

x1

1

w11

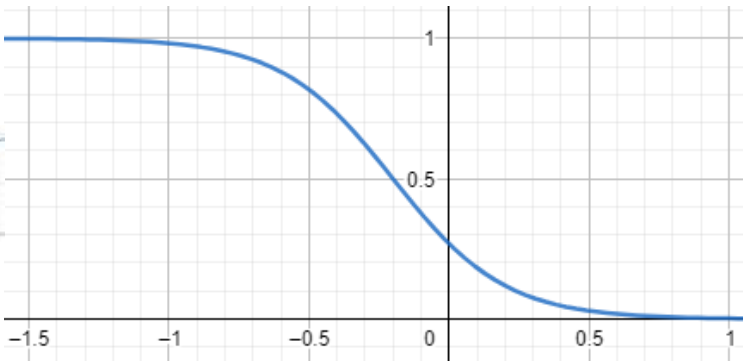
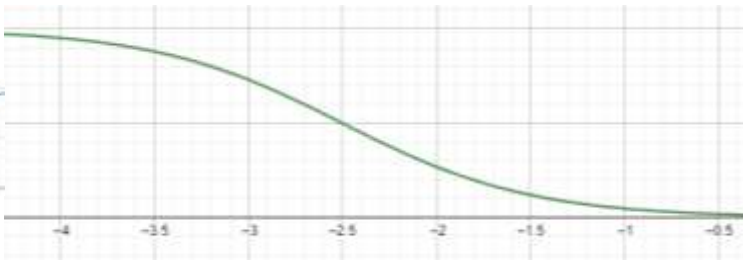
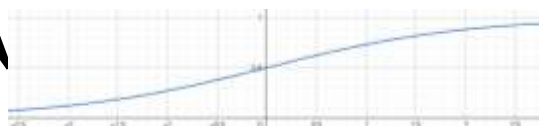
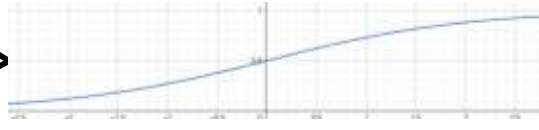
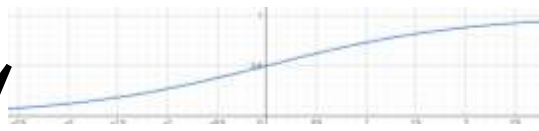
w12

w21

w22

w31

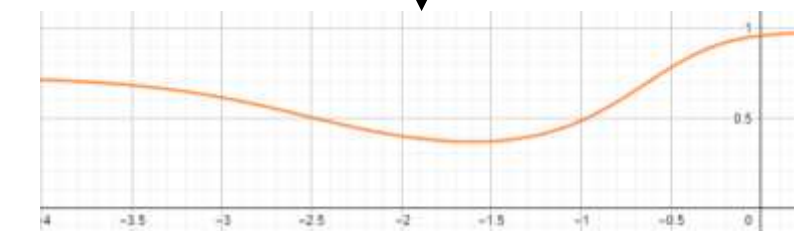
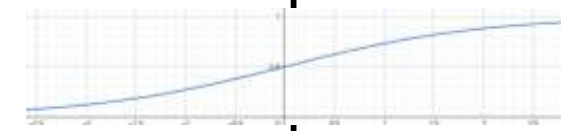
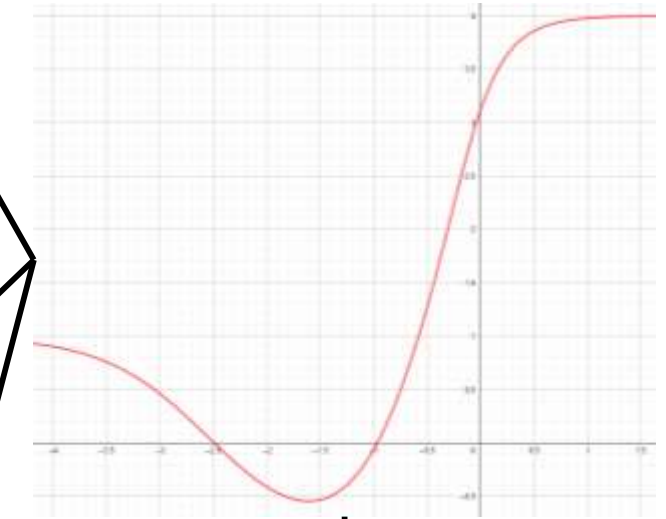
w32

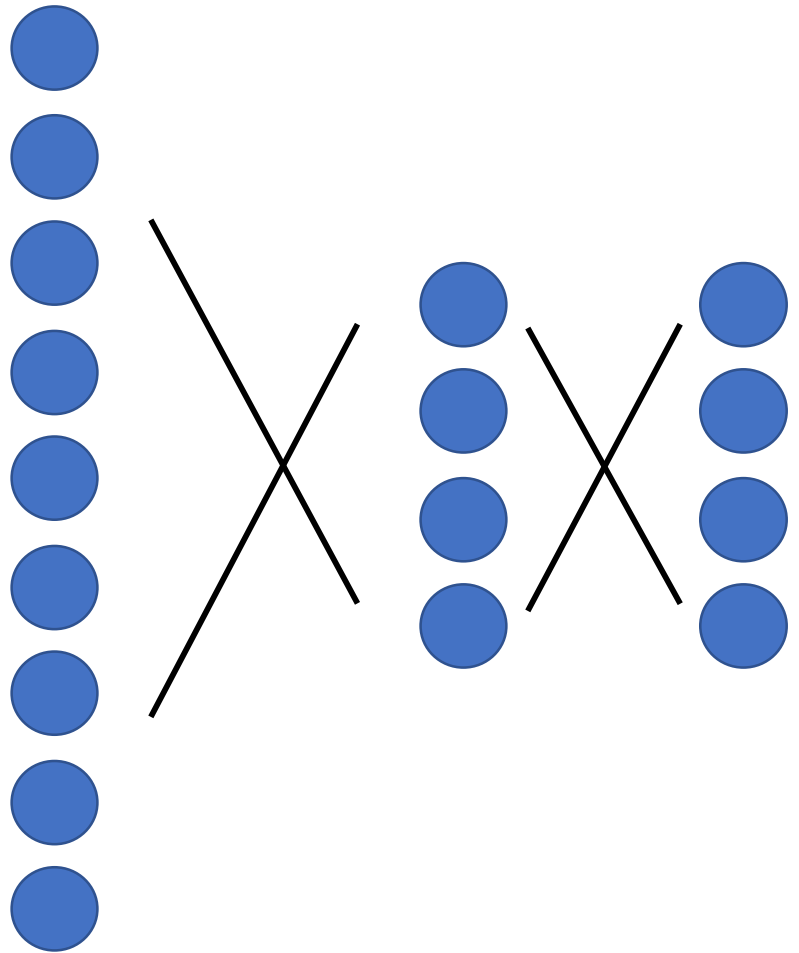
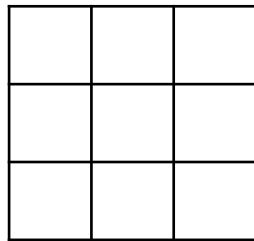
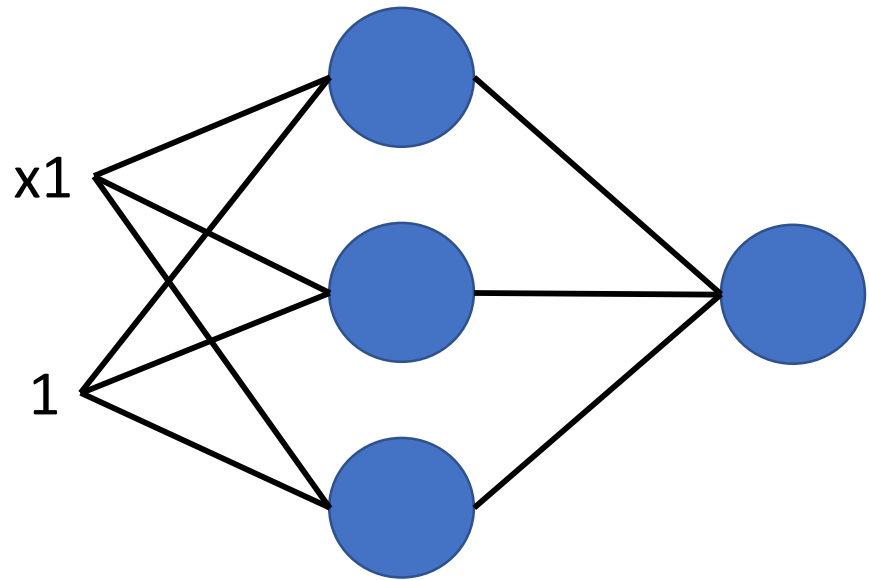


W1

W2

W3







https://twitter.com/robertmaguire_/status/1056365216009740289



右鉤拳柴犬



肌肉狐狸

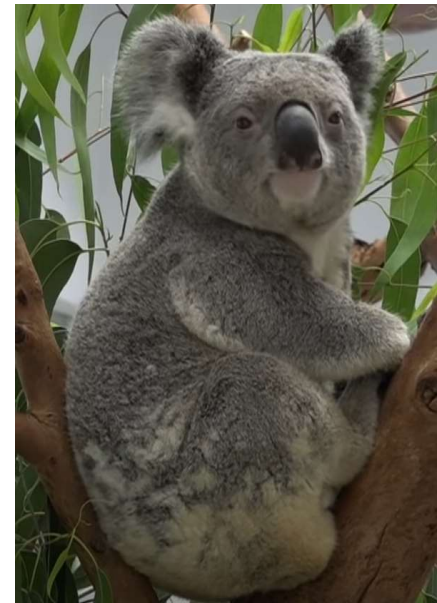


<https://www.facebook.com/kiwildlifepark/photos/a.695281337188191/4108311802551777/>

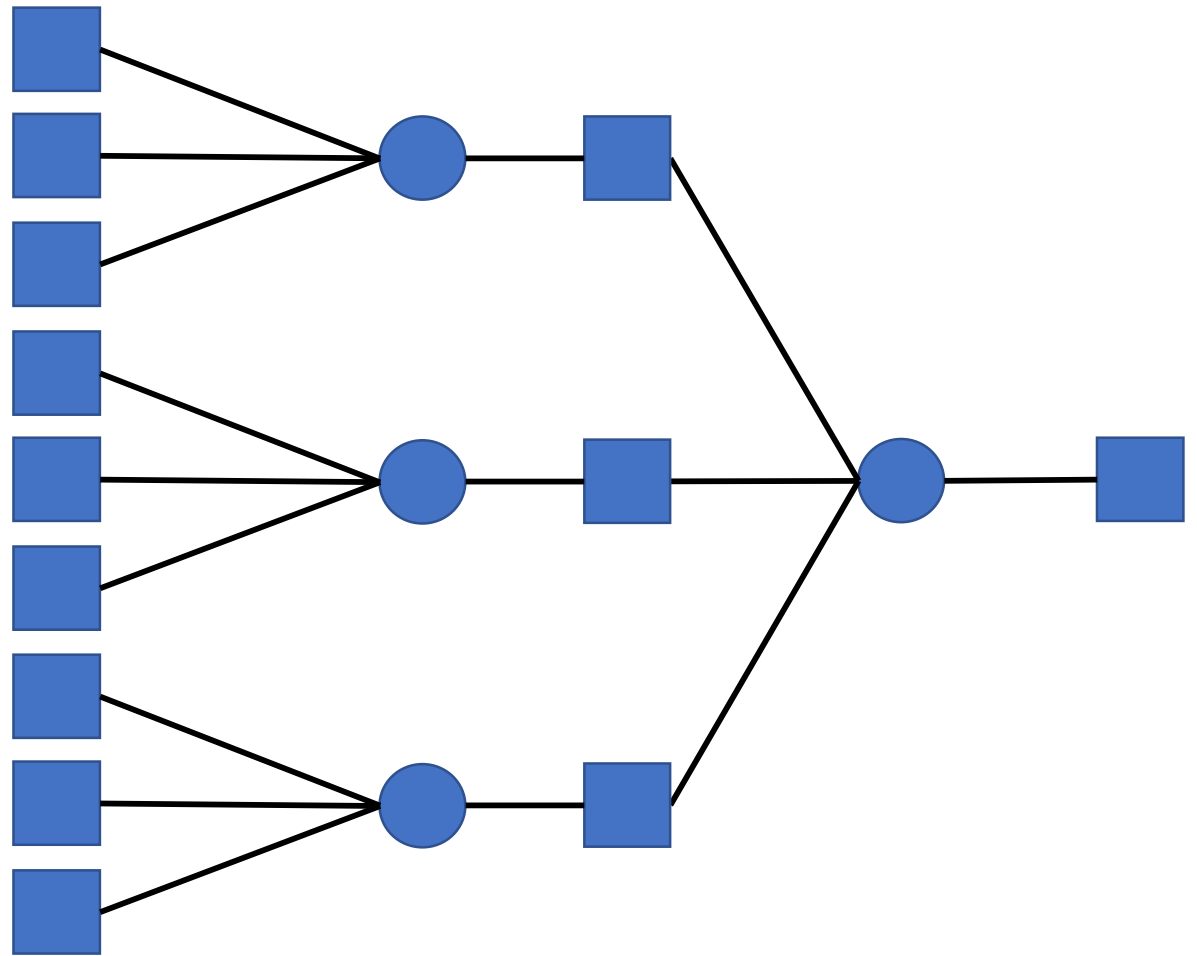
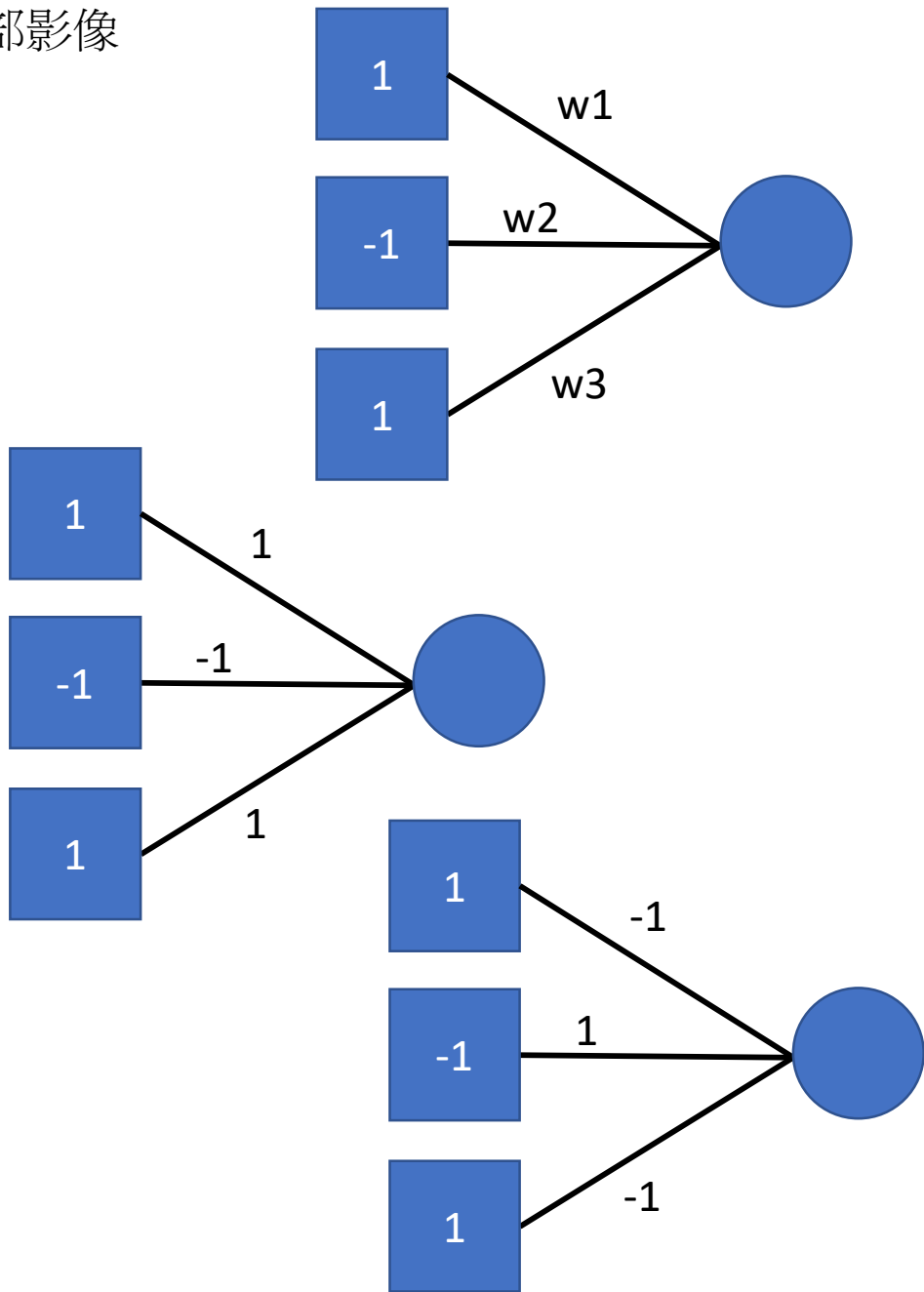
局部影像
不同位置
縮放



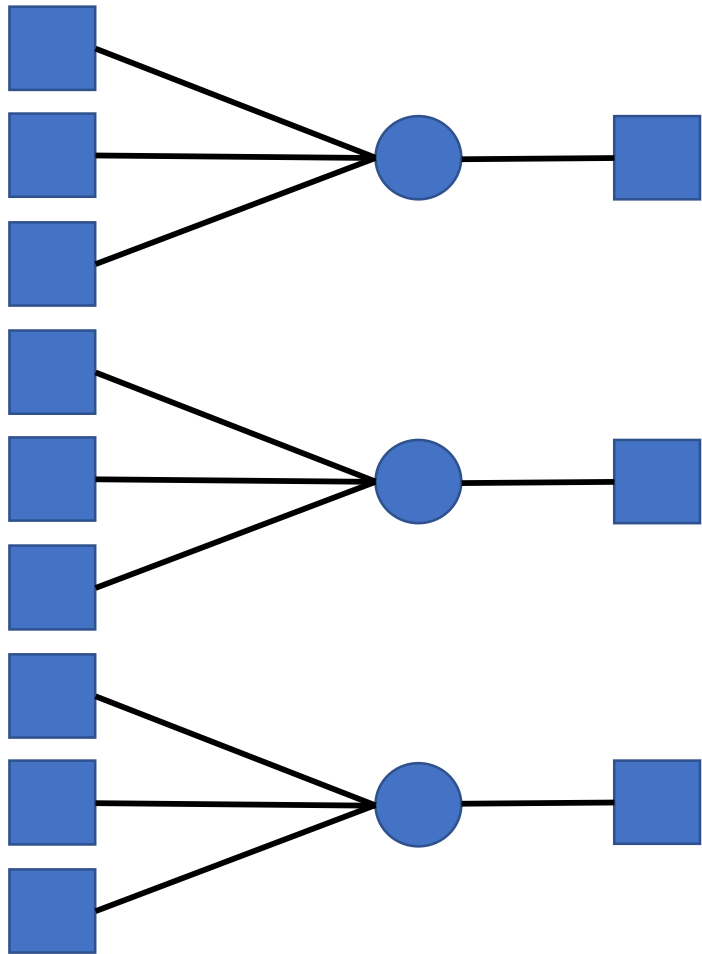
臺北市立動物園 <https://www.youtube.com/watch?v=3MNQTVDCc8E>



局部影像

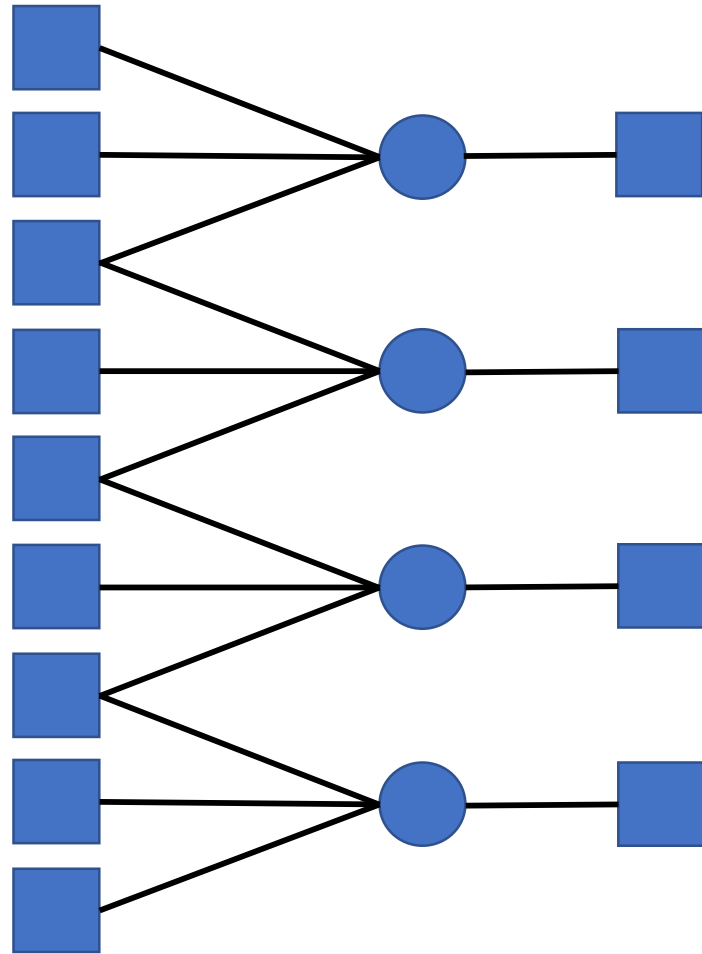


局部影像



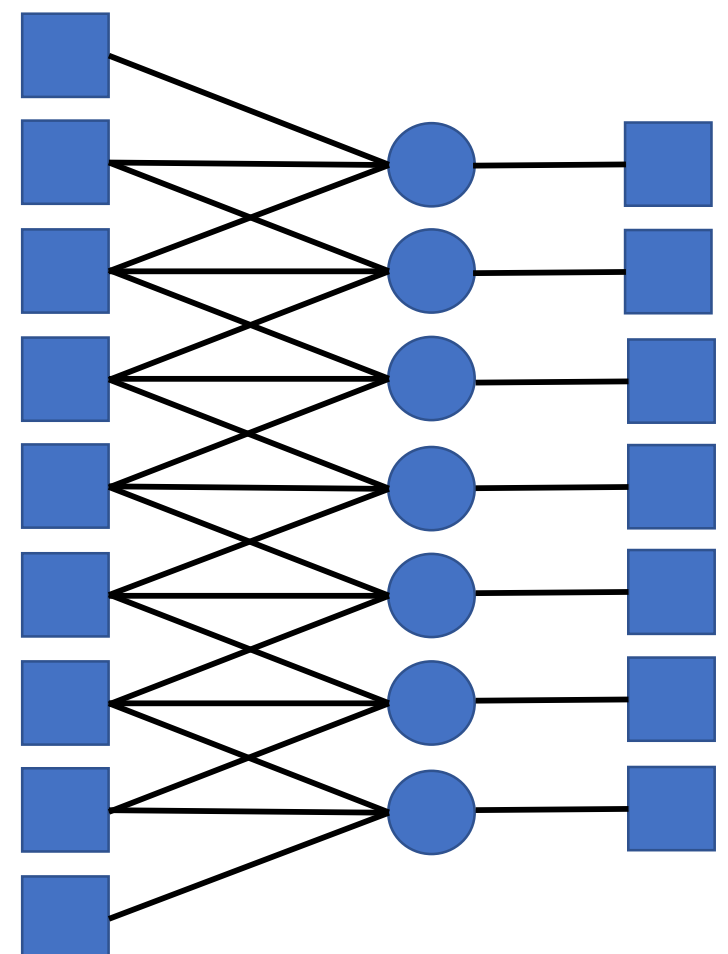
9

Stride = 3



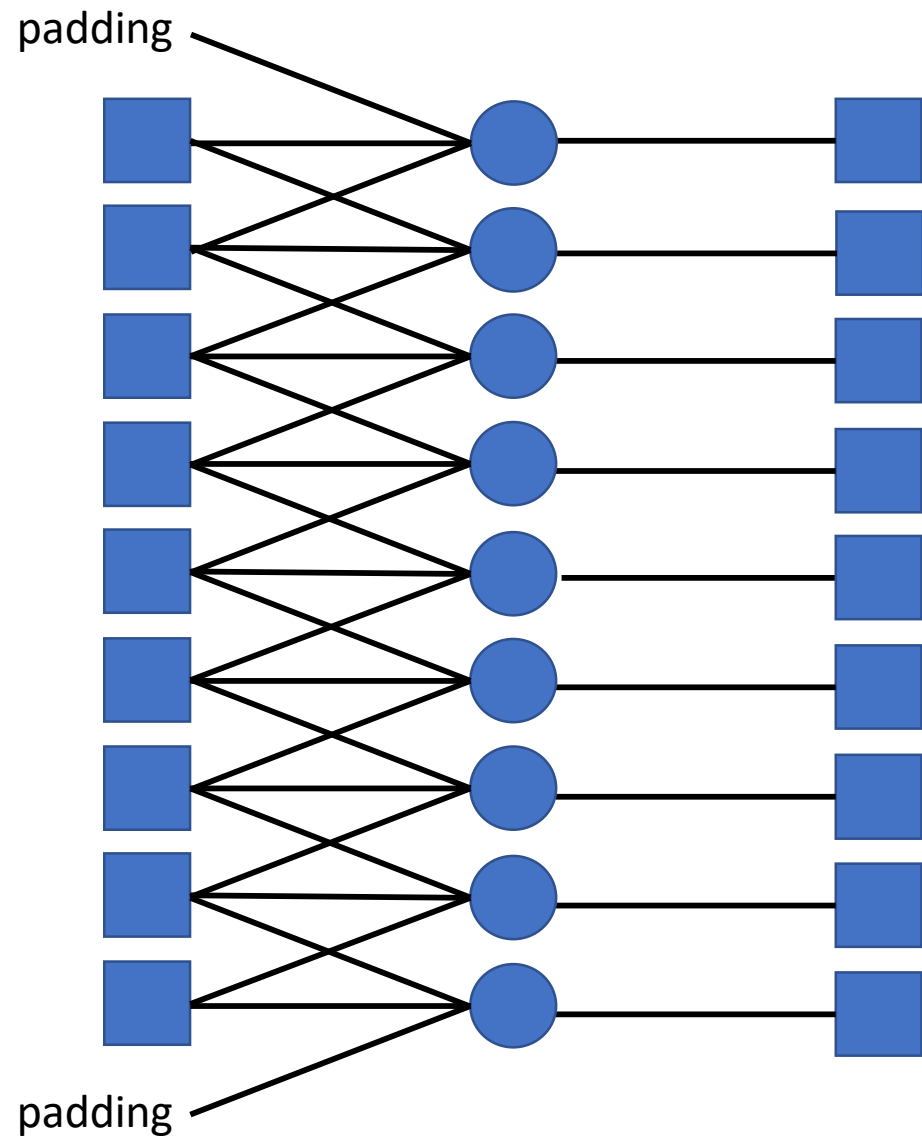
12

Stride = 2

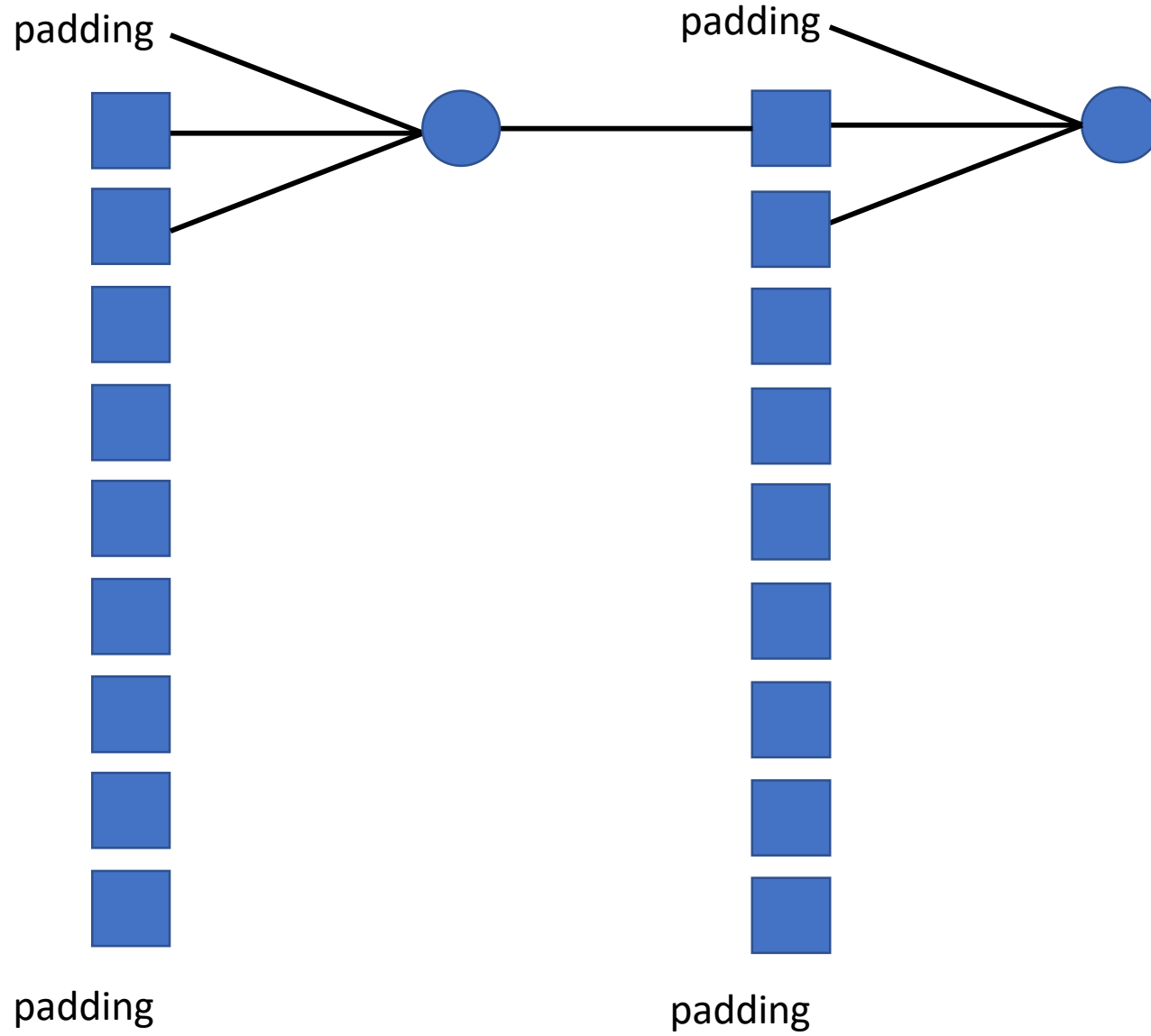


21

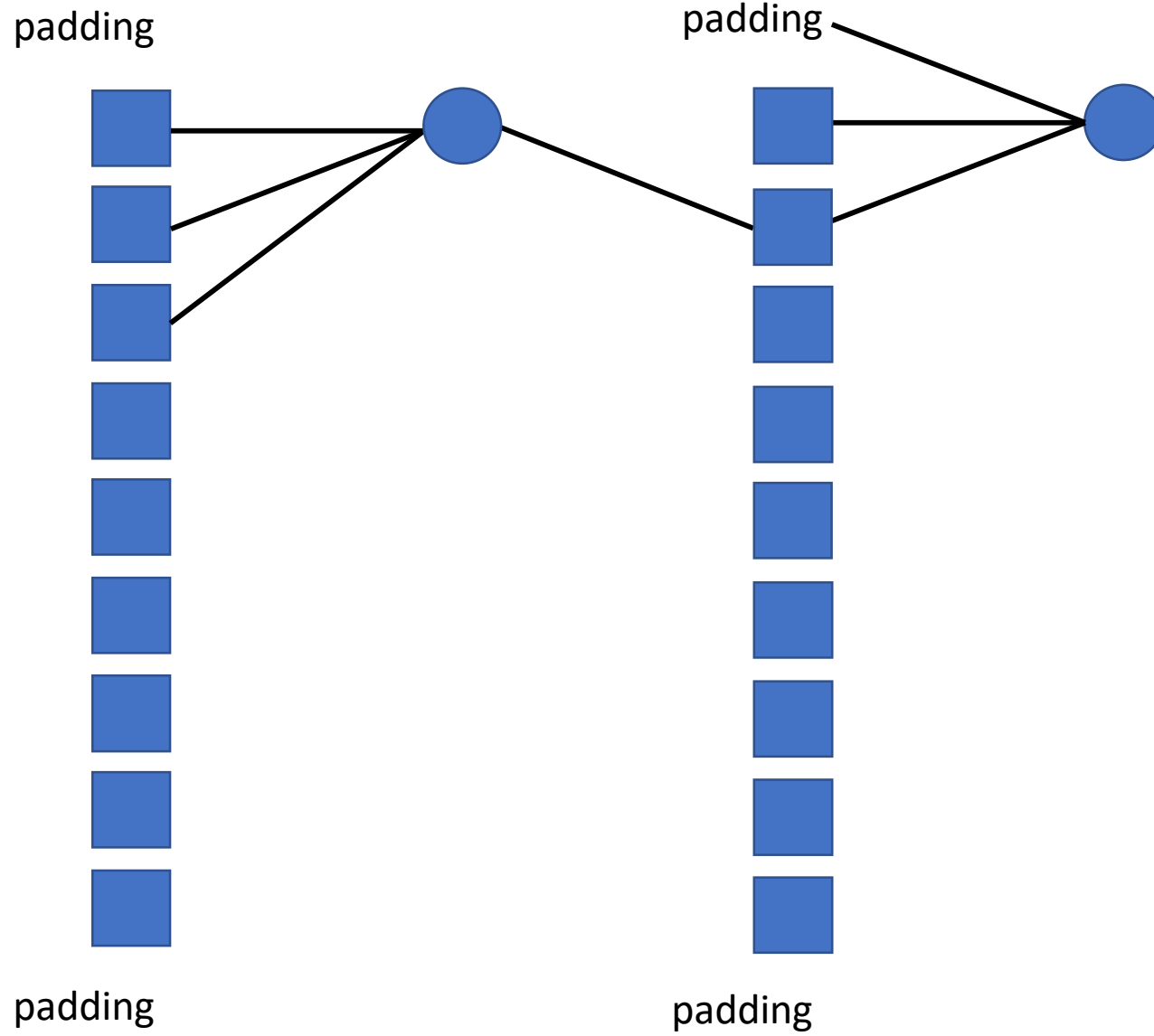
Stride = 1



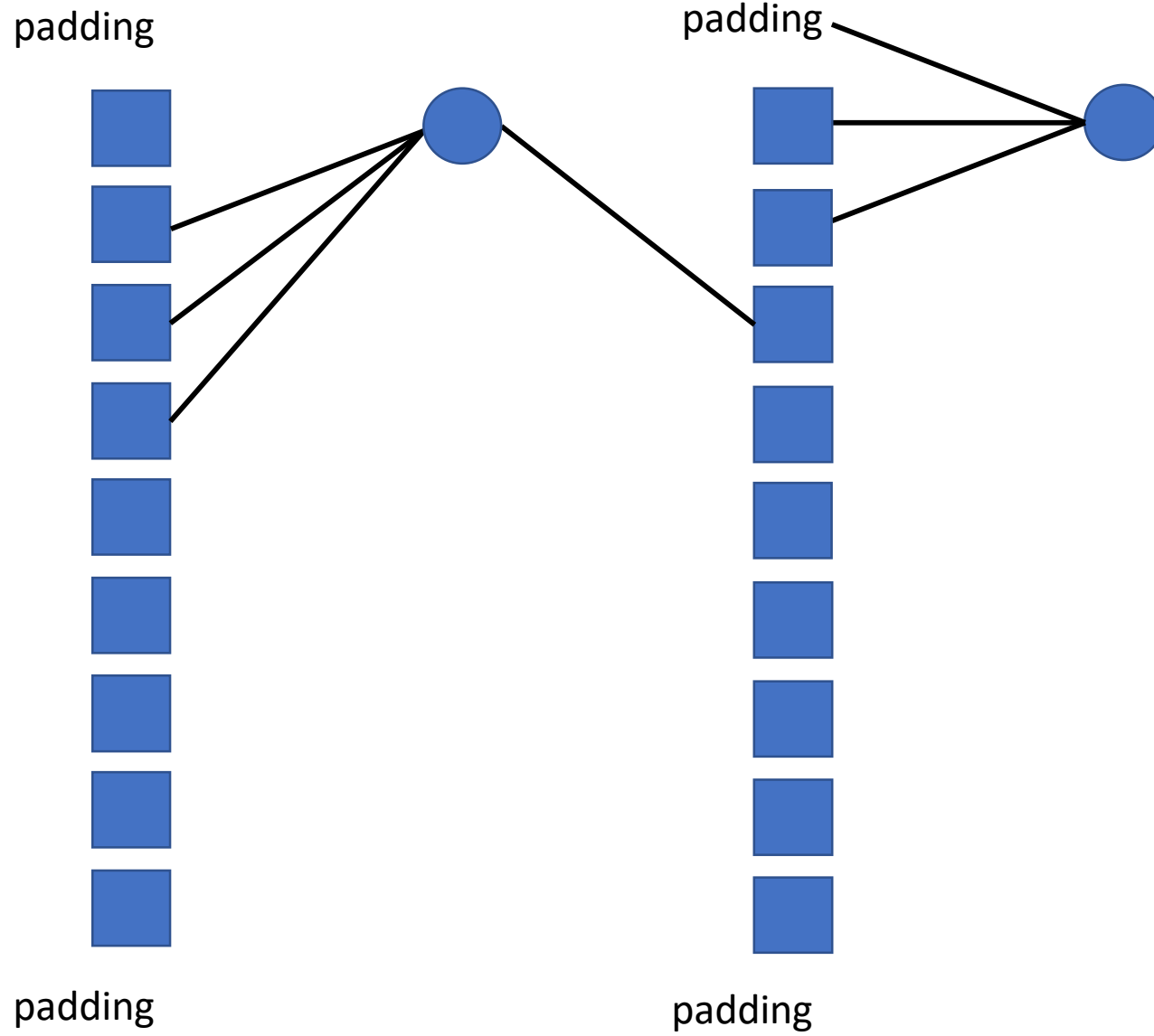
不同位置

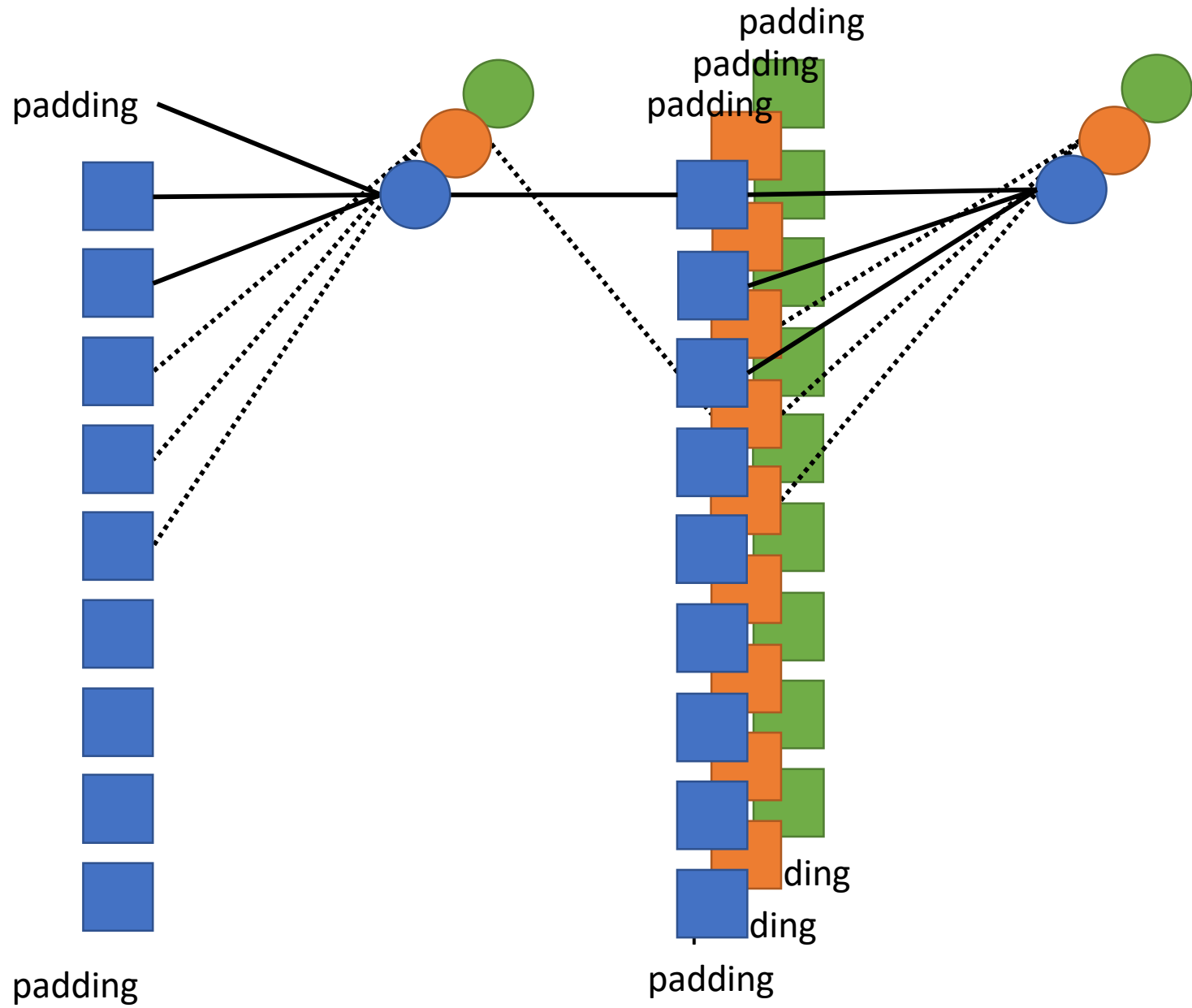


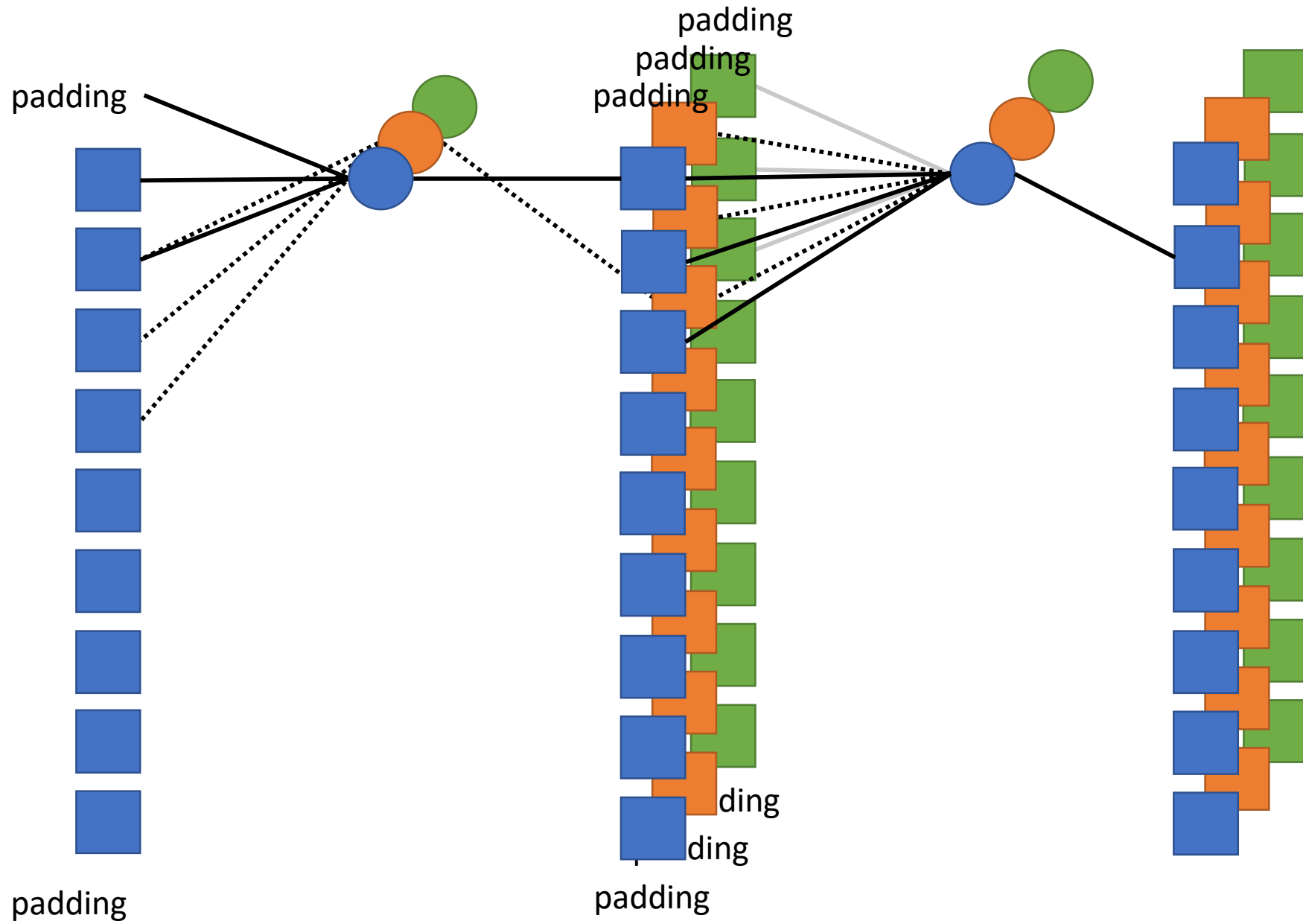
不同位置

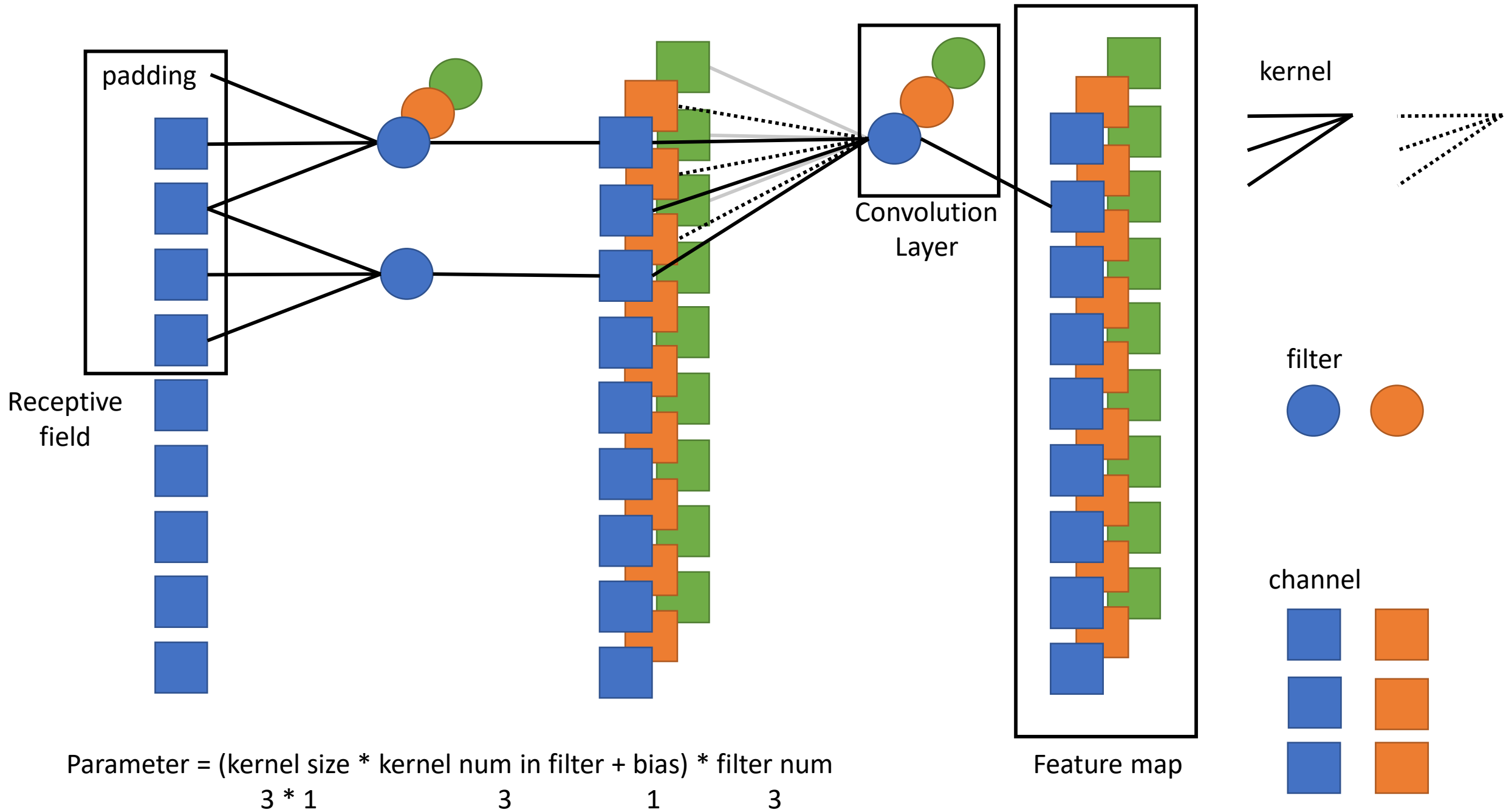


不同位置



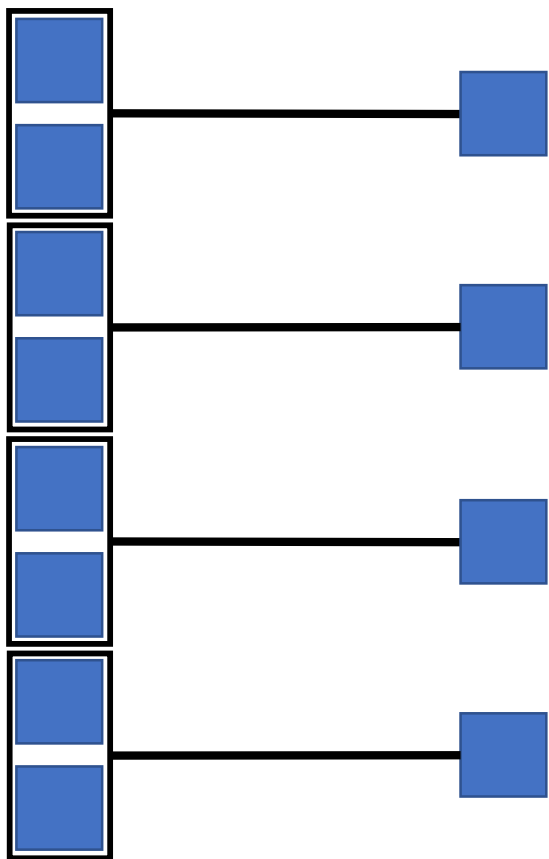




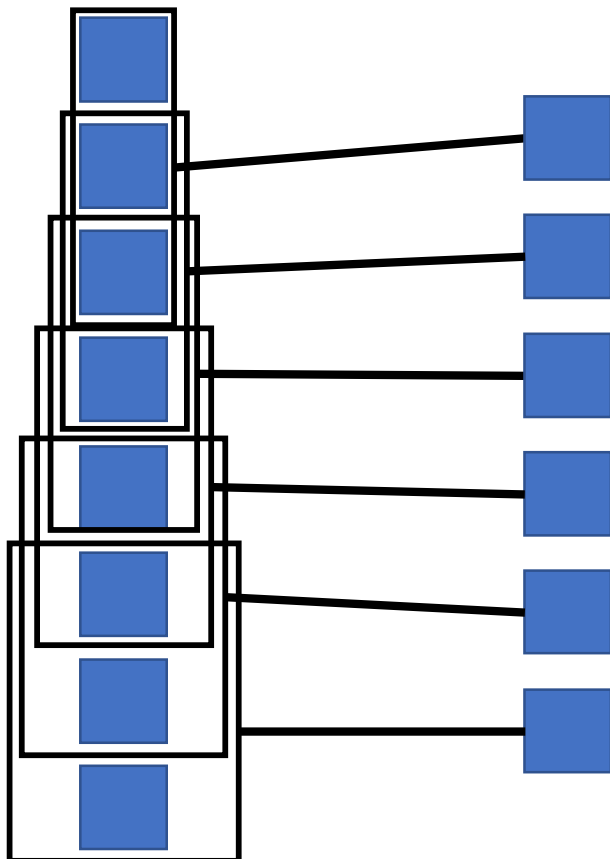


縮放

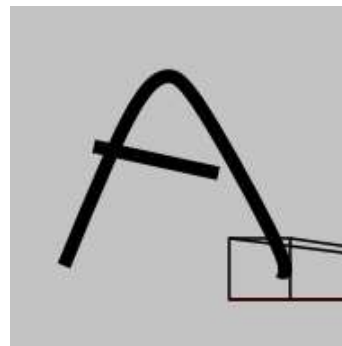
Max, Average



Size = 2
Stride = 2

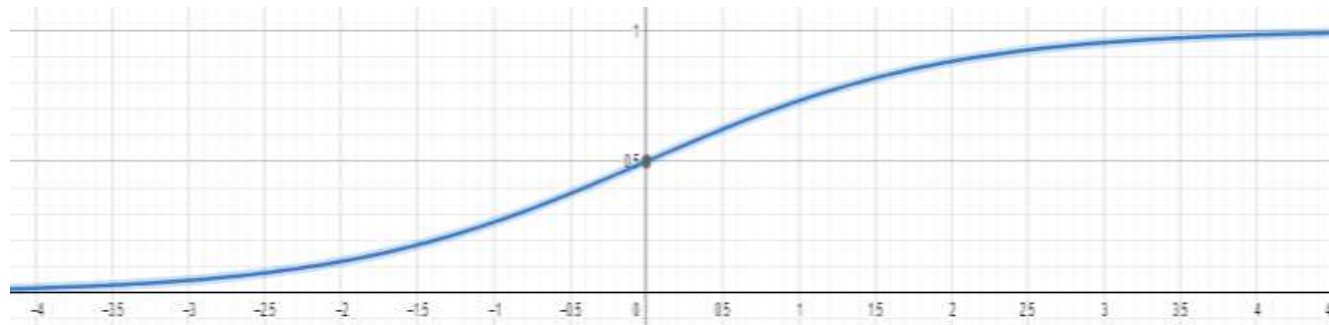
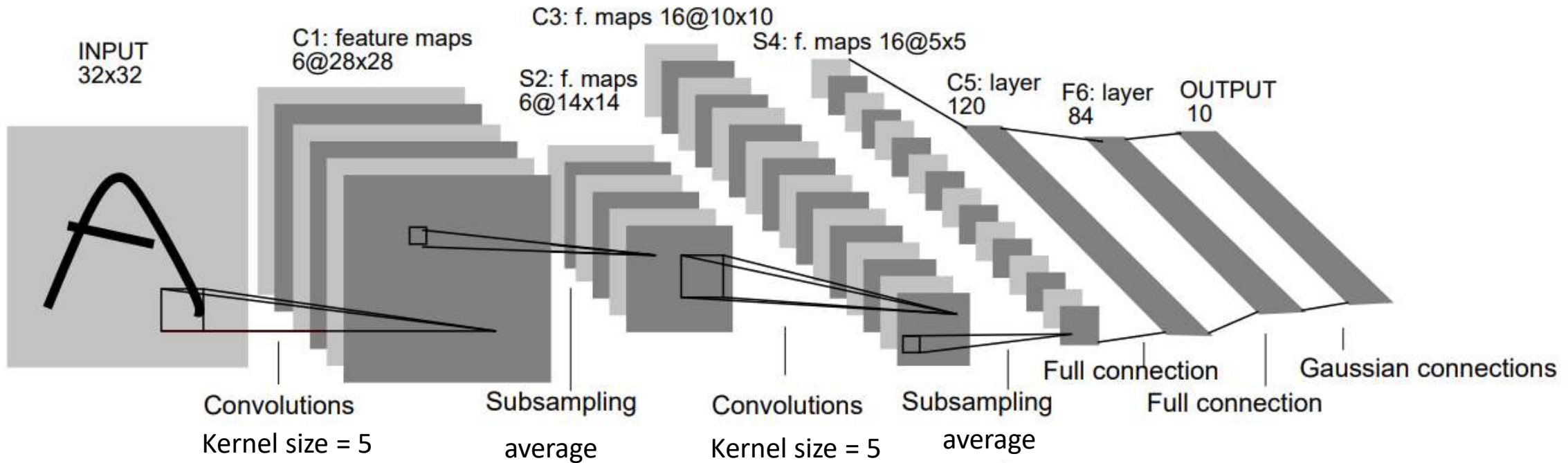


Size = 3
Stride = 1



Gradient-Based Learning Applied to Document Recognition

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner



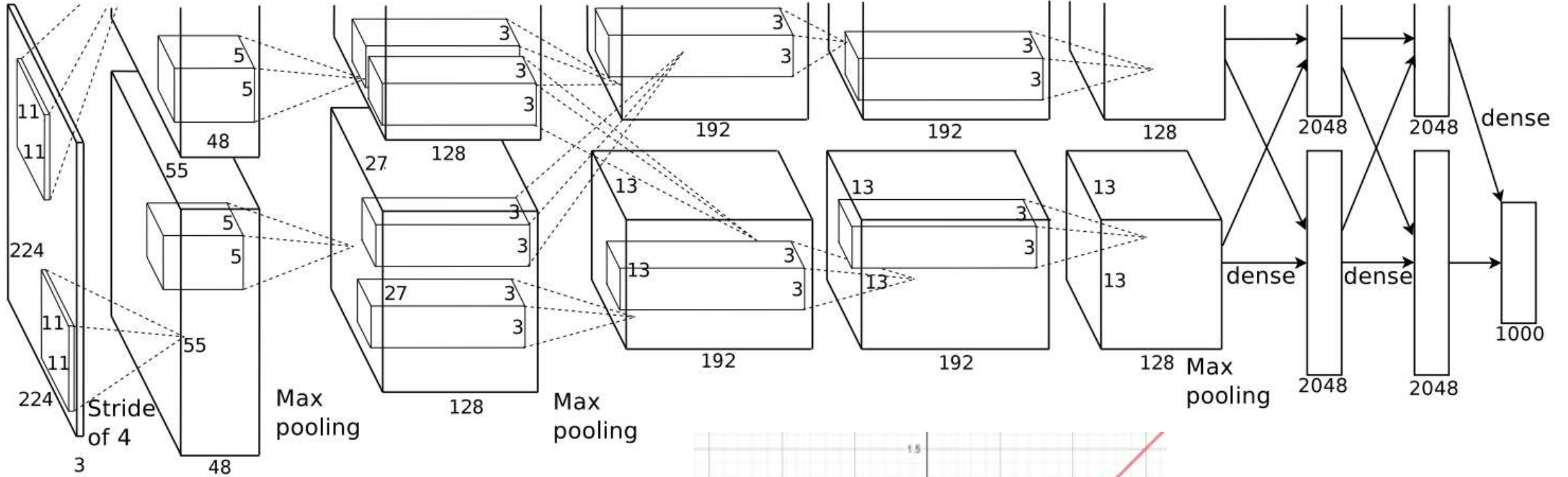
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3		X	X	X			X	X	X	X			X		X	X
4			X	X	X			X	X	X	X		X	X		X
5				X	X	X			X	X	X	X		X	X	X

ImageNet Classification with Deep Convolutional Neural Networks

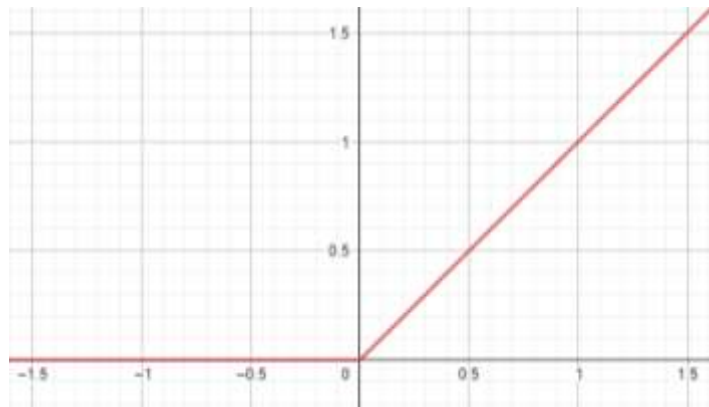
Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

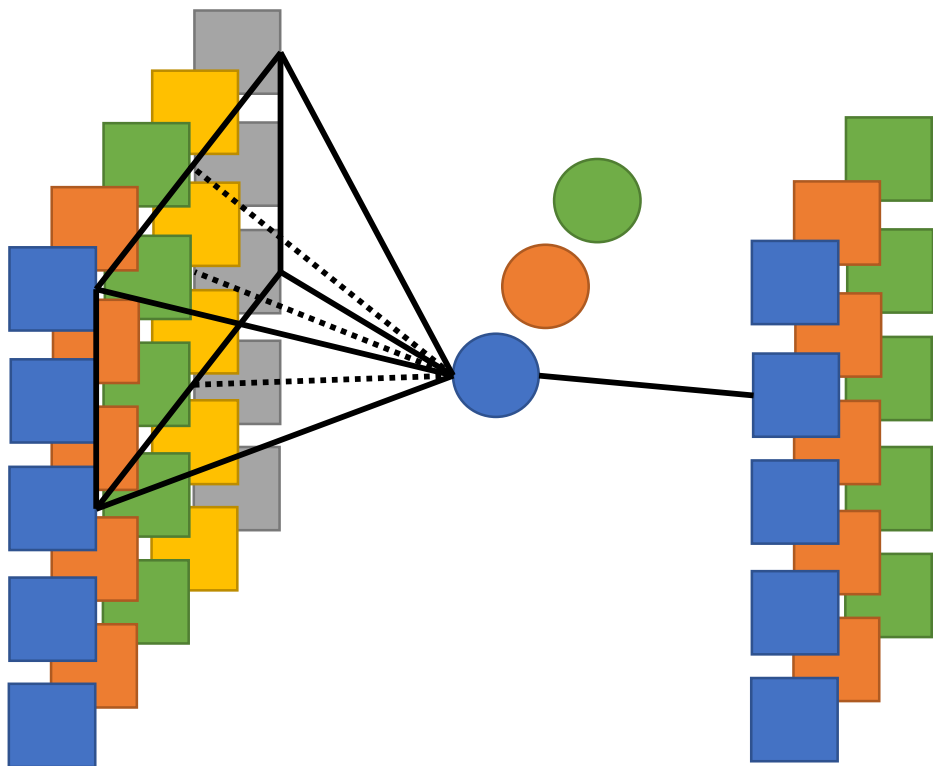
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca



- Relu
- Max pooling
- Overlapping pooling
- Dropout
- Data augmentation



Local Response Normalization



$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{j=\max(0, i-n/2)}^{\min(N-1, i+n/2)} (a_{x,y}^j)^2 \right)^\beta$$

$k = 2, n = 5, \alpha = 10^{-4},$ and $\beta = 0.75.$

mean activity. Response normalization reduces our top-1 and top-5 error rates by 1.4% and 1.2%, respectively. We also verified the effectiveness of this scheme on the CIFAR-10 dataset: a four-layer CNN achieved a 13% test error rate without normalization and 11% with normalization³.

VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION

Karen Simonyan* & Andrew Zisserman⁺

Visual Geometry Group, Department of Engineering Science, University of Oxford

{karen, az}@robots.ox.ac.uk

All hidden layers are equipped with the rectification (ReLU (Krizhevsky et al., 2012)) non-linearity. We note that none of our networks (except for one) contain Local Response Normalisation (LRN) normalisation (Krizhevsky et al., 2012): as will be shown in Sect. 4, such normalisation does not improve the performance on the ILSVRC dataset, but leads to increased memory consumption and computation time. Where applicable, the parameters for the LRN layer are those of (Krizhevsky et al., 2012).