1. Make a brief introduction about a variant of Transformer. (2 pts)
   Conformer is a variant of Transformer which introduces convolutional neural networks layers in between the self-attention layers. This new feature improves the performance on speech recognition tacks by the nature that transformer is good at capturing global properties while CNN is good at capturing local properties, which means they can complement each other.

2. Briefly explain why adding convolutional layers to Transformer can boost performance. (2 pts)
   As question one suggests, transformer is good at capturing global properties while CNN is good at capturing local properties, which means they can complement each other. As a result, the model combined with transformer and convolutional layers is capable of featuring more informative knowledge of the data.