

Xception_net

- Inception 모듈에 대한 고찰로 탄생한 모델이다.
- Cross-channel correlations 와 Spatial correlations 를 독립적으로 계산하기 위해 고안된 모델이다.
- 이를 위해 새로운 Inception 모듈을 제안한다.

이미지의 여러 채널 간의 상관 관계를 의미한다.

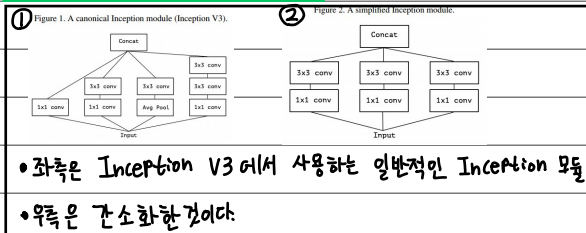
• 한마디로, Inception 모듈은 원래 적은 파라미터로 충분성능 나오는데?

아니야 Xception 모듈처럼 수정하면 더 좋고 작아져

• 이미지의 픽셀 간의 공간적 상관 관계

• 공간적 상관 관계를 통해 모델은 객체의 형태나 텍스처와 같은 공간적 특성을 인식한다.

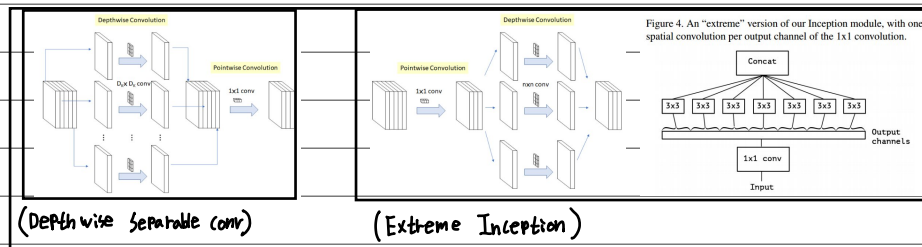
Inception 모듈에 대한 저자의 해석



- 좌측은 Inception V3 에서 사용하는 일반적인 Inception 모듈
- 우측은 간소화한 것이다

- 저자는 (1x1)conv는 Cross-channel Correlation 을 계산
- (3x3)conv는 Spatial Correlation 을 수행하는 것이다.
- 저자는 위에처럼 Cross-channel 과 Spatial 을 잘 분해해서 좋은 성능이 나온다고 한다.

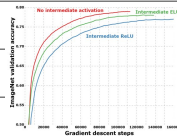
Extreme Inception (Modified Depthwise separable convolution)



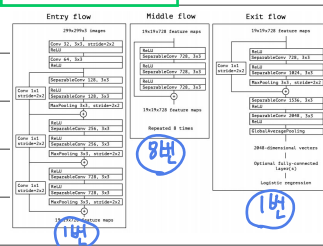
(Depthwise separable conv)

(Extreme Inception)

- Depthwise separable convolution 을 수정해서 inception 모듈 대신에 사용한다.
- Pointwise conv 이후에 depthwise conv 를 사용한다.
- (1x1)conv 를 먼저 하게 되면 채널 수는 n 개의 segment 로 나뉜다.
- 하여 퍼 파라미터
- 아래처럼 Inception 모듈은 (1x1)conv 이후에 ReLU 비선형 함수 수행 하지만 Xception 모듈은 비선형 함수를 사용하지 않는다.
- 신경망이 깊을 때는 비선형 함수가 도움이 되지만 Xception 모듈같은 1개의 채널을 사용하는 depthwise separable 은 비선형 함수가 성능을 해친다. 저자는 정보손실 때문에 그럴꺼라 생각한다.



이펙티브넷 실험



- 비슷한 수의 파라미터를 가지는 Inception V3 에 비해 정확도 향상

Table 1. Classification performance comparison on ImageNet (single crop, single model). VGG-16 and ResNet-152 numbers are only included as a reminder. The version of Inception V3 being benchmarked does not include the auxiliary tower.

	Top-1 accuracy	Top-5 accuracy
VGG-16	0.715	0.901
ResNet-152	0.770	0.933
Inception V3	0.782	0.941
Xception	0.790	0.945

Inception 이전은
제일 좋다