Human detection in aerial images for search and rescue operations

Nayee Muddin Khan Dousai and Sven Lončarić

Department of Electronic Systems and Information Processing
Faculty of Electrical Engineering and Computing
Image Processing Group
University of Zagreb, Croatia

1. INTRODUCTION

• Much research has been done in the field of object recognition in the last decade, but the recognition of objects in aerial images is still a challenging problem

• Object recognition in aerial images depends on various factors such as altitude, low visibility, object of interest, pose and scale variation, camouflaged environment with rocks and tree, and high-resolution aerial images

• To overcome the highly expensive and time-consuming methods, we can adopt the use of consumer drones in SAR operations as they are easily available in the market to purchase and much cheaper

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 [2] architecture

• The Methodology can be stated as below steps:

❖ Regenerate the HERIDAL dataset into 512, 640 and 1024 image resolution

❖ 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
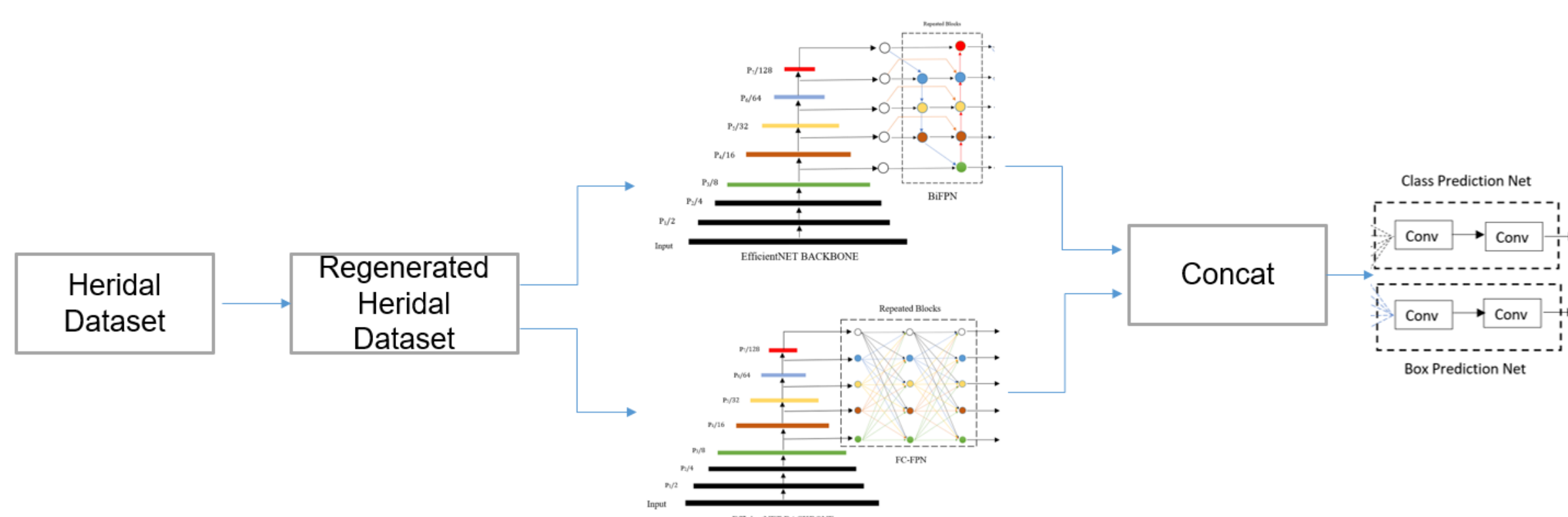
4. EXPERIMENTS & RESULTS

• The experiments on HERIDAL dataset are carried on 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	HERIDAL dataset with 512 image resolution	HERIDAL dataset with 640 image resolution	HERIDAL dataset with 1024 image resolution
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 FC-FPN step 1	91.46%	90.47%	88%
mAP based on EfficientDET with FC-FPN step 2	93.31%	91.86%	89.45%
mAP based on EfficientDET with Ensemble learning	**95.11%**	**92.63%**	**90.06%**

Ensemble learning results for various HERIDAL dataset image resolutions

Object detection model	mAP Calculated
Kundid Vasić et al.	68.89%
Božić-Štulić et al.	88.9%
mAP based on EfficientDET with Ensemble learning on 1024 image resolution	90.06%
mAP based on EfficientDET with Ensemble learning on 640 image resolution	92.63%
mAP based on EfficientDET with Ensemble learning on 512 image resolution	95.11%

Comparing results with other papers

Detecting humans in HERIDAL dataset

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]. Tan, Mingxing, and Quoc V. Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." arXiv preprint arXiv:1905.11946 (2019)

Fakultet elektrotehnike i računarstva