The University of Queensland
AUSTRALIA
CREATE CHANGE

Pattern Analysis

Convolutional Neural Networks

Dr Shekhar “Shakes” Chandra
shekhar.chandra@uq.edu.au

V1.0

> “The desire to dominate is not correlated with intelligence, in fact, we have many examples of this... in the world. It's not the smartest of us that necessarily wants to be the chief.”
>
> Yann LeCun (link)
> (1960-)

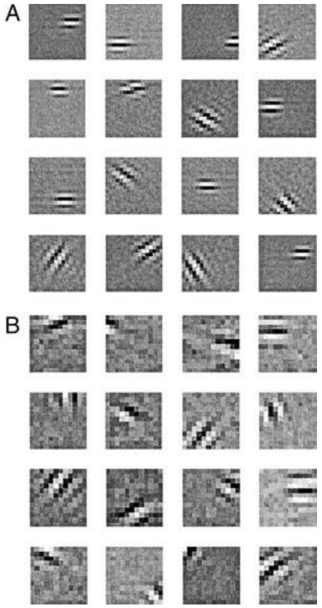
Receptive Fields

- Mammalian vision system uses response functions called receptive fields
- These fields are spread out across the visual inputs and been linked to Gabor filters [1] and controls the neural firing within the visual cortex
- We compute them using convolution

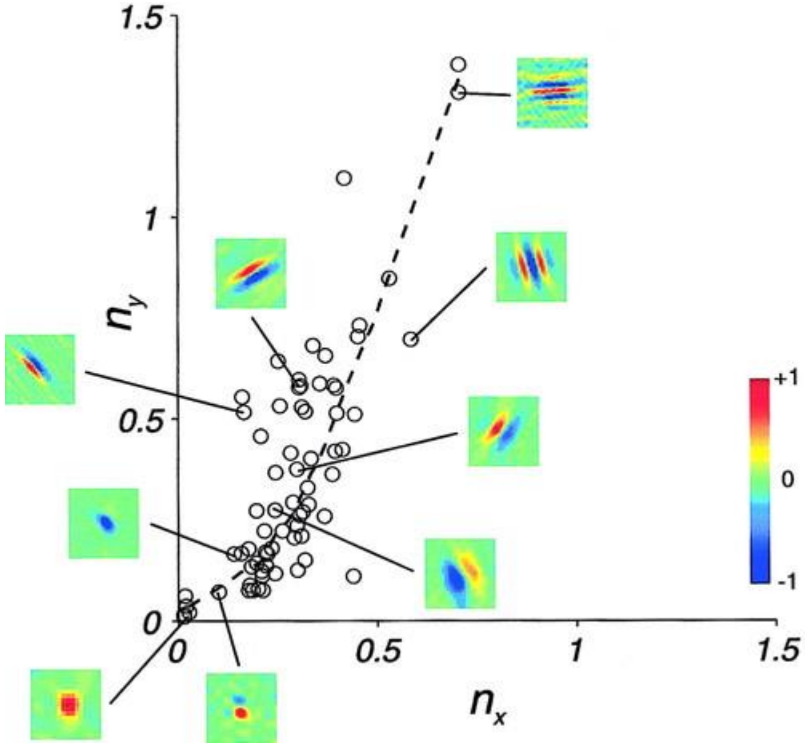
Modelled Field

Gabor Filters

Actual Field



Example of the receptive fields of a Macaque's primary visual cortex and their distribution with respect to spatial parameters [2]

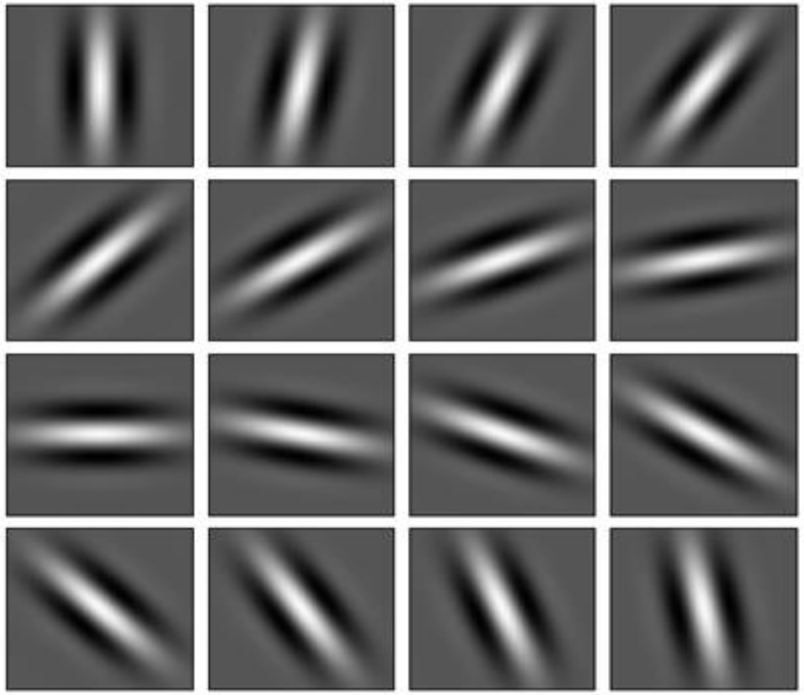


[1] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, p. 607, Jun. 1996.
[2] D. L. Ringach, "Spatial Structure and Symmetry of Simple-Cell Receptive Fields in Macaque Primary Visual Cortex," *Journal of Neurophysiology*, Jul. 2002.

https://medium.com/@anuj_shah/through-the-eyes-of-gabor-filter-17d1fdb3ac97

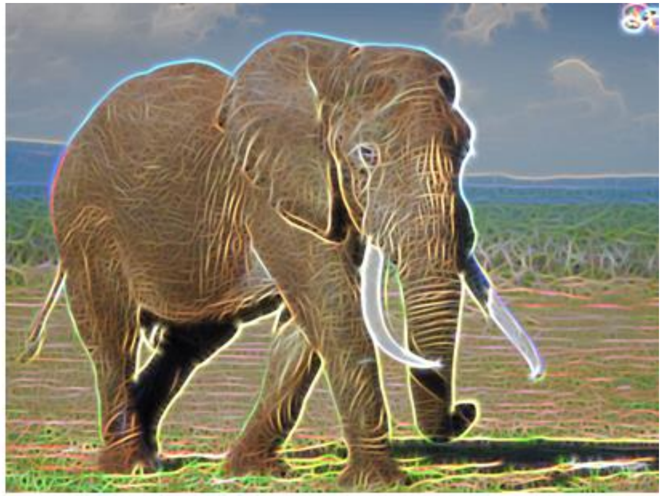
Receptive Fields

- We compute them using convolution



Gabor Filters

a

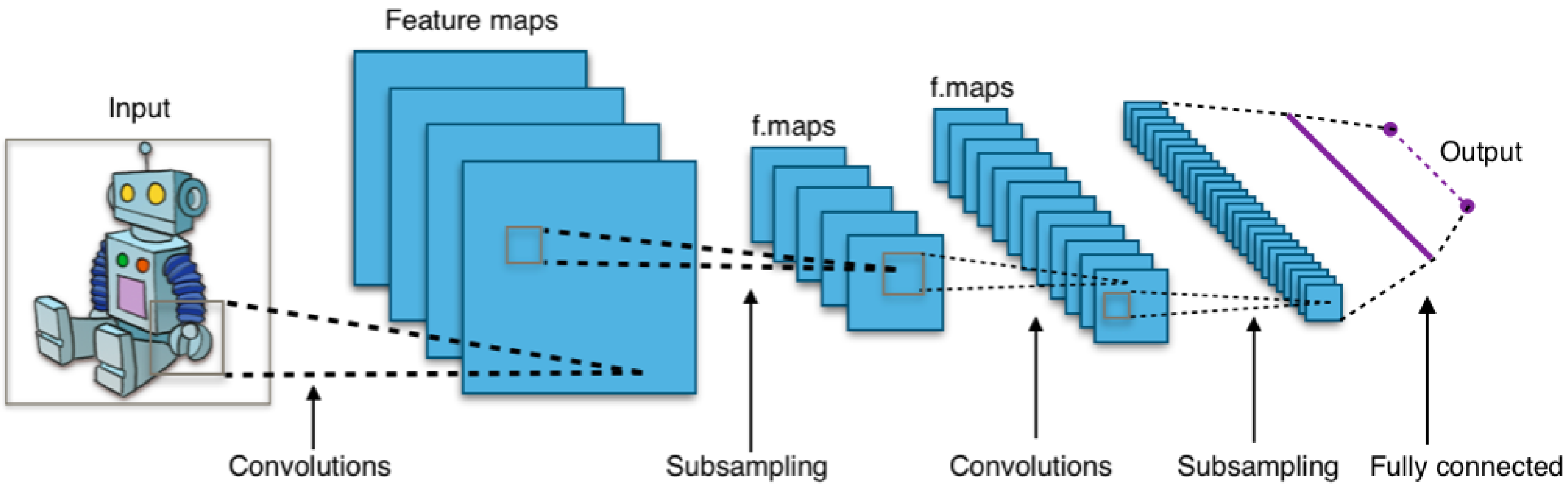


b

https://en.wikipedia.org/wiki/Convolutional_neural_network

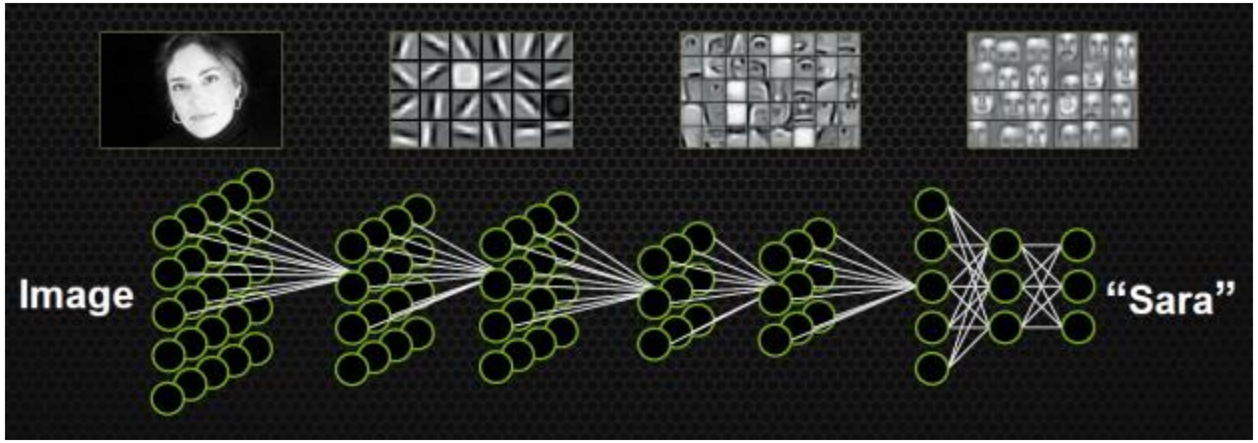
Convolution Neural Networks (CNNs)

- The CNN mimics the human vision system by computing a set of receptive fields across the visual inputs at multiple scales called layers.
- Responses from a previous layer with smaller scales are passed on to the subsequent layer with a larger scale by averaging or pooling the responses at this layer.
- In this way, the initial layers pass on smaller finer details for the higher layers to receive and respond.
- The higher layers respond to these pooled or aggregated responses resulting in responses to larger scale structures.



Convolution Neural Networks

https://devblogs.nvidia.com/accelerate-machine-learning-cudnn-deep-neural-network-library/



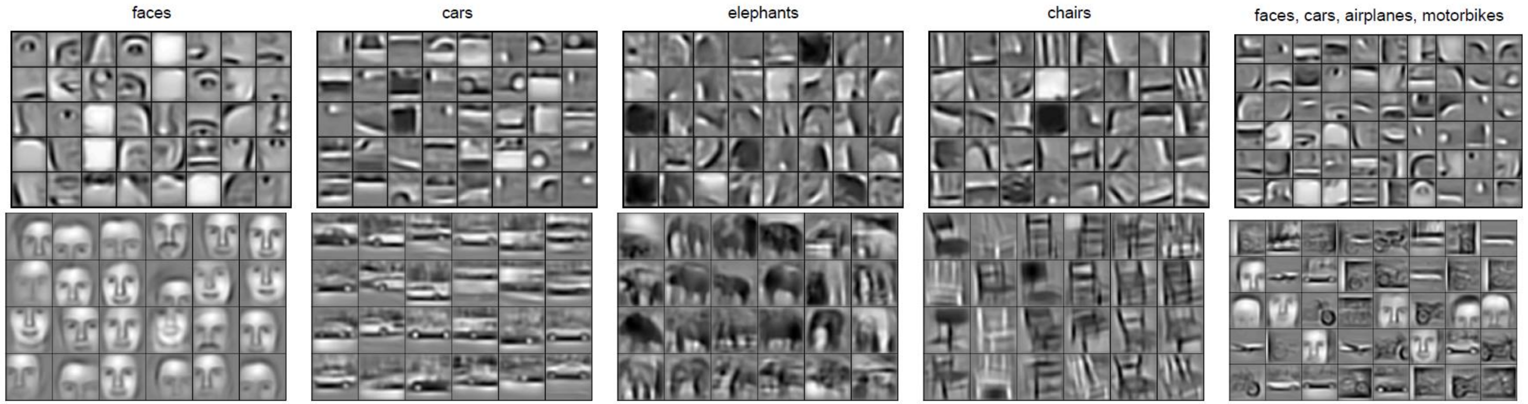
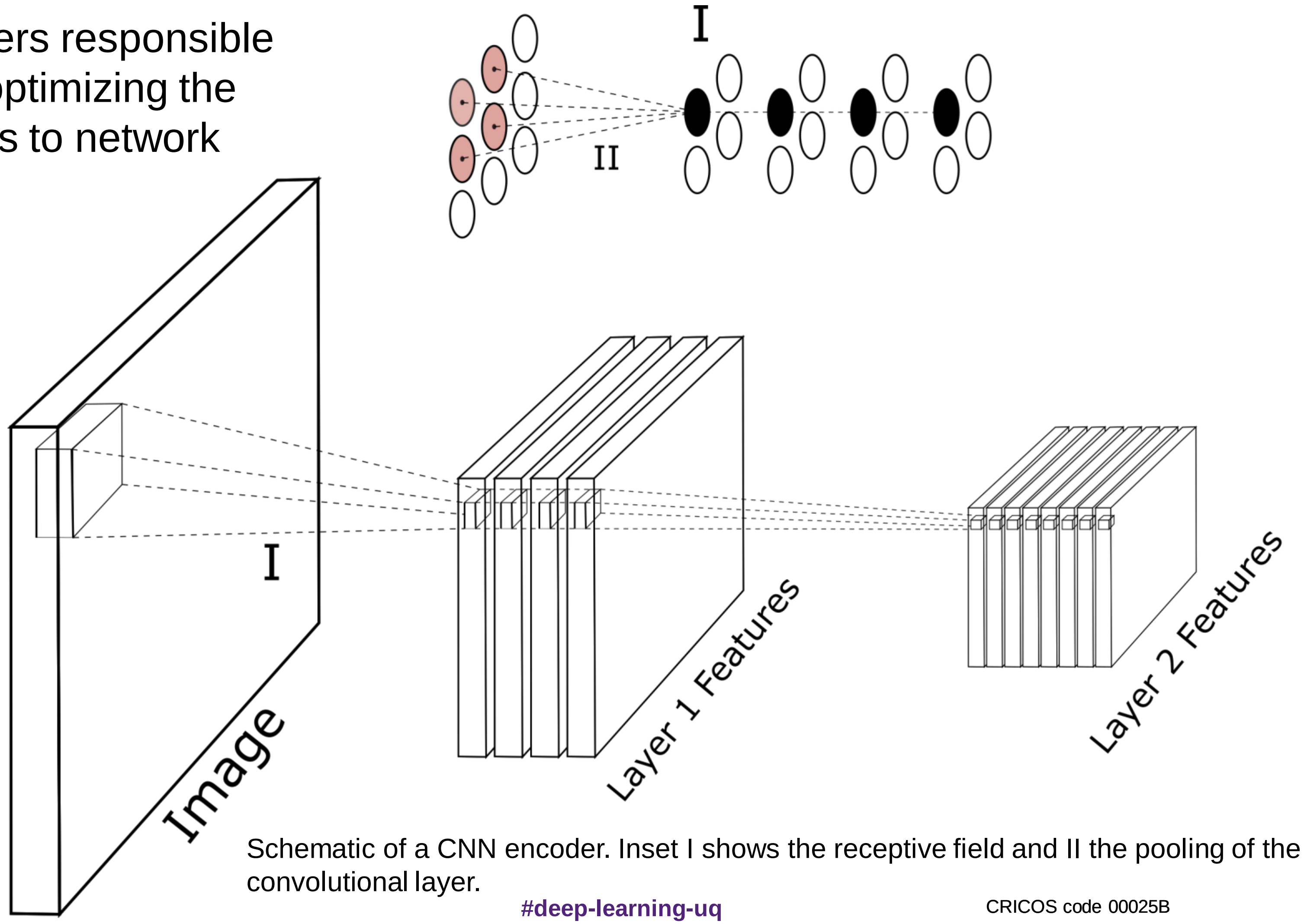


Figure 3. Columns 1-4: the second layer bases (top) and the third layer bases (bottom) learned from specific object categories. Column 5: the second layer bases (top) and the third layer bases (bottom) learned from a mixture of four object categories (faces, cars, airplanes, motorbikes).

Convolution Neural Networks

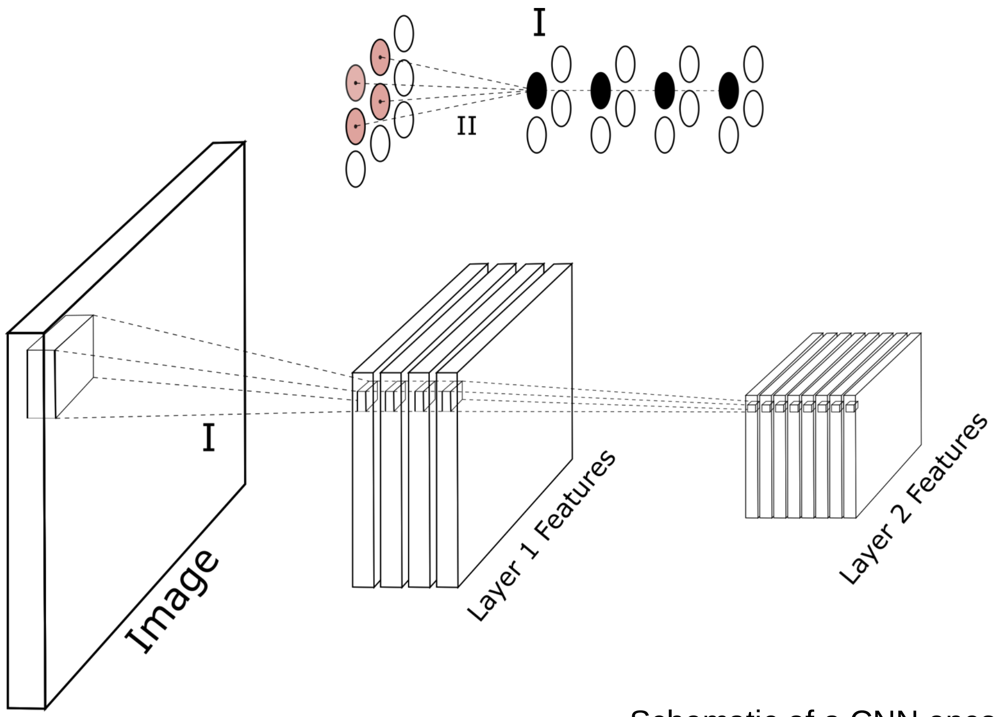
We learn the optimal filters responsible for our desired task by optimizing the receptive field responses to network input and output.



Schematic of a CNN encoder. Inset I shows the receptive field and II the pooling of the responses into the subsequent convolutional layer.

Convolution Neural Networks

We learn the optimal filters responsible for our desired task by optimizing the receptive field responses to network input and output.



- Compute the 2D Convolution

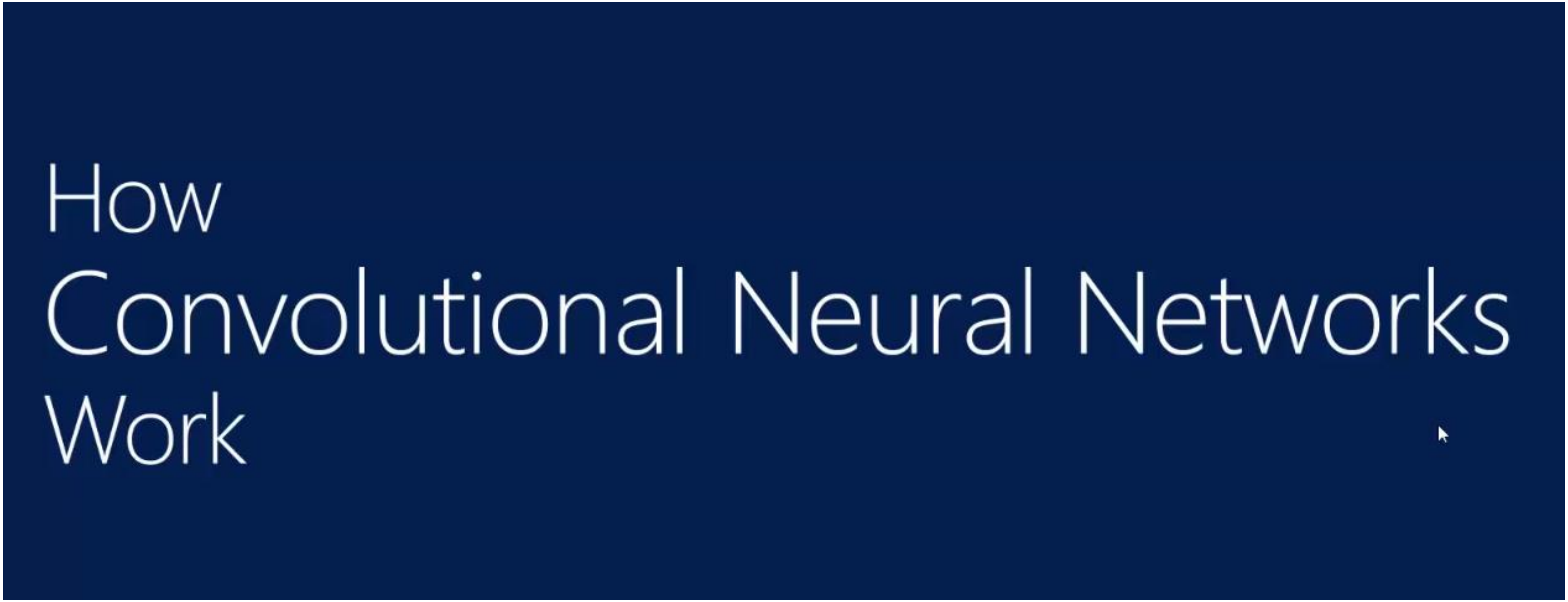
$$O(i,j) = \sum_{u=-2k-1}^{2k+1} \sum_{v=-2k-1}^{2k+1} F(u,v) I(i-u, j-v)$$

- Here F is the filter, I is the input image and O is the output
- We then compute these for a fixed number of filters to capture features called the depth D

$$O(i,j) = \sum_{d=1}^{D} \sum_{u=-2k-1}^{2k+1} \sum_{v=-2k-1}^{2k+1} F_d(u,v) I_d(i-u, j-v)$$

- The result then has the activation function applied to it

Schematic of a CNN encoder. Inset I shows the receptive field and II the pooling of the responses into the subsequent convolutional layer.


How
Convolutional Neural Networks
Work

Conclusion

- Convolution of Receptive fields model the mammalian vision system for natural images
- We can model this in a hierarchical fashion by using a multi-scale, layered set of convolutions with different filters to capture all relevant features in an image
- We can then determine the optimal filters for a given depth to have the network learn the vision task
- Linking different CNN network architectures give us tools to attack various recognition tasks

What’s Next?

We will find out what type of architectures and network designs can be used for various different pattern recognition tasks…

THE UNIVERSITY OF QUEENSLAND AUSTRALIA | CREATE CHANGE

Thank you

Dr Shekhar “Shakes” Chandra | Lecturer
School of Information Technology and Electrical Engineering
shekhar.chandra@uq.edu.au
+61 7 3365 8359

@shakes76
shekhar-s-chandra