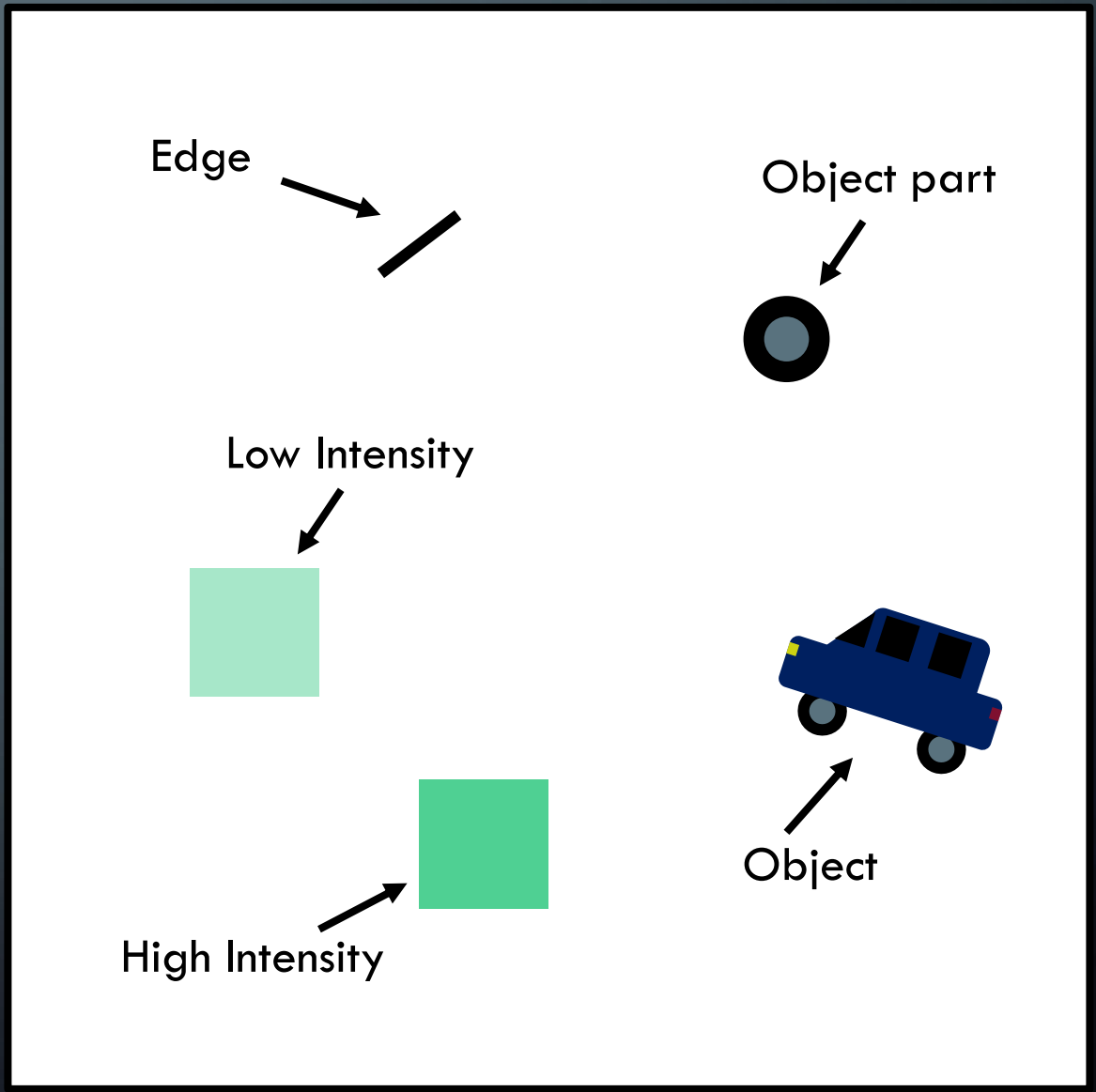
CONVOLUTIONAL DEEP BELIEF NETWORKS FOR SCALABLE UNSUPERVISED LEARNING OF HIERARCHICAL REPRESENTATIONS

LEE, GROSSE, RANGANATH, AND NG (2009)

PRESENTER: ZOE KENDALL

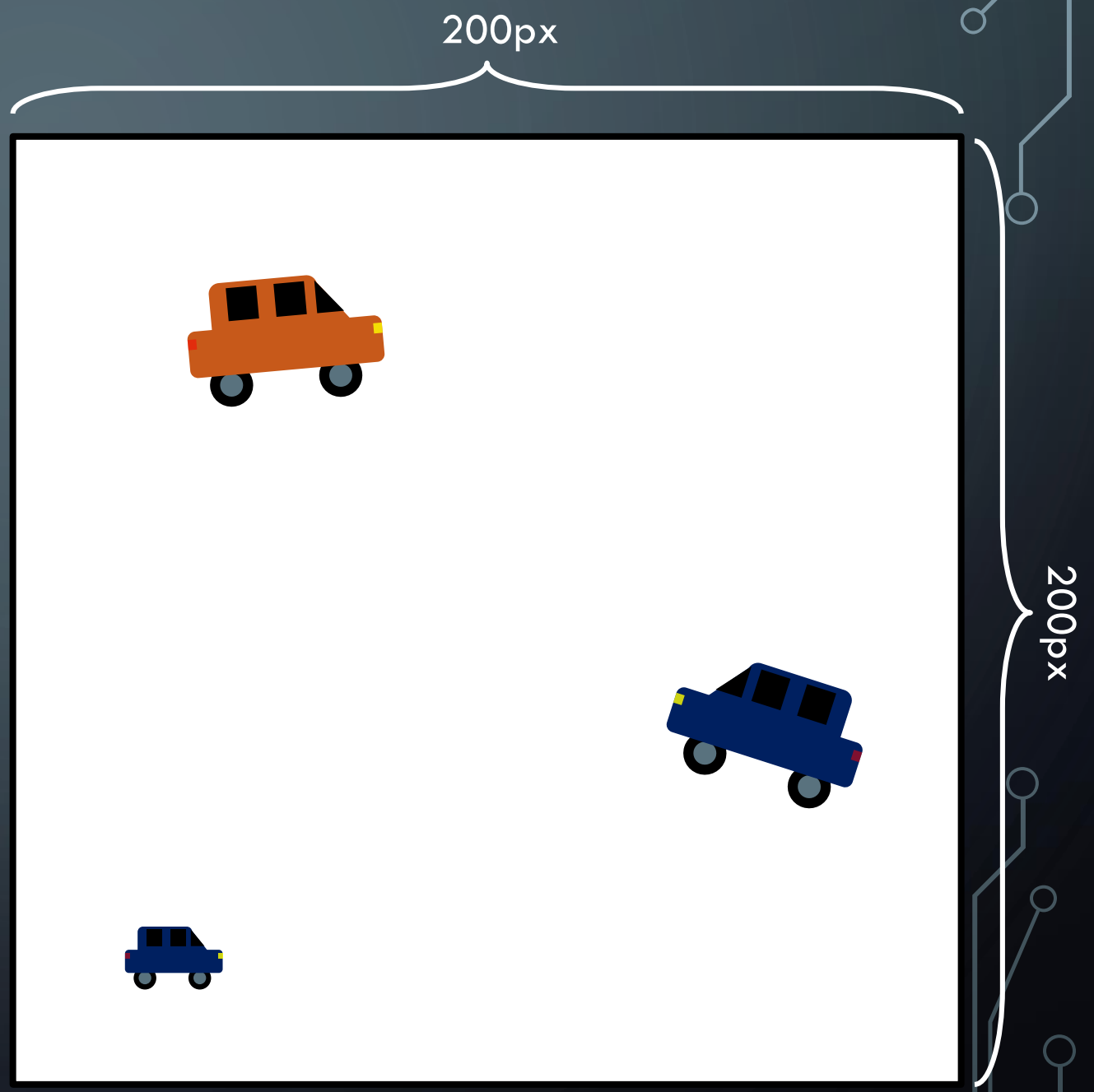
MOTIVATION

- Images contain many layers of complexity.



MOTIVATION

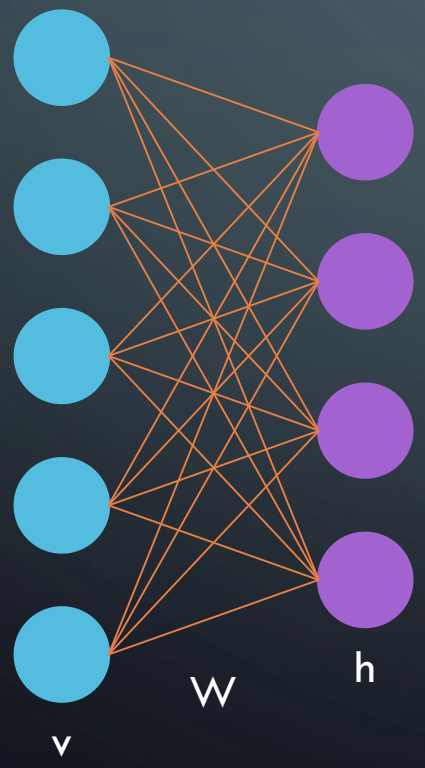
- Images contain many layers of complexity.
- Current issues: scalability and location of features



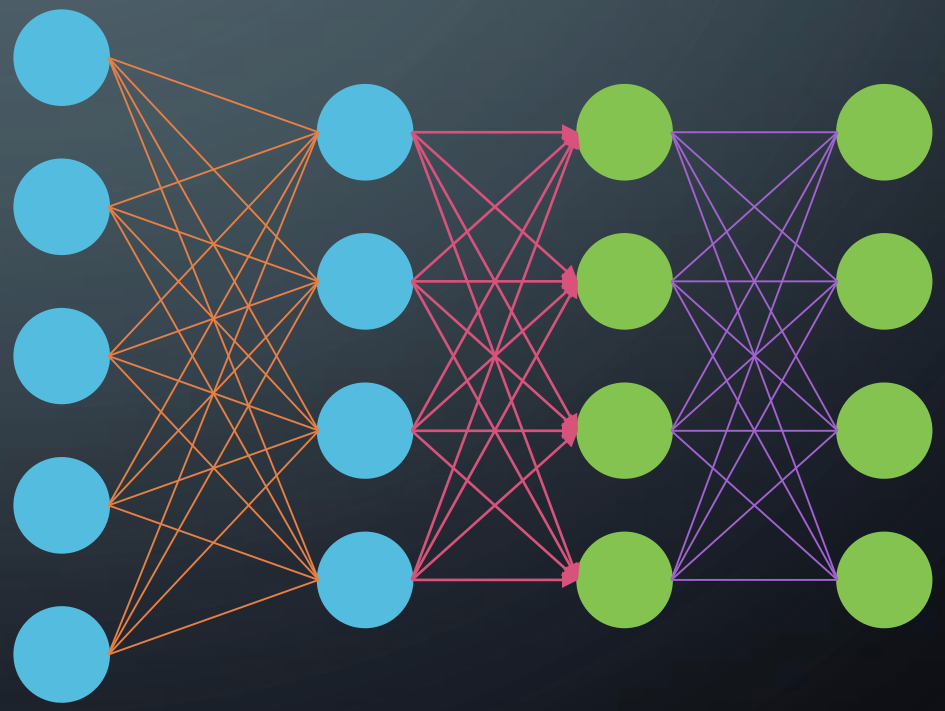
KEY TERMINOLOGY

- Deep Learning (DL) – a type of machine learning inspired by the human brain

Restricted Boltzmann Machine (RBM)



Deep Belief Network (DBN)



CONVOLUTIONAL DEEP BELIEF NETWORK WITH MAX-POOLING

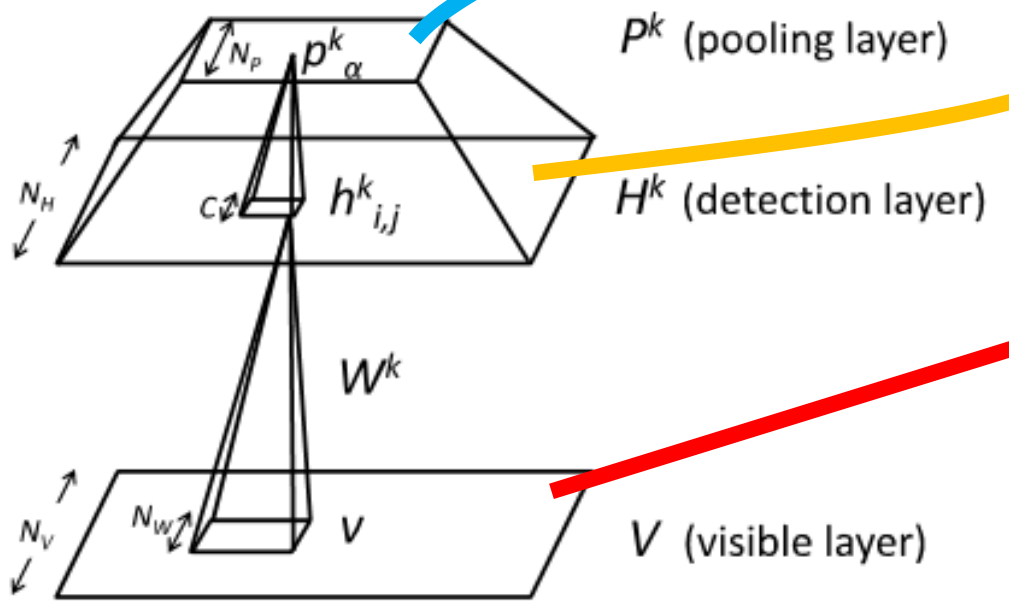
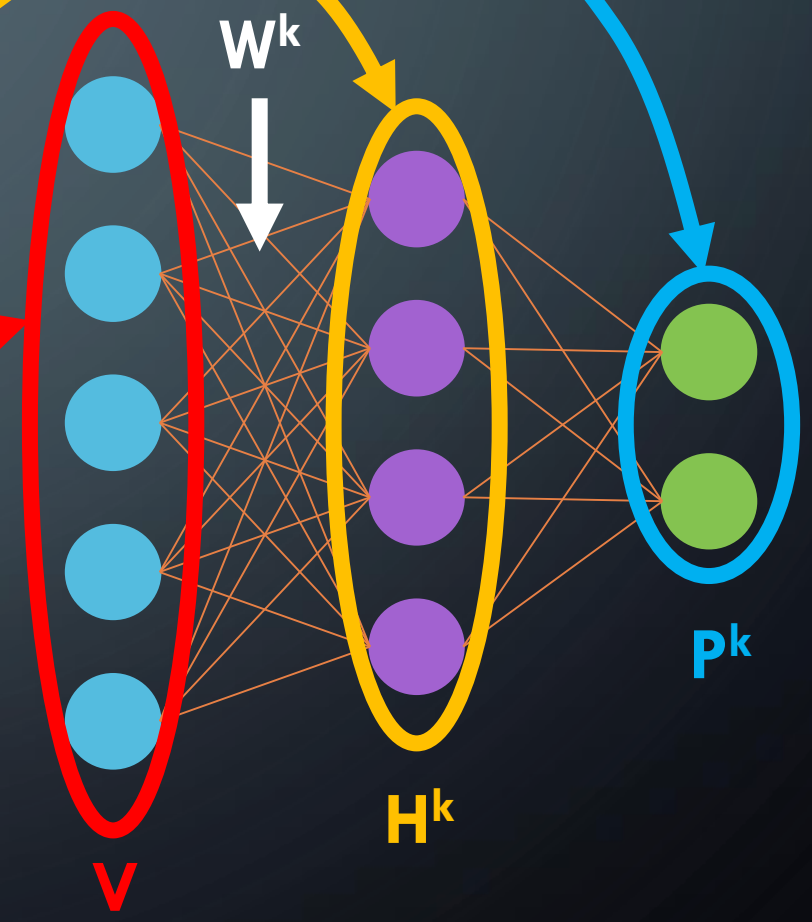
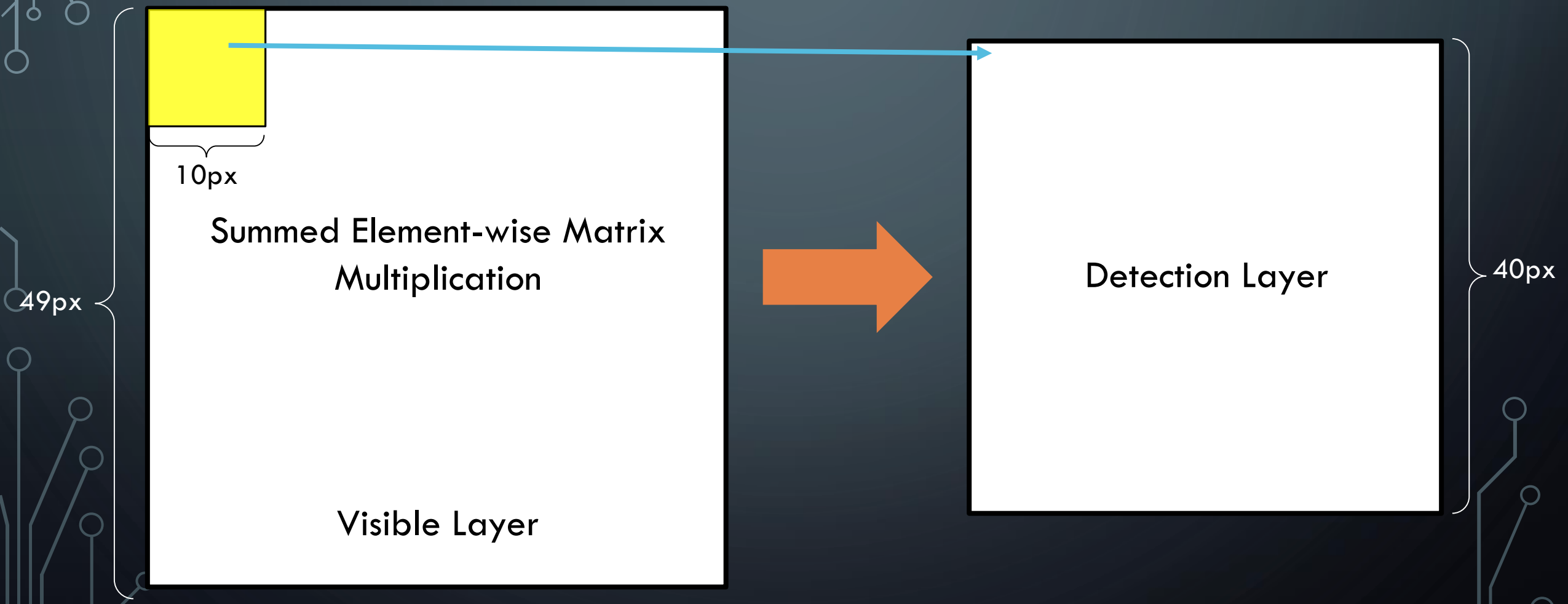


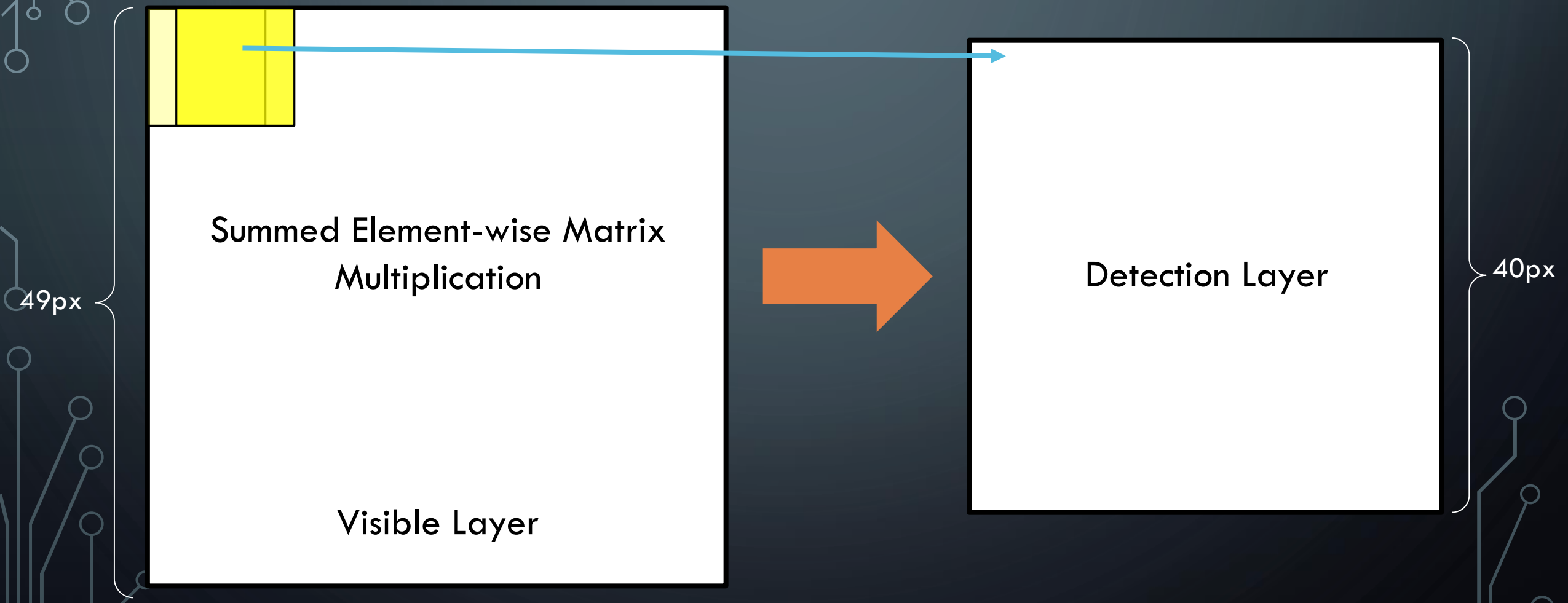
Figure 1. Convolutional RBM with probabilistic max-pooling. For simplicity, only group k of the detection layer and the pooing layer are shown. The basic CRBM corresponds to a simplified structure with only visible layer and detection (hidden) layer. See text for details.



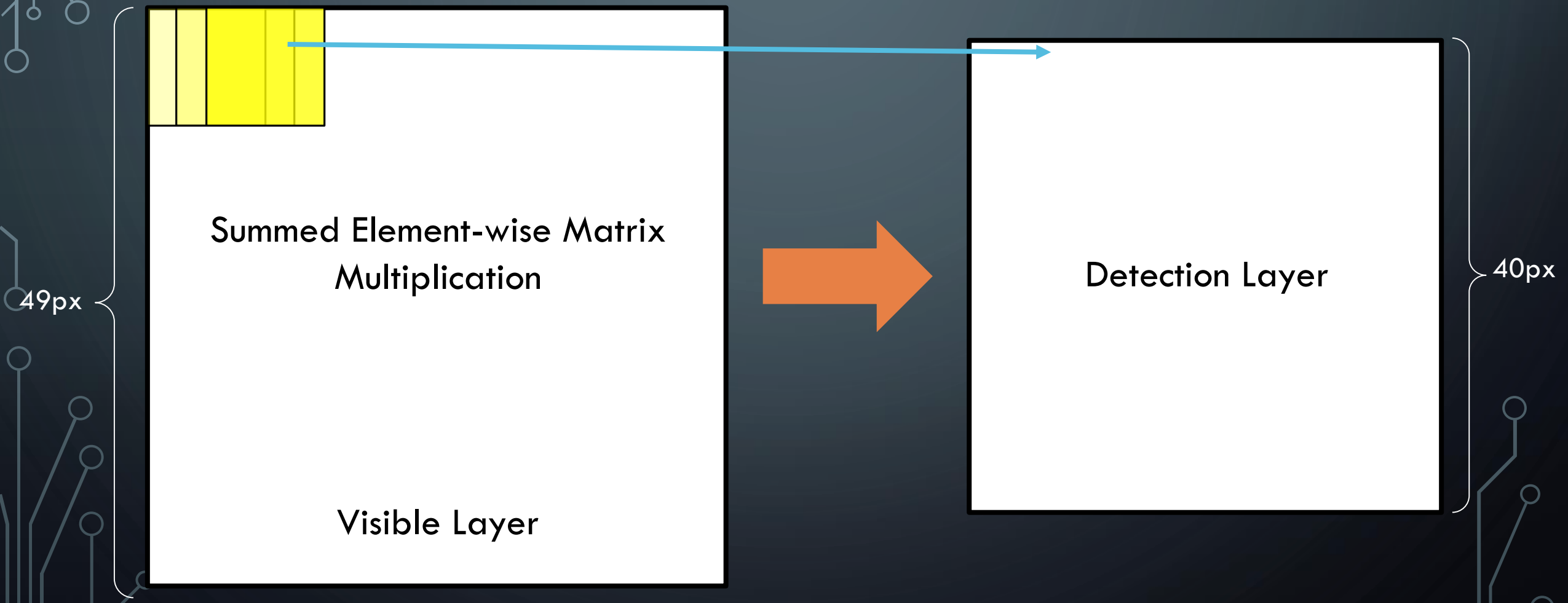
CONVOLUTION FROM VISIBLE TO DETECTION



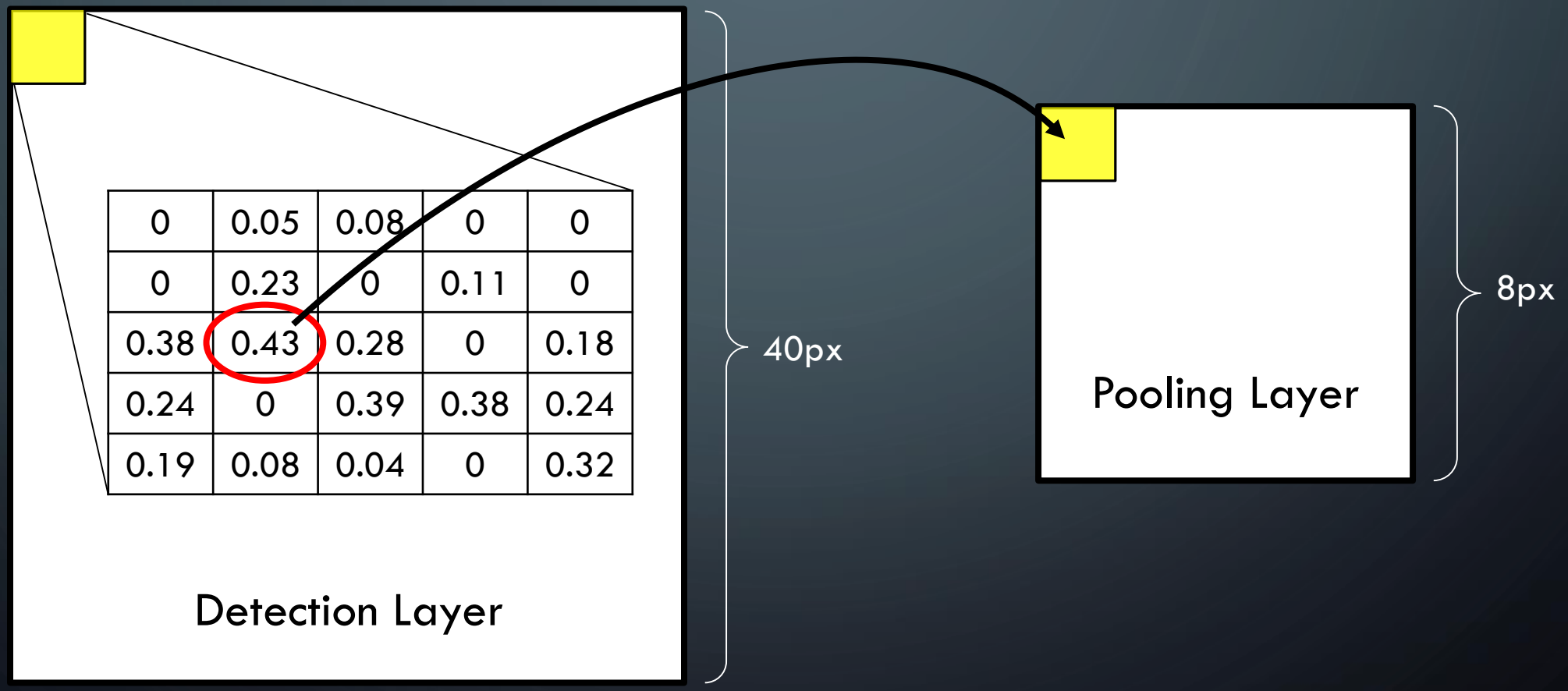
CONVOLUTION FROM VISIBLE TO DETECTION



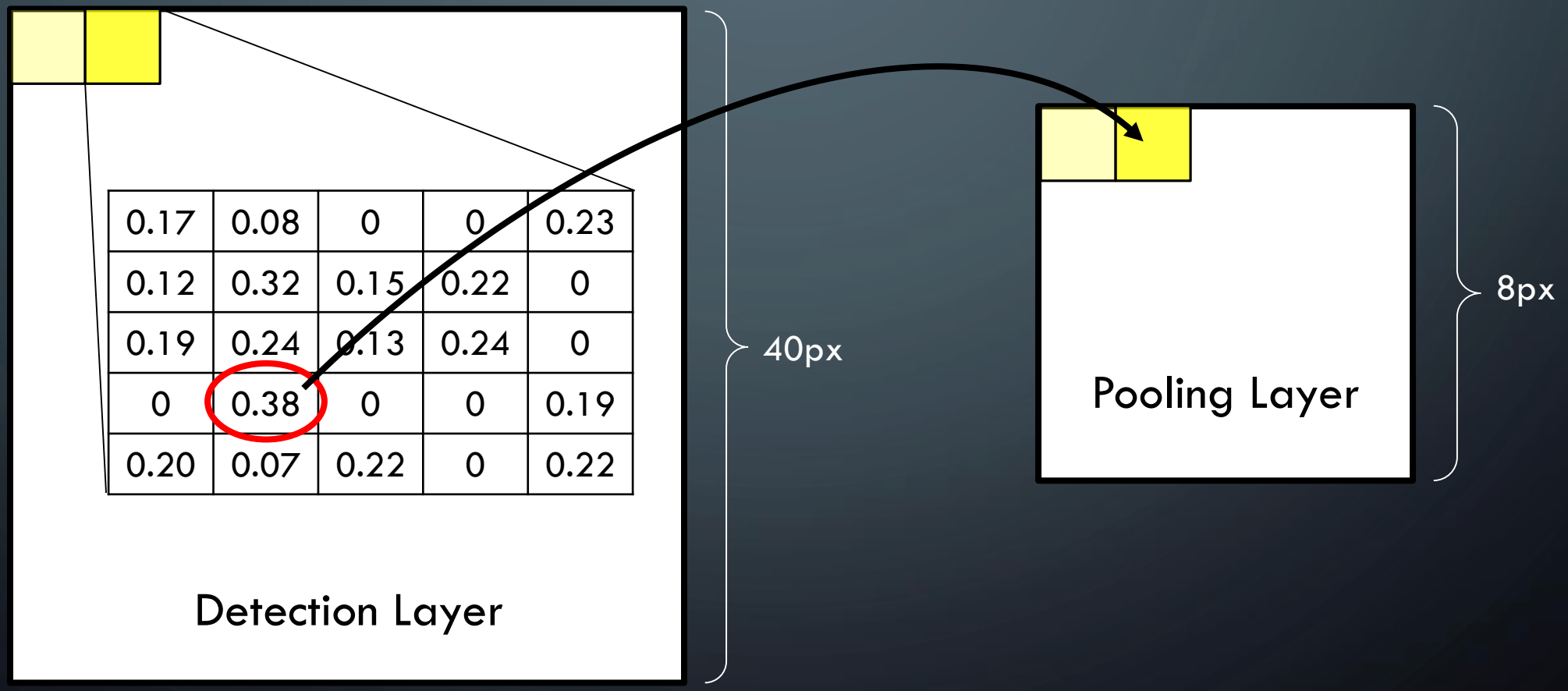
CONVOLUTION FROM VISIBLE TO DETECTION



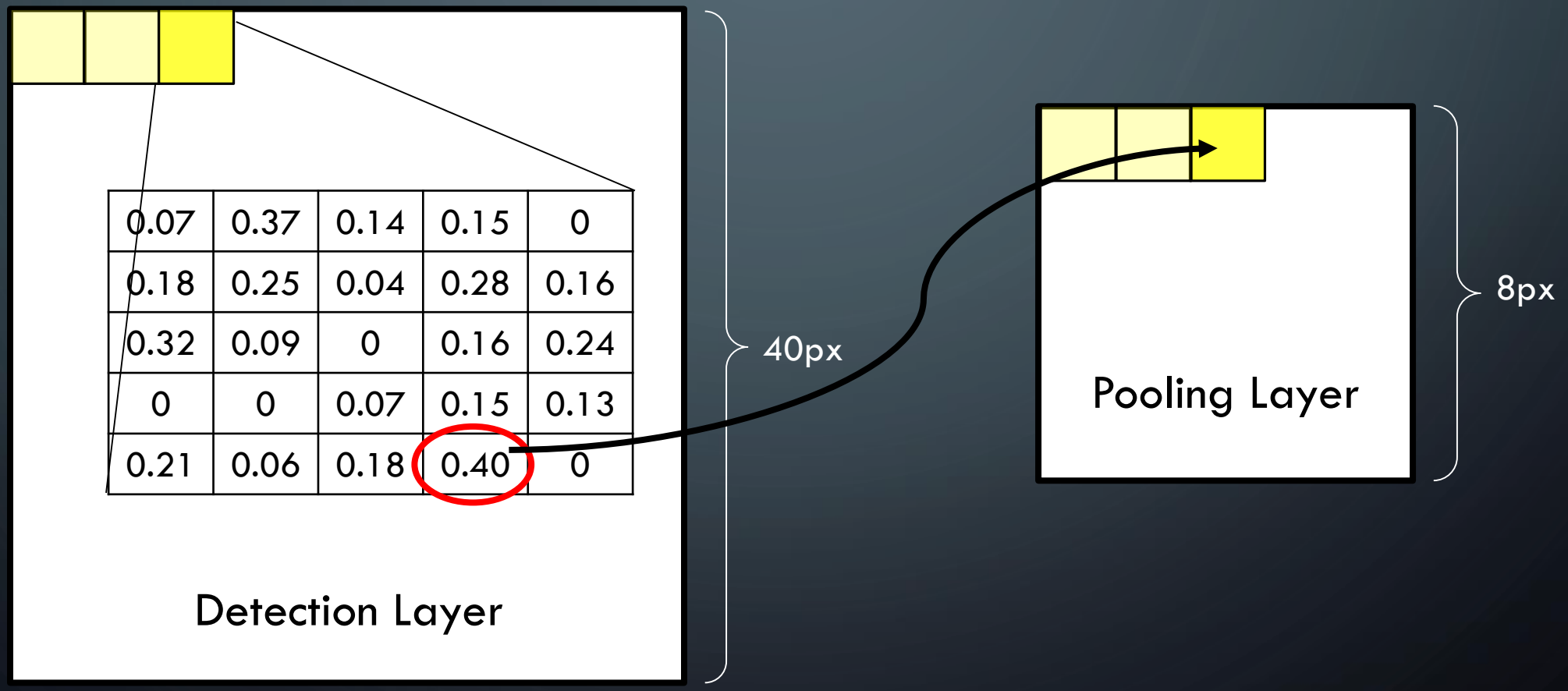
MAX-POOLING


0 0.05 0.08 0 0
0 0.23 0 0.11 0
0.38 0.43 0.28 0 0.18
0.24 0 0.39 0.38 0.24
0.19 0.08 0.04 0 0.32
Detection Layer
40px
Pooling Layer
8px

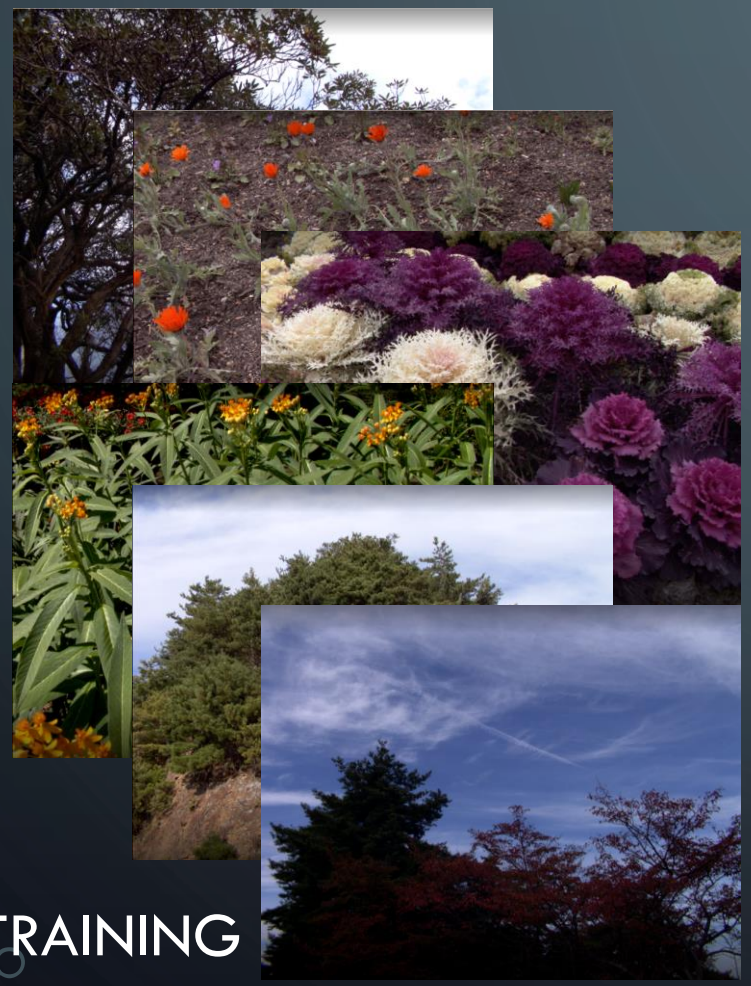
MAX-POOLING


0.17 0.08 0 0 0.23
0.12 0.32 0.15 0.22 0
0.19 0.24 0.13 0.24 0
0 0.38 0 0 0.19
0.20 0.07 0.22 0 0.22
Detection Layer
40px
Pooling Layer
8px

MAX-POOLING


0.07 0.37 0.14 0.15 0
0.18 0.25 0.04 0.28 0.16
0.32 0.09 0 0.16 0.24
0 0 0.07 0.15 0.13
0.21 0.06 0.18 0.40 0
Detection Layer
40px
Pooling Layer
8px

ABILITY TO CLASSIFY IMAGES IN SEMI-SUPERVISED SETTING



TRAINING

TESTED ON THE CALTECH-101 DATASET

Table 1. Classification accuracy for the Caltech-101 data

Training Size	15	30
CDBN (first layer)	53.2±1.2%	60.5±1.1%
CDBN (first+second layers)	57.7±1.5%	65.4±0.5%
Raina et al. (2007)	46.6%	-
Ranzato et al. (2007)	-	54.0%
Mutch and Lowe (2006)	51.0%	56.0%
Lazebnik et al. (2006)	54.0%	64.6%
Zhang et al. (2006)	59.0±0.56%	66.2±0.5%

UNSUPERVISED OBJECT RECOGNITION

CALTECH-101 DATASET

Faces | Cars | Elephants | Chairs | Combination

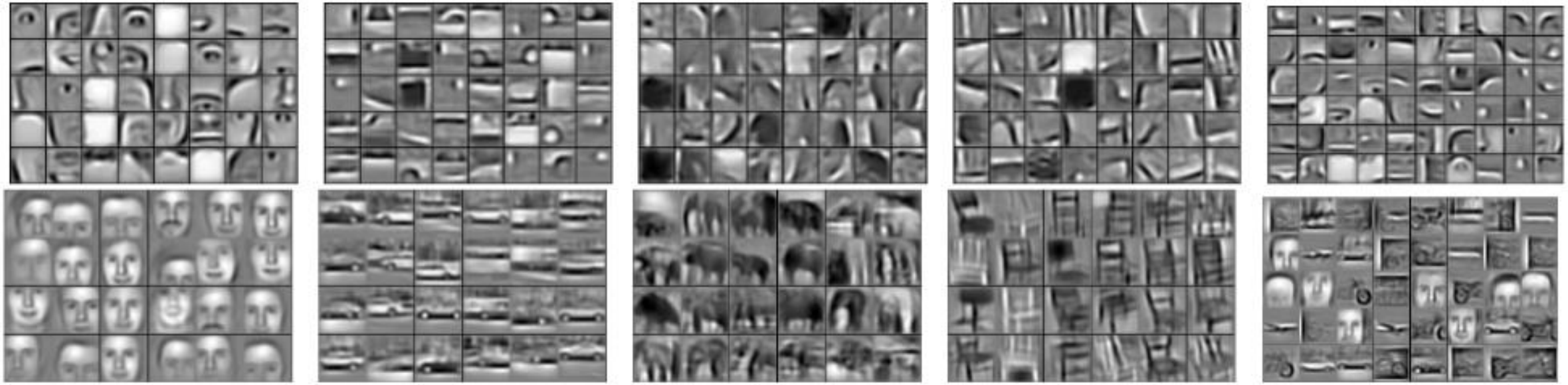


Figure 3. Columns 1-4: the second layer bases (top) and the third layer bases (bottom) learned from specific object categories. Column 5: the second layer bases (top) and the third layer bases (bottom) learned from a mixture of four object categories (faces, cars, airplanes, motorbikes).

HIERARCHICAL PROBABILISTIC INFERENCE

Figure 6. Hierarchical probabilistic inference. For each column: (top) input image. (middle) reconstruction from the second layer units after single bottom-up pass, by projecting the second layer activations into the image space. (bottom) reconstruction from the second layer units after 20 iterations of block Gibbs sampling.

- Model combines bottom-up input and top-down context
- This improves representation.

CONCLUSIONS

- The authors' convolutional DBN was scalable to larger images.
- It could learn hierarchical representations from unlabeled images.
- Future work: use on high-dimensional, complex data