Title: How to train your d…detector, Computer vision and object detection

Speaker: Dr. Gilberto Ochoa-Ruiz, ITESM Guadalajara

Abstract: Object detection, as of one the most fundamental and challenging problems in computer vision, has received great attention in recent years. Its development in the past two decades can be regarded as an epitome of computer vision history. In this tutorial we analyze object detection in the light of its technical evolution, spanning over a quarter-century's time. Several topics will be covered, including the milestones detectors in history, detection datasets, metrics, fundamental building blocks of the detection system, speed up techniques, and the recent state of the art detection methods. We will provide a thorough survey of recent advances of object detection based on deep learning architectures and discuss some important detection applications, such as pedestrian detection, face detection, text detection and carry out and in-detph analysis of their challenges as well as technical improvements in recent years. We conclude the tutorial with the implementation of a couple of modern object detectors using Google Collab and Tensorflow on a custom dataset.

Outline (Duration 3 hours)

1. Object Detection in the last 20 years: an overview (30 mins)
   1.1. Traditional object detectors
   1.2. CNN-based two stage detectors
   1.3. CNN-based one stage detectors
2. Object detection datasets and metrics (15 mins)
   2.1. Datasets: Pascal VOC, ILSVRC, MS-COCO, Open Images
   2.2. Metrics: FPPW, AP, IoU, mAP
3. Technical Evolution in Object Detection (15 mins mins)
   3.1. Early naïf approaches and multi-scale object detection
   3.2. Detection with object proposals
   3.3. Deep learning-based object detection
4. Technical progresses on deep learning-based object detection (30 mins)
   4.1. Evolution on box regression, context priming, non-maximum suppression
   4.2. Developments for speeding up object detection
   4.3. Improvements on feature extraction
   4.4. Beyond the sliding window, improvements on localization
   4.5. Modern examples: text detection on the wild, traffic signs, among others
5. Introduction to Object Detection using Google Collab and Tensorflow (1 hour)
   5.1. Working with Jupyter Notebooks, Tensorflow and Python utilities
   5.2. What is Google Collab? Setting the environment
   5.3. Importing the dataset to Google Cloud
   5.4. Importing a pretrained model (YOLO, MobileNet, RetinaNet)
   5.5. Training the pretrained model using a custom dataset
   5.6. Inference
6. Creating you custom generic object detector (1 hour)
   6.1. Annotating the dataset
   6.2. Training the model of choice (YOLO, RetinaNet, Focal Loss)
   6.3. Testing the model of choice
   6.4. Evaluating the model

Target Audience: Students or professionals interested in image processing and computer vision applications, especially those in which object detection is a fundamental block (object tracking, text detection and retrieval, scene text recognition for augmented reality, among others). The attendants are expected to have some image processing and computer vision notions, but the tutorial is intended to be accessible for a broad audience, so even non-experts can leverage the power of modern object detection algorithms for creating working applications. The tutorial can serve as a means for practitioners on the domain to update their knowledge in this particular area.

Bio: Researcher and lecturer in several post-graduate programs accredited by the CONACYT PNPC program, geared towards Computer Science and Communication and Information Technologies Since Sept. 2019, he is part of the academic staff of Instituto Tecnológico Estudios Superiores de Monterrey (ITESM), in its Guadalajara campus, affiliated to the Advanced Artificial Intelligence Group

He coursed a bachelor's degree in Electronics and Communications Engineering by the Universidad de Guadalajara and a specialty degree in Digital and SoC Design at the CINVESTAV Guadalajara. Afterwards, he obtained a master's degree in computer Vision and Robotics by Heriot-Watt-University (Edinburgh, UK) and a PhD in Electronic Imaging and Computer Vision by the Université de Bourgogne (France).

His research areas and interests are focused on investigating and implementing novel HW/SW architectures for a variety of applications. He is particularly interested in using reconfigurable technologies and systems (FPGAs and ACAPs) for designing real-time, heterogenous, edge computing applications. More recently, he has focused my research to the application of reconfigurable devices in robotics, computer vision and IoT applications (artificial intelligence at the edge) for applications in smart cities. In this sense, he currently carries out research on the implementation of end-to-end deep learning-based computer vision applications on constrained devices, making use of deep learning models amenable for customized embedded and mobile deployments.

Contact address:

Dr. Gilberto Ochoa Ruiz
ITESM, Campus Guadalajara
Escuela de Ingeniería y Ciencias
Avenida General Ramón Corona 2514
Nuevo México, 45138 Zapopan, Jal, Mexico
Tel.: +52 (1) 33 3669 3000}
Email: gilberto.ochoa@tec.mx

Technical Requirements:

The tutorial will be executed on a Jupyter Notebook mounted on Google Collab and Cloud. The attends need only to bring a laptop, Tensorflow, OpenCV, Python 3.6 and the associated libraries (Numpy, Pandas, Matplotlib) are accessible through the notebook. Google Colab provides access to GPUs and TPUs which are enough for the purposes of this tutorial.