Bottle Detection in the Wild Using Low-Altitude Unmanned Aerial Vehicles

Jinwang Wang Wei Guo Ting Pan Huai Yu Lin Duan
Wen Yang

Wuhan University, Wuhan, China

2018.07.10-2018.07.13 Fusion 2018

武漢大學 WUHAN UNIVERSITY

dsp.whu.edu.cn 武漢大學 信号处理研究室

Good afternoon, everyone. I am Jinwang Wang from Wuhan University. In the next few minutes, I will introduce our “Bottle Detection in the Wild Using Low-Altitude Unmanned Aerial Vehicles”.

Outline

1. Introduction
2. UAV-Bottle Dataset
3. Baselines and Methods
4. Conclusion and Future Work

I will talk my work through the following sections.

Introduction

First, introduction.

Motivation

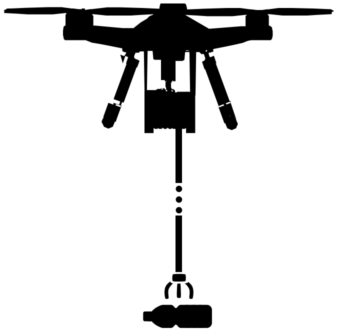
It's very dangerous for the sanitation workers who pick up rubbish on the mountain. Can we help them? Yes!

The sanitation workers on the Huashan Mountain.

As well we know, It's very dangerous for the sanitation workers who pick up rubbish on the mountain. So can we help them? Yes! We can use UAV to pick bottles.

Motivation

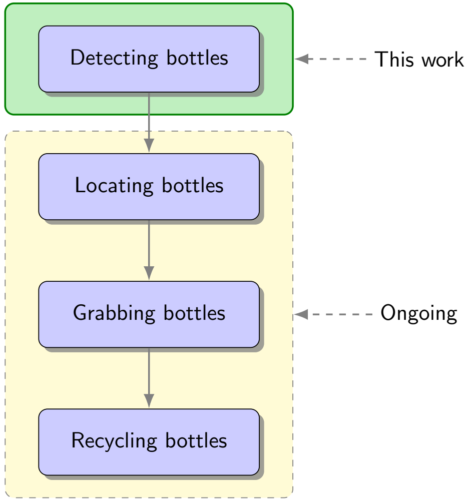
It's very dangerous for the sanitation workers who pick up rubbish on the mountain. Can we help them? Yes!



UAVs can help them!

As well we know, It's very dangerous for the sanitation workers who pick up rubbish on the mountain. So can we help them? Yes! We can use UAV to pick bottles.

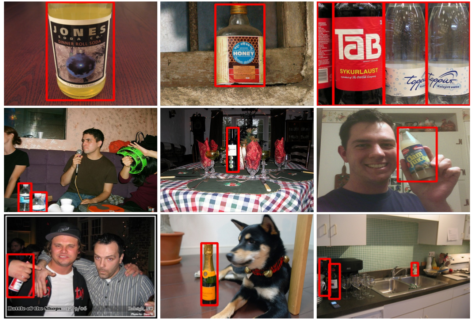
Pipeline



This work mainly focuses on bottles detection. The whole pipeline is shown as below. After bottles detection, bottles location, grab bottles and recycle is going on.

challenges

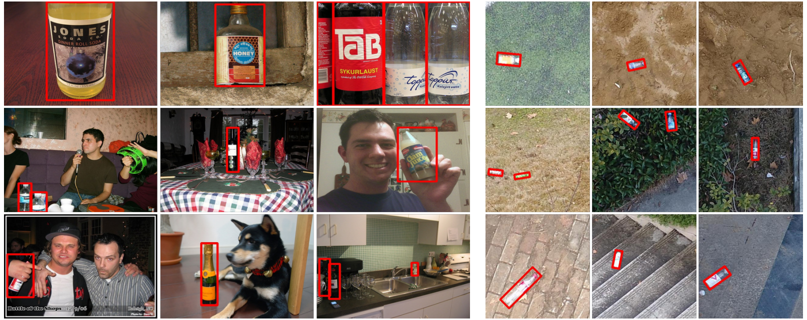
In UAV images, the bottles look completely different from the bottles in datasets such as PASCAL VOC, Microsoft COCO, etc.



In order to detect bottles on UAV, a high quality bottle dataset is needed for training deep learning based detection models. There are exactly bottles in the PASCAL VOC, Microsoft COCO, etc. datasets, why don't we use these bottles to training models?

challenges

In UAV images, the bottles look completely different from the bottles in datasets such as PASCAL VOC, Microsoft COCO, etc.



Because in UAV images, the bottles look completely different from the bottles in PASCAL VOC etc. datasets. As to UAV image, as you can see, detecting bottles exists several unique challenges. First, bottles are very small, generally less than 50×50 pixels; Second, the backgrounds of the bottles are very complex. Third, in contrast to conventional object detection dataset, the bottles in UAV images often appear with arbitrary orientations. Fourth, plastic bottles are often transparent, thus the background will can be seen through the bottles.

challenges

As to UAV images, detecting bottles exists several unique challenges.

- The size is very small, generally less than 50×50 pixels.

challenges

As to UAV images, detecting bottles exists several unique challenges.

- The size is very small, generally less than 50×50 pixels.
- Backgrounds are very complex.

challenges

As to UAV images, detecting bottles exists several unique challenges.

- The size is very small, generally less than 50×50 pixels.
- Backgrounds are very complex.
- Arbitrary orientations.

challenges

As to UAV images, detecting bottles exists several unique challenges.

- The size is very small, generally less than 50×50 pixels.
- Backgrounds are very complex.
- Arbitrary orientations.
- Plastic bottles are often transport, increaseing the difficulty of detection.

UAV-Bottle Dataset

Dataset

To build a baseline for bottle detection in UAV images, we establish a large scale bottle detection dataset, we call UAV-Bottle Dataset(UAV-BD) and benchmark.

To build a baseline for bottle detection in UAV images, we establish a large scale bottle detection dataset, we call UAV-Bottle Dataset(UAV-BD) and benchmark.

Dataset Collection

We used DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions:

- Wide range of scale and aspect ratios.

We use DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions. First, collecting images including bottles a wide range of scale and aspect ratio. Second, collecting images including different backgrounds. Third, collecting images including bottles different orientations. Third, collecting as many types of bottles as possible.

Dataset Collection

We used DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions:

- Wide range of scale and aspect ratios.
- Different backgrounds.

We use DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions. First, collecting images including bottles a wide range of scale and aspect ratio. Second, collecting images including different backgrounds. Third, collecting images including bottles different orientations. Third, collecting as many types of bottles as possible.

Dataset Collection

We used DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions:

- Wide range of scale and aspect ratios.
- Different backgrounds.
- Different orientation.

We use DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions. First, collecting images including bottles a wide range of scale and aspect ratio. Second, collecting images including different backgrounds. Third, collecting images including bottles different orientations. Third, collecting as many types of bottles as possible.

Dataset Collection

We used DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions:

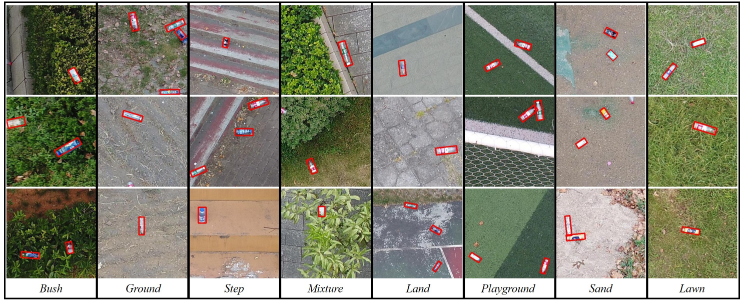
- Wide range of scale and aspect ratios.
- Different backgrounds.
- Different orientation.
- As many types of bottles as possible.

We use DJI Phantom 4 Pro to collect images. The resolution of the captured images are 5472×3078 pixels. At the same time, we follow four key suggestions. First, collecting images including bottles a wide range of scale and aspect ratio. Second, collecting images including different backgrounds. Third, collecting images including bottles different orientations. Third, collecting as many types of bottles as possible.

Dataset Collection

8 background scenes are chosen and annotated in our UAV-BD:

- Bushforest land
- Waste land
- Step
- Forest land
- Flat land
- Plastic stadium
- Sand land
- Grassland



Eight background scenes are chosen and annotated in our UAV-BD, including Bush forest land, Waste land, Step, Forest land, Flat land, Plastic stadium, Sand land, Grassland.

Annotation Method

A common description of horizontal bounding boxes is (c_x, c_y, h, w) , where (c_x, c_y) is the center location, h, w are the height and width.

However, horizontal bounding box cannot accurately or compactly outline oriented instances such as the bottles in UAV images. So we use (c_x, c_y, h, w, θ) to describe bounding boxes, where θ is the angle from the horizontal direction of the horizontal bounding box.

A common description of horizontal bounding boxes is (c_x, c_y, h, w) , where (c_x, c_y) is the center location, h, w are the height and width. However, horizontal bounding box cannot accurately or compactly outline oriented instances such as the bottles in UAV images. So we use (c_x, c_y, h, w, θ) to describe bounding boxes, where θ is the angle from the horizontal direction of the horizontal bounding box.

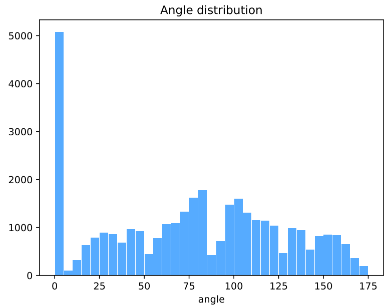
Dataset Statistics

The size of original image(OI) in UAV-BD is 5472×3078 pixels, they are too large to be trained for CNN-based algorithms. So we segment each OI into 144 small subimages(SIs), and the size of SIs is 342×342 pixels. Images and instances number in UAV-BD are shown in the table below.

Scenes	OI	SI	Instances in OI	Instances in SI
Bushforest land	230	4134	1812	3047
Wasteland	379	7598	4355	5800
Step	135	2691	1325	2106
Forest land	285	5724	3702	4891
Flat land	134	2803	1538	2142
Plastic stadium	336	6807	4180	4998
Sand land	249	5570	2704	4008
Grassland	456	9029	5778	7787
Total	2204	44356	25394	34779

In our dataset, the size of original image is 5472×3078 pixels, they are too large to be trained for CNN-based algorithms. So we segment each original image into 144 small subimages, and the size of subimages is 342×342 pixels. Images and instances number in UAV-BD are shown in the table below.

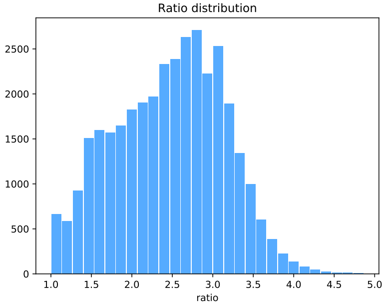
Dataset Statistics



Angle distribution of UAV-BD. The angle ranges from 0 to $180°$.

For better training the CNN-based algorithms, we counted some dataset's data distribution. Including angle distribution, ratio distribution, size distribution.

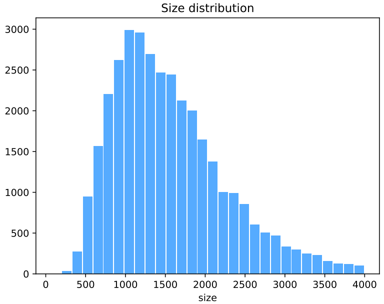
Dataset Statistics



Ratio distribution of UAV-BD. The ratio ranges from 1.0 to 5.0.

For better training the CNN-based algorithms, we counted some dataset's data distribution. Including angle distribution, ratio distribution, size distribution.

Dataset Statistics



Size distribution of UAV-BD. The size ranges from 0 to 4000.

For better training the CNN-based algorithms, we counted some dataset's data distribution. Including angle distribution, ratio distribution, size distribution.

Baselines and Methods

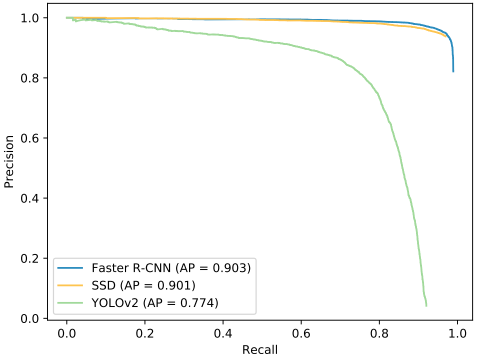
Dataset Split

We randomly select 64%, 16% and 20% of the UAV-BD as the training, validation and testing data. So the whole UAV-BD contains $16,258$ images with $22,211$ instances for training, $4,055$ images with $5,624$ instances for validation and $5,081$ images with $6,944$ instances for testing.

We randomly select 64%, 16% and 20% of the UAV-BD as the training, validation and testing data. So the whole UAV-BD contains $16,258$ images with $22,211$ instances for training, $4,055$ images with $5,624$ instances for validation and $5,081$ images with $6,944$ instances for testing.

Baselines with Horizontal Bounding Boxes(HBB)

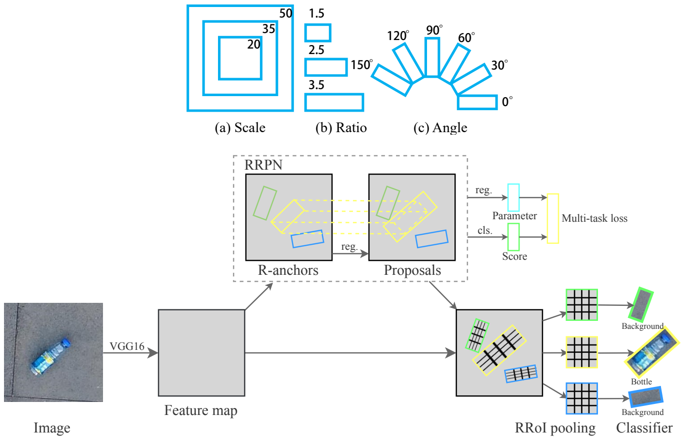
We select Faster R-CNN, SSD, YOLOv2 as our baseline for horizontal object detection. The experimental results of HBB prediction are shown in figure below.



We select Faster R-CNN, SSD, YOLOv2 as our baseline for horizontal object detection. The experimental results of HBB prediction are shown in figure below.

Baseline with Oriented Bounding Boxes(OBB)

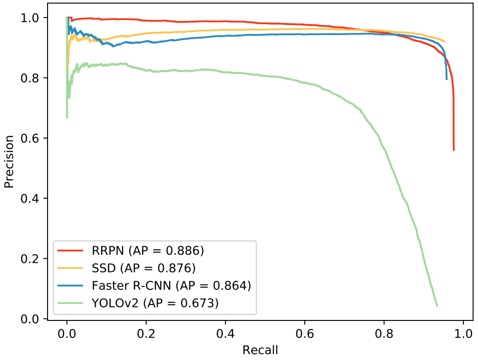
For oriented object detection, wo modify the original Rotation Region Proposal Network(RRPN) algorithm to predict properly oriented bounding boxes. RRPN's network structure is shown in figure below.



For oriented object detection, wo modify the original Rotation Region Proposal Network(RRPN) algorithm to predict properly oriented bounding boxes. RRPN's network structure is shown in figure below. The scale, ratio and angle settings are came from dataset's distribution.

Baseline with Oriented Bounding Boxes(OBB)

The experimental results of OBB prediction are shown in figure below.

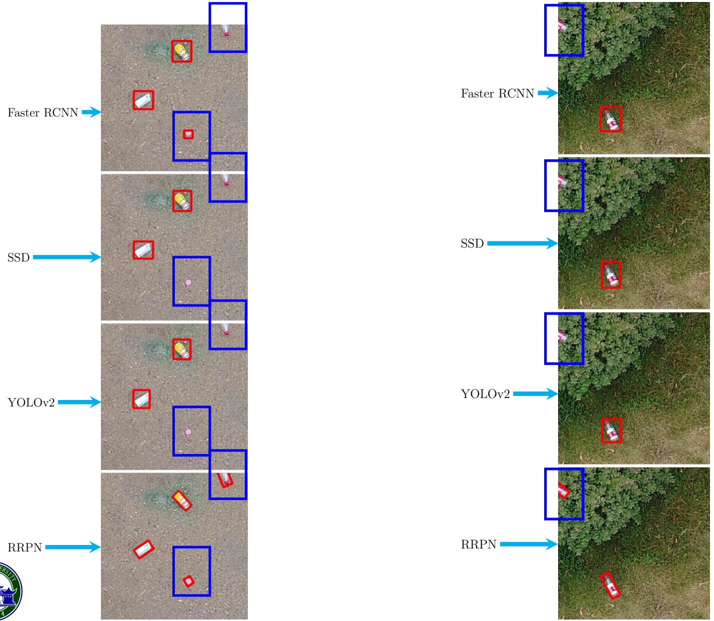


The experimental results of OBB prediction are shown in figure below.

Experimental Analysis

Bottles detection results in our dataset. As you can see, RRPN's performance is highest.

Some Details



Bottles detection results in our dataset. As you can see, RRPN's performance is highest.

Conclusion and Future Work

Conclusion and Future Work

Contribution

- Built a large-scale dataset for bottle detection in UAV images named UAV-BD.

Future Work

Our contributions including built a large-scale dataset for bottle detection in UAV images named UAV-BD, annotated a huge number of well-distributed bottles with oriented bounding boxes, established a benchmark for bottle detection. In the future, we will focus on locating and recycling bottles in the real-world using UAV.

Conclusion and Future Work

Contribution

- Built a large-scale dataset for bottle detection in UAV images named UAV-BD.
- Annotated a huge number of well-distributed bottles with oriented bounding boxes.

Future Work

Our contributions including built a large-scale dataset for bottle detection in UAV images named UAV-BD, annotated a huge number of well-distributed bottles with oriented bounding boxes, established a benchmark for bottle detection. In the future, we will focus on locating and recycling bottles in the real-world using UAV.

Conclusion and Future Work

Contribution

- Built a large-scale dataset for bottle detection in UAV images named UAV-BD.
- Annotated a huge number of well-distributed bottles with oriented bounding boxes.
- Established a benchmark for bottle detection.

Future Work

Our contributions including built a large-scale dataset for bottle detection in UAV images named UAV-BD, annotated a huge number of well-distributed bottles with oriented bounding boxes, established a benchmark for bottle detection. In the future, we will focus on locating and recycling bottles in the real-world using UAV.

Conclusion and Future Work

Contribution

- Built a large-scale dataset for bottle detection in UAV images named UAV-BD.
- Annotated a huge number of well-distributed bottles with oriented bounding boxes.
- Established a benchmark for bottle detection.

Future Work

- We will focus on locating and recycling bottles in the real-world using UAV.

Our contributions including built a large-scale dataset for bottle detection in UAV images named UAV-BD, annotated a huge number of well-distributed bottles with oriented bounding boxes, established a benchmark for bottle detection. In the future, we will focus on locating and recycling bottles in the real-world using UAV.

Get UAV-Bottle Dataset

UAV-Bottle Dataset and Development Kit can be downloaded on Google Drive and Github.

UAV-Bottle Dataset

https://jwwangchn.github.io/UAV-BD/

Development Kit

https://github.com/jwwangchn/UAV-BD.git

You can download our dataset from Google Drive and download develop kit from Github.

References

- Xia G S, Bai X, Ding J, et al.
  DOTA: A large-scale dataset for object detection in aerial images
  IEEE CVPR, 2018
- Ma J, Shao W, Ye H, et al.
  Arbitrary-oriented Scene Text Detection via Rotation Proposals
  IEEE Transactions on Multimedia, 2018
- Ren S, He K, Girshick R, et al.
  Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks
  Advances in Neural Information Processing Systems, 2015
- Liu W, Anguelov D, Erhan D, et al.
  SSD: Single Shot Multibox Detector
  ECCV, 2016
- Redmon J, Farhadi A.
  YOLO9000: Better, Faster, Stronger
  IEEE CVPR, 2017

There are some references.

Thanks!