

Streaming End-to-End Speech Recognition Method, Apparatus and Electronic Device

Abstract

The embodiments of the present application disclose a method, an apparatus, and an electronic device for streaming end-to-end speech recognition. The method includes: extracting and encoding speech acoustic features of a received voice stream in units of frames; performing block processing, and predicting a number of activation points included in a same block that need to be encoded and outputted; determining position(s) of activation point(s) that need(s) to be decoded and outputted according to a prediction result, to a decoder to perform decoding at the position(s) of the activation point(s) and output a recognition result. Through the embodiments of the present application, the robustness of a streaming end-to-end speech recognition system to noise can be improved, thereby improving the performance and the accuracy of the system.

Figure 1