

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. Jindřich Helcl	Oponent
Pracoviště	Ústav formální a aplikované lingvistiky	

K celé práci

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Obtížnost zadání	X	X		
Splnění zadání	X			
Rozsah práce <i>... textová i implementační část, zohlednění náročnosti</i>	X			
<p>Práce se zabývá pravidlovým generováním obrázků (koláží) k uživatelem zadanému popisku. Jde o pěkný nástroj, který v daném textu rozpozná vztahy mezi elementy, stáhne pro každý element odpovídající obrázek z internetu a obrázky poskládá do výsledné koláže s využitím informací o vzájemných pozicích.</p> <p>Podobně jako ve strojovém překladu, kde nejlepší modely stojí na využití syntetických dat, i zde autor provádí řadu experimentů, ve kterých využívá obrázky vygenerované svým programem jako syntetická data pro úlohu image captioningu.</p> <p>Rozsah práce hodnotím velmi pozitivně. Provedené experimenty přináší práci výzkumnou přidanou hodnotu. Vytyčené cíle se autorovi práce podařilo splnit.</p>				

Textová část práce

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Formální úprava <i>... jazyková úroveň, typografická úroveň, citace</i>		X	X	
Struktura textu <i>... kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>		X		
Analýza	X	X		
Vývojová dokumentace		X		
Uživatelská dokumentace		X		
<p>Práce je v českém jazyce, je psaná srozumitelně, strukturovaná je logicky a přehledně. Autor dostatečně zdůvodňuje rozhodnutí, která v rámci práce činil. Hlavním (ne však zásadním) nedostatkem je jazyková úroveň. Práci by více slušela angličtina; autor nepřekládá nejen zaběhlé anglické pojmy, ale často i ty, jejichž překlad do češtiny je relativně běžný („encoder“ -> enkodér, „offset“, „Levenshtein“ -> Levenštejn). Často se taky objevují nepřirozené přívlastkové konstrukce („Winforms knihovna“) a jiné jazykové nešváry („nejvíce optimální“). Rovnice pro Levenštejnovu vzdálenost by si zasloužila vysázet lépe (písmenka s diakritikou jsou psaná odlišným fontem, celé by to mělo být vysázené jako <code>\text</code>).</p>				

Implementační část práce

lepší OK horší nevyhovuje

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

Z uživatelského pohledu jde o velmi jednoduchý a stabilní program s přehledným uživatelským rozhraním. Asi by bylo lepší upravit chování obrázkové mezipaměti, protože pokud si napřed stáhnu nějaký obrázek bez image captioningu a pak pro ten stejný popisek pustím generování s image captioningem, tak se použijí uložené obrázky, místo toho aby se stáhly lepší.

Co se týče pravidlového generování, byla učiněna řada omezení, kterými si autor možná ubližuje. Ignorovat jeden z elementů u „or“ nebo „but“ může někdy pomoci, ale existují i příklady, kdy to výsledný obrázek zpřesní. Taky mi není jasné, jak se ignorují negované elementy. Program tak například nedokáže správně pochopit popisky jako „Sky without clouds“ nebo „A clown but without makeup“. Tyto problémy však pramení z rozhodnutí nevěnovat se vztahům ve větě, které mění nebo upřesňují významy jednotlivých elementů, což by vyžadovalo mnohem náročnější způsob zpracování.

V předzpracování bude zřejmě bug, protože věta „bear with two cubs“ vede k vygenerování obrázku bez medvěda, kdežto věta „a bear with two cubs“ už medvěda zobrazí. Na příkladu předzpracování mě zaujalo převedení idiomu „side by side“ na text „the side by the side“ – pomáhá to, škodí to, anebo je to jedno?

Obzvlášť pozitivně hodnotím přítomnost experimentální části, bez které by to na bakalářskou práci stačilo také. K experimentům mám několik připomínek: Používat BLEU-3 jako hlavní metriku je zvláštní. A jelikož rozdíl v BLEU v těchto hodnotách stejně o ničem nevypovídá, bylo by lepší držet se standardního BLEU-4. Zajímalo by mě, jak by si modely vedly v porovnání s referenčními popisky. Tahle evaluace v práci chybí.

Celkové hodnocení Výborně

Práci navrhuji na zvláštní ocenění Ne

Datum

Podpis