

Convolution Neural Networks in Practice

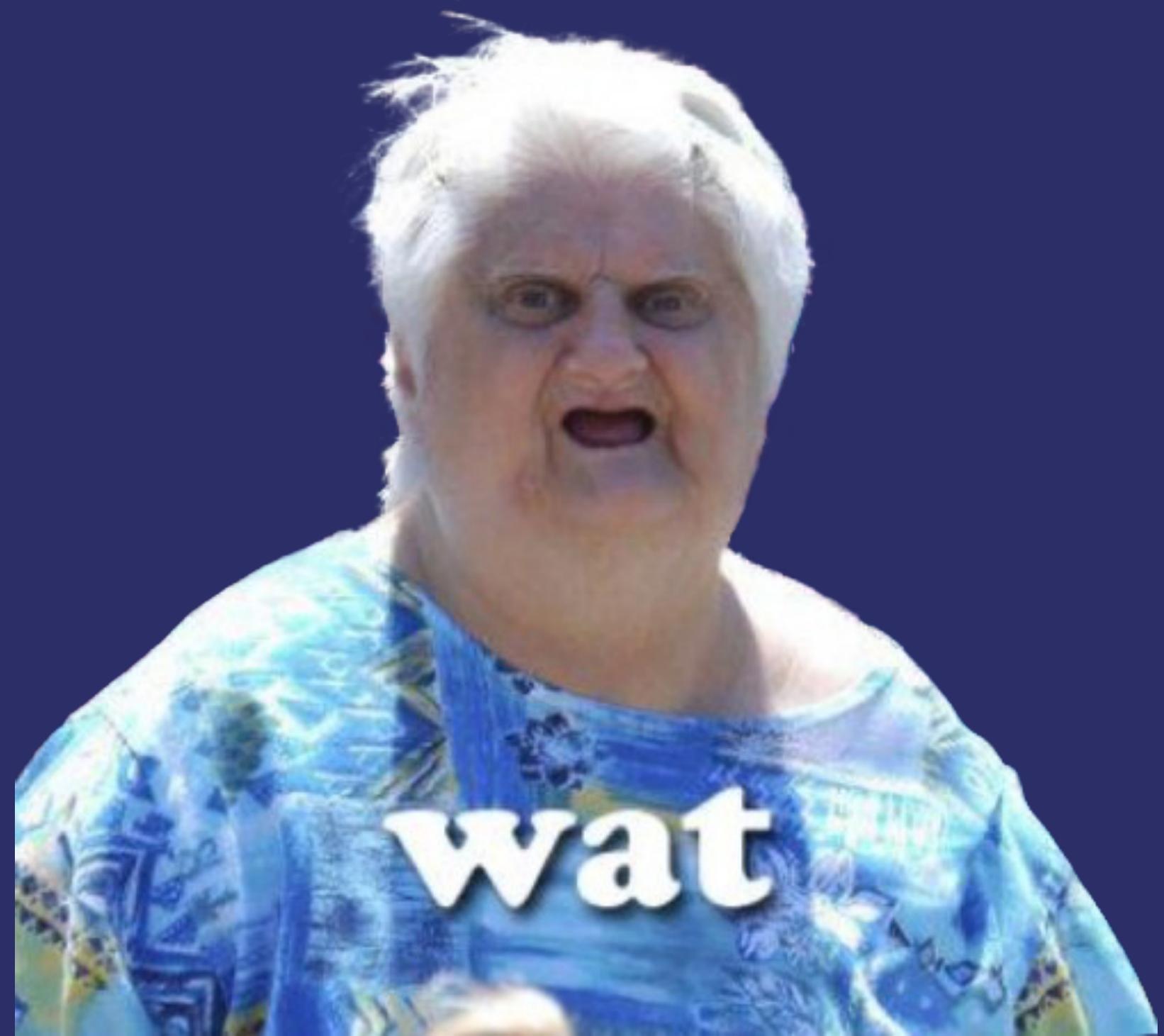
Cassidy Williams

Software Engineer & Developer Evangelist
at **Clarifai**

Machine Learning → Artificial Intelligence

“A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.”

Tom Mitchell, 1997

















**So, how does
clarifai do it?**

*Neural
Networks!*

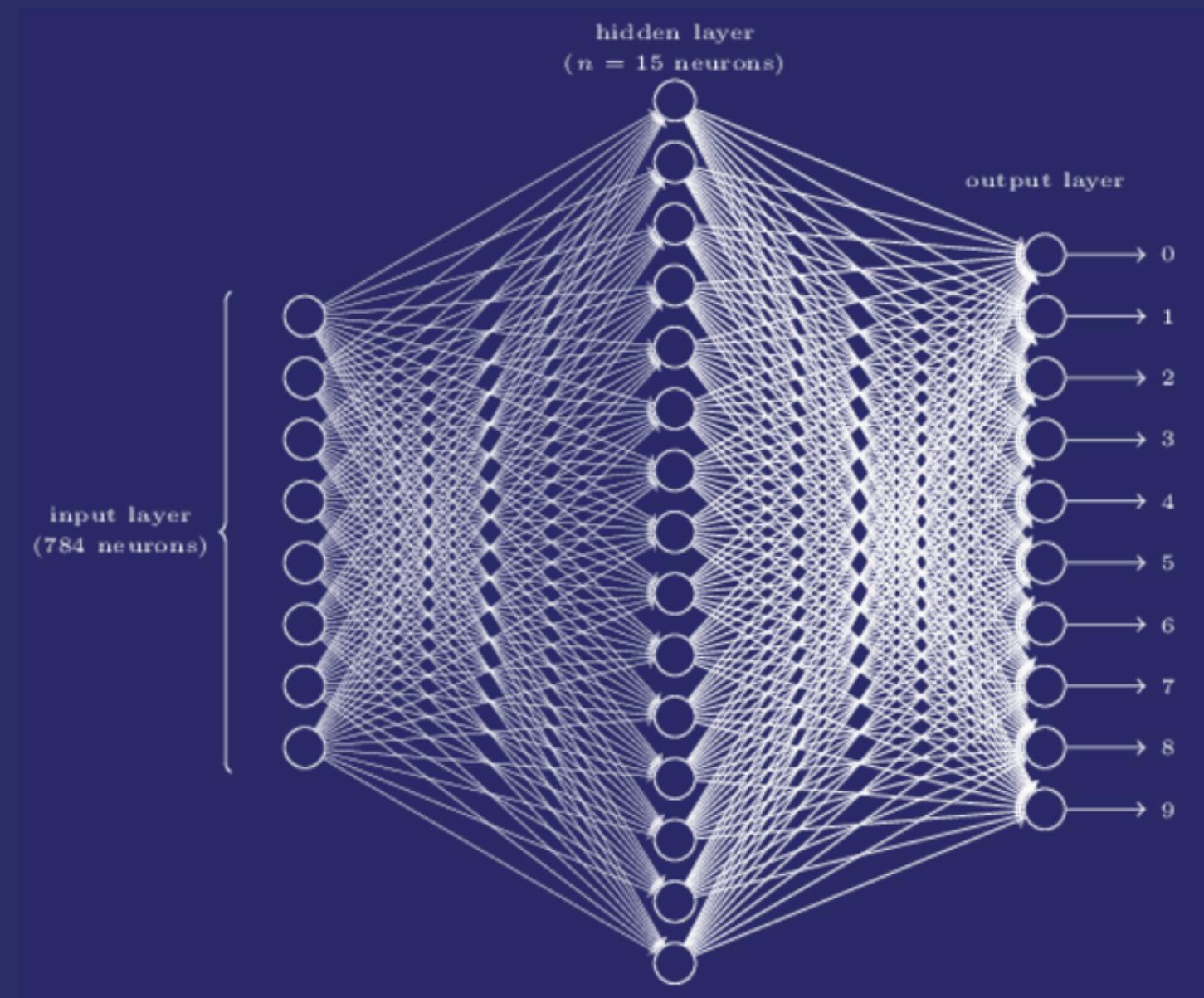
Training phase

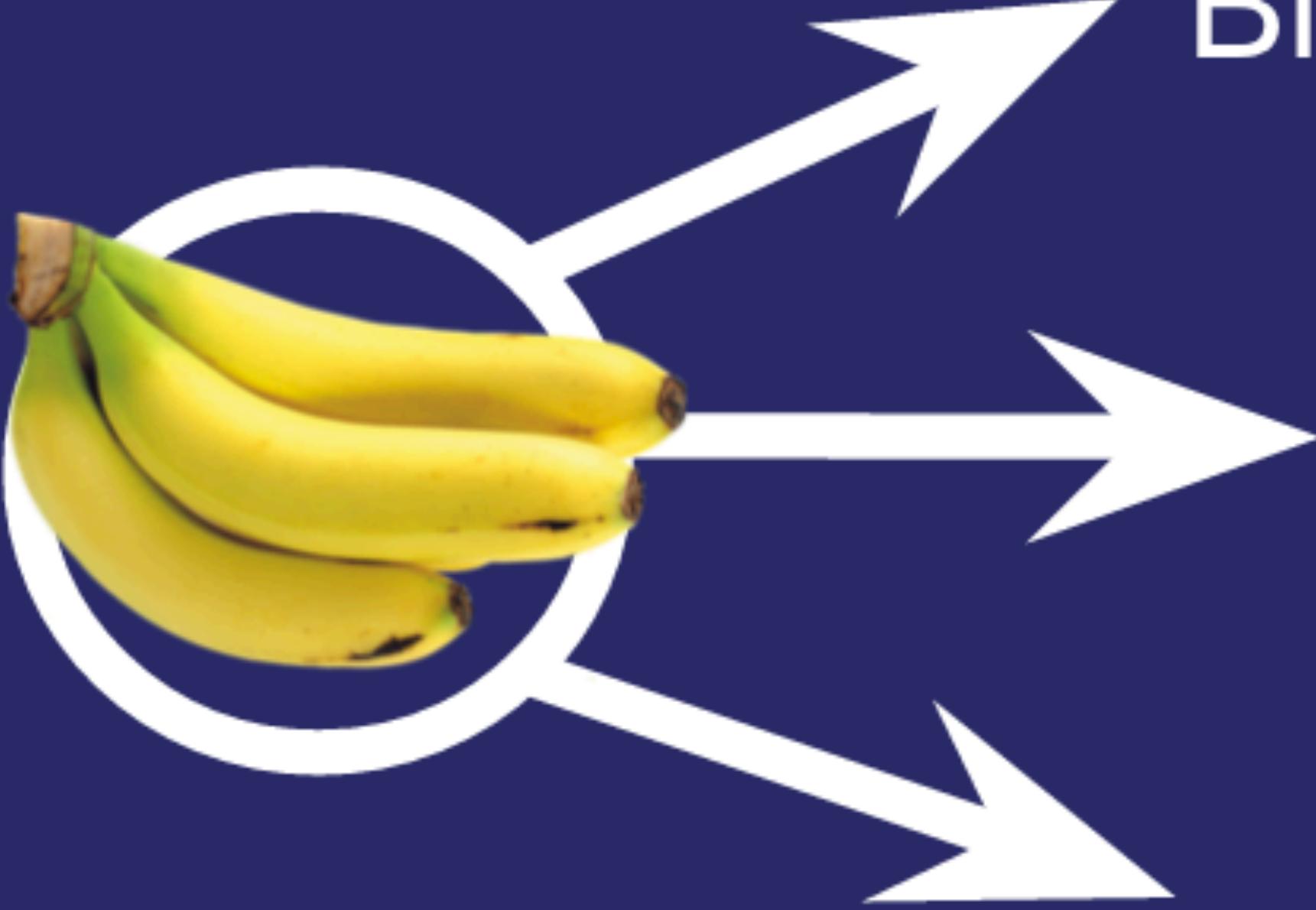
Take pairs of input data and desired output that has been collected beforehand

**The easy to understand
way...**

(Clarifai does not do this)

Adjacenct Neural Networks



A photograph of a bunch of ripe yellow bananas is positioned on the left side of the image. A white magnifying glass icon is placed over the bananas, with its handle pointing towards the top-left and its lens centered on the bunch.

Blue

Red

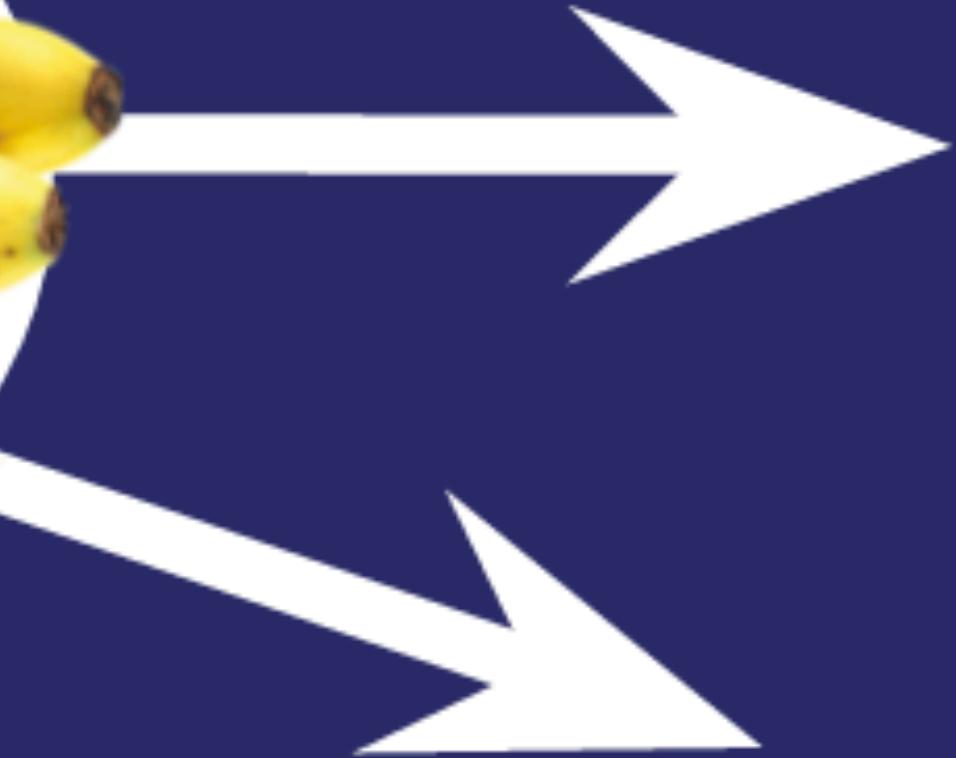
Yellow



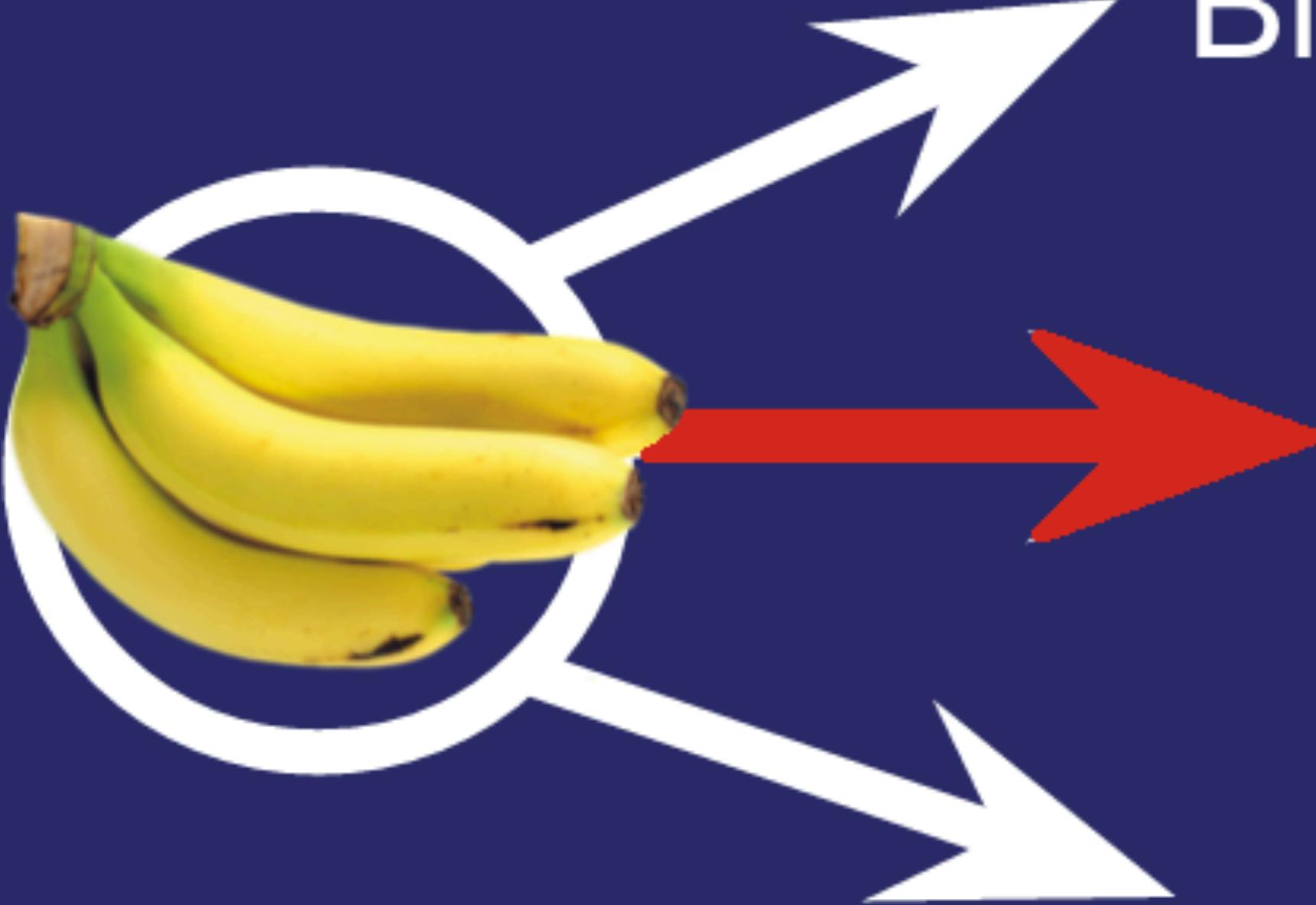
Blue



Red



Yellow



Blue

Red

Yellow



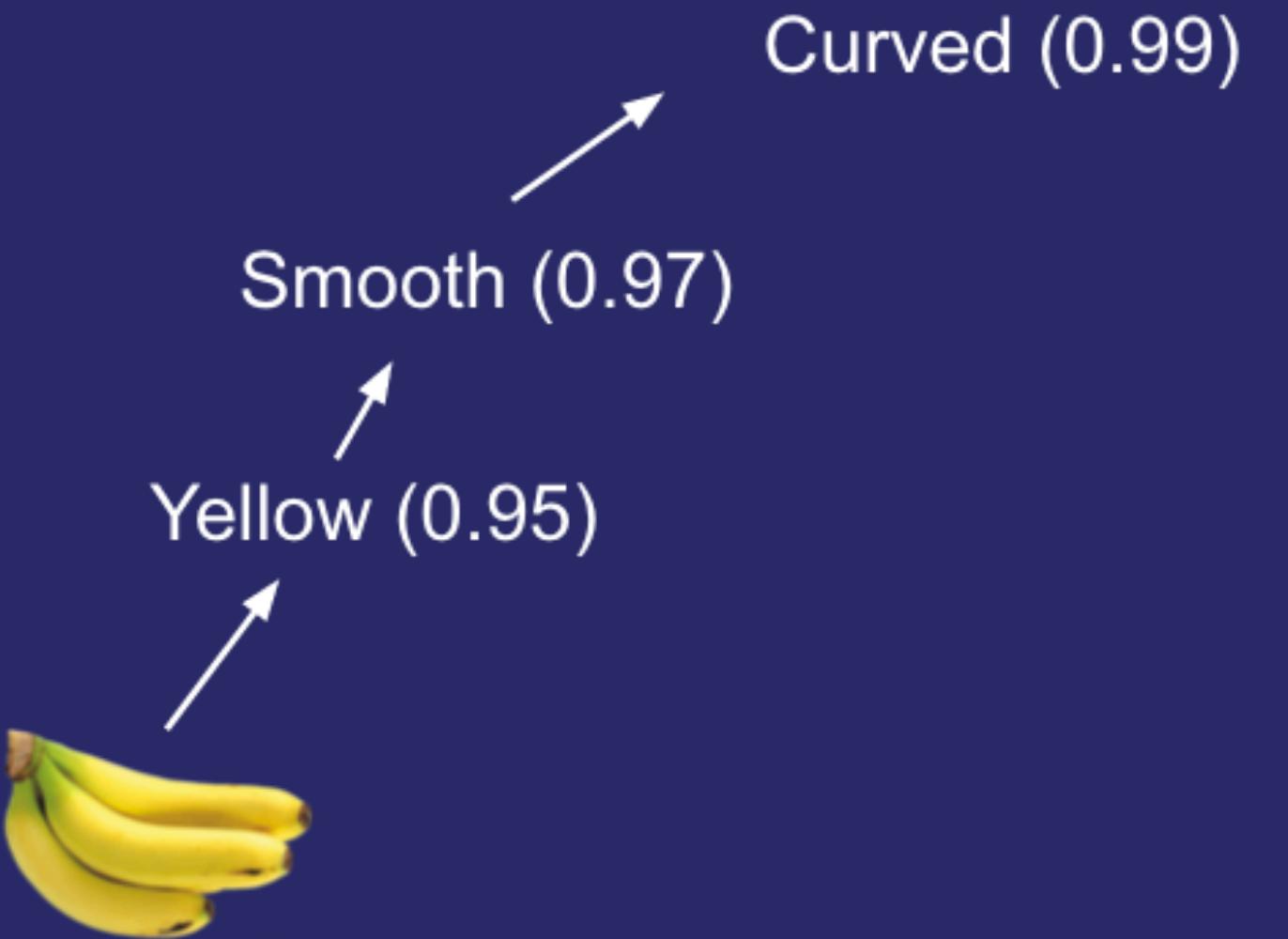
Blue



Red



Yellow



There is a 97%
chance that it is a
BANANA

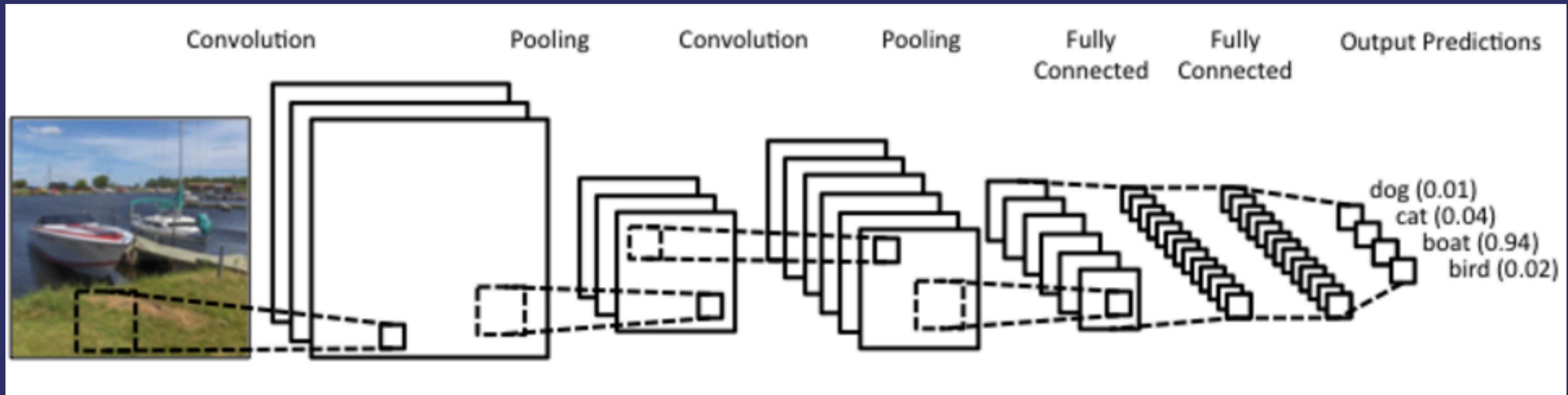
with the tags YELLOW,
SMOOTH, and CURVED

**Then there's the
AWESOME WAY.**

(Clarifai does do this)

convolution
Neural
Networks

Convolution Neural Networks



Local Receptive Fields

a window on the input of pixels

Feature Map

mapping from input layer to hidden layer

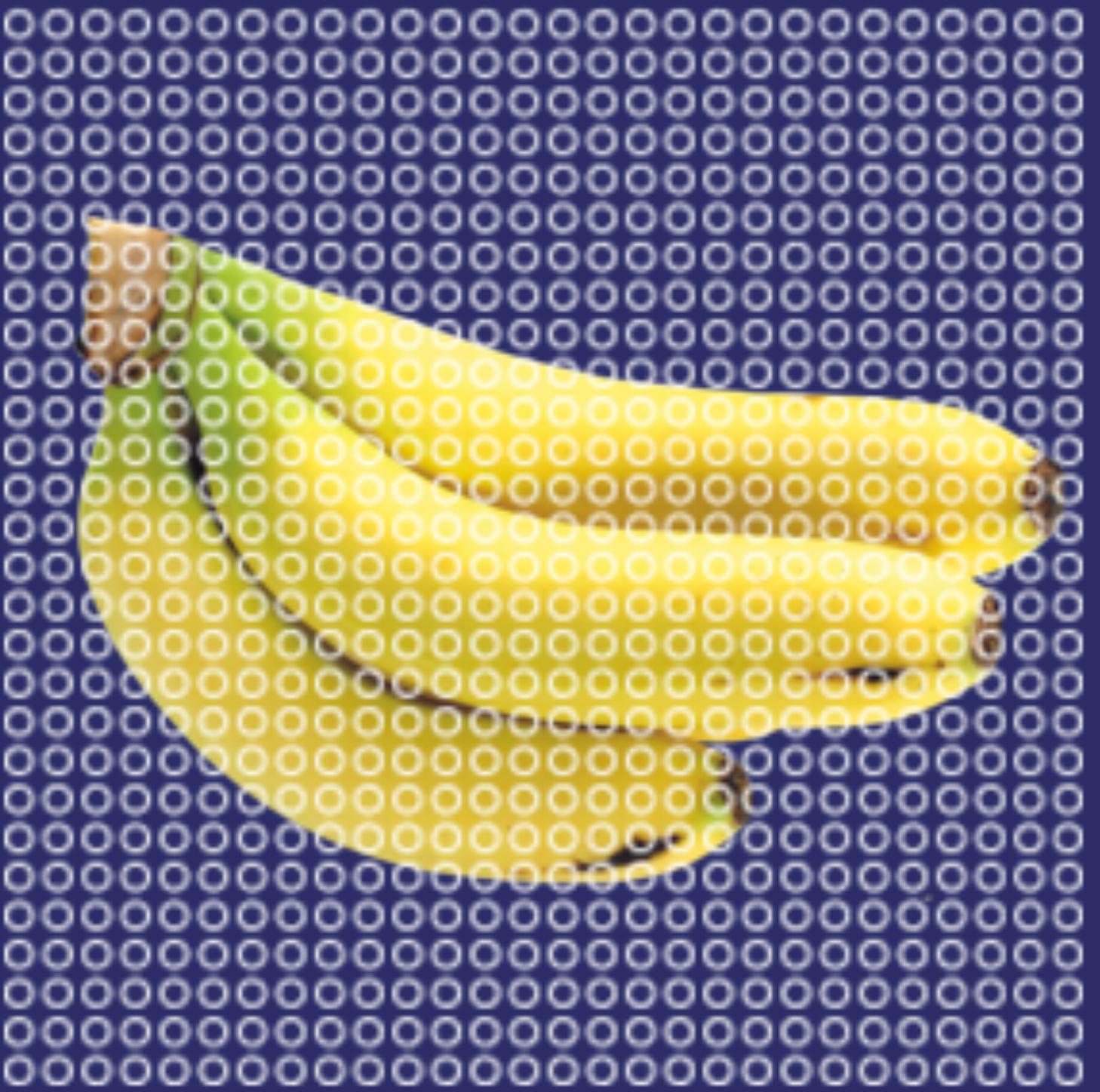
Shared Weights

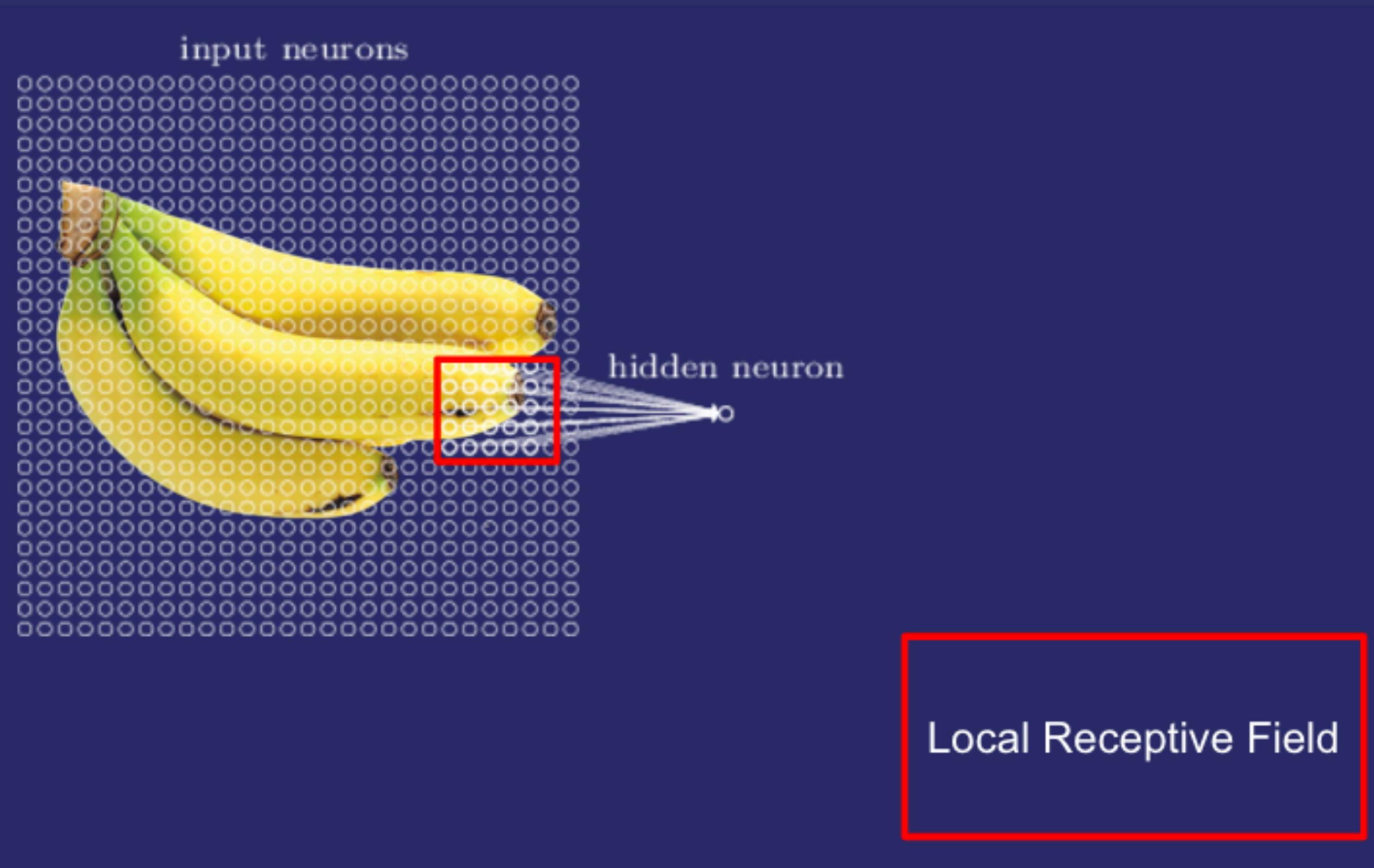
positive or negatives on a feature map

Pooling

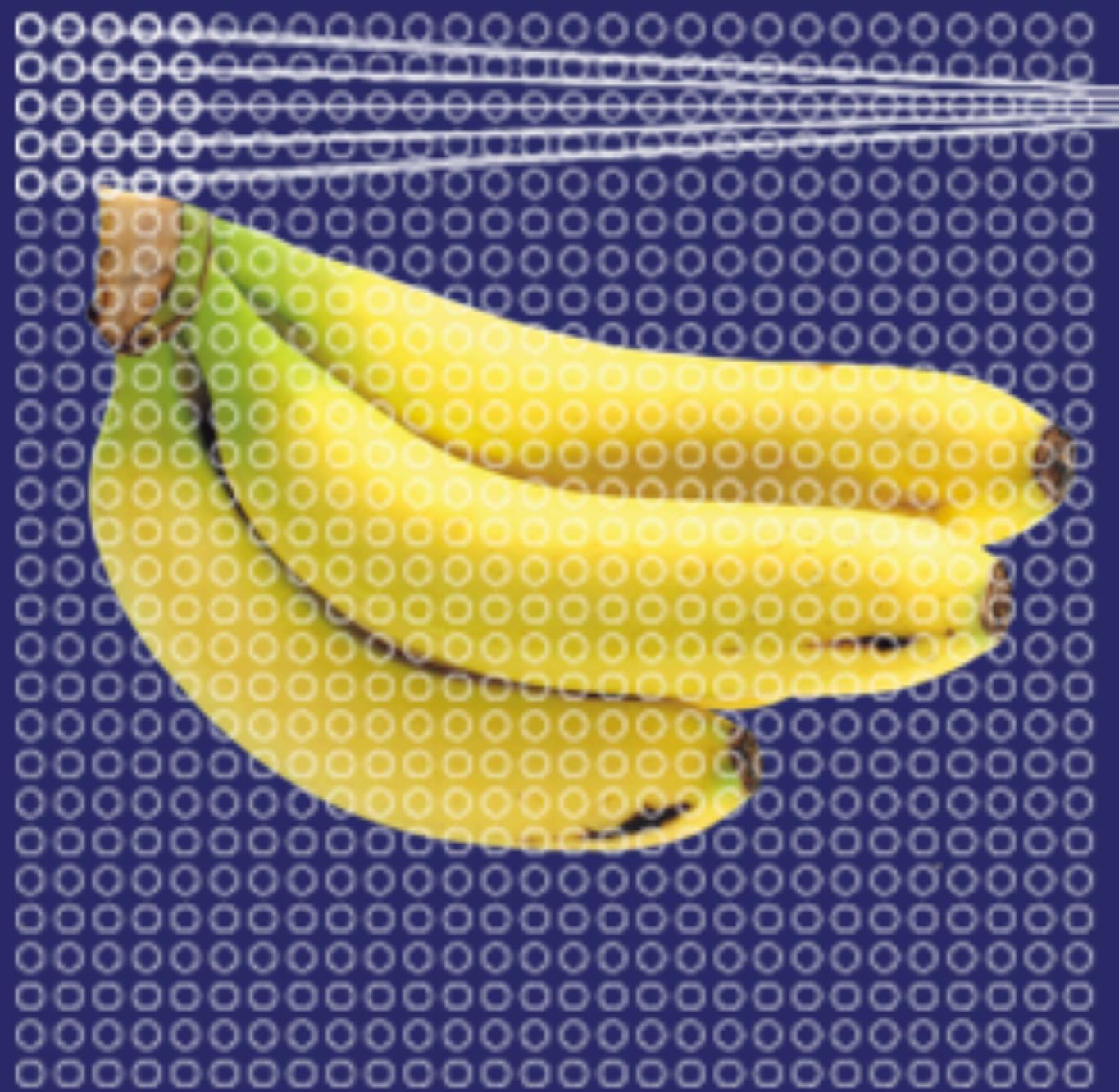
simplifying the information from the feature maps

input neurons

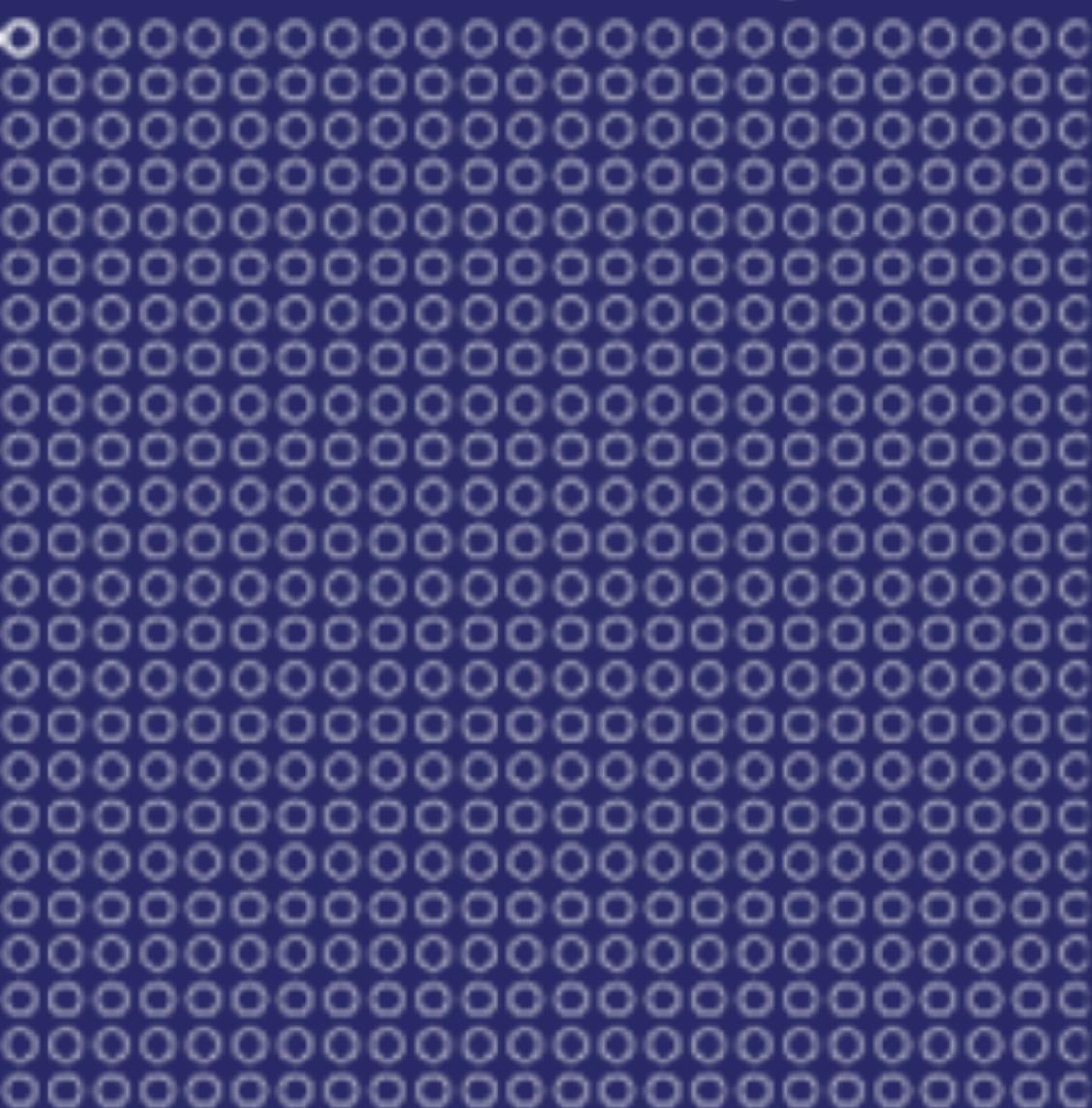




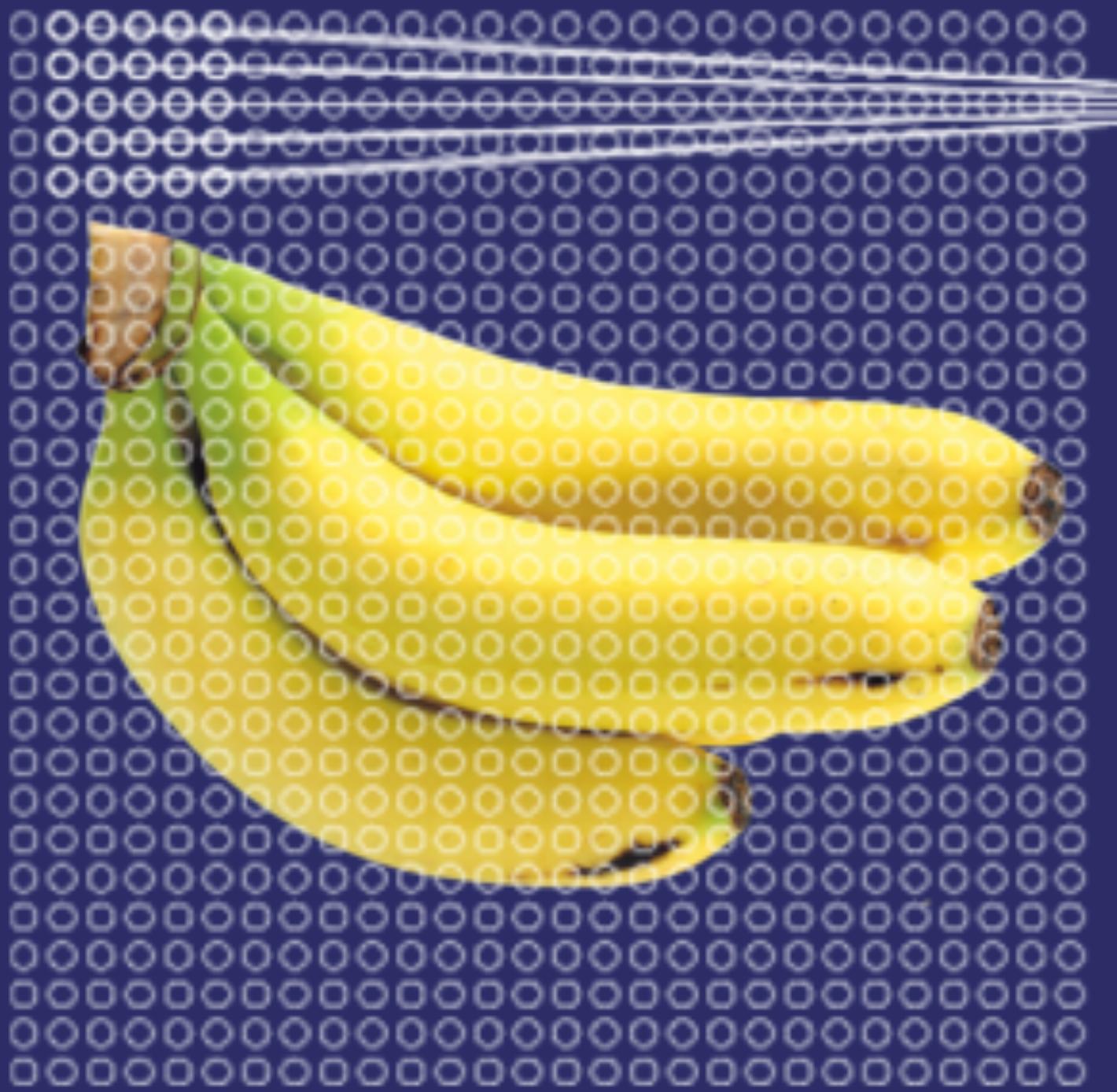
input neurons



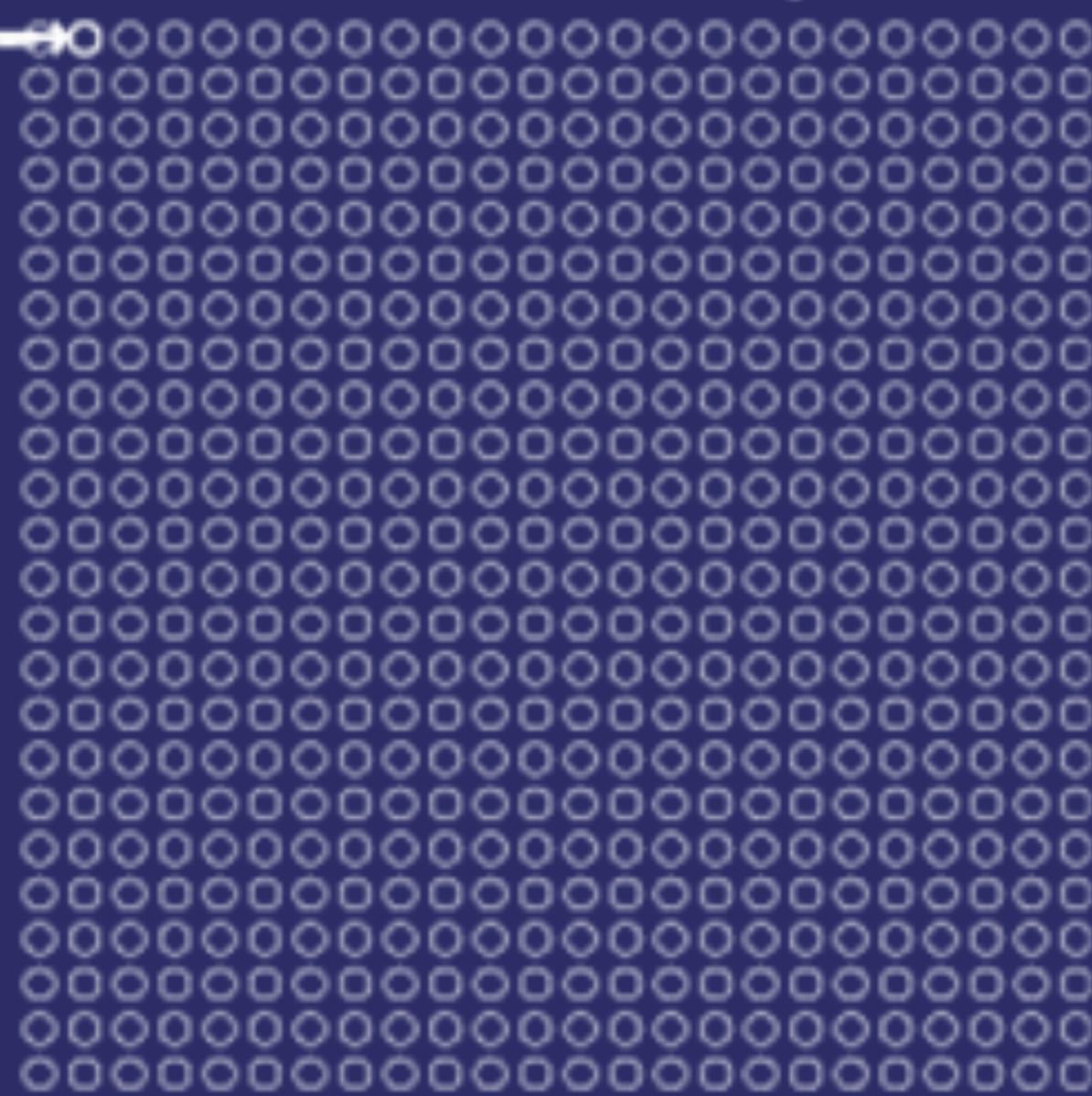
first hidden layer

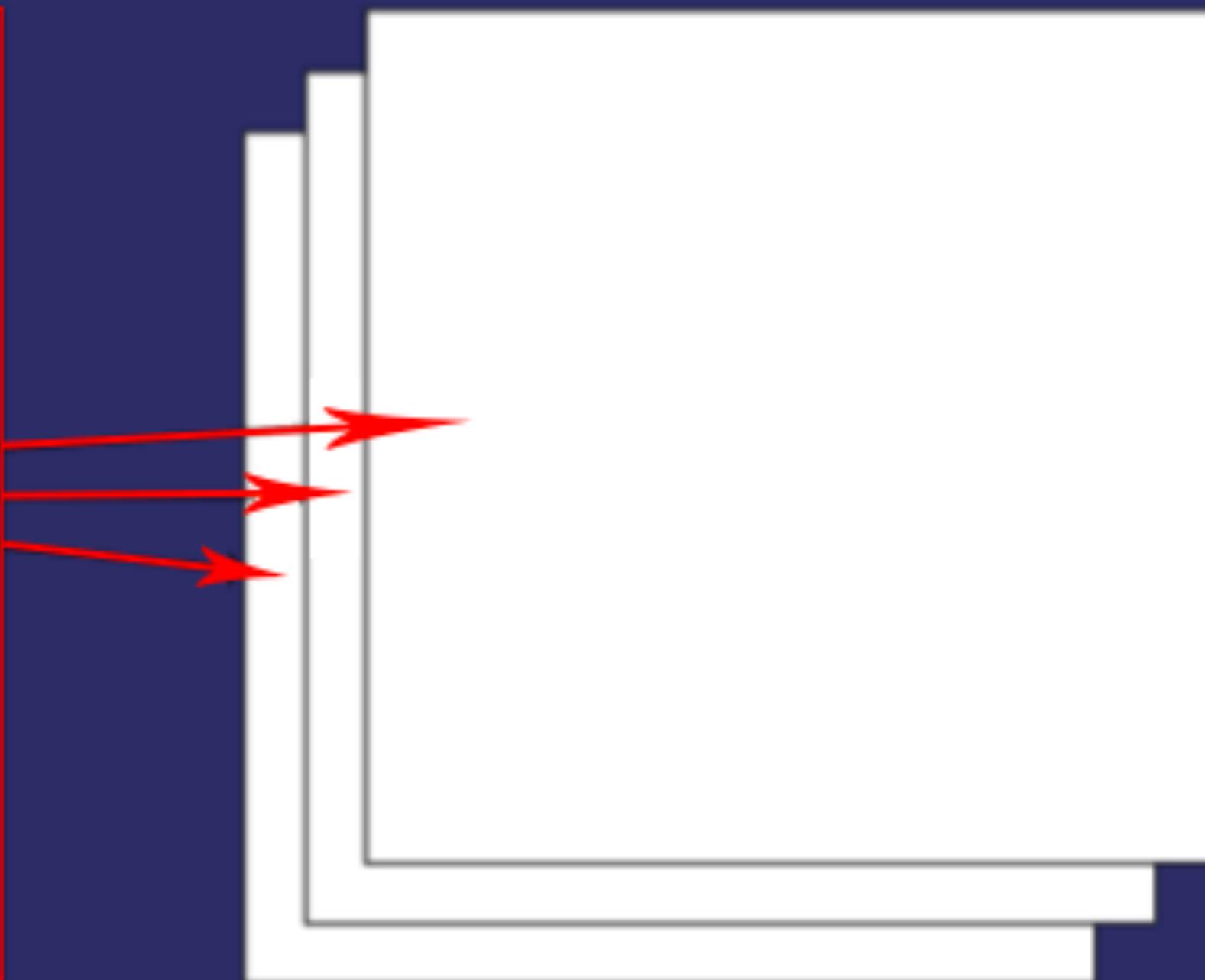


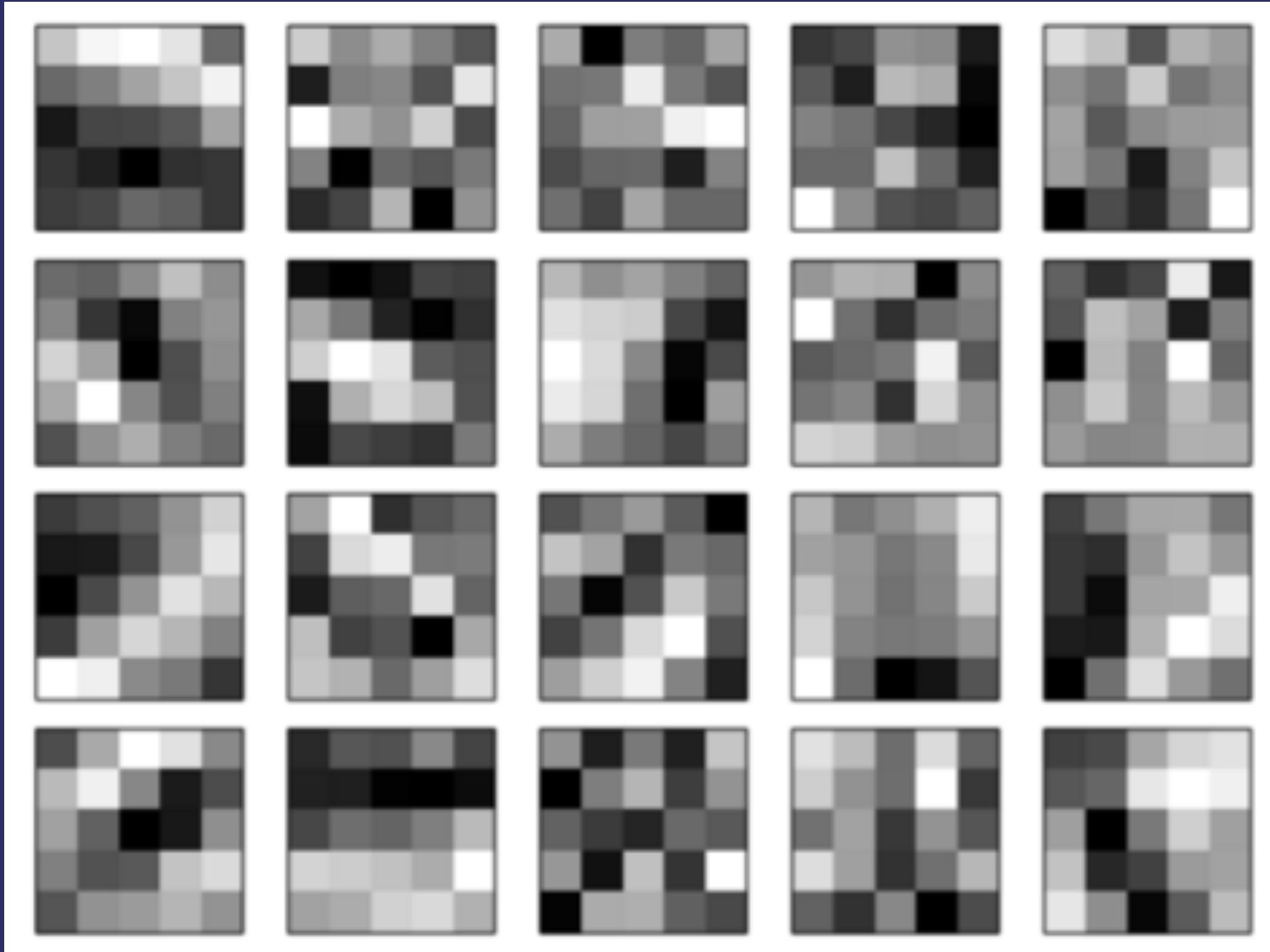
input neurons



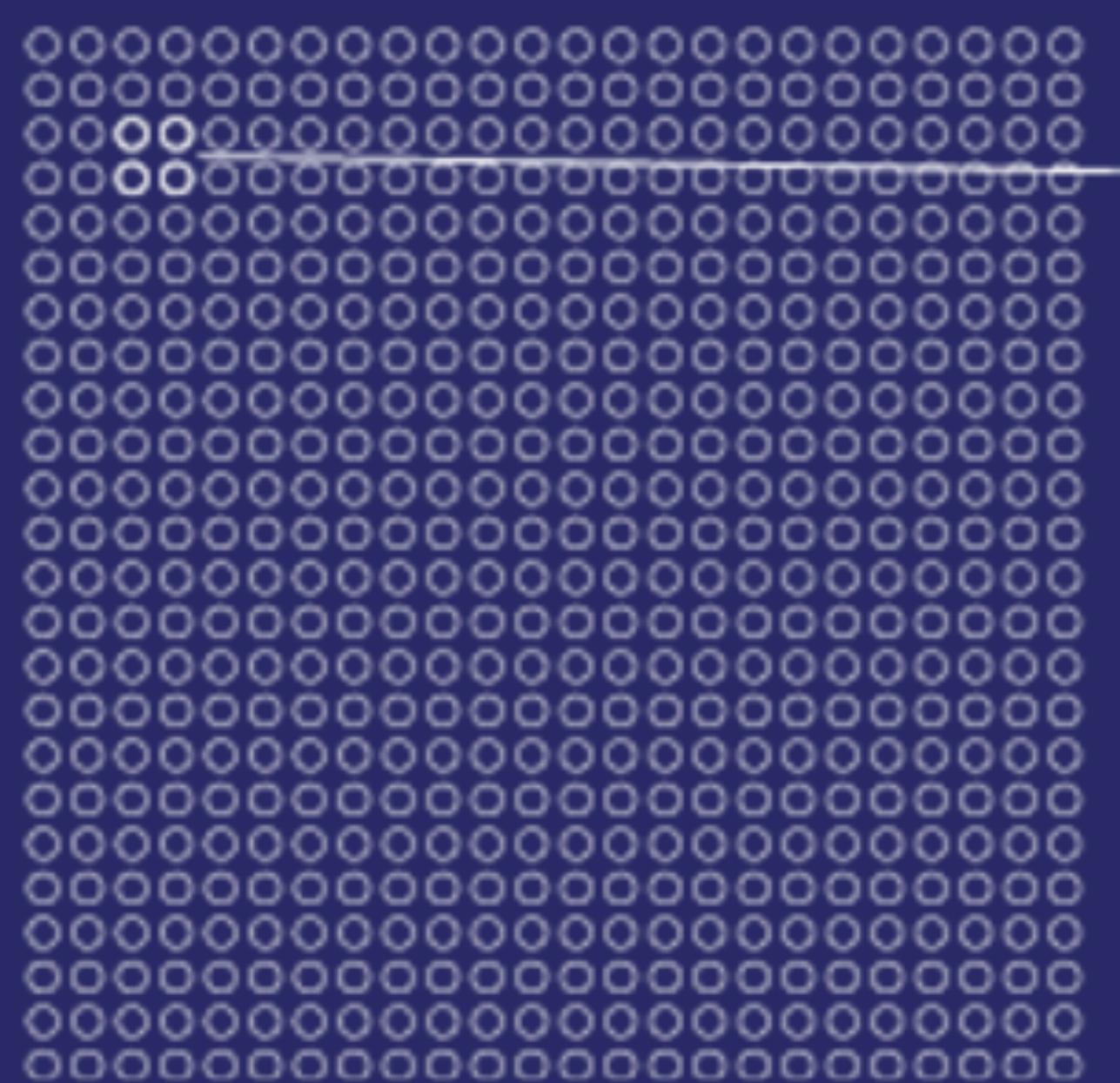
first hidden layer



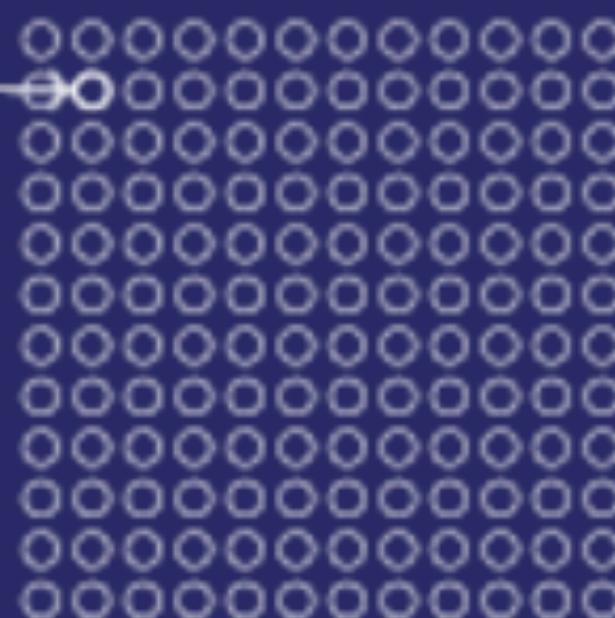




hidden neurons (output from feature map)



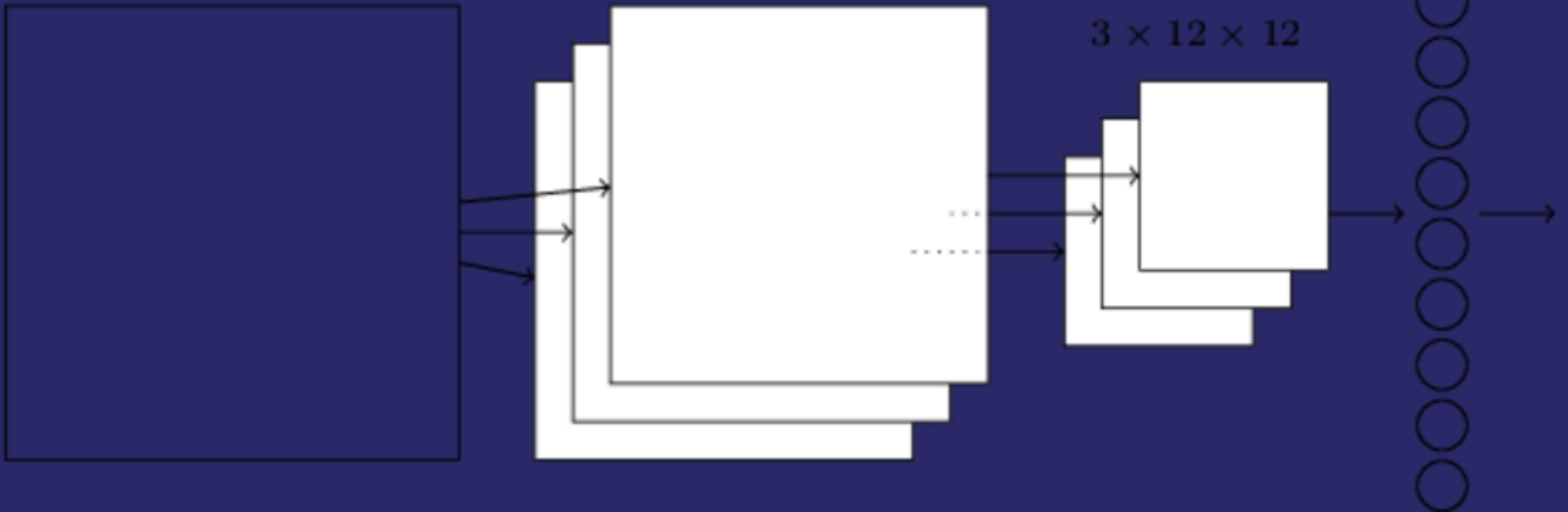
max-pooling units



28×28

$3 \times 24 \times 24$

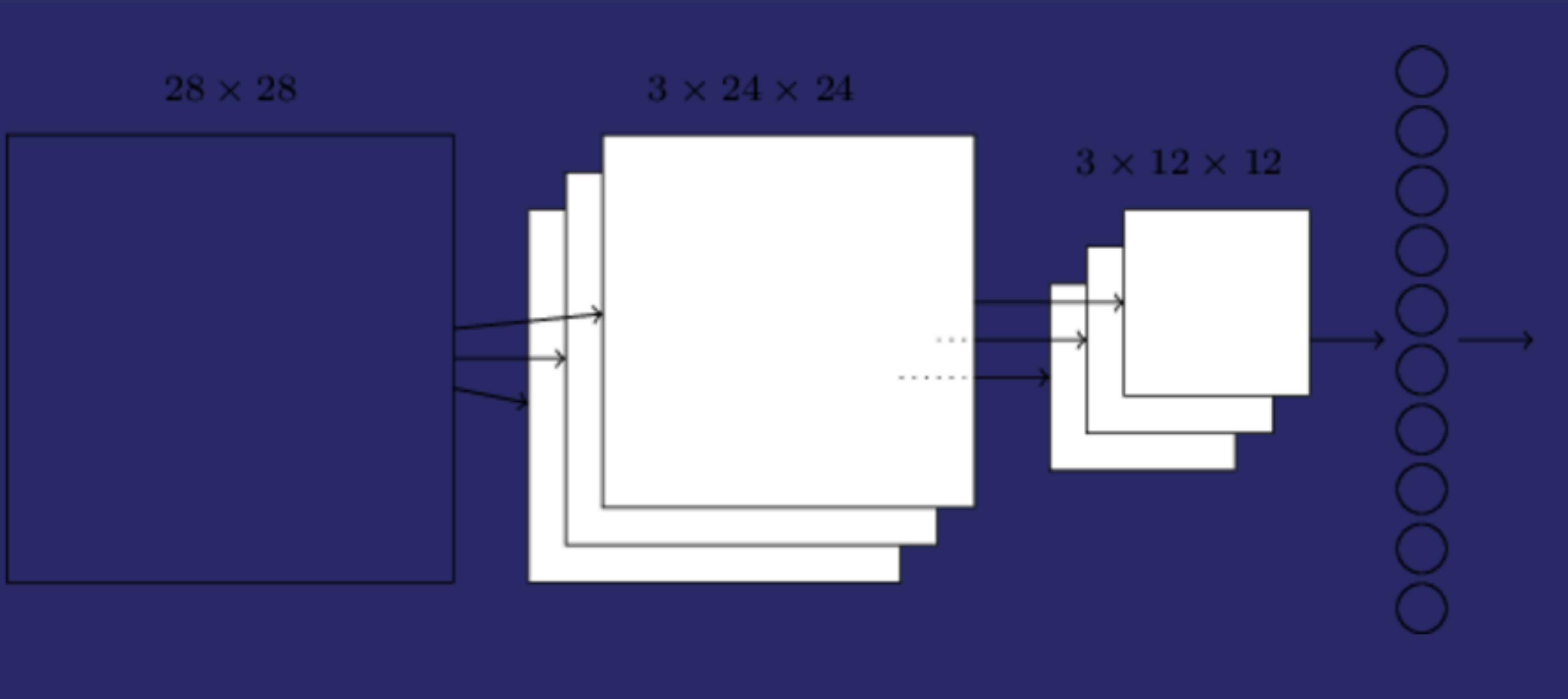
$3 \times 12 \times 12$



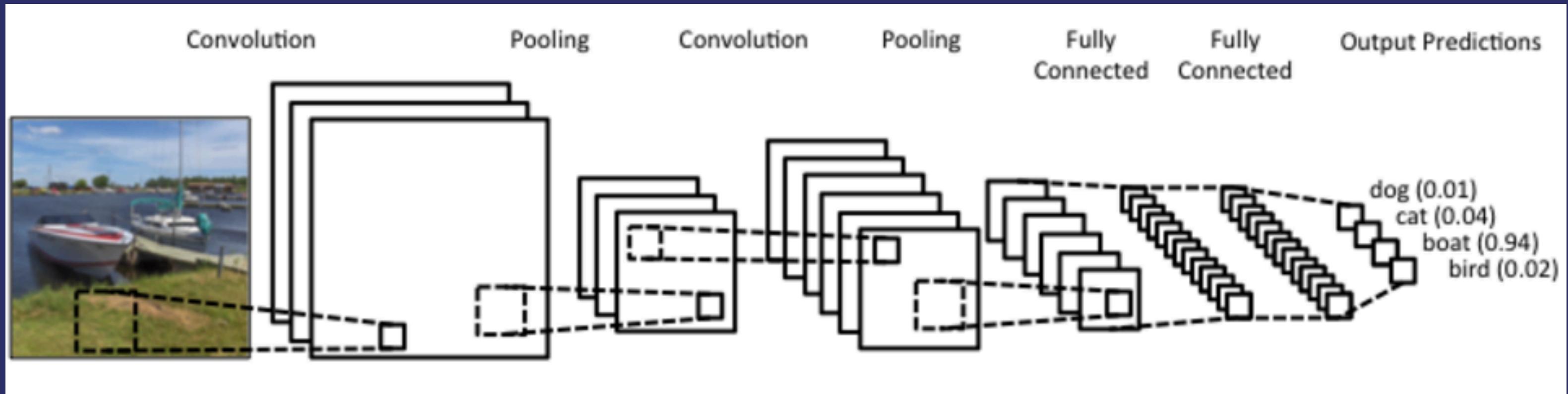
28×28

$3 \times 24 \times 24$

$3 \times 12 \times 12$



Convolution Neural Networks



Convolution vs. Adjacent

- Convolution: Fast to train, multiple items found
- Adjacent: No recognition of spacial structure, great for finding a single item
- Both: Create multilayer neural network

**Did you get
that?**

clarifi

Thank you!

@cassidoo