accounted, decoder network projects features to classifier. For generalization to unseen tasks, parameters in all these modules are updated with MAML. Meta-learner provides initial values for these parameters in base

ing samples. In our approach, the parameters of the model are explicitly trained such that a small number of gradient steps with a small amount of training data from a new task will produce good generalization performance on that task. In

veloping new methods and analyzing existing ones. We then proposed a model-agnostic meta-learning algorithm, or MAML, that embeds gradient descent into the meta-learner, aiming to find an initial representation such that one or a few gradient steps leads to effective generalization. This meta-learner, by construction, will acquire consistent learn-

MAML optimizes for generalization, akin to cross-validation.