Immersivaudio: audio generation based on video features.

Michele Vitale
ist1111558

Daniele Avolio
ist1111559

Teodor Chakarov
ist1111601

I. INTRODUCTION

II. PROBLEM DESCRIPTION

III. PROPOSED SOLUTION

IV. ARCHITECTURE

V. DEPLOYMENT

The deployment of the system is done using Python 3.10.12 on Google Colab [1]. The first intention was to whole process run on a local machine, but it quickly became clear that this is not possible due to the high hardware requirements that audio generation models have. The next idea was to serve the audio generation module over an API, loading it on the machine provided at IST. This hypotesis got aswell discarded for lack of sufficient resources.
The front-end has been developed using the Gradio [2] library, which allows the fast development of web GUIs, especially for machine learning models.

VI. IMPLEMENTED FEATURES

VII. EVALUATION

VIII. CONCLUSIONS AND FUTURE WORK

REFERENCES

[1] https://colab.research.google.com/
[2] https://www.gradio.app/