Journal of the Korea Academia-Industrial cooperation Society
Vol. 25, No. 6 pp. 54-62, 2024

https://doi.org/10.5762/KAIS.2024.25.6.54
ISSN 1975-4701 / eISSN 2288-4688

잡음에 강건한 이미지 분류 모델을 위한 다양한 필터 크기를 가진 다중 합성곱 신경망 앙상블 학습 연구

구기범
국방기술품질원

A Research on Multiple CNN Ensemble Learning with Various Filter Sizes for Training a Noise-robust Image Classification Model

Ki-Beom Ku
Defense Agency for Technology and Quality

요 약 본 논문에서는 노이즈 이미지에 대한 강건한 분류 모델 작성을 위하여 다양한 필터크기를 가진 다중 합성곱 신경망 앙상블 모델을 연구하였다. 다중 합성곱 신경망 앙상블 모델은 2개의 다중필터 합성곱 신경망 앙상블 모델의 소프트 보팅(soft voting) 조합으로 구성되는데, 각각의 다중필터 합성곱 신경망 앙상블 모델은 다시 5개의 단일 합성곱 신경망의 소프트 보팅 조합으로 구성된다. 단일 합성곱 신경망들은 서로 다른 필터 크기가 적용되었으며, 그 필터 크기는 임의의 크기인 3, 5, 7, 11, 13이다. 단일 합성곱 신경망들은 학습 과정에서 특별한 연산이 적용되었는데, 각 단일 합성곱 신경망에서 마지막 완전 연결 계층의 출력이 서로 적절히 조합되어, 다시 각각의 단일 합성곱 신경망 모델의 소프트맥스 및 교차 엔트로피 오차(softmax-with-Loss) 계층으로 보내져 순전파 연산을 수행한다. 그 후, 오차역전파 연산을 수행하며 가중치를 갱신하도록 함으로써 단일 합성곱 신경망 모델들이 학습 과정에서 상호보완 될 수 있도록 하였다. 작성된 다중 합성곱 신경망 앙상블 학습모델의 강건성 평가는 MNIST(Modified National Institute of Standards and Technology) 손글씨 숫자 이미지 데이터로 수행되었다. 숫자 이미지 데이터는 1개의 정상 데이터 셋과, 3개의 비정상 잡음 데이터 셋으로 구성되어 있으며, 정상/비정상 데이터 상관없이 학습 및 테스트하여 모델의 성능을 평가하였다. 그 결과, 다중 합성곱 신경망 앙상블 모델의 분류 성능이 단일 합성곱 신경망 모델보다 전반적으로 향상된 것을 확인하였다.

Abstract A multiple CNN ensemble model was developed using various filter sizes to train a noise-robust image classification model. The multiple CNN ensemble model consisted of a soft voting combination of two multiple-sized filter CNN ensemble models. The multiscale filter CNN ensemble model consisted of a soft voting combination of five single CNN models. Single CNN models were made using 3, 5, 7, 11, and 13-sized filters. During the modeling, a special calculation process was applied to CNNs, outputs of last affine layer in each CNN model are combine properly. The combined outputs were sent to softmax-with-loss layer to conduct rest of forward propagation. Lastly, by progressing progress back propagation calculation, CNN models made it possible to complement each other's weight variables. Evaluation of the robustness of the multiple CNN ensemble model was performed using an MNIST handwritten number image data set, 1 normal data set, and 3 noise added data set. These image data sets were used to model and test the performance of the classification model regardless of normal or noise. The resulting multiple CNN ensemble model had better classification performance than the single CNN model.

Keywords : Convolution Neural Network, Ensemble Learning, Voting, MNIST, Robust model

*Corresponding Author : Ki-Beom Ku(Defense Agency for Technology and Quality)
email: slal123@dtaq.re.kr
Received April 5, 2024 Revised May 14, 2024
Accepted June 7, 2024 Published June 30, 2024

1. 서론

신경망 학습 기법은 인식 분야에서 사용되는 대표적인 통계 모델 중 하나로, 공학, 군사, 의료[1-4] 등등 다양한 연구 분야에서 흔히 볼 수 있을 정도로 널리 사용되고 있다. 또한, 사용하는 데이터에 따라 인공 신경망(Artificial Neural Network, ANN), 합성곱 신경망(Convolutional Neural Network, CNN), 재귀 신경망(Recursive Neural Network, RNN)과 그 신경망들을 응용한 다양한 모델들이 존재한다.

통계 모델의 주요한 성능 중 하나인 강건성(robustness)은 작성한 모델 자체가 높은 정확도를 가져야 함과 동시에 다양한 데이터에 대응하여 일반화된 성능을 발할 수 있어야하는 성질이다. 최근에는 민간 또는 군사 분야에서 딥러닝 모델을 이용한 영상이나 이미지 등의 데이터를 기반으로 객체나 대상을 식별하는 분류/평가 모델이 많이 쓰이고 있다. 이러한 모델들은 다양한 환경 조건에서 영상이나 이미지 데이터에 대한 일반화된 성능을 발휘할 수 있어야 한다. 예컨대, 군사 분야에서 활용되는 피아식별 분류 모델은 다양한 전장 환경에서 잡음(noise)이나 적대적 공격(adversarial attack)[5-7] 등에 노출될 수 있다. 만일 분류 모델이 강건한 성능을 발휘하지 못할 경우, 아군의 피해로 직결되는 매우 중대한 문제를 맞이할 수 있다.

이러한 의도치 않은 적대적 공격에 대한 중요성이 강조됨에 따라, 최근에는 적대적 공격에 대응하기 위한 딥러닝 모델 작성, 적대적 공격패턴 분석 등에 관한 연구가 진행되고 있다[8,9].

또한, 수집된 데이터로 모델링을 갱신하는 기능이 있는 AI 모델이 탑재된 무기체계일 경우, 수집된 데이터에 잡음이나 적대적 공격으로 인한 왜곡이 포함됐을 때 모델의 성능이 영향을 받을 가능성이 크다. 따라서 무기체계에 적용되는 AI 모델들은 다양한 형태의 데이터로 학습 되도 높은 수준의 성능을 발휘할 수 있어야 한다.

본 연구에서는 기존 단일 합성곱 신경망 모델보다 분류 성능이 높으면서, 잡음이 섞인 데이터를 이용한 학습에도 강건한 신경망 학습모델의 확보를 위한 연구를 수행하였다. 합성곱 신경망은 같은 이미지를 분류하더라도 필터 크기에 따라 학습된 모델의 분류 성능이 다르다. 이 점을 이용하여 서로 다른 필터 크기를 가진 합성곱 신경망들이 상호 보완하며 앙상블 학습을 수행하면 강건한 분류 모델을 얻을 수 있을 것으로 판단하여 다양한 필터를 가진 다중 합성곱 신경망 앙상블 모델을 작성하였다. 다중 합성곱 신경망 앙상블 모델은 주어진 이미지 데이터에 대하여, 5개의 단일 합성곱 신경망의 출력을 소프트 보팅으로 조합한 앙상블 모델인 다중필터 합성곱 신경망 앙상블 모델 2개의 출력을 다시 소프트 보팅으로 조합하여 이미지를 최종 분류하는 방식의 앙상블 모델이다. 단, 5개의 단일 합성곱 신경망 모델들은 모델링 과정에서는 마지막 완전 연결 계층의 출력을 특정 조건으로 조합한 후, 다음 순전파 연산 및 오차역전파를 수행함으로써 모델링 과정에서 상호보완 할 수 있도록 설계하였다.

이렇게 작성된 다중 합성곱 신경망 앙상블 모델의 성능은 정상 및 잡음이 섞인 MNIST 숫자 손글씨 이미지 데이터 셋을 활용하여 확인하였다. 같은 이미지 데이터 셋을 적용한 단일 합성곱 신경망 모델과의 분류 정확도를 비교함으로써 다중 합성곱 신경망 앙상블 모델이 모델 및 테스트 데이터 종류에 상관없이 강건한 모델인지 검증하였다.

2. 본론

2.1 MNIST 숫자 손글씨 이미지 데이터

본 연구에서 사용된 이미지 데이터는 MNIST 손글씨 숫자 이미지 데이터[10]로, 합성곱 신경망의 학습 데이터로 널리 쓰이는 이미지 데이터이다. 본 연구에서 사용하는 숫자 이미지 데이터 셋은 일반적인 정상 MNIST 숫자 이미지 데이터(이하 Normal) 1개 셋(Fig. 1)과, AWGN (Additive White Gaussian Noise, 이하 Noise1), Motion Blur(이하 Noise2), Reduced Contrast and AWGN(이하 Noise3)이 각각 적용된 3가지의 노이즈 데이터 셋[11](Fig. 2)의 총 4가지 데이터 셋으로 모델링 및 모델의 분류 성능 시험에 활용하였다.

모델링 데이터의 수는 정상 및 잡음 데이터의 4종류 모두 각각 60,000개이며, 테스트 데이터는 각각 10,000개씩 활용하였다. 이미지는 28×28 크기의 도트 이미지 데이터이며 흑백으로 구성되어 있다. 정답 레이블은 숫자 0~9에 해당하는 출력 값을 갖는다.

Fig. 1. MNIST handwritten digit data(normal)

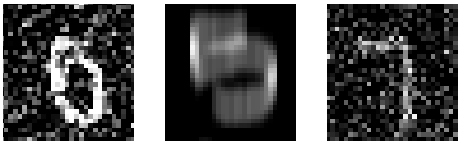

Fig. 2. MNIST with Noise
(a) AWGN(left), (b) Motion Blur(center),
(c) Reduced Contrast and AWGN(right)

2.2 학습모델 기법

2.2.1 합성곱 신경망

합성곱 신경망(Convolution Neural Netwrok, CNN)은 주로 이미지 인식과 음성 인식 등을 위한 딥러닝에 적용되는 모델링 기법이다. 인공 신경망(Artificial Neural Network, ANN)과의 차이점은 완전연결(affine) 계층에 더하여 입력 데이터 행렬에 필터를 합성곱 연산하는 합성곱(convolution) 계층과, 합성곱 연산의 출력 데이터인 출력 특징 맵(output feature map)의 가로세로 공간을 줄이는 풀링(pooling)계층이 추가되어 신경망을 구성한다는 점이다[12](Fig. 3).

또한, 합성곱 신경망 특유의 변수들도 등장하는데, 입력 데이터와 함께 합성곱 연산에 적용되는 필터, 합성곱 연산을 여러 번 수행하는 딥러닝에서 출력 특징 맵의 크기를 조정하는 목적으로 사용되는 패딩(padding), 그리고 합성곱 연산에 필터를 적용하는 위치의 간격을 조정하는 스트라이드(stride) 등이 있다.

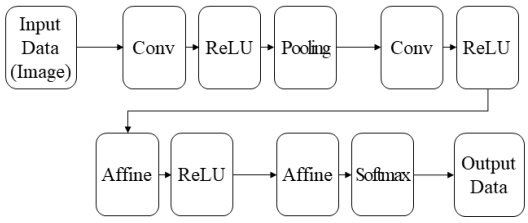

Fig. 3. An example of CNN structure

2.2.2 앙상블 학습

앙상블 학습(ensemble learning)은 개별적으로 학습시킨 여러 모델의 출력에 가중치를 부여하여 조합하는 학습형태로, Eq. (1)과 같이 표현된다[13].

$$f(y|x,\pi)=\sum_{m\in M} w_m f_m(y|x) \tag{1}$$

여기서 w_m은 조율이 가능한 모수로, m번째 모델의 출력에 대한 가중치 값이며 f_m은 m번째 모델의 출력이다. 여기서 가중치는 아래와 같은 Eq. (2)로 추정이 가능하다.

$$\hat{w}=\arg max \sum_{i=1}^{N} L\left(y_i, \sum_{m=1}^{M} w_m f_m(x)\right) \tag{2}$$

각 가중치로 조합된 모델의 결과와 정답 간의 손실 함수(loss function)가 최소화되는 가중치 변수 w의 모수 추정을 통해 앙상블 학습이 수행된다.

한편, 개별적으로 학습시킨 여러 모델의 출력을 보팅(voting), 배깅(bagging) 하는 등의 단순한 조합방식으로 앙상블 학습을 수행하는 방법도 존재한다[14,15]. 보팅은 같은 데이터를 여러 개의 분류모델에 독립적으로 학습시킨 후 각 모델의 출력을 조합하는 학습 방식으로, 출력계층의 분류 값을 평균하여 가장 큰 값을 지닌 레이블을 결과로 출력하는 소프트 보팅(soft voting)이 대표적이다. 배깅은 데이터를 부분 집합으로 나누어 여러 개의 모델에 독립적으로 학습시킨 후 그 결과를 평균화하여 최종 결과를 추론하는 학습방식이다.

2.3 다중 합성곱 신경망 앙상블 모델

2.3.1 단일 합성곱 신경망 구성 및 학습 조건

다중 합성곱 신경망 앙상블 모델을 구성하는 앙상블 모델인 다중필터 합성곱 신경망 앙상블 모델은 5개의 단일 합성곱 신경망 모델로 구성되어 있다. 단일 합성곱 신경망 모델은 Fig. 4의 구조로 되어있다. 각 단일 합성곱 신경망 모델은 다른 필터 크기를 가지는데, 필터 크기는 각각 임의의 크기인 3X3, 5X5, 7X7, 11X11, 13X13로 설정하였다.

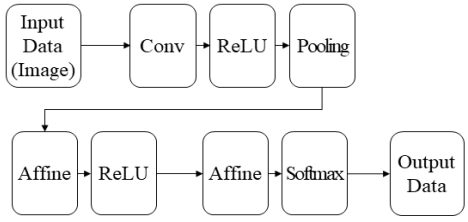

Fig. 4. CNN layer structure

모델링 및 테스트용으로 사용되는 MNIST 숫자 이미지 데이터 셋은 28X28의 픽셀 크기를 갖는 흑백 데이터를 입력변수로 가지며, 0~9를 나타내는 레이블을 출력변수로 가지고 있다. 본 연구에서 작성하는 모델은 MNIST 숫자 데이터 셋에 맞춘 입출력 뉴런 수를 가져야 하므로,

신경망 모델은 728개의 입력 뉴런을 갖는다. 한편, 모델의 출력은 입력 데이터에 대하여 0~9의 각 레이블로 분류되는 확률을 나타내야 하므로 출력계층은 10개의 뉴런을 갖도록 구성하였다.

신경망의 학습 조건은 에포크(epoch) 수 20, 배치 크기는 100으로 설정하였다. 아울러 필터의 수는 30개, 패딩은 0, 스트라이드는 1로 입력하였다. 신경망의 매개변수는 아담(Adaptive Moment Estimation, Adam)을 이용하여 최적화하였다[16]. 아담은 모멘텀(momentum)과 AdaGrad(Adaptive Gradient)를 조합한 최적화 기법으로, 하이퍼파라미터(학습률)의 편향 보정과 최적화 수렴 속도 조절이 가능한 것이 특징이다. 본 연구에서는 아담 알고리즘의 입력 변수에 학습률(learning rate)은 0.001을, beta1 및 beta2는 각각 0.9, 0.999를 적용하여 신경망 매개변수의 최적 값 탐색을 수행하였다.

2.3.2 다중필터 합성곱 신경망 앙상블 모델 구성

다중필터 합성곱 신경망 앙상블 모델은 5개의 합성곱 신경망 모델을 Fig. 5와 같이 구성한 형태인데, 각 합성곱 신경망 마지막의 완전 연결 계층의 출력을 Eq. (3)에 따라 평균 조합하여 소프트맥스 계층으로 보내 최종 분류하도록 하였다. 여기서 x_{ji}는 i번째 합성곱 신경망에서 해당 입력 이미지 데이터가 j로 분류될 확률을 의미한다. 다중 필터 합성곱 신경망 앙상블을 구성하는 합성곱 신경망의 학습 조건은 단일 합성곱 신경망의 학습 조건과 동일하게 설정하였다.

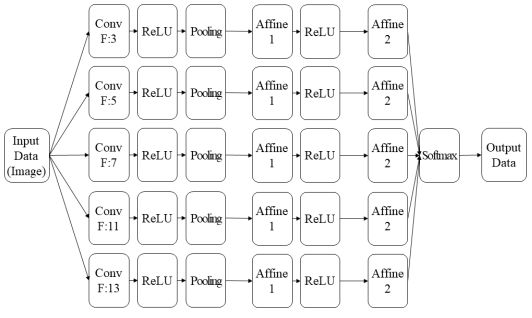

Fig. 5. Multiple-size filter CNN ensemble model

$$y_j = \frac{1}{5}\sum_{i=1}^{5}(x_{ji}) \quad (\text{단}, j = 0....9) \tag{3}$$

2.3.3 다중필터 합성곱 신경망 앙상블 모델 작성

다중필터 합성곱 신경망 앙상블 모델을 만들기 위한 각각의 단일 합성곱 신경망은 개별 모델이 독립적으로 학습되는 대신, 학습 과정에서 마지막 완전 연결 계층(Affine 2)의 출력이 소프트맥스 및 교차 엔트로피 오차(cross entropy error) 계층으로 넘어갈 때 특별한 연산을 수행함으로써 각각의 합성곱 신경망의 모델이 학습과정에서 서로의 가중치를 상호 보완하는 앙상블 모델을 작성하였다. 특별한 연산은 2가지로, 각 특별한 연산 방법은 각각의 다중 필터 합성곱 신경망 앙상블 모델에 적용되었다. 이러한 특별한 연산을 수행하는 이유는 같은 입력 데이터 및 계층 구성이라도 필터 크기에 따라 정답률이 영향을 받기 때문이다.

첫 번째 특별한 연산은 각 신경망의 출력에 대한 표준편차를 계산하여, 표준편차가 가장 큰 신경망의 출력은 그대로 두고 나머지 4개의 신경망에 대한 출력은 보강하는 연산이다. 연산순서는 다음과 같다. 1) 각각의 단일 합성곱 신경망의 완전 연결 계층 2에서 출력된 데이터에 대한 표준편차를 계산한다. 2) 5개의 합성곱 신경망에서 계산된 표준편차를 비교하여 가장 높은 합성곱 신경망은 자신의 완전 연결 계층 2의 출력을 그대로 가져가게 된다. 나머지 4개의 합성곱 신경망은 기존의 계산된 완전 연결 계층 2의 출력 값 대신, 5개의 각 완전 연결 계층 2의 출력 값을 Eq. (3)에 따라 조합된 결과로 대체한다. 3) 2)에 따라 수정된 완전 연결 계층 2의 출력 값은 표준편차가 가장 높은 합성곱 신경망을 제외한 4개의 각 합성곱 신경망의 교차 엔트로피 오차 계층으로 보내고, 오차역전파를 통해 가중치 변수들을 갱신한다.

두 번째 특별한 연산은 첫 번째의 특별한 연산과 마찬가지로 각 신경망의 출력에 대한 표준편차를 이용하는 연산이나, 표준편차가 가장 작은 합성곱 신경망의 출력을 보강하고 나머지 4개의 신경망 출력은 그대로 두는 연산이다. 연산순서는 다음과 같다. 1) 각각의 합성곱 신경망의 완전 연결 계층 2에서 출력된 값의 표준편차를 계산한다. 2) 5개의 합성곱 신경망 중에서 계산된 표준편차 중 가장 낮은 값을 가진 합성곱 신경망은 자신의 완전 연결 계층 2의 출력 값 대신, 5개의 완전 연결 계층 2의 출력 값을 Eq. (3)에 따라 조합한 결과로 대체한다. 나머지 4개의 합성곱 신경망은 각각 자신의 완전 연결 계층 2의 기존 출력 값을 가져간다. 3) 2)에 따라 수정된 각 합성곱 신경망의 완전 연결 계층 2의 출력 값은 각 합성곱 신경망의 교차 엔트로피 오차 계층으로 보내지고 오차역전파를 통해 가중치 변수들을 갱신한다. Fig. 6 및 Fig. 7은 이러한 합성곱 신경망 조합 모델의 완전 연결 계층 2의 출력을 조합하는 과정을 도식한 것이다.

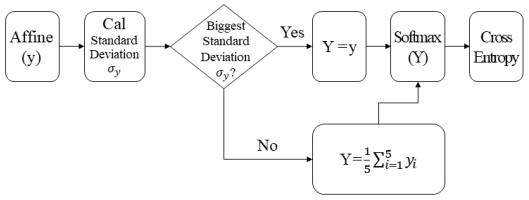

Fig. 6. Output combination process of each affine layer 2 (maximum standard deviation method)

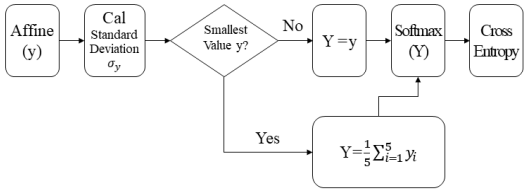

Fig. 7. Output combination process of each affine layer 2 (minimum standard deviation method)

위와 같이, 합성곱 신경망 앙상블 학습에 가중치가 아닌 각 단일 합성곱 신경망 모델의 완전 연결 계층 2의 출력에 대한 표준편차를 고려한 이유는 MNIST 숫자 이미지 분류를 위한 합성곱 신경망의 분류 결과와 정답률 간의 상관관계 때문이다. Fig. 8은 학습된 단일 합성곱 신경망에 테스트 데이터로 분류한 결과를 바탕으로 레이블 간 분류 확률의 표준편차와 정확도를 데이터 1,000개 단위 평균으로 도식한 것이다. 그리고 Fig. 9는 동 모델 분류 결과에서, 1순위로 분류된 레이블의 분류 확률과 2순위로 분류된 레이블의 분류 확률, 그리고 레이블 간 분류 확률 표준편차 사이의 관계를 데이터 1,000개 단위 평균으로 도식한 그래프이다. 학습데이터는 정상 MNIST 숫자 이미지 데이터를 활용하였으며, 테스트는 Noise3 데이터로 수행하였다.

Fig. 8 및 Fig. 9의 그래프에 따르면, 출력계층에서 분류 확률 간 표준편차가 클수록 정확도가 높은 경향을 보이며 표준편차가 높을수록 첫 번째로 분류된 확률과 두 번째로 분류된 확률의 차이가 크다. 이는 모델이 출력계층에서 유력한 정답을 더 편중적으로 선택하게 된다면, 그것이 출력계층의 표준편차 증가로 이어지며 곧 정확도가 높아질 수 있다고 판단할 수 있는 근거가 된다. 따라서 이러한 데이터 예측분류의 정확도를 개선하기 위해서는 1) 출력계층의 노드 간 표준편차가 가장 작은 신경망을 개선하거나 2) 출력계층의 노드 표준편차가 높은 신경망을 제외한 나머지 4개의 신경망을 개선하는 방향으로 모델의 개선이 필요하다고 판단하여, 본 연구에서는 Eq. (1)과 같은 가중치 변수를 이용한 앙상블 학습 대신 표준편차를 고려하여 앙상블 학습모델을 작성하였다.

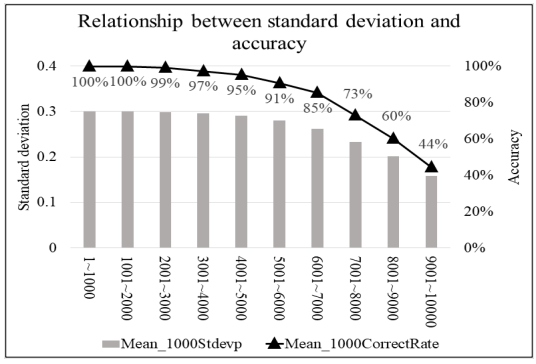

Fig. 8. Relationship between standard deviation and accuracy

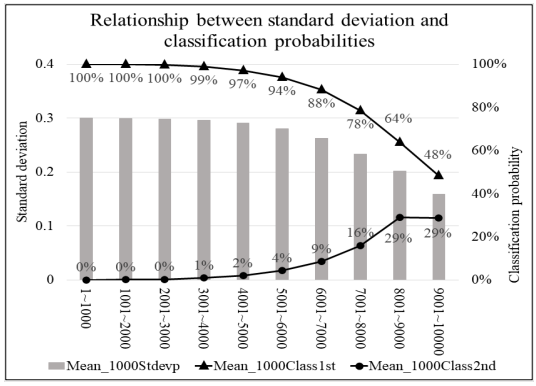

Fig. 9. Relationship between standard deviation and classification probabilities

이러한 과정으로 학습된 5개의 합성곱 신경망 모델을 Eq. (3)에 따라 앙상블 조합함으로써 다중 필터 합성곱 신경망 앙상블 모델을 작성하였다.

2.3.4 다양한 크기의 필터가 적용된 다중 합성곱 신경망 앙상블 모델 작성

2가지의 특별한 연산을 각각 적용하여 학습된 다중필터 합성곱 신경망 앙상블 모델은 다시 소프트 보팅 방식으로 조합되어 다중 합성곱 신경망 앙상블 모델이 작성되었다. 이로써, 다중 합성곱 신경망 앙상블 모델은 입력되는 숫자 이미지 데이터는 단일 합성곱 신경망의 소프트 보팅 조합과 다중 필터 합성곱 신경망 앙상블 모델의 소프트 보팅 조합 과정을 거쳐 최종적으로 이미지를 분류하게 된다.

3. 모델링 결과

본 연구에서 강건한 분류 모델 확보를 위해 작성된 다중 합성곱 신경망 앙상블 모델의 분류 성능은 단일 합성곱 신경망 모델보다 잡음이 없는 정상 학습데이터 뿐만 아니라, 잡음이 섞인 비정상 학습데이터로 모델링 하더라도 기본적인 분류 성능이 뛰어나야 한다. 또한, 그 모델은 분류 대상 데이터에 대하여 잡음 여부에 상관없이 기존의 단일 신경망보다 높은 분류 성능을 나타내야 한다. 이러한 분류 성능을 평가하기 위해 다중 합성곱 신경망 앙상블 모델과 단일 합성곱 신경망 모델에 정상 및 비정상 학습 데이터 셋 4종류 중 1종으로 모델링하고, 정상 및 비정상 4종류의 테스트 데이터 분류한 정확도를 확인함으로써 모델의 강건성을 확인하였다. Table 1은 다중 합성곱 신경망 앙상블 모델의 분류 정확도를, Table 2~6은 필터 크기별 단일 합성곱 신경망 모델의 정확도를 나타낸 것이다(단위: %).

Table 1. accuracy of multi CNN Ensemble model

Ensemble CNN		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	99.30	97.25	96.32	90.03
	Noise1	99.23	98.66	96.85	96.43
	Noise2	98.51	93.01	99.05	74.56
	Noise3	98.98	98.41	97.18	97.46

Table 2. accuracy of single CNN model(filter size: 3X3)

Single CNN (Filter size 3x3)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	98.42	73.32	89.66	52.70
	Noise1	98.26	97.01	91.63	84.54
	Noise2	95.05	64.04	98.60	41.24
	Noise3	97.76	95.76	93.45	94.77

Table 3. accuracy of single CNN model(filter size: 5X5)

Single CNN (Filter size 5x5)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	98.58	90.42	91.19	71.86
	Noise1	98.80	97.82	95.69	93.79
	Noise2	97.06	73.27	98.73	47.06
	Noise3	98.64	97.89	96.00	95.79

Table 4. accuracy of single CNN model(filter size: 7X7)

Single CNN (Filter size 7x7)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	98.89	95.55	94.07	85.79
	Noise1	98.98	98.16	94.32	95.52
	Noise2	98.00	85.45	98.73	63.06
	Noise3	98.78	97.85	96.46	96.47

Table 5. accuracy of single CNN model(filter size: 11X11)

Single CNN (Filter size 11x11)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	98.70	94.26	95.43	82.60
	Noise1	98.96	98.18	95.95	95.84
	Noise2	96.23	88.06	98.41	70.29
	Noise3	98.87	98.29	96.16	96.72

Table 6. accuracy of single CNN model(filter size: 13X13)

Single CNN (Filter size 13x13)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	99.07	94.33	94.95	82.87
	Noise1	98.91	98.04	96.35	95.07
	Noise2	97.88	90.35	98.75	73.61
	Noise3	98.64	97.99	96.95	96.69

전체 16만 건의 테스트 케이스 중 다중 합성곱 신경망 앙상블 모델은 153,123 건의 정답을 맞혀 전체 97.7%의 분류 정확도를 보여주었다. 단일 합성곱 신경망 중에서는 필터 크기가 13인 모델이 151,045 건을 맞혀 94.4%로 가장 높은 분류 정확도를 보였으며, 개별 분류 케이스에서는 다중 합성곱 신경망 앙상블 모델이 필터 크기 13의 단일 합성곱 신경망과 비교하여 0.23~7.13%의 더 높은 정확도를 보여주었다. 5개의 단일 합성곱 모델에서 가장 정확도가 높은 케이스만 모아서 합산하더라도 151,673으로 다중 합성곱 신경망 앙상블 모델에 비해 낮은 94.8%의 분류 정확도를 보여주어 개별 분류 케이스에 대해서도 다중 합성곱 신경망 앙상블 모델이 0.11~4.24% 더 높은 분류 정확도를 보여주었다.

그러나 정확도는 모델의 클래스별 분류 성능을 고려하지 못해 분류 모델의 성능 평가지표로서 부족한 점이 있다. 이러한 정확도의 단점을 보완하기 위해 도입된 것이 F-score이다. F-score는 정밀도(precision)와 재현율(recall)을 통해 계산되는 분류 모델의 성능 평가지표로, 다중 클래스 분류 모델의 클래스별 분류 성능을 고려하

여 분류 모델의 성능을 평가할 수 있다. F-score는 0~1 값으로 표현되는데, 1에 가까울수록 모델의 분류 성능이 좋은 것으로 평가된다. 본 연구에서는 클래스별 정밀도와 재현율의 조화평균으로 계산되는 macro-averaged F1-score를 이용하여 모델의 성능을 평가하였다. Table 7은 다중 합성곱 신경망 앙상블 모델의 F1-score를 도식하였고, Table 8~12는 필터 크기별 단일 합성곱 신경망 모델의 F1-score를 나타낸 것이다.

Table 7. F1-score of multi CNN Ensemble model

Ensemble CNN		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	0.9929	0.9724	0.9629	0.8992
	Noise1	0.9922	0.9866	0.9683	0.9639
	Noise2	0.9850	0.9289	0.9904	0.7362
	Noise3	0.9898	0.9841	0.9718	0.9745

Table 8. F1-score of single CNN model(filter size: 3X3)

Single CNN (Filter size 3x3)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	0.9842	0.7083	0.8879	0.4729
	Noise1	0.9825	0.9699	0.9164	0.8437
	Noise2	0.9499	0.6177	0.9859	0.3771
	Noise3	0.9775	0.9575	0.9339	0.9474

Table 9. F1-score of single CNN model(filter size: 5X5)

Single CNN (Filter size 5x5)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	0.9857	0.9000	0.9081	0.7068
	Noise1	0.9879	0.9780	0.9568	0.9369
	Noise2	0.9703	0.7302	0.9872	0.4545
	Noise3	0.9864	0.9788	0.9600	0.9576

Table 10. F1-score of single CNN model(filter size: 7X7)

Single CNN (Filter size 7x7)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	0.9888	0.9552	0.9395	0.8566
	Noise1	0.9898	0.9815	0.9423	0.9549
	Noise2	0.9799	0.8534	0.9872	0.6087
	Noise3	0.9877	0.9783	0.9638	0.9644

Table 11. F1-score of single CNN model (filter size: 11X11)

Single CNN (Filter size 11x11)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	0.9869	0.9436	0.9542	0.8325
	Noise1	0.9895	0.9816	0.9588	0.9581
	Noise2	0.9620	0.8798	0.9840	0.6893
	Noise3	0.9887	0.9828	0.9611	0.9670

Table 12. F1-score of single CNN model (filter size: 13X13)

Single CNN (Filter size 13x13)		Test Data Set			
		Normal	Noise1	Noise2	Noise3
ModelData Set	Normal	0.9906	0.9437	0.9492	0.8335
	Noise1	0.9890	0.9803	0.9635	0.9505
	Noise2	0.9786	0.9039	0.9874	0.7378
	Noise3	0.9863	0.9798	0.9692	0.9667

F1-score를 비교한 결과, 다중 합성곱 신경망 앙상블 모델은 하나의 케이스를 제외한 모든 경우에 대해 전반적으로 단일 합성곱 신경망보다 높은 분류 성능을 나타냄을 확인할 수 있다. 하나의 케이스는 Noise2 학습 데이터-Noise3 테스트 데이터 조합 셋에서 다중 합성곱 신경망 앙상블 모델의 F1-score가 필터 크기 13의 단일 합성곱 신경망보다 미흡한 분류 성능을 보여주었다. 이는 일부 분류 클래스에서 다중 합성곱 신경망 앙상블 모델의 정밀도가 낮은 결과를 보여 분류 클래스의 F1-score 및 해당 케이스의 F1-score가 낮아진 것으로 확인되었다. 다른 필터 크기에서의 Noise2 학습 데이터-Noise3 테스트 데이터 조합 셋에 대한 F1-score는 다중 합성곱 신경망 앙상블 모델이 앞서며, 다중 합성곱 신경망 앙상블 모델의 16개 조합 셋의 평균 F1-score는 0.9562로 필터 크기 13의 단일 합성곱 신경망의 조합 셋 평균 F1-score인 0.9444 보다 높아 전체 분류 성능은 더 좋은 것으로 평가된다.

따라서, 다중 합성곱 신경망 앙상블 모델이 단일 합성곱 신경망 모델과 비교하여 더 강건한 모델임을 확인할 수 있다.

4. 결론

본 연구에서는 이미지를 분류하는 강건한 신경망 모델

작성을 목표로 다양한 필터 크기를 적용한 다중 합성곱 신경망 앙상블 모델을 작성하였다. 다중 합성곱 신경망 앙상블은 5개의 단일 합성곱 신경망을 학습한 후 소프트 보팅 앙상블 조합하여 다중 필터 합성곱 신경망 앙상블 모델을 작성하고, 다시 2개의 다중필터 합성곱 신경망 앙상블 모델을 소프트 보팅 앙상블 조합하여 작성되었다. 이때, 5개의 단일 합성곱 신경망 학습과정에는 특별한 조합 연산이 포함되었다. 단일 합성곱 신경망의 학습 과정은 일반적인 합성곱 신경망과 대부분 다르지 않으나, 학습 중간에 마지막 완전 연결 계층에서 출력된 값을 서로 적절히 조합하는 연산이 수행됨으로써 학습 과정에서 각각의 단일 합성곱 신경망의 가중치 변수가 상호 보완되도록 하였다.

작성된 다중 합성곱 신경망 앙상블 모델의 분류 성능은 MNIST 손글씨 숫자 이미지 데이터를 이용하여 확인하였다. 잡음이 없는 정상 학습 데이터 1개 셋과 잡음이 섞인 비정상 학습 데이터 3개 셋을 이용하여 모델 학습을 진행하고, 정상 테스트 데이터 1개 셋과 잡음이 섞인 비정상 테스트 데이터 3개 셋으로 이미지 분류에 활용하였다. 이러한 학습 및 테스트 데이터 셋을 이용하여 총 16개 경우의 학습 및 분류 케이스에 대한 모델 분류 성능을 정확도 및 F1-score를 통해 확인하였다. 그 후, 같은 시험 케이스를 단일 합성곱 신경망에도 적용하여 분류 성능을 비교하였다. 그 결과, 다중 합성곱 신경망 앙상블 모델의 전반적인 분류 성능이 단일 합성곱 신경망보다 더 높은 것을 확인하였다.

같은 이미지를 대상으로 다양한 필터 크기를 가진 단일 합성곱 신경망의 분류 성능 확인을 통해, 필터 크기는 모델의 분류 성능에 영향을 미치는 주요 인자임을 확인하였다. 이러한 관점에서 다중 합성곱 신경망 앙상블 모델은 다양한 필터 크기를 가진 여러 개의 단일 합성곱 신경망이 상호 보완을 거쳐 학습이 진행되므로, 다양한 이미지 크기의 객체에 대한 분류를 요구하는 객체 인식 분야에 활용이 가능할 것으로 기대된다. 또한, 하이퍼파라미터 최적화를 통해 활용 목적에 맞게 신경망 모델을 조정할 수 있다면 모델의 분류 성능은 더 높아질 것으로 판단된다.

그러나, 다중 합성곱 신경망 앙상블 모델은 이를 구성하는 단일 합성곱 신경망의 특별한 조합 연산이 마지막 완전 연결 계층의 순전파 이후에 수행되므로 단일 합성곱 신경망들의 가중치 변수 간 직접적인 영향이 제한된다. 또한, 다중 합성곱 신경망 앙상블 모델의 분류 성능이 필터 크기 단일 합성곱 신경망보다 낮은 모델-테스트 데이터 조합도 있어 모델 성능의 개선이 필요하다. 이를 위해, 하나의 합성곱 신경망에 각각 다른 크기의 필터 크기를 가진 합성곱 계층을 적용한다면 가중치 변수 간의 상호 보완성이 강화되어 분류 성능이 개선될 것이라 기대된다.

References

[1] S. J, Kwon, M. S. Kim, "Flaw Evaluation of Bogie connected Part for Railway Vehicle Based On Convolutional Neural Network", *Journal of the Korea Academia-Industrial cooperation Society*, Vol.21, No.11, pp.53-60, Nov. 2020.
DOI: http://doi.org/10.5762/KAIS.2020.21.11.53

[2] S. Zhong, S. Fu, L. Lin, "A novel gas turbine fault diagnosis method based on transfer learning with CNN", *Measurement*, Vol.137, pp.435-453, Apr. 2019.
DOI: https://doi.org/10.1016/j.measurement.2019.01.022

[3] Y. G. Lee, J. M. Ma, "A Comparative Study on the Performance of CNN models for Classification of Artillery Weapon Systems", *Journal of the Korea Academia-Industrial cooperation Society*, Vol.24, No.1, pp.344-350, Jan. 2023.
DOI: http://doi.org/10.5762/KAIS.2023.24.1.344

[4] H. S. Shon, Y. G. Yi, K. O. Kim, E. J. Cha, K. A. Kim, "Classification of stomach cancer gene expression data using CNN algorithm of deep learning", *Journal of Biomedical and Translational Research*, Vol.20, No.4, pp.15-20, Mar. 2019.
DOI: https://doi.org/10.12729/jbtr.2019.20.1.015

[5] H. S. Kim, M. J. Kwon, J. Y. Byun, C. I. Kim, "Transfer-based Adversarial Attack using Camera Viewpoint Transformation", *Journal of the Institute of Electronics and Information Engineers*, Vol.59, No.7, pp.53-59, Jul. 2022.
DOI: https://doi.org/10.5573/ieie.2022.59.7.53

[6] M. J. Son, M. J. Kwon, S. J. Cho, C. I. Kim, "An Image Warping Method for Improving the Transferability of Adversarial Attacks", *Journal of the Institute of Electronics and Information Engineers*, Vol.59, No.10, pp.152-159, Oct. 2022.
DOI: http://dx.doi.org/10.5573/ieie.2022.59.10.152

[7] Y. Wang, K. Wang, Z. Zhu, F. Y. Wang, "Adversarial attacks on Faster R-CNN object detector", *Neurocomputing*, Vol.382, pp.87-95, Mar. 2020.
DOI: https://doi.org/10.1016/j.neucom.2019.11.051

[8] H. Y. Kim, M. Y. Chung, "Singular Value-Based Randomized Image Reconstruction Method for Adversarial Attack Defense", *The Journal of Korean Institute of Next Generation Computing*, Vol.19, No.4, pp.68-75, Aug. 2023.
DOI: http://doi.org/10.23019/kingpc.19.4.202308.007

[9] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, D. Mukhopadhyay, "A survey on adversarial attacks and defences", *CAAI Transactions on Intelligence Technology*, Vol.6, No.1, pp.25-45, Mar. 2021.
DOI: https://doi.org/10.1049/cit2.12028

[10] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, et el., "ImageNet: A large-scale hierarchical image database", *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp.248-255, Aug. 2009.
DOI: https://doi.org/10.1109/CVPR.2009.5206848

[11] S. Basu, M. Karki, S. Ganguly, R. DiBiano, S. Mukhopadhyay, et el., "Learning sparse feature representations using probabilistic quadtrees and deep belief nets", *Neural Processing Letters*, Vol.45, pp.855-867, Sep. 2016.
DOI: https://doi.org/10.1007/s11063-016-9556-4

[12] S. Goki, "Deep Learning from Scratch", p.312, Hanbit Media, Inc., 2017, pp.227-254

[13] K. P. Murphy, "Machine Learning: A Probabilistic Perspective", p.1286, acorn publishing Co., 2017, pp.698-699

[14] J. H. Hwang, H. S. Park, J. D. Kim, "A Study on Forecasting initial Provisioning Based on Ensemble Learning", *Journal of the Korea Academia-Industrial cooperation Society*, Vol.28, pp.431-438, Aug. 2022.
DOI: https://doi.org/10.5762/KAIS.2022.23.8.431

[15] S. Y. Park, "Malicious Insider Detection Using Boosting Ensemble Methods", *Journal of the Korea Institute of Information Security & Cryptology*, Vol.32, pp.267-277, Apr. 2022.
DOI: https://doi.org/10.13089/JKIISC.2022.32.2.267

[16] D. P. Kingma, J. Ba, "Adam:A Method for Stochastic Optimization", arXiv preprint arXiv:1412.6980, Dec. 2014.
DOI: https://doi.org/10.48550/arXiv.1412.6980

구 기 범(Ki-Beom Ku) [정회원]

- 2016년 8월 : 경희대학교 일반대학원 기계공학과 (공학석사)
- 2017년 9월 ~ 현재 : 국방기술품질원 선임연구원

〈관심분야〉
통계학, 빅 데이터, 최적화 설계