

Posudek oponenta diplomové práce

Student: Hamerník Pavel, Bc.**Téma:** Využití hlubokého učení pro rozpoznání textu v obrazu grafického uživatelského rozhraní (id 22520)**Oponent:** Špaňhel Jakub, Ing., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Cílem této práce bylo vytvořit neuronovou síť pro rozpoznávání textu v obraze grafického uživatelského rozhraní. Z důvodu nutnosti implementace detekční i klasifikační sítě řadím toto zadání k obtížnějším.
- 2. Splnění požadavků zadání** **zadání splněno**
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
Závěr práce na straně 45.
- 4. Prezentací úroveň předložené práce** **70 b. (C)**
Práce má logickou strukturu, jednotlivé kapitoly na sebe rozumně navazují a je celkem dobře pochopitelná, přestože jsou některé části textu trochu zmatené. Kapitola Experimenty by měla být rozsáhlejší. Aktuálně zabírá pouze tři strany.
- 5. Formální úprava technické zprávy** **60 b. (D)**
 - Práce obsahuje malé množství překlepů a gramatických chyb.
 - Některé větné konstrukce jsou příliš kostrbaté a díky tomu mnohdy nepochopitelné.
 - Text obsahuje velké množství jednopísmenných předložek/spojek na koncích řádků.
 - Stále jsou však v práci i přes předchozí upozornění použity rastrové obrázky v místech, kde mohou být vektorové, což sráží celkový dojem z textové části práce. U některých obrázků není zřejmé, zda se jedná o původní tvorbu či nikoliv.
- 6. Práce s literaturou** **75 b. (C)**
Autor cituje rozumné publikace v oblasti detekce a rozpoznávání textu. Citovaných zdrojů by mohlo být více. Citace jsou však v textu použity korektně.
- 7. Realizační výstup** **70 b. (C)**
Výsledkem práce jsou dva modely konvoluční neuronové sítě - detekční model a model pro rozpoznávání textu založený na rekurentních neuronových sítích. Detekční model funguje dle vyhodnocení celkem obstojně (přesnost 87%). Naopak výsledky rozpoznávacího modelu jsou více jak podprůměrně (přesnost 16%).
- 8. Využitelnost výsledků**
Po dotrénování rozpoznávacího modelu bude možné řešení používat i v praxi. V současném stavu by však při celkové přesnosti/úspěšnosti systému 14% nebylo řešení příliš použité. Samostatný detektor je použitelný již nyní.
- 9. Otázky k obhajobě**
 1. V práci uvádíte, že využíváte při učení sítí "nulování gradientu". V části implementace však hovoříte o "resetování learning rate". O jakou techniku se tedy přesně jedná a jaké výhody/nevýhody při trénování přináší?
 2. V závěru práce tvrdíte, že vytvořit OCR systém, který rozpoznává všechny Unicode znaky není příliš reálné. Přitom konkurenční systémy, se kterými se srovnáváte dosahují obstojných výsledků i na textech s diakritikou. Čím si tento rozdíl vysvětlujete?
- 10. Souhrnné hodnocení** **70 b. dobře (C)**
Autor navrhl a implementoval řešení pro detekci a rozpoznávání textu v grafických uživatelských rozhraních. Současné řešení je na dobré cestě, ale kvůli rozpoznávací části není ještě zcela použitelné. Z důvodů uvedených výše hodnotím práci na rozhraní stupňů **C/D**.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 19. srpna 2019

Špaňhel Jakub, Ing.
oponent