
Hodnocení vedoucího diplomové práce

Bc. Vojtěch Frič

Image Similarity Assessment

Diplomová práce Bc. Vojtěcha Friče se zabývá problémem automatického vyhodnocování podobnosti digitálních snímků v kolekci a vychází z dlouhodobého zájmu diplomanta o problematiku strojového vidění a učení (tématu strojového vidění se věnoval již ve své bakalářské práci). Zadání práce zároveň souvisí s vědecko-výzkumnou činností vedoucího, na které se diplomant aktivně podílel.

Diplomant k práci přistupoval po celou dobu velmi příkladně. Iniciativně vyhledával vědecké publikace a konzultoval vhodnost a možnosti implementace v nich uváděných technik. Prostudoval řadu poměrně značně komplikovaných postupů, přičemž k jejich úplnému pochopení musel nabýt dalších znalostí značně překračujících rámec navazujícího magisterského studia.

Nakonec implementoval dvě perspektivní metody posuzování podobnosti digitálních snímků, a to na základě neuronové sítě s hlubokým učením (Deep Belief Network, DBN) a na základě waveletové transformace. Tyto metody pak porovnává se (též vlastnoručně implementovanými) základními metodami postavenými na měření prosté eukleidovské vzdálenosti snímků a PCA.

Autor se rozhodl napsat průvodní text práce z vlastního popudu v anglickém jazyce, čímž vedoucího velice potěšil. Také oceňují odvalu autora, s níž se pustil do práce s DBN, což je technika zcela nová a zatím nepřilíš prozkoumaná.

Diplomant pravidelně docházel na konzultace s vedoucím, vždy vzorně připraven. Průběh diskusí zejména v pozdějších fázích práce svědčil o vynikajícím přehledu ve všech dotčených oblastech. Na vznesené požadavky či drobné výtky reagoval prakticky okamžitě a k naprosté spokojenosti vedoucího je řešil.

Práce je původní, autor vycházel z řady vědeckých publikací, a to jak knižních, tak aktuálních článků z nedávno proběhlých konferencí. Bibliografie je rozsáhlá (23 titulů) a relevantní. Citace v textu i bibliografie na konci práce jsou provedené v souladu s požadavky.

Implementační část předloženého díla je zdrojem značné kontroverze: Na jedné straně stojí autorem uváděné dosažené výsledky, které jsou bezesporu vynikající. Na straně druhé pak skutečnost, že software na přiloženém CD je v takovém stavu, že jej nelze bez vynaložení značného úsilí ani spustit. Autor v příloze B.1 uvádí seznam požadavků, přičemž jen instalace součástí z tohoto seznamu je skutečné dobrodružství – vedoucí práce strávil přípravou běhového prostředí a instalací všech potřebných komponent řadu hodin a byl nucen celkem 4× přeinstalovat operační systém (což ovšem nelze klást za vinu autorovi). V přílohách B.2 a B.3 pak autor popisuje, jak software spustit a otestovat. Bohužel na CD chybí některé skripty, které je třeba spustit, aby aplikace fungovala.

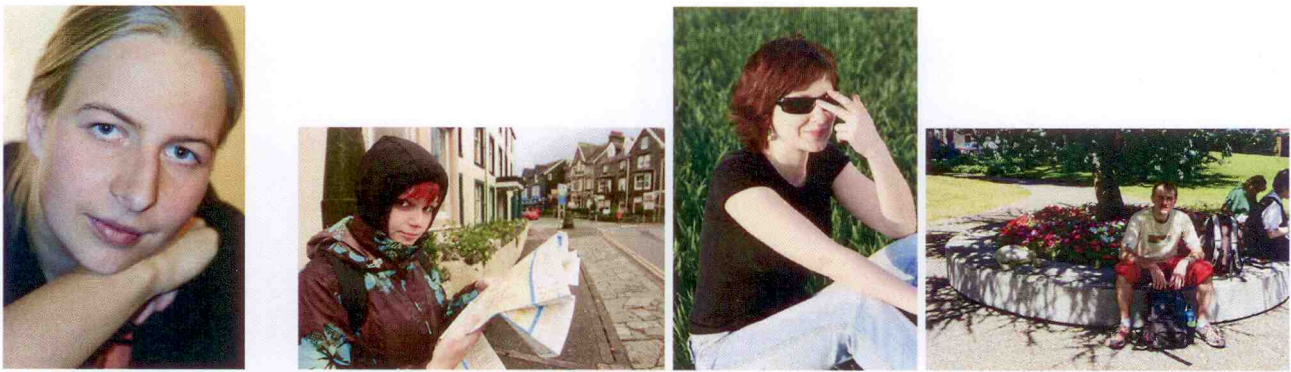
K vývoji byl použit jazyk Python a operační systém Linux. Zdrojový kód programového řešení je zapsán ve shodě se zvyklostmi, je celkem dobře čitelný a přehledný. Je ovšem poměrně málo komentovaný. Na několika místech zdrojového kódu byly chyby, které si vedoucí musel sám opravit, aby mohl aplikaci spustit a otestovat.

Výsledky, které aplikace poskytla při testování vedoucím, se neshodují s těmi uvedenými autorem zejména ve smyslu spolehlivosti indikace podobných snímků. Při testování vykazovala aplikace velice diskutabilní výkony: Nalezené snímky z kolekce, které se měly podobat testovanému snímku, byly značně nepodobné, viz obr. 1.

Textová část díla je přiměřeně rozsáhlá – má 65 stran, je napsána vynikající technickou angličtinou a vysázena v L^AT_EXu. Grafická úroveň je zcela profesionální a celkově působí velmi harmonickým dojmem. Text práce je srozumitelný a čtivý, autorovy vyjadřovací schopnosti jsou velmi dobré. Gramatické ani typografické chyby se v práci prakticky nevyskytují.

Text je doplněn kvalitními ilustracemi, vzorci a výpisy kódu, které text žádoucím způsobem obohacují a přispívají k pochopení vykládané problematiky.

Práce je dobře logicky strukturovaná, poměr jednotlivých částí textu bohužel není zcela vyvážený: Teoretická část výrazně převažuje nad popisem implementace, ovšem na druhé straně v tomto případě (výzkumný prototyp) nejsou detaily implementace až tak zásadní. Nicméně s ohledem na níže zmíněný zásadní přínos práce lze tento fakt považovat za největší nedostatek textové části.



Obrázek 1: Testovací snímek (zcela vlevo), k němuž aplikace hledala v kolekci podobné; nalezené „podobné“ snímky z kolekce (vpravo od testovacího) se snižujícím se skóre podobnosti.

Největším přínosem této práce je nepochybně její demonstrační hodnota: Ukazuje totiž velmi podrobně, jak používat knihovny NumPy a SciPy pro jazyk Python a zejména pak rozsáhlou knihovnu Theano, která je prakticky nezbytným nástrojem při zkoumání neuronových sítí s hlubokým učení.

Všechny body zadání byly splněny s následujícími výhradami: V bodu 1 zadání je diplomantovi uloženo seznámit se s existujícími algoritmy a metrikami pro posuzování podobnosti digitálních snímků. To sice učinil, ale dosti stručně, a opominul některé významné, běžně používané metody, jako např. fázovou kongruenci, atp.

V bodě 2 zadání byl formulován požadavek na přehledné vyjádření podobnosti snímků. Autorovo řešení tohoto požadavku je redukováno na výpis prvků pole, které obsahuje indexy hodnocených snímků v pořadí od největší po nejmenší míru shody. To nepovažuji za dostatečně přehledné vyjádření podobnosti snímků v kolekci.

Nicméně práce představuje hodnotné dílo zejména proto, že autor úspěšně aplikoval techniku DBN za pomoci knihovny Theano, tzn. práci lze použít i jako návod, jak s touto knihovnou pracovat. O úplné řešení zadaného problému v podobě a kvalitě použitelné pro nasazení do ostrého provozu se ovšem v žádném případě nejedná. Autor však projevils značnou snahu a píli a výsledek lze akceptovat jako výzkumný prototyp.

Práci **doporučuji k obhajobě** a i přes uvedené výhrady zejména s ohledem na výkon, který musel autor podat při studiu techniky DBN a při zprovoznění a integraci knihovny Theano, hodnotím ještě klasifikačním stupněm

„velmi dobře“.

Kamil Ekštejn
 Ing. Kamil Ekštejn, Ph.D.
 KIV FAV ZČU

V Plzni dne 27. 5. 2014

Doplňující otázky:

1. Podle jakého klíče jste vybíral metody uvedené v předhledu existujících metod (kap. 2)?
2. Co by mohlo být příčinou tak mizerného výkonu metody MRIQ při testování na kolekci PARMAL-Faces01, provedeném vedoucím práce?

Západočeská univerzita v Plzni
 Fakulta aplikovaných věd
 katedra informatiky a výpočetní techniky

①

Oh!