Convolutional Neural Networks

Lecture # 8

Convolutional Neural Networks

“ConvNet” architecture originally proposed for images

Convolutional Neural Networks

"ConvNet" architecture originally proposed for images

The assumption was that spatially local information is very important in images

Convolutional Neural Networks

"ConvNet" architecture originally proposed for images

The assumption was that spatially local information is very important in images

Objects are "groups" of pixels that are nearby

Convolutional Neural Networks

ConvNet's propose the concept of a *filter* (also known as a *kernel*) - which acts like a **feature detector**

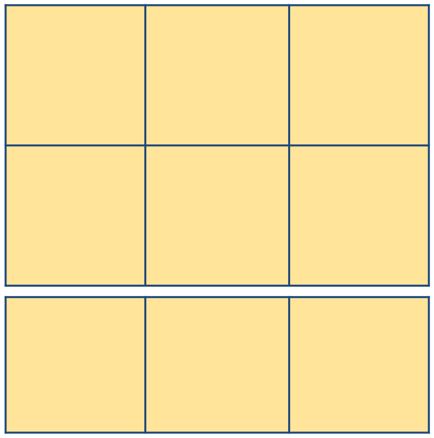
Convolutional Neural Networks

ConvNet's propose the concept of a *filter* (also known as a *kernel*) - which acts like a **feature detector**

It does the same job as a *neuron* in the networks we've seen so far - each *filter* considers some aspect of the input and its outputs are a measure of how much the filter supports the particular aspect/feature

Convolutional Neural Networks

Filters are defined by some width and height, and are usually square



We slide a filter over the entire image, and therefore produce a corresponding second image from the activations of the filter

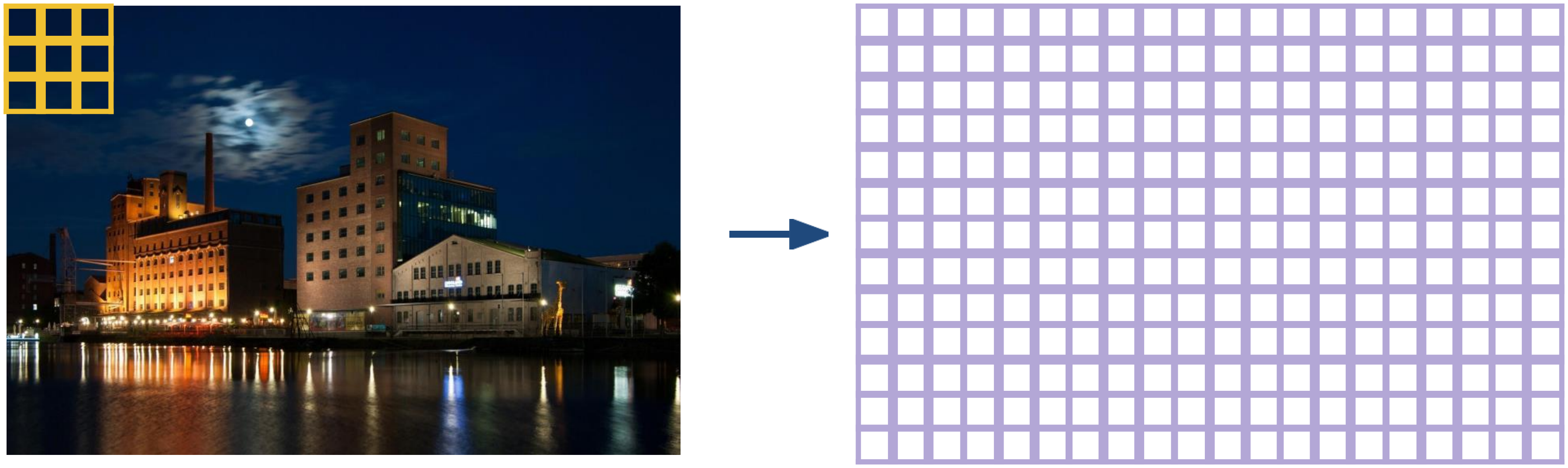
Convolutional Neural Networks

Let’s look at a concrete example

Consider an image of size 15x22 pixels

Convolutional Neural Networks

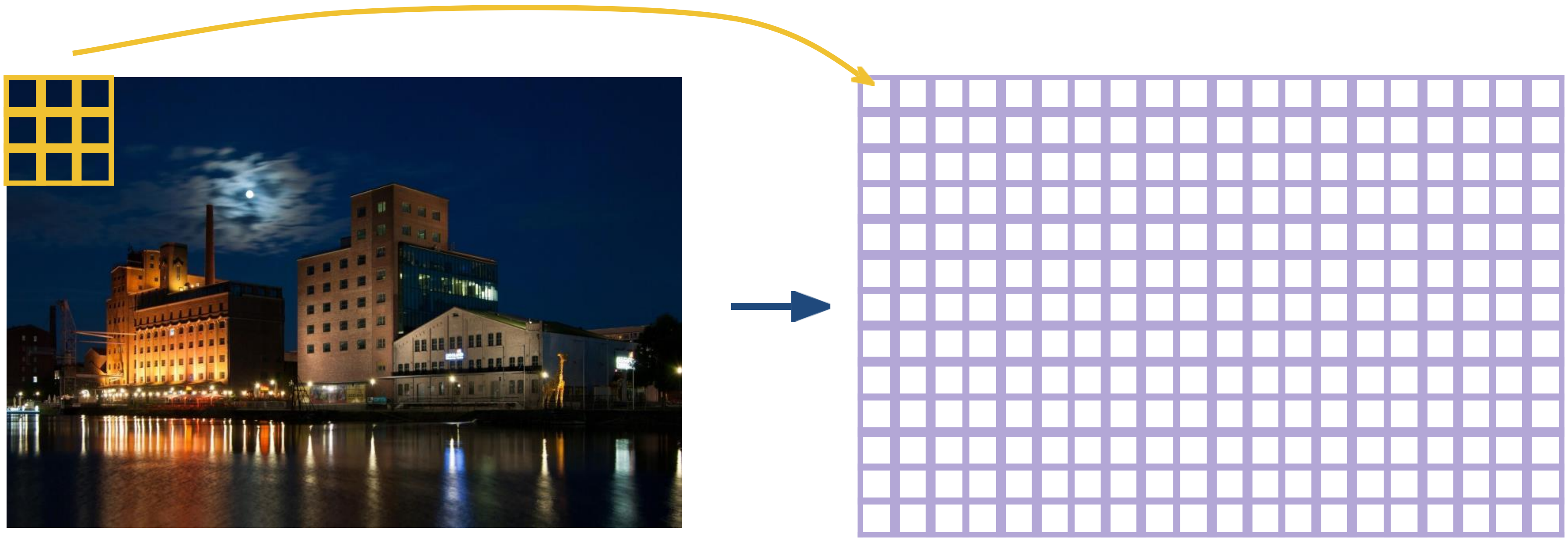
Consider a small 3x3 filter that has learned to detect buildings



“Activation map”
Generated Image

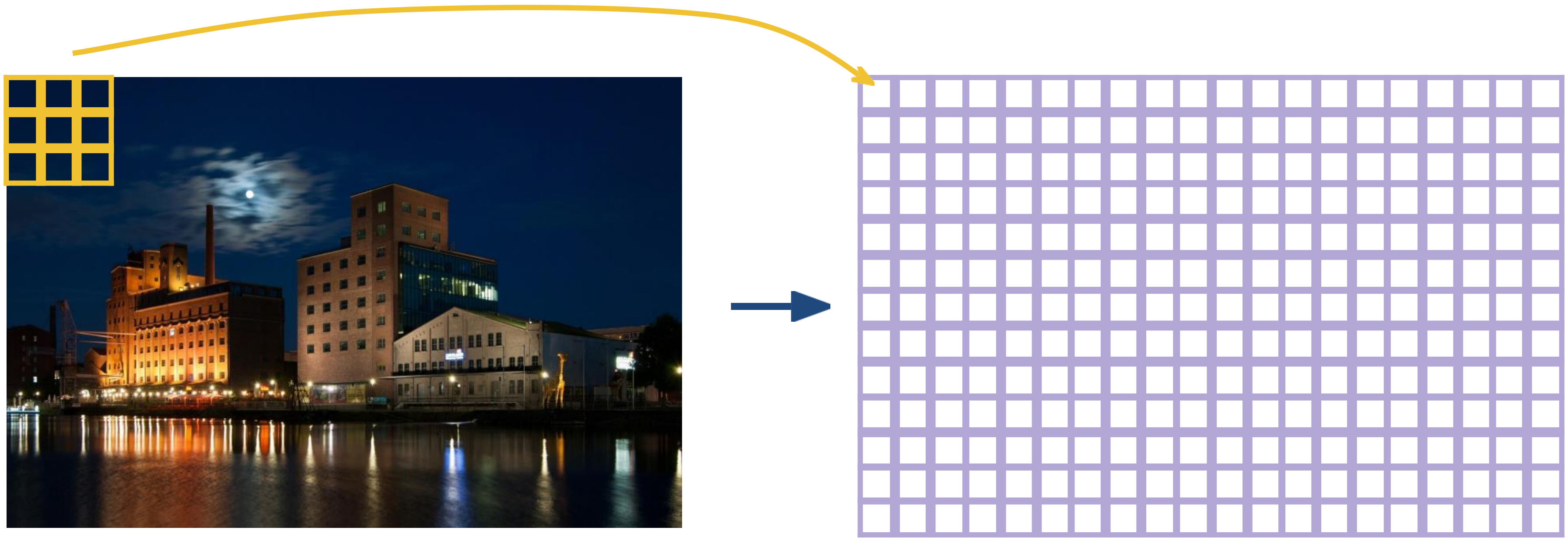
Convolutional Neural Networks

The filter will look at the 9 pixels underneath it,
and output **one** value



Convolutional Neural Networks

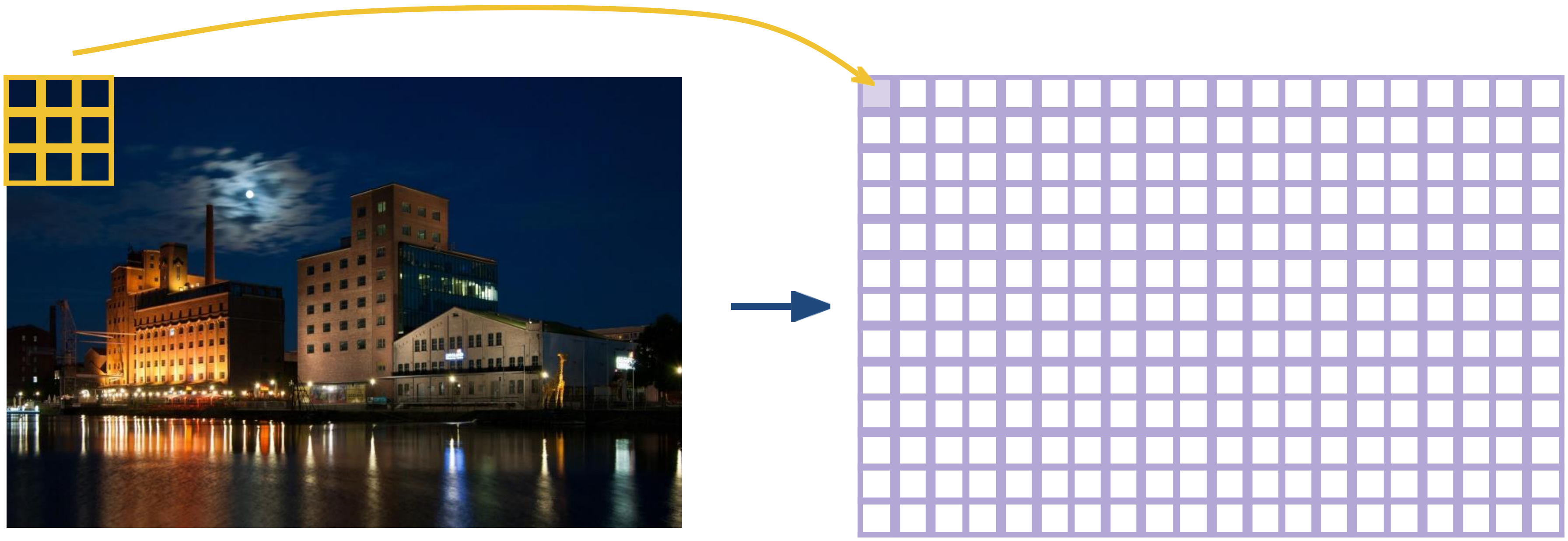
This value can be considered as its “support” for the feature it is learning, in this case “buildings”



“Activation map”
Generated Image

Convolutional Neural Networks

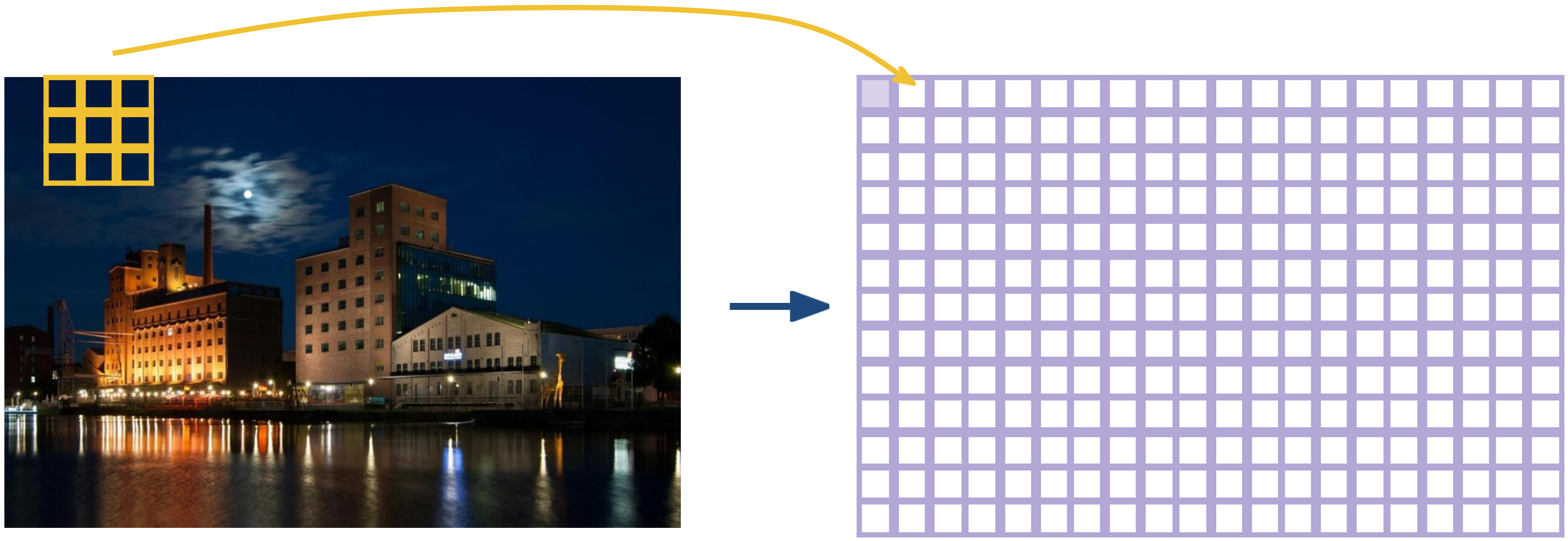
Since there is no building underneath the filter, it will output a low value in this case



“Activation map”
Generated Image

Convolutional Neural Networks

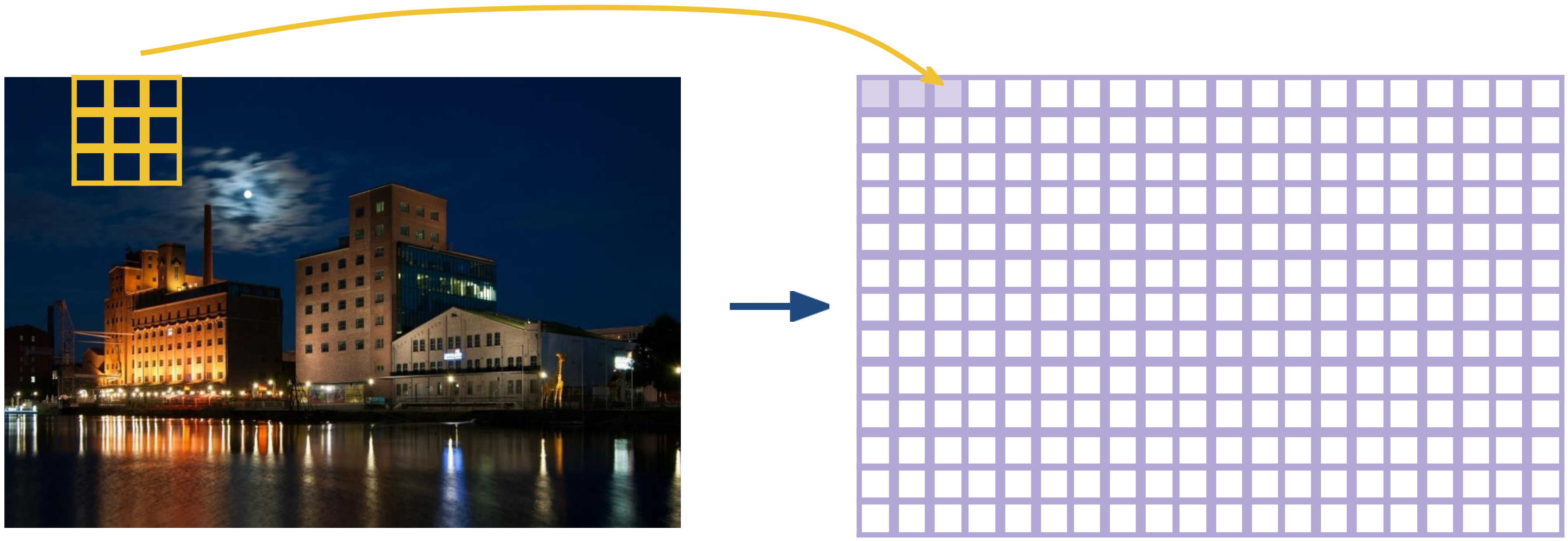
We then slide the window **by a pixel** to the right,
and repeat the process



"Activation map"
Generated Image

Convolutional Neural Networks

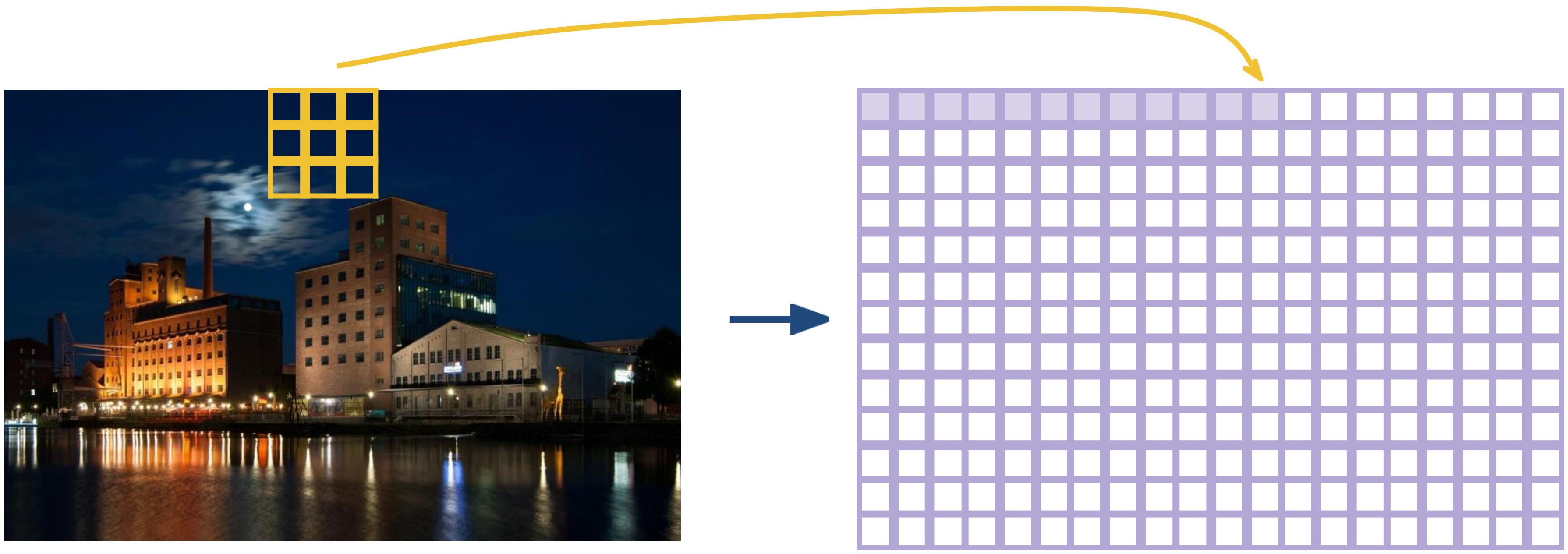
We then slide the window **by a pixel** to the right,
and repeat the process



"Activation map"
Generated Image

Convolutional Neural Networks

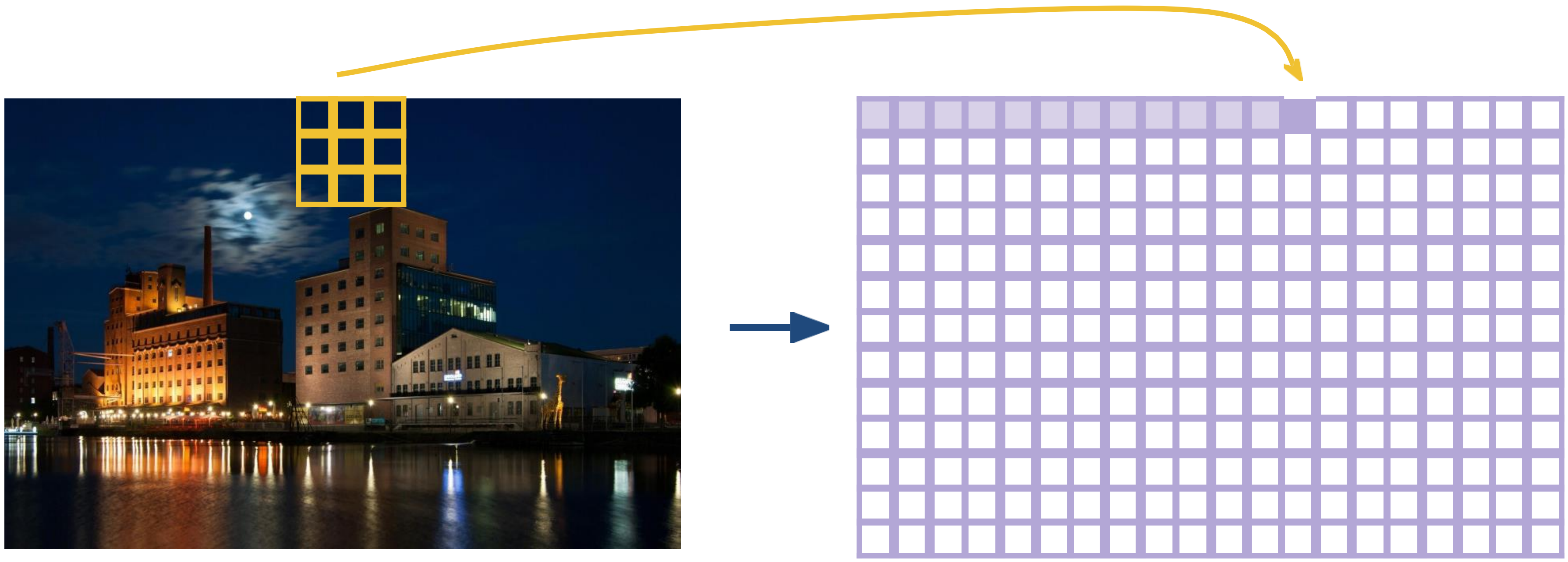
We then slide the window **by a pixel** to the right,
and repeat the process



“Activation map”
Generated Image

Convolutional Neural Networks

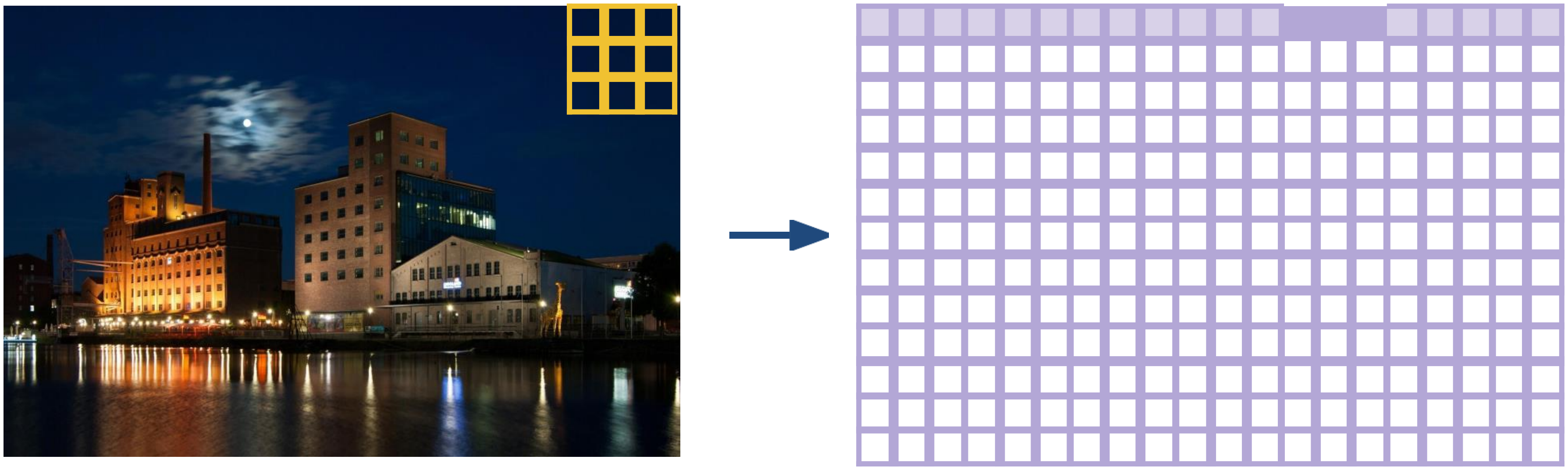
In this case, the filter finally sees some portion of some building, so its output will be slightly higher



“Activation map”
Generated Image

Convolutional Neural Networks

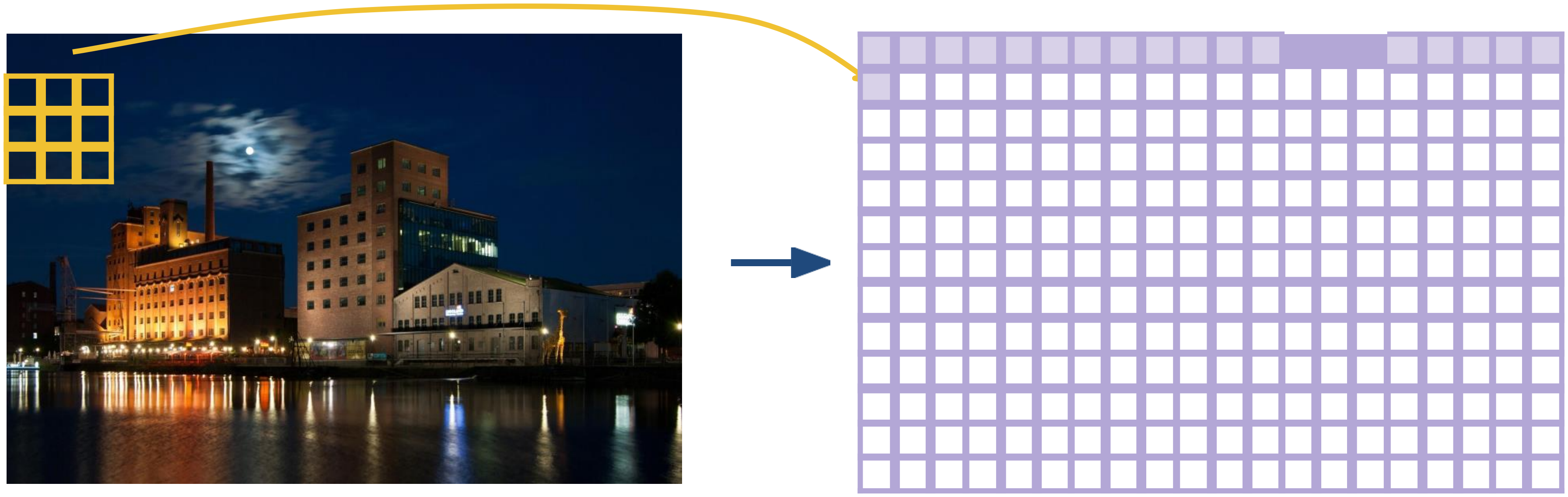
We repeat the process until we reach the end of the row



“Activation map”
Generated Image

Convolutional Neural Networks

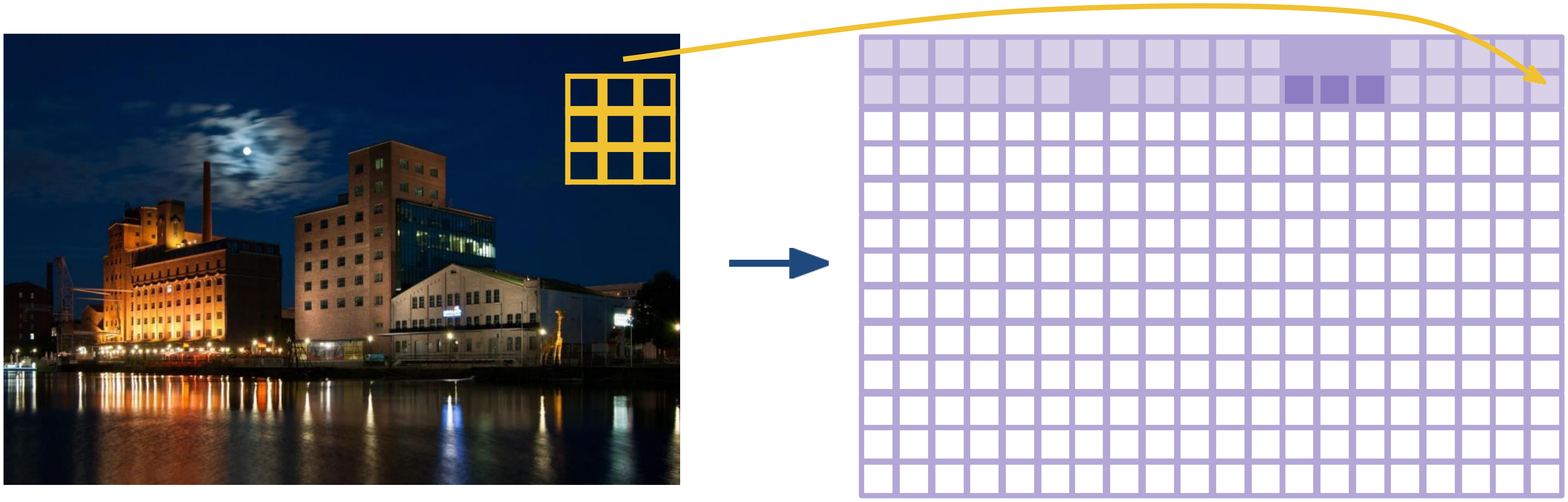
We shift **one pixel down** and repeat the entire process, filling in the second row of the "activation map"



"Activation map"
Generated Image

Convolutional Neural Networks

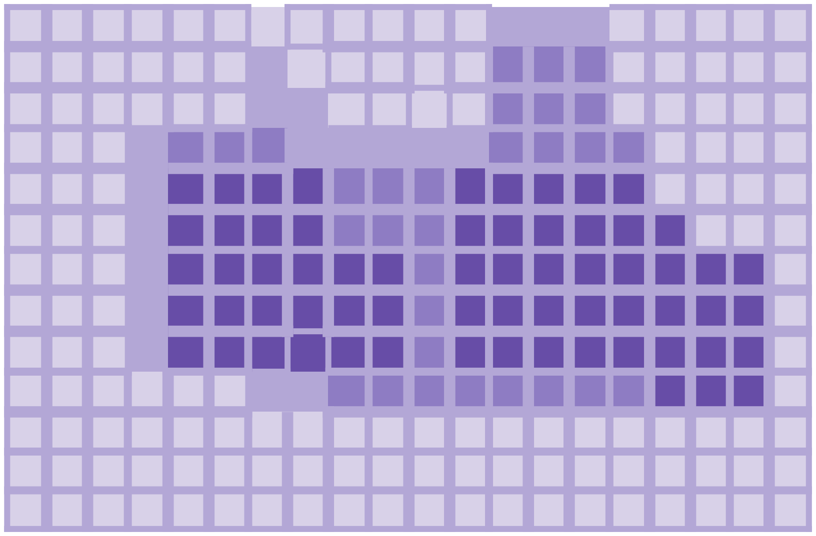
We continue until the end of the row



“Activation map”
Generated Image

Convolutional Neural Networks

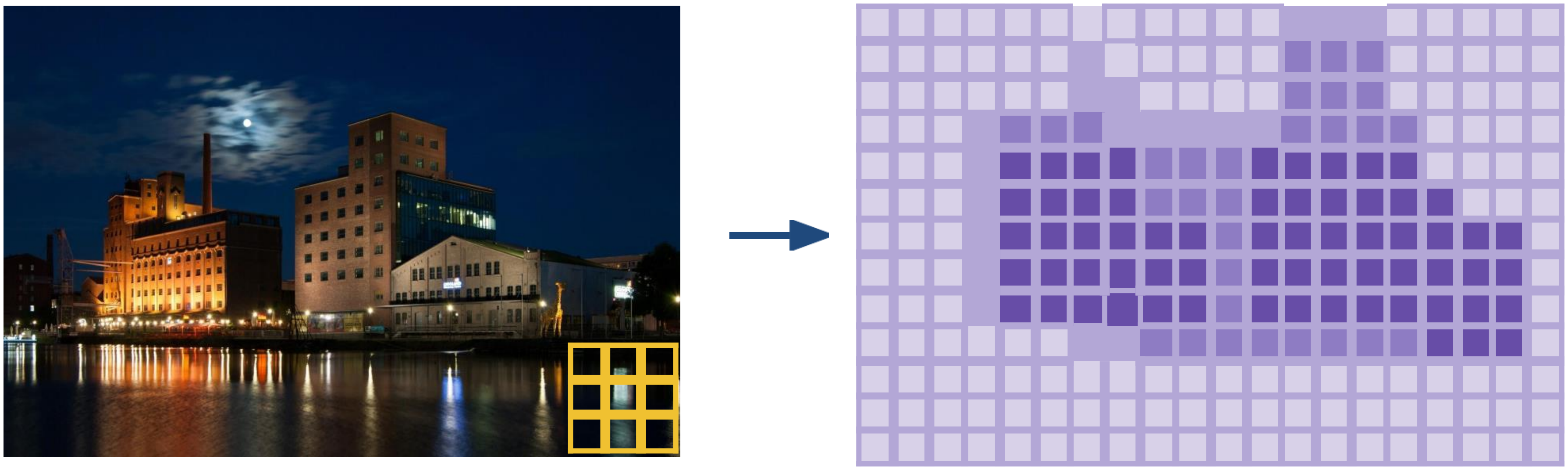
And repeat the entire process for the whole image



"Activation map"
Generated Image

Convolutional Neural Networks

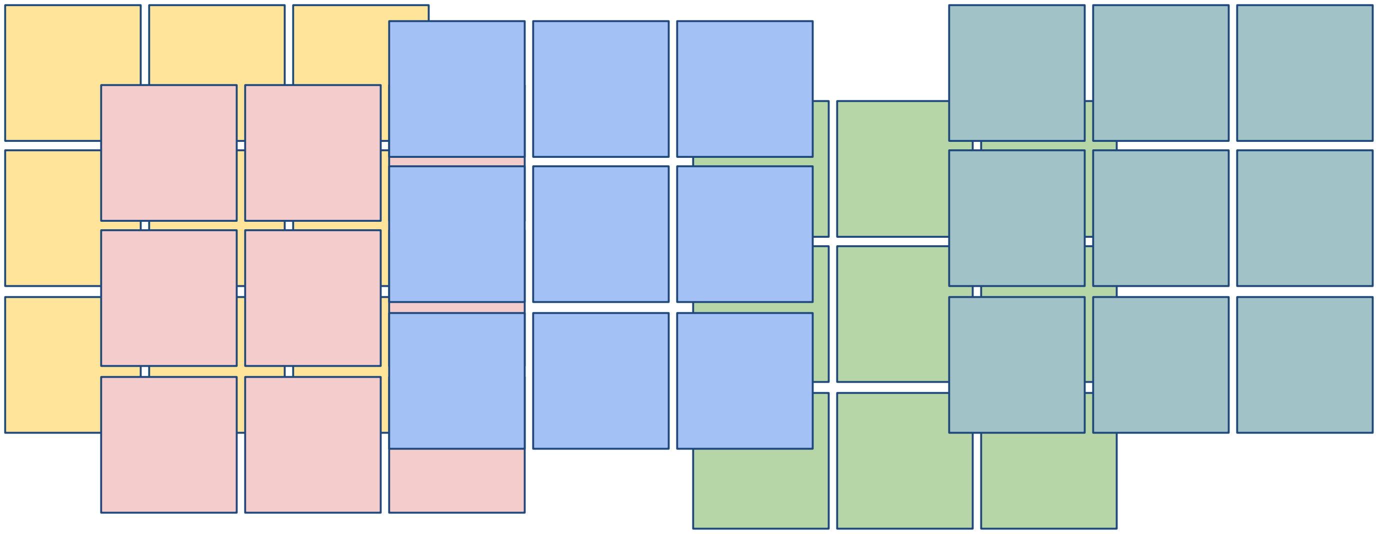
The higher layers can now use this activation map to create richer features based on “buildings information”



“Activation map”
Generated Image

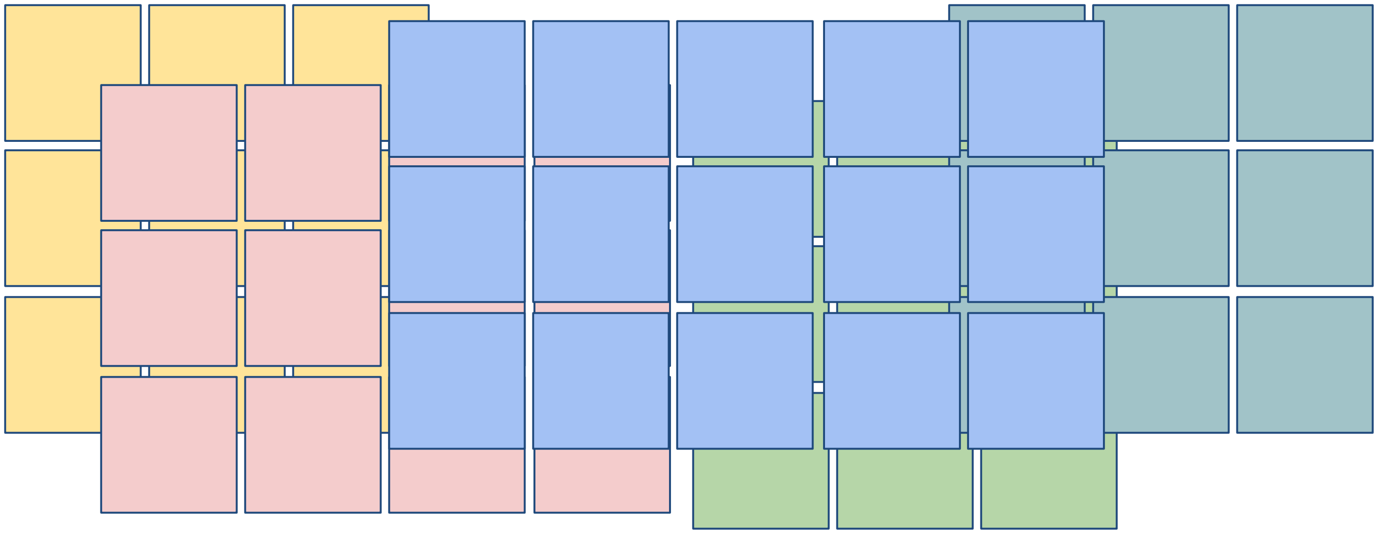
Convolutional Neural Networks

Usually, we have a lot of filters, so each of them can learn different features from the input image



Convolutional Neural Networks

Usually, we have a lot of filters, so each of them can learn different features from the input image



Although not common in images, for text, we can also have filters of different sizes!

Convolutional Neural Networks

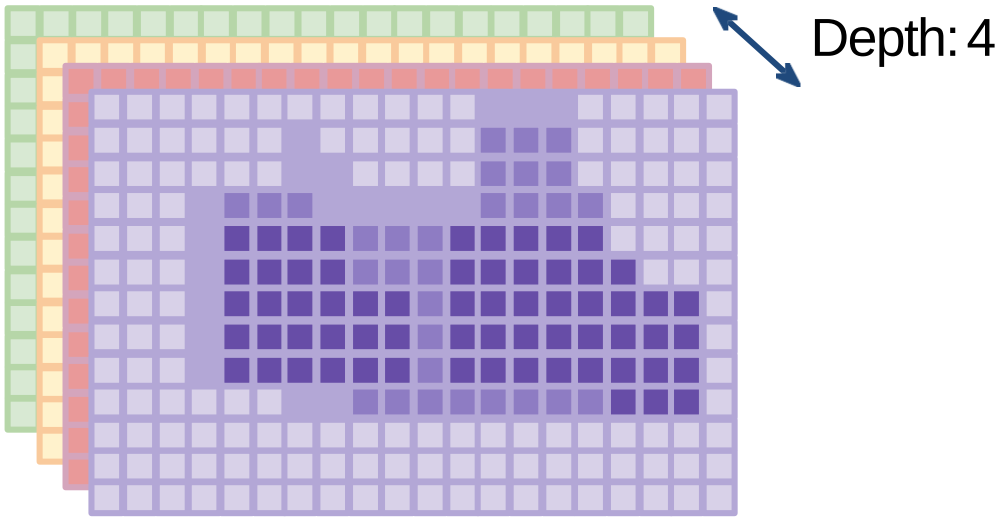
Some terminologies:

- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps

Convolutional Neural Networks

Some terminologies:

- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps



Convolutional Neural Networks

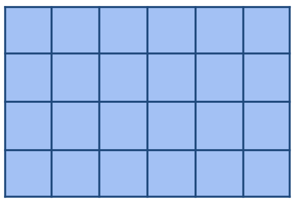
Some terminologies:

- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller

Convolutional Neural Networks

Some terminologies:

- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller

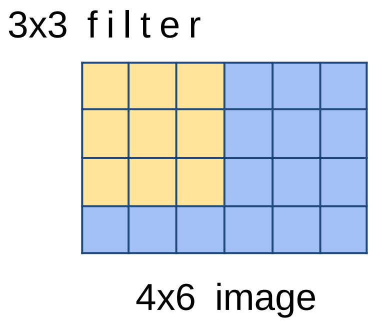


4x6 image

Convolutional Neural Networks

Some terminologies:

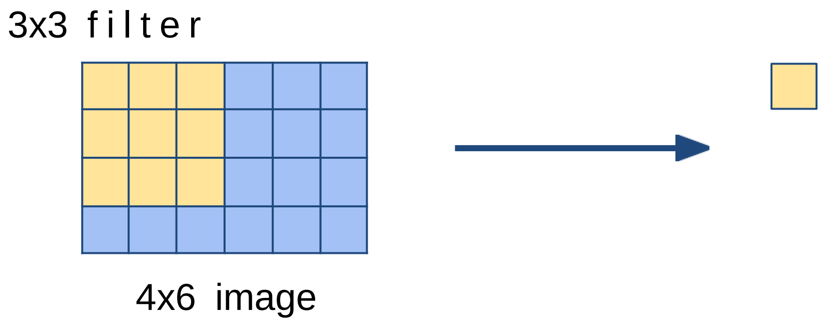
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

Some terminologies:

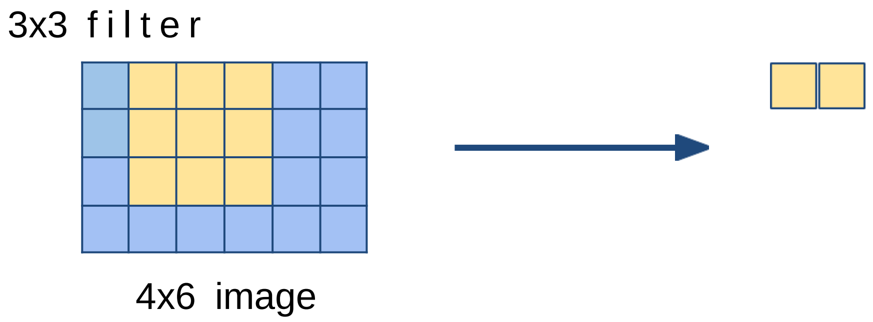
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

Some terminologies:

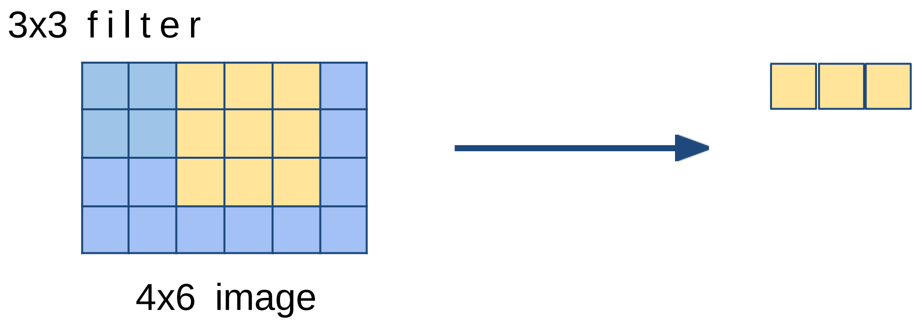
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

Some terminologies:

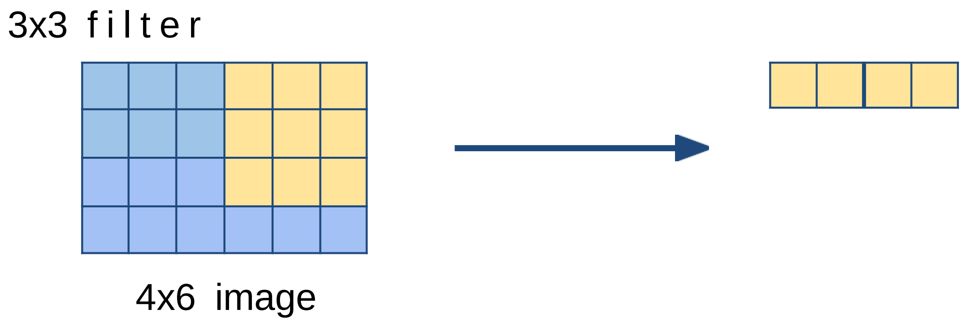
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

Some terminologies:

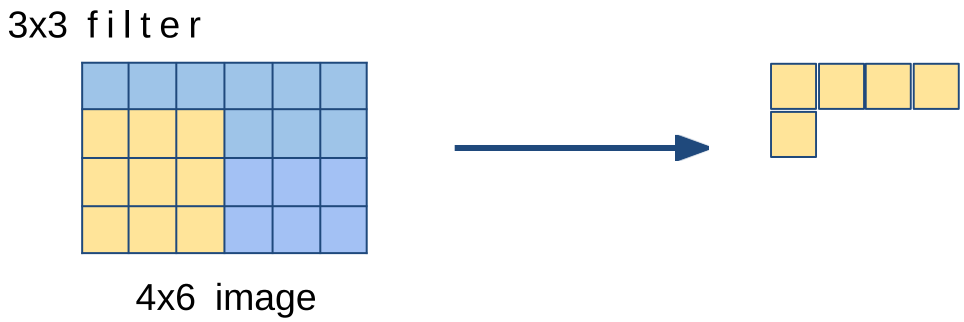
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

Some terminologies:

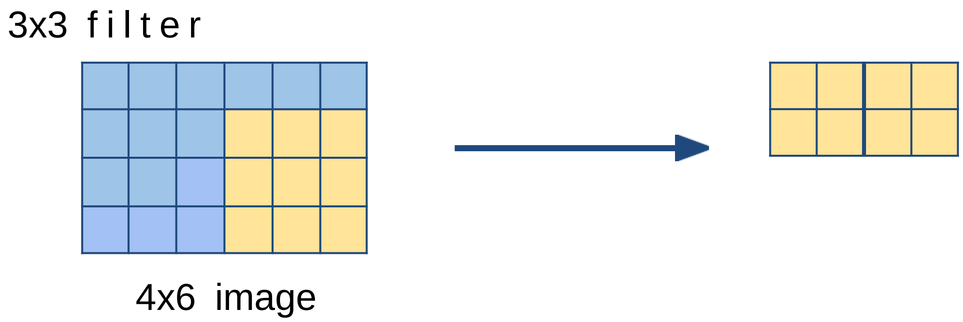
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

Some terminologies:

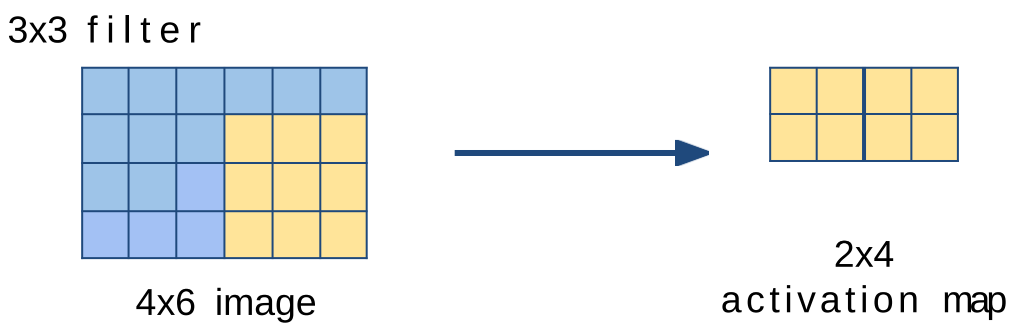
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

Some terminologies:

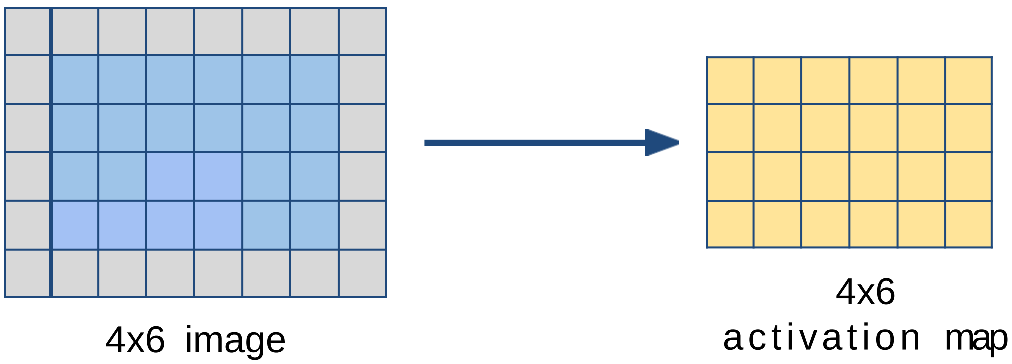
- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Since we are sliding the filters within the image boundaries, the activation map will be smaller



Convolutional Neural Networks

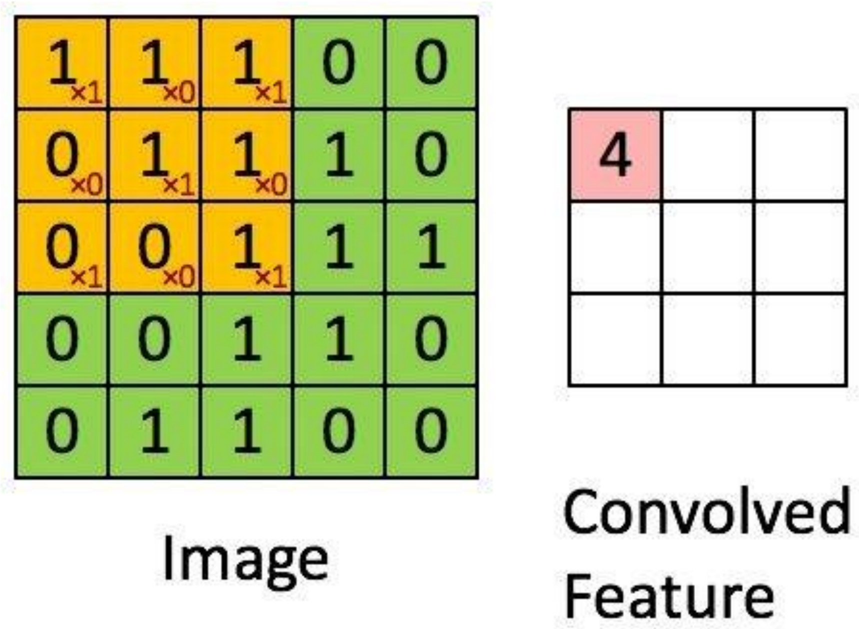
Some terminologies:

- **Stride:** The number of pixels to shift at each step
- **Depth:** The depth of a Conv layer is the number of filters in it - If we have *N* filters, we will have *N* activation maps
- **Padding:** Each image is padded with "zero" pixels to maintain the image size in the map



Convolutional Neural Networks

- Example of an actual computation



http://deeplearning.stanford.edu/wiki/index.php/Feature_extraction_using_convolution

Convolutional Neural Networks

Convolutional layers in text are quite similar to the ones in images, but we work only in one dimension!

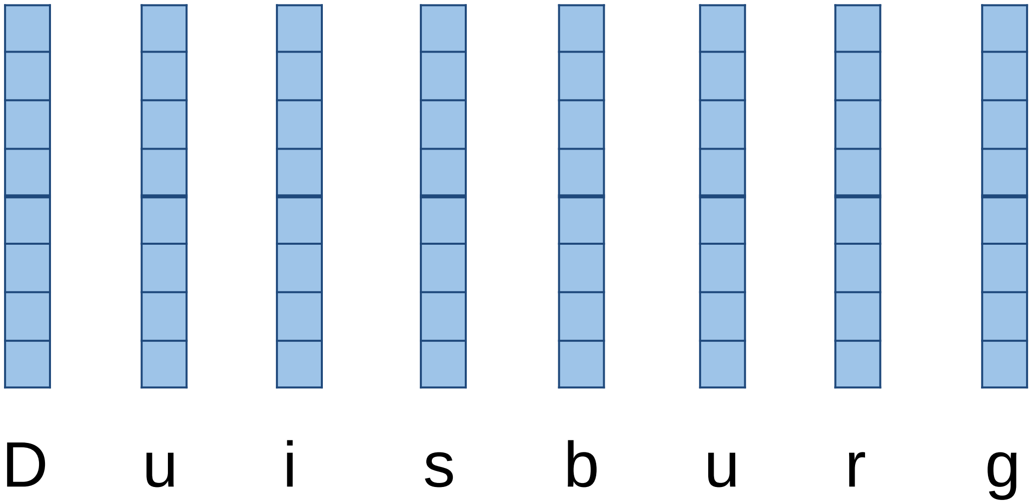
2 Dimensions

John is driving a car

1 Dimension

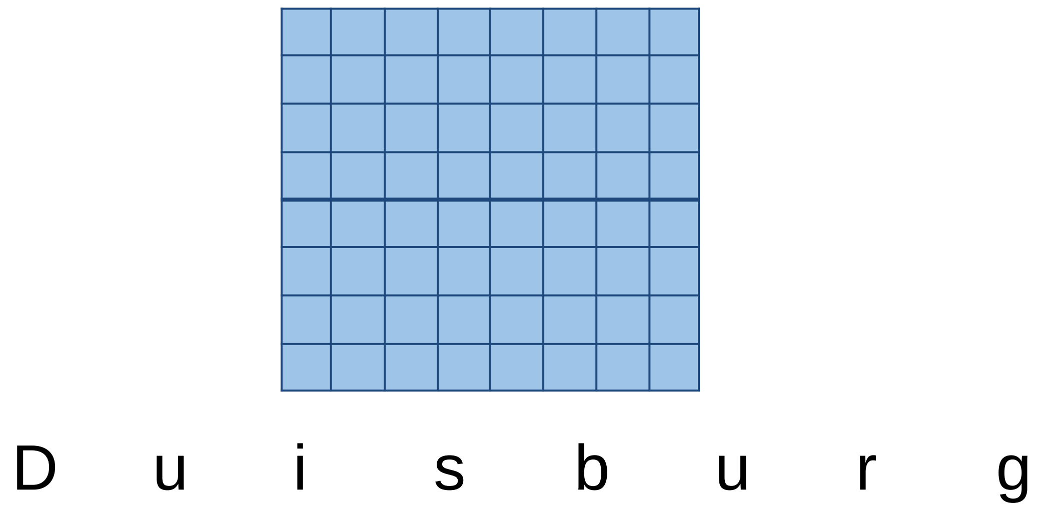
Convolutional Neural Networks

We “slide” our *filters* on the embeddings of the characters



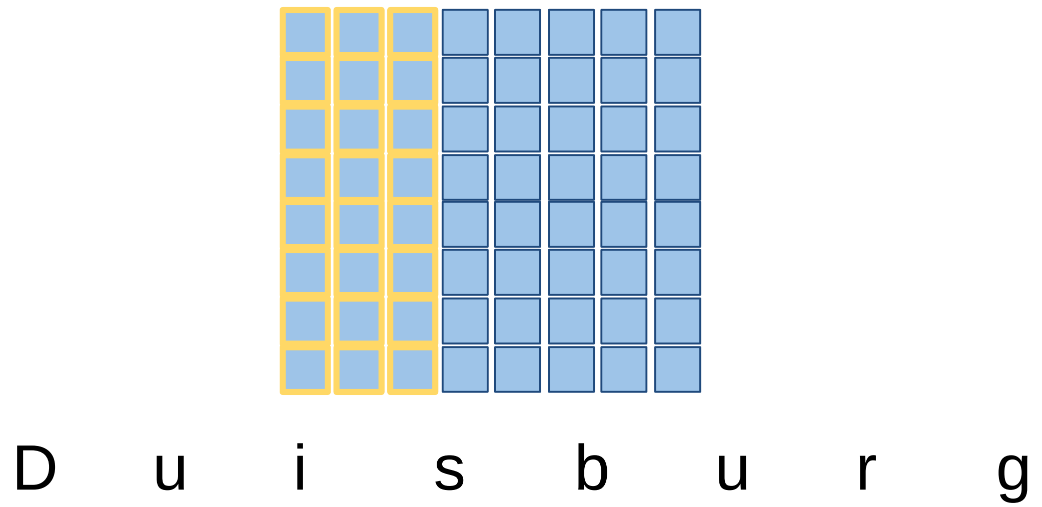
Convolutional Neural Networks

We “slide” our *filters* on the embeddings of the characters



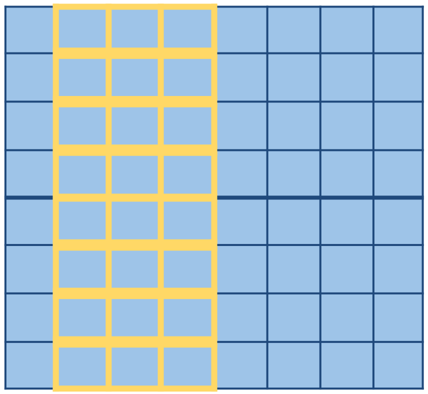
Convolutional Neural Networks

We “slide” our *filters* on the embeddings of the characters



Convolutional Neural Networks

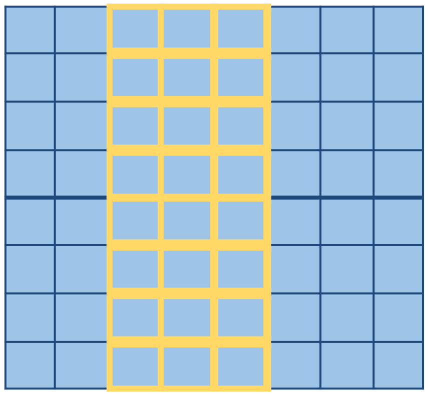
We "slide" our *filters* on the embeddings of the characters



D u i s b u r g

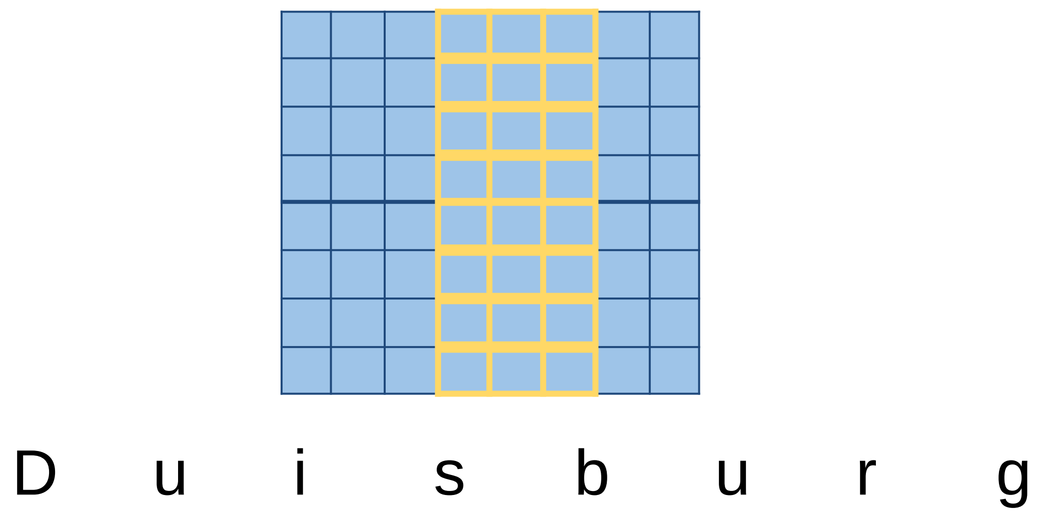
Convolutional Neural Networks

We “slide” our *filters* on the embeddings of the characters



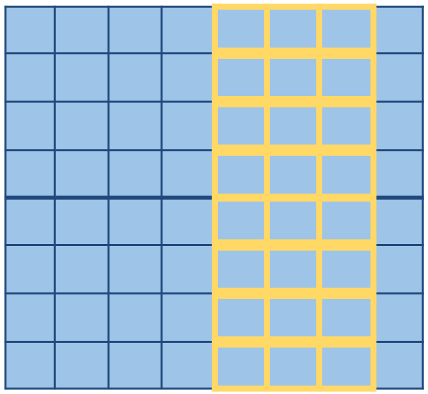
Convolutional Neural Networks

We "slide" our *filters* on the embeddings of the characters



Convolutional Neural Networks

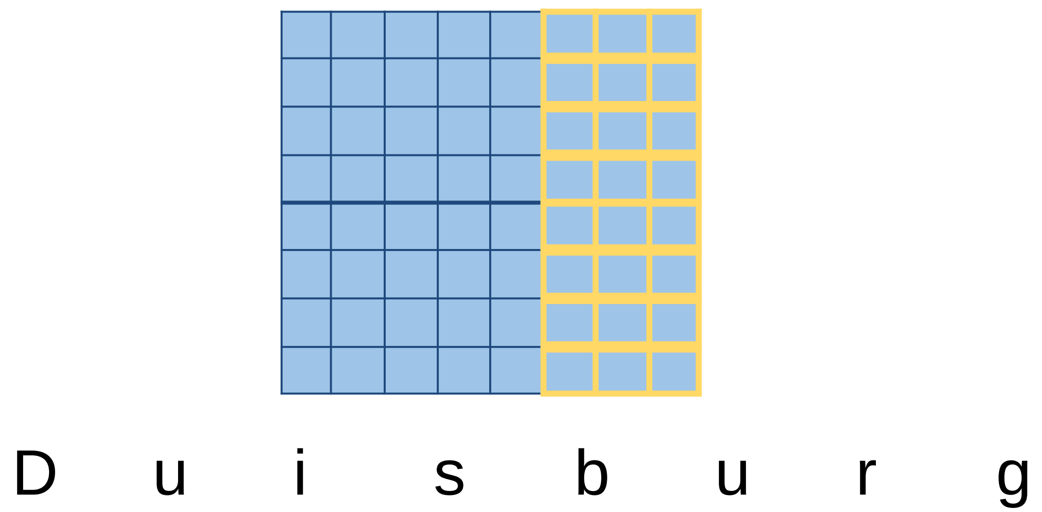
We “slide” our *filters* on the embeddings of the characters



D u i s b u r g

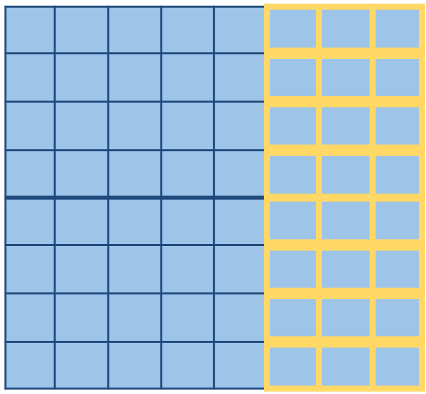
Convolutional Neural Networks

We "slide" our *filters* on the embeddings of the characters



Convolutional Neural Networks

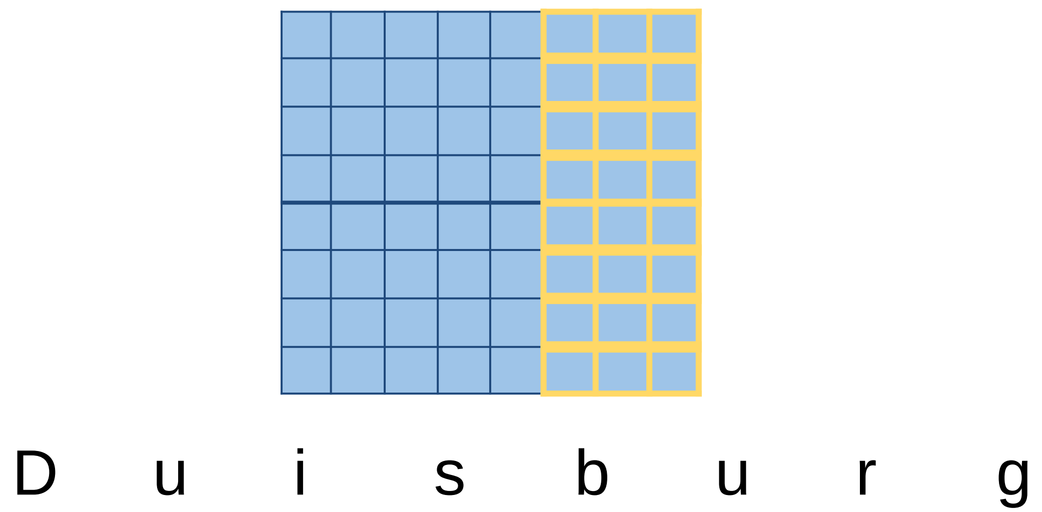
Filters are usually as tall as the embedding layer, and sliding is in one dimension only



D u i s b u r g

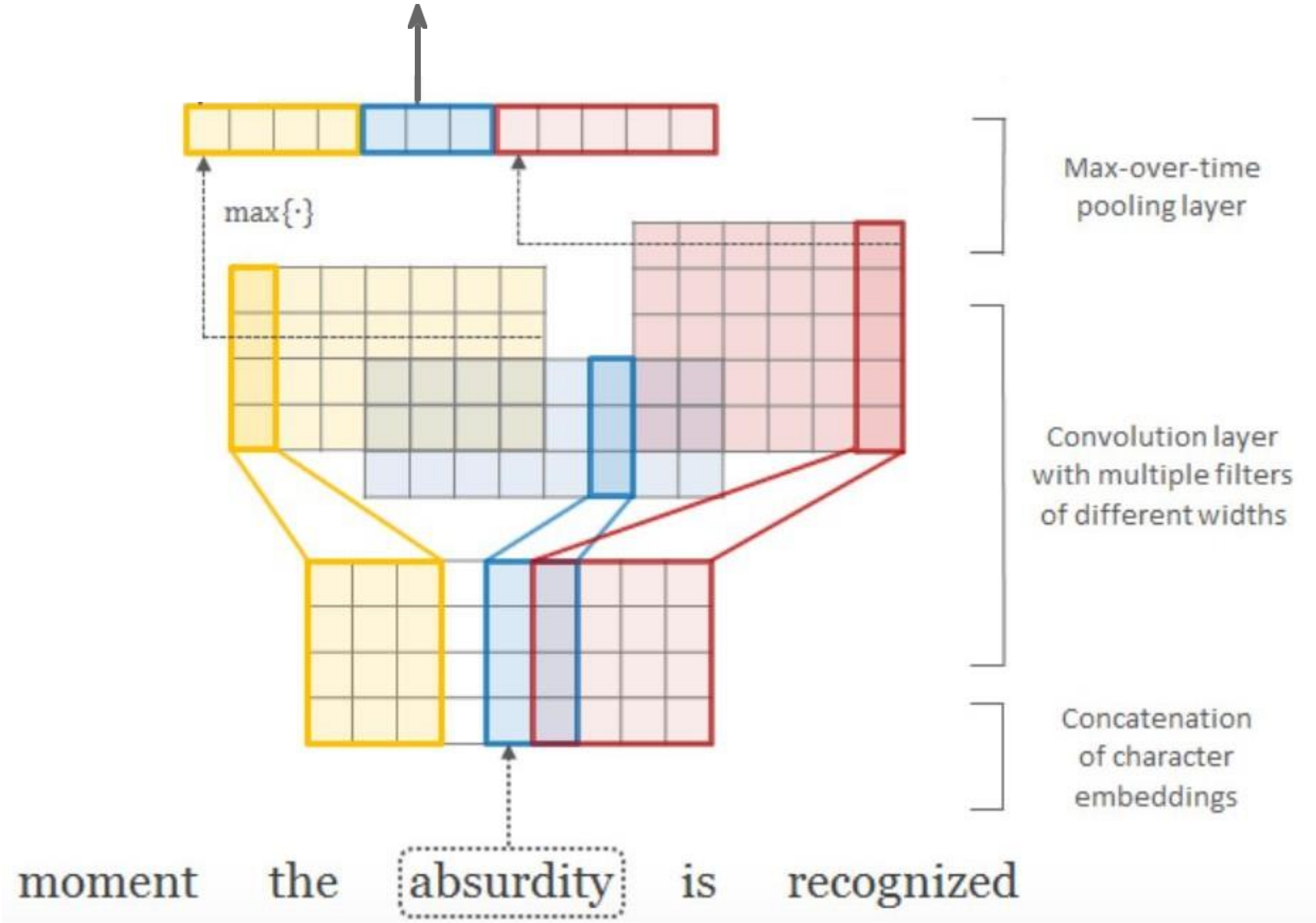
Convolutional Neural Networks

Filters are usually as tall as the embedding layer, and sliding is in one dimension only



The output in this case will be a 6-vector

Character CNN



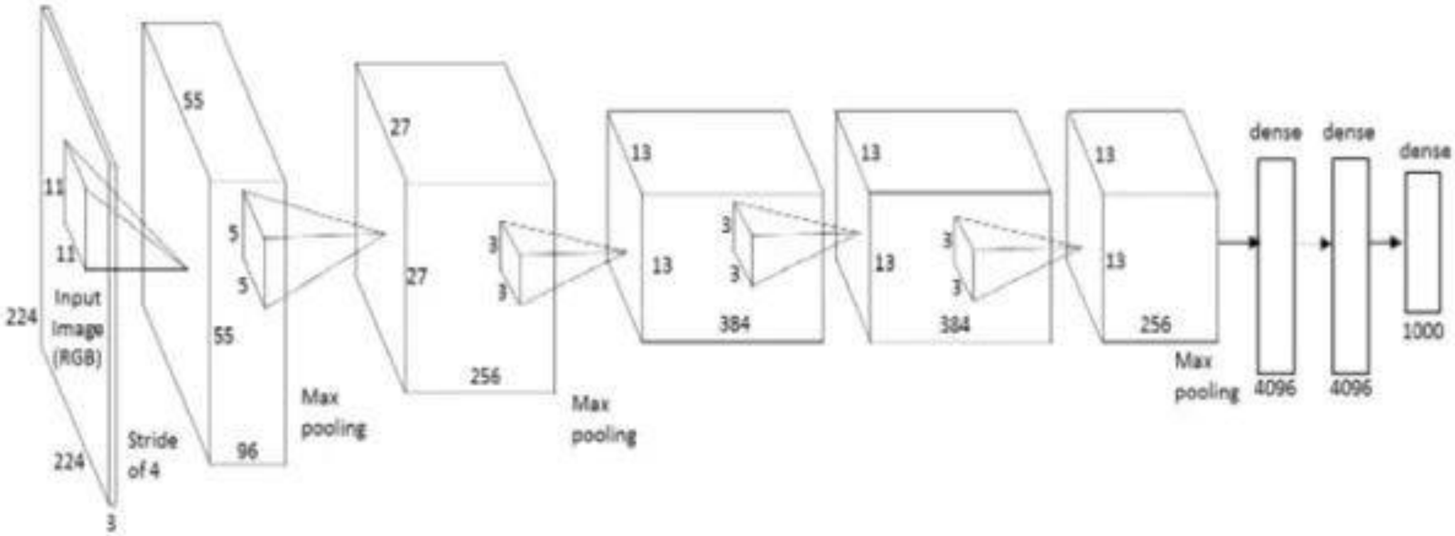
Kim, Jernite, Sontag, Rush, *Character-Aware Neural Language Model*, AAAI 2016

Convolutional Neural Networks

- ConvNets are quite popular in text at character level
- Used to build word embeddings for a larger network

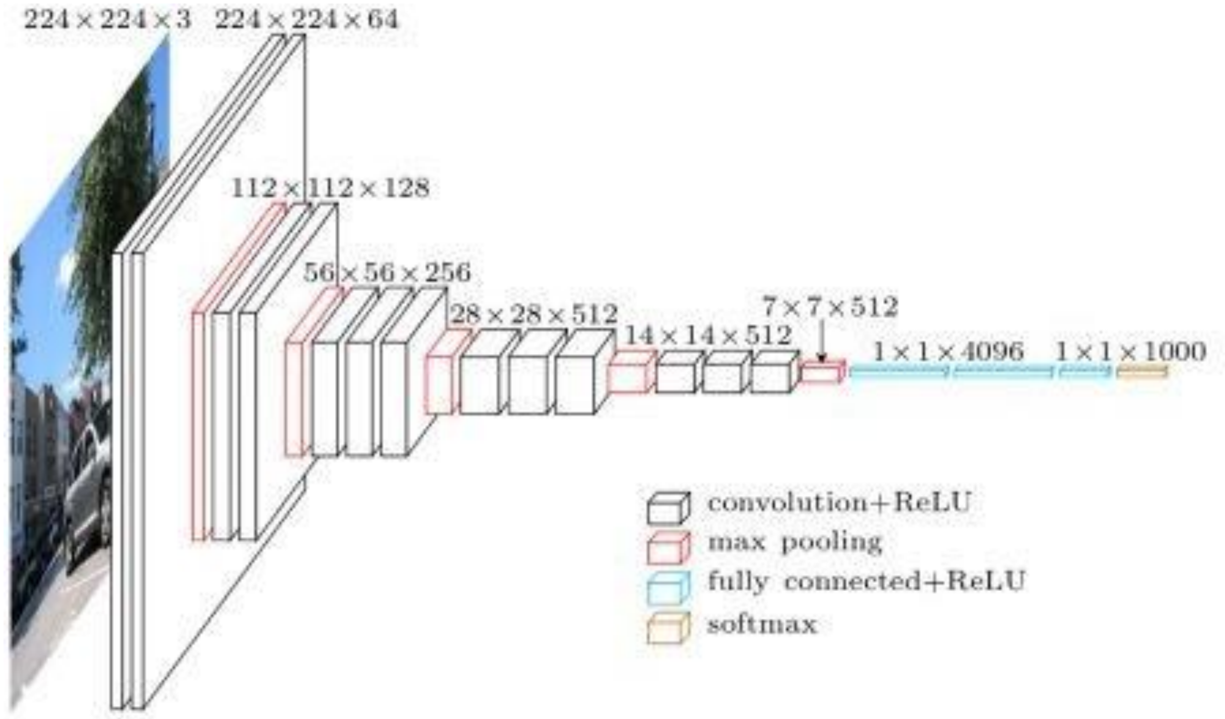
Convolutional Neural Networks

- Atleast in vision, some of the deepest architectures involve ConvNets



Convolutional Neural Networks

- Atleast in vision, some of the deepest architectures involve ConvNets



Convolutional Neural Networks

- Atleast in vision, some of the deepest architectures involve ConvNets

ResNet