Prototypical Networks for Few-shot Learning: Supplementary Material

Jake Snell
University of Toronto*
Vector Institute

Kevin Swersky
Twitter

Richard Zemel
University of Toronto
Vector Institute
Canadian Institute for Advanced Research

1 Additional Omniglot Results

In Table 1 we show test classification accuracy for Prototypical Networks using Euclidean distance trained with 5, 20, and 60 classes per episode.

Table 1: Additional classification accuracy results for Prototypical Networks on Omniglot. Configuration of training episodes is indicated by number of classes per episode ("way"), number of support points per class ("shot") and number of query points per class ("query"). Classification accuracy was averaged over 1,000 randomly generated episodes from the test set.

		Train Episodes			**5-way Acc.**		**20-way Acc.**	
Model	**Dist.**	Shot	Query	Way	1-shot	5-shot	1-shot	5-shot
PROTOTYPICAL NETWORKS	Euclid.	1	15	5	97.4%	99.3%	92.0%	97.8%
PROTOTYPICAL NETWORKS	Euclid.	1	15	20	98.7%	99.6%	95.4%	98.8%
PROTOTYPICAL NETWORKS	Euclid.	1	5	60	98.8%	99.7%	96.0%	99.0%
PROTOTYPICAL NETWORKS	Euclid.	5	15	5	96.9%	99.3%	90.7%	97.8%
PROTOTYPICAL NETWORKS	Euclid.	5	15	20	98.1%	99.6%	94.1%	98.7%
PROTOTYPICAL NETWORKS	Euclid.	5	5	60	98.5%	99.7%	94.7%	98.9%

2 Additional *mini*ImageNet Results

In Table 2 we show the full results for the comparison of training episode configuration in Figure 2 of the main paper.

We also compared Euclidean-distance Prototypical Networks trained with a different number of classes per episode. Here we vary the classes per training episode from 5 up to 30 while keeping the number of query points per class fixed at 15. The results are shown in Figure 1. Our findings indicate that construction of training episodes is an important consideration in order to achieve good results for few-shot classification. Table 3 contains the full results for this set of experiments.

*Initial work done while at Twitter.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

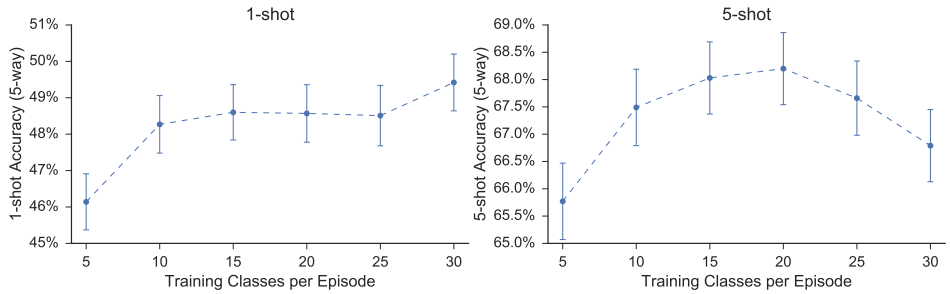

Figure 1: Comparison of the effect of training "way" (number of classes per episode) for Prototypical Networks trained on *mini*ImageNet. Each training episode contains 15 query points per class. Error bars indicate 95% confidence intervals as computed over 600 test episodes.

Table 2: Comparison of Matching Networks and Prototypical Networks on *mini*ImageNet under cosine vs. Euclidean distance, 5-way vs. 20-way, and 1-shot vs. 5-shot. All experiments use a shared encoder for both support and query points with embedding dimension 1,600 (architecture and training details are provided in Section 3.2 of the main paper). Classification accuracy is averaged over 600 randomly generated episodes from the test set and 95% confidence intervals are shown.

		Train Episodes			**5-way Acc.**	
Model	**Dist.**	Shot	Query	Way	1-shot	5-shot
MATCHING / PROTO. NETWORKS	Cosine	1	15	5	$38.82 \pm 0.69\%$	$44.54 \pm 0.56\%$
MATCHING / PROTO. NETWORKS	Euclid.	1	15	5	$46.61 \pm 0.78\%$	$59.84 \pm 0.64\%$
MATCHING / PROTO. NETWORKS	Cosine	1	15	20	$43.63 \pm 0.76\%$	$51.34 \pm 0.64\%$
MATCHING / PROTO. NETWORKS	Euclid.	1	15	20	$49.17 \pm 0.83\%$	$62.66 \pm 0.71\%$
MATCHING NETWORKS	Cosine	5	15	5	$46.43 \pm 0.74\%$	$54.60 \pm 0.62\%$
MATCHING NETWORKS	Euclid.	5	15	5	$46.43 \pm 0.78\%$	$60.97 \pm 0.67\%$
MATCHING NETWORKS	Cosine	5	15	20	$46.46 \pm 0.79\%$	$55.77 \pm 0.69\%$
MATCHING NETWORKS	Euclid.	5	15	20	$47.99 \pm 0.79\%$	$63.66 \pm 0.68\%$
PROTOTYPICAL NETWORKS	Cosine	5	15	5	$42.48 \pm 0.74\%$	$51.23 \pm 0.63\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	5	$44.53 \pm 0.76\%$	$65.77 \pm 0.70\%$
PROTOTYPICAL NETWORKS	Cosine	5	15	20	$42.45 \pm 0.73\%$	$51.48 \pm 0.70\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	20	$43.57 \pm 0.82\%$	$68.20 \pm 0.66\%$

Table 3: Effect of training "way" (number of classes per training episode) for Prototypical Networks with Euclidean distance on *mini*ImageNet. The number of query points per class in training episodes was fixed at 15. Classification accuracy is averaged over 600 randomly generated episodes from the test set and 95% confidence intervals are shown.

		Train Episodes			5-way Acc.	
Model	**Dist.**	Shot	Query	Way	1-shot	5-shot
PROTOTYPICAL NETWORKS	Euclid.	1	15	5	$46.14 \pm 0.77\%$	$61.36 \pm 0.68\%$
PROTOTYPICAL NETWORKS	Euclid.	1	15	10	$48.27 \pm 0.79\%$	$64.18 \pm 0.68\%$
PROTOTYPICAL NETWORKS	Euclid.	1	15	15	$48.60 \pm 0.76\%$	$64.62 \pm 0.66\%$
PROTOTYPICAL NETWORKS	Euclid.	1	15	20	$48.57 \pm 0.79\%$	$65.04 \pm 0.69\%$
PROTOTYPICAL NETWORKS	Euclid.	1	15	25	$48.51 \pm 0.83\%$	$64.63 \pm 0.69\%$
PROTOTYPICAL NETWORKS	Euclid.	1	15	30	$49.42 \pm 0.78\%$	$65.38 \pm 0.68\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	5	$44.53 \pm 0.76\%$	$65.77 \pm 0.70\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	10	$45.09 \pm 0.79\%$	$67.49 \pm 0.70\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	15	$44.07 \pm 0.80\%$	$68.03 \pm 0.66\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	20	$43.57 \pm 0.82\%$	$68.20 \pm 0.66\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	25	$43.32 \pm 0.79\%$	$67.66 \pm 0.68\%$
PROTOTYPICAL NETWORKS	Euclid.	5	15	30	$41.38 \pm 0.81\%$	$66.79 \pm 0.66\%$