



## 다층신경망의 다양한 연결구조 제안 및 분석

김미숙\*, 강태원\*\*

### Proposal and Analysis of Various Link Architectures in Multilayer Neural Network

Mi-Suk Kim\*, Tae-Won Kang\*\*

#### 요 약

신경망은 생물학적 뇌 구조와 동작을 모사한 계산모델이다. 가장 흔하게 사용하는 신경망은 여러 개의 레이어 즉, 층으로 구성된 다층 전진전파 신경망이고 학습은 오류역전파 알고리즘을 사용한다. 기존 다층신경망의 경우 레이어가 깊어질수록 가중치 수정량의 변화에 의해 학습성능이 나빠진다. 본 논문에서는 입력신호가 신경망 전체에 전달될 수 있도록 인접하지 않은 레이어의 뉴런들끼리 연결이 허용되는 신경망 연결구조를 제안하고 분석한다. 분석결과, 학습속도에 영향을 주지 않을 정도로 연결을 추가하여도 제안한 구조의 깊은 신경망이 그렇지 않은 경우보다 우수한 학습성능을 보임을 확인하였다.

#### Abstract

Neural networks are computational models that simulate biological brain structures and behaviors. The most commonly used neural network is a multilayer forward propagation neural network composed of several layers, ie layers, and learning uses error propagation algorithms. In the case of existing multilayer neural networks, the learning performance deteriorates due to the change in the weight modification amount as the layer becomes deeper. In this paper, we propose and analyze a neural network connection structure in which non adjacent layers neurons are allowed to connect to each other so that the inputs can be transmitted to the entire neural network. As a result of the analysis, it was confirmed that the deep neural network of the proposed structure shows better learning performance even if the connection is added so as not to affect the learning speed.

#### Keywords

multilayer neural network, vanishing gradient, connection structure, machine learning

\* 강릉원주대학교 컴퓨터공학과

- ORCID: <https://orcid.org/0000-0002-0762-4276>

\*\* 강릉원주대학교 컴퓨터공학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0003-4343-5517>

• Received: Nov 12, 2017, Revised: Mar. 25, 2018, Accepted: Mar. 28, 2018

• Corresponding Author: Tae-Won Kang

Dept. of Computer Science & Engineering, Gangneung-Wonju National Univ.

150, Namwon-ro, Heungeop-myeon, Wonju-si, Gangwon-do, 26403

Tel.: +82-33-760-8666, Email: [twkang@gwnu.ac.kr](mailto:twkang@gwnu.ac.kr)

## 1. 서론

신경망(Neural Network)은 생물학적 뇌 구조와 동작을 모사한 계산모델이다. 뇌의 전기화학적 신경신호 전달물질과 전달경로를 통하여 학습시킴으로써 지능적으로 행동할 수 있도록 한다[1]. 뇌는 뉴런(Neuron)들 사이에 수많은 시냅스(Synaptic)가 복잡한 그물처럼 연결되어있다. 뉴런의 연결강도에 따라 연결구조가 변하는 것을 뇌 가소성(Neuro-Plasticity) 이라한다. 뇌 가소성은 학습과 관계가 있다[2]-[4]. 생물학적 뇌는 그림 1과 같이 뉴런들 사이에 레이어의 구분이 없다. 또한, 각 레이어에 위치한 뉴런은 인접한 레이어의 뉴런들로만 연결되어 있지 않다. 하나의 뉴런에서 여러 가지 돌기(Dendrite)가 뻗어 나와 다른 뉴런으로 연결되는 복잡한 연결구조를 가진다.

신경망은 주어진 학습을 통하여 스스로 지식을 획득한다. 가장 흔하게 사용하는 다층신경망(Multilayer Neural Network)은 여러 개의 레이어 즉, 층으로 구성된 전진전파(Feed Forward) 신경망이고 오류역전파(Backpropagation) 학습알고리즘을 사용한다.

다층신경망 네트워크는 입력레이어, 은닉레이어, 출력레이어로 구성되어있다. 각 레이어 내의 연결과 출력레이어에서 입력레이어로 직접적인 연결은 존재하지 않는 전진전파 네트워크이다. 오류역전파 알고리즘은 각 레이어의 뉴런과 뉴런 사이의 연결강도 값을 조정한다. 각 레이어의 연결강도 값은 편미분의 연쇄법칙(Chain Rule)[5]을 이용한다.

또한, 다층신경망은 그림 2와 같이 계산상의 편의성을 위하여 인접한 뉴런들과 연결되는 완전연결 구조(Fully Connected)를 가지고 있다.

다층신경망은 복잡한 비선형문제를 해결하기 위하여 층이 많은 깊은 레이어를 사용한다. 레이어가 깊어질수록, 가중치 수정량(Gradient)이 미분에 의해 빠른 속도로 수렴한다. 수렴 속도가 빠를 경우, 학습성능이 저하되는 문제(Vanishing Gradient)가 발생한다[6]. 또한 허용오차 부근이나 오차 곡선이 깊은 계곡 형태인 경우, 실제 데이터에 대한 오차가 커져 학습성능이 저하되는 문제(Overfitting)가 발생된다

[7][8]. 그러나 학습된 신경망은 연결가중치에 오류가 생겨도 나름의 결과를 내놓는다[9].

본 논문에서는 입력신호가 네트워크 전체에 고르게 전달하기 위해 학습이 가능하고 인접하지 않은 레이어의 뉴런들과 연결이 허용되는 신경망 연결구조를 제안하고 분석한다. 2장에서 신경망의 연결구조 관련연구를 살펴본다. 3장과 4장에서는 제안한 연결구조를 알아보고 다층신경망과 비교하여 학습 성능을 분석한다. 마지막 5장에서는 연구의 성과 및 향후 방향을 제시한다.

## II. 관련연구

연결구조에 따른 기존 연구에서, 은닉 뉴런간의 연결구조에 따른 시계열 학습과 동역학을 분석한 연구[10]와 ResNet(Residual Network)[11][12]이 있다. 동역학 분석 연구의 경우, ESN(Echo State Network)을 기준으로 네트워크 모델에 따라 은닉 뉴런간의 연결구조로 구성되어있다. ESN의 구성은 일반적인 RNN(Recurrent Neural Network)과 동일하다. RNN의 경우 시간에 따른 Recurrent 레이어가 동일한 가중치를 가지므로 학습의 대상이 아니다.

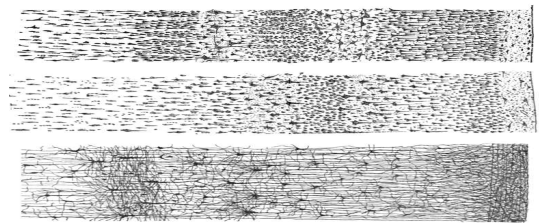


그림 1. 뇌 신경망피질  
Fig. 1. Cajal cortex drawing

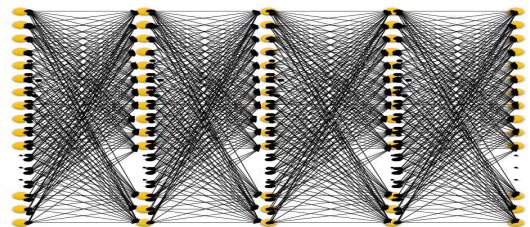


그림 2. MLP 모델의 신경망 구조  
Fig. 2. Neural network architecture of MLP model

또한, 최근 가장 많이 연구되고 있는 ResNet은 CNN(Convolution Neural Network)을 기반으로 연결 구조를 바꾼 경우이다. 일반적인 CNN은 입력 신호를 컨볼루션(Convolution) 레이어와 서브 샘플링(Subsampling) 레이어를 거쳐 출력을 내는 반면 ResNet은 출력 값과 입력 신호의 차이를 얻을 수 있도록 학습하게 된다. 즉, 입력 신호가 컨볼루션 레이어와 서브 샘플링 레이어를 건너뛰어 곧바로 출력으로 연결되는 연결선을 생성한다. 그러나 이러한 연결선은 가중치의 변화 없이 전달되므로 일반적인 신경망 연결과는 다르다. 즉, 학습의 대상이 아니다.

본 연구에서는 모든 연결이 학습이 가능하고 입력신호를 네트워크 전체에 효과적으로 전달되는 연결 구조를 제안한다.

### III. 먼 레이어 뉴런의 연결을 허용하는 구조

입력신호를 신경망 네트워크 전체에 고르게 전달하기 위해 인접하지 않은 레이어의 뉴런들과도 연결이 허용되는 신경망 연결구조를 제안한다. 데이터는 표준으로 사용되는 MNIST(Modified National Institute of Standards and Technology database) 데이터셋과 HandWritten 데이터셋을 사용하였으며 제안한 깊이별 연결구조는 아래와 같다.

#### 3.1 다양한 연결구조

제안한 연결구조는 다층신경망에 연결구조를 추가하여 구성하였다. 비교분석을 위하여, 기존 다층신경망을 SLA(Standard Link Architecture), 제안한 신경망 연결구조는 MLA(Modify Link Architecture)로 명명하였다. MLA 연결구조는 크게 두 가지로 분류하였다. 하나는 척도 없는 네트워크[12]-[15]의 허브(Hub)를 응용한 연결구조이다. 척도 없는 네트워크는 연결선수가 멱함수(Power Law)를 따르며 다수의 연결 수를 가진 허브가 존재하는 네트워크이다.

또 다른 하나는 뉴런 즉, 노드의 선택을 무작위 네트워크를 응용한 연결구조이다. 무작위 네트워크는 확률에 의해 연결선이 생성되는 통계학적 모델이다. 도수의 분포함수가 푸아송 분포를 따르는 Erdős-Rényi 모델이라고도 한다.

예를 들어, 입력과 출력레이어를 제외한 은닉레이어가 5개인 경우, SLA 연결구조를 제외한 제안한 다양한 MLA 연결구조는 다음과 같다.

- SMLA : SLA + 입력레이어로부터 각 은닉레이어의 첫 번째 노드에 연결
- SSMLA : SLA + 입력레이어로부터 다음 은닉레이어의 첫 번째 노드에 연결
- RMLA : SLA + 입력레이어로부터 다음 은닉레이어의 노드를 무작위로 선택한 연결
- RRMLA : SLA + 각 레이어로부터 다음 은닉레이어의 노드를 무작위로 선택한 연결

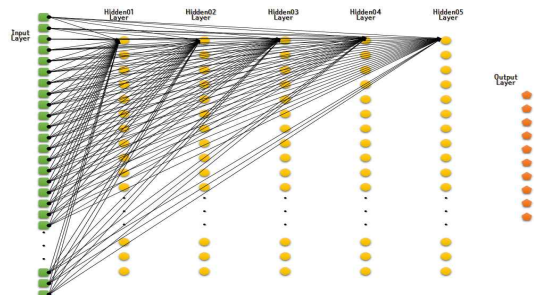


그림 3. SMLA 연결구조  
Fig. 3. SMLA link architecture

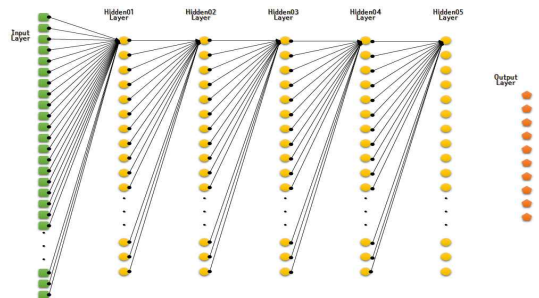


그림 4. SSMLA 연결구조  
Fig. 4. SSMLA link architecture

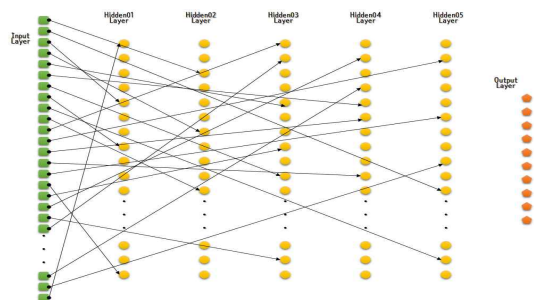


그림 5. RMLA 연결구조  
Fig. 5. RMLA link architecture

## 14 다층신경망의 다양한 연결구조 제안 및 분석

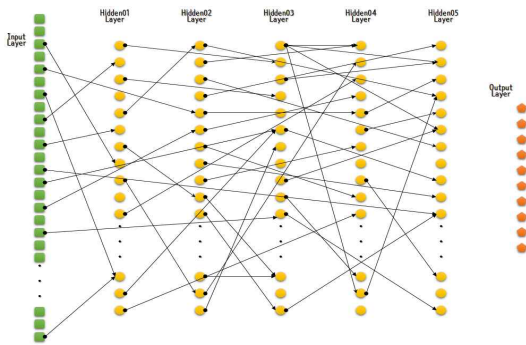


그림 6. RRMLA 연결구조  
Fig. 6. RRMLA link architecture

### 3.2 평가를 위한 문제 및 환경

제안한 연결구조의 성능을 분석하기 위한 데이터는 UCI Machine Learning Repository[16]의 Semeion Handwritten Digit Dataset[17]과 Gisette 데이터셋[18]을 비교분석하였다.

Semeion Handwritten Digit Dataset은 그레이스케일의 픽셀 16×16 경계상자를 사용하였다. 각 픽셀은 고정된 임계값을 사용하여 부울(0/1)값으로 스케일링되었다.

Gisette 데이터셋은 MNIST 데이터셋이라고도 불리며 다양한 이미지 처리 시스템을 교육하기 위해 일반적으로 사용되는 필기 자릿수 데이터베이스이다[19][20]. MNIST 데이터베이스는 기계 학습 분야의 교육 및 테스트에도 널리 사용되고 있다[21]. 또한, 그림 7과 같이 MNIST 흑백이미지는 픽셀 28×28 경계상자에 이미지 가장자리를 부드럽게 처리하는 안티앨리어싱(Anti-aliased)으로 정규화되어 있다.

연결구조의 입력은 각 숫자 이미지의 픽셀 값들의 벡터로 주어지고 출력은 0부터 9까지 10개의 숫자를 인식하는 문제이다. 각각의 훈련 데이터는 HandWritten 데이터셋은 1593개를 사용하였으며 테스트 데이터는 100개의 데이터를 사용하여 분석하였다. MNIST 데이터셋은 대용량의 데이터베이스로 60,000개를 사용하였으며 테스트 데이터는 10,000개의 데이터를 사용하였다. 각 깊이별 학습성능이 저하되는 문제를 분석하기 위하여 SLA와 MLA 연결구조를 학습률(Learning Rate), 모멘텀(Alpha)을 다양하게 바꾸어 10회 테스트의 평균으로 분석하였다.

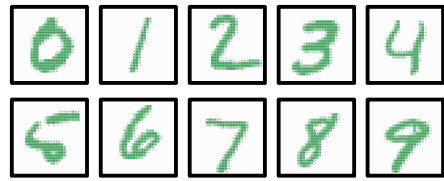


그림 7. MNIST Data 예  
Fig. 7. MNIST data example

레이어가 깊어질수록 학습성능이 저하되는 문제를 분석하기 위하여 은닉레이어를 다양하게 실험하여 보았다. 일반적으로, MNIST 데이터셋을 사용한 모델에서 학습성능이 유지되는 경우는 6인 레이어까지 사용한 경우이다[22]. 제안한 연결구조의 학습성능을 실험하기 위하여 은닉레이어를 5부터 10까지 깊이별로 분석하였다.

또한, 실제 데이터에 대한 오차가 커져 학습성능이 저하되는 문제를 개선하기 위해 훈련 데이터와 테스트 데이터를 달리하여 분석하였다.

## IV. 연결구조 학습성능 분석

제안한 MLA 연결구조는 기존 다층신경망 연결구조(SLA 연결구조)에 추가된 가중치 연결로 이루어져 있다. 각 연결구조는 은닉레이어의 개수별 평균 학습성능과 평균 학습속도를 비교분석하였다. 비교분석을 위하여 RMLA, RRMLA, SLA, SMLA, SSMLA 연결구조순으로 열거하였으며 은닉레이어는 레이어 5부터 10까지로 나누어 깊이별 평균 학습성능과 평균 학습속도를 실험하였다.

### 4.1 데이터셋에 따른 깊이별 학습성능

기존 다층신경망인 SLA 연결구조와 제안한 연결구조들의 데이터셋을 달리하여 보았다. 제안한 연결구조들의 HandWritten, MNIST 데이터셋의 깊이별 학습성능과의 관계를 알아보았다.

표 1은 HandWritten 데이터셋에서 은닉레이어의 깊이가 5인 경우에서 은닉레이어 10인 경우의 평균 학습성능을 비교하여보았다. SLA 연결구조의 경우 44.1%를 보였으며 RMLA 연결구조의 경우 레이어가 94.5%의 높은 학습성능 평균을 보였다.



표 1. HandWritten의 각 깊이별 평균 학습성능

Table 1. HandWritten, average learning performance for each layer

Link Architecture	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10	평균
RMLA	1.00	0.98	0.99	0.93	0.92	0.85	94.5%
RRMLA	0.98	0.95	0.94	0.87	0.32	0.51	76.0%
SLA	0.98	0.92	0.43	0.11	0.10	0.10	44.1%
SMLA	0.99	0.98	0.93	0.90	0.78	0.86	90.9%
SSMLA	0.99	0.98	0.95	0.90	0.87	0.91	93.4%

표 2. MNIST의 각 깊이별 평균 학습성능

Table 2. MNIST, average learning performance for each layer

Link Architecture	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10	평균
RMLA	0.86	0.88	0.86	0.81	0.84	0.82	84.6%
RRMLA	0.89	0.89	0.89	0.88	0.88	0.87	88.3%
SLA	0.87	0.86	0.77	0.42	0.10	0.10	52.0%
SMLA	0.87	0.87	0.84	0.80	0.79	0.72	81.5%
SSMLA	0.90	0.88	0.72	0.49	0.10	0.14	54.0%



그림 8. HandWritten의 각 깊이별 평균 학습성능 변화  
Fig. 8. HandWritten, average learning performance change by each layer

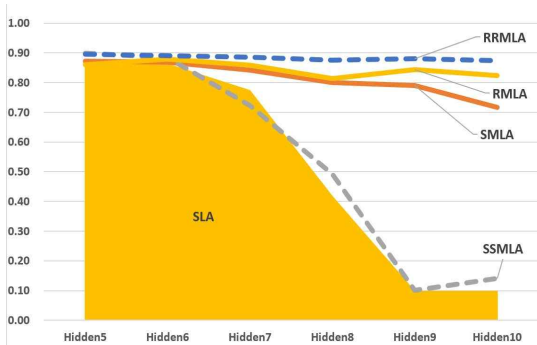


그림 9. MNIST의 각 깊이별 평균 학습성능 변화  
Fig. 9. MNIST, average learning performance change by each layer

표 2는 MNIST 데이터셋에서 은닉레이어의 깊이가 5인 경우에서 은닉레이어 10인 경우의 평균 학습성능을 비교하여보았다. SLA 연결구조의 경우 52.0%를 보였으며 RRMLA 연결구조의 경우 레이어가 88.3%의 높은 학습성능 평균을 보였다.

그림 8과 그림 9를 살펴보면 레이어가 깊어질수록 기존 다층신경망인 SLA 연결구조는 학습성능이 현저하게 떨어진다. 반면, 제안한 연결구조의 경우 학습성능이 완만한 연결구조를 보이고 있다.

특히, 입력신호가 네트워크 전체에 고르게 전달 되기 위해 학습이 가능하고 인접하지 않은 레이어의 뉴런들과도 연결이 허용되는 신경망 연결구조 중 RMLA 연결구조는 HandWritten, MNIST 데이터셋 모두 우수한 학습성능을 보였다.

#### 4.2 데이터셋에 따른 깊이별 SLA와 RMLA 학습성능

데이터에 따른 깊이별 학습성능 분석결과, 기존 다층신경망 SLA 연결구조와 비교하여 데이터셋을 달리한 경우 RMLA 연결구조가 모두 우수한 학습성능을 보였다. 이를 토대로 HandWritten, MNIST 데이터셋의 매개변수 학습률을 0.1/0.5로 차이를 크게 하여 실험하였으며, 모멘텀(Alpha)은 0.5/0.8로 달리하여 은닉레이어 깊이별 학습성능을 알아보았다.

표 3과 표 4에서 보면 HandWritten 데이터셋의 SLA 연결구조는 은닉레이어의 깊이가 깊어질수록 매개변수와 무관하게 현저하게 떨어지는 것을 알 수 있었다. 반면, RMLA 연결구조는 매개변수를 달리하여도 평균 학습성능과는 무관함을 보였다.

표 3. HandWritten\_SLA의 매개변수와 평균 학습성능  
Table 3. HandWritten\_SLA, parameter and average learning performance for each layer

Learning Rate	Alpha	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10
0.1	0.5	0.9902	0.9858	0.6749	0.1083	0.1069	0.1051
0.1	0.8	0.9824	0.9205	0.4315	0.1082	0.1034	0.0997
0.5	0.5	0.8897	0.9208	0.4631	0.0966	0.1043	0.1065
0.5	0.8	0.4617	0.3517	0.1019	0.1018	0.1018	0.1018

16 다층신경망의 다양한 연결구조 제안 및 분석

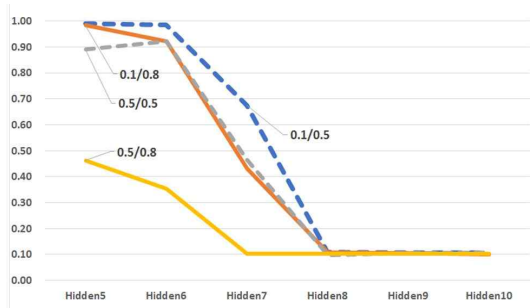


그림 10. HandWritten\_SLA의 매개변수와 평균 학습성능  
Fig. 10. HandWritten\_SLA, parameter and average learning performance

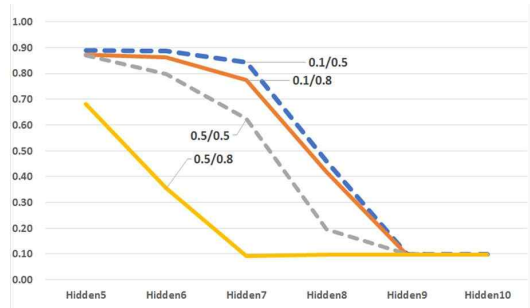


그림 12. MNIST\_SLA의 매개변수와 평균 학습성능  
Fig. 12. MNIST\_SLA, parameter and average learning performance

표 4. HandWritten\_RMLA의 매개변수와 평균 학습성능  
Table 4. HandWritten\_RMLA, parameter and average learning performance

Learning Rate	Alpha	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10
0.1	0.5	0.9899	0.9695	0.9830	0.9470	0.8640	0.8087
0.1	0.8	0.9955	0.9779	0.9868	0.9341	0.9222	0.8545
0.5	0.5	0.9733	0.9934	0.9608	0.9699	0.9711	0.9290
0.5	0.8	0.7472	0.8587	0.9170	0.9386	0.8913	0.7700

표 6. MNIST\_RMLA의 매개변수와 평균 학습성능  
Table 6. MNIST\_RMLA, parameter and average learning performance

Learning Rate	Alpha	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10
0.1	0.5	0.8791	0.8708	0.8781	0.8606	0.8277	0.8082
0.1	0.8	0.8602	0.8785	0.8591	0.8136	0.8428	0.8238
0.5	0.5	0.8770	0.8587	0.8536	0.8486	0.8255	0.8503
0.5	0.8	0.7530	0.7303	0.8095	0.7883	0.8060	0.8153

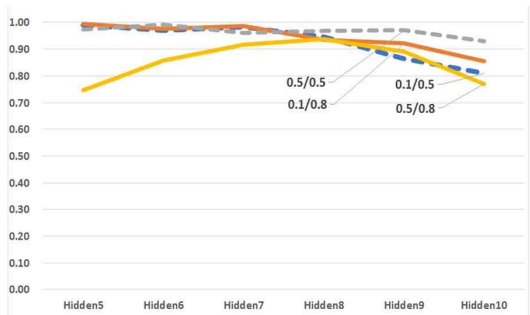


그림 11. HandWritten\_RMLA의 매개변수와 평균 학습성능  
Fig. 11. HandWritten\_RMLA, parameter and average learning performance

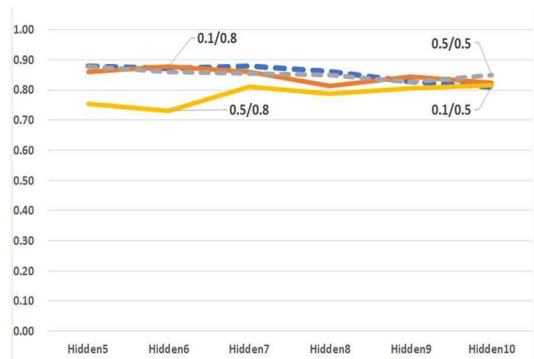


그림 13. MNIST\_RMLA의 매개변수와 평균 학습성능  
Fig. 13. MNIST\_RMLA, parameter and average learning performance

표 5. MNIST\_SLA의 매개변수와 평균 학습성능  
Table 5. MNIST\_SLA, Parameter and Average learning performance

Learning Rate	Alpha	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10
0.1	0.5	0.8880	0.8858	0.8428	0.4578	0.0991	0.0982
0.1	0.8	0.8720	0.8616	0.7729	0.4172	0.0980	0.0981
0.5	0.5	0.8708	0.7978	0.6233	0.1941	0.0986	0.0976
0.5	0.8	0.6814	0.3548	0.0910	0.0972	0.0979	0.0977

표 5와 표 6에서 보면 HandWritten 데이터셋과 유사하게 MNIST 데이터셋의 SLA 연결구조는 은닉 레이어의 깊이가 깊어질수록 매개변수와 무관하게 현저하게 떨어지는 것을 알 수 있었다. 반면, RMLA 연결구조는 매개변수를 달리하여도 평균 학습성능과는 무관함을 보였다.

### 4.3 가중치 연결 수에 따른 학습속도

제안한 연결구조는 SLA 연결구조에 추가된 가중치 연결로 이루어져있다. MNIST 데이터셋의 경우 입력신호가 28×28(784개)의 입력신호를 가지므로 연결구조에 따라 추가된 연결선의 수는 많아지게 된다. MNIST 데이터셋에서 SLA 연결구조를 기준으로 추가된 연결 수와 학습속도 관계를 알아보았다.

표 7. 각 깊이별 평균 가중치 연결수  
Table 7. Average weighted link for each layer

Link Architecture	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10
RMLA	67,984	72,080	76,176	80,272	84,368	88,464
RRMLA	67,984	72,080	76,176	80,272	84,368	88,464
SLA	67,200	71,296	75,392	79,488	83,584	87,680
SMLA	71,120	76,000	80,880	85,760	90,640	95,520
SSMLA	68,240	72,400	76,560	80,720	84,880	89,040

표 8. 각 깊이별 평균 학습속도(Second)  
Table 8. Average learning performance timefor each layer

Link Architecture	Hidden 5	Hidden 6	Hidden 7	Hidden 8	Hidden 9	Hidden 10
RMLA	3,432	4,150	3,891	4,104	3,943	4,427
RRMLA	3,282	3,485	3,620	4,701	4,567	4,253
SLA	3,399	3,620	3,653	3,888	3,993	4,193
SMLA	3,813	3,567	3,950	4,425	4,419	4,690
SSMLA	3,572	3,565	3,648	3,867	4,660	4,324

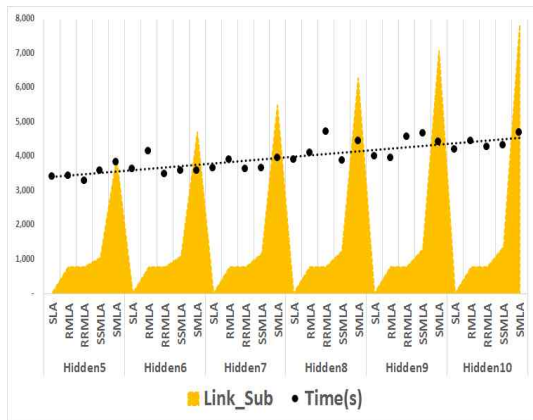


그림 14. 각 깊이별 가중치 연결 수와 학습속도  
Fig. 14. Number of weight connections and learning speed for each layer

표 7은 SLA 연결구조를 기준으로 각 레이어 깊이별 제안한 연결구조의 연결선 수를 나타낸다. 각 레이어별로 보면 SMLA 연결구조가 기존 SLA 연결구조에 비하여 많은 학습 연결선으로 표현된다.

표 8은 가중치 연결 수에 따른 학습 완료시간을 초(Second) 단위로 비교하여 보았다.

그림 14는 제안한 MLA 연결구조의 가중치 연결 수에 따른 각 깊이별 추가된 연결 수(Link\_Sub)와 학습속도(Time(s))의 관계를 나타낸다. 레이어가 깊어질수록 SLA를 기준으로 가중치 연결 수는 각 연결구조에 따라 증가하게 된다. 반면, 각 깊이별로 학습속도는 증가된 가중치 연결 수에 비하여 일정한 분포를 이루고 있다. 학습속도면에서, 추가된 연결 수는 기존 다층신경망의 학습속도와 무관하였다.

### 4.4 가중치 연결 수에 따른 학습성능

SLA 연결구조와 제안한 MLA 연결구조들의 평균 학습성능을 실험하였다. 먼저, 추가된 가중치 연결수와 학습성능의 관계를 알아보았다. 특히, 레이어의 깊이와 학습성능의 관계를 확인하기 위해 레이어 깊이를 다르게 하여 분석하였다.

그림 15는 제안한 MLA 연결구조의 가중치 연결 수에 따른 각 깊이별 연결 수(Link)와 학습성능(Average)의 관계를 나타낸다. SLA 연결구조의 경우, 은닉레이어가 8부터 급격하게 학습성능이 떨어진다. 반면, RMLA와 RRMLA 연결구조는 레이어가 깊어져도 추가된 가중치 연결 수에 비해 학습성능이 떨어지지 않았다.

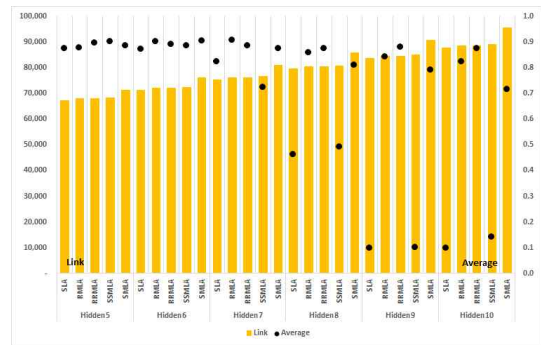


그림 15. 각 깊이별 가중치 연결 수와 학습성능  
Fig. 15. Number of weighted connections and learning performance for each layer

실험결과, 다층신경망에 가중치 연결수를 추가하더라도 학습속도와는 무관하였다. 또한, 가중치 연결을 무작위로 추가한 경우 레이어가 깊어져도 학습성능에 변화를 주지 않았다.

## V. 결론 및 향후 과제

본 논문에서는 입력신호를 네트워크 전체에 고르게 전달하기 위해 학습이 가능하고 인접하지 않은 레이어의 뉴런들이 서로 연결될 수 있는 신경망 연결구조를 제안하고 분석하였다.

제안한 다양한 연결구조는 기존 다층신경망에 (상대적으로 극히 적은) 연결을 추가한 것이다. 학습속도와 학습성능을 분석한 결과 제안한 연결구조는 기존 다층 신경망에 견주어 레이어가 깊어져도 학습속도와 무관하였다. 또한, 가중치 연결을 무작위로 연결한 경우 레이어가 깊어져도 우수한 학습성능을 유지하였다.

본 논문에서는 기존 다층신경망의 학습성능을 높이기 위해 신경망의 추가된 연결구조가 학습성능에 영향을 미칠 수 있다는 점에 의의를 찾아볼 수 있다. 추후, 다양한 활성화함수를 사용하여 성능을 분석하고 특히 기울기 소실 문제를 심도 있게 분석할 것이다. 또한, 딥러닝을 위한 다양한 데이터와 신경망 연결구조에 적용해 볼 것이다.

## References

- [1] Sung-Kwun Oh, "Computational Intelligence by Programming focused on Fuzzy, Neural Networks, and Genetic Algorithms", Naeha Press, pp. 225, Aug. 2002.
- [2] J. T. Wall, J. Xu, and X. Wang, "Human brain plasticity: an emerging view of the multiple substrates and mechanisms that cause cortical changes and related sensory dysfunctions after injuries of sensory inputs from the body", Brain Research Reviews, Vol. 39, No. 2-3, pp. 181-215, Sep. 2002.
- [3] R. G. Coss, J. G. Brandon, and Albert. Globus, "Changes in morphology of dendritic spines on honeybee calycal interneurons associated with cumulative nursing and foraging experiences", Brain Research, Vol. 192, No. 1, pp. 49-59, Jun. 1980.
- [4] F. H. Gage, "Structural plasticity: cause, result, or correlate of depression", Biol Psychiatry, Vol. 48, No. 8, pp. 713-714, Nov. 2000.
- [5] Robert G. Bartle and Donald R. Sherbert, "Introduction to REAL ANALYSIS", John Wiley & Sons, Inc., p. 208, 2006.
- [6] E. K. Blum, "Approximation of Boolean Functions By Sigmoidal Networks: Part I: XOR and other two-variable functions", Neural Computation, Vol. 1, No. 4, pp. 532-540, Winter 1989.
- [7] M. K. Weir, "A Method for Self-Determination of Adaptive Learning Rates in Back Propagation", Neural Networks, Pergamon Press, Vol. 4, No. 3, pp. 371-379, Dec. 1991.
- [8] Mohandes, M., Codrington, CW., and Gelfand S.B., "Two Adaptive Step-size Rules for Gradient Descent and their Application to the Training of Feedforward Artificial Neural Networks", Proceedings of IEEE International Conference on Neural Network, pp. 555-560, Jul. 1994.
- [9] Tae-Won Kang, "Fault Tolerance Analysis of the Neural Networks of Scale-Free Network Architecture", JKIIIT, Vol. 11, No. 7, pp. 151-157, Jul. 2013.
- [10] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology", Computer Communications Review 29, pp. 251-262, Sep. 1999.
- [11] Sangwoong Yoon and Byung-Tak Zhang, "Analyzing Performance and Dynamics of Echo State Networks Given Various Structures of Hidden Neuron Connections", KIISE Transactions on Computing Practices, Vol. 21, No. 4, pp. 338-342, Apr. 2015.
- [12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and



Jian Sun. "Deep residual learning for image recognition", arXiv preprint arXiv:1512.03385, 2015b.

- [13] Mark Buchanan, "Ubiquity: the science of history. or why the world is simpler than we think", Amazon, 2001.
- [14] D. J. Watts, "Small world: the dynamics of networks between order and randomness", Princeton, 1999.
- [15] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law", Contemporary Physics, Vol. 46, No. 5, pp. 323-351, Sep. 2005.
- [16] UCI Machine Learning Repository, <http://archive.ics.uci.edu/ml/index.php>, [Accessed: Jan. 10. 2018]
- [17] Semeion Handwritten Digit Dataset, <https://archive.ics.uci.edu/ml/datasets/semeion+handwritten+digit>, [Accessed: Jan. 10. 2018]
- [18] Gisette Dataset, <https://archive.ics.uci.edu/ml/datasets/Gisette>, [Accessed: Jan. 10. 2018]
- [19] X. Zhang and J. Zhou, "Support vector machines speed pattern recognition - Vision Systems Design", Vision Systems Design, Aug. 17. 2013.
- [20] S. Gangaputra, "Handwritten digit database", Retrieved 17, Aug. 2013.
- [21] Q. Yu, "THE MNIST DATABASE of handwritten digits", Retrieved 18, Aug. 2013.
- [22] MNIST database Error Rates, [https://en.wikipedia.org/wiki/MNIST\\_database](https://en.wikipedia.org/wiki/MNIST_database), [Accessed: Oct. 5. 2017]

### 강 태 원 (Tae-Won Kang)



1985년 : 연세대학교 수학과 (이학사)  
1988년 : 고려대학교 전산과학과 (이학사)  
1991년 : 고려대학교 수학과 (이학석사)  
1996년 : 고려대학교 컴퓨터학과

(이학박사)

1997년 ~ 현재 : 강릉원주대학교 컴퓨터공학과 교수  
관심분야 : 복잡계, 인공생명, 인공지능, 소프트 컴퓨팅

### 저자소개

### 김 미 숙 (Mi-Suk Kim)



2000년 2월 : 강릉원주대학교  
컴퓨터과학과(이학석사)  
2018년 2월 : 강릉원주대학교  
컴퓨터공학과(공학박사)  
2006년 8월 ~ 2017년 6월 :  
(유)코디아이 대표이사  
관심분야 : 인공지능, 인공생명,

빅데이터