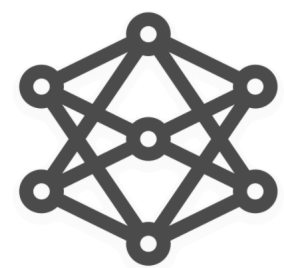


A

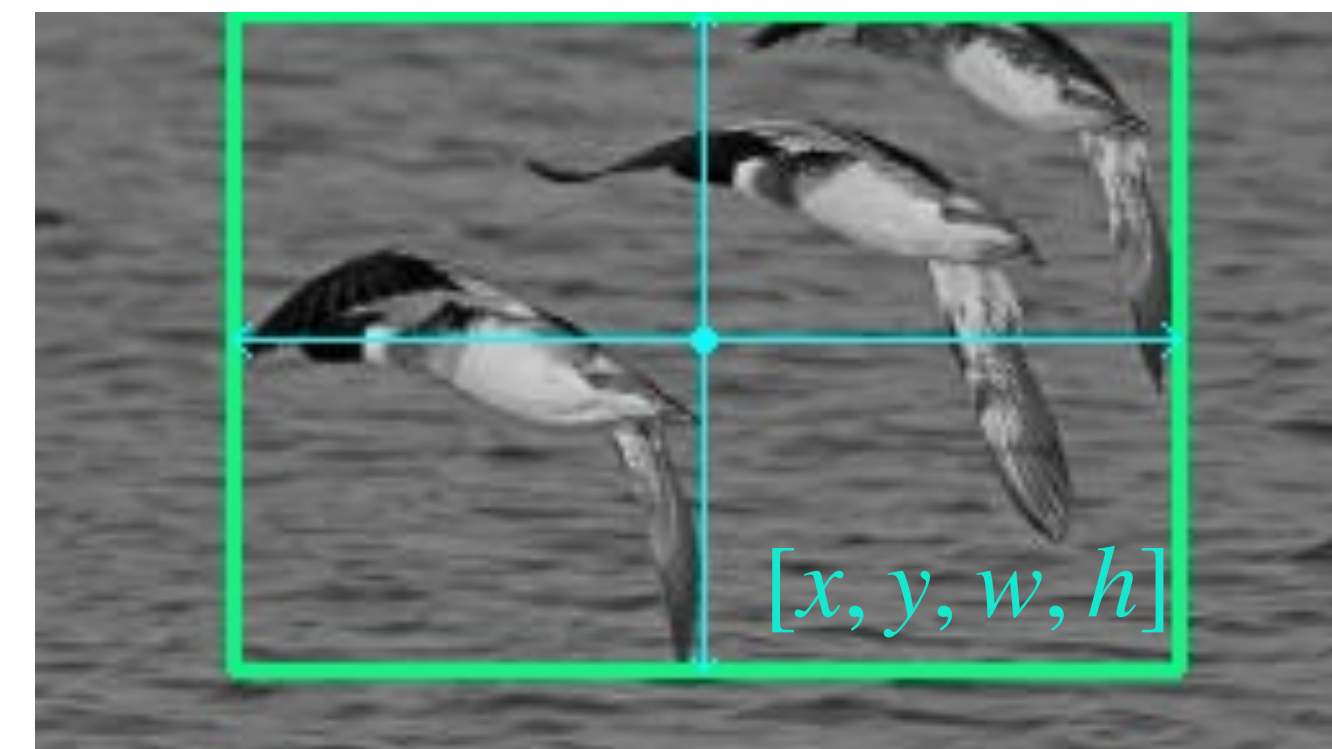
IMAGENET
ILSVRC 2017



Large
YOLOv5l

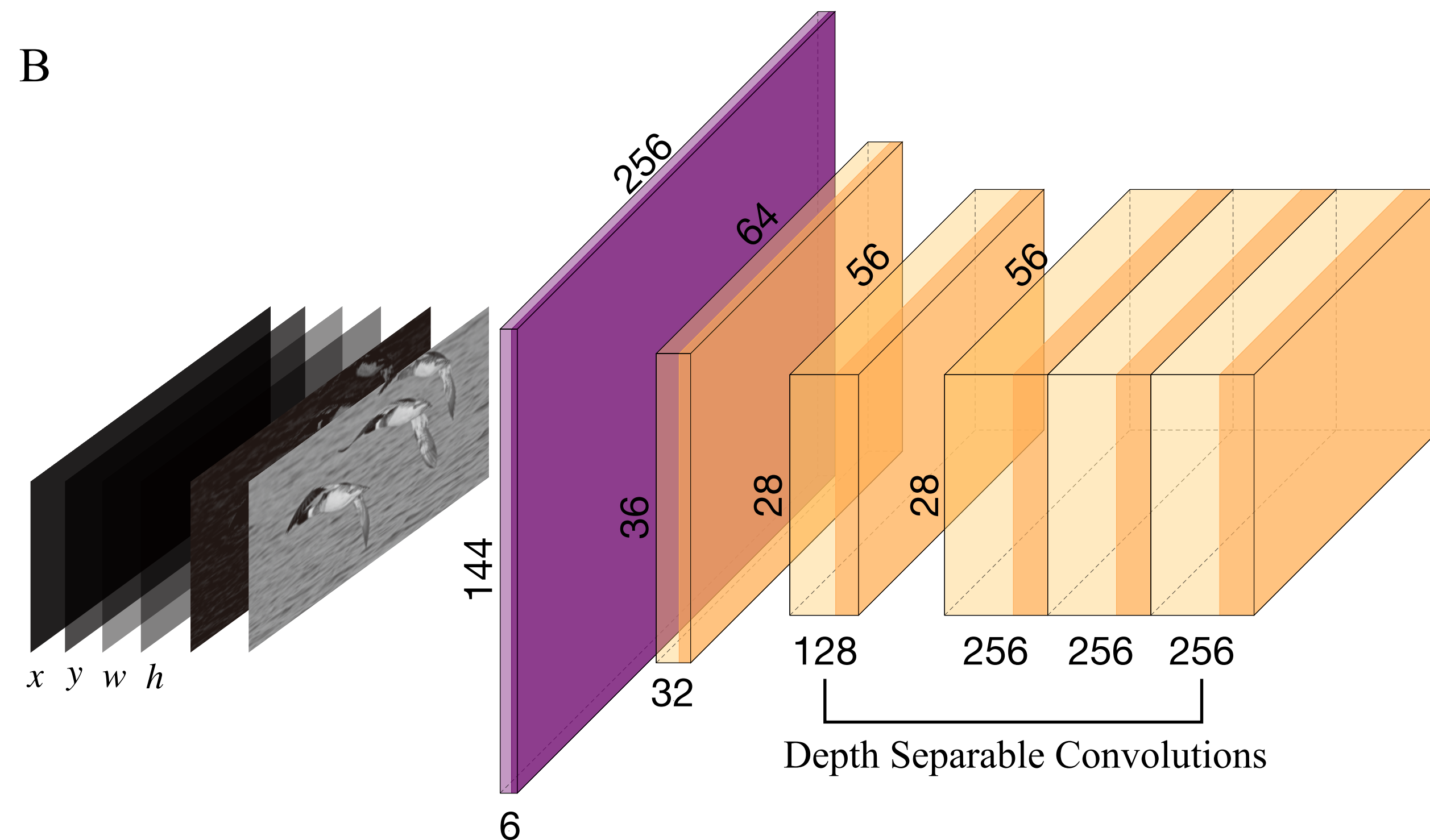


Label objects



Merge bounding boxes

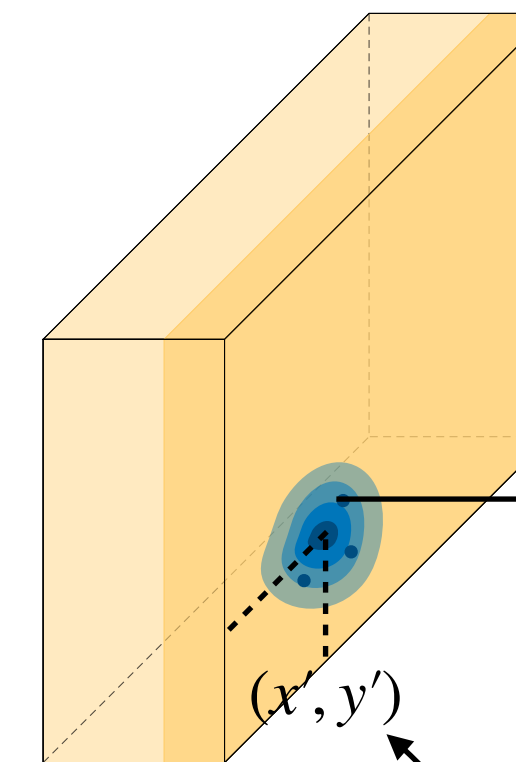
B



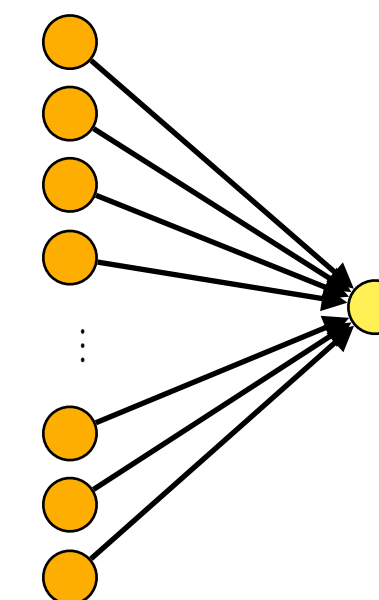
6-channel input

Convolution core

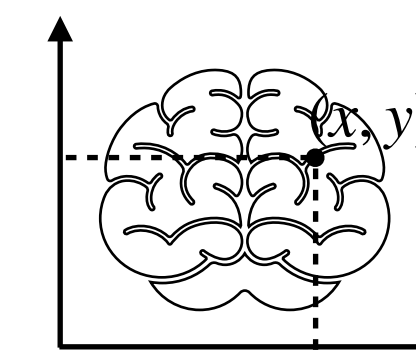
Extracted features



256-d feature



Prediction for
1 neuron



Gaussian readout



Readout position
network