Deep CNNs

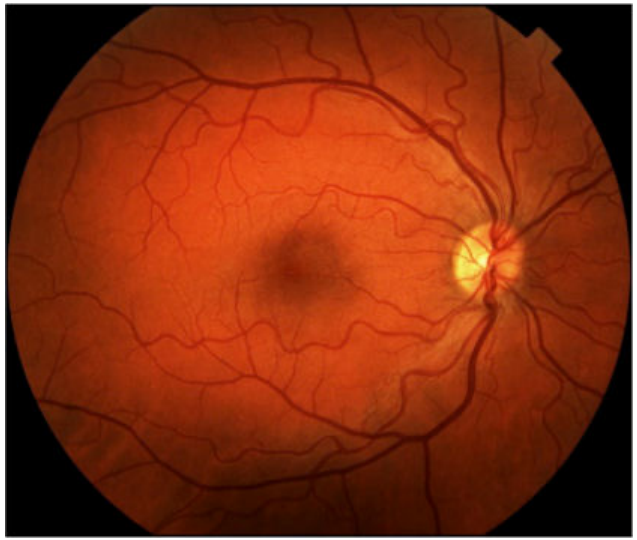
Matthew Engelhard

Many slides created by Tim Dunn

CNNs Take Advantage of Repeated, Hierarchical Structure in Images

Deep Learning for Image Analysis

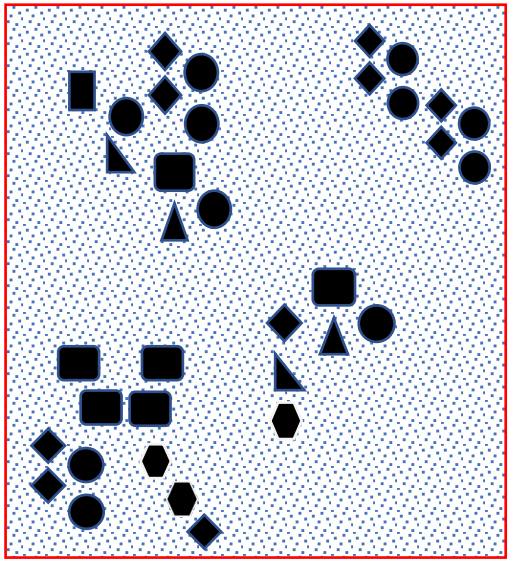
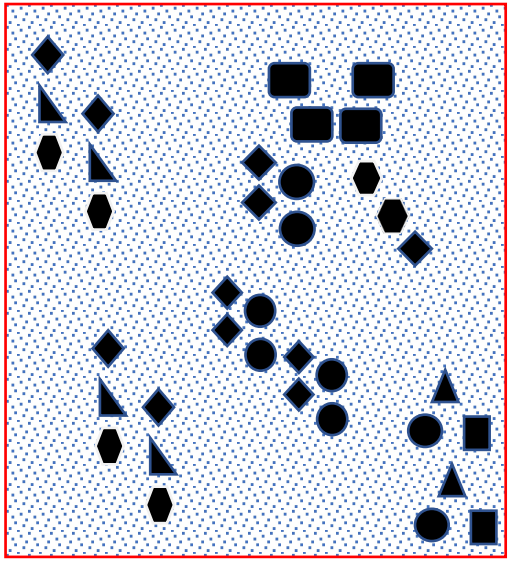
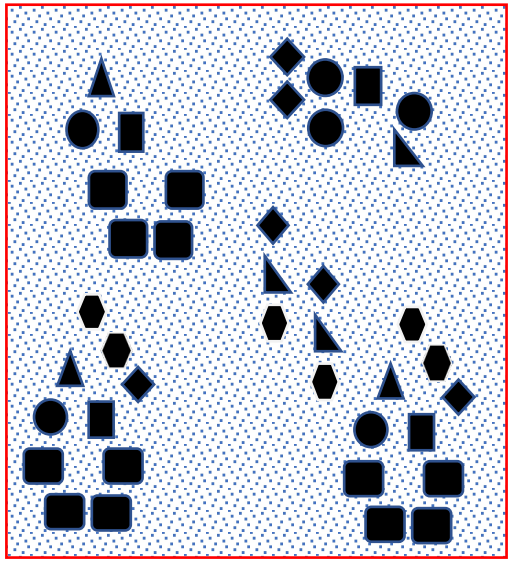
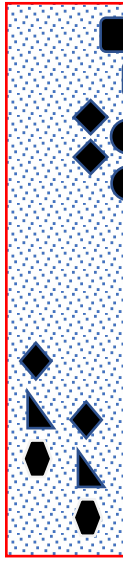
Diabetic Retinopathy
Classification



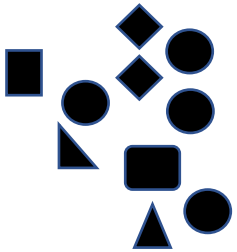
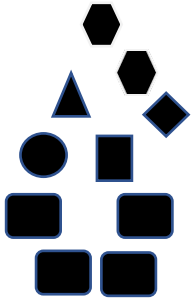
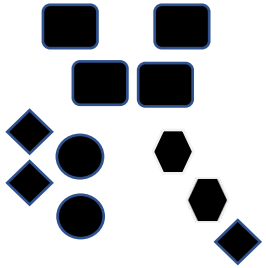
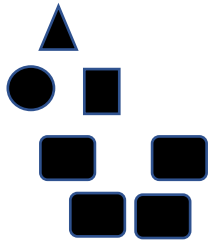
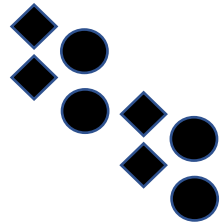
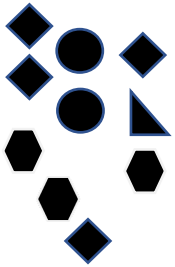
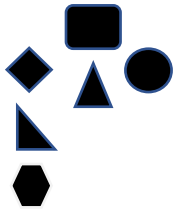
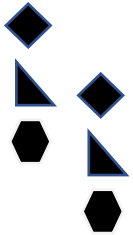
Healthy Retina

Unhealthy Retina

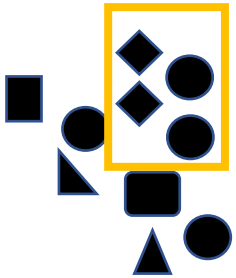
Consider a Set of “Toy” Images,
for illustration of how this structure can be extracted by an algorithm

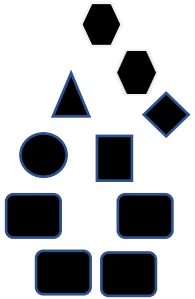





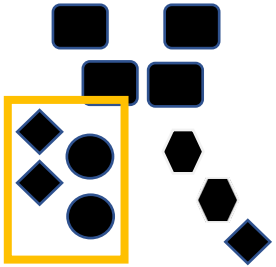
High-Level Motifs/Structure

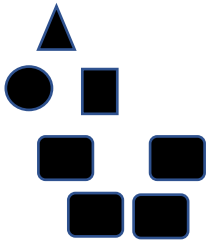









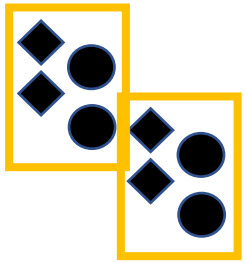
Shared Substructure Within Motifs

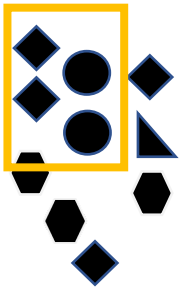


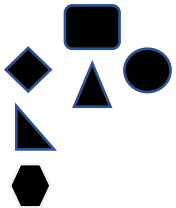


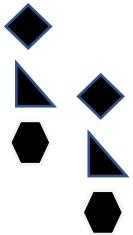


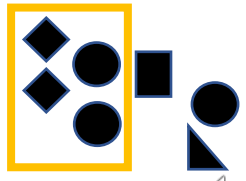




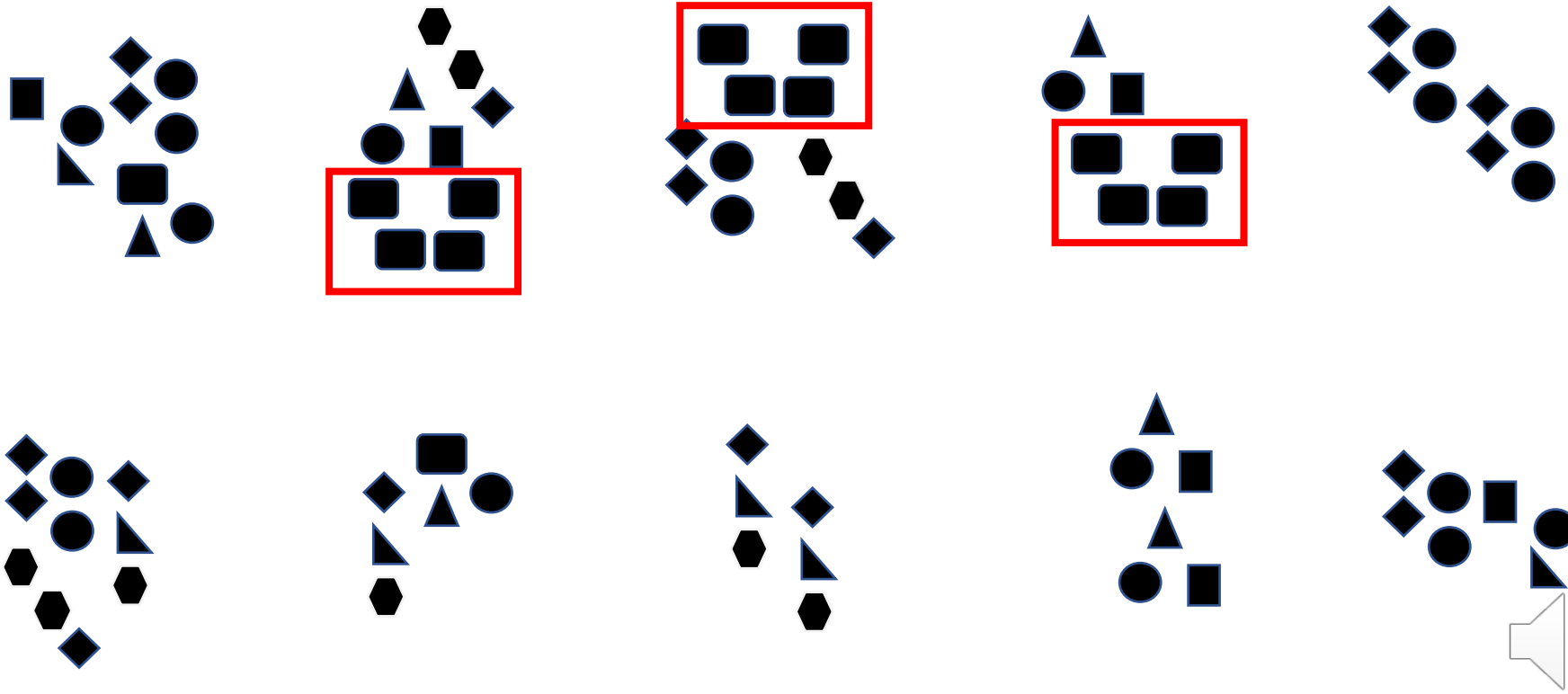




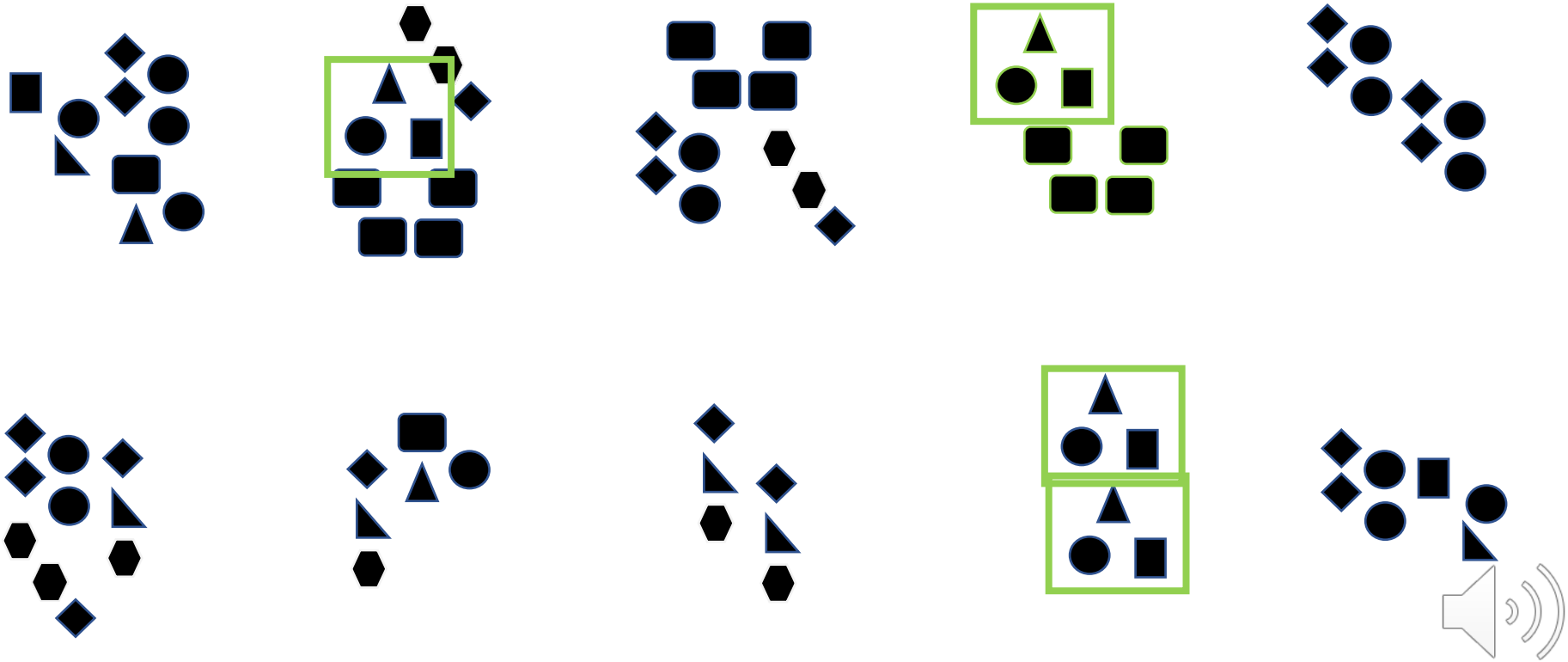




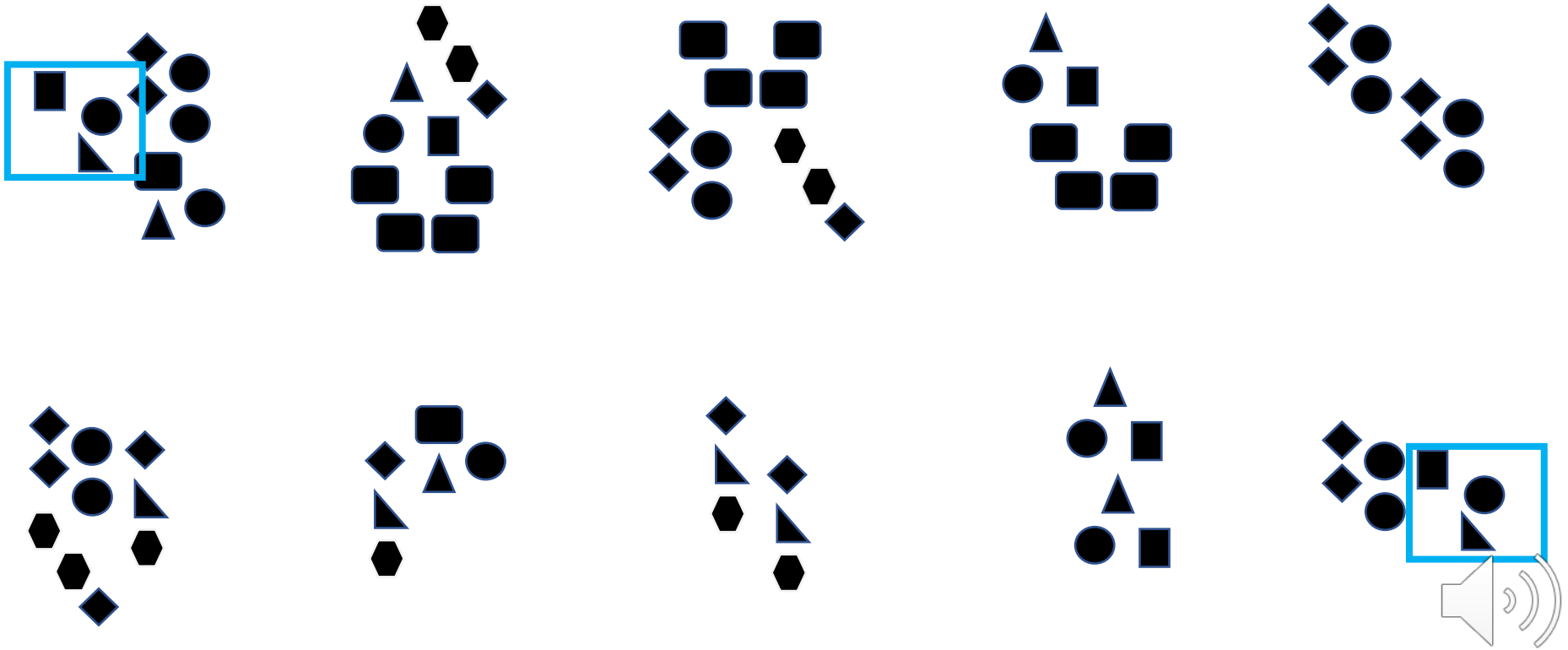
Shared Substructure Within Motifs



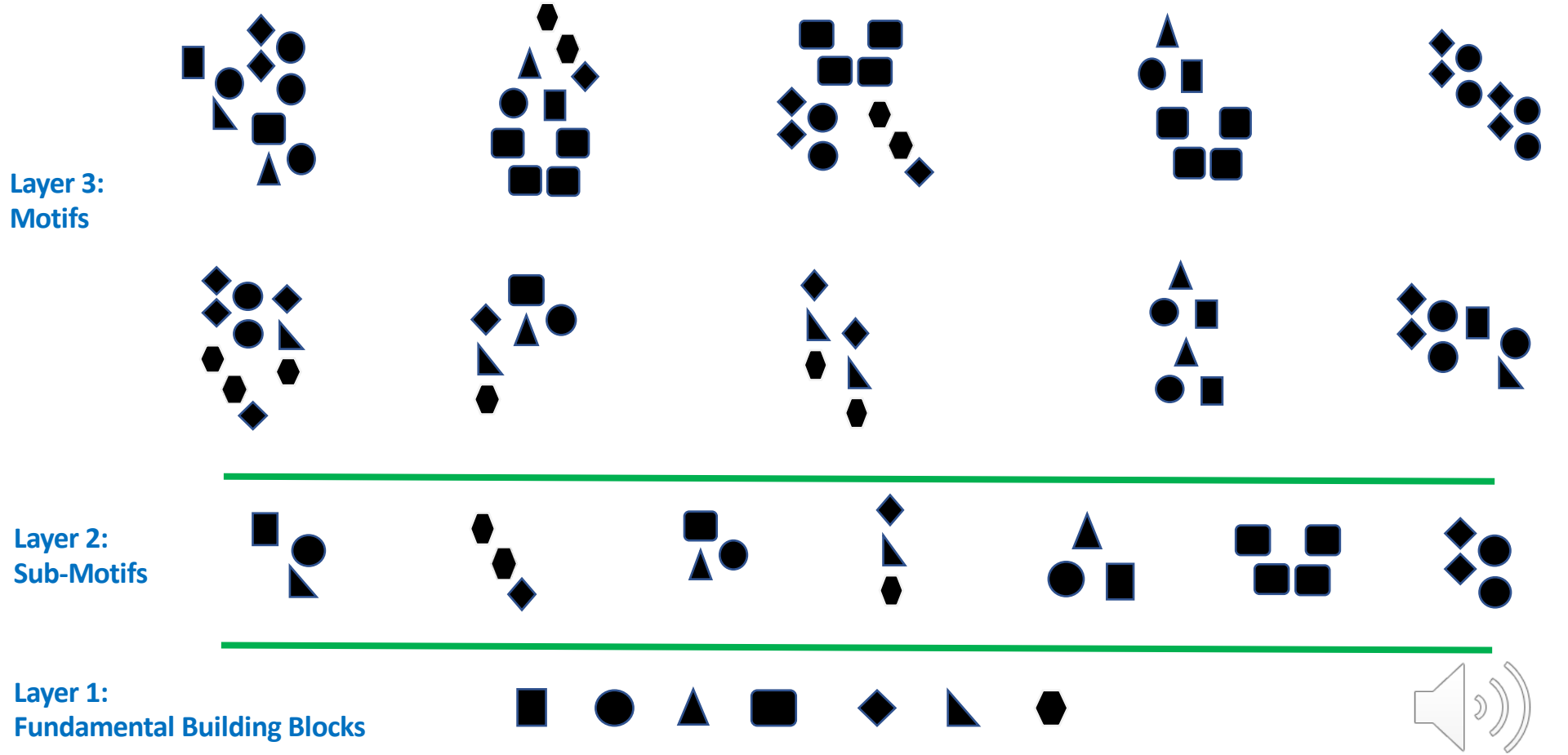
Shared Substructure Within Motifs



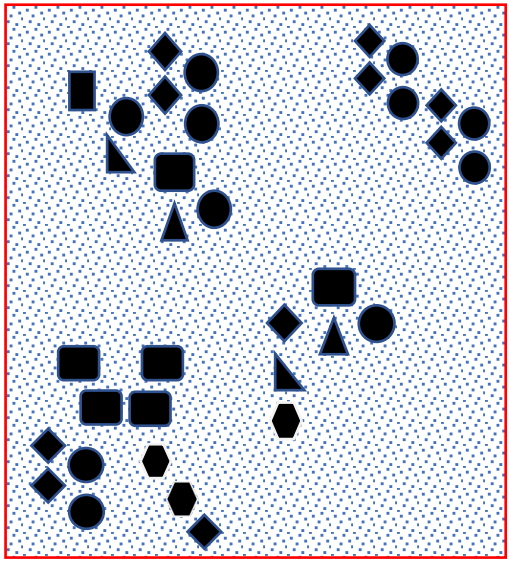
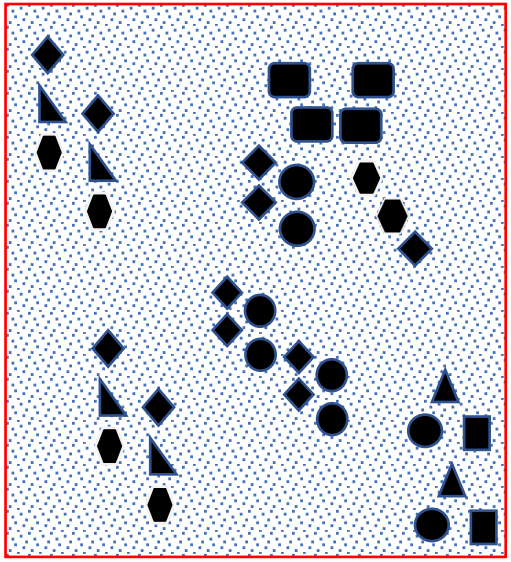
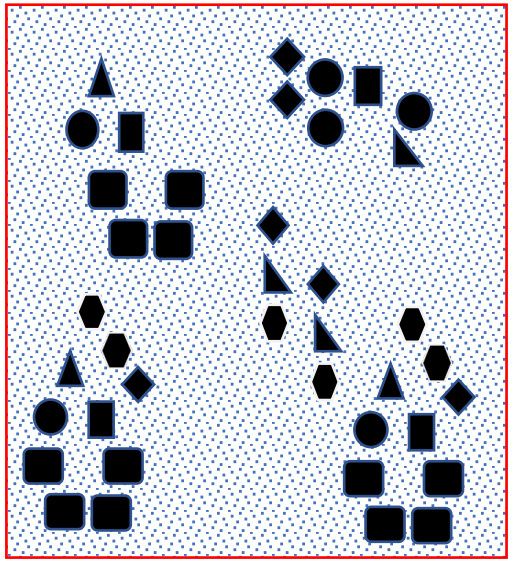
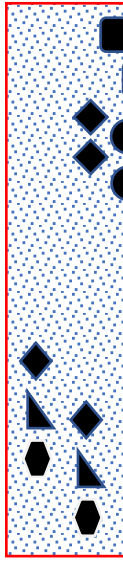
Shared Substructure Within Motifs



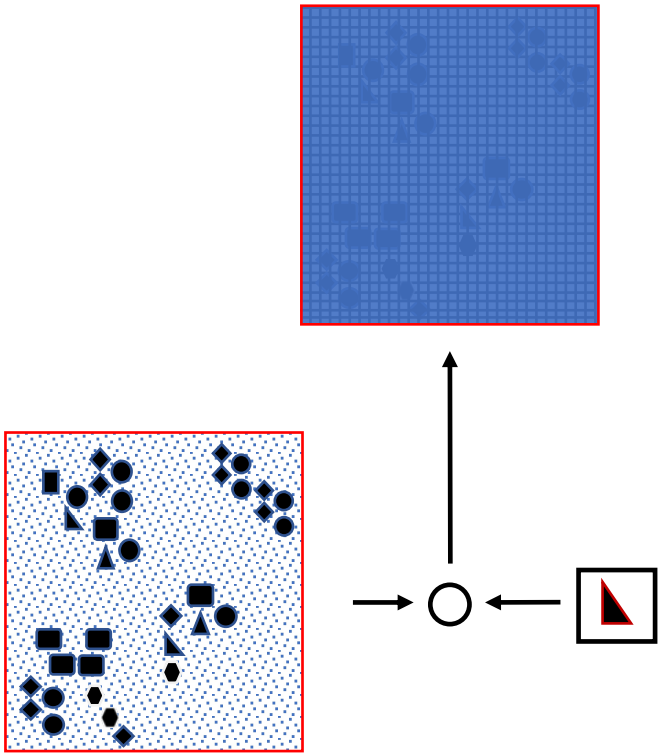
Hierarchical Representation of Images



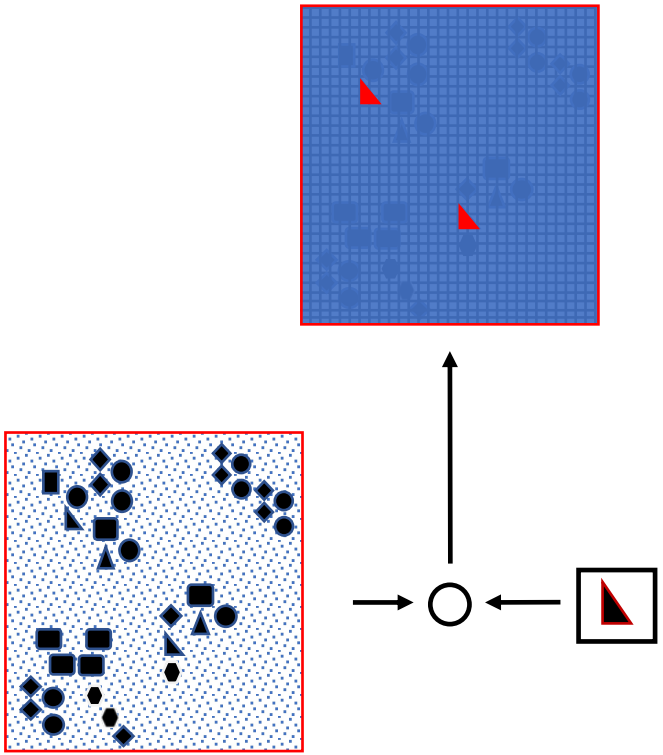
Recall the Data/Images

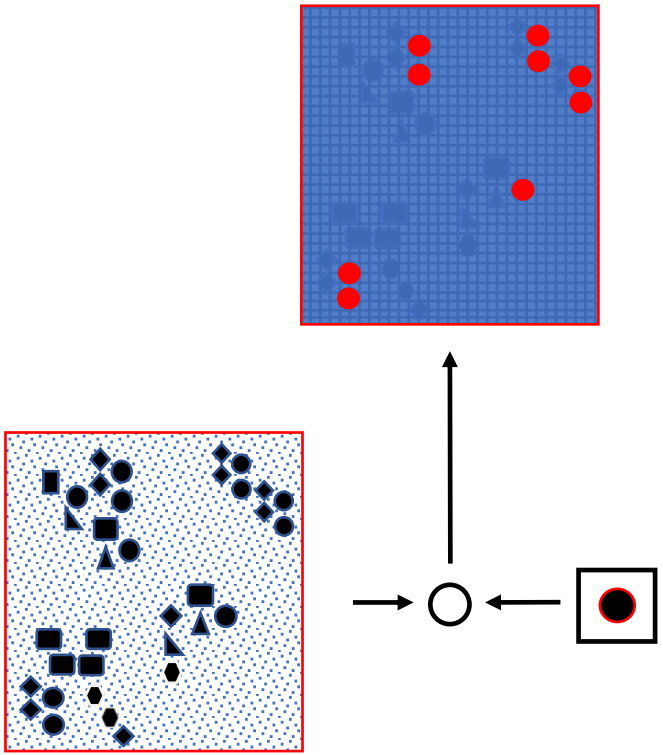
Convolutional Filter



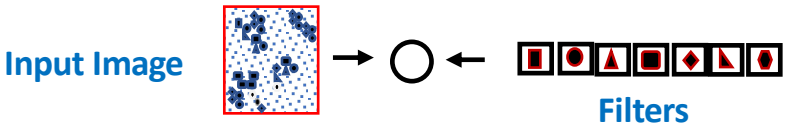
Convolutional Filter

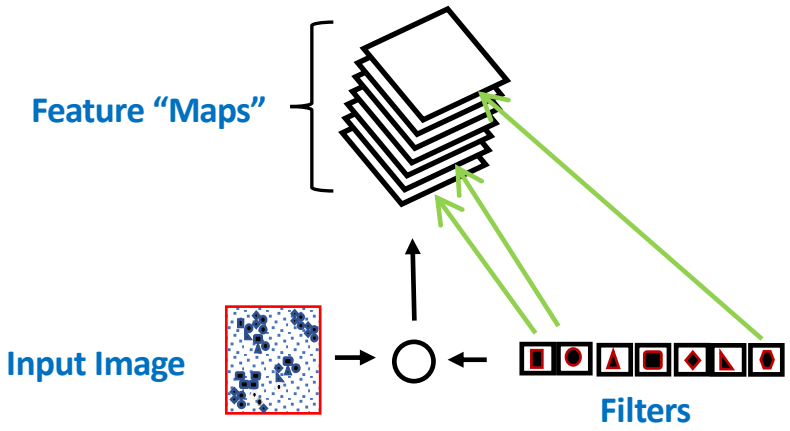


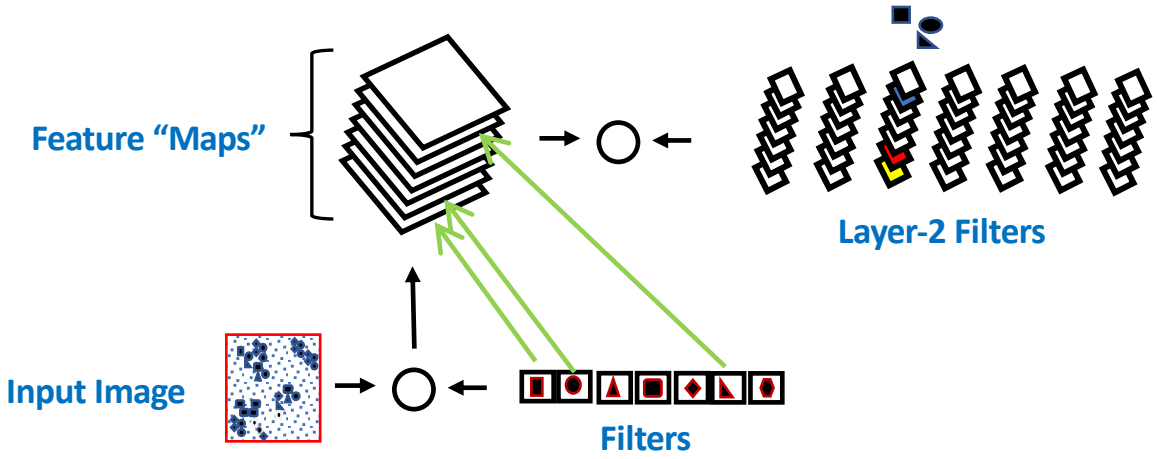
Convolutional Filter

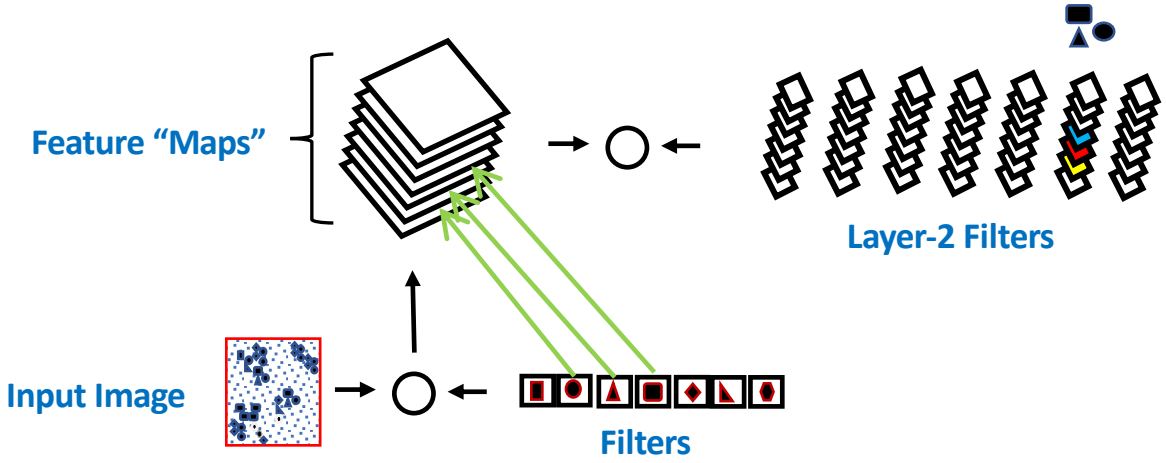


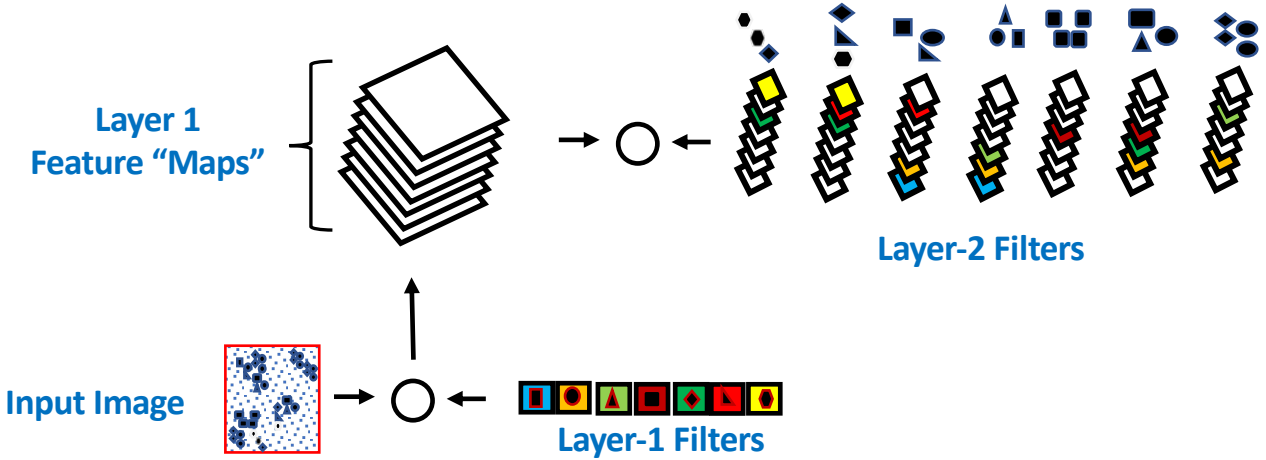
Multiple Filters, One for Each Building Block

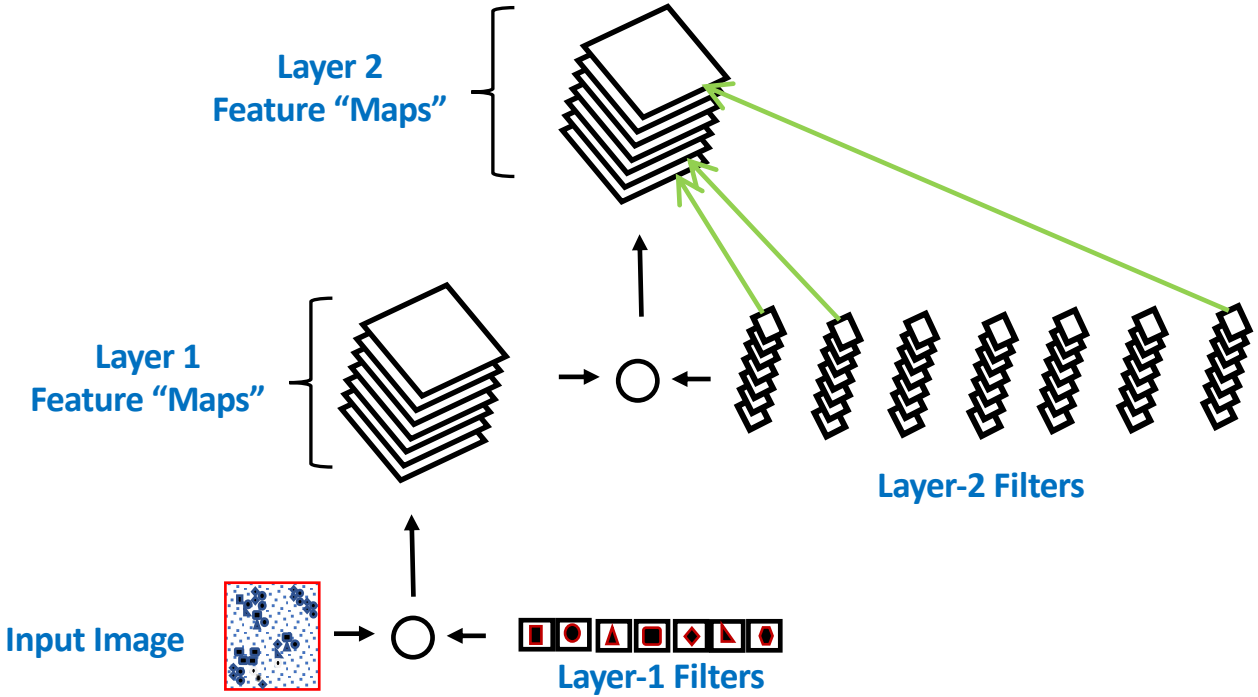


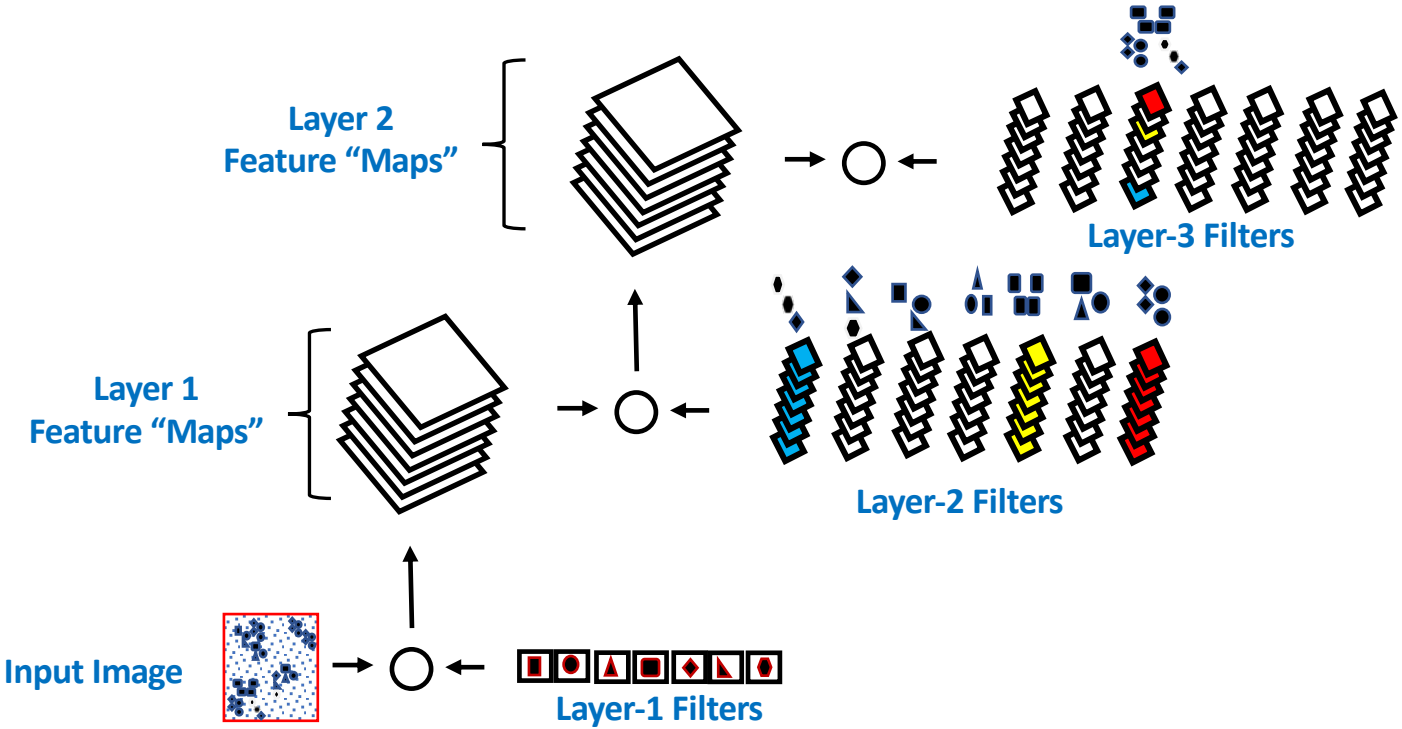

Feature “Maps”
Input Image
Filters

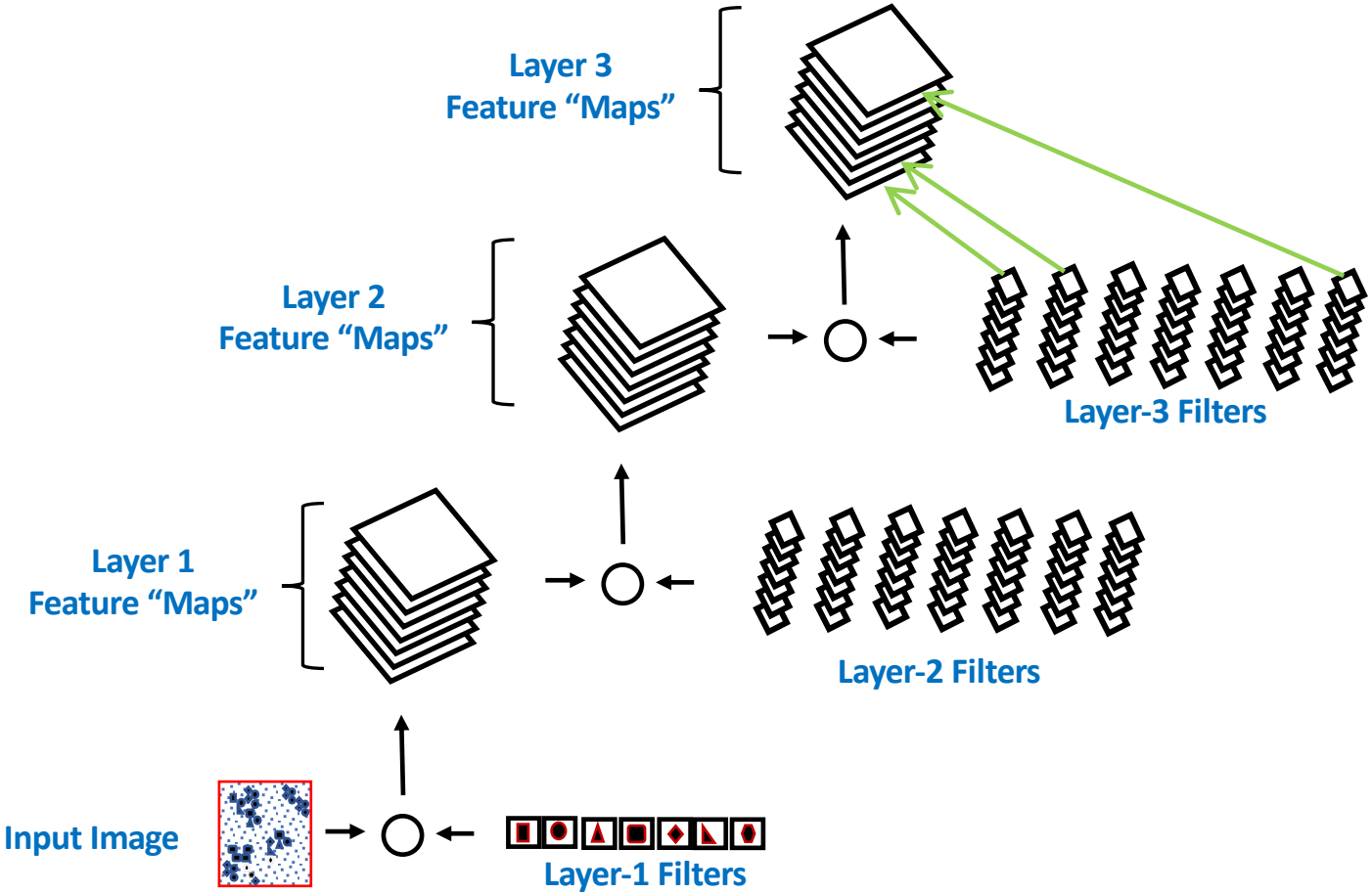

Feature “Maps”
Layer-2 Filters
Input Image
Filters


Feature “Maps”
Layer-2 Filters
Input Image
Filters

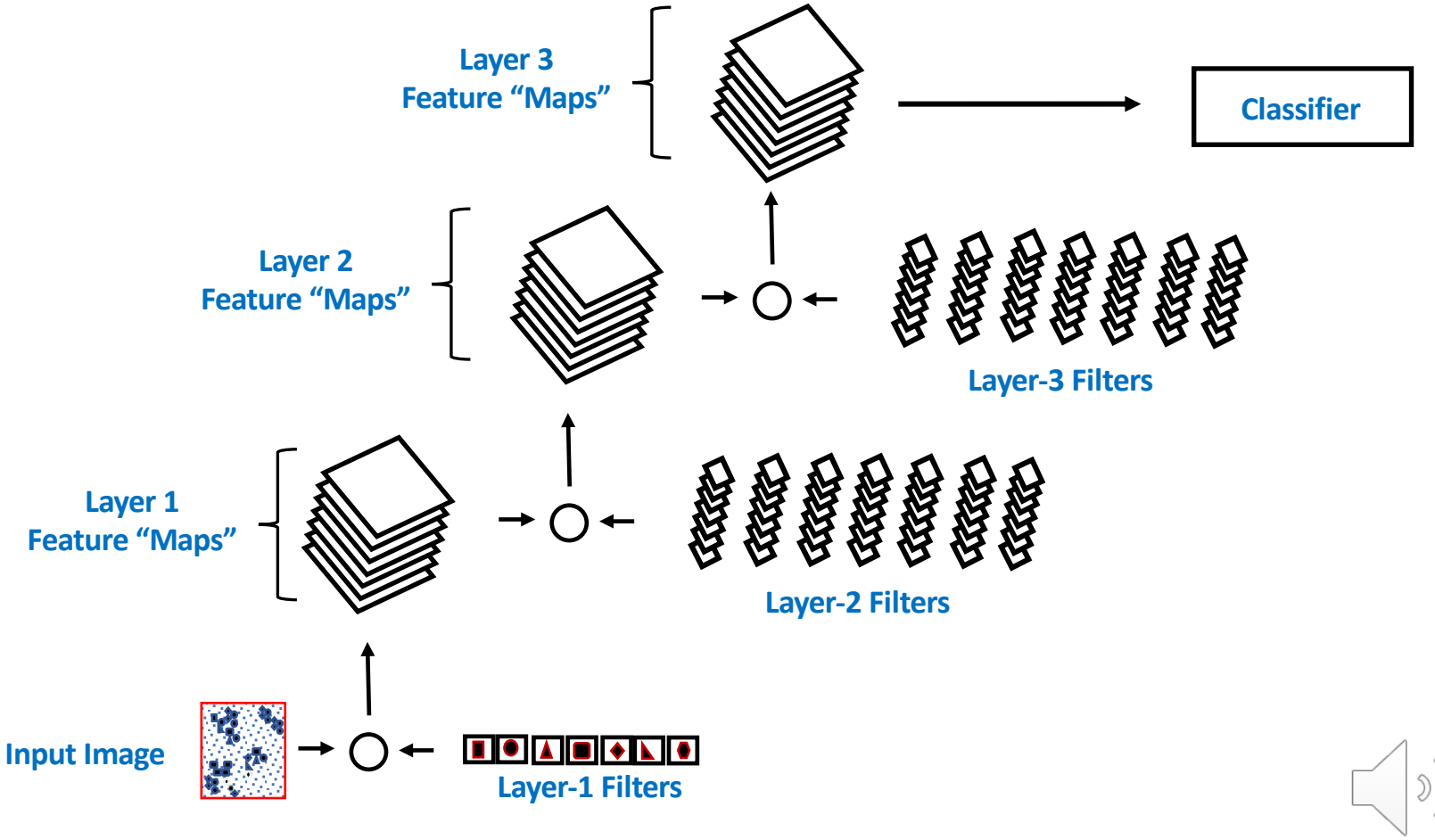

Layer 1
Feature "Maps"
Layer-2 Filters
Input Image
Layer-1 Filters


Layer 2
Feature “Maps”
Layer 1
Feature “Maps”
Layer-2 Filters
Input Image
Layer-1 Filters

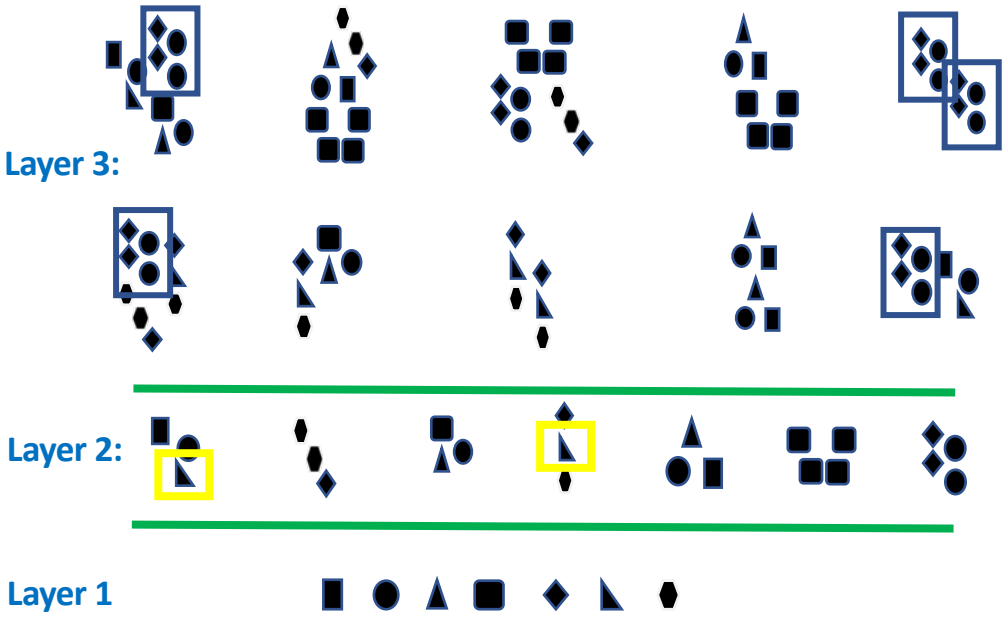

Layer 2
Feature "Maps"
Layer-3 Filters
Layer 1
Feature "Maps"
Layer-2 Filters
Input Image
Layer-1 Filters


Layer 3
Feature "Maps"
Layer 2
Feature "Maps"
Layer-3 Filters
Layer 1
Feature "Maps"
Layer-2 Filters
Input Image
Layer-1 Filters

Deep CNN Architecture



Advantage of Hierarchical Features?



- By learning and sharing statistical similarities within high-level motifs, we better leverage all training data
- If we do not use such a hierarchy, top-level motifs would be learned in isolation of each other

2D Spatial Convolution

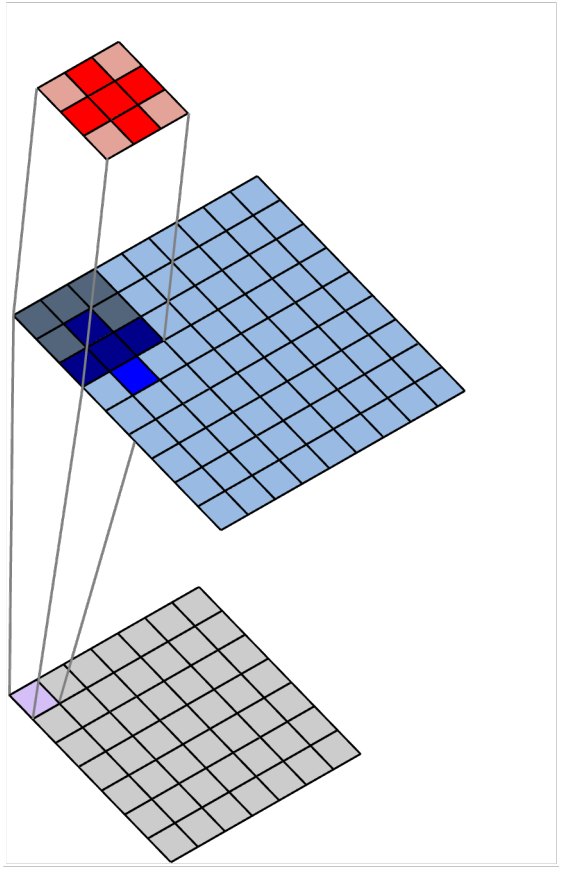
Filter

×

Image

Σ

Convolved Image (Feature Map)



2D Spatial Convolution

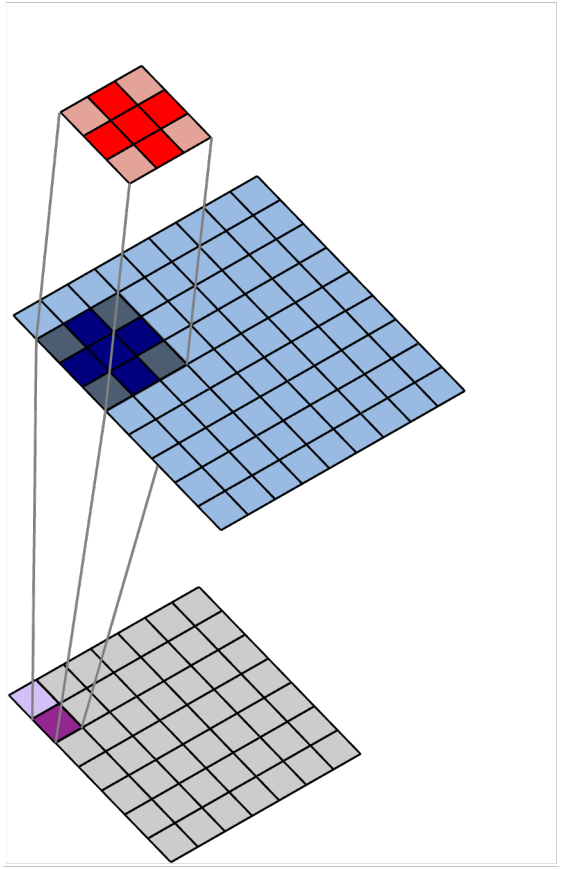
Filter

×

Image

Σ

Convolved Image (Feature Map)



2D Spatial Convolution

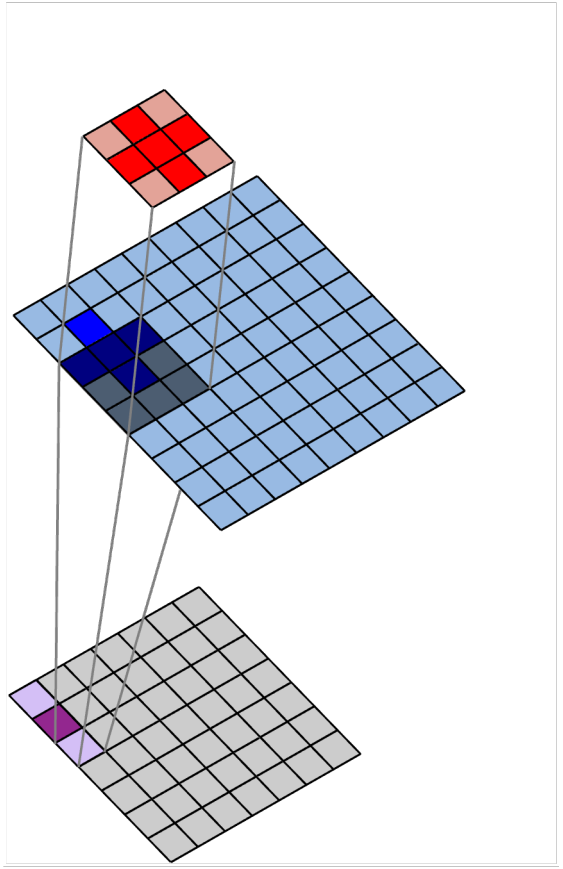
Filter

×

Image

Σ

Convolved Image (Feature Map)



2D Spatial Convolution

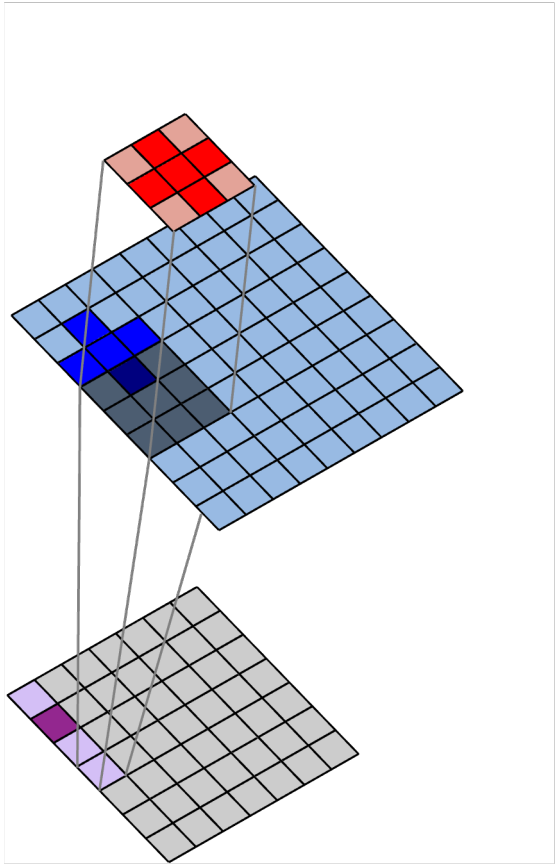
Filter

×

Image

Σ

Convolved Image (Feature Map)



2D Spatial Convolution

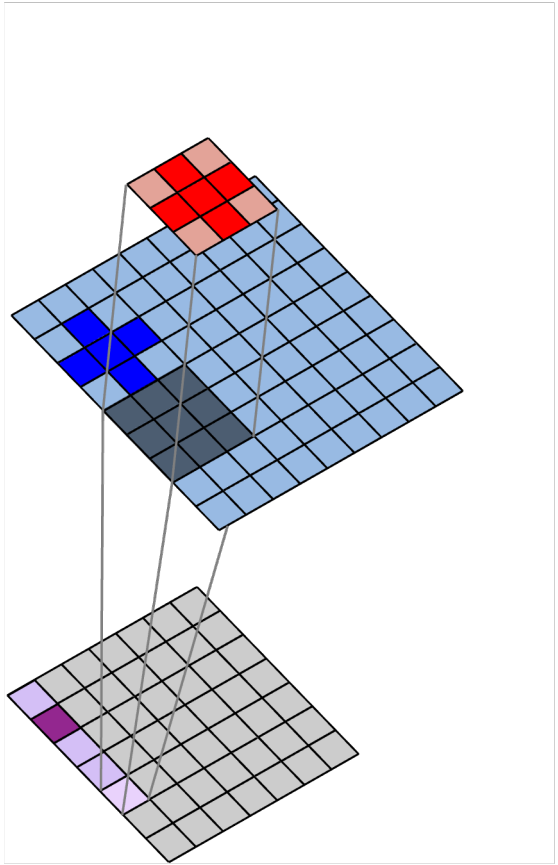
Filter

×

Image

Σ

Convolved Image (Feature Map)

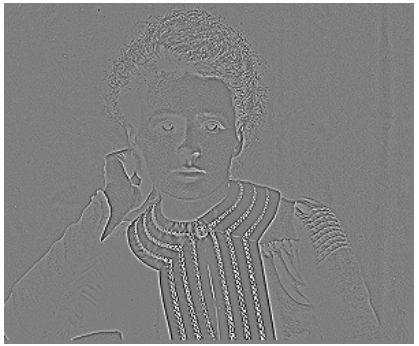


2D Spatial Convolution

Filter (W)

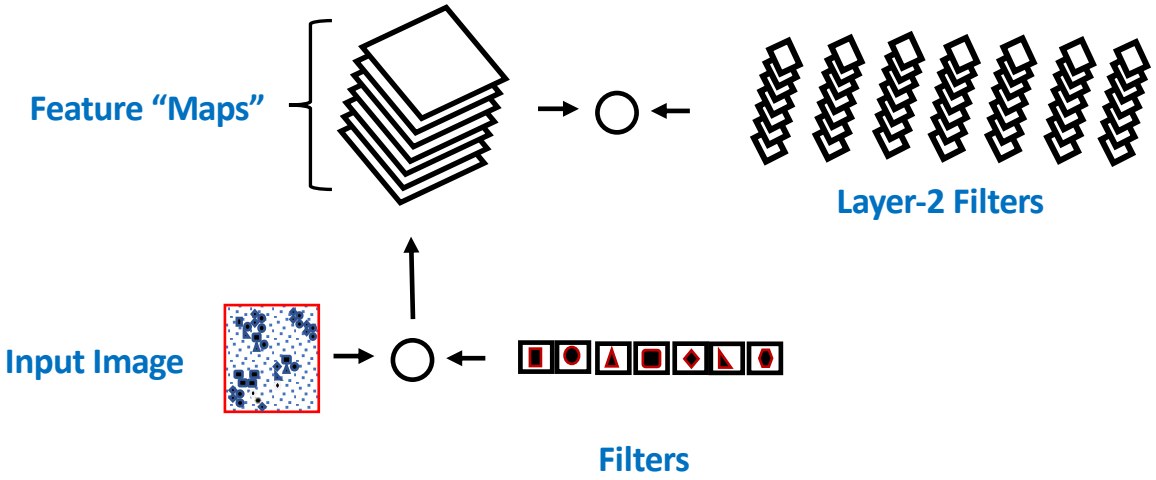
0	1	0
1	-4	1
0	1	0

Input (I)

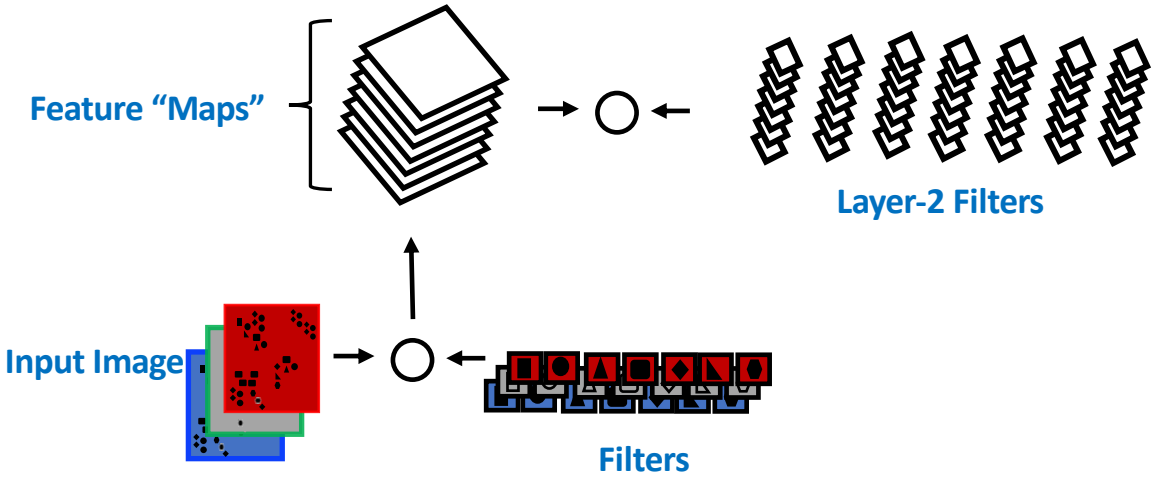


Filters Operate Over Input

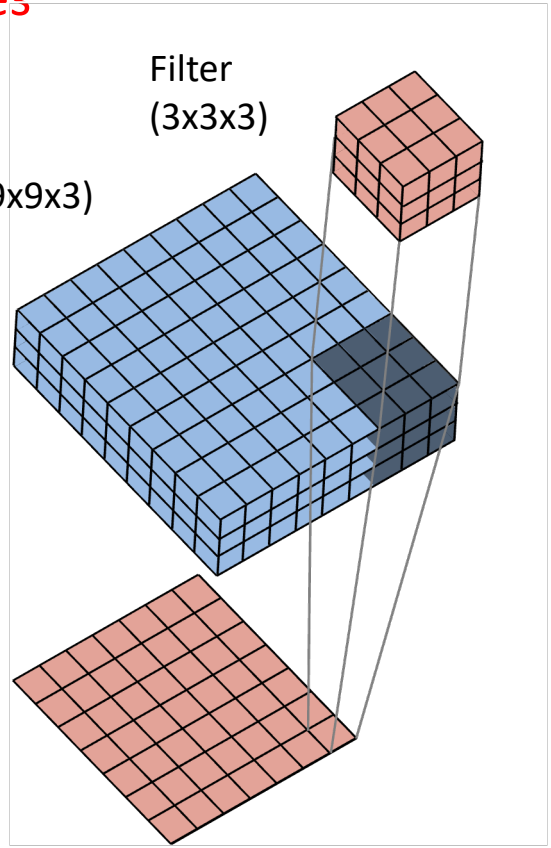
Volumes



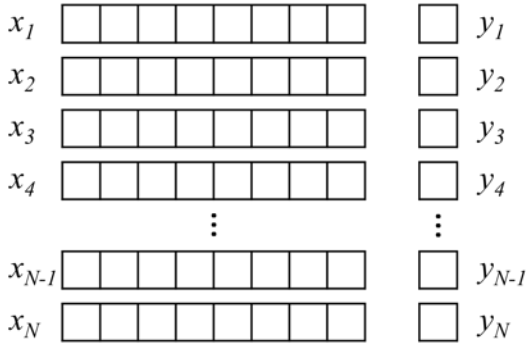
Filters Operate Over Input Volumes



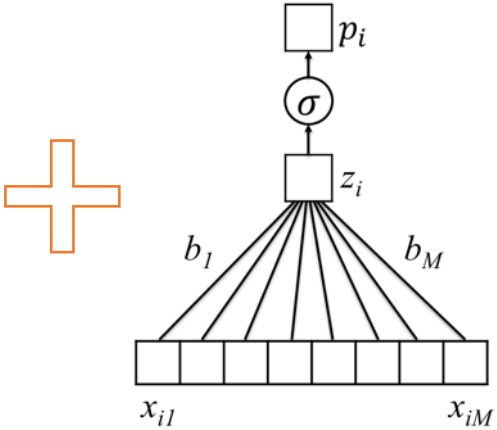
Filters Operate Over Input Volumes



Given Labeled Training Images, How do we Learn the Parameters of the CNN?

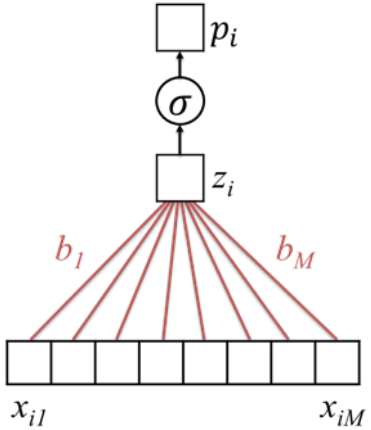


Training Set



$$p_i = \sigma(b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_M x_{iM})$$

Untrained Logistic Regression Model (or "Network")



$$b = (b_0, \ldots b_M)$$

Trained Model (with learned parameters)

Given Labeled Training Images, How do we Learn the Parameters of the CNN?

CNN Architecture

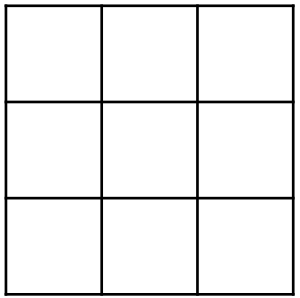
Trained CNN

Architecture (specified) vs Parameters (learned)

Architecture:

- Number of layers
- Layer types (e.g. convolutional, pooling, fully connected)
- Number of filters in each layer
- Shape and size of filters

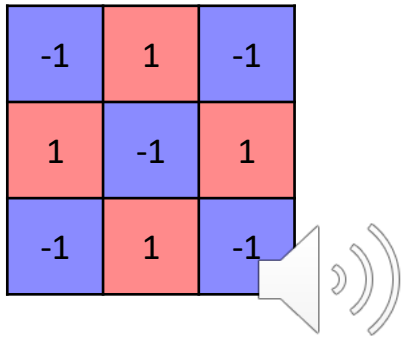
Use 3x3 filters
In layer 1



Parameters:

- Individual Elements of each filter
- Parameters of other layers

Learn values of
Each layer 1 filter



Summary

- Convolutional neural networks learn to recognize **high-level structure** in images by building **hierarchical representations of features**
- Features are extracted via spatial convolutions with **filters**
- Filters are learned by minimizing a loss function just like in the other models we've discussed
- Convolutional neural networks have shown capabilities beyond human performance for image analysis