

Hodnocení vedoucího práce

Jméno studenta: Milan Troller

Název práce: Practical OCR system based on state of art neural networks

Návrh klasifikace: A (výborně)

Vedoucí práce: Mgr. Petr Baudiš

Student ve své bakalářské práci zpracovává problematiku OCR systémů, přičemž se zaměřil na prozkoumání a evaluaci OCR konkrétního jednořádkového výřezu stránky v kontextu systému na extrakci informací z faktur.

Dlužno zdůraznit, že v průběhu práce na projektu bylo nutné posunout původní zadání - nejdříve jsme chtěli rozšířit otevřený OCR systém Tesseract o metody založené na moderních neuronových sítích a hlubokém učení, začátkem roku 2017 ovšem vyšla nová verze Tesseractu, která přesně toto již implementovala. Aby výsledný projekt byl stále užitečný, zaměřil se student proto na speciálnější úlohu mimo kontext systému Tesseract. Vycházel z dříve navržené architektury neuronové sítě a implementačního základu (příloha frameworku Keras). Těžiště práce tak bylo v důkladné rešerši a aplikaci neuronové sítě na konkrétní realistický dataset, tedy vizuálnímu posunu syntetických trénovacích dat co nejbližší reálným scanům stránek.

Hlavním výstupem práce je systém, kterým se podařilo výrazně překonat předchozí výsledky dosažené Tesseractem. Nutno je též ocenit, že je systém k dispozici pro další použití jako open source na serveru GitHub. širší čtenářstvo si může najít díky zpracování v anglickém jazyce i samotný text práce.

Student průběžně konzultoval vývoj projektu, zároveň projevoval iniciativu i samostatnost v celém průběhu řešení práce, sám navrhoval cesty k jejímu zdárnému dokončení. Zatímco softwarová stránka byla dotažena ve vhodném předstihu, většinu textu student dopisoval na poslední chvíli, což je místy znát. Text práce je mírně nestandardně strukturován co se týče jednotlivých kapitol a stejnými kapitolami se oproti zvyklostem prolíná výchozí práce a studentova nadstavba (ovšem svůj přínos student vždy explicitně označuje). Naopak nutno vyzdvihnout množství obrázků a vizuálně dobře zpracované diagramy.

Práci navrhuji k úspěšnému obhájení známkou výborně.

Petr Baudiš
v Praze 14. 6. 2017