

인공 신경망의 Catastrophic forgetting 현상 극복을 위한 순차적 반복 학습에 대한 연구

¹최동빈, ^{*2}박용범

¹ 단국대학교, 72180417@dankook.ac.kr
^{*2}교신저자 단국대학교, ybpark@dankook.ac.kr

A study on sequential iterative learning for overcoming catastrophic forgetting phenomenon of artificial neural network

¹Dong-bin Choi, ^{*2}Young-beom Park

¹ Dankook University, 72180417@dankook.ac.kr
^{*2} Corresponding Author Dankook University, ybpark@dankook.ac.kr

요 약

현재 인공신경망은 단일 작업에 대해선 뛰어난 성능을 보이거나, 다른 종류의 작업을 학습하면 이전 학습 내용을 잊어버리는 단점이 있다. 이를 catastrophic forgetting이라고 한다. 인공신경망의 활용도를 높이긴 위해선 이 현상을 극복해야 한다. catastrophic forgetting을 극복하기 위한 여러 노력이 있다. 하지만 많은 노력이 있었음에도 완벽하게 catastrophic forgetting을 극복하지는 못하였다. 본 논문에서는 여러 노력 중 elastic weight consolidation(EWC)에 사용되는 핵심 개념을 이용하여, 순차적 반복학습을 제시한다. 인공신경망 학습에 많이 쓰이는 MNIST를 확장한 EMNIST 데이터 셋을 이용하여 catastrophic forgetting 현상을 재현하고 이를 순차적 반복학습을 통해 극복하는 실험을 진행하였으며, 그 결과 모든 작업에 대해서 학습이 가능하였다.

Abstract

Currently, artificial neural networks perform well for a single task, but NN have the problem of forgetting previous learning by learning other kinds of tasks. This is called catastrophic forgetting. To use of artificial neural networks in general purpose this should be solved. There are many efforts to overcome catastrophic forgetting. However, even though there was a lot of effort, it did not completely overcome the catastrophic forgetting. In this paper, we propose sequential iterative learning using core concepts used in elastic weight consolidation (EWC). The experiment was performed to reproduce catastrophic forgetting phenomenon using EMNIST data set which extended MNIST, which is widely used for artificial neural network learning, and overcome it through sequential iterative learning.

Keywords: Catastrophic forgetting, Artificial neural network, sequential iterative learning, LeNet, EMNIST

I. 서론

* Corresponding Author

Received: Aug. 28, 2018, Revised: Nov. 01, 2018, Accepted: Dec. 27, 2018

Legg and Hunter[1]는 인공지능이 범용으로 사용되기 위해서는 다양한 작업을 학습하고 기억해야한다고 하였다. 하지만 catastrophic forgetting 으로 인해 다양한 작업에 대한 학습이 어렵다. Catastrophic forgetting 이 일어나는 이유는 인공신경망의 메모리라 할 수 있는 가중치들이 매 학습마다 변경되기 때문이다[2].

이를 극복하기 위한 여러 노력들이 있다. Kemker, Abitino, McClure, and Kanan[3]은 5 가지로 분류하고 성능을 테스트 하였다. 하지만 그 결과는 catastrophic forgetting 을 완전하게 극복하지 못한 것으로 나타났다. EWC[4]의 방식이 가장 근접하였다.

베이지안 인공 신경망의 경우 가중치를 확률 변수로 표현한다[5]. 단일 작업을 학습하여 조절된 가중치들 중에 다른 작업에 대해서도 활용이 가능한 경우가 있다. 만약 인공 신경망의 모든 가중치를 다른 작업에도 활용할 수 있도록 설정할 수 있다면, 다중 작업에 대한 학습이 가능하다.

EWC는 기존 학습된 가중치를 다른 작업을 학습해도 공유 되는 방향으로 유도하는 방정식을 고안하여 이를 구현하였다. 하지만 의도한 대로 가중치 이동이 되지 않고, 오히려 가중치 변경에 제약이 걸리기에 마지막으로 학습한 작업에 대해서는 오히려 일반적인 학습을 한 인공신경망보다 낮은 결과를 낸다[5].

가중치에 대한 제약을 가하지 않으면서도 EWC 에서 언급한 작업 간의 공통부분으로 가중치를 조절해 보고자 한다. 그러기 위한 방법으로 순차적 반복 학습을 이용해 볼 것이며, 그 성능을 검증하기 위한 시험을 제시한다.

순차적 반복학습에 쓰인 신경망은 MNIST 인식에 뛰어난 CNN 중 하나인 LeNet[6]을 사용하며, 학습 데이터로는 MNIST 와 유사하나 좀 더 많은 데이터를 보유한 EMNIST 를 사용한다[7].

본 논문의 구성은 다음과 같다. 2 장에서는 순차적 다중학습을 방해하는 catastrophic forgetting 과 이를 극복하기 위한 방법들을 알아본다. 3 장에서는 본 논문에서 제안하는 순차적 반복학습에 대해서 설명하며, 성능을 검증하기 위한 실험은 제안한다. 4 장에서는 실험에 대한 결과를 분석하고 5 장에서 결론을 내린다.

II. 관련연구

2.1. Catastrophic forgetting

인공지능을 범용으로 활용되기 위해서는 각기 다른 작업을 순차적으로 학습이 가능해야한다. 하지만 인공신경망에 순차적으로 각기 다른 작업을 학습 시켰을 시 이전 학습 능력을 잃어버린다. 이를 catastrophic forgetting 이라고 한다[4].

이 현상이 일어난 원인은 새로운 작업에 대한 학습이 진행 되면서 이전에 작업이 맞춰진 가중치들이 변하기 때문이다. 그림 1 에서도 나타나듯이 새로운 작업에 대한 학습이 진행되면서 기존 학습된 내용을 점차 잃어버린다.

위 현상을 극복하기 위해 여러 노력들이 있어 왔다. 그리고 Ronald Kemker et al[3] 표 1 과 같이 그러한 노력들은 크게 5 가지로 분류하고 각 방법을 대표하는 방식을 선정하여 그 성능을 테스트 하였다.

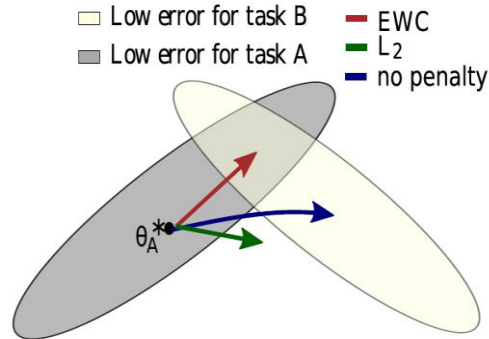
Ronald Kemker et al[3]이 분류한 5 가지 방식과 그 방식을 대표하는 방법은 다음과 같다. 1) Regularization - EWC, 2)Ensemble - PathNet, 3)Reheasl - GeppNet, 4) Dual-Memory - GeppNet+STM 5) Sparse-coding - FEL. 그리고 테스트한 결과는 표 1 과 같다. 결과에서 보듯이 EWC 의 문제 해결 접근 방식이 가장 가능성이 있다.

Table 1. Part of results of Kemker, Abitino, m McCulre and kanan[3]

Model	Multi-Model Avg Acc
MLP	0.600
EWC	0.913
PathNet	0.666
GeppNet	0.275
GeppNet+STM	0.222
FEL	0.453

2.2. Elastic weight consolidation

Catastrophic forgetting 을 극복하기 위해 Kirkpatrick et al[4]이 제안한 방식이다. 주요 내용은 그림 1 에 나와 있다. 신규 학습을 진행할 시 기존 가중치를 유지하면서도 신규 학습에 요구되는 가중치 부분으로 조절하는 방식이다(그림 1). 이를 위해 가중치 조절하는 방식은 수식 1 에 따른다.

**Figure 1.** Weight adjustment method in EWC[4]

Equation 1. EWC's equation

$$L(\theta) = L_B(\theta) + \sum_i \frac{\lambda}{2} F_i (\theta_i - \theta_{A,i}^*)^2$$

손성호, 김지섭, 장병탁[5]은 EWC 를 이용하여 순차적 다중작업 학습을 테스트하였다. 테스트 결과는 일반 인공신경망 보다 기존 작업들 성능을 유지하는 면에서는 뛰어나다. 하지만 어느 정도 유지는 되었더라도, 해당 작업만 학습한 인공신경망의 성능에 비해서는 많이 낮다. 또한 가중치의 변경에 제약을 가하였기에 마지막으로 학습한 작업에 대해서도 일반 인공신경망에 비해 성능이 낮은 편이다.

인공지능을 범용으로 활용되기 위해서는 다중 학습이 가능해야 한다. 이는 인공신경망이 학습한 모든 작업에 대해서 높은 성능을 가져야 한다는 것을 의미한다. EWC 는 비록 catastrophic forgetting 을 어느 정도 극복하기는 하였으나, 모든 작업에 대해서 단일 작업을 학습한 인공 신경망 보다 낮은 성능을 보이기에, 완벽한 해답이라고는 할 수 없다.

III. 순차적 반복 학습을 통한 다중 학습 성능 측정

3.1. 순차적 반복 학습

각기 다른 작업에 대해 학습을 순차적으로 진행하면, catastrophic forgetting 으로 인해 이전 학습에 대한 성능이 하락된다. 그림 2 에서 왼쪽 그래프에서 나타나듯 단일 작업에 대해서 성공적으로 학습하여, 높은 성능을 보인 인공 신경망이지만, 다른 작업을 학습하는 순간 그림 2 오른쪽 그래프처럼 기존 작업에 대한 성능이 점차 하락하는 것을 볼 수 있다.

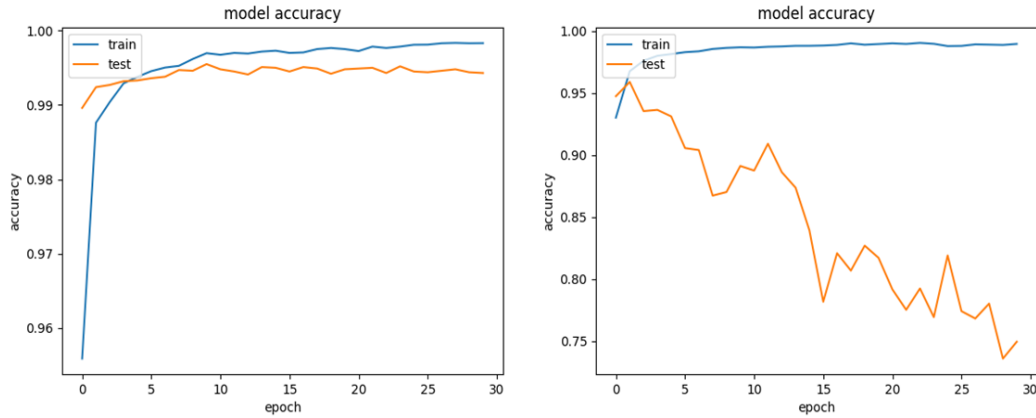


Figure 2. The left figure shows the test and the accuracy test for Task1, and the right is the test for Task1 in the process of learning about Task2. As the learning of Task2 is performed, it can be confirmed that the accuracy of Task1 is lowered.

순차적 반복 학습은 그림 3 에서 보이듯 단일 작업에 대한 학습을 완전하게 진행하고 다음 작업에 대한 학습을 시작한다. 그리고 이를 지속적으로 반복하는 것이다. 순차적 반복학습을 통해서 EWC에서 언급하였듯 인공신경망의 가중치들을 다른 작업과 공유 할 수 있는 부분으로 수렴이 가능하며, 이를 통해서 인공 신경망의 가중치 들이 점차 모든 작업에 활용될 수 있게 변화 된다. 모든 가중치들이 제대로 설정이 되면 다중 학습이 완료 된다

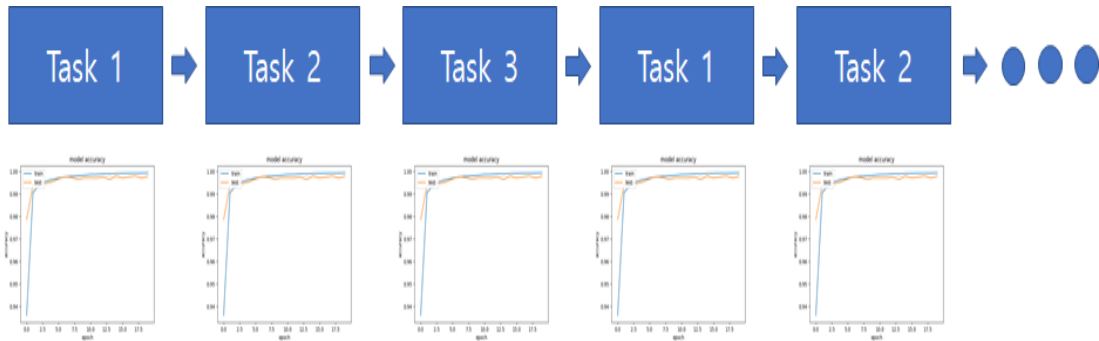


Figure 3. Learn each task completely and then learn the next task.

3.2. 실험 환경

순차적 반복 학습에 대한 성능을 테스트하기 위해 설계한 실험은 다음과 같다.

다른 작업으로 순차적 학습을 구현하기 위해서 MNIST 와 유사하나 데이터를 더 많이 보유한 EMNIST 를 사용하였다[7]. EMNIST 는 알파벳과 숫자로 이루어져 있으나, 본 실험에서는 EMNIST 데이터 중 숫자 데이터를 사용하였다. 순차적으로 각기 다른 학습을 진행해야 하기에 EMNIST 숫자 데이터를 3 그룹으로 분류 하였다.

숫자 데이터를 나눈 방법은 그림 4 와 같이 숫자 1-3, 4-6, 7-9 를 한 그룹씩 총 3 개의 그룹으로 나누었다. 숫자 0 그룹은 각 그룹 구성에 동일성을 주기 위해 제외 하였다. 각 그룹은 72000 개의 학습 데이터, 12000 개의 테스트 데이터를 가지고 있다.

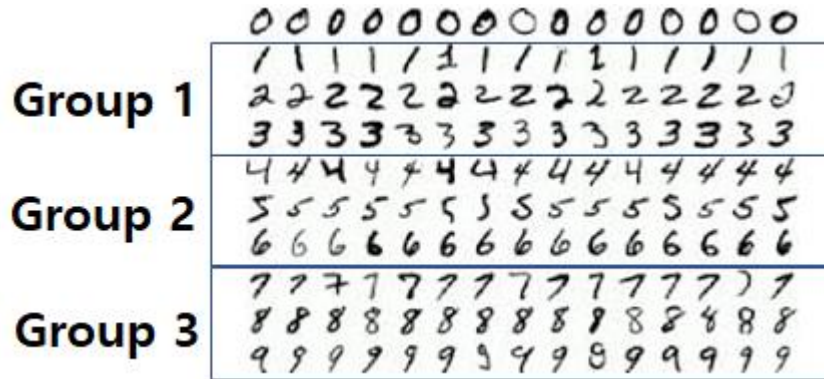


Figure 4. For sequential multi - learning environment, EMNIST numerical data are classified into 3 groups.

인공 신경망이 3 그룹으로 나누어진 실험 데이터 셋을 각기 다른 작업으로 인식하고, 그로 인해 catastrophic forgetting 이 나타나는지 확인하기 위해 순차적으로 학습 시키고, 그룹 3 까지 학습한 인공신경망에게 그룹 1, 2, 3 의 테스트 데이터로 확인해 보았다. 3 번 정도 반복하여 나타난 결과는 표 3 과 같다.

Table 2. Experimental results confirmed that the data sets divided into three groups were recognized as different tasks.

Group 1	Group 2	Group 3
0	0	0.99825
0	0	0.998333
0.002833	0.001333	0.998583

표 3 에서 알 수 있듯이 인공신경망은 각 그룹의 데이터를 각기 다른 작업으로 인식하며, catastrophic forgetting 이 일어난 모습을 확인 할 수 있다.

학습에 사용되는 모델은 MNIST 학습에 뛰어난 성능을 보이는 convolutional neural networks 구조를 가지는 LeNet 을 사용하였다[5]. LeNet 모형이 가진 단일 작업에 대한 성능은 그림 5 와 같다. Catastrophic forgetting 을 극복하게 된다면 각 그룹에 대한 정확도는 단일 작업을 학습 했을 시와 비슷한 정확도를 지니게 될 것이다.

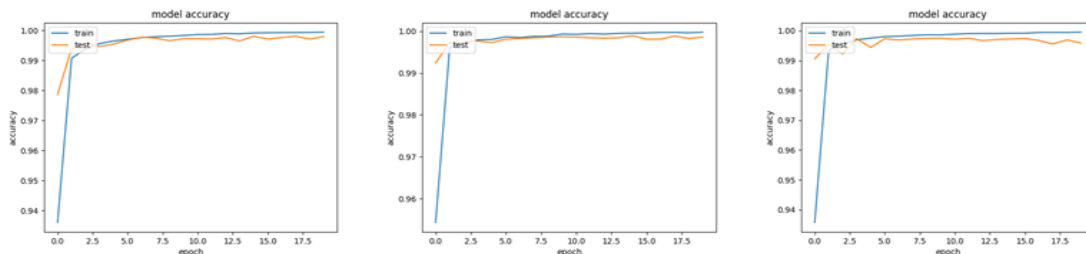


Figure 5. Result of learning each group data with LeNet. All have more than 99% accuracy.

각 그룹에 대해서 EPOCH 는 20, Batch size 는 200 으로 학습하였고, 그룹 1, 그룹 2, 그룹 3 순서로 학습이 진행되었다. 이것을 지속적으로 반복하면 그림 6 과 같은 결과가 나온다.

IV. 실험 결과 및 분석

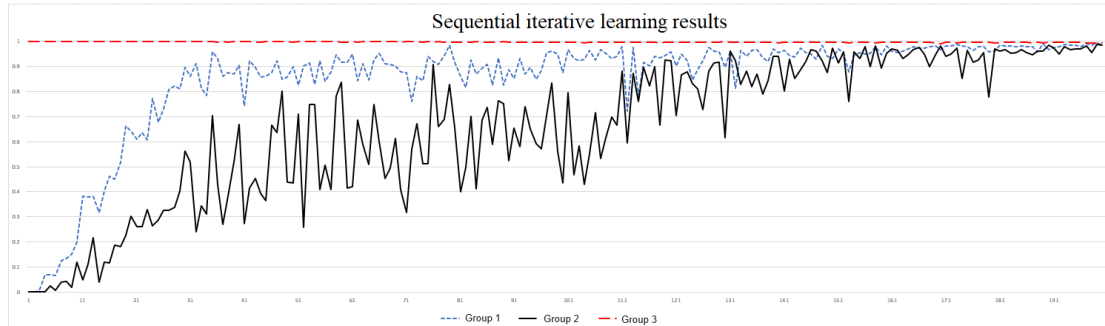


Figure 6. Sequential iterative learning results.

각 그룹에 대한 학습은 그림 5 와 같이 99%이상의 정확도를 보이거나 다음 그룹을 학습하면 이전 그룹의 학습 내용이 catastrophic forgetting 으로 인해 매우 저하 된다. 본 실험에서는 그룹 1, 2 에 대한 표 3 과 같이 정확도가 0 으로 나타났다.

순차적 반복학습을 사용하여 지속적으로 학습하면 가중치가 조절되어 그룹 1, 2 에 대한 정확도가 점차적으로 상승하는 것을 볼 수 있다. 기존 연구들과 다르게 가중치 변동에 대한 제약을 걸지 않았기에 그룹 2 의 학습이 그룹 1 보다 이후에 학습 되었음에도 정확도의 변동이 그룹 1 에 비해 매우 크다.

하지만 가중치에 대한 제약이 없기에 마지막으로 학습하는 그룹 3 에 대한 정확도는 99% 이상을 유지한다. 이는 다른 방법에 비해 순차적 반복 학습이 가지는 장점이라고 할 수 있다. 총 반복 회수는 200 번이며 마지막 5 번의 결과는 표 3 과 같다.

Table 3 Sequential iterative learning results.

Repeat	Group 1	Group 2	Group 3
196	0.96825	0.985917	0.99525
197	0.977083	0.972083	0.995833
198	0.97775	0.948167	0.996917
199	0.988	0.979833	0.995583
200	0.984917	0.965667	0.996333

결과적으로 인공신경망의 가중치들이 작업들의 공통부분으로 수렴하여, 모든 작업에 대해서 높은 성능을 가지게 되었다. 이전 작업에 대한 높은 정확도는 물론 가중치 변경에 대한 제약이 없기 때문에 마지막으로 학습한 작업에 대한 정확도 또한 높다.

V. 결론

인공지능을 범용으로 활용하기 위해선 각기 다른 작업을 순차적으로 활용할 수 있어야한다. 하지만 catastrophic forgetting 을 극복하지 않으면 불가능 하다. 본 논문은 이를 위해 순차적 반복 학습을 제안한다. 순차적 반복 학습으로 각 작업에 대한 높은 수준의 성능을 유지가 가능하다.

또한 catastrophic forgetting 을 극복하기 위해 제시 되었던 다른 방식과 달리 가중치 조절에 제약이 없기에 마지막으로 학습하는 작업에 대해 높은 성능 그대로 유지가 가능하다.

다만, 그룹 3에 대한 높은 정확도는 Babyak MA[8]가 언급한 overfitting 에 대한 우려가 있다. 이에 대해선 추후에 연구가 필요하다. 또한 반복 학습이 진행되면서 catastrophic forgetting 으로 인해 낮아진 성능이 점차 높아진 이유는 EWC 에서 말하는 가중치의 교차점 부분으로 인공 신경망의 가중치가 조절되기 때문이라고 추측은 가능하나, 이에 대한 정확한 근거를 제시하진 못하였다.

그룹 1, 2, 3 순으로 학습이 되기에 정확도 또한 그룹 3, 2, 1 순으로 높을 것이라 예상과 다르게 그룹 2의 정확도가 낮으며, 반복학습 과정에서 변동치도 그룹 1과 다르게 매우 큰 편이다. 학습 도중에 가중치가 어떠한 방식으로 이동되는지에 대한 대답도 확실하진 않다. 이에 대한 설명과 원인 규명 역시 추후 연구에서 이루어 져야 할 것이다.

VI. 감사의 글

이 논문은 2018 년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 지역신산업선도인력양성사업 성과임(No.NRF-2016H1D5A1909989)

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학 ICT 연구센터 육성 지원사업의 연구결과로 수행되었음" (IITP-2018-2017-0-01628)

VII. 참고문헌

- [1] Shane Legg and Marcus Huntter. Universal intelligence : A definition of machine intelligence. *Mind and Machines* 17(4): 391-444, 2007
- [2] Michael McCloskey and Neal J Cohen. Catastrophic interference in connectionist networks: The sequential learning problem. *The psychology of learning and motivation*, 24(109-165):92, 1989
- [3] Ronald Kemker, Angelina Abitino, Marc McClure, and Christopher Kanan. Measuring Catastrophic Forgetting in Neural Networks. arXiv preprint arXiv:1708.02072, 2017.
- [4] Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; Veness, J.; et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proc. of the National Academy of Sciences* 201611835.
- [5] Seongho Son, Jiseob Kim, Byoing-Tak Zhang “Sequential Multitask Learning Optimization Using Bayesian Neural Network” *KIISE Transactions on Computing practices*. Vol. 24. No. 5. Pp. 251-255. 2018. 5
- [6] Y. LeCun and Y. Bengio *Convolutional Networks for Images Speech, and Time-Series, brain theory neural networks*, vol. 3361, 1995.
- [7] Cohen, G., Afshar, S., Tapson, J., & van Schaik, A. (2017). EMNIST: an extension of MNIST to handwritten letters. Retrieved from <http://arxiv.org/abs/1702.05373>
- [8] Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med.* 2004;66(3):411-421