

Posudek práce

předložené na Přírodovědecké fakultě JU

- ☐ posudek vedoucího ☒ posudek oponenta
☐ bakalářské práce ☒ diplomové práce

Autor/ka: **Bc. Miroslav Bartyzal**
Název práce: **Modul do serverové aplikace pro rozpoznávání
identifikačních údajů z osobních dokladů**
Studijní program a obor: **Aplikovaná informatika**
Rok odevzdání: **2018**

Jméno a tituly vedoucího/opponenta: **PhDr. Miloš Prokýšek, Ph.D.**
Pracoviště: **Ústav aplikované informatiky**
Kontaktní e-mail: **prokysek@prf.jcu.cz**

Odborná úroveň práce:

- ☒ vynikající ☐ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Věcné chyby:

- ☒ téměř žádné ☐ vzhledem k rozsahu přiměřený počet ☐ méně podstatné četné ☐ závažné

Výsledky:

- ☒ originální ☐ původní i převzaté ☐ netriviální kompilace ☐ citované z literatury ☐ opsané

Rozsah práce:

- ☒ veliký ☐ standardní ☐ dostatečný ☐ nedostatečný

Grafická, jazyková a formální úroveň:

- ☒ vynikající ☐ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Tiskové chyby:

- ☒ téměř žádné ☐ vzhledem k rozsahu a tématu přiměřený počet ☐ četné

Celková úroveň práce:

- ☒ vynikající ☐ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Slovní vyjádření, komentáře a připomínky vedoucího/oponenta:

Předkládaná práce je velmi kvalitní ve všech klíčových oblastech, tedy struktuře, rešeršní, teoretické i praktické části, včetně evaluace dosažených výsledků. Pokud lze této práci něco vytknout, pak pouze drobné nekonzistence mezi citacemi v textu a v popisech obrázků (např. str. 17), které vznikly zřejmě v důsledku práce s opravdu širokým portfolioem zdrojů a vzhledem k rozsahu práce (116 stran) v některých případech až učebnicový styl, zacházející do velkých podrobností, a to především v úvodním přehledu.

Případné otázky při obhajobě a náměty do diskuze:

Jak hodláte v práci pokračovat a budete její závěry publikovat?

Práci

☒ doporučuji

☐ nedoporučuji

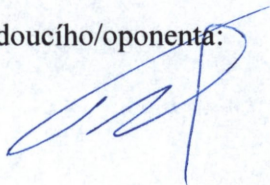
uznat jako diplomovou/bakalářskou.

Navrhuji hodnocení stupněm:

☒ výborně ☐ velmi dobře ☐ dobře ☐ neprospěl/a

Místo, datum a podpis vedoucího/oponenta:

V Českých Budějovicích



Posudek práce

předložené na Přírodovědecké fakultě JU

☐ posudek vedoucího
☐ bakalářské práce

☒ posudek oponenta
☒ diplomové práce

Autor/ka: Bc. Miroslav Bartyzal

Název práce: Modul do serverové aplikace pro rozpoznávání identifikačních údajů z osobních dokladů

Studijní program a obor: Aplikovaná informatika

Rok odevzdání: 2018

Jméno a tituly vedoucího/opponenta: Ing. Jiří Jelínek, CSc.

Pracoviště: ÚAI

Kontaktní e-mail: jjelinek@prf.jcu.cz

Odborná úroveň práce:

☒ vynikající ☐ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Věcné chyby:

☐ téměř žádné ☒ vzhledem k rozsahu přiměřený počet ☐ méně podstatné četné ☐ závažné

Výsledky:

☐ originální ☒ původní i převzaté ☐ netriviální kompilace ☐ citované z literatury ☐ opsané

Rozsah práce:

☐ veliký ☒ standardní ☐ dostatečný ☐ nedostatečný

Grafická, jazyková a formální úroveň:

☐ vynikající ☒ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Tiskové chyby:

☐ téměř žádné ☒ vzhledem k rozsahu a tématu přiměřený počet ☐ četné

Celková úroveň práce:

☒ vynikající ☐ velmi dobrá ☐ průměrná ☐ podprůměrná ☐ nevyhovující

Slovní vyjádření, komentáře a připomínky vedoucího/opponenta:

Tématem práce je vytvoření softwarového řešení pro rozpoznávání údajů z osobních dokladů včetně navržení příslušné zpětné vazby vůči poskytovateli těchto dokladů. Přímo v zadání je z práce vyloučen bezpečnostní aspekt a autor se měl soustředit extrakce dat z obrazů. Z obecného hlediska je možné namítnout, že existují centralizované databáze zaměřené na shromažďování osobních údajů a je tedy otázkou, zda kompletní vyčítání údajů z grafické předlohy je správnou cestou zajišťující dostatečnou spolehlivost a autenticitu, zadání však bylo formulováno jednoznačně.

Členění práce je logické a odpovídá danému tématu, autor postupuje od teorie až po testování konkrétní aplikace. V úvodu se tak zabývá technikami a metodami pro segmentaci obrazu, a to jak pomocí detekce hran tak pomocí detekce celých oblastí. Popis je dostatečný a podává jasný přehled problematiky. V další části jsou uvedeny základy problematiky neuronových sítí použitelných pro extrakci textů. Popis zvolené kategorie sítí je velmi dobře formulován, autor zde však používá i citace celých rozsáhlých bloků textu (odstavců), což není ideální. Obsah kapitoly 2.2.3.3 by mohl být převeden do předchozích částí. Celá teoretická rešerše obsažená v kapitole 2 může sloužit jako úvod do problematiky analýzy obrazu pomocí moderních metod.

Kapitola 3 se následně věnuje problému čtení a lokalizace textu v obrazu. Zejména první podkapitola pak má opět charakter spíše souhrnné informace o postupech využívaných v současnosti, získaných z aktuální literatury. Poslední podkapitola o významu a pokroku v oblasti konvolučních neuronových sítí by mohla být i součástí kapitoly 2.

Kapitola 4 se již zaměřuje na problematiku českých osobních dokladů a detailně popisuje údaje, které jsou z nich extrahovatelné i problémy, které jsou s touto extrakcí spojeny. Zkoumány jsou i související projekty s obdobnou tematikou a jejich nedostatky pro české prostředí. Autor z nich pak vychází při stanovení požadavků na navrhované řešení.

Jeho základní principy jsou definovány v kapitole 5. Uvedeny jsou již konkrétní mechanismy a algoritmy, přičemž z procesu výběru těchto postupů autor prezentuje především finální řešení. To by mělo být podpořeno alespoň stručným popisem procesu výběru a jeho kritérií. Právě kapitolu 5 lze považovat za zásadní přínos autora, neboť popisuje navrhovaný vlastní mechanismus extrakce informací.

Samotná implementace řešení je popsána v kapitole 6. Kladně lze hodnotit použití sofistikovaného generátoru příkladů pro trénovací množinu. Pro tuto a některé další části byl jako hlavní nástroj zvolen poněkud překvapivě systém Matlab. V kapitole 6.1.4 jsou uvedeny poměrně malé počty vzorů testovací množiny, které byly rozpoznávány, ty by bylo vhodné navýšit. Podkapitola 6.3 se pak věnuje popisu celé serverové aplikace, která je vhodně řešena jako samostatný produkt.

V kapitole 7 je následně popsána klientská aplikace. Rozporuplná je zde možnost opravy zjištěných identifikačních údajů, uživatel by mohl načtené údaje změnit, čímž by celá identifikace ztrácela smysl. Jako zbytečně omezující lze hodnotit použití jediného klientského zařízení, na kterém byly testy prováděny.

V kapitole 8 je popsáno testování a ověření řešení v praxi. Součástí jsou i realizovaná vylepšení, jako reakce na praktické zkušenosti. V kapitole 8.2 by neměla být opomíjena část věnující se lokalizaci textu, i když je tato oblast dle autora dále rozvíjena.

Práce používá vhodný jazyk a styl. Text obsahuje menší množství drobných gramatických nedostatků, přičemž je vhodně doplněn řadou obrázků a grafů. Použitá literatura je zcela odpovídající a čítá 126 položek. Práce nemá žádné přílohy, ani CD se zdrojovými kódy aplikace.

Vzhledem k tomu, že téma práce bylo zvoleno po dohodě s externí firmou, měla by práce obsahovat i vyjádření této firmy k dosaženým výsledkům.

Celkově je předložená práce velmi dobrým příkladem práce Mgr. úrovně. Cílů práce bylo dosaženo a autor v ní jasně prokázal svou schopnost integrovat informace z různých zdrojů do smysluplné rešerše a následně získané poznatky prakticky tvůrčím způsobem uplatnil v návrhu

vlastního řešení založeného jak na integraci existujících komponent, tak na jejich vlastní úpravě. Zejména praktické kroky jsou dokumentovány rovněž číselnými údaji umožňujícími udělat si představu o výsledné efektivitě a funkčnosti aplikace. Práci doporučuji k obhajobě a navrhuji hodnotit stupněm výborně.

Případné otázky při obhajobě a náměty do diskuze:

1. Do jaké míry je uživateli umožněno měnit údaje získané pomocí automatické identifikace obrazů?
2. Porovnejte získané výsledky s obdobně zaměřenými aplikacemi na trhu včetně komerčních.
3. Prezentujte online vaše řešení.
4. Jak hodnotí výsledky práce firma, na jejíž popud práce vznikla?

Práci

☒ doporučuji

☐ nedoporučuji

uznat jako diplomovou/bakalářskou.

Navrhuji hodnocení stupněm:

☒ výborně ☐ velmi dobře ☐ dobře ☐ neprospěl/a

Místo, datum a podpis vedoucího/oponenta:
V Českých Budějovicích, 1. 5. 2018

