizáciou dosahovala nižšie hodnoty ako v prípade modelu bez modifikácií. Na základe toho bola pre nasledujúce experimenty vybraná varianta modelu Shared s L2 regularizáciou.

## 5.2 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.

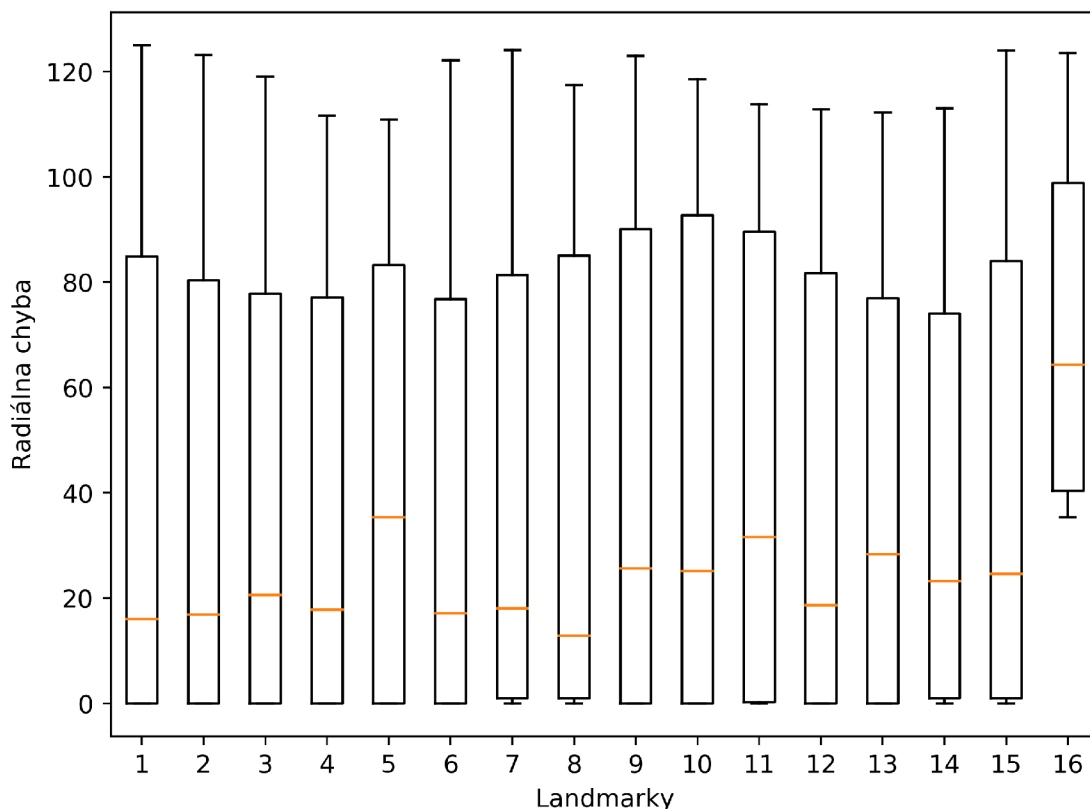
## 5.3 Vyhodnotenie modelov

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.

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



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

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

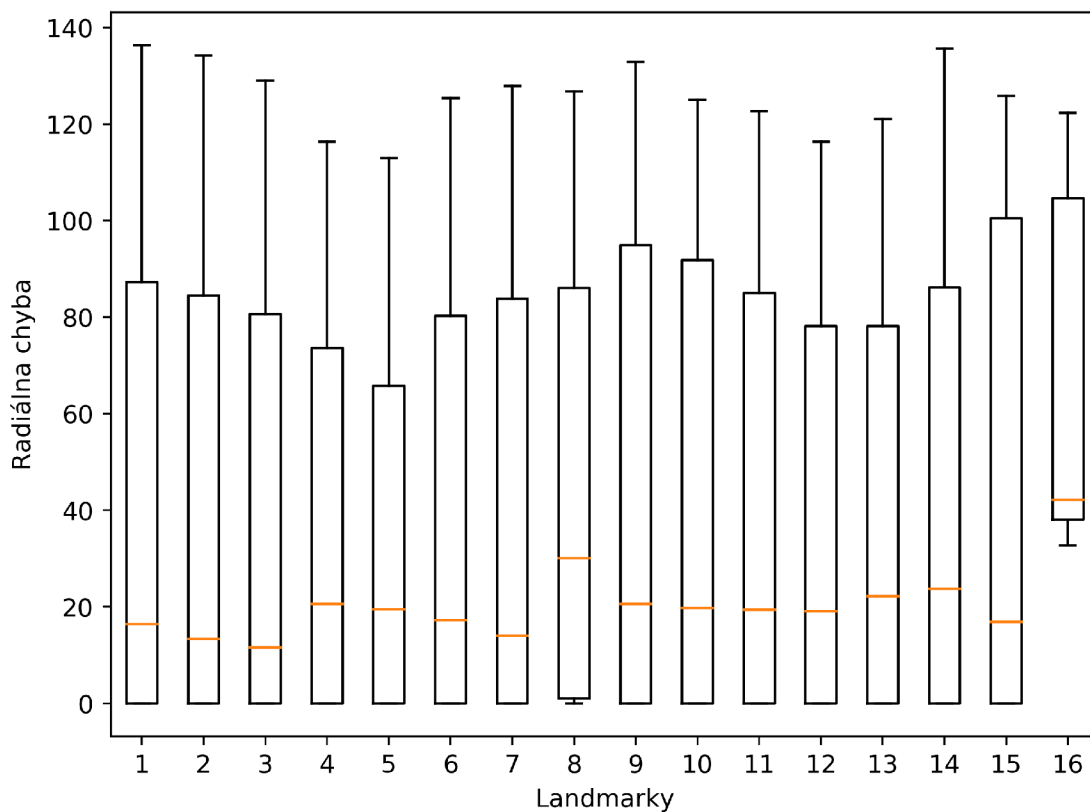
kde  $n$  je dimenzionalita výstupu siete, teda počet landmarkov v jednej snímke. Táto funkcia bola vypočítaná pre každú predikciu zvlášť. Na základe výsledkov bola získaná celková presnosť predpovedí jednotlivých modelov.

Pre obidve metriky boli odhadované pozície landmarkov vypočítané ako súradnice bodu na tepelnej mape s najvyššou amplitúdou, ako bolo navrhnuté v [5]:

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

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

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



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

Model	Presnosť
Lower	0.4618
Upper	0.4768
Shared	0.4694

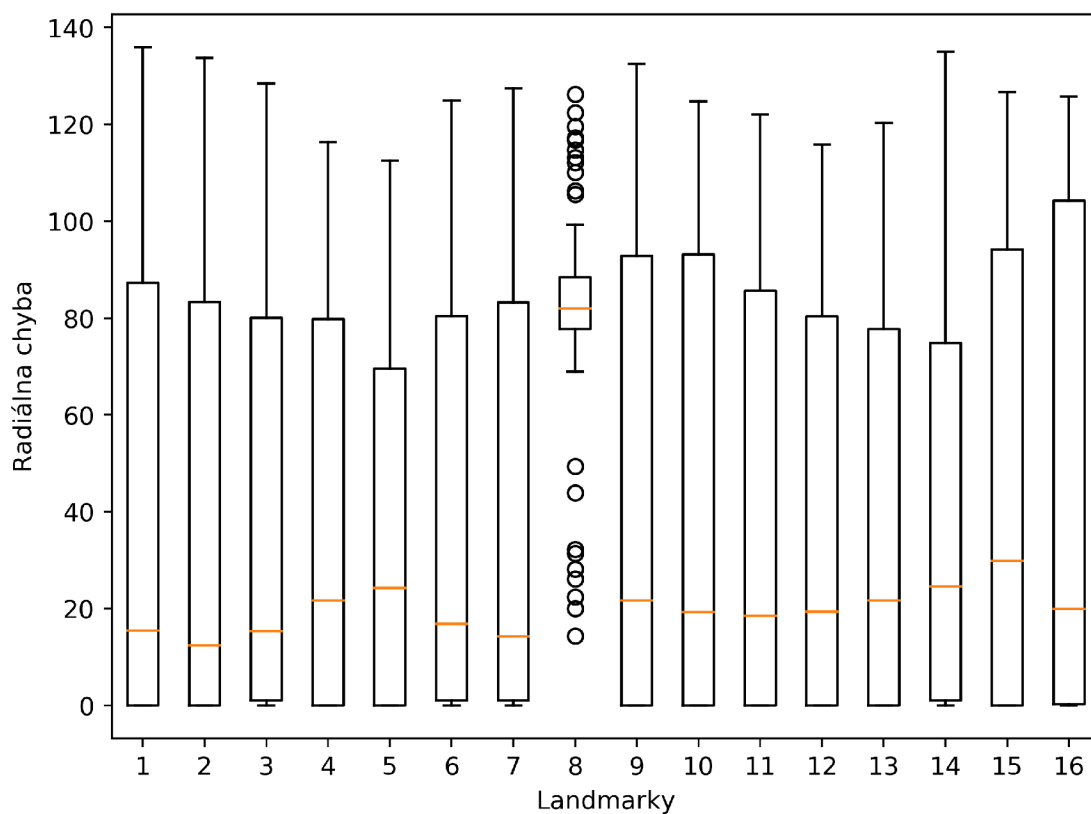
Tabuľka 5.1: Presnosť modelov

## 5.4 Zhodnotenie riešenia

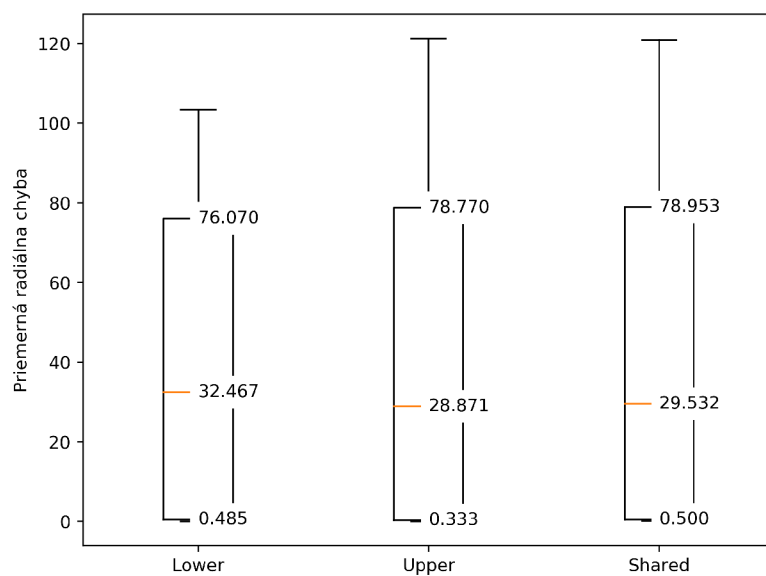
Riešenie navrhnuté v tejto práci rozhodne nie je použiteľné v praxi, keďže všetky natrénované modely dosiahli úspešnosť menšiu ako *frac12*. Navyše, v tejto práci nebolo riešené spätné mapovanie pozícií landmarkov z 2D obrázkov do 3D modelov. Na presnosť modelov mohli mať vplyv viaceré faktory. Ako vidno 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
Spodná čelusť	10	9	3	19	32	80	70	149	7	9	2	20	31	83	76	165
Vrchná čelusť	22	21	10	33	45	63	102	214	25	19	17	40	40	68	118	230

Tabuľka 5.2: Počty snímok, na ktorých chýbajú jednotlivé zuby

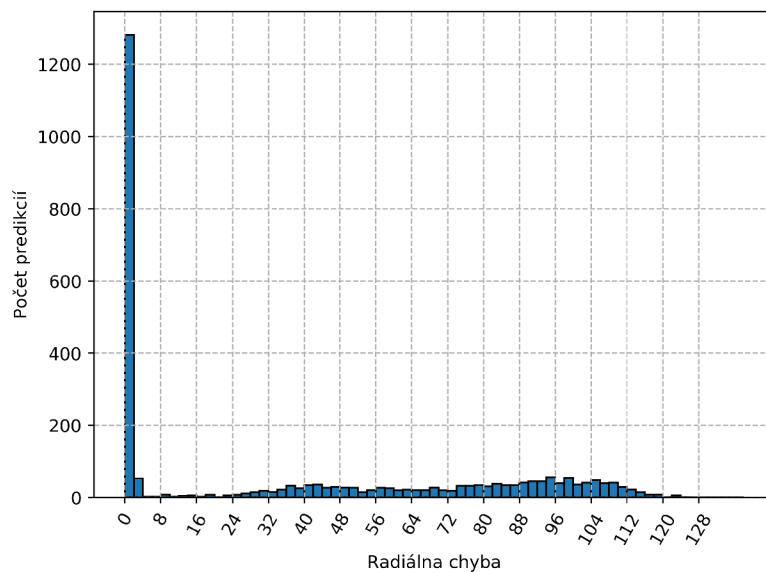


Obr. 5.12: Radiálna chyba modelu Shared na jednotlivých landmarkoch.

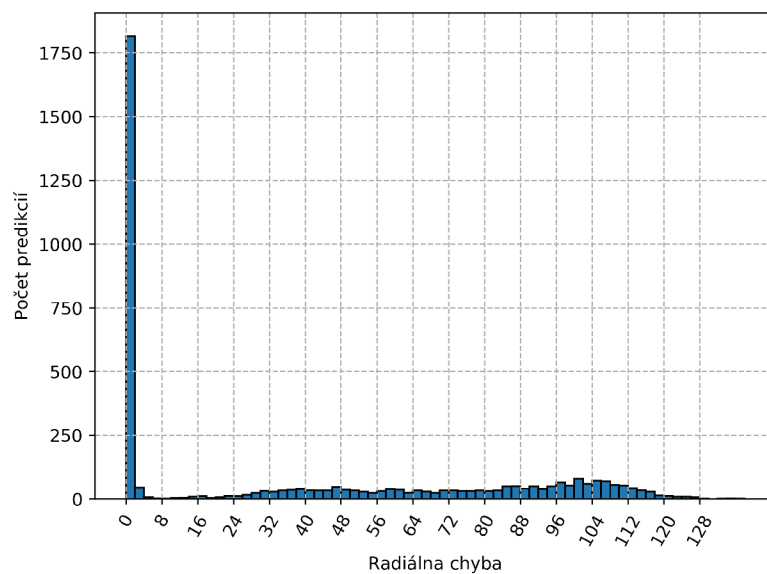


Obr. 5.13: Priemerná radiálna chyba jednotlivých modelov.

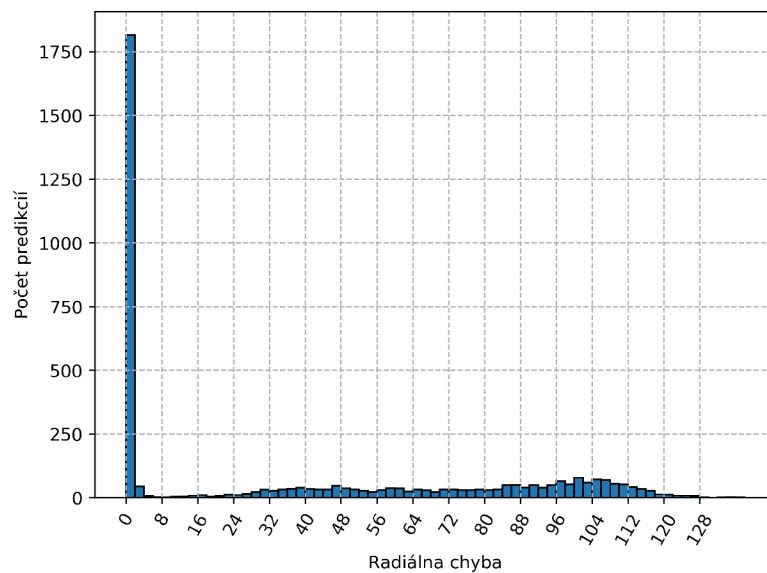
Keď tieto počty porovnáme s radiálnou chybou na daných zuboch, zistíme, že tieto hodnoty sú v niektorých prípadoch korelované, čo najviac vidieť pri landmarku č. 16. Za toto však môže aj malá veľkosť tréningového datasetu. V prípade väčšieho datasetu by bolo možné dosiahnuť väčšiu variabilitu dát a viac obrázkov, na ktorých by sa jednotlivé zuby nachádzali, čo by pomohlo pri predikcii aj často chýbajúcich zubov. Zvláštnosťou predikcií všetkých modelov je rozloženie veľkostí ich chyby. Aj napriek tomu, že, ako je možné vidieť v boxplote 5.13, sa mediány radiálnej chyby jednotlivých modelov pohybujú až okolo hodnoty 30, 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 tepelných máp jednotlivých landmarkov. Tepelné mapy obsahujú 2D Gaussovu krivku so stredom v mieste landmarku, ktorému prislúchajú. 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. Aplikovaná bola L2 regularizácia váh, ktorá mala pozitívny vplyv iba na model Shared, ktorý bol trénovaný na všetkých trénovacích dátach. Druhou aplikovanou optimalizáciou bolo pripojenie dropout vrstiev na výstupy konvolučných vrstiev modelu. Tá nemala pozitívny efekt na žiaden z trénovaných modelov.

Ako najlepšie modely boli teda vybraté model Lower bez vyššie spomínaných optimalizácií, ktorý bol trénovaný na snímkach spodnej čeluste, model Upper tiež bez modifikácií, ktorý bol trénovaný na snímkach vrchných čelustí a model Shared s L2 regularizáciou váh. 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, stačilo by túto prácu rozšíriť o spoľahlivý spôsob mapovania súradníc landmarkov z 2D snímok naspäť do 3D modelov.



# Literatúra

- [1] ANGULU, R., TAPAMO, J.-R. a ADEWUMI, A. Landmark Localization Approach for Facial Computing. In: Marec 2017. DOI: 10.1109/ICTAS.2017.7920664.
- [2] BEUMER, G. M., TAO, Q., BAZEN, A. M. a VELDHUIS, R. N. J. A landmark paper in face recognition. In: *7th International Conference on Automatic Face and Gesture Recognition (FGRO6)*. 2006, s. 6 pp.–78.
- [3] 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>.
- [4] NAGPAL, A. L1 and L2 Regularization Methods. *Towards Data Science*, 13. októbra 2017 [cit. 2020-05-20]. Dostupné z: <https://towardsdatascience.com/l1-and-l2-regularization-methods-ce25e7fc831c>.
- [5] OURSELIN, S., JOSKOWICZ, L., SABUNCU, M. R., ŮNAL, G. B. a WELLS, W., ed. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016 - 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II*. 2016, sv. 9901. Lecture Notes in Computer Science. DOI: 10.1007/978-3-319-46723-8. ISBN 978-3-319-46722-1. Dostupné z: <https://doi.org/10.1007/978-3-319-46723-8>.
- [6] RONNEBERGER, O., FISCHER, P. a BROX, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015.
- [7] 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>.