COMPARE LEARNING: BI-ATTENTION NETWORK FOR FEW-SHOT LEARNING

ICASSP 2020

Li Ke, Meng Pan, Weigao Wen, Dong Li
Alibaba Group
Outline

- Introduction & Motivation
- Definition & Setup
- Different methods in Few-shot learning
- Classic methods and our improvement
Why few-shot learning (FSL)

1. Learning quickly is a hallmark of human intelligence
2. Annotation cost
What is the N-way-K-shot problem

N=2

K=5

Train

task1

task2

Test

New task

Support Set

Mean Accuracy

High Classification Accuracy

High Classification Accuracy

Support Set

New task
Different methods in FSL

1. Data augmentation
2. Meta learning
3. Metric learning
Metric learning in FSL

Prototypical networks learn a metric space in which classification can be performed by computing distances to prototype representations of each class.

Prototypical Networks for Few-shot Learning (NIPS 2017)
Deep Metric learning in FSL

Figure 1: Relation Network architecture for a 5-way 1-shot problem with one query example.

Learning to compare: Relation network for few-shot learning (CVPR2018)
How Relation Network $g_\phi$ compare
Bi-attention Compare Network

\[ f_\theta \]

\[ W^Q \]

\[ W^K \]

\[ W^V \]

\[ Q \]

\[ K \]

\[ V \]

\[ Z \]

Softmax

\[ \sqrt{d_z} \]

Same category or not
Overview of Bi-Attention Network for a 5-way-1-shot image recognition task (ICASSP2020)
Experiments

Table 1. The 5-way, 1-shot and 5-shot classification testing accuracy(%) on miniImageNet dataset.

<table>
<thead>
<tr>
<th>Few-shot method</th>
<th>1 shot</th>
<th>5 shot</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAML</td>
<td>48.70 ± 1.84</td>
<td>63.11 ± 0.92</td>
</tr>
<tr>
<td>ProtoNets</td>
<td>49.42 ± 0.78</td>
<td>68.20 ± 0.66</td>
</tr>
<tr>
<td>RelationNets</td>
<td>50.44 ± 0.82</td>
<td>65.32 ± 0.70</td>
</tr>
<tr>
<td>MetaGAN</td>
<td>52.71 ± 0.64</td>
<td>68.63 ± 0.67</td>
</tr>
<tr>
<td>Our Approach</td>
<td>53.74 ± 0.89</td>
<td>71.90 ± 0.76</td>
</tr>
</tbody>
</table>

Table 2. The 5-way, 1-shot and 5-shot classification testing accuracy(%) on CIFAR100 dataset. MetaGAN has released neither source code nor result of CIFAR100 till now, so it doesn’t appear below.

<table>
<thead>
<tr>
<th>Few-shot method</th>
<th>1 shot</th>
<th>5 shot</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAML</td>
<td>38.10 ± 1.70</td>
<td>50.40 ± 0.99</td>
</tr>
<tr>
<td>ProtoNets</td>
<td>36.70 ± 0.68</td>
<td>56.50 ± 0.71</td>
</tr>
<tr>
<td>RelationNets</td>
<td>36.56 ± 0.70</td>
<td>48.86 ± 0.65</td>
</tr>
<tr>
<td>Our Approach</td>
<td>39.08 ± 0.81</td>
<td>56.89 ± 0.79</td>
</tr>
</tbody>
</table>

Fig. 2. (a)(b) show the results of 1-shot and 5-shot on miniImageNet; (c)(d) show the results of 1-shot and 5-shot on CIFAR100.
Thanks

Q&A

keli.kl@alibaba-inc.com
huafang.pm@alibaba-inc.com