

Posudek oponenta bakalářské práce

Student: Soroka Matej
Téma: Detekce kvádrů-krabic v obraze (id 23247)
Oponent: Bartl Vojtěch, Ing., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Na bakalářskou práci se jedná spíše o obtížnější zadání, jelikož je potřeba nastudovat velké množství informací v oblasti počítačového vidění. Zároveň jde o nepřilíš řešený problém a je tedy třeba vymyslet do jisté míry nový postup při řešení problému.
- 2. Splnění požadavků zadání** **zadání téměř splněno**
Mírné odklonění od zadání spočívá v tom, že se neprovádí detekce objektu (krabic) v pravém slova smyslu. Namísto detekce objektu je prováděna klasifikace jednotlivých částí krabice - rohů a hran. Zároveň datová sada pro trénování je dosti omezená (snímky jen tří krabic z různých úhlů a pohledů) a měla by obsahovat daleko větší počet krabic. Součástí zadání je vytvoření videa prezentujícího práci, které nebylo odevzdáno.
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
Rozsah zprávy je v běžném rozmezí. Některé části jsou ovšem naprosto zbytečné a působí, že mají za cíl jen práci prodloužit. Například úvod kapitoly 2.3, kapitoly 2.3.3 a 2.3.4.
- 4. Prezentační úroveň předložené práce** **45 b. (F)**
Celkově se práce velice špatně čte. Kapitola 3.4 popisující využití technické nástroje působí, že je v podstatě jen přeložena z nějakého internetového zdroje v anglickém jazyce (domněnka podložena zapomenutou anglickou větou v části 3.4.3). Rovněž se v práci píše o různých variantách klasifikace - 10, 8 a 6 tříd, ale popsána je v podstatě jediná varianta a ostatní jsou jen okrajově zmíněny v experimentech, ale nejsou nikterak vysvětleny. Popis modelu neuronové sítě v kapitole 4.3 formou zdrojového kódu je pro neznalého čtenáře naprosto nevyhovující a nicneříkající. Rovněž zmínky o plně propojených konvolučních vrstvách jsou matoucí (nic takového neexistuje). Rovněž se občas vyskytují části textu, které jsou napsány velice krkolomně a téměř nedávají smysl - např. 2.4.4, 3.1.1, 3.1.2 a 4.3.3. Vzorec pro konvoluci v kapitole 2.4.1 je naprosto nevyhovující a nepopisuje to, co by měl. Celá kapitola 6 popisuje metody, které by mohly být použity, ale vůbec použity nejsou, přestože by pravděpodobně dosahovaly lepších výsledků. Tato kapitola by se tak měla objevit spíše někde na začátku práce, kde se řeší daná problematika a existující přístupy. V závěru je popisován experiment, který v samotných experimentech zmíněn vůbec není.
- 5. Formální úprava technické zprávy** **45 b. (F)**
Jak již bylo zmíněno, některé části textu jsou napsány dosti divně a nedávají příliš smysl. Špatně jsou používány téměř veškeré reference. Citace nejsou součástí textu a jsou vyjmuty z vět. Veškeré obrázky i tabulky jsou naprosto bez referencí a někdy je složité se dobrat (především tabulky v kapitole 4) co k čemu správně patří. Objevuje se několik chyb v textu - např. chybějící mezery, nejednotnost použitých výrazů, věty bez sloves, nedokončená věta apod.
V kapitole 4.1.3 je nikam nevedoucí reference na kapitolu. Rovněž se vyskytuje nejednotnost ve formátování textu - někdy nový odstavec, někdy nový řádek u totožného rámce. Stejně tak je nejednotnost ve zvýrazňování pojmů (někdy ano, jindy zase ne). Obrázek 5.8 obsahuje popis obrázku 5.1, přestože znázorňuje něco jiného.
- 6. Práce s literaturou** **50 b. (E)**
Drtivá většina citované literatury se týká obecných základů a principů neuronových sítí nebo používaných nástrojů. Samotného tématu detekce/klasifikace se týká v podstatě jediná reference. Další dvě reference se týkají architektur U-Net a MirrorNet, které jsou v práci sice zmíněny, ale nejsou nijak využity. Z celkového počtu 20 referencí jsou tedy 3 opravdu týkající se samotné problematiky detekce nebo klasifikace. Zbytek jsou různé návody a nástroje, které jsou sice také potřebné, ale daleko více referencí by se mělo týkat samotného problému řešeného prací.
- 7. Realizační výstup** **55 b. (E)**
Výstupem je natrénovaná neuronová síť pro klasifikaci jednotlivých výřezů obrázku krabice. Odevzdaný soubor pro detekci je ovšem nepoužitelný (zahazuje detekce, neprovádí detekci pomocí sliding window). Odevzdané Jupyter Notebooky, které byly použity pro trénování obsahují chyby. Trénovací složka pro YOLO detektor je prázdná. Soubor readme, který by měl řešení nějak popisovat je rovněž prázdný. Při předvedení studentem byla použita funkční verze detektoru (nebyla odevzdána do IS), která na poskytnutých datech vyprodukovala výstup. Ten se ale nedá považovat na příliš kvalitní, kdy většina rohů krabic nebyla správně klasifikována. Celé

vyhodnocení práce je provedeno pouze vyčíslením přesnosti nad validační sadou, která je ovšem vytvořena ze stejných dat jako trénovací sada, takže není zcela jasné, zda nedošlo k přetrénování modelu. Během konzultace nebyl student schopen vysvětlit část kódu, na niž jsem se dotazoval.

8. Využitelnost výsledků

Odevzdaná verze programu je zcela nefunkční. Verze, kterou prezentoval student později na konzultaci již výsledky poskytuje, ale nejsem zcela přesvědčen o funkčnosti řešení (na testovacích datech, které se lišily od trénovacích a validačních, byl výsledek spíše nepoužitelný). Celý přístup klasifikace výřezů pomocí sliding window neřeší samotný problém detekce krabice (pro následnou detekci by byla potřeba nějaká dodatečná logika, která by zpracovala klasifikované body). Zároveň není nijak řešen problém klasifikace při překrývajících se pozicích sliding window.

9. Otázky k obhajobě

- Prováděl jste nějaký experiment, kdy byste zjistil výstup na jiných datech než na validační sadě?
- Zkoušel jste jiný typ augmentace dat nežli jen rotaci?

10. Souhrnné hodnocení

47 b. nevyhovující (F)

Samotný text práce působí, že byl dopisován na poslední chvíli. Obsahuje velké množství různých chyb a velice špatně se čte. Odevzdaný program je zcela nefunkční. Prezentovaná varianta sice funkční byla, ale výsledky byly spíše nepoužitelné. Problém je i v samotném vyhodnocení práce, kdy je použita přesnost na validační sadě, která ovšem vznikla ze stejných dat jako trénovací sada. Bylo nafoceno 272 snímků tří krabic z různých úhlů a z těchto snímků byly extrahovány výřezy rohů a hran, na nichž se poté trénovala neuronová síť pro klasifikaci. Jelikož je trénovací i validační sada ze stejných obrázků, nelze při vysoké přesnosti na validační sadě vyloučit přetrénování. Přestože se jedná o těžší zadání, dosažené výsledky jsou téměř nepoužitelné a samotný text práce je opravdu špatně sepsán. Navržené postupy v poslední kapitole by byly pravděpodobně výrazně vhodnější, ale bohužel k jejich aplikování nedošlo. Zásadní problém rovněž nastal při vytváření datasetu, kdy byly uloženy výřezy namísto anotovaných pozic, což znemožnilo použití dat při jakémkoli jiném přístupu.

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

V Brně dne: 3. června 2021

Bartl Vojtěch, Ing.
oponent