

## CNN 기반 특징맵 사용에 따른 특징점 가시화와 에러율

### Feature Visualization and Error Rate Using Feature Map by Convolutional Neural Networks

진태석<sup>1\*</sup>

Taeseok Jin<sup>1\*</sup>

#### 〈Abstract〉

In this paper, we presented the experimental basis for the theoretical background and robustness of the Convolutional Neural Network for object recognition based on artificial intelligence. An experimental result was performed to visualize the weighting filters and feature maps for each layer to determine what characteristics CNN is automatically generating.

experimental results were presented on the trend of learning error and identification error rate by checking the relevance of the weight filter and feature map for learning error and identification error. The weighting filter and characteristic map are presented as experimental results. The automatically generated characteristic quantities presented the results of error rates for moving and rotating robustness to geometric changes.

*Keywords : CNN, Robustness, Deep learning, Image processing, Sparse Connection*

---

<sup>1\*</sup> 교신저자, 정회원, 동서대학교 메카트로닉스공학과 교수 <sup>1\*</sup> Corresponding Author, Professor Dept. of Mechatronics  
부산대학교 졸업(공학박사) Ph.D Dongseo University  
E-mail: jints@dongseo.ac.kr

## 1. 서론

신경망은 인간 두뇌의 신경 회로의 구조를 모방한 모델로써 문자 인식을 비롯한 물체 검출 방법으로 많이 응용되어왔다. 하지만, 2012년경부터 Deep Learning이라는 새로운 방식이 제안되어 다시 신경망이 주목받게 되었다. Deep Learning은 여러 중간층을 갖는 다층 퍼셉트론이다. 기존의 다층 퍼셉트론에서는 여러 계층으로 구성된 중산층이 충분한 학습은 불가능하였지만 선행학습(pre-training) [1]과 Hinton [3] 등의 제안에 따라 여러 중간층을 가진 퍼셉트론에서도 충분한 학습이 가능하게 되었다. 또한, Deep Learning이 주목된 큰 이유는 학습과정을 통해서 식별에 유효한 특징량을 자동으로 획득할 수 있게 되었다. 기존의 기계학습은 Histogram of Oriented Gradients (HOG) 특징량[4]이나 Scale Invariant Feature Transform (SIFT) 특징량[5] 등의 인식에 의해 유효한 특징량을 시행착오를 통해서 선택해야 했다.

Deep Learning의 선구적인 방법으로는 Convolutional Neural Network (CNN)가 제시되었다[6]. CNN은 각층에서 필터 회선과 풀링을 처리하는 퍼셉트론이다. 각층에서 회선과 풀링하여 각 층간의 결합이 느슨해지고 오차가 확산되는 것을 막아주기 때문에 다층 구성에서도 과학습이 그다지 일어나지 않는다. 신경망 연구는 1940년대부터 연구가 진행되고 있으며 현재까지 여러 차례 연구 패러다임으로 인기를 끌기도 하였다. 1980년에 신경망의 학습 방법중에 오차역전파 방법 [1]이 그 중 하나였으며 퍼셉트론의 출력과 교사 신호로부터 오차를 구하고 파라미터를 갱신하는 학습기법이다. 특히, Deep Learning의 특징량의 학습이 식별기에 어떤 영향을 주고 있는지를 CNN을 이용한 학습결과가 도출되었다.

본 연구에서는 CNN의 회선과 풀링에 의해 자

동 생성된 특징량이 미세한 기하변화에 대하여 어느 정도 불변성을 가지는지에 대한 실험 결과를 제시하였다. 그리고 기존의 신경망의 학습방법으로 사용된 랜덤 학습을 적용하여 CNN이 기하변화에 대한 불변성을 개선할 수 있는지를 테스트하고 랜덤 학습오차와 식별오차를 관찰하여 그에 따른 실험적 결과를 제시하였다[2].

## 2. CNN

### 2.1 시스템 구성

CNN은 필터링 기법을 인공신경망에 적용함으로써 이미지를 더욱 효과적으로 처리하는 기법이다. 다층 퍼셉트론으로 구성되고 중간층에 회선과 풀링 작업으로 필터링 연산작업을 하는 것이 CNN이다. 기존의 필터링 기법은 3행 x 3열 형식의 고정된 필터를 이용하여 이미지를 처리했다. CNN의 기본 개념은 행렬로 표현된 필터의 각 요소가 데이터 처리에 적합하도록 자동으로 학습되게 하는것이다. 이러한 연산방법은 동물의 초기 시각 피질에 특정 기울기를 가진 선에 선택적으로 반응하는 단순 세포와 특정 기울기를 가진 선을 이동시켜도 반응하는 세포의 존재를 기반으로 신경회로 모델로 발전한 것이 그림 1에 표시한 네오코그니트론(Neocognitron)이다[4].

네오코그니트론은 동일한 결합 가중치를 가진 유닛을 병렬로 나열하고, 출력을 상위 계층으로 통합하는 풀링을 수행하는 기하학적 변화에 대한 불변성을 실현하고 있지만, CNN은 네오코그니트론에 오차 역전파법을 이용한 경사하강 최적화 기법을 도입한 방법이다.

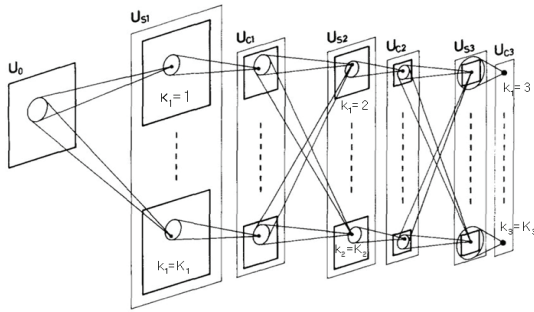


Fig. 1 Neocognitron underlying CNN

## 2.2 특징추출과 식별

특징 추출부에서는 중간 계층의 회선과 풀링 작업을 반복함으로써 특징량을 자동으로 생성한다. 이 절에서는 CNN의 특징 추출부에서 처리되는 회선과 풀링의 처리방법을 제시한다[5][6].

### • 회선 처리

CNN의 회선처리는 가중치 필터와 입력 이미지 또는 가중치 필터와 특징맵에서 내적을 취하고, 라스타스캔(Raster scan)을 통해 반복 회선을 실시하여 아래와 같은 특징 맵의 크기를 계산할 수 있다. 회선 처리에서 이미지와 가중치 필터의 크기를 각각  $n_x \times n_y$ ,  $n_w \times n_w$  할 때 출력되는 특징맵의 크기,  $n'_x, n'_y$ 는 식 (1)과 같다.

$$n'_x = n_x - n_w + 1 \quad (1)$$

$$n'_y = n_y - n_w + 1$$

### • 풀링 처리

풀링은 입력되는 특징맵의 작은 영역에서 값을 출력하고 새로운 특징맵으로 변환하는 작업으로써

여러 개의 값 중에서 대표적인 특징값을 갖는 값을 꺼내서 모아 놓는 것을 뜻한다. CNN에서 자주 사용되는 것은 맥스 풀링이다. 맥스 풀링은 그림 2와 같이 작은 영역 P의 값으로 표현하였고  $h_i$ 에서 최대값을 선택하는 풀링과정을 제시하였다. 맥스 풀링의 출력,  $h'$ 은 식 (2)와 같이 나타낼 수 있다. 여기서  $i$ 는 작은 영역 P의 위치를 나타내고 맥스 풀링 회선 층의 출력과 인접한  $2 \times 2$ 의 유닛의 최대값을 선택하기 위해 출력되는 특징맵의 크기는 식 (3)과 같이 결정한다.

$$h_i = \max_{i \in P} h_i \quad (2)$$

$$n'_x = n_x / 2 \quad (3)$$

$$n'_y = n_y / 2$$

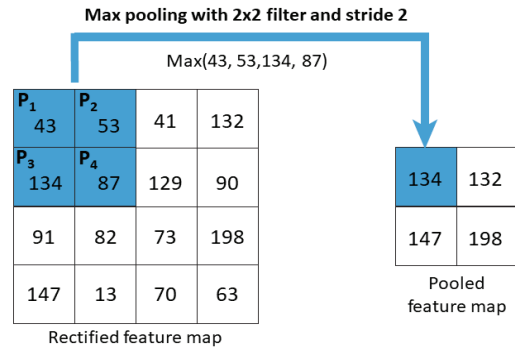


Fig. 2 Pooling process underlying CNN

그림 2와 같이  $4 \times 4$ 로 이루어진 출력 결과에 대해  $2 \times 2$  필터를 사용해서 4개 영역으로 구분한다. 필터가 이동하는 크기인 stride는 필터 크기와 동일하게 2로 설정하게 된다. 이때 출력 맵은  $2 \times 2$ 가 되고 필터가 가리키는 영역에서 가장 큰 값을 선택하게 된다.

### 3. 실험환경

실험을 위하여 MNIST Dataset을 사용하여 CNN의 매개 변수 특징량의 시각화에 따른 학습 오차와 식별오차 값을 각각 제시하였다.

#### 3.1 실험설정

실험을 위한 CNN 유닛은 그림 3과 같이 구성하였다. 입력층과 출력 계층의 유닛수는 다층 퍼셉트론과 같다. 중간 계층은 회선과 풀링에 의해 유닛 수가 각각 식(1)과 식 (2)와 같이 변환하였다.

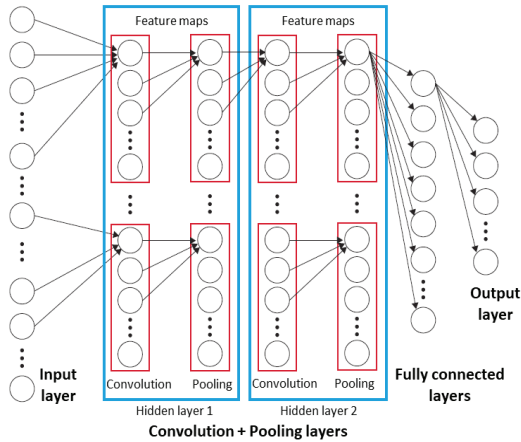


Fig. 3 Structure of CNN unit

회선 계산에서는  $n_w \times n_w$ 의 가중치 필터로 회선처리를 하기 위해  $n_w \times n_w$  범위의 장치에서 출력 응답값을 만들어 냈다. 그 후,  $2 \times 2$ 의 작은 영역에서 풀링을 수행한 후 하나의 유닛에 응답 값을 출력하였다. 식별부는 특징 추출부에서 추출된 특징에서 식별부의 완전연결(Full-Connected)층 단위와 전체 결합하여 확인하였다. 이러한 일련의

처리에서 유닛의 구성은 그림 3과 같은 구조로서 CNN의 층간 결합이 느슨한 (sparse connection) 구조를 갖게 된다.

#### 3.2 학습 매개 변수

실험에서 적용한 CNN 가중치 필터의 크기는  $5 \times 5$ 이고 첫 번째 층은 4(6)장, 두 번째 층은 12 장 이용하였다. 풀링은  $2 \times 2$ 의 작은 영역에서 최대값을 출력하는 맥스 풀링을 사용하였고, 식별부 완전연결층의 단위 수는 500개 사용하였다. 학습 계수는 0.1로 설정하고 미니배치 크기는 10개 또는 20개 미니배치 학습을 실시하였다. 학습오차의 계산은 교차 엔트로피 오차 함수를 이용하고 출력 레이어 단위의 활성화 함수로 소프트 맥스 함수를 사용하였다. 또한 각 실험의 비교실험에서 사용한 다층 퍼셉트론은 중간층이 첫 번째 층에서 유닛수가 1000개의 다층 퍼셉트론을 사용하였다. 학습 계수와 배치 크기는 CNN과 동일하게 0.1과 10로 각각 설정하였고 epoch의 수는 300에서 학습과 오차를 각각 관찰하였다.

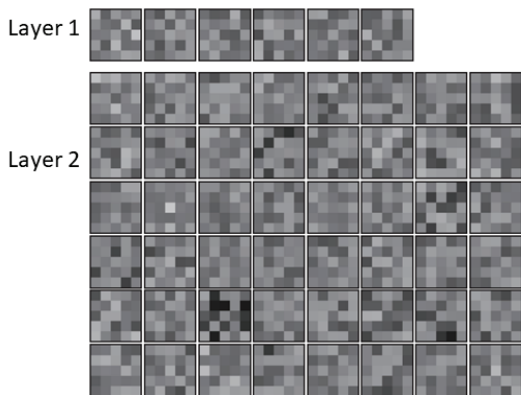
### 4. 실험결과

#### 4.1 실험 개요

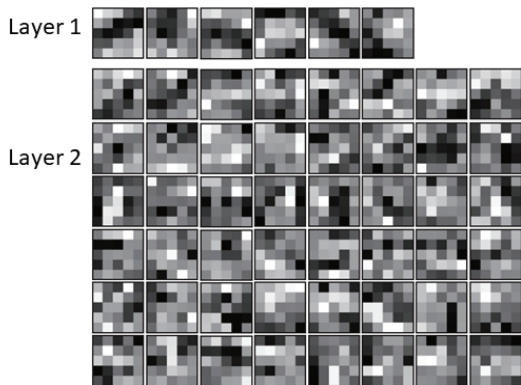
CNN이 어떤 특징량을 자동으로 생성하고 있는 지를 확인하기 위하여 각 층의 가중치 필터와 특징맵을 시각화하였다. 이때 학습오차와 식별오차에 대한 가중치 필터와 특징맵의 관련성을 확인하기 위해 학습오차와 식별율의 추이를 확인하고 가중치 필터와 특징맵이 학습에 의해 변화형태를 실험하였다.

### 4.2 실험 결과

CNN에서 생성된 가중 필터를 가시화 한 결과를 그림 4에 나타내었다.



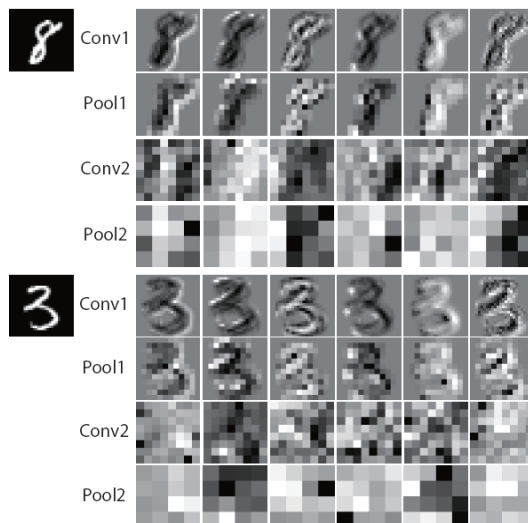
(a) Initial state



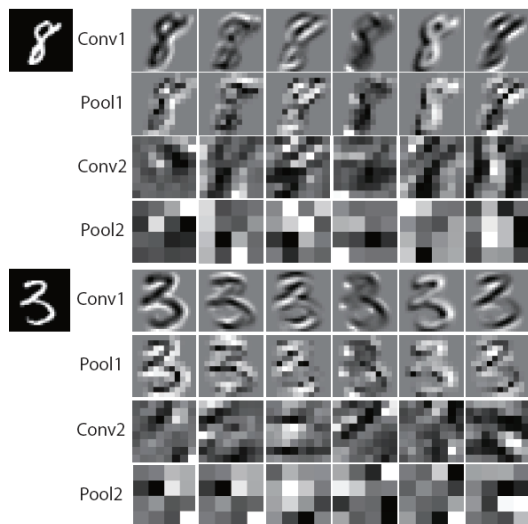
(b) 300 epoch

Fig. 4 Visualized results pattern of weighted filter(kernel)

그림 4(a)는 학습 전에 가중치 필터를 나타낸 것이며, 그림 4(b)는 학습후 가중치 필터를 나타낸 것이다. 실험결과에서와 같은 그림 4는 학습에 의해 가중치 필터의 색조가 변화하고 있는 것을 보여주고 있다.



(a) Initial state for number 8 and 3



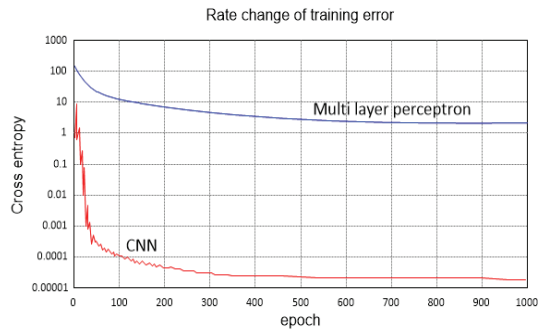
(b) 300 epoch for number 8 and 3

Fig. 5 Visualized results of feature map

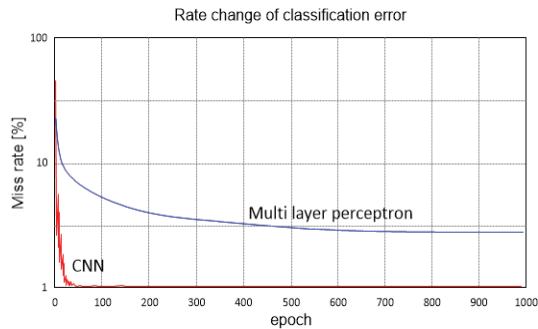
그림 5는 첫 번째 층과 두 번째 층의 회선과 풀링에 의해 도출된 특징 맵을 나타낸다. 그림 5에서 epoch가 증가함에 따라 입력 영상의 에지가 강조되고 있는 것을 알 수 있다. 특히 두 번째 층의 특징 맵은 학습횟수(epoch)가 증가함에 따라

국소적인 가장자리가 크게 강조되고 있는 것을 확인할 수 있다.

그림 6(a)와 그림 6(b)는 학습오차의 추이와 식별오차 추이를 각각 나타낸다. 이 때, 비교 대상으로 다층 퍼셉트론의 학습오차와 식별오차도 제시하였다. 학습오차는 5배정도, 식별오차는 5%정도의 차이로 다층 퍼셉트론보다 오차의 수렴이 빠른 것을 알 수 있다.



(a) Rate change of training error



(b) Rate change of classification error

Fig. 6 Rate change of training error and classification error

## 5. 결론

본 논문에서는 인공지능 기반의 객체 인식을

위한 CNN(Convolutional Neural Network)에 대한 이론적 배경과 강건성에 대한 실험적 근거를 제시하였다. CNN이 어떤 특징량을 자동으로 생성하고 있는지를 확인하기 위한 각 층의 가중치 필터와 특징 맵을 시각화하는 실험적 결과를 수행하였다.

학습오차와 식별오차에 대한 가중치 필터와 특징맵의 관련성 확인을 통한 학습오차와 식별율의 추이에 대한 실험적 결과와 가중치 필터와 특징맵의 학습에 의한 변화형태를 제시하였다. 자동 생성된 특징량은 학습에 의한 식별과 유효 특징량에 변화에 따른 학습오차와 식별오차의 실험적 결과를 제시하였다.

## 사 사

본 논문은 2019년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 임(No. 2016R1D1A1B04932124).

## 참고문헌

- [1] D. Rumelhart, G. Hinton, and R. Williams, "Learning representations by back-propagating errors," *Nature*, pp.533-536, 1986.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems 25*, pp.1106-1114, 2012.
- [3] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *CoRR*, vol. abs/1207.0580, 2012.
- [4] N. Dalal, and B. Triggs, "Histograms of oriented

- gradients for human detection,” International Conference on Computer Vision & Pattern Recognition, vol.2, pp.886-893, 2005.
- [5] K.J Choi, “Optimization of Posture for Humanoid Robot Using Artificial Intelligence,” Journal of the Korean Society of Industry Convergence, vol.22, no.2, pp.87-93, 2019.
- [6] J.P. Lee, J.W. Lee, G.H. Lee, “A Scheme of Security Drone Convergence Service using Cam-Shift Algorithm,” Journal of the Korean Society of Industry Convergence, vol.7, no.5, pp.629-34, 2016.
- 
- (접수: 2020.11.27. 수정: 2021.01.08. 게재확정: 2021.01.11.)