Scene classification using ML algorithms

Dataset :https://www.kaggle.com/puneet6060/intel-image-classification

SVM(Support vector machine

Deep learning

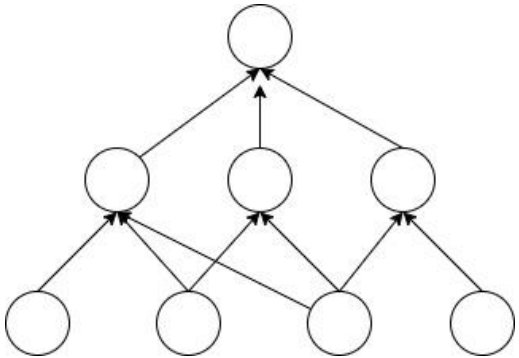
- **Deep learning** is part of a broader family of **machine learning** methods based on artificial neural networks
- Deep learning achieves recognition accuracy at higher levels than ever before
- Deep learning requires large amounts of **labeled data**.
- Deep learning requires substantial **computing power**. High-performance GPUs have a parallel architecture that is efficient for deep learning. When combined with clusters or cloud computing, this enables development teams to reduce training time for a deep learning network from weeks to hours or less.
- Most deep learning methods use **neural network** architectures, which is why deep learning models are often referred to as **deep neural networks**.
- The term “deep” usually refers to the number of hidden layers in the neural network. Traditional neural networks only contain 2-3 hidden layers, while deep networks can have as many as 150.
- Deep learning models are trained by using large sets of labeled data and neural network architectures that learn features directly from the data without the need for manual feature extraction.
- CNN is the most popular type of deep learning.

What's the Difference Between Machine Learning and Deep Learning?

- A machine learning workflow starts with relevant features being manually extracted from images. The features are then used to create a model that categorizes the objects in the image.
- Deep learning performs “end-to-end learning” – where a network is given raw data and a task to perform, such as classification, and it learns how to do this automatically.
-

CNN

They are special kind of multi-layer neural networks, designed to recognize visual patterns directly from pixel images with minimal preprocessing



CNN Architectures

- LeNet:

  7-level convolutional network ,that classifies digits, was applied by several banks to recognise hand-written numbers on checks (cheques) digitized in 32x32 pixel greyscale inputimages. The ability to process higher resolution images requires larger and more convolutional layers, so this technique is constrained by the availability of computing resources.

- AlexNet:

  The network had a very similar architecture as LeNet ,stcked convolution network,AlexNet was trained for 6 days simultaneously on two Nvidia Geforce GTX 580 GPUs which is the reason for why their network is split into two pipelines

- GoogLeNet:

  22 layers deep network, was used to assess its quality in the context of object detection and classification.

- ResNet:

  Residual Neural Network

How to Create and Train Deep Learning Models

1. Training from scratch:
   To train a deep network from scratch, you gather a very large labeled data set and design a network architecture that will learn the features and model.
   This is least used as these models take weeks or days of time to train.

2. Transfer learning:
   Process that involves fine tuning a pretrained model such as AlexNet,GoogleNet etc.
   For small datasets ,we train only the final layers i.e new layers
   For bigdata retain the whole network

3. Feature extraction:

   A slightly less common, more specialized approach to deep learning is to use the network as a **feature extractor**. Since all the layers are tasked with learning certain features from images, we can pull these features out of the network at any time during the training process. These features can then be used as input to a machine learning model such as support vector machines (SVM).

Channels

- How many numbers are used to specify the color of each pixel is the number of **channels** each pixel has.
- A monochrome image that has one number per pixel has one channel.
- A more typical image that has three (R, G, B) numbers per pixel has three channels. Such images are called **RGB** images.

Activation functions

Relu Activation Function (rectified linear unit):

Relu F(x) = max (x, 0), is mostly used deep learning activation function, for hidden layers. A rectified linear unit has output ‘0’ if the input is less than ‘0’ and raw output ‘otherwise’.

. **Tanh activation function :**

Tanh function [tanh(x) = (ex - e -x) / (ex + e-x)] produces output in range of -1 to +1. It is continuous function, which produces output for every ‘x’ value.

Classifiers

Sigmoid classifier:

Sigmoid classifier takes any range of real number and returns the output value which falls in the range of „0" to „1". It produces the curve in 'S' shape.

Softmax classifier:

The softmax classifier squashes the outputs of each unit to be between 0 and 1, just like a sigmoid classifier. But it also divides each output such that the total sum of the outputs is equal to 1. The output of the softmax classifier is equivalent to a categorical probability distribution, it tells you the probability that any of the classes are true.