

## Lokalizace jaterních lézí v CT snímcích

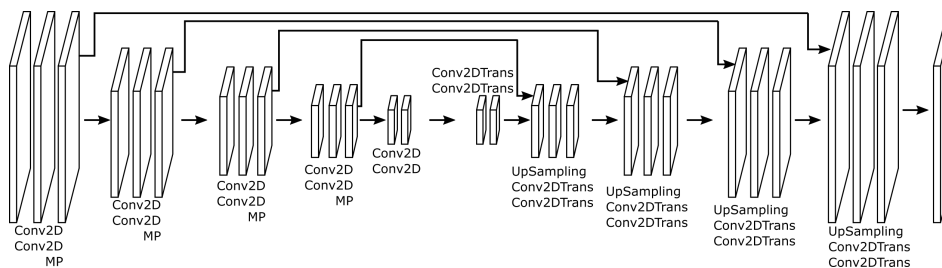
Miroslav Bulka<sup>1</sup>

### 1 Úvod

Významnou úlohou počítačového vidění je segmentace obrazu, která má velké využití v oblasti medicíny, a to například v úloze lokalizace jaterních lézí v CT snímcích. Cílem této práce bylo navrhnout a natrénovat konvoluční neuronovou síť, která by tuto činnost prováděla automaticky. Bylo zapotřebí navrhnout takovou architekturu sítě, která by produkovala na svůj výstup segmentační mapu, kde je každý pixel zařazen do jedné z definovaných tříd.

### 2 Architektura SegNet

Pro takovouto sémantickou segmentaci obrazu byla použita síť, která se inspirovala architekturou SegNet, navrženou v práci Badrinarayanan et al. (2017). SegNet se skládá ze dvou částí, a to z enkodéru a odpovídajícího dekodéru, který má vzhledem k enkodéru zrcadlovou strukturu. Enkodér je zde tvořen pěti za sebou jdoucími bloky obsahujícími několik konvolučních vrstev a na konci jednu max-pooling vrstvu. Navrhovaná architektura byla zjednodušena tak, že obsahovala pouze čtyři bloky a v každém z nich jen dvě konvoluční vrstvy s jednou vrstvou max-pooling. Během návrhu architektury bylo zkoumáno především zorné pole výstupu. Jelikož léze vyskytující se v trénovacích datech byly různých velikostí, bylo žádoucí, aby byla na výstup sítě propagována zorná pole odpovídající často zastoupeným velikostem lézí v datech. To bylo provedeno připojením výstupů vrstev s požadovanými zornými poli na vstup odpovídajících vrstev dekodéru (viz obrázek 1). Tyto způsoby spojení, zvané skip



**Obrázek 1:** Architektura sítě se skip connections.

connections, byly navíc obohaceny o konvoluční vrstvu s velikostí jader  $1 \times 1$ , což je principem tzv. inception modulů (viz Szegedy et al. (2015)). Třetím způsobem bylo využít inception modul o velikosti konvolučních jader  $1 \times 1$  i  $3 \times 3$ , vedených paralelně a zřetěžených před vstupem do vrstvy dekodéru. Byly tedy testovány tři architektury. Pro každou z nich byla navíc testována varianta s přídatnými dvěma konvolučními vrstvami připojenými na výstup dekodéru.

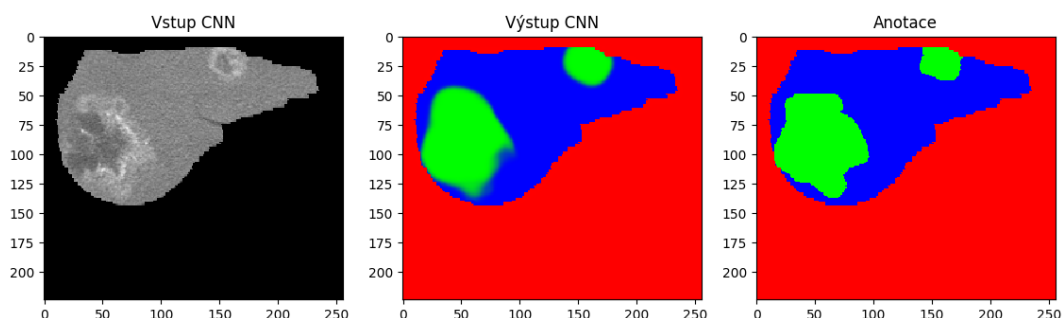
<sup>1</sup> student navazujícího studijního programu Aplikované vědy a informatika, obor Kybernetika a řídicí technika, specializace Umělá inteligence a biokybernetika, e-mail: bulkam@students.zcu.cz

### 3 Hodnotící metody

Pro porovnání výsledků experimentů byla navržena hodnotící metrika, která průměrovala tři hodnoty pravděpodobnosti, s jakou byl výsledný pixel sítě zařazen do dané třídy, pokud do této třídy náležel i ve skutečnosti. Cílem tedy bylo maximalizovat hodnotu této míry  $P$ . Pouze pro účely porovnání s jinými pracemi byly pro výsledky experimentů vyčísleny také hodnoty recall, precision a FPC (False Positive per Case), spočtené metodou využívající Jaccardův index, stejně jako u PASCAL výzvy (viz Everingham et al. (2015)).

### 4 Výsledky

Pro všech šest testovaných architektur byly testovány tři různé optimalizační metody, a to SGD, RMSprop a Adam. Ze všech 18 experimentů vykazovala nejlepší výsledky ( $P = 0.9380$ ) architektura sítě s inception moduly s velikostmi konvolučních jader  $1 \times 1$  i  $3 \times 3$  a s přidavnými dvěma konvolučními vrstvami připojenými na konec dekodéru, přičemž byl využit optimalizátor Adam. Ukázka porovnání skutečné anotace s výstupem sítě pro jeden z řezů je k vidění na obrázku 2. Ohodnocení výsledků pomocí Jaccardova indexu vedlo v případě nejlepšího modelu k hodnotám metrik FPC = 0.1493, recall = 0.6125 a precision = 0.7226.



**Obrázek 2:** Porovnání výstupu sítě a skutečné anotace. Modrá barva představuje zdravou jaterní tkáň, zelená lézi a červená barva značí pozadí.

### Poděkování

Medicínská data byla poskytnuta Fakultní nemocnicí v Plzni v rámci smlouvy o spolupráci mezi FN Plzeň a ZČU v Plzni. Veškerá data byla anonymizována, tj. zbavena veškerých osobních a citlivých údajů (jméno, adresa, rodné číslo).

Tato práce vznikla za podpory projektů CERIT Scientific Cloud (LM2015085) a CESNET (LM2015042) financovaných z programu MŠMT Projekty velkých infrastruktur pro VaVaI.

### Literatura

- Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12), 2481-2495.
- Everingham, Mark, et al. (2015) The pascal visual object classes challenge: A retrospective. *International journal of computer vision*, 111.1: 98-136.
- Szegedy, Christian, et al. (2015) Going deeper with convolutions. In: *CVPR*.