POSUDEK OPONENTA DISERTAČNÍ PRÁCE

Název práce: **Vývoj metod rozpoznání českého znakového jazyka metodami strojového učení**

Autor: **Ing. Jan Šnajder**
Vysoké učení technické v Brně
Fakulta strojního inženýrství
Ústav mechaniky těles, mechatroniky a biomechaniky

Oponent: **Ing. Milan Chlada, Ph.D.**
Ústav termomechaniky AVČR, v.v.i.

Téma a struktura práce

Disertační práce se zabývá vývojem metod strojového učení pro rozpoznávání a překlad českého znakového jazyka. Je adekvátně a logicky členěna. Prvotně formulované cíle jsou po pečlivé rešerši problematiky upřesněny a odpovídají současnému stavu poznání a dostupnosti datových sad. Klíčové kapitoly 5 a 6 se pak věnují samotnému řešení dvou hlavních úloh, konkrétně rozpoznávání a klasifikaci české znakové abecedy, resp. překladu vymezené komunikační oblasti českého znakového jazyka. V závěrečné části jsou přehledně shrnuty dosažené výsledky a přínosy práce spolu s možnými směry dalšího výzkumu.

Splnění cílů práce

Na základě použitých postupů a uváděných výsledků implementovaných metod a modelů lze konstatovat, že vytyčené cíle byly splněny. Silnou stránkou práce je především autorova pečlivá tvorba datasetů, což je klíčovým krokem jakékoliv aplikace strojového učení. Vytvořené vlastní datové sady jsou unikátní nejen svým rozsahem, ale i kvalitou zpracování. Vznikl např. rozsáhlý soubor precizně separovaného a anotovaného videomateriálu předpovědí počasí z on-line archivu České televize, jehož využitelnost přesahuje rámec této disertace a může sloužit i pro další výzkum.

Velmi přínosný je inovativní přístup využívající relativní souřadnice klíčových bodů pózy jako vstup modelů pro klasifikaci znakovaného obsahu. V přesnosti překonává klasicky používané metody založené na přímém zpracování obrazu konvolučními sítěmi a zároveň poskytuje nezávislost na prostředí snímání. Navržené metody jsou tak lépe aplikovatelné, což dokládá vyvinutá mobilní aplikace pro překlad české znakové abecedy, která ověřuje funkčnost a použitelnost navrženého řešení v reálném čase. Tento výstup lze vnímat jako důležitý „prototyp“, jenž může sloužit jak k výuce znakové abecedy a rovněž jako základ pro složitější aplikace.

K dalším originálním výsledkům patří vytvoření dosud neexistující datové sady diakritiky české znakové abecedy a její využití v kombinaci s rekurentními neuronovými sítěmi. Tento příspěvek reflektuje specifikum českého znakového jazyka a představuje významný krok směrem k jeho komplexnímu strojovému zpracování.

Z hlediska problematiky samotného strojového učení považuji za nejvýznamnější část práce zaměřenou na návrh a ověření modelu pro překlad českého znakového jazyka na základě architektury Transformer. Aplikace tohoto modelu mnohdy ztroskotává na nedostatku dat, či obtížném učení. Autor ale po počátečních nezdarech dokázal adaptovat tento moderní model na

podvzorkovaná data a dosáhl srovnatelných výsledků jako v existujících publikacích, avšak pro mnohem delší sekvence.

Celkově lze říci, že autor postupoval v celé práci velmi systematicky, pečlivě analyzoval příčiny případných neúspěchů a hledal cesty k jejich překonání. Díky této důslednosti a schopnosti kritické reflexe působí dosažené výsledky robustně a důvěryhodně.

Poznámky a připomínky

Po formální stránce je disertační práce rovněž velmi zdařilá, jedná se o srozumitelný stabilizovaný text bez překlepů, doplněný o názorné ilustrace a fotografie. Přesto bych měl následující připomínky:

První ohledně české terminologie. Autor používá „novodobé“ výrazy, jako např. „aktivační funkce sigmoid“, zatímco v českém matematickém prostředí je přirozenější označení „sigmoidální přenosová funkce“ či „sigmoida“. Podobně termín „bias“ je ponechán v anglické podobě, ačkoliv se v české literatuře o neuronových sítích tradičně používal výraz „práh“ nebo „prahová hodnota“. Podobně jako v biologii, ze které matematický model jednoduchého neuronu vychází. Toto je ale otázkou udržování tradičních českých termínů nebo přejímání anglických výrazů, nikoliv faktické korektnosti textu.

Další připomínku mám k vizuální stránce práce. Obrázky se schématy dílčích výpočetních jednotek (neuronů) jsou graficky zdařilé a přehledné, nicméně bych ocenil i globálnější pohled na celkovou architekturu propojení uvnitř modelů. Přehledné diagramy by čtenáři umožnily snáze pochopit nejen jednotlivé stavební bloky, ale i vidět celkovou složitost a strukturu modelů. V přehledových tabulkách pak autor používá veličinu „velikost“, kde by terminologicky korektnější a názornější bylo uvést přímo „počet neuronů vrstvy“.

Pro důkladnější posouzení originality autorovy práce ale postrádám detailnější dokumentaci implementačních částí. Zejména tam, kde uvádí vlastní pojetí modelu Transformer, v současné verzi prostředí TensorFlow nepodporovaného. Uvedení ukázek kódu nebo podrobnějších technických specifikací by přispělo k lepší reprodukovatelnosti výsledků a usnadnilo jejich využití dalšími výzkumníky. Vhodným řešením by bylo například zpřístupnit implementační detaily formou repozitáře (např. na platformě GitHub či GitLab), což je běžná praxe v oblasti strojového učení, umožňující širší komunitě ověřit a rozvíjet dosažené výsledky.

Za určitý limit lze rovněž považovat závislost na externích nástrojích (například MediaPipe) při extrakci příznaků. Přestože je tento přístup praktický a inovativní, do budoucna by bylo vhodné uvažovat o metodách, které by umožňovaly větší kontrolu nad procesem extrakce příznaků a nezávislost na třetích stranách.

Tyto připomínky však nijak zásadně nesnižují kvalitu a přínos disertační práce, spíše ukazují směry, kterými by se mohl autorův další výzkum ubírat.

Závěry a doporučení

Předložená disertační práce přináší originální a aplikovatelné výsledky v oblasti automatického rozpoznávání a překladu českého znakového jazyka. Její přínos je nesporný jak z hlediska teoretického (inovativní metodologie, nové datové sady, důkaz účinnosti využití klíčových bodů pózy), tak i praktického (mobilní aplikace, automatizace tvorby dat). Výsledky mají potenciál

významně posunout obor a poskytnout základ pro další aplikace ve vzdělávání, překladatelských technologiích i inkluzivních komunikačních nástrojích. Autor prokázal schopnost samostatné vědecké práce, orientaci v oboru strojového učení a zároveň přinesl výsledky, které mají mezinárodní relevanci.

Rozsah i zpracování předložených tezí disertační práce považuji rovněž za více než dostatečné. Umožňují čtenáři získat ucelený obraz o cílech, metodách i výsledcích samotné práce.

Na základě výše uvedeného **doporučuji** práci k obhajobě a v případě, že tato bude úspěšná, udělit autorovi titul Ph.D.

Dotazy pro obhajobu

1. Jaké praktické zkušenosti má autor s problémy mizejícího či explodujícího gradientu (uvedenými v teoretické části) při použití různých přenosových funkcí? Bylo v modelech potřeba testovat více jejich typů?

2. V pasáži o rekurentních neuronových sítích jsou zmíněny metody inicializace a regularizace vah. Lze tyto postupy (resp. zda je potřeba) aplikovat i na jiné architektury a souvisejí s metodou dropoutu, potažmo dobrou generalizací sítí obecně?

3. Autor v páté kapitole uvádí, že k přetrénování došlo již po dvaceti epochách, což se jeví jako poměrně brzký moment. Jak bylo přetrénování detekováno? Je možné, že to souvisí se způsobem výběru validačních dat, např. validací „jedním poněkud vymykajícím-se člověkem“? Testoval autor i pomocí jiných osob, popř. skupin?

V Praze dne 7.10.2025

Ing. Milan Chlada, Ph.D.