Supplementary Document:Generative Domain-Migration Hashing for Sketch-to-Image Retrieval

Anonymous ECCV submission

Paper ID 1917

Although the main paper stands on its own, it is still worthwhile showing more model details and experimental results. In this supplementary document, we provide:

- Experiments on the Sketchy database
- More illustrative results.
- failure cases and analysis

1 Experiments on the Sketchy database

1.1 Datasets and settings

In this section, we present further experimental results on the recently released Sketchy database[1]. The Sketchy database contains 125 categories with 75,471 sketches of 12,500 object images. In each category, there are 100 object instances. Each instance has one photo and at least 5 corresponding sketches. Following the dataset partitions in [1], we used 90% object instances for training and the rest for testing. Specifically, 11,250 of 12,500 photos and 68,175 of 74,425 sketches are used for training. In the provided test data partition, the gallery set contains 1250 photos (10 photos per category) and the test sketch set contains 6250 query sketches (50 sketches per category).

We compare GDH with 3 existing strongest fine-grained SBIR baselines, including : **AN Siamese:** The base network is AlexNet [2], and it trained with Siamese and classification loss. **GN Siamese:** The base network is GoogLeNet [3] in a heterogeneous architecture. And the network is also trained with Siamese and classification loss. **GN Triplet:** Same GoogLeNet as the base network is used. The network is trained with Triplet ranking loss and classification loss. Moreover, we compare our model performance with human. **Human:** This is obtained by asking human participants to browse through the 1,250 gallery photos to find the best match for a given query sketch.[1]

1.2 Model implementation

The whole model is same as the model presented in the main paper. And the training strategy follows the algorithm presented in the main paper. We use the Adam solver [4] with a batch size of 32. Our balance parameters are set to

$\alpha = 10^{-5}, \beta = 10^{-5}$ and $\lambda = 1$. All networks are trained with an initial learning rate $lr = 0.0002$. After 25 epochs, we decrease the learning rate of the hashing network $lr \rightarrow 0.1lr$ and terminate the optimization after 30 epochs for both datasets. Our method is implemented by Pytorch with dual 1080Ti GPUs and an i7-4790K CPU.

Table 1. Accuracy comparison with different fine-grained SBIR baselines on Sketchy

methods		**Sketchy.acc@1**
Real-valued vectors	AN Siamese [1]	0.2136
	GN Siamese [1]	0.2736
	GN Triplet [1]	0.3710
Binary Codes	GDH @ 64-bit	0.4835
non-numeric	Human[1]	**0.5427**

1.3 Results and Discussions

In Table. 1, we report the top-1 accuracies of GDH over other three methods and human performance on Sketchy dataset for fine-grained SBIR. Compared to the state-of-the-art real-valued GN Triplet, the 64-bit GDH achieves 11.25% improvements. Despite binary hashing codes are used, improved performance over the real-valued state-of-the-art methods can be observed in Table. 1. On the other side, the binary codes in GDH allow much reduced memory costs and retrieval time than the real-valued approaches.

2 Illustrative Results

2.1 Domain-migration networks results on different stage

We show more illustrative domain-migration networks results on QMUL-chairs of different training stages in Fig. 1. The sketch-to-image and image-to-sketch results indicate that our domain-migration networks are capable to transfer domains from both directions.

2.2 Fine-grained SBIR reslults

We show more illustrative fine-grained SBIR results in Fig. 2. Most example queries attain high accuracy among the first five retrieved candidates.

3 failure cases and analysis

We do experience failure cases during our experiments. Recalling the fine-grained results on some unseen query examples, some retrieval results are beyond the top-10 as shown in Fig. 3. We speculate this is because: **quantization error**. The geometrical morphology and detailed instance-level characteristic within a category can be much more difficult to capture with binary hashing codes than the inter-category discrepancies. We further use the feature without quantization to test, results shows that most failure cases change into the top-10.

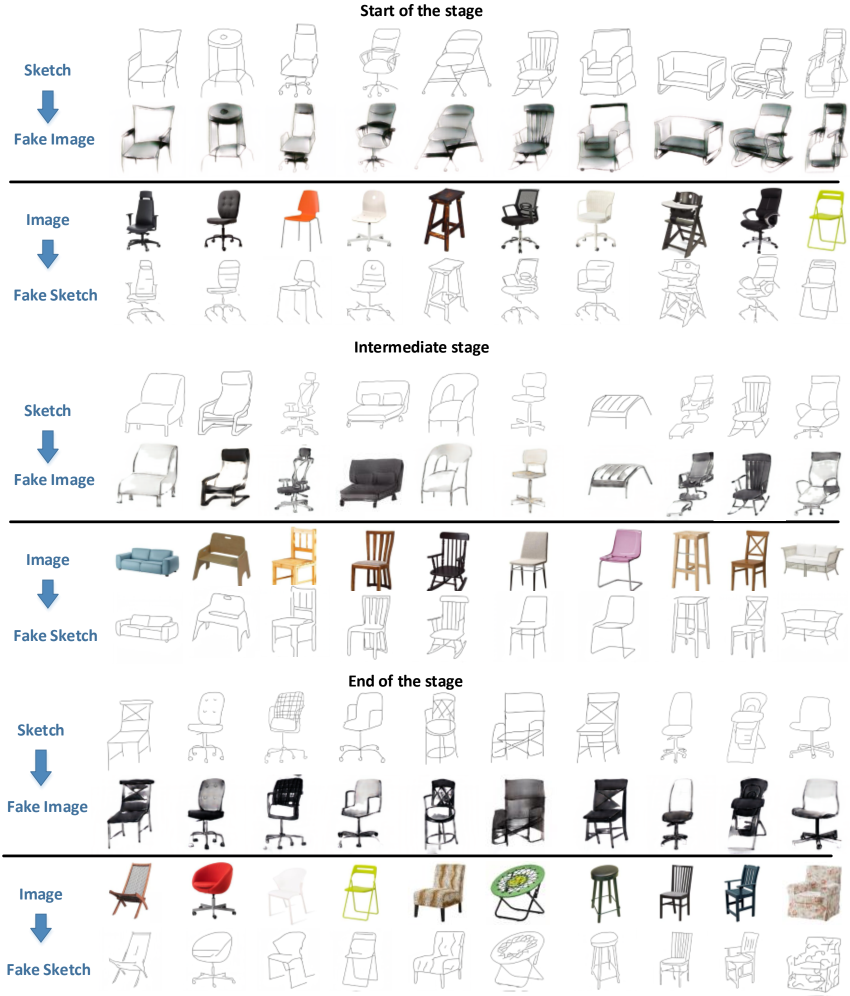

Fig. 1. Visualization of our domain-migration networks on different stages. In each stage, the first two rows are sketch- to-image results and the last two rows are image-to-sketch results, which indicates that our domain-migration networks are capable to transfer domains from both directions.

Fig. 2. Example query sketches with their top-10 retrieval accuracies on the Sketchy dataset by using 128-bit GDH codes. Orange boxes indicate the groundtruth results.

Fig. 3. Failure example query sketches with their top-10 retrieval accuracies on the Sketchy dataset by using 128-bit GDH codes. Top right is the corresponding result and its rank.

References

1. Sangkloy, P., Burnell, N., Ham, C., Hays, J.: The sketchy database: learning to retrieve badly drawn bunnies. ACM Trans. Graph. **35**(4) (2016) 119:1–119:12
2. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States. (2012) 1106–1114
3. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015. (2015) 1–9
4. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. CoRR **abs/1412.6980** (2014)