| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 7, July 2020||

Car Detection System in Aerial images with Occlusion Handling

M. Dhanushree[1], Dr. S. Chitrakala[2]

PG Student, Dept. of Computer Science & Engineering, Anna University, Tamil Nadu, India[1]

Professor, Dept. of Computer Science & Engineering, Anna University, Tamil Nadu, India[2]

ABSTRACT: Object detection in aerial images is an emerging research field in computer vision. Nowadays detection of car in aerial image has wide range of applications related to traffic management like traffic accidents, daily traffic jams and better planning of road construction. Car detection in aerial images has challenges such as variation in orientation of car, partial occlusions, variation in illumination, interclass variations. Particularly partial occlusions were not handled separately in existing literatures. In order to overcome the partial occlusion challenge, two features namely Harris corner feature and Local steering kernel feature are extracted. Then Bag of Visual Words (BOVW) model is used to represent the extracted features which is then fed to Support Vector Machine (SVM) for classification. Experimental results proves that the proposed work outperforms the existing methods.

KEYWORDS: Car detection; Local Steering Kernel (LSK) feature; Harris corner feature; Support Vector Machine (SVM); Bag Of Visual Words (BOVW)

I. INTRODUCTION

In recent years, there is a sudden increase in the number of cars in roads and parking lot. Increase in number of cars increases the traffic jams along with air pollution. This leads to the necessity in detecting the car especially in aerial view because the aerial images has better coverage of land when compared to other images. Thus Car detection in aerial images is one of the growing research fields. Car detection application not just includes the traffic management but also has other wide range of applications such as counting the number of cars in parking lot, property assessment, the unusual movement of vehicles in restricted areas and better planning of road construction.

Aerial images has different challenges than the other images and hence normal detection techniques are not suitable for aerial images. The lighting conditions affect the view of car. Sometimes the view of car is partially occluded by trees and canopy [1]. This paper presents a car detection system which uses bag of visual words model to resolve the challenges of aerial images such as inter class variations and partial occlusions.

II. RELATED WORK

In this section, various car detection techniques are reviewed. The work of Makiuchi et al. [2] restricts the search area to asphalt region by screening step. Then feature extraction by scalar invariant feature transform (SIFT) is performed which identifies a set of keypoints. A car is identified by many keypoints which undergo merging to produce “one keypoint- one car”. In the work of ElMikaty et al. [3], region of interest is extracted and sliding window technique is used for detection. A Support Vector Machine (SVM) classifier is used for classification and then a post processing is carried out which eliminates redundant detection of a single car. Moranduzzo et al. [4] used Harris corner detector to extract features followed by extraction of high density features. Finally heavily overlapped features are clustered and a colour based refinements are made for detection. In the work of Moranduzzo et al. [5], vehicles are detected in disastrous areas. This is done by extracting the road first so that localizing the area of search. Then removal of asphalt and building is done followed by the detection of vehicles. Chen Z et al. [6] presented the study of vehicle detection from high-resolution aerial images. A new super pixel segmentation method was proposed to control the segmentation with a low breakage rate.The meaningful patches are extracted based on the centres of the segmented pixels. Gleason et al. [7] selects the candidate region by detecting shadow. Then Harris corner response map is calculated for that candidate region based on which the top N points are selected for highest corner response. Finally rotational invariant local shape features are extracted and a classifier based on Euclidean distance is used to discriminate the objects. In the work of Wang et al. [8], a screening process is done to identify the asphalt areas. Then histogram of gradients (HOG) features are extracted both horizontally and vertically to yield preliminary detection of cars. In these potential car image patches, orientation estimation is done followed by merging of image points identifying the same car. Kembhavi A et al. [9] used colour probability maps to capture the colour statistics of vehicles and their surroundings. He also used HOG and pair of pixels for feature extraction. Then a feature selection methodology is employed to select the desired patches for detection. Hao Sun et al. [10], proposed a new detection framework which is based on spatial sparse coding Bag of Words model. The methodology uses a spatial mapping approach to encode the geometric information. Finally a linear SVM is used to detect the vehicles. Helmut Grabner, et al. [11] used an extensive search method to detect the car from aerial images. In post-processing a mean shift clustering methodology is used in improving the detection rate. Anna Gąszczak [12] presented a vehicle detection approach which used multiple trained cascaded Haar classifiers. Xianbin Cao [13] proposed a computationally efficient algorithm which uses the extracted Histogram of Oriented Gradients (HOG) in various scales for greater classification accuracy. This speeds up the feature extraction process. The work of Ayane Makiuchi [14], is particularly for detecting vehicles in disastrous situations for rescue operation. They first remove the asphalt and building region from a shadow corrected aerial image followed by canny edge detection. By examining the shape features they determine the vehicle along with its position by using projective transformation. Luo-Wei Tsai,[15] proposed a novel approach of colour transformation model which recognizes vehicle pixels from varying background. By doing so, possible candidates are chosen from which corner, edge and coefficients of wavelet transform features are extracted. Then a cascaded multichannel classifier was constructed for detection. Jongmin Yoon [16], proposed a new feature called Order Relation Feature (ORF) particularly for car detection in aerial images which is a code representation of order relation. Feature extraction is followed by the pose classifier and bin-specific weighted linear discriminative analysis which classifies as car. Most of the above mentioned works does not provide a solution for handling partial occlusions in car detection.

III. PROPOSED ALGORITHM

In this work, car detection system of aerial images is proposed which can handle the variations such as type of cars and partial occlusions. Here two types of features namely Local Steering Kernel (LSK) feature and Harris Corner feature are extracted. The LSK feature presented in [17] is used for car detection to effectively capture the structure information. Moreover the Harris Corner feature is exclusively used for handling the partial occlusions which is one of the biggest challenges in aerial image processing. The extracted features are concatenated and represented using Bag of Visual Words (BOVW) model during training. While testing these features are converted to visual words using the constructed BOVW model which are then converted to spatial histogram. The histogram vector is then used for classification using linear SVM.

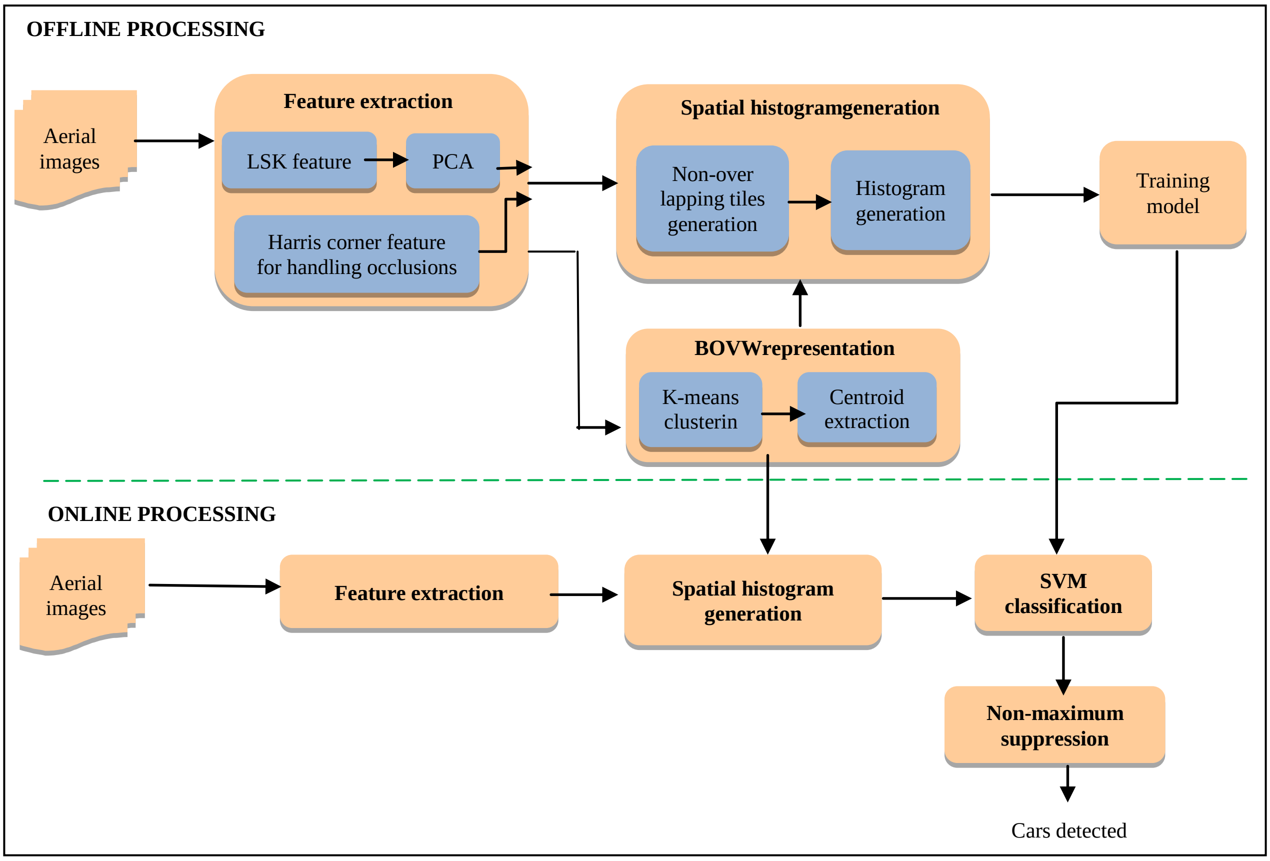

Figure 1. Block diagram of car detection system in aerial images

As shown in Figure 1, basically there are two processing namely offline processing and online processing. Offline processing indicates the training phase while the online processing is the testing of trained machine learning model. There are four modules which are listed below.

- Feature extraction
- BOVW representation
- Spatial histogram generation
- Non-maximum suppression

1. *Feature extraction*

The first step is to extract features from aerial images. Here two features are extracted namely Local steering kernel features and Harris corner feature which are particularly to handle partial occlusions.

a) LSK feature

These are the local descriptors of the image structure which is used to capture both spatial and pixel-value information. They are a non-linear combination of weighted spatial distances between a pixel of an image of size MXN and its surrounding PXP pixels. The distance between an image pixel p and its neighbouring pixel pi is calculated using a weighted Euclidean distance. The weights are the covariance matrix C_i of the image gradients taken along x and y axes. The kernel descriptors are given by the Equation (1).

$$K(x) = \frac{\sqrt{det(C_i)}}{2\pi h^2} exp\left(-\frac{(x_i-x)^T C_i(x_i-x)}{2h^2}\right) \quad (1)$$

Where the covariance matrix Ci is given by the Equation (2)

$$\begin{bmatrix} \sum_{x_j \epsilon w_j} Z_x(x_j).Z_x(x_j) & \sum_{x_j \epsilon w_j} Z_x(x_j).Z_y(x_j) \\ \sum_{x_j \epsilon w_j} Z_y(x_j).Z_x(x_j) & \sum_{x_j \epsilon w_j} Z_y(x_j).Z_y(x_j) \end{bmatrix} \quad (2)$$

The dimensions of the features W is given by the Equation (3)

$$W = [w^1, \ldots\ldots, w^{MN}] \in \mathbb{R}^{P^2 \times MN} (3)$$

Since the features are dense, dimensionality reduction is performed using principle component analysis (PCA) which is given by the Equation (4).

$$V \in \mathbb{R}^{P^2 \times d} (4)$$

Where the d is set to 3 i.e., three principle components are taken.

Algorithm for LSK feature extraction
Input: Aerial image **Output:** LSK feature vectors **Begin** Divide image into 5X5 patches **for** each patch **for** each pixel Find the neighbour cells within the patch Calculate the gradient in both horizontal $Z_x(.)$ and vertical direction $Z_y(.)$ Calculate the covariance matrix Ci of the form given in Equation (2) Calculate the LSK features using the Equation (1) **end** **end** **for** each LSK feature Calculate the normalized feature **end** Apply PCA for the obtained feature vector **End**

b) Harris corner feature extraction

This feature is specially for handling occlusions. Objects when present under a tree is said to be occluded by the tree. When compared to the tree branches, object under consideration has precise corner points. These corner points are extracted using Harris corner feature which computes the corner response function R given by the Equation (5). Then the local maximum of R is taken as corner points.

$$R = Det(H) - k\left(Trace(H)\right)^2 \quad (5)$$

Where H is given by the Equation (6).

$$H(x,y) = \begin{bmatrix} S_{x^2} & S_{xy} \\ S_{xy} & S_{y^2} \end{bmatrix} \quad (6)$$

2. *BOVW representation*

The Bag of Visual words (BOVW) model is used to represent the features. The extracted features are converted to bag of visual words. Every object in an image has certain important features or patterns with which the humans decide as to what the object perceived is. In case of car, the features include windshield, roof, hood, etc.

These features are extracted from all training samples and vector quantization is done using the k-means clustering algorithm with k value set empirically using trial and error method. Here, each cluster represents a feature which can be used to discriminate it from other objects. For each cluster, its corresponding cluster centre is taken and a word is assigned to it. These words i.e. the cluster centres form the bag of visual words. While testing, the extracted features are compared against these bag of visual words. Euclidean distance between the extracted feature and the cluster centres is calculated. The feature is assigned a word for whom the Euclidean distance is minimum.

3. *Spatial histogram generation*

The image is divided into four non-overlapping tiles. Features are extracted from these tiles and visual word is assigned as mentioned in previous step. Now the features are used to generate histogram in each non-overlapping tiles. These histograms are concatenated to form a single feature vector which is fed to SVM for training.

The purpose of spatial histogram is to preserve the structure of the object of interest. By dividing the image into non-overlapping tiles and then concatenating the histograms obtained from those tiles, the structure of the object is preserved. The formation of spatial histogram is shown in Figure 2.

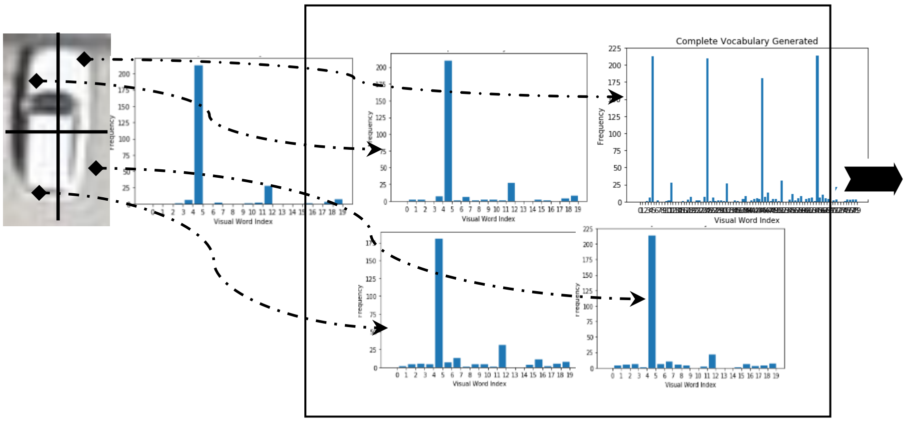

Figure 2. Spatial histogram

4. *Non-maximum suppression*

Since window based scanning method is used, there may be multiple responses for a single object. In order to avoid this non-maximum suppression is implemented which results in single bounding box for each object. SVM output consists of labels and the score. Based on the score the bounding boxes are sorted and selected greedily i.e., the boxes having the maximum score. Then overlapping ratio is calculated which is also used to discard the spurious bounding boxes having ratio more than 30 percentage.

The results of non-maximum suppression are shown in the Figure 3. As we can see that multiple responses of a car is reduced to a single response.

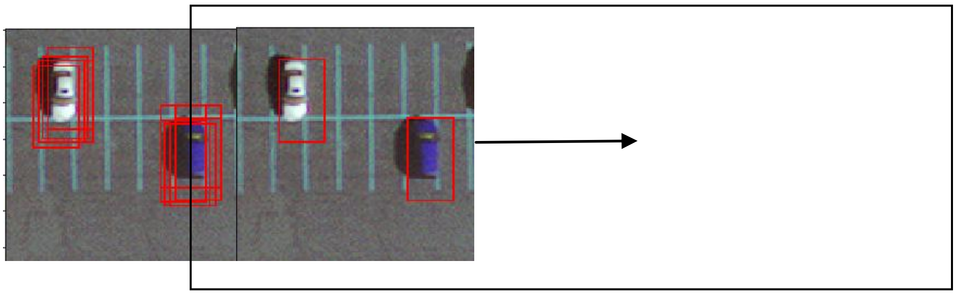

Figure 3. Non-maximum suppression

IV. EXPERIMENTAL RESULTS

A. Dataset description

Dataset for the proposed work should contain enough samples to address the problems of aerial images such as variation in illumination, orientation, type of object and partial occlusions. Thus two datasets were used namely Overhead Imagery Research Data Set (OIRDS) and Vehicle Detection in Aerial Imagery (VEDAI) for object detection.

OIRDS: This dataset [18] is focused on Automatic Target Detection (ATD) task for passenger vehicles. It consists of nearly 400 aerial images with approximately 1800 objects such as cars, trucks and vans.

VEDAI: It [19] consists of 300 aerial images of 1024X1024 dimensions. It is especially created for vehicles. The objects are small enough having variety of vehicles and has partially occluded objects. Thus the dataset has all challenges that can be addressed. Sample data shown in Figure 4 and Figure 5 which is from OIRDS and VEDAI dataset respectively.

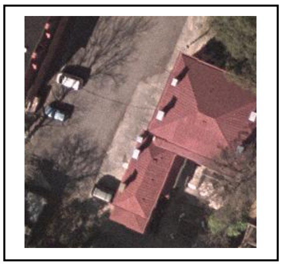

Figure 4. OIRDS

Figure 5. VEDAI

B. Metrics used

The following metrics are used for evaluating the performance of the proposed object detection system.

- Precision
- Recall
- F-Measure

Precision - Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$\boldsymbol{Precision = TP/(TP+FP)} \qquad (7)$$

Using Equation (7), precision gives the measure of ratio of the correctly classified patches over the total detected cars.

Recall (Sensitivity) - Recall is the ratio of correctly predicted positive observations to the all observations in actual class as in Equation (8).

$$\boldsymbol{Recall = TP/(TP+FN)} \qquad (8)$$

F score - F Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account as in Equation (9).

$$\boldsymbol{F\ Score = 2*(Recall * Precision)\ /\ (Recall + Precision)} (9)$$

TN / True Negative: case was negative and predicted negative

TP / True Positive: case was positive and predicted positive

FN / False Negative: case was positive but predicted negative

FP / False Positive: case was negative but predicted positive

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 7, July 2020||

C. Output

The aerial images are pre-processed, the necessary features are extracted and fed to SVM for training. For localization of the car, sliding window approach is used. Multiple responses for a single car is reduced using non-maximum suppression. Table 1 shows the input and the output of the occlusion aware car detection system in aerial images.

Table 1. Output of the car detection system in aerial images

S.NO.	Type of variation	Input aerial image	Output image with cars detected
1.	**Variation in type of object**		
2.	**Partial occlusions**		

The aerial images are subject to challenges such as variation in type of object and partial occlusions. The corresponding input and output are shown in the Table 1. The first challenge is that aerial images is that there may be different types of car models with varying sizes, shapes and colour. This is addressed by the BOVW model which represents the car as distinct features such as windshield, roof, hood, etc.

Next, one of the major challenges of aerial images is handling the occlusions. Objects may be covered by other objects such as canopy, tree branches, etc. These are addressed by the feature extraction methods which helps to find the partially occluded cars effectively.

The above mentioned test cases are taken and the evaluation is done for each test case based on precision, recall and f-score. For each test case, the ground truth along with false positive and true positive are shown in Table 2.

Table 2. Performance score

Test cases	Precision (%)	Recall (%)	F score (%)
Variation in type of object	89.28	94.34	91.74
Partial occlusions	86.95	80.00	83.33

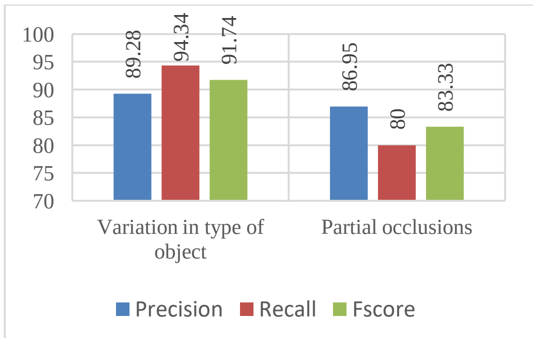

Figure 6. Test cases analysis

Thus from the Figure 6, it can be seen that the test cases are evaluated individually with the metrics such as precision, recall and F score. The test case, variation in type of object has a high F score of 91.74 % due to the extraction and appropriate representation of vital features from the object car. Other test case partial occlusions also showed a good performance with the F score of 83.33 %. It is inferred that the proposed car detection system proves its strength in detecting a car in aerial images.

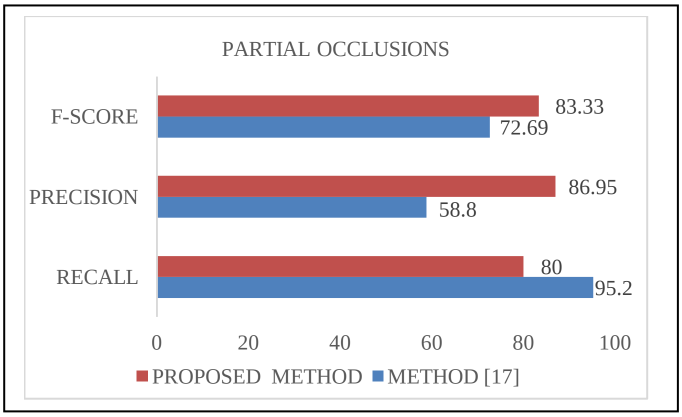

Figure 7. Result comparison for partial occlusions

In particular the combination of Harris corner feature and LSK feature extraction handles the partial occlusions effectively. The Figure 7 shows the comparison of performance measure of existing method [17] and proposed system for occlusion handling. The proposed detection method performs better in terms of occlusion handling with an increase in 28.15% percentage of accuracy and an increase of 10.64 % increase in F-score when compare to method [17]. Thus the proposed car detection system in aerial images has high detection rate when compared to the existing system.

V. CONCLUSION AND FUTURE WORK

Thus the proposed car detection system successfully detects the car present in aerial images despite the challenges such as variation in type of objects and partial occlusions in aerial images. It gives better performance when compared to existing models in occlusion handling. The future work aims at further enhancement of the car detection system is to achieve robustness of the system and also increase its overall performance.

REFERENCES

1. Zhao, Tao, and Ram Nevatia. "Car detection in low resolution aerial images." Image and Vision Computing 21, no. 8 (2003): 693-703.
2. Makiuchi, Ayane, and Hitoshi Saji. "Vehicle Detection Using Aerial Images in Disaster Situations." In International Conference on Global Research and Education, pp. 187-194. Springer, Cham, 2018.
3. El Mikaty, Mohamed, and Tania Stathaki. (2017), 'Car Detection in Aerial Images of Dense Urban Areas', IEEE Transactions on Aerospace and Electronic Systems 54.1: 51-63.
4. Moranduzzo, Thomas, and Farid Melgani. (2014), 'Detecting cars in UAV images with a catalog-based approach', IEEE Transactions on Geoscience and Remote Sensing, 52.10: 6356-6367.
5. Moranduzzo, T., & Melgani, F. (2013), 'Automatic car counting method for unmanned aerial vehicle images'. IEEE Transactions on Geoscience and Remote Sensing, 52(3), 1635-1647.
6. Chen Z, Wang C, Wen C, Teng X, Chen Y, Guan H, Luo H, Cao L, Li J. (2015), 'Vehicle detection in high-resolution aerial images via sparse representation and superpixels', IEEE Transactions on Geoscience and Remote Sensing, 54(1):103-16.
7. Gleason, Joshua, et al. (2011), 'Vehicle detection from aerial imagery', IEEE International Conference on Robotics and Automation, pages 2065-2070.
8. Wang, Sheng. (2011), 'Vehicle detection on aerial images by extracting corner features for rotational invariant shape matching', IEEE 11th International Conference on Computer and Information Technology, pages 171-175.
9. Kembhavi A, Harwood D, Davis LS. (2011), 'Vehicle detection using partial least squares', IEEE Transactions on Pattern Analysis and Machine Intelligence,33(6):1250-65.

10. Sun, H., Sun, X., Wang, H., Li, Y., & Li, X. (2011), ‘Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model’ IEEE Geoscience and Remote Sensing Letters, 9(1), 109-113.
11. Grabner, Helmut, Thuy Thi Nguyen, Barbara Gruber, and Horst Bischof, ‘On-line boosting-based car detection from aerial images.’ ISPRS Journal of Photogrammetry and Remote Sensing 63, no. 3 (2008): 382-396.
12. Gaszczak, Anna, Toby P. Breckon, and Jiwan Han. ‘Real-time people and vehicle detection from UAV imagery.’ In Intelligent Robots and Computer Vision XXVIII: Algorithms and Techniques, vol. 7878, p. 78780B. International Society for Optics and Photonics, 2011.
13. Cao, Xianbin, Changxia Wu, Jinhe Lan, Pingkun Yan, and Xuelong Li. ‘Vehicle detection and motion analysis in low-altitude airborne video under urban environment.’ IEEE Transactions on Circuits and Systems for Video Technology 21, no. 10 (2011): 1522-1533.
14. Makiuchi, Ayane, and Hitoshi Saji. ‘Vehicle Detection Using Aerial Images in Disaster Situations.’ In International Conference on Global Research and Education, pp. 187-194. Springer, Cham, 2018.
15. Tsai, Luo-Wei, Jun-Wei Hsieh, and Kao-Chin Fan. ‘Vehicle detection using normalized color and edge map.’ IEEE International Conference on Image Processing 2005. Vol. 2. IEEE, 2005.
16. Yoon, Jongmin, InHan Kim, WonTaek Chung, and Daijin Kim. ‘Fast and accurate car detection in drone-view.’ In 2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), pp. 1-3. IEEE, 2016.
17. Zhou, Hailing, et al. (2018), ‘Robust vehicle detection in aerial images using bag- of-words and orientation aware scanning’, IEEE Transactions on Geoscience and Remote Sensing 56.12:7074-7085.
18. Razakarivony, Sebastien, and Frederic Jurie. ‘Vehicle detection in aerial imagery: A small target detection benchmark.’ Journal of Visual Communication and Image Representation 34 (2016): 187-203.
19. F. Tanner et al., ‘Overhead imagery research data set — an annotated data library & tools to aid in the development of computer vision algorithms,’ 2009 IEEE Applied Imagery Pattern Recognition Workshop (AIPR 2009), Washington, DC, 2009, pp. 1-8