

Posudek oponenta diplomové práce

Student: Pavlica Jan, Bc.
Téma: Re-identifikace graffiti tagů (id 19536)
Oponent: Beran Vítězslav, Ing., Ph.D., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Zadání vyžaduje využití konvolučních neuronových sítí. Tím je zvýšena jeho obtížnost, a to jak v potřebě zajištění vhodné datové sady (pro tyto úlohy nemůže obsahovat jen pár desítek obrázků), tak náročnost úkolu vhodné zvolit architekturu a natrénovat úspěšný model na datech, která jsou spíše obtížného charakteru.
- 2. Splnění požadavků zadání** **zadání splněno**
- 3. Rozsah technické zprávy** **splňuje pouze minimální požadavky**
Experimenty jsou popsány skoro na 20ti stránkách technické zprávy. S ohledem na typ práce je to v pořádku, ale více prostoru a pečlivosti by si zasloužily informace nejen o re-identifikaci, ale např. i předzpracování a augmentaci datové sady, představení vybraných architektur ResNet a MobileNet, jejich základní rozdíl apod.
- 4. Prezentací úroveň předložené práce** **70 b. (C)**
Struktura technické zprávy je logická, ovšem značně nevyvážená. Řada kapitol má bez odůvodnění jednu kapitolu 3tí úrovně a není jasné, k čemu toto strukturování slouží. Důležitost podkapitol (měřeno prostorem, který je částem textu věnován) je velmi nevyvážená, některé podkapitoly mají 5 stran (např. 2.2), jiné 6 řádků (např. 3.1). Jinak je text, až na několik drobných poznámek níže, sepsán odborně a čtivě.
Dělení "extrakce příznaků" ve vrstvách hluboké konvoluční sítě, založené na autorovu popisu vizualizace (obr. 2.11 převzatý z [29]) je poněkud odvážné. Nevhodné je i přílišné zobecnění popisu cíle algoritmu učení na "nastavení vah w " (tamtéž), kdy poněkud zaniká konvoluční princip, tedy že se jedná o koeficienty jader konvolučních filtrů, zejména pojednává-li kapitola o klasických a konvolučních sítích. V kap. 3.3 nazvat vzdálenostní funkce (výsledek je skalár) *funkcemi spojovacími* (kde by se očekával výsledek vektor) je matoucí. Autor běžně používá slovo *query*, ačkoliv termín *dotaz* je běžně používaným a zcela adekvátním termínem v české terminologii. V popisu obrázků 5.7 a 5.9, a 5.8 a 5.10 by bylo vhodné uvést i síť, na které byly výsledky získány. Některé části zprávy by si zasloužili více pozornosti.
- 5. Formální úprava technické zprávy** **85 b. (B)**
Typografická úroveň zprávy je velmi dobrá (práce je vysázená v LaTeXu), text je bez chyb a autor se vyjadřuje srozumitelně a odborně. Některé obrázky by bylo vhodné přeložit a zejména zajistit v lepším rozlišení.
- 6. Práce s literaturou** **65 b. (D)**
Autor se odkazuje na 33 studijních pramenů, které jsou k tématu práce relevantní. Problematické je ale využití některých zdrojů, např. z řady zdrojů byl použit pouze jeden obrázek, např. 5, 11, 15, 17, 31 a pravděpodobně další. Teoretický přehled neuronových sítí autor čerpá se své vlastní bakalářské práce ([19]). Zde by bylo na místě odkazovat na původní zdroje. Řada zdrojů také není uvedena kompletně (např. 10, 11, 12 a další).
- 7. Realizační výstup** **90 b. (A)**
Práce je experimentálního charakteru. Autor provádí řadu experimentů, kdy mění celou řadu parametrů (od architektury sítí, přes použité metriky a loss funkce, až po způsoby augmentace datasetu). Parametry jsou variovány s rozmyslem a důsledně vyhodnoceny. Ve výsledku tak nachází vhodné parametry, které dosahují, s ohledem na obtížnost úlohy, velmi dobrých výsledků. Těch bylo dosaženo výběrem vhodných architektur sítí, přípravou a vhodným použitím vlastní datové sady a vytvořením řady relevantních a technicky kvalitně napsaných skriptů pro: trénování, testování, vyhodnocení, loss funkce a augmentaci datové sady.
- 8. Využitelnost výsledků**
Výsledkem práce je nejen sada skriptů využitelných pro vývoj dalších re-identifikačních úloh, ale zejména vytvořená datová sada. Autor si sice správně stěžuje, že je pořád ještě *malá*, ale i tak je to velmi cenný základ pro případný další vývoj.
- 9. Otázky k obhajobě**
 - Jaký je rozdíl mezi architekturami ResNet a MobileNet? V čem to je, že s MobileNet lze v této úloze dosáhnout lepších výsledků?
 - Při modifikaci (augmentaci) datové sady využíváte vcelku málo možností. Jaký je důvod, že jste nevyužil více možností, zejména když trpíte nedostatkem dat? Např. geometrické transformace (afinní, perspektivní) či jasové a barevné transformace (posun barev, kompresní šum, strukturální šum, nelineární korekce atd.).

10. Souhrnné hodnocení

88 b. velmi dobře (B)

Pan Pavlica vyřešil úlohu re-identifikace graffity tagů s využitím konvoluční neuronové sítě MobileNet. Úspěšného řešení dosáhl díky vlastní dostatečně rozsáhlé a kvalitní datové sadě a vhodnou volbou parametrů při trénování různých variant modelů v celé řadě experimentů. Autor v závěru uvádí, že hlavním výsledkem je práce je jednoduchá aplikace s GUI. S tím nelze souhlasit, protože hlavním výsledkem je funkční model neuronové sítě, sada vhodných parametrů a její úspěšné natrénování! S ohledem na obtížnost úlohy je to velmi zdařilý výsledek.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 6. srpna 2020

Beran Vítězslav, Ing., Ph.D.
oponent