SJTU AI Circuit Design Contest 2019

Yongfu Li
Dept. of Micro/Nano Electronics
Shanghai Jiao Tong University

Outline

Lesson #1 (12:30 – 1:30)

- Liu Ting
- Neural Network and Model Compression

Lesson #2 (1:30 – 5:00)

- Sun Sizhen, 孙思侃
- Digilent Pynq FPGA

Lesson #3 (5:00 – 6:00)

- Li Yongfu
- Design Contest

Digilent 培训日程

- PYNQ项目及PYNQ-Z1开发板介绍（~30mins）
- 第一部分：PYNQ开发环境及基础外设（SWs，LEDs）实验（~45mins）
- 第二部分：PYNQ Overlay介绍及Pmod OLED实验（~30 mins）
- 第三部分：PYNQ逻辑工具简介及实验（~45mins）
- 第四部分：PYNQ自定义Overlay设计实验（~60mins）

Important!!!

与本次比赛获奖相关的其他事项包括：

a. 获二等奖（含）及以上的上海交通大学参赛本科生队伍，将直接获得次年参加全国大学生集成电路创新创业大赛的学校推荐名额；

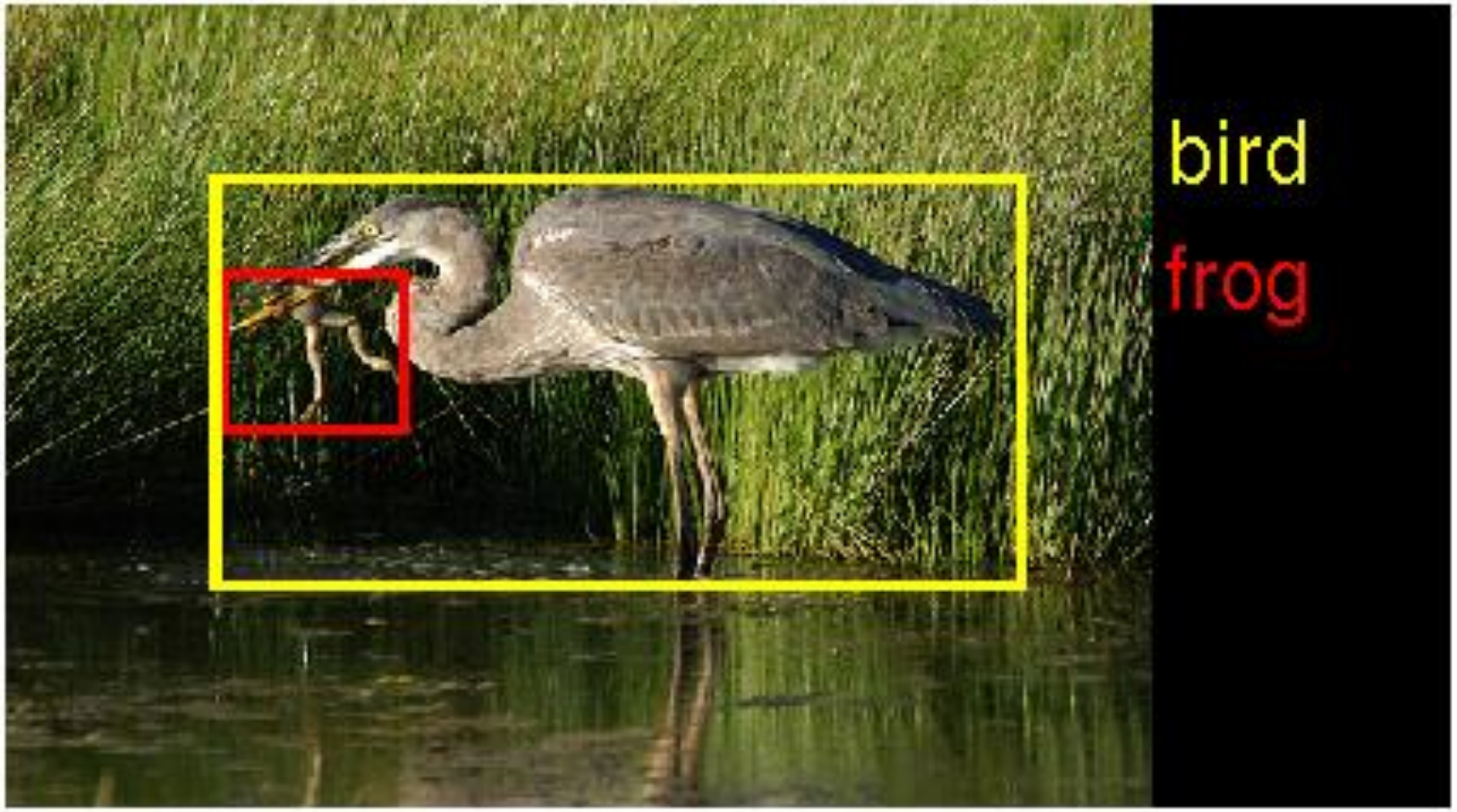
b. 获二等奖（含）及以上的队伍，可申请参加国际低功耗图像识别挑战赛相应资助；

c. 对于所有完成比赛的队伍均可获得对应项素拓分。

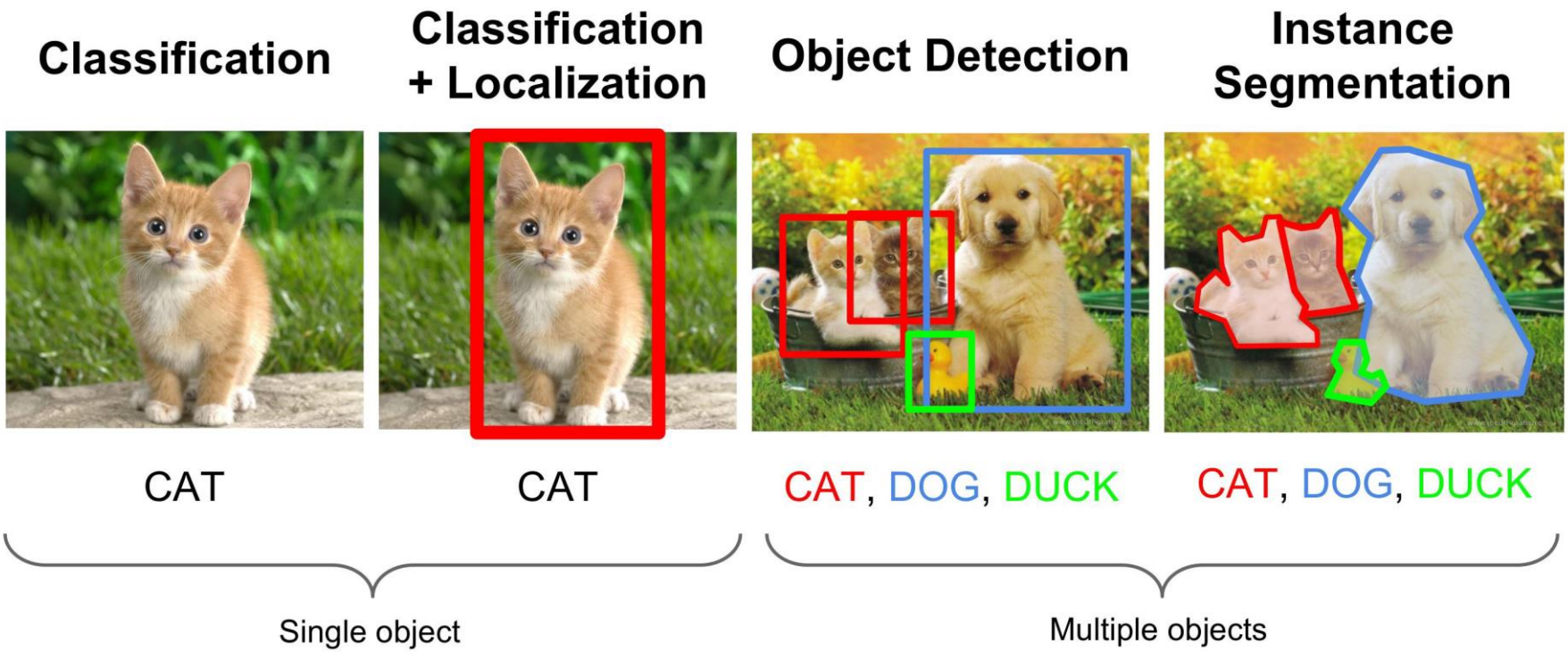
Visual Dataset

Classification & Object Detection

- Example ILSVRC2014 images:



Classification & Object Detection



- Image classification: classify images into a single category
- Object detection: Identify multiple objects in a single image

Visual Datasets

The PASCAL Visual Object Classes Homepage
host.robots.ox.ac.uk › pascal › VOC ▾

COCO dataset
cocodataset.org ▾

ImageNet
www.image-net.org ▾
ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. Currently we have an average of over five hundred images per node.

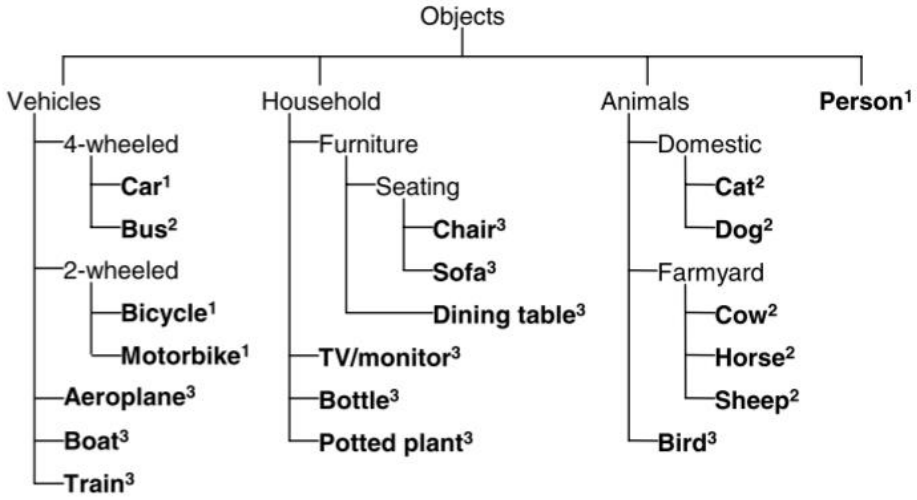
ImageNet Large Scale Visual Recognition Competition ...
www.image-net.org › challenges › LSVRC ▾
The ImageNet Large Scale Visual Recognition Challenge (**ILSVRC**) evaluates ... When reporting results of the challenges or using the **datasets**, please cite:.

K. Gauen *et al.*, "Comparison of Visual Datasets for Machine Learning," *2017 IEEE International Conference on Information Reuse and Integration (IRI)*, San Diego, CA, 2017, pp. 346-355.

PASCAL Visual Object Classification

The PASCAL Visual Object Classes Homepage
host.robots.ox.ac.uk › pascal › VOC ▾

- 20 categories: reference dataset in the **object detection problem**
- 8 different challenges spanning from 2005 to 2012
- 10 000 images for training and validation containing bounding boxes with objects



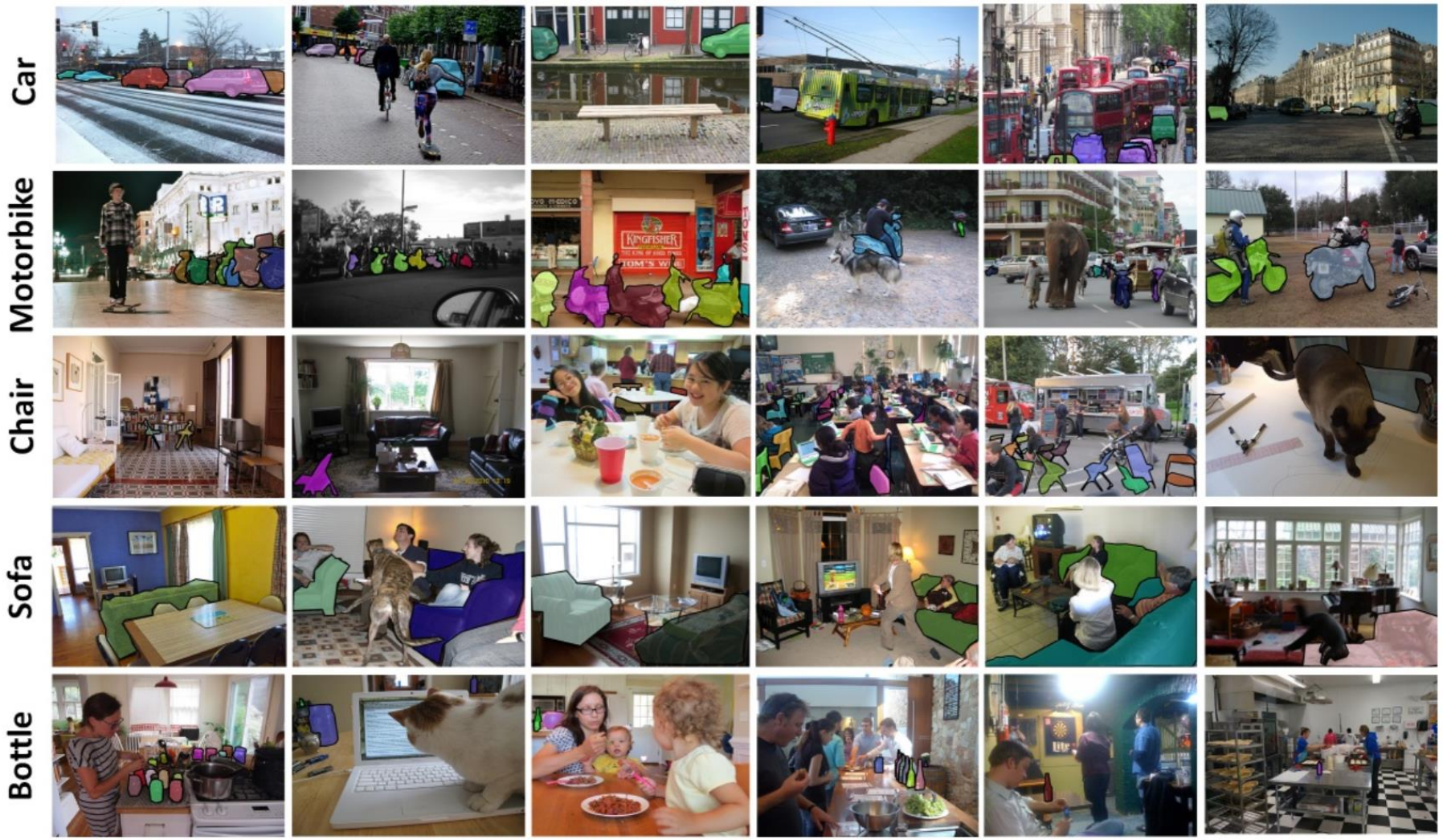
COCO Object Detection Dataset

COCO dataset
cocodataset.org

- Microsoft in 2015
- **caption generation, object detection, key point detection and object segmentation.**

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 330K images (>200K labeled)
- 1.5 million object instances
- 80 object categories
- 91 stuff categories
- 5 captions per image
- 250,000 people with keypoints

COCO Object Detection Dataset



ImageNet Object Detection Dataset

- First release in 2013 with bounding boxes.
- Around 500 000 images with 200 categories.
- Rarely used because the size and required computational power for training.
- The high number of classes complicates the object recognition task.

International Design Contest

Low-Power Image Recognition Challenge

Low-Power Computer Vision Workshop 2019

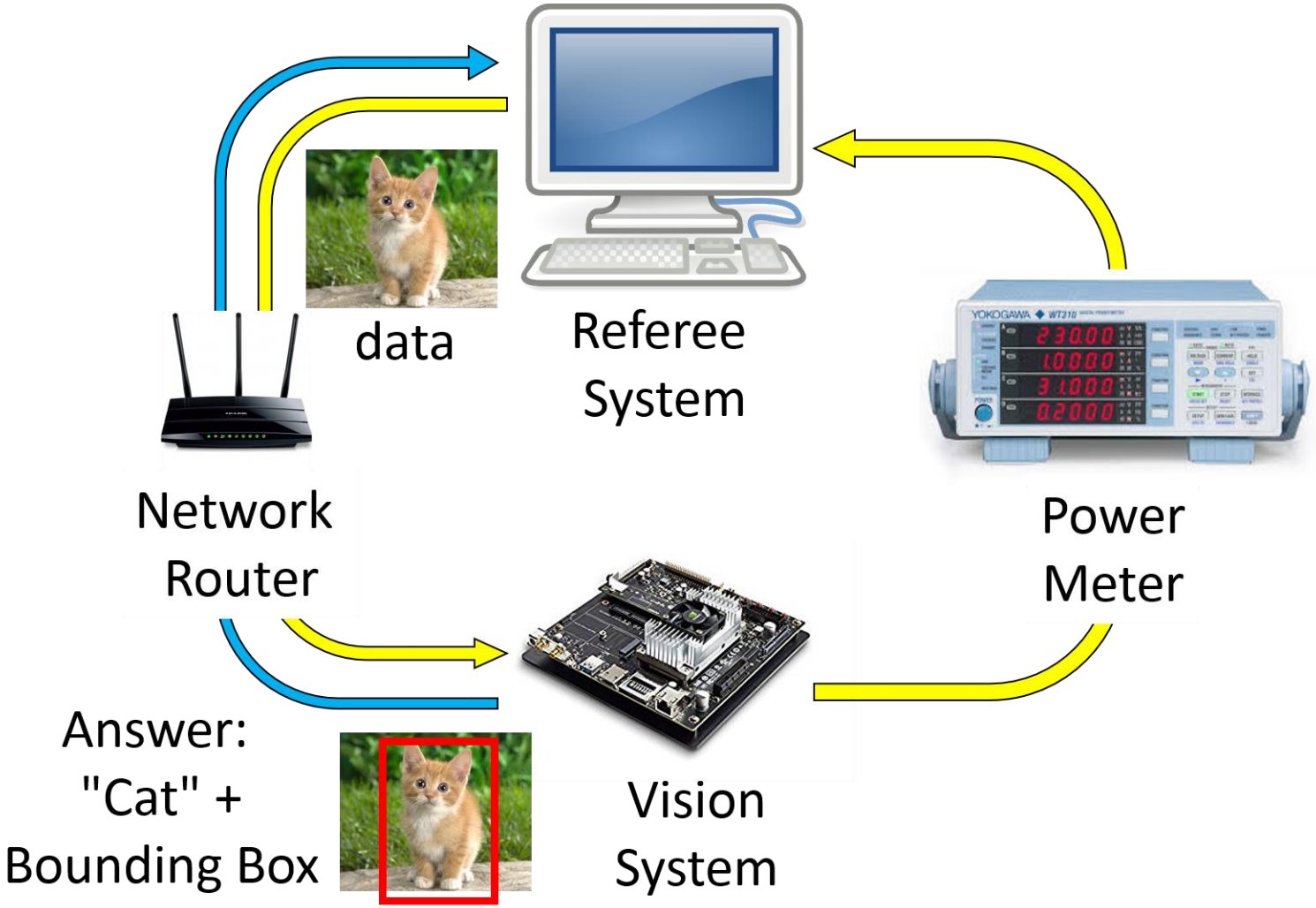
2019 ICCV Workshop
Seoul, Republic of Korea
Monday, 28 October 2019

Read about recent LPIRC competitions

- Started in 2015
- Track 1: TfLite Model on Mobile Phones (Software)
- Track 2: Caffe 2 and TX 2 (Hardware)
- Track 3: Onsite, No Restriction

Low-Power Image Recognition Challenge

Track 2: Caffe 2 and TX 2 (Hardware)



DAC System Design Contest

The 2019 System Design Contest features embedded system implementation of neural network based object detection for drones. Contestants will receive training dataset provided by our industry sponsor DJI, and a hidden dataset will be used to evaluate the performance of the designs in terms of accuracy and power. Contestants will compete in two different categories: FPGA and GPU, and grand cash awards will be given to the top three teams in each category. In addition, our industry sponsor Xilinx and Nvidia will provide a limited number of teams successfully registered with a free design kit (on a first-come-first-served basis). The award ceremony will be held at 2019 IEEE/ACM Design Automation Conference.

IEEE Videos

http://ieeetv.ieee.org

- Designing Efficient On-Device AI
- Classifying attention in Pivotal Response Treatment Videos
- Visual Wake Words Challenge
- Quantization Without Fine-Tuning
- Co-Design of Algorithms & Hardware for DNNs
- Deep Learning & Machine Learning Inference
- The Art of MobileNet Design

SJTU Design Contest

Github Content

SJTU-AI-Circuit-Design-Contest-2019

https://github.com/yongfu-li/SJTU-AI-Circuit-Design-Contest-2019

- Demo ← Example of pynq code
- dataset ← Training & Testing Dataset
- docs
- images ← Example of different category
- README.md
- contest-info.pdf

Our SJTU Challenge!!!

Classification + Localization

CAT

Contest Objectives

- **Accuracy:** Identify and localize the object
- **Energy:** Power measurement
- **Speed:** Frame rate

Reduced PASCAL Visual Object Dataset

The PASCAL Visual Object Classes Homepage

host.robots.ox.ac.uk › pascal › VOC ▾

- Image numbers for each category:
  - Aeroplane: 508
  - Car: 600
  - Sofa: 485
  - TV monitor: 400

- Image size:
  - width – 320 pixels
  - Height – 176 pixels

Dataset

- Train/test list's construction(label information):

  [[[image path,[label information]],[...],[...]],[...],[...]]

- Details:
  - Label information:
    - Contains 4 numbers: $\frac{\text{xmin}}{width}, \frac{\text{ymin}}{height}, \frac{\text{xm}ax}{width}, \frac{ym}{height}$ represent the ratio of bounding box location to image size.
  - Image path: "path to the image file".
  - [] separate the different image information.
  - [] separate the different categories.
  - [] contain the whole dataset.

Goal for computer visions

Winners = Low energy + High accuracy + High speed

- **Power:** 0.1 Watt
- **Speed:** > 100 images at high resolutions (>12MP) per sec.
- **Accuracy:** > 99.99%
- **Objects:** > 100 objects in each image from 1,000 different categories

SJTU Design Contest Evaluation

Preliminary

Submission format:

- Word Report

Content:

- Python neural network model
- Accuracy score

Final

Submission format:

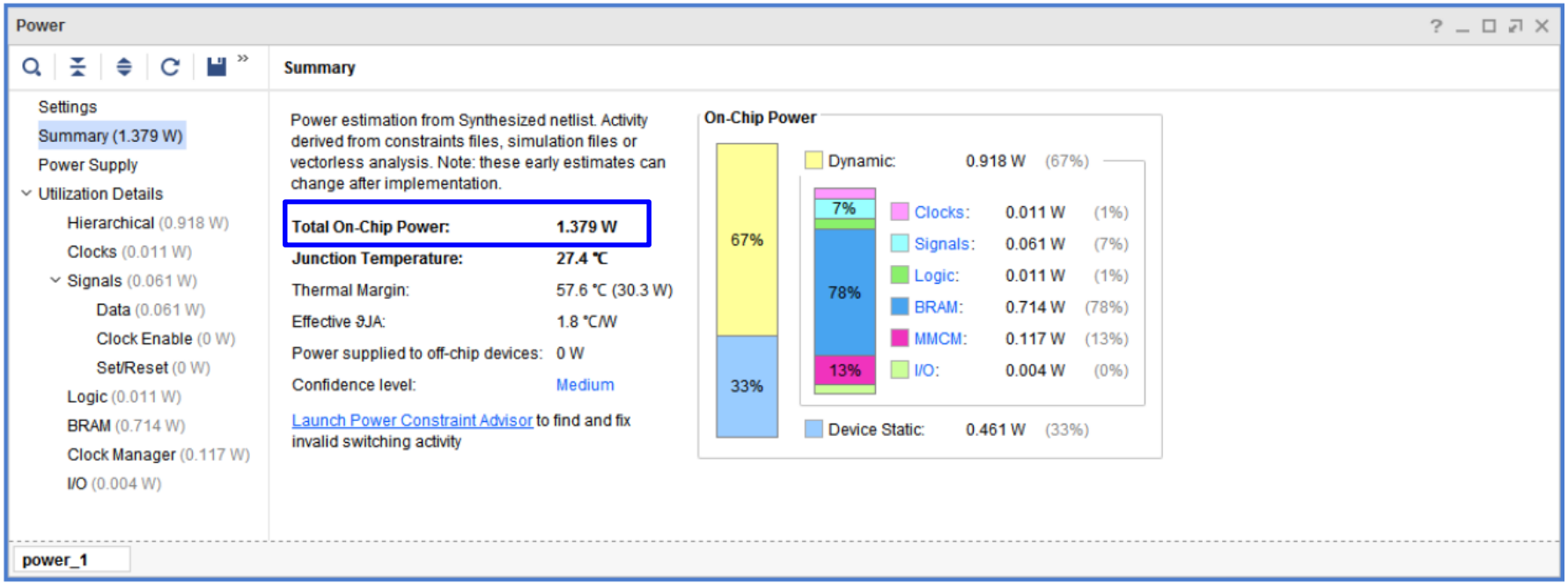
- Word Report
- Live Demonstration
- QnA session

Content:

- FPGA implementation
- Min. Requirement:
  - Energy, Frame Rate
- Benchmark Requirement:
  - Accuracy

Evaluating the Power Consumption Sim.

- Power report can be generated by vivado design suite as follow:



User guide: Github /docs/ug997-vivado-power-analysis-optimization-tutorial.pdf

Evaluating the Power Consumption Meas.

LPIRC Contest Standard (Bonus for SJTU contest)

- Yokogawa WT310 Digital Power Meter, and the scores are reported in watt-hours.
- **mean Average Precision (mAP)** to measure the accuracy of object detection methods
- Our final evaluation metric is the **mean average precision** divided by the **total energy consumption** in the **10 minute interval.**

Evaluating the Throughput

IEEE DAC Design Contest Standard

- Min. Speed requirement is 20FPS

Evaluating the Accuracy

IEEE DAC Design Contest Standard

- Intersection over Union (IoU)

$$IoU_i = \frac{Area\ of\ Overlap}{Area\ of\ Union} = \frac{DetectionResult \cap GroundTruth}{DetectionResult \cup GroundTruth}$$

- Detection Results should be in following format:

[image name,

$$[\frac{xmin_{detection}}{width}, \frac{ymin_{detection}}{height}, \frac{xmax_{detection}}{width}, \frac{ymax_{detection}}{height}]],$$

[…],[…],…

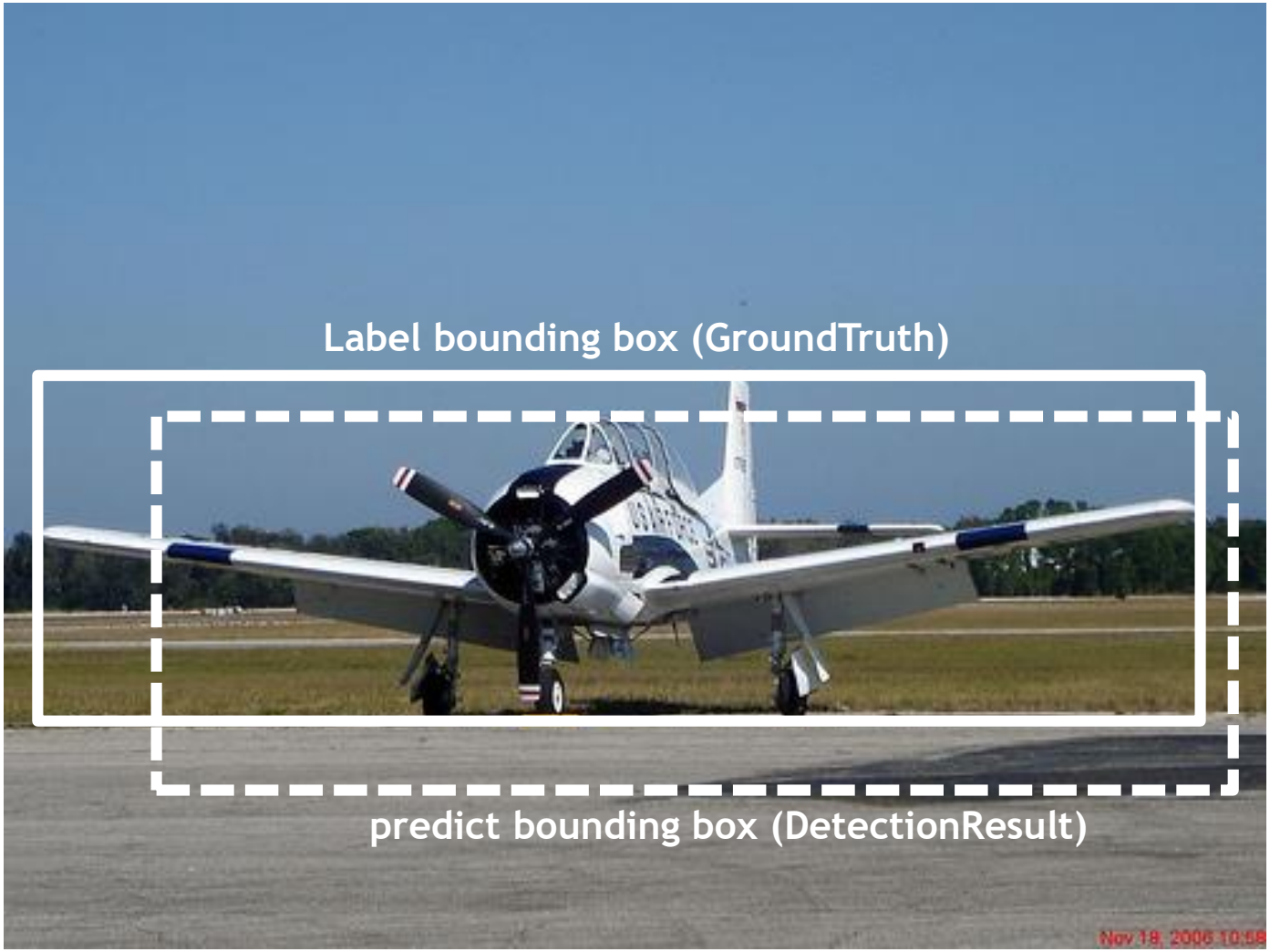
- Scores of team *i*:

$$\boldsymbol{score_{accuracy} = \frac{\sum_k^K IoU_k}{K}}, \qquad \boldsymbol{K = total\ number\ of\ image}$$

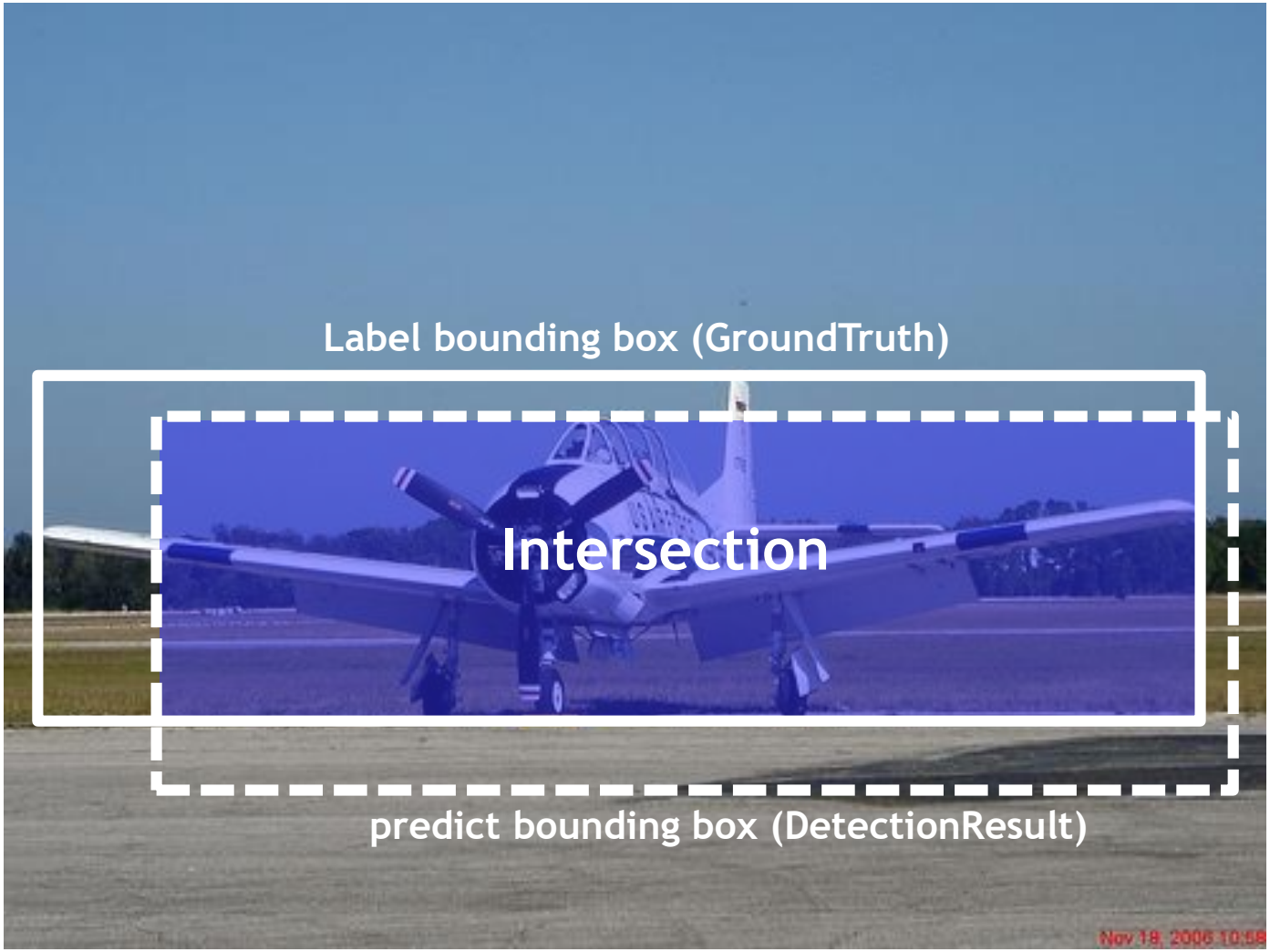
Example


xmin
xmax
Label bounding box(GroundTruth)
ymax
ymin
height
width

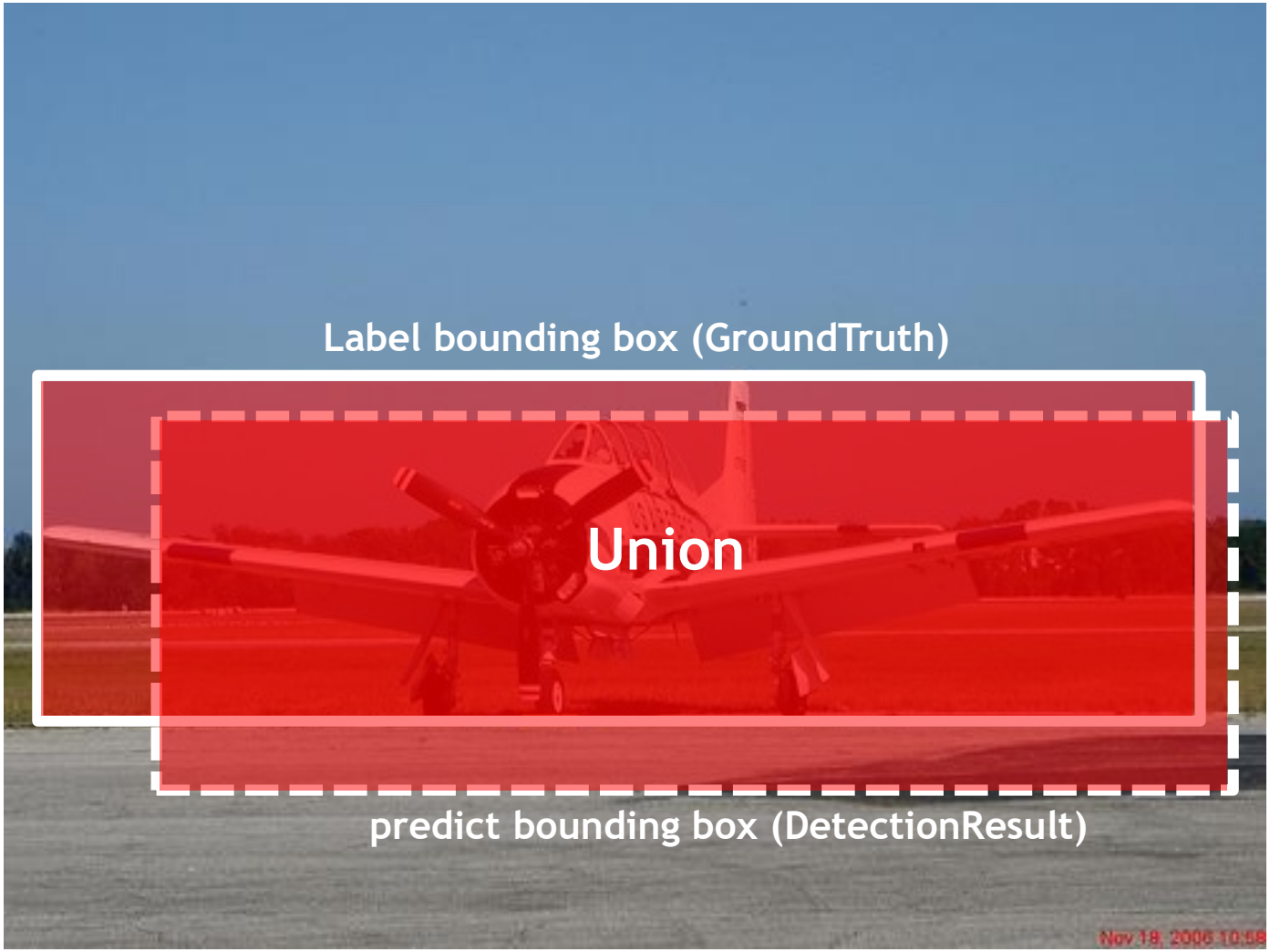
Evaluation (Label)



Evaluation (Intersection)



Evaluation (Union)



SJTU Contest Evaluation [Preliminary]

(1) Accuracy

$$score_{accuracy} = \frac{\sum_{k}^{K} IoU_k}{K}, \qquad K = total\ number\ of\ image$$

(2) Memory Size??

- $M_{avg} = \frac{\sum_{i=1}^{I} M_i}{I}$ (I = total number of teams)

$$score_{memory} = \max\left\{0, 0.2 \times log_2 \frac{M_{avg}}{M_i}\right\}$$

Total scores

$$score_{total} = score_{accuracy} \times (score_{memory} + 1)$$

SJTU Contest Evaluation [Final]

(1) Power consumption

- $E_{avg} = \frac{\sum_{i=1}^{I} E_i}{I}$ (I = total number of teams)

$$score_{energy} = \max\left\{0, 0.2 \times log_2 \frac{E_{avg}}{E_i}\right\}$$

(2) Throughput

- $T_{avg} = \frac{\sum_{i=1}^{I} T_i}{I}$ (I = total number of teams)

$$score_{throughput} = \max\left\{0, 0.2 \times log_2 \frac{T_i}{T_{avg}}\right\}$$

SJTU Contest Evaluation [Final]

(3) Accuracy

$$score_{accuracy} = \frac{\sum_{k}^{K} IoU_k}{K}, \qquad K = total\ number\ of\ image$$

Total scores

$$score_{total} = score_{accuracy} \times (score_{energy} + 1) \times (score_{throughput} + 1)$$

Neural Network Performance

Model	mAP	FPS	Real Time speed
Fast YOLO	52.7%	**155**	Yes
YOLO	**63.4%**	45	Yes
YOLO VGG-16	66.4%	21	No
Fast R-CNN	70.0%	0.5	No
Faster R-CNN VGG-16	73.2%	7	No
Faster R-CNN ZF	62.1%	18	No

Neural Network Compression

To reduce the size, energy consumption, and overtraining of deep neural networks

Methods

- **Parameter pruning:** selective removal of weights based on a particular ranking
- **Low-rank factorization:** using matrix/tensor decomposition to estimate informative parameters
- **Transferred/compact convolutional filters:** special structural convolutional filters to save parameters
- **knowledge distillation:** training a compact network with distilled knowledge of a large network