Learning Graph-Based Priors for Generalized Zero-Shot Learning

Colin Samplawski[1,2], Jannik Wolff[1,3], Tassilo Klein[1], Moin Nabi[1]
[1]SAP ML Research Berlin, [2]UMass Amherst, [3]TU Berlin

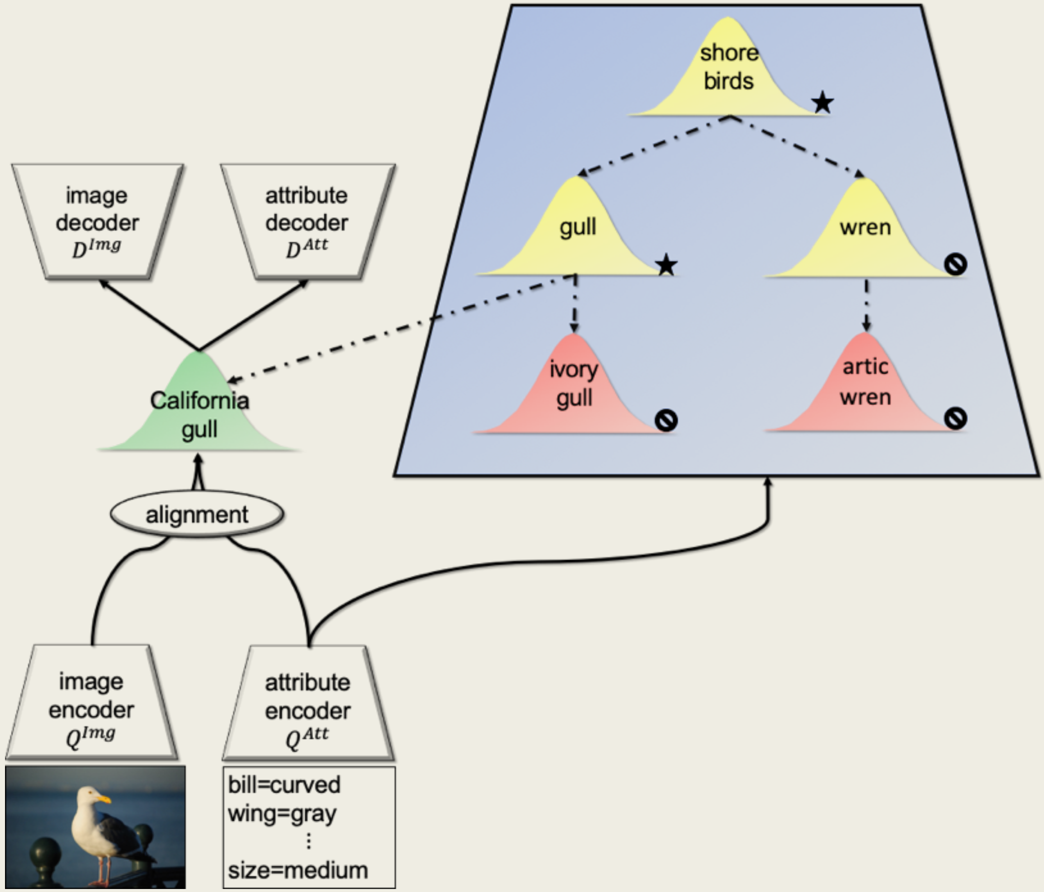
Background and Approach

- Recent approaches to GZSL learn a pair of aligned VAEs, one for each data modality
- A classifier can then be trained using generated latent samples
- We would like to imbue this shared latent space with known graphical information
- We achieve this by embedding each node in the graph as a Gaussian distribution, trained via:

$$L_m(\boldsymbol{c}_i, \boldsymbol{c}_{pos}, \boldsymbol{c}_{neg}) =$$
$$\max\left(0, m + D_{KL}(\boldsymbol{c}_i || \boldsymbol{c}_{pos}) - D_{KL}(\boldsymbol{c}_i || \boldsymbol{c}_{neg})\right)$$

- We then treat this set of embeddings as a prior for an aligned VAE model, replacing the $\mathcal{N}(0,1)$ prior with:

$$L_m(Q^M(\boldsymbol{x}^M), \boldsymbol{c}_{pos}, \boldsymbol{c}_{neg}) \text{ for } M \in \{Img, Att\}$$



Results and Future Work

- We saw improved results on two benchmarks over strong baselines (CUB and SUN)
- Qualitatively, the shared latent space respects the graph structure
- Future work includes investigating datasets which have a more general graph structure, potentially with multiple relations
- Also would like to investigate embedding into a hyperbolic space since many types of graphs can be more efficiently embedded there