Small-Object Detection in Satellite Images

Debojyoti Biswas and Dr. Jelena Tešić

Department of Computer Science

http://DataLab12.github.io/

Motivation

- **Aerial image dataset do not conform to the consumer image dataset assumptions in the analysis de jour**
- **Variations in image captioning conditions (lighting, weather, altitude, content, changes in scenery) render simple domain adaptation impossible**
- **State-of-the-art analysis struggles with the small and dense objects in aerial object detection.**

Contributions

New pipeline for small object detection in satellite images

1. Robust backbone for extracting and preserving small object features.
2. Difficulty scoring module
3. Custom focal loss function designed for small objects

Datasets

DIOR dataset

23,462 images + 192,472 object annotations

- A range of viewpoint angles
- A range of object sizes, ~1000 times difference in pixel size
- Various geographical areas captures
- Images captured in different weather conditions.
- High inter-class similarity and intra-class diversity.

Training set: 22,450 images

Test set: 1012 images.

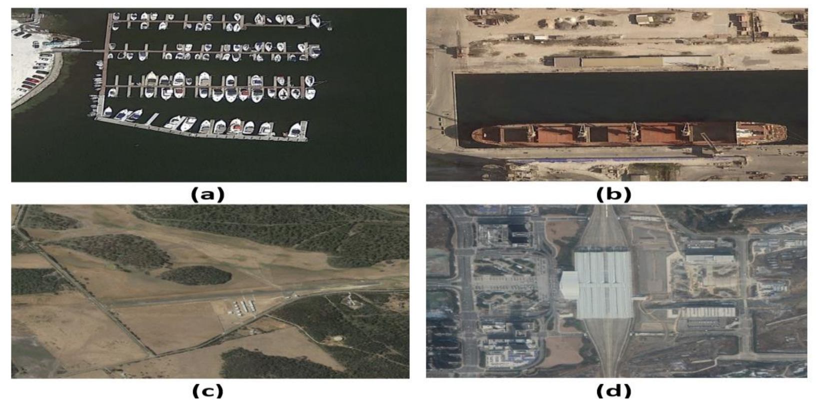

Figure 1. DIOR

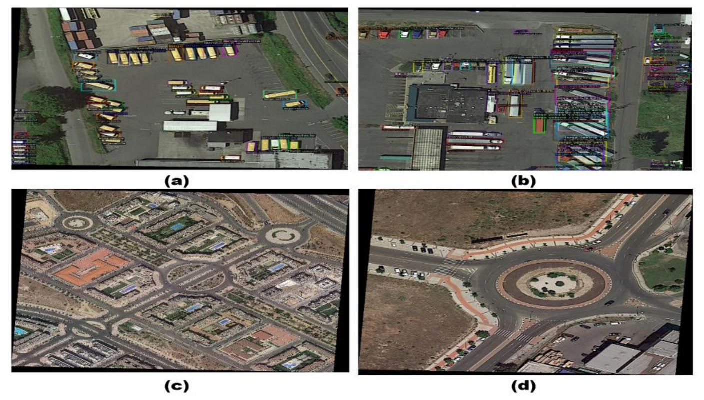

Figure 2. DOTA2.0

DOTA2.0 ***dataset***

- 2,430 overhead images collected from several satellites.
- 1,793,658 annotated objects
- 18 classes.

Training set: 12,700 images

Test set: 4,543 images.

Challenges

- Object with small size.
- Densely packed objects.
- Number of objects per image.
- Large variety in object orientation.
- High Global Spatial Distance(GSD).
- Imbalance Easy and Hard Examples
- Uniform features across the object.

Figure 3. Consumer and aerial image examples

Baseline Model

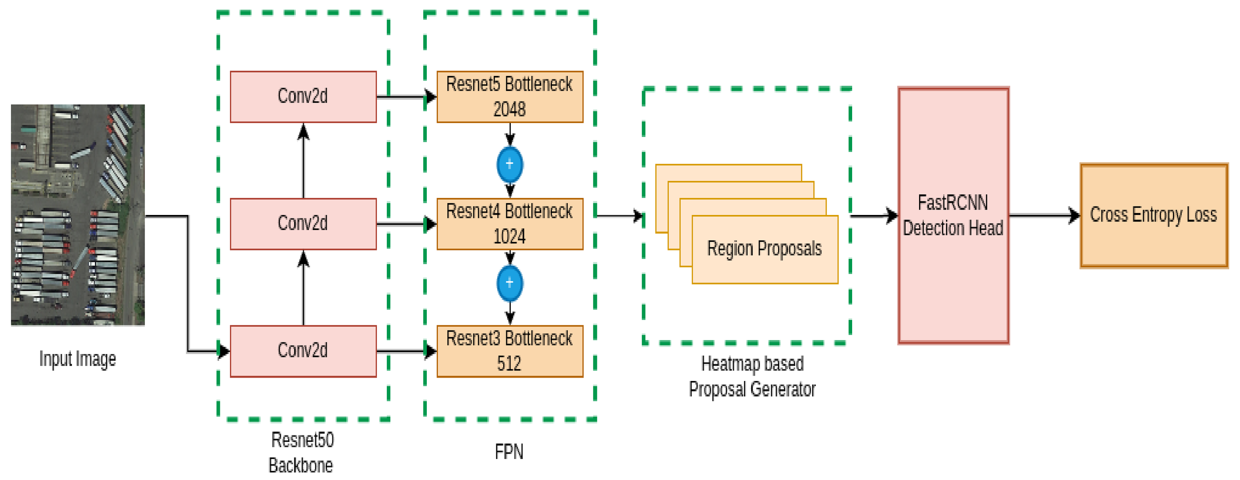

Figure 4. Baseline architecture: CenterNet2

System Specification

System	Configuration
Operating System	18.04
CPU	11th Gen Intel® Core™ i9-11900K @ 3.50GHz × 16
GPU	NVIDIA Corporation GP102 [TITAN Xp]
GPU Memory	12GB
RAM	125GB

Table 1. System Specifications

Small-Object Detection (SOD) Pipeline

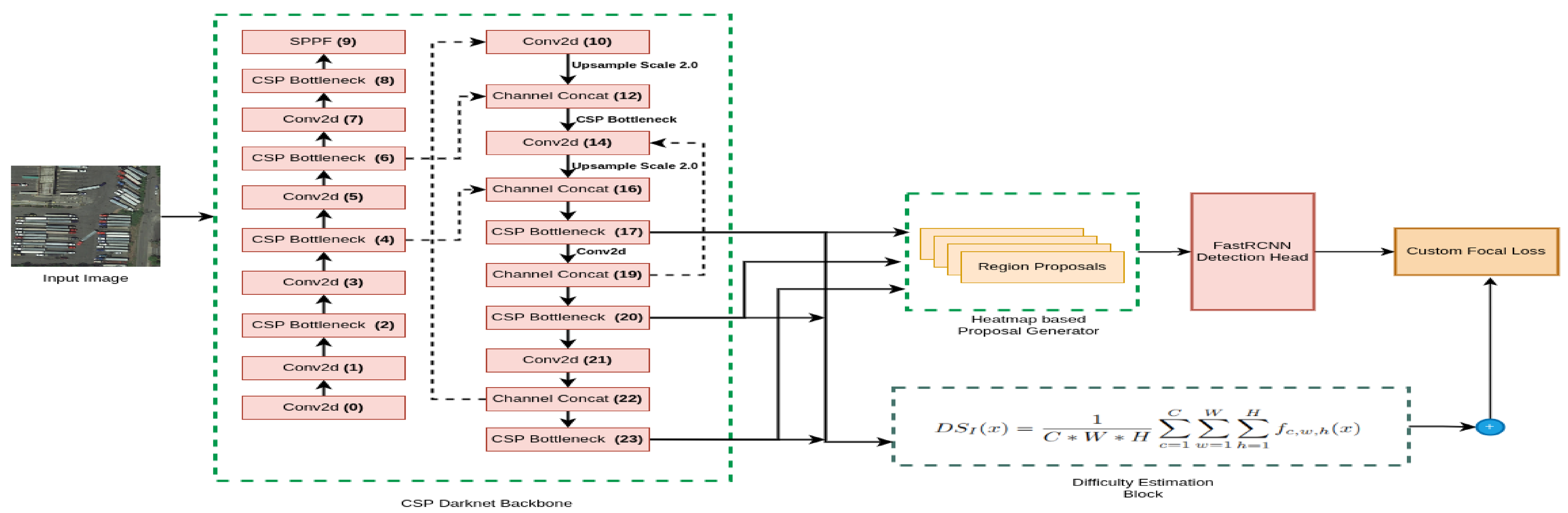

Figure 5. SOD architecture with darknet backbone and difficulty module

Findings

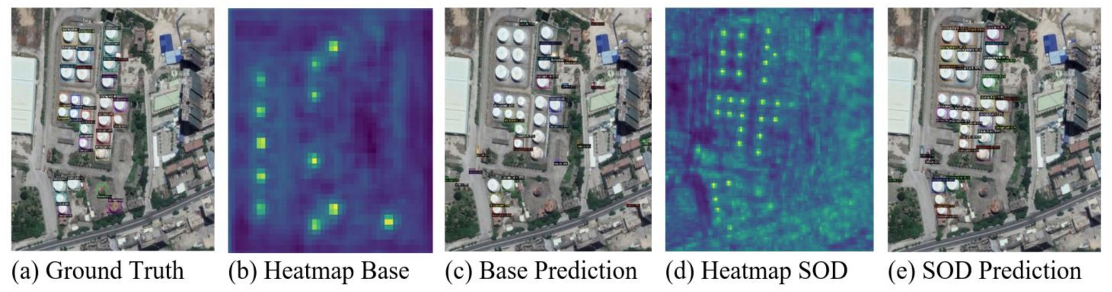

Figure 6. Detection from DIOR dataset

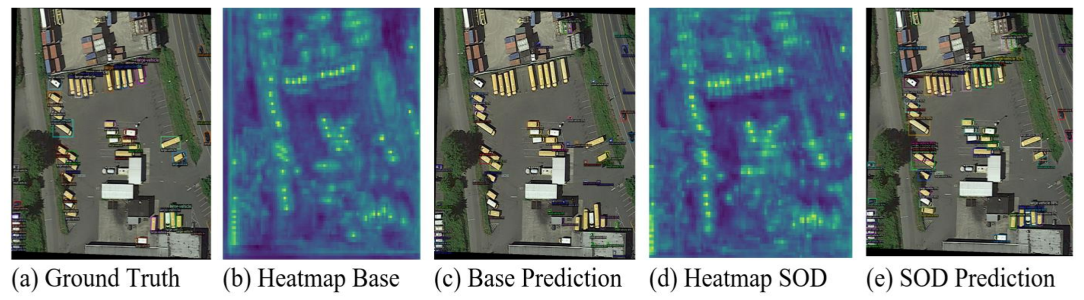

Figure 7. Detection from DOTA dataset

Experiments

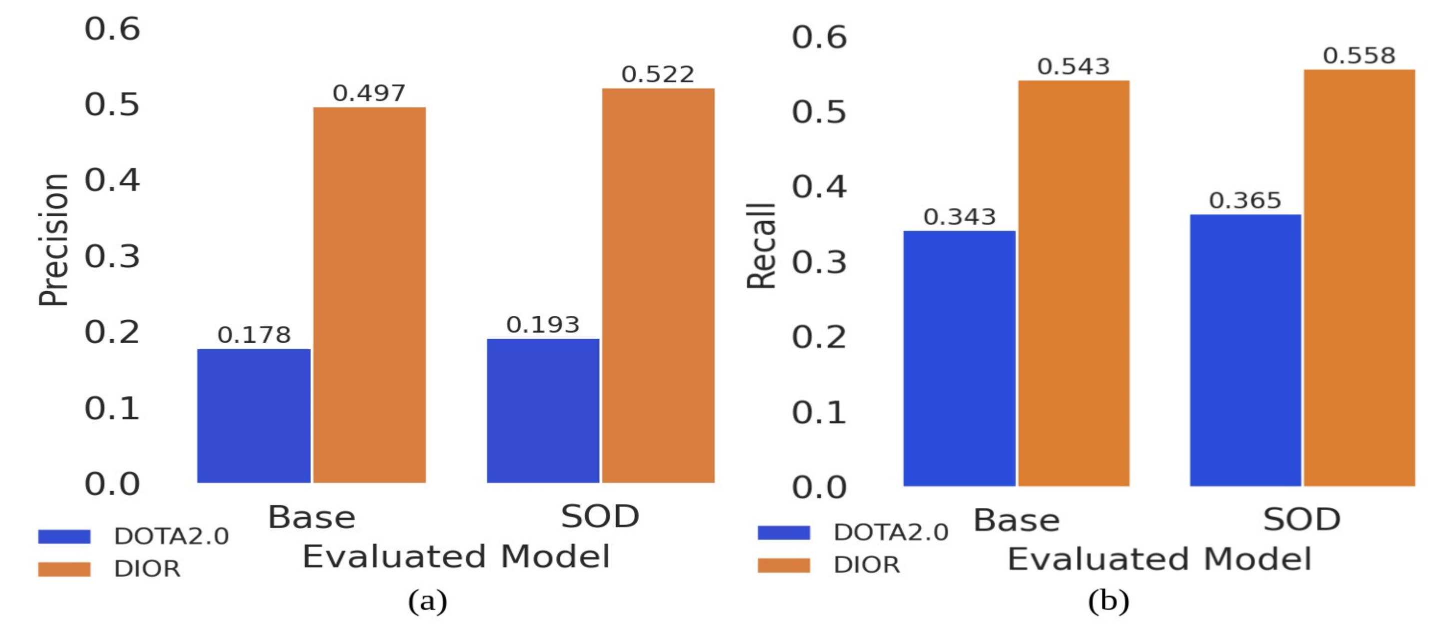

Figure 8. (a) Precision(IOU=0.50:0.95) and (b) Recall(IOU=0.50:0.95) comparison from different models vs. different datasets.

Class Label	mAP	Bridge	Service Area	Harbor	Ship	Storage Tank	Track	Station	Tennis Court	Overpass	Airplane	Dam	Airport	Toll Station
Num. Ann.	NA	207	67	259	2494	2629	154	58	580	163	844	33	56	67
Base	49.6	22.86	54.63	35.03	52.14	42.32	52.60	27.14	74.75	34.92	65.43	29.30	53.73	42.72
SOD	51.9	24.84	58.85	39.72	55.47	44.81	54.25	31.22	76.27	37.51	68.32	31.18	58.12	45.61

Class Label	mAP	Plane	Bridge	Small Vehicle	Large Vehicle	Ship	Basketball	Storage Tank	Roundabout	Harbor	Helicopter	Crane	Helipad	Airport
Num. Ann.	NA	3792	634	53660	6739	17650	240	3045	214	3689	86	28	4	89
Base	17.1	36.18	8.61	10.14	21.68	21.23	21.78	18.13	14.32	19.58	10.36	0.00	0.00	11.35
SOD	18.9	38.23	10.33	11.74	21.82	22.94	22.88	20.21	15.10	21.06	12.11	2.41	1.98	14.11

Table 2. DIOR and DOTA2.0 AP scores for small and difficult classes

Conclusion and Future Work

- DNN object detectors perform well if
  - Training dataset contains enough annotated
  - Feature extraction does not miss small object characteristics
- Heatmap Based proposal generator performs well for small objects.
- Difficulty module and the custom focal loss improve the detection performance with hard and soft example mining.
- In the Future, we plan to perform domain adaptation across multiple aerial datasets.

Acknowledgments

The work has been supported by NAVAIR, NVIDIA @ Data Lab (DataLab12.github.io) TXST