

Hodnocení vedoucího bakalářské práce

Autor/autorka práce: **Marluce Will Pires Dos Santos Quaresma**

Název práce: **Porozumění vizuálním dokumentům**

Bakalářská práce se zabývá úlohou porozumění vizuálním dokumentům (Visual Document Understanding – VDU). Úkolem VDU je zpracovat obrazové dokumenty a získat z nich relevantní informace a ty dále vhodnou formou prezentovat. VDU má několik dílčích úloh, jako např. získávání klíčových informací, klasifikace typu dokumentu atd. Zmíněné úlohy jsou většinou řešeny ve dvou krocích. Nejprve se provede převod obrazového dokumentu na text pomocí OCR metod a následně se textová forma využije pro řešení příslušné dílčí úlohy. V poslední době vznikají tzv. end-to-end metody, které úlohu řeší v jednom kroku. Obrazový dokument je tedy přímo převeden např. na soubor typu json, který obsahuje získané klíčové informace.

Cílem práce bylo prostudovat dostupné metody, vybrat alespoň dvě implementace a ty otestovat a porovnat na vhodných datech.

Nejprve byla provedena rešerše metod, která obsahuje popis čtyř metod pro porozumění vizuálním dokumentům. V sekci Datasets byly shrnuty dostupné datasety pro úlohu VDU a také dataset nomenklatur, který byl dodán vedoucím práce pro otestování na nestandardních datech. Dále jsou popsány metriky používané pro dílčí VDU úlohy. Sekce experiments pak přináší porovnání výsledků dvou zvolených metod na vybraných datasetech. Pro dataset nomenklatur bylo ukázáno, že end-to-end metoda Donut dosahuje výborných výsledků a výrazně překonává dříve publikované výsledky.

Implementační část byla napsána v Pythonu a je založena zejména na dostupných implementacích dvou použitých metod a jejich modifikaci pro účely otestování na jiných datových sadách. Rozsah programátorských prací je spíše menší, což je dáno zejména prací se složitými neuronovými sítěmi, kde se nepředpokládá jejich výraznější modifikace.

Práce je psána v angličtině. Text je přehledný a rozumně členěný. Použité zdroje jsou odpovídajícím způsobem citovány.

Přístup k řešení byl aktivní a postup byl pravidelně konzultován s vedoucím. Při řešení některých dílčích úkolů, zejména při zprovoznování trénování neuronových sítí, by byla vhodná větší samostatnost, ale vzhledem k poměrně náročné práci a minimem zkušeností s využitím frameworků pro neuronové sítě hodnotím postup práce pozitivně.

Zadání bylo splněno bez výhrad

Navrhuji hodnocení známkou **výborně** a práci doporučuji k obhajobě.

V Plzni 24. 5. 2024

Ing. Ladislav Lenc, Ph.D.