

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	3D Object Detection for Autonomous Cars Weakly Supervised by 2D Cues
Jméno autora:	Jan Škvrna
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Kybernetiky
Oponent práce:	Ing. Patrik Vacek
Pracoviště oponenta práce:	Katedra Kybernetiky (FEL)

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Náplň práce bylo využití existujících supervizovaně naučených modelů pro instance segmentaci a detekci obrázků a využít je pro detekci v 3D doméně v oblasti autonomních aut. Aplikace vyžadovala spuštění hotových modelů, ale i vlastní naučení existujících detekčních sítí a hlavně implementaci vlastní supervize na základě existujících metod registrace bodů.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Student splnil všechny body zadání, prozkoumal velké množství způsobů, jak využít segmentované 3D point cloudy po kalibraci. Veškeré přístupy jsou zhodnoceny a diskutovány z hlediska přesnosti, výpočetní náročnosti i robustnosti, což považuji za velmi přínosné.	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Všechny metody jsou z mého pohledu rozumné, i intuitivní (což je správně). Využití registrace podle šablony objektu je správný postup k získání trénovacích dat pro 3D sítě, kde geometrie je hlavním diskriminativním příznakem. Cením též dotknutí se temporální složky v podobě trackování objektů a oblasti viditelnosti point cloudu. Některé metody by stálo za to prozkoumat více, např. zmíněnou oblast viditelnosti, kde diskutovaný problém je unifikovaný CAD model pro všechny automobily, nicméně v rozsahu diplomové práce je hloubka naprosto dostačující.	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Student čerpal z předních konferencí i časopiseckých článků a využil nejnovější metody z posledních let výzkumu. Data použil ze standartního srovnávacího datasetu pro autonomní řízení. Srovnání metod je přehledné v tabulkách v sekci experimenty.	

Formální a jazyková úroveň, rozsah práce	A - výborně
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je obsáhlá, strukturovaná a jazykové správná. Rovnice jsou řádně číslovány.	

Výběr zdrojů, korektnost citací	A - výborně
<i>Vyjádrěte se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	

viz. Odborná úroveň

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Práci hodnotím velmi kladně, obzvláště vyhodnocení experimentální části a je třeba vyzdvihnout nadstandartní výsledky, které mají potenciál k publikaci na menších konferencích. Následující dotazy se mimo prvního týkají spíše možných rozšíření.

1) Věty typu **“without using any human labels”** jsou z mého pohledu zavádějící (obzvláště když vlastní práci klasifikujete jako weakly-supervised). Chápu, že se naráží na to, že anotace nejsou ve 3D, ale MS-COCO dataset obsahuje 12876 ručně anotovaných snímků [https://cocodataset.org/#explore], kde se nachází jedno nebo více automobilů z různých sekvencí na kterých se staví. Zavádí se i mírný prior při výběru nejlepší validační backbone sítě (RegNetY) na 2D KITTI obrázcích. Pokud je tedy cíl využít foundation modely, tedy využití již existujících ručně anotovaných snímků, jaké jsou proti-argumenty k použití hotových kamerových detektorů (nebo i 3D detektorů) z vyjmenovaných datasetů (Waymo a nuScenes), které jsou specializované na dopravní situace?

2) Na straně 43 se diskutuje vliv outlierů na chamfer distance. Jaký by byl efekt při použití tzv. “Truncated chamfer distance”, tedy výpočtu lossu pouze na bodech vzdálených do nějaké omezené vzdálenosti?

3) Předpokládá se určitý geometrický prior ve formě velikosti bounding boxu napočítaný z aut KITTI datasetu. Co způsobí tento prior při chybě sítě z detectronu2, kde například chybně zamění velké nákladní auto za osobní automobil?

4) Může být Vámi navržené řešení aplikováno i na další agenty na silnicích, například Lidi a cyklisty? Může nastat nějaký problém v těchto dvou kategoriích?

5) Na obrázků 5.9 a), kde je spatně matchnutý automobil, ale správná trajektorie, zkoušel jste zavést nějaký loss, který by dokázal řešit tento případ?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 16.1.2024

Podpis: