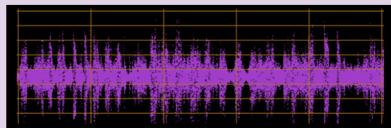


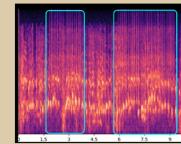
Target Sound Detection

Query: When did the keyboard sound happen in the audio?



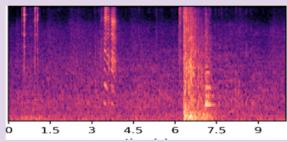
Call Target Sound Detection Model

Response: The keyboard sound happened in this audio from 1.9 to 3.8 seconds and from 5.9 to 9.3 seconds.



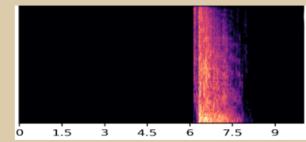
Sound Extraction

Query: I want to extract the explosive sound from the mixture



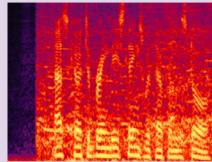
Call Sound Extract Model

Response: The extracted audio is saved on ["audio/786tfc3.wav"](#).



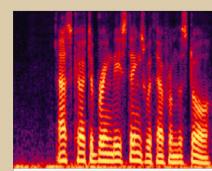
Speech Enhancement

Query: Please help me to denoise this audio.



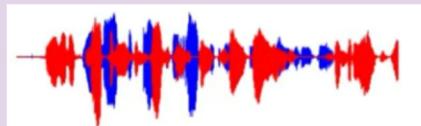
Call Speech Enhancement Model

Response: The denoised audio is saved on ["audio/34cffogh.wav"](#).



Speech Separation

Query: Can you help me to separate the mixture speech into two speaker speech ?



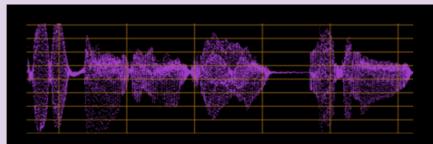
Call Speech Separation Model

Response: The mixture audio includes two speakers, the separated audios are saved on ["audio/fhb32il.wav"](#) and ["audio/iugp67w.wav"](#)



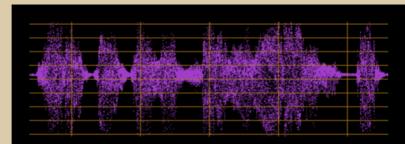
Speech Translation

Query: Translate the English speech into Spanish speech



Call Speech Translation Model

Response: The generated audio is saved on ["audio/83jkoo9.wav"](#)



Content: **it can be found in algeria lebanon portugal and spain**

Content: **se encuentra en argelia libano portugal y españa**