

Neurónová sieť typu U (U-Net)

Lukáš Novoročník

Abstract — V súčasnosti je pomerne veľký dopyt po neurónových sieťach, ktoré by boli schopné spoľahlivo fungovať, resp. vygenerovať uspokojivý výsledok aj pri menšom počte dát, na ktorých by sa neurónová sieť mohla natrénovať. Jedným z možných využití takejto neurónovej siete by mohla byť obrazová segmentácia a práve na tento účel sa zameriava neurónová sieť typu U, ktorá je popísaná v tomto dokumente.

Kľúčové slová— Neurónová sieť, Konvolučná neurónová sieť, Neurónová sieť typu U

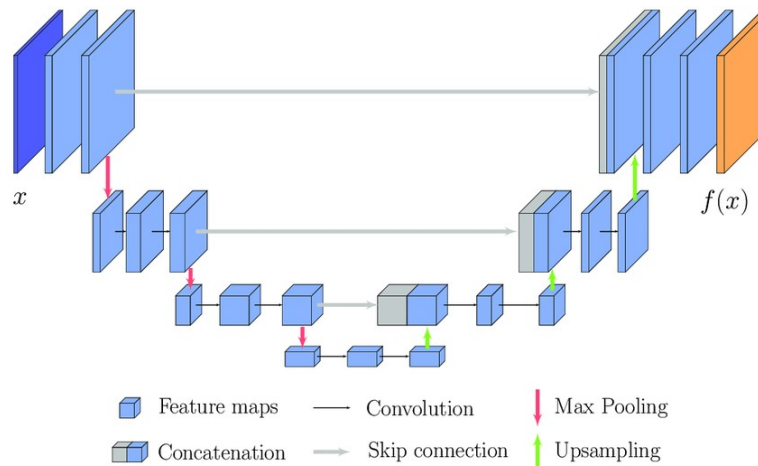
I. UVOD

V posledných rokoch sa výkonnosť neurónových sietí a špeciálne hlbokých neurónových sietí výrazne zvýšila. Neurónové siete sa čoraz viac zväčšujú a prehľbujú, a sú schopné zohľadniť väčšie množstvo parametrov. Veľký prielom uskutočnili Krizhevsky et al., ktorý využili učenie s učiteľom na veľkej neurónovej sieti s 8 vrstvami s až 1 miliónom parametrov, ktorú aplikovali na ImageNet datasete, ktorý pozostáva z 1 milióna tréningových snímok.

V prípade využitia neurónových sietí je typické použitie týchto sietí na klasifikačné účely, kde sa na výstupe očakáva jedna rozpoznaná trieda. V mnohých prípadoch sa však môžeme stretnúť s potrebou rozpoznania viacerých tried naraz. Jedným z takýchto prípadov je využitie neurónových sietí pre biomedicínske účely, kde je pre nedostatok dostupných tréningových dát potrebné priradovať triedu každému pixelu zvlášť. Jedným z takýchto prístupov je predpovedanie triedy každého pixelu tak, že ako vstup pre neurónovú sieť sa použije konkrétna oblasť okolo daného pixelu. Takýto prístup je však nepraktický napríklad kvôli časovej náročnosti, pretože neurónová sieť musí bežať pre každú oblasť, pričom pri takomto prístupe vzniká množstvo redundancií pri prekrývaní jednotlivých oblastí. Okrem toho je pri takomto prístupe problém aj s nájdením správneho pomeru výkonu tréningovania a veľkosťou oblasti okolo pixelu, nakoľko pri príliš malých oblastiach je možná strata kontextu. V tomto dokumente je popísaná neurónová sieť typu U, ktorá sa niektoré z týchto problémov pokúša vyriešiť.

II. ARCHITEKTÚRA SIETE

Na obrázku 1 je znázornená architektúra siete typu U. Táto sieť sa vo všeobecnosti skladá z dvoch častí. Jednu časť predstavuje „zmenšovacia“ (contracting) cesta siete, ktorá je na obrázku 1 znázornená na ľavej strane, druhá časť sa označuje ako „zväčšovacia“ (expanding) cesta, tá je na obrázku znázornená na pravej strane. Zmenšovacia cesta má charakteristiku klasickej konvolučnej siete. Tá pozostáva z opakovanej aplikácie dvoch 3x3 konvolúcií, za ktorými nasleduje aplikovanie ReLU jednotky a následne je aplikovaná 2x2 max-pooling operácia s krokováním (parameter stride) 2 pre downsamplingovanie. Pri každom kroku downsamplingovania procesu sa znásobí počet kanálov. Pri zväčšovacej časti sa v každom kroku vykonáva upsamplingovanie mapy príznakov, za ktorým nasleduje 2x2 konvolúcia, ktorá zníži počet kanálov na polovicu. V tomto kroku sa rovnako deje spojenie s upravenou mapou príznakov zo zmenšovacej časti siete a následne sa vykonajú dve 3x3 konvolúcie, pričom za každou konvolúciou nasleduje aplikácia ReLU jednotky. Pri procese spracovania obrazu sieťou dochádza k miernemu orezaniu obrazu. Toto orezanie je výsledkom spracovania obrazu konvolučnými vrstvami. V poslednej vrstve je použitá 1x1 konvolúcia, ktorá je potrebná na mapovanie všetkých 64-komponentových príznakových vektorov na požadovaný počet tried na výstupe.



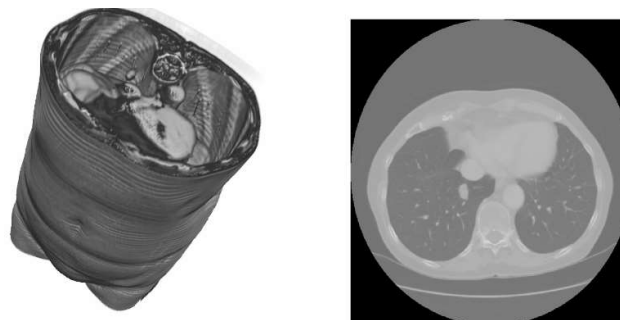
Obrázok 1: Znáozornenie architektúry siete typu U

III. POUŽITÉ METÓDY

A. Dataset a jeho príprava

Pre účely testovania neurónovej siete typu U bol použitý CT-ORG dataset. Tento dataset sa skladá z 140. spracovaných tomografických skenov, pričom na každom z poskytnutých skenov je označených 5 orgánov, ktoré zahŕňajú pľúca, kosti, pečeň, obličky a močový mechúr. Všetky poskytnuté skeny sú v 3D a v 3D sú aj označené. Na minimálnom počte skenov sa nachádza aj mozog, pričom v prípade, že sa na skene nachádza aj mozog, je poskytnuté aj jeho označenie. Obrázky nachádzajúce sa v datasete, pochádzajú z viacerých zdrojov, ktoré zahŕňajú celotelové skeny, skeny z oblasti brucha, skeny s vyšším kontrastom aj skeny s nižším kontrastom a vysokodávkové aj nízкодávkové CT skeny. Z poskytnutých 140 skenov je 131 skenov výhradne CT skenov, zvyšných 9 skenov predstavujú komponenty prebrané zo skúšky PET-CT. V datasete sa nachádzajú veľké a ľahko rozpoznateľné orgány, ako napríklad pľúca rovnako ako menšie orgány, ktoré sú zložitejšie na rozpoznanie ako napríklad žalúdok.

Kvôli komplikovanosti používania 3D obrázkov pre obrazovú segmentáciu a výpočtovej náročnosti, ktorá je pri takomto type datasetu potrebná bol spomínaný dataset upravený. Z niekoľkých náhodne vybraných 3D zväzkov z datasetu boli vytvorené nové snímky. Každý 3D zväzok sa skladá z mnohých menších obrázkov (vrstiev) a preto je možné tieto jednotlivé obrázky rozdeliť a uložiť ako samostatné dáta. Na tento účel bola použitá slučka, ktorá iterovala cez všetky vrstvi určeného 3D zväzku a následne bola každá vrstva uložená ako nový obrázok vo formáte .tif. Rovnaký postup musel byť samozrejme aplikovaný aj na masky, resp. označené dáta, aby sa neskôr dali správne použiť a korešpondovali s vybranými tréningovými obrázkami.



Obrázok 2: Originál 3D zväzok (naľavo) a konkrétna vrstva (napravo)

B. Model U-net siete

Ako model pre experimenty bol použitý vlastný model siete vytvorený podľa vyššie uvedenej architektúry. Model bol vytvorený s pomocou knižnice *keras*. Popis aj postupnosť jednotlivých vrstiev je opäť popísaná vyššie. Ako aktivačná funkcia pre výstup bola použitá funkcia *softmax*. Softmax je normalizovaná exponenciálna funkcia. V neurónovej sieti je využitá na normalizáciu výstupu na rozdelenie pravdepodobnosti predpokladaných tried.

C. Modely s využitím knižnice *segmentation_models*

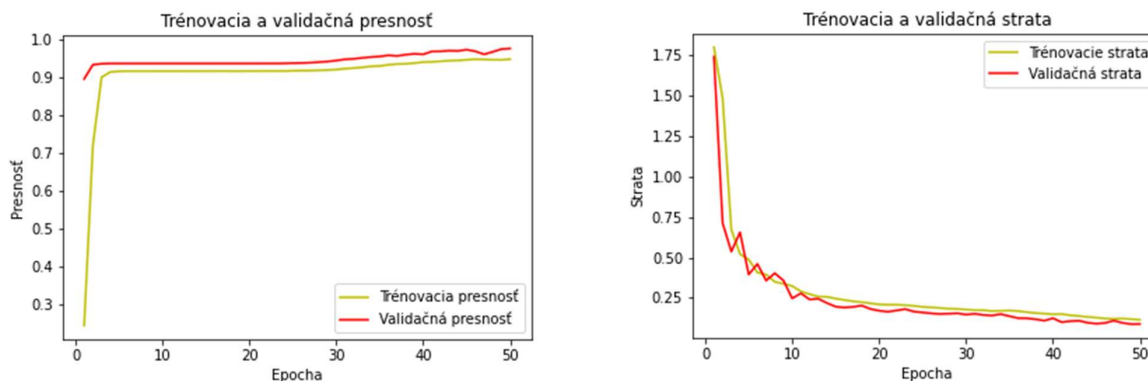
Okrem vlastnej U-net siete je možné využiť aj predpripravenú sieť z knižnice *segmentation_models*. Táto knižnica obsahuje niekoľko rôznych neurónových sietí vrátane siete typu U. V experimentoch bola použitá aj táto implementácia siete typu U, pričom boli využité rôzne typy backboneu siete a to konkrétne *resnet34*, *vgg16* a *inceptionv3*. Jednotlivé backboney siete sa odlišujú v rôznom usporiadaní vrstiev siete a v riadení prúdenia dát v sieti.

IV. DIZAJN EXPERIMENTOV

Experimenty boli vykonávané so všetkými vyššie popísanými modelmi, resp. metódami. Pre každý model boli použité dáta z rovnakého datasetu. Dáta boli pri všetkých experimentoch rozdelené rovnako a to v pomere 80:20, kde 20% predstavovali dáta určené na testovanie. Následne boli vytvorené inštancie jednotlivých modelov. Každému modelu boli predané rovnaké dáta. V prípade modelov z knižnice *segmentation_models* boli dáta pred použitím v modeli spracované pre daný typ modelu, resp. backboneu. Po spracovaní dát boli dáta predané jednotlivým modelom. Pri modeloch z knižnice *segmentation_models* bola pre validáciu použitá matrika *Intersection over Union* (ďalej IoU), čo je metrika, používajúca sa v obrazovej segmentácii a popisuje pomer medzi oblasťou pretnutia zaznačenej oblasti objektu v maske a skutočným predpokladaným ohraničením objektu, a oblasti, ktorá je spojením týchto dvoch predošlých oblastí. V prípade vlastného modelu siete je ako metrika použitá *Accuracy*, avšak je vyrátaná aj metrika IoU.

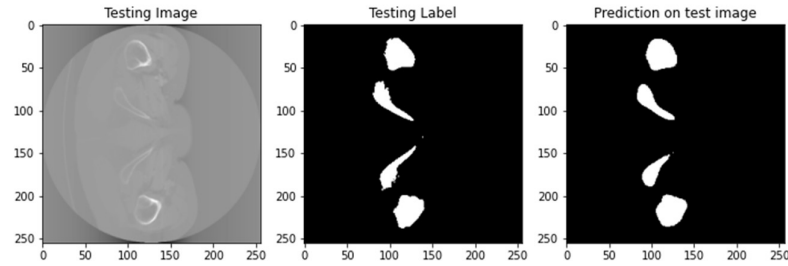
V. VÝSLEDKY EXPERIMENTOV

V prípade vlastného modelu bolo vykonaných 50 epoch pri tréovaní modelu. Metrika *Accuracy* mala po aplikovaní metódy priemerne (z 5 pokusov) *fit()* hodnotu 97%. Priebeh tréovanie je znázornený na nasledujúcich grafoch.



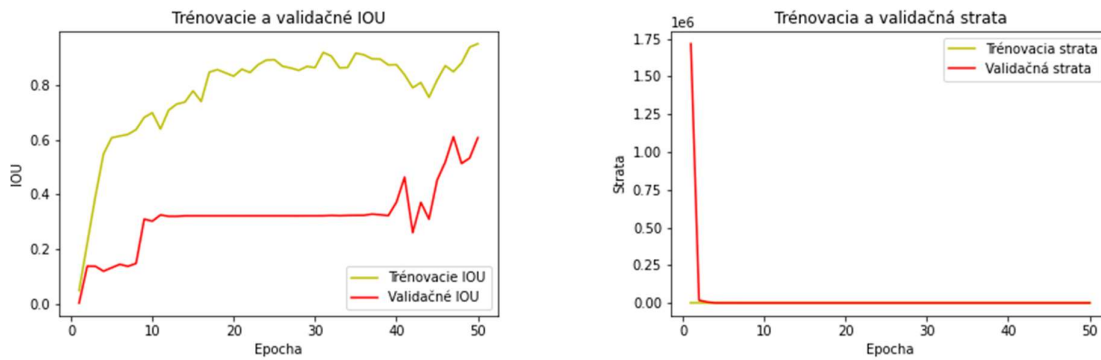
Obrázok 3: Výsledky tréovania pre vlastnú u-net

Metrika priemerného IoU mala hodnotu 33 zo 100, avšak hodnoty pre jednotlivé triedy boli 97 pre triedu 1, 0 pre triedy 2-4 a 69 pre triedu 4. Za pomerne nízku hodnotu priemerného IoU by potenciálne mohla byť zodpovedná absencia všetkých tried v spracovávanom obrázku. Výsledné aplikovanie natrénovaného modelu na náhodný obrázok je znázornený na nasledujúcom obrázku. Z porovnania masky a predikovaného obrázku je vidieť, že sú takmer rovnaké.

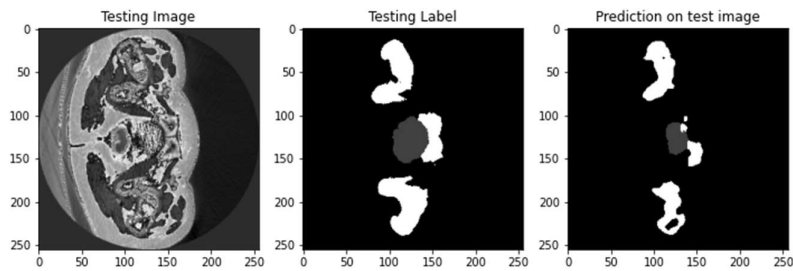


Obrázok 4: Porovnanie originálneho obrázku, masky a predpovede vlastnej U-net

V prípade použitia modelu z knižnice segmentation_models s backbone **resnet34** boli výsledky nasledovné:



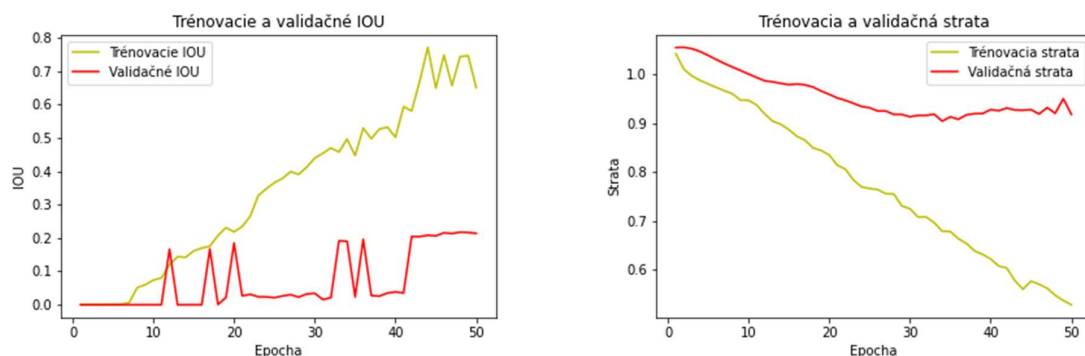
Obrázok 5: Výsledky tréovania pre backbone resnet34



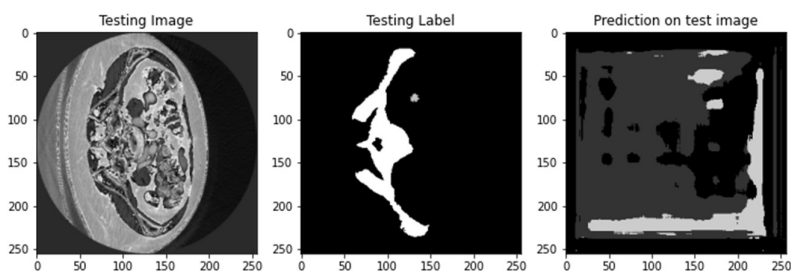
Obrázok 6: Porovnanie originálneho obrázku, masky a predpovede U-net s backbone resnet34

Na porovnávacích obrázkoch je možné vidieť, že model s backbone resnet34 rozpoznal jednotlivé triedy, avšak presnosť rozpoznávaných tried v obrázku nie je dokonalá.

Model s backbone *inceptionv3* vygeneroval nasledujúce výsledky:



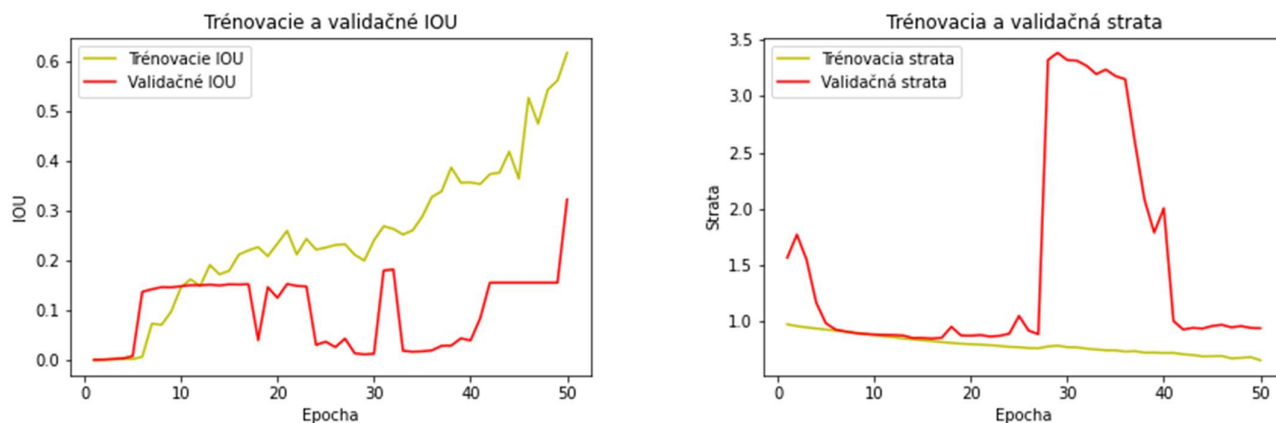
Obrázok 7: Výsledky modelu s backbone *inceptionv3*



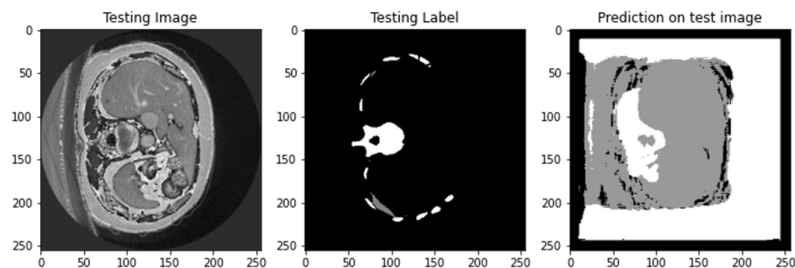
Obrázok 8: : Porovnanie originálneho obrázku, masky a predpovede U-net s backbone *inceptionv3*

Model s backbone *inceptionv3* vykazoval pomerne sľubné dáta, avšak vo finálnom porovnaní nastali problémy a tento model nebol s daným postupom a dátami schopný správne rozpoznať obrázok.

Posledný testovaný model s backbone *vgg16* vygeneroval nasledujúce výsledky:



Obrázok 9: Výsledky modelu s backbone *vgg16*



Obrázok 10: Porovnanie originálneho obrázku, masky a predpovede U-net s backbone vgg16

Model s backbone vgg16 mal podobné problémy ako model s backbone inceptionv3. V praxi model nevedel rozpoznať jednotlivé triedy z obrázka. Problém mohol nastať v spracovaní dát. Pri rozdeľovaní 3D zväzkov boli jednotlivé obrázky vrstiev ukladané vo formáte .tif, avšak dáta obrázkov boli uložené v type uint8, čím mohlo dôjsť k strate niektorých informácií z pôvodných obrázkov.

VI. ZÁVER

Neurónová sieť typu U má vo svete pomerne dobrú reputáciu v sémantickej obrazovej segmentácii. Jej výhody by mali byť pomerne jednoduchá implementácia a presnosť aj pri menších množstvách dát. Z vyššie uvedených experimentov je možné skonštatovať, že pri správnom spracovaní dát vie byť neurónová sieť typu U pomerne efektívna a presná. Zaujímavým experimentom by mohlo byť využiť neurónovú sieť pre tréning a rozpoznávanie na 3D zväzkoch. To by však vyžadovalo vyšší výkon pre spracovanie a taktiež komplikovanejšie spracovanie dát. Celkovo ma však jednoduchá architektúra a pomerne vysoká efektívnosť siete zaujali.