



Basic Concepts of Continual Learning

Author: Chaeun Ryu

Reference:

1. <https://paperswithcode.com/task/continual-learning>
2. <https://ffighting.tistory.com/112>
3. <https://engineering-ladder.tistory.com/94>
4. <http://dmgm.korea.ac.kr/activity/seminar/266>
5. <https://www.slideshare.net/NaverEngineering/lifelong-learning-for-dynamically-expandable-networks>
6. <https://noteforstudy.tistory.com/m/entry/Continual-Learning-새로운-데이터로-모델을-재학습해야한다면>

1. Definition

Continual Learning (also known as **Incremental Learning**, **Life-long Learning**) is a concept to learn a model for a **large number of tasks sequentially without forgetting knowledge obtained from the preceding tasks**, where the data in the old tasks are not available anymore during training new ones.

2. Background (Why CL is needed!)

“기존 학습된 정보는 유지하면서 새로운 정보를 학습할 수는 없을까?”

2.1 “기존 학습된 정보는 유지하면서”

= Prevent Catastrophic Forgetting & Semantic Drift

2.1.1 = Prevent **Catastrophic Forgetting!**

→ 그래서 incremental learning을 *Learning without Forgetting*이라고 부르기도 함!



Catastrophic Forgetting이란?

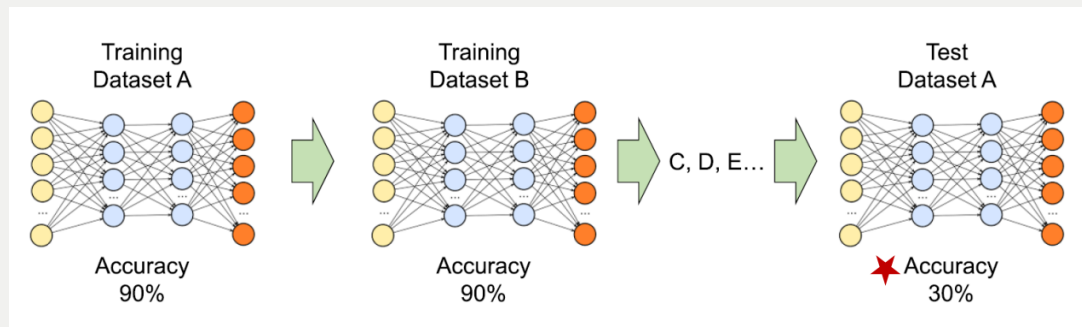
새로운 task에 대해 학습하게 되면 신경망모델이 이전에 배운 task에 대해서는 네트워크가 까먹는 현상으로 Transfer Learning에서 대표적으로 자주 발생함.

Catastrophic Forgetting Example 1)

Race Classification(인종 구분)에 사용했던 모델을 Animal Classification(동물 종류 구분)에 fine-tuning했을 때, 이 모델은 Race Classification을 더이상 효과적으로 해결 못함.

Catastrophic Forgetting Example 2)

Car classification model을 만든다고 할 때, 처음에 Toyota에서 나온 차종에 대해 훈련시킨 후 Tesla에서 나온 차종에 대해 훈련시키면, Toyota에 나온 차종에 대한 분류 성능이 감소함.



Ref: [4]

2.1.2 = Prevent Semantic Drift!



Semantic Drift란?

새로 학습하는 과정에서 pre-trained weights가 과도하게 조정될 경우 Node나 Weight의 의미가 변하는 현상.

Example 1) 도로의 정보를 담고 있던 Node가 새로 학습하는 과정에서 자동차에 대한 정보를 처리하도록 바뀜

2.2 “새로운 정보 학습”의 필요성

= Continual Enlargement of Data! (현실에서 데이터는 급속도로 계속 성장한다.)



1) 현실에서는 연구의 방향이나 시장의 수요에 따라 데이터/클래스가 세분화 혹은 계속적으로 증가하고, 이에 따라 새로운 task도 추가된다.

이때, building individual model for each newly added task is too expensive!!

<Task Update History>

[2013.01]

Task1: 빌딩 유무 Classification

[2020.04]

Task1: 빌딩 유무 Classification

Task2: 신호등 유무 Classification

[2021.01]

Task1: 빌딩 유무 Classification

Task2: 신호등 유무 Classification

Task3: 감시 카메라 유무 Classification

[2022.02]

Task1: 빌딩 유무 Classification

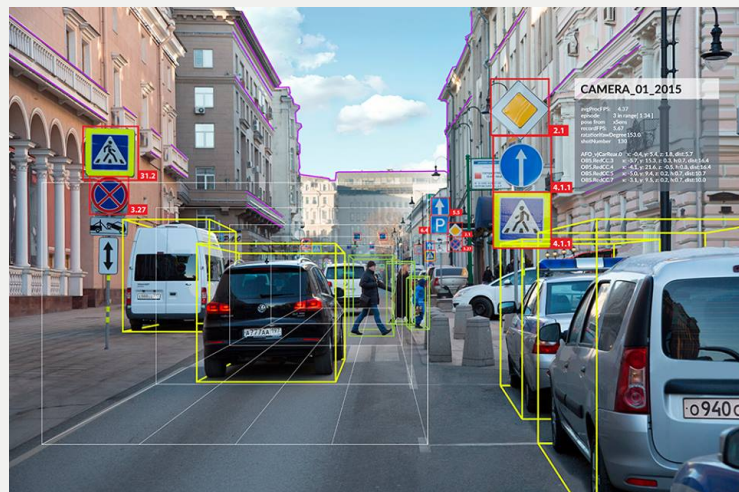
Task2: 신호등 유무 Classification

Task3: 감시 카메라 유무 Classification

·
·
·

Task999: 수소차 유무 Classification

→ ~~build 999+ models...~~?



3. Continual Learning Setting

1. **전체 데이터로 재학습 불가능:** 학습에 사용된 subset데이터는 저장에 불가능하여 더 이상 사용할 수 없음. (=과거 데이터 재활용 불가능!)
2. 여러 task를 하나의 모델에 **순차적으로 학습하여 최종적으로 모든 task의 수행이 가능한** 모델을 학습

4. Essentials of continual learning

- **Efficiency!**
 - Memory
 - Computation cost
 - Time
- **Prevent Negative-transfer!**
 - Multi-task를 해결 할 때 Task끼리 서로 안 좋은 영향을 주는 것(=negative transfer)을 막는다
- **Remember previously learned information!**