Contrastive Prototype Learning with Augmented Embeddings for Few-Shot Learning (Supplementary material)

Yizhao Gao[1] **Nanyi Fei**[2] **Guangzhen Liu**[2] **Zhiwu Lu**[*1] **Tao Xiang**[3]

[1]Gaoling School of Artificial Intelligence, Renmin University of China, Beijing, China
[2]School of Information, Renmin University of China, Beijing, China
[3]University of Surrey, Guildford, Surrey, United Kingdom

In this document, we provide the results with alternative shuffling order used in our CPLAE model, and also show more visualization results to demonstrate the effectiveness of our CPLAE model.

Table 1: Comparative results by adopting different shuffling orders for our full CPLAE model on *mini*ImageNet (with Conv4-64 as the backbone).

Method	5-way 1-shot	5-way 5-shot
ProtoNet†	52.79 ± 0.45	71.23 ± 0.36
ProtoNet†+AE	55.89 ± 0.43	73.43 ± 0.35
CPLAE (no shuffling)	56.04 ± 0.44	73.75 ± 0.35
CPLAE (shuffling)	56.83 ± 0.44	$\mathbf{74.31 \pm 0.34}$
CPLAE (shuffling*)	$\mathbf{57.14 \pm 0.44}$	74.16 ± 0.34

1 RESULTS WITH ALTERNATIVE SHUFFLING ORDER

As we have stated in Section 3.3 of the main paper, we shuffle the order of the three augmentations and obtain a shuffled augmented embedding $\hat{f}_\phi(x_i)$ of each query sample $x_i \in \mathcal{Q}$ (see Eq. (5) in the main paper) for our Contrastive Prototype Learning (CPL). In our experiments, We have found that in the shuffled concatenation, once the original image's embedding $f_\phi(x_i)$ remains in the first place of $\hat{f}_\phi(xi)$, how exactly the other three embeddings are shuffled makes little difference. To demonstrate this more clearly, we thus adopt a different shuffling order to obtain an alternative shuffled augmented embedding for each $x_i \in \mathcal{Q}$:

$$\hat{f}^*_\phi(x_i) = A\left(f_\phi(x_i), f^{(3)}_\phi(x_i), f^{(1)}_\phi(x_i), f^{(2)}_\phi(x_i)\right). \quad (1)$$

Let 'shuffling' and 'shuffling*' denote the shuffling order used in the main paper and that in Eq. (1), respectively. We list the comparative results on the *mini*ImageNet dataset in Table 1, where Conv4-64 is adopted as the feature extractor. It can be observed from Table 1 that: (1) Both CPLAE (shuffling) and CPLAE (shuffling*) achieve performance improvements over CPLAE (no shuffling), which demonstrates the importance of the shuffling operation for our CPL. (2) The results obtained by CPLAE (shuffling) and CPLAE (shuffling*) are of little difference, indicating that the shuffling order of the augmented embeddings has little influence on our CPLAE model.

2 MORE VISUALIZATION RESULTS

Similar to Section 4.4 of the main paper, we provide more visualization results in Figure 1 and Figure 2. Concretely, for each row in Figure 1, the first four columns present the data distributions of the original meta-test episode and its three extended episodes, while the last column presents the visualization results using the integrated embeddings. For each row in Figure 2, we visualize the data distributions of the same meta-test episode obtained by five FSL models: ProtoNet†, ProtoNet†+AE, CPLAE (w/o Proto, w/o Proj, no shuffling), CPLAE (no shuffling), and our CPLAE. From these two figures, we can observe that: (1) The integrated embedding has the best data clustering structure, indicating that the implicit alignment of the original episode and its three extended ones is crucial for FSL with only few shots. (2) Our CPLAE model results in better cluster structure than the four simplified versions. Overall, similar observations can be made as in Section 4.4 of the main paper, validating the effectiveness of our CPLAE for standard FSL.

Acknowledgments

This work was supported in part by National Natural Science Foundation of China (61976220 and 61832017), Beijing Outstanding Young Scientist Program (BJJWZYJH012019100020098), and Open Project Program Foundation of Key Laboratory of Opto-Electronics Information Processing, CAS (OEIP-O-202006).

[*]Corresponding author.

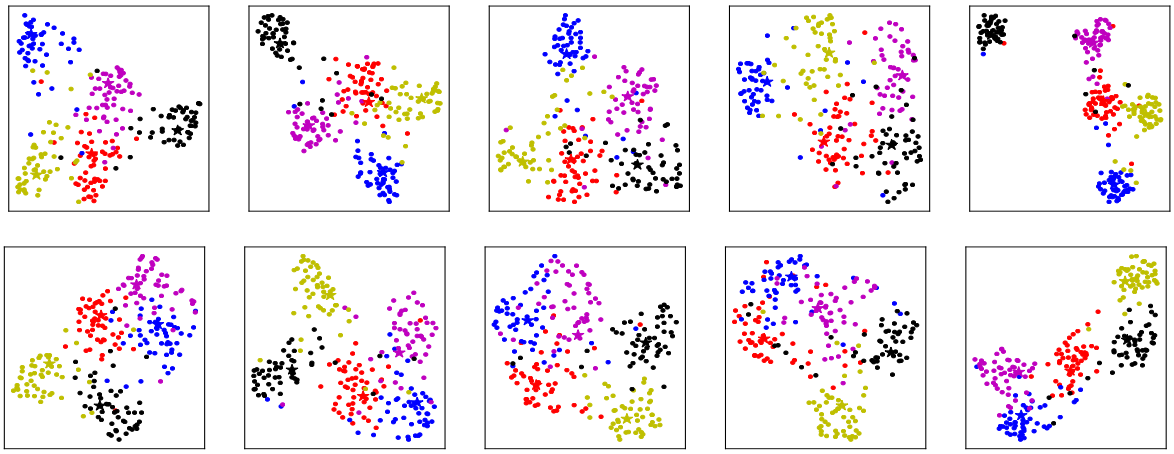

Figure 1: Data distribution visualization results of two sets of episodes (one row for one set of five episodes) on *mini*ImageNet using the UMAP algorithm [McInnes et al., 2018]. For each row, the first four columns present the results of one original meta-test episode and its three extended ones, respectively. The last column presents the visualizations using integrated embeddings. The 5-way 5-shot setting is considered, with Conv4-64 as the feature extractor.

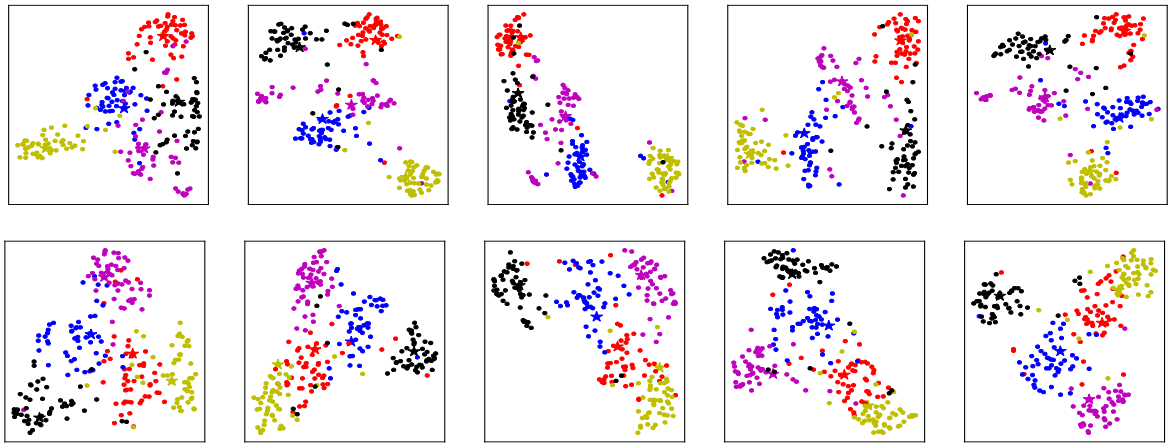

Figure 2: Visualizations of data distributions of two meta-test episodes (one row for one episode) from *mini*ImageNet using the UMAP algorithm [McInnes et al., 2018] for five FSL models (from left to right): ProtoNet†, ProtoNet†+AE, CPLAE (w/o Proto, w/o Proj, no shuffling), CPLAE (no shuffling), and our CPLAE. Each row presents one test episode. The 5-way 5-shot setting is considered, with Conv4-64 as the backbone.

References

Leland McInnes, John Healy, and James Melville. UMAP: Uniform manifold approximation and projection for dimension reduction. *preprint arXiv:1802.03426*, 2018.