

Posudek disertační práce

Autor: Ing. Pavel Svoboda, Fakulta informačních technologií, VUT Brno
Název: Image Restoration Based on Convolutional Neural Networks

K posouzení mi byla předložena práce výše uvedeného názvu a autora. Práce má 99 číslovaných stran a člení se do sedmi kapitol (1. Introduction, 2. Engineered Image Restoration, 3. Convolutional Neural Networks, 4. Image Processing Based on Neural Networks, 5. CNN image restoration, 6. Experiments, 7. Conclusion).

Motivací pro práci je představa, že mnoho metod pro rekonstrukci obrazu lze nahradit unifikovaným přístupem založeným na využití konvoluční neuronové sítě. Dříve obvyklou detailní teroretickou analýzu jednotlivých dílčích případů poškození obrazu a stanovení odpovídající rekonstrukční metody lze pak nahradit stanovením architektury sítě, shromážděním dostatečného množství vzorků a naučením sítě (samozřejmě včetně nevyhnutelného vyhodnocení výsledků). S takto autorem formulovanou motivací nelze než souhlasit. Lze ji považovat za hlavní přínos konvolučních neuronových sítí. Téma aplikace konvolučních neuronových sítí lze tak dnes považovat za velmi aktuální, prodávající další vývoj a obecně užitečné. Existence do značné míry univerzálního aparátu konvolučních neuronových sítí činí ovšem disertační práci současně také poněkud obtížněji realizovatelnou, pokud se přidržíme názoru, že by disertační práce měla přinášet nové poznatky, např. poznatky teoretické povahy.

Strukturu práce považuji za poněkud neobvyklou (byť, s přihlédnutím ke zvolenému tématu, za i do určité míry pochopitelnou). Značný objem práce je věnován kapitolám, které bych označil jako průpravné (kapitoly 2, 3, 4 a konec konců i významné části kapitoly 5), přičemž jen některé informace a odkazy v uvedených kapitolách je možné považovat za popis současného stavu v oblasti, v níž se autor hodlá pohybovat, či za zavedení nebo rekapitulaci terminologie. Vzato striktně, lze vlastní přínos autora práce nalézt až v kapitole 6 Experimenty, což považuji za onu neobvyklost. Autor se zde zaměřuje na dva problémy, a to na zaostření obrázků obsahujících registrační značky vozidel a na odstranění artefaktů způsobených kompresí JPEG. Třebaže kapitolu 6 nelze označit za krátkou (má cca 20 stran), přeci jen bych si představoval větší zaměření se na vlastní výsledky, a to tím spíš, že úvodní kapitoly byly pojednány vcelku bohatě (ne vše, co v nich bylo prezentováno, bylo nakonec k pochopení vlastní práce autora nezbytně nutně zapotřebí).

V první aplikaci konvolučních sítí se pan disertant zaměřil na odstranění pohybového rozmazání se zvláštním zřetelem na obrazy registračních značek automobilů (jak soudím z obrázků, které jsou na odpovídajícím místě práce uvedeny, a také z korespondující publikační položky). Architektura sítě byla převzata z [35] s odůvodněním, že tam se již osvědčila. Výsledky experimentu jsou doloženy jednak vizuálně (obrázek 6.7) a také srovnáním s algoritmem podle [5]. Autor uzavírá, že jím v experimentech dosažené výsledky jsou lepší než výsledky prezentované v [5] a dodává, že i z hlediska výpočetní náročnosti je metoda použitelná. S tím je zřejmě možné souhlasit. Za jistý problém zde ovšem považuji míru přínosu pana disertanta. Jestliže architektura sítě byla převzata z [35] (a jestliže situaci chápu správně), pak se do značné míry jedná o experiment potvrzující práci [35], byť zřejmě za poněkud modifikovaných podmínek. Aby uvedenou situaci kompenzoval, mohl se pan disertant ptát, jaké další informace/výsledky by mohl odborné komunitě sdělit/předat. Mohl se třeba více zabývat architekturou sítě (je použita architektura skutečně optimální v tom smyslu, že jakákoli změna parametrů by vedla ke zhoršení výsledků?), mohly být diskutovány otázky rychlosti trénování a vlastní rekonstrukce, mohlo být naznačeno, jaký software byl použit, či snad mohly být výsledky odborné komunitě dány k dispozici dokonce ve formě odpovídajícího software a použitých testovacích obrázků. Bylo by pěkné, kdyby pan disertant mohl do doby obhajoby na uvedené přimomínky/návrhy nějak reagovat.

V oblasti odstraňování artefaktů při JPEG kompresi navrhuje pan disertant použití konvolučních sítí s architekturami, které označuje jako L4, L5 a L8. V textu předložené práce jsem nenalezl ostré explicitní prohlášení, že by se pan disertant považoval za autora myšlenky použití těchto konkrétních architektur pro daný problém. Soudím ale, že jím pravděpodobně je. Jeho architektury L4 a L8 jsem zahlédl citovány v některých člancích z tohoto roku, což hodnotím velmi kladně. Nejméně ve dvou případech byly pak i zahrnuty do testování prezentovaného v oněch člancích. Na mém pozitivním hodnocení v tomto případě nic nemění fakt, že v nových člancích autoři uváděli dosažené parametry nepatrně lepší než u architektur pana disertanta. Směřem k panu disertantovi si naopak neodpustím poznámku, že fakt, že jej lze považovat za autora myšlenky použití architektur L4, L8 měl v práci velmi zřetelně explicitně vyjádřit (není povinností oponenta, aby to za něj dohledával) a zejména tato část práce pak mohla být rozsáhlejší (a třeba i prezentována jako první, před zaostřením registračních značek).

Práce je napsána anglicky se solidní formální úpravou. Jazykových prohrěšků (pokud já mohu posoudit) jsem zahlédl vcelku minimálně (např. předložky atd.). Text práce je dobře srozumitelný a poskytuje současně i dobrý všeobecný přehled o oblasti, v níž se pan autor pohyboval.

V databázi SCOPUS jsem pro pana disertanta našel celkem 4 publikační položky (z toho 3 články z konferencí). Ve dvou případech byl pan disertant prvním autorem (CNN for license plate motion deblurring, ICIP 2016; Compression artifacts removal using convolutional neural networks, WSCG 2016). Citace, o nichž jsem se dříve zmiňoval, jsem našel ve člancích, které zatím ještě zřejmě nejsou ve SCOPUS uvedeny, jedná se o citace nevlastní. Počet publikací prezentujících disertabilní jádro práce lze sice považovat za hraničně nízký, avšak s přihlédnutím k výše zmíněným citacím si jej dovolím označit za akceptovatelný. Dále také předpokládám, že pan disertant byl platným členem autorských kolektivů i u ostatních článků, jak zajisté u obhajoby upřesní.

Závěr: Předložená práce prokazuje, že se pan disertant podrobně seznámil se vcelku náročnou a současně velmi aktuální problematikou konvolučních neuronových sítí. U části zaměřené na zaostření znaků lze za přínos pana disertanta považovat nejspíš jen provedení odpovídajících experimentů. U části zaměřené na odstranění artefaktů způsobených kompresí se zdá, že je pan disertant odbornou komunitou považován za autora odpovídajících architektur a je komunitou i citován a tedy i akceptován, což považuji za velmi významé. Text práce mohl být strukturován s větším akcentem na vlastní práci a vlastní přínos pana disertanta. Po pečlivém zvážení všech okolností *doporučuji práci k obhajobě.*

V Ostravě 28. 6. 2017

doc. Dr. Ing. Eduard Sojka
FEI, VŠB-TU Ostrava