



Edge Enhanced Neural Network For High Accuracy Image Classification

Sangmin Suh*

Department of Information & Telecommunication Engineering, Gangneung-Wonju National University

A B S T R A C T

Among the several deep learning research areas, image classification is a fundamental area and has been widely applied to many practical applications. Image classification is to determine which category the given image belongs to. Since the image classification is a typical supervised learning, test images and the corresponding answers, i.e., labels are also given. And, with the given test images and labels, a neural network is trained to minimize a loss function defined by error between the label and the inference result. Therefore, as the loss decreases, the inference accuracy increases. As a result, the accuracy is a criteria of the performance of the neural network in image classification. In this paper, a new method for high accuracy image classification is suggested. Authors think that recognizing things mean seeing the shape of the things. With the intuition, in the proposed method, additional edge information is applied to the image and around 3% accuracy improvement is achieved in the experiment. In order to clarify the improvement, compared feature maps in the hidden layers are visualized and analyzed. And, it is also confirmed that the feature map of the proposed method is more clear and sharp than that of the conventional one. Another merit of the proposed method is that this method can easily improve the accuracy of the conventionally existing neural networks through the transfer learning because the proposed method just modifies the first layer of the conventional neural networks.

© 2020 KKITS All rights reserved

KEYWORDS : Artificial intelligence, Deep learning, Edge enhancement, Feature map, Image classification, Neural network, Supervised learning

ARTICLE INFO: Received 17 April 2020, Revised 8 June 2020, Accepted 11 June 2020.

*Corresponding author is with the Department of Information & Telecommunication Engineering, Gangneung-Wonju National University, 150, Namwon-ro,

Wonju-si, 26403, KOREA.

E-mail address: sangminsuh@gwnu.ac.kr

1. 서론

인공 지능 분야는 최근에 가장 활발히 연구되고 있는 분야 중에 하나이며, 이 인공 지능은 인공 신경망을 근본으로 한다. 최초의 인공 신경은 인간의 신경인 뉴런(neuron)의 동작 원리를 모방하여 설계 제작되었다[1]. 그 후 다중 레이어 신경망 구조(MLP: multi-layer perceptron)[2]와 역전파 방법(back propagation) [3]이 개발되어 현실적으로 사용이 가능한 인공 신경망을 설계 할 수 있게 되었다. 그러나 이러한 인공 신경망을 훈련시키고 실제 사용하기 위해서는 무엇보다도 많은 연산을 빠르게 수행 할 수 있는 계산기 (computer)가 필요한데, 2000년대 초반까지만 하여도 이러한 계산기가 없어 실생활에서 적용을 하지 못하고 있었다. ILSVRC (ImageNet Large-Scale Visual Recognition)은 인공 지능 분야 에서도 가장 기본이 되는 이미지 분류 (image classification) 대회이다. 그런데 2012년 이 대회에서, Alex Krizhevsky가 기존에 비하여 거의 10%에 가까운 성능 향상을 얻은 논문을 발표하였다[4]. 성능 향상의 요인으로 활성화 함수 개선과 과적합 (overfitting) 문제를 해결하기 위한 방법론도 제시하였지만, 가장 큰 요인은 병렬 처리가 가능한 GPU (graphics processing unit)를 사용한다는 것이었다. 기존과 다르게, GPU를 사용함으로써 많은 연산을 빠르게 처리하는 것이 가능해졌고, 이를 통하여 많은 훈련이 가능하여 궁극적으로 커다란 성능 향상을 얻게 되었다.

이후 인공지능 분야는 많은 분야에 적용이 되었는데, 영상 의학 데이터를 분류하는 데 사용되었고 [5], 농작물의 상태를 분석하는 분야에도 사용되었다 [6]. 또한, 추론 정밀도 향상 외에도 고속 인공지능 구현을 위한 방법도 제안되었다[7].

이 논문에서는, 인공 지능 이미지 분류기의 성능 향상을 위한 새로운 방법을 제시한다. 이 논문의

아이디어로서 저자는 우리가 사물을 바라보고 그것이 무엇인지를 파악할 때 사물의 외곽을 보고 판단한다는 점에 착안하여 그 정보를 신경망에 적용하였으며, 신경망 은닉 레이어 (hidden layer)의 특징 맵 (feature map)을 상대 비교함으로써 이 방법이 실제 유효함을 보인다. 성능 비교 평가를 위하여 텐서플로우 (tensorflow)를 이용하여 인공신경망을 구현하였으며 테스트 이미지 세트로는 10 종류의 이미지를 분류하는 CIFAR-10 이미지를 사용하였다[8]. 동일한 조건의 실험에서 제안하는 방법을 사용하여 약 3%의 성능 개선을 이룰 수 있음을 보인다. 제안하는 구조의 또 하나의 장점으로서는 기존 신경망을 바꾸지 않고도 단지 입력 레이어 만을 변경하여 성능을 올릴 수 있다는 것이다. 이는 거의 동일한 가중치 (weight)나 바이어스 (bias) 수를 유지할 수 있으며, 이것은 추가적인 메모리나 단일 곱셈누산기 (MAC: multiplier-accumulator)의 증가가 거의 없음을 의미한다. 다시 말해서, 단 소수점 이하 몇 퍼센트의 성능을 높이기 위하여 백층이 넘게 레이어 쌓는 것에 비하면 [9], 제안하는 방법은 대단히 단순하면서도, 작은 크기의 효과적인 인공 신경망을 만들 수 있는 방법이다.

이 논문의 구성은 다음과 같다. 제 2장에서는 신경망 설계와 성능 향상을 위한 아이디어와 그 구현을 제시하며, 제 3장에서는 실험 환경과 특징 맵들을 시각화한 결과를 비교하여 성능 개선이 명확히 됨을 보이고, 마지막으로 결론을 기술한다.

2. 고성능 이미지 분류기 설계

2.1 제안하는 이미지 분류기의 개념

우리가 사물을 보고 그것이 무엇인지 판단할 때 우리의 의식의 흐름은, 우선 사물의 외곽을 보고, 두 번째는 색상과 재질을 보게 된다. 예를 들어,

우리가 고양이를 인식할 때 더욱 명확해 진다. 처음 고양이의 외곽을 보고 이것이 고양이임을 알 수 있고, 둘째 색상과 재질을 보고난 후 만약 색깔이 빨강거나 반짝이는 재질. 즉 우리가 일반적으로 알고 있는 기존 정보와 다른 색상이나 재질일 경우, 우리는 이것을 고양이이긴 하나, 가짜 고양이 혹은 장난감 고양이로 인식하게 된다.

이 논문에서는 이러한 사물의 외곽 정보를 신경망에 적용하여 이미지 분류기의 성능을 높이고자 한다.

2.2 인공 신경망 설계

이미지 분류기에서 인공 신경망의 입력은 영상이며 일반적으로 빨강 (Red), 녹색 (Green), 파랑 (Blue) 3가지 색상의 정보를 가지고 있으며 아래의 <그림 1>과 같은 구조이다.

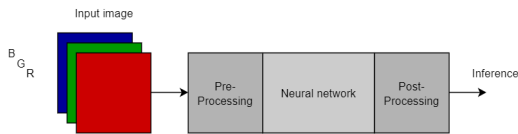


그림 1. 기존의 이미지 분류기
Figure 1. Conventional image classification

입력으로는 3 채널 이미지를 받고, 그 이미지를 정제하는 전처리 (pre-processing) 과정을 거친 후, 인공 신경망 (neural network) 과 결과를 정리하는 후처리 (post-processing) 과정을 거쳐, 주어진 이미지에 대한 최종 추론을 하게 된다.

이 논문에서는 이미지 분류기의 정밀도를 향상하고자, 기존 이미지 정보에 사물의 외곽 (Edge) 정보를 추가한다.

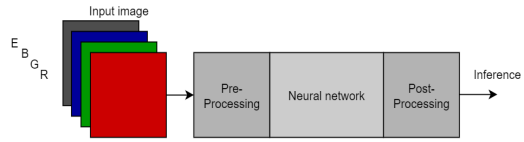


그림 2. 제안하는 이미지 분류기
Figure 2. Proposed image classification

<그림 2>는 제안하는 이미지 분류기 구조이며 기존의 구조에서 이미지의 외곽 정보를 추가하였다. 이것은, 우리가 사물을 인식할 때 외곽 부분을 눈여겨본다는 데에서 착안하였고, 그 개념을 이미지 분류기의 입력으로 추가한 것이다.

이제 제안하는 방법의 유용성을 검증 및 확인하기 위하여 인공 신경망을 설계한다. 신경망 설계를 위한 프레임워크는 텐서플로우 2.1[10]을 사용하며, 이미지 테스트 세트로는 10가지 이미지를 분류하는 CIFAR-10 세트를 사용한다.



그림 3. 채널별 이미지 (왼쪽 위에서 시계 방향으로, RGEGB)
Figure 3. Channel images (from left-top clockwise, RGEGB)

3개의 채널 RGB를 갖는 기존의 이미지에 외곽 정보를 포함하는 새로운 채널을 추가적으로 만들어야 한다. 외곽 정보를 포함하는 채널을 만드는 방법은 여러 가지가 있을 수 있지만, 이 논문에서는 영상 샤프닝 (image sharpening) 방법 중 하나인 라플라시안 (laplacian) 방법을 사용한다[11]. <그림 3>은 표준 테스트 이미지[12,13]를 이용하여 RGB를 포함한 외곽영상 채널 (E)까지 추출한 그림의 예시이다.

E채널에서는 이미지의 외곽만이 표현되고 있다.

E채널이 만들어 졌으니, 이제 전체 인공 신경망을 설계한다. 제안하는 방법의 유용성을 빠르게 확인하기 위하여 전체 신경망은 아래와 같이 단순한 구조로 설계하였다.

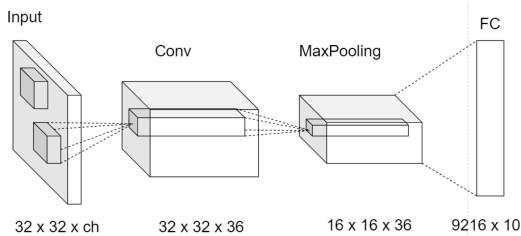


그림 4. 신경망 구조 설계
Figure 4. Design if neural network

CIFAR-10의 테스트 이미지의 크기는 32 (height) × 32 (width) × 3 (channel)이며, 이 논문이 제안하는 방법을 사용하면 이미지의 크기는 32×32×4가 된다. 그림에서 ‘Conv’는 컨볼루션 레이어, ‘MaxPooling’은 최대 풀링 레이어 (max pooling layer)를, 그리고 ‘FC’는 완전 연결 층 (fully connected layer)를 나타낸다. 그리고 이 논문에서 사용된 파라미터 값들을 <표 1>에 정리하였다.

표 1. 설계된 신경망의 파라미터

Table 1. parameters of the designed neural network

Layer name	Parameters
Conv	kernel size: 3×3, stride:1
MaxPooling	kernel size: 2×2, stride:2
optimizer	stochastic gradient descent method
epoch	64
weight/bias initializer	Xavier normal initializer

3. 실험

설정된 파라미터 값을 이용하여 훈련을 진행하였고 텐서보드 (tensor board)[10]를 이용하여 성능을 비교하였다. 훈련에서 에포크 (epoch)가 진행됨에 따라서 변화하는 손실 값은 아래 <그림 5>와 같이 얻어졌다.



그림 5. 손실 값 비교
Figure 5. Compared loss values

여기서 ‘logs_wo’은 에지 채널 E를 인가하지 않은 상태의 손실 값의 변화를 나타내며, ‘log_w’은 에지 채널 E를 인가하였을 때의 손실 값의 변화를 나타낸다. E 채널을 인가하였을 때 손실 값이 더 빠르게 떨어짐을 분명히 알 수 있다.

손실 값이 이렇게 빠르게 줄어든다는 것은 이미 지 분류기의 정밀도가 더 빠르게 향상된다는 것을 의미하며, 그 비교된 정밀도를 <그림 6>에 나타내었다.

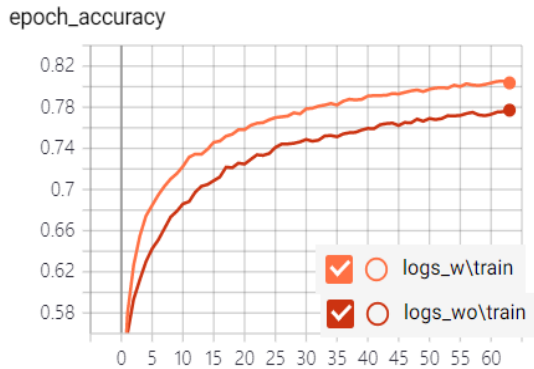


그림 5. 정밀도 비교
Figure 5. Compared accuracies

훈련이 진행 될수록, 기존 방식에 비하여 제안한 방법에서의 정밀도가 더욱 증가함을 알 수 있다. 이 실험에선 기존 방식은 77%의 정밀도를 보였으며, 제안하는 방식에서는 80%의 정밀도를 보이며, 제안하는 방법으로 3%의 성능 향상을 얻을 수 있었다.

이제 왜 이러한 결과가 나온 것인지 좀 더 구체적으로 신경망 내부를 알아보려고 한다. 이 논문에서는 신경망 내부를 보기 위하여 은닉층 (hidden layer)의 특징 맵 (feature map)을 비교한다. 그 비교를 위하여, 이미 훈련이 끝난 두 신경망에 <그림 6>과 같은 하나의 실험 이미지를 인가한다.



그림 6. 은닉층 비교를 위한 실험 이미지
Figure 6. Test image to compare hidden layers

이 실험 이미지의 해상도는 640×426이므로 32×32로 다운 샘플링 (down sampling)하여, 이미 구현된 신경망의 입력 크기에 맞도록 이미지의 크기를 조절하였다. 이 이미지를 <그림 4>의 입력으로 인가하고 ‘Conv (3×3×36)’레이어의 출력 특징 맵 (output feature amp)을 시각화하여 비교하였다. 출력 채널이 총 36개 층이 있고 모든 층을 시각화 하였다.

<그림 7>은 기존 방식으로 얻어진 특징 맵이고 <그림 8>은 새롭게 제안한 방식으로 얻어진 특징 맵이다. 그리고 각 이미지의 위치를 행렬의 원소 위치와 동일하다고 생각한다. 두 특징 맵을 비교해 볼 때, 일부는 다른 부분도 있지만, ‘전체적’으로 보면 <그림 8>이 <그림 7>보다 더욱 선명하다. 특히 이미지(1,8)과 이미지(2,8)에선 더욱 차이가 많이 나며, 다른 특징 맵에서도 거의 동일한 현상이 나타난다. 그리고 이러한 선명한 특징 맵은 최대 풀링과 완전 연결 층들을 거쳐 소프트 맥스 층 (softmax layer)으로 진행되면서 더 많은 정보를 주게 되고, 최종적으로 더 높은 성능을 얻을 수 있게 된다.

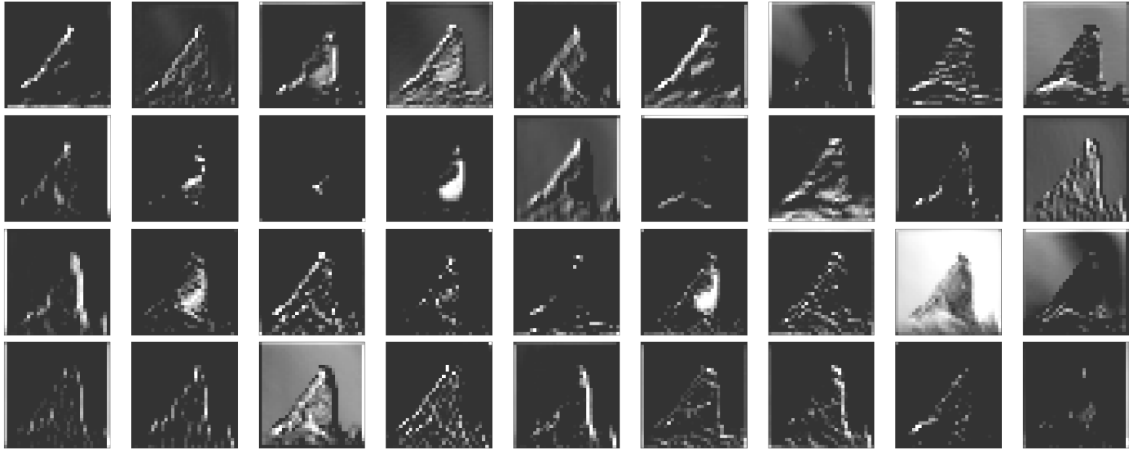


그림 7. 기존 방식에서 컨볼루션 레이어의 특징 맵
Figure 7. Feature maps of the convolution layer in the conventional method

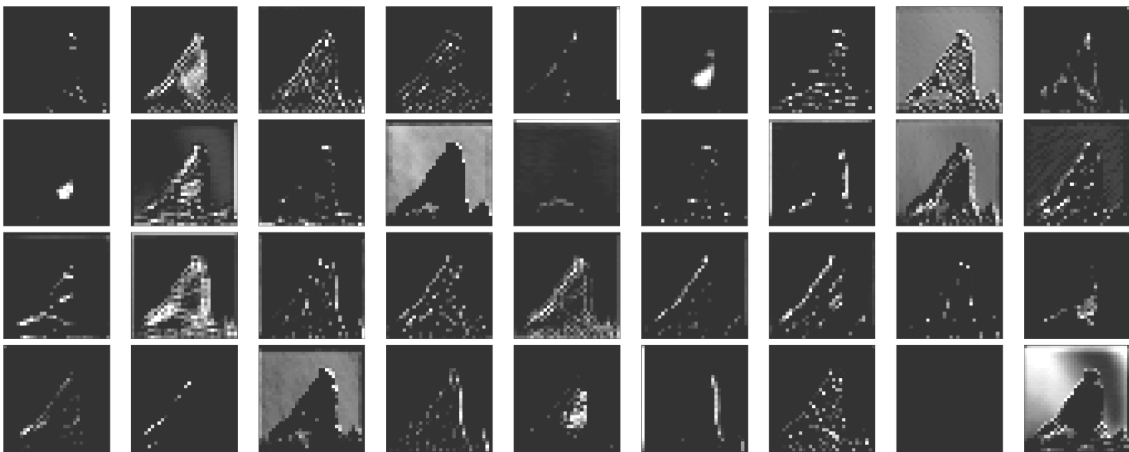


그림 8. 제안한 방식에서 컨볼루션 레이어의 특징 맵
Figure 8. Feature maps of the convolution layer in the proposed method

3. 결론

이 논문에서 이미지 분류기의 성능 향상을 위한 새로운 방법을 제시하였다. 사물을 인식한다는 것은 사물의 형상을 인식하다는 직관으로부터, 이를 구체화하기 위하여 신경망의 입력에 사물의 외곽 채널을 추가하였고, 내부 특징 맵을 상대 비교하여 왜 이런 결과가 얻어졌는지를 분석하였다.

실험을 통하여, 동일한 조건에서 이미지 분류기의 성능이 약 3% 향상됨을 보였다.

제안하는 방법의 또 다른 장점으로는, 기존에 이미 구현된 신경망에서 단지 입력층만을 수정한 후 전이 학습 (transfer learning)을 통하여 성능을 더욱 올릴 수 있다는 것이다. 이것은 처음부터 새롭게 훈련해야 하는 방법보다는 분명히 시간과 노력을 줄여준다.

References

- [1] F. Rosenblatt, *The perceptron: a probabilistic model for information storage and organization in the brain*, Cornell Aeronautical Laboratory, Psychological Review, Vol. 65, No. 6, pp. 386-408, 1958.
- [2] M. Minsky, and S. A. Papert, *Perceptrons: an introduction to computational geometry*, The MIT Press, 1969.
- [3] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, *A learning algorithm for Boltzmann machines*, Cognitive science, Vol. 9, No. 1, pp. 147-169, 1985.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Imagenet classification with deep convolutional neural networks*, <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-network-s.pdf>, Sep. 2019.
- [5] S. Shin, H. Kim, and S-H. Park, *Recent trend of biomedical image classification methods*, Journal of Knowledge Information Technology and Systems(JKITS), Vol. 12, No. 5, pp. 765-776, 2017.
- [6] S-H. Jeong, M-H. Lee, and H. Yoe, *Fruit classification system using deep learning*, Journal of Knowledge Information Technology and Systems(JKITS), Vol. 13, No. 5, pp. 589-595, 2018.
- [7] S. Suh, *Effective implementation for fast deep learning algorithm*, Journal of Knowledge Information Technology and Systems(JKITS), Vol. 14, No. 5, pp. 553-561, 2019.
- [8] <https://www.cs.toronto.edu/~kriz/cifar.html>, Apr. 2020.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, <https://arxiv.org/abs/1512.03385>, Apr. 2020.
- [10] <https://www.tensorflow.org/>, Apr. 2020.
- [11] R. C. Gonzalez, R. E. Woods, *Digital image processing*, Pearson, 2018.
- [12] <https://www.ece.rice.edu/~wakin/images/>, Apr. 2020.
- [13] <https://homepages.cae.wisc.edu/~ece533/images/>, Apr. 2020.

고정밀 이미지 분류기를 위한 엣지 강화 신경망

서상민

강릉원주대학교 정보통신공학과 조교수

요약

여러 인공지능 분야에서도 이미지 분류기는 가장 근본이 되는 분야이며 또한 가장 널리 사용되는 분야이다. 이 논문은 이미지 분류기의 정밀도 성능을 올리기 위한 새로운 방법을 제안한다. 사물을 인식할 때, 사물의 형상을 보고 판단한다는 직관으로부터 새로운 신경망 구조를 구현하였고 은닉층의 특징 맵을 상대 비교함으로써, 필연적으로 성능향상이 이루어질 수 있음을 보였다. CIFAR-10 실험 세트를 이용하여 약 3% 정도의 성능향상을 보여, 제안하는 방법이 실제 유효함을 보였다.



Sangmin Suh received B.S., M.S., and Ph.D. degree in the electronics engineering in the Hanyang university of Seoul Korea in 1991, 1994, and 2003 respectively. From

1994 to 1999, he has been with Daewoo telecommunication. From 2003 to 2019, he was in Samsung electronics. He currently holds assistant professor in the department of information and telecommunication engineering at Gangneung-Wonju national university. His current research interests include artificial intelligence, robotics, and control.

E-mail address: sangminsuh@gwnu.ac.kr