

## I. IDENTIFIKAČNÍ ÚDAJE

<b>Název práce:</b>	<b>Metody pro hierarchickou klasifikaci histopatologických obrazů</b>
<b>Jméno autora:</b>	<b>Aleš Král</b>
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra kybernetiky
<b>Oponent práce:</b>	Jan Šochman
<b>Pracoviště oponenta práce:</b>	Katedra kybernetiky, FEL, ČVUT

## II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

<b>Zadání</b>	<b>náročnější</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Cílem práce bylo pochopit, popsat a (případně) naimplementovat pět metod popsaných ve specifikovaných publikacích a porovnat je na problému klasifikace histopatologických obrazů. Bylo možno použít existující kód, pokud je k dispozici (dvě metody mají kód na github.com [4, 5], jedna je vytvořena školitelem [1] a ke dvěma kód veřejně dostupný není [2, 3]).	
Vzhledem k množství metod hodnotím zadání spíše jako náročnější. I když by se podařilo spustit všechny tři dostupné metody bez problémů (nepředpokládám), tak naimplementovat dvě metody podle článku není úplně snadné.	

<b>Splnění zadání</b>	<b>splněno s většími výhradami</b>
<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>	
V práci je popsána implementace dvou metod bez dostupného kódu [2, 3], je použita vlastní implementace [4] a jsou popsány experimenty s metodou dostupnou na github [5]. Metoda vytvořená školitelem [1] není prací ani citována, ani není podáno zdůvodnění, proč nebyla vyhodnocena. Vzhledem k náročnosti práce by se toto ale dalo pochopit.	
Co se týče porovnání metod, zde vidím větší nedostatky. Práce postupuje při jejich implementaci a vyhodnocení velmi nemetodologicky. Není jasně ohraničené, kde se práce odchyluje od původně popsané metody a navržené modifikace nejsou vyhodnoceny zvlášť, ale vždy až jako celek. Tím se ztrácí možnost zjistit, která komponenta, či která modifikace může za výsledné zhoršení výsledků popsané v experimentech.	

<b>Zvolený postup řešení</b>	<b>správný</b>
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Postup je poměrně jasně daný snahou o implementaci daných metod a jejich aplikaci na histopatologických datech. Zde je vše v pořádku.	

<b>Odborná úroveň</b>	<b>C - dobře</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce obsahuje vysvětlení velkého množství algoritmů a metod. Bohužel, značná část z nich je popsána tak, že jsem z textu nedokázal pochopit, jak fungují (Otsu's algorithm, MIL, Data preprocessing v sekci 3.3.1, pro pochopení CLAM bych potřeboval nejlépe obrázek, rovnice (3.4), sekce Clustering network, ...). Důvodem jsou často příliš neformální styl vysvětlení, nepřesné formulace, občas jsou použité rovnice, které nedávají smysl a význam jednotlivých symbolů se mění občas i odstavec od odstavce.	
Další nedostatek je v popisu vlastní implementace. Autor činí spoustu rozhodnutí, jak metody měnit a jen velmi zřídka toto rozhodnutí řádně zdůvodní a experimentálně ověří jako správné. To se následně projeví tím, že navržené implementace dosahují výrazně horších výsledků, než referenční metody. Ideálně bychom chtěli mít metodu implementovanou co nejdříve a od té se následně odchylovat a zjišťovat, jestli nějaká změna nepřinese lepší výsledky.	

V detailním popisu metody pak není dostatečně zdůrazněné, kde jde o implementaci metody jako takové a kdy jde o navrženou změnu. Ideálně bych si představoval sekci Contributions, kde by bylo přehledně shrnuto, co se konkrétně podařilo a co ne.

**Formální a jazyková úroveň, rozsah práce**

**B - velmi dobře**

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

Našel jsem několik rovnic, které vyložene nedávají smysl. Např. rovnice (4.7) a (4.8) jsem nepochopil, rovnice (4.9) a (4.10) lze zredukovat na  $I_C = I_C / \ln(10)$  a je tedy nejasné, proč se o nich mluví jako o dekonvoluci. Také nevím, co je „norm“ v rovnici (4.12). Není také úplně ideální v rovnicích používat dvoupísmenná značení veličin (např.  $OD_C$ ,  $ODM$ ,  $BL$ ,  $CL$ )

Práce používá nedostatečný systém odkazování se na rovnice, obrázky či kapitoly jen formou čísla. To je značně matoucí (např. „One was to pass the feature vectors into a MIL 2.5 classifier and the other...“, kde 2.5 není verze MIL, ale číslo kapitoly). Správně je všude používat: Section 2.5, equation (3), Figure 4.1, Table 1.3.

Zcela nedostatečné a matoucí je popisovat sítě pomocí matic vah (např. „...the clustering network  $W_{inst,m} \in R^{2 \times 512}$  ...“). Neuronové sítě mají typicky nějaký vstup, nějaký výstup, počet vrstev, nonlinearity, pooling vrstvy, ...

Částečně matoucí je i stejné pojmenování datové struktury (Quadtree) a následného využití této reprezentace v celém frameworku zahrnujícím i CNN (zase se na to odkazuje jako na Quadtree).

V jednu chvíli (Fig 4.1) je dokonce kus textu vysvětlující navrženou změnu nahrazen jen obrázkem bez dalšího komentáře.

**Výběr zdrojů, korektnost citací**

**B - velmi dobře**

*Vyjáďte 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.*

Zdroje jsou relevantní a citace používány korektně. Jen u několika by stálo za to dohledat původní článek a ne jen poslední použití myšlenky (MIL, quadtree).

Také sekce referencí by si zasloužila trochu poklidit (několikrát „et al.“, podivné zkratky konferencí i časopisů, kapitálky, odkaz na arxiv, ...).

**Další komentáře a hodnocení**

*Vyjáďte 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.*

Z prezentovaných výsledků vyplývá, že vzniklo několik implementací inspirovaných metodami popsány v článkách zmíněných v zadání. Ani jedna z metod bohužel nedosahuje přesnosti prezentované v původních článkách. Důvodem je pravděpodobně volba přístupu v místech, kde jsou navrženy modifikace – buď proto, že z článku nebylo úplně zřejmé, jak to udělat, nebo i bez nějakého jasného důvodu. V práci úplně chybí jakákoliv „ablation study“.

Původní myšlenka porovnání známých metod na stených datech se tedy nepodařila zcela zrealizovat, nicméně je vidět, že za prací je spousta úsilí a snahy pochopit implementované metody.

### 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.*

Hlavními problémy práce jsou formální úroveň popisu a nedostatečné experimentální vyhodnocení modifikovaných částí implementovaných metod. Jelikož bylo ale zadání dle mého názoru spíše obtížnější, volím i přes všechny výše zmíněné nedostatky mírnější hodnocení B.

U obhajoby prosím o zodpovězení těchto otázek:

1. Poprosil bych o vysvětlení světle modré krabičky ve Fig 3.2, která, zdá se, naznačuje, že se používá multiple-instance *learning* i v inferenci. Neměl by tou dobou už být klasifikátor naučený a zafixovaný?
2. Recurrent visual attention model (Sekce 4.2) v sobě přirozeně skrývá rekurenci, která je ale nakonec řešena jen posláním stavu z kroku  $t$  do zpracování sítí v kroku  $t+1$ . Figure 4.1 naznačuje, že takovýchto „rekurzivních“ kroků tam je více. Očekával bych tedy, že začne být problém s vanishing gradients. Uměl byste vysvětlit, proč tento problém nenastává?
3. Mohl byste přehledně vysvětlit rovnice (4.7) a (4.8) včetně všech použitých symbolů?
4. V sekci 4.3.2 se používá pro generování příznaků síť naučená na ImageNet datasetu. Nicméně, data, která tato síť dostává jsou diametrálně odlišná. Mohl byste tuto volbu okomentovat? Proč to je dobrá volba?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 1.6.2022

Podpis: