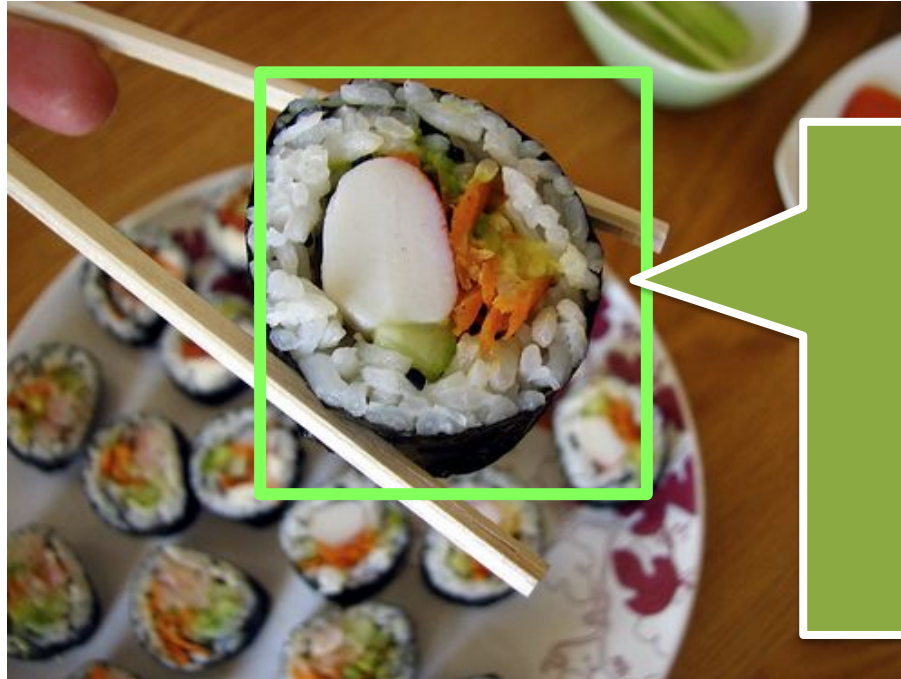


# CSE455 - Recitation

Large Scale Recognition



## California Roll

Ingredients: Rice, Seaweed,  
Crab, Cucumber, Avocado

Calories: 40

Fat: 7g

Carb: 40g

Protein: 5g

**Gluten Free**



## **Amanita phalloides**

[http://en.wikipedia.org/wiki/Amanita\\_phalloides](http://en.wikipedia.org/wiki/Amanita_phalloides)

**TOXIC. DO NOT EAT**





## Mountain Lion

### DO NOT RUN

Raise arms to appear larger.

Show your teeth



**IKEA POANG Chair**  
**ON SALE**  
\$29.00 at [ikea.com](https://www.ikea.com)



## **Mornonga (Japanese flying squirrel)**

Inhabits sub-alpine forests in Japan.  
Nocturnal. Eats seeds, fruit, tree leaves  
(Wikipedia)

I wish my computer could recognize  
EVERYTHING





Surveillance



Robotics



Assistive tools



Wearable devices



Driverless cars



Smart photo album

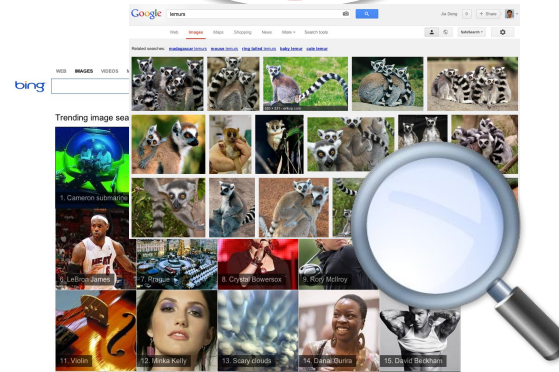


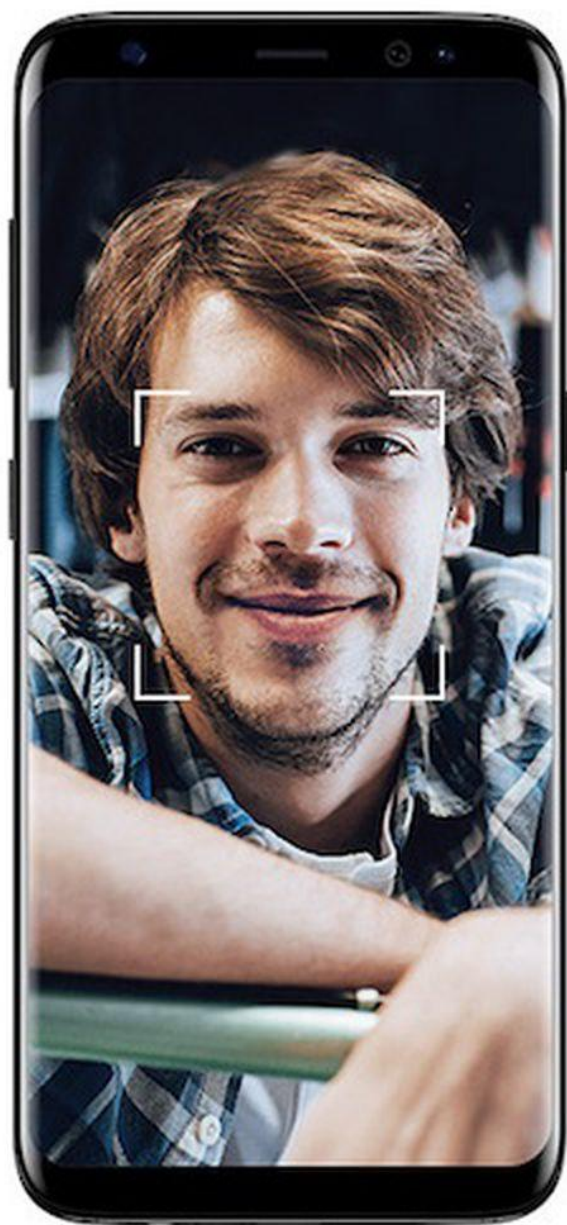
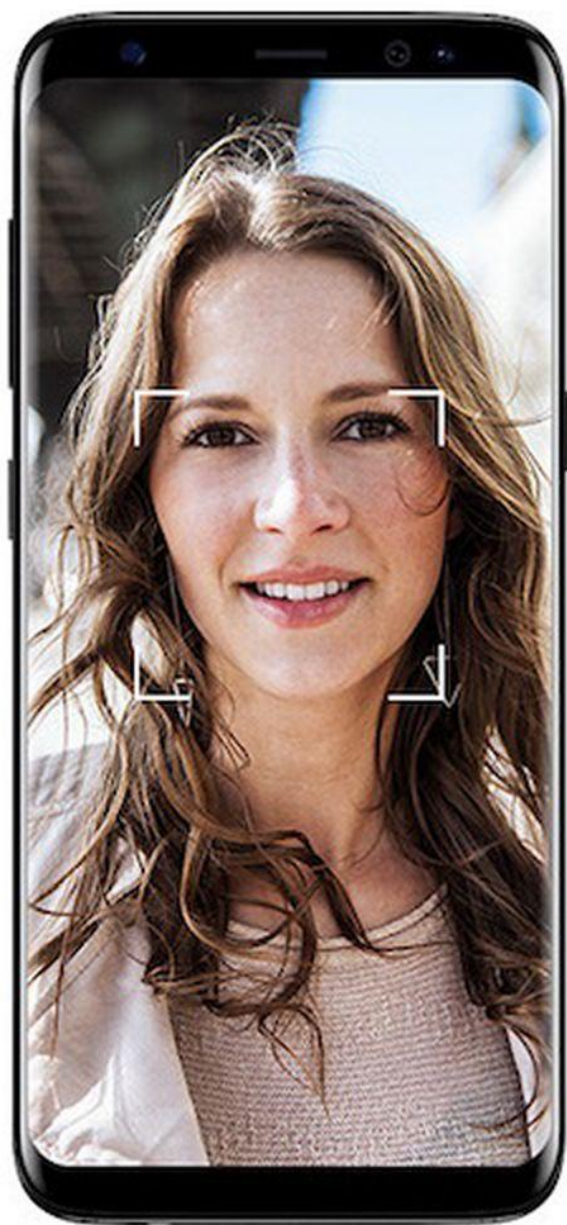
Image search



Mining social media

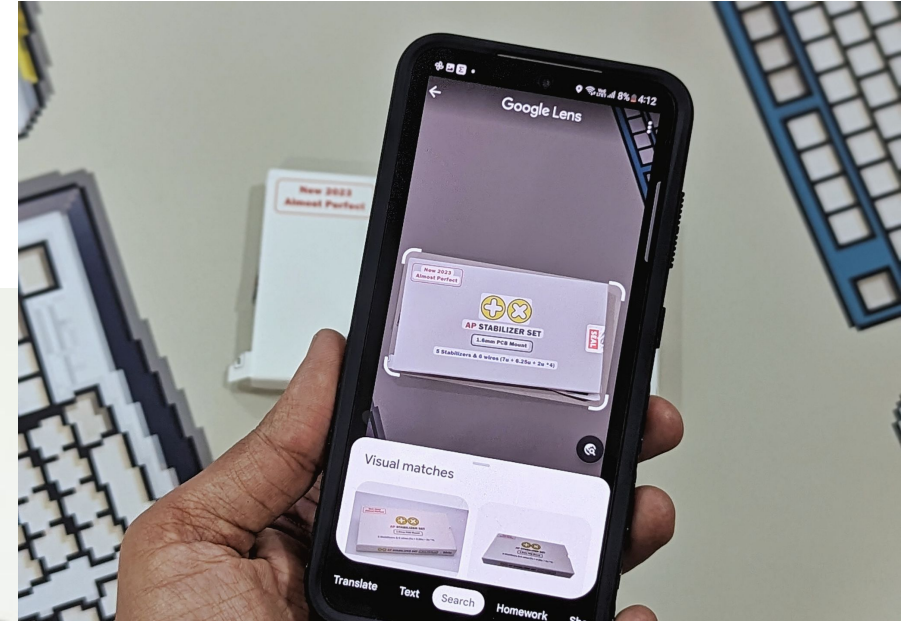
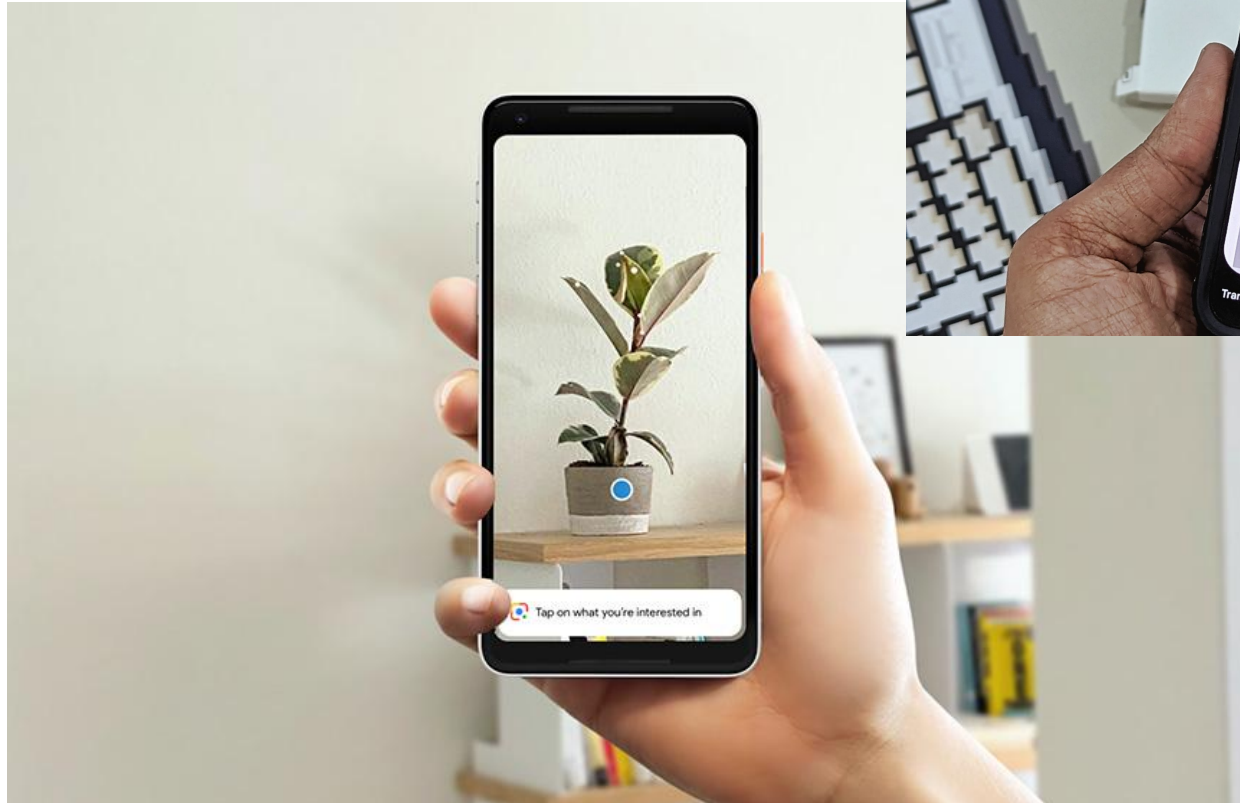


**What can computers already recognize?**





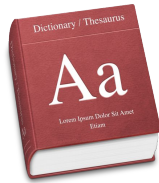
Google Lens





Let's work on recognizing **EVERYTHING**

# How many things are there?



10K+  
[Biederman '87]



60K+  
product  
categories



80K+  
English nouns  
[Miller '95;  
Fellbaum '98]

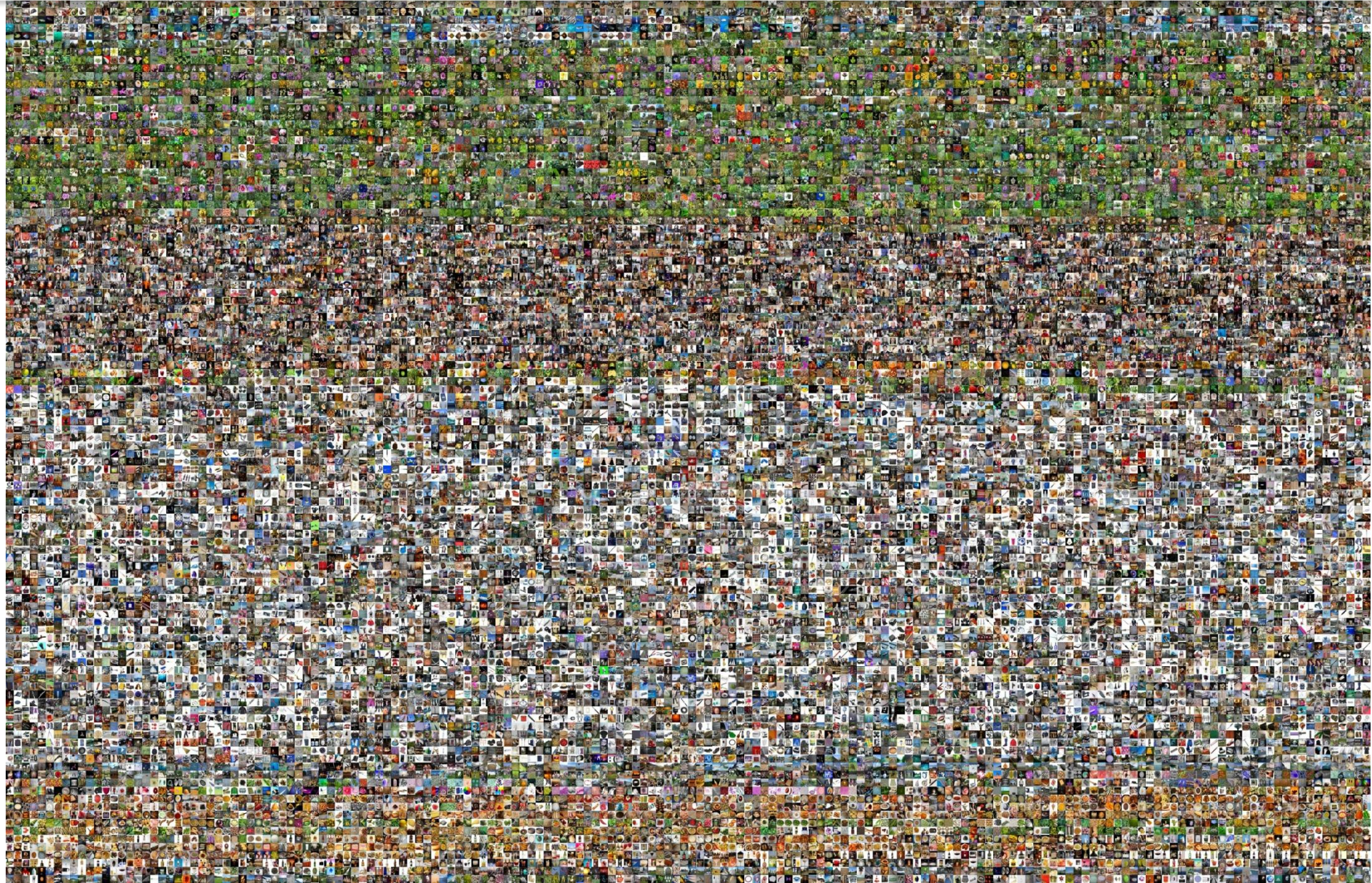


3.5M+  
unique tags  
[Sigurbjörnsson &  
Zwol '08]



4.1M+  
articles

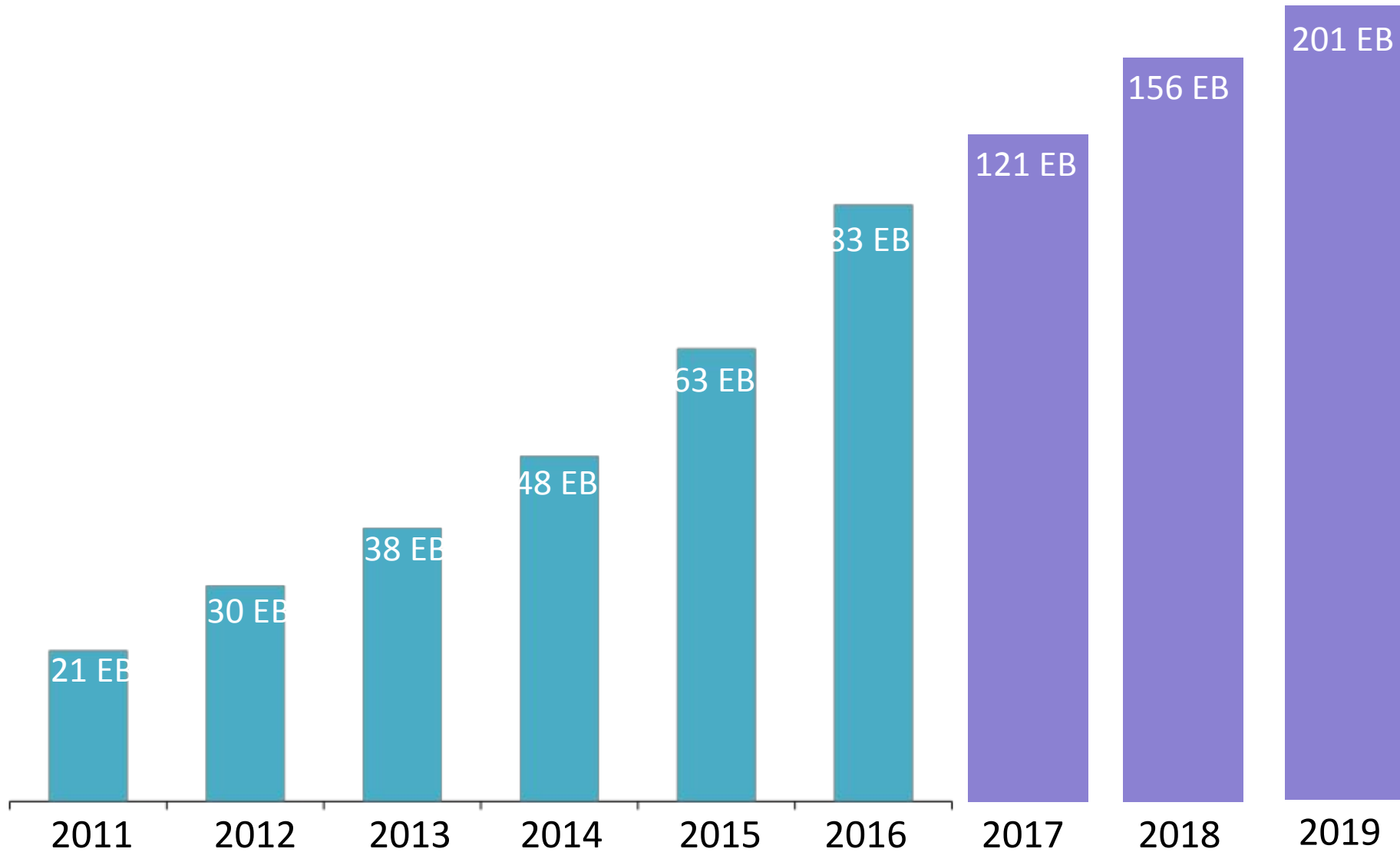
From 20 classes to Millions?







# Global Consumer Internet Traffic Per Month



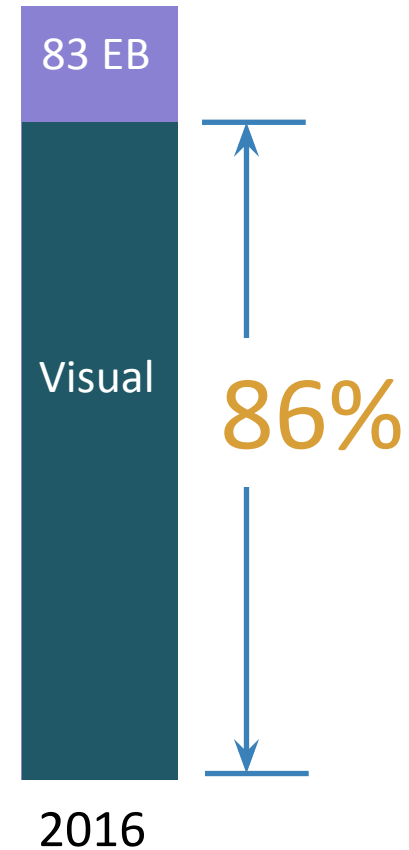
Source: Cisco



72 hours of videos / min



300 million images / day



Source: Cisco

# Big Data from the Internet

- The Internet can teach  
**EVERYTHING**

Google

pitbullfrog



### Evolution Gone Wild

*Future plants and animals*

<http://www.worth1000.com/contests/12705/contest>



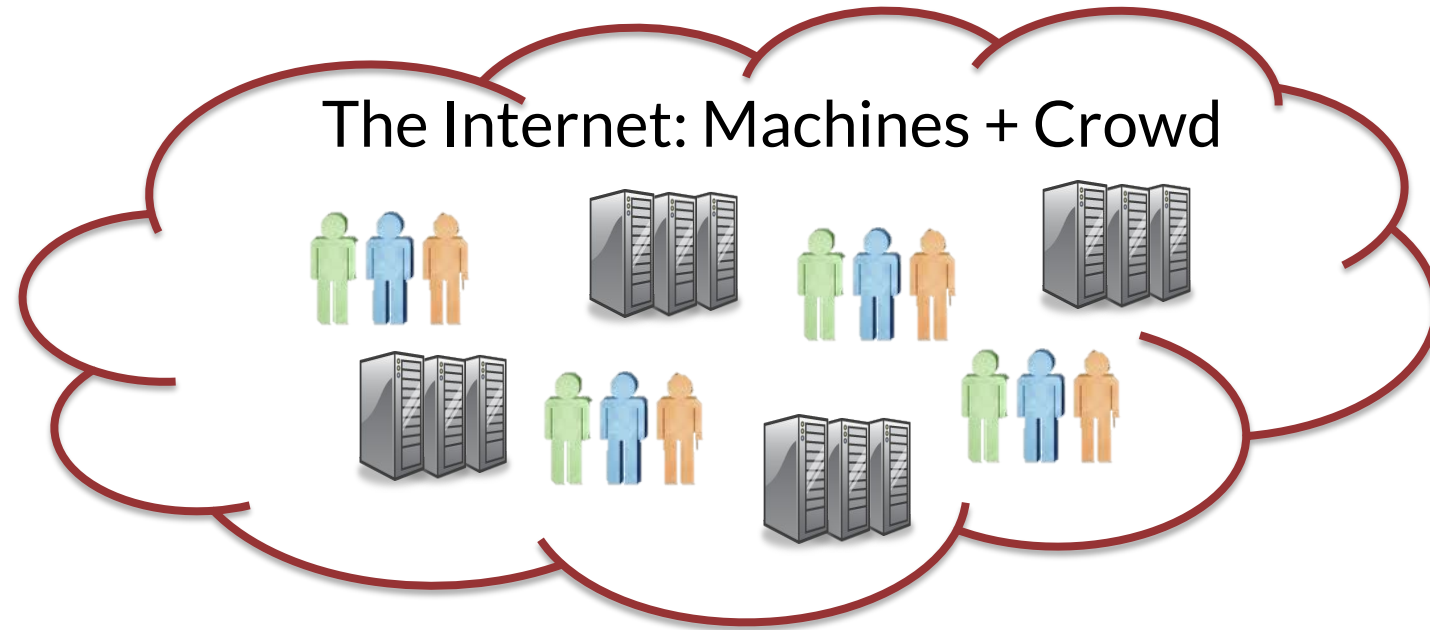
**Anon User**

2 votes by Anon User and Anon User

It looks like a Northern Trust Visa, which would make sense given his public disclosures report a banking relationship with the firm:



**What kind of credit card is President Obama using in this video of him donating to his campaign?**



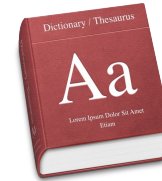
Teach machines to recognize **EVERYTHING**

# PASCAL VOC



20

[Everingham et al.'06-'12]



10K+

[Biederman '87]

Goal: Build a recognition engine on ~~EVERYTHING~~  
10K classes



IMAGENET [Deng et al. 2009]

[www.image-net.org](http://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





# Number of Labeled Images

SUN, 131K  
[Xiao et al. '10]

LabelMe, 37K  
[Russell et al. '07]

PASCAL VOC, 30K  
[Everingham et al. '06-'12]

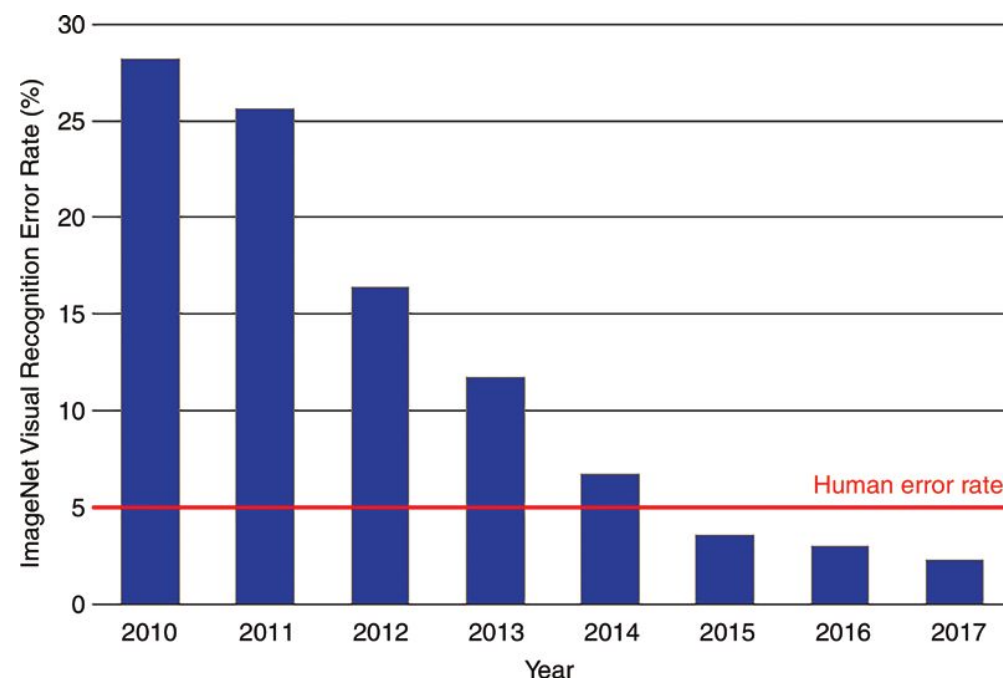
Caltech101, 9K  
[Fei-Fei, Fergus, Perona, '03]

ImageNet, 14M  
[Deng et al. '09]



# Learn to Classify 10K Classes

Model	Time	Accuracy	Num. of Parameters	Num. of FLOPs	Num. of Layers
AlexNet [21]	2012	57.2%	60 M	720 M	8
VGGNet [22]	2014	71.5%	138 M	15,300 M	16
GoogleNet [23]	2014	69.8%	6.8 M	1,500 M	22
ResNet [24]	2015	78.6%	55 M	2,300 M	152
DenseNet [25]	2017	79.2%	25.6 M	1,150 M	190
SENet [26]	2017	82.7%	145.8 M	42,300 M	–
NASNet [27]	2018	82.7%	88.9 M	23,800 M	–
SqueezeNet [29]	2016	57.5%	1.2 M	833 M	–
MobileNet [30]	2017	70.6%	4.2 M	569 M	28
ShuffleNet [31]	2018	73.7%	4.7 M	524 M	–
ShiftNet-A [32]	2018	70.1%	4.1 M	1,400 M	–
FE-Net [33]	2019	75.0%	5.9 M	563 M	–



Other ways to use these models?



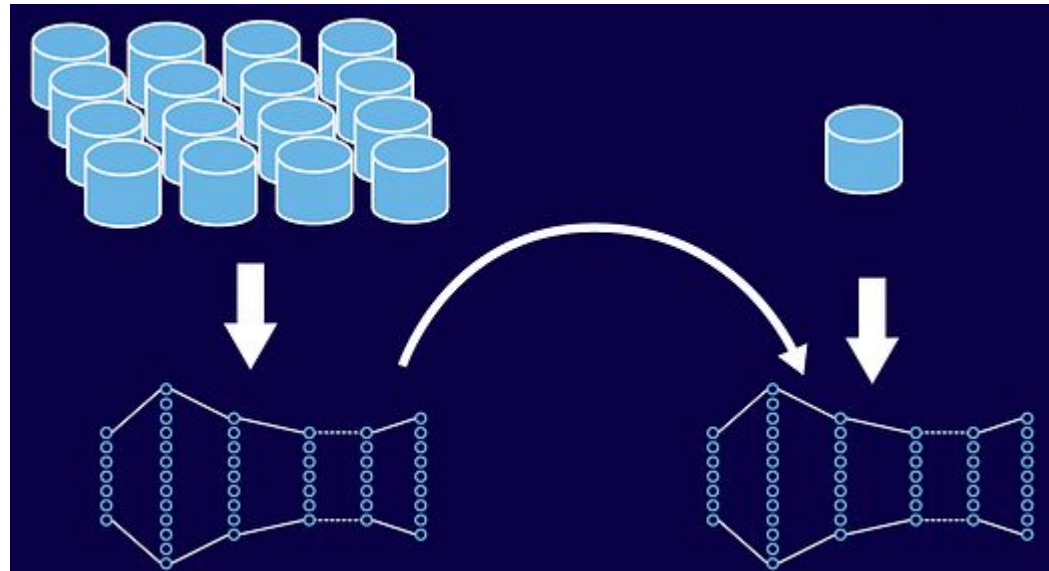
# Transfer Learning

- What is Transfer Learning?
  - Leveraging a model pre-trained on a large dataset (ImageNet) to solve a different but related task.
  - Utilizes learned features (e.g., edges, textures) that are generic across many vision tasks.
- Why Use ImageNet Pre-trained Models?
  - Efficient Use of Data
  - Robust Feature Extraction
  - Reduced Training Time
  - Improved Performance

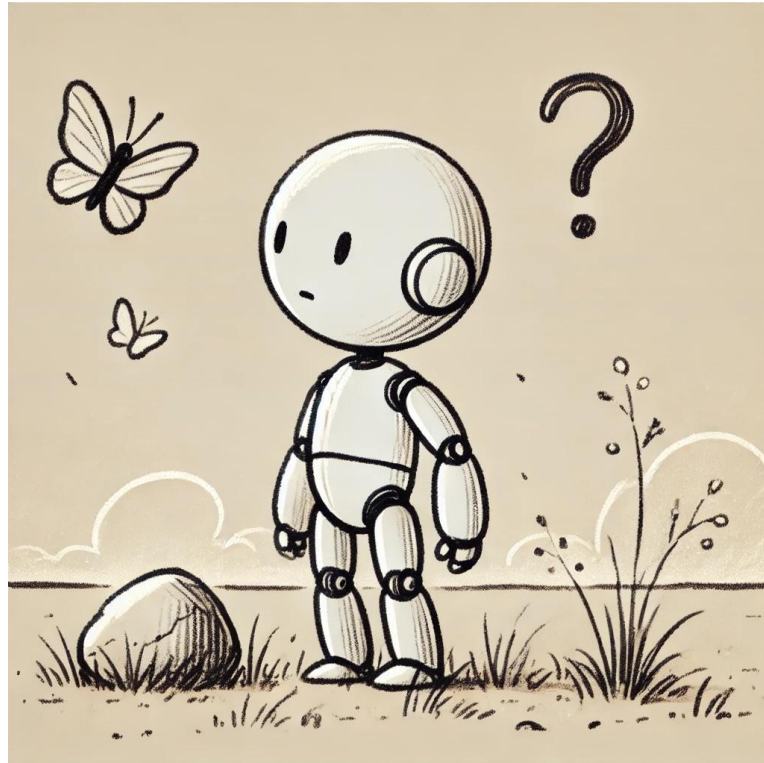


# Transfer Learning

- Methods of Transfer Learning:
  - Feature Extraction: Use as a backbone for a new model
  - Fine-Tuning: Train some of the later layers
- Real-World Examples:
  - Medical Imaging
  - Object Detection



Enough technicalities, where is the fun?



# Mornonga



What do you see in this picture?



I see three adorable baby squirrels perched on a tree branch. They have light brown fur with white bellies and black eyes. The baby squirrels are huddled together, with one in the center holding onto the branch, another to its right, and a third partially visible in the background. They appear to be very young and fluffy, with their tiny paws gripping the branch. The background is blurred, showing a mix of brown and white, with a hint of blue sky visible in the upper right corner.



ChatGPT 4o >



What do you see in this picture?

This image features several small, fluffy rodents clinging to a tree, likely a species of flying squirrel. They have large, dark eyes, soft grayish fur, and white bellies, which are characteristic of certain types of flying squirrels, such as the Japanese dwarf flying squirrel (*Pteromys momonga*) or the Siberian flying squirrel (*Pteromys volans*). These nocturnal creatures are known for their ability to glide between trees using a special membrane called the patagium.

Would you like any further details about them?

# Amanita phalloides



What do you see in this picture?



I see a close-up view of a forest floor with two mushrooms as the main focus. The mushrooms have white stems and light yellowish-white caps. One mushroom is standing upright, while the other is tilted to the left. The ground is covered with a mix of brown and green leaves, some of which appear wet. There are also some twigs and branches scattered around. On the right side of the image, I can see some green leaves, and in the upper left corner, there are two green leaves with serrated edges. The overall scene gives a sense of a damp, natural forest environment.



ChatGPT 4o >



What do you see in this picture?

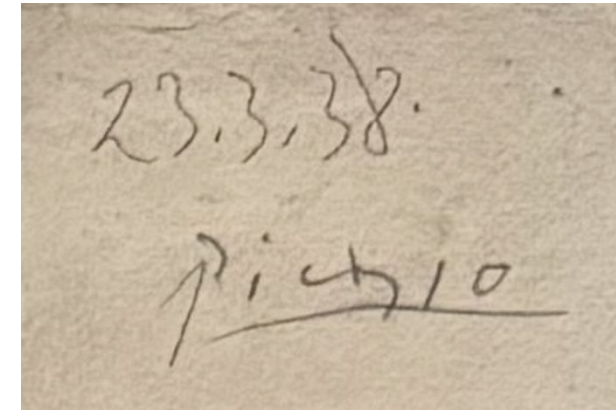
The image shows two mushrooms growing in a forested environment with damp, decaying leaves, twigs, and ferns surrounding them. The mushrooms have a pale yellowish cap with a slightly slimy or wet appearance, and they have white gills and stems.

These mushrooms resemble *Amanita phalloides* (commonly known as the death cap), which is highly toxic. However, precise identification requires careful examination of features like spore print color, habitat, and other microscopic characteristics.

If you were considering foraging, I strongly advise against consuming any wild mushrooms unless identified with certainty by an expert. Let me know if you need help with identification!



# Painting



What do you see in this picture?

The image shows a striking black and white pencil drawing of a rooster. The rooster is depicted in a side profile, facing left, with its beak wide open as if crowing. Its feathers are intricately detailed, with a prominent crest on its head and a long, flowing tail. The bird's feet are clearly visible, showing three toes on each foot.

The drawing is framed in a thin black frame with a white mat, hanging on a gray wall. The background of the drawing itself is a light brown or tan color, which contrasts nicely with the dark lines of the rooster.

In the upper right corner of the drawing, there's a signature that reads "23338" and "124150", which might be the artist's name and a date or series number.

The overall composition is simple yet powerful, with the rooster as the clear focal point against the lighter background. It's an excellent example of detailed pencil work capturing the essence and energy of a rooster.



ChatGPT 4o >



What do you see in this picture?

This image appears to be an expressive, abstract drawing of a rooster, rendered in black charcoal or ink on paper. The exaggerated features, such as the wide-open beak, large eyes, and dynamic, rough strokes, give the piece a dramatic and intense feel. The signature in the top right corner suggests that it could be a work by Pablo Picasso, as it exhibits his characteristic distortion of form and bold, gestural lines.

Would you like to know more about this piece, such as its historical context or artistic significance?

# Privacy & Offline

# Questions?

