
Posudek oponenta bakalářské práce

Marluce Quaresma Porozumění vizuálním dokumentům

Obsah práce

Cílem předložené bakalářské práce je otestovat a analyzovat end-to-end modely pro Visual Document Understanding (VDU). Oceňuji, že se studentka společně s vedoucím zaměřili na nejmodernější přístupy k řešení této úlohy (popsáno ve druhé kapitole). Kapitola 3 se věnuje datovým sadám relevantní pro tuto práci. Teoretickou část práce hodnotím pozitivně a žádné zásadní výhrady k ní nemám. Praktickou část práce včetně prezentace výsledků popisují kapitoly 4 a 5. Zde mám naopak řadu výhrad (viz níže).

Kvalita řešení a dosažených výsledků

Většinu dodaných zdrojových kódů jsem spustil. Jako uživatel dodaného SW bych očekával nicméně **výrazně lepší způsob distribuce SW**. Uživatel si musí sám pro každý model vytvořit prostředí (*environment*). Uvítal bych řešení prostřednictvím nějakého kontejneru (např. *Docker*) či balíčku, kde už budou vyřešené závislosti a knihovny ve správné verzi.

Uživatelská dokumentace je pro mě lehce nesrozumitelná a příliš mi nepomohla k otestování dodaného SW. Místo jednoho spouštěcího skriptu, kde argumentem určím požadovaný model, musím složitě zjišťovat, jak se který model používá, co má za rozhraní a co očekává za argumenty. Toto je zcela určitě největší výtka, kterou k práci mám.

Domnívám se, že studentka ve velké míře využila veřejné implementace modelů z dostupných repozitářů a spustila již vyvinutý zdrojový kód na různých datasetech. Netroufám si tedy hodnotit úroveň a formu zápisu zdrojových kódů, protože není zřejmé, jaké zdrojové kódy pocházejí od přímo od autorky práce.

Vzhledem k tomu, že práce se věnuje porozumění vizuálním dokumentům, očekával bych v dokumentu práce i nějaké vizualizace natrénovaných modelů, které podpoří celkové dosažené výsledky a budou součástí analýzy a zhodnocení výsledků.

Formální úroveň

Práce obsahuje celkem 47 stran a je psána anglicky. Obsahuje překlepy v malé míře a některé nesoulady např. v prohlášení „V Plzni, on 1 May 2024“. Mám k formální úrovni následující připomínky a výhrady: chybí číslování rovnic / vzorečků (str. 21), obrázky 2.1, 2.3 příliš malé, poměrně nepřehledná tabulka 3.1., horší kvalita obrázků 4.3, 5.1, 5.2, 5.3 – použití rastrových obrázků pro grafy.

Další zásadní výhradu vidím v prezentaci výsledků. Použití nekvalitních rastrových obrázků pro prezentaci výsledků je nevhodné, obzvlášť když existují způsoby jak vytvořit grafy ve vektorové grafice případně použít k prezentaci tabulky.

Nepochopil jsem použití metriky *accuracy* – v kapitole 4 není vysvětlena. Naopak F1 score, které je vysvětleno i pomocí vzorečků ve výsledcích není. Dále je nevhodné dle mého názoru používat metriku Levenshteinovy vzdálenosti, když jednotlivé datasety obsahují různě dlouhé texty. Je evidentní, že

model udělá nominálně více chyb v delších textech než v textech velice krátkých. K evaluaci rozpoznávaného textu bych spíše použil chybovost na úrovni znaků či slov (a tedy rozsah 0 – 1, což pak lze vyjádřit v procentech).

V neposlední řadě bych určitě použil jinou formulaci při zhodnocení výsledků modelu Doc2graph: „*The average accuracy was approximately 60.7%, which is neither classied as good nor bad performance.*“

Reference a práce s literaturou

V poměrně velké míře se mezi referencemi objevují citace z portálu arXiv ačkoliv u řady z nich (např. [1], [6] či [15]) lze dohledat oficiálně vydané (recenzované) verze ve sbornících nebo časopisech. Osobně bych upřednostnil tyto citace. Jiné výhrady k práci s literaturou nemám.

Splnění zadání

Ze zadání a zásad pro vypracování vyplývá, že úkolem je provést rešerši, otestování a analýzu výsledků. Toto bylo dle mého názoru naplněno a zadání považuji za splněné.

Závěr a hodnocení

Práci i přes všechny výhrady hodnotím známkou „dobře“ a **doporučuji ji k obhajobě**.

Dotazy k práci

1. Dokážete změřit nebo aspoň odhadnout (např. procentuálně) množství kódu/zdrojových souborů, které jste sama vytvořila?
2. Vysvětlíte prosím metriku *accuracy* v prezentaci výsledků. Co touto metrikou měříte?

V Plzni dne 22. května 2024

Ing. Jiří Martínek, Ph.D.
(oponent BP)