TASK - 4

1. **What is Transfer Learning ?**

Transfer Learning is a process of re-using a pre-trained model on a particular task which will be used again for training a similar task.

Ex: In Deep learning, model trained to identify vehicles can be used for training and identifying cars.

1. **List the various Transfer Learning Architectures.**

Various architectures used for Transfer learning are:

* LeNet
* AlexNet
* VGG
* GoogleNet
* ResNet

1. **Give Description of each architecture.  
   (No. of layers, type of layers, Data used for training, required Input Image shape, Output shape(Number of Classes), Loss function, Disadvantages, and applications.**

Above Architectures are explained briefly below:

**LeNet**

LeCun et al in 1998 submitted a paper called “[Gradient-Based Learning Applied to Document Recognition](http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf" \t "https://www.pyimagesearch.com/2016/08/01/lenet-convolutional-neural-network-in-python/_blank).” which primarily concentrated on OCR(Optical character recognition) in documents. This paper led to the implementation of first CNN(Convolutional Neural Network ). It has an **error rate of 28.2 %**.

It was a simple architecture which used less memory that can run on the CPU without the need of GPU leading to steppingstone for Neural networks.

**LeNet**  has 7 layers out of which 3- Convolutional layer, 2-pooling layer, 1 - fully connected layer and 1-output layer. It used Tanh as activation function while the input size of image was 32x32 (grayscale) and the output was 0-10 values. LeNet uses *MNIST dataset* which thousands of handwritten digits from 0-9.

Demerits: LeNet led to Overfitting in some cases.

Application: It was used in US postal code recognition system which uses hand written digits image as input.

**AlexNet**

AlexNet was the winner of 2012 ImageNet LSVRC-2012 competition which has error rate of **15.3 %.** AlexNet improved its architecture by using ReLu instead of Tanh like used by LeNet while using dropout instead of regularisation for handling overfitting issue .

*AlexNet* used 8 layers with an input image size of 256x256 RGB with ReLu as activation function. It has an output of 1 vector(predicting cat of dog). Out of 8 layers, 5- Convolutional layers while 3- fully connected layers. It used ImageNet as Dataset which consists of more than 15 million labeled images and 22 thousand classes.

Demerits: It had a drawback of non standard filter size leading to modifications.

Application: Since ReLu is used which led to faster image recognition and classification.

**VGG**

VGG stands for Vector Geometry Group which was proposed by *[Karen Simonyan](http://www.robots.ox.ac.uk/~karen/" \o "www.robots.ox.ac.uk" \t "https://www.quora.com/_blank)* and *[Andrew Zisserman](https://en.wikipedia.org/wiki/Andrew_Zisserman" \o "en.wikipedia.org" \t "https://www.quora.com/_blank)*of [Oxford Robotics Institute](https://en.wikipedia.org/wiki/Oxford_Robotics_Institute" \o "en.wikipedia.org" \t "https://www.quora.com/_blank) in the the year 2014. It has error rate of **7.3 %** and uses the same *ImageNet dataset.*

VGG-16 has a total of 16 layers with an input image size of 224x224x3. It uses ReLu as activation function. Among 16 layers, 13-Convolutional Layers, 5-max pooling layers and 3-Dense layers.

Demerits: VGG uses high computational power.

Application: It is used for identifying objects from whole image.

**GoogleNet**

GoogleNet was from Google whose main idea was to run the model through a smart-phone. It mainly introduced the Inception Model and was the winner of ILSVRC 2014 with a top 5 **error rate of 6.7%**. It uses the same ImageNet dataset but using the structured layers lead to improved performance.

GoogleNet used 22 layers with input image size of 224x224x3 with ReLu as activation function.

Demerit: Used very high computational power.

Application: Faster Image recognition/classification which can be implemented in Computer Vision like Automation of driving.

**ResNet**

ResNet was introduced from Microsoft in 2015 which became popular since it used high number of layers which was a truly deep network architecture. ResNet won ILSVRC 2015 with an error rate of 3.6% while human error was around 5-10% which was tremendous achievement for an artificial model.

ResNet used 152 layers with input image size of 224x224x3 with ReLu as activation function. It uses residual previous values to understand and build model while skip connections was used to improve accuracy and performance.

Demerits: High demand on computational power and time consuming for running the model.