

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 alfa 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

Posudek oponenta bakalářské práce

Autor: **Milan Troller**

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

Posudek vypracoval oponent práce Ing. Vít Listík

Tématem práce je vytvoření OCR systému založeného na neuronových sítích. Tento postup je v posledních letech velmi často používaný.

Práce shrnuje postupný vývoj OCR systémů a končí u systému Tesseract. Ten je v práci uveden jako referenční řešení. Dále se práce věnuje ukázce zveřejněné v knihovně Keras, jejímu popisu a možným vylepšením. Velkým těžištěm práce je také generování vlastních dat, které je podrobně popsáno před zhodnocením implementovaného řešení.

Práce je formálně a jazykově v pořádku až na drobné výjimky jako například: Chybějící nadpis kapitoly 4, Chybějící citace grafů [3.3, 3.6, 3.7], Stejně grafy 3.6 a 3.7 a také občasné překlepy. Práce má logickou strukturu a jednotlivé kapitoly srozumitelně navazují.

Student správně cituje všechny použité zdroje. Většina citací je z internetových zdrojů.

Těžiště práce je ve dvou bodech. Prvně se student seznamoval s knihovnou Keras a teorií spojenou s ukázkovým řešením, které následně upravil pro potřeby definované v práci s účelem zlepšení výkonnosti systému. Druhý bod se věnuje problematice generování dat, které se snaží co nejvíce podobat reálným datům, zde student přidává i vlastní způsoby úprav generovaných dat. Implementační část odpovídá rozsahu bakalářské práce.

Ve zhodnocení práce bych očekával podrobnější analýzu jednotlivých úprav systému (přidání generovaných dat, doplnění informace o regionu výřezu).

Zadání mluví o praktickém open-source systému pro OCR založeném na neuronových sítích. Práce se od počátku věnuje podúloze, tedy OCR pro čtení neslovníkových dat (částky, datumy, bankovní spojení). Výsledná implementace je samozřejmě přiložena k práci, ale o publikování zdrojových kódů, které je zmíněno v zadání nikde nemluví (možná proto, že základ práce již je zveřejněn v knihovně Keras). To je ale pouze malý nedostatek, vzhledem k dosaženému výsledku práce, tedy zlepšení o 15% proti referenčnímu řešení na interní datové sadě založené na reálných datech. Student tedy zadání splnil a proto doporučuji práci k obhajobě a hodnotím ji stupněm A - výborně.

Otázky:

- Je nějaký důvod nechávat v systému max pooling vrstvy, když je pooling nastaven na 1? [3.3.1]
- Zkoušel jste použít hierarchii struktur nebo například jen nadřazené struktury (například datum)? Je možné, že by taková informace vnášela do systému méně šumu. [3.3.3]
- Jaké je srovnání s původním modelem zveřejněným v knihovně Keras a Vámi upraveným modelem (accuracy)?

V Praze 3.6.2017

Ing. Vít Listík