Table 3: Per-class accuracy (%) comparisons over unseen classes for zero-shot learning between our method and baselines on five datasets. Parts of the results are from (**?**).

Methods	**SUN**	**CUB**	**AWA1**	**AWA2**	**aPY**
DEVISE (**?**)	56.5	52.0	54.2	59.7	39.8
ALE (**?**)	58.1	54.9	59.9	62.5	39.7
SYNC (**?**)	56.3	55.6	54.0	46.6	23.9
SAE (**?**)	40.3	33.3	53.0	54.1	8.3
RN (**?**)	-	55.6	68.2	64.2	-
GAFE (**?**)	62.2	52.6	67.9	67.4	**44.3**
APNet(ours)	**62.3**	**57.7**	68.0	**68.0**	41.3

DEM (**?**) achieved similar performance compared to our APNet on AWA1. However, the generalization ability of DEM, which is measured by the performance in the setting of generalized zero-shot learning, has space for improvements. This is probably because DEM treats visual space as the embedding space for nearest neighbor search. Comparatively, we mainly do transformation over the semantic embeddings by propagation on a shared graph of training and test classes and use a parametric KNN classifier for a semantic-visual joint similarity prediction. The visual space used in DEM has high bias and non-i.i.d. problems and might be difficult to generalize when considering predictions over both seen classes and unseen classes. GAFE (**?**) achieved better performance on aPY dataset but also suffers from the generalization ability problem. Similar to DEM, they also focused on the visual space and proposed a reconstruction regularizer on the visual feature representations. In remains open if the regularizer can generalize on the unseen classes without their visual feature representations during training stage.

Ablations and Variants of APNet.

We did ablation study and developed some variants of APNet in Table 4. In the ablation of propagation ($\times$ in "Graph Hierarchy"), we skip the propagation step and directly use the node feature representations for similarity comparison and prediction. Propagation can achieve above 1 point improvement on H under different training strategies.

Two variants of APNet are developed by adjusting the training strategy to traditional minibatch training (denoted by $\times$ in "Meta training") and replacing the graph hierarchy used in propagation with the hierarchy defined in WordNet. The meta-learning style training strategy we used can bring $2 \sim 3\%$ accuracy improvements on unseen accuracy and $2 \sim 3$ points improvement on H while keeping the rest components the same, which verifies that training on different tasks in every iteration can get better model generality. When the hierarchy is defined by WordNet, the propagation graph is assumed to be fully connected and the value of the adjacency matrix for node i and node j (A_{ij}) is defined as $1/d_{ij}$, where d_{ij} is the number of hops between node i and node j on WordNet. Predefined graph hierarchy can bring improvements but is sensitive to the training strategy. The reason is probably that the exact/optimal formulation for generating the adjacency matrix based on the distance between nodes is unclear even though intuitively closer nodes should have higher values for its adjacency matrix. Dedicated designs on how to integrate the hierarchy information more effectively have the potential to boost the performance.

Table 4: Ablations and performance comparisons on variants of APNet on AWA2.

Meta Training	**Graph Hierarchy**	**S**	**U**	**H**
$\times$	$\times$	82.8	50.2	62.5
$\times$	WordNet	83.9	50.4	62.9
$\times$	Learned	81.2	52.6	63.9
✓	$\times$	**84.4**	53.4	65.4
✓	WordNet	82.6	**55.7**	**66.6**
✓	Learned	83.9	54.8	**66.4**

6 Acknowledgments

This research was funded by the Australian Government through the Australian Research Council (ARC) under grants 1) LP160100630 partnership with Australia Government Department of Health and 2) LP180100654 partnership with KS Computer. We also acknowledge the support of NVIDIA Corporation and Google Cloud with the donation of GPUs and computation credits.

Incidunt quos totam, unde consectetur vel, rem tenetur corrupti dignissimos commodi blanditiis modi earum incidunt, impedit perferendis reiciendis excepturi iure laboriosam accusamus?Porro doloribus vero molestiae eius adipisci veritatis illo, mollitia culpa reprehenderit debitis officiis corporis impedit nam, at saepe alias voluptatem voluptate veniam eaque asperiores fuga corrupti voluptas harum, perferendis nesciunt accusamus modi unde.Itaque repellendus quia illum adipisci fugit velit nihil, voluptatem est adipisci tempore vero nobis molestiae modi minus maiores, nobis ea non ratione atque iste recusandae tenetur?Cumque voluptatem officiis porro blanditiis praesentium doloremque corporis id culpa, fugit facere ab nam iure explicabo, sit sunt consectetur consequatur voluptatem, nulla pariatur temporibus maiores possimus adipisci a eius, delectus ad voluptate sapiente dicta ipsam amet dolorem.Facilis explicabo blanditiis inventore deserunt delectus reprehenderit, magnam nemo adipisci ut dolorum, voluptatem hic esse dicta accusamus fuga ipsum quibusdam quos, accusamus praesentium atque quam est asperiores assumenda culpa incidunt blanditiis expedita dicta, perferendis dicta impedit pariatur inventore.Sit officia animi, ipsum eos alias assumenda sint, exercitationem doloribus est laborum, saepe magni molestias necessitatibus quibusdam illo error fugit tenetur magnam beatae, sapiente consequatur