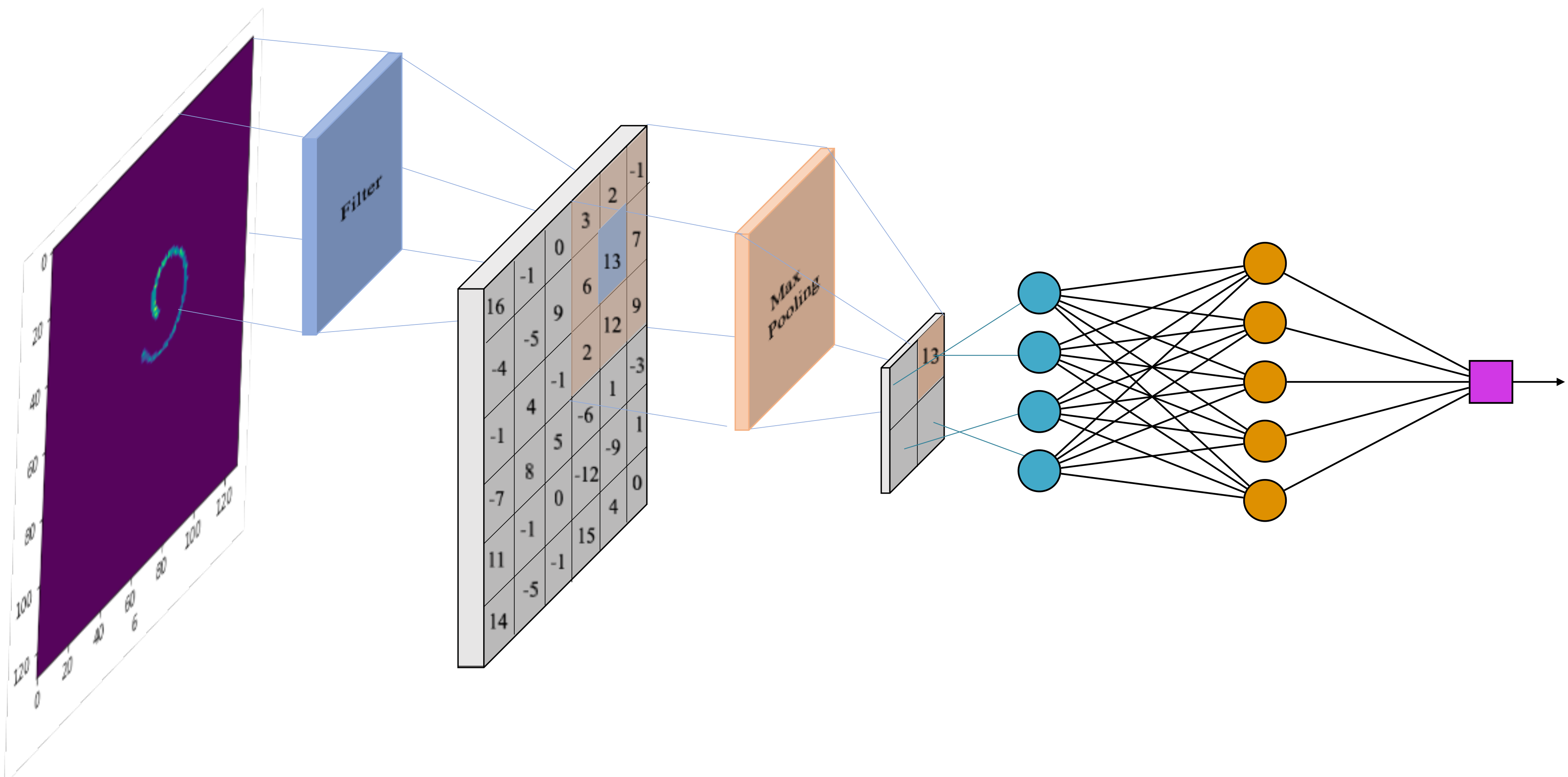
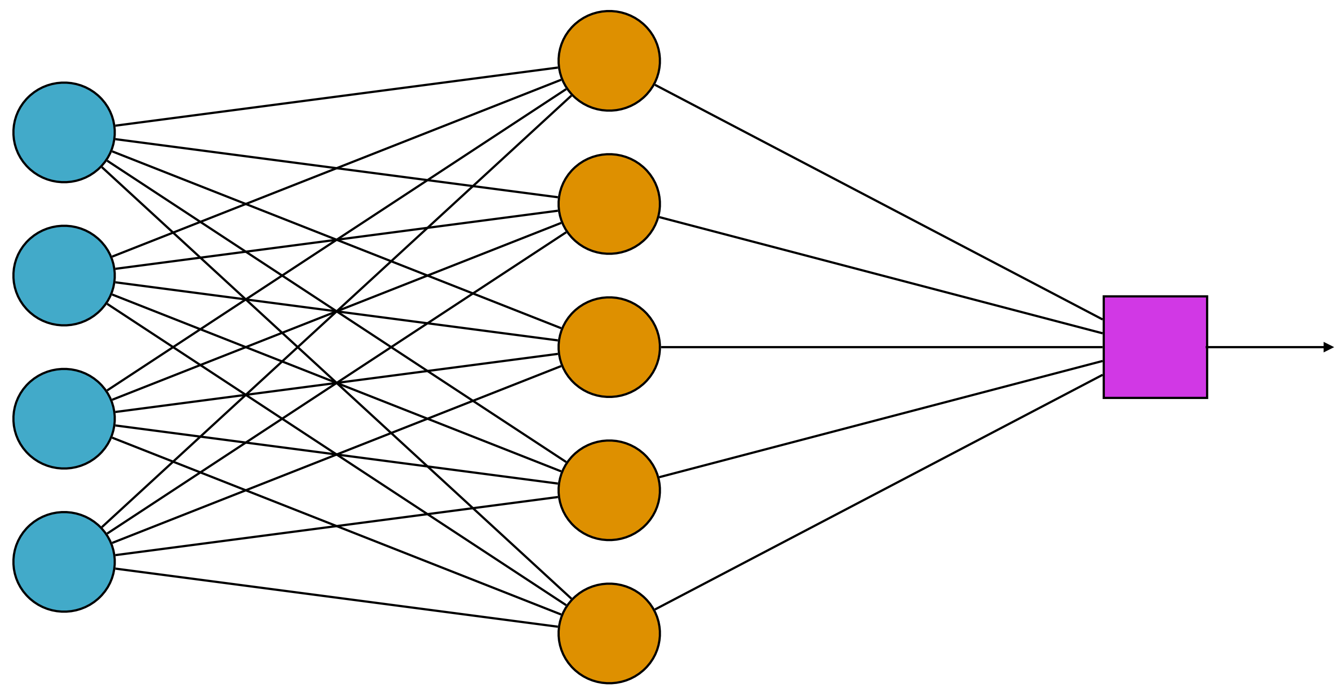
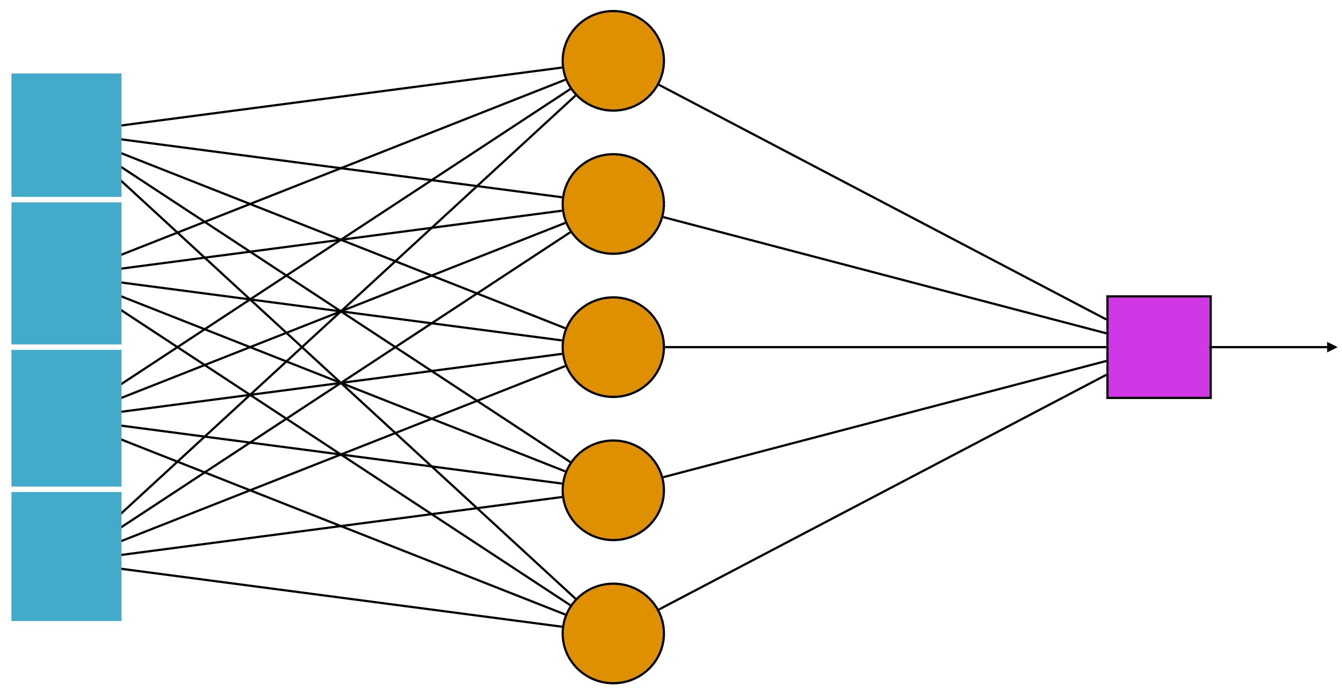
CONVOLUTIONAL NEURAL NETWORKS: ARCHITECTURE AND TRAINING

CONVOLUTIONAL NEURAL NETWORKS

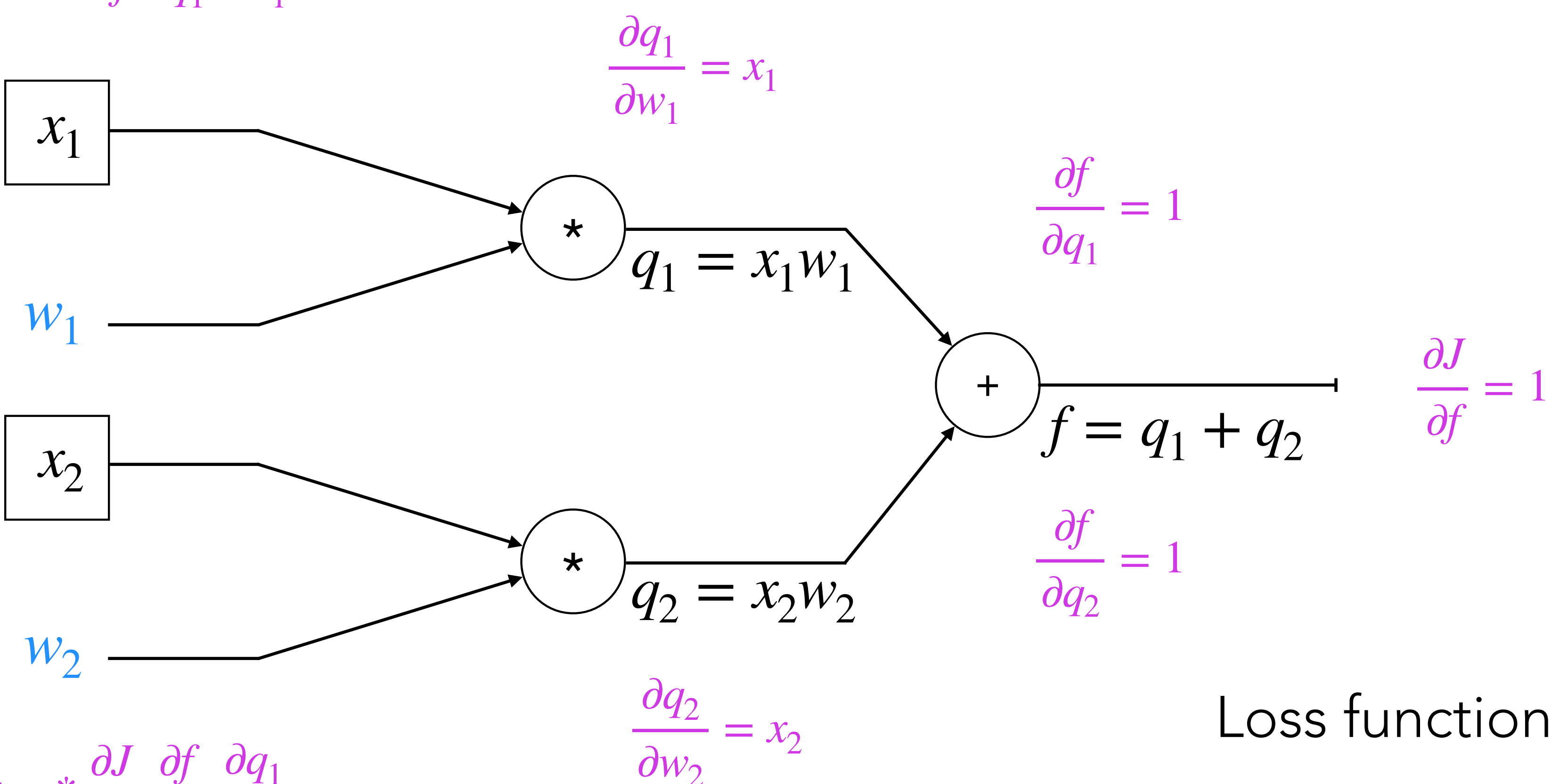






BACKPROPAGATION

$$w_1 = w_1 + \eta * \frac{\partial J}{\partial f}\frac{\partial f}{\partial q_1}\frac{\partial q_1}{\partial w_1}$$



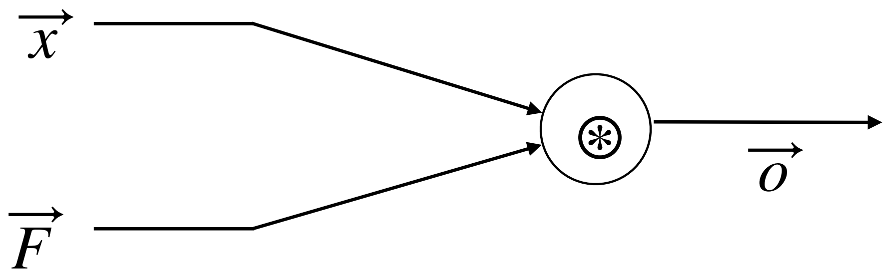
$$w_2 = w_2 + \eta * \frac{\partial J}{\partial f}\frac{\partial f}{\partial q_1}\frac{\partial q_1}{\partial w_2}$$

Loss function

$$J(w) = f - \hat{f}$$

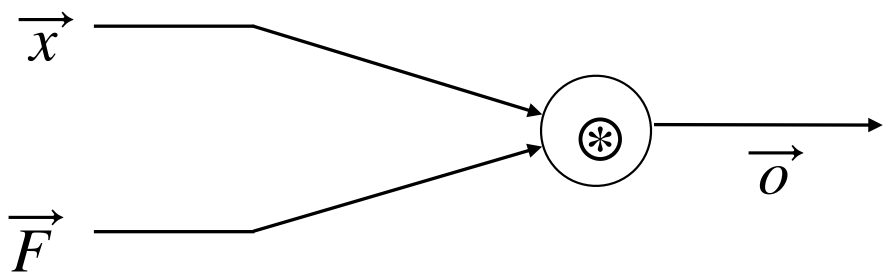
BACKPROPAGATION

$$\vec{o} = \vec{x} \circledast \vec{F}$$

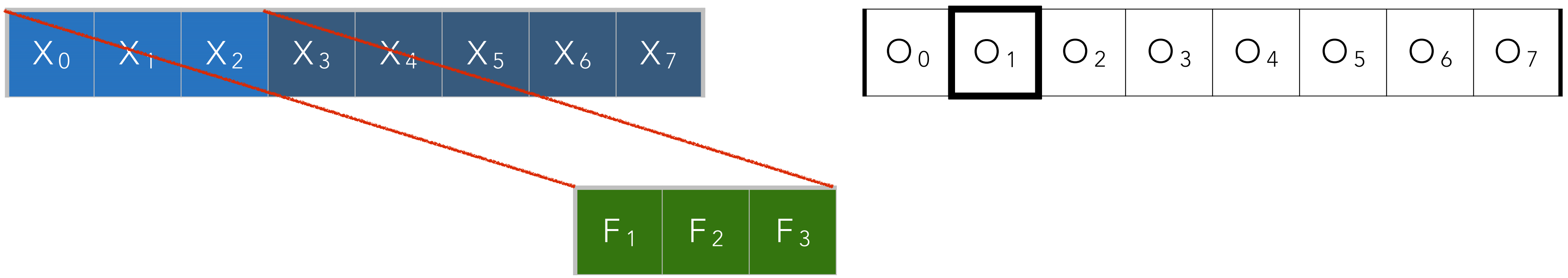


BACKPROPAGATION

$$\vec{o} = \vec{x} \circledast \vec{F}$$



$$o[n] = (x \circledast F)[n] = \sum_{i=-\omega}^{\omega} x[i+n+\omega] * F[i+\omega]$$

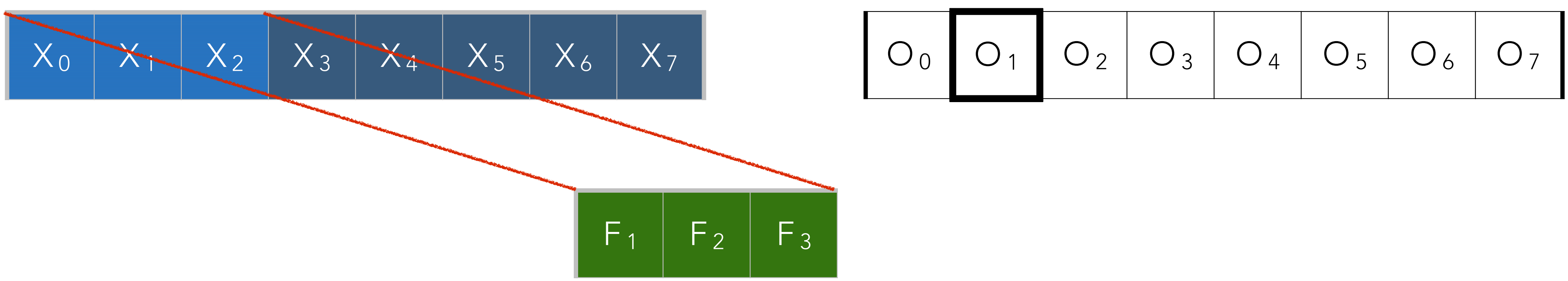


BACKPROPAGATION

$$F_i \leftarrow F_i - \eta \frac{\partial J}{\partial F_i}$$

$$\frac{\partial J}{\partial F_i} = \sum_{k=1}^{M} \frac{\partial J}{\partial o_k} \frac{\partial o_k}{\partial F_i}$$

$$o[n] = (x \circledast F)[n] = \sum_{i=-\omega}^{\omega} x[i+n+\omega] * F[i+\omega]$$

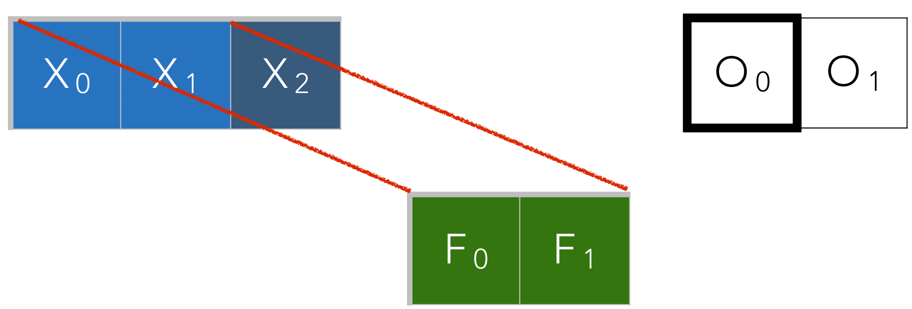


BACKPROPAGATION

$$F_i \leftarrow F_i - \eta \frac{\partial J}{\partial F_i}$$

$$\frac{\partial J}{\partial F_i} = \sum_{k=1}^{M} \frac{\partial J}{\partial o_k} \frac{\partial o_k}{\partial F_i}$$

$$o[n] = (x \circledast F)[n] = \sum_{i=-\omega}^{\omega} x[i+n+\omega] * F[i+\omega]$$



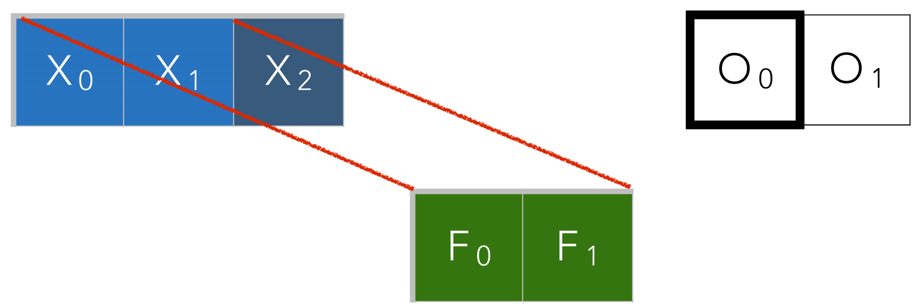
BACKPROPAGATION

$$F_i \leftarrow F_i - \eta \frac{\partial J}{\partial F_i}$$

$$\frac{\partial J}{\partial F_1} = \frac{\partial J}{\partial o_0}\frac{\partial o_0}{\partial F_1} + \frac{\partial J}{\partial o_1}\frac{\partial o_1}{\partial F_1}$$

$$\frac{\partial J}{\partial F_i} = \sum_{k=1}^{M} \frac{\partial J}{\partial o_k}\frac{\partial o_k}{\partial F_i}$$

$$o[n] = (x \circledast F)[n] = \sum_{i=-\omega}^{\omega} x[i+n+\omega] * F[i+\omega]$$



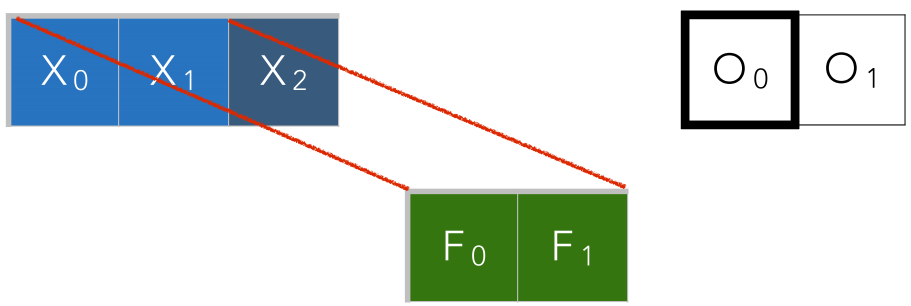
BACKPROPAGATION

$$\frac{\partial J}{\partial F_1} = \frac{\partial J}{\partial o_0}\frac{\partial o_0}{\partial F_1} + \frac{\partial J}{\partial o_1}\frac{\partial o_1}{\partial F_1}$$

$$F_i \leftarrow F_i - \eta\frac{\partial J}{\partial F_i}$$

$$\frac{\partial J}{\partial F_i} = \sum_{k=1}^{M} \frac{\partial J}{\partial o_k}\frac{\partial o_k}{\partial F_i}$$

$$o[n] = (x \circledast F)[n] = \sum_{i=-\omega}^{\omega} x[i+n+\omega] * F[i+\omega]$$



BACKPROPAGATION

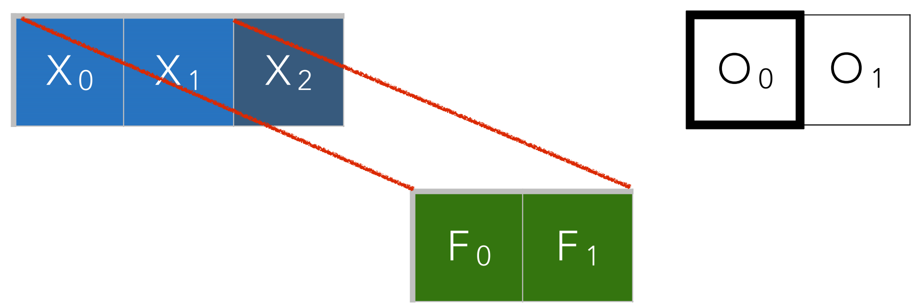
$$\frac{\partial J}{\partial F_1} = \frac{\partial J}{\partial o_0}\frac{\partial o_0}{\partial F_1} + \frac{\partial J}{\partial o_1}\frac{\partial o_1}{\partial F_1}$$

$$o_0 = F_0 x_0 + F_1 x_1$$
$$o_1 = F_0 x_1 + F_1 x_2$$

$$F_i \leftarrow F_i - \eta \frac{\partial J}{\partial F_i}$$

$$\frac{\partial J}{\partial F_i} = \sum_{k=1}^{M} \frac{\partial J}{\partial o_k}\frac{\partial o_k}{\partial F_i}$$

$$o[n] = (x \circledast F)[n] = \sum_{i=-\omega}^{\omega} x[i+n+\omega] * F[i+\omega]$$



BACKPROPAGATION

$$\frac{\partial J}{\partial F_1} = \frac{\partial J}{\partial o_0}\frac{\partial o_0}{\partial F_1} + \frac{\partial J}{\partial o_1}\frac{\partial o_1}{\partial F_1}$$

$$F_i \leftarrow F_i - \eta \frac{\partial J}{\partial F_i}$$

$$\frac{\partial J}{\partial F_i} = \sum_{k=1}^{M} \frac{\partial J}{\partial o_k}\frac{\partial o_k}{\partial F_i}$$

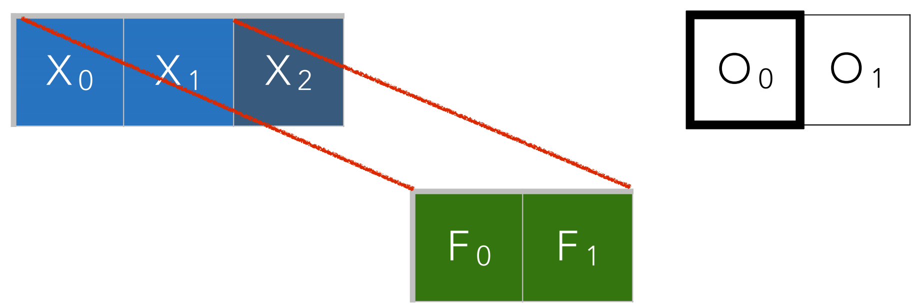
$$o_0 = F_0 x_0 + F_1 x_1$$

$$o_1 = F_0 x_1 + F_1 x_2$$

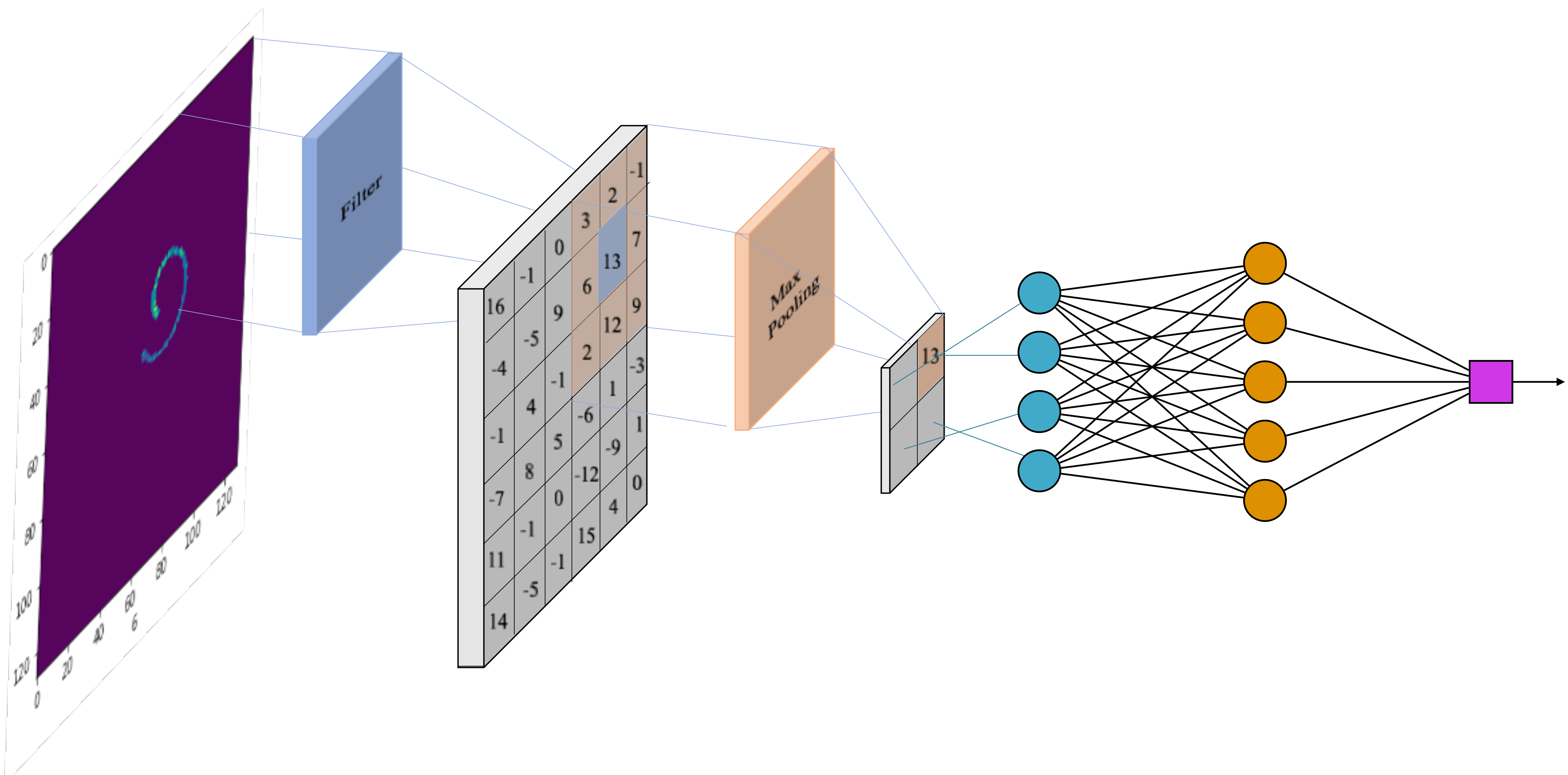
$$o[n] = (x \circledast F)[n] = \sum_{i=-\omega}^{\omega} x[i+n+\omega] * F[i+\omega]$$

$$\frac{\partial o_0}{\partial F_1} = x_1$$

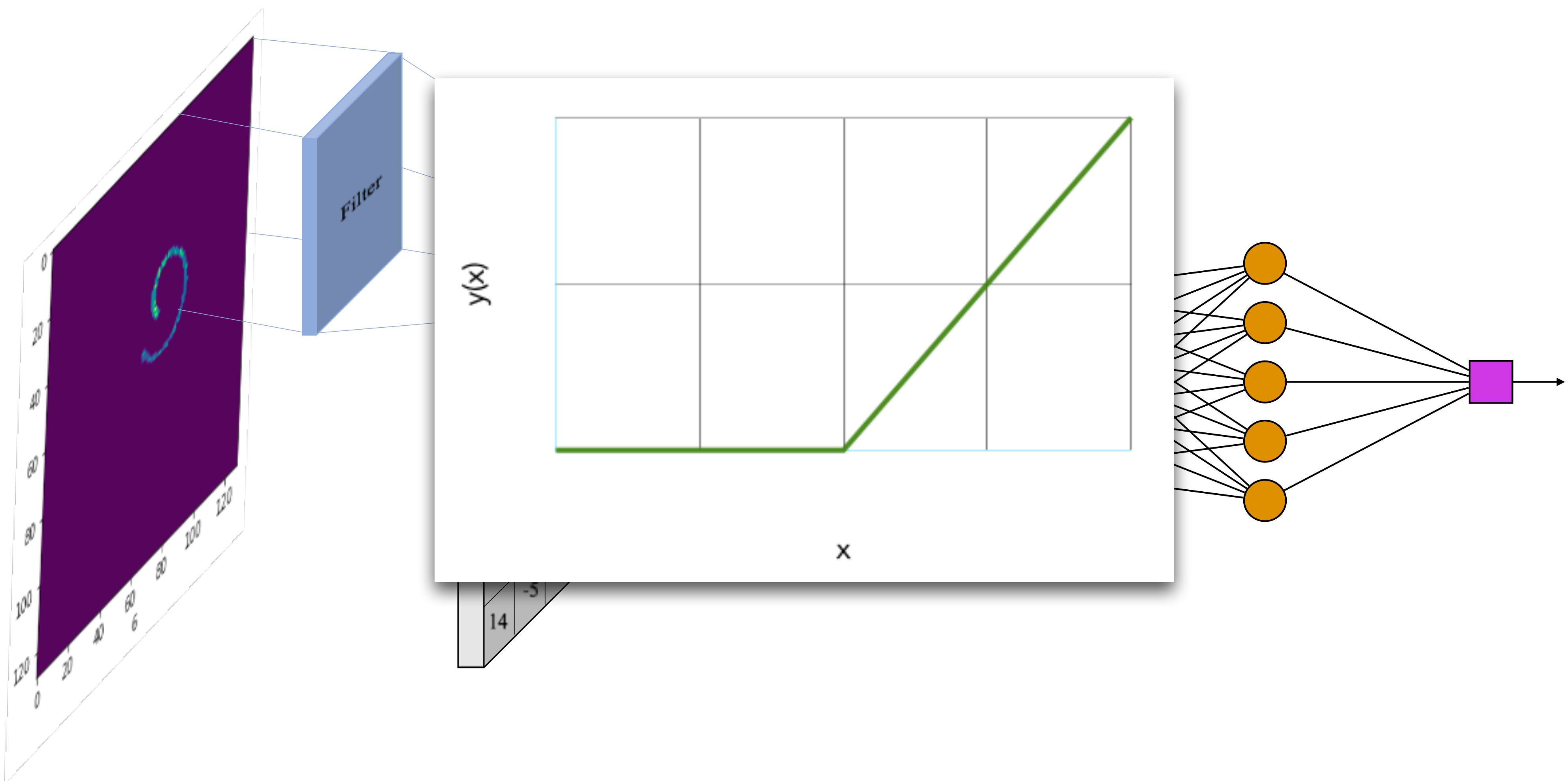
$$\frac{\partial o_1}{\partial F_1} = x_2$$



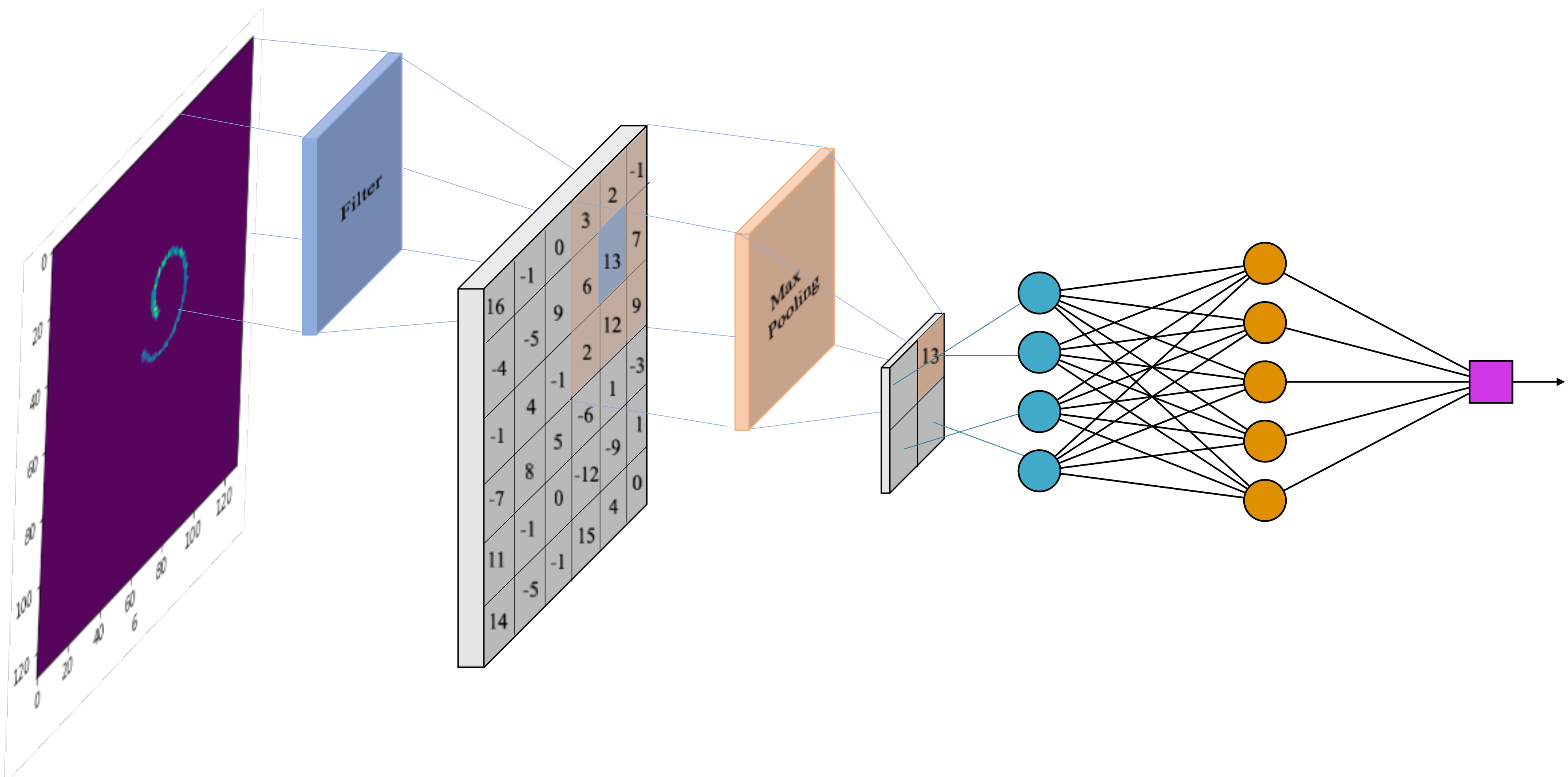
CONVOLUTIONAL NEURAL NETWORKS



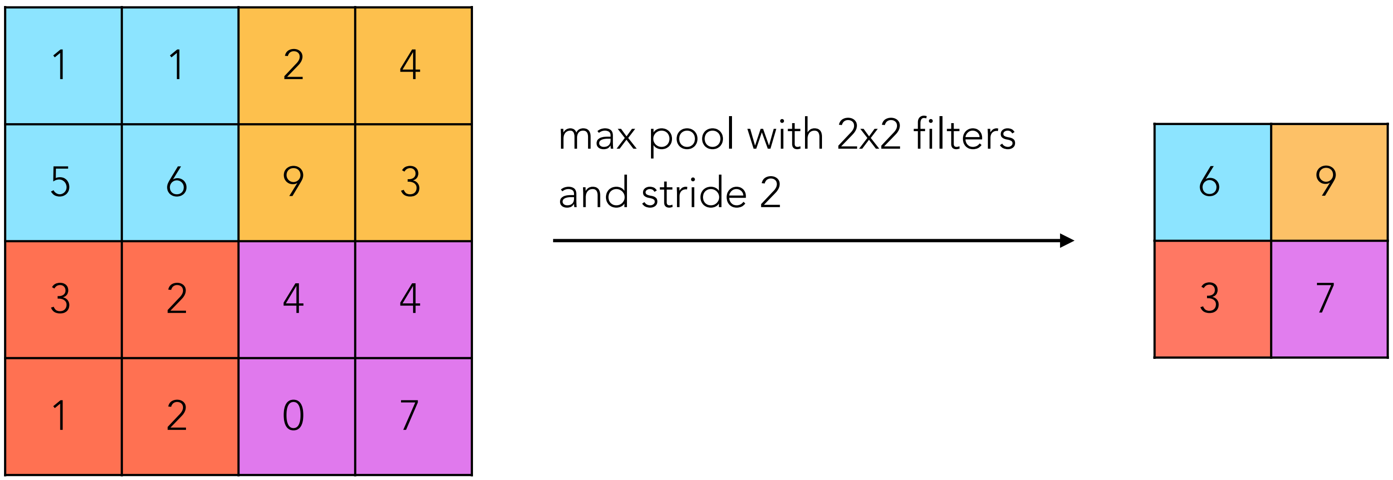
CONVOLUTIONAL NEURAL NETWORKS



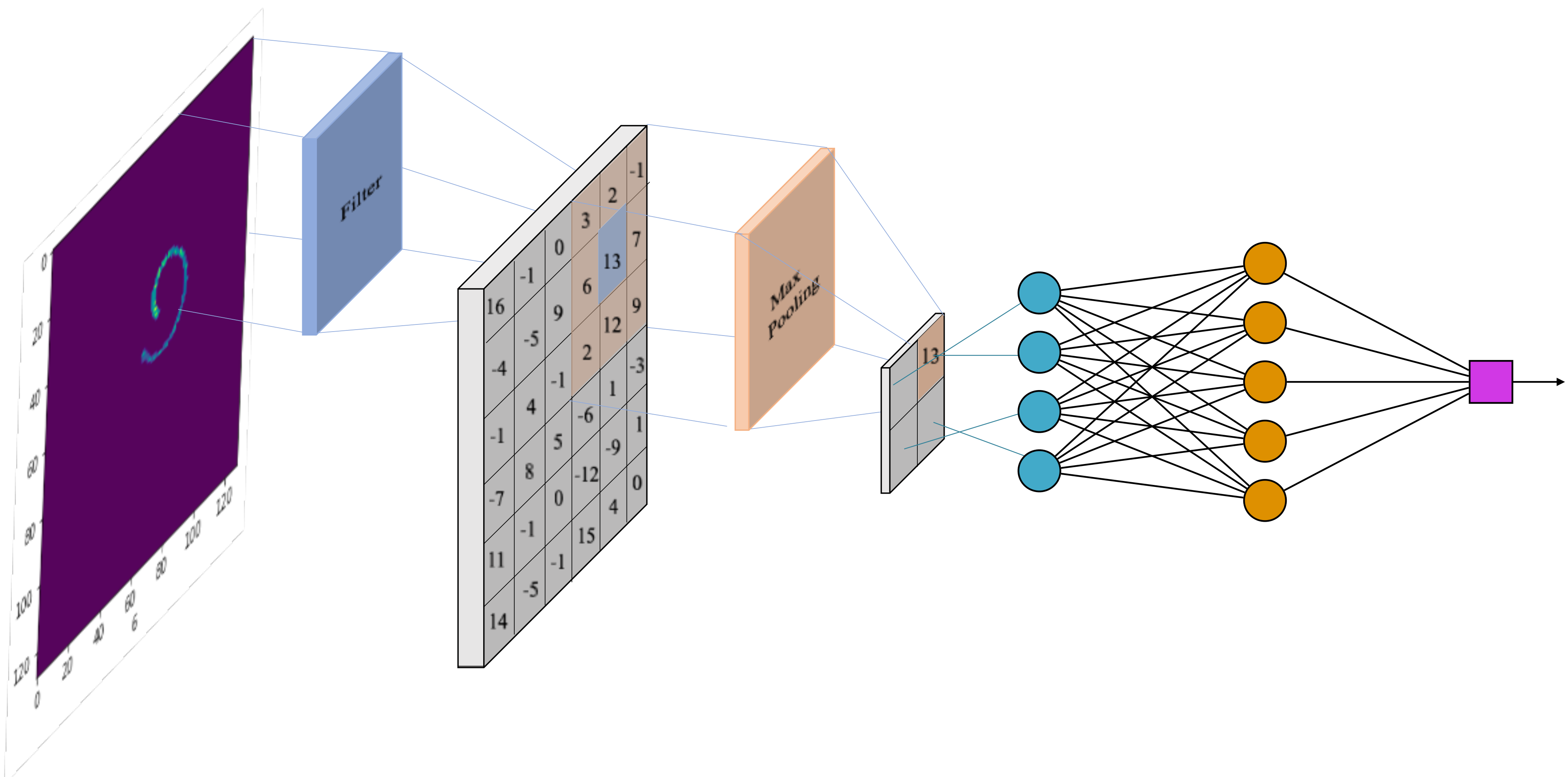
CONVOLUTIONAL NEURAL NETWORKS



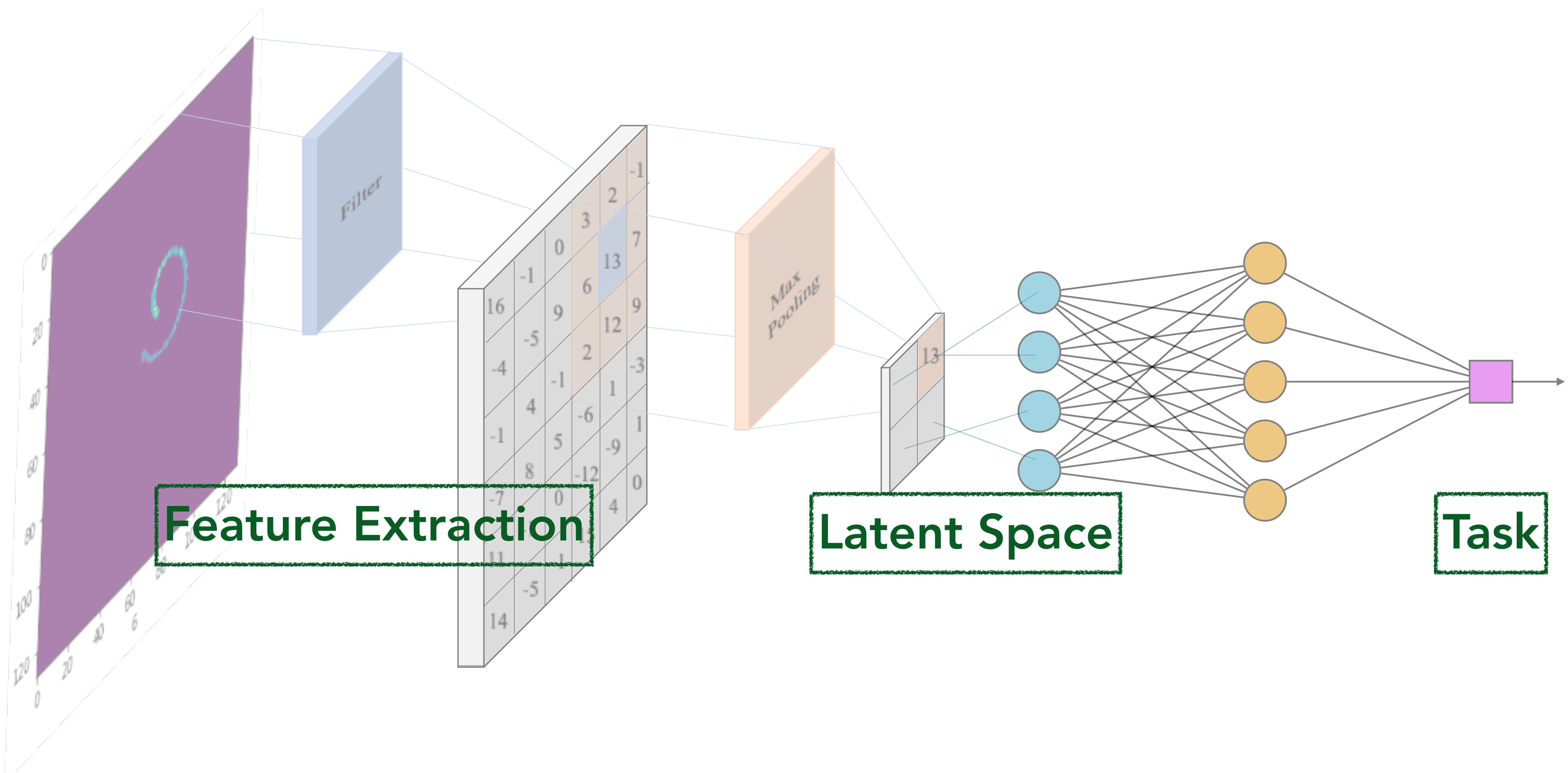
POOLING


1 1 2 4
5 6 9 3
3 2 4 4
1 2 0 7
max pool with 2x2 filters
and stride 2
6 9
3 7

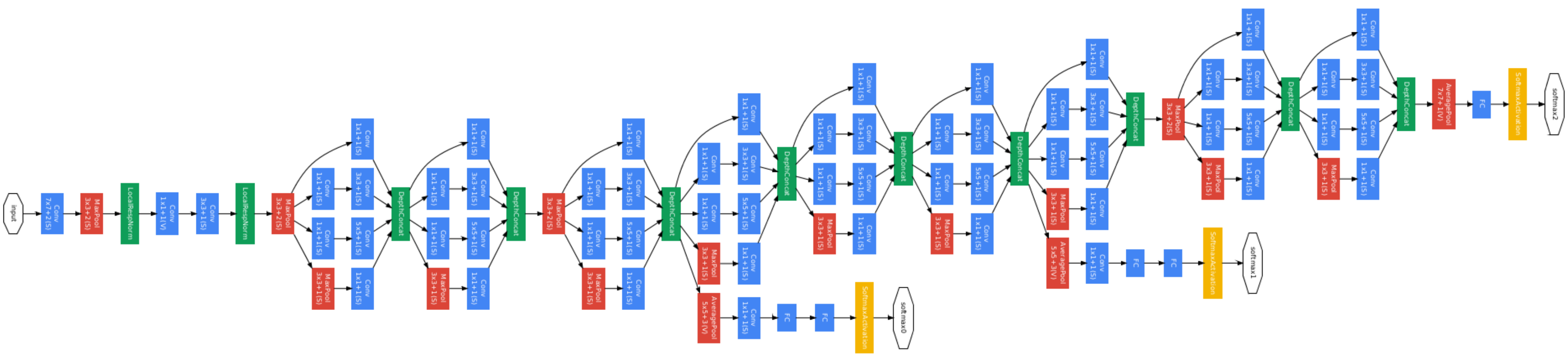
BACKPROPOGATION



CONVOLUTIONAL NEURAL NETWORKS



CONVOLUTIONAL NEURAL NETWORKS



CHRISTIAN SZEGEDY ET. AL. GOING DEEPER WITH CONVOLUTIONS.

CHOOSING AN ARCHITECTURE

HOW MANY LAYERS?

HOW MANY NODES PER LAYER?

LEARNING RATE

DROPOUT?

WHAT ACTIVATION FUNCTION(S)?

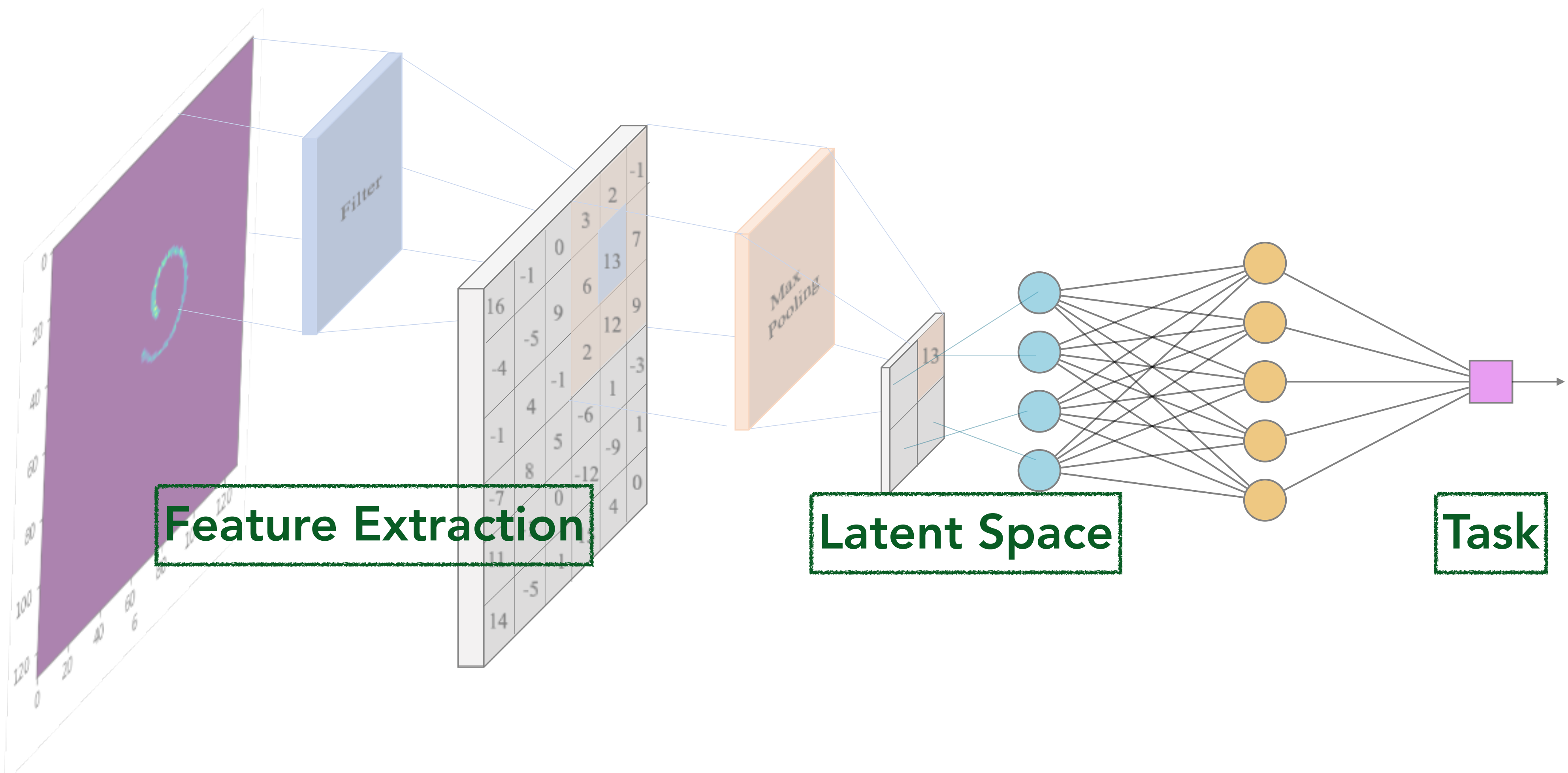
HOW MANY CONVOLUTION LAYERS?

FILTER SIZE?

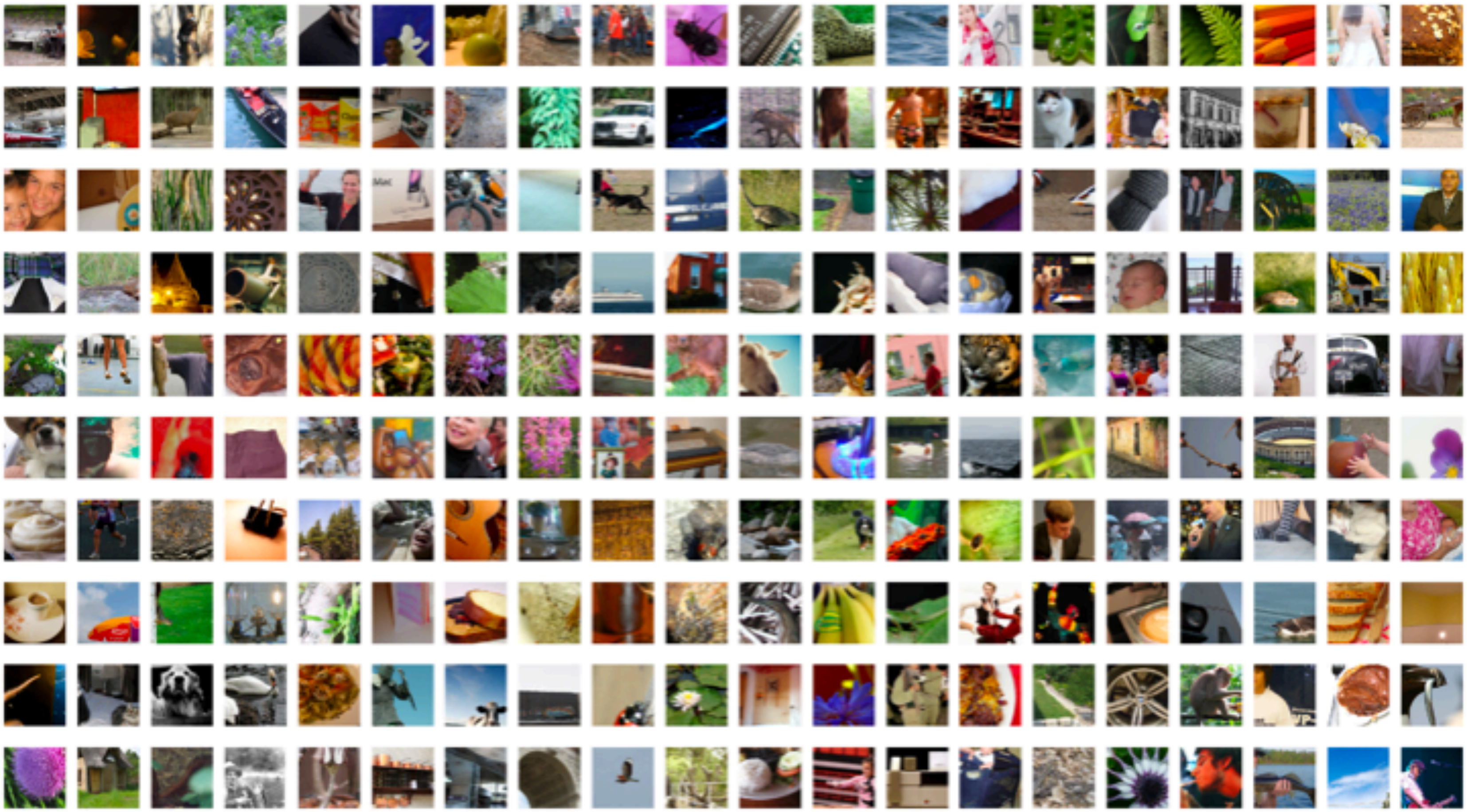
STRIDE?

POOLING?

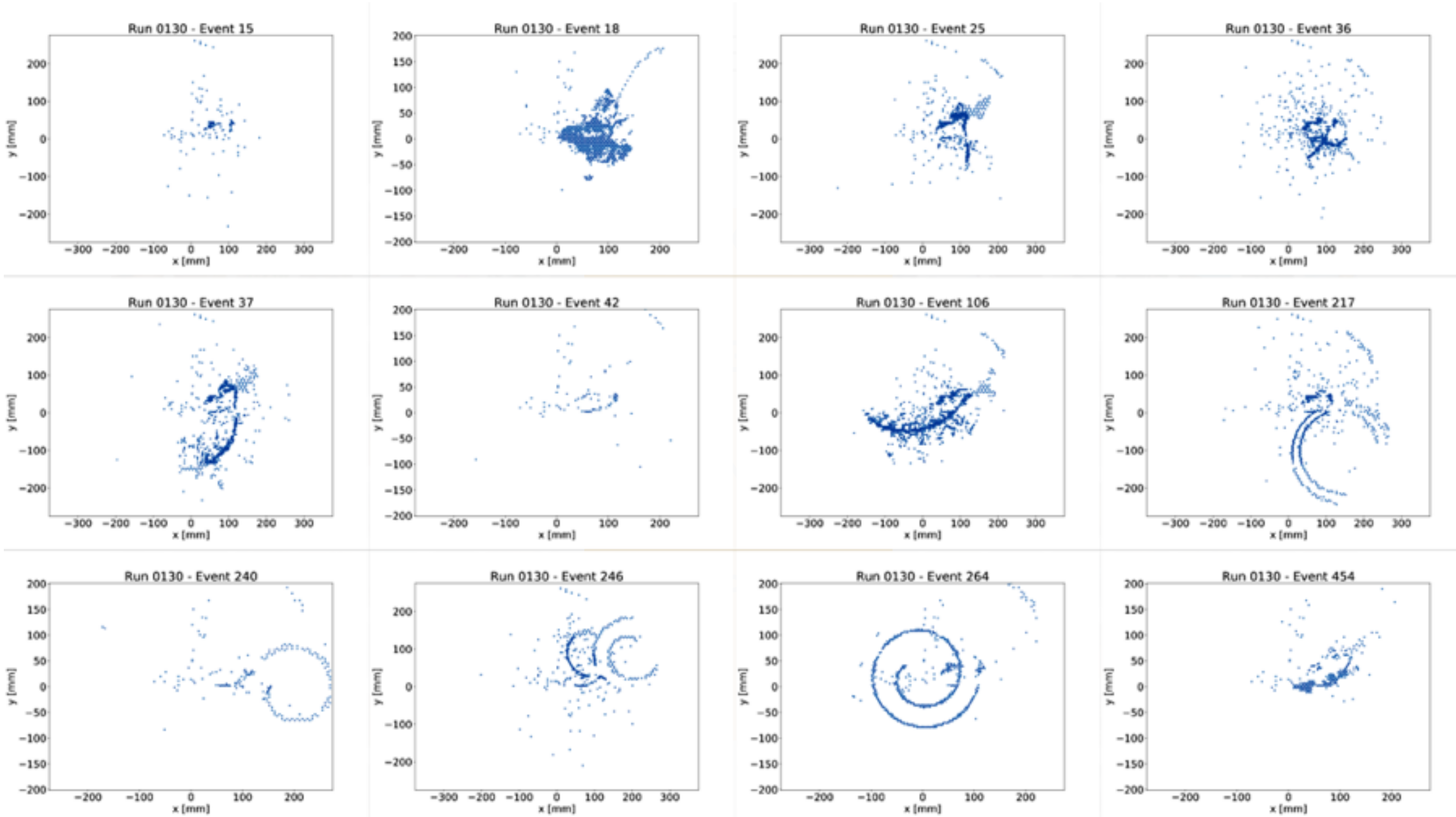
PRE-TRAINED MODELS

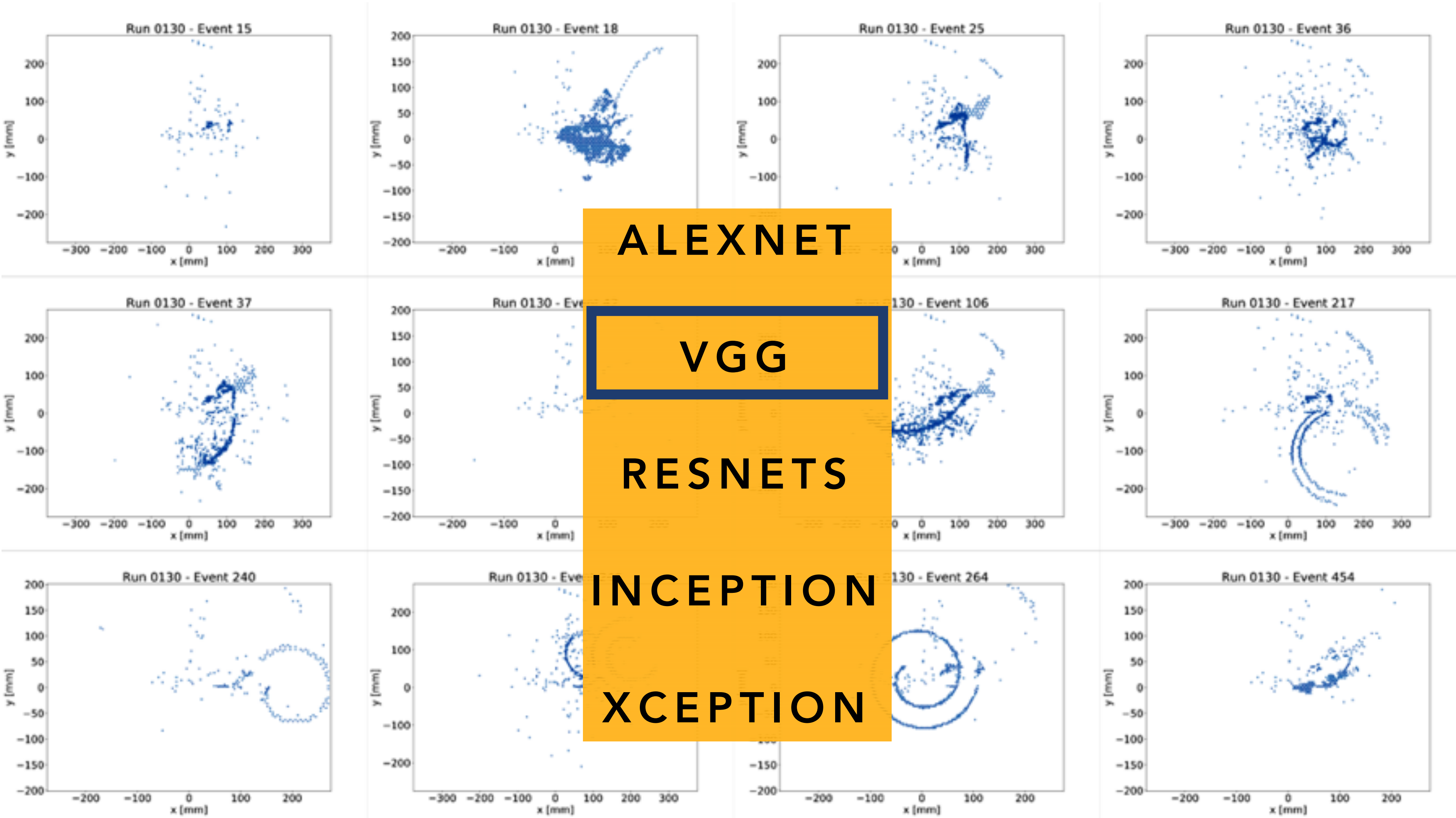


PRE-TRAINED MODELS



PRETRAINED MODELS


Run 0130 - Event 15
Run 0130 - Event 18
Run 0130 - Event 25
Run 0130 - Event 36
Run 0130 - Event 37
Run 0130 - Event 42
Run 0130 - Event 106
Run 0130 - Event 217
Run 0130 - Event 240
Run 0130 - Event 246
Run 0130 - Event 264
Run 0130 - Event 454
x [mm]
y [mm]



ALEXNET

VGG

RESNETS

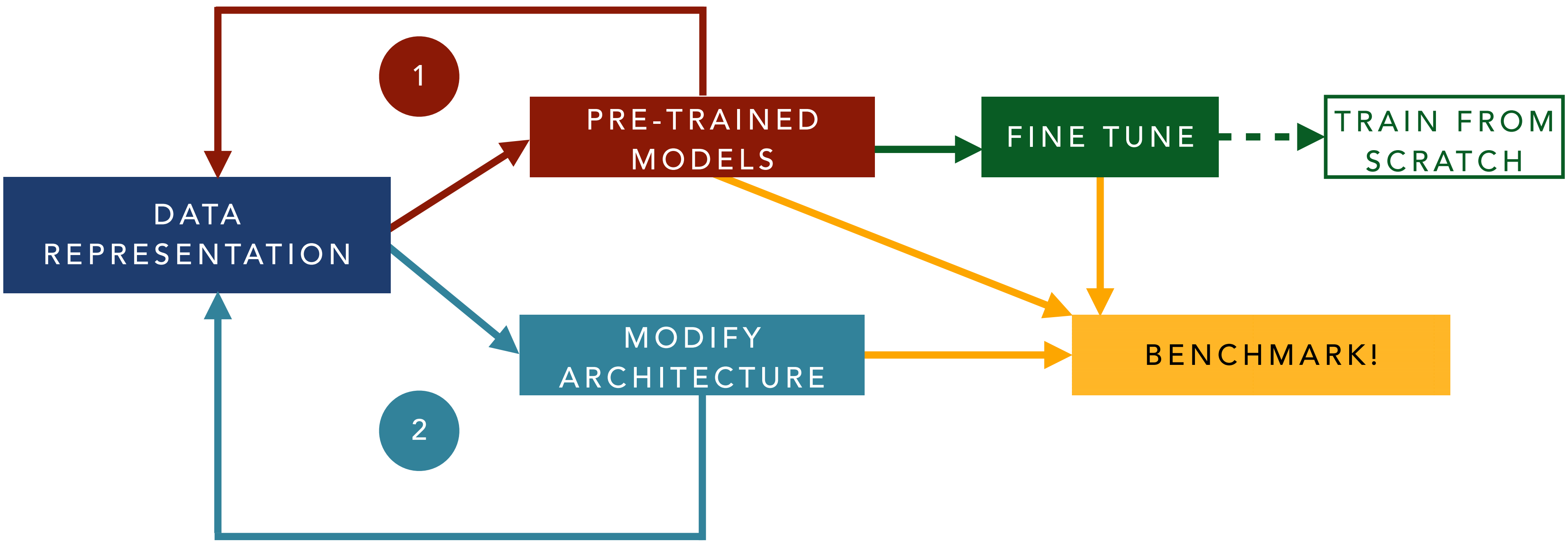
INCEPTION

XCEPTION

J. Z. TAYLOR, HONOR'S THESIS, DAVIDSON COLLEGE

M.P. KUCHERA, ET. AL. NUCLEAR INSTRUMENTATION AND METHODS A (2019)

EXAMPLE WORKFLOW



EXAMPLE WORKFLOW