Convolution Neural Network

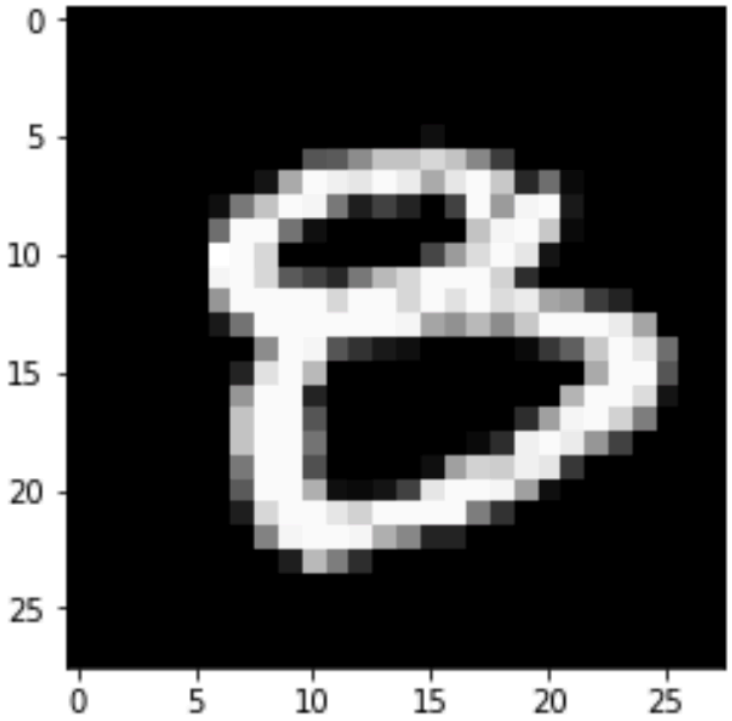
Agenda

Agenda

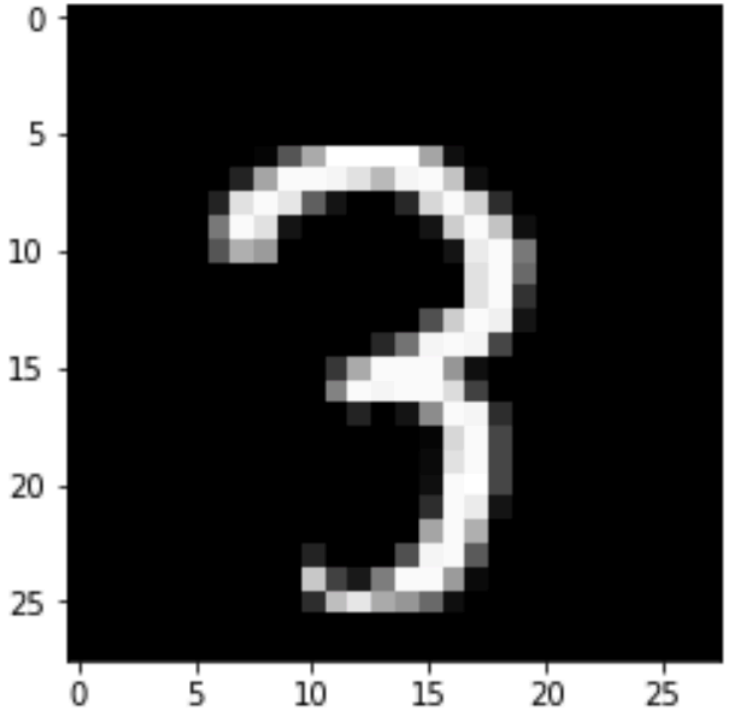
- Limitations of Feed Forward Network with Image Data
- Convolutions for image abstraction
- Pooling
- Fully Connected Layer
- Complete CNN Architecture
- Building a Convolution Network with Tensorflow

FFN & Images

Limitation of Feed Forward Network




```
predicted label: 0
real label: 8
```



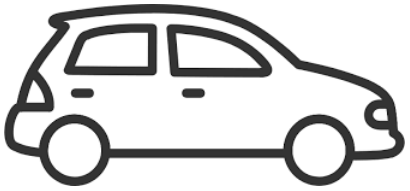

```
predicted label: 9
real label: 3
```

Limitations

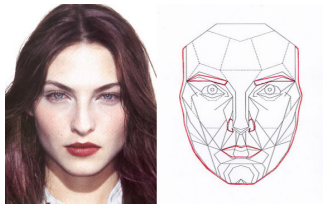
- Simple feed forward network uses pixel values as features
- Does not make use of spatial relationship
- Does not make use of image abstraction*

Convolutions for Image Abstractions

What do we see when we see a Car



… or a face



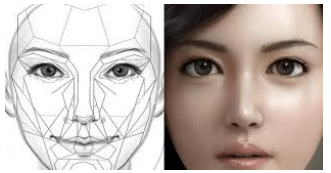
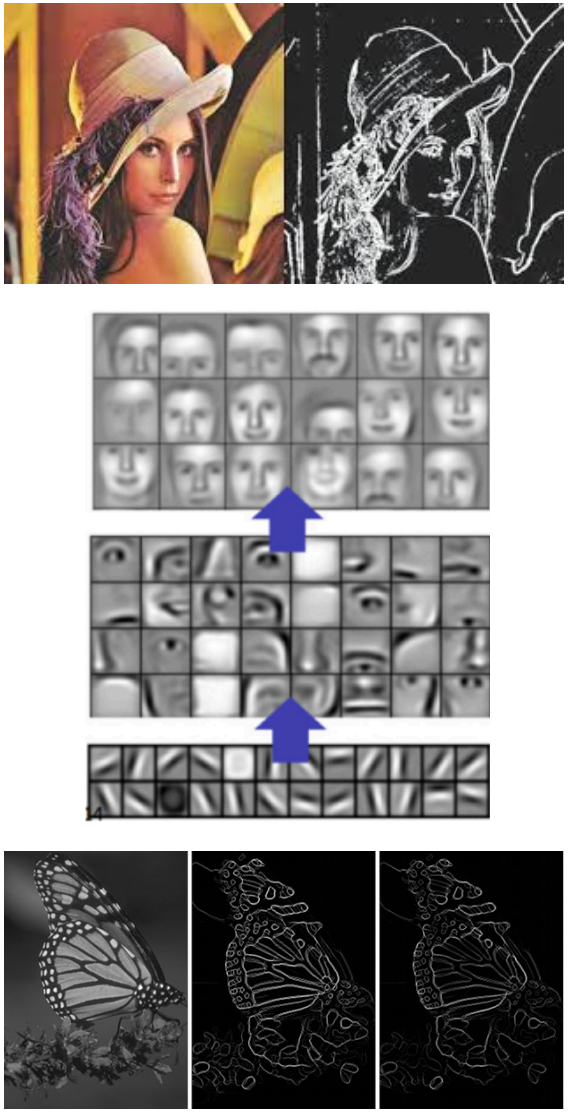


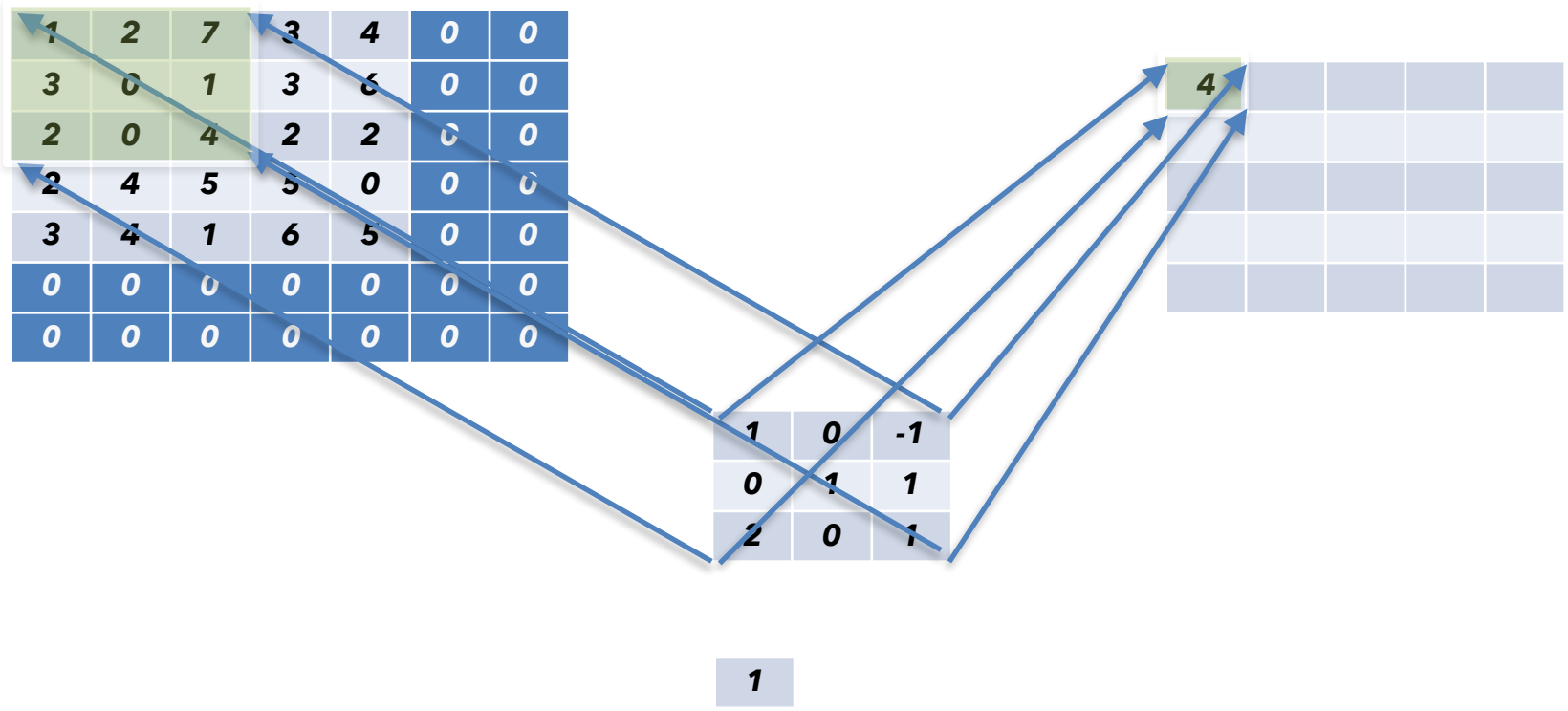
Image Abstractions

- When humans see an image, we don't make use of all the details to identify the object
- We rather focus on some key areas which is then matched with their general representation in our brain
- This is one of the main reason when beginner artists fail to draw exactly what they see (we end up drawing the general idea of the object rather than exactly what is in front of us)
- Any object can have multiple such general representation associated with them
- We call them convolutions

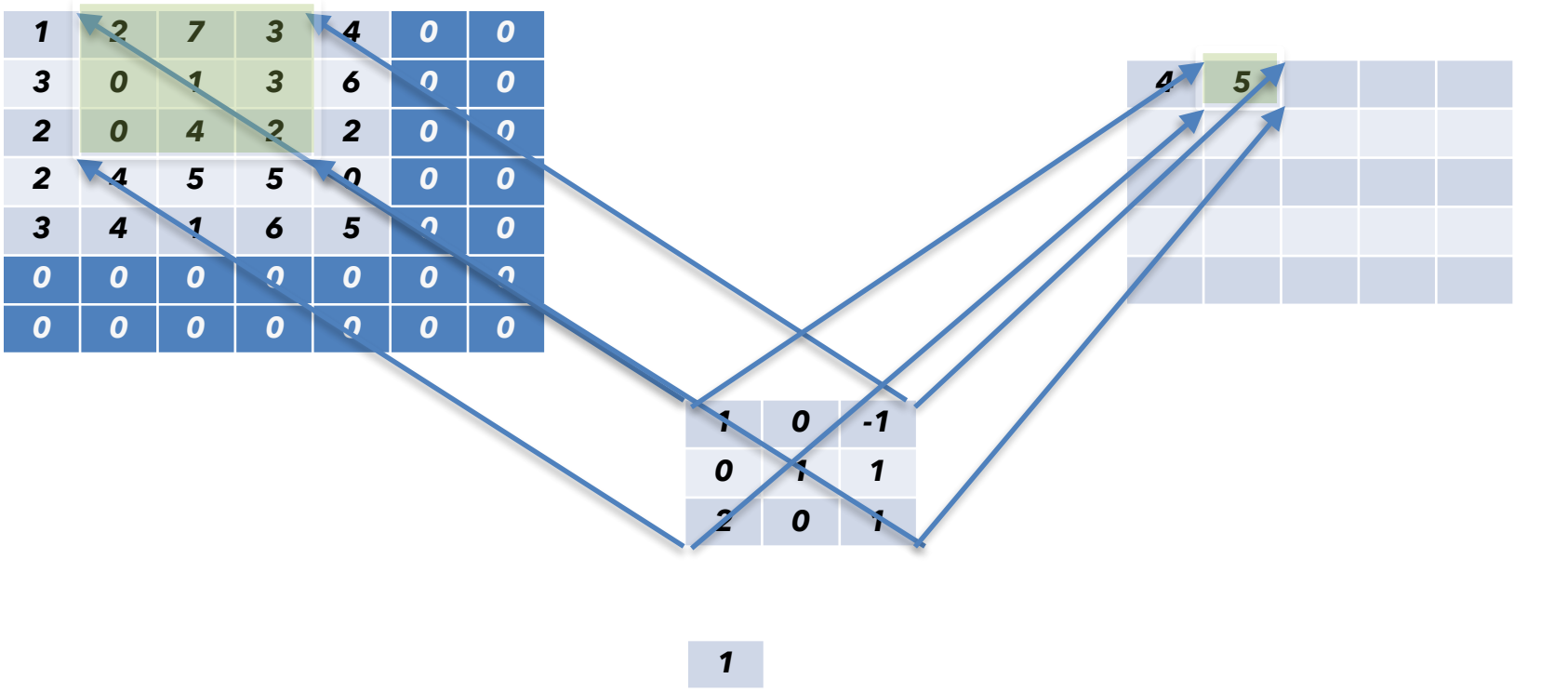
Example of Convolutions



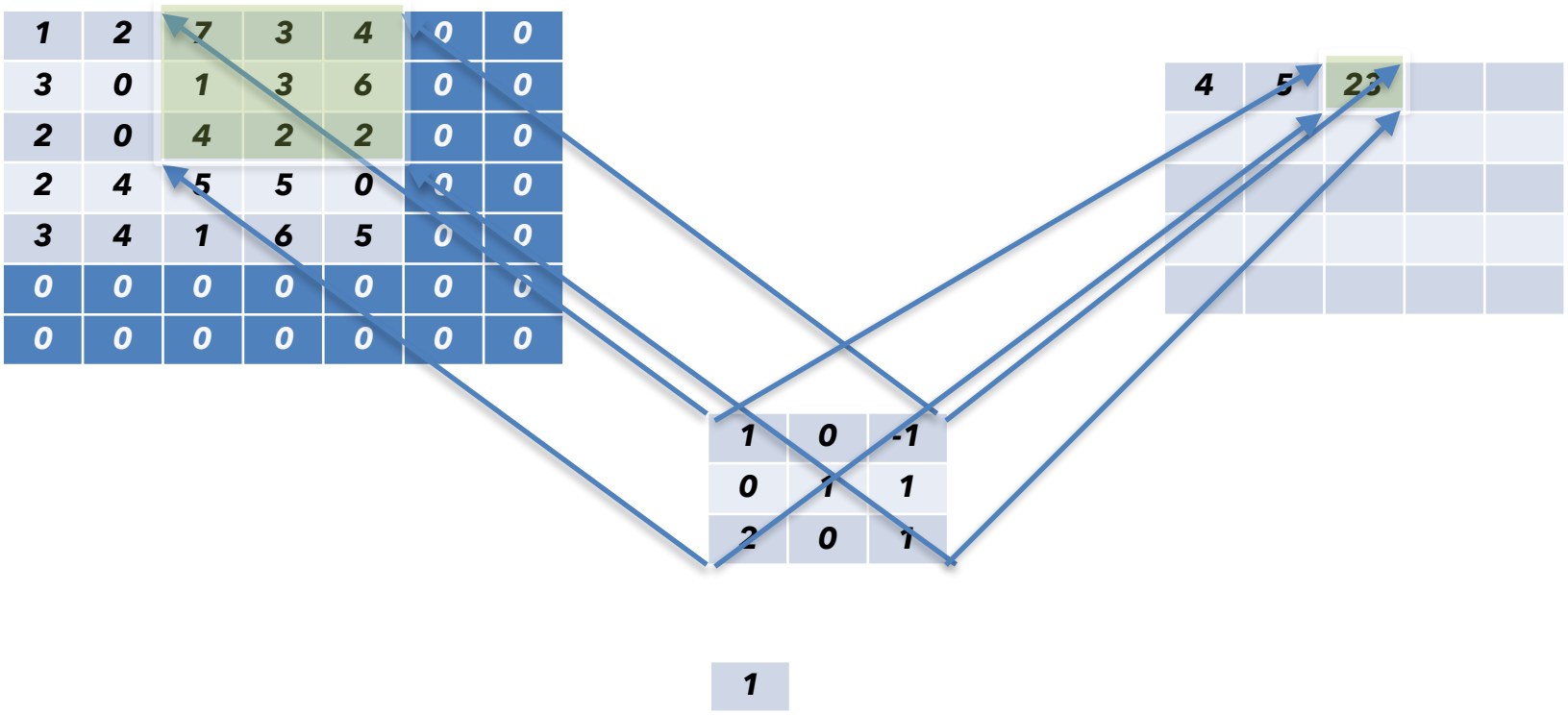
Convolutions as a mathematical operation



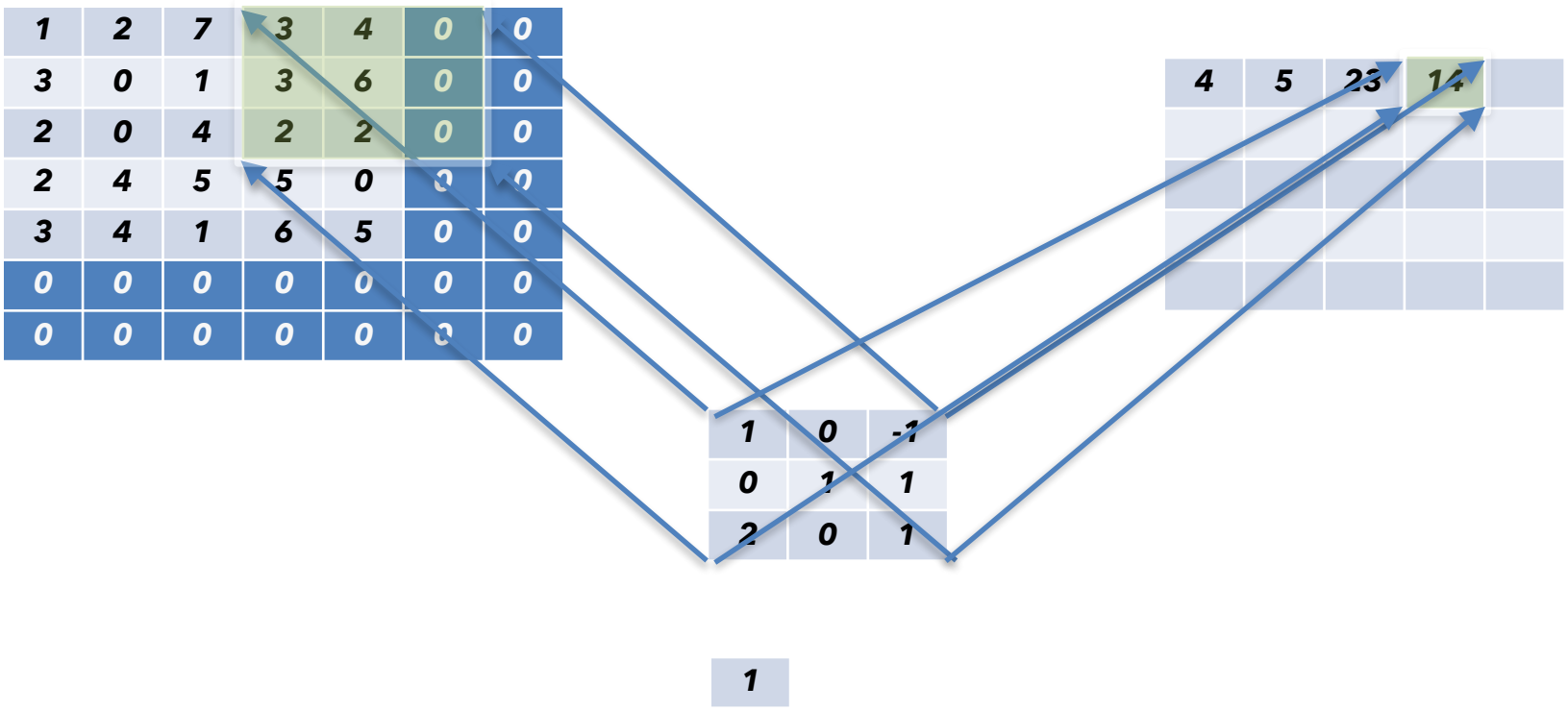
Convolutions as a mathematical operation



Convolutions as a mathematical operation

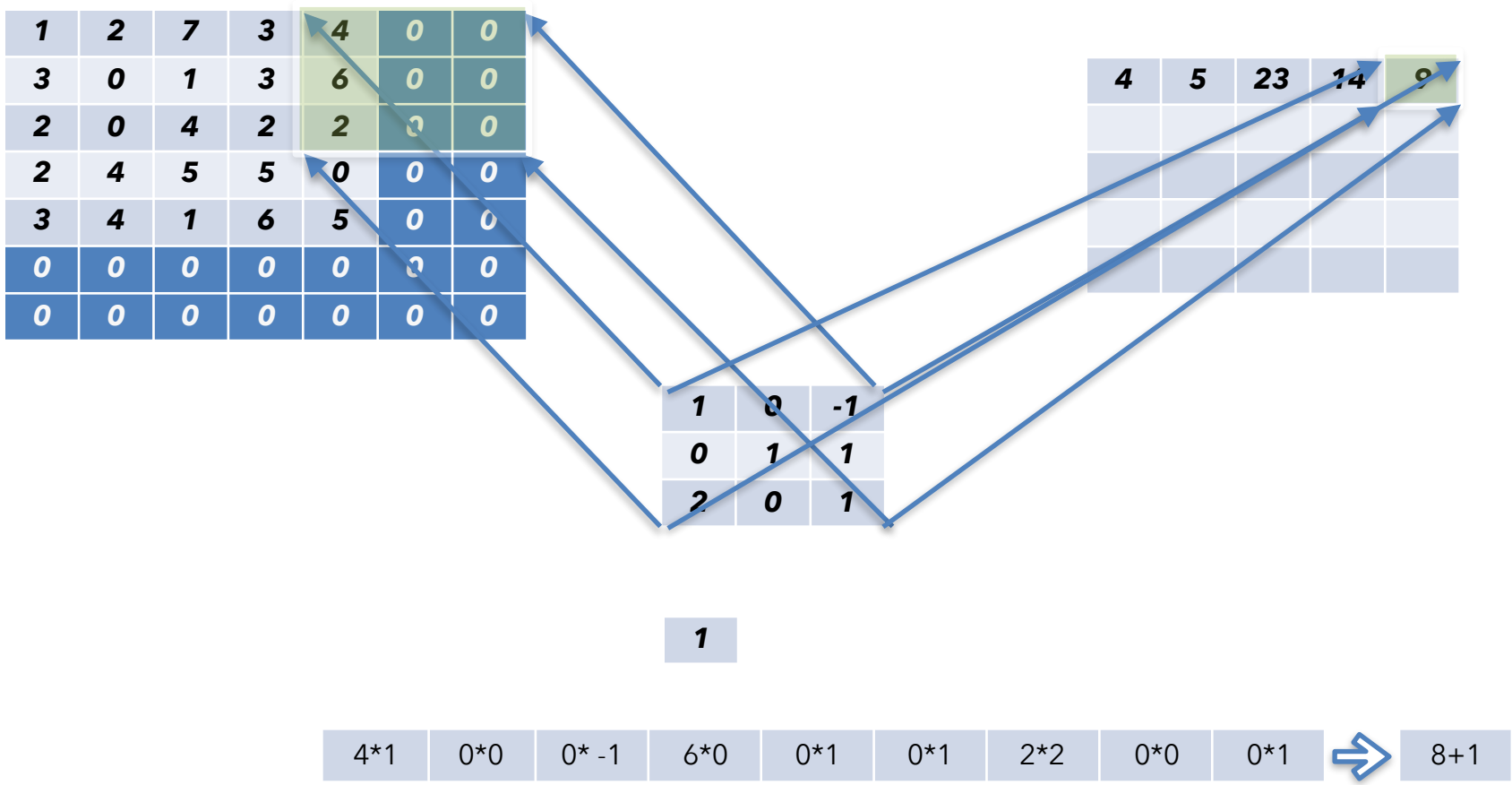


Convolutions as a mathematical operation

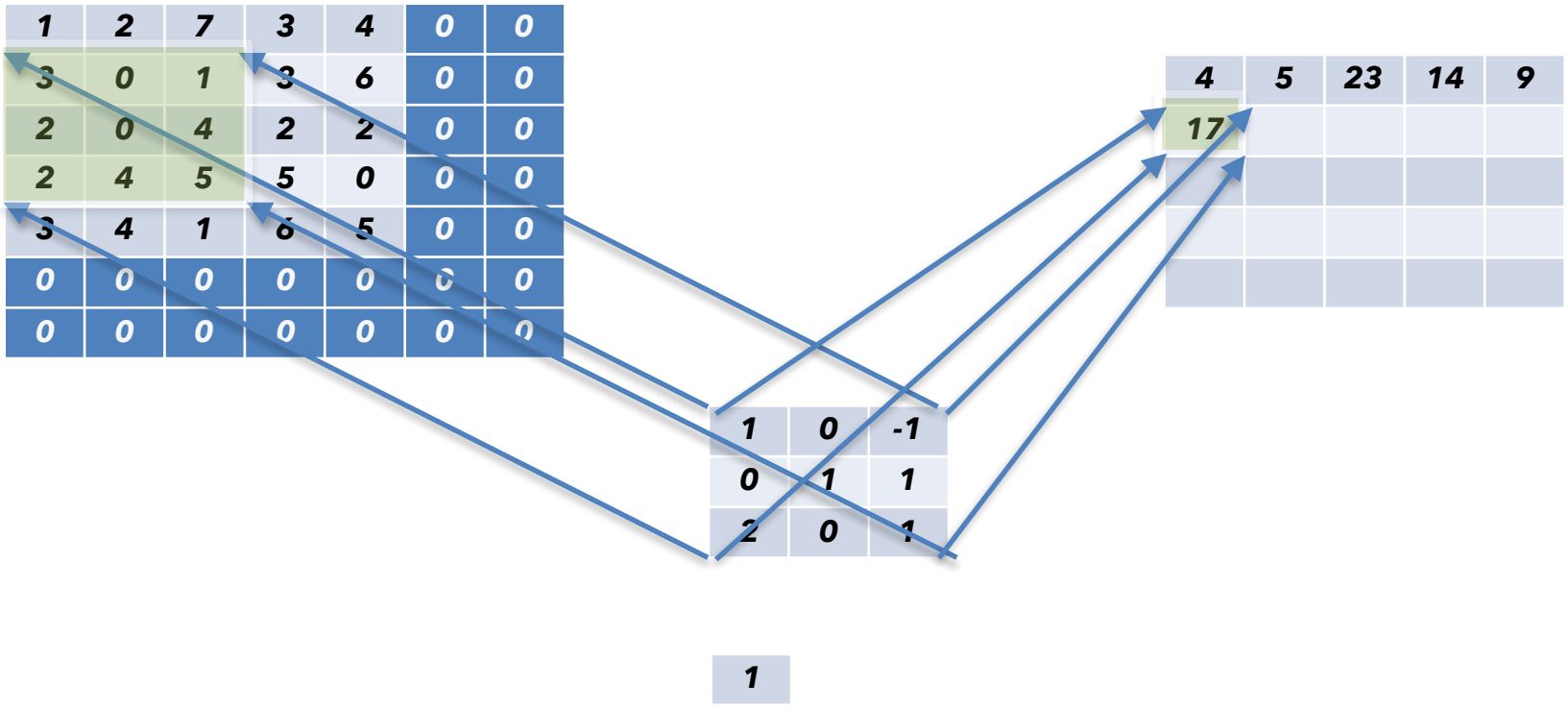


3*1	4*0	0* -1	3*0	6*1	0*1	2*2	2*0	0*1	13+1

Convolutions as a mathematical operation

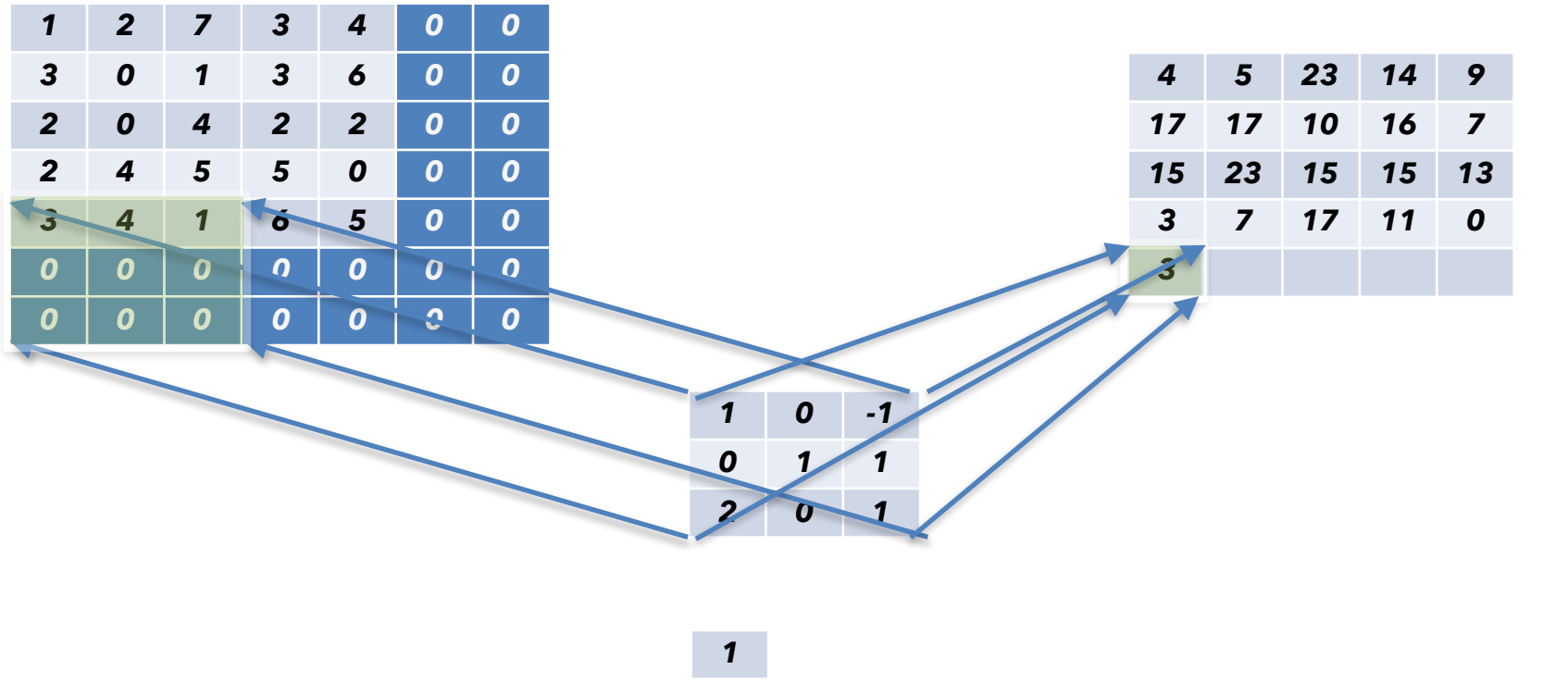


Convolutions as a mathematical operation



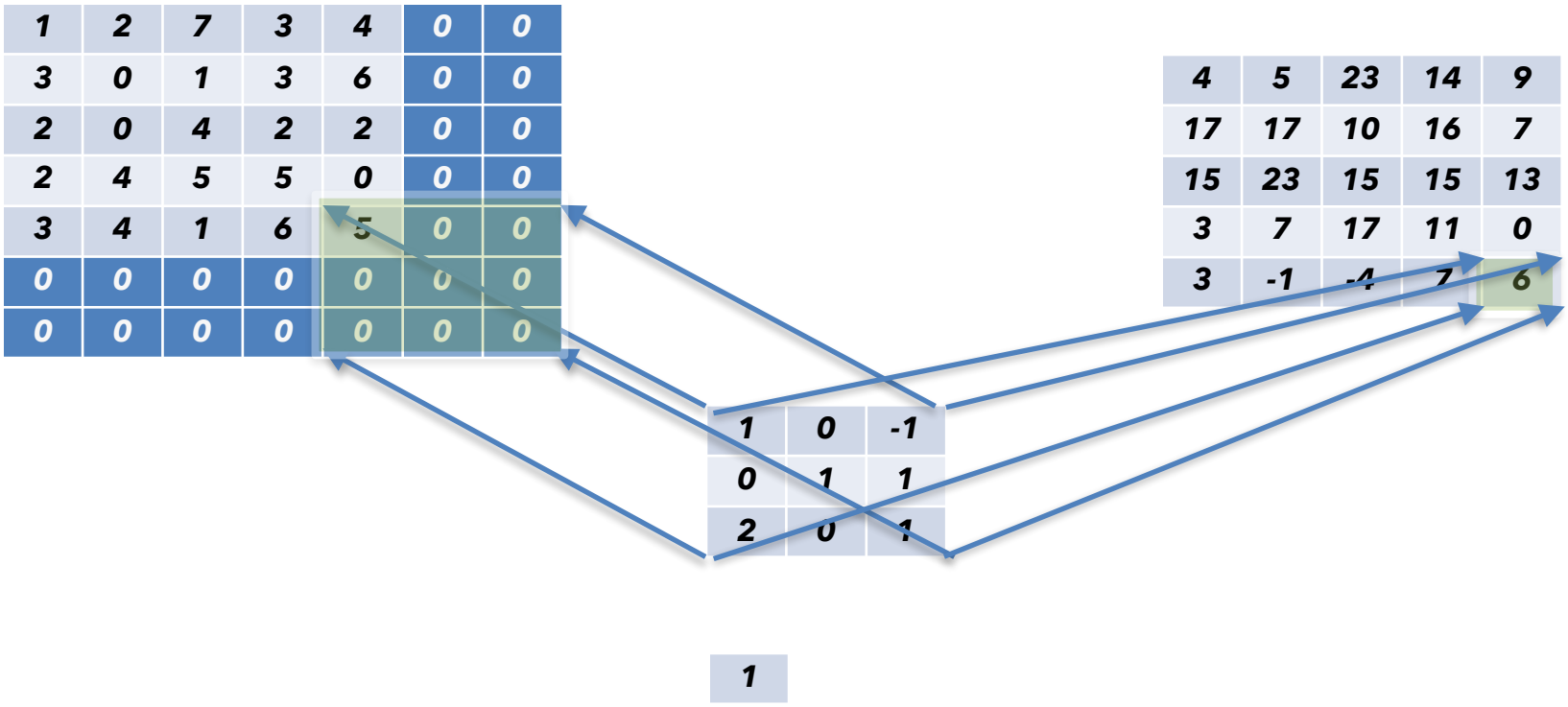
3*1	0*0	1* -1	2*0	0*1	4*1	2*2	4*0	5*1	⇨	16+1

Convolutions as a mathematical operation



3*1	4*0	1* -1	0*0	0*1	0*1	0*2	0*0	0*1	⇨ 2+1

Convolutions as a mathematical operation



5*1	0*0	0* -1	0*0	0*1	0*1	0*2	0*0	0*1	⇨ 5+1

What Convolution Kernel to use?

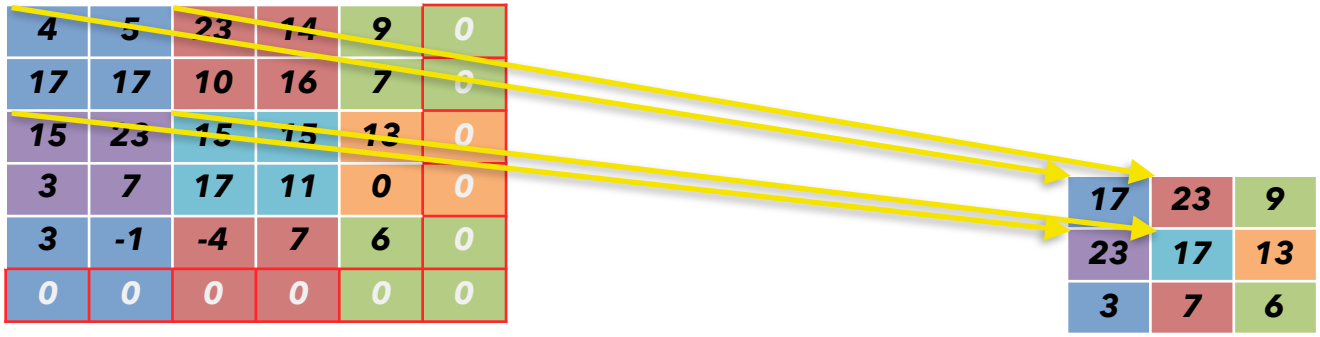
- Before advent of DL and convolution network in general , Computer Vision used to use some generic fixed kernels to get image abstractions
- In CNN we'll rather randomly initialise fixed size kernels and treat them as network parameters to be trained by back propagation
- This enables the algorithm to learn appropriate convolution kernels as per the data
- Multiple kernels are used , which mimic , multiple generic abstractions of a single image

Max Pooling

Purpose of Pooling

- Reduces spatial dimensions (number of parameters), introduces zero parameters
- Picks most pronounced features
- Provides slight invariance to translation (positional shifts of features)
- Large pooling windows however lead to destructions of features due to wide spread omissions and might not work that great
- In practice, strides are equal to pooling window size . However pooling windows can be overlapping as well (used with larger pooling windows)
- There are other varieties of pooling as well : Average and sum (are not used a lot)

Max Pooling as a mathematical operation



- 2X2 pooling with stride 2 and padding 'same'

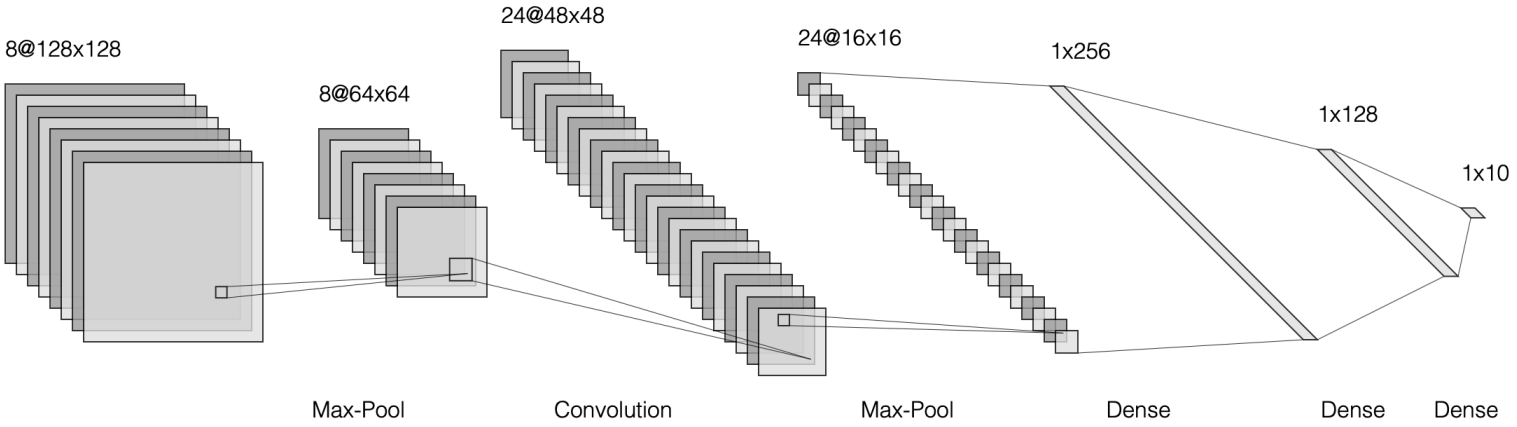
Flattening & Fully Connected Layer

Flattening & Fully Connected Layer

- Output layer is simply a flat layer as we saw in simple feed forward network
- Multi-Dimensional output of convolution and pooling operations need to be flattened to connect with output layer
- We can further introduce more flat layers in between , instead of simply connecting the flattened output to target layer
- These are called fully connected or dense layers

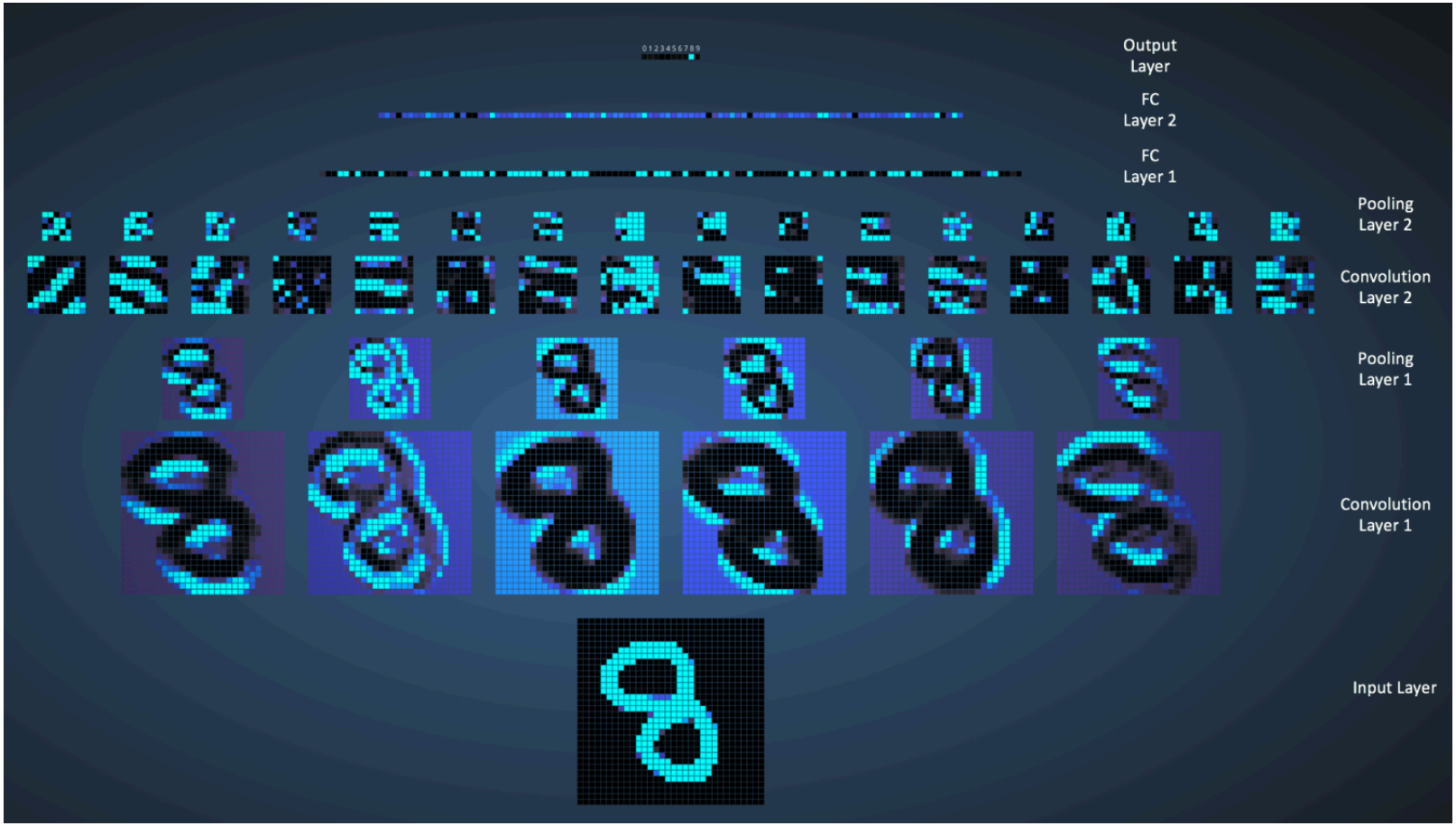
Complete CNN Architecture

CNN Architecture



- Here input is 128X128 pixel image
- First result is obtained by convolution using 8 kernels
- Then a 2X2 max pool window is applied with stride of 2
- Then a convolution is applied with 24 kernels (without padding , this results in altered size of convolution results)
- Then 3X3 max pool window is applied with stride 3
- Output is internally flattened and is connected to Fully connected layer of size 256
- Another fully connected layer of size 128 then connects to output layer

Visualising Effect on a single image



Source : https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/