

Posudek oponenta diplomové práce

Student: Benda Tomáš, Bc.
Téma: Rozpoznávání tváří (id 19775)
Oponent: Hradiš Michal, Ing., Ph.D., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Jedná se o úlohu spíše náročnou vzhledem k použitým metodám, práci s daty i náročnosti experimentů.
- 2. Splnění požadavků zadání** **zadání splněno**
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
- 4. Prezentační úroveň předložené práce** **61 b. (D)**

Práce v některých ohledech není rozvržená vhodně, a pochopitelnost a přístupnost textu pro čtenáře je poněkud horší. Podobně trpí návaznost textu, takže čtenář často musí jen hádat, proč čte určité informace a jak souvisí s řešením studenta. Některé informace se také v práci zbytečně opakují. V práci je mnoho nepřesností. Následují některé mé poznámky ke struktuře práce:

- V úvodu by mohl být explicitně popsán přínos práce a vlastní práce studenta.
- Název Kapitoly 2 "Vyhledávání obličejů ve videu" je poněkud zavádějící, ale možná jen kvůli rozcházejícím se představám o významu slova "vyhledávání".
- V Kap. 2 chybí nějaký obecný úvod a přehled. Proč je hned text o odečítání pozadí? Details o GMM jsou zbytečné. I druhý algoritmus na odečítání pozadí nebyl použit. Použit byl jiný, který ale tady popsán není.
- Není jasná souvislost K-means s detekcí obličeje. Stejně tak i proč je zde "Klasifikace" a SVM.
- V 2.4.3. nejsou popsány jen architektury sítí, ale spíše celkově přístupy k rozpoznávání lidí podle obličeje.
- Proč Kap. 3 začíná výpisem knihoven a jazyků, když ještě není jasné, co bylo vytvořeno?
- Celkově je popis funkce rozpoznávání hůře pochopitelný a pomíchaný.
- Využití MongoDB, JSON a DBSCAN je v práci dvakrát.

Některé poznámky k faktickému obsahu:

- GMM není "rekurzivní modelovací technika pro modelování prostředí". V popisu jsou nepřesnosti.
- 2.2 občas nedává smysl.
- Rovnice 2.10 nedává smysl. Co je g_x , g_z ?
- 2.4.1 Hodnoty vstupu neuronu nemusí být jen binární (0,1) nebo bipolární (-1, 0, +1), "u" není koeficient učení. Perceptron neklasifikuje nadrovinu do dvou nadrovin!
- Backpropagation: Síť není inicializována z intervalu $\langle -0,5; +0,5 \rangle$.
- GD není "založen na postupném snižování gradientu spolu s chybovou funkcí".
- Konvoluční síť spíše neřeší invariance "tak, že obraz postupně rozdělují na menší části a sdílejí váhy". Popis konvoluční vrstvy je dost na hraně. ReLU je aktivační funkce!
- Rovnice 2.22 nezobrazuje "postup klasifikace K-NN".
- 2.4.4 - Výstupem neuronových sítí není vždy 128 dimenzionální vektor a jejich vzdálenost není "definovaná jako euklidovská".

- 5. Formální úprava technické zprávy** **57 b. (E)**

Práce je po typografické stránce průměrná s několika nedostatky. Po jazykové stránce je práce ale velmi špatná. Text obsahuje velké množství základních pravopisných chyb i úplně chybně vystavěných vět, které nedávají smysl. Nejvyšší počet chyb na stránce byl asi osm. Je překvapivé, kolik je v práci chyb ve shodě podmětu s přísudkem.

Překvapivá pro mě byla špatná jazyková (i faktická) úroveň kapitoly 2.2, ale po nahlédnutí do odkazovaného zdroje musím studenta ocenit, protože úroveň tohoto zdroje je ještě horší a často hraničí s úplnou nesmyslností. Nechápu, jak mohl někdo takový text přijmout na konferenci a vydat a obdivuji, že z něj student něco pochopil. Mezi typografické prohřešky patří nízká kvalita mnoha obrázků, obr. 2.8 obsahující rovnici, a dokonce rastrové tabulky, které jsou uvedeny jako obrázky (obr. 2.18 a 2.19). Také jsou v textu některé části rovnic vysázené obyčejným textem. Ukázky některých chybných vět z textu práce:

- Úvod: "..., se budu zabývat algoritmy používaných při sekreci pohybujiícího se popředí od statického pozadí videa, ..."
- s. 39 "Z celkového počtu 56ti osob 2 osoby nebyli nalezeny pomocí detektoru hledající obličej ze předu a to z toho důvodu, že po celou dobu záznamu byli tyto osoby natočeny"
- "Klíčové body byli vybírány tak, aby popisovali obličej ..."

6. Práce s literaturou

62 b. (D)

Práce odkazuje 24 zdrojů, které jsou většinou kvalitní a dobře pokrývají řešenou oblast. Z neznámého důvodu u mnoha z nich chybí základní informace - někdy informace o typu, někdy autor, někdy dokonce samotný název díla (4x). V textu jsou odkazy použity nevhodně na koncích odstavců a sekcí, a někdy tak není jasné, odkud které informace pocházejí. Někdy zdroje chybí - 2.4.5 DeepFace, s. 34 u triplet loss a obrázku, ...

Občas z textu není patrné, co je dílem studenta a co převzal. Důvodem je často i nepoužívání první osoby. Například v závěru je napsané: "V práci byly použity obě tyto sítě. První síť byla natrénovaná na obličejích natočených směrem ke kameře." Student ale tuto síť netrénoval, přičemž text to silně evokuje. V práci je mnoho nepřesností, což napovídá, že student pochopil uvedené zdroje spíše jen povrchně.

7. Realizační výstup

75 b. (C)

Použitý přístup studenta je celkově vhodný a má smysl. Výsledné úspěšnosti ale nejsou příliš vysoké, což může být způsobeno nedokonalou realizací některých částí, malými datovými sadami, nebo i chybami ve vyhodnocení. Vyhodnocení by celkově mělo být standardnější a důkladnější.

8. Využitelnost výsledků

Vytvořená datová sada by mohla být použitelná pro testování. Natrénované neuronové sítě a software pravděpodobně další uplatnění nenajdou.

9. Otázky k obhajobě

- Jak jste získal 1910 videozáznamů pro 47 osob z YouTube faces, když v datové sadě je maximálně 6 záznamů pro jednu osobu?
- Proč jste nevyhodnotil Váš systém na celé YouTube faces standardním způsobem?
- Proč používáte MongoDB na ukládání descriptorů obličejů ve formátu JSON? Není to krajně neefektivní a zbytečné?
- Jste si jistý, že nastavení prahů při identifikaci na EER bod verifikace je vhodné? Jaká bude pak FAR, když už budete mít v databázi například 100 lidí?

10. Souhrnné hodnocení

63 b. uspokojivě (D)

Práce se ubírala vhodným směrem a student řešil zajímavý aspekt rozpoznávání lidí podle obličejů. Vytvořil funkční systém a odvedl kus práce, ale zpracování včetně vyhodnocení má nedostatky a text působí uspěchaně a jeho kvalita je horší.

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

V Brně dne: 9. června 2017

.....
podpis