Analysis of Utilization of Generative Models to Increase Image Quality for Object Detection in Satellite Imagery

Jason Zisheng Chang, Yifei Wang, Ruiqi Wang
Master in Interdisciplinary Data Science (MIDS), Duke University

Introduction

The goal of this project is to observe the performance gains of object detection using a generative adversarial networks (GAN) to increase the resolution of satellite imagery, and determine the difference in performance of object detection on 0.3 m/px and 1.2 m/px resolution data. Deep learning models built are a **You Only Look Once (YOLO)** system for object detection, and a **Super Resolution Generative Adversarial Networks (SRGAN)** system for super resolution (SR). Results showed that object detection performance of YOLO can be increased by SRGAN.

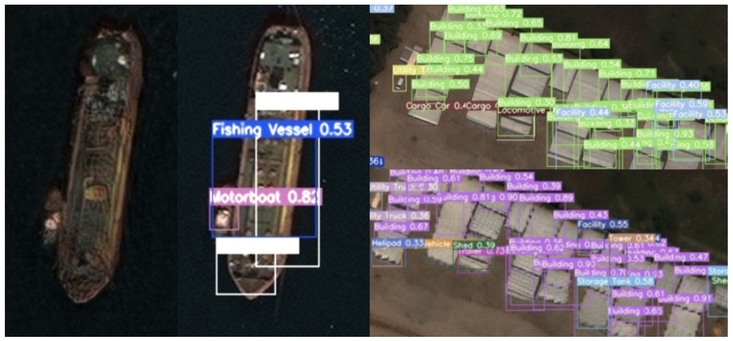


Figure 1. Images synthesized using SRGAN and original dataset.
Left column: left is SRGAN, right picture is original.
Right column: top is original, bottom is SRGAN.

Motivation

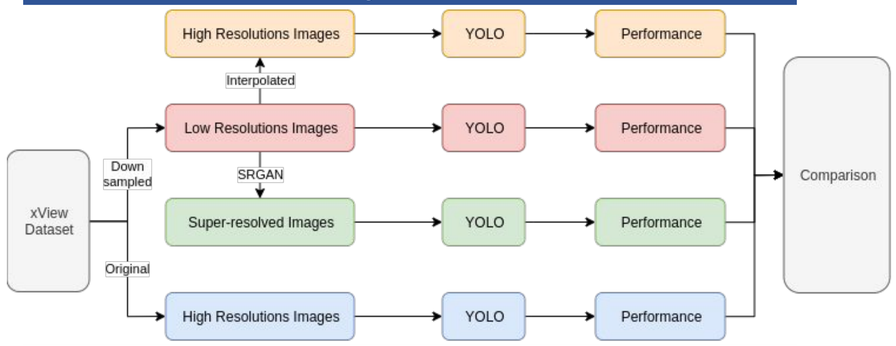
Object detection algorithms often struggle with low-resolution objects. SR seems to be an appropriate solution for this problem, but most approaches lack a quantitative measurement for SR images. In this project we quantitatively analyzed using a a SRGAN model to improve the performance of the baseline object detection model.

Data

Satellite imageries from the xView Dataset [1]. 50 training, 11 testing. Each images is around 10,000px x 10,000px in size. (3 km^2)

Class Label (train)	# of Ground Truth	Class Label (test)	# of Ground Truth
Building	59919	Building	7076
Small Car	1951	Small Car	840
Bus	1254	Truck	314
Cargo Truck	805	Bus	238
...	...	...	...

Project Flow



Models - YOLO

We utilize YOLO for unified object detection. YOLO detection system first resizes the input image into a grid, and thresholds the resulting detections by the models confidence. (Figure 2.)

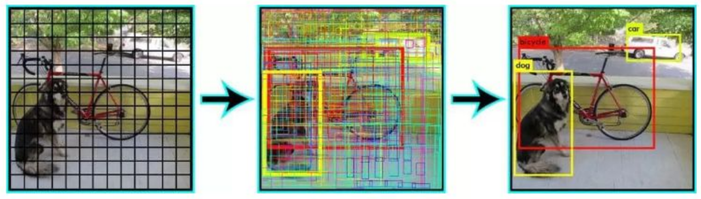


Figure 2. A demonstration of YOLO detection system [2]

Models - SRGAN

A generative adversarial network (GAN) is a model that contains two neural networks: a generator and discriminator). The GAN learns to generate new data with same statistics as the training set. A super resolution generative adversarial network (SRGAN) is a GAN that increases the resolution of low-resolution (LR) images by training on corresponding high-resolution (HR) images. By synthesizing sub pixel information in LR imagery a SRGAN generates super-resolved (SR) images.

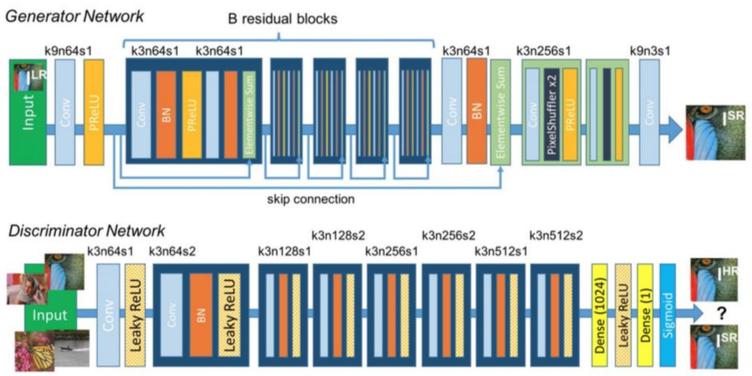


Figure 3. The network architecture of SRGAN system. [3]

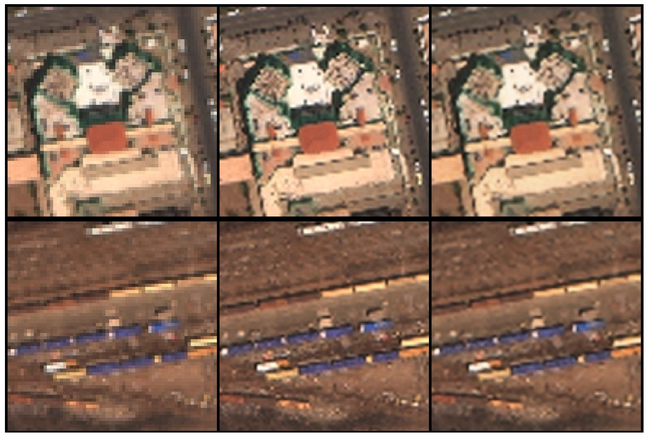


Figure 4. From left to right: the low resolution images, the original high resolution images, the super resolved images by the SRGAN.

Test Results

	#	Upsampled mAP	Original mAP	Generated mAP	Downsampled mAP
Plane	32	0.308709	0.590898	0.609714	**0.788502**
Building	7076	0.239417	0.190386	**0.328532**	0.18349
Yacht	86	0.0344961	**0.238569**	0.0307962	0.0523159
Car	840	0.0999877	**0.207925**	0.0285936	0

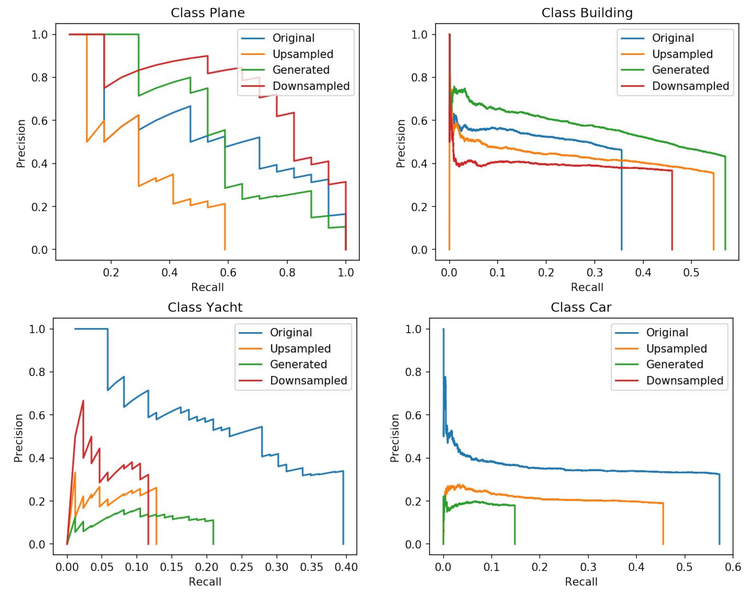


Figure 5. PR curves of YOLO detection for different classes of objects.

Contact

yifei.wang828@duke.edu
ruiqi.wang@duke.edu
zisheng.chang@duke.edu

References

[1] Darius Lam, Richard Kuzma, Kevin McGee, Samuel Dooley, Michael Laielli, Matthew Klaric, Yaroslav Bulatov, and Brendan McCord. xview: Objects in context in overhead imagery. 02 2018.

[2] Image source: Efficient Implementation of MobileNet and YOLO Object Detection Algorithms for Image Annotation. https://hackernoon.com/efficient-implementation-of-mobilenet-and-yolo-object-detection-algorithms-for-image-annotation-717e867fa27d

[3] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo- realistic single image super-resolution using a generative adversarial network. pages 105–114, 07 2017.