Learning to Adapt With Mem
Probabilistic Few-Shot Lea

Lei Zhang, *Member, IEEE*, Liyun Zuo, Yingjun Du, and Xi

***Abstract*— Few-shot learning has recently generated increasing popularity in machine learning, which addresses the fundamental yet challenging problem of learning to adapt to new tasks with the limited data. In this paper, we propose a new probabilistic framework that learns to fast adapt with external memory. We model the classifier parameters as distributions that are inferred from the support set and directly applied to the query set for prediction. The model is optimized by formulating as a variational inference problem. The probabilistic modeling enables better handling prediction uncertainty due to the limited data. We impose a discriminative constraint on the feature representations by exploring the class structure, which can improve the classification performance. We further introduce a memory unit to store task-specific information extracted from the support set and used for the query set to achieve explicit adaption to individual tasks. By episodic training, the model learns to acquire the capability of adapting to specific tasks, which guarantees its performance on new related tasks. We conduct extensive experiments on widely-used benchmarks for few-shot recognition. Our method achieves new state-of-the-art performance and largely surpassing previous methods by large margins. The ablation study further demonstrates the effectiveness of the proposed discriminative learning and memory unit.**

***Index Terms*— Few shot learning, external memory, variational inference.**

I. Introduction

HUMANS have the instinct to effortlessly learn new concepts from a few examples and show great generalization ability to new tasks. However, existing machine learning models, e.g., deep neural networks (DNNs) [54], [59], rely heavily on large-scale annotated training data in order to achieve satisfactory performance. The huge gap between human intelligence and DNNs motivates us to try and progress

Manuscript received September 11, 2020; revised November 12, 2020 and December 20, 2020; accepted January 6, 2021. Date of publication January 19, 2021; date of current version October 28, 2021. This work was supported in part by the Natural Science Foundation of China (NSFC) under Grant 61871016 and Grant 61976060, in part by the Department of Education of Guangdong Province under Grant 2018KCXTD019, and in part by the Open Research Project of the State Key Laboratory of Industrial

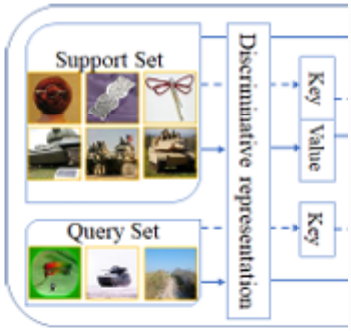

Fig. 1. Illustration of the pr
few-shot classification. For each
set is stored in the memory. T
attention mechanism to achieve

the task of learning fro
learning [19], [27], [40].

Learning to recognize
samples pose great challe
ing models [32]. In or
learning, meta-learning
larity in the machine le
knowledge learned from
meta-learning provides a
Generally speaking, meta
general knowledge from
quickly adapt to new ta
learners. In addition, the
tain since the model is
samples. To tackle the ur
been explored recently, s
learning [11], [15]. We in
framework for few-shot l
the memory mechanism.

Neural networks with e
generated increasing atten
nity, which is inspired by
intelligence. External mem