CLAIMS

What is claimed is:

1. A method for detecting audio, comprising:

acquiring audio file data;

determining attribute detection data corresponding to the audio file data, wherein the attribute detection data comprises at least two of: user rating data, classification probability data, and voiceprint feature data, the user rating data being indicative of a user rating, the classification probability data being indicative of a classification probability corresponding to a voice violation, and the voiceprint feature data being indicative of a voiceprint feature corresponding to the audio file data; and

generating a voice detection result corresponding to the audio file data by voice violation detection on the attribute detection data by a pre-trained fully connected network model.

2. The method according to claim 1, wherein determining the attribute detection data corresponding to the audio file data comprises:

acquiring at least two frames of audio time-domain information by slicing the audio file data;

acquiring magnitude spectrum feature data and the voiceprint feature data by feature extraction on the at least two frames of audio time-domain information;

generating feature vector data by splicing the magnitude spectrum feature data and the voiceprint feature data; and

acquiring the classification probability data as the attribute detection data by voice classification on the feature vector data by a pre-trained voice classification model.

3. The method according to claim 2, wherein acquiring the magnitude spectrum feature data by the feature extraction on the at least two frames of audio time-domain information comprises:

acquiring audio frequency-domain information by frequency-domain transformation on the at least two frames of audio time-domain information; and

acquiring the magnitude spectrum feature data by magnitude spectrum feature extraction based on the audio frequency-domain information.

4. The method according to claim 1, wherein determining the attribute detection data corresponding to the audio file data comprises:

acquiring at least two frames of audio time-domain information by slicing the audio file data;

acquiring audio frequency-domain information by frequency-domain transformation on the at least two frames of audio time-domain information;

acquiring first fixed-length data by averaging the audio frequency-domain information; and

acquiring voiceprint feature data as the attribute detection data by voiceprint feature extraction by a pre-trained neural network model based on the first fixed-length data.

5. The method according to claim 4, further comprising:

acquiring audio file data to be trained from a predetermined training set;

acquiring frame time-domain information by slicing the audio file data to be trained by using a predetermined moving window;

acquiring frame frequency-domain information by frequency-domain transformation on the frame time-domain information;

acquiring second fixed-length data by averaging the frame frequency-domain information; and

acquiring the neural network model by training according to a predetermined neural network algorithm based on the second fixed-length data and label data corresponding to the audio file data.

6. The method according to claim 1, further comprising:

acquiring history behavior data of a target user, wherein the history behavior data comprises at least one of: history login data, user consumption behavior data, violation history data, and recharge history data; and

acquiring the user rating data as the attribute detection data based on the history behavior data.

7. The method according to any one of claims 1 to 6, wherein generating the voice detection result corresponding to the audio file data by the voice violation detection on the attribute detection data by the pre-trained fully connected network model comprises:

inputting the attribute detection data to the fully connected network model for detection; and

determining a detection result output by the fully connected network model as the voice

detection result.

8. The method according to claim 7, further comprising:

determining that the audio file data comprises violation voice data in response to the voice detection result being a voice violation detection result; and

prohibiting the transmission or playback of the violation voice data, or shielding a voice input from a user corresponding to the violation voice data.

9. The method according to claim 2 or 3, further comprising:

acquiring frame time-domain information by slicing acquired audio file data to be trained by using a predetermined moving window;

acquiring magnitude spectrum feature training data and voiceprint feature training data by feature extraction on the frame time-domain information, wherein the feature extraction comprises: magnitude spectrum feature extraction and voiceprint feature extraction;

acquiring third fixed-length data by averaging the magnitude spectrum feature training data;

generating feature vector training data by splicing the magnitude spectrum feature training data and the voiceprint feature training data; and

acquiring the voice classification model by training the third fixed-length data and the feature vector training data.

10. The method according to any one of claims 1 to 6, further comprising:

acquiring attribute detection data to be trained; and

acquiring a fully connected network model by training the attribute detection data to be trained.

11. An apparatus for detecting audio, comprising:

an audio file data acquiring module, configured to acquire audio file data;

an attribute detection data determining module, configured to determine attribute detection data corresponding to the audio file data, wherein the attribute detection data comprises at least two of: user rating data, classification probability data, and voiceprint feature data, the user rating data being indicative of a user rating, the classification probability data being indicative of a classification probability corresponding to a voice violation, and the voiceprint feature data being indicative of a voiceprint feature corresponding to the audio file

data; and

a voice violation detecting module, configured to generate a voice detection result corresponding to the audio file data by voice violation detection on the attribute detection data by a pre-trained fully connected network model.

12. A device, comprising:

a processor; and

a memory storing at least one instruction therein;

wherein the at least one instruction, when loaded and executed by the processor, causes the device to perform the method for detecting audio as defined in any one of claims 1 to 10.

13. A computer-readable storage medium storing at least one instruction therein, wherein the at least one instruction, when loaded and executed by a processor of a device, causes the device to perform the method for detecting audio as defined in any one of claims 1 to 10.