DnA: Improve Few-Shot Transfer Learning with Low-Rank Decompose and Align

Ziyu Jiang\textsuperscript{1,2†}, Tianlong Chen\textsuperscript{3}, Xuxi Chen\textsuperscript{3}, Yu Cheng\textsuperscript{1}, Luowei Zhou\textsuperscript{1}, Lu Yuan\textsuperscript{1}, Ahmed Awadallah\textsuperscript{1}, and Zhangyang Wang\textsuperscript{3‡}

\textsuperscript{1} Microsoft Corporation
\textsuperscript{2} Texas A&M University
\textsuperscript{3} University of Texas at Austin

This supplementary material contains the following details that we could not include in the main paper due to space restrictions.

– (Sec. 1) Details of the computing infrastructure.
– (Sec. 2) Comparison with LoRA \cite{1}.

1 Details of computing infrastructure

Our codes are based on Pytorch \cite{2}, and all models are trained with NVIDIA A100 Tensor Core GPU.

2 Comparison with LoRA \cite{1}

LoRA, a closely related work, shows that formalizing the weight changing as a low-rank matrix can also improve the fine-tuning performance. Therefore, we compare with Align+LoRA to verify the effectiveness of the proposed Decomposition method. As illustrated at Table 1, by applying the LoRA with Align, the performance could improve by 0.4%. However, it is still weaker than the proposed DnA with an obvious margin of 0.5%.

Table 1: Compare with LoRA \cite{1} in terms of the 5-shot performance on iNaturalist-1k.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Align</td>
<td>47.4</td>
</tr>
<tr>
<td>Align+LoRA</td>
<td>47.8</td>
</tr>
<tr>
<td>DnA (Ours)</td>
<td>48.3</td>
</tr>
</tbody>
</table>

\textsuperscript{†} Work done during an internship at Microsoft Corporation
\textsuperscript{‡} Correspondence to: Zhangyang Wang (atlaswang@utexas.edu)
References
