

Computer Vision I

2022. 7

정 준 수 PhD

Evolution's Big Bang



[This image is licensed under CC-BY 2.5](#)



[This image is licensed under CC-BY 2.5](#)



[This image is licensed under CC-BY 3.0](#)

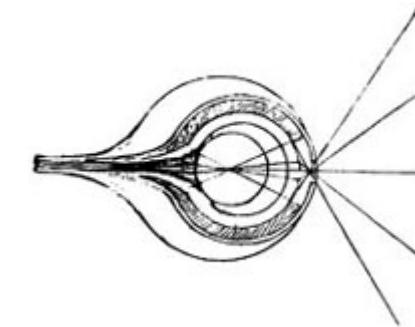
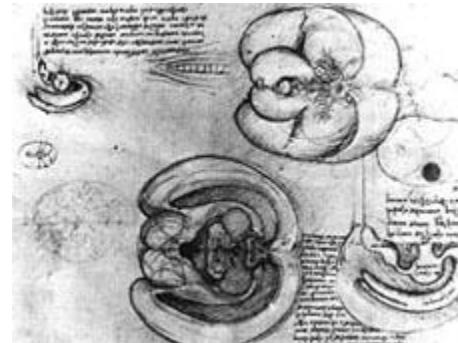
543million years, B.C.

Men of vision

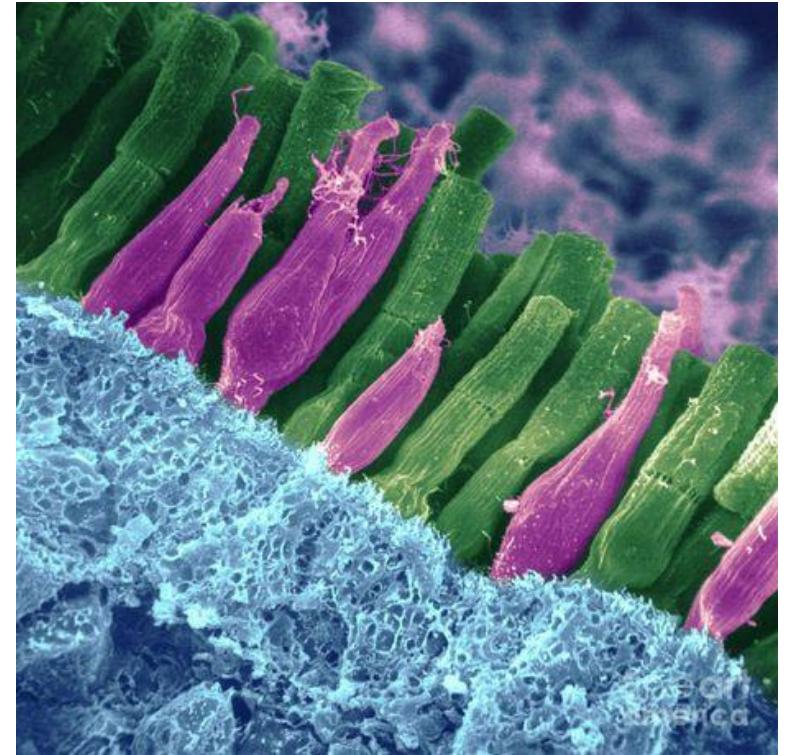
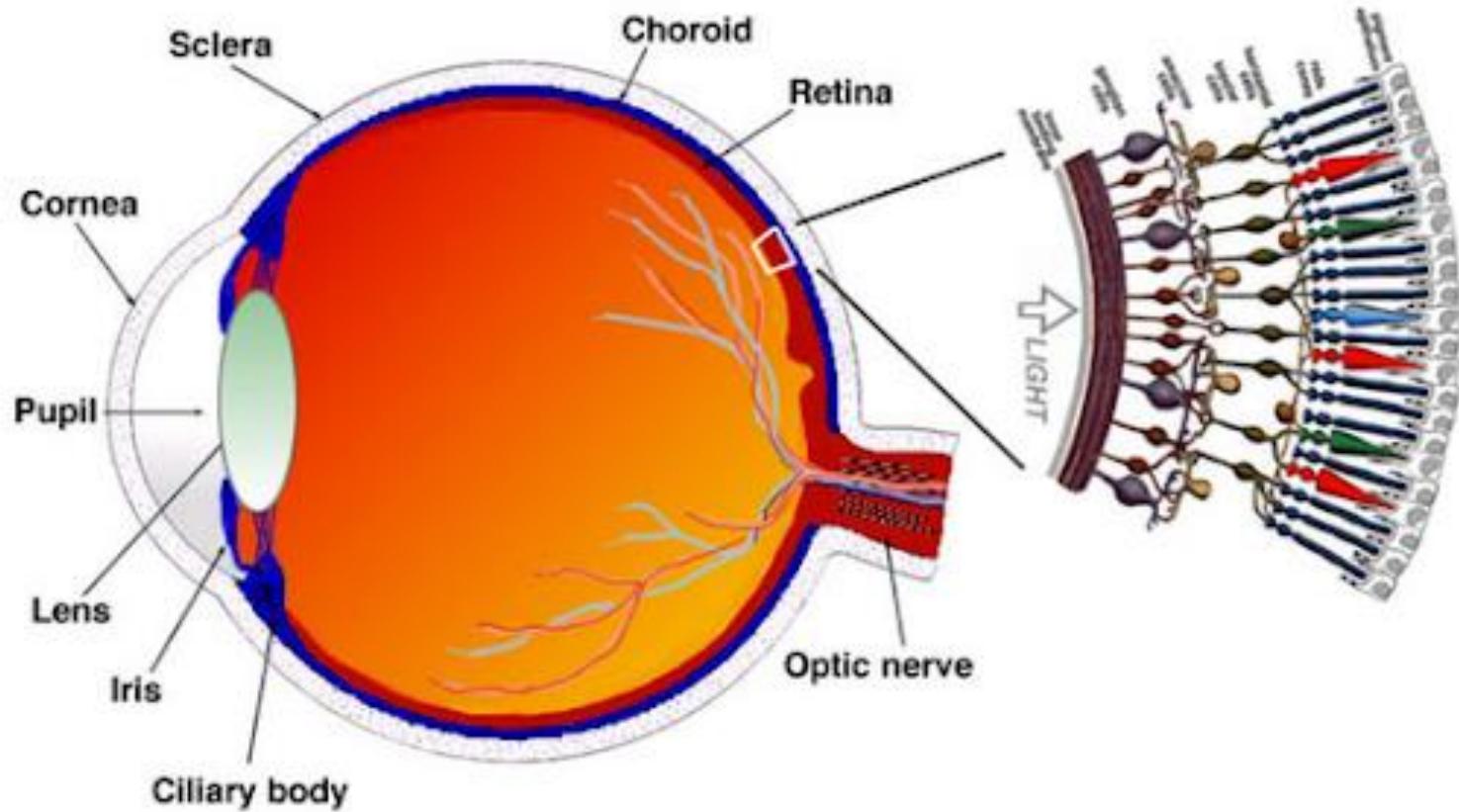
The study of vision has attracted many great thinkers over the centuries, not all of them scientists in the strictest sense, but we can call them all 'visionaries'

The Early Modern Period

Leonardo da Vinci proposed a theory that the inverted retinal image was re-inverted inside the brain. Below you can see two of Leonardo's drawings, dating from 1490 and 1506 respectively. the third image demonstrates his theory of the re-inversion of an inverted retinal image.



[출처] <https://www.college-optometrists.org/the-college/museum/online-exhibitions/virtual-eye-and-vision-gallery/men-of-vision.html>

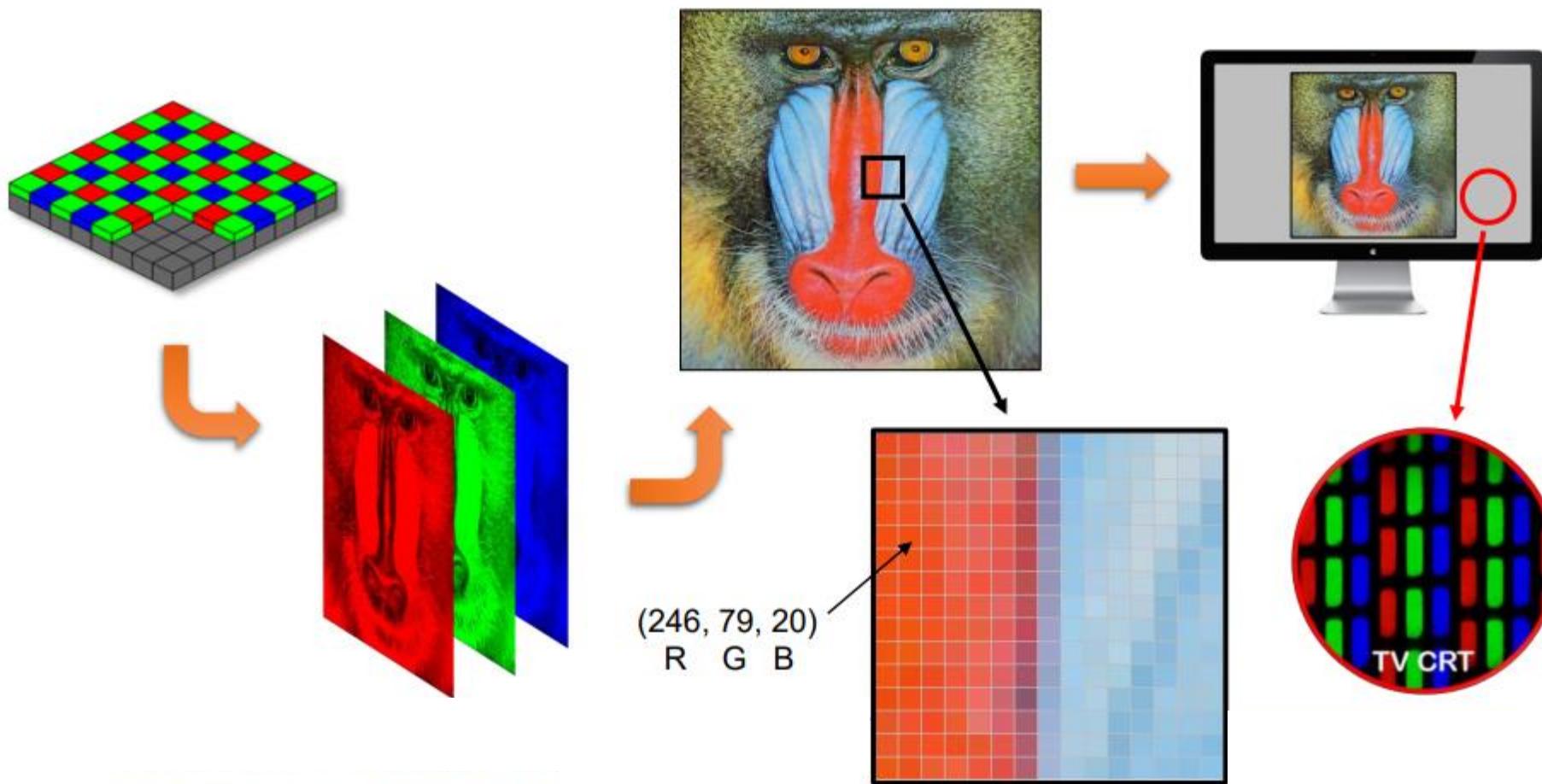


- 원추체 (색상)
- 간상체 (어둡고 밝) 물체의 모양

영상표현 방법

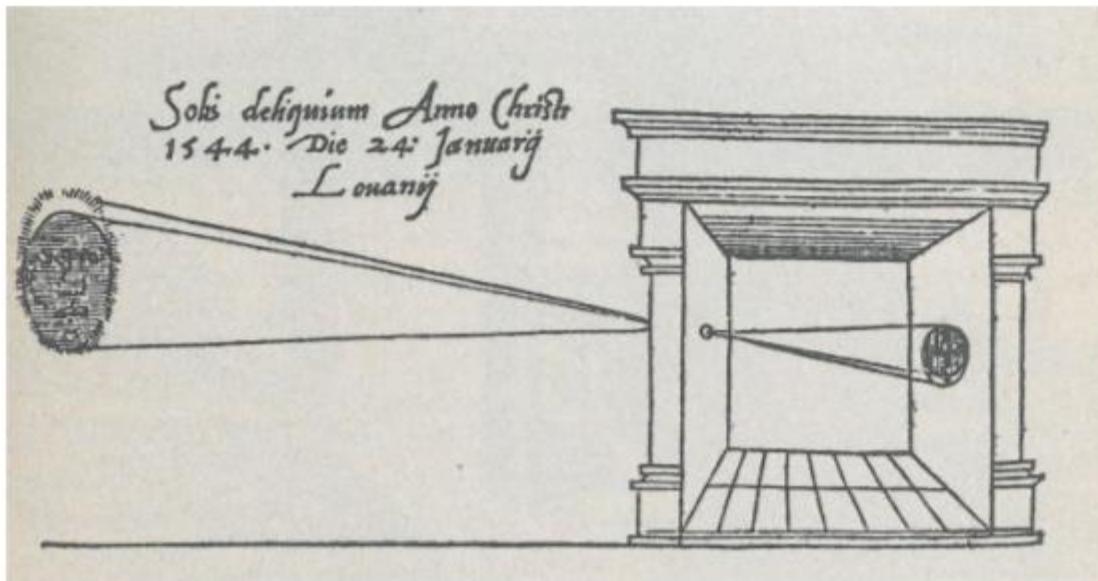
- 영상(image)이란?

- 픽셀(pixel)이 바둑판 모양의 격자에 나열되어 있는 형태 (2차원 행렬)

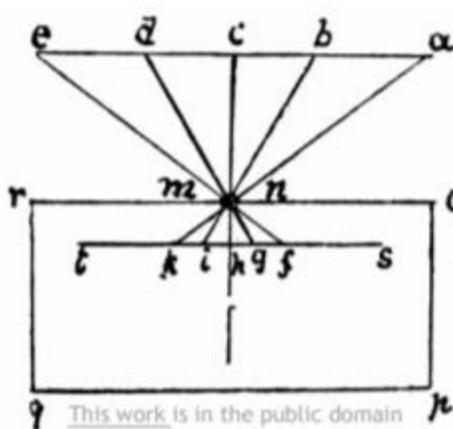
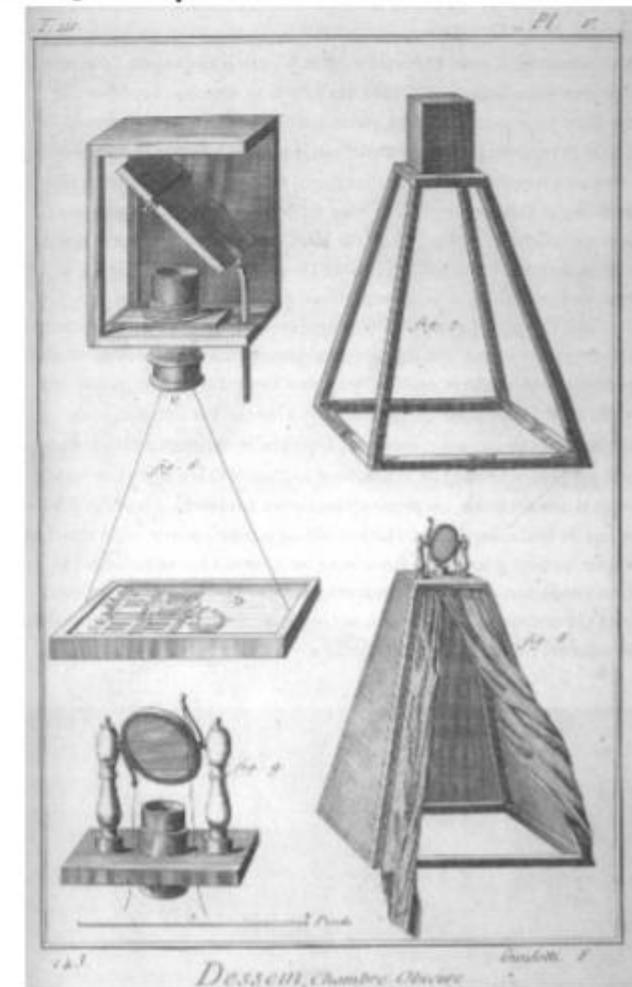


Camera Obscura

Gemma Frisius, 1545



Encyclopedie, 18th Century



Leonardo da Vinci,
16th Century AD

A Nobel Partnership: Hubel & Wiesel

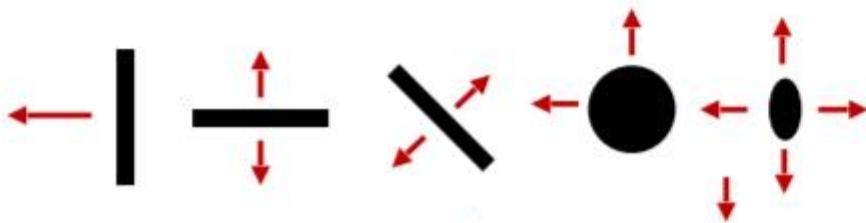


Hubel & Wiesel Come to Harvard

David Hubel and Torsten Wiesel came to Harvard from Johns Hopkins University with Steven Kuffler in the early 1960s to establish the Department of Neurobiology at Harvard Medical School. Their breakthrough discoveries about the visual system and visual processing earned them the Nobel Prize for Physiology or Medicine in 1981.

Hubel and Wiesel recorded electrical activity from individual neurons in the brains of cats. They used a slide projector to show specific patterns to the cats and noted that specific patterns stimulated activity in specific parts of the brain. Such single-neuron recordings were an innovation at the time, enabled by Hubel's earlier invention of a special recording electrode. They systematically created a map of the visual cortex with these experiments. The original film projector, light filters and slides, are held at the Warren Anatomical Museum at the Countway Library of Medicine.

<https://braintour.harvard.edu/archives/portfolio-items/hubel-and-wiesel>



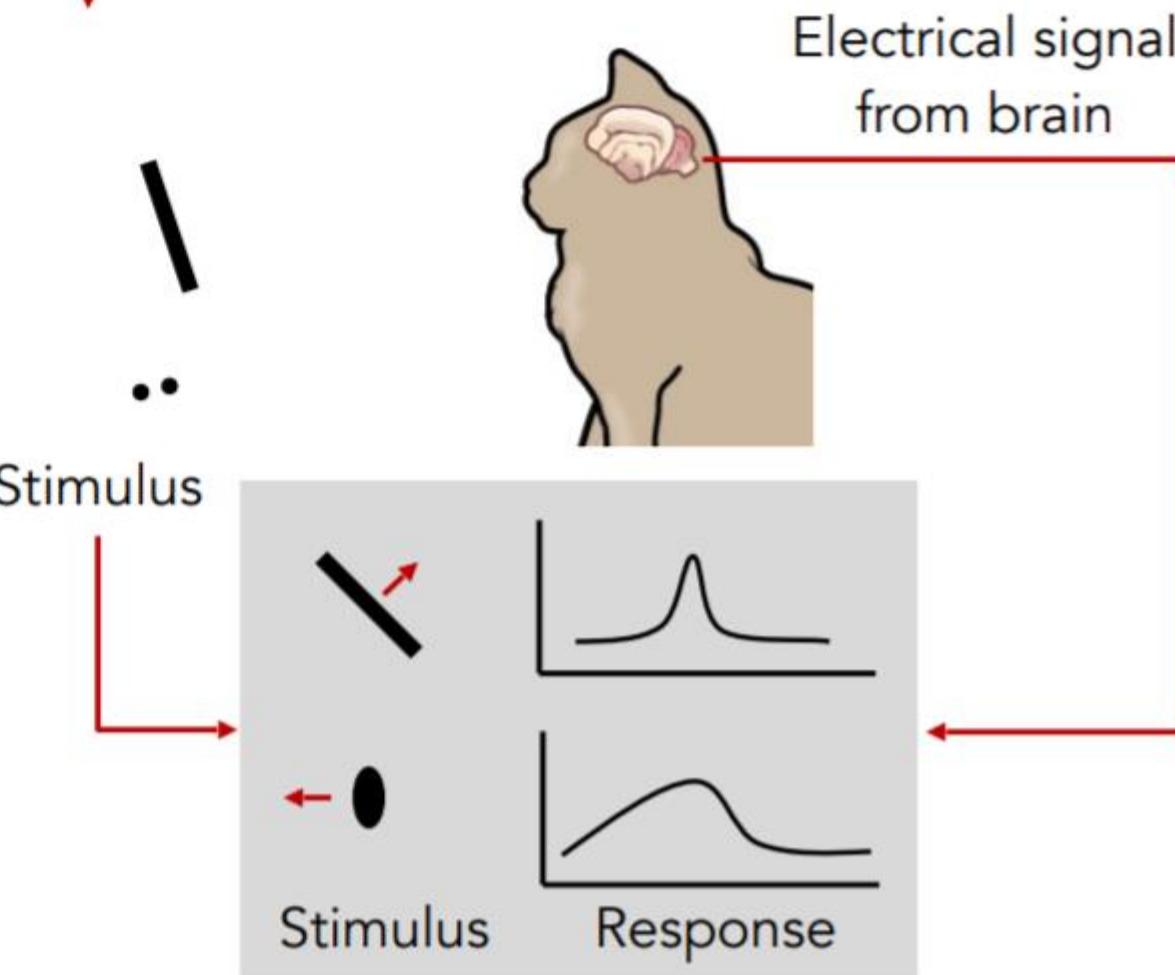
Simple cells:
Response to light orientation

Complex cells:
Response to light orientation and movement

Hypercomplex cells:
Response to movement with end point

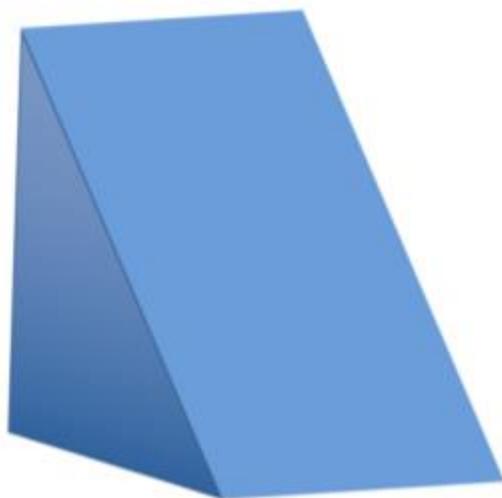


Hubel & Wiesel, 1959

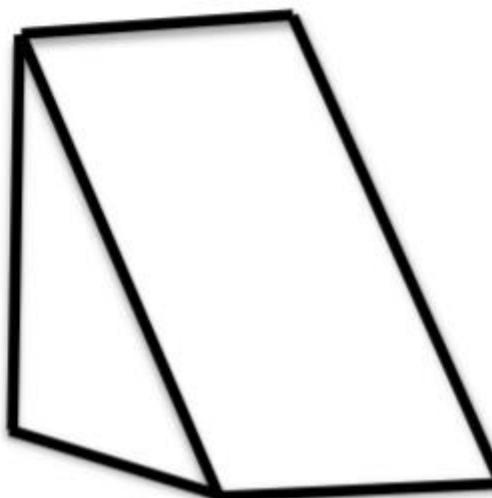


Block world

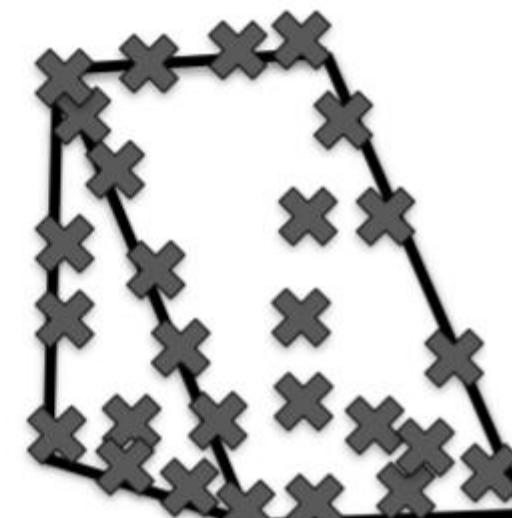
Larry Roberts, 1963



(a) Original picture



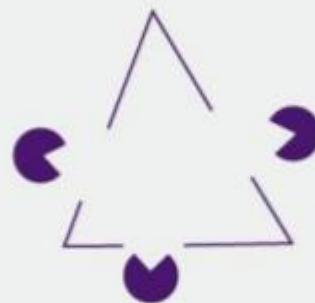
(b) Differentiated picture



(c) Feature points selected

Copyrighted Material

VISION



David Marr

FOREWORD BY
Shimon Ullman

AFTERWORD BY
Tomaso Poggio

Copyrighted Material

David Marr, 1970s

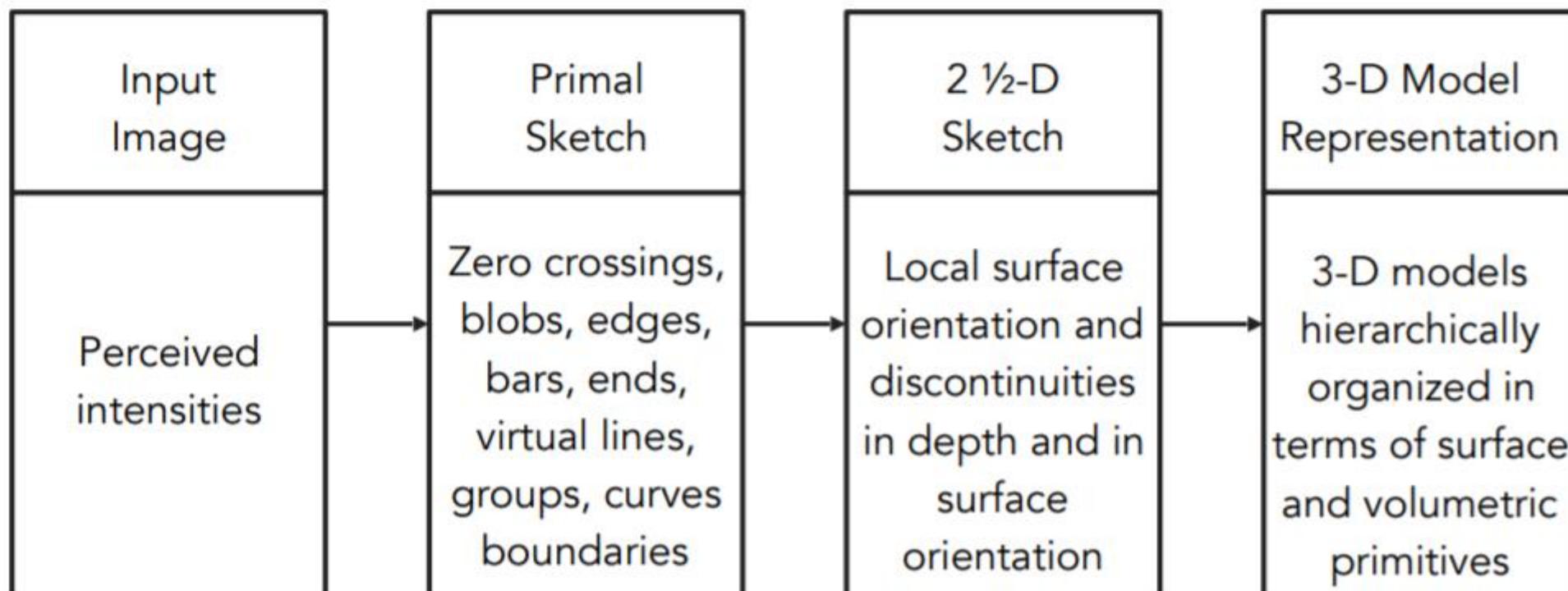
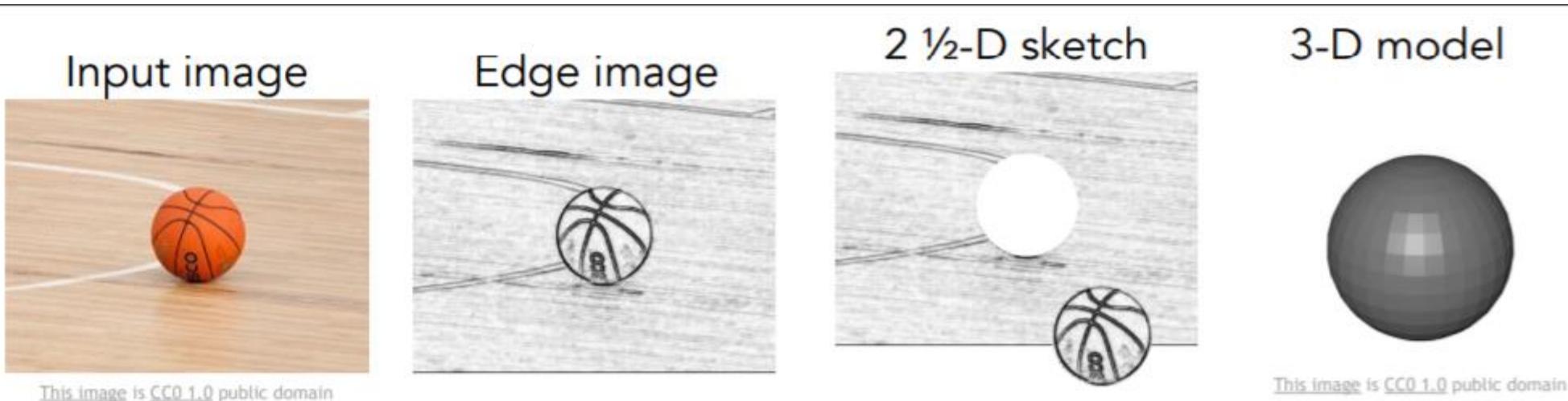
David Courtenay Marr (19 January 1945 – 17 November 1980) was a British neuroscientist and physiologist. Marr integrated results from psychology, artificial intelligence, and neurophysiology into new models of visual processing. His work was very influential in computational neuroscience and led to a resurgence of interest in the discipline.

인지심리학적 접근을 통해 새로운 시각적 대상재인 모형을 제시(Ellis & Young. 1996)

Marr(1982)의 시지각 이론을 바탕으로,

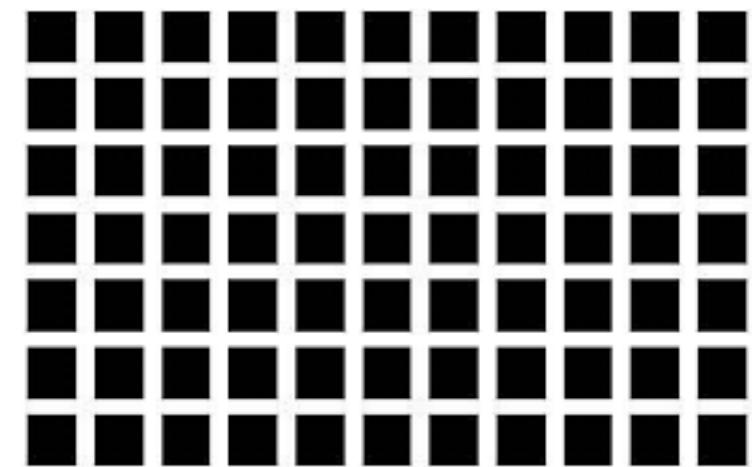
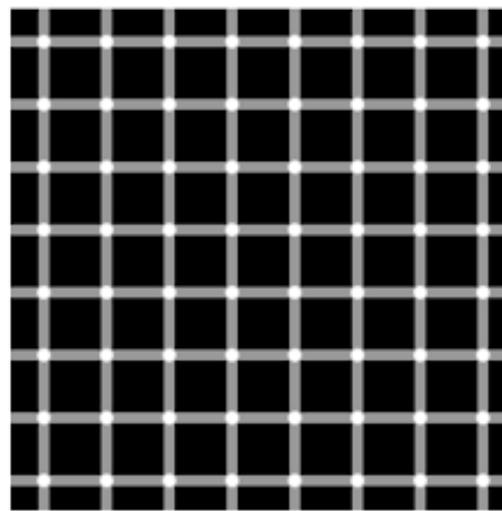
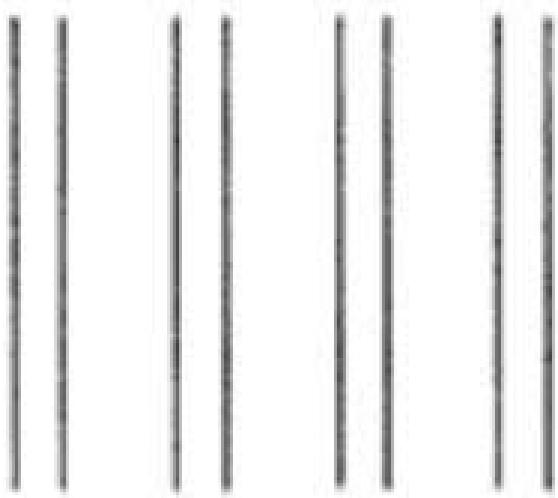
- 1) 망막에 맺힌 2D 영상으로부터 통합된 일차 스케치 생성(경계선, 윤곽, 명암 등)
- 2) 관찰자 중심의 관점에서 전체적인 형태, 윤곽 정보를 포함한 2.5D 영상의 생성
- 3) 대상항등성과 지각적 범주화를 포함한 3D표상의 생성(관찰자의 위치와 무관한 진정한 대상재인)

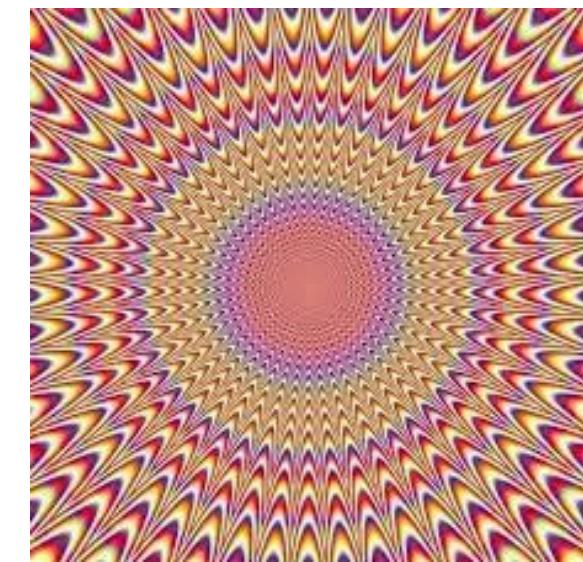
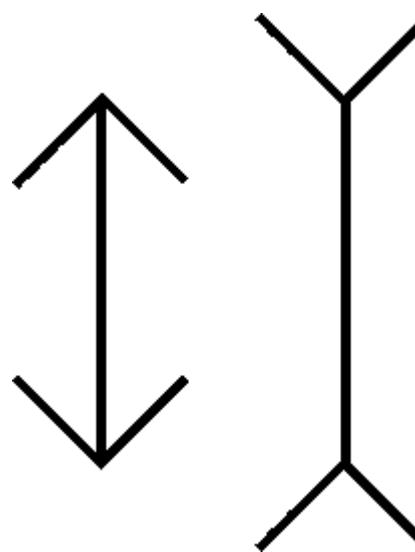
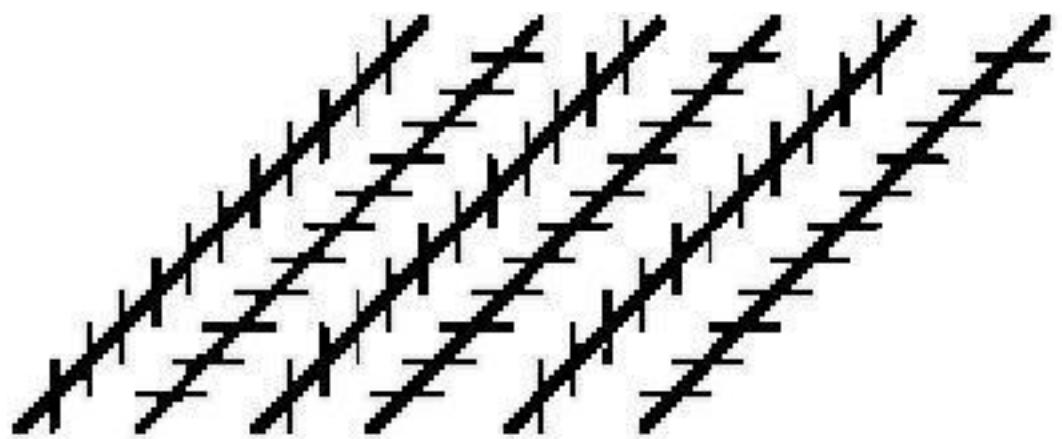
<http://www.aistudy.com/pioneer/Marr.D.htm->



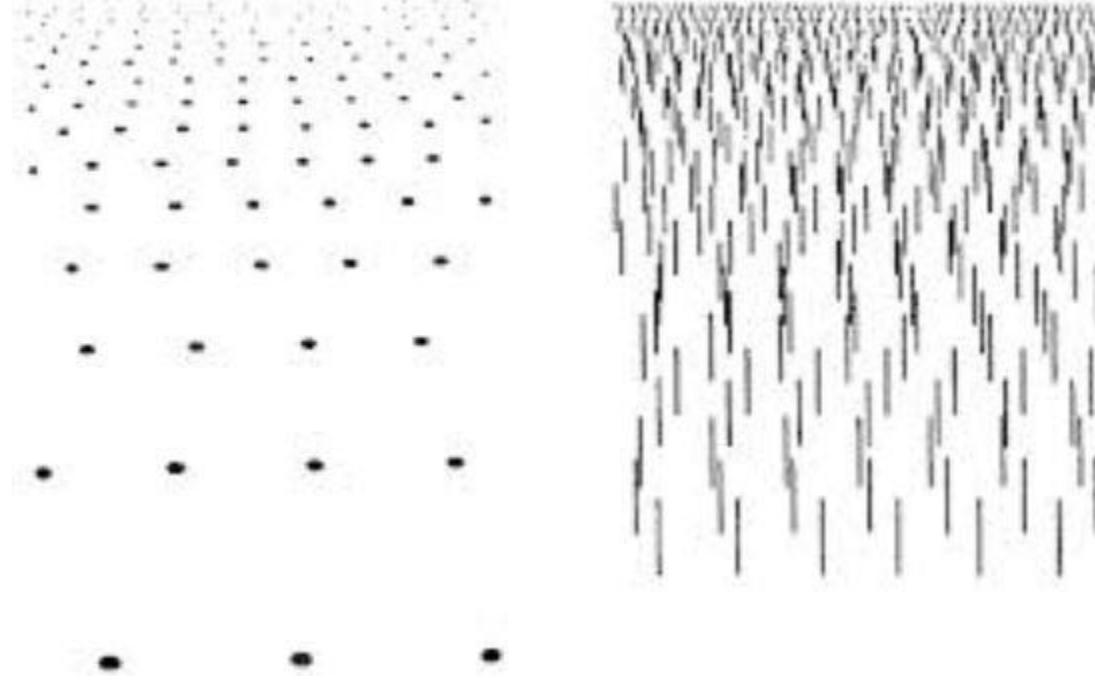
Stages of Visual Representation, David Marr, 1970s

착시 현상의 원인은?





깊이와 표면 지각



결기울기는 지각된 깊이에 대한 강력한 결정요인이며, 결기울기에서의 불연속은 다양한 결 표면들간의 공간적 관계에 대한 또 다른 정보를 제공한다. 따라서 왼쪽 그림의 결밀도 변화는 위로 올라가는 경사면의 느낌을 주고 오른쪽 그림에서의 변화는 갑작스러운 낙하, 즉 "시각절벽"의 느낌을 준다.

매우 강력한 일련의 깊이단서가 결기울기에 의해 제공되는데, 지각심리학에서 매우 영향력 있는 이론가인 James J. Gibson(1904~1979)은 결기울기의 중요성을 강조했다. 결기울기는 궁극적으로 조망에 의해 생성되며, 우리가 길에 깔린 자갈들을 보거나 초원에 널린 풀무더기들을 볼 때 어떤 것이 눈에 들어올까 생각해 보면 이해가 빠르다. Gibson은 이런 물체들이 망막에 투사된 상은 관계된 표면들의 공간적 배열에 따른 연속적인 변화 즉 결기울기를 반드시 보여야 한다고 지적함.

Yann André LeCun

(/la'kvn/ French pronunciation: born July 8, 1960) is a French computer scientist working primarily in the fields of machine learning, computer vision, mobile robotics, and computational neuroscience.

He is well known for his work on optical character recognition and computer vision using convolutional neural networks (CNN), and is a founding father of convolutional nets.

548

LeCun, Boser, Denker, Henderson, Howard, Hubbard, and Jackel

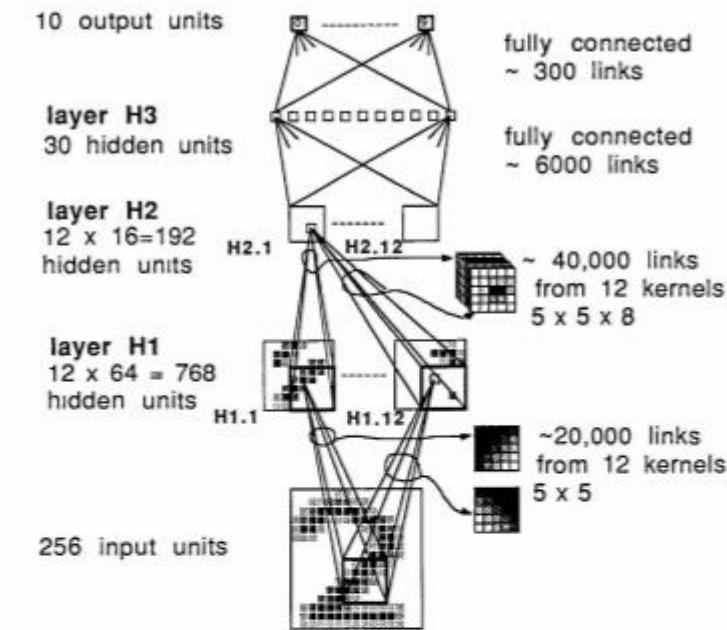


Figure 3 Log mean squared error (MSE) (top) and raw error rate (bottom) versus number of training passes

training set, 8.1% misclassifications on the test set, and 19.4% rejections for 1% error rate on the remaining test patterns. A full comparative study will be described in another paper

5.1 Comparison with Other Work. The first several stages of processing in our previous system (described in Denker *et al.* 1989) involved convolutions in which the coefficients had been laboriously hand designed. In the present system, the first two layers of the network are constrained to be convolutional, but the system automatically learns the coefficients that make up the kernels. This "constrained backpropagation" is the key to success of the present system: it not only builds in shift-invariance, but vastly reduces the entropy, the Vapnik-Chervonenkis dimensionality, and the number of free parameters, thereby proportionately reducing the amount of training data required to achieve a given level

시각작업기억의 처리 과정에는 기억 표상이 형성되고 유지되는 과정 뿐 아니라 기억된 항목에 대한 재인과 관련된 비교 과정이 포함된다(Hyun, Woodman, Vogel et al., 2006)

초기 시각 정보 처리

- 시상
- 물체지각
- 형태주의 체계화 원리 근접성의 원리 연속성의 원리 – 외측 슬상핵
- 2차원 그림 만들기
- 3차원 (깊이와 표면지각) 단안단서 결기울기 입체시 운동시차
- 비더만(Biederman) 물체재인 32개 components
- 형판맞추기

표상 (REPRESENTATION)

원래의 것과 같은 인상을 주는 이미지 또는 형상

- 정신적 표상은 정신 안에서 비교적 일관되게 재생산되는 의미 있는 사물이나 대상에 대한 지각을 일컫는다.
- 관념적 표상은 사고나 생각의 토대를 제공하는 정신적 표상으로서, 실질적으로는 정신적 표상과 동일하다.
- 본능적 표상은 자기 표상 안에 존재하는 개인의 욕동(원본능) 측면들을 말한다.

표상은 자아의 하부 구조를 구성하며 자아 내용물의 일부로 간주된다.

표상 학습(Representation learning)

표상 학습(representation learning) 혹은 특성 학습(feature learning)은 직접적인 데이터 대신, 유용한 정보를 더 쉽게 추출할 수 있게 만들어진 데이터의 표상(representation)을 통해 분류기나 다른 예측 기계를 학습시키는 것을 뜻한다.

확률적으로 좋은 표상은 입력을 설명할 수 있는 내재적인 설명 요인들의 posterior 분포를 포함할 때가 많다. 대화 인식, 신호 처리, 물체 인식, 자연어 처리와 같은 분야들에서 표상 학습 방법론은 실증적인 성공을 이루어 냈으며, 다양한 확률 모델과 인공 신경망, 그리고 딥러닝이 자동으로 표상 학습을 위한 특성 추출에 이용된다.

정보처리적 패러다임의 인지과학은 마음에 대한 보는 틀을 이와 같이 상정하고 나서, 정보처리체계로서의 마음의 작용을 감각, 지각, 학습, 기억, 언어, 사고, 정서 등의 여러 과정으로 나눈 다음, 각 과정에서 어떠한 정보처리가 일어나는가, 각 과정들은 어떻게 상호 작용 하는가를 묻고, 각 과정에서 어떠한 정보(지식)구조, 즉 표상(표현)구조가 관련 되는가를 규명하려 한다. 따라서 마음의 현상, 즉 심리적 사건은 정보의 내용 및 정보를 처리하는 사건으로 개념화 되어지는 것이다.

표상(표현: representation)의 강조 (Fodor, 1975)

인간과 컴퓨터가 자극 정보를 어떠한 상징으로 기억에 저장한다는 것은 자극 자체를 저장하는 것이 아니라 자극에 대한 표상(표현)을 저장하는 것이며 이는 마음과 컴퓨터 모두가 자극의 정보를 내적 기호(상징)로 변화시켜 기억에 보유한다는 것이다. 따라서 무엇을 안다는 것은 이들 **표상간의 연관을 찾거나 새로운 관계성을 만들어 낸다는 것을 의미한다.** 따라서 암의 과정에 대한 연구는 자극들이 어떻게 상징(기호) 표상들로 전환되고 또 활용 되는가를 연구하는 것이라 하겠다. 즉 **인지과학의 핵심 연구주제는 마음이나 컴퓨터에서의 표상의 처리과정(계산)과, 표상의 본질 및 그 구조적 특성의 연구라고 할 수 있다.**

핵심어 -> 연관성과 추상화

게슈탈트 심리학([독일어](#): Gestaltpsychologie, Gestalt 心理學)은 [심리학](#)의 한 학파이다. 인간의 정신 현상을 개개의 감각적 부분이나 요소의 집합이 아니라 하나의 그 자체로서 전체성으로 구성된 구조나 갖고 있는 특질에 중점을 두고 이를 파악한다. 이 전체성을 가진 정리된 구조를 독일어로 게슈탈트(Gestalt)라고 부른다. 박은정 교수에 따르면 이 단어의 의미는 전체 형태의 모양, 배열인데 지금 이 순간의 경험이다. 한편 형태심리학(形態心理學)을 나타내는 독일어이다.

HUMAN COGNITIVE NEUROPSYCHOLOGY

A Textbook with Readings

Andrew W. Ellis
and
Andrew W. Young

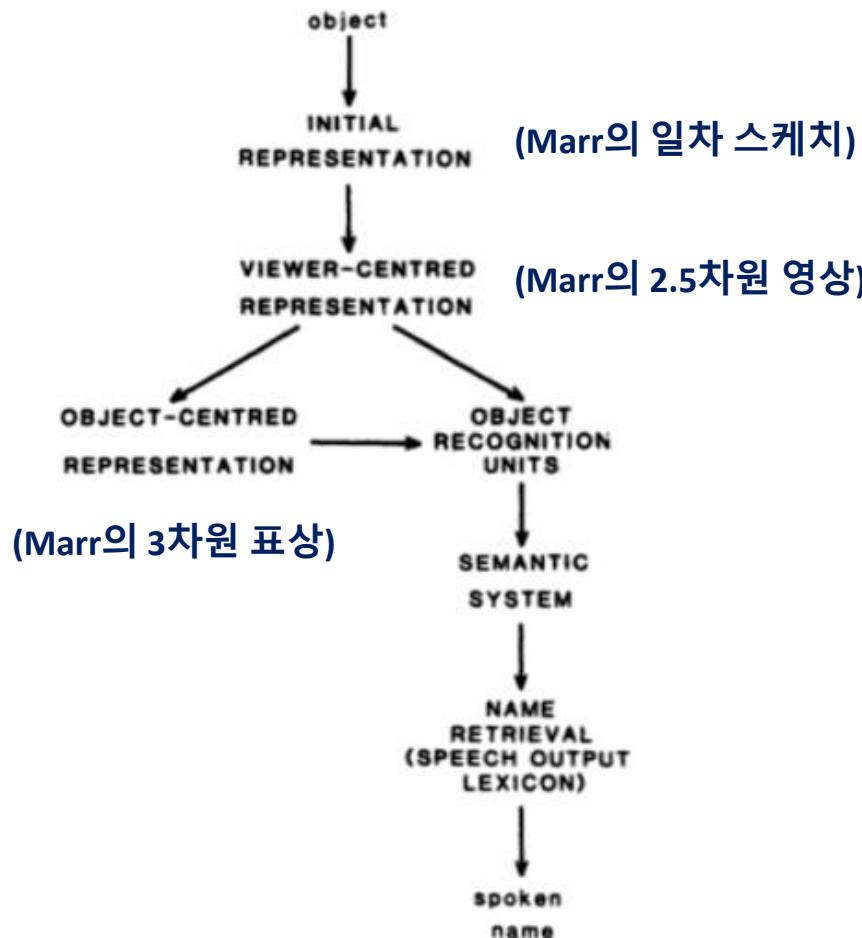


FIG. 2.1. Functional model for object recognition.

James Jerome Gibson (/ˈgɪbsən/; January 27, 1904 – December 11, 1979), was an [American psychologist](#) and one of the most important contributors to the field of [visual perception](#). Gibson challenged the idea that the nervous system actively constructs conscious visual perception, and instead promoted [ecological psychology](#), in which the mind directly perceives environmental stimuli without additional cognitive construction or processing.

생태학적 지각이론

생태학적 지각이론(Ecological Theory of Perception)은 김슨(James J. Gibson)에 의해 체계적으로 발전된 이론으로, 지각을 주변환경 안에서 일어나는 생존을 위한 적극적 행위로 인식한다.

즉 인간의 감각기관들을 외부의 자극에 대하여 수동적이기만한 수용기(受容器)로 보지 않으며, 정보탐색을 위한 적극적인 지각 시스템으로 본다.

이 이론은 인간과 환경이 서로 깊은 영향을 주고받으며, 상호보완하고 의존하는 모든 환경 내의 지각대상물들을 토탈 시스템의 차원에서 인식하는 시스템적 접근방식을 취한다.

AARON BEN-ZEEV*

J. J. GIBSON AND THE ECOLOGICAL
APPROACH TO PERCEPTION

"Whom he loves he rebukes" (*Proverbs*, 3, 12)

THE CONTRIBUTION of J. J. Gibson to psychology spans half a century; his first article was published in 1929, and his last book appeared in 1979. In this paper, I deal with his mature theory as it is expressed in his two last books: *The Senses Considered as Perceptual Systems* (1966), and *The Ecological Approach to Visual Perception* (1979); occasionally, earlier claims will be mentioned (especially from his book *The Perception of the Visual World* [1950]).¹ The aim of this paper is to examine the conceptual foundations, rather than the empirical claims of Gibson's view. Concerning the empirical level, Gibson's contributions to the understanding of perceptual processes are extremely significant, but this is not the place to describe them. His conceptual contribution has been less often evaluated, and there has been little agreement over such evaluation. I attempt, in this paper, to clarify some theoretical issues concerning his view.

Gibson's approach already has advocates who clarify and elaborate different aspects and implications of this approach. Due to limitations of space, I will not deal with their views separately, but merely discuss them when they clarify Gibson's own position.²

In the first section of this paper, I compare Gibson's approach to perception with the traditional ones; I indicate that Gibson suggests solutions to traditionally basic difficulties. Then I deal with some basic conceptual issues in

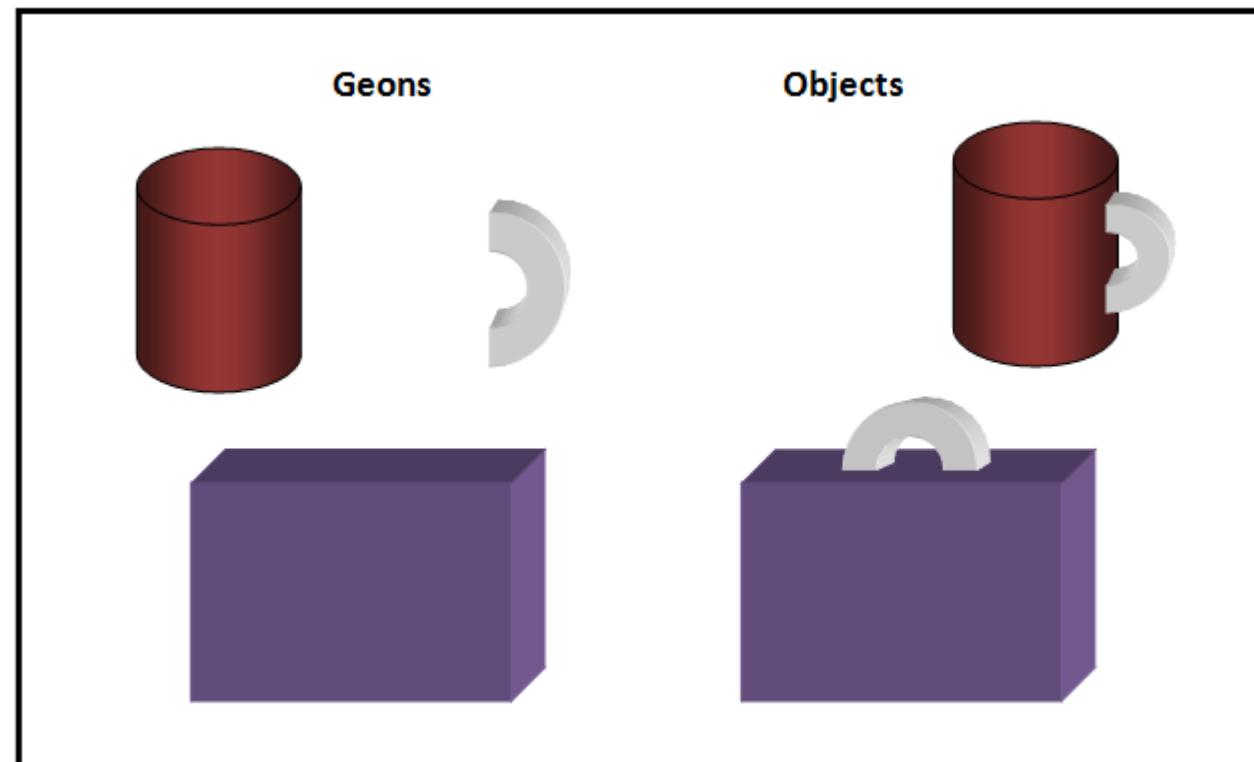
*Department of Philosophy, University of Haifa, Mt. Carmel, Haifa 31999, Israel.

'All these works published by Houghton Mifflin, Boston. These books will be referred to by date in the text. On the development of Gibson's ideas from 1929 to 1973 see T. J. Lombardo, *J. J. Gibson's Ecological Approach to Visual Perception: Its Historical Context and Development* (Unpublished doctoral dissertation, University of Minnesota, 1973).

²The papers most relevant for our discussion were written by R. E. Shaw and his colleagues: R. E. Shaw and J. Bransford, 'Introduction: Psychological Approaches to the Problem of Knowledge', *Perceiving, Acting and Knowing: Toward an Ecological Psychology*, R. E. Shaw and J. Bransford (Eds) (Hillsdale: Lawrence Erlbaum, 1977); R. E. Shaw, M. T. Turvey and W. Mace, 'Ecological Psychology: The Consequence of a Commitment to Realism', *Cognition and Symbolic Processes*, II, W. Weimer and D. Palermo (Eds) (Hillsdale: Lawrence Erlbaum, in press); R. E. Shaw and M. T. Turvey, 'Coalitions as Models for Ecosystems: A Realist Perspective on Perceptual Organization', *Perceptual Organization*, M. Kubovy and J. Pomerantz (Eds) (Hillsdale: Lawrence Erlbaum, in press); M. T. Turvey and R. E. Shaw, 'The Primacy of Perceiving: An Ecological Reformulation of Perception for Understanding Memory', *Perspectives on Memory Research*, L. G. Nilsson (Ed) (Hillsdale: Lawrence Erlbaum, 1979).

지각을 환경에 적응하려는 인간의 행태를 지원하는 기능으로 보며, 인간의 움직임에 따라 주변조건이 변화하는 역동적인 지각현상으로 해석하고, 인간과 환경간의 상호관계를 인간의 지각과 행동의 측면에서 설명한다. 즉, 생태학적 지각이론에서의 지각은 인간이 주변 환경에 적응하여 활동하기 위해 이루어진 환경과의 직접적인 관계이며, 주변 환경을 구성하는 구체적 사건, 사물 혹은 배치상태에 관한 정보를 환경으로부터 직접 습득하는 과정이다. 그리고 이러한 지각을 통해 얻는 정보의 핵심은 환경 내에서의 원활한 활동을 위한 수단을 얻고자 하는 것이며, 주변환경이 지각자에게 제공하는 행태지원성(affordance)이라는 것이다. 따라서 인간은 생태적 환경 내의 다양한 행태지원성을 지각하고 이러한 지각을 바탕으로 환경 내에서의 행위를 결정하게 된다. 즉, 지각은 행태를 지원하며, 행태는 지각에 필요한 정보를 제공하는 상호작용인 것이다. 이는 지각이 주변 환경으로부터 다양한 행태지원성에 대한 정보를 습득하는 것과 아울러 자신의 행동능력과 환경이 가지고 있는 행태지원성 사이에서 적합성을 찾아내려는 시도임을 의미한다. 생태학적 지각이론은 공간디자이너들로 하여금 어떠한 공간을 구성할 때, 그 기저에 시각적 복잡성과 같은 물리적 구성의 차원을 넘어 지원성과 같은 생태적 특성이 존재함을 인식시켰고, 이에 기반하여 인간과 환경 또는 공간의 상호관계에 적합성을 제공할 수 있는 디자인이 수행되어야 함을 이론적으로 뒷받침하였다. 또한 생태학적 지각이론은 인간행태를 일상적인 환경 하에서의 지각현상으로 설명하기 때문에, 환경디자인 실무에 적용 가능한 실제적 이론으로 발전되었다.

The **recognition-by-components theory**, or **RBC theory**,^[1] is a bottom-up process proposed by [Irving Biederman](#) in 1987 to explain [object recognition](#). According to RBC theory, we are able to recognize objects by separating them into **geons** (the object's main component parts). Biederman suggested that **geons** are based on basic 3-dimensional shapes (cylinders, cones, etc.) that can be assembled in various arrangements to form a virtually unlimited number of objects

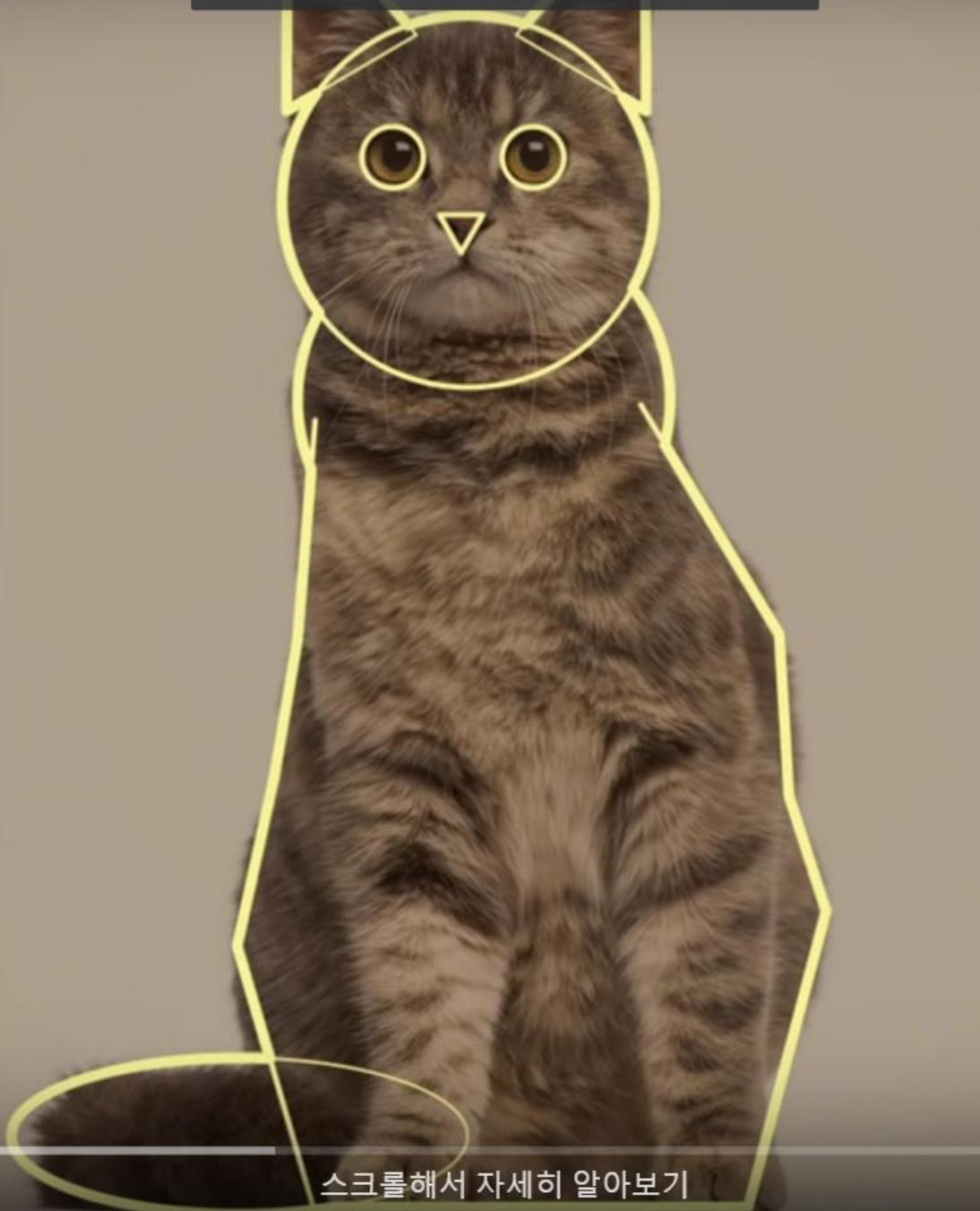


전체화면을 종료하려면 Esc 을(를) 누르세요.



스크롤해서 자세히 알아보기

전체화면을 종료하려면 Esc 을(를) 누르세요.



스크롤해서 자세히 알아보기

형태재인이란?

표상(Representation): 외부의 시각 정보를 받아 들여 어떠한 방식으로 우리 내부에서 재현

형태재인: 표상을 이미 내부에 저장되어 있는 시각적 사물들에 대한 기억이나 기준의 표상들과 대조하는 과정

- 자극 분석, 비교, 기업 탐색, 의사결정 등의 과정을 포함하는 복잡한 과정
- 시각, 청각 등의 여러 감각 정보에 대해 이루어질 수 있으며, 정보 처리도 2차원적인 것과 3차원적인 것 모두 일어날 수 있음

형태재인의 고전적 모형

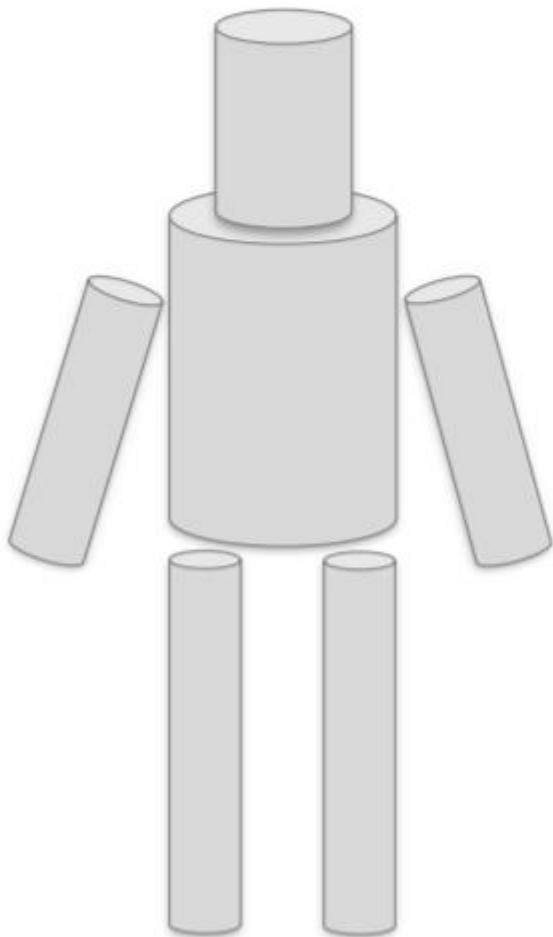
- 형판 맞추기(Template matching)모형
- 세부특징 분석모형(Feature detectors)

형태재인의 현재 모형

- 공간주파수 분석 접근
- 계산적 접근(Computational approach): David Marr의 시지각 계산모형

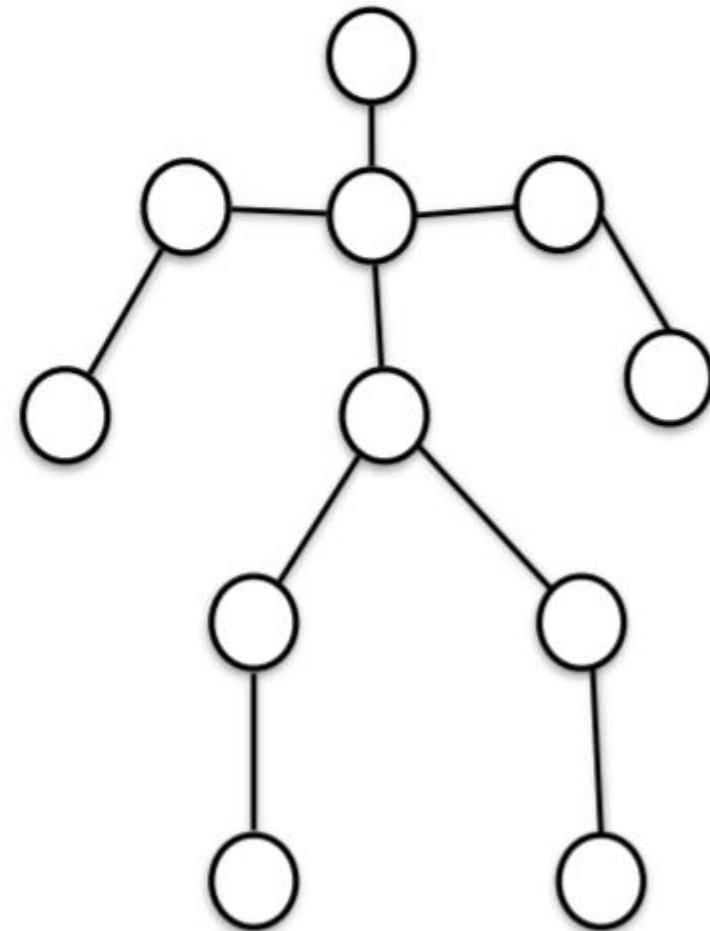
- Generalized Cylinder

Brooks & Binford, 1979

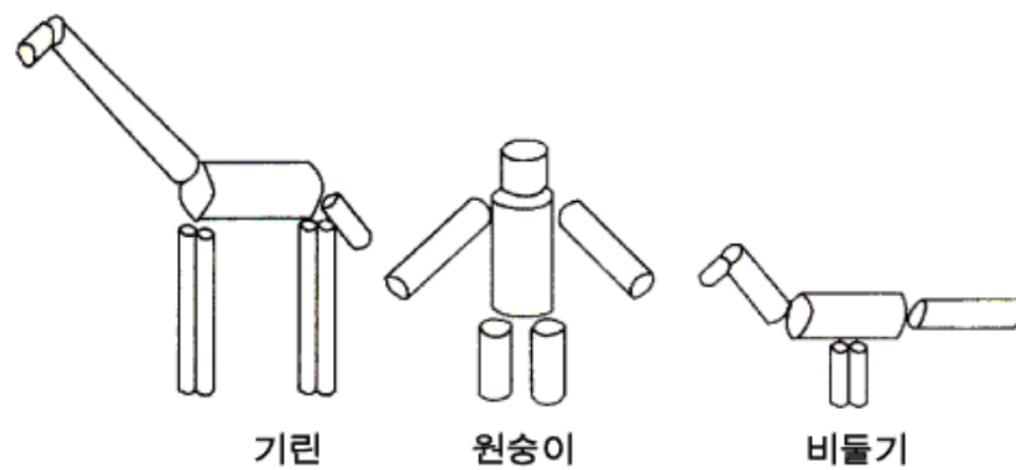
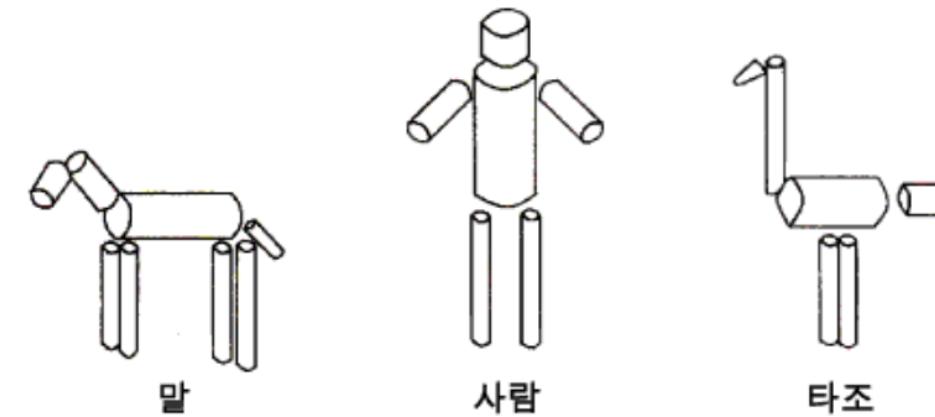


- Pictorial Structure

Fischler and Elschlager, 1973

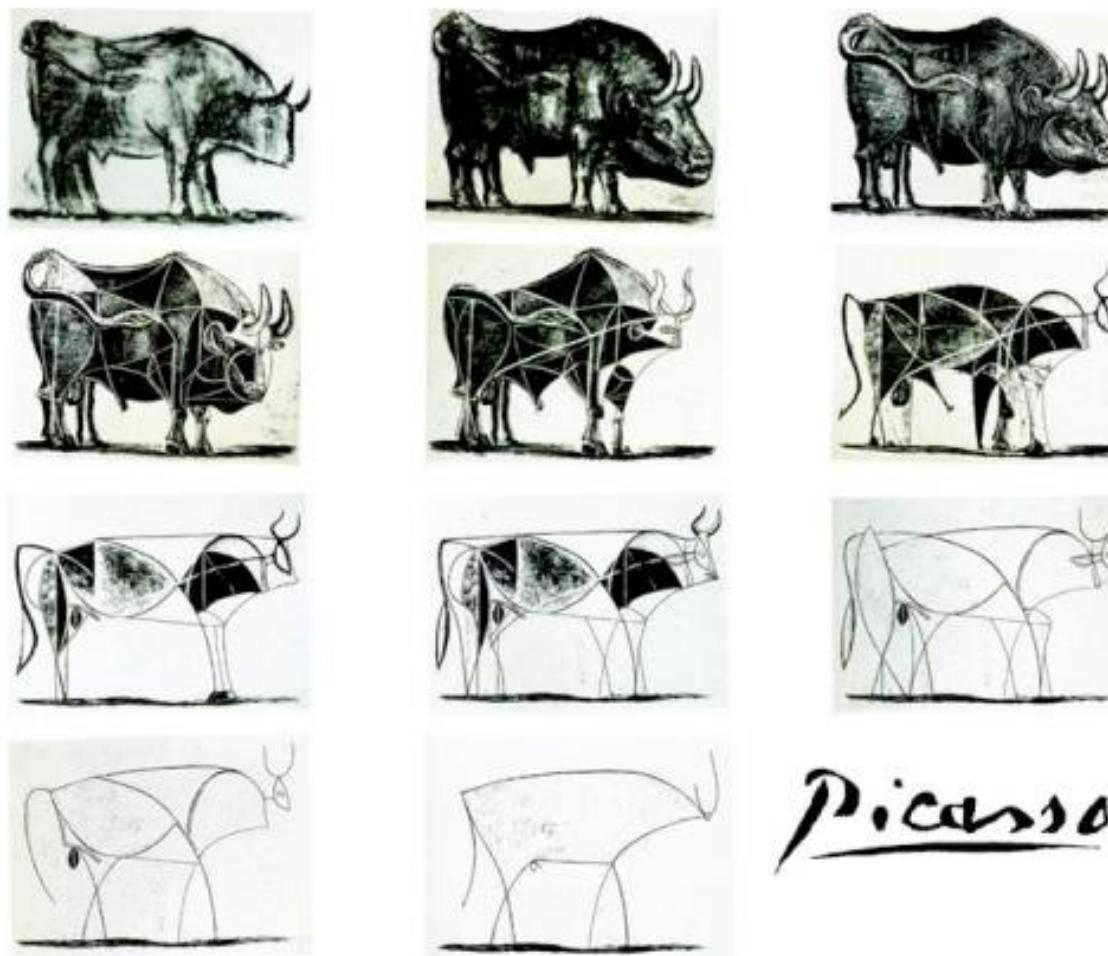


◆ 사람의 물체(시각) 재인 방식



친숙한 물체들을 기본 원통 형태로 분할한 것. (Marr & Nishihara. 1978)

◆ 추상화 (Abstract)



Pablo Picasso, Bull (plates I - XI) 1945



Image is CC BY-SA 4.0

David Lowe, 1987

Normalized Cut (Shi & Malik, 1997)

[Image](#) is CC BY 3.0



[Image](#) is public domain



[Image](#) is CC-BY SA 3.0







[Image](#) is public domain

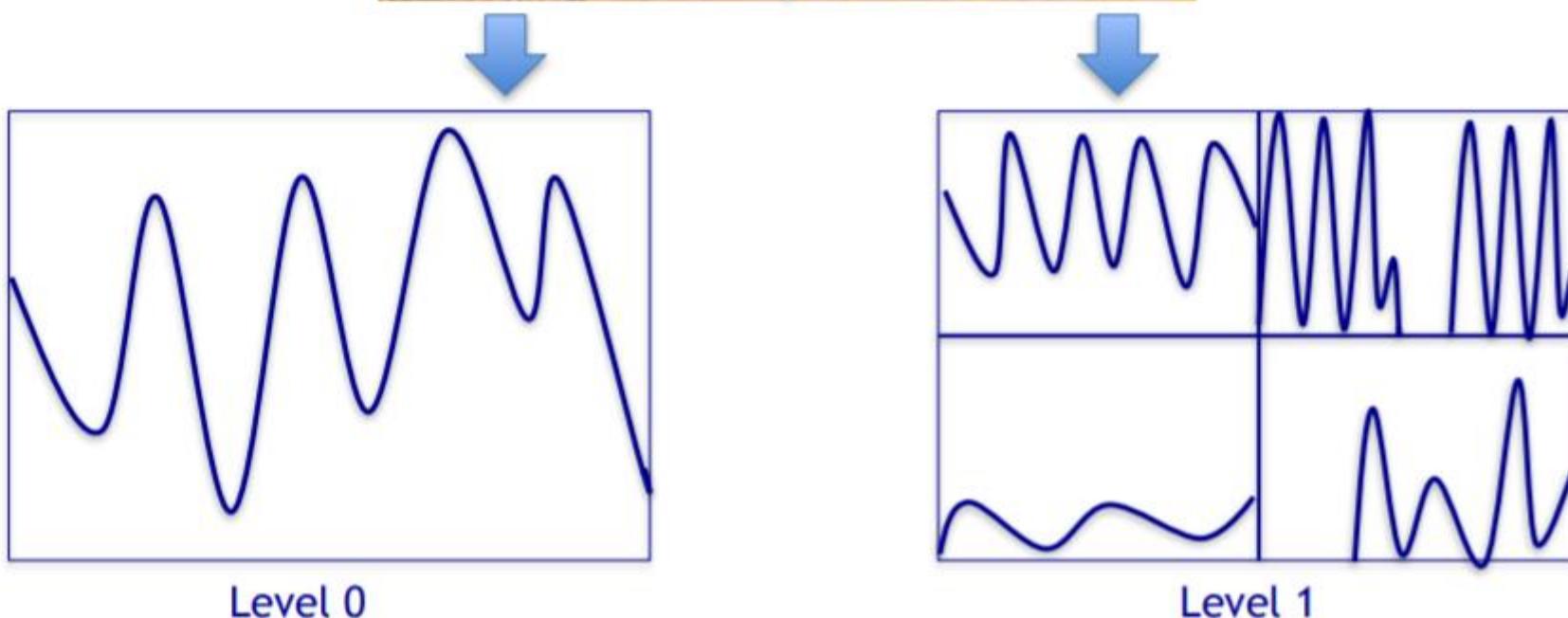


[Image](#) is CC BY-SA 2.0

"SIFT" & Object Recognition, David Lowe, 1999

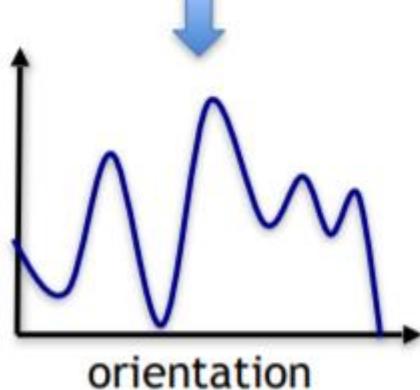


Image is CC0 1.0 public domain

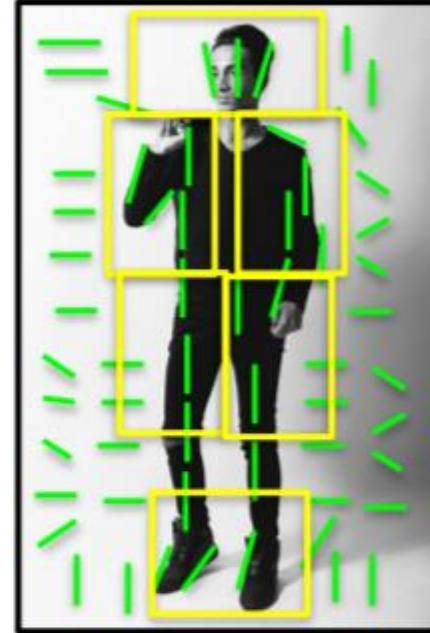


Spatial Pyramid Matching, Lazebnik, Schmid & Ponce, 2006

<https://www.youtube.com/watch?v=hPCTwxF0qf4>



Histogram of Gradients (HoG)
Dalal & Triggs, 2005



Deformable Part Model
Felzenswalb, McAllester, Ramanan, 2009

PASCAL Visual Object Challenge

(20 object categories)

[Everingham et al. 2006-2012]

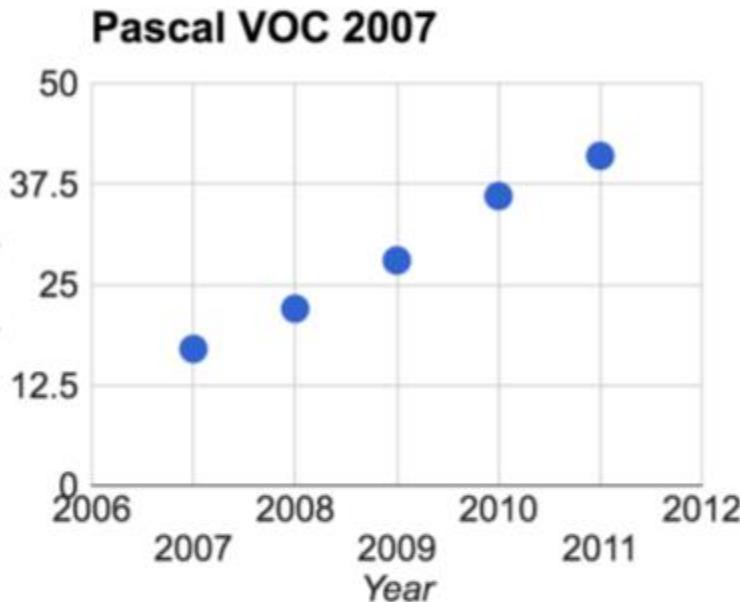
[Image](#) is CC BY-SA 3.0



[Image](#) is CC0 1.0 public domain

[This image](#) is licensed under
CC BY-SA 2.0; changes made

Mean Average Precision
(mAP)





IMAGENET

www.image-net.org

22K categories and **14M** images

- Animals
 - Bird
 - Fish
 - Mammal
 - Invertebrate
- Plants
 - Tree
 - Flower
 - Food
 - Materials
- Structures
 - Artifact
 - Tools
 - Appliances
 - Structures
- Person
- Scenes
 - Indoor
 - Geological Formations
- Sport Activities



IMAGENET Large Scale Visual Recognition Challenge

Steel drum

The Image Classification Challenge:

1,000 object classes

1,431,167 images



Output:
Scale
T-shirt
Steel drum
Drumstick
Mud turtle

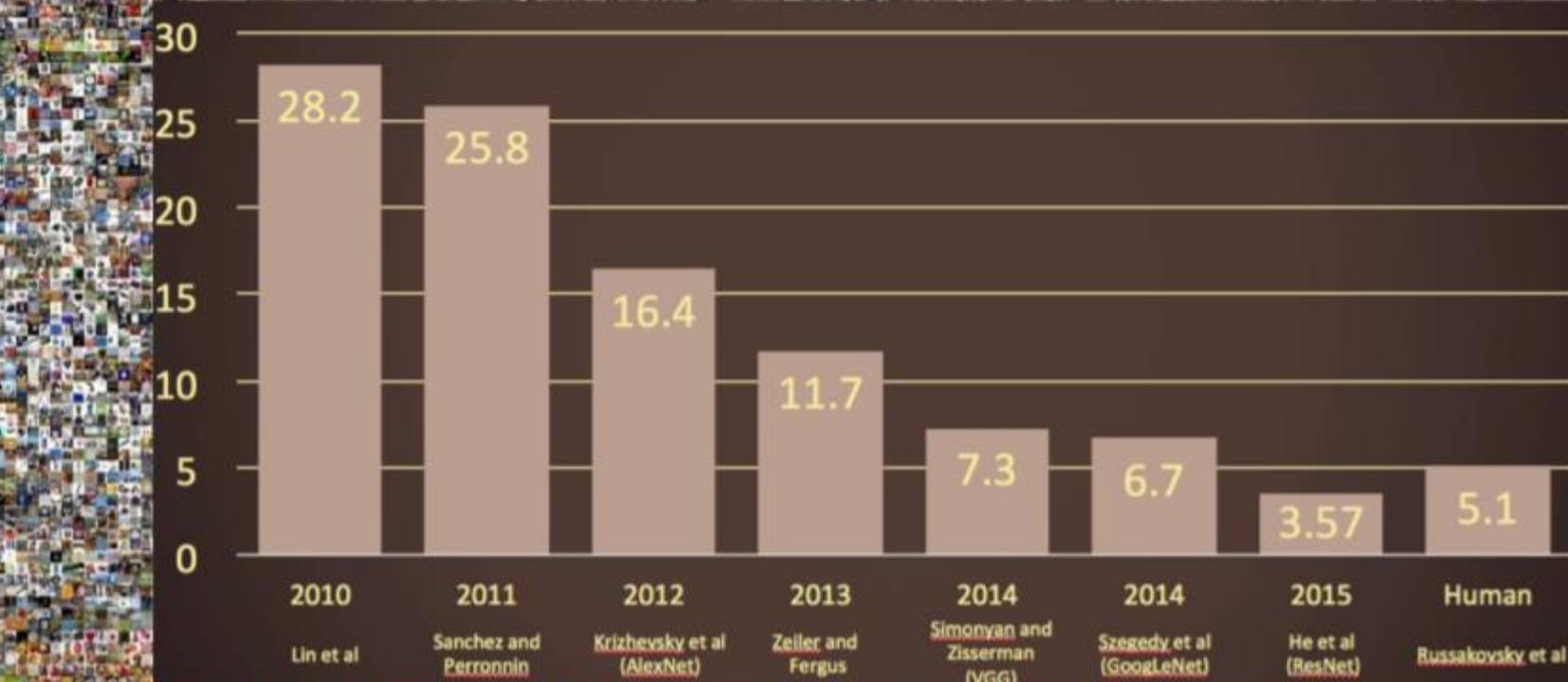


Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle



IMAGENET Large Scale Visual Recognition Challenge

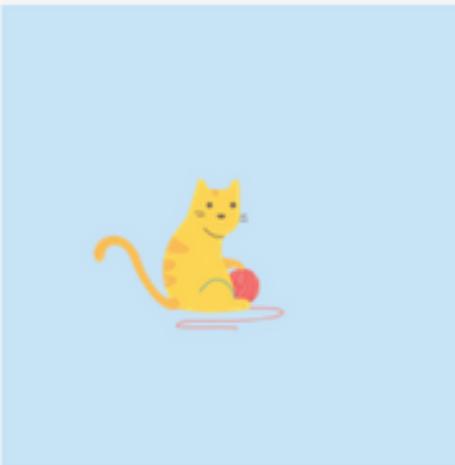
The Image Classification Challenge:
1,000 object classes
1,431,167 images



Vision Tasks

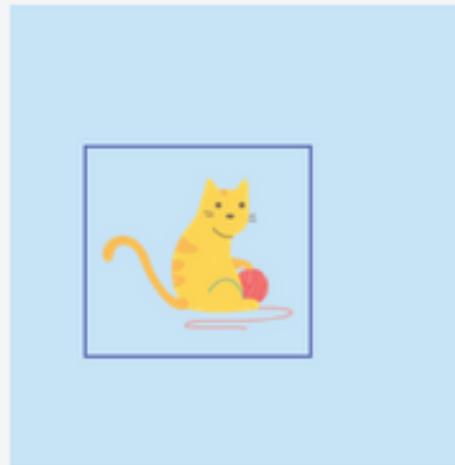
Computer Vision Problem Types

Classification



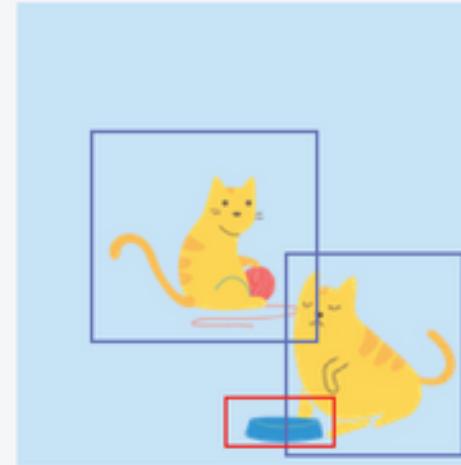
CAT

Classification
+ Localization



CAT

Object Detection



CAT, CAT, BOWL

Single Object

Semantic
Segmentation



CAT, CAT, BOWL

Multiple Objects

roboflow

제조 현장에서의 컴퓨터 비전 적용 사례

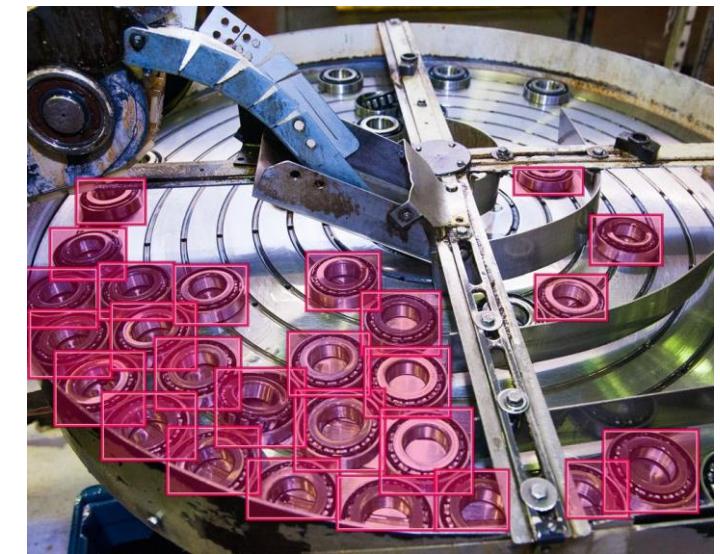
Predictive maintenance



Defect detection



Assembling products and components

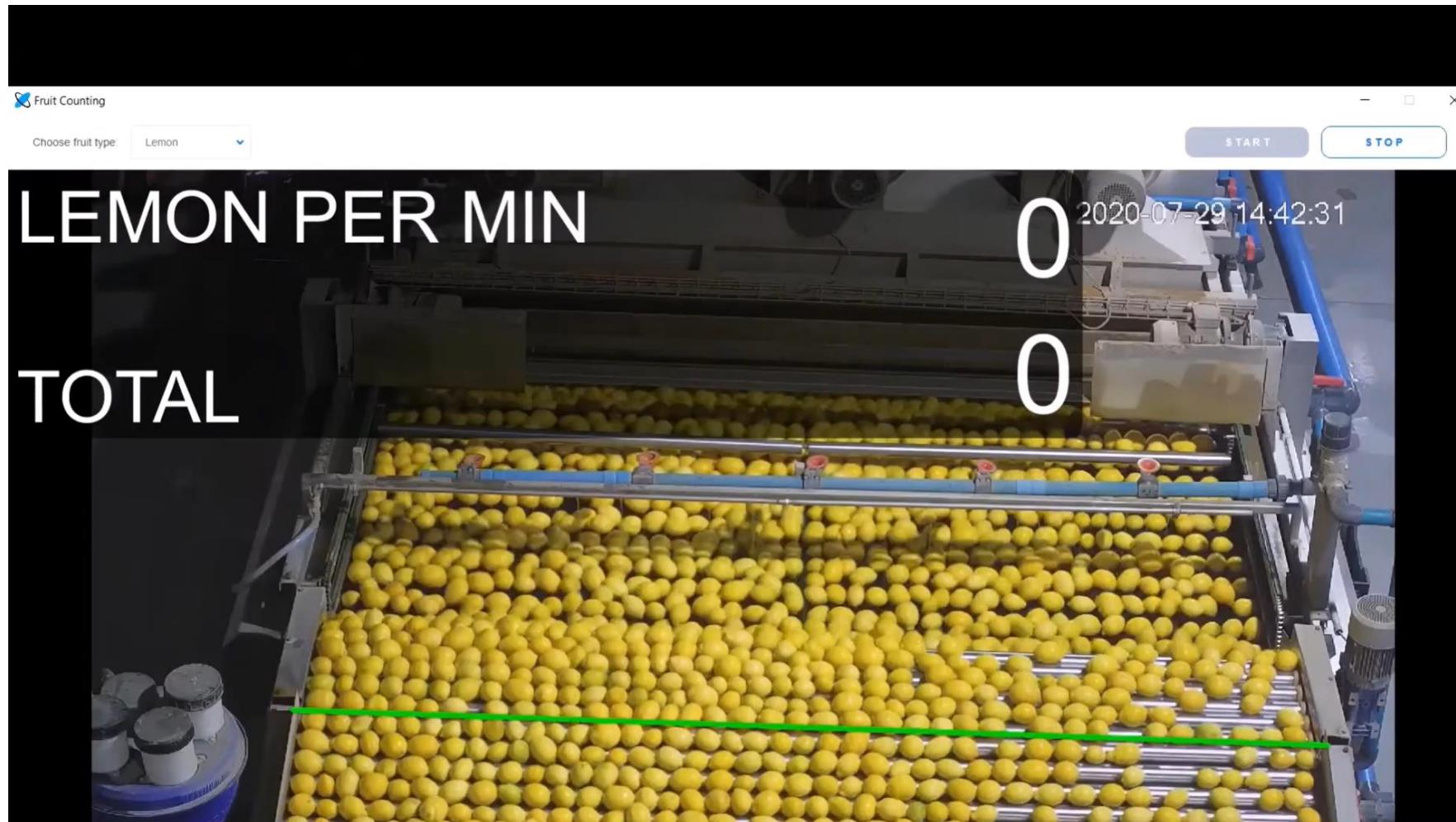


[출처] <https://blog.roboflow.com/computer-vision-use-cases-for-manufacturing/>

Object Counting

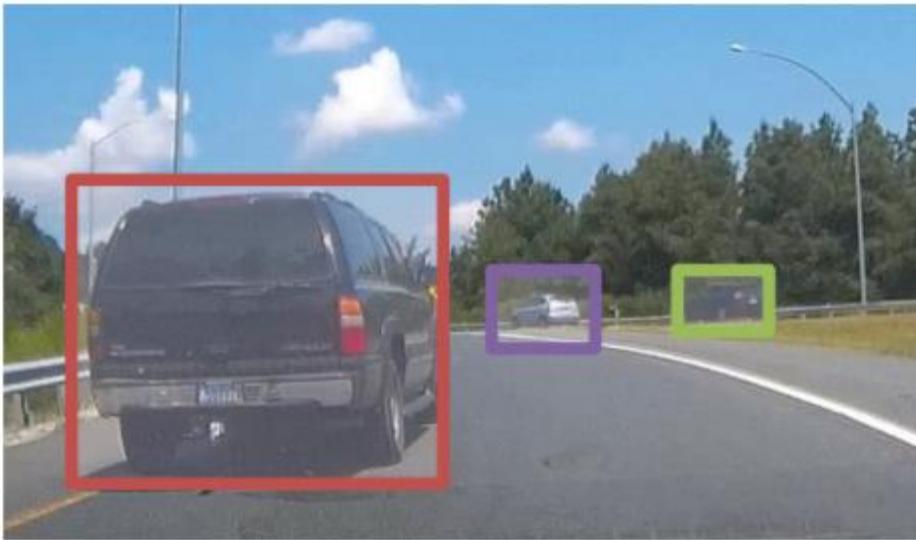


Object Counting

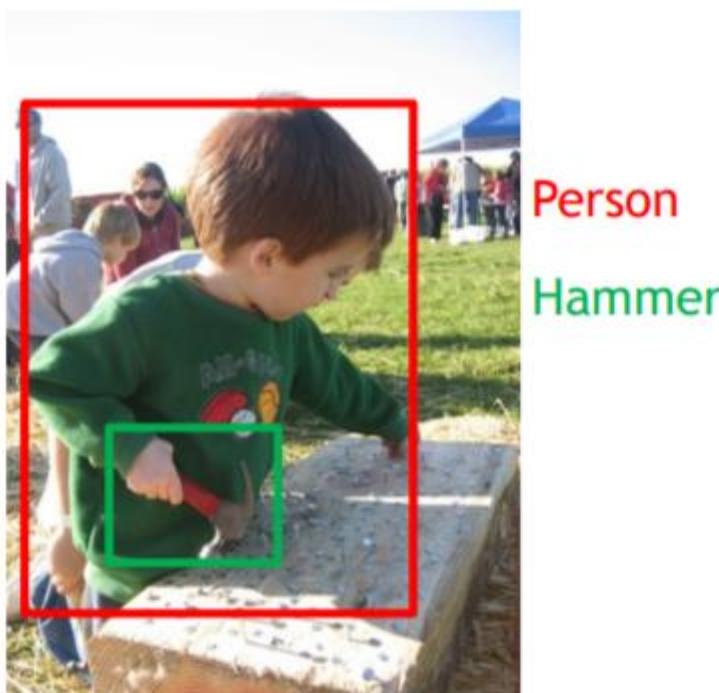


Convolutional Neural Networks (CNN)
were not invented overnight

Convolutional Neural Networks (CNN) have become an important tool for object recognition



[This image](#) is licensed under CC BY-NC-SA 2.0; changes made

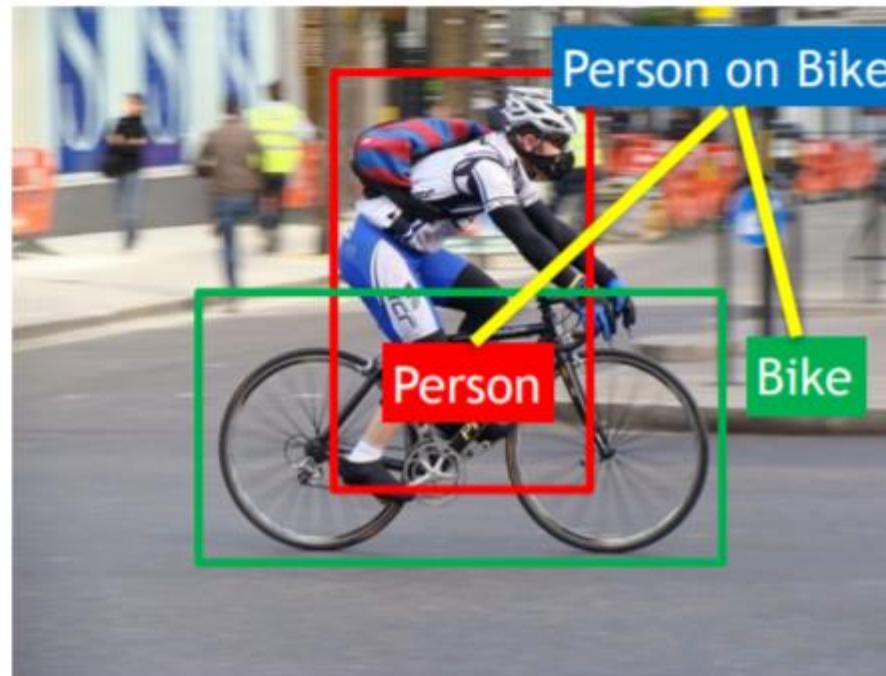


Person

Hammer

[This image](#) is licensed under CC BY-SA 2.0; changes made

- Object detection
- Action classification
- Image captioning
- ...

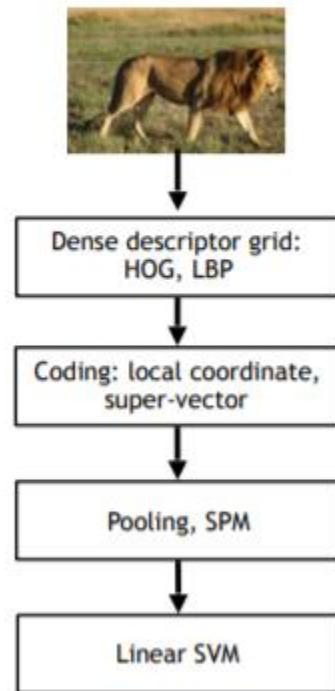


[This image](#) is licensed under CC BY-SA 3.0; changes made

IMAGENET Large Scale Visual Recognition Challenge

Year 2010

NEC-UIUC

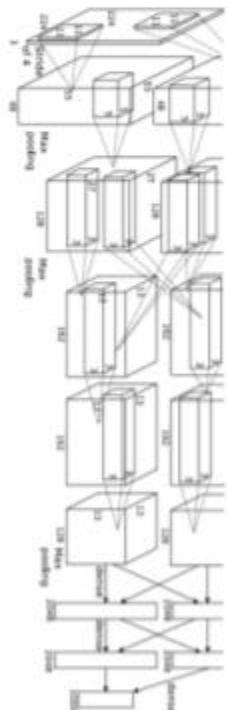


[Lin CVPR 2011]

Lion image by Swissfrog is licensed under CC BY 3.0

Year 2012

SuperVision



[Krizhevsky NIPS 2012]

Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

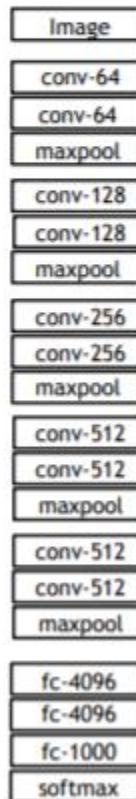
Year 2014

GoogLeNet

- Pooling
- Convolution
- Softmax
- Other



VGG



[Szegedy arxiv 2014]

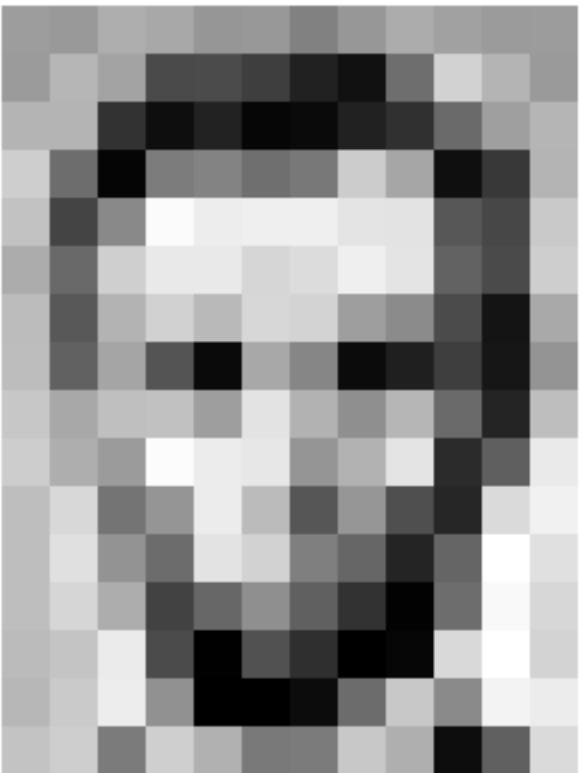
[Simonyan arxiv 2014]

Year 2015

MSRA

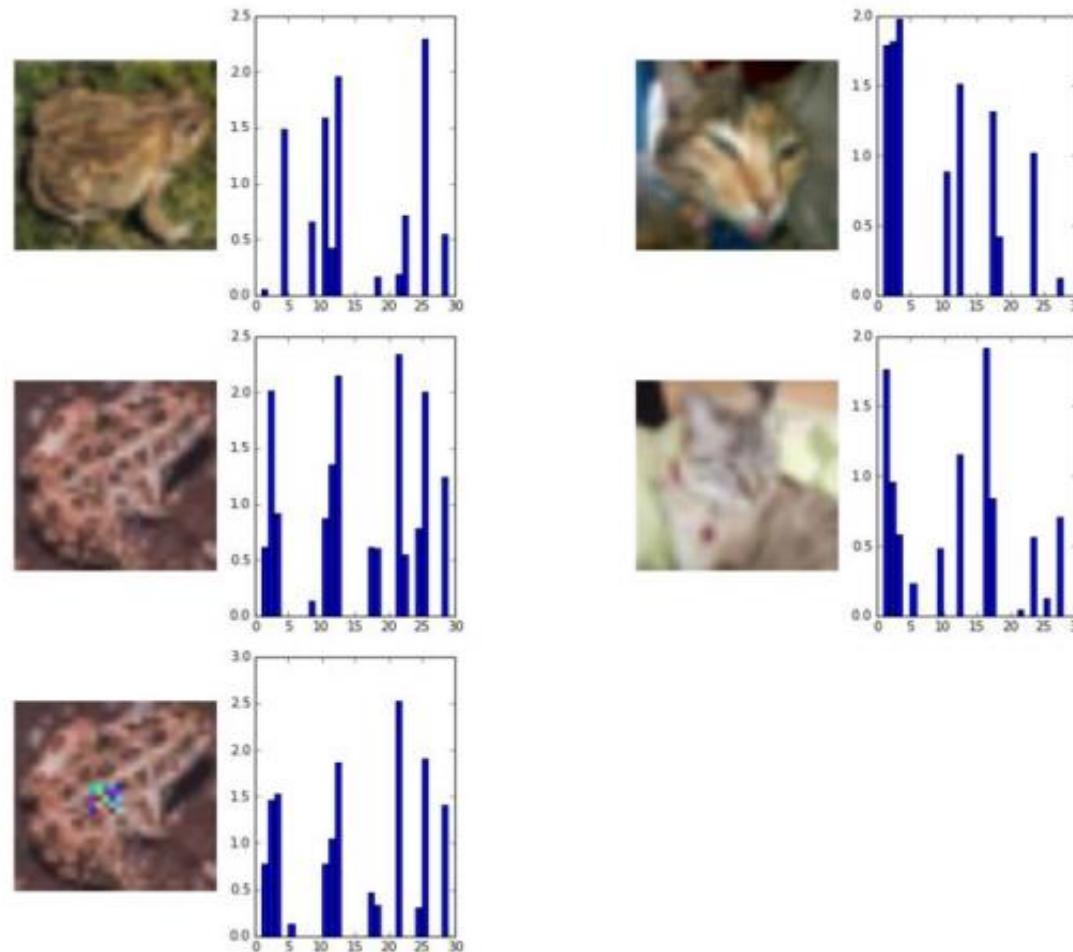


[He ICCV 2015]

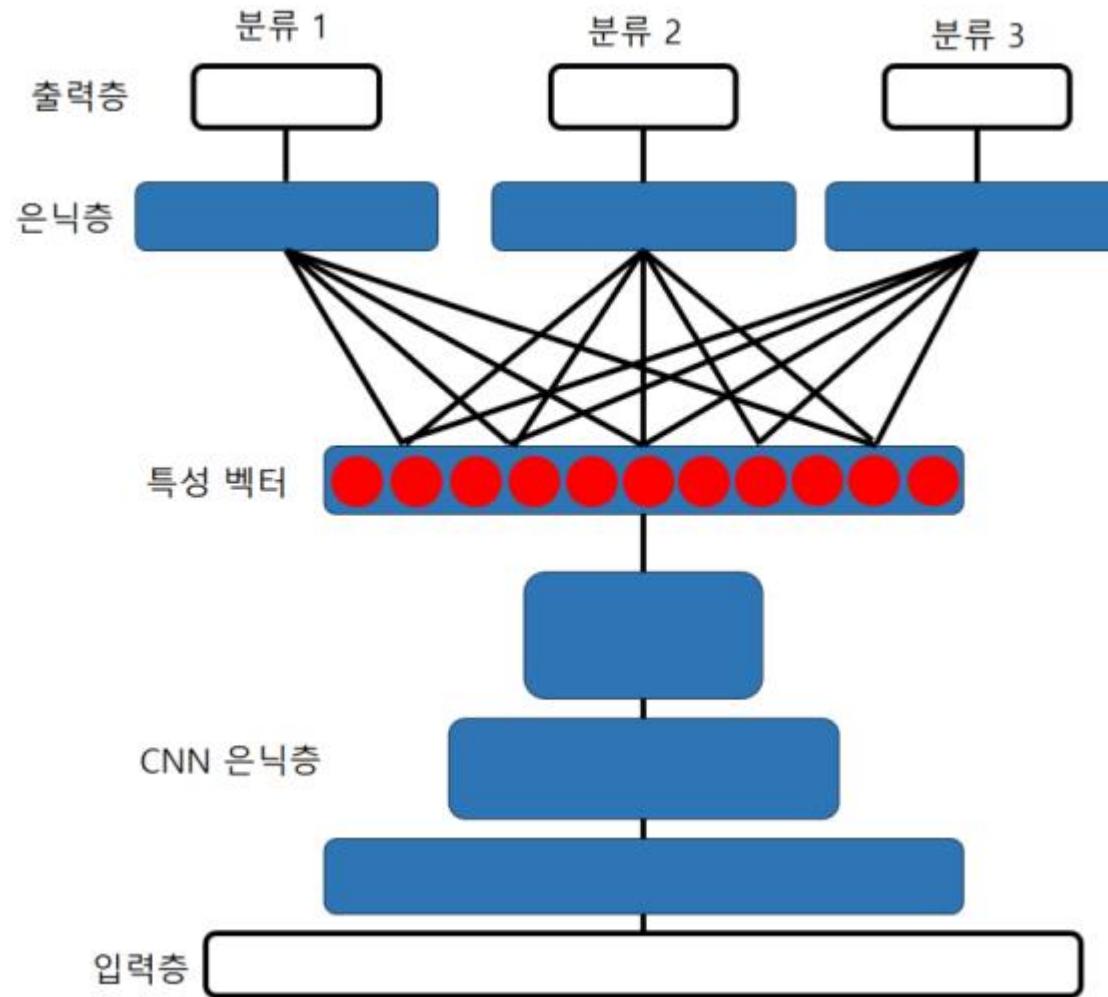


157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	105	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	105	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	105	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218



서로 다른 클래스의 물체들과 딥러닝 방식을 통해 추출한 특성 벡터. 이미지 내의 선, 모양, 무늬들을 감지해 이것을 값들의 집합으로 나타낸다. 딥러닝으로 추출된 벡터는 이미지의 범주나 구도 등에 따라 상이한 결과가 나오며, 여러에 관한 내성이 존재한다.

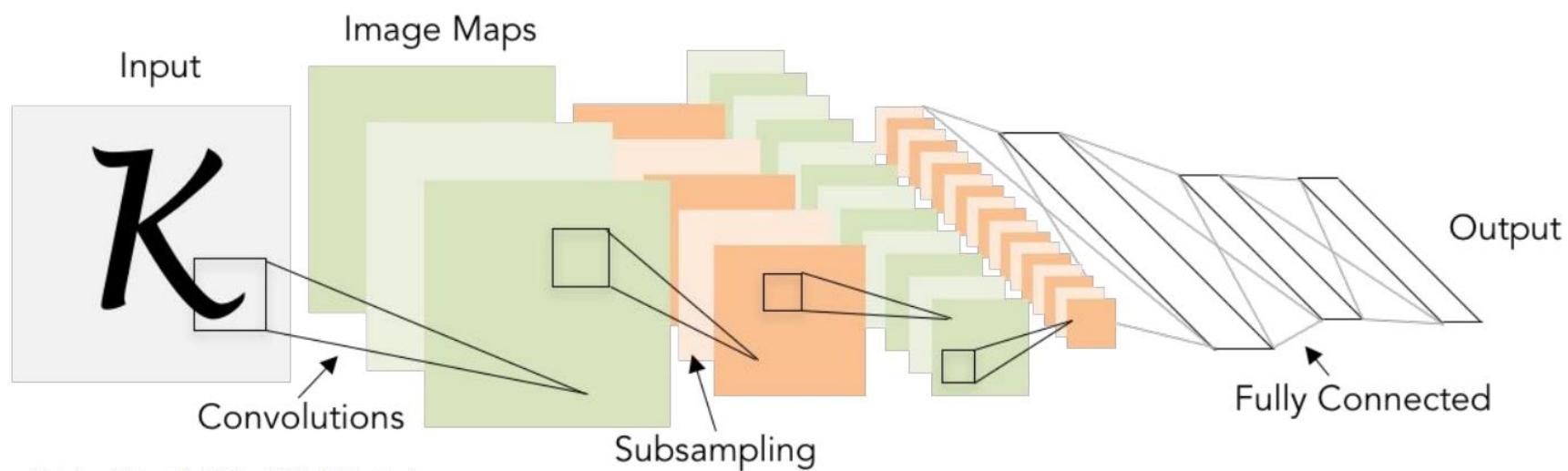


(예시) 다목적 이미지 분류를 위한 전체적인 구조도로서 학습된 딥러닝 구조를 통해 특성 벡터를 추출하고, 특성 벡터를 인공 신경망의 입력으로 삼아 다양한 분류기를 구성한다.

CNN (Convolution neural networks)

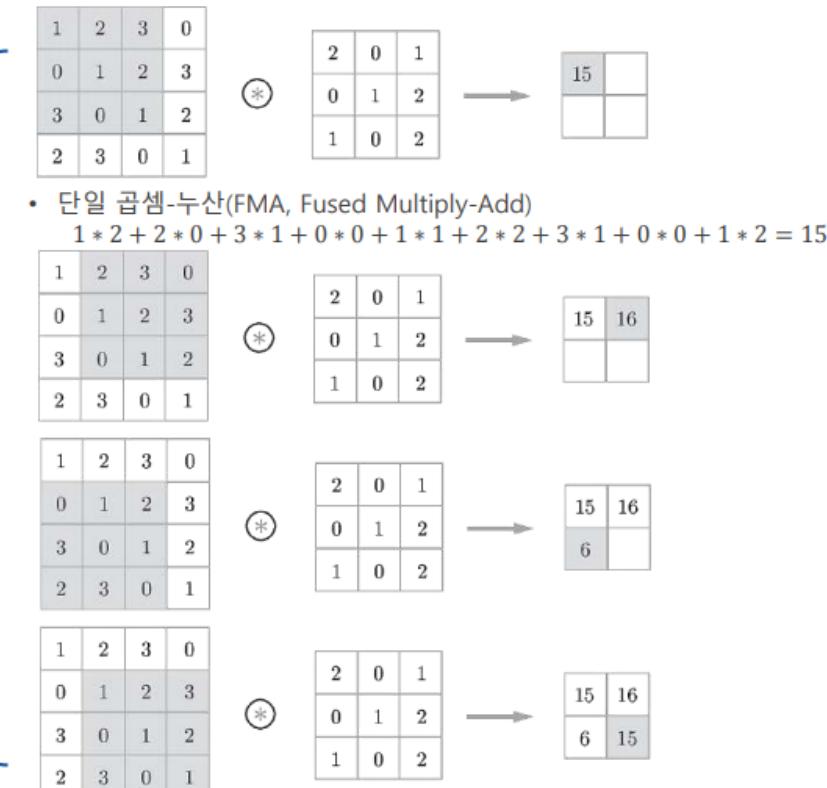
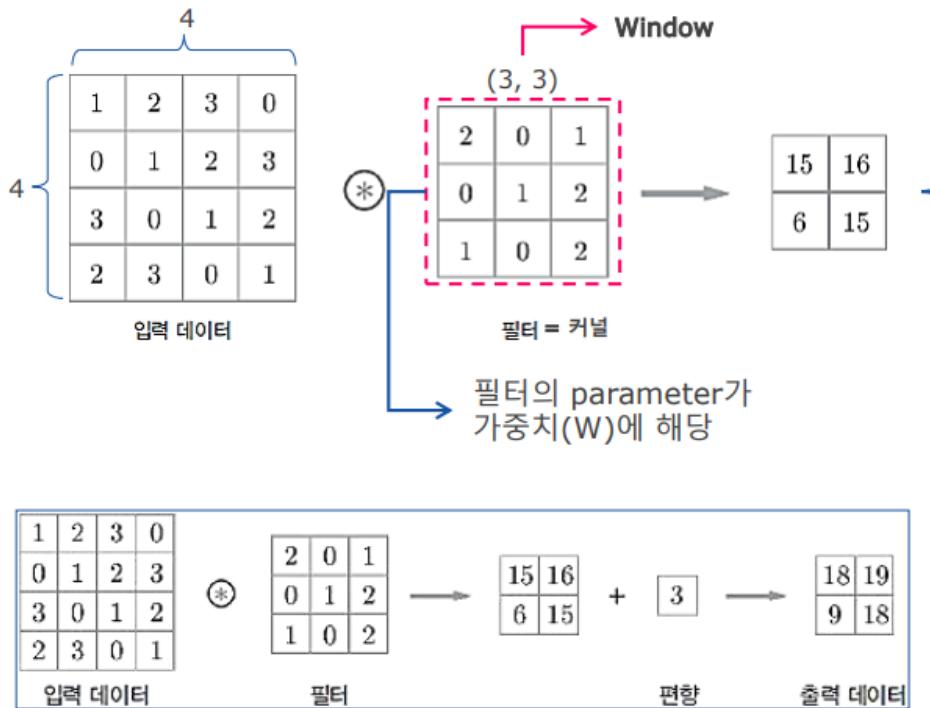
Gradient-based learning applied to document recognition

[LeCun, Bottou, Bengio, Haffner 1998]

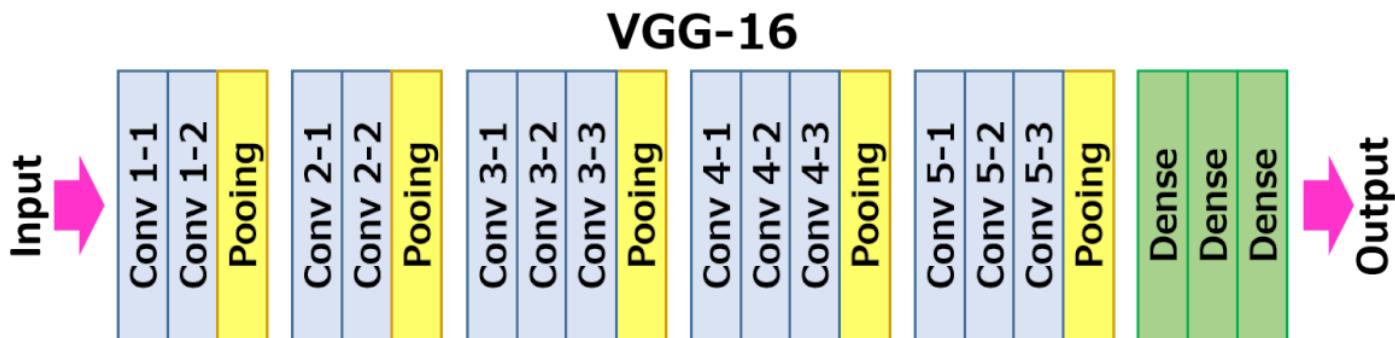
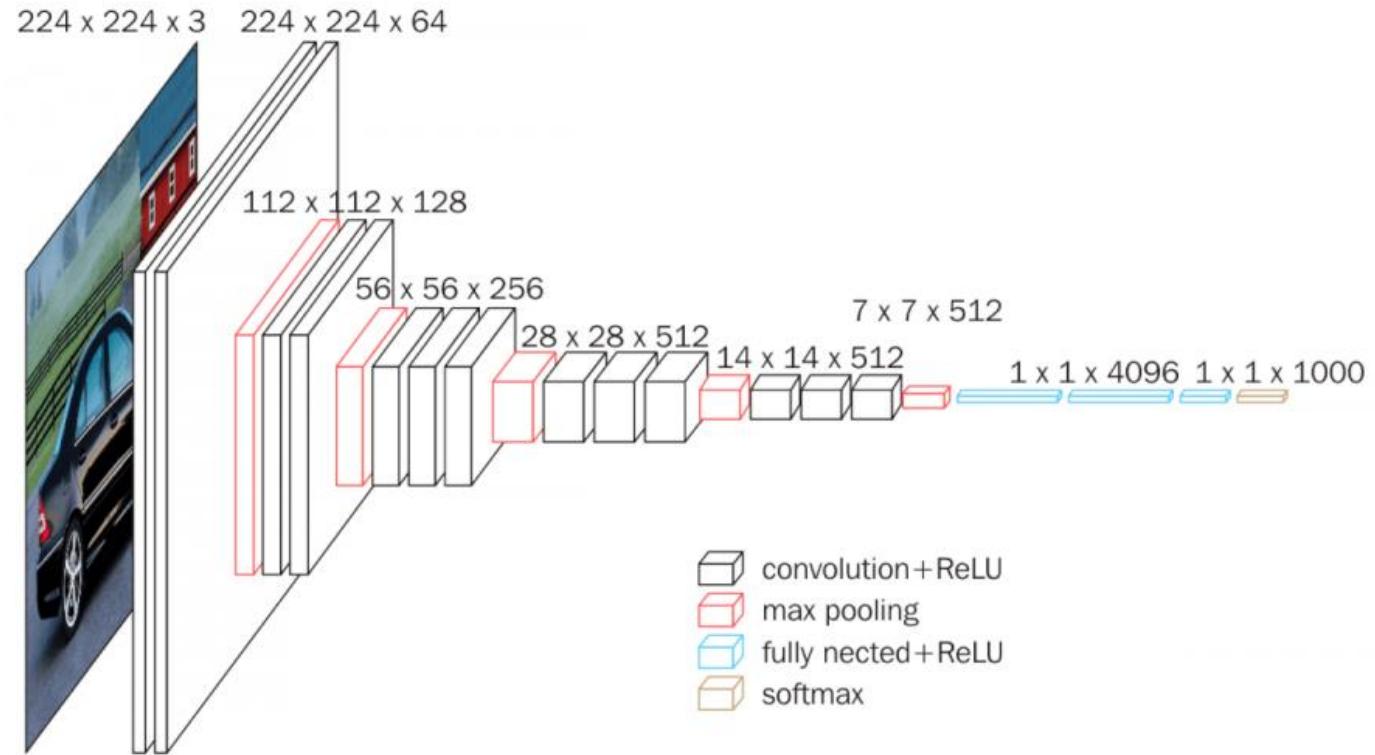


합성곱 계층 - 연산

- ✓ 합성곱 계층에서 연산 수행 → 필터(커널) 연산
- ✓ 데이터와 필터의 형상을 (높이height, 너비width)로 표기
- ✓ 윈도우window 를 일정 간격(Stride)으로 이동하며 계산

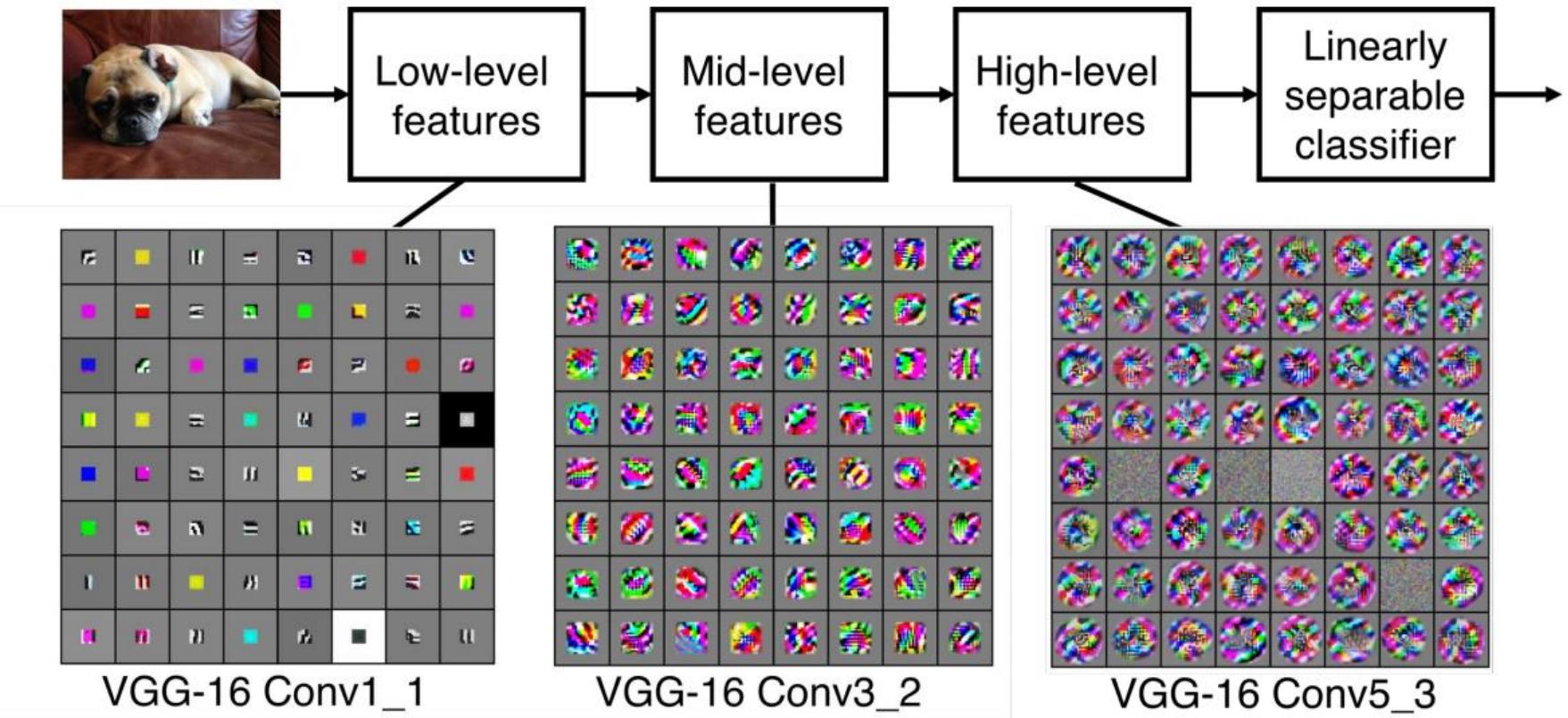


VGG-16 Pre-Trained Model 구성



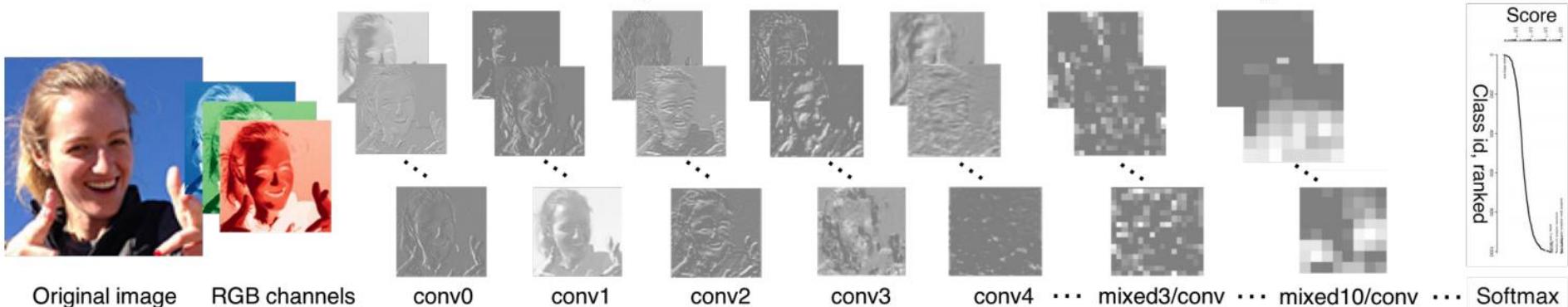
Layer 별 Features

Preview

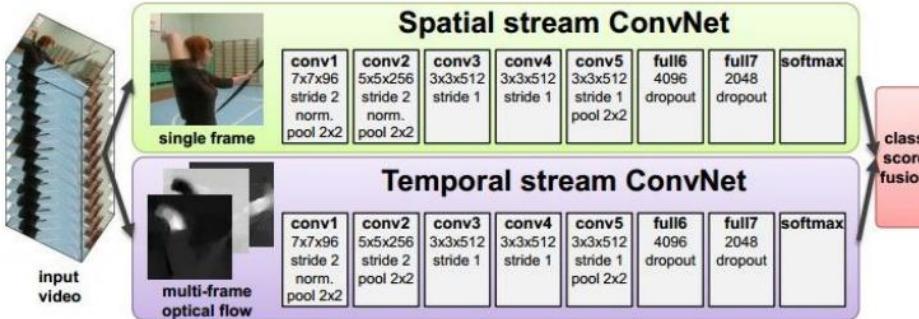


Activation functions

Fast-forward to today: ConvNets are everywhere

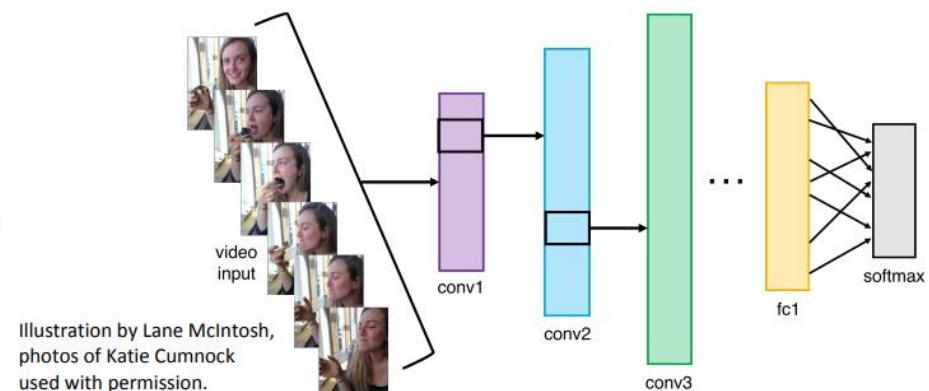


Activations of [inception-v3 architecture](#) [Szegedy et al. 2015] to image of Emma McIntosh, used with permission. Figure and architecture not from Taigman et al. 2014.

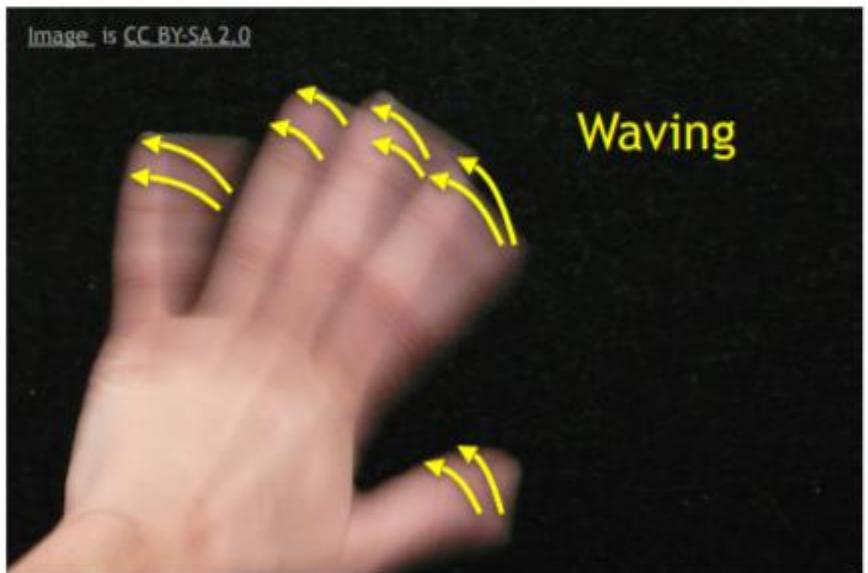
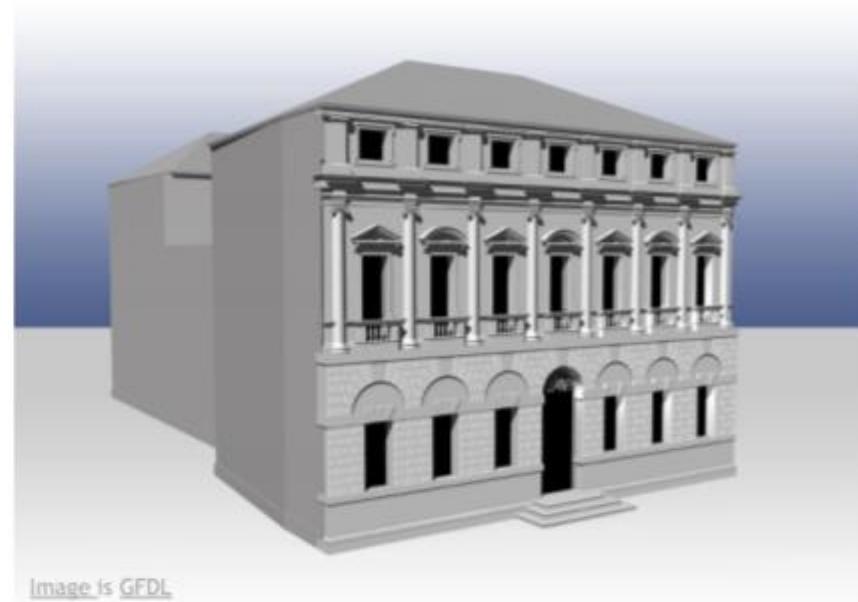
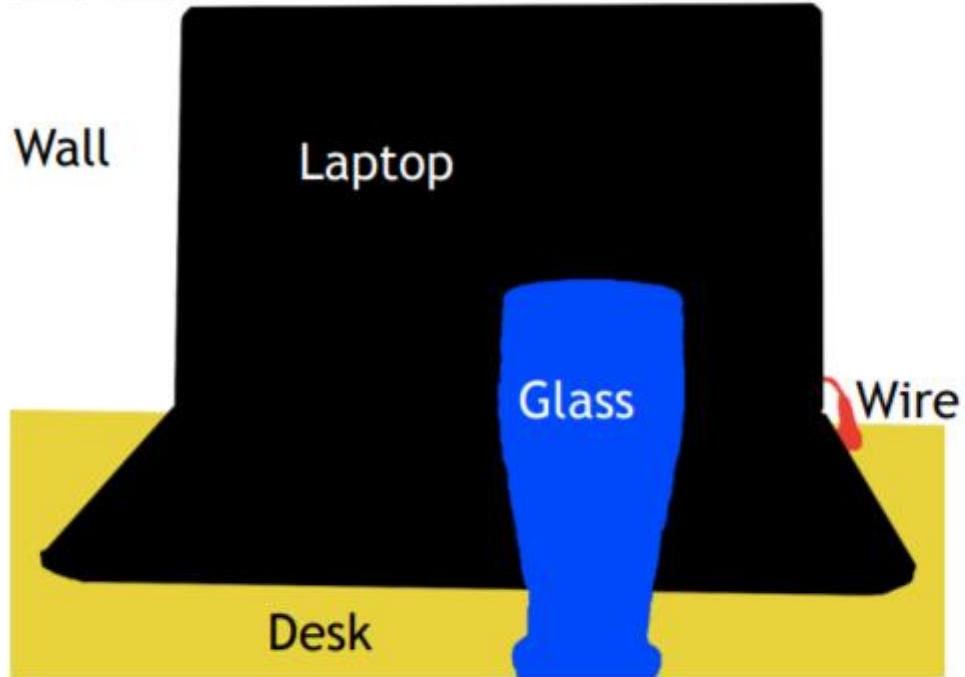


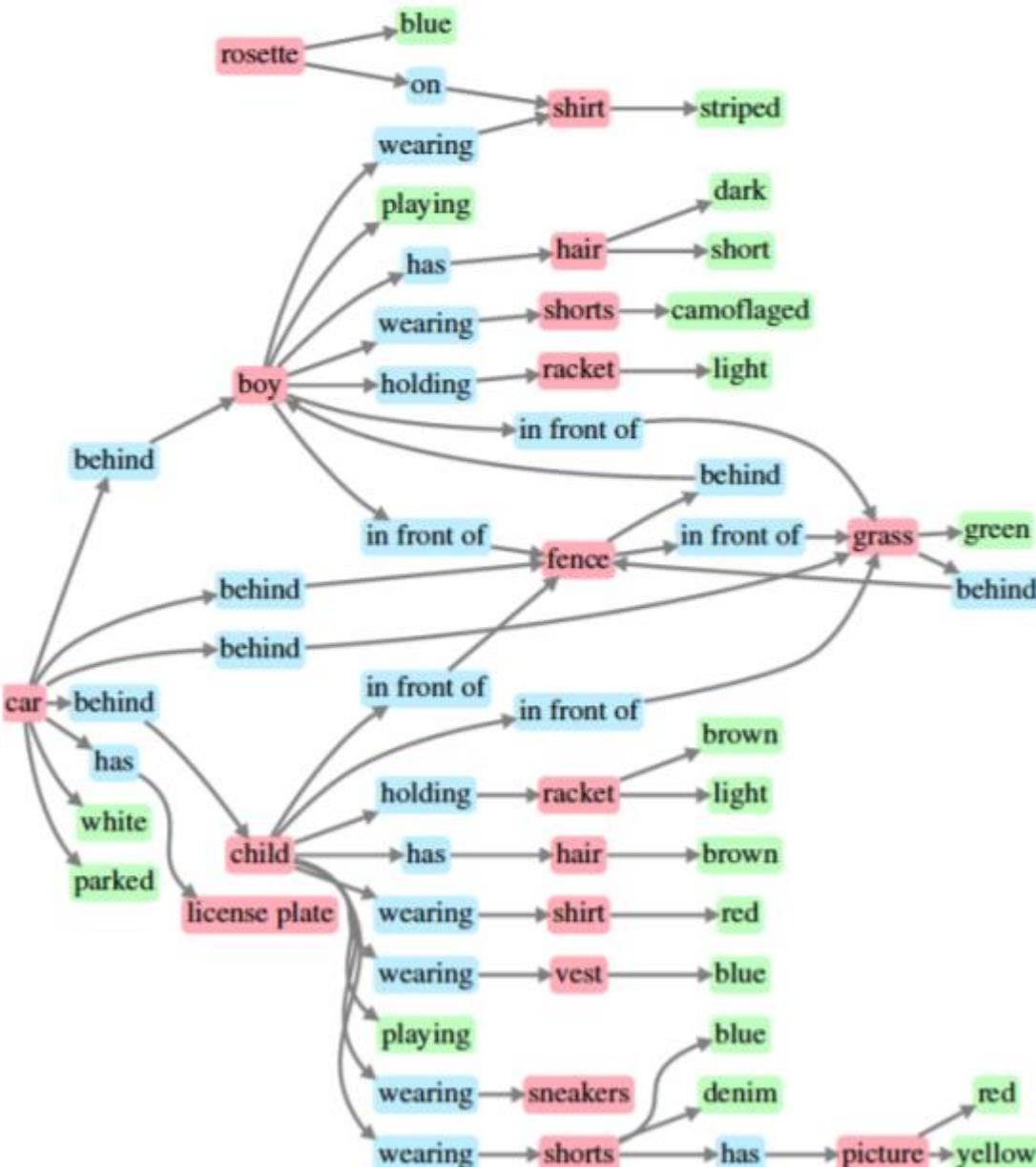
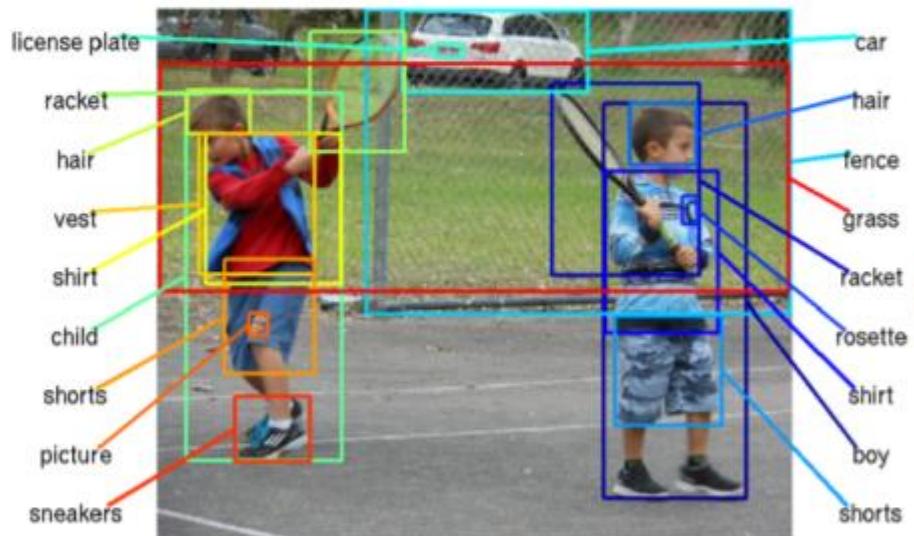
[Simonyan et al. 2014]

Figures copyright Simonyan et al., 2014.
Reproduced with permission.



The quest for visual intelligence
goes far beyond object recognition...





Johnson et al., "Image Retrieval using Scene Graphs", CVPR 2015

Figures copyright IEEE, 2015. Reproduced for educational purposes

PT = 500ms



Some kind of game or fight. Two groups of two men? The man on the left is throwing something. Outdoors seemed like because i have an impression of grass and maybe lines on the grass? That would be why I think perhaps a game, rough game though, more like rugby than football because they pairs weren't in pads and helmets, though I did get the impression of similar clothing. maybe some trees? in the background. (Subject: SM)



Outside border images, clockwise, starting from top left:

[Image by Pop Culture Geek](#) is licensed under CC BY 2.0; changes made

[Image by the US Government](#) is in the public domain

[Image by the US Government](#) is in the public domain

[Image by Glogger](#) is licensed under CC BY-SA 3.0; changes made

[Image by Sylenius](#) is licensed under CC BY 3.0; changes made

[Image by US Government](#) is in the public domain

Inside four images, clockwise, starting from top left:

[Image](#) is CCO 1.0 public domain

[Image by Tucania](#) is licensed under CC BY-SA 3.0; changes made

[Image by Intuitive Surgical, Inc.](#) is licensed under CC BY-SA 3.0; changes made

[Image by Oyundari Zorigtbaatar](#) is licensed under CC BY-SA 4.0



정 준 수 / Ph.D (heinem@naver.com)

- 前) 삼성전자 연구원
- 前) 삼성의료원 (삼성생명과학연구소)
- 前) 삼성SDS (정보기술연구소)
- 現) (사)한국인공지능협회, AI, 머신러닝 강의
- 現) 한국소프트웨어산업협회, AI, 머신러닝 강의
- 現) 서울디지털재단, AI 자문위원
- 現) 한성대학교 교수(겸)
- 전문분야: Computer Vision, 머신러닝(ML), RPA
- <https://github.com/JSeong-me/>