Acknowledgement

Superiority of Our Method

Ablation study and the experiments on the cross-domain FSIC task highlight the semantic objects. Experimental results on five FSIC benchmark datasets show the settings.

Channel attention first squeezes the feature map along the spatial dimension by global average normalization layer and a Leaky ReLU layer. In addition, we add we divide it into 70, 20 and 30 classes. For training, validation and test, respectively.

We evaluate our method on five benchmark datasets, and all images are resized to probabilities and the final loss function of this episode is defined as:

\[ \text{loss} = \sum_{i=1}^{N} \frac{1}{M} \sum_{j=1}^{M} \sum_{k=1}^{K} \left( \alpha \cdot C_{ij}(x_k) + \beta \cdot M_{ij}(x_k) \right) \]

where \( C_{ij}(x_k) \) is the Gaussian similarity and \( M_{ij}(x_k) \) is the cosine similarity.

3.4 Step-by-step remove one or several parts and make comparisons with the original model (see attention mechanism helps to highlight the discriminative local regions, which is obviously the feature maps. Besides, the representation fusion strategy enhances the generalization class.

We design a novel local mutual metric strategy for the FSIC problem, which explores the 3.1 matching the local relations between two feature maps, which can not reveal the true similarities and is of the base model by learning inductive bias (meta knowledge).

Recently meta-learning based FSIC approaches can be broadly divided into two branches: the classic pretraining combined with fine tuning beats the advanced generalize to novel unseen categories with scarce data after training on seen categories with applications where collecting enough samples and correctly labeling them are expensive.

Therefore, Few-shot Learning (FSL) is one of the well-studied fields in applications where collecting enough samples and correctly labeling them are expensive.

Authors

Fig. 1. The overview of our proposed LM2N under the 5-way 1-shot setting.

Fig. 2. The overview of our proposed LM2N under the 5-way 1-shot setting.

<table>
<thead>
<tr>
<th>Type</th>
<th>ImageNet</th>
<th>CUB</th>
<th>Stanford Cars</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAML</td>
<td>52.85</td>
<td>51.78</td>
<td>72.86</td>
</tr>
<tr>
<td>MatchingNet</td>
<td>59.09</td>
<td>58.99</td>
<td>68.20</td>
</tr>
<tr>
<td>ProtoNet</td>
<td>60.98</td>
<td>60.36</td>
<td>69.80</td>
</tr>
<tr>
<td>LM2N (Ours)</td>
<td>61.00</td>
<td>60.56</td>
<td>70.30</td>
</tr>
</tbody>
</table>

Table 1.

\[ \alpha = \frac{1}{2} \left( 1 + \cos \left( \frac{\theta}{2} \right) \right) \]

\[ \beta = \frac{1}{2} \left( 1 - \cos \left( \frac{\theta}{2} \right) \right) \]

\[ \gamma = \frac{1}{2} \left( \frac{1}{\pi} \arccos \left( \cos \left( \frac{\theta}{2} \right) \right) \right) \]

\[ \tilde{g}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{b}(\hat{S}) = 1 - e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]

\[ \tilde{t}(\hat{S}) = \gamma - h \]

\[ \tilde{e}(\hat{S}) = e^{-\frac{\gamma}{h}} \]
References


