Peer review 1. Project: **OSM: One-Shot Multi-speaker**

1. Project objectives are clear.

2. The baseline solution picked by the team is relevant to the problem. The original paper is frequently cited and has ready-to-go implementation which makes it a good pick for the project.

3. The report doesn't mention anything about evaluation. I would suggest taking a look at ESPnet toolkit that has many pre-trained solutions for TTS as well as utilities for evaluation.

4. The github repository doesn't contain any implementation yet so it is hard to guess which framework the team is going to use to make the project easily extendable. Though their report suggests that they're going to use "Real-Time-Voice-Cloning" as a starting point, extend it to make it modular, and finally add several more Speaker Encoders to the project.

5. The list of possible improvements

   a. The picture the team has in the repository could've been added to the report for more clarity

   b. The team described pros and cons of the "Real-Time-Voice-Cloning" implementation, but they didn't say anything about pros and cons of the methods in the repository

   c. Maybe it was worth adding some details about how baseline solutions work in the report to better illustrate the understanding of the problem

   d. The report is fine but I would recommend either deciphering abbreviations or using citations along with them. It is not clear what LDE and TDNN stand for from the report.

6. The team's repository is easy to follow and satisfies all the requirements mentioned in the assignment.