Image Hashing via Linear Discriminant Learning

Weixiang Hong∗
National University of Singapore
weixiang.hong@outlook.com

Yu-Ting Chang
UC Merced
ychang39@ucmerced.edu

Haifang Qin∗
Peking University
qhfpku@pku.edu.cn

Wei-Chih Hung
UC Merced
whung8@ucmerced.edu

Yi-Hsuan Tsai
NEC Labs America
wasidennis@gmail.com

Ming-Hsuan Yang
UC Merced/Google
mhyang@ucmerced.edu

1. Ablation Study on ResNet-50

We show an ablation study on CIFAR-10 to validate the effectiveness of loss functions proposed in our method using the ResNet-50 as the backbone. In Table 1, without using the proposed inter-class loss, the performance drops significantly.

Table 1. With/Without the LDA loss using the ResNet-50 backbone on CIFAR-10.

<table>
<thead>
<tr>
<th>mAP</th>
<th>12 bits</th>
<th>24 bits</th>
<th>32 bits</th>
<th>48 bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without Inter Loss + Intra Loss</td>
<td>83.1</td>
<td>84.5</td>
<td>85.6</td>
<td>86.3</td>
</tr>
<tr>
<td>With Inter Loss + Intra Loss</td>
<td>86.9</td>
<td>87.2</td>
<td>88.3</td>
<td>88.1</td>
</tr>
</tbody>
</table>

2. Sensitivity Analysis

In Figure 1, we show the sensitive analysis of the loss weights $\alpha$ and $\beta$ on CIFAR-10. We use grid search to determine the value of hyper-parameters $\alpha$ and $\beta$, and fix $\alpha = 0.01$ and $\beta = 0.001$ for all the experiments.

Figure 1. mAP on the CIFAR-10 dataset.

3. Implementation Details

During training, we train LDH for 164 epochs and divide the learning rate by 10 at epoch 81 and 122. In addition, the train/test loss curve w.r.t. epoch is illustrated in Figure 2, while the change of train/test accuracy w.r.t. epoch is provided in Figure 3. With a single Nvidia Tesla v100 GPU, it takes around 40 minutes for training on the CIFAR-10 dataset, 6 hours on the ImageNet dataset, and 11 hours on the NUS-WIDE dataset.

∗Work done at Google.
Figure 2. The train/test loss of LDH. From left to right, we show the results from CIFAR-10, NUS-WIDE and ImageNet.

Figure 3. The train/test accuracy of LDH. From left to right, we show the results from CIFAR-10, ImageNet and NUS-WIDE.

4. More Results and Analysis

In this section, we provide more retrieval examples in Figure 4, 5 and 6. The LDH algorithm is able to retrieve images that share the same semantic labels with the input query. In addition, we evaluate the performances of binary code using a recently proposed metric, $m$AP for unseen classes [1]. As shown in Table 2, our LDH achieves promising $m$AP for unseen classes on the CIFAR-10 dataset.

Table 2. $m$AP of unseen classes, with 16-bit binary code.

<table>
<thead>
<tr>
<th>methods</th>
<th>CCA-ITQ</th>
<th>DHN</th>
<th>DPSH</th>
<th>HashNet</th>
<th>LDH (Ours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mAP</td>
<td>15.4</td>
<td>17.9</td>
<td>18.5</td>
<td>19.8</td>
<td>20.4</td>
</tr>
</tbody>
</table>
Figure 4. Retrieval results on the CIFAR-10 dataset. We use red rectangles to denote false positives.
Figure 5. Retrieval results on the NUS-WIDE dataset. We use red rectangles to denote false positives.
Figure 6. Retrieval results on the ImageNet dataset. We use red rectangles to denote false positives.
References