



Posudek oponenta diplomové práce

Student: Bc. Kajetán Poliak

Název práce: Odhad pozice rukou během řízení automobilu

Předložená práce si klade za cíl trasovat pohyby rukou řidiče během řízení automobilu v nasnímané videosekvenci. Používá dva druhy dat, jednak data z běžné barevné kamery, jednak data doplněná o hloubkovou informaci, jaká poskytuje např. zařízení Kinect.

Práce je rozdělena na šest kapitol plus nečíslovaný úvod, závěr a dodatek. První kapitola obsahuje rešerši stávajících metod strojového učení zaměřených na segmentaci, detekci a odhad pozice rukou. Druhá kapitola se věnuje teorii neuronových sítí, zejména zpětnému šíření sítí a problému mizejícího gradientu. Třetí kapitola představuje metody segmentace obrazu a detekce objektu v obrazu. Odhad 2D pozice ruky je rozebírán ve čtvrté kapitole. Jsou zde také představeny existující datasety k odhadu pozice rukou a popis vytváření vlastních datových sad. Pátá kapitola navazuje odhadem prostorové pozice ruky. Poslední kapitola rozebírá aplikaci nastudovaných metod na data z trenážeru řízení.

Z předložené práce je zřejmé, že student věnoval mnoho úsilí splnění zadání. Oceňuji hlavně lidsky přístupný popis použitých hlubokých konvolučních neuronových sítí, jejich architektury a trénování.

Pokud mám práci něco vytknout, tak jsou to drobnosti v podobě několika neobratných formulací a překlepů. Oceňuji, že student alespoň použil automatickou kontrolu pravopisu, jak je z povahy chyb zřejmé.

Pro úplnost uvádím pouze připomínky, které nesouvisí s jazykem:

- Obrázek 1.1 - obdélník “detekce”: takhle zkraje práce není jasné, o detekci čeho jde.
- Oddíl 2.2: “... jak ilustruje obrázek 15.” V práci žádný obrázek 15 není. Myslí se tím obrázek 2.2?
- Oddíl 2.3: “ e_j je vektor s jedničkou na j -té pozici.” Co má vektor na ostatních pozicích?
- Oddíl 2.5: “Potom vážená suma $z = \sum_i w_i x_i + b$ bude mít gaussovské rozdělení se střední hodnotou nula a variací $\sqrt{501} \approx 22,4$.” Myslím si, že střední hodnota bude b a $\sqrt{501}$ bude standardní odchylka, nikoli variace.

- Oddíl 3.1.1: "... množiny A, B splňující $A \cap B = V \wedge A \cup B = \emptyset$..." To nelze pro neprázdný graf $G = (V, E)$ splnit. Má to být $A \cup B = V \wedge A \cap B = \emptyset$?
- Oddíl 3.1.1: " $\|F(i) - F(j)\|_2^2$ je euklidovská vzdálenost mezi $F(i)$ a $F(j)$." Obvykle se tak značí kvadrát euklidovské vzdálenosti. Euklidovská vzdálenost je $\|F(i) - F(j)\|_2$.
- Oddíl 3.2.1: "... definujeme dilataci B o S jako $B \oplus S = \bigcup_{b \in B} S_b$." V obou případech by měla být použita stejná značka, tedy pravděpodobně \oplus .

I přes výše uvedené připomínky jsem přesvědčen, že student zadání své diplomové práce splnil a předloženou práci navrhuji hodnotit známkou **A** (výborně).

Dotaz k obhajobě:

v práci není dostatečně vysvětlena motivace úlohy. Mohl by student k tomu něco říci?

V Praze dne 12. 5. 2021

Ing. Tomáš Suk, DSc.