

The background is a dark blue gradient with a subtle pattern of white dots. Overlaid on the left side are several concentric circles and arcs in a lighter blue color. Some of these arcs have degree markings, such as 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, and 260. There are also small white arrows pointing in various directions, suggesting a sense of motion or data flow.

# LIVE-EYE PROJECT PROPOSAL

BIG DATA ECOSYSTEMS

BY : RASHMI PALIWAL

SOHAM TALUKDAR

TARUN AKIRALA

# LARGE SCALE RECOGNITION

## Data Diversity

Data needs to have sufficient variety to represent the real world.



## Data Annotation

Data needs to contain enough information for algorithms to learn more.



## Algorithm Complexity

Algorithms need to be powerful enough to learn from the data.



## Algorithm Speed

Algorithms need to be fast enough to process the large data.



# OBJECTIVE

Aim : Object detection and classification (out of 80 predefined object classes) in the given set of images using deep learning methods.

- Object Segmentation
- Recognition in Context
- Multiple objects per image



## MS COCO 2015

Early 2015

- 80-100 object categories
- 330k images
- 2M+ instances (700k people)
- Every instance segmented

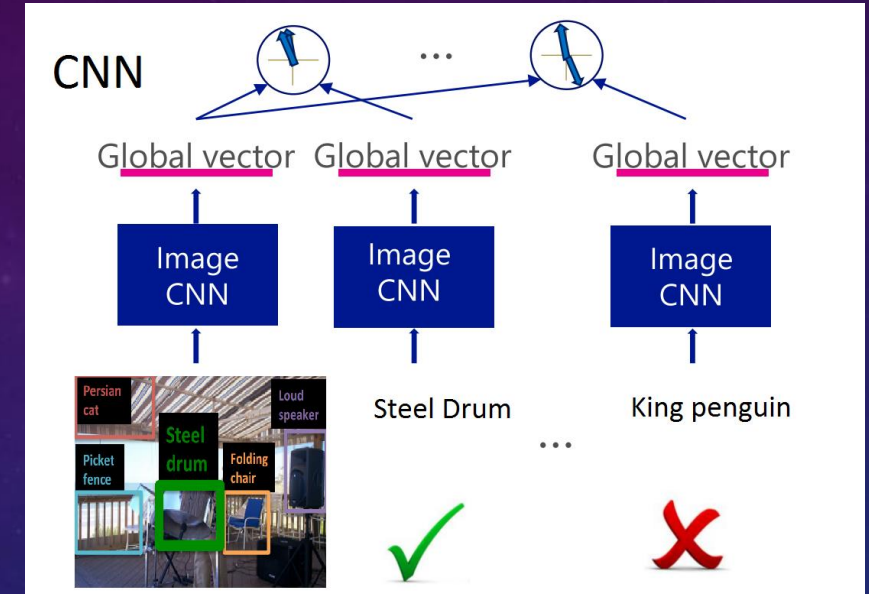
<http://mscoco.org>



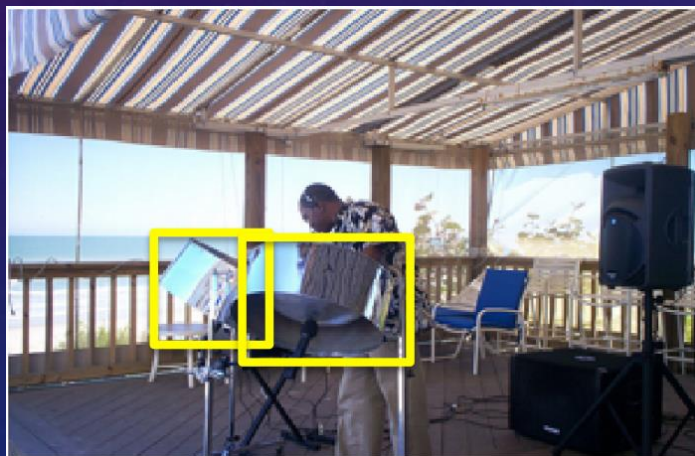


# DEEP LEARNING - USING CNN?

Convolutional neural network makes sure that the incorrect local detections do not propagate in further layers.



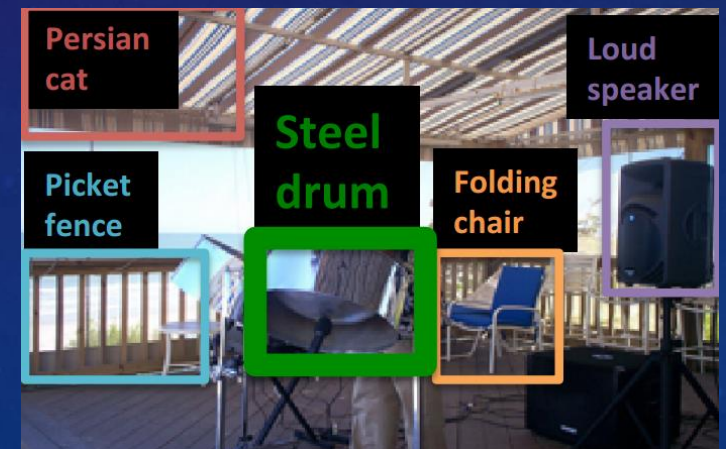
Raw Image



Prediction i



Prediction j



# TOOLS EMPLOYED

- Programming Language
  - Python
- Framework
  - Spark
- Running Environment
  - AWS/Google cloud/ Super computer
- Deep Learning Framework
  - Caffe/Torch/Theanu/Tensorflow



# DATA SET

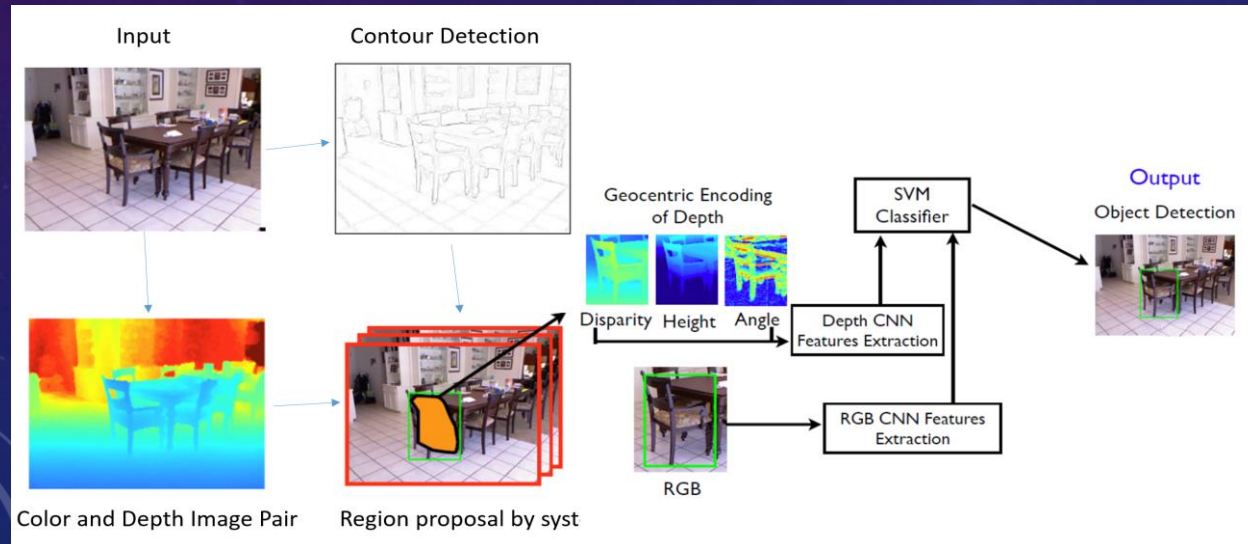
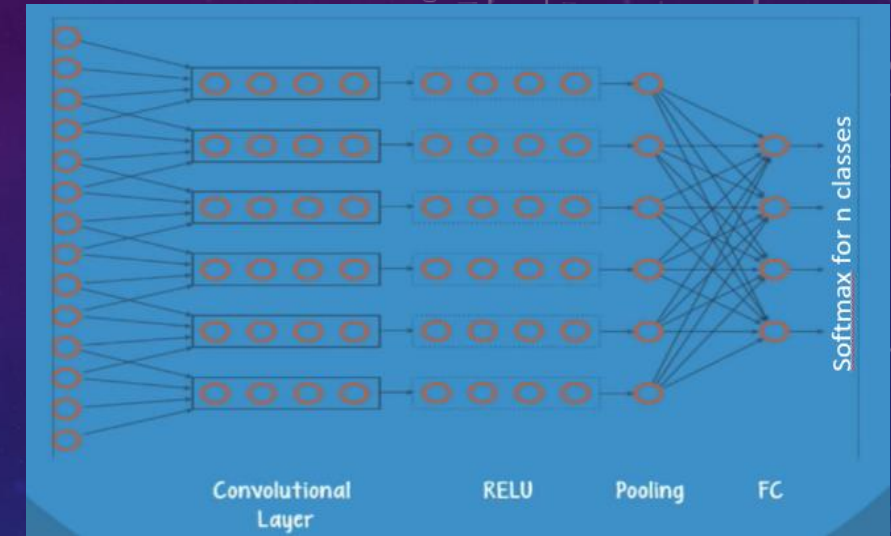
- The training, validation and test sets contain more than 200,000 images and 80 object categories.
- Training set - 13GB/80K
- Validation set-6.2GB/40k
- Testing Set1-6.2GB/40k
- Testing Set2-12.4GB/80k

```
annotation{  
  "id" : int,  
  "image_id" : int,  
  "category_id" : int,  
  "segmentation" : RLE or [polygon],  
  "area" : float,  
  "bbox" : [x,y,width,height],  
}
```



# ALGORITHM AND ARCHITECTURE

- use any of the methods, preferably Selective search to generate a sample of (in our case, 1000) category-independent sample regions of the input image.
- Extract a fixed length feature vector from each sample region using CNNs. Features can be extracted by forward-propagating the image through a number of convoluted and fully connected layers.
- Classify each region with category specific linear SVMs.



# CHALLENGES

- Diversity of Images to be identified. (Size, Texture)
- Identify the various images in any setting. Are we done?
- Classification together with localization may change the perception of the image.
- Man made deformable objects are difficult to train.





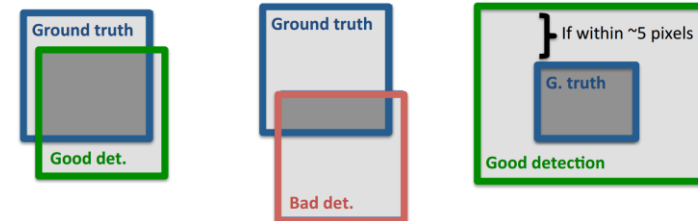
# ERROR RATE CONTROL

- Algorithm outputs a list of bounding box detections with confidences.
- A detection is considered correct if overlap with ground truth is big enough.
- Evaluated by average precision per class.
- Global detection will overwrite the incorrect local detections.
- Most object categories will be from the predefined categories (which is 80 in the dataset considered).
- All instances of all target object classes expected to be localized in all test images.

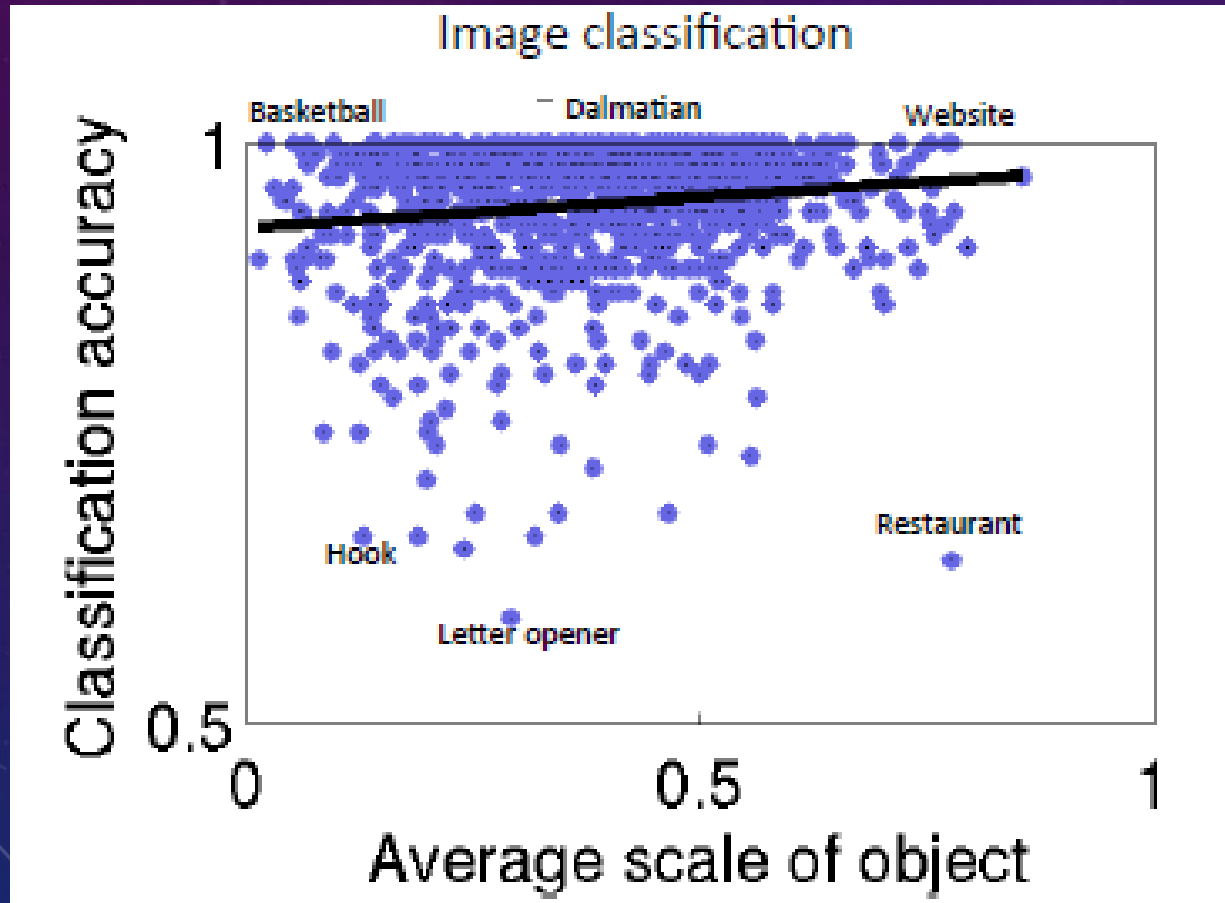
## Threshold for Correct Detection

$$\frac{\text{Intersection}}{\text{Union}} \geq \min\left(0.5, \frac{w_{gt} h_{gt}}{(w_{gt}+10)(h_{gt}+10)}\right)$$

*Only matters for small boxes (< ~25 pixels), which is about 5.5% of cases*



# SMALLER OBJECTS ARE NOT NECESSARILY HARD TO CLASSIFY.



- Each Dot is an Object Class
- X-axis: Average fraction of image area occupied by an instance of that class on a validation set.
- Y-axis : Highest accuracy achieved by any method.

WHEN A USER TAKES A PHOTO,  
THE APP SHOULD CHECK WHETHER  
THEY'RE IN A NATIONAL PARK...

SURE, EASY GIS LOOKUP.  
GIMME A FEW HOURS.

... AND CHECK WHETHER  
THE PHOTO IS OF A BIRD.

I'LL NEED A RESEARCH  
TEAM AND FIVE YEARS.



IN CS, IT CAN BE HARD TO EXPLAIN  
THE DIFFERENCE BETWEEN THE EASY  
AND THE VIRTUALLY IMPOSSIBLE.

xkcd: Tasks

"The Virtually Impossible"

# THANK YOU!!