An Asymmetric Augmented Self-Supervised Learning Method for Unsupervised Fine-Grained Image Hashing

Supplementary Materials

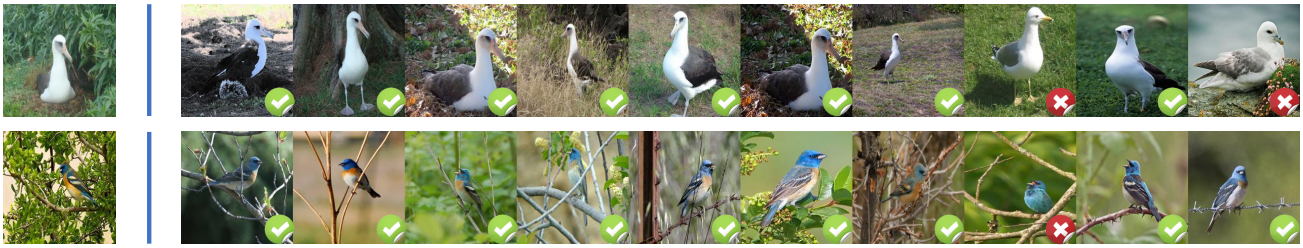
In the supplementary materials, we first introduce the details of five fine-grained benchmark datasets used in the paper. Then, we present more examples of top-10 retrieval results on these fine-grained datasets.

1. Datasets Details

In the paper, we utilize five fine-grained benchmark datasets for empirical evaluation. Specifically, *CUB200-2011* is one of the most popular used fine-grained datasets. It contains 11,788 bird images from 200 bird species and is officially split into 5,994 images for training and 5,794 images for test. *Oxford Flowers* contains 8,189 images spanning 102 flowers with 1,020 for training, 1,020 for validation and 6,149 for test. The *Stanford Cars* consists of 196 classes of cars with a total of 16,185 images, taken from the rear. The data is divided into almost a 50-50 train/test split with 8,144 training images and 8,041 testing images. The *Stanford Dogs* dataset contains 20,580 images of 120 classes of dogs from around the world, which are divided into 12,000 images for training and 8,580 images for testing. For large-scale datasets, *Food101* contains 101 kinds of foods with 101,000 images, where for each class, 250 test images are checked manually for correctness while 750 training images still contain a certain amount of noise.

2. Visualization

In Fig. 1 within the supplementary materials, we showcase additional examples of top-10 retrieval results achieved by A^2-SSL using 48-bit hash codes on five fine-grained datasets. These examples highlight the model's effective retrieval performance across diverse subordinate categories, notably when capturing identical species with variations against distinct backgrounds. Nonetheless, instances of retrieval failure persist, often attributed to extremely subtle visual distinctions (*e.g.*, arising from different perspectives) between the query image and the retrieved images.



(a) CUB200-2011

(b) Oxford Flowers

(c) Stanford Dogs

(d) Stanford Cars

(e) Food101

Figure 1. Examples of top-10 retrieved images on five fine-grained datasets of 48-bit hash codes by our A^2-SSL. The first column in each sub-figure is the query image, and the rest are returned results.