

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Generating of the Training Data for 2D Segmentation
Jméno autora:	Jan Ferbr
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Vedoucí práce:	Michal Polic
Pracoviště vedoucího práce:	Český institut informatiky, robotiky a kybernetiky

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	mimořádně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání práce navazuje na existující vývoj v rámci EU H2020 grantu ARTwin. Konkrétně měl student nasnímat laboratoř B-670 pomocí několika v projektu používaných kamer, tj. HoloLens, Azure Kinect a S10e; vytvořit 3D model z nasnímaných RGB-D fotografií; doplnit chybějících částí rekonstrukce pomocí SPSG; manuálně segmentovat 3D model na jednotlivé objekty; vykreslit jednotlivé objekty do 2D snímků společně s příslušnými maskami a použít tyto data pro trénování 2D segmentace.	

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 zadání v plném rozsahu.	

Aktivita a samostatnost při zpracování práce	A - výborně
<i>Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatně tvůrčí práce.</i>	
Student pracoval aktivně a prokázal schopnost samostatně řešit zadané cíle. Na pravidelné konzultace docházel připraven.	

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>	
Vzhledem k diskusi v rámci konzultací se domnívám, že student nabyl řadu znalostí, které se probírají až v navazujícím inženýrském studiu. Konkrétně se jedná o princip řídké a husté rekonstrukce (předmět TDV), základy lineární algebry, transformace mezi souřadnicovými systémy a možnosti reprezentace rotačních matic (předměty PRO a GVG). Dále pak rozšířil své znalosti ve softwarovém vývoji pod několika platformami a v několika programovacích jazycích. Nabyté znalosti samostatně používal pro řešení dílčích úkolů a doplnění chybějící částí SPSG algoritmu.	

Formální a jazyková úroveň, rozsah práce	C - dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Vzhledem k chybějící dokumentaci a částem kódu SPSG se hlavní tvůrčí činnost na projektu překlápěla do softwarově zaměřené bakalářské práce, čemuž odpovídá text práce. Přínosem pro CV komunitu je samotný návod a implementace pomocných funkcí, tak, aby byl algoritmus SPSG spustitelný na vlastních datech a použitelný pro další vývoj. V tomto ohledu přináší práce zásadní přínos pro vědeckou komunitu. Zadání přesně nespécifikuje, že by měl student matematicky popsat jednotlivé metody, a proto se domnívám, že je formální úroveň a rozsah práce odpovídající. Vytknul bych nepřesnosti ve vyjadřování (např. při popisu neuronové sítě), chybějící vysvětlení všech pojmů (např. chybí popis modelu kamery použitého v COLMAPu, vysvětlení, co je TSDF apod.) a matematický popis (např. chyby ve vzorcích 5.7 a 5.8).	

Výběr zdrojů, korektnost citací

C - dobře

Vyjádřete 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.

Student nastudoval relevantní publikace a uvedl je v práci. Jako nedostatek lze vytknout neúplnost údajů v a formátování uvedené literatury.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Vložte komentář (nepovinné hodnocení).

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Zadání práce bylo navrženo tak, aby student naplno využil znalostí, které již v době zadání ovládal on sám nebo jeho kolegové. Díky pravidelným konzultacím, jeho péči a aktivnímu přístupu vyřešil téměř všechny úkoly kromě spuštění SPSPG na vlastních datech během prvních 4 týdnů. Na spuštění SPSPG na vlastních datech jsme ve spolupráci s doktorskými studenty z jiných univerzit a autory SPSPG aktivně pracovali po dobu minimálně 6 měsíců. Z tohoto důvodu hodnotím finální výsledek, tj. spuštění SPSPG, dopsání chybějících kódů, dopsání dokumentace a rozchození této metody na vlastních datech jako zásadní přínos pro CV komunitu. Zprovoznění SPSPG se podařilo těsně před odevzdáním práce a nezbyl čas na důkladnou opravu textů. Na základě pravidelných konzultací se domnívám, že student problematice porozuměl, ale nebyl ji schopen v rychlém čase kvalitně popsat.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 2.6.2022

Podpis:

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Generation of Training Data for 2D Segmentation
Jméno autora:	Jan Ferbr
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Radim Šára
Pracoviště oponenta práce:	Katedra kybernetiky

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	mimořádně náročné
---------------	--------------------------

Hodnocení náročnosti zadání závěrečné práce.

Zadání se týká automatické tvorby anotovaných trénovacích množin pro segmentaci instancí objektů v obrazech. Princip je založen na tom, že se roboticky nasnímají data z reálného prostředí, pomocí vhodného systému pro 3D vnímání se zrekonstruuje 3D model scény a rozpoznají a nasegmentují se v ní objekty ve 3D prostoru (existujícími metodami). Taková segmentace je snazší úloha než segmentace z pouhého barevného obrazu. Lze tedy uvažovat o jejím použití jako učitele. Ze získaného 3D modelu scény se vygenerují pohledy do virtuálních kamer a 3D segmentace se promítnou do vygenerovaných obrazů. Z výsledku lze sestavit anotovanou trénovací množinu pro učení segmentace obrazů bez 3D informace. Protože ale automaticky získaný 3D model obsahuje prázdná místa (díry), bylo hlavním cílem této práce prokázat, že s pomocí nedávno publikovaného, ale nevyzkoušeného algoritmu SPSG pro doplňování prázdných míst, k němuž je k dispozici jen autorský výzkumný software, lze docílit lepších výsledků při než bez takového doplňování dat. Výsledky naznačují opak, ale experiment je neprůkazný. Hlavní problém je ten, že zadání je postaveno na předpokladu, že proces automatického generování trénovací množiny lze vytvořit na základě předdimenzovaného a komplikovaného systému 3D vidění s několika různorodými senzorickými systémy, kde každý z nich má vlastní sadu problémů a jejichž výsledky jsou zatíženy různými artefakty. V počítačovém vidění nelze takovýto mechanický postup použít, protože žádná metoda není a nemůže být dokonalá. Takový projekt je možno svěřit jen expertovi, který se orientuje v tom, jaké mají jednotlivé metody nedostatky a umí je kvalifikovaně obejít, případně některé dílčí metody zcela odmítnout. Student přitom musel uvést do provozu řadu komponent, které by vyžadovaly netriviální spojení. Zadání tak svým rozsahem a nutností expertní znalosti vysoce přeceňuje schopnosti běžného studenta na úrovni bakalářského studia.

Splnění zadání	splněno
-----------------------	----------------

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.

Student se v rámci omezeného času vyrovnal s požadavky zadání tak, jak je formulováno. Bonusový bod zadání již nestačil splnit, ale to nebylo povinnou součástí zadání.

Zvolený postup řešení	částečně vhodný
------------------------------	------------------------

Posuďte, zda student zvolil správný postup nebo metody řešení.

Proti postupu mám výhrady uvedené výše, ale tyto výhrady nejsou výtkami vůči studentovi. Otázkou do diskuse je, jak často se se svým vedoucím scházel, jak průběžně modifikovali původní plány a kolikrát byl text práce po konzultaci s vedoucím revidován.

Odborná úroveň	D - uspokojivě
-----------------------	-----------------------

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Zasazení práce do kontextu výzkumu ve strojovém učení chybí. Ale ani samotný výběr metod není zdůvodněn, ani ověřen. Vhodnost a funkcionality jednotlivých systémů není posouzena. Například na str. 27 se tvrdí *all the HoloLens 2 frames were not as blurry as the Azure Kinect images*. Proč? Co bylo uděláno, aby obrázky z Kinectu rozmazané nebyly?

Práce je na mnoha místech neúplná nebo nepřesná. Co se míní pod *poincloud fusion*? Jak funguje? Jak přesná je fáze registrace pro fúzi? Jak je omezena redundance? Není důležité, že fúze používá knihovnu Open CV, ale je důležité, na jakém algoritmu je tato konkrétní metoda založena, protože to určuje její vlastnosti. Ale je nutno být přesný: To, že úloha registrace mraků bodů byla formulována jako *procrustes* algorithm jako vysvětlení nepostačuje, protože tento algoritmus má různé varianty, které se při velkých rotacích mohou chovat jinak, dokonce při velkých rotacích selhávat.

Není vysvětleno proč bylo použito tolik senzorů. Očekával bych, že u každého bude uvedeno proč je součástí projektu a čím je jeho funkcionality unikátní a nepostradatelná. Také není jasné, na základě jakých potřeb a podle jakého protokolu byla nasnímaná experimentální data.

Jak bylo rozhodnuto o vnitřním modelu kamery (sekce 5.3.3)? Je koeficient radiálního (?) zkreslení 0.05 velký nebo malý (sekce 5.3.3)? Otázku nelze rozhodnout, když model zkreslení není popsán.

Co se považuje za *SoTA approach* v sekci 5.4?

Matice v (5.1) není projekční matice kamery, ale matice euklidovské transformace ve trojrozměrném prostoru (prvek $SE(3)$), reprezentující transformaci ze světové souřadné soustavy do vnitřní souřadné soustavy kamery. Střed C kamery vyjádřený ve světových souřadnicích je pak dán inverzí této matice a není třeba složitější odvozování.

Není jasné co znamená notace ve vzorci (5.5). Vzorec nedává žádný smysl. Transformace působí na vektor zleva, ne zprava.

Jaké dva body byly zvoleny pro odhad rotace v sekci 5.4.6?

Vzorec (5.8) je chybný, platí za předpokladu, že w je jednotkový vektor, ale tak to není, viz (5.6). **Vzorec (5.7) je také chybný**, skalární součin definuje kosinus úhlu, ne úhel samotný. Jsou v implementaci také tyto chyby?

Chybí vysvětlení důležitých konceptů: například co jsou *non-local prototype masks* u YOLACT? Co je *PV camera frame*? Co je *TSDV function*? Co je to *[datagen] principal axis*?

V práci je řada drobných nepřesností typu *neural network is a function created as a concatenation of simpler functions*, konvoluční maska nebo digitální obrázek má jak velikost, tak dimenzi, ale tyto pojmy nelze zaměňovat, dataset není totéž co formát dat.

Formální a jazyková úroveň, rozsah práce

C - dobře

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce je psaná v anglickém jazyce a je po jazykové stránce srozumitelná, s malým počtem jazykových chyb. Struktura textu je ale nevhodná, text je zatížen velkým počtem nepodstatných technikalit typu v jakém adresáři je jaký soubor a čtenáři uniká cíl a hlavní myšlenky zvoleného postupu, které navíc nejsou dostatečně jasně prezentovány, jak diskutováno výše. Implementační detaily by patřily do přílohy. Naopak, formální popis metod v práci chybí.

Proč je popis senzorů v textu na dvou místech (v kapitolách 2 a 4)?

V práci jsou stylistické chyby:

- Odkazy na obrázky, tabulky, atd se píše s velkým písmenem, tedy např. Figure 1, Section 2.5, nikoliv figure 1, section 2.5.
- Pokud jsou ve větě všechna čísla od nuly do devíti, píše se slovy.
- Dopředné reference se nedělají, jen ve velmi vyjimečných případech. Viz str 44.

Výběr zdrojů, korektnost citací

D - uspokojivě

Vyjádřete 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.

Výběr pramenů odpovídá potřebám a zdroje jsou citovány v potřebném rozsahu. Vlastní výsledky a úvahy jsou snadno rozpoznatelné. Velkým nedostatkem je ale chybný formát bibliografických referencí, které nejen že nejsou dle normy, ale ani neobsahují povinné údaje. Navíc, seznam bibliografických referencí je uveden v příloze, kam nepatří.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Kladně hodnotím sebereflexi, kterou student uvádí na konci práce. S jeho hodnocením výsledků projektu souhlasím.

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.

Zadání se bohužel ukázalo jako nevhodné pro studenta na úrovni bakalářského studia. Úspěšné splnění takového rozsáhlého úkolu vyžaduje expertní znalost počítačového vidění. Student byl vržen do naprosto neznámého terénu. Jeho práce se redukovala na zápas s rozličnými nesourodými systémy, jejichž volba a redundance je navíc sporná. Pro velký rozsah nutných technických prací pravděpodobně nezbyl dostatek času na zpracování textu práce. Ta potom trpí základními nedostatky: Neposkytuje ani matematický popis řešených úloh, ani algoritmickou dokumentaci. Softwarovou dokumentaci poskytuje jen částečně. Podle mého názoru si z tohoto projektu student neodnáší porozumění alespoň některým problémům z oboru počítačového vidění. Je otázkou jaký pedagogický cíl byl splněn. Snad jen ten, že v inženýrském životě se nesmí věřit ani práci prezentované na nejprestižnější konferenci v oboru, dokud nejsou k dispozici kladné reference nezávislých autorů.

Otázky uvedené výše jsou míněny spíše jako zpětná vazba. Do technické diskuse mám ještě následující otázku:

Proč přesně nefungovala segmentace objektů na reálných obrázcích z Kinectu? Kvůli rozmazání zmíněném výše?

S přihlédnutím k tomu, že zadání bylo obtížné a na samotnou prezentaci pravděpodobně nezbyl čas, předloženou závěrečnou práci hodnotím klasifikačním stupněm **D - uspokojivě**.

Datum: 30.5.2022

Podpis: