

# **AUDIO PROCESSING METHOD, MODEL TRAINING METHOD AND APPARATUSES**

## **ABSTRACT**

The present disclosure provides audio processing methods, model training methods and apparatuses,  
5 devices, and a storage medium. The method includes: inputting mixed audio data into an encoder of a  
pre-trained neural audio codec (NAC) to obtain a first embedding of the mixed audio data, where the  
mixed audio data is a mixture of multiple pieces of audio data from multiple acoustic sources, and the  
first embedding is a representation of an audio feature of the mixture; where the pre-trained NAC is  
trained with a set of sample mixed audio data for audio separation; performing audio separation on the  
10 first embedding with a pre-trained separator to obtain at least one second embedding, where the at least  
one second embedding corresponds to at least one of the multiple pieces of audio data one-by-one, each  
of the at least one second embedding is a representation of an audio feature of corresponding audio data,  
and the at least one second embedding is used for reconstruction of the at least one of the multiple pieces  
of audio data with a decoder of the pre-trained NAC.

15

**FIG. 3**