

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Automatická detekce metastáz v histologických obrázcích lymfatických uzlin pomocí hlubokých neuronových sítí
Jméno autora:	Koutecká Pavlína
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Jan Šochman
Pracoviště oponenta práce:	Katedra kybernetiky, FEL, ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání obsahuje takřka přímý postup řešení a požadované výstupy. Vzhledem k dostupnosti podobných řešení a dnešní jednoduchosti učení CNN architektur považuji zadání za rozumné a průměrně náročné.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Cílem bylo implementovat CNN řešení pro úlohu detekce/klasifikace metastáz definované dvěma online soutěžemi (Kaggle a CAMELYON16). Práce toto dále rozšiřuje o účast v soutěži CAMELYON17, která byla navržena jako nepovinné další směřování. Zadání práce tedy považuji za splněné.	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Pro všechny tři úlohy jsou v práci zvolené současně běžně používané a úspěšné architektury, které dosahují nejlepších výsledků i v uvažovaných soutěžích. Při závěrečné klasifikaci závažnosti onemocnění, kde není jasné, která metoda funguje nejlépe, jsou testovány dva možné algoritmy.	
Několik dotazů k postupu uvádím v závěrečné části posudku, nicméně, zvolený postup je určitě pro BP rozumný.	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Velice bych chtěl vyzdvihnout prezentovaný úvod do problému detekce a léčby rakoviny. Úvodní kapitola popisující problém z medicínského pohledu je velice informativní a dobře napsaná.	
Ostatní části jsou popsány také dostatečně. Snad je u DeepLabV3 bych ocenil trochu hlubší vysvětlení, ale nebyl problém to dohledat.	

Formální a jazyková úroveň, rozsah práce

A - výborně

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce je napsaná výbornou angličtinou, je velmi čtivá a velmi pečlivě strukturovaná. Až na pár (jednotky) drobných chybiček jsem nenašel nic, co by bránilo v pochopení textu či bylo formálně nesprávně. Práce je dobře vyvážená co se týká motivace, popisu problému a následného řešení a experimentů.

Jediné vylepšení, které mě napadá je trochu jiná organizace kapitol. Pro čtenáře by bylo o něco snazší, kdyby se tolik neprolínaly popisy těch tří problémů, ale mohl se soustředit vždy na vysvětlení jednoho problému od přípravy dat, teorie, implementace, přes výsledky až po závěrečnou diskusi. Takto jsem musel mít práci otevřenou na mnoha místech zároveň, pokud jsem chtěl zhodnotit, jestli navržená architektura opravdu funguje.

A také není formálně zcela korektní nazývat výstup segmentační sítě pravděpodobností. Je to jen skóre.

Výběr zdrojů, korektnost citací

B - velmi dobře

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Práce velmi pečlivě cituje původní zdroje. Co se týče kompletnosti použitých zdrojů, tak bych možná navrhl se podívat i mimo uvažované soutěže. Určitě existují podobné problémy, kde jsou metody lépe zdokumentovány. Mohlo by se tak objevit více vzhledů do toho, co funguje, co je zásadní a co naopak nedůležité.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Dosažené výsledky odpovídají zadání práce, kde cílem bylo seznámit se s problémem a navrhnout základní řešení. Přestože tedy jsou výsledky o něco horší, než mnohé konkurenční metody, pro splnění zadání jsou dostatečné. Text práce dále uvádí možné směry vylepšení, které by mohly výsledky dále posunout.

Snad jen pár komentářů:

Křivky z průběhu učení ukazují, že trénovací chyba klesá, zatímco validační už od začátku stagnuje (Fig 6.4 – 6.6). Je to zdůvodněno přefitováním, ale to by validační křivka stoupala. Spíš to vypadá jako by distribuce na trénovací a validační sadě byla rozdílná... Kdybych toto viděl u svého experimentu, hledal bych chybu. Zajímá mě váš komentář k tomuto jevu.

Dále bych velice ocenil shrnující tabulku či graf s porovnáním navržené metody se state-of-the-art a případně i s úspěšností patologa. Tyto informace v práci jsou, ale díky zvolené struktuře textu jsou roztroušené po jednotlivých sekcích, takže není úplně snadné si je k sobě přiřadit.

Vizuální výsledek v Fig 6.10 vypadá velice dobře. Jestli jsem to dobře pochopil, tak je to „agregovaný“ výstup z Fig 6.9. Tam ale výsledky vypadají výrazně hůř. Jak je toto možné? Co znamená ta agregace?

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Práci hodnotím velmi kladně. Je dobře napsaná, zadání je splněno a dokonce rozšířeno a výstupy odpovídají očekávání. Při obhajobě bych poprosil o zodpovězení následujících otázek:

1. I přes vynikající text a vizuální prezentaci se mi nepodařilo získat vhled do toho, jak vypadá metastáze a jak ji odlišit od zdravé tkáně. Na prezentovaných obrázcích není vůbec zřejmé, jak jedno od druhého odlišit. Pravděpodobně to tak nějak bude i v realitě. Nicméně, dal by se ten rozdíl nějak vizualizovat? Bez tohoto vhledu se jen velmi těžko hodnotí kvalitativní výstupy...
2. Prosím, okomentujte průběh trénovacích a validačních křivek učení (viz výše).
3. Prosím okomentujte rozdíl ve výsledku ve Fig 6.9 a 6.10 (viz výše).
4. Pro CAMELYON soutěž je oproti Kaggle použita větší backbone síť (ResNet101 namísto ResNet50), která tedy má mnohem více parametrů. Text ale říká, že se v tomto případě nepoužije předtrénování na jiné datové sadě (pro Kaggle se použilo). Očekával bych tedy pomalejší učení, které bude vyžadovat více epoch a je tu i větší riziko přeučení se na tento problém. Dalo by se něco takového v experimentech vysledovat? Případně, jak tomu bylo zabráněno. Například Fig 7.5 vypadá, jako že se síť naučila jednoduché vzorky, ale ty složitě ne...
5. Využití „Initial learning rate searching process“: Toto není úplně standardní postup. Hledal jsem k tomu nějaké reference a půlka lidí říká, že to pomáhá, druhá, že vůbec. Běžný postup to v komunitě rozhodně není. Zkoušeli jste pro váš problém pustit učení s/bez tohoto? Opravdu to pomáhá? A jaký byl vlastně použit optimalizační algoritmus?
6. Proč byl použit grid search algoritmus pro hledání velikosti nejmenšího nádoru pro zahazení falešných detekcí, když tuto hodnotu lze snadno získat z ground truth?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 24.8.2020

Podpis: