# Lecture 9: Machine Learning Overview

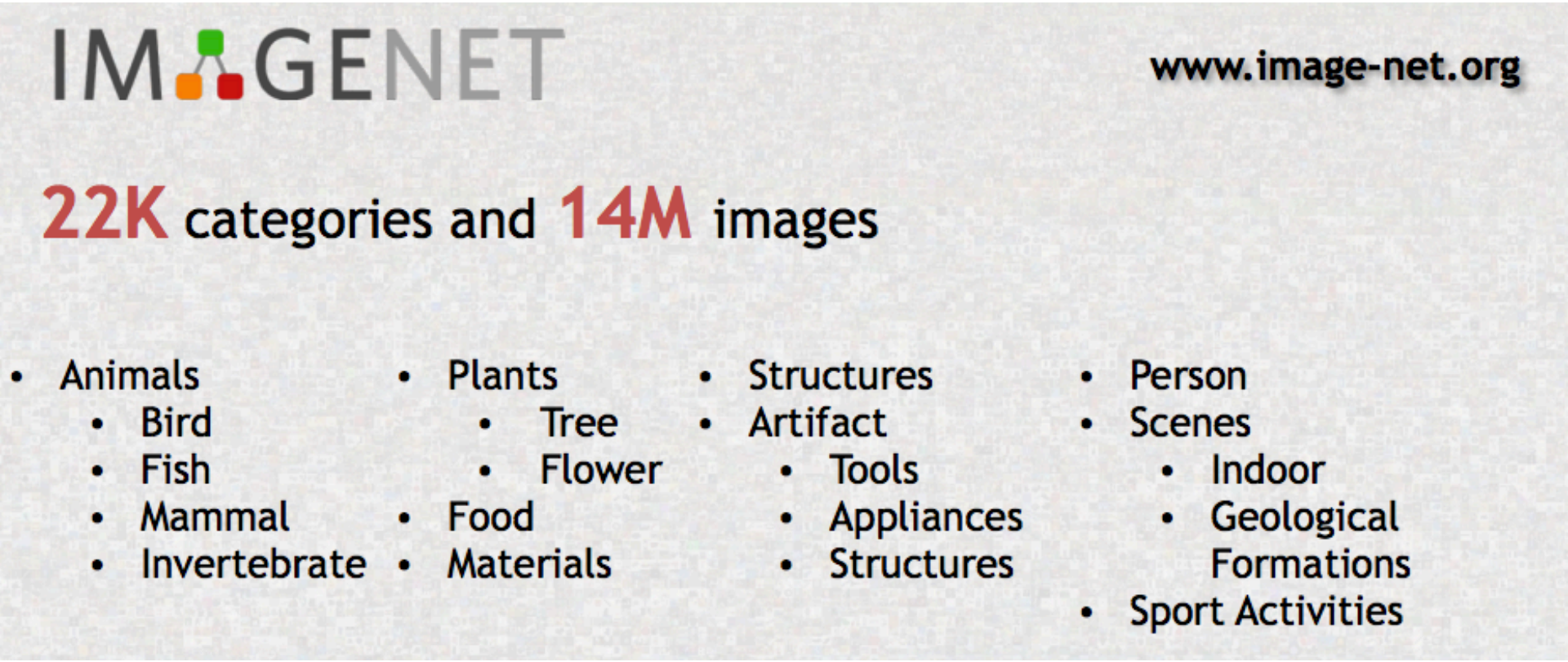
## AP Computer Science Principles

# Machine Learning

- **Machine learning(statistical learning, artificial intelligence)** is a large set of tools and algorithms that are used to process and understand data.
- Two types of machine learning:
  - **supervised learning:** algorithms that learn patterns/structures from labeled data(training examples) and then make predictions on new data. (e.g. neural network, linear/logistic regression)
  - **unsupervised learning:** algorithms that learn to find patterns/structures from unlabeled data.
    - reinforcement learning
- Successful machine learning: not about algorithms. Algorithms have been around since the 80's and 90's.
  - data, data, data(labeled data)
  - computing power(GPU)
  - infrastructure(software)

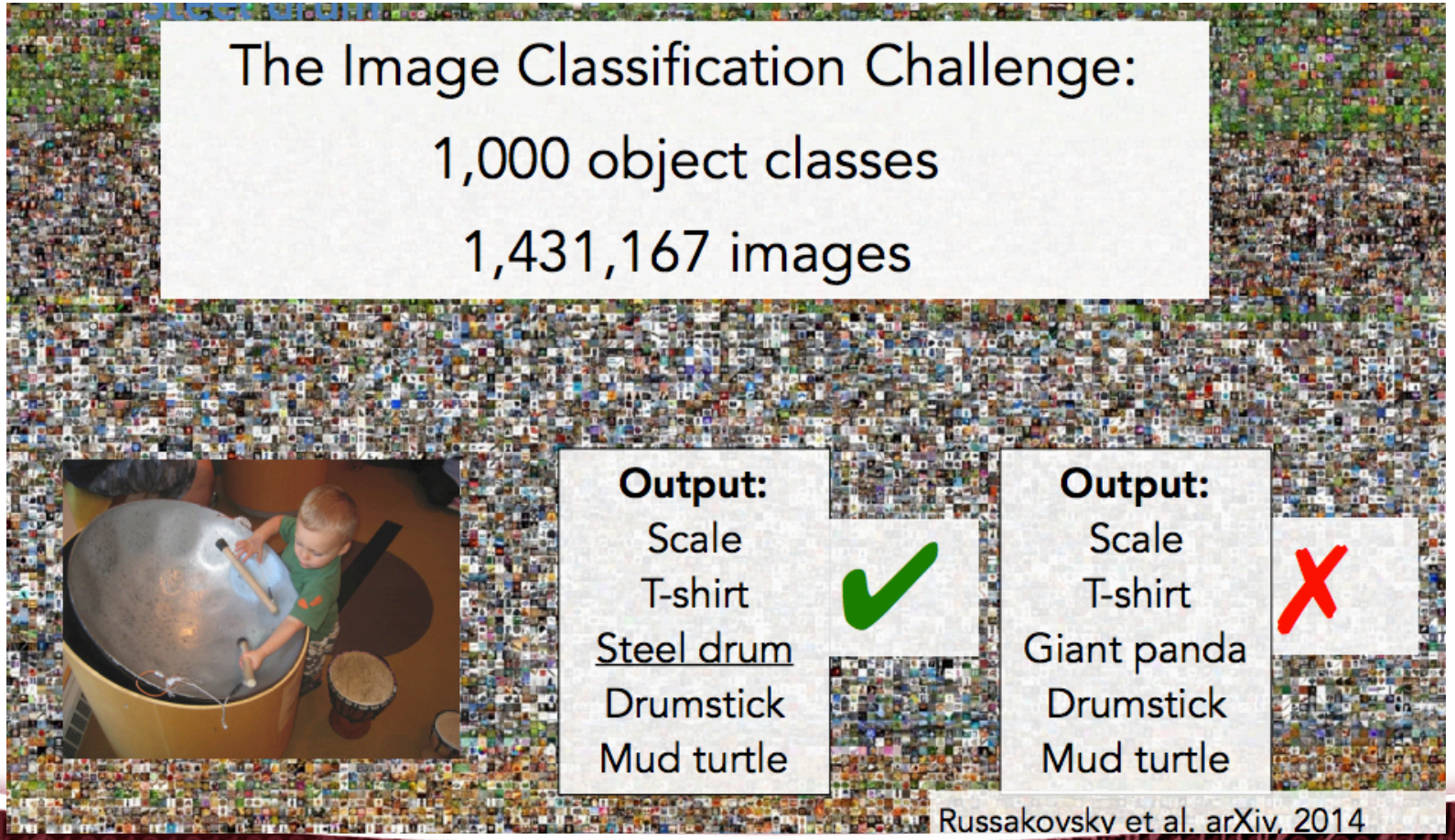
# IMAGENET

- IMAGENET: database that provides "clean", labeled images for machine learning algorithms.



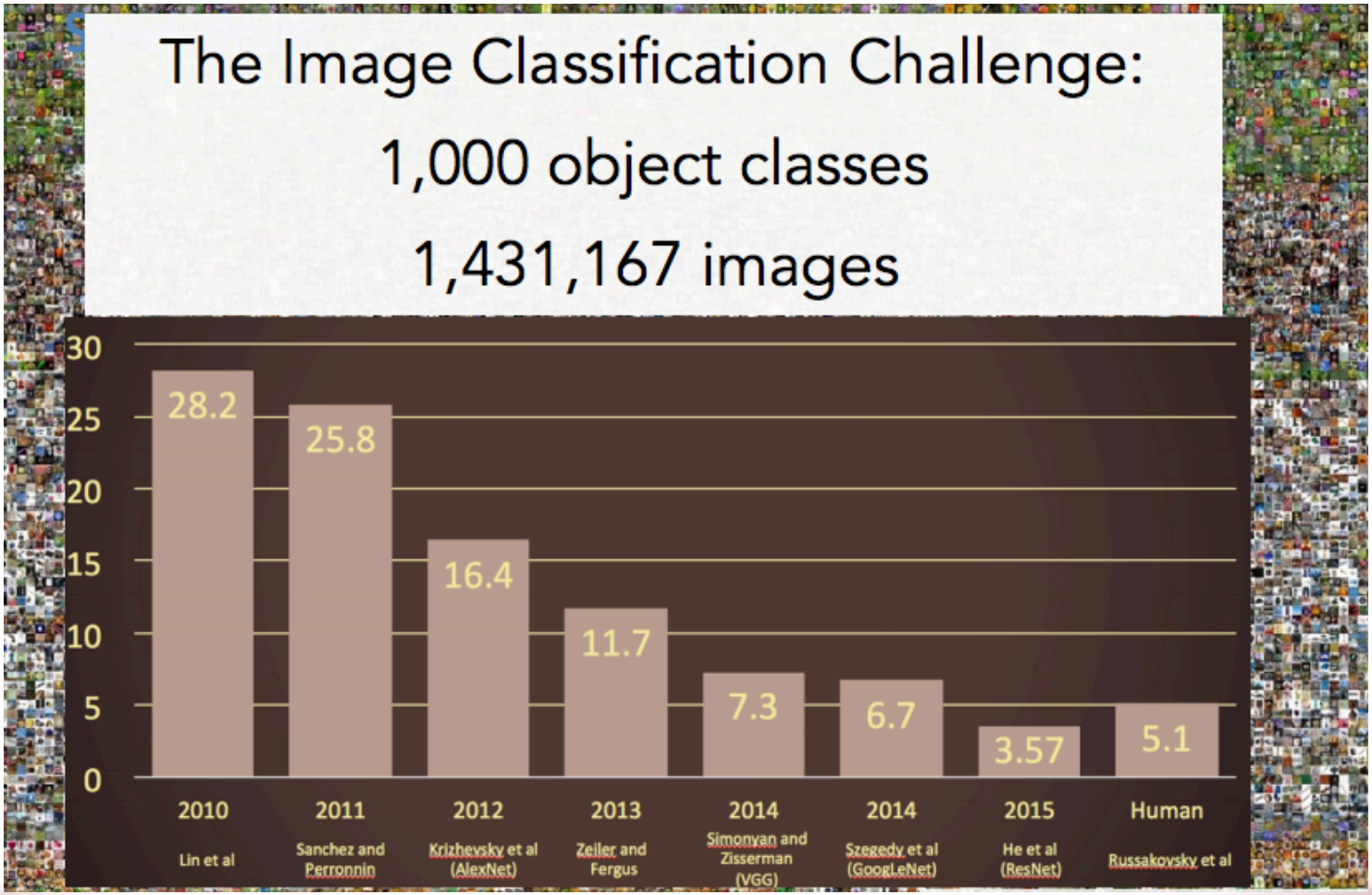
# Image Classification Challenge

- competition that encourages researchers to compare progress on image recognition algorithms.



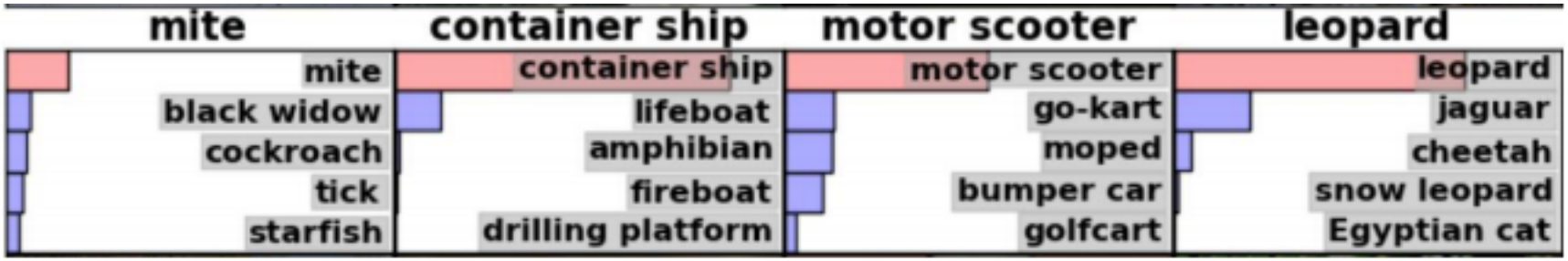
# Error Rates

- A major breakthrough occurred in 2012. A group of research scientist submitted a convolutional neural network(CNN) that demolishes the competition.



# Identify these objects

- **Can you beat a computer by correctly identifying these objects?**



# Machine Learning

- Important machine learning problems:
  - autonomous driving
  - natural language processing:
    - language translation
    - sentiment analysis(rate reviews on Amazon)
    - document similarity(search engines)
      - word2vec(Google)
      - GloVe(Stanford)
    - speech recognition(Siri, Alexa, Watson)
    - semantic analysis
    - question/answering(Siri, Alexa)
    - spam detection
    - name-entity extraction
      - Teddy Roosevelt was a US President.
      - Teddy bears are on sale!

# Machine Learning

- Important machine learning problems:
  - image recognition
    - face detection/recognition(Facebook)
    - image captioning
    - segmentation
    - object detection
  - medical prognosis
  - product recommendation(Amazon, Netflix)
  - credit card fraud detection

# Identify the task

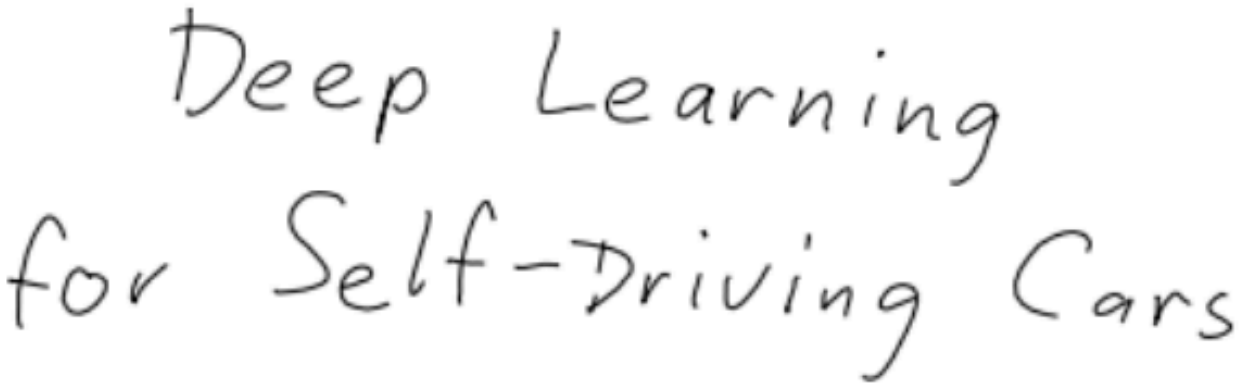
- Identify the following computer task

**Input:**

**Text** --- up to 100 characters, lower case letters work best

Deep Learning for Self Driving Cars

**Output:**




Answer: handwriting generation

# Identify the task

- Identify the following computer task

Life Is About The Weather!
Life Is About The (Wild) Truth About Human-Rights
Life Is About The True Love Of Mr. Mom
Life Is About Where He Were Now
Life Is About Kids
Life Is About What It Takes If Being On The Spot Is Tough
Life Is About… An Eating Story
Life Is About The Truth Now

The meaning of life is literary recognition.
The meaning of life is the tradition of the ancient human reproduction

Answer: (natural language processing: sentence completion)

# Identify the task(Generating Shakespeare)

```
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:
Well, your wit is in the care of side and that.
```

- **Notice the model learns the character/dialogue format of Shakespeare. Andrej Karpathy of Stanford trained the neural network for several hours. (Recurrent Neural Network)**

# RNN Generating Shakespeare

- This can be done using:

1) N-grams:(easy)

2) RNN(recurrent neural network): much harder math

This text is generated **one letter at a time** by an RNN.

```
Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:
Come, sir, I will make did behold your worship.

VIOLA:
I'll drink it.
```

```
KING LEAR:
O, if you were a feeble sight, the courtesy of your law,
Your sight and several breath, will wear the gods
With his heads, and my hands are wonder'd at the deeds,
So drop upon your lordship's head, and your opinion
Shall be against your honour.
```

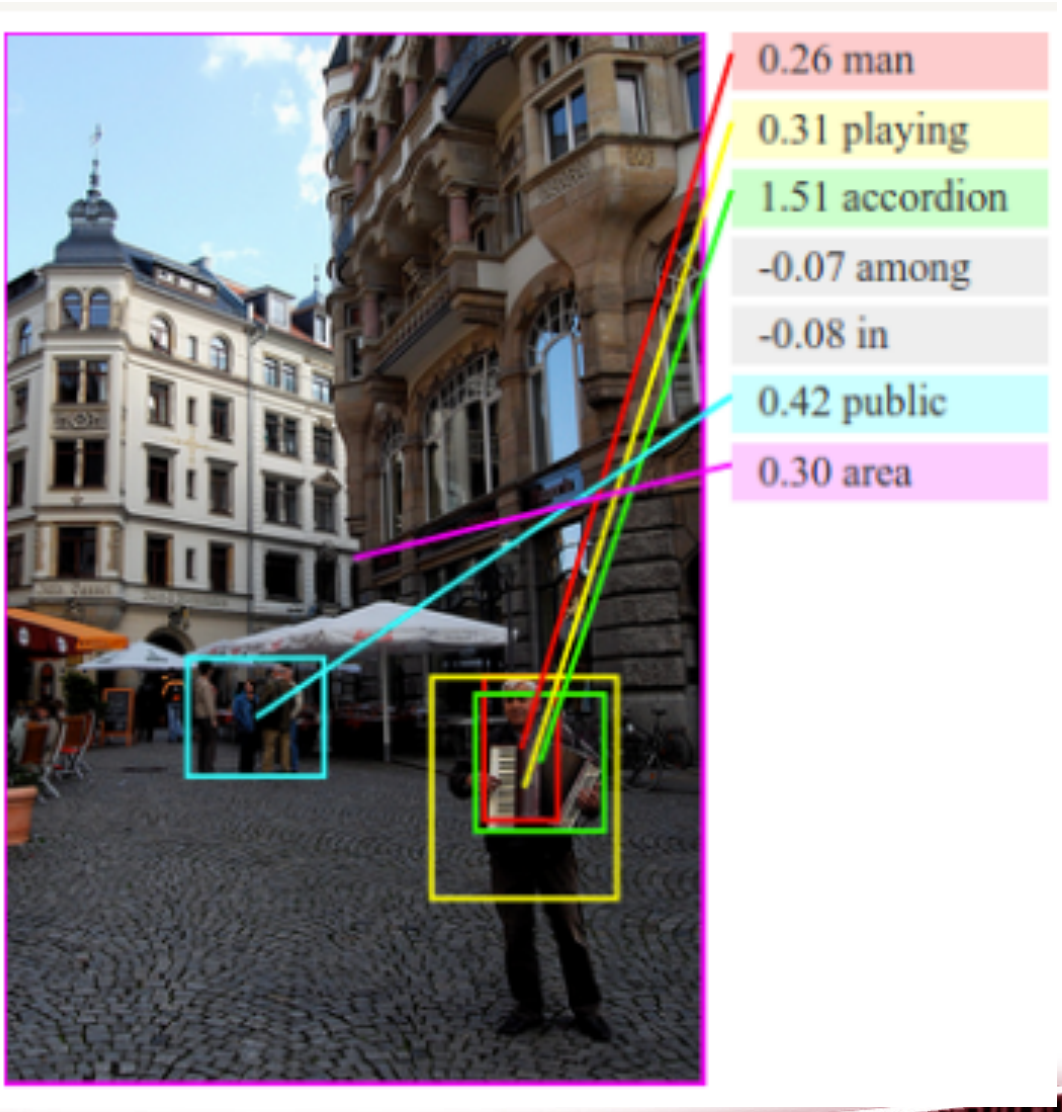
# RNN Generating Wikipedia

- A recurrent neural network was fed the entire Wikipedia corpus.
- It generates sample text mimicking its language.
- **Amazingly, similar to the previous example with Shakespeare, it does this one letter at a time!!**

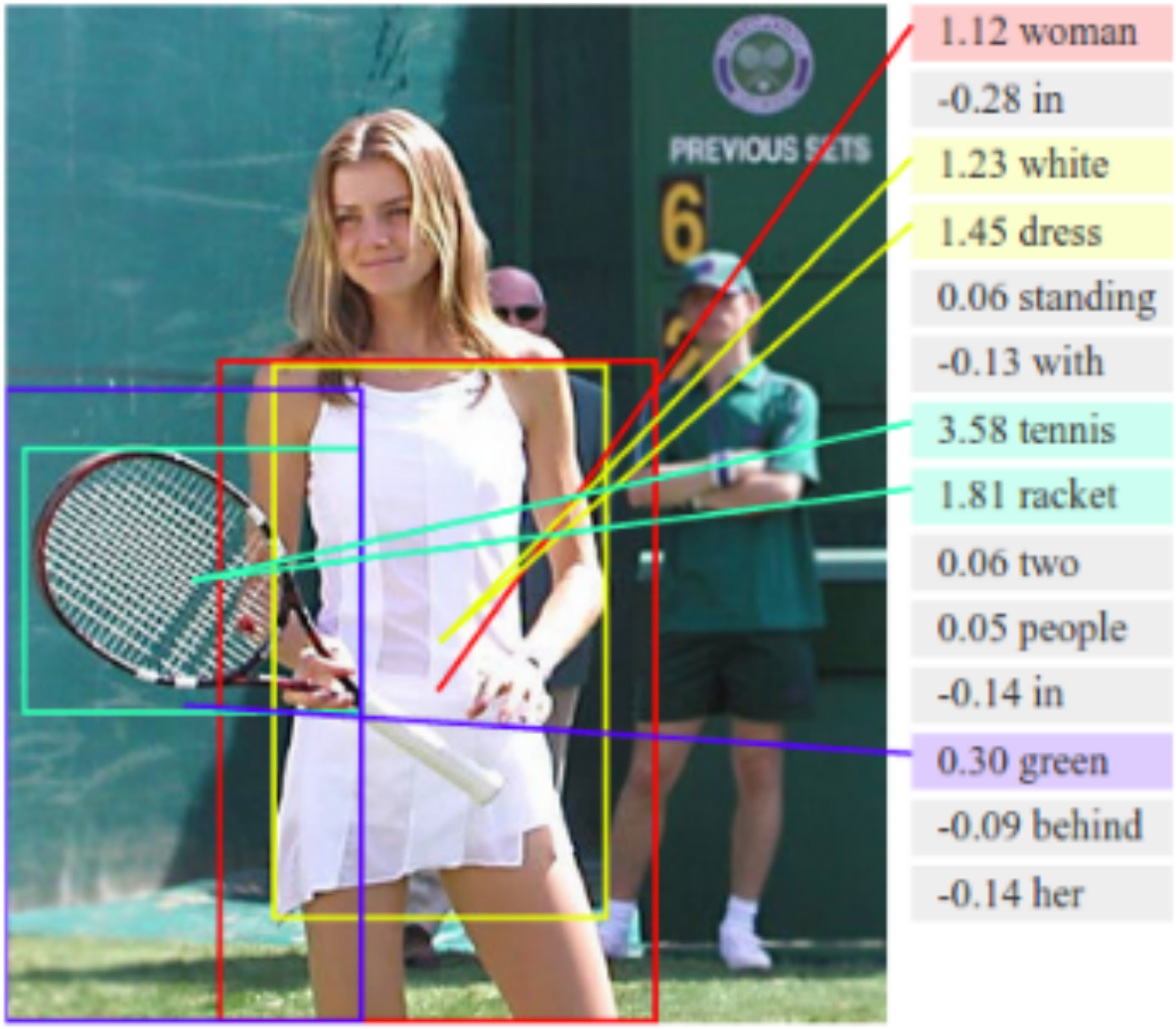
> Naturalism and decision for the majority of Arab countries' capitalide was grounded by the Irish language by [[John Clair]], [[An Imperial Japanese Revolt]], associated with Guangzham's sovereignty. His generals were the powerful ruler of the Portugal in the [[Protestant Immineners]], which could be said to be directly in Cantonese Communication, which followed a ceremony and set inspired prison, training.

# Identify the Task

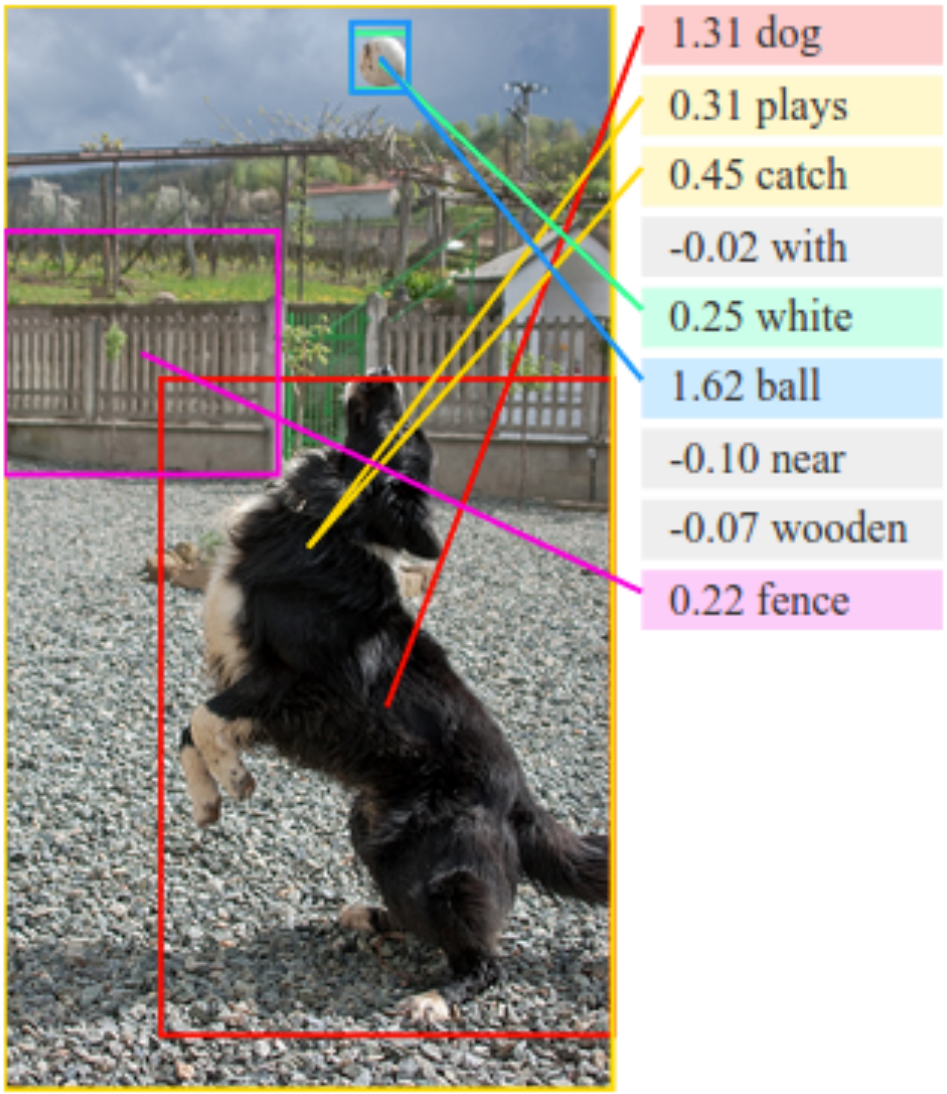
- Objects detection.



# Amazing!



# Another one



# Identify the Task

- image captioning.

"little girl is eating piece of cake."

"baseball player is throwing ball in game."

"woman is holding bunch of bananas."

"black cat is sitting on top of suitcase."

# Image Captioning

- Image captioning: objects recognition + natural language generation

"girl in pink dress is jumping in air."

"black and white dog jumps over bar."

# Image Captioning

"man in black shirt is playing guitar."

"construction worker in orange safety vest is working on road."

"man in blue wetsuit is surfing on wave."

# Image Captioning

- It doesn't always get it right! Uses some probability.

"a young boy is holding a baseball bat."

"a cat is sitting on a couch with a remote control."

"a woman holding a teddy bear in front of a mirror."

"a horse is standing in the middle of a road."

# Image Captioning

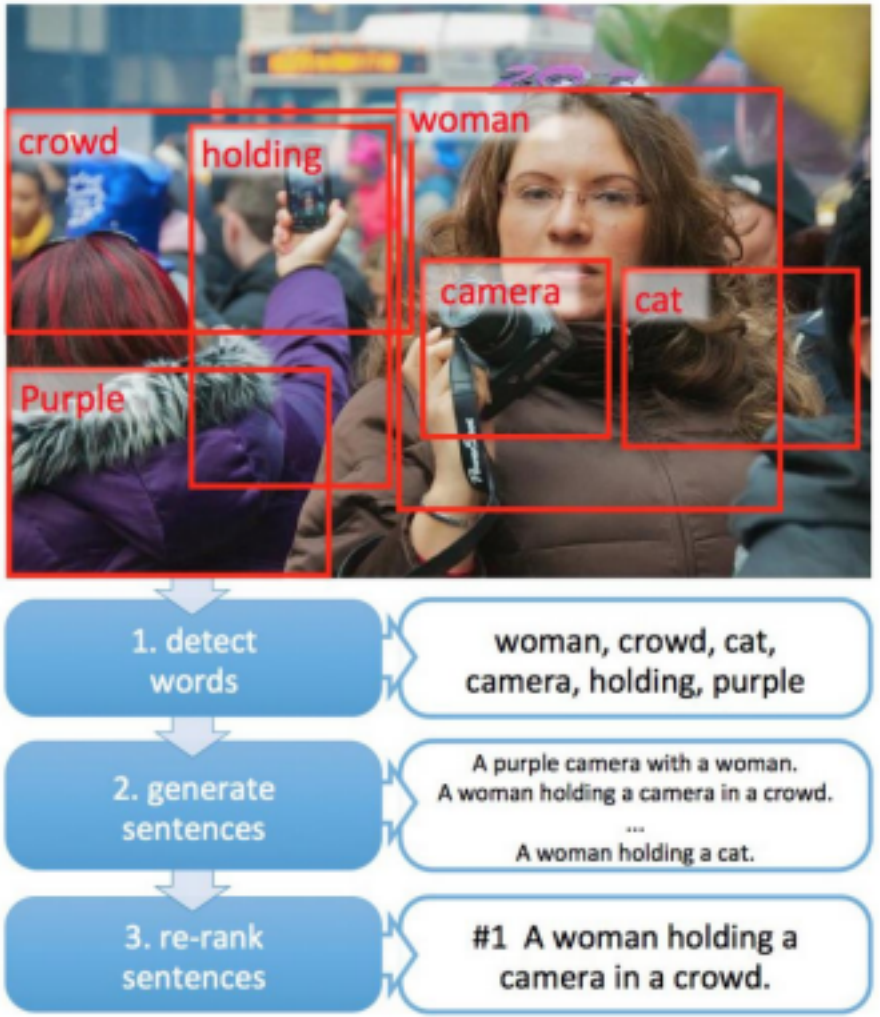
- It doesn't always get it right! Uses some probability.

"two young girls are playing with lego toy."

"boy is doing backflip on wakeboard."

# How does image captioning work?

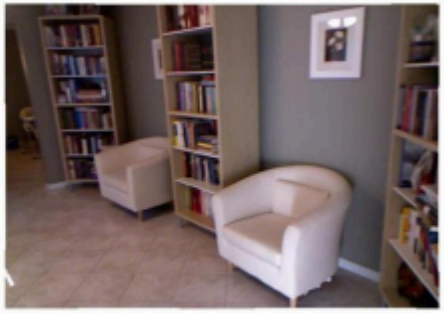
1) Detect objects.
2) Recognize and generate corresponding words.
3) Generate sentences and assign probabilities based on how likely that this is an English sentence based on:
   a) grammar
   b) syntax
   c) semantic meaning
4) Highest probability wins.



# Image Question Answering

COCOQA 33827
**What is the color of the cat?**
Ground truth: black
IMG+BOW: black (0.55)
2-VIS+LSTM: black (0.73)
BOW: gray (0.40)

COCOQA 33827a
**What is the color of the couch?**
Ground truth: red
IMG+BOW: red (0.65)
2-VIS+LSTM: black (0.44)
BOW: red (0.39)



DAQUAR 1522
**How many chairs are there?**
Ground truth: two
IMG+BOW: four (0.24)
2-VIS+BLSTM: one (0.29)
LSTM: four (0.19)

DAQUAR 1520
**How many shelves are there?**
Ground truth: three
IMG+BOW: three (0.25)
2-VIS+BLSTM: two (0.48)
LSTM: two (0.21)

COCOQA 14855
**Where are the ripe bananas sitting?**
Ground truth: basket
IMG+BOW: basket (0.97)
2-VIS+BLSTM: basket (0.58)
BOW: bowl (0.48)

COCOQA 14855a
**What are in the basket?**
Ground truth: bananas
IMG+BOW: bananas (0.98)
2-VIS+BLSTM: bananas (0.68)
BOW: bananas (0.14)

DAQUAR 585
**What is the object on the chair?**
Ground truth: pillow
IMG+BOW: clothes (0.37)
2-VIS+BLSTM: pillow (0.65)
LSTM: clothes (0.40)

DAQUAR 585a
**Where is the pillow found?**
Ground truth: chair
IMG+BOW: bed (0.13)
2-VIS+BLSTM: chair (0.17)
LSTM: cabinet (0.79)

# Have we solved the image recognition problem?

- We have a long way to go.
- Computers currently only "understand" images superficially.
- true understanding:
  - prior knowledge
  - social contexts, cues
  - natural language comprehension
  - depth perception, object permanence, periphery vision(challenges of autonomous driving)
- Consider the image on the next slide. Humans will immediately appreciate the humor. Computers won't for a while.

# Can computers “understand” this picture?

# More Challenges

- Moravec's Paradox: easy problems are "hard".
  - walking, motion planning, opening a door knob, getting out of a car.
  - https://youtu.be/1L0TKZQcUtA?t=4289
  - computers can learn to play a video game by only looking at pixels on the screen.
    - very hard for robots to do basic tasks
    - playing soccer is really hard!
- "Hard" problems are easy:
  - playing Chess, Go.
  - Google uses "deep learning", neural network to create a program that can play at the superhuman level Chess, Go and Shogi. (AlphaZero, 2017)

# Homework

1)Read and reread these lecture notes.

2)Watch(Required):

a) Machine Learning(PBS Studios Crash Course in CS)
https://www.youtube.com/watch?v=z-EtmaFJieY
b) Natural Language Processing(PBS Studios Crash Course in CS)
https://www.youtube.com/watch?v=fOvTtapxa9c

# References

1)Justin Johnson, Andrej Karpathy. CS231N. Convolutional Neural Network for Visual Recognition.
http://cs231n.stanford.edu/

2)Lex Fridman, Introduction to Deep Learning and Self-Driving Cars. https://www.youtube.com/watch?v=1L0TKZQcUtA