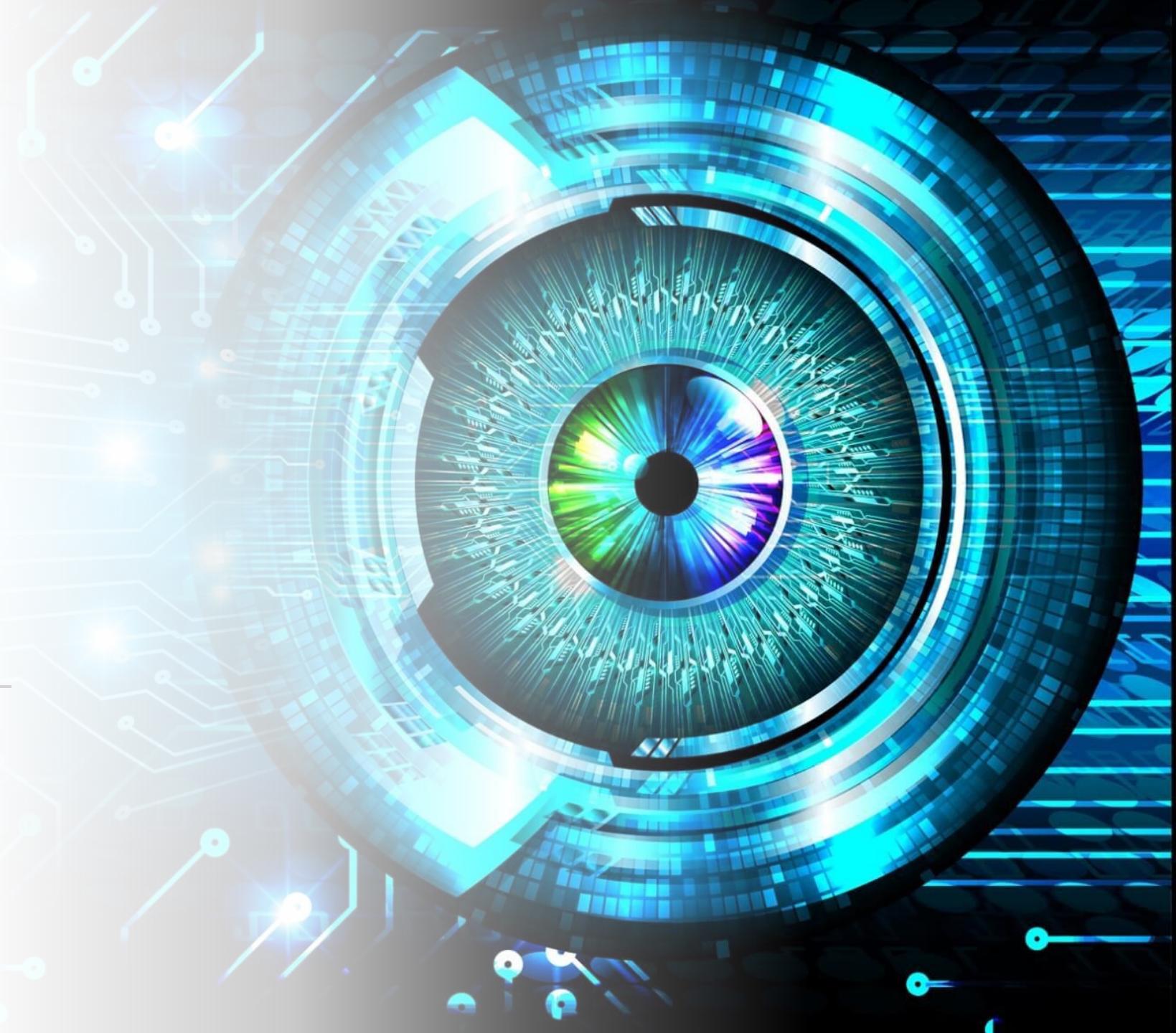


Computer Vision

Paweł Pozorski

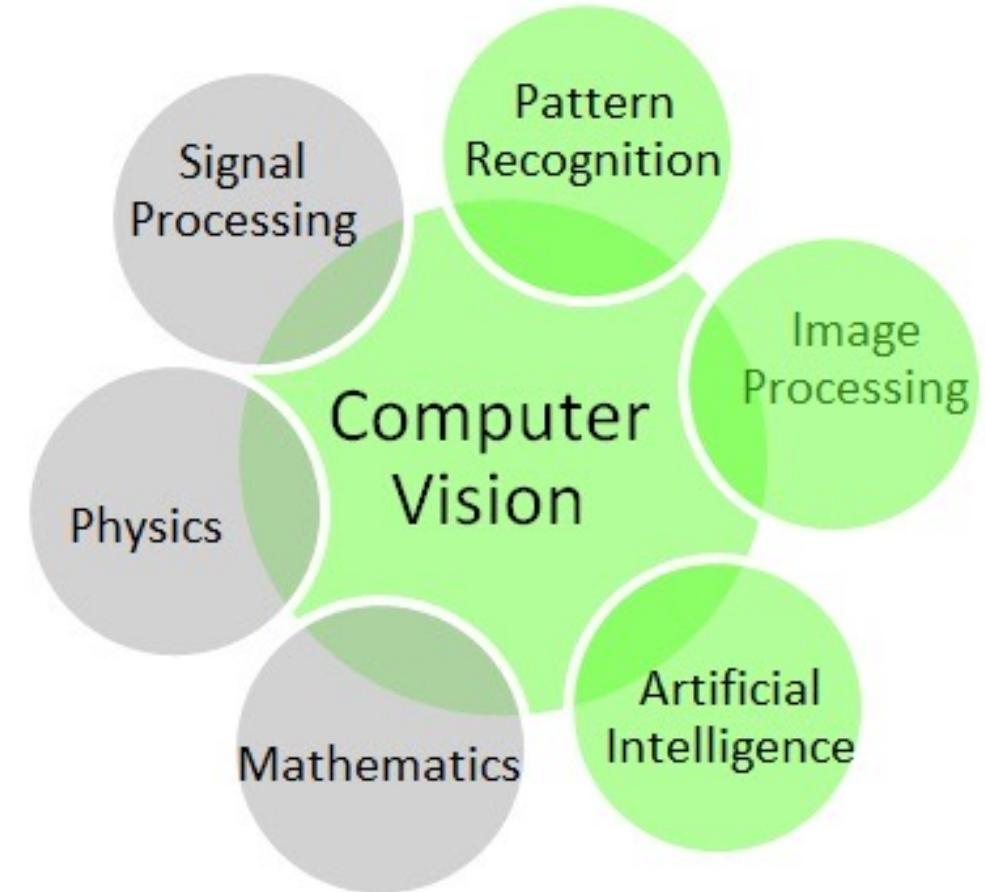


What is Computer Vision?



Orange
Tap
Fruit
Water Outlet
White Table

Computer vision history traces the evolution of technology aimed at enabling machines to interpret and understand visual information. Beginning with early image processing in the 1960s, the field progressed through rule-based systems in the 1970s, reaching a significant milestone in the 1980s with the development of edge detection and feature extraction techniques. The late 1990s and early 2000s witnessed advancements in machine learning, leading to breakthroughs like object recognition, and more recently, deep learning has become a dominant paradigm, driving remarkable achievements in image classification, object detection, and facial recognition.





You

how many images have been approximately taken by humans so far?

ChatGPT

It's challenging to provide an exact number of images taken by humans globally because the rate at which images are captured is continually increasing, and accurate tracking is difficult. However, as of my last knowledge update in January 2022, it was estimated that billions of photos are taken every day worldwide.

Assuming an average image size of 3 megabytes (MB), which accounts for a mix of high-resolution photos and compressed images typically shared on social media, we can use the following calculation:

$$\text{Total size} \approx \text{Number of images} \times \text{Average size per image}$$

Let's assume 1 trillion images have been taken globally:

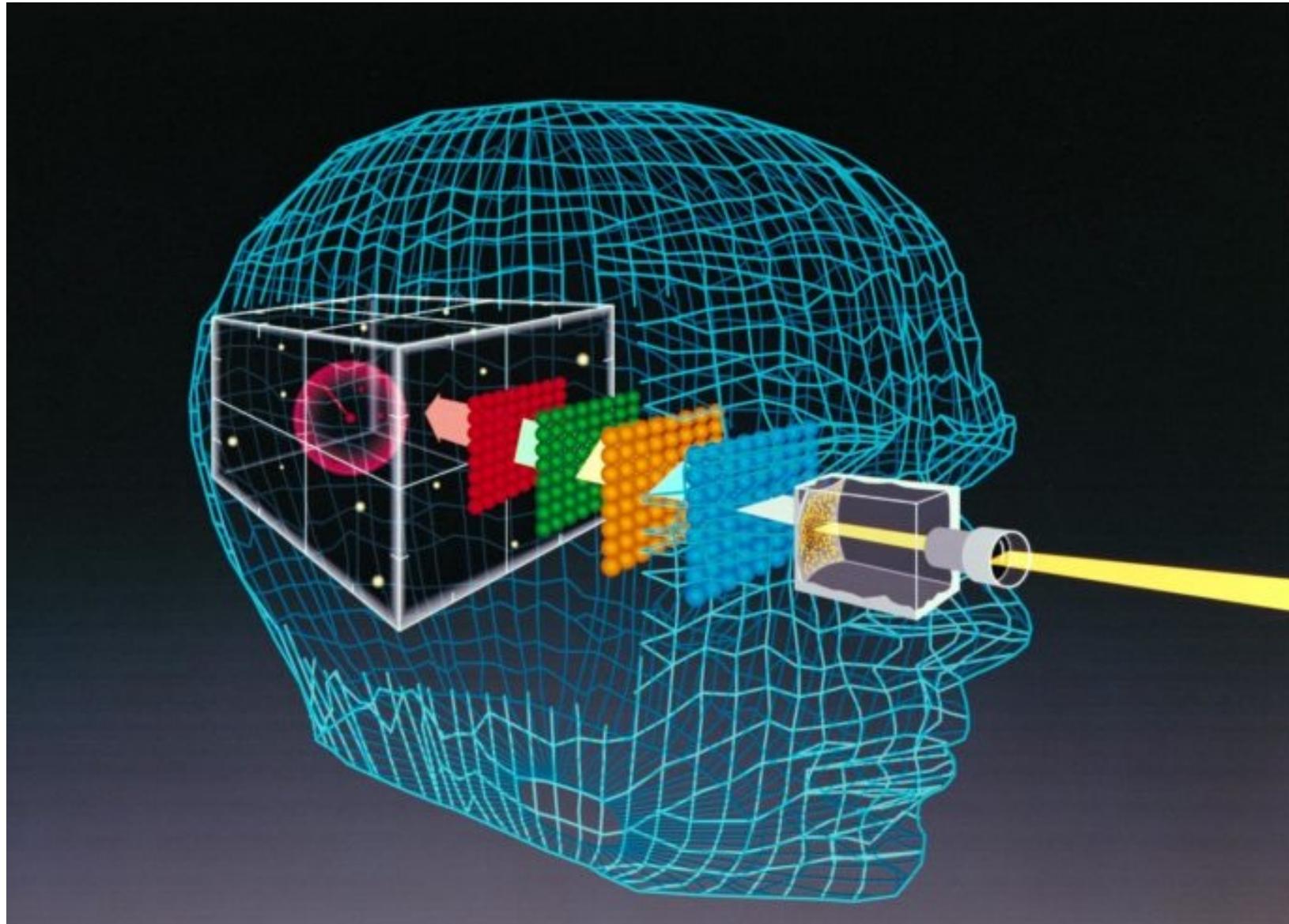
$$\text{Total size} \approx 1 \text{ trillion images} \times 3 \text{ MB per image}$$

$$\text{Total size} \approx 3 \text{ petabytes (PB)}$$



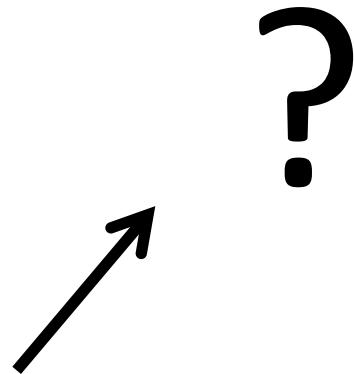

200.0	199.0	198.0	197.0	197.0	197.0	196.0	195.0	195.0	194.0	194.0	193.0	193.0	192.0	192.0	191.0	191.0	191.0	190.0	189.0	190.0					
202.0	201.0	200.0	200.0	200.0	200.0	201.0	200.0	198.0	197.0	198.0	198.0	196.0	197.0	196.0	196.0	195.0	194.0	195.0	194.0	193.0	193.0				
206.0	206.0	205.0	206.0	205.0	206.0	204.0	204.0	203.0	202.0	201.0	200.0	200.0	199.0	198.0	198.0	197.0	197.0	196.0	196.0	196.0	196.0				
207.0	208.0	207.0	208.0	206.0	206.0	210.0	210.0	209.0	207.0	205.0	203.0	204.0	203.0	202.0	202.0	202.0	201.0	201.0	200.0	201.0	199.0	199.0			
211.0	210.0	210.0	210.0	211.0	210.0	209.0	209.0	208.0	211.0	210.0	208.0	208.0	208.0	208.0	208.0	208.0	208.0	208.0	208.0	208.0	208.0				
214.0	214.0	213.0	213.0	212.0	212.0	211.0	211.0	208.0	214.0	215.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0				
216.0	216.0	215.0	214.0	215.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0	214.0				
220.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0	219.0				
222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0	222.0				
219.0	193.0	253.0	11.0	25.0	33.0	21.0	35.0	33.0	234.0	44.0	98.0	35.0	230.0	79.0	123.0	48.0	137.0	228.0	236.0	214.0	230.0	214.0			
164.0	158.0	32.0	9.0	17.0	23.0	27.0	26.0	30.0	229.0	47.0	47.0	146.0	148.0	7.0	36.0	43.0	146.0	167.0	185.0	139.0	121.0	139.0			
140.0	178.0	11.0	23.0	22.0	27.0	19.0	21.0	26.0	53.0	47.0	61.0	42.0	92.0	7.0	9.0	97.0	100.0	118.0	115.0	142.0	89.0	199.0	218.0		
148.0	12.0	10.0	20.0	26.0	28.0	19.0	26.0	28.0	41.0	39.0	5.0	6.0	6.0	10.0	57.0	112.0	248.0	147.0	144.0	142.0	136.0	135.0	134.0	138.0	134.0
39.0	25.0	6.0	22.0	28.0	30.0	33.0	31.0	28.0	35.0	6.0	28.0	5.0	5.0	41.0	40.0	169.0	166.0	159.0	164.0	163.0	162.0	150.0	151.0	160.0	158.0
26.0	26.0	5.0	25.0	24.0	25.0	30.0	25.0	28.0	27.0	26.0	19.0	19.0	6.0	8.0	179.0	159.0	151.0	150.0	154.0	163.0	163.0	159.0	160.0	159.0	
24.0	24.0	5.0	10.0	19.0	25.0	27.0	28.0	15.0	23.0	32.0	12.0	20.0	6.0	164.0	162.0	168.0	165.0	160.0	167.0	152.0	159.0	149.0	164.0	151.0	164.0
28.0	26.0	4.0	3.0	6.0	7.0	4.0	5.0	29.0	26.0	177.0	155.0	151.0	159.0	28.0	59.0	155.0	166.0	151.0	153.0	158.0	148.0	153.0	154.0	156.0	151.0
26.0	29.0	5.0	4.0	4.0	5.0	5.0	23.0	22.0	143.0	148.0	172.0	153.0	154.0	87.0	121.0	163.0	135.0	163.0	153.0	166.0	149.0	165.0	166.0	144.0	138.0
27.0	30.0	28.0	25.0	28.0	26.0	28.0	20.0	154.0	179.0	160.0	138.0	136.0	162.0	164.0	147.0	152.0	149.0	156.0	146.0	154.0	160.0	146.0	176.0	162.0	140.0
24.0	29.0	28.0	27.0	19.0	24.0	7.0	33.0	156.0	161.0	159.0	162.0	154.0	164.0	149.0	131.0	144.0	172.0	149.0	170.0	157.0	171.0	141.0	149.0	155.0	167.0
28.0	26.0	27.0	28.0	26.0	22.0	16.0	35.0	168.0	167.0	148.0	178.0	158.0	95.0	163.0	155.0	133.0	122.0	148.0	149.0	187.0	176.0	167.0	125.0	130.0	185.0
24.0	28.0	28.0	29.0	28.0	32.0	30.0	33.0	166.0	169.0	137.0	160.0	166.0	215.0	150.0	153.0	129.0	145.0	131.0	143.0	140.0	138.0	166.0	138.0	154.0	143.0
24.0	25.0	27.0	28.0	28.0	10.0	32.0	33.0	146.0	153.0	146.0	162.0	164.0	57.0	159.0	149.0	137.0	149.0	152.0	153.0	160.0	139.0	168.0	104.0	148.0	162.0
25.0	25.0	24.0	26.0	31.0	131.0	19.0	30.0	120.0	127.0	133.0	137.0	149.0	49.0	127.0	133.0	143.0	148.0	163.0	203.0	162.0	167.0	164.0	140.0	169.0	143.0
21.0	21.0	26.0	28.0	28.0	121.0	131.0	33.0	185.0	160.0	163.0	134.0	103.0	168.0	40.0	54.0	142.0	136.0	138.0	143.0	133.0	105.0	132.0	157.0	180.0	167.0
24.0	24.0	25.0	30.0	31.0	116.0	126.0	6.0	150.0	131.0	113.0	143.0	122.0	160.0	124.0	217.0	168.0	143.0	136.0	159.0	244.0	132.0	120.0	130.0	148.0	153.0

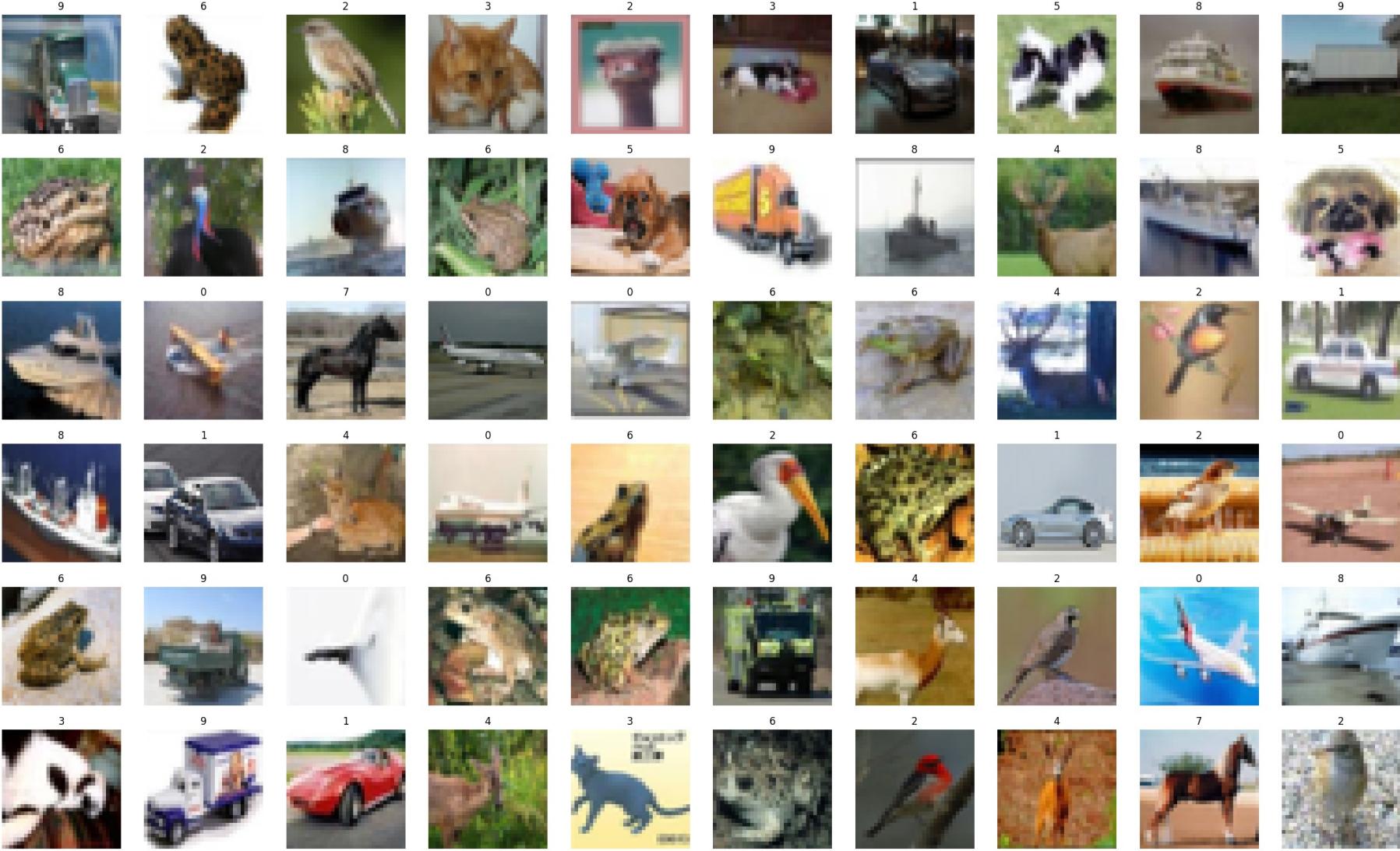
Resolution	32-bit RGB	8-bit BW
144p	147 KB	36 KB
240p	410 KB	77 KB
360p	922 KB	173 KB
480p	1,745 KB	327 KB
720p HD	3,686 KB	691 KB
1080p Full HD	8,294 KB	1,554 KB
1440p Quad HD	14,746 KB	2,765 KB
4K UHD	33,178 KB	6,221 KB
8K UHD	132,710 KB (or 130 MB)	24,883 KB (or 24 MB)

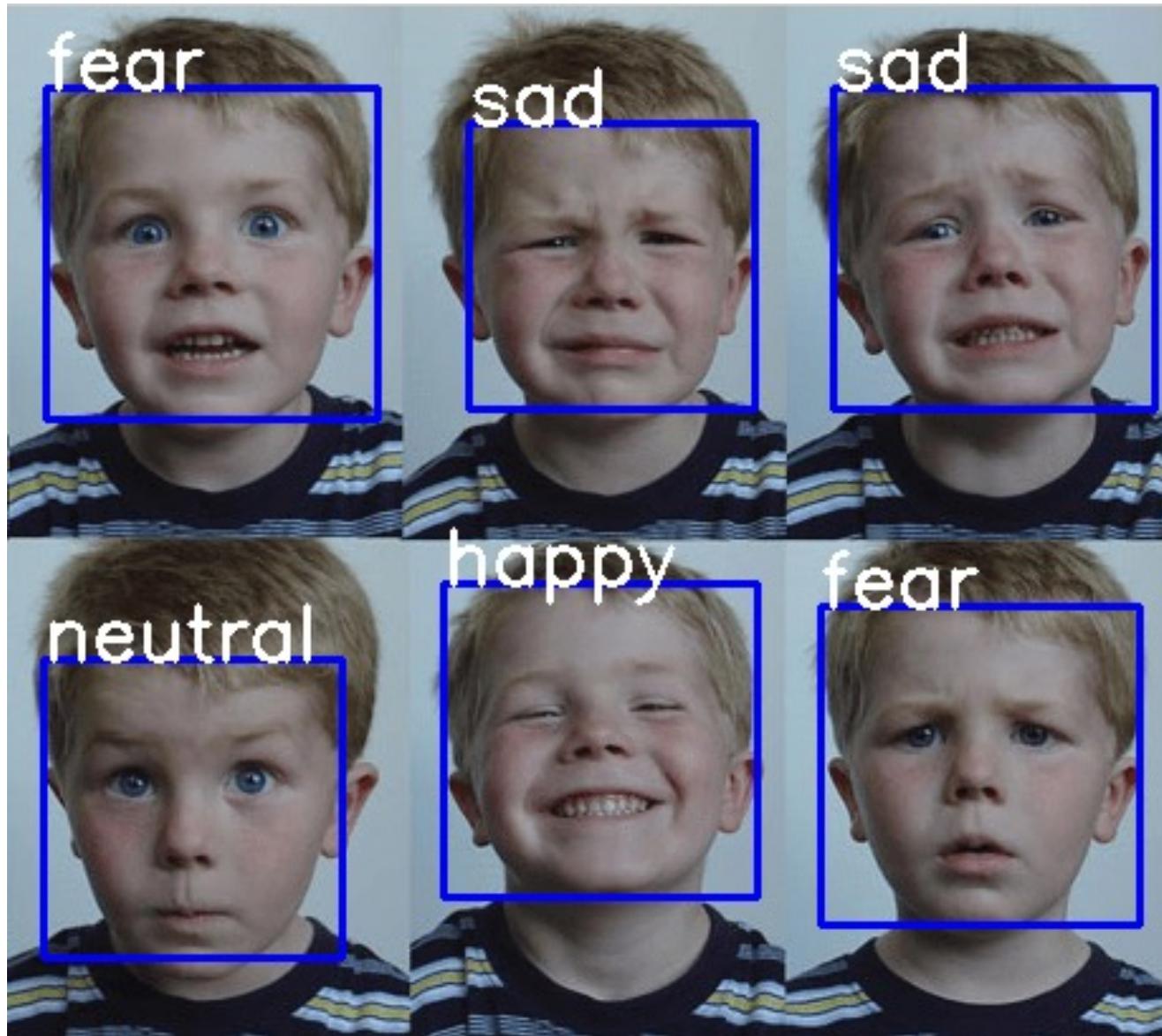


Resolution	32-bit RGB	8-bit BW
144p	~ 3.5 MB	~ 0.8 MB
240p	~ 9 MB	~ 2.2 MB
360p	~ 20 MB	~ 4.5 MB
480p	~ 35 MB	~ 8 MB
720p HD	~ 80 MB	~ 18 MB
1080p Full HD	~ 180 MB	~ 40 MB
1440p Quad HD	~ 350 MB	~ 80 MB
4K UHD	~ 750 MB	~ 170 MB
8K UHD	~ 3 GB (or 300 MB)	~ 700 MB (or 0.7 GB)

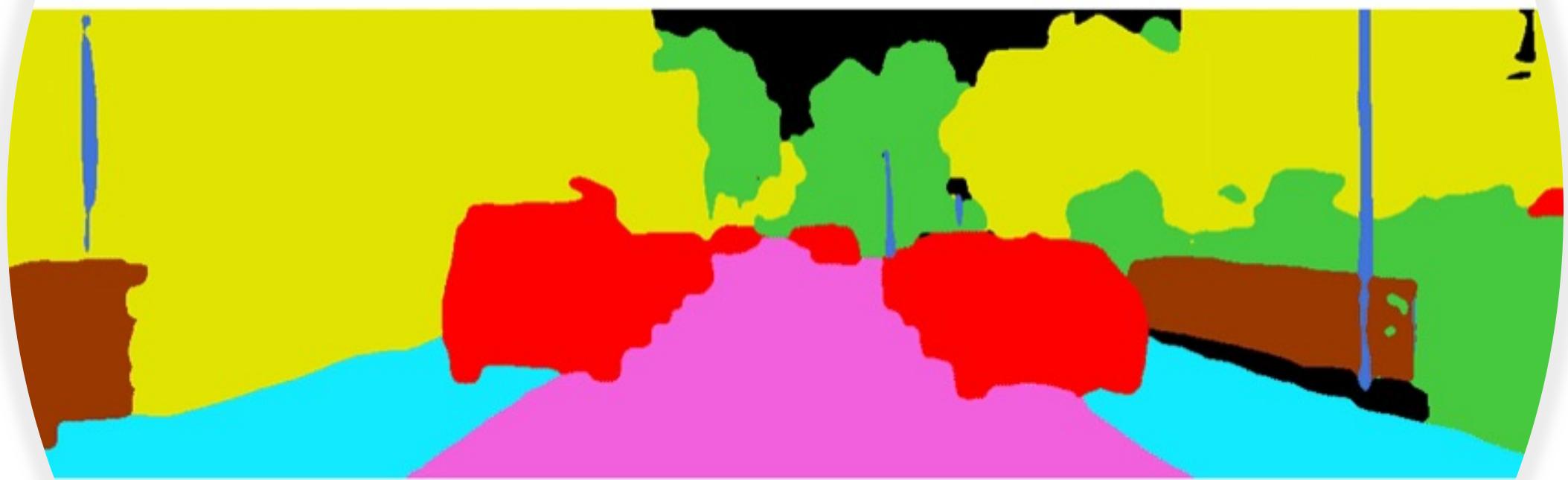
What do we expect?











Road

Sidewalk

Building

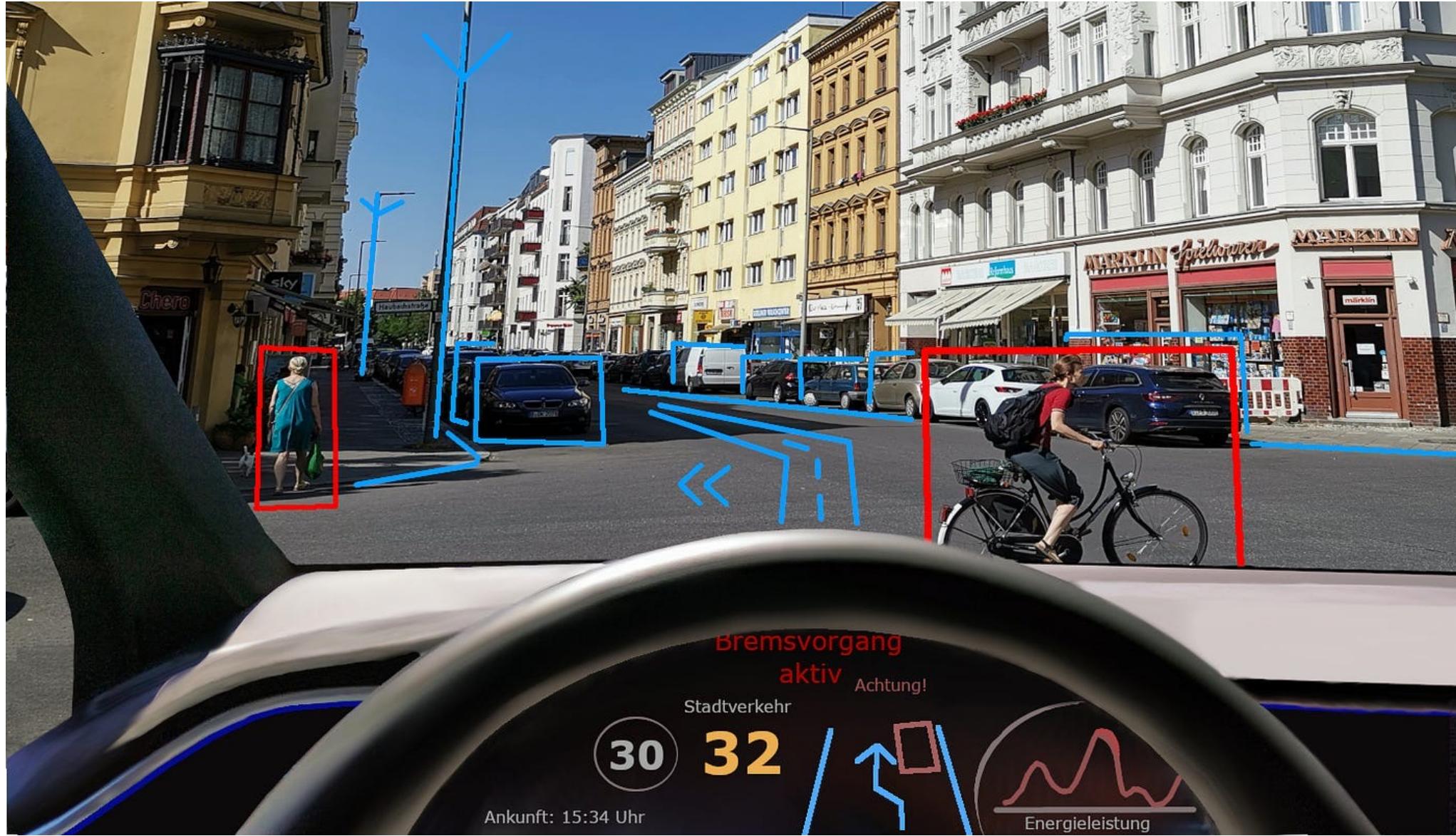
Fence

Pole

Vegetation

Vehicle

Unla'



Applications

- Facial Recognition
- Gesture Recognition
- Medical Imaging
- Autonomous Vehicles
- Augmented Reality (AR)
- Robotics
- Video Analysis
- Quality Control in Manufacturing
- Biometrics
- Document Analysis
- Retail Analytics
- Entertainment



Applications

- **Image Classification:**

Automatically categorizing objects within images, a fundamental task in computer vision, has applications in various domains, such as healthcare, manufacturing, and security.

- **Object Detection:**

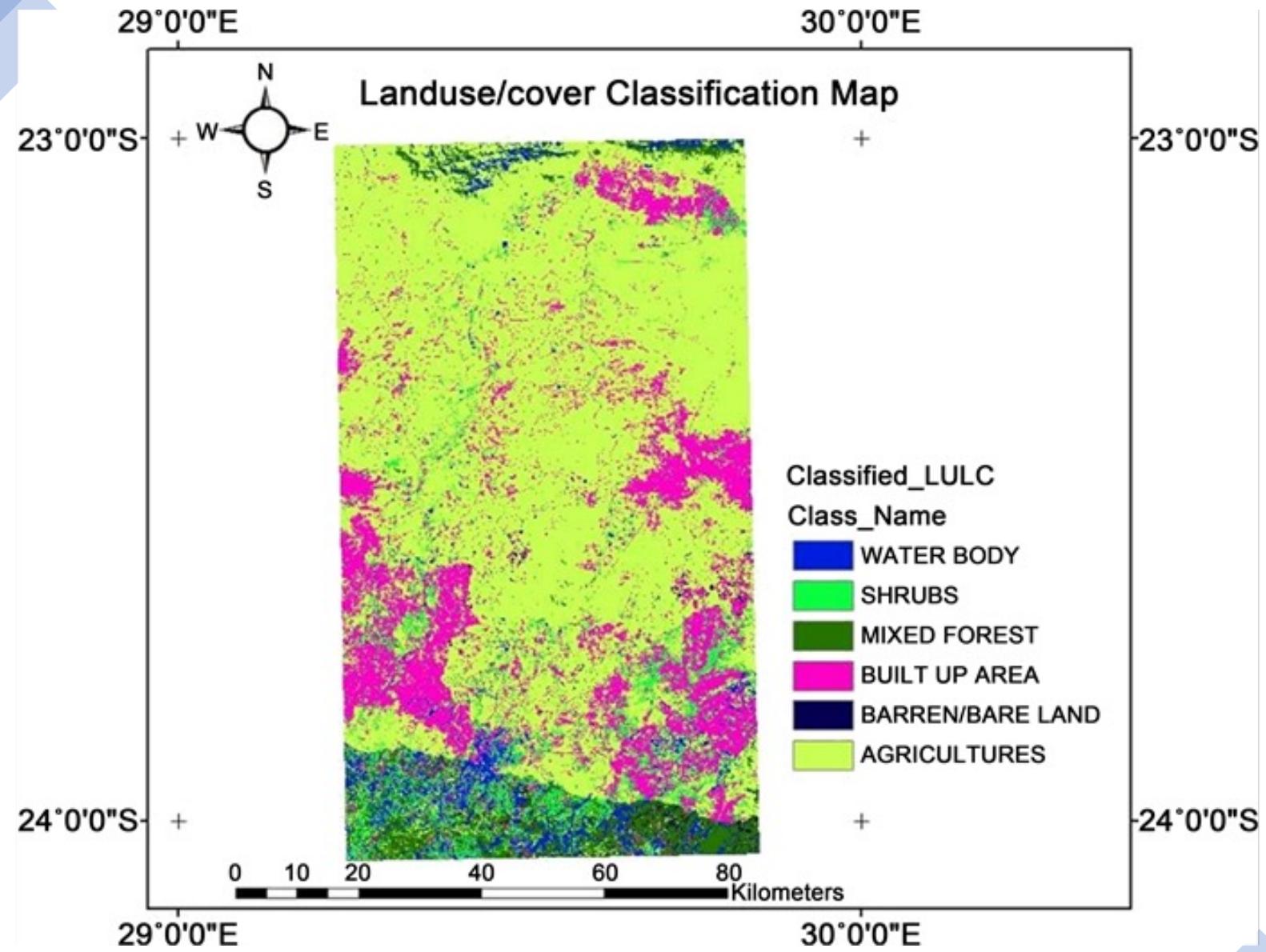
Identifying and locating multiple objects within an image, often used in surveillance, autonomous vehicles, and retail for inventory management.

- **Agricultural Monitoring:**

Monitoring crop health, detecting diseases, and optimizing farming practices using aerial and satellite imagery.

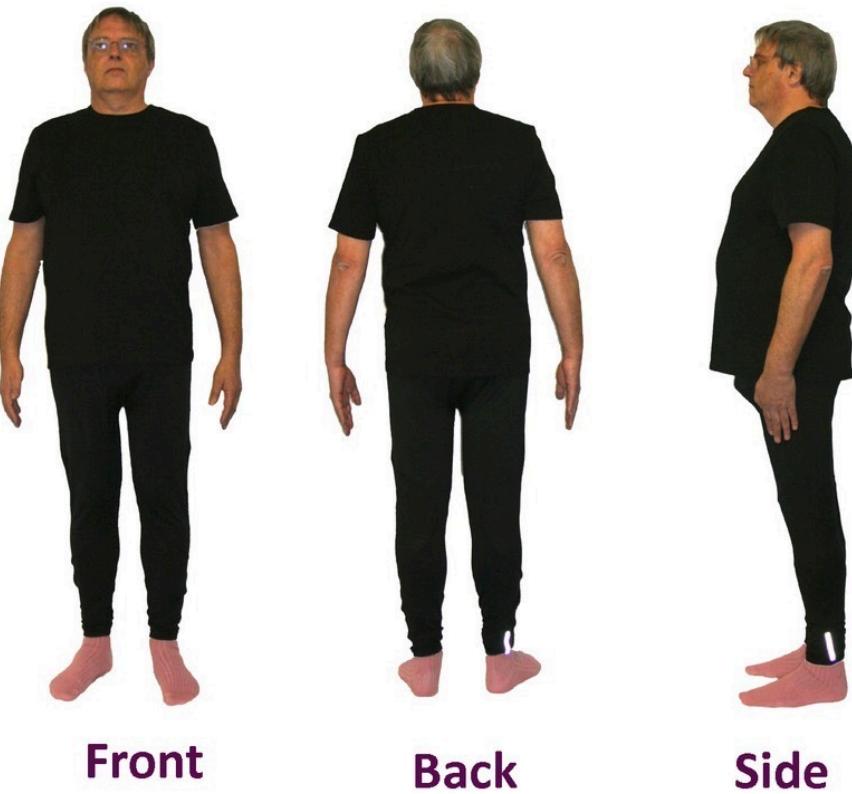
- **Environmental Monitoring:**

Analyzing satellite imagery to monitor and study environmental changes, including deforestation, urban growth, and climate patterns.

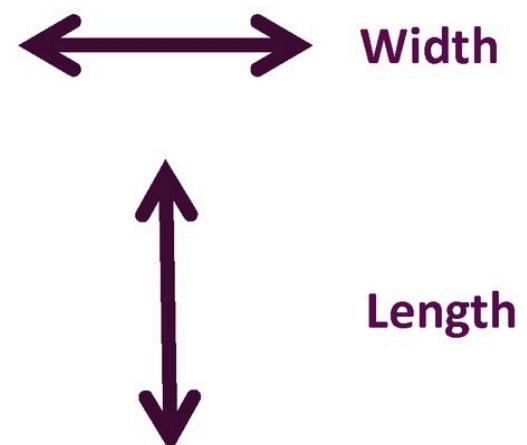


Challenges – viewpoint variations

Three Viewpoints



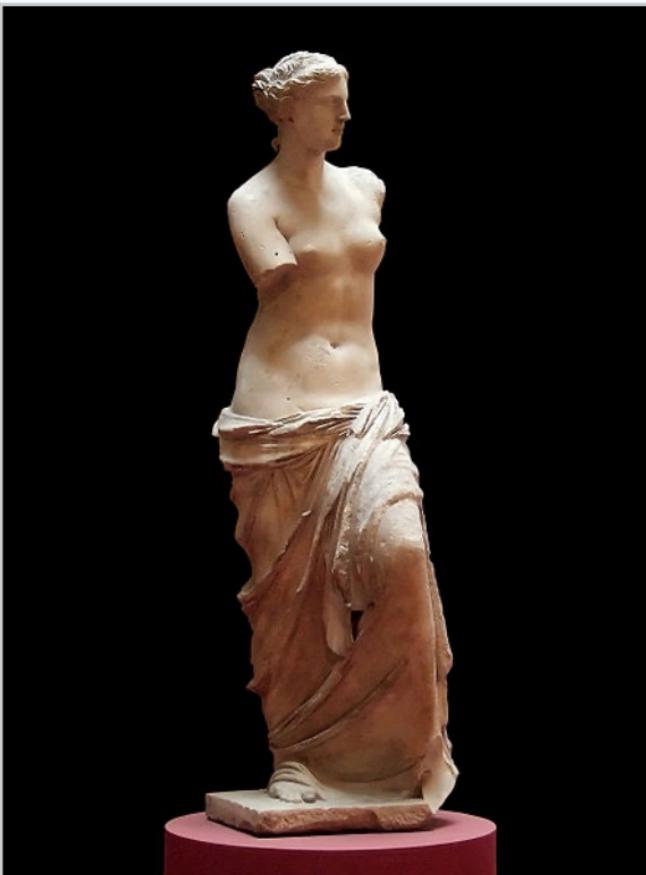
Two Dimensions



Challenges – viewpoint variations



Front view

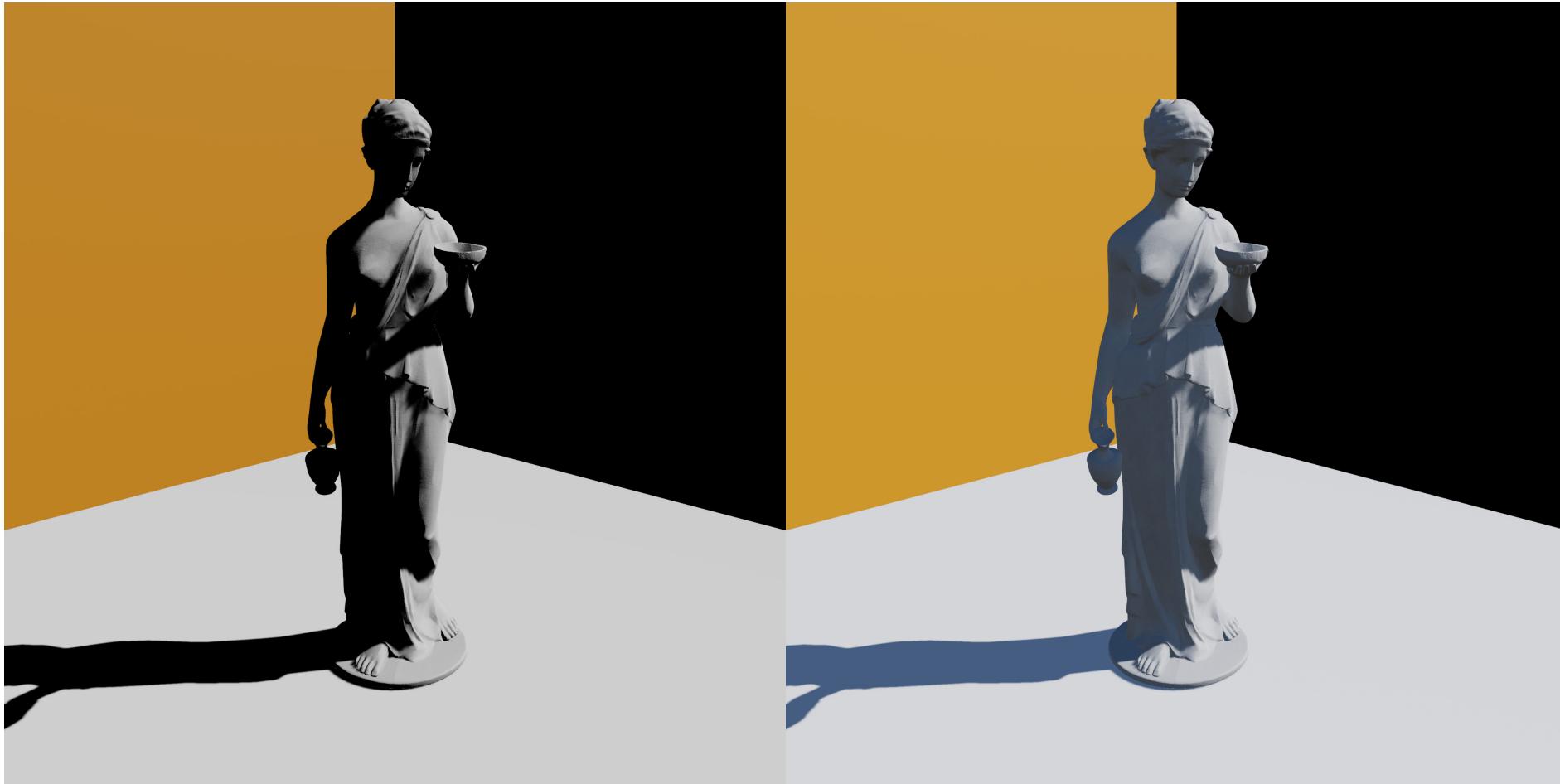


Three-quarter view



Back view

Challenges – illumination



Challenges – scale



Challenges – visibility and overlapping



Challenges – occlusion



Challenges – deformation



Challenges – background clutter



Challenges – motion



Challenges – in-class variations





ILSVRC Competition (ImageNet)



You

Describe in 2 sentences dataset provided in ILSVRC. Provide dataset size, minimum and maximum image resolution and number of classes

ChatGPT

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) dataset comprised over a million labeled images with a diverse set of objects distributed across 1,000 categories. The images varied in resolution, with sizes ranging from a minimum of a few hundred pixels to a maximum of several thousand pixels, providing a comprehensive benchmark for assessing computer vision algorithms.

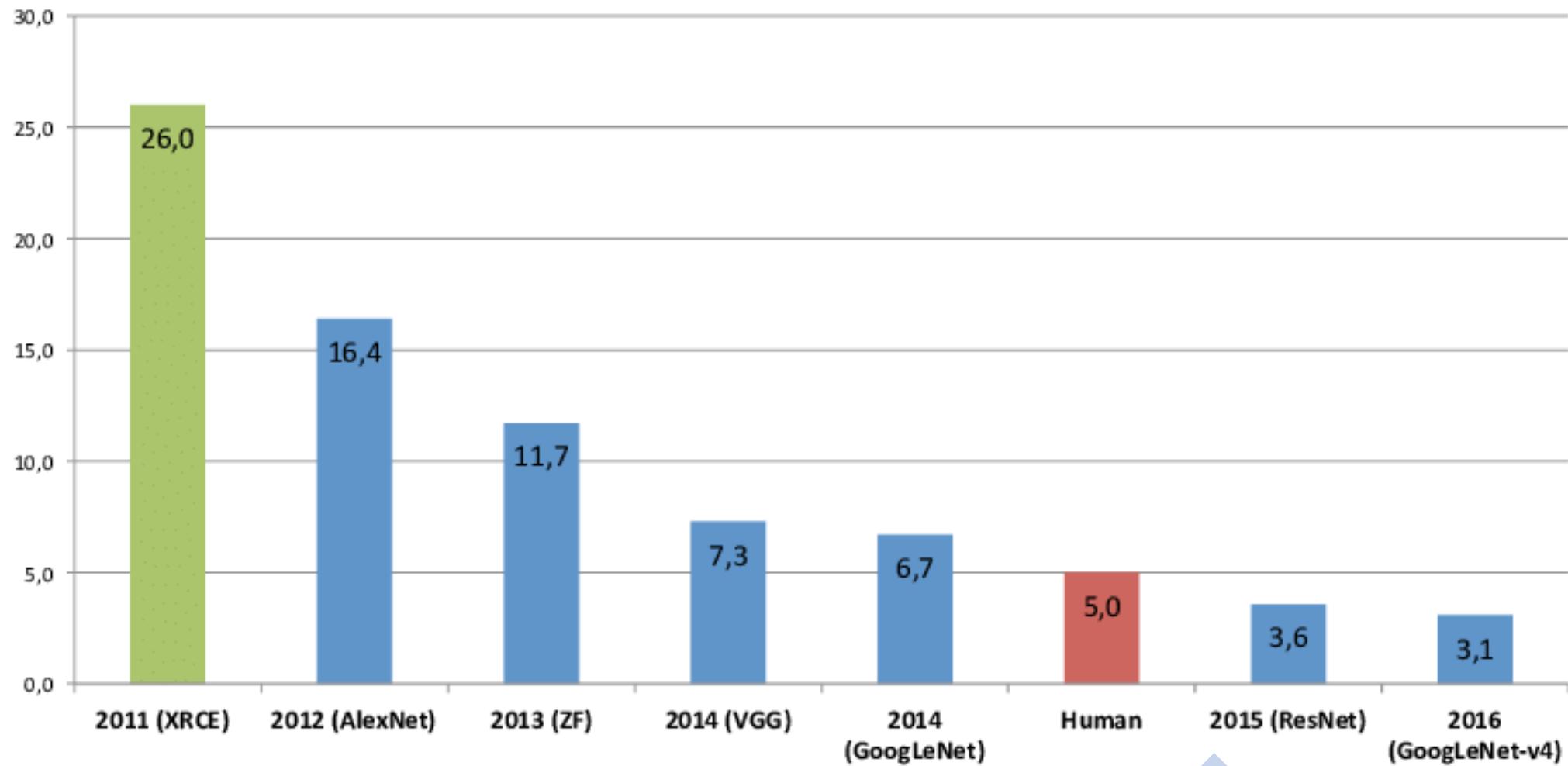
You

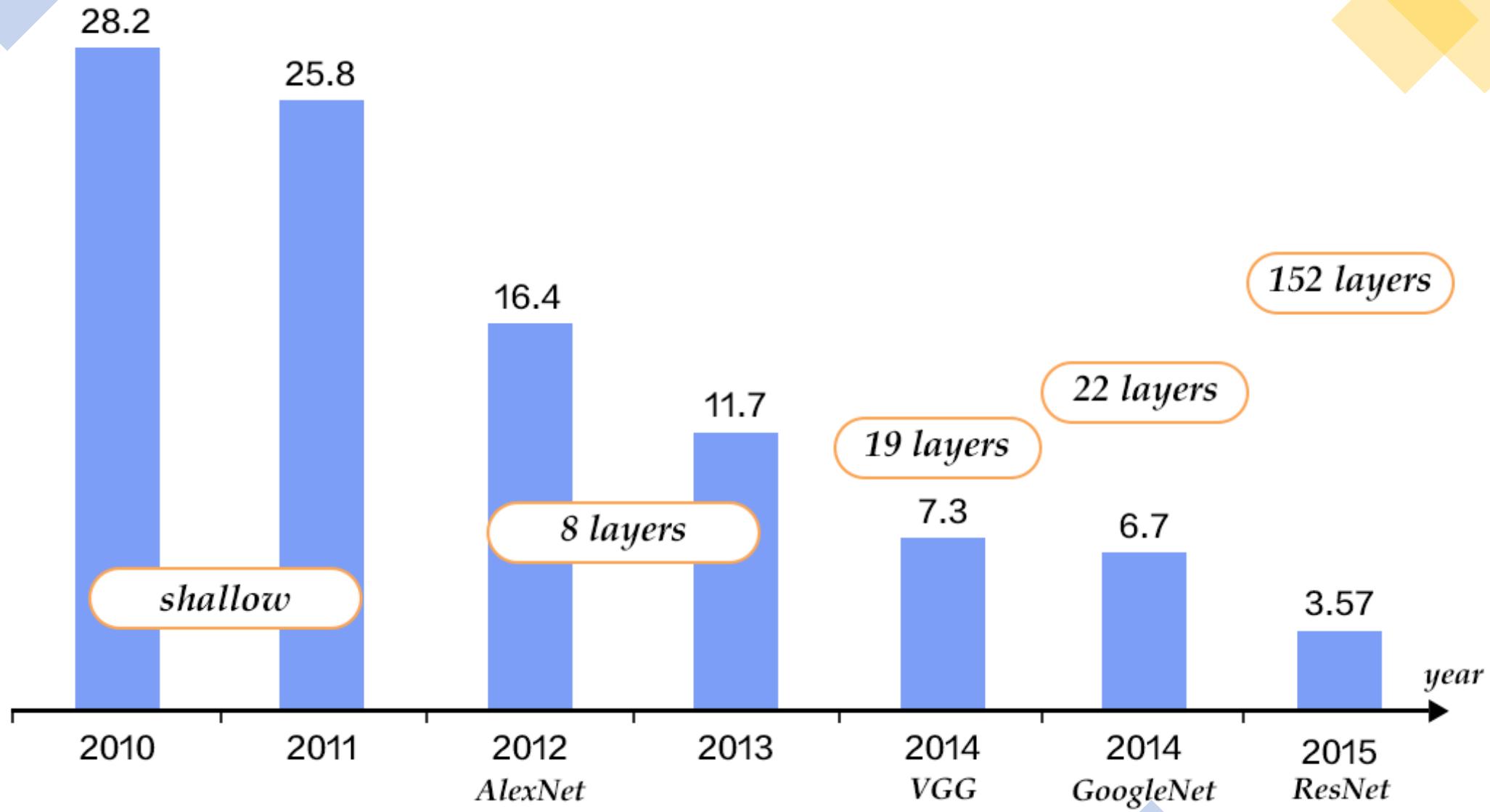
ImageNet Competition. Short description

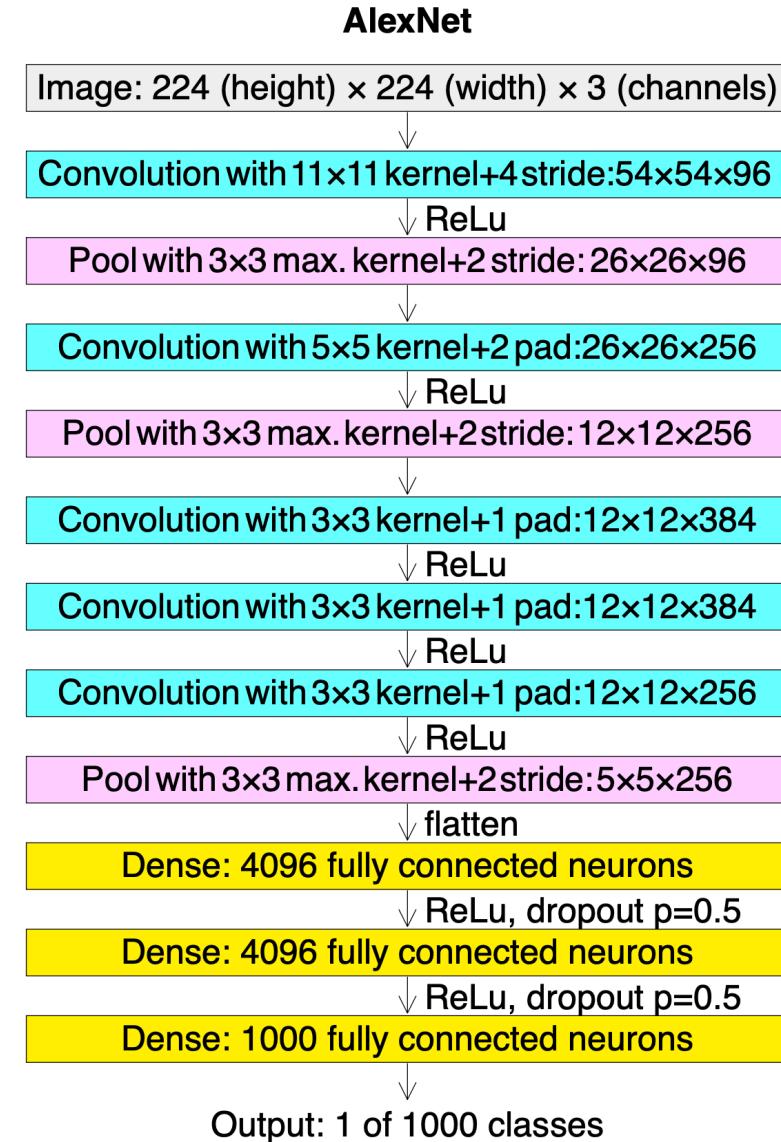
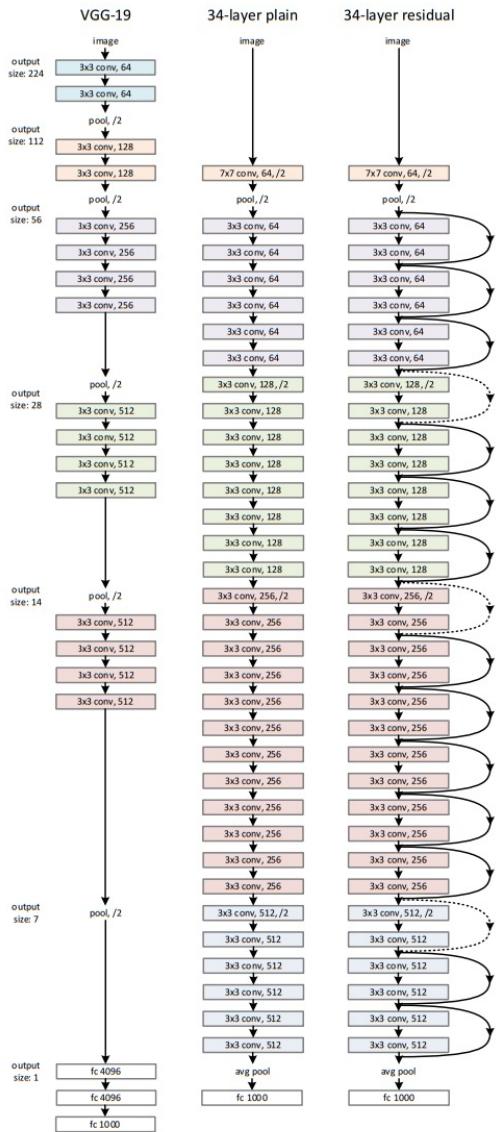
ChatGPT

The ImageNet competition, formally known as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), was an influential annual competition in the field of computer vision. It took place from 2010 to 2017 and aimed to advance the state-of-the-art in image classification and object detection. Participants were tasked with developing algorithms capable of accurately categorizing images from a vast dataset of millions, spanning thousands of object categories. The competition played a pivotal role in the rise of deep learning, with convolutional neural networks (CNNs) achieving breakthroughs in image recognition tasks. The ImageNet competition significantly contributed to the development and popularization of deep learning methods in the broader field of artificial intelligence.

ImageNet Classification Error (Top 5)







Sources

- http://vision.stanford.edu/teaching/cs231a_autumn1213_internal/lecture/lecture1_introduction_cs231a.pdf
- <https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/>
- https://upload.wikimedia.org/wikipedia/commons/c/cc/Comparison_image_neural_networks.svg
- <https://images.iphonemod.net/wp-content/uploads/2018/11/face-id-for-iphone-x-trips-1.jpg>
- https://upload.wikimedia.org/wikipedia/commons/thumb/d/d7/Google_Translate_logo.svg/1200px-Google_Translate_logo.svg.png
- https://blogs.ubc.ca/communicatingscience2018w109/files/2018/11/35477618781_4445e53d96_o-1024x538.jpg
- https://githubraw.cdn.bcebos.com/PaddlePaddle/book/develop/03_image_classification/image/lsvrc.png?raw=true
- https://devopedia.org/images/article/172/9880_1559385994.png
- https://irralves.net/img/2017/11/C134_SupercomputingSpeedsUpDeepLearning.jpg
- <https://i0.wp.com/maxembodied.com/wp-content/uploads/2012/12/computer-vision-autovpr.jpg?w=377>
- <https://haarbook.github.io/img/computer-vision.jpg>
- <https://en.wikipedia.org/wiki/File:Autonomous-driving-Barcelona.jpg>
- https://www.researchgate.net/figure/Example-of-2D-semantic-segmentation-Top-input-image-Bottom-prediction_fig3_326875064
- https://medium.com/@VK_Venkatkumar/3d-reconstruction-basic-terminology-traditional-computer-vision-e148496f389
- <https://www.facebook.com/grettvinnnoise/>
- <http://www.ayelim.co.il:8080/ieins.php?cid=181&kvs=ip+surveillance&e=1>
- <https://www.motherjones.com/politics/2013/10/timeline-technology-silicon-valley-government-subsidies/>
- <https://i1.wp.com/sefiks.com/wp-content/uploads/2018/01/kid-expressions-cover.png?resize=459%2C409&ssl=1>
- https://www.frontiersin.org/files/Articles/493453/fosve-10-02805-HTML-r1/image_m/fosve-10-02805-e001.jpg
- https://en.wikipedia.org/wiki/Venus_de_Milo
- <https://forum.unity.com/attachments/slymanger-combined-img.1124798/>
- http://upload.wikimedia.org/wikipedia/commons/f/f0/Ferrari_Formula_1_lineup_at_the_Nurburgring.jpg
- https://live.staticflickr.com/5256/5492866132_ab169a002f_n.jpg
- https://lh6.googleusercontent.com/-B25wVlOpRL4/TnqoKYoXmVI/AAAAAAAABDs/4q_IO_YKPTo/s0/Rene%25CC%2581%2520Magritte%2520-%2520The%2520Blank%2520Check.jpg
- https://en.wikipedia.org/wiki/Venus_of_Willendorf
- <http://upload.wikimedia.org/wikipedia/commons/3/37/Jackson's-chameleon-3.jpg>
- https://live.staticflickr.com/5564/14548397539_d650124b9f_b.jpg
- <https://i.stack.imgur.com/kllfc.jpg>
- https://i.stack.imgur.com/65535/50472054311_0936c57d7f_b.jpg
- https://live.staticflickr.com/4008/4321311483_89bae0f423_n.jpg
- <https://kecute.files.wordpress.com/2007/10/cats-on-couch.jpg?w=455>
- https://upload.wikimedia.org/wikipedia/en/9/91/jemima_Cats.jpg
- https://images-wixmp-ed30a8b68c4ca88773594c2.wixmp.com/f/2a55962a-2ee8-426d-b8f0-b95c8e350066/d2kk13d-16f592d3-34cc-47fb-8ed5-280e11e3c84b.jpg/v1/fill/w_900,h_510,q_75,strp/warrior_cats_into_the_wild_by_eraizen.jpg?token=evj0eXAiOjKV1QjLCJhbGciOiJIUzI
- https://upload.wikimedia.org/wikipedia/commons/6/6f/Malayan_Tiger_Cubs.jpg
- <https://html.scirp.org/file/14-2801413x5.png>