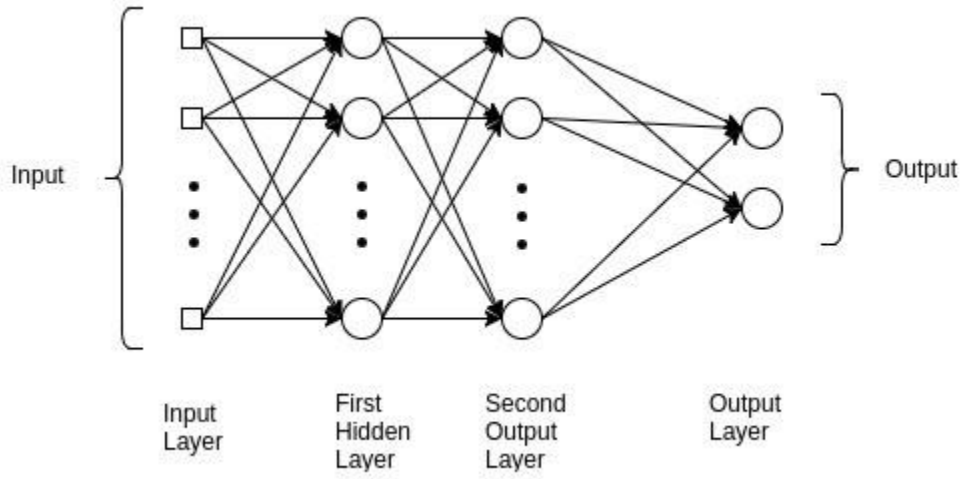
CSL 7020 Assignment 5

Ashutosh Kumar Jatav (B16CS004)

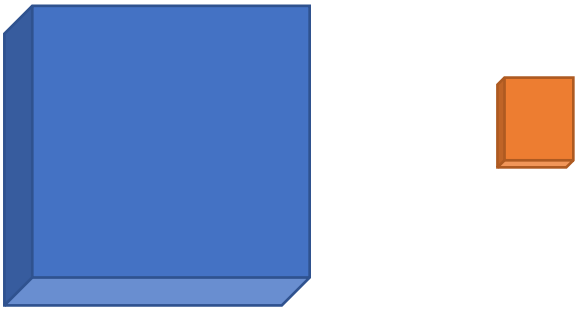
Multilayer Perception:



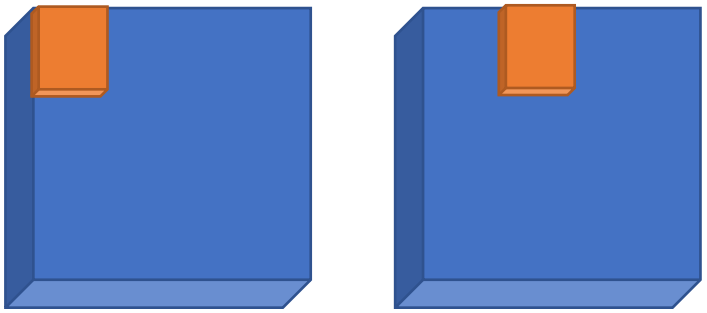
- Has fully connected layers
- Total number of parameters grow very high
- In perceptron algorithm we multiply weights and add bias
- weight = weight + lr * (expected – predicted)*x
- Lets take example of 3 layer network, first layer Is input layer, second hidden layer and last output layer, every neural of 1 layer is connected to every neuron of next layer
- It discards spatial information
- Inefficient because there is redundancy in such high dimensional parameters

Convolutional Neural Network:

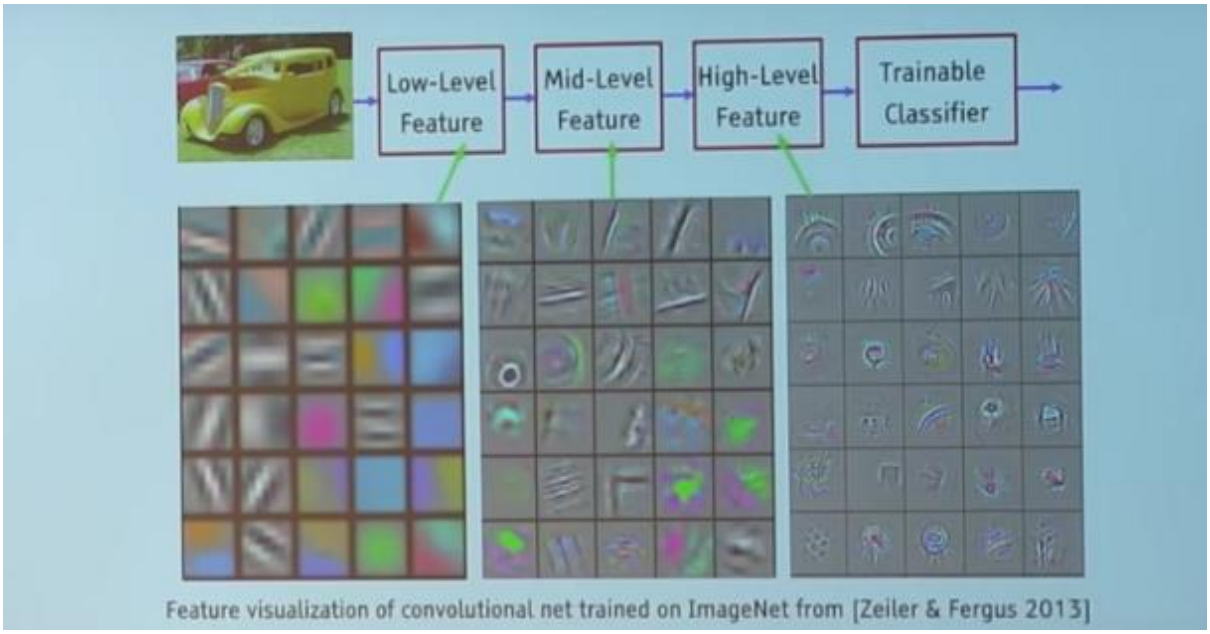
- CNN operated over volume
- Unlike MLP where input was a vector here input is a multi-channelled image
- **Convolution Layer**:



- This orange coloured kernel will slide over the input image and take dot product between kernel and chunks of the input image

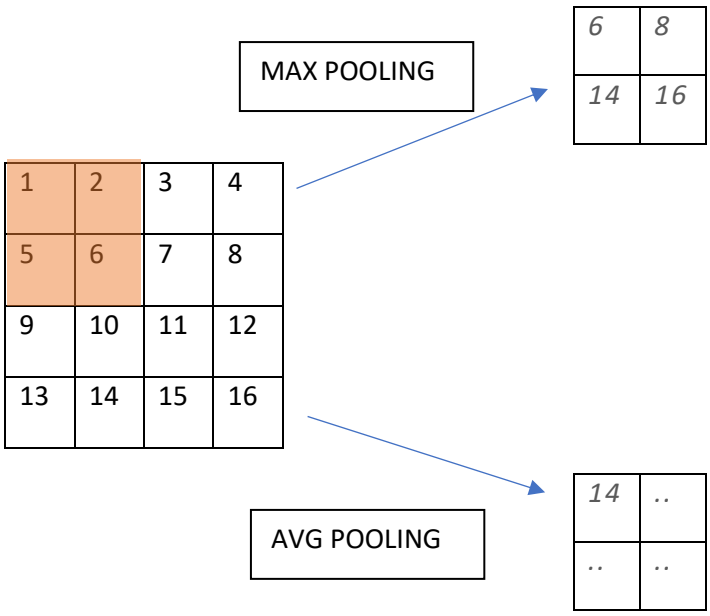


- Convolution layer consists of set of independent filters, each filter is independently convolved with image
- All these filter are initialized randomly and become our parameters which will be learned by the network.
- Like the example given below is the CNN trained over ImageNet network though
- backpropagation they tuned themselves to become such shapes or edges with colours

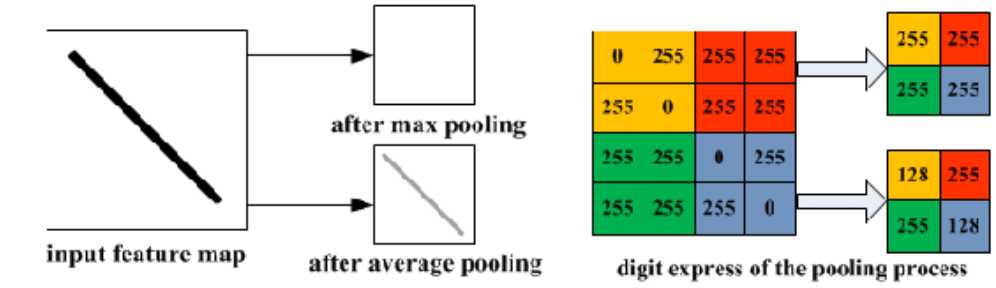


Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

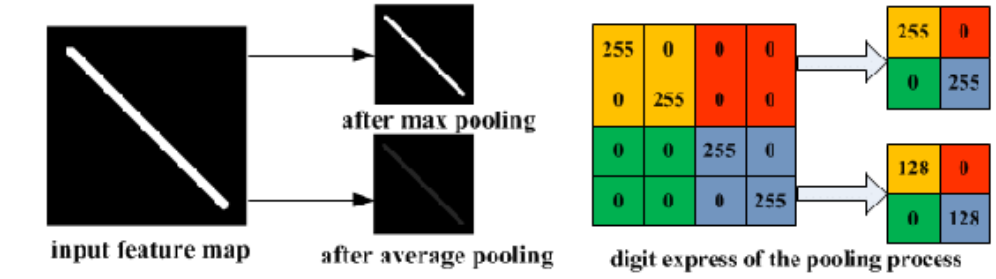
- **Pooling Layer**:



- Pooling above is done with a 2*2 kernel where in max pooling maximum value has been selected out of the kernel overlapped and in average pooling average value has been selected
- Max pooling extracts most important features like edges whereas average pooling extracts features so smoothly
- So which is better?



(a) Illustration of max pooling drawback



(b) Illustration of average pooling drawback

Source: https://www.researchgate.net/publication/300020038_Mixed_Pooling_for_Convolutional_Neural_Networks/figures?lo=1

- So max pooling help model to learn edges and average pooling prevents it.

- **Stride**
    - Till now we were sliding the kernel by a value of one but we can also move it with some value and that term is known as stride so in place of pooling we can also use stride for down sampling, stride 2 means down sampling by 2.

Why CNN better than MLP?

- **Parameter Sharing**
    - Parameter sharing is sharing of weights by neuron in a particular feature map
    - This works in a way by assuming that if one feature is useful to compute at some spatial position (x,y) then it should also be useful to compute at a different position (x', y') in other words lets say a volume size of [x, y, d] has d depth slice each of [x,y]. neurons in each depth slice are constrain to use the same weights and bias so there would be only d unique weights in this case instead of x*y*d.

- **Local Connectivity:**
    - Each neuron is connected to only a subset of input image unlike MLP where each neuron is connected to all the neuron of previous layer
    - This helps to reduce the number of parameters of the system and make computation efficient

- **Spatial Information is maintained in terms of filters**

Preformation comparison of different models

- Accuracy comparison of MLP AND CNN for character recognition

	Learning rate (%)	Classification rate (%)
MLP	70.72	43.4
caffe-cnn (Lenet)	88	88.39
caffe-cnn (Lenet-5)	86.23	85.53
caffe-cnn (SPnet)	**89.90**	**90.56**

Source : A comparison study between MLP and Convolutional Neural Network models for character recognition Syrine Ben Driss, Mahmoud Soua, Rostom Kachouri, Mohamed Akil

- Parameters comparison: IMAGE CLASSIFICATION
  - CNN : 528054
  - MLP: 3590642
  - Accuracy mlp: 39.42%
  - Accuracy cnn: 67.90%
  - Source: https://www.peculiar-coding-endeavours.com/2018/mlp_vs_cnn/