

Posudek oponenta diplomové práce

Autor/autorka práce: **Josef Baloun**

Název práce: Segmentace stran rukopisných dokumentů

Obsah práce

Cílem práce je navrhnout a implementovat software pro segmentaci rukopisných obrazových dokumentů. Zpracovávají jsou zejména kroniky obsažené v portálu Porta fontium. Předmětem segmentace je zde oddělení textových bloků, obrázků a pozadí.

V práci je uveden přehled metod používaných pro segmentaci se zaměřením zejména na konvoluční neuronové sítě. V kapitole „Data“ je představen portál Porta fontium a také dostupné datové sady, které je možné pro úlohu segmentace použít. V další kapitole je popsáno vytvoření vlastní datové sady z dat z Porta fontium. Následuje popis architektury neuronové sítě, která byla v práci použita a postup předzpracování vstupních obrázků.

V rozsáhlé experimentální části jsou srovnány výsledky s použitím rozšíření trénovací datové sady (data augmentation), s využitím syntetických dat, vytvořených vložením obrázků stran s textem, a také rozšířením trénovací množiny o tištěná data. Zajímavým přínosem je navržení váhované chybové funkce, která by měla umožnit lepší oddělení textových bloků (npř. oddělení dvou sousedních odstavců textu). Na základě přínosu jednotlivých technik byl vytvořen finální model, který využívá některé z vše uvedených technik. U váhované chybové funkce je těžké zhodnotit přínos jen na základě poskytnutých metrik. Bylo by zde vhodné využít evaluaci na základě výsledných regionů, která zohledňuje i chybné spojení nebo rozdělení jednotlivých bloků. Je zde předpoklad, že taková evaluace by ukázala významnější přínos této techniky.

Všechny experimenty byly provedeny s nově vytvořenou datovou sadou. Je tedy těžké porovnat dosažené výsledky s literaturou. Zajímavé by zde bylo srovnání prezentované metody s ostatními na některé ze standardních datových sad, např. RDCL.

Rovněž by bylo vhodné prezentovat výsledky dosažené s originální architekturou sítě U-Net.

Spuštění SW z příloženého DVD je bezproblémové dobře popsáné v uživatelské příručce.

Text práce je přehledný a je psán vhodnou formou. V přehledu neuronových sítí bych volil jiný název pro architekturu FCN [Wick a Puppe], protože i síť U-Net patří do skupiny plně konvolučních sítí.

Práce s literaturou je na vysoké úrovni a použité prameny jsou v práci dostatečně citovány.

Zadání bylo splněno bez výhrad.

Dotazy k práci

Zkoušel jste porovnat výsledky originální architektury U-Net s výsledky prezentované upravené architektury?

Navrhuji hodnocení známkou **v ý b o r n ě** a práci doporučuji k obhajobě.

V Plzni 8.6.2020

Ing. Ladislav Lenc, Ph.D.