

Speech Recognition: Conformer

Model Overview

Conformer-CTC (around 120M parameters) is trained on ASRSet with over 1000 hours of English(en-GB) speech. The model transcribes speech in lower case English alphabet along with spaces and apostrophes.

Model Architecture

Conformer-CTC [1] model is a non-autoregressive variant of Conformer model [1] for Automatic Speech Recognition which uses CTC loss/decoding instead of Transducer. You may find more info on the detail of this model here: [Conformer-CTC Model](#).

Training

The model was trained on various proprietary and open-source datasets. These datasets include variety of accents, domain specific data for various domains, spontaneous speech and dialogue, all of which contribute to the model's accuracy. This model delivers WER that is better than or comparable to popular alternate Speech to Text solutions for a range of domains and use cases.

How to Use this Model

To use this model, we can use Riva Skills Quick start guide, it is a starting point to try out Riva models. Information regarding Quick start guide can be found: [here](#). To use Riva Speech ASR service using this model, this [document](#) has all the necessary information.

Input

Audio sample that is to be transcribed

Output

This model provides transcribed speech as a string for a given audio sample.

References

[1] [Conformer: Convolution-augmented Transformer for Speech Recognition](#)

License

By downloading and using the models and resources packaged with Riva Conversational AI, you would be accepting the terms of the Riva [license](#)