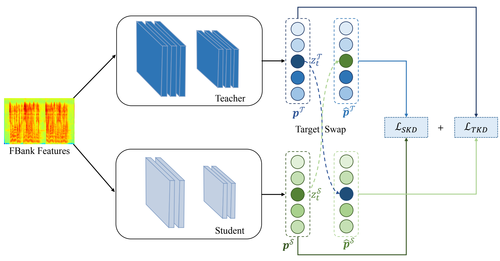
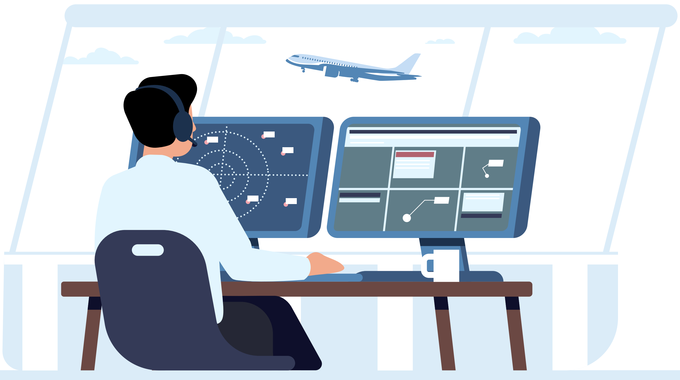
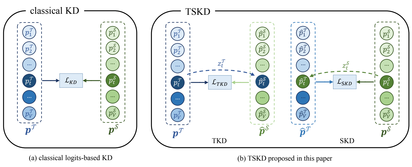
A Lightweight Speech Recognition Method with target-swap knowledge distillation for Mandarin air traffic control communications



Miscommunications between air traffic controllers (ATCOs) and pilots in air traffic control (ATC) may lead to catastrophic aviation accidents. To allow ATCOs and pilots sufficient time to respond instantly and effectively, the automatic speech recognition (ASR) systems for ATC must have both superior recognition performance and low transcription latency. This paper introduces knowledge distillation into the ASR for Mandarin ATC communications to enhance the generalization performance of the light model.



We propose a simple yet effective lightweight strategy, named Target-Swap Knowledge Distillation (TSKD), which swaps the logit output of the teacher and student models for the target class. It can mitigate the potential overconfidence of the teacher model regarding the target class and enable the student model to concentrate on the distillation of knowledge from non-target classes.



(a) classical logits-based KD

(b) TSKD proposed in this paper

The experimental results reveal that the generated lightweight ASR model achieves a balance between recognition accuracy and transcription latency.