

Průběh obhajoby – A21N0127P VLADAŘ Lukáš

- Syntéza řeči (TTS - Text to speech)
- Neurální syntéza řeči
- Otázky řešené v rámci DP
- Přehled natrénovaných modelů
- Hodnocení natrénovaných modelů
- Objektivní míra MCD (Mel Cepstral Distortion)
- Poslechové testy MUSHRA
- Statistické vyhodnocení experimentů
- Výsledky experimentů
- Náměty na další výzkum

OTÁZKY:

Dělal jste si nějakou podrobnější analýzu u experimentů s neočekávaným výsledkem, kdy např. model trénovaný na větším množství dat získal nižší MUSHRA hodnocení? Např. použití 6 vs 1,5 hodiny dat v obr. 6.2 nebo 3 vs 1.5 u obr. 6.3, případné kolísání kvality v obrázcích 6.6/6.7? Něco, co se již do práce nedostalo nebo nebylo možné jednoduše vysvětlit? Například poslechem promptů s cílem identifikace příčiny nízkého hodnocení (výskyt nějakého nepřírozeného jevu?). Jsou výsledky MUSHRA testů v souladu s Vaším subjektivním hodnocením v těchto případech?

Mohla by být nižší kvalita modelu neprofesionálního řečníka "fine-tuněná" na modelu profesionálního řečníka způsobena odlišností charakteru použitých hlasů? Fine-tuning na multi-speaker modelu již dosahuje obecně lepších výsledků. To je nicméně v rozporu s profesionálním řečníkem (i multi-speaker model byl z profesionálů), kde jsou ale výsledky poměrně protichůdné.

Popište blíže krabicový graf, který jste použil v prezentaci.

Klasifikace – výborně

Datum obhajoby – 7.9.2023