# A Scalable Logo Recognition Model with Deep Meta-Learning

Mark de Blaauw[1,2], Diederik M. Roijers[1], and Vesa Muhonen[2]

[1] Vrije Universiteit, Amsterdam, the Netherlands
[2] Mobiquity inc, Amsterdam, the Netherlands

Our goal is to create a deep learning (DL) logo classifier that is widely applicable for conferences. We train this classifier prior to a conference for logos of organisations that will be present. However, DL models typically require many samples to learn such a classification task, whereas deep meta-learning models can learn different but related tasks with 1-10 training samples per class. This reduces data collection costs and increases deployment speed. In this thesis [1], we create a meta-learning classifier that can learn different sets of 35 logo classes with only five training samples per class.

| Approach | Model | 5-shot 5-way | 5-shot 20-way |
|---|---|---|---|
| DML | PN | 87.96 ± 0.68% | 76.25 ± 0.42% |
| | GPN | **89.68 ± 0.67%** | **79.97 ± 0.40%** |
| | Relation | 87.82 ± 0.70% | **80.22 ± 0.39%** |
| IL | MAML | 82.22 ± 0.65% | 62.25 ± 0.48% |
| | Reptile | **88.56 ± 0.70%** | 66.88 ± 0.50% |

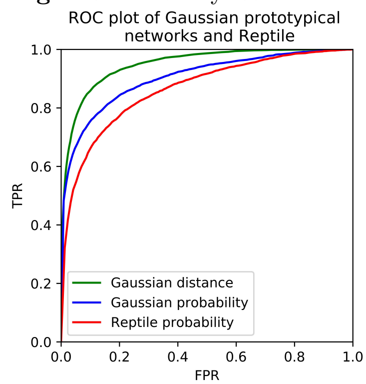
**Table 1.** Accuracy results from the small-logo test set. The 95% CIs are computed over 500 tasks.

Current techniques in the meta-learning domain focus primarily on accuracy, although for AI-driven applications, multiple factors play an important role: (i) robustness of accuracy to different tasks and (ii) robustness to in- and out-domain unknown class samples. We show for (i) that *deep metric learning (DML)* approaches are more robust to tasks with more classes than *initialisation learning (IL)* approaches and for (ii) that Gaussian prototypical networks (GPN) [2] are more robust to in- and out-domain unknown class samples than Reptile [3].

We create two datasets, the small- and large logo dataset, and compare different popular deep meta-learning models. GPN gives the best performance on the small logo dataset. We propose a simpler version of GPN that we call *Mahalanobis prototypical networks (MPN)*, and use the small logo dataset to show that it performs equally well.

We define a conference with $K$ logo classes as a task $t^K$. A classifier is trained using 5 training samples per class, i.e., a 5-shot $K$-way learning task. We construct tasks by uniformly sampling $K$ logo classes from a logo dataset, and form a set of tasks $T^K$ by repeating this procedure. We define three quality metrics for logo classifiers: (i) the accuracy, i.e., the average accuracy across the tasks in set $T^K$, (ii) the rejection accuracy, i.e., the proportion of correct rejections of in- and out-domain unknown class samples, and (iii) robustness, i.e., a low variance in accuracy across tasks from $T^K$ (within-robustness) and a low variance in accuracy across sets with more or less than $K$ logo classes (between-robustness).

**Fig. 1.** 5-shot 5-way ROC-curves.




We first test popular models from IL and DML to investigate accuracy and robustness with the small logo dataset. We present the test set results of these models in Table 1 and we observe that GPN has the highest accuracy or shares this on every learning task. We also see that GPN and the relation model have the best between-robustness when we compare 5-way to 20-way learning. The within-robustness seems to be roughly the same for every model when we compare the 95% confidence intervals (CIs).

Standard DL models, as used in IL, and DML use different decision boundaries in representation space. DML can assign inputs to a class with a probabilistic or distance threshold. We hypothesise that DML has a higher rejection accuracy. We test the rejection accuracy with out-domain unknown class samples from the mini-ImageNet dataset resulting in the ROC curve in Figure 1. We further show that GPN also has a better rejection accuracy with the probabilistic distance for in-domain unknown class samples. We conclude that DML, and especially GPN, has a higher accuracy, rejection accuracy, and better between-robustness than IL. We therefore use GPN as a foundation to create a scalable logo classifier.

GPN uses an estimated covariance matrix to adjust the prototypes and to use the Mahalanobis distance metric. We empirically show that only the Mahalanobis distance metric contributes to the performance increase of GPN over prototypical networks (PN). We test this with the small logo dataset on a set of 5-shot 20-way learning tasks and see that only using the adjusted prototypes gives an accuracy of $76.37 \pm 0.43\%$, whereas only using the Mahalanobis distance metric gives an accuracy of $79.74 \pm 0.43\%$. Therefore, we introduce the MPN model that only uses the Mahalanobis distance metric.

We use the large logo dataset to investigate different CNN architectures for MPN on a set of 5-shot 35-way learning tasks. Furthermore, we find that ResNet-18 gives the highest accuracy and therefore we use MPN with ResNet-18 to get performance metrics on the large logo test set. This results in an accuracy of 89.70%, a threshold accuracy of 77.88% and rejection accuracy of 88.21%.

MPN seems to be the most appropriate technique to use for this conference application based on accuracy, rejection accuracy and between-robustness. Although our MPN model is a simpler method than GPN, we show it has comparable performance. We further note that rejection accuracy and between-robustness can be important metrics for practical meta-learning such as logo classification.

## References

1. De Blaauw, M.: A Scalable Logo Recognition Model with Deep Meta-Learning. Master's thesis, Vrije Universiteit (2019)
2. Fort, S. Gaussian Prototypical Networks for Few-Shot Learning on Omniglot. arXiv:1708.02735 (2017)
3. Nichol, A., Achiam, J., Schulman, J.: On First-Order Meta-Learning Algorithms. arXiv:1803.02999 (2018)