TITLE OF THE INVENTION: LANGUAGE PROCESSING MODEL TRAINING METHOD, APPARATUS, ELECTRONIC DEVICE, AND READABLE STORAGE MEDIUM

ABSTRACT

The embodiments of the present disclosure propose a language processing model training method, an apparatus, an electronic device, and a readable storage medium, relating to the field of speech synthesis technology. The method includes: acquiring training text data and ground-truth speech token data and inputting them into a prediction module of a language processing model to obtain predicted speech token data; extracting style features from both the ground-truth and predicted speech token data respectively; constructing a total loss function based on the ground-truth speech token data, the predicted speech token data, the ground-truth style features, and the predicted style features; and then performing optimization training on the prediction module according to the total loss function. By incorporating the style feature loss into the total loss function, the language processing model is enabled to learn style information, thereby allowing the language processing model to more accurately reproduce speech styles. This improves the performance of the language processing model in handling speech style information. Since the speech processing model is a key component of a speech synthesis model, the performance of a large speech synthesis model in handling speech style information can thus be enhanced.

Fig. 1