Toward Multimodal Model-Agnostic Meta-Learning

Risto Vuorio[1] **Shao-Hua Sun**[2] **Hexiang Hu**[2] **Joseph J. Lim**[2]
SK T-Brain[1] University of Southern California[2]

Abstract

Gradient-based meta-learners such as MAML [5] are able to learn a meta-prior from similar tasks to adapt to novel tasks from the same distribution with few gradient updates. One important limitation of such frameworks is that they seek a common initialization shared across the entire task distribution, substantially limiting the diversity of the task distributions that they are able to learn from. In this paper, we augment MAML with the capability to identify tasks sampled from a multimodal task distribution and adapt quickly through gradient updates. Specifically, we propose a multimodal MAML algorithm that is able to modulate its meta-learned prior according to the identified task, allowing faster adaptation. We evaluate the proposed model on a diverse set of problems including regression, few-shot image classification, and reinforcement learning. The results demonstrate the effectiveness of our model in modulating the meta-learned prior in response to the characteristics of tasks sampled from a multimodal distribution.

1 Introduction

Recent advances in meta-learning offer machines a way to learn from a distribution of tasks and adapt to a new task from the same distribution using few samples [11, 31]. Different approaches for engaging the task distribution exist. Optimization-based meta-learning methods offer learnable learning rules and optimization algorithms [21, 2, 19, 1, 8], metric-based meta learners [11, 31, 26, 25, 27] address few-shot classification by encoding task-related knowledge in a learned metric space. Model-based meta-learning approaches [4, 32, 17, 15] generalize to a wider range of learning scenarios, seeking to recognize the task identity from a few data samples and adapt to the tasks by adjusting a model's state (*e.g.* RNN's internal states). Model-based methods demonstrate high performance at the expense of hand-designing architectures, yet the optimal strategy of designing a meta-learner for arbitrary tasks may not be obvious to humans. On the other hand, model-agnostic gradient-based meta-learners [5, 6, 9, 12, 7] seek an initialization of model parameters such that a small number of gradient updates will lead to fast learning on a new task, offering the flexibility in the choice of models.

While most existing gradient-based meta-learners rely on a single initialization, different modes of a task distribution can require substantially different parameters, making it infeasible to find a common initialization point for all tasks, given the same adaptation routine. When the modes of a task distribution are disjoint and far apart, one can imagine that a set of separate meta-learners with each covering one mode could better master the full distribution. However, this not only requires additional identity information about the modes, which is not always available or could be ambiguous when the task modes are not clearly disjoint, but also eliminates the possibility of associating transferable knowledge across different modes of a task distribution. To overcome this issue, we aim to develop a meta-learner that acquires a prior over a multimodal task distribution and adapts quickly within the distribution with gradient descent.

To this end, we leverage the strengths of the two main lines of existing meta-learning methods: model-based and gradient-based meta-learning. Specifically, we propose to augment gradient based meta-learners with the capability of generalizing across a multimodal task distribution. Instead of

Workshop on Meta-Learning at 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

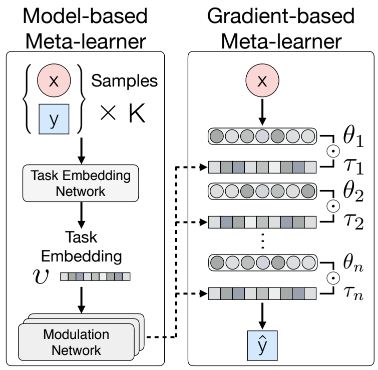


Figure 1: Model overview.

Algorithm 1 META-TRAINING PROCEDURE.

Input: Task distribution $P(\mathcal{T})$, Hyper-parameters α and β
Randomly initialize θ and ω.
while not DONE **do**
  Sample batches of tasks $\mathcal{T}_j \sim P(\mathcal{T})$
  for all j **do**
    Infer $\tau = g(\{x, y\}; \omega)$ with K samples from $\mathcal{D}^{train}_{\mathcal{T}_j}$
    Evaluate $\nabla_\theta \mathcal{L}_{\mathcal{T}_j}\left(f(x; \theta, \tau); \mathcal{D}^{train}_{\mathcal{T}_j}\right)$ w.r.t the K samples
    Compute adapted parameter with gradient descent: $\theta'_{\mathcal{T}_j} = \theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_j}\left(f(x; \theta, \tau); \mathcal{D}^{train}_{\mathcal{T}_j}\right)$
  end for
  Update $\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\mathcal{T}_j \sim P(\mathcal{T})} \mathcal{L}_{\mathcal{T}_j}\left(f(x; \theta', \tau); \mathcal{D}^{val}_{\mathcal{T}_j}\right)$
  Update $\omega \leftarrow \omega - \beta \nabla_\omega \sum_{\mathcal{T}_j \sim P(\mathcal{T})} \mathcal{L}_{\mathcal{T}_j}\left(f(x; \theta', \tau); \mathcal{D}^{val}_{\mathcal{T}_j}\right)$
end while

learning a single initialization point in the parameter space, we propose to first estimate the mode of a sampled task by examining task related samples. Given the estimated task mode, our model then performs a step of *model-based adaptation* to modulate the meta-learned prior in the parameter space to fit the sampled task. Then, from this model adapted meta-prior, a few steps of *gradient-based adaptation* are performed towards the target task to progressively improve the performance on the task. This main idea is illustrated in Figure 1.

2 Method

We aim to develop a Multi-Modal Model-Agnostic Meta-Learner (MUMOMAML) that is able to quickly master a novel task sampled from *a multimodal task distribution*. To this end, we propose to leverage the ability of model-based meta-learners to identify the modes of a task distribution as well as the ability of gradient-based meta-learners to consistently improve the performance with a few gradient steps. Specifically, we propose to learn a model-based meta-learner that produces a set of task specific parameters to modulate the meta-learned prior parameters. Then, this modulated prior learns to adapt to a target task rapidly through gradient-based optimization. An illustration of our model is shown in Figure 1.

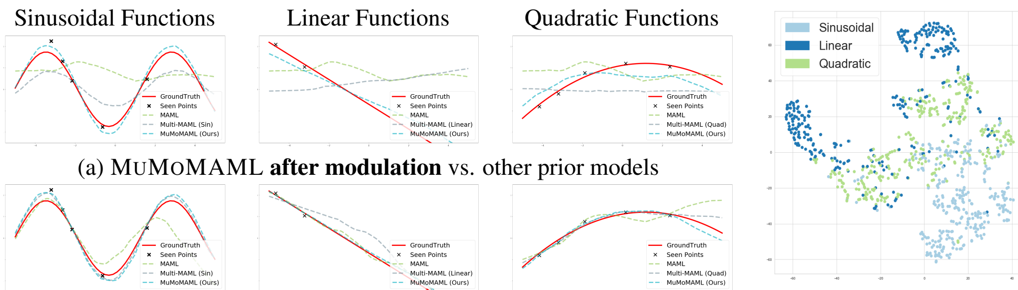
The **gradient-based meta-learner**, parameterized by θ, is optimized to quickly adapt to target tasks with few gradient steps by seeking a good parameter initialization similar to [5]. For the architecture of the gradient-based meta-learner, we consider a neural network consisting of N blocks where the i-th block is a convolutional or a fully-connected layer parameterized by θ_i. The **model-based meta-learner**, parameterized by ω, aims to identify the mode of a sampled task from a few samples and then modulate the meta-learned prior parameters of the gradient-based meta-learner to enable rapid adaptation in the identified mode. The model-based meta-learner consists of a *task embedding network* and a *modulation network*.

Given K data points and labels $\{x_k, y_k\}_{k=1,...,K}$, the task embedding network f learns to produce an embedding vector υ that encodes the characteristics of a task according to $\upsilon = f(\{x_k, y_k\}_{k=1,...,K}; \omega_f)$. The modulation network g learns to modulate the meta-learned prior of the gradient-based meta-learner in the parameter space based on the task embedding vector υ. To enable specialization of each block of the gradient-based meta-learner to the task, we apply the modulation block-wise to activate or deactivate the units of a block (*i.e.* a channel of a convolutional layer or a neuron of a fully-connected layer). Specifically, modulation network produces the modulation vectors for each block i by $\tau_1, ..., \tau_N = g(\upsilon; \omega_g)$, forming a collection of modulated parameters τ. We formalize the procedure of applying modulation as: $\phi_i = \theta_i \odot \tau_i$, where ϕ_i represents the modulated prior parameters for the gradient-based meta-learner, and $\odot$ represents a general modulation function. In the experiments, we investigate some representative modulation operations including attention-based modulation [16, 30] and feature-wise linear modulation (FiLM) [18].

Training The training procedure for jointly optimizing the model-based and gradient-based meta-learners is summarized in Algorithm 1. Note that τ is not updated in the inner loop, as the model-based meta-learner is only responsible for finding a good task-specific initialization through modulation. The implementation details can be found in Section C and Section D.

Table 1: Model's performance on the **multimodal 5-shot regression** with two or three modes. Gaussian noise with $\mu = 0$ and $\sigma = 0.3$ is applied. The three mode regression is in general more difficult (thus higher error). In Multi-MAML, the GT modulation represents using ground-truth task identification to select different MAML models for each task mode. MuMoMAML (wt. FiLM) outperforms other methods by a significant margin.

Configuration		**Two Modes (MSE)**		**Three Modes (MSE)**	
Method	Modulation	Post Modulation	Post Adaptation	Post Modulation	Post Adaptation
MAML [5]	-	15.9255	1.0852	12.5994	1.1633
Multi-MAML	GT	16.2894	0.4330	12.3742	0.7791
MuMoMAML (ours)	Softmax	3.9140	0.4795	0.6889	0.4884
MuMoMAML (ours)	Sigmoid	1.4992	0.3414	2.4047	0.4414
MuMoMAML (ours)	FiLM	1.7094	**0.3125**	1.9234	**0.4048**



(a) MuMoMAML **after modulation** vs. other prior models

(b) MuMoMAML **after adaptation** vs. other posterior models

(c) Task embeddings

Figure 2: Few-shot adaptation for the multimodal regression task. (a): Without any gradient update, MuMoMAML (blue) fits target functions by modulating the meta-learned prior, outperforming the prior models of MAML (green) and Multi-MAML (gray). (b): After five steps of gradient updates, MuMoMAML outperforms MAML and Multi-MAML on all functions. More visualizations in Figure 8 and Figure 9. (c): tSNE plots of the task embeddings υ produced by our model from randomly sampled tasks; marker color indicates different types of functions. The plot reveals a clear clustering according to different task modes, showing that MuMoMAML is able to infer the mode from a few samples and produce a meaningful embedding. The distance among distributions aligns with the intuition of the similarity of functions (*e.g.* a quadratic function can sometimes be similar to a sinusoidal or a linear function while a sinusoidal function is usually different from a linear function).

3 Experiments

To verify that the proposed method is able to quickly master tasks sampled from multimodal task distributions, we compare it with baselines on a variety of tasks, including regression, reinforcement learning, and few-shot image classification [1].

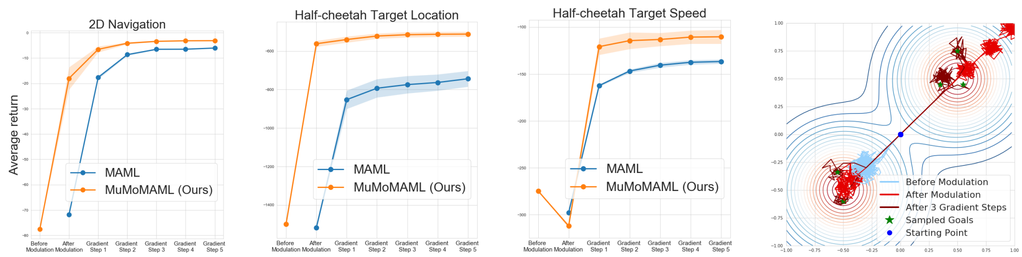
3.1 Regression

We investigate our model's capability of learning on few-shot regression tasks sampled from multimodal task distributions. In these tasks, a few input/output pairs $\{x_k, y_k\}_{k=1,...,K}$ sampled from a one dimensional function are given and the model is asked to predict L output values $y_1^q, ..., y_L^q$ for input queries $x_1^q, ..., x_L^q$. We set up two regression settings with two task modes (sinusoidal and linear functions) or three modes (quadratic functions added). Please see Section D for details.

As a baseline beside MAML, we propose Multi-MAML, which consists of M (the number of modes) separate MAML models which are chosen for each query based on ground-truth task-mode labels. This baseline serves as an upper-bound for the performance of MAML when the task-mode labels are available. The quantitative results are shown in Table 1. We observe that Multi-MAML outperforms MAML, showing that MAML's performance degrades on multimodal task distributions. MuMoMAML consistently achieves better results than Multi-MAML, demonstrating that our model is able to discover and exploit transferable knowledge across the modes to improve its performance. The marginal gap between the performance of our model in two and three mode settings indicates that MuMoMAML is able to clearly identify the task modes and has sufficient capacity for all modes.

We compared attention modulation with Sigmoid or Softmax and FiLM modulation and found that FiLM achieves better results. We therefore use FiLM for further experiments. Please refer to Section A for additional details. Qualitative results visualizing the predicted functions are shown in Figure 2.

[1] Due to the page limit, the results of few-shot image classification are presented and discussed in Section B



(a) 2D Navigation (b) Target Location (c) Target Speed (d) 2D Navigation Results

Figure 3: (a-c) Adaptation curves for MUMOMAML and MAML baseline in 2D navigation and half-cheetah environments. The "after modulation" step represents the rewards of the modulated policy for MUMOMAML and the initial rewards for MAML. MUMOMAML outperforms MAML across the gradient update steps given a single extra trajectory. (d) Visualized trajectories sampled using MUMOMAML in the 2D navigation environment. The contours represent the probability density of the goal distribution (red: high probability; blue: low probability). The trajectories demonstrate the effect of modulation and the subsequent fine tuning with gradient steps. Additional trajectory visualizations can be found in Figure 10.

Figure 2 (a) shows that our model is able to identify tasks and fit to the sampled function well without performing gradient steps. Figure 2 (b) shows that our model consistently outperforms the baselines with gradient updates. Figure 2 (c) plots a tSNE [14], showing the model-based module is able to identify the task modes and produce embedding vectors υ. Additional results are shown in Section E.

3.2 Reinforcement Learning

We experiment with MUMOMAML in three reinforcement learning (RL) environments to verify its ability to learn to rapidly adapt to tasks sampled from multimodal task distributions given a minimum amount of interaction with an environment. [2]

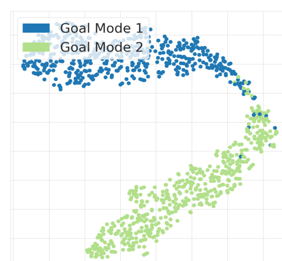


Figure 4: A tSNE plot of task embeddings of randomly sampled tasks in Target Location environment capturing the bimodal task distribution.

2D Navigation. We utilize a 2D navigation environment with bimodal task distribution to investigate the capabilities of the embedding network to identify the task mode based on trajectories sampled from RL environments and the modulation network to modulate a policy network. In this environment, the agent is rewarded for moving close to a goal location. The model-based meta-learner is able to identify the task modes and modulate the policy accordingly, allowing efficient fast adaptation. This is shown in the agent trajectories and the average return plots presented in Figure 3 (a) and (d), where our model outperforms MAML with any number of gradient steps.

Half-cheetah Target Location and Speed. To investigate the scalability of our method to more complex RL environments we experiment with locomotion tasks based on the half-cheetah model. In the target location and target speed environments the agent is rewarded for moving close to the target location or moving at target speed respectively. The targets are sampled from bimodal distributions. In these environments, the dynamics are considerably more complex than in the 2D navigation case. MUMOMAML is able to utilize the model-based meta-learner to effectively modulate the policy network and retain an advantage over MAML across all gradient update steps as seen from the adaptation curves in Figure 3 (b) and Figure 3 (c). A tSNE plot of the embeddings in Figure 4 shows that our model is able to produce meaningful task embeddings υ.

4 Conclusion

We presented a novel meta-learning approach that is able to leverage the strengths of both model-based and gradient-based meta-learners to discover and exploit the structure of multimodal task distributions. With the ability to effectively recognize the task modes as well as rapidly adapt through a few gradient steps, our proposed MUMOMAML achieved superior generalization performance on multimodal few-shot regression, reinforcement learning, and image classification.

[2]Please refer to Section D for details on the experimental setting.

References

[1] Marcin Andrychowicz, Misha Denil, Sergio Gomez Colmenarejo, Matthew W. Hoffman, David Pfau, Tom Schaul, and Nando de Freitas. Learning to learn by gradient descent by gradient descent. In *Advances in Neural Information Processing Systems*, 2016.

[2] Samy Bengio, Yoshua Bengio, Jocelyn Cloutier, and Jan Gecsei. On the optimization of a synaptic learning rule. In *Preprints Conf. Optimality in Artificial and Biological Neural Networks*, 1992.

[3] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[4] Yan Duan, John Schulman, Xi Chen, Peter L Bartlett, Ilya Sutskever, and Pieter Abbeel. Rl ^{2}: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint arXiv:1611.02779*, 2016.

[5] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, 2017.

[6] Chelsea Finn, Kelvin Xu, and Sergey Levine. Probabilistic Model-Agnostic Meta-Learning. In *Advances in Neural Information Processing Systems*, 2018.

[7] Erin Grant, Chelsea Finn, Sergey Levine, Trevor Darrell, and Thomas Griffiths. Recasting Gradient-Based Meta-Learning as Hierarchical Bayes. In *International Conference on Learning Representations*, 2018.

[8] Sepp Hochreiter, A Steven Younger, and Peter R Conwell. Learning to learn using gradient descent. In *International Conference on Artificial Neural Networks*, pages 87–94. Springer, 2001.

[9] Taesup Kim, Jaesik Yoon, Ousmane Dia, Sungwoong Kim, Yoshua Bengio, and Sungjin Ahn. Bayesian Model-Agnostic Meta-Learning. In *Neural Information Processing Systems*, 2018.

[10] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

[11] Gregory Koch, Richard Zemel, and Ruslan Salakhutdinov. Siamese neural networks for one-shot image recognition. In *Deep Learning Workshop at International Conference on Machine Learning*, 2015.

[12] Yoonho Lee and Seungjin Choi. Gradient-Based Meta-Learning with Learned Layerwise Metric and Subspace. In *International Conference on Machine Learning*, 2018.

[13] Zhenguo Li, Fengwei Zhou, Fei Chen, and Hang Li. Meta-SGD: Learning to Learn Quickly for Few-Shot Learning. *arXiv preprint arXiv:1707.09835*, 2017.

[14] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 2008.

[15] Nikhil Mishra, Mostafa Rohaninejad, Xi Chen, and Pieter Abbeel. A Simple Neural Attentive Meta-Learner. In *International Conference on Learning Representations*, 2018.

[16] Volodymyr Mnih, Nicolas Heess, Alex Graves, et al. Recurrent models of visual attention. In *Advances in Neural Information Processing Systems*, 2014.

[17] Tsendsuren Munkhdalai and Hong Yu. Meta networks. In *International Conference on Machine Learning*, 2017.

[18] Ethan Perez, Florian Strub, Harm de Vries, Vincent Dumoulin, and Aaron C. Courville. Film: Visual reasoning with a general conditioning layer. In *Association for the Advancement of Artificial Intelligence*, 2018.

[19] Sachin Ravi and Hugo Larochelle. Optimization as a Model for Few-Shot Learning. In *International Conference on Learning Representations*, 2017.

[20] Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, and Timothy Lillicrap. Meta-Learning with Memory-Augmented Neural Networks. In *International Conference on Machine Learning*, 2016.

[21] Jurgen Schmidhuber. Evolutionary principles in self-referential learning. on learning now to learn: The meta-meta-meta...-hook. Master's thesis, Technische Universitat Munchen, Germany, 1987.

[22] Jürgen Schmidhuber, Jieyu Zhao, and Nicol N Schraudolph. Reinforcement learning with self-modifying policies. In *Learning to learn*. 1998.

[23] Jürgen Schmidhuber, Jieyu Zhao, and Marco Wiering. Shifting inductive bias with success-story algorithm, adaptive levin search, and incremental self-improvement. *Machine Learning*, 1997.

[24] John Schulman, Sergey Levine, Philipp Moritz, Michael I. Jordan, and Pieter Abbeel. Trust Region Policy Optimization. In *International Conference on Machine Learning*, 2015.

[25] Pranav Shyam, Shubham Gupta, and Ambedkar Dukkipati. Attentive recurrent comparators. *arXiv preprint arXiv:1703.00767*, 2017.

[26] Jake Snell, Kevin Swersky, and Richard S. Zemel. Prototypical Networks for Few-shot Learning. In *Advances in Neural Information Processing Systems*, 2017.

[27] Flood Sung, Yongxin Yang, Li Zhang, Tao Xiang, Philip H.S. Torr, and Timothy M. Hospedales. Learning to compare: Relation network for few-shot learning. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[28] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems*, 2014.

[29] Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *International Conference On Intelligent Robots and Systems*, 2012.

[30] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. In *Advances in Neural Information Processing Systems*, 2017.

[31] Oriol Vinyals, Charles Blundell, Tim Lillicrap, Daan Wierstra, et al. Matching networks for one shot learning. In *Advances in Neural Information Processing Systems*, 2016.

[32] Jane X Wang, Zeb Kurth-Nelson, Dhruva Tirumala, Hubert Soyer, Joel Z Leibo, Remi Munos, Charles Blundell, Dharshan Kumaran, and Matt Botvinick. Learning to reinforcement learn. *arXiv preprint arXiv:1611.05763*, 2016.

[33] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 1992.

[34] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. In *International Conference on Machine Learning*, 2015.

[35] Zichao Yang, Xiaodong He, Jianfeng Gao, Li Deng, and Alex Smola. Stacked attention networks for image question answering. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[36] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-Attention Generative Adversarial Networks. *axXiv preprint arXiv:1805.08318*, 2018.