Bridging Domain Gaps for Fine-Grained Moth Classification Through Expert-Informed Adaptation and Foundation Model Priors

Supplementary Material

8. Embedding-Based Clustering for Train/Test Splitting

To obtain train/test splits from AMI data that are decoupled, we first embed each AMI image using a ResNet-50 network pretrained on ImageNet-1K; this places visually similar images near one another in the embedding space. For every moth class, we then perform agglomerative hierarchical clustering with cosine distance, setting the number of clusters to $\max(K_{\text{min}}, \sqrt{N})$, where N is the number of images in the class and $K_{\text{min}} = 5$. After clustering, we randomly permute the cluster order and walk through this shuffled list, adding images to the test split until the specified amount of test examples are collected. This selective part is illustrated in Figure 4.

Figure 4. List of image embeddings with each showing its cluster number. Cluster groups have been shuffled. Each element are coloured and labelled by a unique cluster ID. The test set is partitioned as the first images in this list, with the remainder as train.

This approach enforces semantic separation between the train/test split, reducing bias and providing a realistic assessment of model generalisation.

9. Dataset Composition

Target-domain Mix (%)	Target-domain (AMI) Contribution	Source-domain (GBIF) Contribution	Total Dataset Size
0%	0	18573	18573
1%	187	18573	18760
5%	997	18573	19550
10%	2063	18573	20636
20%	4643	18573	23216
25%	6191	18573	24764
33%	9147	18573	27720
50%	10100	10100	20200

Table 2. Statistics of each domain-mixed training dataset at various amounts of target-domain percentages. Contributions from the target-domain (AMI) images and the source domain (GBIF) images vary to meet the desired domain mix fraction.