2018 IEEE International Conference on Big Data (Big Data)

# YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers

Rachel Huang*
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, United States
rachuang22@gmail.com

Jonathan Pedoeem*
Electrical Engineering
The Cooper Union
New York, United States
pedoeem@cooper.edu

Cuixian Chen
Mathematics and Statistics
UNC Wilmington
North Carolina, United States
chenc@uncw.edu

***Abstract*—This paper focuses on YOLO-LITE, a real-time object detection model developed to run on portable devices such as a laptop or cellphone lacking a Graphics Processing Unit (GPU). The model was first trained on the PASCAL VOC dataset then on the COCO dataset, achieving a mAP of 33.81% and 12.26% respectively. YOLO-LITE runs at about 21 FPS on a non-GPU computer and 10 FPS after implemented onto a website with only 7 layers and 482 million FLOPS. This speed is 3.8× faster than the fastest state of art model, SSD MobilenetvI. Based on the original object detection algorithm YOLOV2, YOLO-LITE was designed to create a smaller, faster, and more efficient model increasing the accessibility of real-time object detection to a variety of devices.**

***Index Terms*—object detection; YOLO; neural networks; deep learning; non-GPU; mobile**

## I. Introduction

In recent years, object detection has become a significant field of computer vision. The goal of object detection is to detect and classify objects leading to many specialized fields and applications such as face detection and face recognition. Vision is not only the ability to see a picture in ones head but also the ability to understand and infer from the image that is seen. The ability to replicate vision in computers is necessary to progress day to day technology. Object detection addresses this issue by predicting the location of objects through bounding boxes while simultaneously classifying each object in a given image [1], [2], [3].

In addition, with recent developments in technology such as autonomous vehicles, precision and accuracy are no longer the only relevant factors. A model's ability to perform object detection in real-time is necessary in order to accommodate for a vehicle's real-time environment. An efficient and fast object detection algorithm is key to the success of autonomous vehicles [4], augmented reality devices [5], and other intelligent systems. A lightweight algorithm can be applied to many everyday devices, such as an Internet connected doorbell or thermostat. Currently, the state-of-the-art object detection algorithms used in cars rely heavily on sensor output from expensive radars and depth sensors. Other techniques that are solely computer based require immense amount of GPU power and even then are not always real-time, making them impractical for everyday applications. The general trend in computer vision is to make larger and deeper networks to achieve higher accuracy [6], [7], [8], [9]. However, such improvement in accuracy with heavy computational cost may not be helpful to face the challenge in many real world applications which require real-time performance carried out in a computationally limited platform.

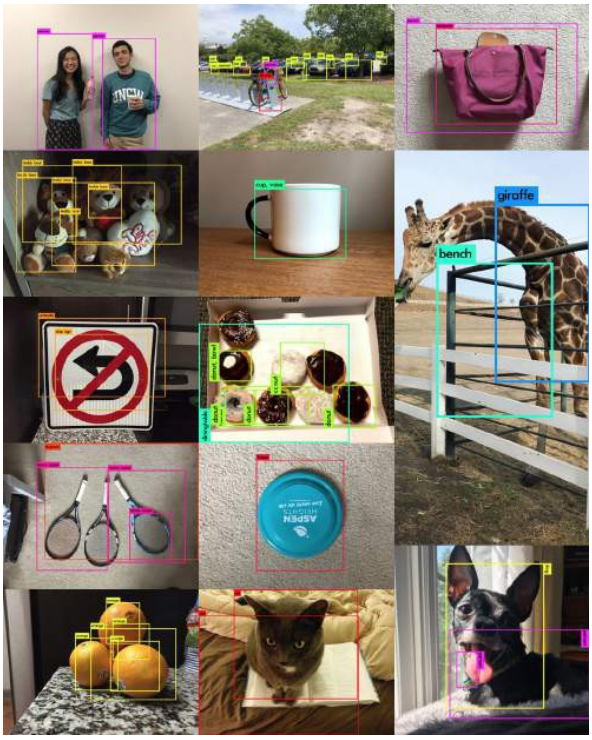


Fig. 1. Example images passed through our YOLO-LITE COCO model.

Previous methods, such as You-Only-Look-Once (YOLO) [10] and Regional-based Convolutional Neural Networks (R-CNN) [11], have successfully achieved an efficient and accurate model with high mean average precision (mAP); however, their frames per second (FPS) on non-GPU computers render

*equal authorship

978-1-5386-5035-6/18/$31.00 ©2018 IEEE