A Brief Overview Of ZFNet Architecture.

Azeem - I · Follow
4 min read · Nov 9, 2023

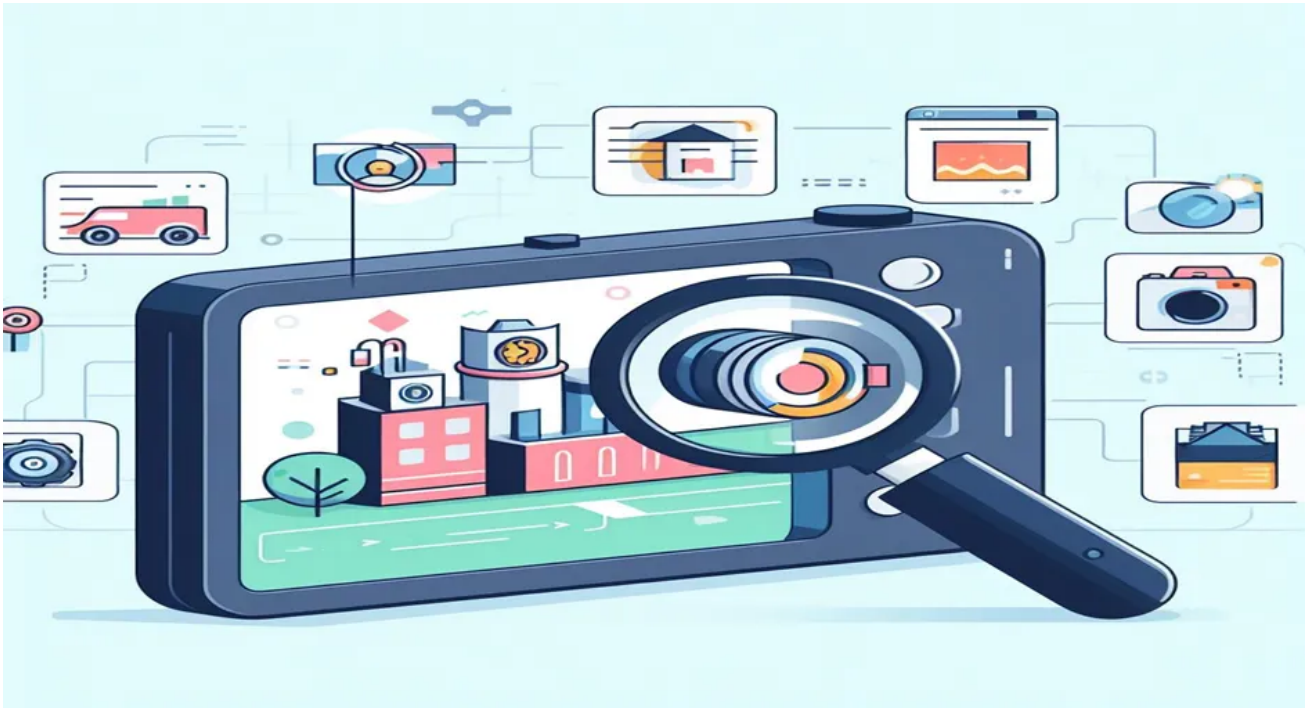


image generated with Microsoft Bing

To understand ZFNet First we have to take a look into what is Convolutional neural network and Imagenet.
A **convolutional neural network** is a special kind of multi neural network designed to extract visual patterns from a given image. Since images are formed of pixels so a convolutional neural network tries to capture the important pixels values by a process known as convolution.

ImageNet Project is a large database designed for use in visual object recognition software research.
The ImageNet Project runs an annual competition ImageNet Large Scale Visual

Recognition Challenge (ILSVRC) where software programs compete to correctly classify the objects and scenes.

It is imperative to know about **AlexNet** before coming to ZFNet. AlexNet was primarily designed by Alex Krishevsky. It was published with Ilya Sutskever and Geoffry Hinton.It won first place in ImageNet large Scale Visual recognition Challenge in 2012 by achieving error of 15.3%.This was 10.8% lower than that of runner up. AlexNet was considered as a massive jump in the accuracy of neural networks.

ZFNet

Rob Fergus and Matthew D.Zeiler introduced ZFNet. ZFNet is named after their surname Zeiler and Fergus. ZFNet was a slight improvement over AlexNet .The 2013 ILSVRC was won by ZFNet.It actually visualized how each layer of AlexNet performs and what parameters can be tuned to achieve greater accuracy.

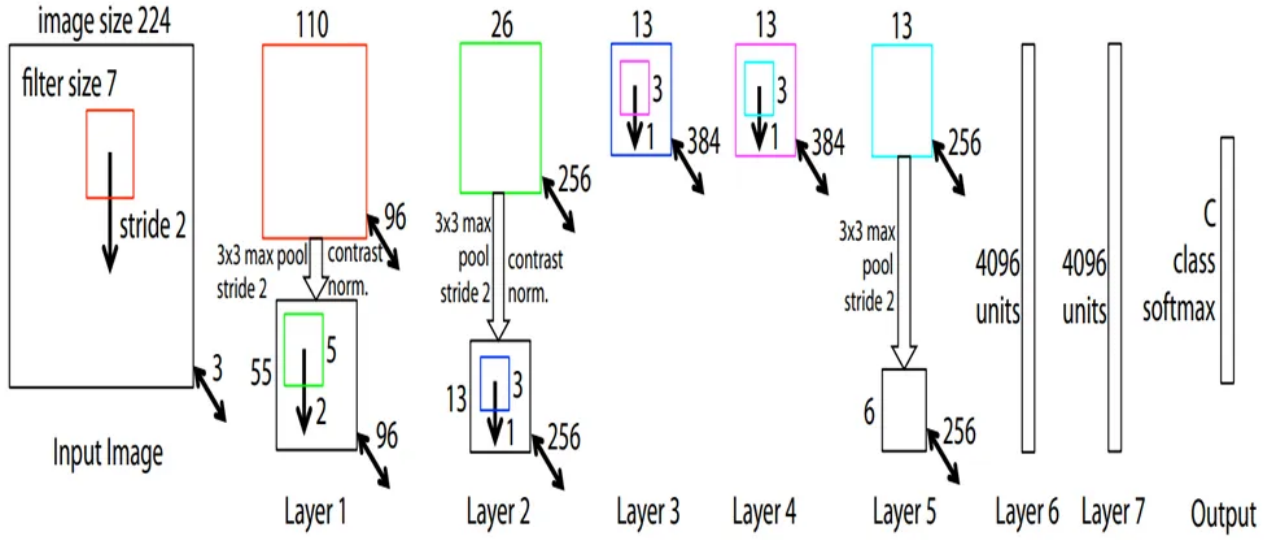


image from original paper-https://arxiv.org/pdf/1311.2901.pdf

Some Key Features of ZFNet architecture

· Convolutional layers:

In these layers convolutional filters are applied to extract important features,ZFNet consists of multiple convolutional layers to extract important features.

· MaxPooling Layers:

MaxPooling Layers are used to downsample the spatial dimensions of feature map in.It consist of aggregation function known as maxima.

· **Rectified Linear Unit:**

Relu is used after each convolution layer to introduce non linearity into the model which is crucial for learning complex patterns. It rectifies the feature map ensuring the feature maps are always positive.

· **Fully Connected Layers:**

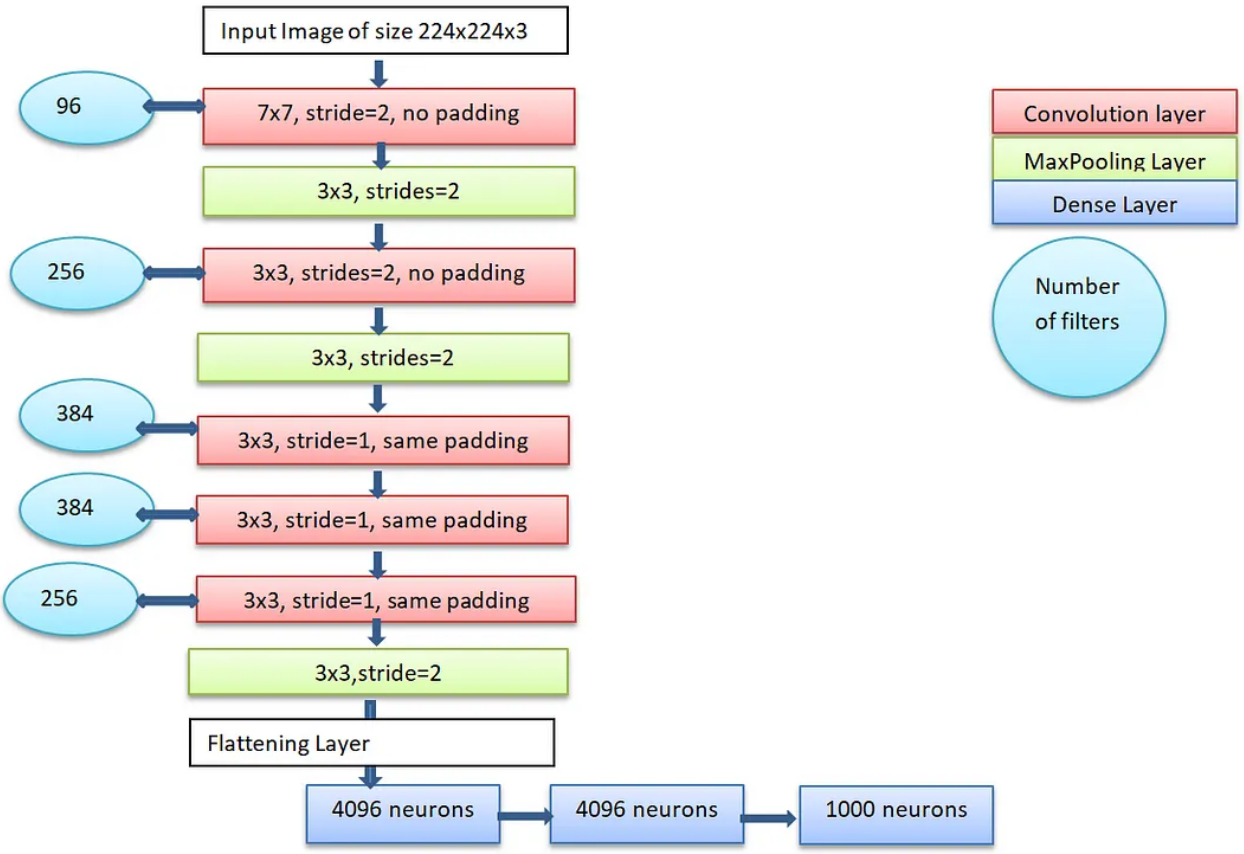
In the latter part of ZFNet architecture fully connected dense layers are used to extract patterns from features .The activation function used in the neurons is relu.

· **SoftMax Activation:**

SoftMax activation is used in the last layer to obtain the probabilities of the image belonging to the 1000 classes.

· **Deconvolution Layers:**ZFNet introduced a visualization technique involving deconvolutional layers(Transposed Layers) .These layers provide insights into what network has learned by projecting feature activations back into input pixel space.

Architecture:



ZFNet Architecture (image by me)

Input

· The input image is of size 224x224x3.

First Layer

· In the first layer 96 filters of size 7x7 and stride of 2 are used to convolve followed by relu activation.
The output feature map is then passed through Max Pooling Layer with pool kernel of 3x3 and stride of 2 .Then the features are contrast normalized.

Second layer

· In the second layer 256 filters are applied of size 3x3 with stride of 2. Again the obtained feature map is passed through MaxPooling layer with pooling kernel of 3x3 with stride of 2.After that features are contrast normalized.

Third layer and Fourth Layer

· The third and fourth layers are identical with 384 kernels of size 3x3 and padding is kept as same and stride is set to 1.

Fifth Layer

· In the fifth layer 256 filters of size 3x3 are applied with stride 1. After then the MaxPooling kernel of size 3x3 is applied with stride of 2 .Then the features are contrast normalized.

Sixth Layer and Seventh Layer

· The sixth and seventh layers are fully connected dense layers with 4096 neurons each.

Eighth Layer

· The last layer is dense layer with 1000 neurons(number of classes).

Differences Between AlexNet and ZFNet

Architecture

· AlexNet consists of eight layers, five convolutional layers followed by three fully connected layers. ZFNet retained basic architecture of AlexNet but made some a[illegible] m[illegible] in[illegible] la[illegible]

Filters

· AlexNet used 11x11,5x5 and 3x3 filter sizes while ZFNet used 7x7 filter size in the first layer only and 3x3 in the latter layers only.

· There is stride of 4 in the first layer of AlexNet while in ZFNet there is stride of 2 used.

Normalization

· AlexNet used Local Response Normalization while ZFNet used Local Contrast Normalization