Continual Learning

- **Proper Noun**
  - Task Incremental Learning
  - Class Incremental Learning
    - *Incremental Implicitly-Defined Classification (IIDC)*
    - *Fine-chel class-incremental learning*
- **Experiments**
- **Awesome Expression**
  - Abstract
    - Paradigm systems of continual learning consider sequential deployment data of classes in the same order to be learned. Recent advances in continual learning have witnessed various benefits of using various regularization, rehearsal, or parameter isolation techniques. We propose a simple continual learning method by combining two complementary ideas, namely, Bayesian weight regularization and sparse rehearsal, which jointly attack catastrophic forgetting in deep neural networks. We show that our proposed method achieves competitive results on various benchmarks.
  - Introduction
  - Related Work

Background

- **Problem**
  - 持续学习研究的是人类能力的重要特性，即不会忘记旧知识，而在学习新知识的同时，还能运用先前学习来提高学习效率。
- **Major challenge**
  - 目前最常见的方法要么存储先前数据用于学习，要么要求在未来学习新任务时，先前学习的数据可用。这两种方法都会带来不可持续的内存开销问题。
- **Scope**
  - Only consider the real-world sequential setting, in which mini-batch data arrives one class at a time, and stream SGD—a single epoch.
  - 只考虑实时场景，即每次学习一个类别，一个mini-batch数据到来，只训练一个epoch。

Approaches

Regularization-based methods

Prior-focused

- Online your Hamiltonian Series Weight Regularization for Continual Learning and Transfer Learning, 2019, 深度哈密顿网络的在线权重正则化方法
  - 2019 · 2019 · Online
  - 1. Hamiltonian Dynamics for Weight Regularization
  - 2. Online Hamiltonian Weight Regularization
  - 3. Transfer Hamiltonian Weight Regularization
- Memory Aware Synapses: Learning What (not) to Forget, 2018, 记忆感知突触：学习什么该忘记
  - 2018 · 2017
  - $\Omega_k = \frac{1}{N}\sum_{i=1}^{N} \left\| \frac{\partial \left(F(x_i;\theta + \Delta\theta) - F(x_i;\theta) \right)}{\partial \Delta\theta} \right\|_{\Delta\theta=0}$
  - $\Omega_k = \frac{1}{N}\sum_{i=1}^{N} \left\| \frac{\partial F(x_i;\theta)}{\partial \theta} \right\|$
  - Learning to learn without forgetting by maximizing transfer and minimizing interference, 2019, 通过最大化迁移和最小化干扰来学习不忘记的学习
  - 2019 · 2018 · Meta
  - $\sum_t \sum_{s<t} \gamma^{t-s} \left[\nabla_\theta \log(p(x^s;\theta)) \cdot \nabla_\theta \log(p(x^t;\theta)) \right]$
  - $\sum_t \sum_{s<t} \gamma^{t-s} \left[\nabla_\theta \log(p(x^s;\theta)) \cdot \nabla_\theta \log(p(x^t;\theta)) \right]$
  - Problem Setting: Meta-CL
  - Meta-Learning Objective
  - Meta-Learning Objective
  - Previous Task
  - Current Task
  - Meta-Learning Objective
  - Meta-Learning Objective
  - Meta-Learning Objective