

Introducing Journal

Continual Learning on Deep Generative Replay

Continual Learning with Deep Generative Replay

Part of [Advances in Neural Information Processing Systems 30 \(NIPS 2017\)](#)

[Bibtex](#)[Metadata](#)[Paper](#)[Reviews](#)

Authors

Hanul Shin, Jung Kwon Lee, Jaehong Kim, Jiwon Kim

Author

Jung Kwon Lee (MIT)

Jaehong Kim (SK T-Brain)

Jiwon Kim (SK T-Brain)

Hanul Shin (SK T-Brain)

Publisher : NIPS(Advances in Neural Information Processing Systems)

Citation : 896

CONTENTS

1

Introduction

2

Related Work

3

Generative Replay

01 INTRODUCTION

Background

Continual Learning (CL)은 인간과 대형 영장류에서 나타나는 특성

살아가는 동안 지속적으로 배우고 학습하며 지식을 축적하는 특성

이를 딥러닝에 적용할 때, “Catastrophic Forgetting(CF)”의 문제점

이전 task를 학습한 후, 다음 task를 학습할 때, 이전 task에 대한 성능이 낮아지는 현상
본 논문에서는 MNIST 1~10 숫자와 Street View House Dataset(SVHN)의 1~10 숫자를 연속적으로 학습

Task가 달라지면 Parameter가 바뀜

Artificial NN에서, $\text{Output} = W * \text{Input}$ 을 만족
task마다 W가 달라지기 때문에, MNIST를 학습한 W에 SVHN을 연속학습하면, MNIST의 성능이 낮아짐

Experience Replay(Episodic Memory)의 한계점

이전 Task에 대한 데이터를 저장 후 학습하는 방법으로
데이터 저장 용량 및 privacy의 한계점 존재
특히 의료 Domain에서는 Data Privacy가 중요

01 INTRODUCTION

Idea

인간과 영장류의 차이 : 신피질과 해마의 연결성

해마 : 정보를 Encoding하여 흘려 보내고, 잠자는 동안 기억을 활성화 함
이후 저장할 Encoding된 기억은 여러번 반복하여 동기화

신피질 : 최종적으로 기억할 내용을 신피질에 통합

이 내용이 RL에서 Agent 학습할 때 나온다고 함

이 과정에서, 해마의 특징이 추가로 발견

해마가 과거 기억을 단순 저장하는 것이 아닌, 살짝 **왜곡**하여 반복 활성화함
NN에서도 단순 반복보단, DGM으로 약간의 왜곡을 주어 학습을 주면 더 좋을 것으로 예상

기대 효과

DGR은 실제 데이터가 아닌 Generated 데이터이므로, 데이터 Privacy의 해결이 가능할 것이다.

02 Related Work

CF 완화 연구

모델의 가중치 제한

Elastic Weight Consolidation

이전 task에 대한 중요성에 기반해 특정 weight를 지키도록 학습

학습 억제

LwF

가중치마다 Learning Rate를 줄여 해당 파라미터의 학습 자체를 억제

=====

논문 이외의 방법

모델 수정 방법

Dynamically Expandable Network(DEN); (Yoon et al. 2017. KAIST)

데이터 저장 방법

Gradient Episodic Memory(Lopez-Paz and Ranzato.2017)

02 Related Work

Complementary Learning System (CLS) Theory

이전 데이터에 접근 불가능 할 때, 이전 task에서 얻어지는 가짜 input과 가짜 output을 가지고 함께 학습하는 방법

→ 쉬운 task는 랜덤 노이즈랑 결과를 주고 어려운 task는 해마랑 유사한 구조를 가지게 하는 등의 연구

실제 세계에서 고해상도와 관련된 연구는 없음

Deep Generative Models

생성 모델

Generation을 하면 생성 모델
VAE, GAN 등 모두 Generative Models

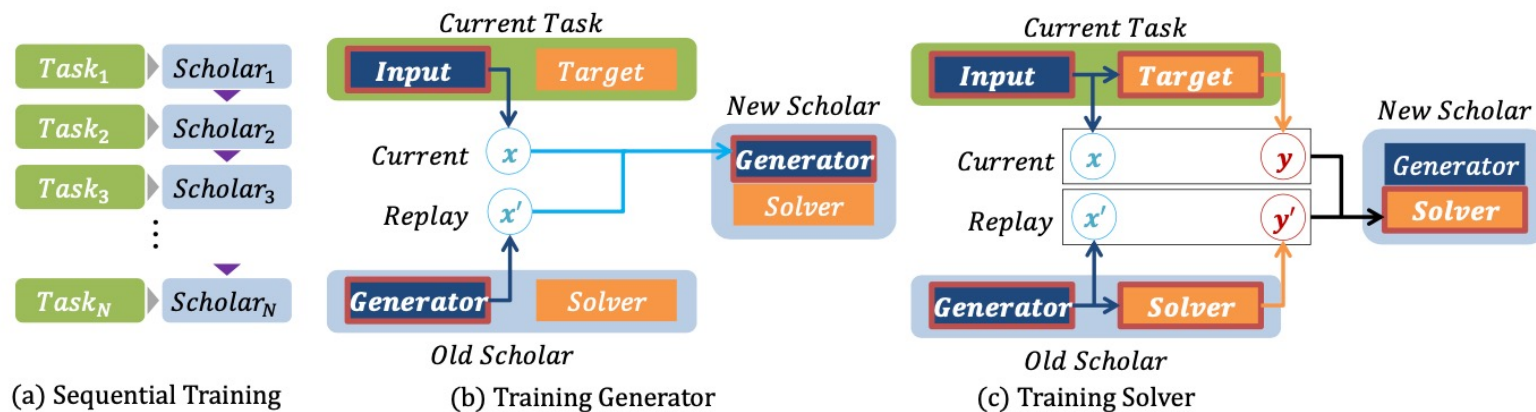
GAN

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p(z)} [\log(1 - D(G(z)))]$$

생성망과 판별망을 이용하여 제로섬 게임

Loss = 실제 이미지를 진짜라고 할 확률 + (1 - 가짜 이미지를 진짜라고 할 확률) => 얼마나 잘 맞추는가

03 Generative Replay



(a) Sequential Training

T : Task, 학습용 샘플들을 가져올 수 있도록 함
Scholar : 우리 모델, 새로운 Task를 학습하고 이를 다른 network에 전달

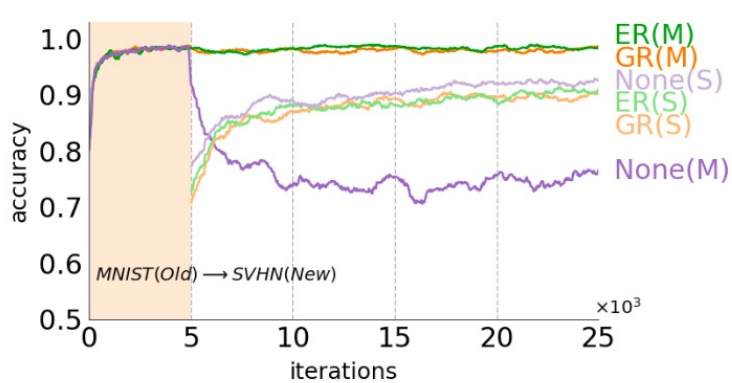
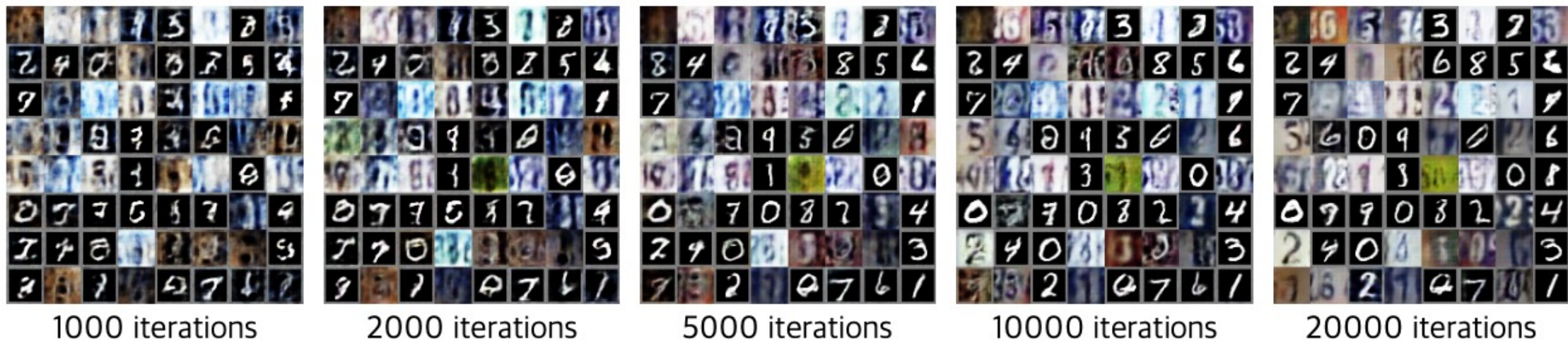
(b) Training Generator

개별 Scholar 내부에서의 학습 과정
현재 task의 input x 와 이전 task의 replay x' 을 받고, 중요도를 비율로 조정

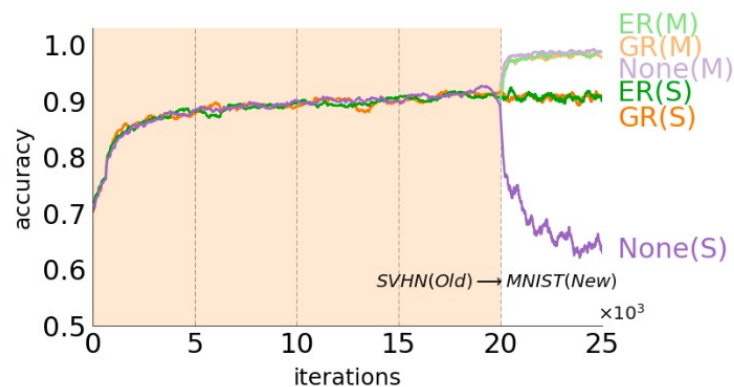
(c) Training Solver

실제 데이터와 Generated Input 데이터의 조합으로 Solver 학습

03 Generative Replay



(a) MNIST to SVHN



(b) SVHN to MNIST

GR
ER
None

: 본 논문의 구조를 따르는 모델
: Generator가 충분히 좋을 때의 모델
: Solver가 Naïve하게 학습된 모델