

Posudek Oponenta Diplomové Práce

Autor diplomové práce a její název:

Jiří Těžký: Aktivní detekce obětí s RGB-D-T senzorem

Oponent a autor tohoto dokumentu:

Ing. David Hurych, PhD.
Software Engineer - Research Engineer
Valeo R&D CDV Prague
Office Park Hloubětín
Poděbradská 55/88
198 00 Praha 9
Czech Republic
tel: +420 225 355 338

Shrnutí diplomové práce:

Cílem diplomové práce bylo efektivně využít data z RGB-D-T snímků získaných z nahrávek ze záchranného robota (výzkumný projekt EU) k detekci potenciálních lidských obětí ve scénáři záchranné operace. Konkrétně se jedná o návrh vhodné struktury konvoluční neuronové sítě (CNN) s využitím z dat z termo kamery a hloubkových dat. Součástí zadání práce bylo i experimentální vyhodnocení úspěšnosti detekce finální varianty CNN a její implementace pro ROS, který běží na záchranném robotovi.

Téma této práce je velmi zajímavé a naráží na hranice state of the art. Hlavní částí diplomové práce byly experimenty s učením a architekturami neuronových sítí pro úlohu detekce obětí ve scénáři záchranné operace.

Hodnocení:

Teoretický rozbor v kapitole 3.1 je sice stručný, ale dostačující k porozumění základních principů učení neuronových sítí. Využití předtrénované neuronové sítě se zdá být rozumnou volbou vzhledem k počtu trénovacích příkladů, které měl autor k dispozici.

Text obsahuje mnoho překlepů, chybějících slov a vágních tvrzení navíc podložených irelevantními odkazy na literaturu. Například v kapitole 3.1 strana 16 je tvrzení, že „Pokud $L(W)$ není konvexní funkce může SGD skončit v lokálním minimu, v praxi ale většinou pracuje velmi dobře [18]“. Fce $L(W)$ je téměř vždy nekonvexní a když se podíváme na web tutorial schovaný pod odkazem [18] zjistíme, že tam nejsou žádné výzkumné výsledky (teoretické ani praktické), které by nám říkaly něco o tom že „SGD v praxi pracuje velmi dobře“. SGD v podstatě vždy skončí v lokálním minimu a vlastně vůbec nevíme jak daleko jsme od globálního minima a tento odkaz nám odpověď nedává.

Kapitola 3.3 strana 21, u parametrů, které byly subjektem testování (γ , iter, power, base_lr, etc.) mi chybí konkrétní odkazy na rovnice v odvození učení neuronových sítí v kapitole 3.1. Velmi by to usnadnilo orientaci.

Hlavní částí práce jsou experimenty se strukturou sítě, nastavení parametrů učení a způsobem připojení kanálů s hloubkovou informací a obrazem z termo kamery. Autor postupně otestoval několik inkrementálně vylepšených struktur sítě a vždy srovnal výsledky na testovací sadě obrázků a někdy i na té části testovací sady, která sestává z reálných dat. Testy na reálných datech (obr. 3.9 a 3.17) poukazují na fakt, že naučené sítě špatně generalizují – problém bude pravděpodobně v rozdílnosti generovaných dat a dat reálných, jak autor správně zhodnotil v kapitole 3.3.

Chybí mi tu snaha pátrat po reálné příčině rozdílnosti výsledků, když přidáme či odebereme tu kterou vrstvu. Veškeré hodnocení výsledků je složeno z pouhých odhadů, které mohou, ale nemusejí být pravdivé. Ale bohužel takový je současný trend deep learningu konvolučních neuronových sítí, který nás často nutí používat deep learning frameworky s přednaučenými sítěmi jako black box bez bližšího porozumění co se v těchto vysoce nelineárních klasifikátorech děje a přístupem pokus omyl stavět sítě nové. Toto není kritika autora či této práce, protože bohužel tento trend vidíme i v prestižních vědeckých publikacích. Přitom ani netušíme jak velkou a jak komplexní neuronovou síť jsme schopni naučit s dostupnými daty a bez overfittingu což je dle mého názoru nutná podmínka k tomu abychom vůbec mohli začít srovnávat různé struktury sítě.

Návrh greedy algoritmu v kapitole 4 pro řízení pan-tilt jednotky s termo kamerou mi přijde rozumný a tím že jej autor implementoval pro ROS a funkčnost otestoval v reálném experimentu splnil tuto konkrétní část zadání. Experiment srovnávající greedy algoritmus a fixed camera jednoznačně ukazují, že greedy algoritmus dosahuje lepších výsledků.

Finální shrnutí:

Text práce je dobře strukturován a obsahuje vše potřebné k porozumění řešeného problému včetně rozboru state of the art, přehledné prezentace struktur testovaných sítí a experimentálních výsledků. Bohužel je text plný překlepů, chybějících slov a některých nepodložených tvrzení. Značení členů v matematických vzorcích (převážně typograficky) se nedrží zavedených konvencí. Celkově je práce špatně čitelná. Naopak obrazová schémata jsou přehledná a velmi napomáhají porozumění. Experimentální část vyhodnocující přesnost jednotlivých struktur sítí byla pro mě nejzajímavější. Je škoda, že snaze správně porozumět rozdílnosti výsledků nebyl v práci dán větší prostor a to i z důvodu, že úkolem autora nebylo samotné učení naprogramovat, ale hlavně odladit strukturu sítí. Finální část zakomponování CNN do ROS na robotovi byla bezpochyby úspěšná. Celkově hodnotím tuto diplomovou práci jako dobrou (známka C).

Otázky k obhajobě:

1. Jak byla vygenerována hloubková a termovizní data ke snímkům, do nichž byly uměle vloženy oběti? Při srovnání snímku z termovize 3.1 (uměle vložená oběť) a snímku z termovize 3.2 (reálná oběť) tak vidíme výrazný rozdíl v rozložení teploty na oběti a na podlaze. Nepoškodily uměle připravené termovizní obrázky učení konvoluční sítě?
2. V práci se zmiňujete, že Recall a Precision je lepší míra než false positive rate (FPR), false negative rate (FNR) či jejich opačné hodnoty true positive rate (TPR) a true negative rate (TNR). Přičemž $\text{Recall} = \text{TPR} = 1 - \text{TNR}$ čili jediný rozdíl nastává v Precision vs FPR. Jak byste použití $\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$ obhájil, když vezmeme v potaz, že se změnou počtu dat

v negativní sadě se mi adekvátně změní i FP a tím pádem se nutně musí změnit i Precision aniž bychom jakkoliv změnili parametry našeho detektoru či neuronové sítě. Přece přesnost našeho detektoru by měla být hodnota neměnná za předpokladu, že nijak nezměním parametry tohoto detektoru a Precision tuto podmínku nespĺňuje, jelikož je závislá na počtu dat v jednotlivých třídách. Tento jev u FPR, FNR nenastává. Nedává hodnota Precision prostor pro neadekvátní manipulaci s výsledky?

3. Jak byste po zkušenostech s učením CNN vypočítal/odhadnul nezbytný počet trénovacích příkladů v závislosti na komplexnosti (počtu parametrů) sítě?

V Praze, červen 2016

David Hurych