

Speech and Language Recognition with Low-rank Adaptation of Pretrained Models

Amrutha Prasad



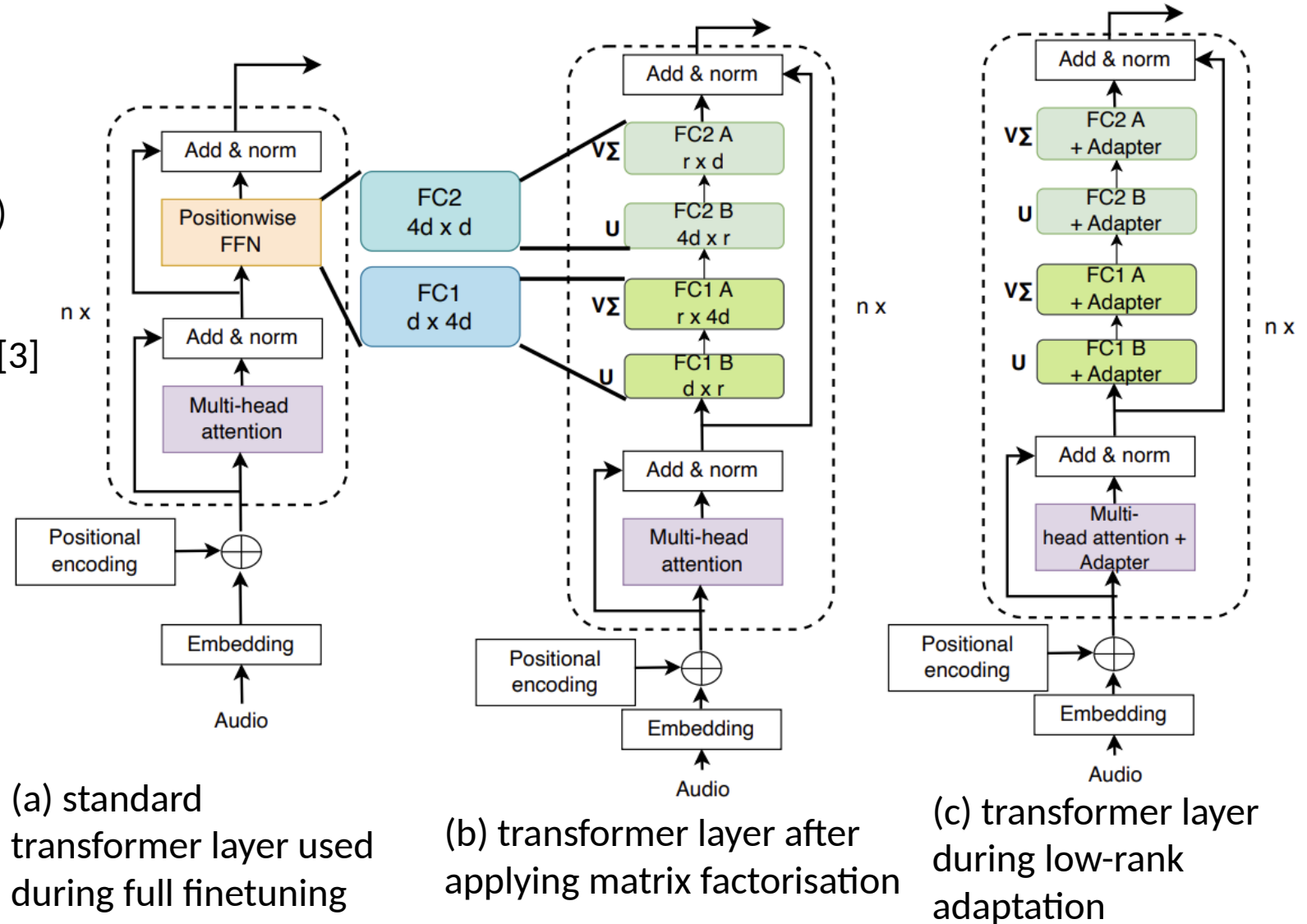
Introduction

- ♦ Current approach: Fine-tune large pretrained models to achieve state-of-the-art performance for downstream tasks
- ♦ Main challenge:
 - ♦ In resource constrained environment during fine-tuning and inference
- ♦ Research focuses on reducing the parameter budget
 - ❖ Freeze the weight matrix during training and learn the updates through an additional set of parameters (adapters)
 - ⊗ Drawback: During inference, no reduction in the computational requirements
- ♦ We apply Singular Value Decomposition (SVD) to the linear layer weights [1]
 - ☑ Parameter reduction during both training and inference

Low-rank Adaptation

- Tasks: Automatic Speech Recognition (ASR) and Language Identification (LID)
- Pre-trained models: XLS-R[2] and Whisper[3]

We observed no performance degradation with a parameter reduction of 28% for XLS-R and 23% for Whisper



[2] A. Babu et al., "XLS-R: Self-supervised cross-lingual speech representation learning at scale," in Proc. of Interspeech, 2022, pp. 2278–2282

[3] A. Radford et al., "Robust speech recognition via large-scale weak supervision," in International Conference on Machine Learning. PMLR, 2023, pp. 28 492–28 518.

Thank you!
