

딥러닝 입문

(모델의 성능 끌어올리기)

Histroy

Date	Ver	내용
2020.12.28	v01	7.3 모델의 성능 끌어올리기

목차

- ▶ 01 배치 정규화(normalization)
- ▶ 02 깊이별 분리 합성곱(depthwise separable convolution)
- ▶ 03 하이퍼 파라미터 최적화
- ▶ 04 하이퍼 파라미터 최적화 과정
- ▶ 05 모델 앙상블

01 배치 정규화

- 머신 러닝 모델에 주입되는 샘플들을 균일하게 만드는 광범위한 방법
- 2015년 아이오페와 세게디가 제안한 층의 한 종류
- 케라스에서는 BatchNormalization 클래스로 사용 가능

01 배치 정규화

- Gradient Vanishing/Gradient Exploding이 일어나지 않도록 하는 아이디어 중의 하나.

(1) 불안정한 변화가 일어나는 원인 중의 하나가 Network의 각 층이나 Activation마다 입력의 분포가 달라지는 현상이 있다.

▶ 입력의 분포를 평균 0, 표준편차 1인 입력으로 정규화시키는 방법.
이는 Whitening의 방법으로 해결 가능.

01 배치 정규화 - Whitening

- Whitening은 기본적으로 들어오는 입력의 특징(feature)을
 - (1) 상관성 없게(uncorrelated)게 만들어주고,
 - (2) 각각의 분산을 1로 만들어주는 작업.

▶ 이 방법은 계산 량이 많고, 일부 파라미터의 영향이 무시된다.

▶ [해결] Whitening를 해결하기 위한 트릭이 바로 **배치 정규화**

01 배치 정규화(Batch Normalization)

- 각 레이어마다 정규화하는 레이어를 두어, 변형된 분포가 나오지 않도록 조절한다.
- 미니 배치의 평균과 분산을 이용하여 정규화를 수행, scale 및 shift를 감마(γ)값, 베타(β) 값을 통해 실행.

감마와 베타 값은 학습을 통해 구할 수 있음.

- 배치 정규화는 단순히 평균과 분산을 구하는 것 뿐만 아니라 감마(scale), 베타(shift)를 통한 변환을 통해 비선형 성질을 유지하면서 학습이 될 수 있게 해준다.

02 깊이별 분리 합성곱

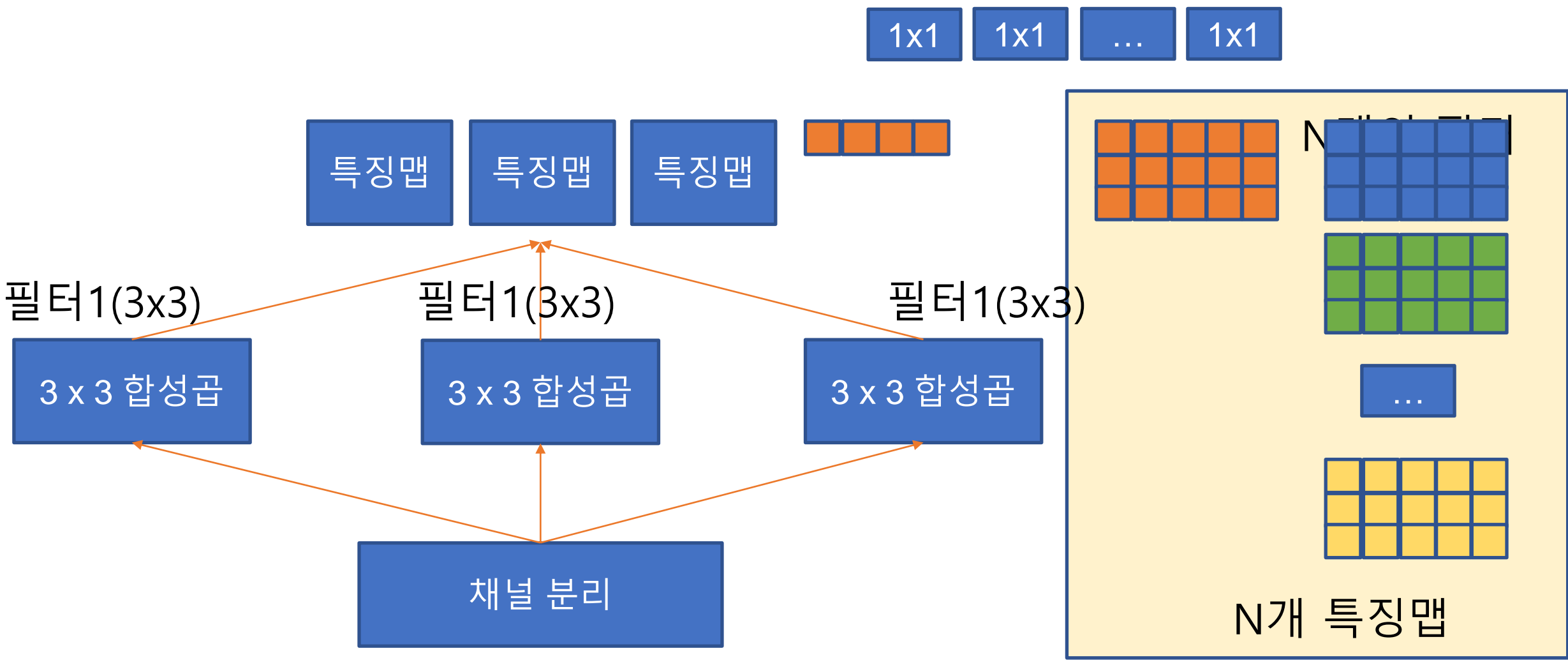
01. 1차적으로 채널별로 합성곱을 수행한다.

02. 각각의 결과물을 만든다.

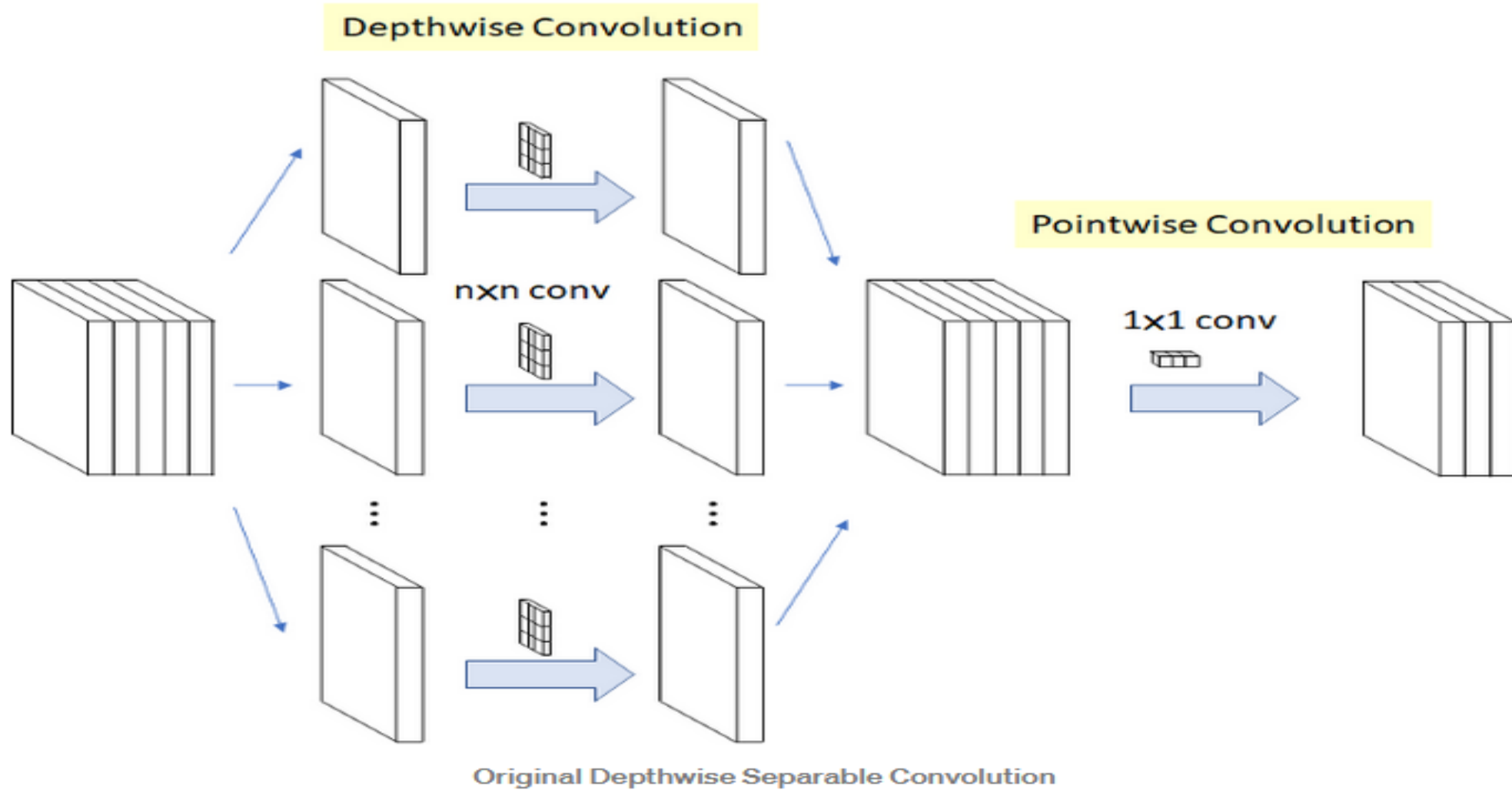
03. 1×1 필터를 이용하여 각각의 특징맵을 만든다. (N개의 필터)

* 케라스에서 SeparableConv2D를 이용하여 수행할 수 있다.

02 깊이별 분리 합성곱



02 깊이별 분리 합성곱



(참조) <https://towardsdatascience.com/review-xception-with-depthwise-separable-convolution-better-than-inception-v3-image-dc967dd42568>

03 하이퍼 파라미터 최적화

- 얼마나 많은 유닛이나 필터를 두어야 할까?
- relu 활성화 함수를 사용해야 할까?
- 어떤 층 뒤에 BatchNormalization을 사용해야 할까?
- Dropout은 얼마나 해야 할까?
- 얼마나 층을 쌓아야 할까?

04 하이퍼 파라미터 최적화 과정

- (1) 일련의 하이퍼 파라미터를 (자동으로) 선택합니다.
- (2) 선택된 하이퍼 파라미터로 모델을 만든다.
- (3) 훈련 데이터로 학습하고 검증 데이터에서 최종 성능을 측정합니다.
- (4) 다음으로 시도할 하이퍼 파라미터를 (자동으로) 선택합니다.
- (5) 이 과정을 반복합니다.
- (6) 마지막으로 테스트 데이터에서 성능을 측정합니다.

* Hyperopt와 Hyperas 라이브러리를 이용하여 하이퍼 파라미터 최적화를 수행할 수 있다.

05 모델 앙상블

- 앙상블은 여러 개 다른 모델의 예측을 합쳐서 더 좋은 예측을 만든다.
- 아주 뛰어난 단일 모델보다 여러 개를 합친 앙상블이 성능이 좋습니다.
- 장님과 코끼리에 관한 우화를 생각해 보자.
 - => 서로 코끼리의 다른 부분을 만지고 각자의 관점으로 이해한 정답을 이야기합니다.
 - 이들의 관점을 모으면 훨씬 더 일반화된 코끼리를 만들 수 있습니다.
- 여러 모델의 예측을 합치는 가장 쉬운 방법은 예측의 평균을 내는 것.

05 모델 앙상블

- 앙상블의 핵심은 **다양한 모델을 만드는 것**이 중요합니다.
- 최대한 다르면서 좋은 모델을 앙상블해야 합니다.
- 실전에서 잘 동작하는 한 가지 방법은 트리 기반 모델(랜덤 포레스트나 그래디언트 부스팅 트리)나 심층 신경망을 앙상블하는 것.
- 딥러닝과 얇은 모델을 섞은 넓고 깊은 모델 사용. 심층 신경망과 선형 모델을 함께 훈련.