

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Construction of a Neural Networks model for translation of recorded sign language
Jméno autora:	Tomáš Děd
Typ práce:	bakalářská práce
Fakulta:	Fakulta jaderná a fyzikálně inženýrská (FJFI)
Katedra:	Katedra matematiky
Oponent práce:	Ing. Adam Novozámský, Ph.D.
Pracoviště oponenta práce:	Ústav teorie informace a automatizace AV ČR, v.v.i.

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Práce svým zadáním, rozsahem i řešením splňuje spíše parametry diplomové práce.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání.</i>	
Zadání práce splněno bez výhrad.	

Zvolený postup řešení	vhodný s výhradami
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Pokud měla práce sloužit k praktickému vyzkoušení hlubokého učení, splnila práce svůj cíl. Pokud mělo jít o vytvoření nástroje, který bude dobře fungovat v praxi, myslím, že by bylo dobré použít novější technologie, než jednu z prvních sítí LeNet, která už je téměř ¼ století stará. Což je v našem oboru pravěk.	

Odborná úroveň	výborná
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Odbornou úrovní předbíhá studium a řadí se spíše k diplomovým pracím.	

Formální a jazyková úroveň	výborná
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Text práce je logicky členěn, psán srozumitelně a nenarazil jsem ani na překlepy, špatné číslování nebo jiné nedostatky. Jediné co bych vytknul je nadužívání některých slov: 62x however a 34x nevertheless. Rekord pak bylo použití slova however 3x v jediném odstavci.	

Výběr zdrojů, korektnost citací	průměrné
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce.</i>	
Citace jsou provedeny správně, mám ale výhradu k aktuálnosti citovaných materiálů. Překlad znakového jazyka je široce řešené a diskutované téma v Computer Vision komunitě a proto mi v citacích chybí SOTA články (např. transformer-based) než jsou ty uvedené do roku 2020. Pěkné přehledové články k tématu: https://doi.org/10.1007/s10209-023-00992-1 https://doi.org/10.1016/j.eswa.2022.118993 ; https://doi.org/10.3390/electronics12122678	

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce.

Rešerše v první kapitole mi přijde nedostatečná - (viz předchozí komentář). V druhé kapitole jsou pěkně shrnuty teoretické základy strojového učení z různých zdrojů. Nejsilnější část práce je součástí kapitoly 4, kde se autor věnuje ladění hyperparametrů sítě. Tato kapitola je ukončená podkapitolou 4.7 – Comparative Analysis in State of the Art. Kterou shledávám, na druhou stranu, jako nejslabší část práce:

1) Srovnávací metody nejsou SOTA – 2016, 2017, 2019 a 2020

2) Table 4.25 – Autor srovnává jablka a hrušky. Výsledky v této tabulce jsou pouze převzaté výsledky citovaných autorů na zcela jiných datech s jiným počtem tříd znaků a s jiným zastoupením v daných třídách. Proto nemá tato tabulka a dané srovnání vypovídající hodnotu. Pokud takové srovnání dělám, musím mít zaručeno, že trénování a testování probíhalo na identických datech pro všechny metody. Pokud toto nemám zaručeno. Je potřeba získat od autorů danou implementaci, popřípadě si ji sám implementovat podle článku.

3) Pokud jsem správně pochopil, jak autor pracuje s datasetem a jeho rozdělením na trénovací, validační a testovací množinu. Tak se v tabulce 4.18 testovací množina = validační. Tím, že validační množina vznikla pouhým rozdělením datasetu na trénování a validaci, tak jsou znaky od stejného člověka v trénovací i testovací množině daných tříd. To velmi zvyšuje úspěšnost pro testovací data. Autor sice vytvořil ještě jednu množinu dat od člověka, kterého naučený model neviděl. Bohužel dataset o jednom člověku je příliš malý.

Škoda, že autor nepoužil nějaký volně stažitelný dataset, kde by už měl relevantní porovnání se SOTA algoritmy.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Celkově se mi práce líbila. Z textu jde znát, že autor problematice rozumí a umí ji dobře vysvětlit. Problémům v rešerši a vyhodnocení metod, které jsem v posudku sepsal, by se autor v navazujících pracích mohl jednoduše vyhnout a proto je nevidím jako zásadní. Přece jen se jedná o bakalářskou práci.

Zde jsou nějaké další otázky a podněty:

1) Zkoušel jste Váš model na nějakém datasetu používaném v literatuře?

2) ASL má přes 5000 různých znaků – je Vámi navržený model schopen takové generalizace? Např. obrázek 4.11 – pokud model vidí jen oprahovaný snímek s hranami, tak se divím, že to u této dvojice znaků nesehává více.

3) Byli lidé z Vašeho datasetu „native signers“, „second language signers“ nebo lidé, kteří znakovku neznají a snažili se jen napodobit figuranta na displeji?

4) „switch hands throughout the process to account for both left- and right-hand sided signers“ – nestačilo by jen obraz osově převrátit?

5) „the colours typically carry much less information than the hand position and shape itself“ – Doporučil bych nějakou transformaci COLOUR >>one-channel, při které by pixely s barvou ruky byly dominantní a naopak by bylo potlačeno pozadí.

6) „The classification performance did not vary much, nevertheless with higher batch size the models seemed to overfit slightly.“ – Nepřijde mi, že by to z tabulky 4.14 nějak plynulo.

7) Pokud je v tabulkách používáno zvýraznění řádku, pro zdůraznění nejlepší metody nebo nastavení dané metody, tak nechápu, proč je třeba v tabulce 4.3 zvýrazněn poslední řádek. Podobně jsem na tom i u jiných tabulek.

8) Doporučoval bych změnit typ/velikost písma pro nadpisy tabulek a obrázků - takhle při tak dlouhém popisu se člověk ztrácí, jestli je ještě u nadpisu nebo v samotném textu práce (viz např. tabulka 4.9).

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

Datum: 8.8.2023

Podpis: