

Review of Bachelor's Thesis

Student: Kocurek Pavel
Title: Visual Question Answering (id 22598)
Reviewer: Ondřej Karel, Ing., DCGM FIT BUT

1. **Assignment complexity** **average assignment**
2. **Completeness of assignment requirements** **assignment fulfilled with enhancements**
Nad rámec zadání bylo provedeno hodnocení aplikace s komunitou zrakově postižených, vytvoření datové sady a nad ní provedená analýza.
3. **Length of technical report** **in usual extent**
Práce je v rozsahu 51 vysázených stran s větším počtem obrázků vhodně doplňujících text.
4. **Presentation level of technical report** **85 p. (B)**
Práce je čtivá a jednotlivé části textu na sebe logicky navazují. Vytkl bych pouze:
 - Struktura první kapitoly zastřešuje celkem 4 témata z různých oblastí, přičemž sítě Transformers jsou v této kapitole chybně uvedeny jako podkapitola rekurentních sítí.
 - Kapitola 3.2 místo zvolené architektury Pythia popisuje architekturu nazvanou up-down. Pythia z této architektury vychází, ale není stejná.
 - Nekonzistentní značení (pravděpodobně převzaté z původních článků) a jeho nedostatečné vysvětlení komplikuje pochopení modelů. Např. různé značení množiny všech reálných čísel na stranách 27-28, definování "tanh layer" na straně 27 a dále v textu vystupují další dvě vrstvy pojmenované "non-linear layer" bez jasného propojení s touto definicí.
5. **Formal aspects of technical report** **85 p. (B)**
Práce obsahuje malý počet pravopisných a typografických chyb. Student několikrát začíná větu malým písmenem.
6. **Literature usage** **95 p. (A)**
Student pracuje s adekvátní literaturou vzhledem k tématu práce. Reference obsahují několik odkazů na github.
7. **Implementation results** **90 p. (A)**
Výsledky práce je prototyp aplikace, její vyhodnocení potencionálními uživateli z komunity zrakově postižených a datová sada čítající 111 anotovaných obrázků. Samotné zdrojové soubory jsou nedostatečně komentované a obsahují větší počet absolutních cest.
8. **Utilizability of results**
Na základě této práce může vzniknout unikátní mobilní aplikace, nebo rozšíření do již existujících aplikací, která by usnadňovala každodenní činnosti uživatelům se zrakovým postižením.
9. **Questions for defence**
 1. V čem se liší architektura sítě Pythia od architektury up-down?
 2. Srovnajte síť Pythia se sítí OSCAR.
10. **Total assessment** **85 p. very good (B)**
Práce je na nadprůměrné úrovni včetně podstatných rozšíření a zaměřuje se na vyhodnocení použitelnosti VQA v praxi. Nicméně není jasné zda student pochopil princip sítě Pythia, která je jádrem implementovaného systému, proto navrhuji hodnotit práci stupněm B.

In Brno 2 June 2021

Ondřej Karel, Ing.
reviewer