MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
A Constituent Institution of Manipal University

Department of Computer Science and Engineering

Subject: Deep Learning and Applications

Subject Code: CSE – 5214

Presentation Topic: Understanding Xception Net

Presented by:

230913003	Pasupuleti Rohithsaidatta
230913004	Amit Kumar
230913025	Aston Dsouza
230913021	Prajwal Purushotham Varekar
230913017	Abhay S Gad
230913026	Akshay C

Under the guidance of:

Dr. Siddalinga Swamy P. C.

Introduction

Xception is an advanced version of the Inception model, specialized in image recognition. Unlike its predecessors, Xception achieves superior results through more effective parameter utilization rather than merely increasing their quantity.

Previous Successes
Inception models like V1, V2, and V3 laid the groundwork for understanding images before the emergence of Xception.

Effective Parameter Usage
Xception's superiority lies in its efficient parameter usage, distinguishing it from previous models.

Performance
Xception surpasses Inception V3 on ImageNet and demonstrates remarkable progress on larger datasets.

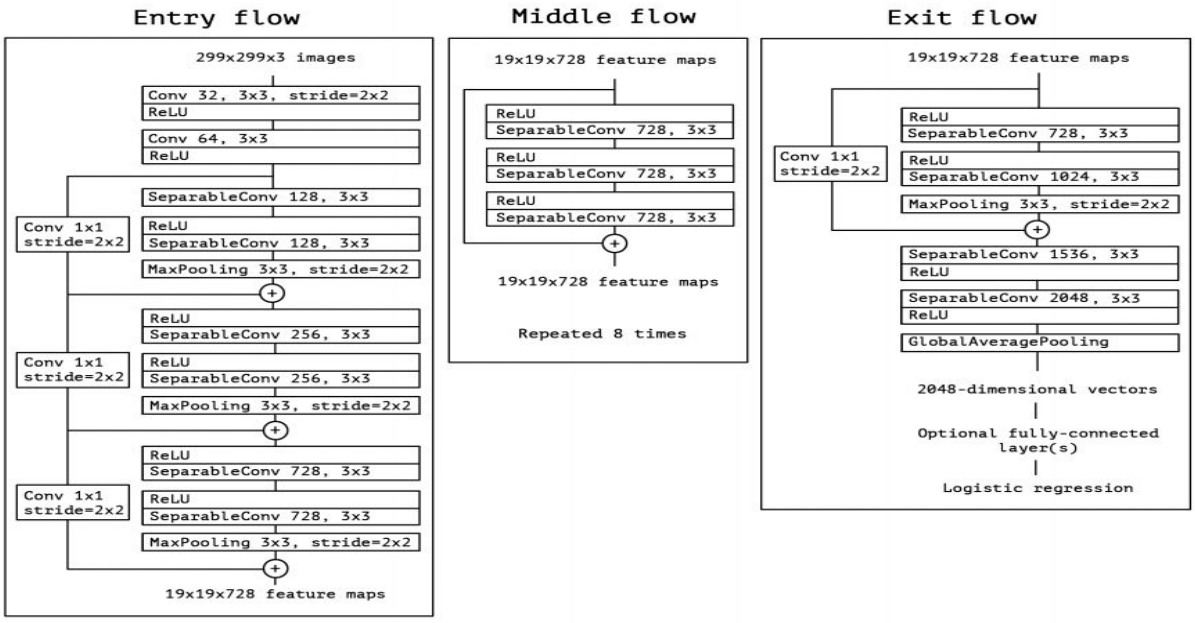
Innovative Approaches
Xception extends the Inception concept by treating color/features and shape separately, utilizing separable convolution to enhance learning efficiency.

Xception vs. Traditional Architectures

Xception surpasses traditional architectures like VGGNet and ResNet in complexity and depth. However, this complexity poses challenges:

- Training Difficulty: Xception's complexity makes it harder to train, especially in resource-limited environments.
- Computational Power Requirement: Xception demands significant computing power, making it unsuitable for devices with limited resources.
- Training Time: Due to its complexity, training Xception models is time-consuming, which may not be suitable for real-time tasks.

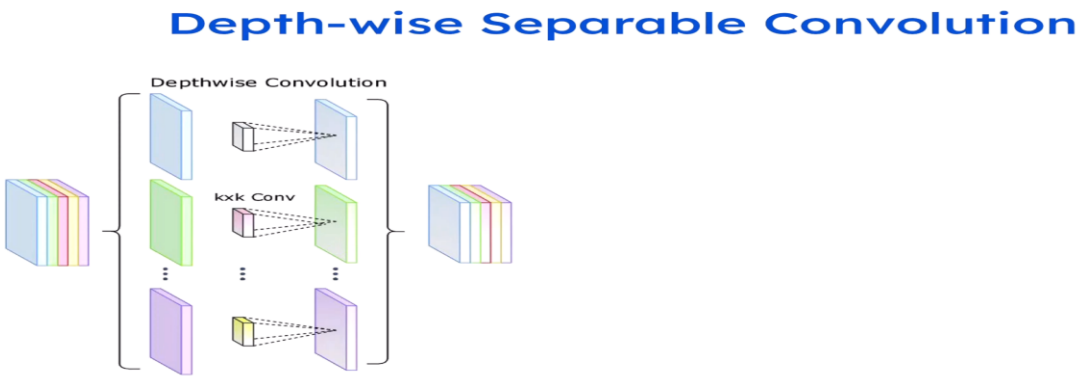
Xception Net Architecture



The Xception architecture comprises three main components: the entry flow, middle flow (repeated eight times), and exit flow. Each convolutional layer is followed by batch normalization.

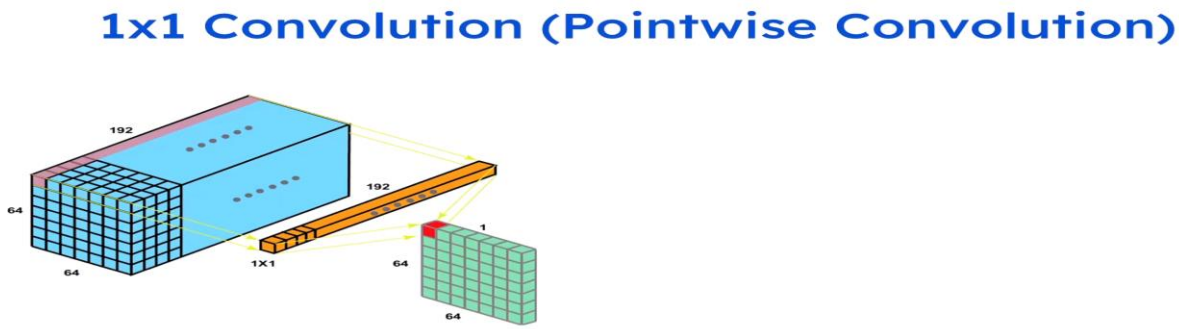
Standard Convolution vs Depthwise Separable Convolution

- Standard Convolution: Involves sliding a filter/kernel over the input image, leading to high computational cost.
- Depthwise Separable Convolution: Decomposes standard convolution into depthwise and pointwise convolutions, reducing computational cost significantly.



Xception Net Working (Depthwise Separable Convolution + Pointwise Convolution)

- Pointwise Convolution: Applied after depthwise convolution, involving a 1x1 convolutional kernel.
- Combining Output Channels: Combines output channels produced by depthwise convolution.



Applications

Advantages:

- Reduced Parameters:
  Xception uses fewer parameters, making it faster and more memory-efficient.
- Enhanced Performance:
  Despite its simplicity, Xception performs exceptionally well on image tasks.
- Scalability:
  Xception can adapt to different hardware setups, making it versatile.

Disadvantages:

- Complexity in Training:
  Training Xception can be tricky, especially for beginners.
- High Computational Requirement:
  It needs a lot of computing power, limiting its use on less powerful devices.
- Long Training Time:
  Training Xception models takes a long time due to their complexity.

Comparison of ImageNet: Xception Model with other Models

Xception surpasses traditional models like VGGNet and ResNet in both efficiency and performance on ImageNet benchmarks, owing to its reduced parameter count and enhanced computational capabilities. Its streamlined architecture maintains top-tier accuracy while offering scalability across diverse hardware platforms, ensuring versatility in deployment.

VGGNet – 1st Runner Up in ILSVRC 2014

ResNet – Winner in ILSVRC 2015

Inception-v3 – 1st Runner Up in ILSVRC 2015

	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**