

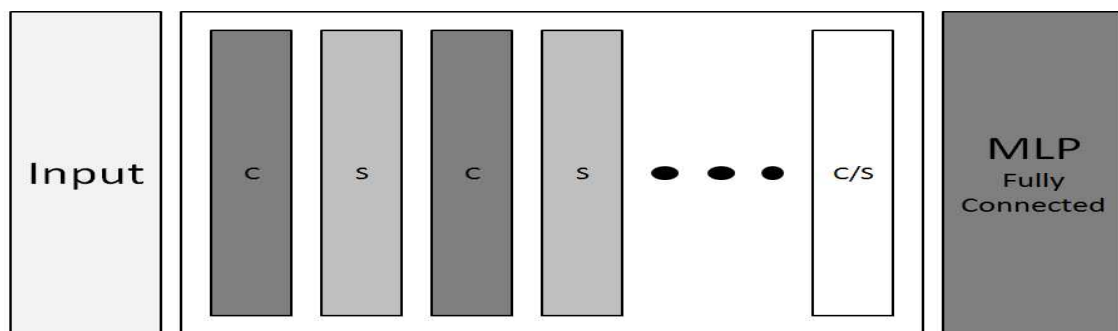
o Convolutional Neural Network를 이용한 필기체 인식 시스템

- Convolutional Neural Network 모델

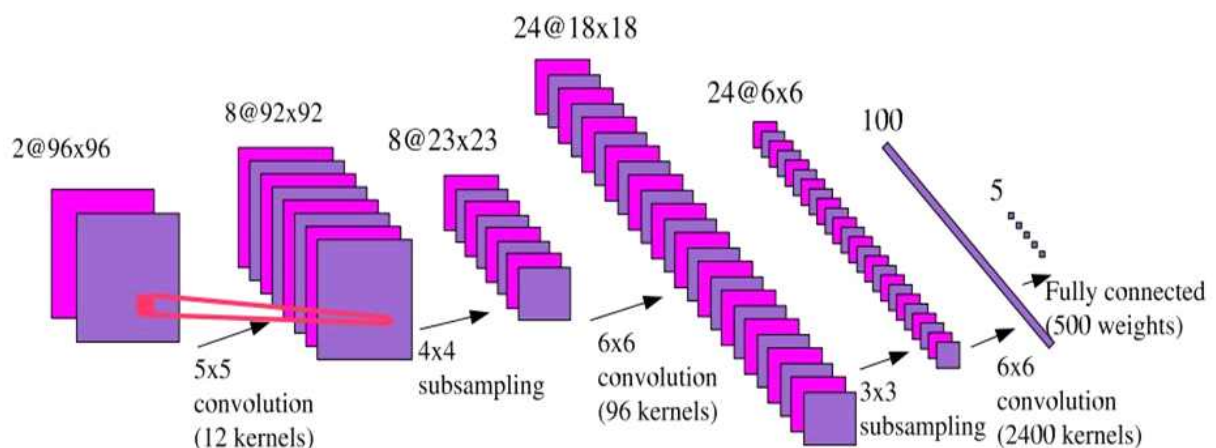
CNN은 Feed-Forward Neural Network 모델의 한 종류로서 각 뉴런이 연결되어 시각 영역을 중복적으로 응답하는 방식으로 된 구조이며 입력과 출력사이의 오차가 최소가 되도록 설계된 다층 퍼셉트론 형태로서 이미지에 관련한 패턴인식에서 사용되는 대표적인 Supervised Learning 방식의 Deep Learning 종류임.

- CNN의 Layer는 Convolutional Layer, SubSampling Layer로 분리됨. Abstract Feature는 Convolutional, SubSampling Layer의 순서로 추출됨(Pooling). 이는 Rotation, Scale에 Robust 한 특징을 추출하기 위한 연산임.

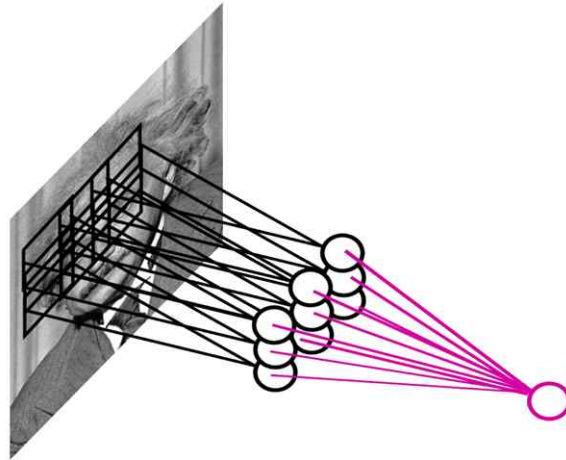
- CNN은 이미지 특정 필터 영역에서 픽셀의 Image Structure를 사용함. 이는 공유 웨이트를 가지며 이미지 Feature로 사용됨.



Basic Convolutional Neural Network Structure



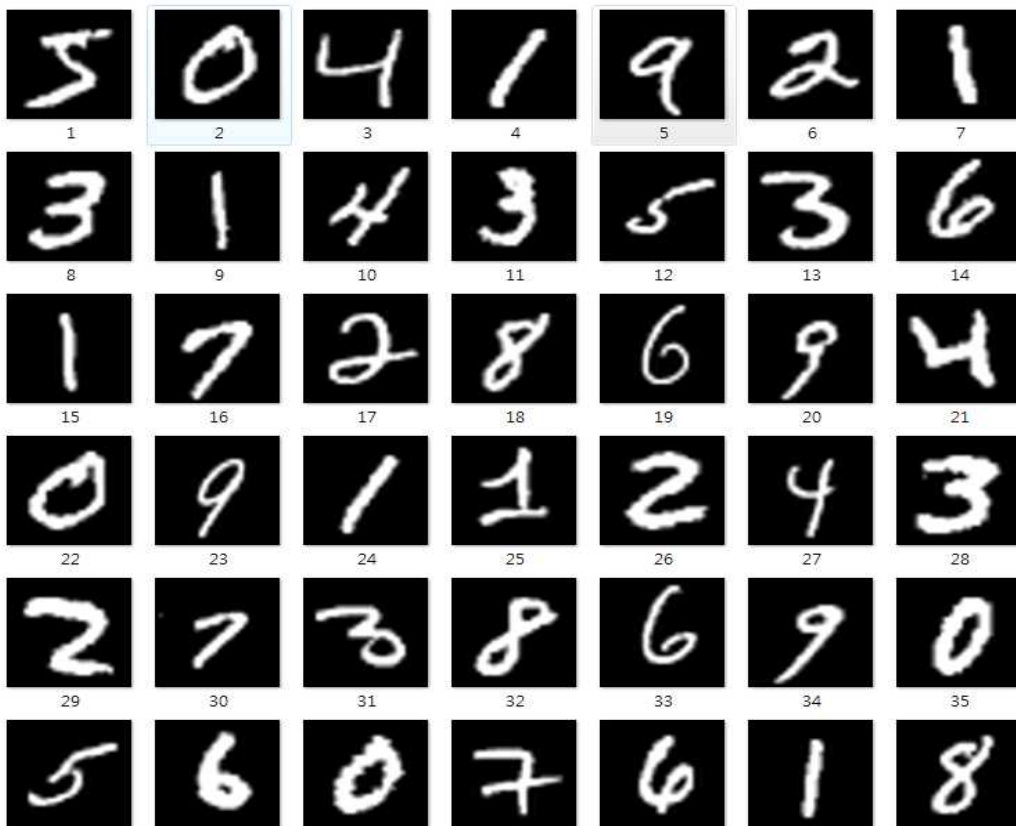
Convolutional Neural Network Structure



Abstract Feature

#### o 학습 데이터

- 범용적인 학습데이터를 이용하여 학습해야 범용적인 미들웨어로서의 가치를 보일 수 있고 성능 또한 높일 수 있음. LeCun 교수팀이 NIST 데이터베이스를 수정한 MNIST 숫자 데이터베이스를 사용함. 학습데이터 셋은 총 60,000개로 구성되어 있으며 이에 따라 실험은 1 Epoch 당 60000 Iteration 으로 구성됨.



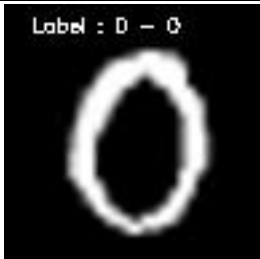
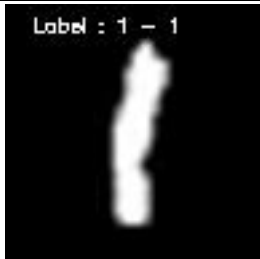
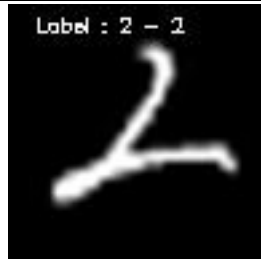
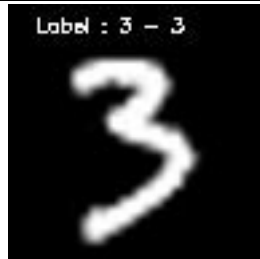
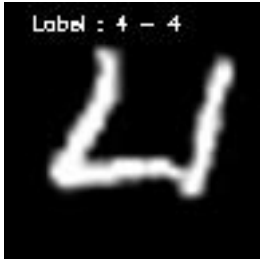
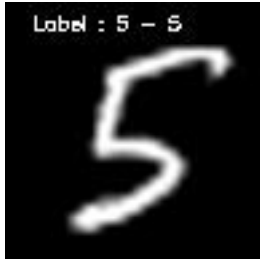
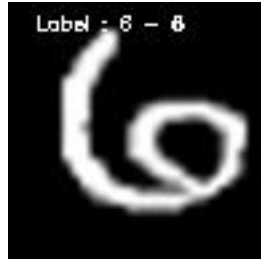
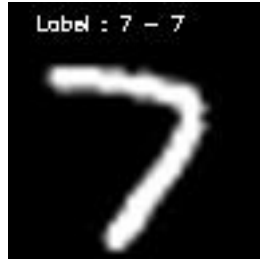

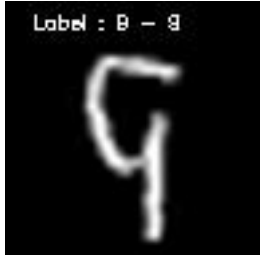
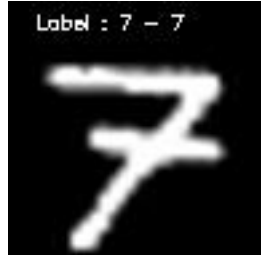
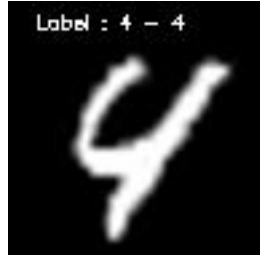
MNIST 숫자 학습 필기체 데이터 세트

o 실험

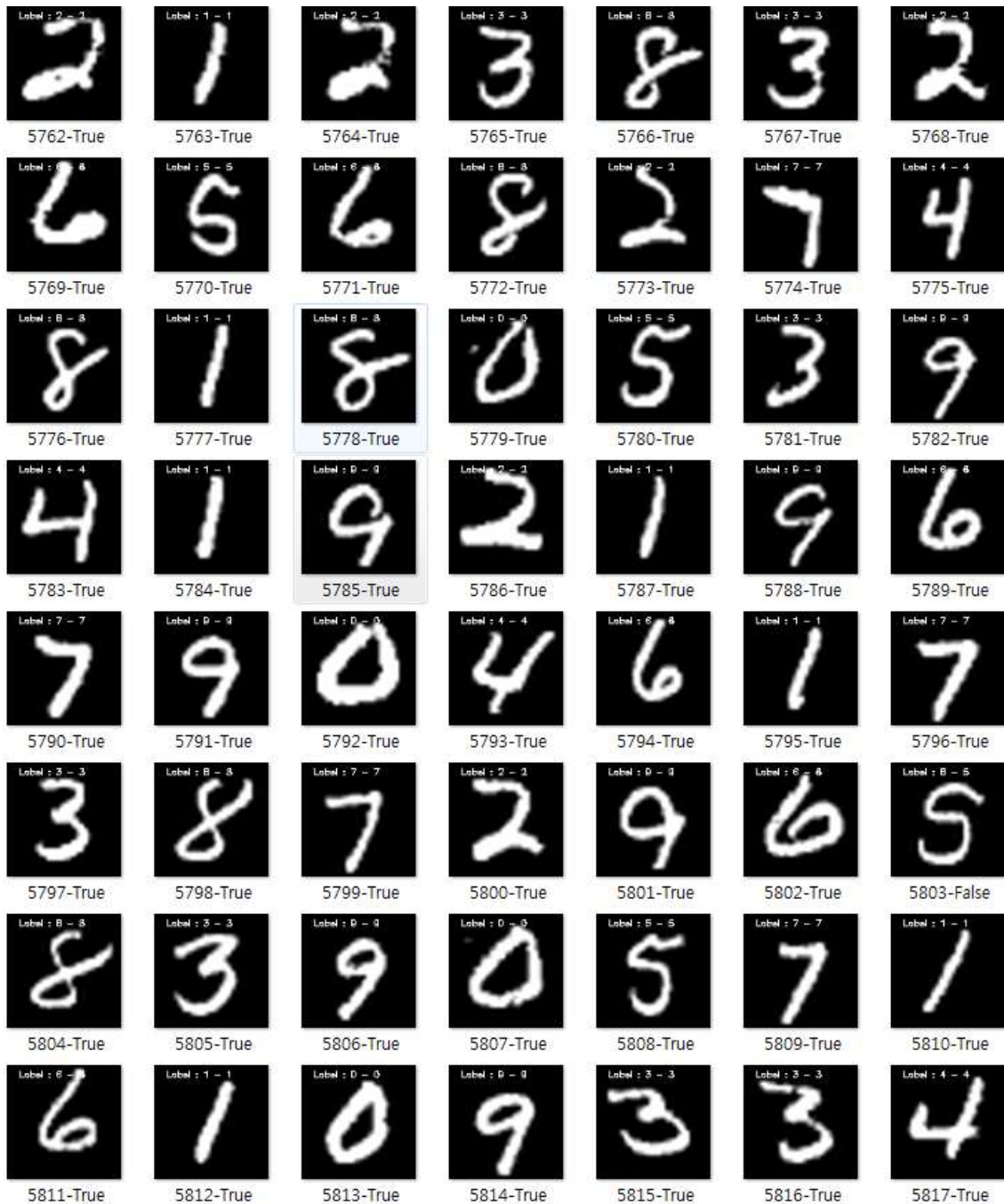
- CNN Layer의 구성은 Input Layer 1개, CNN Layer 8 - 8 - 24 - 24 - 100, Output Layer 10개로 구성함.
- Epoch 수 100회, 카테고리는 10개(0~ 9), Learning Gain 은 0.02, Error Rate는 0.002
- 실험은 MNIST 필기체 데이터베이스 중 10000개의 테스트 데이터베이스로 실험을 진행하였고 실험결과 True Percentage는 90.020001로 높은 인식률을 보였다.

7 True Count Number = 9002  
Percentage = 90.020001

테스트 결과 화면

|  |  |   |  |
|--|--|---|--|
| Label : 0 - 0<br>  | Label : 1 - 1<br>  | Label : 2 - 2<br>  | Label : 3 - 3<br>  |
| Label : 4 - 4<br> | Label : 5 - 5<br> | Label : 6 - 6<br> | Label : 7 - 7<br> |
| Label : 8 - 8<br> | Label : 9 - 9<br> | Label : 7 - 7<br> | Label : 4 - 4<br> |

실험 결과 샘플



실험 결과 샘플 - 2

#### o 결론

- CNN 학습 결과 60,000 개의 필기체 학습 데이터를 이용한 실험에서도 좋은 성능을 보였을 뿐 아니라 Convolution, SubSampling Layer(Pooling) 연산으로 인해 스케일, 회전 등에 강인한 특징을 추출해냄으로써 10,000개의 학습 데이터와 다른 복잡한 필기체 데이터베이스에서도 좋은 성능을 낼 수 있었음. 10,000개의 데이터 중에서 998개의 오류가 있었음을 알 수 있었고 실제 미들웨어로써 사용화에도 문제가 없을 것 으를 판단됨.

- 이 외에 Multi Layer Perceptron을 이용한 실험을 진행 했는데 이 실험에선 에러율이 떨어지지 않을 뿐 아니라 인식률 또한 10 Percentage 이하로 전혀 학습이 되지 않았음. 이유는 60,000개의 이미지를 사용된 MNIST 데이터셋 학습은 다양한 인풋 패턴이 존재하며 학습이 진행될 수록 신경망 최적화 문제인 Local Extrema Problem이 발생되는 것을 확인 할 수 있었음. (Error Rate는 떨어지나 가중치의 비중이 한쪽으로 치우쳐지는 현상)

#### o 향후과제

- 범용적인 필기체 인식 미들웨어를 제작하기 위해선 숫자 데이터뿐 아니라 영문자 알파벳과 한글, 한자 등 다양한 문자에 대한 복합적인 패턴을 갖는 데이터에 대한 처리가 요구됨.
- 문제는 숫자 '0' 과 알파벳 'o' , 'O' 를 구분하기 위한 작업이 단일 CNN 모델에서 적용됐을 때 Output Neuron이 많을수록 인식률이 떨어질 뿐 아니라 Output Neuron Delta가 많을수록 최적화된 Weight를 찾기가 힘들어 질 수 있을 것으로 판단됨.
- 이를 위해선 단일 CNN 모델이 아닌 이를 Ensemble 구조를 갖는 CNN 모델을 고안하여 각 문자 카테고리에 대해 CNN 모듈별 개별적 학습과 이 모델을 Ensemble 구조로 구성하여 성능을 높일 수 있는 Ensemble CNN 혹은 Improved CNN을 고안하고자 함.

#### o 참고문헌

- [1] J. Yang, X.Zeng, "Effective Neural Network Ensemble Approach for Improving Generalization Performance" , IEEE Trans on NN, vol.24, No.6 June 2013
- [2] "3D Convolutional Neural Networks for Human Action Recognition" , IEEE Trans on Pattern Recognition, vol.35,no.1, Jan 2013
- [3] A.Calderon, S.Roa, "Handwritten Digit Recognition using Convolutional Neural Networks and Garbor filters" , ICCCI, 2003
- [4] W. Xu, M. Yang, "3D Convolutional Neural Networks for Human Action Recognition" , Machine Inteligence IEEE Trans on Vol.35. Issue 1, Jan 2013