

# [논문리뷰] Xception [2017]

☀ 해당 논문은 Convolution과 depthwise separable 사이의 Conv2d 연산을 수행하는, 중간 단계로 Convolution 신경망의 Inception Module에 관하여 고찰합니다.

논문에서 설명할 **Xception Model**은 InceptionV3의 성능을 증가합니다.

**Xception**은 InceptionV3보다 parameter의 개수가 증가하는 것이 아니라, 파라미터를 효율적으로 쌓아 사용합니다.

## Introduction

### The Inception hypothesis

☀ 해당 논문에서는 **XceptionNet**을 설명하기 전, **InceptionV3**를 되짚어봅니다.

Figure 1. A canonical Inception module (Inception V3).

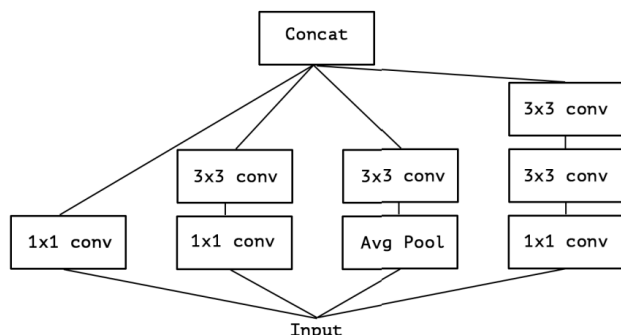
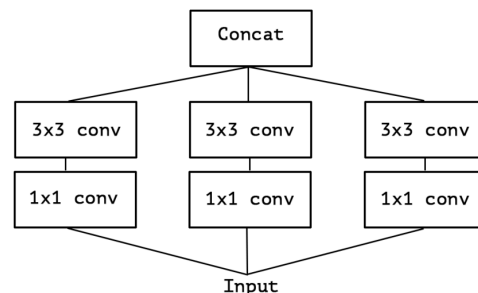


Figure 2. A simplified Inception module.



**Inception Module**은 합성곱과는 비슷하더라도, 경험적으로는 매개변수를 덜 사용하면서 더 풍부한 representation을 학습하는 것이 가능합니다. 그렇다면, **어떻게 작동하고, 합성곱과 어떻게 다른 것인가?**

**Inception** 이후에 어떤 설계 전략이 있는가?

한 가지의 합성곱(3x3)을 사용하며, **평균 풀링을 포함하지 않는다**. **Inception Module**은 **크기가 큰 1x1 합성곱 하나와 출력 채널에서 겹치지 않는 부분에 대해 연산을 진행할 공간 합성곱으로 재구성**하는 것이 가능합니다.

채널 간 상관관계와 공간 간 상관관계가 완벽하게 분리되어 매핑될 수 있다는 인셉션의 가설보다 더 강력한 가설을 합리적으로 세울 수 있다고 생각되어 만들어진 **단순한 Inception Module이 좌측 Module**입니다.



정리하자면, 공간 및 채널 간의 상관관계가 Pooling을 하는 도중 작아지기 때문에, 상관관계를 최대화하기 위해서, 채널과 공간을 완벽하게 분리하여 Mapping한 후, 합치는 것이 상관관계가 낮아지는 것을 지킬 수 있다고 봅니다.

## The continuum between convolutions and separable convolutions

Figure 3. A strictly equivalent reformulation of the simplified Inception module.

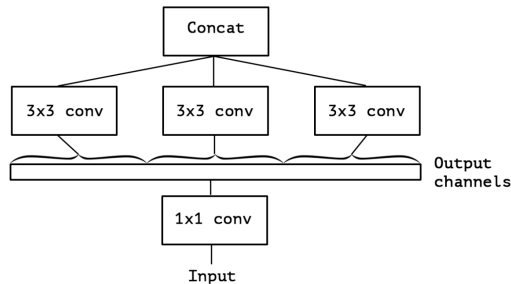
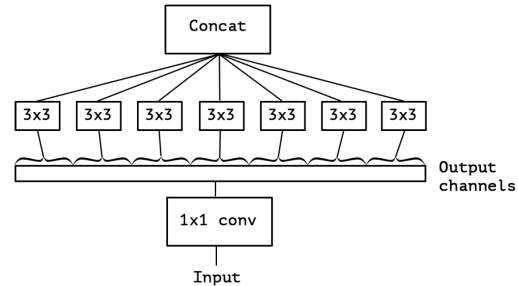


Figure 4. An “extreme” version of our Inception module, with one spatial convolution per output channel of the 1x1 convolution.



Stronger hypothesis에 기반한 “Extreme” version의 Inception Module은 cross-channel correlations를 Mapping하기 위해 1x1 Convolution을 사용하며, 분리해서 mapping합니다. 모든 Output Channels의 공간 상관관계를 별도로 Mapping합니다.

Figure4의 방식은 Depthwise Separable Convolution이라고 부릅니다.

두 가지 미묘한 차이

Extreame Inception Module 과 Depthwise seprable Convolution에는 차이가 존재합니다.

- The order of the operations
  - Xception : Kernel Size (1x1 → 3x3) 을 적용합니다.
  - Inception : Kernel Size (3x3 → 1x1) 을 적용합니다.
- The presence or absence of non-linearity after the operation (ReLU 활성화 함수 부재)
  - Xception : 1x1 → ReLU → 3x3
  - Depthwise : 1x1 → Spatial Convolution 사이에는 ReLU가 들어가지 않습니다.

## The Xception Architecture

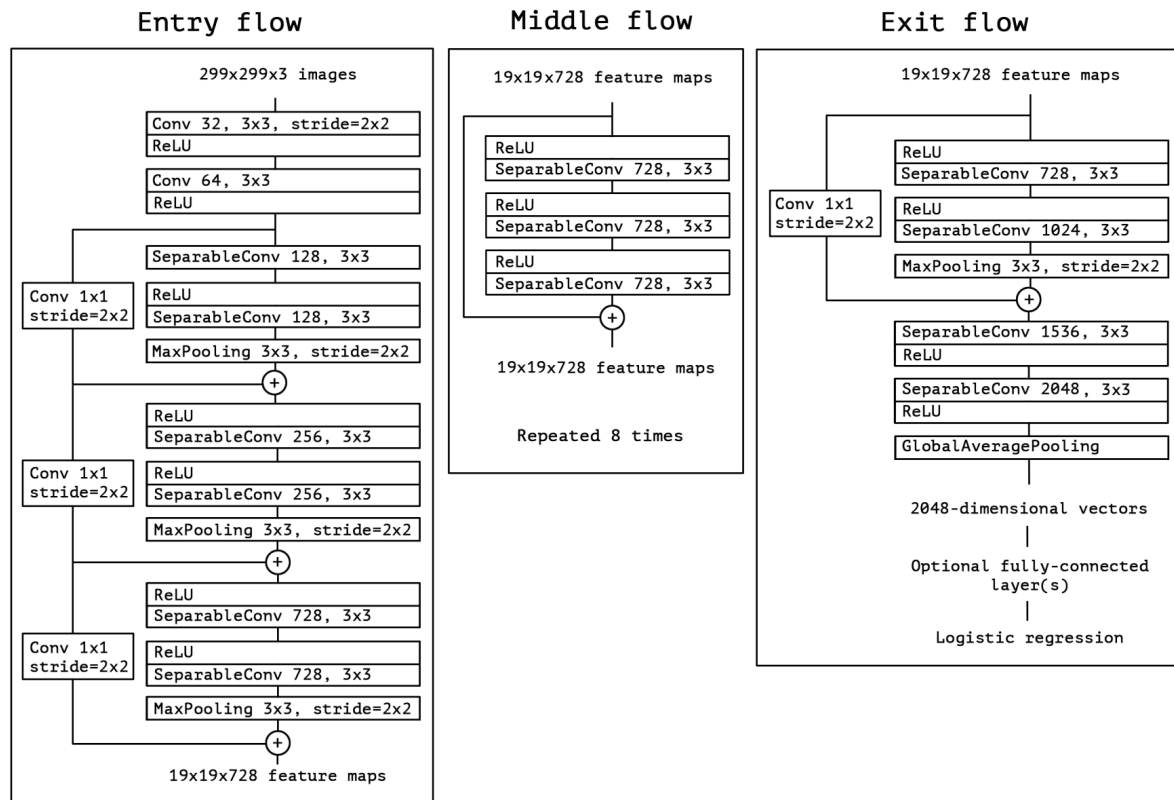


Depthwise Separable Convolution layer에 기반을 둔 Convolutional Architecture를 제안합니다.

공간(Spatial-Correlations)과 채널(Cross-channels)의 Mapping 관계를 완벽히 분리합니다.

이러한 아이디어를 제안한 모델들 Xception이라고 부르기로 하였습니다. [Extream Inception]

Figure 5. The Xception architecture: the data first goes through the entry flow, then through the middle flow which is repeated eight times, and finally through the exit flow. Note that all Convolution and SeparableConvolution layers are followed by batch normalization [7] (not included in the diagram). All SeparableConvolution layers use a depth multiplier of 1 (no depth expansion).



Xception은 Depthwise Separable Convolution을 residual Connection과 함께 linear Stack 으로 쌓은 것으로 볼 수 있습니다. Residual을 더하기 전의 연산 모듈들을 nn.Sequential을 활용하여 Block으로 생성한 것입니다.

## Conclusion

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

Table 2. Classification performance comparison on JFT (single crop, single model).

	FastEval14k MAP@100
Inception V3 - no FC layers	6.36
Xception - no FC layers	6.70
Inception V3 with FC layers	6.50
Xception with FC layers	6.78

Xception은 Convolution과 Depthwise가 어떻게 분리되어 연산하는지 보여주었습니다. Convolution은 discrete spectrum의 양 극단에 놓여 있으며, Inception Module은 중간지점에 위치합니다.

Xception은 Inception V3와 상당한 parameter를 가지고 있으며, 약간의 더 좋은 성과를 Classification에서 보여줍니다. 연구진들은 Depthwise Separable이 CNN의 Cornerstone 기대하고 있습니다.

아직은 regular convloution layer이 더 쉽지만, 추후 Inception Module과 유사한 목적을 가진 Block들이 발전될 것이라고 전망합니다.

## Reference


### Xception: Deep Learning with Depthwise Separable Convolutions

We present an interpretation of Inception modules in convolutional neural networks as being an intermediate step in-between regular convolution and the depthwise separable convolution operation (a depthwise convolution followed by a pointwise convolution). In this light, a depthwise separable convolution can be understood as an  
<https://arxiv.org/abs/1610.02357>

### [딥러닝 논문 리뷰] Xception: Deep Learning with Depthwise Separable Convolutions (feat.Pytorch)(1)

안녕하세요. 저번 Depthwise Separable Convolution 기법에 대해 글을 올렸습니다. 오늘은 이 기법을 사용한 Xception 논문에 대해 리뷰하도록 하겠습니다.

[https://openaccess.thecvf.com/content\\_cvpr\\_2017/html/Chollet\\_Xception\\_Deep\\_Learning\\_CVPR\\_2017\\_paper.ht](https://openaccess.thecvf.com/content_cvpr_2017/html/Chollet_Xception_Deep_Learning_CVPR_2017_paper.html)

 <https://coding-yoon.tistory.com/78>