

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/317589208>

Pedestrian Detection for UAVs Using Cascade Classifiers and Saliency Maps

Conference Paper · May 2017

DOI: 10.1007/978-3-319-59147-6_48

CITATIONS

13

READS

156

7 authors, including:



Wilbert G. Aguilar

Universidad de las Fuerzas Armadas-ESPE

50 PUBLICATIONS 307 CITATIONS

[SEE PROFILE](#)



Marco Andrés Luna

Universidad de las Fuerzas Armadas-ESPE

8 PUBLICATIONS 30 CITATIONS

[SEE PROFILE](#)



Vanessa Abad

University of Barcelona

14 PUBLICATIONS 42 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



VisualNavDrone [View project](#)



Group-Robot Interaction [View project](#)

Pedestrian Detection for UAVs using Cascade Classifiers and Saliency Maps

Wilbert G. Aguilar^{1,3,4,*}, Marco A. Luna², Julio F. Moya², Vanesa Abad⁵,
Humberto Parra^{1,6}, Hugo Ruiz^{1,7}, Cecilio Angulo⁴

¹ Departamento de Seguridad y Defensa DESP, Universidad de las Fuerzas Armadas ESPE,
Sangolquí, Ecuador

(wgaguilar@espe.edu.ec)

² Departamento de Eléctrica y Electrónica DEEE, Universidad de las Fuerzas Armadas ESPE,
Sangolquí, Ecuador

³ Centro de Investigación Científica y Tecnológica del Ejército CICTE, Universidad de las
Fuerzas Armadas ESPE, Sangolquí, Ecuador

⁴ Research Group on Knowledge Engineering GREC, Universitat Politècnica de Catalunya
UPC-BarcelonaTech, Barcelona, Spain

⁵ Universitat de Barcelona UB, Barcelona, Spain

⁶ Universidad Politécnica de Madrid UPM, Madrid, Spain

⁷ Purdue University PU, Indiana, United States

Abstract. We propose in this paper an algorithm and provide a dataset for pedestrian detection in applications with micro multi rotors UAV. Training dataset is obtained by capturing images from surveillance cameras at different angles and altitudes. We propose a method based on HAAR-LBP cascade classifiers with Adaboost training and, additionally we combine cascade classifiers with saliency maps for improving the performance of the pedestrian detector. We evaluate our dataset by the implementation of the HOG algorithm with Adaboost training and, finally, algorithm performance is compared with other approaches from the state of art. The results shows that our dataset leads to accurate pedestrian detection in UAVs, HAAR-LBP behaves better than HAAR like features and the use of saliency maps improves the performance of detectors due to the elimination of false positives in the image.

Keywords: HAAR, HOG, LBP, Saliency Maps, People Detection, Cascade Classifiers, UAVs.

Introduction

In the field of computer vision, there are several applications in object detection, pedestrian detection among them, used on surveillance [1], [2], driver assistance systems – particularly in PPSs (pedestrian

protection systems) [3], [4], robotics [5], [6] and others. Multiple feature extraction algorithms working with machine learning and datasets exist dealing with this problem.

Developments in computer vision have been introduced for UAVs (unmanned aerial vehicles) [7], [8], [9]. Pedestrian detection can be performed with UAVs, considering that complex dynamic of the vehicle and altitude variation add extra challenges to the detection [10]. Conventional classifiers fail when altitude increases generating more false positives.

Our proposal for pedestrian detection in UAVs considers the altitude and introduces the CICTE-PeopleDetection dataset with images captured from surveillance cameras. We use two trained algorithms: The first one based on a combination of the feature extraction methods HAAR-LBP, and the second one based on HOG. Both algorithms use cascade classifiers with Adaboost training. In addition our proposal use a Saliency Maps based algorithm to provide detection robustness. Our proposal is evaluated in images captured from UAVs in different scenarios.

This paper is organized as follows: Section II describes the related work on pedestrian detection. Next, our proposal for pedestrian detection, the creation of dataset and video stabilization are described in Section III. In Section IV we present the experimental results, followed by the summary. Finally conclusions and future works are presented in section V.

Related works

In the literature, several research groups have created different datasets and methods for pedestrian detection. INRIA introduced a first method [11], with training based on Histograms of Oriented Gradients (HOG). Widely used datasets are Caltech Pedestrian Dataset [12] and KITTI [13], due to they are comparatively large and challenging. According to [14], [15] there are two types of datasets: photo datasets and video datasets. Photo datasets like MIT [16], CVC [4], NICTA[17] aboard the classification problem: train binary classification algorithms. Video

datasets as ETH [18], TUD-Brussels [19] or Dalmier (DB) [20] are focused on the detection problem: design and test full image detection systems and human locomotion modeling.

Two popular algorithms exist for pedestrian detection and object detection in general: Haar-like features [21] by Viola and Jones, and Dalal and Triggs algorithm called HOG [11]. Both algorithms have generated over 40 new approaches [14]. Several methods for pedestrian detection includes feature extraction algorithms: HAAR [21], HOG [11], [22], HOG-HAAR [23] and HOG-LBP [24]; working with machine learning approaches based on SVMs [11], [25] or Adaboost[4], [20].

The applications of pedestrian detection in UAVs are manifold: Human safety [26], rescue and monitoring missions [27], [28], track people systems [25], [29], and others. One of the challenges of pedestrian detection in UAVs is the camera perspective variations that deform the images. In [30], [31], they use thermal imagery combined with cascade classifiers to perform the detection. Few papers like [28] works on altitudes around five meters. In this paper, authors propose post-disaster victims detection with cascade classifier methods. In UAVs, the use of saliency maps is widely used to object and motion detection in aerial images [28], [32], [33]. Works like [27] use saliency maps to reduce the search space and detect people from 10 to 40 meters of altitude.

Our Approach

1 Dataset Creation

One of the reasons for introducing our dataset is the requirement to detect people from UAV cameras. The main problem in pedestrian detection is the high altitude, where people images have deformation of their characteristics. The main difference of the proposed CICTE-PeopleDetection with previous photo datasets is the location and

perspective of the cameras that emulate the onboard camera perspective of the UAV. We use surveillance cameras for photo dataset creation due to UAVs videos are stable and comparable with fixed cameras. There are approximately 100 cameras (we can not specify the exactly number of cameras for security reasons) with D1 resolution located in the University between 2.3 m and 5 m of height looking down as shown in Figure 1.

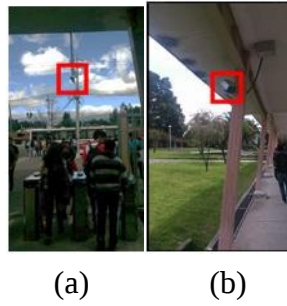


Fig. 0. Location of the cameras in the campus (a) 5m height. (b) 2.3 m

For training we provide positive and negative images. Positive images are images that contain the object to be detected, in our case pedestrians. Negative images are frames without pedestrians. Our dataset is composed of 3900 positive images and 1212 negative ones. The positives images were captured in the Universidad de las Fuerzas Armadas ESPE during the day and the night in different scenarios, and contain entire and partial occluded people samples.

2 Training Process

Our approach consists in the combination of two algorithms for extraction of the feature set: Local Binary Patterns (LBP) and Haar-like features. We use Adaptive Boosting (AdaBoost) as training algorithm and a combination of Haar-LBP features since they are algorithms of low computation time. To train our Haar-LBP algorithm we divided the overall set of images in 70% for training and the other 30% for testing.

Then, we use the tuned algorithm with UAV images in different scenarios. Additionally, we train a HOG cascade classifier and compare it with OpenCV HOG to validate our dataset. The training processes are shown Figure 2.

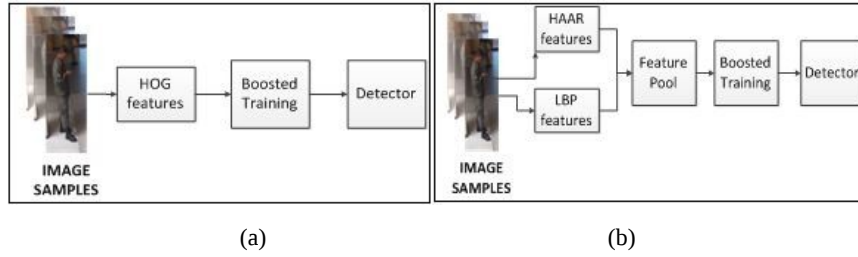


Fig. 0. Pedestrian detection training. (a) HOG features with Adaboost. (b) Haar-LBP features with Adaboost

The methods used for training the cascade classifiers are described as follows:

Local Binary Patterns

This feature extractor was presented in [34] as a texture descriptor for object detection, and compares a central pixel with the neighbours. The window to be examined is separated into cells of 16x16 pixels. 8 neighbours are considered for each pixel inside the cell, the central pixel value is the threshold. A value of 1 is assigned if the neighbour is greater or equal to the central pixel, otherwise the value is 0.

Haar-like Features

Viola and Jones uses a statistical approach for the tracking and detection problem, describing the ratio between light and dark areas within a defined kernel. This algorithm is robust regarding to noise and lighting changes. The method uses simple feature sets similar to Haar basis functions [21], [35].

Histogram of Oriented Gradients (HOG)

This algorithm is a feature descriptor for object detection focused on pedestrian detection. The image window is separated into smaller parts called cells. For each cell, we accumulate a local 1-D histogram of gradient orientations of the pixels in the cell. Each cell is discretized into angular bins according to the gradient orientation and each pixel of the cell contributes with a gradient weight to its corresponding angular bin. The adjacent cells are grouped in special regions called blocks and the normalized group of histograms represents the block histogram.

Adaboost

Adaboost is a machine learning algorithm [36] that initially keeps uniform distribution of weights in each training sample. In the first iteration the algorithm trains a weak classifier using a feature extraction method or a mix of them achieving a higher recognition performance for the training samples. In the second iteration, the training samples, misclassified by the first weak classifier, receive higher weights. The new selected feature extraction methods should be focused in these misclassified samples.

3 People Detection Algorithm

In order to get a better performance of the classifier we implement a combination of cascade classifier with saliency maps, an algorithm presented in [37]. The purpose of saliency maps is to locate prominent areas at every location in the visual field. The areas with high saliency correspond to objects or places they are most likely to be found, and the areas with lower saliency are associated to the background [38]. The saliency maps algorithms are deduced by convolving the function f by an isotropic bi-dimensional Gaussian function [39]:

$$S(X) = f(X) G_{\sigma}(X) \quad (0)$$

where σ is the standard deviation of the Gaussian function. The standard deviation depends on the experimental setup (size of the screen and viewing distance). To eliminate the false positives in the image we obtain the salient region; we consider a threshold from the salient map and we create a mask where values greater than threshold will belong to salient map. Additionally, this region was dilated to give it robustness. This algorithm is shown in Figure 3.



Fig. 0. Saliency maps algorithm. (a) Saliency map (b) Saliency region.

Once it has been obtained the salient region, our algorithm proposes as true positives only the cascade classifier detections inside this region. Hence, we define the salient region as Region of Interest (ROI). To determinate if a detection bounding box is inside the salient region, we compute the center point of the bounding box:

$$x_m = x + \left(\frac{w}{2}\right); y_m = y + \left(\frac{h}{2}\right) \quad (0)$$

where x and y are the horizontal and vertical coordinates of the top left of the bounding box, x_m and y_m are the coordinates of the central point and w, h are the width and height. We take the center point as reference to avoid false positives that could have small parts of their bounding box in salient regions. Our proposal is presented graphically in Figure 4.

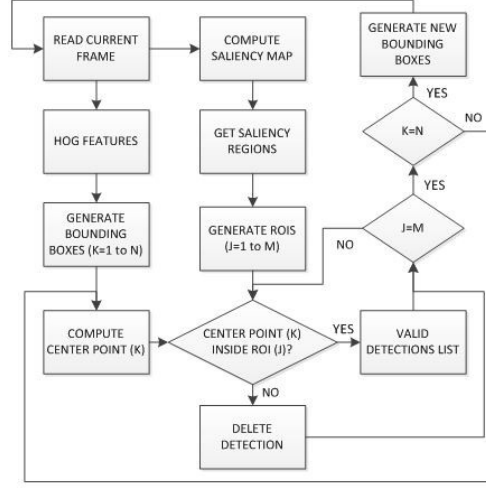


Fig. 0. Algorithm for people detection using HOG Cascade classifier and Saliency Maps

The results of the application of this algorithm are presented in the Section 4.

Results and Discussion

1 Dataset and Training Evaluation

The metric of evaluation for our approach is based on the sensitivity (true positive rate-TPR) and the miss rate (False negative rate-FNR). Defined as follows:

$$TPR = \frac{TP}{TP + FN} * 100 ; FNR = \frac{FN}{TP + FN} * 100 \quad (SEQ equation * MERGEFORMAT 3)$$

For the dataset evaluation we have trained the cascade classifier based on HOG features and compared this classifier with the OpenCV HOG cascade classifier. We tested the cascade classifier with videos captured from UAVs. Experimental results are presented in Table 1.

Table 0. Dataset Training performance

Algorithm	Sensitivity	Miss rate
-----------	-------------	-----------

	(%)	(%)
HOG-CICTE	55.71%	44.28%
HOG [25]	35.71%	64.28%

In this table, two cascade classifiers are compared: HOG-CICTE PeopleDetection and a HOG cascade classifier with the Adaboost training from the OpenCV library. Result shows our approach has better performance, the miss rate of our proposal is 20% lower than the conventional classifier miss rate, and the sensitivity is higher. ROC curves for comparing both algorithms are presented in Figure 5.

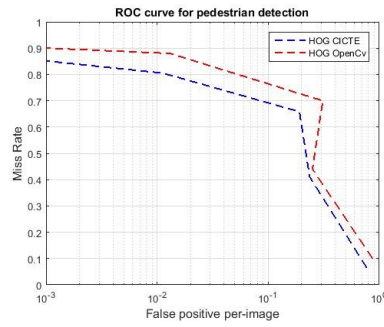


Fig. 0. Comparison of ROC curves for HOG from OpenCV cascade classifier and HOG-CICTE cascade classifier trained for pedestrian detection in UAVs.

In Figure 5, HOG-CICTE classifier has a better performance than HOG from OpenCV cascade classifier in videos captured from UAVs

2 Algorithm Evaluation

For the algorithm evaluation we are using 3 scenarios with 3 different altitudes. We compare HAAR-LBP features and HOG features (trained with CICTE-PeopleDetection) with respect to other cascade classifiers. Results are presented in Table 2.

Table 0. Cascade Classifiers Performance

Algorit	Altitude	Sensitivity	Miss rate
---------	----------	-------------	-----------

hm	(m)	(%)	(%)
HAAR-LBP	2	10.66%	89.34%
	3	7.18%	92.82%
	4	5.43%	94.57%
HAAR [36]	2	3.27%	96.73%
	3	2.28%	97.72%
	4	1.68%	98.32%
LBP [35]	2	41.98%	58.02%
	3	33.18%	66.82%
	4	23.09%	76.91%
HOG CICTE	2	67.23%	32.76%
	3	63.72%	36.27%
	4	60.44%	39.56%

In Table 2, the combination of HAAR-LBP features has low sensitivity compared with the other methods; however the proposal is higher than HAAR features. With altitude increasing, sensitivity decrease in all cascade classifiers. Performance curves are presented in the Figure 6.

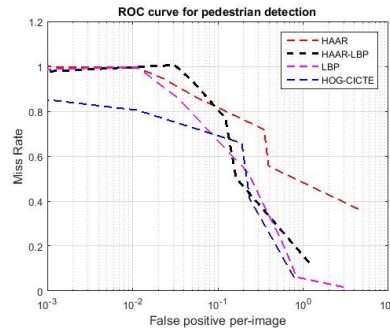


Fig. 0. Comparison of ROC curves for different approaches.

In the Figure 6 can be appreciated that the performance of the HAAR-LBP features algorithm is better than HAAR individually applied. HAAR-LBP features generate a lower rate of false positives. True positive rate of HAAR-LBP features is higher than HAAR features but

lower than LBP. Nevertheless, the HOG-CICTE cascade classifier still has the best performance due to its higher true positives rate and lower false positives rate.

3 Cascade Classifier- Saliency Maps Combination

Based on the results of performance we choose HOG CICTE cascade classifier to implement our algorithm. Graphical results are shown in the Figure 7 and video results are provided at https://www.youtube.com/watch?v=KN_hVgp1_t4

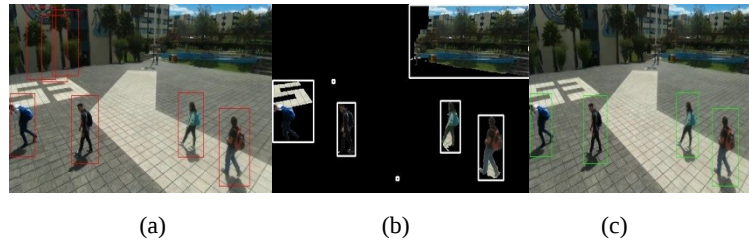


Fig. 0. Combination of Saliency Maps and Cascade classifier. (a) Cascade classifier result (b) Saliency regions (c) Final Result

As we can see in the Figure7, the use of saliency region helps to eliminate the false positives in the images. For the evaluation we take an additional metric of evaluation that is precision or positive predictive value (PPV), given by:

$$PPV = \frac{TP}{TP + FP} * 100$$

where TP are the true positive values and FP are the false positives. The results for precision of the detector with the application of the saliency region algorithm (SR) are shown in Table 3.

Table 0. Comparison of precision between algorithms

Algorithm	Precision (%)
-----------	---------------

HOG-CICTE	72.23%
HOG-CICTE +SR	92.1%

In Table 3, it can be observed that the application of the proposed saliency region algorithm improves the precision of detection in 20% approximatively, this denote an improvement in the perfomance too. The performance curves for both algorithms are shown in Figure 7.

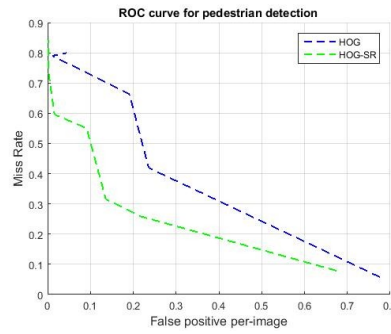


Fig. 0. ROC curves for HOG and HOG-SR

It is showed that the use of the saliency region algorithm improves the detection performance by eliminating false positives.

Conclusions and Future Work

Our proposal for pedestrian detection based on HOG features presents a higher performance than OpenCV HOG respect to sensitivity and miss rate (with an improvement of 20%), as shown in Table 1 and Figure 5, because the images used for training emulate UAVs perspective.

In order to improve the HAAR algorithm performance we combine two algorithms (HAAR and LBP). The sensitivity increased and the miss rate decreased as shown in Table 2 and Figure 6. However, performance is lower in comparison with HOG-CICTE and LBP algorithm.

When the altitude increased from 2 to 4 meters, the sensitivity decreased in the four algorithms. Comparing HAAR-LBP and HAAR , HAAR-LBP has a better performance even in the altitude of 4 meters.

The use of saliency maps improves the performance detectors, saliency map helps to eliminate background regions even in mobile cameras like UAVs, and these regions may contain objects that confuse the classifier that is important to decrease the number of false positives.

In the future, detection should be further improved. We will train new classifiers with images captured from UAVs, taking into consideration other human body parts like face, head, shoulders, etc. In addition, a robust of detection could be used for many applications like people tracking or people avoidance systems.

Acknowledgement

This work is part of the project Perception and localization system for autonomous navigation of rotor micro aerial vehicle in gps-denied environments, VisualNavDrone, 2016-PIC-024, from the Universidad de las Fuerzas Armadas ESPE, directed by Dr. Wilbert G. Aguilar.

References

1. H. Torresan, "Advanced surveillance systems: combining video and thermal imagery for pedestrian detection," *Proc. SPIE*, pp. 506–515, 2004.
2. L. Z. L. Zhang, B. W. B. Wu, and R. Nevatia, "Pedestrian Detection in Infrared Images based on Local Shape Features," *2007 IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 0–7, 2007.
3. D. M. Gavrila, "Pedestrian Detection from a Moving Vehicle," in *Proceedings of the 6th European Conference on Computer Vision*, 2000, vol. 1843, pp. 37–49.
4. D. Gerónimo, A. M. López, A. D. Sappa, and T. Graf, "Survey of Pedestrian Detection for Advanced Driver Assistance Systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1239–1258, 2010.

5. O. H. Jafari, D. Mitzel, and B. Leibe, "Real-time RGB-D based people detection and Tracking for mobile robots and head-worn cameras," *Proc. - IEEE Int. Conf. Robot. Autom.*, no. April 2016, pp. 5636–5643, 2014.
6. M. Kobilarov, G. Sukhatme, J. Hyams, and P. Batavia, "People tracking and following with mobile robot using an omnidirectional camera and a laser," *IEEE Int. Conf. Robot. Autom. 2006. ICRA 2006.*, no. May, pp. 557–562, 2006.
7. W. G. Aguilar and C. Angulo, "Real-time video stabilization without phantom movements for micro aerial vehicles," *EURASIP J. Image Video Process.*, vol. 2014, no. 1, p. 46, 2014.
8. W. G. Aguilar and C. Angulo, "Estabilización robusta de vídeo basada en diferencia de nivel de gris," in *Memorias del VIII Congreso de Ciencia y Tecnología ESPE 2013*, 2013.
9. W. G. Aguilar and C. Angulo, "Robust video stabilization based on motion intention for low-cost micro aerial vehicles," in *Multi-Conference on Systems, Signals Devices (SSD), 2014 11th International*, 2014, pp. 1–6.
10. P. Rudol and P. Doherty, "Human Body Detection and Geolocalization for UAV Search and Rescue Missions Using Color and Thermal Imagery."
11. N. Dalal and W. Triggs, "Histograms of Oriented Gradients for Human Detection," *2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR05*, vol. 1, no. 3, pp. 886–893, 2004.
12. P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: A benchmark," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009*, 2009, pp. 304–311.
13. A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361.
14. B. Benenson, Rodrigo; Omran, Mohamed; Hosang, Jan; Schiele, "Ten Years of Pedestrian Detection, What Have We Learned?," *Proc. Comput. Vision-ECCV 2014 Work.*, pp. 613–627, 2014.
15. P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: an evaluation of the state of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34(4), no. 4, p. 743–61, Apr. 2012, 2012.
16. C. Papageorgiou and T. Poggio, "Trainable system for object detection," *Int. J. Comput. Vis.*, vol. 38, no. 1, pp. 15–33, 2000.
17. G. Overett, L. Petersson, N. Brewer, L. Andersson, and N. Pettersson, "A new pedestrian dataset for supervised learning," *IEEE Intell. Veh. Symp. Proc.*, pp. 373–378, 2008.
18. A. Ess, B. Leibe, K. Schindler, and L. van Gool, "Robust multiperson tracking from a mobile platform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 10, pp. 1831–1846, 2009.
19. C. Wojek, S. Walk, and B. Schiele, "Multi-Cue onboard pedestrian detection," *2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009*, pp. 794–801, 2009.
20. M. Enzweiler and D. M. Gavrilu, "Monocular pedestrian detection: Survey and experiments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2179–2195, 2009.
21. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Conf. Comput. Vis. Pattern Recognit.*, pp. 1–9, 2001.
22. Q. Zhu, S. Avidan, M. C. Yeh, and K. T. Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," *IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2, pp. 1491–1498, 2006.

23. C. Wojek and B. Schiele, "A Performance Evaluation of Single and Multi-feature People Detection," *Jt. Pattern Recognit. Symp.*, pp. 82–91, 2008.
24. X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," *Comput. Vision, 2009 IEEE 12th Int. Conf.*, no. Iccv, pp. 32–39, 2009.
25. Y. Imamura, S. Okamoto, and J. H. Lee, "Human Tracking by a Multi-rotor Drone Using HOG Features and Linear SVM on Images Captured by a Monocular Camera," vol. I, pp. 8–13, 2016.
26. A. Lioulemes, G. Galatas, V. Metsis, G. L. Mariottini, and F. Makedon, "Safety Challenges in using AR . Drone to collaborate with humans in indoor environments," *Proc. 7th Int. Conf. Pervasive Technol. Relat. to Assist. Environ.*, p. 33, 2014.
27. P. Blondel, A. Potelle, C. Pegard, and R. Lozano, "Human detection in uncluttered environments: From ground to UAV view," *2014 13th Int. Conf. Control Autom. Robot. Vision, ICARCV 2014*, pp. 76–81, 1997.
28. M. Andriluka, P. Schnitzspan, J. Meyer, S. Kohlbrecher, K. Petersen, O. Von Stryk, S. Roth, and B. Schiele, "Vision Based Victim Detection from Unmanned Aerial Vehicles," *Intell. Robot. Syst. (IROS), 2010 IEEE/RSJ Int. Conf.*, no. October, pp. 1740–1747, 2010.
29. F. De Smedt, D. Hulens, and T. Goedeme, "On-board real-time tracking of pedestrians on a UAV," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2015–Octob, pp. 1–8, 2015.
30. P. Rudol, P. Doherty, and I. Science, "Human Body Detection and Geolocalization for UAV Search and Rescue Missions Using Color and Thermal Imagery .," *Aerosp. Conf. 2008 IEEE*, pp. 1–8, 2008.
31. A. Gąszczak, T. P. Breckon, and J. Han, "Real-time People and Vehicle Detection from UAV Imagery," *IS&T/SPIE Electron. Imaging*, no. January, pp. 8–11, 2011.
32. J. Sokalski, T. P. Breckon, and I. Cowling, "Automatic Salient Object Detection In UAV Imagery," *Proc. 25th Int. Unmanned Air Veh. Syst.*, p. 11.1-11.12, 2010.
33. M. Siam and M. Elhelw, "Robust autonomous visual detection and tracking of moving targets in UAV imagery," *Int. Conf. Signal Process. Proceedings, ICSP*, vol. 2, no. December, pp. 1060–1066, 2012.
34. L. Wang and D. He, "Texture classification using texture spectrum," *Pattern Recognit.*, 1990.
35. C. P. Papageorgiou and M. Oren, "A general framework for object detection," *Comput. Vision, IEEE Int. Conf.*, vol. 0, no. January, pp. 555–562, 1998.
36. R. E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," *Mach. Learn.*, vol. 37, no. 3, pp. 297–336, 1999.
37. L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. pattern Anal.*, 1998.
38. F. Moosmann, D. Larlus, and F. Jurie, "Learning saliency maps for object categorization," *Work. ...*, 2006.
39. O. Le Meur and T. Baccino, "Methods for comparing scanpaths and saliency maps: strengths and weaknesses," *Behav. Res. Methods*, 2013.