

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Lukáš Chaloupský	
Název práce	Automatické generování obrázků a jejich využití jako trénovacích dat	
Rok odevzdání	2020	
Studijní program	Informatika	
Studijní obor	Obecná informatika	
Autor posudku	Mgr. Rudolf Rosa, Ph.D.	Vedoucí
Pracoviště	Ústav formální a aplikované lingvistiky	

K celé práci

lepší OK horší nevyhovuje

Obtížnost zadání	X			
Splnění zadání		X		
Rozsah práce <small>... textová i implementační část, zohlednění náročnosti</small>		X		

Problém automatického generování obrázků dle textového popisu je problémem poměrně obtížným, i nejlepší aktuálně dostupné systémy nepodávají zcela uspokojivé výsledky. Jedná se tedy o náročné téma, přičemž návržení a realizace řešení, které by fungovalo opravdu dobře, by odpovídalo spíše dizertační práci. V tomto kontextu je tedy nutné odevzdanou práci nahlížet. Obrázky, které vytvořená aplikace generuje, obvykle technicky vzato odpovídají vstupnímu popisu, spíše než o generování souvislých obrázků se ale jedná o jakési strukturované koláže.

Jeden z velmi obtížných aspektů autorova pojetí problému je snaha o zcela obecné řešení, které není omezené jen na obrázky konkrétního typu, které by zahrnovaly jen konkrétní objekty či omezovaly jazyk použitý pro popis obrázku. Součástí úlohy je tak potřeba porozumění významu obecného textu, což je samo o sobě těžká a dosud nevyřešená úloha. Při sestavování samotného obrázku se pak naráží na nejeznoznačnost jazyka i na nedostupnost znalostních databází potřebných pro správné sestavení koláže; těžkým problémem například je neznalost obvyklých relativních velikostí jednotlivých objektů, potřebná pro správné naškálování jednotlivých objektů (aneb je těžké určit, že např. myš má být 10x menší než pes a pes má být 5x menší než kůň).

V rámci možností se domnívám, že autor se s problémem vypořádal velmi dobře, kdy výsledné obrázky sice nejsou obzvlášť krásné či dokonalé, avšak řešení je skutečně do značné míry univerzální – takřka pro jakýkoliv textový vstup aplikace vygeneruje obrázek, který do nějaké míry odpovídá vstupnímu popisku. Tímto aspektem řešení vyniká nad jiná existující řešení daného problému, která jsou typicky omezena použitými trénovacími daty, za jejichž hranici nejsou schopna zobecňovat.

(pokračování na další straně)

Jakkoliv samotné výstupy aplikace nejsou zcela ohromující, zajímavé a potenciálně užitečné je jejich využití pro syntézu trénovacích dat pro image captioning, tedy úlohu automatického generování popisku obrázku. Tímto způsobem je možné uměle zvětšit trénovací data pro tuto úlohu a tím zvýšit její úspěšnost, jak ukazují provedené a důsledně vyhodnocené experimenty. Ještě zajímavější potenciál pak spatřuji v možnosti syntetizovat trénovací data pro doménu, pro niž ručně vytvořená data neexistují – image captioning je jako klasická úloha řešená strojovým učením omezený na typ obrázků vyskytujících se v trénovacích datech (běžně například lidi, domy a dopravní prostředky), zatímco nástroj vytvořený v rámci této bakalářské práce umožňuje syntetizovat obrázky i jiného typu, a tím rozšířit typy obrázků, které image captioning pokrývá (například o přírodní scény zahrnující interakce různých zvířat).

Textová část práce

lepší OK horší nevyhovuje

Formální úprava <i>... jazyková úroveň, typografická úroveň, citace</i>	X			
Struktura textu <i>... kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	X			
Analýza	X			
Vývojová dokumentace		X		
Uživatelská dokumentace		X		

Text práce je nadprůměrně kvalitní, vhodně strukturovaný, typograficky dobře zpracovaný, přehledný a s odpovídající mírou detailu.

Práce obsahuje rešerši souvisejících prací, pro niž autor nastudoval několik odborných článků v anglickém jazyce, a seznámil se i s obtížným konceptem Generative adversarial networks, které se pro řešení problému generování obrázků běžně používají (byť autor nakonec v práci z mnoha dobrých důvodů volí jiný postup). Autor důsledně a zodpovědně porovnává jím zvolený postup s postupy jiných autorů, přičemž vyváženě uvádí výhody i nevýhody jak těchto postupů, tak svého postupu.

Na práci oceňuji velké množství příkladů, diagramů a ukázek, díky nimž je text dobře srozumitelný a názorný.

Implementační část práce

lepší OK horší nevyhovuje

Kvalita návrhu <i>... architektura, struktury a algoritmy, použité technologie</i>	X			
Kvalita zpracování <i>... jmenné konvence, formátování, komentáře, testování</i>	X			
Stabilita implementace		X		

Autor používá vhodné programátorské postupy a techniky, včetně šikovného zapojení REST API jak pro získávání obrázků z Flickru, tak pro zapojení image captioningu. Aplikace je napsána multiplatformně v jazyce C#. Kód je velmi přehledně strukturován, zdrojové kódy jsou velmi úhledné, pokryté dokumentačními komentáři, orientace v nich je velmi snadná. Důležité komponenty jsou pokryty unit testy.

Celkové hodnocení Výborně

Práci navrhuji na zvláštní ocenění Ne

Datum

Podpis