

Supervisor assessment of Master's Thesis

Student: Lukáč Peter, Bc.
Title: Speaker Verification without Feature Extraction (id 23746)
Supervisor: Mošner Ladislav, Ing., DCGM FIT BUT

1. Assignment comments

Zadání považuji za spíše obtížnější. Práce se zabývá replikací modelu RawNet, který extrahuje charakteristické vektory řečníků ("embedding") přímo z časového průběhu signálů. Současně přináší analýzu různých částí modelu a vliv jejich změn na výslednou přesnost systému pro verifikaci řečníků se snahou o zlepšení. Modifikace se týkají objektivní funkce a architektury sítě (reziduální bloky, squeeze & excitation, SincNet, hloubka sítě). Ze zmíněného tak vyplývá, že složitost práce tkvěla zejména v pochopení a aplikování několika technik využívaných v neuronových sítích odpovídajících současnému stavu poznání. Množství existujících implementací pak pomohlo zmírnit přílišnou náročnost. Výsledná práce splnila zadání a jsem s ní spokojen.

2. Literature usage

Student vycházel z doporučené literatury. Nad její rámec však samostatně dohledával odborné články autorů výchozí publikace pro diplomovou práci za účelem získání většího přehledu. To zahrnovalo i článek vylepšující RawNet prezentovaný na konferenci Interspeech v říjnu 2020 a z velké části korejsky psaný popis jedné z technik (alfa-FMS). Současně student dohledával literární prameny týkající se dalších využitých technik.

3. Assignment activity, consultation, communication

Diplomová práce začala aktivní prací v prvním semestru a vlastní iniciativou týkající se návrhu experimentů. Následovala výraznější pauza v polovině řešení. Druhý semestr však byl opět ve znamení aktivity a značného posunu práce. Student se dokázal velmi rychle zorientovat v problematice a začal tak se samotnou prací brzy. Iniciativně si zajistil výpočetní prostředky na Google Cloud pro první část práce. Kromě pauzy probíhaly konzultace každý týden. Na ně byl student připraven a prezentoval pokroky. Dohodnuté termíny byly dodržovány.

4. Assignment finalisation

Závěr práce stále doprovázela experimentální činnost přispívající k úplnosti DP. Ta s sebou ale přinesla hektičtější dokončování textu. Díky zvýšené intenzitě práce byl text předán v takovém předstihu, který umožňoval přečtení a předání zpětné vazby.

5. Publications, awards

Publikační činnost není vedoucím známa. Kód práce je však umístěn ve fakultním GIT repositáři s dobrým popisem. Domnívám se proto, že případné zveřejnění by bylo přímočaré.

6. Total assessment

very good (B)

Práce přinesla kvalitní analýzu modelu RawNet. Vliv jednotlivých komponent na výsledný systém může být zajímavý pro širší publikum (např. pro nebo i nad rámec skupiny Speech@FIT). Současně bylo v práci dosaženo zlepšení oproti původnímu systému a vyzkoušení fúze s jiným modelem založeným na extrakci příznaků. Student během řešení ukázal schopnost porozumění moderním metodám využívaným v rámci strojového učení a využití existujících dílčích řešení pro vytvoření funkčního výsledku. Souhrnně tak hodnotím jak technickou, tak textovou práci za nadprůměrnou.

In Brno 8 June 2021

Mošner Ladislav, Ing.
supervisor