

Deep learning-based detection of humans for search and rescue operations



Nayee Muddin Khan Dousai

(nayee.dousai@fer.hr)

Mentor: Prof. dr. sc. Sven Lončarić

University of Zagreb Faculty of Electrical Engineering and Computing



1. Introduction and Challenges

- Much research has been done in the field of object detection in the last decade, but the detection of objects in aerial images is still a challenging problem
- Object detection in aerial images depends on various factors such as altitude, low visibility, object of interest, pose and scale variation, camouflaged environment with rocks and trees, and high-resolution aerial images
- To overcome the expensive and time-consuming SAR traditional methods, we will adopt the implementation of deep learning-based object detection models

2. Heridal Dataset

- HERIDAL dataset[1] consists of almost 1500 well-labelled training images and about 500 testing images
- Every captured image for training and testing is 4000 X 3000 image resolution
- The images in this dataset are captured from real-world using drones mainly for search and rescue operations (SAR)

3. Methodology

- Our research is based on ensemble learning using EfficientDET architecture
- The human detection algorithm can be explained in below steps:
 - ❖ Regenerate the HERIDAL dataset into various image resolutions such as 512, 640 and 1024
 - ❖ Train each image resolution separately on BiFPN and FC-FPN
 - ❖ Concatenate the best features from the above two various experiments
 - ❖ Train the network to detect and localize humans at the end

4. Experiments and Results

- The experiments on the HERIDAL dataset are carried on a few stages as plotted in below table
- ❖ At first, we will train the EfficientDET network with Bi-FPN for all the respective image resolutions by freezing the backbone in step 1 and unfreeze in step 2
- ❖ In the second step, we do the same with FC-FPN

Experiments	Image resolution of 512	Image resolution of 640	Image resolution of 1024
mAP based on EfficientDET with BiFPN step 1	91.27%	91.05%	88.07%
mAP based on EfficientDET with BiFPN step 2	93.29%	91.52%	89.56%
mAP based on EfficientDET with FCFPN step 1	91.46%	90.47%	88%
mAP based on EfficientDET with FCFPN step 2	93.31%	91.86%	89.45%
mAP based on EfficientDET and Ensemble learning	95.11%	92.63%	90.06%

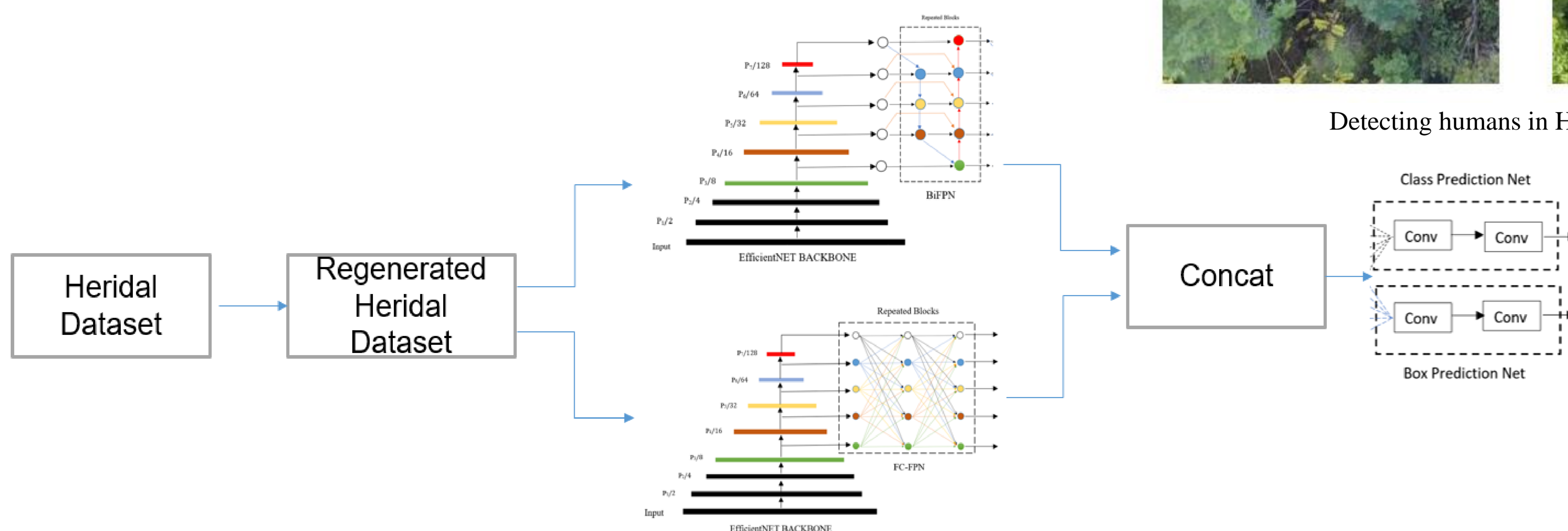
Ensemble learning results for various HERIDAL dataset image resolutions [2]

Object detection model	Calculated mAP
Kundid Vasić et al.	68.89%
Božić-Štulić et al.	88.9%
Nayee Muddin Khan et al. [2]	93.29%
mAP based on EfficientDET and Ensemble learning on 1024 resolution	90.06%
mAP based on EfficientDET and Ensemble learning on 640 resolution	92.63%
mAP based on EfficientDET and Ensemble learning on 512 resolution	95.11%

Comparing results with other papers



Detecting humans in HERIDAL dataset



Proposed architecture

5. References

- [1]. Božić-Štulić, Dunja, Željko Marušić, and Sven Gotovac, "Deep learning approach in aerial imagery for supporting land search and rescue missions." *International Journal of Computer Vision* 127.9 (2019), pp. 1256
- [2]. Dousai, Nayee Muddin Khan, and Sven Lončarić. "Detecting Humans in Search and Rescue Operations Based on Ensemble Learning." *IEEE Access* 10 (2022): 26481-26492.

6. Project Acknowledgement

This work was supported by the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant under Agreement 764951

