

Hodnocení vedoucího bakalářské práce

Daniel Schnurpfeil

Klasifikátor architektonických slohů ze snímků budov

Bakalářská práce pana Daniela Schnurpfeila se zabývá zadáním ze třídy v poslední době velmi populárních problémů, a to tvorbou aplikace pro výkonná mobilní zařízení (smartphony, apod.) se schopností rozpoznávat či klasifikovat nějaká vstupní data, která pořizuje prostřednictvím svých senzorů přímo předmětné zařízení. Takových aplikací jsou již k dispozici tisíce a více či méně úspěšně řeší různé specifické úlohy od klasifikace fotoaparátů zařízení snímaných květin, dopravních značek, registračních značek vozidel, přes různé rozpoznávače zvuků, detektory počtu ušlých kroků, atp.

Student Daniel Schnurpfeil si zvolil téma klasifikace architektonických slohů budov. Cílem práce tedy bylo na výkonném počítači natrénovat dostatečně přesný a spolehlivý klasifikátor, který na vstupu přijímá snímek budovy a na výstupu poskytuje identifikátor třídy odpovídající jednomu konkrétnímu slohu, ve kterém je budova na snímku vystavěna. Tento klasifikátor je pak provozován (již natrénovaný) na výrazně méně výkonném mobilním zařízení v rámci uživatelské aplikace, která dovoluje snímky budov pořizovat či zavádět z jiného zdroje než je fotoaparát zařízení, a následně uživateli zobrazuje N nejpravděpodobnějších slohů, které klasifikátor detekoval ve snímku.

Autor si zadání vybral sám, resp. na základě vlastního zájmu dodefinoval obecné zadání vedoucího práce. Všechny dílčí cíle práce se podařilo dosáhnout: K dispozici je tak trénovací procedura, která dovoluje optimálně nastavit parametry klasifikátoru na bázi hluboké konvoluční neuronové sítě (CDNN), a dále dvě uživatelské aplikace, které tento klasifikátor využívají – aplikace pro mobilní zařízení s operačním systémem Android a webová aplikace typu *single page*. Obě aplikace poskytují shodné výsledky (neboť samozřejmě využívají stejný klasifikátor), přičemž úspěšnost klasifikace se pohybuje nad 80 %, což lze pro tento typ úlohy považovat za rozumný, běžný a očekávatelný výsledek.

Odevzdaná mobilní aplikace prošla poměrně komplikovaným vývojem, během něhož autor vyzkoušel několik různých vývojových prostředků (Embarcadero RAD Studio, Laz4Android, Python a Android Studio). Výsledkem je dobře vypadající a stabilně fungující aplikace pro OS Android, která bez zbytečných okras spolehlivě provádí to, co má, tedy dovoluje pořídít fotoaparát mobilního zařízení (nebo nahrát z úložiště) snímek, který následně klasifikuje a uživateli zobrazí tři nejpravděpodobnější stavební slohy, ve kterých by vyfocená budova mohla být postavena. Výhodou je, že používá předtrénovanou CDNN jako klasifikátor přímo na mobilním zařízení, což sice mírně zvyšuje nároky na jeho výkon (ale nikoli nad dnes již naprosto běžnou úroveň), ale zase není nutná datová konektivita.

Sesterská webová single-page aplikace funguje také správně, je stabilní a jednoduše ovladatelná. Dovoluje pouze upload snímku (jiný způsob nemá u generické webové aplikace ani smysl) v některém z běžných grafických formátů, a ten následně klasifikuje a zobrazí opět tři nejpravděpodobnější stavební slohy, ve kterých by budova zachycená na snímku mohla být postavena. Design webové aplikace je bohužel viditelně slabší než je tomu v případě mobilní verze, nicméně funkčnost je stejná.

Obě aplikace jsou naprogramované minimalisticky za použití příslušných aplikačních frameworků, takže vlastního zdrojového kódu zapsaného autorem práce je celkem málo. Mobilní aplikace napsaná v Javě v Android Studiu sestává z 543 řádek kódu a webová aplikace napsaná v Javě za využití frameworku Spring má celkem 312 řádek kódu. Kvalita kódu je na běžné úrovni, neboť je zřejmě více ovlivněna použitými vývojovými prostředky než přímo autorem zdrojového textu.

Sada skriptů v jazyce Python pak slouží k přípravě trénovacích dat (jejich pořízení crawlingem a následná augmentace) a k natrénování CDNN klasifikátoru a má spíše charakter prototypu či pomocného kódu – nepředpokládá se, že by s tímto kódem pracovali koncoví uživatelé aplikace, takže lze tolerovat jeho více pracovní či experimentální charakter. Jde o celkem 6324 řádek kódu v jazyce Python, kvalita zápisu a čitelnost kódu je kolísavá. Kód není prakticky vůbec komentován. Struktura (a třeba i pojmenování identifikátorů) odpovídá faktu, že se jedná o mnohokrát přepracovávaný experimentální kód.

K zamýšlenému účelu je výsledný produkt bez námitek použitelný. Spolehlivost klasifikace snímků budov do jednotlivých tříd podle stavebních slohů se pohybuje mezi 80 % a zhruba 88 %, což je hodnota očekávatelná při použití klasifikační technice a naprosto přijatelná pro nasazení aplikací v rámci zábavně-vzdělávacích a volnočasových, tedy neprofesionálních, aktivit případných uživatelů.

zásadními technickými chybami. Nicméně i tak lze místy nalézt některé drobné odchylky od normy či zvyklostí sazby (např. kolize čísel podkapitol příloh s názvy v obsahu, chybějící 1/6em mezery před znakem '%', užívání písmene 'x' namísto znaku 'x', sazba procent v tab. 7.1, aj.), které ale nenarušují zásadním způsobem čtenářský zážitek. Asi největším technickým nedostatkem je nízké rozlišení řady rastrových obrázků (např. obr. 2.3, obr. 3.2 či tab. 4.1).

Větším problémem jsou autorovy schopnosti literární. Některé, zejména popisné, úseky textu jsou hůře srozumitelné a obtížně čitelné. Autor měl místy zřejmě skutečné těžkosti vyjádřit v přijatelné technické češtině své úvahy, a čtenář pak tok jeho myšlenek sledovat. Nicméně je naprosto evidentní, že se upřímně snažil sepsat dílo co nejlépe. Velice oceňuji snahu autora všechna svá strategická rozhodnutí (jako je výběr knihoven, vývojových prostředků, klasifikačních tříd, apod.) v textu dobře odůvodnit.

Občas autor používá ne zcela běžné názvosloví (např. na str. 12 „diskrétní maticová konvoluce“) či mírně matoucí výroky, také některé poznámky pod čarou pochopení textu spíše komplikují, než že by mu napomáhaly. Nicméně po obsahové stránce je dokument celkem dobře vyvážen a pokrývá v rozumné míře jak teoretická prerekvizitní témata, tak samotný návrh a implementaci aplikací a sady trénovacích skriptů.

Autor v práci cituje celkem 46 zdrojů, které jsou k předmětnému tématu relevantní. V textu jsou citace provedeny správně. V seznamu zdrojů jsou některé zdroje uvedeny dle mého názoru v ne zcela úplné formě, ovšem týká se to výhradně webových stránek, u kterých patrně skutečně nebylo možné dohledat původce textu či datum prvního zveřejnění. Autorovi tedy nelze v tomto ohledu nic vytknout.

Zadání bylo bez výhrad splněno ve všech bodech. Výsledkem práce jsou dvě funkční aplikace (pro mobilní zařízení a webová) a sada trénovacích a přípravných skriptů. Software dobře slouží zamýšlenému účelu a funguje bez závad s úspěšností klasifikace vhodnou pro neprofesionální zábavně-vzdělávací použití. K úrovni průvodního dokumentu mám sice řadu připomínek, nicméně autorem odvedenou práci popisuje dostatečně a nedostatky nejsou závažné. Práci proto **doporučuji k obhajobě** a hodnotím klasifikačním stupněm

„velmi dobře“.

Ing. Kamil Ekštejn, Ph.D.
KIV FAV ZČU

V Plzni dne 18. května 2022

Doplňující otázky:

1. Proč jste při vyhodnocování vhodnosti jednotlivých knihoven opomenul vedoucím navrhovaný Dlib?