Motivation

➢ We hypothesize that: CNNs descriptors can boost the performance of state-of-the-art hashing methods, compared to hand-crafted descriptors.

➢ “How can we further enhance discrimination of the features for hashing purpose?”

- Address the problem of producing compact but very discriminative features.

Contributions

➢ Propose to embeds feature vector into a low-dimensional vector and, simultaneously, enhances the discriminative property of features before passing them into hashing.

Gaussian Mixture Model Embedding (Gemb)

1. Dimensionality reduction

- Reduce computational cost
- Produce efficient hashing codes.
  - Balance hash code.
  - Bits are pairwise independent.

- Accomplished by PCA:independent

\[ X = \{ x^{(1)}, \ldots, x^{(m)} \} \in \mathbb{R}^d \rightarrow \hat{X} = \{ \bar{x}^{(1)}, \ldots, \bar{x}^{(m)} \} \in \mathbb{R}^D \]

Where \( D \ll d \).

2. Posterior probability as embedding features

- Assume data belonging to a Gaussian Mixture Model (GMM): \( \lambda = \{ w_i, \mu_i, \Sigma_i \}_{i=1}^N \).

- The embedding feature of a sample \( x^{(t)} \):

\[
P ( j | x^{(t)} , \mu_j, \Sigma_j ) = \frac{w_j p ( x^{(t)} | \mu_j, \Sigma_j )}{\sum_{i=1}^{N} w_i p ( x^{(t)} | \mu_i, \Sigma_i )}
\]

- The posterior probability captures the strength of relationship between a sample \( x^{(t)} \) and a Gaussian model \( N ( \mu_j, \Sigma_j ) \).

3. “Unsparsifying” by Power Normalization

- Observation:
  - Number of Gaussians \( \sim \) Embedding features sparsity.
  - Global descriptors discrimination \( \sim \) Embedding features sparsity.
- Unsparsifying: \( f ( x ) = \text{sign} ( x ) | x | \)

Results

➢ VGG-FC7 descriptors achieve higher hashing performances than GIST 512-D descriptors.

➢ Gemb clearly helps to boost performance of BA and ITQ significantly.

Table 1. Retrieval performance (mAP).

<table>
<thead>
<tr>
<th>Methods</th>
<th>CIFAR-10</th>
<th>LabelMe-12,5K</th>
<th>GIST 512D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mAP</td>
<td>mAP</td>
<td>mAP</td>
</tr>
<tr>
<td>16</td>
<td>12.88</td>
<td>12.71</td>
<td>12.99</td>
</tr>
<tr>
<td>32</td>
<td>14.46</td>
<td>13.13</td>
<td>15.88</td>
</tr>
<tr>
<td>64</td>
<td>15.34</td>
<td>16.86</td>
<td>17.74</td>
</tr>
<tr>
<td>SH</td>
<td>16.59</td>
<td>17.42</td>
<td>18.02</td>
</tr>
<tr>
<td>BA</td>
<td>20.80</td>
<td>22.20</td>
<td>22.45</td>
</tr>
<tr>
<td>ITQ</td>
<td>21.36</td>
<td>22.44</td>
<td>22.59</td>
</tr>
<tr>
<td>Gemb+BA</td>
<td>21.36</td>
<td>22.44</td>
<td>22.59</td>
</tr>
<tr>
<td>Gemb+ITQ</td>
<td>21.36</td>
<td>22.44</td>
<td>22.59</td>
</tr>
</tbody>
</table>

Reference

➢ Link to code: https://github.com/hnanhtuan/Gemb

Fig 1: Our proposed Gemb method:

- \( + \) The inputs are image global descriptors such as GIST or CNN.
- \( + \) The outputs are the embedding features for hashing.

Fig 2. Histogram of embedding features (in log scale) for CIFAR-10.

Fig 3. Visualizing GIST descriptors of a subset of MNIST dataset.