

Review of Master's Thesis

Student: Piwowarski Lukáš, Bc.
Title: Neural Networks for Automatic Table Recognition (id 24864)
Reviewer: Španěl Michal, Ing., Ph.D., DCGM FIT BUT

- 1. Assignment complexity** **considerably demanding assignment**

Téma má výzkumný charakter. Řešitel se zaměřil na obtížnou úlohu automatického rozpoznání tabulek v ručně psaných historických dokumentech, která vyžaduje pečlivé studium existujících přístupů publikovaných ve vědeckých článcích a současně dobré pochopení problematiky konvolučních a grafových sítí, protože modely pro rozpoznávání tabulek nejsou zrovna triviální.
- 2. Completeness of assignment requirements** **assignment fulfilled**

Všechny požadované body zadání řešitel splnil. Navrhnul a realizoval vlastní metodu pro rozpoznání ručně psaných tabulek ve skenech historických dokumentů. Řešitel také navrhuje novou metriku pro porovnání úspěšnosti rozpoznávání, která kvantifikuje výsledky lépe než v současnosti používané metriky.
- 3. Length of technical report** **in usual extent**

Rozsah technické zprávy se pohybuje na spodní hranici obvyklého rozmezí. Text však neobsahuje zbytečné kapitoly s převzatou teorií obvyklé pro jiné práce. Autor se věnuje tématu a věcem potřebným pro pochopení navrhovaných metod.
- 4. Presentation level of technical report** **85 p. (B)**

Technická zpráva je přehledná a dobře pochopitelná. Snad jen kapitoly 3 a 4 se zvláště prolnají a mohly být strukturovány jinak.

 - Pro kompletní představu o řešení bych uvítal detailnější informaci o trénování CNN pro extrakci vizuálních příznaků. Jde o důležitou část metody a pouhý odkaz do literatury mi tu nepřipadá vhodný.
 - Kromě číselného vyhodnocení bych ocenil i více ukázek konkrétních výstupů sítí a jejich vyhodnocení různými metrikami. Jak pro ilustraci výsledků, tak pro porovnání metrik.
- 5. Formal aspects of technical report** **85 p. (B)**

Technická zpráva je sepsána dobrou angličtinou. K typografické a jazykové stránce nemám zásadních připomínek. Pouze bych doporučoval více využívat trpný rod namísto první osoby množného čísla. a místo "we fit the graph's nodes" psát "the graph's nodes are fitted." Občas se v textu objevují "osiřelé" citace mimo větu až za tečkou.
- 6. Literature usage** **95 p. (A)**

Řešitel prostudoval velké množství vědeckých publikací a na základě informací načtených z článků realizoval vlastní metodu, která je kombinací publikovaných architektur a vlastních postupů (např. závěrečný krok zpracování výstupu grafové sítě).
Řešitelův přehled existujících metod je pěkný. Trochu lépe mohl shrnout výhody a nevýhody jednotlivých architektur (např. proč jsou sítě typu transformer výhodné proti jiným, nebo jaké použití mají různé typy GNN vrstev).
- 7. Implementation results** **90 p. (A)**

Programové řešení tvoří rozsáhlé Python skripty jak pro anotaci datových sad a práci s nimi, tak pro trénování a implementaci dílčích kroků algoritmu pro rozpoznávání. Nešlo jen o pouhé spouštění standardních modelů. Kódy jsou dobře strukturované, pečlivě komentované a nechybí ani komentáře ohledně instalace potřebných knihoven a kořenový skript pro spouštění jednotlivých částí.
Řešitel vyhodnotil celkovou úspěšnost své metody pomocí standardní i vlastní metriky. Provedl porovnání s jinými metodami na veřejné datové sadě. Výsledky jsou povzbudivé.
Navrhovanou metodu tvoří sekvence navazujících kroků, je škoda, že se nepokusil vyhodnotit jaký vliv na výsledek mají tyto dílčí kroky - kvalita OCR, Node Discovery algoritmus, zpracování výstupu GNN, atd.
- 8. Utilizability of results**

Výsledky pravděpodobně najdou uplatnění v rámci výzkumných projektů na FIT a mají i publikační potenciál.
- 9. Questions for defence**
 - Jak byly trénovány CNN pro extrakci vizuálních příznaků pro buňky a hrany? Využíváte nějak informaci o klasifikaci buněk z trénovacích dat?
 - Výstupy OCR využíváte jako vstup Vaší metody i pro tvorbu ground truth anotací. Jak kvalitní jsou výstupy

OCR a jaký mají dopad na přesnost rozpoznání tabulek? Zvládá kupříkladu vertikální text? Jak si poradí s textem přetékájícím do sousedních buněk?

- Dle tabulky 6.2 se zdá, že detekce strukturovaných záhlaví je problematická a síť buňky záhlaví klasifikuje jako datové. Zvažoval jste řešení pro tento problém?
- Problematická je i detekce prázdných buňek. Jsou prázdné buňky významné pro další zpracování rozpoznávaných tabulek? Jak by vypadaly výsledky Vaší metody, pokud byste prázdné buňky při výpočtu metrik vyloučil?

10. Total assessment

93 p. excellent (A)

Pan Piwowarski je schopný výzkumník. Výborně se vypořádal s obtížným tématem. Dokázal se dobře orientovat v problematice a navrhnout vlastní smysluplný postup pro automatické rozpoznání tabulek, který nemá vůbec špatné výsledky. Pečlivě zpracoval i programové řešení a technickou zprávu, kterou sepsal v anglickém jazyce.

In Brno 5 June 2022

Španěl Michal, Ing., Ph.D.
reviewer