

변형된 잔차블록을 적용한 CNN

곽내정[†], 신현준^{**}, 양종섭^{***}, 송특섭^{****}

CNN Applied Modified Residual Block Structure

Nae-Joung Kwak[†], Hyeon-Jun Shin^{**}, Jong-Seop Yang^{***}, Teuk-Seob Song^{****}

ABSTRACT

This paper proposes an image classification algorithm that transforms the number of convolution layers in the residual block of ResNet, CNN's representative method. The proposed method modified the structure of 34/50 layer of ResNet structure. First, we analyzed the performance of small and many convolution layers for the structure consisting of only shortcut and 3×3 convolution layers for 34 and 50 layers. And then the performance was analyzed in the case of small and many cases of convolutional layers for the bottleneck structure of 50 layers. By applying the results, the best classification method in the residual block was applied to construct a 34-layer simple structure and a 50-layer bottleneck image classification model. To evaluate the performance of the proposed image classification model, the results were analyzed by applying to the cifar10 dataset. The proposed 34-layer simple structure and 50-layer bottleneck showed improved performance over the ResNet-110 and Densnet-40 models.

Key words: CNN, Residual Learning, Bottleneck, ResNet

1. 서 론

ICT 기술의 급속한 발전과 제 4차 산업혁명 시대의 도래는 사회 전반적인 변화를 일으키고 환경에 영향을 미치고 있다. 이와 더불어 고속 프로세서들의 개발과 복잡한 연산을 빠르게 처리하는 주변 장치들의 발전은 이전에는 불가능했던 복잡한 연산과 알고리즘들을 처리할 수 있도록 도와주고 있다. 인공지능망은 이러한 기술의 진보를 힘입어 딥러닝이라는 진보된 알고리즘으로 성장했고 다양한 분야에 적용되고 있다. 딥러닝 기법 중 CNN(Convolutional Neural

Network, 합성곱 신경망)은 이미지의 특징을 추출하고 분류하는데 사용된다[1]. CNN은 이미지 특징을 추출하는 컨볼루션 레이어로 구성된 특징 추출 부분과 추출된 특징을 이용하여 이미지를 분류하는 부분으로 구성된다. 특징 추출 부분은 컨볼루션 레이어를 깊게 쌓아 학습할수록 학습 데이터에 존재하는 대표적인 특징을 잘 추출할 수 있어 학습 결과가 좋아진다[2]. 하지만 네트워크의 깊이가 너무 깊어지면 기울기(Gradient)가 너무 작아지거나 커지는 상황으로 인해 학습이 제대로 이뤄지지 않는 기울기 소실 및 증폭(Vanishing/Exploding) 현상과 파라미터 수의

※ Corresponding Author: Teuk Seob Song, Address: (353-49) 88, Doanbuk-ro, Seo-gu, Daejeon, Korea, TEL: +82-42-829-7630, FAX: +82-42-822-8431, E-mail: teukseob@mokwon.ac.kr

Receipt date: Dec. 16, 2019, Revision date: May. 19, 2020
Approval date: Jun. 10, 2020

[†] Dept. of Cyber security, Paejae University
(E-mail: knj0125@pcu.ac.kr)

^{****} Div. of Convergence Computer and Media, Mokwon University (E-mail: y2k4388@gmail.com)

^{****} Div. of Convergence Computer and Media, Mokwon University (E-mail: sheepserverdown@gmail.com)

^{****} Div. of Convergence Computer and Media, Mokwon University

※ This research was supported by Basic Science Research Program through the Korea National Research Foundation of Korea (NRF) funded by the Ministry of Education in 2018(NRF-2018R1D1A1B07048675)

증가에 의한 에러의 증가 문제로 성능이 저하된다[3, 4]. 문제점들을 해결하기 위해 초기화 값의 적절한 선택[5] 더 나은 최적화기법[6], 스킵 연결[7], 지식 전달[8] 및 레이어 별 훈련[9]과 같은 더 깊은 신경망 교육을 가능하게 하는 다양한 기술이 제안되었다. ResNet(Residual Network)[10]은 잔차학습(Residual Learning)을 도입하여 깊은 신경망에서도 기울기의 영향을 줄여주어 학습을 가능하게 할 수 있도록 하였다. 또한 네트워크의 깊이가 깊어지면 파라미터의 증가로 속도가 저하되는 문제를 해결하고자 1×1 컨블루션 구조[11]를 도입한 병목구조(Bottleneck) 구조를 적용하였다. ResNet의 잔차블록은 2개의 컨블루션 레이어를 쌓아 구성되며 입력과 출력의 차이가 최적이 되도록 학습하는 잔차학습의 개념과 입력을 그대로 출력에 더하는 숏컷(Shortcut) 연결을 추가하여 구성된다. 숏컷 연결은 오차 역전파시에 기울기가 잘 흘러갈 수 있게 해주며 덧셈 연산의 추가만으로도 구현이 가능하여 추가적인 파라미터가 많이 필요하지 않다. 그러나 숏컷 연결을 사용하는 경우 기울기가 네트워크를 통해 흐르기 때문에 잔차블록 가중치를 통과하지 않아도 되며 훈련 중에는 아무것도 학습하지 않아도 되므로 유용한 특징을 추출하는 블록이 거의 없거나 많은 블록이 최종 특징추출에 작은 영향을 주는 정보를 거의 공유하지 않을 수 있다[12]. 따라서 잔차블록에서 더 많은 특징을 추출할 수 있는 방법이 연구되어야 한다. 또한 네트워크의 레이어가 너무 얇으면 이미지 분류 성능이 좋지 않고 레이어가 깊어지면 이미지 분류 성능이 개선되기도 하지만 과적합이나 기울기 소실문제 등이 발생한다. 따라서 적절한 레이어에서 이미지 분류 성능 또한 만족할 만한 네트워크가 필요하다.

본 논문에서는 숏컷 연결의 단점을 개선하고자 이미지의 특징을 더 잘 추출할 수 있도록 잔차블록 내 기존 ResNet의 2개의 3×3 컨블루션 층의 수를 늘려 이미지 분류 성능을 개선한 잔차블록과 이를 적용한 병목구조 블록을 제안한다. 또한 변형된 잔차블록을 적용하여 이미지 분류 성능을 개선한 34/50레이어의 이미지 분류 모델을 제안하고 성능을 평가한다. 본 논문은 2장에서는 기존방법인 ResNet에 대해 설명하고 3장에서는 잔차블록의 변형 구조에 대해 기술한다. 4장은 변형된 잔차블록을 이용하여 본 논문에서 제안하는 이미지 분류모델을, 5장에서는 제안모델

의 성능 평가를 기술하며 6장에서는 결론을 맺는다.

2. ResNet

ResNet은 2015년 ILSVRC에서 1위를 차지한 네트워크로 CNN의 기울기 소실 및 증폭(Gradient vanishing/exploding) 문제와 층이 깊어지면 파라미터의 수가 비례적으로 증가하여 과적합(Overfitting)의 문제가 아니라도 에러가 커지는 현상이 발생하는데 이를 해결하는 방법을 제안하였다. 이때 도입된 것이 잔차학습(Residual learning)으로 기존 CNN의 컨블루션 레이어의 파라미터는 출력 $H(x)$ 가 최적의 출력이 되도록 학습되나 잔차학습은 입력과 출력 $(H(x))$ 의 차이 $(F(x))$ 가 최적이 되도록 학습하는 구조로서 CNN의 출력 $H(x)$ 를 다음과 같이 재정의한다.

$$F(x) = H(x) - x \tag{1}$$

$$H(x) = F(x) + x \tag{2}$$

잔차 $F(x)$ 가 최적이 된다는 것은 $F(x)$ 이 0이 되는 것이며 잔차를 학습한다는 의미에서 잔차학습이라 하며 이와 같이 잔차를 학습하므로 잔차블록이라 한다. Fig. 1은 기존의 CNN 네트워크와 잔차학습(Residual learning)의 블록도이다.

잔차블록은 잔차블록의 입력 x 가 컨블루션 레이어 2개를 건너 출력 $H(x)$ 에 더해지는 숏컷 연결이 추가된다. 숏컷 연결은 잔차블록의 입력 x 가 출력에 연결되므로 파라미터의 수에 영향이 없으며 덧셈을 제외하면 연산량의 영향도 없다. 또한 SGD나 오차 역전파에 의한 학습도 가능하다. 그러나 숏컷 연결을 사용하는 경우 기울기가 네트워크를 통해 전달되므로 잔차 블록 가중치를 통과하지 않아도 되며 훈련 중에는 아무것도 학습하지 않아도 되므로 유용한 특

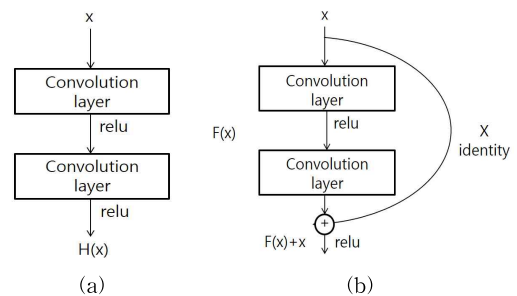


Fig. 1. Convolution layer and residual learning block. (a) CNN (b) residual learning block.

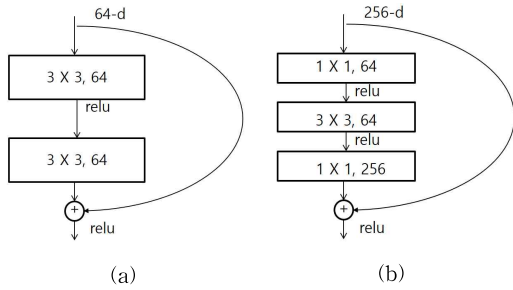


Fig. 2. Configuration of residual block, (a) 34-layer (b) bottleneck architecture.

징을 추출하는 블록이 거의 없거나 많은 블록이 최종 특징추출에 작은 영향을 주는 정보를 거의 공유하지 않을 수 있다.

ResNet은 네트워크의 레이어가 깊어질수록 파라미터의 수가 증가하고 학습에 걸리는 시간이 증가하는 문제를 해결하고자 50층, 101층, 152층 네트워크에 대해 Fig. 2(a)의 잔차블록 내 컨블루션 층 2개를 Fig. 2(b)의 1×1, 3×3, 1×1의 3층으로 변형하였다. 1×1 컨블루션 층은 입력과 출력의 차원을 감소하고 증가시키는 역할을 하고 이 구조가 병목처럼 보인다고 하여 병목구조라 한다.

ResNet은 잔차블록과 병목구조를 도입하여 네트워크를 구성하여 네트워크 구조를 깊게 하였음에도 빠른 속도로 이미지 분류에 성공하였으며 기울기 소실 문제도 해결하였다. 224×224 크기로 7×7 크기의 필터 64개를 적용하여 스트라이드(stride) 2로 처리하며 ResNet은 몇 개의 잔차블록을 그룹화(ConvX)하여 동일 그룹의 입력과 출력의 크기 및 필터의 개수를 동일하게 하였다. 또한 층이 깊어지는 그룹은 특징 맵의 크기를 1/2씩 줄이고 필터의 개수를 2배로 늘려 처리 그룹마다 연산량을 동일하게 하였다. 또한 연산량을 감소시키기 위해 첫 번째 컨블루션 층을 제외하고는 맥스 풀링(max-pooling), 드랍아웃(drop-out) 등을 사용하지 않았고 하나의 완전연결층(fully-connected layer)만 사용하였다.

3. 컨블루션 레이어의 증가를 이용한 변형된 잔차블록 및 병목블록 구조

ResNet은 영상의 특징을 잔차블록의 그룹으로 구성된 컨블루션 블록(ConvX)을 이용하여 추출한다. 이때 잔차블록의 숏컷 연결은 기울기 소실 문제를

해결하지만 잔차블록의 특징 추출을 저해할 수 있다. 따라서 잔차블록의 특징 추출을 증가시킬 수 있는 방법이 필요하며 본 논문에서는 컨블루션 레이어가 쌓일수록 이미지의 특징 추출 성능이 좋아지는 특성을 이용하여 잔차블록 내 컨블루션 레이어를 쌓아 이미지 분류성능을 개선한다. 먼저 잔차블록 내 컨블루션 레이어를 쌓는 기본구조로 본 논문에서는 Fig. 2(a)의 ResNet 기본 구조로 컨블루션 레이어를 쌓아 특징 추출 성능을 보이고 병목구조에 적용한다.

3.1 제안하는 잔차블록 구조

Fig. 3은 본 논문의 네트워크 모델을 구성하는 잔차블록 구조를 보여준다. Fig. 3(a)는 기본구조로 잔차블록은 3×3 컨블루션 필터를 m개 갖는 컨블루션 층 k개가 하나의 숏컷 연결로 구성되며 하나의 컨블루션 그룹 내에 잔차블록은 l개가 있다. Fig. 3(b)는 병목구조에서의 구조로 잔차블록은 1×1 컨블루션 필터 사이에 3×3 컨블루션 필터 m개로 구성되며 n=4×m 이다.

3.2 잔차블록 및 병목블록을 이용한 모델 구성

잔차블록으로 구성된 컨블루션 그룹 외의 네트워크 구조는 ResNet의 기본구조와 동일하게 구성하였다. ResNet은 Fig. 4(a)의 구조로 구성되며 224×224 크기의 입력 영상을 Conv1 그룹에서 처리한다. Conv1은 7×7 크기의 64개 필터가 특징을 추출하며

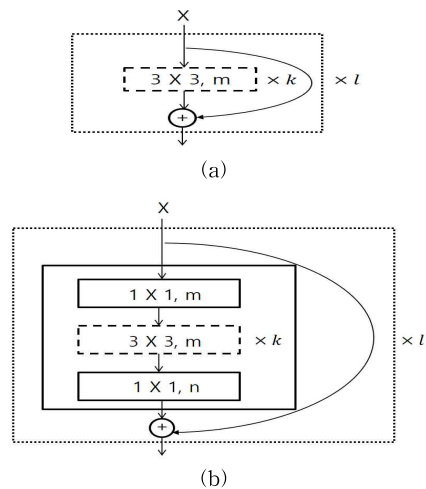


Fig. 3. Proposed residual block structure, (a) stack architecture (b) bottleneck architecture.

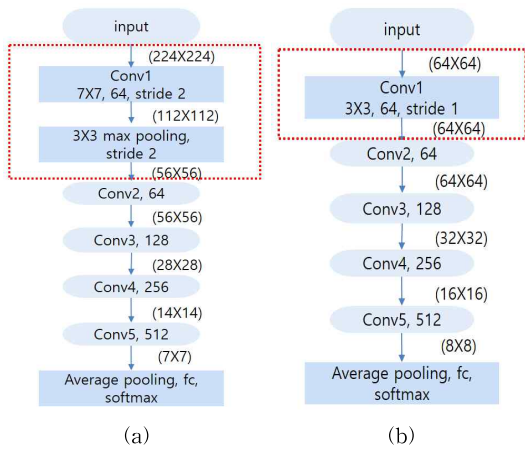


Fig. 4. Model configuration according to input size. (a) 224×224 (b) 64×64.

스트라이드 2로 처리하여 잔차블록 그룹 Conv2 블록으로 입력된다. Conv2의 입력부에서 스트라이드 2가 적용된 3×3 맥스 풀링을 거쳐 다수의 잔차블록으로 이루어진 Conv2 블록에서 특징이 추출된다. Conv2~Conv5 다수의 잔차블록을 출력의 크기 및 필터 개수에 따라 그룹으로 묶은 처리 그룹이다. 특징 추출 후 평균 풀링(Average pooling) 층과 완전연결층을 거쳐 이미지를 분류한다. 본 논문에서는 입력영상의 크기가 224×224인 경우보다 작은 경우를 처리할 경우를 고려하여 224×224 이상의 이미지는 ResNet 기본 구조와 동일하게 처리하며 더 작은 크기의 영상을 처리할 수 있도록 Fig. 4(b)와 같이 변형된 구조를 설계하였다. 입력의 크기가 224×224 보다 작은 경우 64×64 크기로 하여 처리하도록 구성하였으며 설선 박스 부분이 변형된 부분이다. Fig. 4에서 ()는 영상의 크기이며 컨블루션 층의 필터의 개수 m은 ResNet과 동일하게 Conv2는 64, Conv3는 128, Conv4는 256, Conv5는 512로 하였다.

본 논문에서는 잔차블록 구성을 컨블루션 레이어와 숏컷 연결로 구성된 단순구조(SS:Simple Structure)와 잔차블록이 병목구조(BS:Bottleneck Structure)로 이루어진 경우로 나누어 모델을 구성하였으며 단순구조는 34 레이어와 50 레이어의 모델을 구성하였고 병목구조는 50 레이어 모델을 구성하였다. 단순구조의 경우 한 컨블루션 블록(ConvX) 내 컨블루션 레이어가 동일한 개수가 포함이 되도록 구성하였다. 34 레이어의 경우 각 컨블루션 블록 내 컨블루션

레이어가 $8(k \times l = 12)$ 레이어, 50 레이어인 경우 12 ($k \times l = 12$) 레이어로 구성하였다. 이것은 전체 네트워크 레이어의 수와 ResNet의 기본 처리 구조인 Fig. 4의 구조를 적용하기 위한 것이다. 또한 ResNet에서 50레이어에 병목구조를 도입한 것은 파라수의 증가로 인한 속도저하 때문이므로 본 논문에서는 단순구조 구성은 잔차블록 내 컨블루션 레이어를 너무 많이 쌓아 속도 저하가 일어나는 것을 방지하기 위해 50 레이어로 제한하였고 제한방법을 병목구조에 적용하였다. 병목구조인 경우는 1×1 컨블루션 층 사이에 3×3 컨블루션 레이어를 반복적으로 쌓아 전체 네트워크를 50 레이어로 구성하도록 하였다. Table 1은 Fig. 2의 잔차블록 내 컨블루션 레이어의 수 k 와 컨블루션 블록(ConvX) 내 잔차블록의 수 l 를 이용하여 모델 구성을 정리한 표로 k 와 l 은 컨블루션 블록 Conv2~Conv5에 동일하게 적용하였다.

3.3 실험 및 결과분석

제안방법의 성능을 분석하기 위한 서버 환경 구성은 Xeon Processor E5-2620 v4 8/16 2.10Ghz 20M, 우분투 리눅스 서버 16.01, GPU는 NVIDIA TITAN V, 파이썬 3.6 텐서플로 1.8, Keras 2.2.4 환경에서 실험하였다. 데이터 셋은 Caltech-101 데이터셋[13]과 CIFAR-10 데이터 셋[14]를 사용하였다. 실험 환경은 배치 사이즈 64, 옵티마이저는 Adam, 학습률은 0.0001에서 시작해서 40에폭마다 10%씩 감소하도록 하였으며 데이터 전처리와 데이터 증식은 적용하지 않았다.

Table 2은 각 모델의 전체 파라미터수와 1 에폭 당 평균처리시간이다. 34 레이어 및 50 레이어 단순구조의 경우 파라미터수가 동일하며 1에폭 처리시간이 차이가 거의 없다. 병목구조이 경우도 모델별로

Table 1. Proposed model structure

model type	variable	SS		BS
		34	50	50
T1	k	1	2	1
	l	8	6	4
T2	k	2	3	2
	l	4	4	3
T3	k	4	4	4
	l	2	3	2

Table 2. The number of parameters and processing time for each model type

Network layer		model type	total parameter	training parameter	processing time(sec)	
					Caltech-101	CIFAR-10
SS	34	T1	23,744,898	23,727,618	21	157
		T2	23,744,898	23,727,618	20	156
		T3	23,744,898	23,727,618	20	156
	50	T1	36,297,858	36,272,898	26	211
		T2	36,297,858	36,272,898	27	210
		T3	36,297,858	36,272,898	26	209
BS	50	T1	25,891,586	25,837,698	34	77
		T2	29,358,786	29,310,658	34	77
		T3	32,825,986	32,783,618	33	78

파라미터의 수는 증가하지만 1에폭당 처리시간에서는 차이가 많지 않음을 볼 수 있다.

3.3.1 단순구조(SS)의 성능분석

Fig. 5는 에폭을 100으로 했을 때 데이터셋의 훈련 데이터와 검증데이터의 정확도 그래프이다.

Fig. 5에서 T1~T3의 training/validation 데이터셋이 Caltech-101 데이터 셋은 40에폭 근처에서, CIFAR-10 데이터 셋은 80 에폭 근처에서 수렴함을

알 수 있다.

Table 3은 에폭을 150으로 하였을 경우 테스트 셋의 정확도(accuracy)를 측정된 결과를 보여준다. Table 3의 결과는 단순구조의 경우 잔차블록 내 콘블루션 레이어를 증가했을 때 성능이 증가함을 보여준다.

3.3.2 병목구조의 성능분석

Fig. 6은 에폭을 100으로 했을 때 데이터 셋의 훈

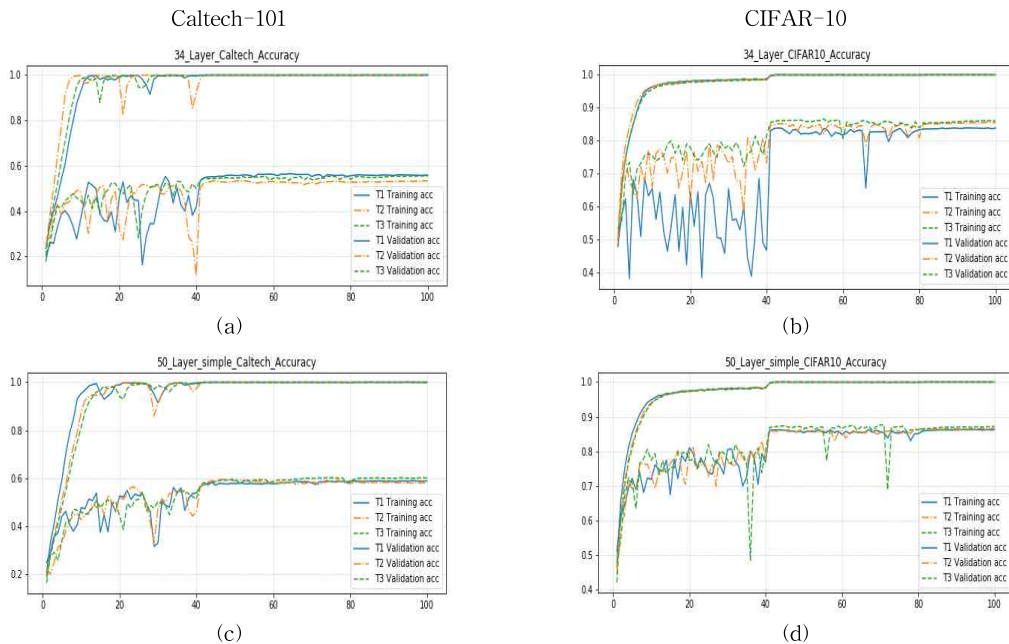


Fig. 5. Train/validation accuracy of caltech-101 and CIFAR-10 at epoch 100. (a) 34 layer(Caltech-101) (b) 34 layer(CIFAR-10) (c) 50 layer(Caltech-101) (d) 50 layer(CIFAR-10).

Table 3. Performance of model type

layer of network	model type	CIFAR10	Caltech-101
34	T1	81.35%	54.83%
	T2	81.86%	46.22%
	T3	83.17%	51.10%
50	T1	81.51%	51.63%
	T2	82.67%	54.93%
	T3	83.75%	54.93%

런데이터와 검증데이터의 정확도 그래프이다.

Fig. 6에서 Caltech-101와 CIFAR-10 데이터 셋의 대부분이 T1~T3의 훈련 데이터 및 검증 데이터의 정확도가 40에폭 근처에서 수렴하나 Caltech-101은 T1 모델의 검증데이터는 80에폭 근처에서, CIFAR-10은 T3 모델의 검증 데이터 정확도가 80 에폭 근처에서 수렴한다.

Table 4는 병목구조에 콘블루션 레이어를 다르게 구성하였을 때의 테스트 셋의 정확도를 보여준다. 병목구조에서도 단순구조와 동일하게 콘블루션 레이어가 많을 경우 성능이 증가함을 볼 수 있다.

Table 3와 Table 4 및 Fig. 6의 결과는 단순구조와 병목구조 모두에서 잔차블록 내 콘블루션 레이어의 수가 많은 경우가 이미지 분류 성능이 증가함을 볼 수 있다.

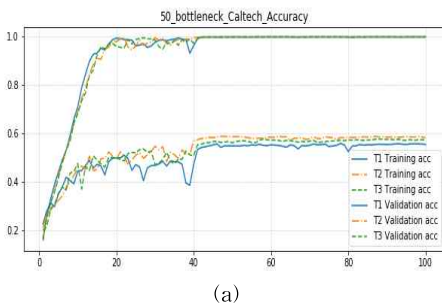


Table 5. The number of / of convolution blocks

layer name \ model	SS	BS
	34	50
conv2_x	1	1
conv3_x	2	2
conv4_x	3	3
conv5_x	2	2

4. 변형된 잔차블록을 이용한 CNN

4.1 변형된 잔차블록을 이용한 이미지 분류 네트워크

얕은 레이어의 CNN은 이미지 분류성능이 저조하며 깊은 레이어의 CNN은 기울기 소실문제나 과적합 등의 문제가 발생한다. 본 논문에서는 적절한 수의 레이어로 구성되면서 개선된 이미지 분류 성능을 보이는 CNN모델을 제안한다. 이 모델은 잔차블록에 콘블루션 레이어를 연속적으로 쌓았을 때 4개의 콘블루션 블록을 쌓은 경우가 단순구조와 병목구조 모두에서 가장 좋은 성능을 보인 3.3절의 결과를 반영하여 Fig. 4의 구조를 이용하여 네트워크를 구성하며 100 레이어 이하로 구성하도록 하였다. 이러한 조건에 맞는 네트워크 구성 시 50 레이어가 넘어가면 처리 속도의 문제로 인해 병목구조를 적용하는 것이 효율적이다. 따라서 본 논문에서는 응용에 따라 적용할 수 있도록 약간 얕은 구조를 갖는 34 단순구조와 좀 더 깊은 구조를 갖는 50 bottleneck 구조를 제안한

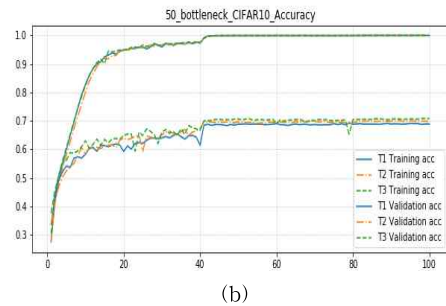


Fig. 6. Train/validation accuracy of caltech-101 and CIFAR-10. (a) Caltech-101 (b) CIFAR-10.

Table 4. Performance of model type of bottleneck structure

layer of network	model type	CIFAR10	Caltech-101
50 (bottleneck)	T1	63.46%	55.25%
	T2	63.60%	56.30%
	T3	66.28%	58.71%

Table 6. Performance of methods (CIFAR10)

Method	Depth	Param	Error Rate(%)
ResNet[10]	32	0.46M	7.51
	44	0.66M	7.17
	56	0.85M	6.97
	110	1.7M	6.43
DenseNet(k=12)[15]	40	1M	5.24
DenseNet-BC(k=12)[15]	100	0.8M	4.51
Proposed-SS34	34	23.7M	4.95
Proposed-BS50	50	32.8M	5.12

다. 이때 각 콘블루션 블록의 잔차블록과 병목 블록의 반복횟수(l)는 Table 6에 있다. Table 5의 conv2_x~conv5_x의 l의 수는 실험에 의해 최적의 값으로 결정된 것이다.

4.2 실험결과 및 분석

제안 방법의 성능을 분석하기 위한 서버 환경 구성과 프로그래밍 환경은 3.3절과 동일하며 배치 사이즈 64, 옵티마이저는 Adam, 학습률은 0.0001에서 시작해서 40에폭마다 10%씩 감소하도록 하였다. 제안 방법의 모델은 Table 5의 3개의 모델을 구성하여 ResNet 모델, DenseNet 모델[15], 제안방법의 결과를 비교한다. 데이터 셋은 CIFAR10과 Caltech-101을 사용했으며 데이터 전처리로 입력 영상에서 평균을 빼고 표준편차로 나누어 정규화 하고 데이터 증식을 적용했다.

Table 6은 CIFAR10 데이터 셋을 기존방법과 제안방법으로 테스트한 결과로 제안방법인 단순구조 34 레이어(Proposed-SS34)와 병목구조 50 레이어(Proposed-BS50) 모두 ResNet 110 레이어와 DenseNet(k=12)인 스택 구조 40레이어 보다 좋은 결과를 보인다. 그러나 DenseNet (k=12)의 100레이어 병목

구조보다는 낮은 성능을 보인다.

Table 7은 Caltech-101 데이터 셋을 기존방법과 제안방법으로 테스트한 결과를 보여준다. 제안방법 단순구조 34 레이어(Proposed-SS34)와 병목구조 50 레이어(Proposed-BS50)는 ResNet50 레이어와 DenseNet40 레이어 보다 좋은 결과를 보인다. 그러나 DenseNet121 레이어 보다는 낮은 성능을 보인다.

Table 6과 Table 7의 결과에서 DensetNet의 100 레이어 이상의 깊은 레이어의 결과에 비해 제안방법의 성능은 약간 저조하나 레이어가 유사한 DenseNet40보다 개선된 성능을 보인 것은 제안방법의 성능이 이미지 분류 성능을 개선했음을 보여준다. 또한 네트워크의 레이어가 깊지 않은 34 레이어와 50 레이어인 경우임에도 탁월한 이미지 분류성능을 보여준다. 이와 같은 결과는 본 논문의 단순구조 34 레이어와 병목구조 50 레이어가 이미지 분류 성능을 개선함을 보여준다.

5. 결론

본 논문은 CNN의 대표적인 방법인 ResNet의 잔차블록 내 콘블루션 레이어의 수를 변형하여 구성된 이미지 분류 모델을 제안한다. 제안방법은 ResNet의 잔차블록과 병목구조로 이루어진 네트워크 중 34/50 레이어의 구조를 변형하였다. 먼저 콘블루션 레이어와 숏컷 연결로 이루어진 34/50 레이어에서 콘블루션 레이어의 수를 증가시켜 Caltech-101과 CIFAR10의 데이터 셋에 대해 이미지 분류 성능을 비교하였다. 또한 이미지 분류성능이 가장 좋은 4개의 콘블루션 레이어를 쌓는 방법을 적용하여 34 레이어 단순구조와 50 레이어 병목구조의 이미지 분류 모델을 구성

Table 7. Performance of methods (Caltech-101)

Method	Depth	Error Rate(%)
ResNet	34	11.32
	50	9.18
DenseNet	40	8.65
	121	7.33
Proposed-SS34	34	7.93
Proposed-BS50	50	8.12

하였고 기존의 방법과 이미지 분류 성능을 비교하였다. 제안한 34 레이어 단순구조와 50 레이어 병목구조는 CIFAR10 과 Caltech 101 데이터 셋에 대해 ResNet 110 이하 모델과 DenseNet-40 모델에 대해서도 개선된 이미지 분류 성능을 보여준다. 이러한 결과는 잔차블록 내 콘블루션 레이어 수를 변형하여 레이어가 50 이하인 모델로서 이미지 분류 성능을 개선한 모델임을 나타낸다.

REFERENCE

- [1] Y.L. Cun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based Learning Applied to Document Recognition," *Proceeding of the IEEE*, pp. 2278-2324, 1989.
- [2] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-scale Image Recognition," *Proceeding of the International Conference on Learning Representations*, pp. 1-14, 2015.
- [3] S. Hochreiter, "The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions," *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, Vol. 6, No. 2, pp. 107-115, 1998.
- [4] X. Glorot and Y. Bengio, "Understanding the Difficulty of Training Deep Feed Forward Neural Networks," *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, pp. 249-256, 2010.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-level Performance on Imagenet Classification," *Proceedings of the International Conference on Computer Vision*, pp. 1026-1034, 2015.
- [6] I. Sutskever, J. Martens, G.E. Dahl, and G.E. Hinton, "On the Importance of Initialization and Momentum in Deep Learning," *Proceedings of the 30th International Conference on Machine Learning*, pp. 1139 - 1147, 2013.
- [7] C.Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised Nets," *arXiv:1409.5185v2*, 2014
- [8] A. Romero, N. Ballas, S.E. Kahou, A. Chas-sang, C. Gatta, and Y. Bengio, "FitNets: Hints for Thin Deep Nets," *arXiv:1412.6550 v4*, 2015
- [9] J. Schmidhuber, "Learning Complex, Extended Sequences Using the Principle of History Compression," *Neural Computation*, Vol. 4, No. 2, pp. 234-242, 1992.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016.
- [11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., "Going Deeper with Convolutions," *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-9, 2015.
- [12] R.K. Srivastava, K. Greff, and J. Schmidhuber, "Highway Networks," *arXiv:1505.00387v2*, 2015.
- [13] L. Fei-Fei, R. Fergus, and P. Perona, "Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories," *Computer Vision and Image Understanding*, Vol. 106, No. 1, pp. 59-70, 2007.
- [14] Canadian Institute for Advanced Research (2012), <https://www.cs.toronto.edu/~kriz/cifar.html> (accessed June 11, 2020).
- [15] G. Huang, Z. Liu, L.V.D. Maaten, and K.Q. Weinberger, "Densely Connected Convolutional Networks," *arXiv:1608.06993v5*, 2008
- [16] J.T. Lee, H. Kang, K.T. Lim, "Moving Shadow Detection using Deep Learning and Markov Random Field," *Journal of Korea Multimedia Society*, Vol. 18, No. 12, pp. 1432-1438, 2015.



곽 내 정

2005년 2월: 충북대학교 정보통신 공학과 공학박사
2005년 3월~2006년 2월 목원대학교 정보통신 공학과 프로그래밍 전문강사
2006년 3월~2009년 2월 목원대학교 정보통신 공학과 강

의전임

2009년 3월~2020년 2월 충북대학교, 배재대학교 시간 강사

2020년 3월~현재 배재대학교 AI·SW 창의융합대학 사이버보안학과 강의전임

관심분야 : 영상처리, 멀티미디어 프로그래밍, 컴퓨터 비전, 딥러닝, 행동인식 및 상황인지



신 현 준

2019년 8월: 목원대학교 융합컴퓨터미디어학부 공학사
관심분야 : 딥러닝, 객체 인식, 활동 인식, 컴퓨터 비전, 머신러닝



양 종 섭

2020년 2월 목원대학교 융합컴퓨터미디어학부 공학사
2020년 3월~현재 (주)KW이노베이션
관심분야 : 사물인터넷, 데이터베이스, 딥 러닝 네트워크, 클라우드 솔루션



송 특 섭

2001년 2월 연세대학교 수학과 (이학박사)
2006년 2월 연세대학교 컴퓨터과학과(공학박사)
2006년 3월~현재 목원대학교 융합컴퓨터미디어학부 교수

관심분야 : 웹환경 어노테이션, 딥러닝, 센서네트워크