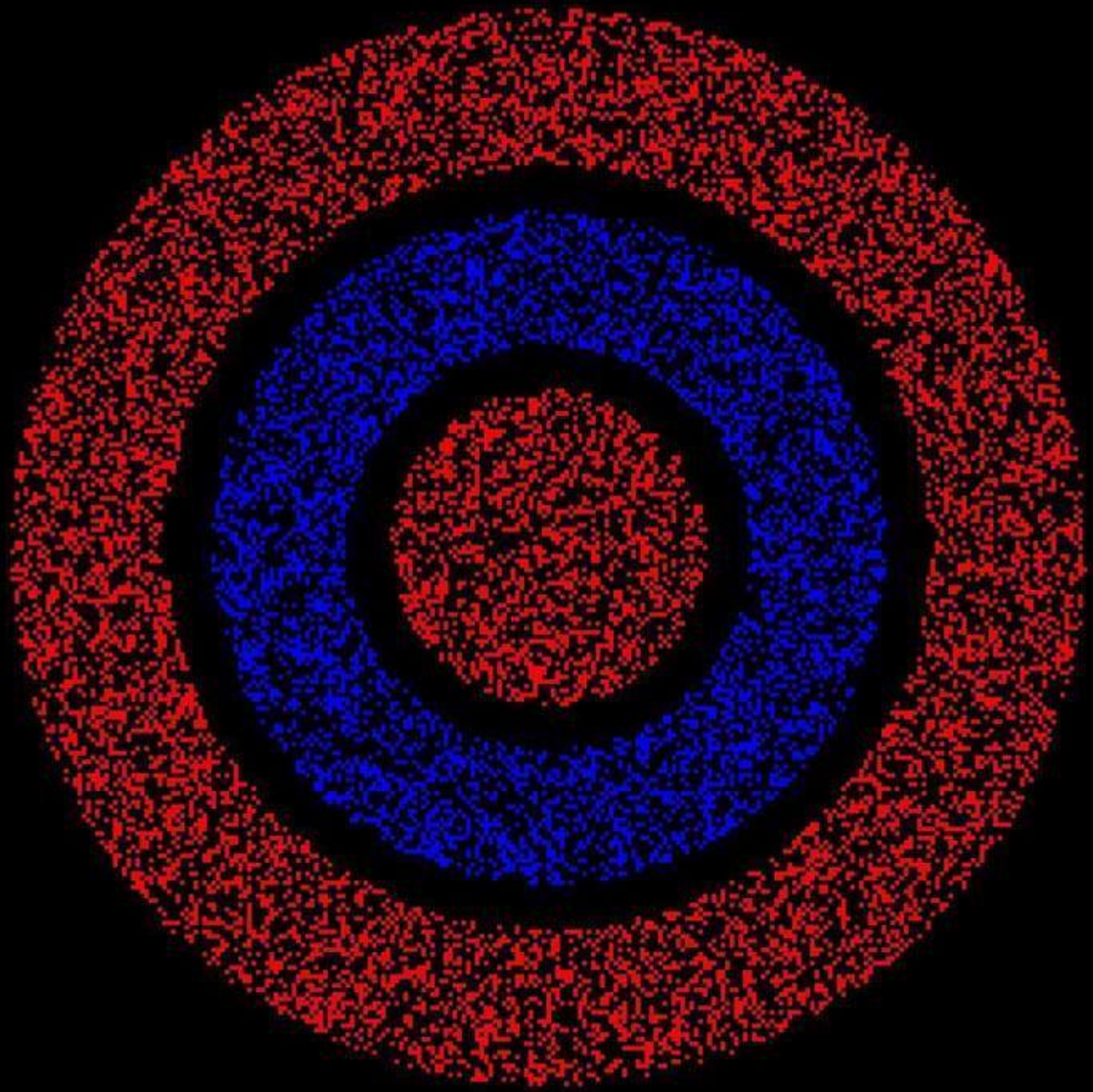
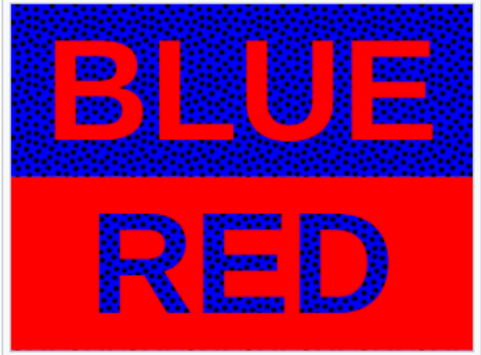


Chromostereopsis

From Wikipedia, the free encyclopedia

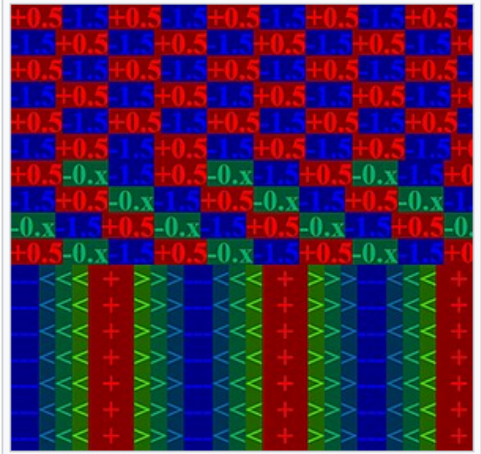
Chromostereopsis is a visual illusion whereby the impression of depth is conveyed in two-dimensional color images, usually of red–blue or red–green colors, but can also be perceived with red–grey or blue–grey images.[1][2] Such illusions have been reported for over a century and have generally been attributed to some form of chromatic aberration.[3][4][5][6][7]



Blue–red contrast demonstrating depth perception effects

Chromatic aberration results from the differential refraction of light depending on its wavelength, causing some light rays to converge before others in the eye (longitudinal chromatic aberration or LCA) and/or to be located on non-corresponding locations of the two eyes during binocular viewing (transverse chromatic aberration or TCA).

Chromostereopsis is usually observed using a target with red and blue bars and an achromatic background. Positive chromostereopsis is exhibited when the red bars are perceived in front of the blue and negative chromostereopsis is exhibited when the red bars are perceived behind the blue.[8] Several models have been proposed to explain this effect which is often attributed to longitudinal and/or transverse chromatic aberrations.[6] However, some work attributes most of the stereoptic effect to transverse chromatic aberrations in combination with cortical factors.[1][5][7]

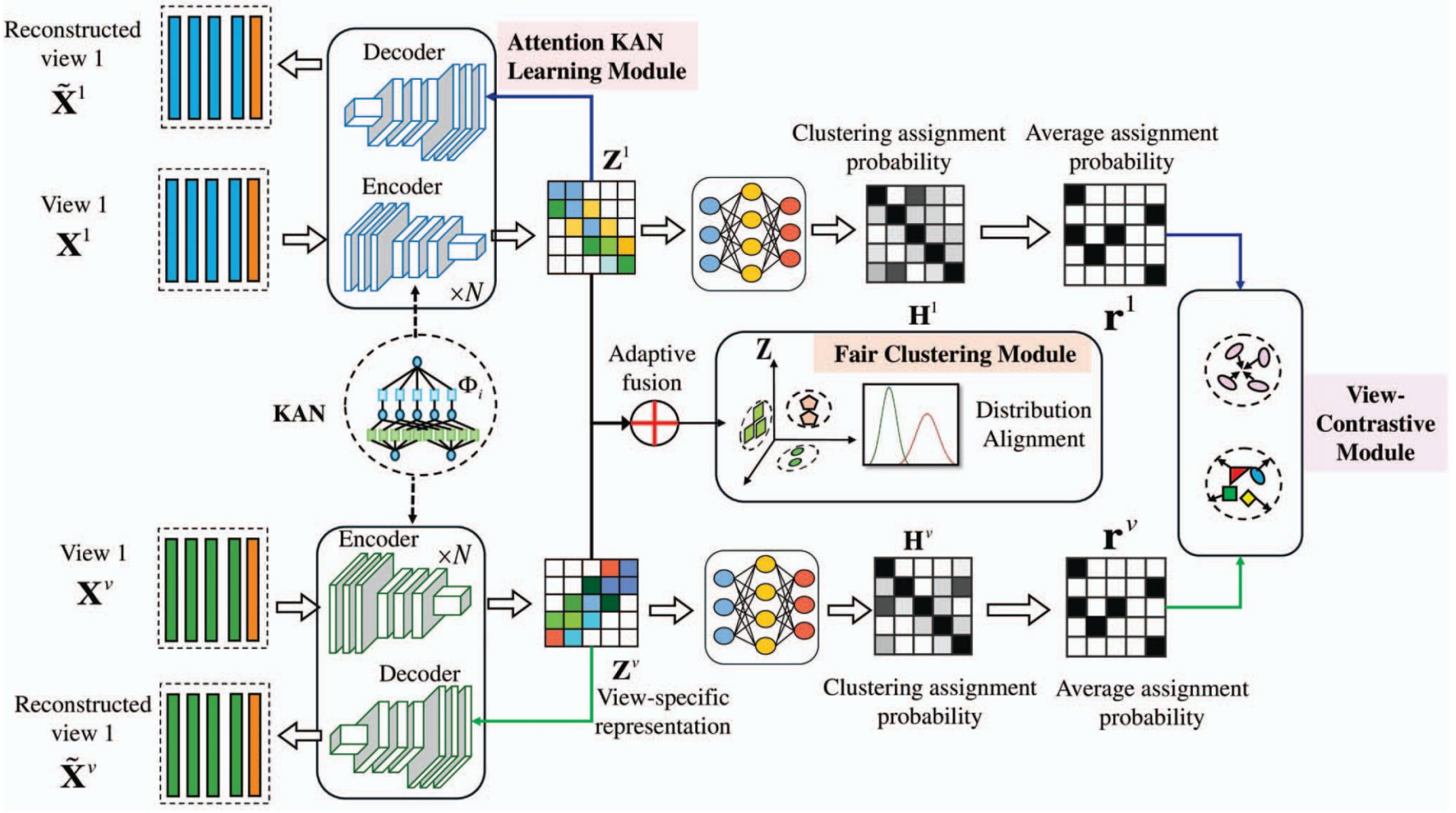


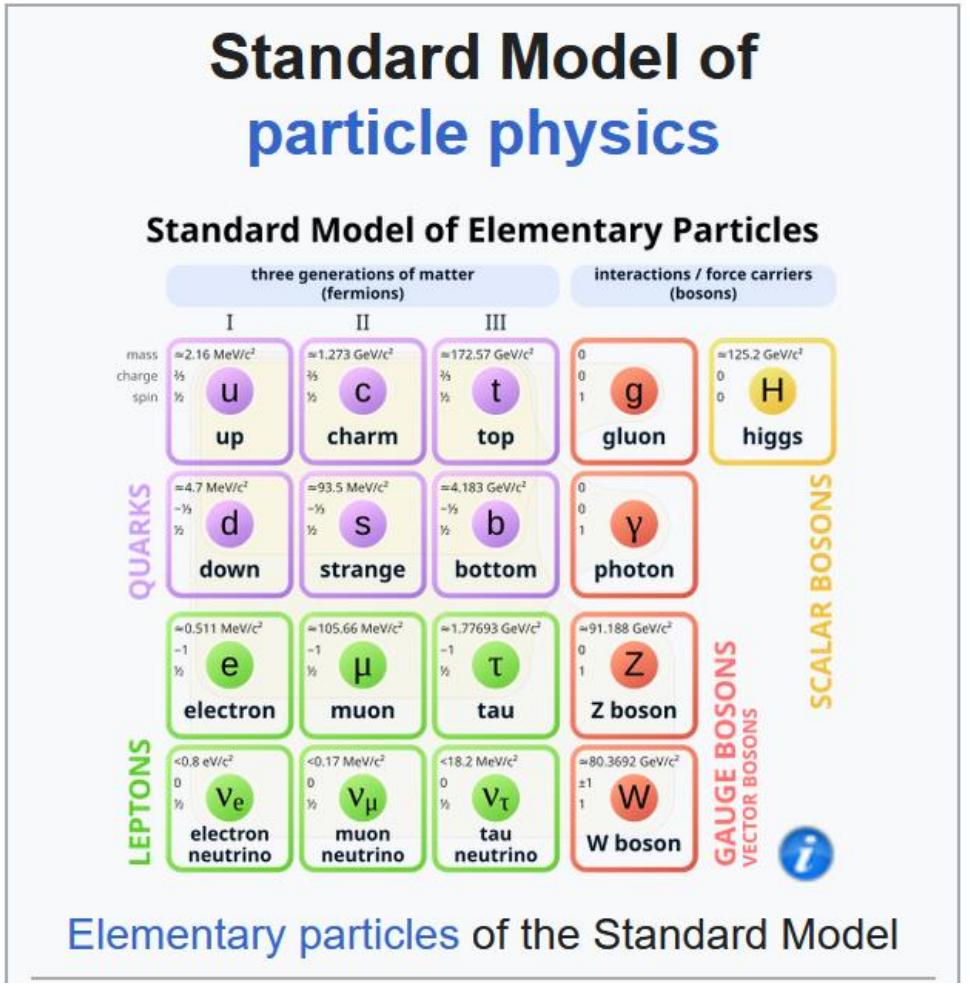
3 Layers of depths "Rivers, Valleys & Mountains"

It has been proposed that chromostereopsis could have evolutionary implications in the development of eyespots in certain butterfly species.

The perceived differences in color's optical power span about 2 diopter (Blue: −1.5, Red

What's going on inside deep networks?





- "Young man, if I could remember the names of these particles, I would have been a botanist.“ Fermi
- "Had I foreseen that, I would have gone into botany.“ Pauli

Object Detectors Emerge in Deep Scene CNNs

Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba. 2014

CNN for Object Recognition

Large-scale image classification result on ImageNet

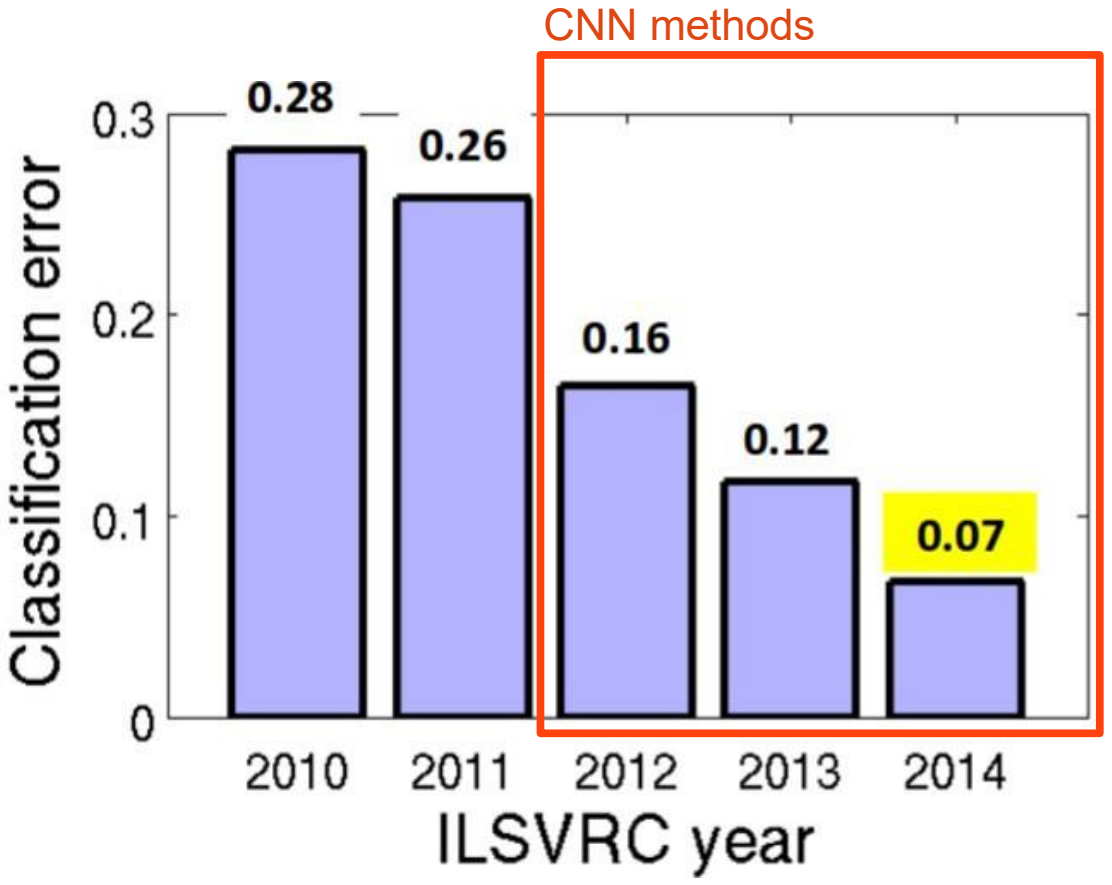
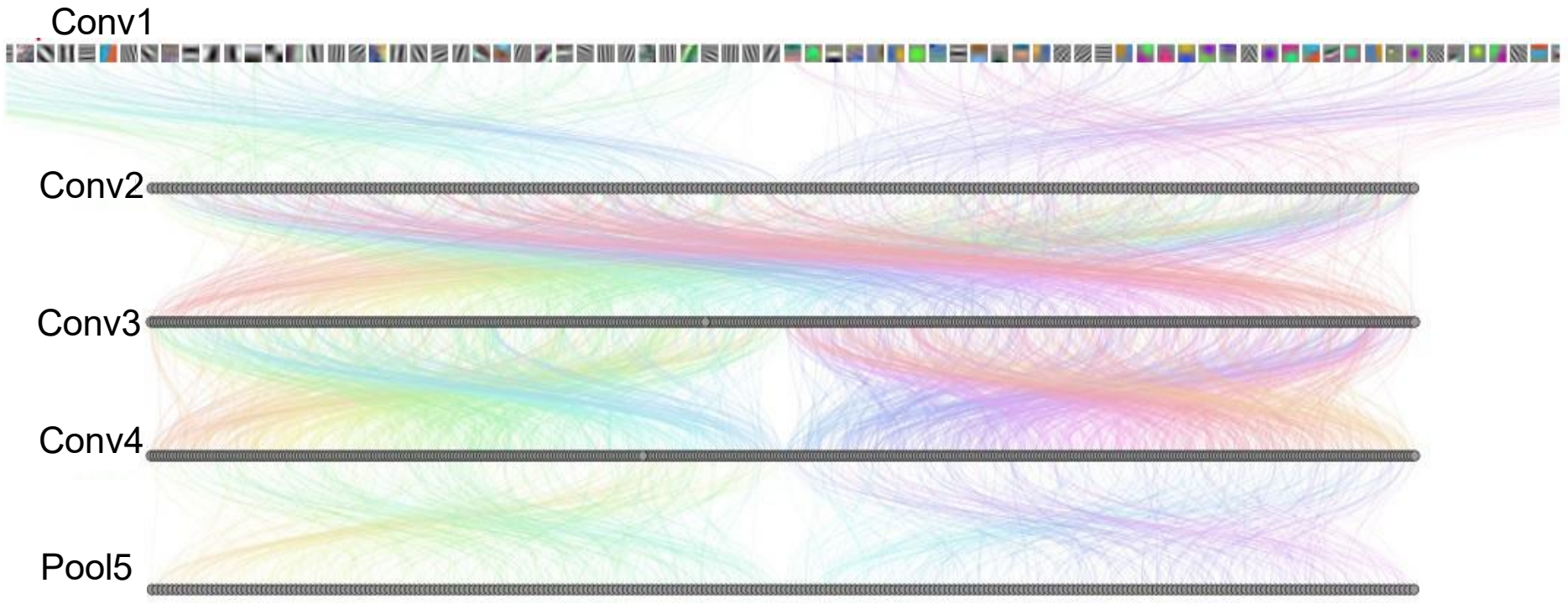


Figure from Olga Russakovsky ECCV'14 workshop

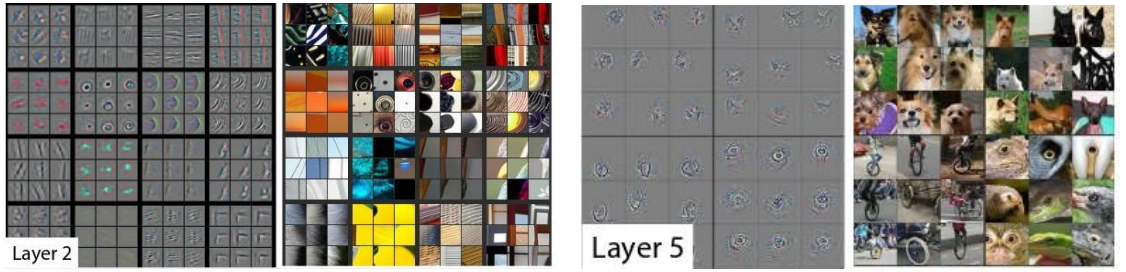
How Objects are Represented in CNN?



DrawCNN: visualizing the units' connections

How Objects are Represented in CNN?

Deconvolution

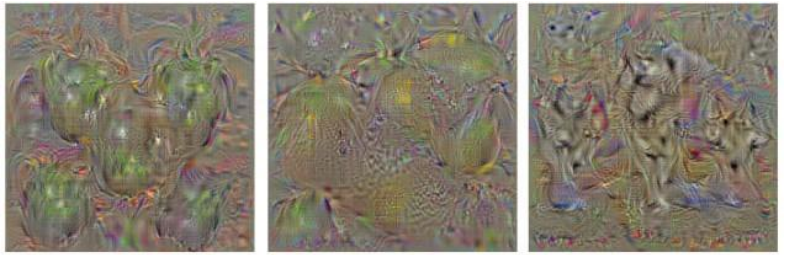


Zeiler, M. et al. Visualizing and Understanding Convolutional Networks,ECCV 2014.

Strong activation image

Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accu-rate object detection and semantic segmentation. CVPR 2014

Back-propagation



Simonyan, K. et al. Deep inside convolutional networks: Visualising image classification models and saliency maps. ICLR workshop, 2014

Another CNN interpretation method: Simplifying Scenes While Maintaining Classifier Decision

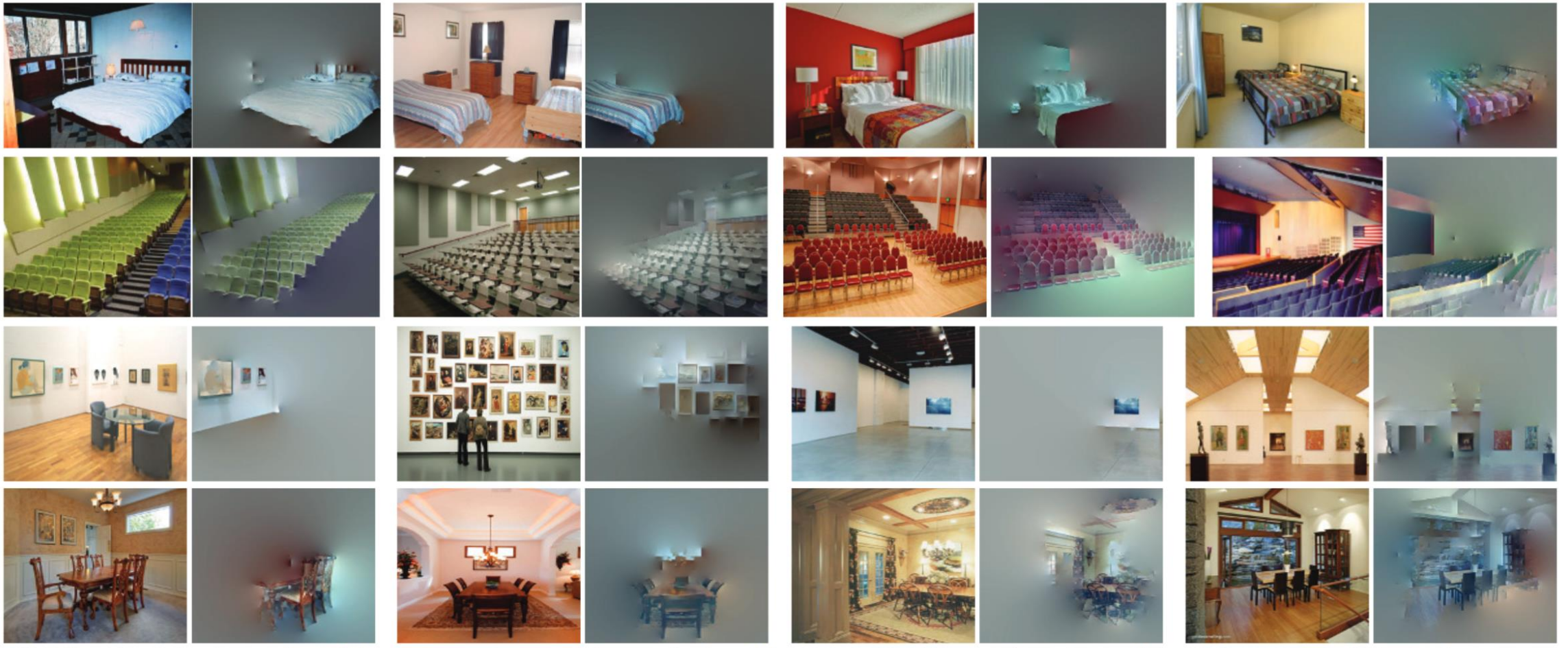


Figure 2: Each pair of images shows the original image (left) and a simplified image (right) that gets classified by the Places-CNN as the same scene category as the original image. From top to bottom, the four rows show different scene categories: bedroom, auditorium, art gallery, and dining room.

Scene Recognition

Given an image, predict which place we are in.


Bedroom


Harbor

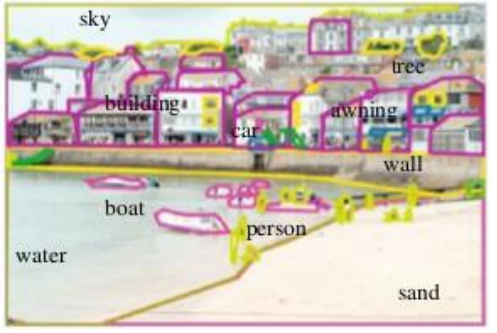
Learning to Recognize Scenes

bedroom

mountain

Possible internal representations:

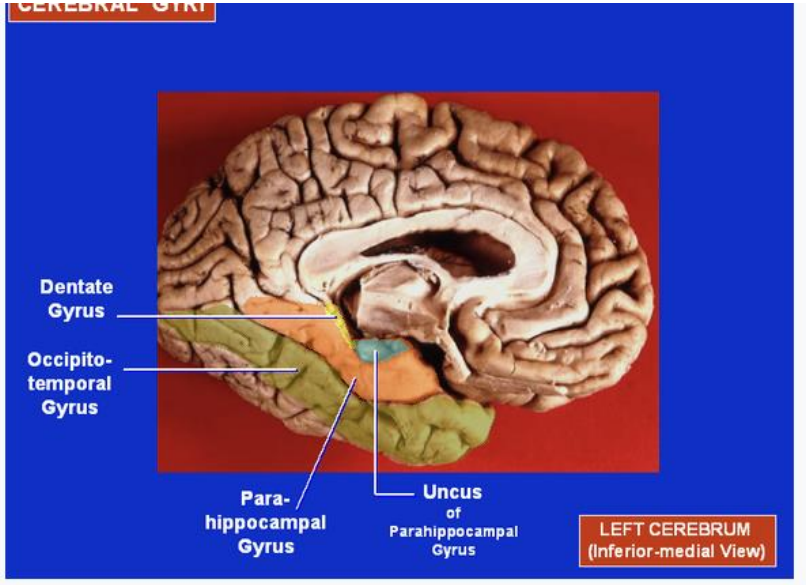
- Objects (scene parts?)
- Scene attributes
- Object parts
- Textures



Scene recognition [edit]

The **parahippocampal place area** (**PPA**) is a sub-region of the parahippocampal cortex that lies medially in the inferior temporo-occipital cortex. PPA plays an important role in the encoding and recognition of environmental scenes (rather than faces). fMRI studies indicate that this region of the brain becomes highly active when human subjects view topographical scene stimuli such as images of landscapes, cityscapes, or rooms (i.e. images of "places"). Furthermore, according to work by Pierre Mégevand et al. in 2014, stimulation of the region via intracranial electrodes yields intense topographical visual hallucinations of places and situations.[4] The region was first described by Russell Epstein and Nancy Kanwisher in 1998 at MIT,[5] see also other similar reports by Geoffrey Aguirre[6][7] and Alumit Ishai.[8]

Damage to the PPA (for example, due to stroke) often leads to a syndrome in which patients cannot visually recognize scenes even though they can recognize the individual objects in the scenes (such as people, furniture, etc.). The PPA is often considered the complement of the fusiform face area (FFA), a nearby cortical region that responds strongly whenever faces are viewed, and that is believed to be important for face recognition.



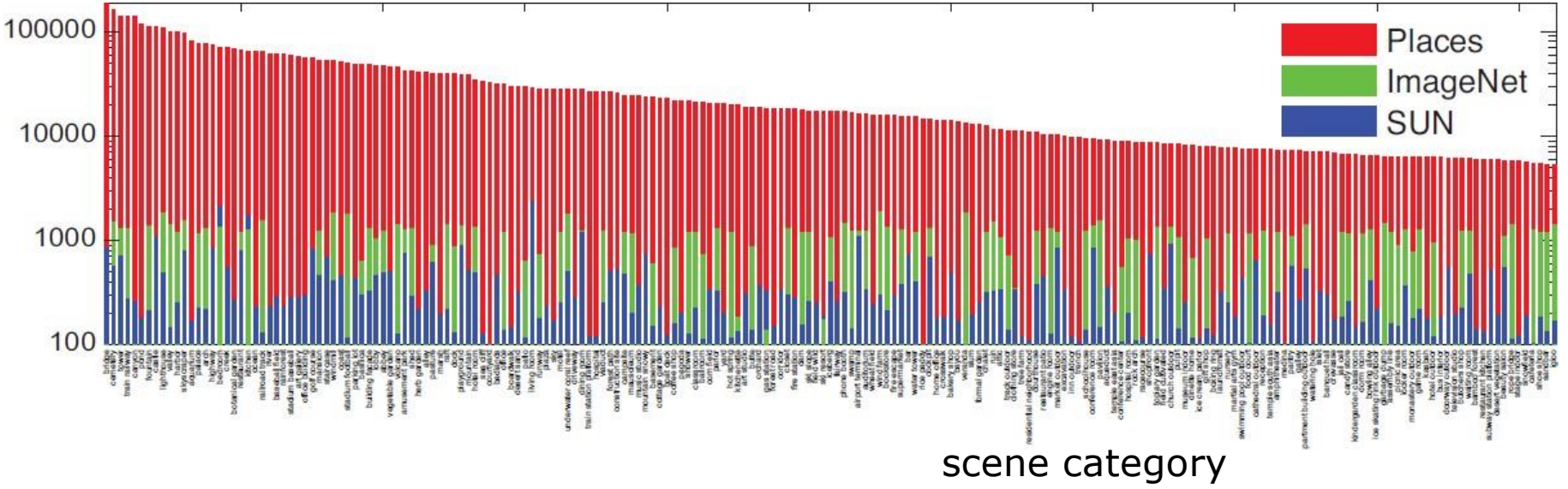
Medial view of left cerebral hemisphere. Parahippocampal gyrus shown in orange.

Details	
Identifiers	
Latin	*gyrus parahippocampalis*
MeSH	D020534
NeuroNames	164
NeuroLex ID	birnlex_807
TA98	A14.1.09.234
TA2	5515
FMA	61918

Anatomical terminology

CNN for Scene Recognition

Places Database: 7 million images from 400 scene categories



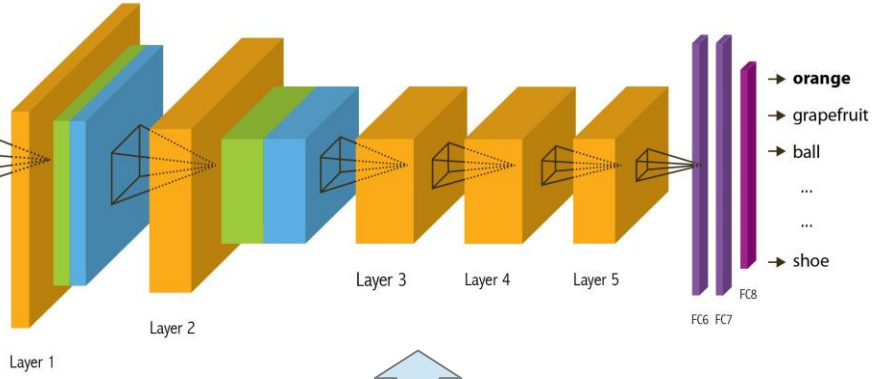
Places-CNN: AlexNet CNN on 2.5 million images from 205 scene categories.

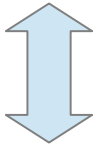
	Places 205	SUN 205
Places-CNN	**50.0%**	**66.2%**
ImageNet CNN feature+SVM	40.8%	49.6%

Zhou, et al. NIPS, 2014.

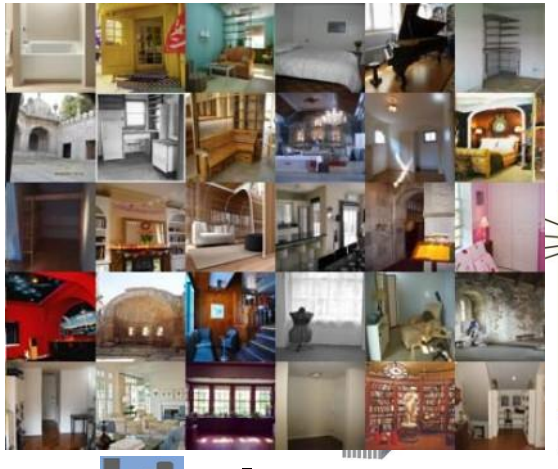
ImageNet CNN and Places CNN

ImageNet CNN for Object Classification

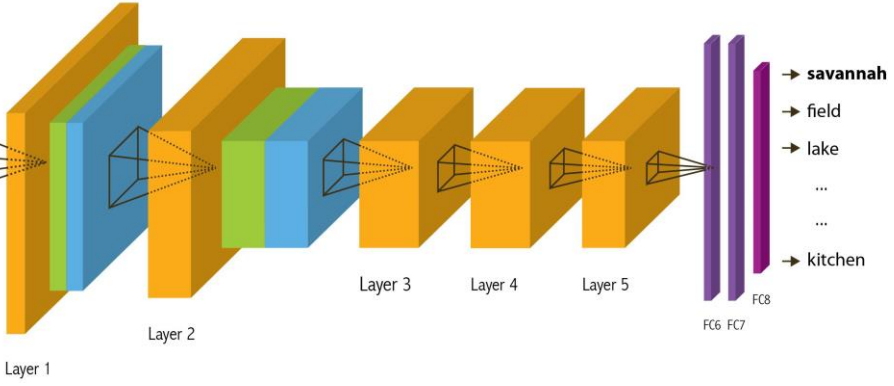




Same architecture: AlexNet

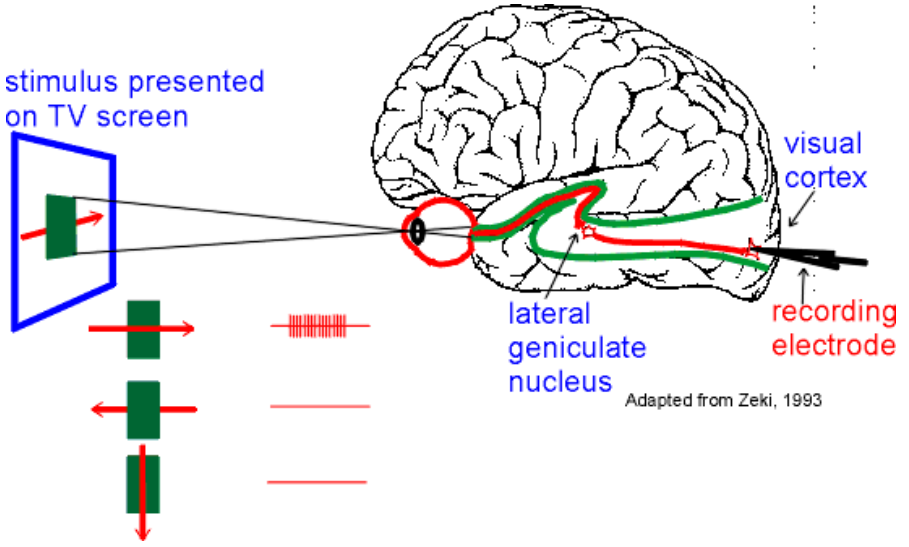


Places CNN for Scene Classification

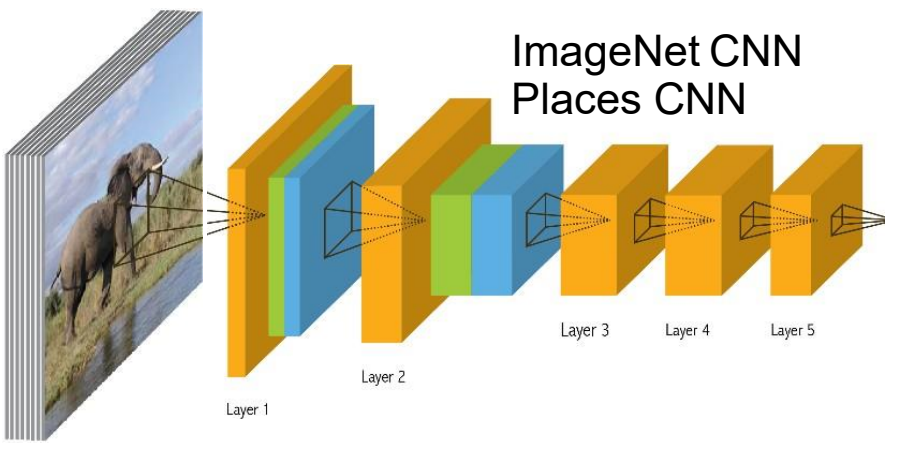


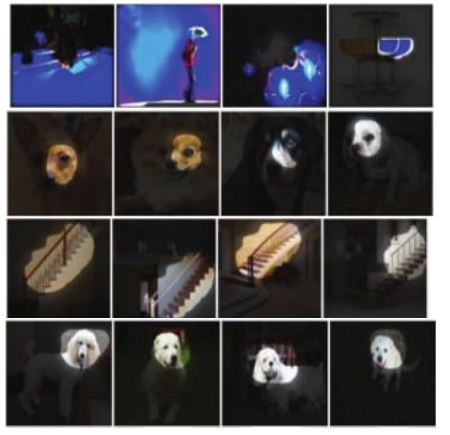
Data-Driven Approach to Study CNN

Neuroscientists study brain

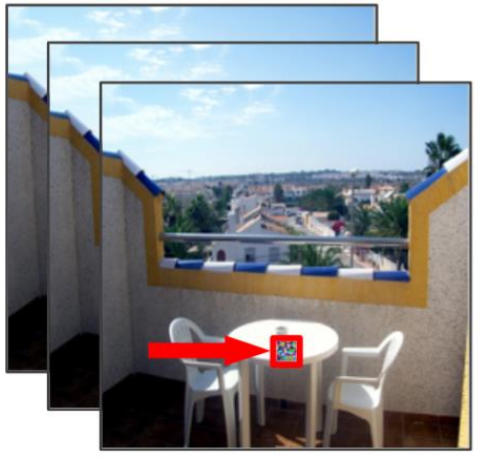


200,000 image stimuli of objects and scene categories (ImageNet TestSet+SUN database)



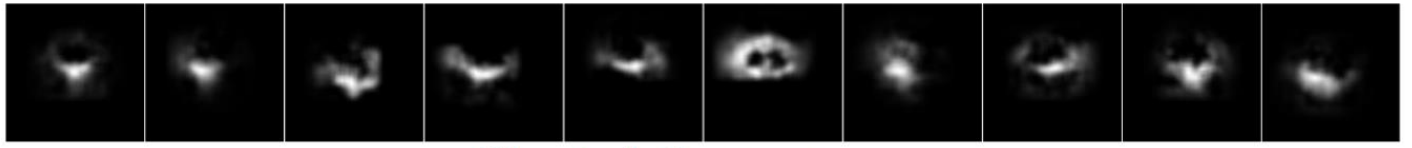


Estimating the Receptive Fields

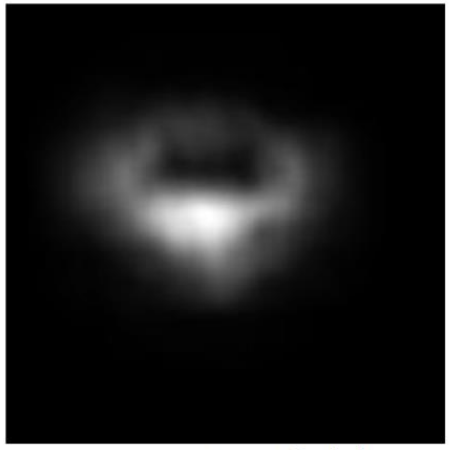


sliding-window stimuli

discrepancy maps for top 10 images



calibrated discrepancy maps



receptive field

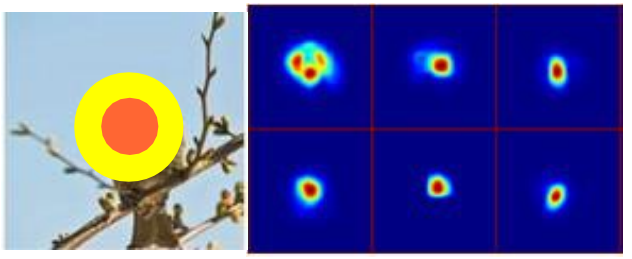
Estimating the Receptive Fields

Estimated receptive fields

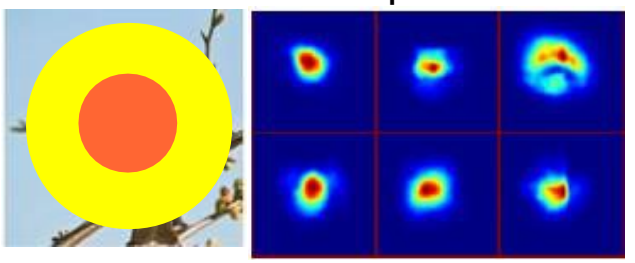
Actual size of RF is much smaller than the theoretic size

pool1

conv3



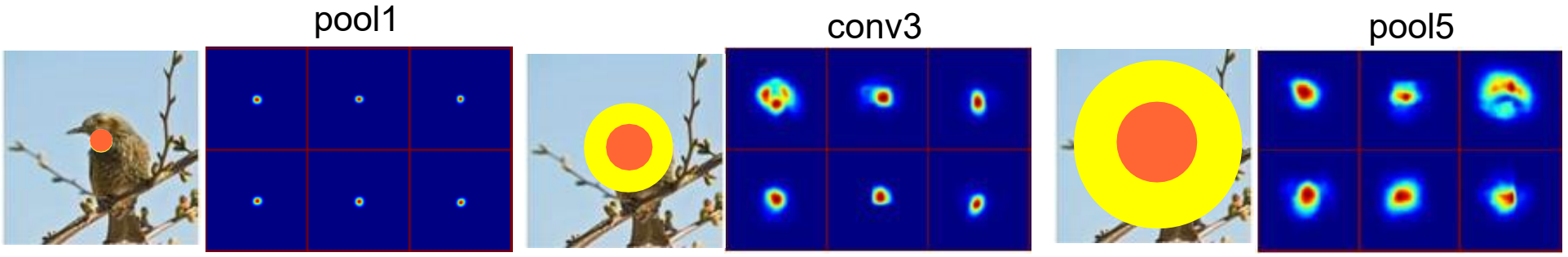
pool5



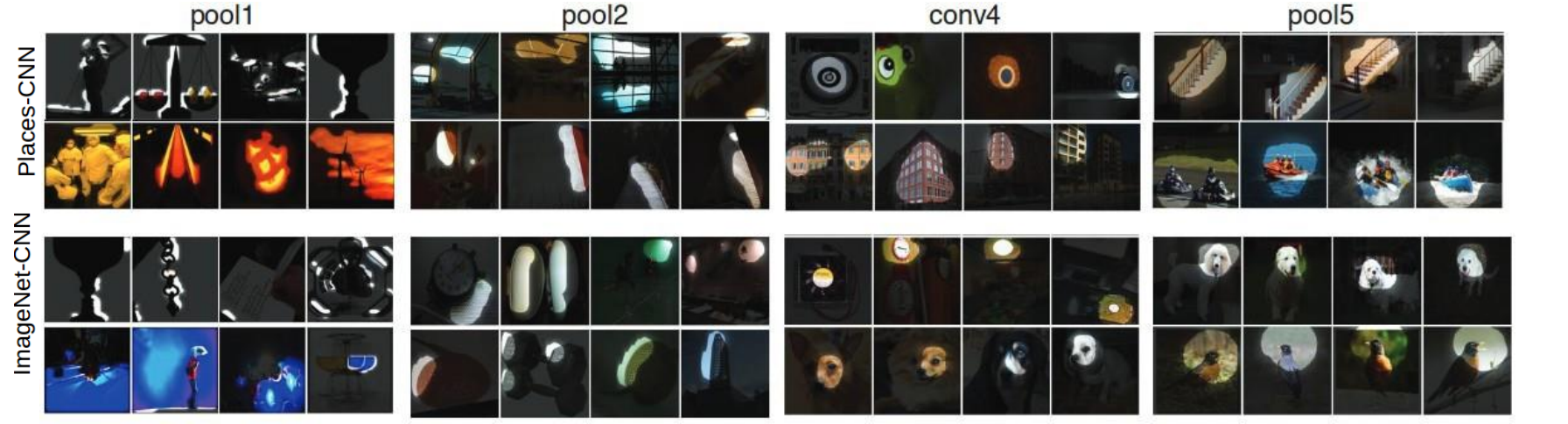
Estimating the Receptive Fields

Estimated receptive fields

Actual size of RF is much smaller than the theoretic size

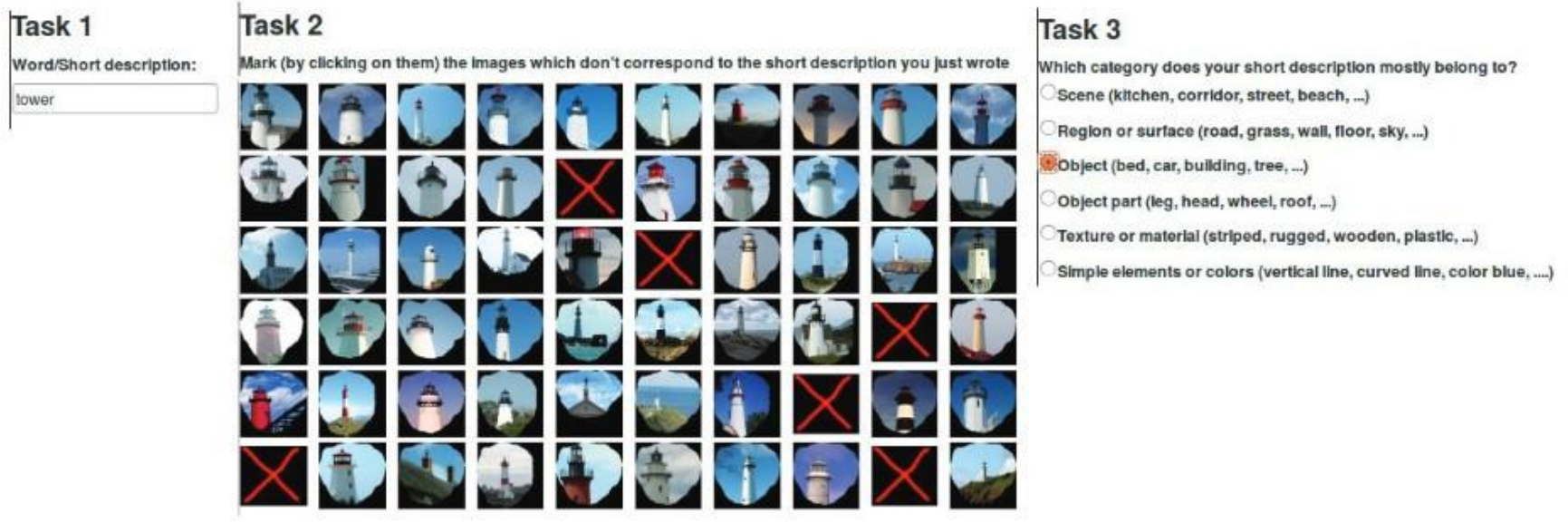


Segmentation using the RF of Units



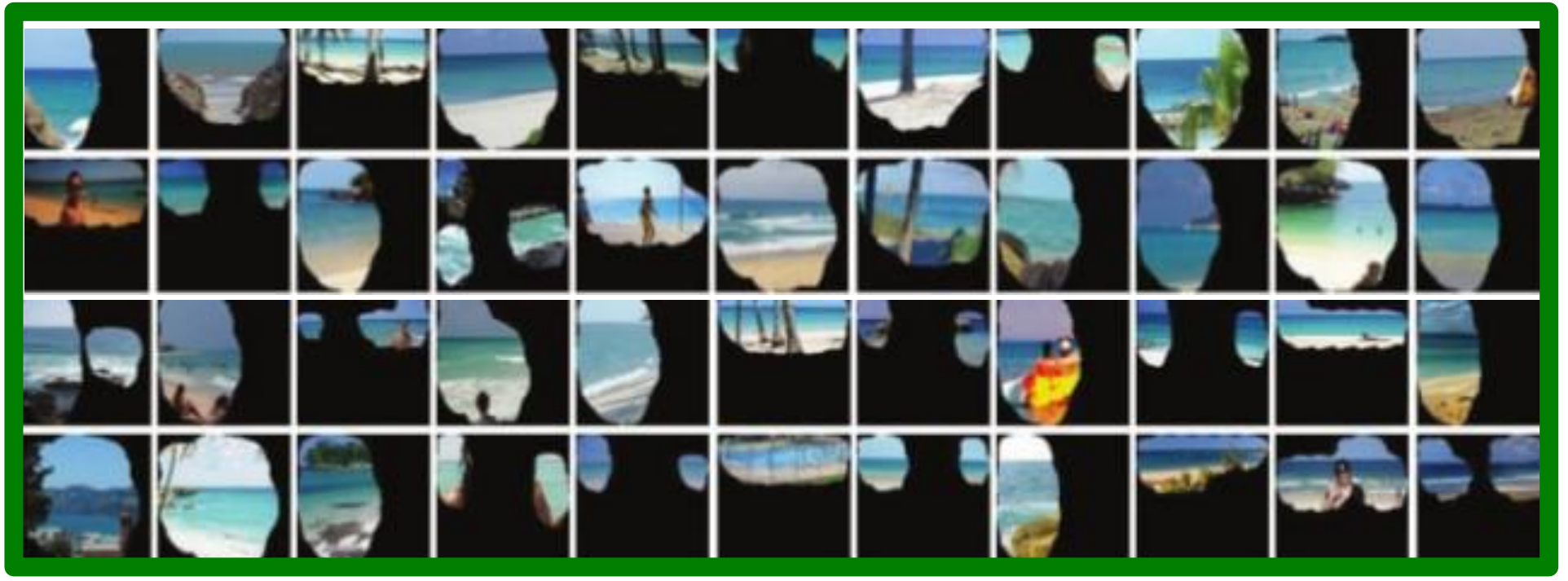
Annotating the Semantics of Units

Top ranked segmented images are cropped and sent to Amazon Turk for annotation.



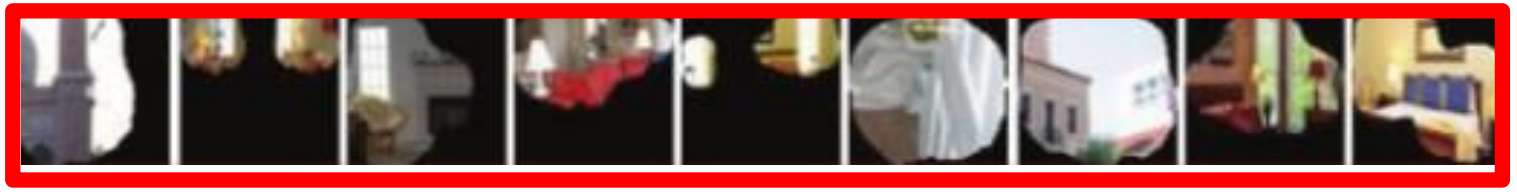
Annotating the Semantics of Units

Pool5, unit 76; Label: ocean; Type: scene; Precision: 93%



Annotating the Semantics of Units

Pool5, unit 13; Label: Lamps; Type: object; Precision: 84%

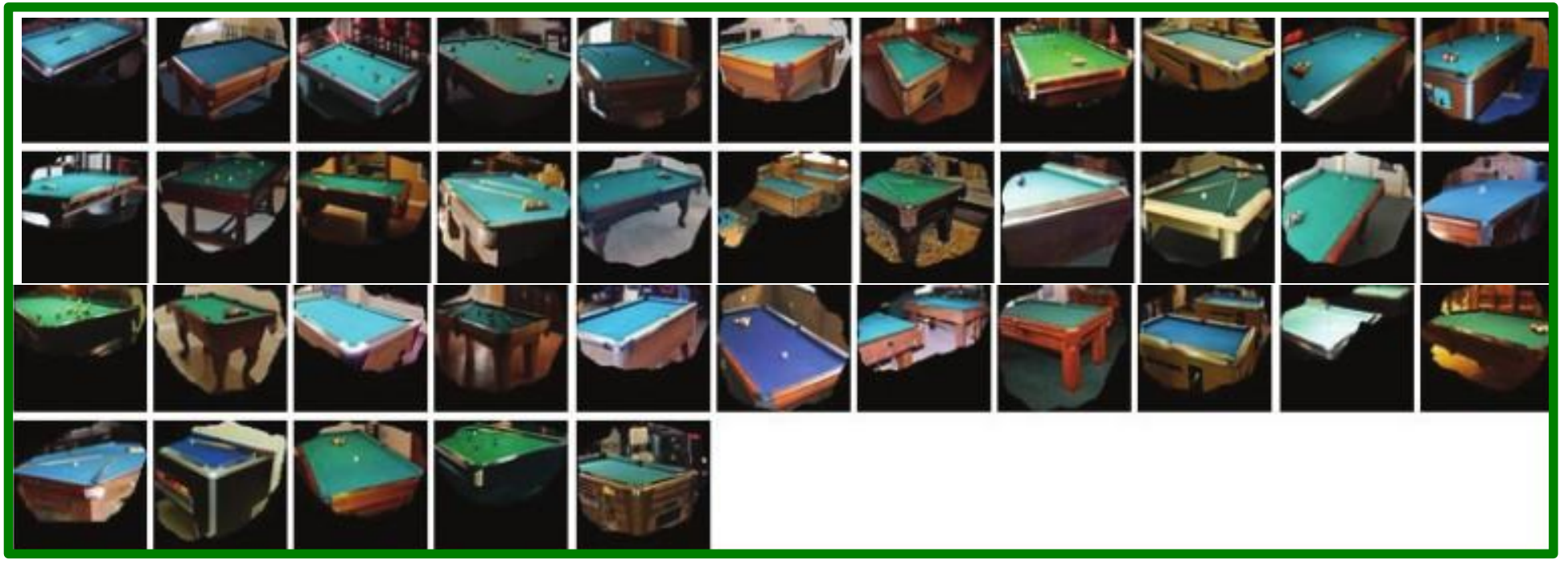


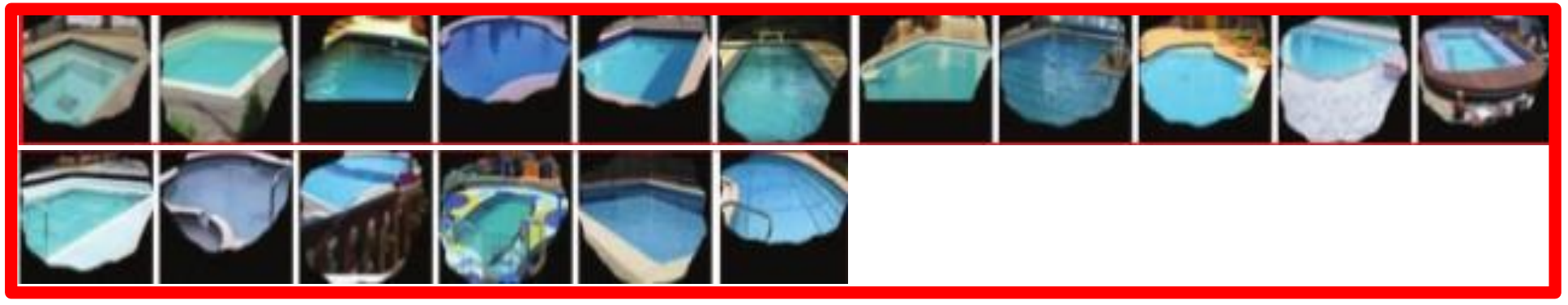
Annotating the Semantics of Units

Pool5, unit 77; Label:legs; Type: object part; Precision: 96%

Annotating the Semantics of Units

Pool5, unit 112; Label: pool table; Type: object; Precision: 70%





Annotating the Semantics of Units

Pool5, unit 22; Label: dinner table; Type: scene; Precision: 60%

Wedding-cake style

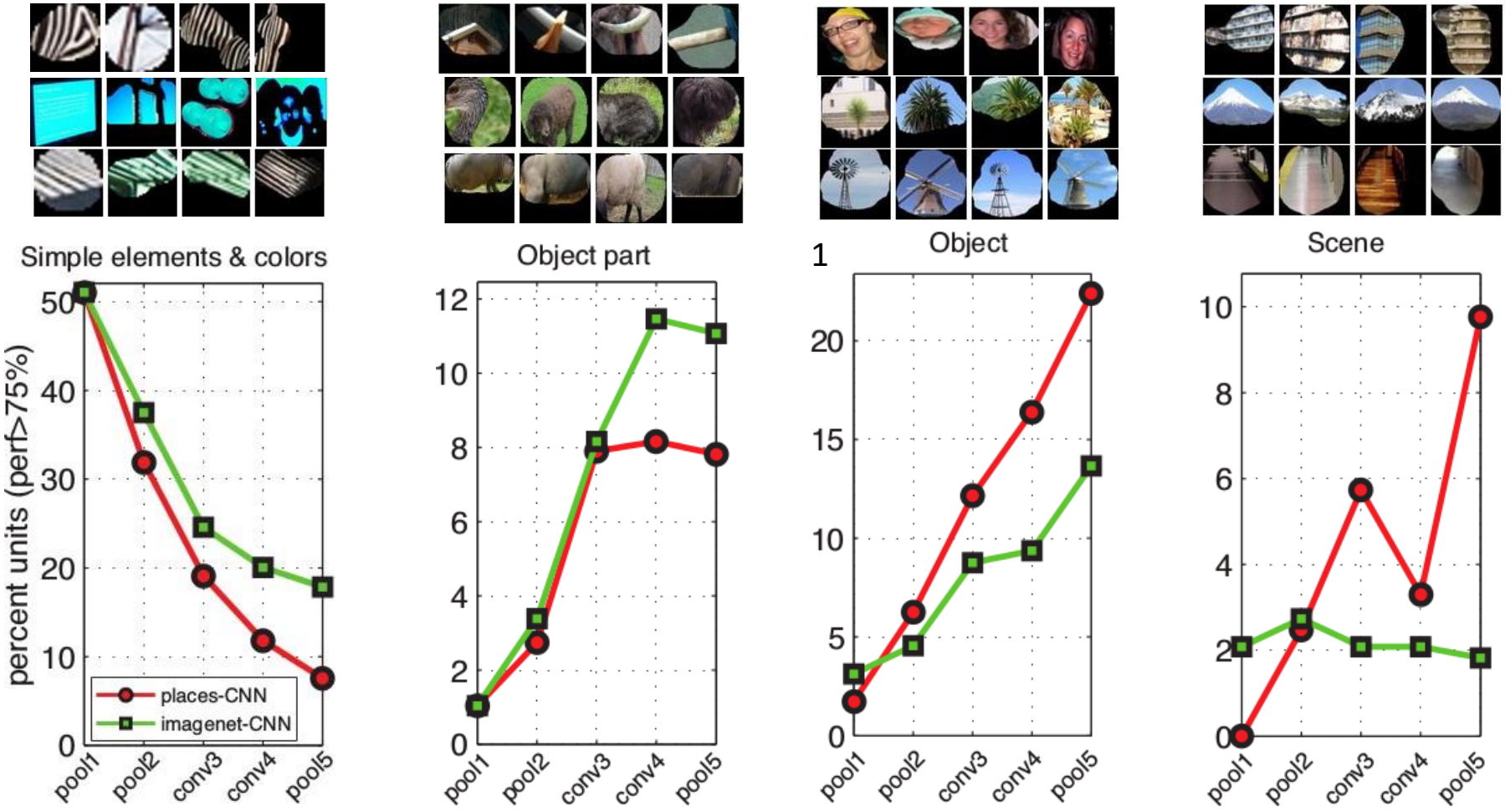
From Wikipedia, the free encyclopedia

In architecture, a **wedding-cake style** is an informal reference to buildings with many distinct tiers, each set back from the one below, resulting in a shape like a wedding cake, and may also apply to buildings that are richly ornamented, as if made in sugar icing.

- In Italy, the Monument to Vittorio Emanuele II is in wedding cake style.
- The British wedding-cake style was created by Sir Christopher Wren, who often placed a steeple at the top of a series of classically detailed diminishing lower stages as with St. Paul's Cathedral.
- In the United States, the style has been predominant in New York City, thanks to the 1916 Zoning Resolution,[1] a former zoning code which forced buildings to reduce their shadows at street level by employing setbacks, resulting in a ziggurat profile.[2] Many iconic New York buildings in the Art Deco style were designed to comply with these zoning regulations, and those designs were highly influential on building projects in other cities. The wedding cake design subsequently became a common feature of many Art Deco buildings around the world. The dome of the United States Capitol in Washington, D.C. is also described as being of wedding-cake style.
- In Russia, the wedding-cake style supercharged with boldly scaled classical detailing is a typical feature of Stalinist architecture.

120 Wall Street in New York, a skyscraper from 1930, is an archetype of wedding-cake architecture.

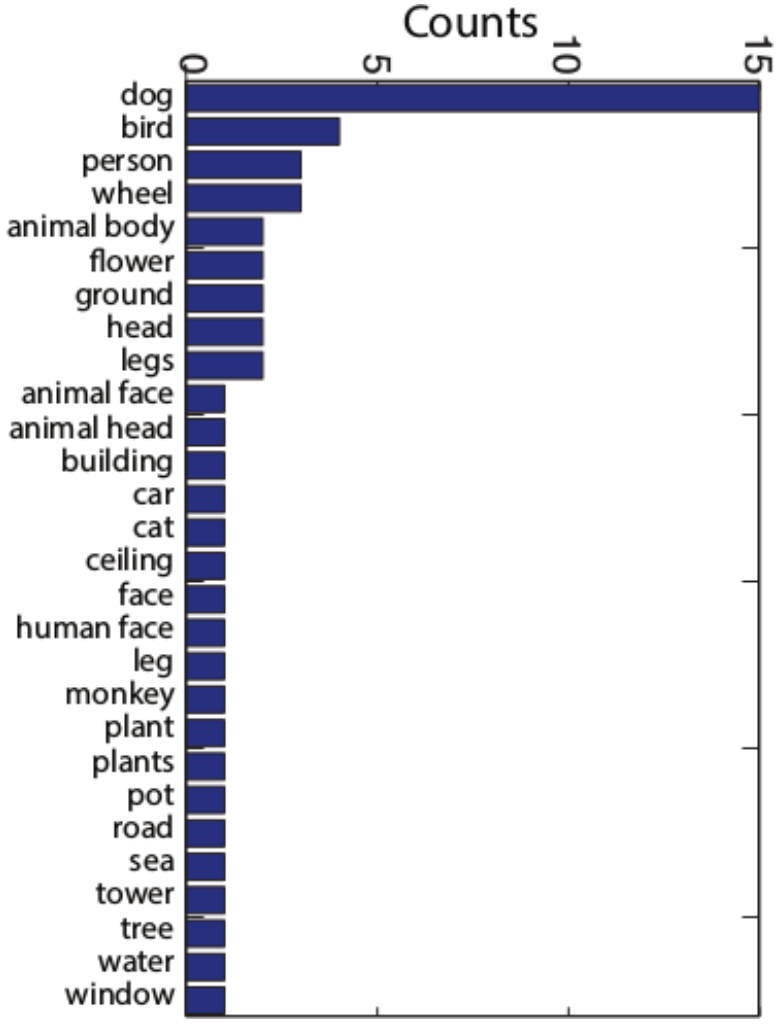
Distribution of Semantic Types at Each Layer



Object detectors emerge within CNN trained to classify scenes, without any object supervision!

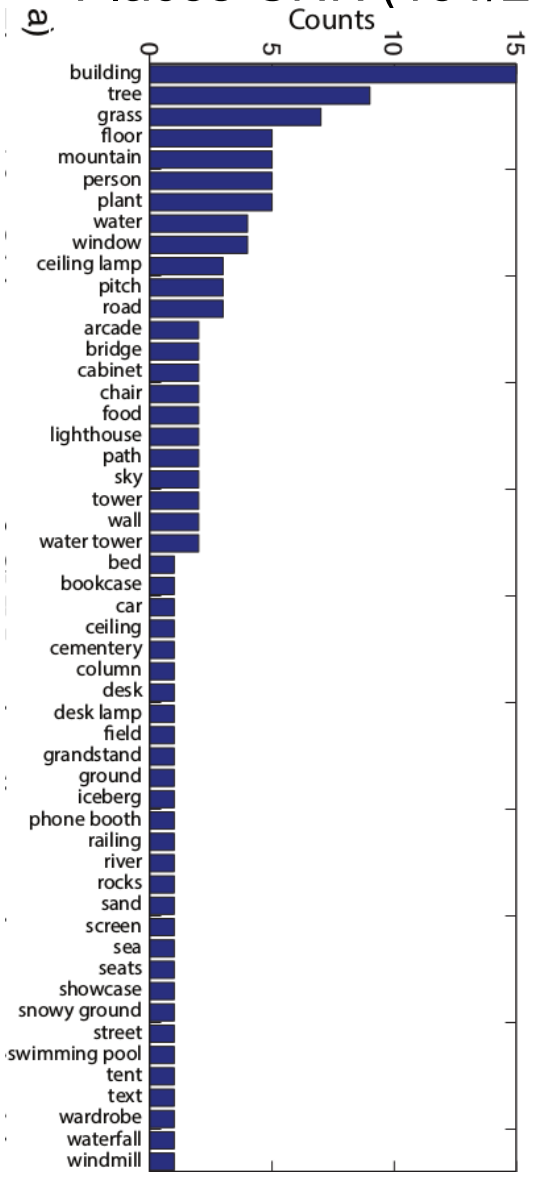
Histogram of Emerged Objects in Pool5

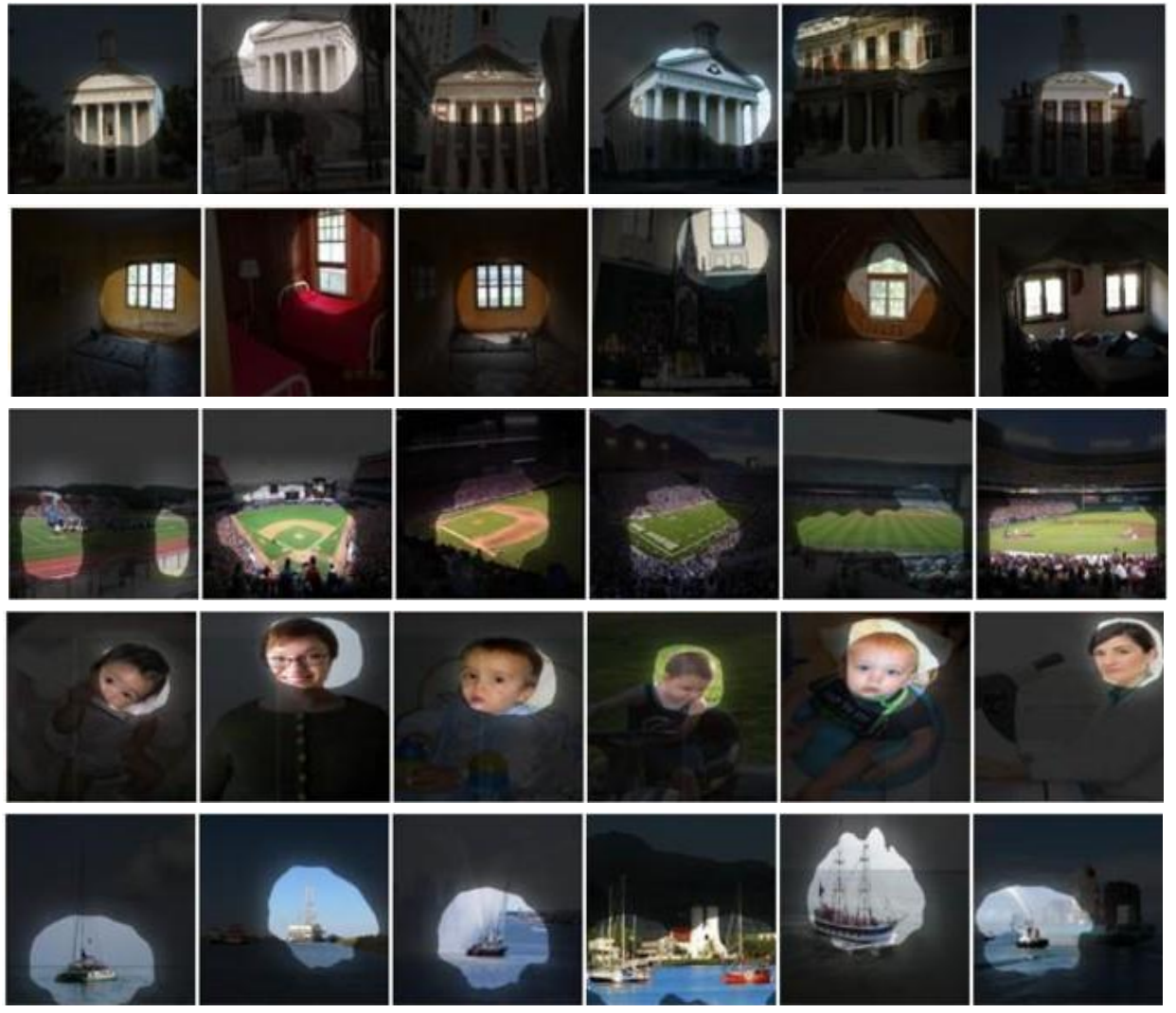
ImageNet-CNN (59/256)



Histogram of Emerged Objects in Pool5

Places-CNN (151/256)





Buildings

56) building

120) arcade

8) bridge

123) building

119) building

9) lighthouse

Furniture

18) billard table

155) bookcase

116) bed

38) cabinet

85) chair

People

3) person

49) person

138) person

100) person

Lighting

55) ceiling lamp

174) ceiling lamp

223) ceiling lamp

13) desk lamp

Nature

195) grass

89) iceberg

140) mountain

159) sand

Wrap up

- There are many ways to visualize what a neural network has learned
- Networks learn smaller receptive fields than the “theoretical” receptive field.
- As you go deeper in the network, the hidden activations correspond more to high-level semantic concepts
- Object detectors emerge inside a CNN trained to classify scenes, without any object supervision.