

배치 정규화와 CNN을 이용한 개선된 영상분류 방법☆ An Improved Image Classification Using Batch Normalization and CNN

지 명 근¹ 전 준 철^{1*} 김 남 기¹
Myunggeun Ji Junchul Chun Namgi Kim

요 약

딥 러닝은 영상 분류를 위한 여러 방법 중 높은 정확도를 보이는 방법으로 알려져 있다. 본 논문에서는 딥 러닝 방법 가운데 합성곱 신경망 (CNN: Convolutional Neural Network)을 이용하여 영상을 분류함에 있어 배치 정규화 방법이 추가된 CNN을 이용하여 영상 분류의 정확도를 높이는 방법을 제시하였다. 본 논문에서는 영상 분류를 더 정확하게 수행하기 위해 기존의 뉴럴 네트워크에 배치 정규화 계층 (layer)를 추가하는 방법을 제안한다. 배치 정규화는 각 계층에 존재하는 편향을 줄이기 위해 고안된 방법으로, 각 배치의 평균과 분산을 계산하여 이동시키는 방법이다. 본 논문에서 제시된 방법의 우수성을 입증하기 위하여 SHREC13, MNIST, SVHN, CIFAR-10, CIFAR-100의 5개 영상 데이터 집합을 이용하여 영상분류 실험을 하여 정확도와 mAP를 측정한다. 실험 결과 일반적인 CNN 보다 배치 정규화가 추가된 CNN이 영상 분류 시 보다 높은 분류 정확도와 mAP를 보임을 확인할 수 있었다.

☞ 주제어 : 배치 정규화, 합성곱 신경망, 영상 분류, 딥 러닝

ABSTRACT

Deep learning is known as a method of high accuracy among several methods for image classification. In this paper, we propose a method of enhancing the accuracy of image classification using CNN with a batch normalization method for classification of images using deep CNN (Convolutional Neural Network). In this paper, we propose a method to add a batch normalization layer to existing neural networks to enhance the accuracy of image classification. Batch normalization is a method to calculate and move the average and variance of each batch for reducing the deflection in each layer. In order to prove the superiority of the proposed method, Accuracy and mAP are measured by image classification experiments using five image data sets SHREC13, MNIST, SVHN, CIFAR-10, and CIFAR-100. Experimental results showed that the CNN with batch normalization is better classification accuracy and mAP rather than using the conventional CNN.

☞ keyword : Batch Normalization, Convolutional Neural Network, Image Classification, Deep Learning

1. 서 론

영상 분류 및 영상 검색에 있어 인공지능 기술인 딥 러닝 (Deep Learning) 기반 학습 방법이 주류가 된 까닭은 2012년 ILSVRC (Large Scale Visual Recognition Challenge) 에서 딥 러닝에 기반 한 AlexNet (SuperVision) [1]이 기존의 방법인 SIFT나 SVM을 이용한 ISI [2], VGG [3]에 비하여 우수한 분류 결과를 보이면서 영상 분류에서 딥 러닝은 주요 방법으로 인식되고 있다. [4] 본 논문에서는 딥 러닝의 여러 방법 중 영상 분류에 주로 사용되는 CNN (Con-

volutional Neural Network)의 성능을 향상시키기 위하여 배치 정규화 (Batch Normalization)를 적용한 CNN을 이용한 영상 분류 방법을 제시하고 기존의 CNN과 배치 정규화가 적용된 개선된 CNN에 의한 영상 분류 결과를 제시하여 제안된 방법의 효율성을 제시한다. 본 연구에서 제시된 방법의 효율성을 입증하기 위하여 SHREC13, MNIST, SVHN, CIFAR-10 및 CIFAR-100의 서로 다른 5개의 영상 데이터 집합을 이용한 영상 분류 실험을 진행하였다. 실험 결과 배치 정규화를 적용한 CNN 모델이 일반적인 CNN을 이용하여 영상을 분류한 결과 보다 성능이 우수함을 입증 하였다.

본 논문은 다음과 같이 구성된다. 먼저 2장에서는 본 논문에서 제시한 방법과 관련된 딥 러닝 기반의 영상 분류 및 영상 검색 방법을 소개하고, 제안된 접근법을 통하여 해결하고자 하는 문제를 설명한다. 3장에서는 본 논문에서 제안하는 배치 정규화가 적용된 CNN 모델을 제시한

1 Department of Computer Science, Kyonggi University, Gyeonggi-do, 443-760, Korea.

* Corresponding author (jcchun@kyonggi.ac.kr)

[Received 26 December 2017, Reviewed 1 January 2018 (R2 21 February 2018, R3 22 March 2018), Accepted 4 April 2018]

☆ 본 연구는 2018학년도 경기대학교 대학원 연구원장학생 장학금 지원에 의하여 수행되었음.

다. 4장에서는 제안된 방법에 의한 실험 결과를 소개하며, 마지막으로 5장에서는 결론을 맺는다.

2. 관련 연구

영상 데이터베이스를 기반으로 영상과 비디오의 객체를 분류하고 탐지하는 연구는 ILSVRC 에서 딥 러닝이 방법이 소개되면서 해당분야에 중요한 수단으로 발전하는 계기가 되었다.[4]

2012년 ILSVRC의 딥 러닝에 기반 한 AlexNet이 오차를 16%로 2등의 26%와 약 10% 차이를 벌리며 우승한 뒤로 ILSVRC의 순위권은 모두 딥 러닝 기반의 방법이 차지하게 되었다. 2013년의 ILSVRC도 딥 러닝 방법을 쓴 ZF Net[5]가 12%의 정확도를 보이며 우승을 차지했으며, 2014년의 ILSVRC도 GoogLeNet[6]이 7%의 정확도로 우승을, 2015년 ILSVRC도 ResNet[7]이 4%의 정확도로 우승을 차지했다. 2016년 ILSVRC에서는 Trimps-Soushen 팀이 3%의 정확도를 보였다. 이렇듯 딥 러닝 기반의 영상 분석 방법은 기존의 방법들보다 빠른 속도로 정확도를 높이고, 이를 바탕으로 뇌 영상 분석, 유방암 진단 등 의료영상 분석이나[8], 페이스북의 얼굴 인식 기술인 DeepFace[9]등의 서비스에도 사용되고 있다.

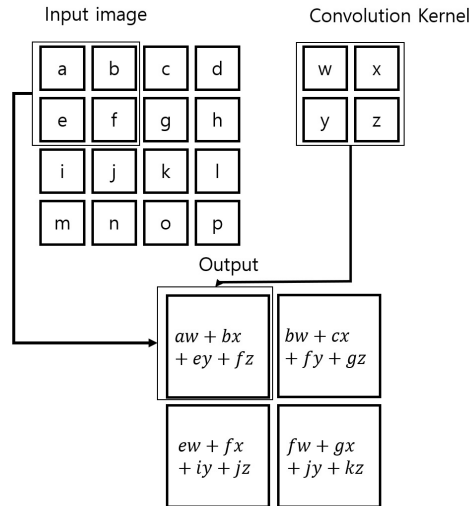
이처럼 우수한 분류 결과를 나타내는 이면에는 딥 러닝 구조를 설계하는 데에 있어 딥 러닝의 문제인 그라디언트 소실(Vanishing Gradient), 과적합(Overfitting)같은 문제를 해결할 필요가 있었다. 이러한 문제를 해결하기 위한 방법들로, 그라디언트 소실 문제에는 활성화 함수를 Sigmoid에서 ReLu(Rectified Linear Unit)[10]등으로 변경하는 방법이 사용되었으며, 과적합을 해결하기 위해서 드롭아웃[11]의 방법 등이 적용되었다.

3. 배치 정규화와 CNN을 결합한 영상 분류 방법

3.1 CNN (Convolution Neural Network)

CNN 이란 영상을 학습하기 위해 주로 사용되는 딥 러닝 네트워크의 한 종류이다. 일반적인 딥 러닝은 한 층의 노드가 그 다음 층의 노드와 전부 연결되는 완전연결 형태인데, 이는 과도한 파라미터를 만드는 원인이 되어 학습 속도를 느려지게 하고, 과적합의 원인이 된다. 이러한 문제를 해결하기 위해 합성곱 연산을 도입한 딥 러닝인 CNN이 개발되었다.[1] CNN의 컨볼루션 계층은 (그림 1)과 같이 영상의 부분에 해당하는 원소들과 컨볼루션

커널의 원소들을 하나씩 곱하고, 그들을 모두 더한 값을 출력 원소의 값으로 사용한다. (그림 1)에서 입력영상 I 에 컨볼루션 커널 F 가 적용된 출력 영상 G 는 다음의 식 (1) 과 같이 정의된다.

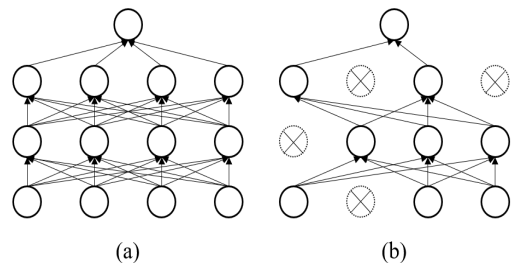


(그림 1) 컨볼루션 레이어의 작동 원리

(Figure 1) The Procedure of Convolution Layer

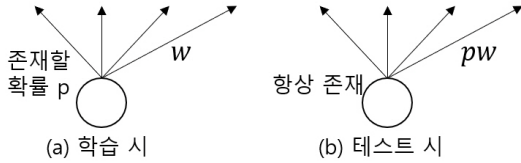
$$G[i, j] = \sum_{u=-k}^k \sum_{v=-k}^k I[u, v] F[i - u, j - v] \quad (1)$$

CNN의 학습과정에 있어 제기되는 문제점 가운데 일반적으로 딥 러닝 네트워크가 지나치게 학습 데이터에 집중되는 현상인 과적합을 막기 위하여 드롭아웃(Dropout)의 방법이 적용된다.[11] (그림 2-a)는 일반적인 딥 러닝 네트워크, (그림 2-b)는 드롭아웃이 적용된 딥 러닝 네트워크를 나타낸다.



(그림 2) 딥 러닝 네트워크와 드롭아웃이 적용된 딥 러닝 네트워크
(Figure 2) Deep Learning Network and Deep Learning Network with Dropout

(그림 3)은 드롭아웃이 적용된 노드의 학습과 테스트 시 작동 방식을 나타낸다. 각 노드는 동작할 확률 p 를 가지며, 학습 시에는 p 의 확률로 노드가 학습될지 또는 그렇게 되지 않을지를 결정하고, 테스트 시에는 노드의 가중치에 p 를 곱하는 방식으로 작동한다. 이는 모델 여러 개를 학습시켜 그 결과에 투표를 적용한 방법과 같은 효과를 보인다.



(그림 3) 드롭아웃이 적용된 노드의 학습 시와 테스트 시 작동 방식
(Figure 3) Dropout node's training and testing working process

3.2 배치 정규화 레이어 추가 적용 방법

배치 정규화[12]란 네트워크의 Internal Covariate Shift 문제를 해결하기 위해 제안된 방법 중 하나로, 각 어파인 계층(Affine layer)을 통과한 미니배치 단위 출력을 표준정규분포로 정규화 하는 작업이다. Internal Covariate Shift란 네트워크 각 층 입력의 분포가 달라지는 현상을 의미한다. 이러한 현상을 방지하기 위하여 각 계층 입력의 분포를 평균이 0, 표준편차가 1이 되도록 일 반화를 시키게 된다. 그러나 입력의 평균을 이동시키고, 표준편차를 변화시킬 때에 일부 파라미터들의 영향이 무시될 수 있다. 이러한 단점을 상쇄시키기 위하여 추가적으로 스케일, 쉬프트 연산을 추가한다. 이 때 스케일, 쉬프트 하는 값은 일반 파라미터와 같이 네트워크 학습 시 같이 학습되게 된다.

일반화 시 평균과 표준편차는 학습 과정에서 사용하는 미니배치의 평균과 표준편차를 이용한다. 테스트 시에는 이 미니배치들의 평균과 표준편차의 평균을 테스트 시의 평균과 표준편차로 이용하게 된다. 다음 식은 이상에 설명된 배치 정규화 과정에서 미니배치의 평균, 미니배치의 분산을 계산하고, 그 값을 기준으로 정규화 하고, 스케일과 쉬프트 연산을 수행하는 과정 수행하는 과정을 보여주고 있다.

$$\begin{aligned} \text{Input: } B &= \{x_1 \dots x_m\} \\ \text{Output: } \{y_i &= BN_{\gamma, \beta}(x_i)\} \end{aligned} \quad (2)$$

배치 정규화의 입력 x 는 입력 영상, 혹은 이전 레이어의 출력이다. 출력의 $BN_{\gamma, \beta}(x_i)$ 은 입력 x 가 정규화 되고, 스케일, 쉬프트 연산을 거친 결과이다.

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad (3)$$

식 (3)과 같이 미니배치의 평균 μ_B 을 계산한다.

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (4)$$

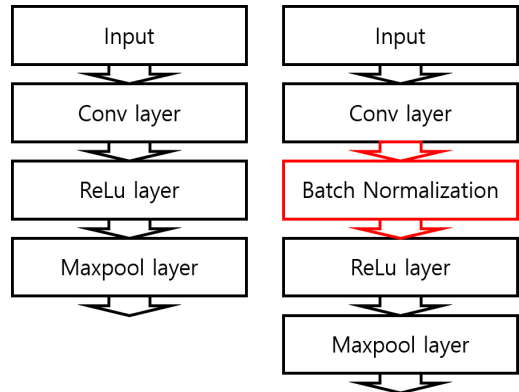
식 (4)와 같이 미니배치의 분산 σ_B^2 을 계산한다.

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (5)$$

식 (5)는 식 (3)에서 계산한 미니배치의 평균 μ_B 과 식 (4)에서 계산한 미니배치의 분산 σ_B^2 을 이용하여 미니배치를 정규화 한 값 \hat{x}_i 을 계산한 값이다. ϵ 은 분모가 0인 상황을 대비하여 추가된 아주 작은 값이다.

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma, \beta}(x_i) \quad (6)$$

식 (6)은 식 (5)에서 정규화가 끝난 값에 스케일(γ), 쉬프트(β) 연산을 수행한 결과이다.



(그림 4) 배치 정규화가 추가된 CNN 모델
(Figure 4) CNN Model with Batch Normalization

본 논문에서 제시하는 배치 정규화가 추가된 CNN의 구조는 (그림 4)와 같다. (그림 4)의 좌측 다이어그램은 기

존 CNN 구조이며, 우측 다이어그램은 본 논문에서 제안한 CNN 네트워크 구조에 배치 정규화가 적용된 구조이다. 배치 정규화는 그림과 같이 컨볼루션 계층과 ReLU 계층 사이에 위치하게 된다.

4. 실험 결과

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (7)$$

$$Training\ error = \frac{F_p + F_n}{T_p + T_n + F_p + F_n}$$

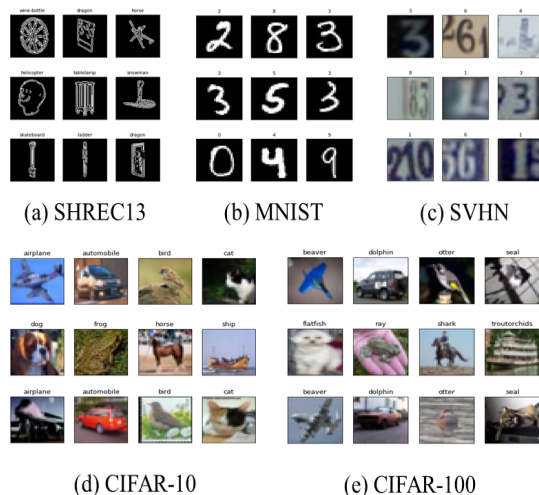
$$Precision = \frac{T_p}{T_p + F_p}$$

$$Recall = \frac{T_p}{T_p + F_n}$$

제안된 배치 정규화를 적용한 CNN의 성능을 실험하기 위해 총 5가지 영상 데이터 집합을 이용하여 실험을 진행하였다. 실험은 전통적인 CNN 구조 네트워크와 배치 정규화를 추가한 CNN 구조 네트워크에 각각의 데이터 집합 별로 학습 단계-트레이닝 에러 곡선(Training error-training step curve)과 학습 단계-정확도 곡선(Training accuracy-training step curve), 학습 완료 시 정확도, mAP(Mean Average Precision), 정확도-재현율 곡선(Precision-Recall)의 5가지 요소를 측정하여 성능을 평가하였다. 각각의 요소를 계산하기 위해 식 (7)의 공식을 사용하였다. 식 (7)에서 사용한 T_p, T_n, F_p, F_n 의 의미는 다음과 같다. 임의의 테스트 영상의 라벨 하나를 참, 거짓으로 구분할 때, 참인 라벨을 참으로 분류한 경우를 T_p , 거짓인 라벨을 거짓으로 분류한 경우를 T_n , 참인 라벨을 거짓으로 분류한 경우를 F_n 라고 하며, 거짓인 라벨을 참으로 분류한 경우를 F_p 라고 한다.

다음 (그림 5)는 영상 분류 실험에 사용된 영상 데이터셋의 샘플이다.

실험에 사용된 영상 데이터의 특성은 다음과 같다. SHREC13[13]은 3D 모델을 사람이 그린 2D 스케치로 검색하기 위하여 제공된 영상 데이터 셋으로 3D 모델을 서로 다른 두 각도에서 촬영한 영상의 윤곽선을 추출한 영상을 데이터 집합으로 사용했다. 90개 클래스를 가지고 있으며, 클래스 별 이미지가 8개에서 368개까지 불균형하게 존재한다. 학습 시 2246개의 영상을 사용하였으며, 테스트에서는 클래스 별로 3개씩의 영상을 선택한 총 270



(그림 5) 영상분류에 사용된 영상 데이터 셋 (Figure 5) Image Data Set for Image Classification

개의 영상을 테스트 영상으로 사용하였다.

MNIST는 필기체 숫자 인식을 위한 데이터 집합으로, 0부터 9까지 총 10개 클래스가 존재하며, 학습 시 60000개의 영상을 사용했으며 테스트 시에는 10,000개의 영상을 사용하였다.

한편 SVHN[14]은 구글의 스트리트 뷰에 찍힌 집 주소 영상 데이터 집합이다. MNIST 데이터와 마찬가지로 0부터 9까지 10개의 클래스가 존재하며, 학습 시 600,000개의 영상, 테스트 시에는 26,000개의 영상을 사용하여 실험 하였다.

CIFAR-10[15]은 작은 크기의 자연 영상 데이터 집합이다. 비행기, 자동차, 새, 고양이, 사슴, 개, 개구리, 말, 배, 트럭의 10가지 클래스가 존재하며, 학습 시 50,000개의 영상을 사용했으며, 테스트 시에는 10000개의 영상을 사용하여 실험 하였다.

CIFAR-100[15]은 CIFAR-10과 비슷하지만 클래스의 종류가 10개에서 100개로 늘어난 영상 데이터 집합이다. 각 클래스 당 600개의 영상이 존재하여 총 60,000개의 영상이 존재하며, 학습 시 50,000개의 영상을 사용하고, 테스트 시에는 10,000개의 영상을 사용하여 실험 하였다.

본 연구에서 사용한 네트워크 구조는 (그림 6)에 나타나 있으며 전통적인 CNN 네트워크 구조와는 다르게 배치 정규화를 추가한 CNN 네트워크 구조는 컨볼루션 계층, 전체 연결 계층과 ReLU 계층 사이에 배치 정규화 계

SHREC13		MNIST		SVHN		CIFAR-10		CIFAR-100	
CNN	CNN with BN	CNN	CNN with BN	CNN	CNN with BN	CNN	CNN with BN	CNN	CNN with BN
13x13 conv 32개		3x3 conv 32개		3x3 conv 16개		3x3 conv 16개		5x5 conv 64개	
	BN		BN		BN		BN		BN
ReLU		ReLU		ReLU		ReLU		ReLU	
Maxpool		Maxpool		Maxpool		Maxpool		Maxpool	
7x7 conv 64개		3x3 conv 64개		3x3 conv 32개		3x3 conv 32개		5x5 conv 64개	
	BN		BN		BN		BN		BN
ReLU		ReLU		ReLU		ReLU		ReLU	
Maxpool		Maxpool		Maxpool		Maxpool		Maxpool	
3x3 conv 256개		fc 625		3x3 conv 32개		3x3 conv 32개		5x5 conv 64개	
	BN		BN		BN		BN		BN
ReLU		ReLU		ReLU		ReLU		ReLU	
Maxpool		Maxpool		Maxpool		Maxpool		Maxpool	
fc 90		fc 10		3x3 conv 32개		3x3 conv 32개		fc 256	
					BN		BN		ReLU
				ReLU		ReLU		Maxpool	
				Maxpool		Maxpool		fc 10	
				ReLU		ReLU			
				Maxpool		Maxpool			
				fc 10		fc 10			

(그림 6) 배치 정규화가 추가된 CNN을 이용한 실험 (Figure 6) Experiment using CNN with Batch Normalization

층을 추가하였다. 다음 (표 1)은 영상데이터 셋에 따른 정확도 및 mAP를 나타낸다.

(표 1) 정확도 및 mAP (Table 1) Accuracy and mAP

영상 데이터	평가 척도	CNN	배치 정규화 + CNN
SHREC13	Accuracy	0.59	0.71
	mAP	0.52	0.77
MNIST	Accuracy	0.95	0.99
	mAP	0.90	0.98
SVHN	Accuracy	0.66	0.89
	mAP	0.70	0.95
CIFAR-10	Accuracy	0.24	0.68
	mAP	0.42	0.65
CIFAR-100	Accuracy	0.18	0.36
	mAP	0.14	0.36

본 연구의 실험 환경으로는 AMD Ryzen 7 1700 3GHz CPU, 16GB RAM, NVIDIA GeForce GTX 1080 Ti GPU, Windows 10 사양의 컴퓨터를 기반으로 Jupyter notebook 환경에서 텐서플로우를 사용하였다. (그림 7)은 제안된 방법과 기존 CNN을 이용한 영상분류를 Accuracy, Training Error, Precision에 따라 나타난 결과이다. 그림에 표시되는 Accuracy, Training Error의 범위로 각 결과의 최대,

최소값을 이용하였다.

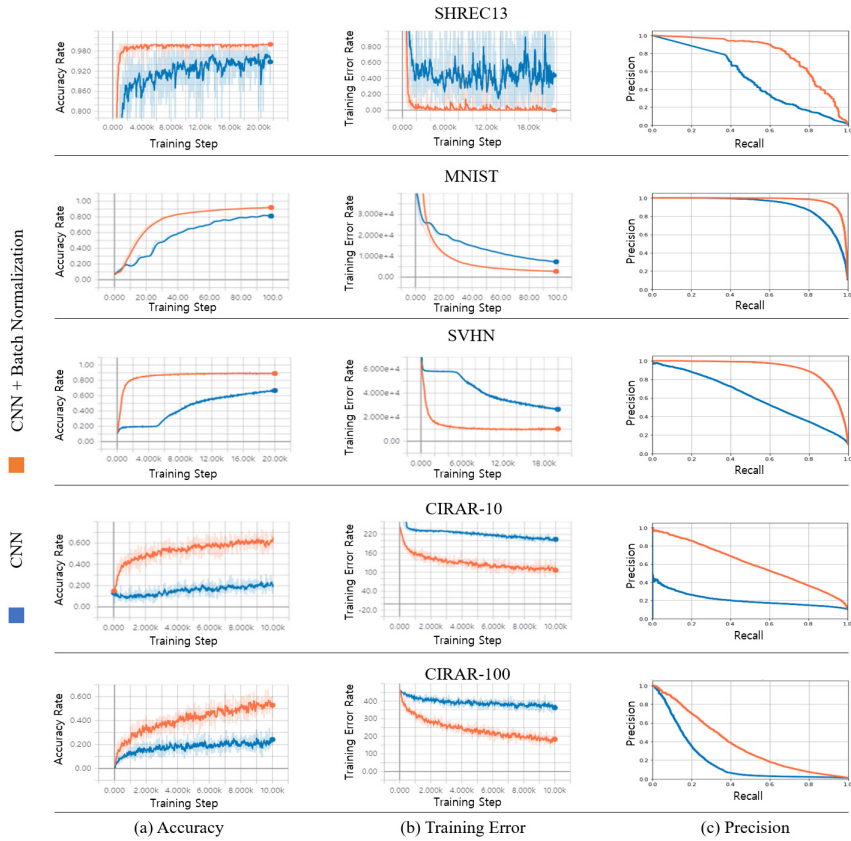
(그림 7)의 SHREC13 데이터 셋의 결과에서 학습 단계에 따른 정확도 곡선은 제안한 배치 정규화를 추가한 CNN 구조의 결과가 기존 CNN 구조보다 전반적으로 높은 정확도를 보이고, 학습 단계에 따른 트레이닝 에러 곡선에서는 전반적으로 더 낮은 에러를 보이는 것을 볼 수 있다. 이는 배치 정규화가 적용되었을 경우 CNN의 학습 속도가 빨라지고, 정확도가 높아지는 것을 의미한다. 또한 제안한 구조의 결과가 기존 구조보다 값이 안정적으로 변함을 알 수 있는데, 이는 배치 정규화를 적용되었을 경우 학습이 보다 안정적이게 됨을 알 수 있다. 이는 PR-곡선을 보았을 때에도 제안한 구조의 결과가 정밀도의 하락이 더 낮음으로 알 수 있다. 최종 테스트 결과도 제안한 구조의 정확도가 0.71, mAP가 0.77로 기존 CNN의 0.59, 0.52 보다 0.12, 0.25 만큼 높은 것을 볼 수 있다.

CIFAR-100 데이터 셋의 결과는 제안한 구조와 기존 CNN 구조의 정확도가 각각 0.36, 0.18, mAP가 0.36, 0.14로 모두 높지 않지만 그림에도 제안한 배치 정규화를 적용하면 정확도와 mAP 모두 높아지는 것을 볼 수 있다.

나머지 MNIST, SVHN, CIFAR-10의 데이터 셋 실험 결과에서도 제안한 구조는 기존 CNN 구조보다 정확도 곡선에서 전반적으로 더 높은 정확도를 보였고, 트레이닝 에러 곡선에서는 전반적으로 더 낮은 에러를 보이는 것을 볼 수 있었다. 이를 통하여 배치 정규화는 특정 데이터 셋이나 네트워크 구조가 아니라 여러 데이터 셋, 네트워크 구조에 적용하여도 좋은 효과를 보임을 볼 수 있다. 이는 PR-곡선을 보았을 때 제안한 구조의 결과가 정밀도의 하락이 더 낮은 것과, 정확도와 mAP 모두 최소 0.05에서 최대 0.44 높은 것을 통하여 알 수 있다.

5. 결 론

본 논문에서는 CNN에 의한 영상을 분류함에 있어 배치 정규화 방법이 추가된 CNN을 이용하여 영상 분류의 정확도를 높이는 방법을 제시하였다. 이를 위하여 본 논문에서는 기존의 뉴럴 네트워크에 배치 정규화 계층을 추가하는 방법을 제안하였다. 배치 정규화는 각 계층에 존재하는 편향을 줄이기 위해 고안된 방법으로, 각 배치의 평균과 분산을 계산하여 이동시키는 방법이다. 제시된 방법의 우수성을 입증하기 위하여 5개의 영상 데이터 집합을 이용한 영상분류 실험 결과 기존 CNN 보다 배치 정규화가 추가된 CNN이 영상 분류 시 데이터 셋, 네트워크



(그림 7) 기존 CNN 구조와 배치 정규화를 추가한 CNN 구조의 정확도, 트레이닝 에러, PR-Curve (Figure 7) Accuracy, training error and PR-Curve for CNN and Batch Normalized CNN

크 구조에 제한을 받지 않고 보다 빠른 학습 속도와 높은 분류 정확도, 안정성을 보임을 확인 할 수 있었다.

본 논문에서는 영상 분류의 정확도를 개선시키기 위하여 기존 CNN 구조에만 배치 정규화 기법을 적용시켰지만 CNN 구조 이외의 다른 네트워크 구조에서도 성능 향상을 기대할 수 있을 것이다.

참고문헌(Reference)

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, May 2017.

<https://doi.org/10.1145/3065386>

[2] Harada, Tatsuya, and Yasuo Kuniyoshi. "Graphical Gaussian vector for image categorization." *Advances in Neural Information Processing Systems*, pp. 1547-1555, 2012.

<https://dl.acm.org/citation.cfm?id=2999307>

[3] C. Szegedy et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

<https://doi.org/10.1109/cvpr.2012.6248018>

[4] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211-252, Apr. 2015.

- <https://doi.org/10.1007/s11263-015-0816-y>
- [5] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," in *Computer Vision - ECCV 2014*, Springer International Publishing, pp. 818 - 833, 2014.
https://doi.org/10.1007/978-3-319-10590-1_53
- [6] C. Szegedy et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
<https://doi.org/10.1109/cvpr.2015.7298594>
- [7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
<https://doi.org/10.1109/cvpr.2016.90>
- [8] G. Litjens et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, Dec. 2017.
<https://doi.org/10.1016/j.media.2017.07.005>
- [9] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
<https://doi.org/10.1109/cvpr.2014.220>
- [10] Nair, V., & Hinton, G. E. "Rectified linear units improve restricted boltzmann machines." In *Proceedings of the 27th international conference on machine learning*, 2010.
<https://dl.acm.org/citation.cfm?id=3104425>
- [11] SRIVASTAVA, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." *Journal of machine learning research*, vol.15, pp. 1929-1958, Jun 2014.
<https://dl.acm.org/citation.cfm?id=2670313>
- [12] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *International Conference on Machine Learning*, vol.37, pp. 448-456, Jul 2015.
<https://dl.acm.org/citation.cfm?id=3045167>
- [13] GODIL, Afzal A., et al. SHREC'13 Track: Large Scale Sketch-Based 3D Shape Retrieval, pp. 89-96, 2013.
<http://dl.acm.org/citation.cfm?id=2601321>
- [14] Netzer, Yuval, et al. "Reading digits in natural images with unsupervised feature learning." *NIPS workshop on deep learning and unsupervised feature learning*. Vol. 2011, No. 2, pp. 5, 2011.
http://dl.stanford.edu/housenumbers/nips2011_housenumbers.pdf
- [15] Krizhevsky, Alex, and Geoffrey Hinton. "Learning multiple layers of features from tiny images." 2009.
<https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>

◎ 저 자 소 개 ◎



지 명 근(Myunggeun Ji)

2017 B.S. in Computer Science, Kyonggi University, Suwon, Korea

2017~Present : M.S. Student in Computer Science, Kyonggi University, Suwon, Korea

Research Interests : Computer Vision, Augmented Reality

E-mail : jmg2968@gmail.com



전 준 철(Junchul Chun)

1984 B.S. in Computer Science, Chung-Ang University, Seoul, Korea

1986 M.S. in Computer Science(Software Engineering), Chung-Ang University, Seoul, Korea

1992 M.S. in Computer Science and Engineering (Computer Graphics), The Univ. of Connecticut, USA

1995 Ph.D. in Computer Science and Engineering (Computer Graphics), The Univ. of Connecticut, USA

2001.02~2002.02 Visiting Scholar, Michigan State Univ. Pattern Recognition and Image Processing Lab.

2009.02~2010.02 Visiting Scholar, Univ. of Colorado, Wellness Innovation and Interaction Lab.

Research Interests : Augmented Reality, Computer Vision, Human Computer Interaction

E-mail : jchun@kgu.ac.kr



김 남 기(Namgi Kim)

1997 B.S. in Computer Science and Engineering, So-Gang University, Seoul, Korea

2002 M.S. in School of Computing, KAIST, Daejeon, Korea

2005 M.S. in School of Computing, KAIST, Daejeon, Korea

Research Interests : Communication system, Network

E-mail : ngkim@kgu.ac.kr