

Posudek oponenta bakalářské práce

Student: Košina Simon
Téma: Robustní odšumování a dereverberace audia (id 21381)
Oponent: Skácel Miroslav, Ing., UPGM FIT VUT

- Náročnost zadání** **průměrně obtížné zadání**
Cílem této bakalářské práce bylo navrhnout a vytvořit model na odstraňování šumu z nahrávek letového provozu. Zadání je **průměrně obtížné**, vyžaduje nastudování technik odšumování, architekturu neuronových sítí a metod pro trénování modelů. Zadání je **adekvátní** pro bakalářskou práci.
- Splnění požadavků zadání** **zadání splněno**
Jednotlivé body zadání **byly splněny**.
- Rozsah technické zprávy** **je v obvyklém rozmezí**
Rozsah práce **splňuje** všechny požadavky.
- Prezentační úroveň předložené práce** **80 b. (B)**
Prezentační úroveň práce je **velmi dobrá**. Text je pochopitelný i pro čtenáře neznalého problematiky.
 - Jednotlivé kapitoly na sebe logicky **navazují**.
 - Všechny obrázky i tabulky jsou pečlivě **popsány**.
 - V práci je použita **jednotná notace**, což umožňuje snazší pochopení problematiky.
 - V textu je popis několika technik, které **nejsou** později použity ani se na ně neodkazuje. Př. hyperbolický tangens, softmax, STOI, MAE, atd. Toto **nepovažují** za chybu, ale může to odvádět pozornost čtenáře od řešení problematiky.
 - V textu **chybí** detailnější popis extrakce příznaků pro trénování sítě. Je zde zmíněna Fourierova transformace. Parametry extrakce jako šířka okna, překryv nebo váhovací funkce lze však zjistit pouze z příloženého zdrojového kódu.
 - Podobně **chybí** konkrétní hodnoty pro trénování model jako třeba použitý *learning rate* pro generátor a diskriminátor, velikost *batche*, nebo počet epoch pro všechny experimenty. Opět lze zjistit z příloženého kódu.
 - Výstup z modelu byl hodnocen metodou *Mean Opinion Score* (MOS), tedy metodou, kde posluchači hodnotí kvalitu nahrávek poslechem. V textu **není uvedeno**, kolik posluchačů se podílelo na hodnocení nebo třeba kolik nahrávek poslechem hodnotili.
- Formální úprava technické zprávy** **90 b. (A)**
Formální stránka práce je **velmi dobrá**. V práci je jen **pár typografických chyb**. Práce je psaná ve slovenštině. Gramatickou stránku tedy **nemohu** ohodnotit, protože nejsem rodilý mluvčí.
- Práce s literaturou** **95 b. (A)**
Literární zdroje jsou **vhodně** zvoleny vzhledem k zadání práce. Počet literárních zdrojů je **velmi rozsáhlý**.
- Realizační výstup** **80 b. (B)**
Student vytvořil model, který **umožňuje** potlačit šum v silně zašuměných nahrávkách komunikace mezi pilotem a řídicí věží. Vyzkoušel několik technik trénování a augmentace dat a vybral tu nejvhodnější kombinaci pro řešení daného problému. Pro ohodnocení výstupu z modelů by **bylo vhodné** použít i další metriky pro skórování kvality nahrávky, jako třeba poměr signálu k šumu nebo v textu zmíněné metriky PESQ a STOI.
- Využitelnost výsledků**
Vytvořený model **lze určitě využít** pro potlačení šumu pro jakoukoliv nahrávku, tedy nevztahuje se pouze na letovou komunikaci.
- Otázky k obhajobě**
Popište, jak probíhalo měření kvality nahrávek poslechem metodou Mean Opinion Score (MOS). Kolik posluchačů se zúčastnilo měření?
- Souhrnné hodnocení** **85 b. velmi dobře (B)**
Realizační výstup práce je **velmi dobrý**. Technická zpráva obsahuje **minimum** nedostatků jak po **strukturální** tak i po **formální** stránce. Student naimplementoval model a provedl řadu experimentů, kde našel ideální kombinaci dat a trénování pro **zlepšení** původní verze modelu. I přesto, že student nedosáhl výrazného vylepšení kvality u silně zašuměných nahrávek, hodnotím práci **velmi pozitivně**. U takto silně

zašuměných dat dochází k poškození řeči v takové míře, že nahrávku často nelze významně vylepšit. Navrhuji práci ohodnotit stupněm **B**.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 2. června 2022

Skácel Miroslav, Ing.
oponent