

Posudek bakalářské práce

Praha, 31. května, 2021

Název: Detekce modelu autonomních aut F1/10 na soutěžním okruhu
Autor: Lukáš Beneda
Vedoucí: Ing. Joel Matějka
Datum přijetí: 26/05/2021

Práce se zabývá vývojem a testováním detektoru autonomního závodního vozidla v obrazu kamery jiného vozidla na téže dráze. První kapitola motivuje celou práci. Druhá kapitola představuje soutěž F1/10 a podává stručný přehled neuronových sítí. Technická kapitola 3 prezentuje vytvořenou anotovanou databázi a dvě vlastní navržená řešení: (1) metoda klasifikátoru obrazu rozděleného na segmenty a (2) použitím standardního detektoru architektury Faster RCNN. Kapitola 4 se zabývá vyhodnocením obou metod z hlediska přesnosti a výpočetní rychlosti. Kapitola 5 diskutuje možná budoucí vylepšení. Kapitola 6 uzavírá práci.

Motivace práce je zřejmá a adekvátní. Vytvořená databáze s cca 53 tisíci obrázky (a 16,352 pozitivně) anotovanými obrázky je asi dostatečně velká pro následující experimenty. Celý proces vytváření anotace a databáze je pečlivě zdokumentován a jistě dobře poslouží dalším výzkumníkům. Zvolené řešení s Faster RCNN je, zdá se, funkční a dosahuje vcelku očekávaných výsledků. Autorovi se zjevně podařilo sestavit celý řetězec, od sběru anotovaných data, po návrh či úpravu existující architektury neuronové sítě, trénování, až po kvantitativní vyhodnocení výsledků. S výhledy budoucí práce lze souhlasit.

Dovolím si okomentovat zvolená řešení. Metoda rozřezání obrázku na okna a následná klasifikace je pokus spíše naivní. Jeho přínos vidím v tom, že se autor seznámil s použitím neuronových sítí. Této metodě se obvykle říká klasifikátor ve skenovacím okně, což je možný přístup, ale příliš výpočetně náročný s neuronovou sítí jako klasifikátorem. Navíc autor vůbec neuvažoval možná různá měřítka, tzn. velikosti auta v obraze. Ve skutečnosti by síť segmentů pro klasifikaci musela být komplikovanější, než rozdělení obrazu na obdélníky. Druhá metoda Faster RCNN je poměrně nový detektor, ale pracuje na principu postupné evaluace automaticky generovaných kandidátních pozic. Část neuronové sítě je vyhodnocována pro všechny kandidátní pozice. Je to detektor přesný, funguje dobře i u komplikovaných scén s vzájemně překrývajícími se objekty, ale není nejrychlejší. Pokud je rychlost detekce důležitým parametrem, doporučuji použít detektor typu YOLO (“You only look once”). Tento detektor vyhodnocuje obrázek najednou. Existuje několik vylepšení původního článku a taky verze v PyTorch. Další možnost zmíněná autorem je DPM. Není uvedena žádná citace, ale pokud to znamená “Deformable part model”, je to klasická dnes už zastaralá metoda detekce objektu. Pokud je k dispozici GPU, neuronové sítě jsou výrazně lepší.

Dále uvedu několik negativních stránek práce:

1. Neformální popis v teoretické části. Popis neuronových sítí a zejména jejich učení je vykládáno poněkud neformálním jazykem. Například by bylo vhodné uvést aspoň optimalizační problém, tj. co se minimalizuje a přes jaké proměnné. V sekci 2.4.1 o trénovacích algoritmech jsou sice vzorce uvedeny, ale už není vysvětleno, co jednotlivé symboly znamenají. Styl někdy až poněkud připomíná populární a popularizující texty určené pro neodbornou veřejnost, což pro bakalářskou práci není dobré.

2. Chybějící popis detektoru. Nepovažuji za úplně nezbytné dnes obecně popisovat, z čeho se skládají neuronové sítě. Bylo by mnohem přínosnější popsat architekturu použitého Faster RCNN detektoru. Obávám se, že autor možná neví, co je uvnitř tohoto detektoru a jak funguje.
3. Absence důležitých detailů. Některé podstatné detaily nebyly uvedeny. Například v jakém poměru byla rozdělena sebraná datová sada na trénovací, validační a testovací části a jakým způsobem. Taky, zda uvedené výsledky byly měřeny na nezávislé testovací sadě. Jaká tam byla přesně kamera. Na jakém stroji byly měřené časy. Uvést jen GPU 1.5 GHz opravdu nestačí.
4. Ne zcela přesvědčivé vyhodnocení. Vyhodnocení prezentuje jen jedinou statistiku “precision” (v procentech). Není ale jasné, jak se počítají správné detekce. Je to podle překryvu bounding-boxů, např. nad 50%? Nestačí uvádět jen precision, je třeba uvádět i recall. Velmi přesný detektor může být málo citlivý. Bylo by dobré vykreslovat Precision-Recall křivku. Standardní statistika je “mean average precision” (mAP), což je integrál pod precision-recall křivkou průměrovaný přes všechny třídy, pokud jich je víc.
5. Chybějící rešerše a přehled relevantní (state-of-the-art) literatury a obecně špatná práce s literaturou. Práce se ani nepokouší hledat přímo související práce. Úplně chybí citace článku o Faster RCNN [1], potom taky např. Yolo [2]. Citace z textu jsou nesprávně umístěny až za konec věty.

Uvážením výše uvedeného, navrhuji hodnocení

B – velmi dobře.

Ing. Jan Čech, Ph.D.

Otázky k obhajobě:

1. Používal jste nějakou metodu “data augmentation”? Snímky v trénovací sadě se (spolu s anotovaným bounding-boxem) geometricky nebo fotometricky transformují. Minimálně se používá zrcadlení z leva do prave. Nebo se obrázky rotují nebo affinně deformují. Intenzitu obrazu je možné taky transformovat (jas, kontrast, barevný tón). Tím získáme virtuálně větší množství dat, typicky omezíme přefitování a získáme větší robustnost detektoru.
2. Práce se zabývá detekcí jiného soutěžního auta v obraze kamery vlastního auta. Máte představu, jakým způsobem z detekce v obraze zjistíte vzájemnou polohu vozidel? To je jistě potřeba dále k plánování a řízení...

Reference

- [1] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” in *Proc. NeurIPS*, 2015.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proc. IEEE CVPR*, 2016.