

Review of Master's Thesis

Student: Válek Lukáš, Bc.
Title: Automated Identification of Graphical UI Elements for Robotic Quality Assurance (id 25178)
Reviewer: Herout Adam, prof. Ing., Ph.D., DCGM FIT BUT

- 1. Assignment complexity** **average assignment**

Cílem práce bylo vyhledávat v obraze z kamery snímající uživatelské rozhraní jednotlivé prvky UI tak, aby bylo možné testovat funkčnost aplikace.
Řešitel vyšel z některých hotových řešení dodaných firmou YSoft, natrénovával klasifikační neuronovou síť standardní architektury a vytvořil dílčí konvenční řešení pro detekci objektů.
- 2. Completeness of assignment requirements** **assignment fulfilled**
- 3. Length of technical report** **in usual extent**
- 4. Presentation level of technical report** **75 p. (C)**

Struktura a obsah technické zprávy jsou rozumné.
Poněkud zarážející je, že v době dostupných hotových a výkonných detektorů objektů založených na hlubokém učení řešitel tyto ve svém textu zcela ignoruje a svoji práci staví na značně "retro" pojetí, kdy konvenčními přístupy (extrakce hran, shlukování, vyhledávání spojitých komponent, ...) identifikuje potenciální objekty a ty pak klasifikuje klasifikační neuronovou sítí.
Z textu je patrné, že řešitel využívá některé již hotové moduly fungující ve stávajícím řešení téhož problému ve firmě YSoft (rozpoznání, lokalizace textu). V celkovém popisu systému není posuzovateli úplně patrné, co tedy řešitel vlastně sám vyřešil a kde se spolehl na dodané řešení a na dodaná data.
Práce působí dojmem, že řešitel dílčím způsobem rozvíjel stávající řešení dodané firmou YSoft a navazoval na jeho architekturu. Tím možná propásl šanci v rámci diplomové práce na zadaný problém zaútočit novým, aktuálním přístupem.
Teoretická část textu se věnuje mnoha informacím, které jsou pro samotnou práci nerelevantní nebo zcela nebyly využity, či byly ignorovány. Jsou popisovány existující datasety uživatelských rozhraní mobilních aplikací - řešitel je nevyužívá. Jsou citovány datasety jako CIFAR a MNIST, které pro řešenou práci nejsou přínosné a obsahují obrázky v rozlišení 32x32.
- 5. Formal aspects of technical report** **75 p. (C)**

Technická zpráva je psána velice dobrou angličtinou prakticky bez chyb a velice dobře srozumitelnou. Formátování a typografie jsou průměrné.
Otravné jsou obrázky typicky vektorové povahy, které jsou vloženy jako rastry v nevelkém rozlišení.
- 6. Literature usage** **85 p. (B)**

Technická zpráva cituje velké množství kvalitních zdrojů.
Práce propásla svoji příležitost podívat se na problém způsobem, který by odpovídal poslednímu stavu poznání - zdá se, že řešitel si osvojil pojetí řešení dodané firmou YSoft, které vzniklo již před lety, a toto inkrementálně vylepšoval.
- 7. Implementation results** **65 p. (D)**

Realizačním výstupem je sada skriptů v jazyce python a provedená měření. Řešitel používal datovou sadu dodanou firmou YSoft.
Zdrojové kódy jsou pramálo nebo vůbec komentované - rozhodně chybí komentáře v docstring, které umožní znovupoužitelnost a eventuelní navázání na práci řešitele.
V kapitole 6.7 řešitel v zásadě vysvětluje, že není možné vytvořit jasná data ground-truth a jeho úlohy je potřeba hodnotit subjektivně a ručně. Toto je značně pochybné tvrzení. Podle toho ovšem vypadá vyhodnocení řešitelovy práce - v zásadě jde jen o tabulku 6.2, kde se míchají počty charakterizující dataset a několik málo číselných procentuálních údajů. Detekci i klasifikaci je přitom zvykem hodnotit v závislosti na nastavitelném prahu pomocí křivek ROC, PRC, DET apod. I čísla reportovaná v tabulce jsou pochybná - řešitel před tím tvrdí, že ground-truth neexistuje a že ignoruje ta data o třídách, která k dispozici má "the class information of the element was ignored during the analysis."
Neuronová síť, kterou řešitel použil, pracuje s obrázky 240x240, klasifikované výřezy s prvky UI jsou však ve výrazně menším rozlišení a jsou pro síť zvětšovány - to je pravděpodobně značně suboptimální.

8. Utilizability of results

Řešitel zřejmě vytvářel svoji práci pro firmu YSoft, která tuto funkčnost již využívá a řešitel měl zřejmě existující řešení vylepšit.

Tím, že řešení není řádně vyhodnoceno, je zřejmě obtížné říci, zda k vylepšení skutečně došlo.

Práce pravděpodobně také propásla možnost přistoupit k řešení novým, čerstvým způsobem, ale je postavená na existujících myšlenkách vtělených do stávajícího řešení.

9. Questions for defence

- Opravdu lze tvrdit, že systém, jako je ten váš, není možné vyhodnotit objektivně, není možné získat GT data a hodnocení musí být ruční a subjektivní?
- Nemohly by současné detektory založené na hlubokých CNN být v detekci a klasifikaci úspěšnější než konvenční nástroje pro vytipování míst v obraze a jejich následná klasifikace?
- Použitá síť EfficientNetB1 pracuje s obrázky 240x240. Typické výřezy, které má ve vaší práci klasifikovat, jsou zřejmě menší. Jaké je rozložení vašich výřezů? Nemohlo by být efektivnější použít síť, která by byla lépe přizpůsobená vašim datům?

10. Total assessment

70 p. good (C)

Řešitel řešil nejjednodušší problém.

Vytvořené řešení je dost "retro" založení (extrakce a shlukování hran). Když řešitel použil soudobou neuronovou síť, učinil tak zřejmě značně suboptimálně (síť je přizpůsobená pro obrázky ve výrazně větším rozlišení, než vyžaduje aplikace).

Není úplně jednoznačně patrné, co jsou myšlenky a řešení samotného řešitele a kdy postupoval podle pokynů firmy a využíval jejich už hotových řešení.

Technická zpráva je napsána velice dobrou angličtinou.

In Brno 1 June 2022

Herout Adam, prof. Ing., Ph.D.
reviewer