ISSN: 2251-9246
EISSN: 2345-6221

pp. 127:133

Two-level Ensemble Deep Learning for Traffic Management using Multiple Vehicle Detection in UAV Images

Zeinab Ghasemi Darehnaei[1], Seyed Mohammad Jalal Rastegar Fatemi*[2],

Seyed Mostafa Mirhassani[3], Majid Fouladian[4]

[1,2,4] Department of Electrical Engineering, College of Engineering, Islamic Azad University Saveh Branch, Saveh, Iran
[3] Department of Electrical Engineering, Islamic Azad University Shahrood Branch, Shahrood, Iran
* mohammadjalalrastegarfatemi@gmail.com

Abstract

Environmental monitoring via vehicle detecting using unmanned aerial vehicle (UAV) images is a challenging task, due to small-size, low-resolution, and large-scale variation of the objects. In this paper, a two-level ensemble deep learning (named 2EDL) based on Faster R-CNN (regional-based convolutional neural network) introduced for multiple vehicle detection in UAV images. We use three CNN models (VGG16, ResNet50, and GoogLeNet) that have already pre-trained on huge auxiliary data as feature extraction tools, combined with five learning models (KNN, SVM, MLP, C4.5 Decision Tree, and Naïve Bayes), resulting 15 different base learners in two levels. The final class is obtained via a majority vote rule ensemble of these 15 models into five vehicle classes (car, van, truck, bus, and trailer) or "no-vehicle". Simulation results on the AU-AIR dataset of UAV images show the superiority of the proposed 2EDL technique against existing methods, in terms of the total accuracy, and FPR-FNR trade-off.

Keywords: Deep transfer learning; ensemble learning; multiple object detection; unmanned aerial vehicles.
Article history: Received 06-July-2021; Revised 12-July-2021; Accepted 18-Aug-2021. Article Type: Research Paper

http://dx.doi.org/10.30495/ijsee.2021.684014

1. Introduction

Unmanned aerial vehicles (UAVs) are cost effective tools, widely used for capturing remote sensing images, especially for traffic monitoring [1-3]. With recent price drop of the UAV products, these tools are becoming more prominent in transportation safety, planning, and management [4]. Advantages of UAVs against ground-based traffic sensors include mobility, wide field of view, fast and on-demand imaging, great maneuverability, safety, and zero impact on the ground traffic [2,5].

To apply UAVs in traffic monitoring, an essential task is vehicle detection, which is challenging due to varying illumination conditions, background motions due to UAV movements, complicated scenes, and different traffic conditions [6]. Many classical and machine learning based object detection methods have been applied for vehicle detection in UAV images. These techniques mainly rely on handcrafted features for building the classification system. Other types of methods based on deep learning have recently significantly outperformed the handcrafted features.

Deep learning, which also known as feature learning, is based on learning high-representative features automatically from the input data (e.g., image). Convolutional neural network (CNN) is the most widely used deep learning for object detection [7-9]. In the case of vehicle detection, Chen et al. [10] proposed a method based on sliding windows and deep CNN. The idea behind this technique is to replicate the convolution layers at different scales, allowing the deep network to detect cars at different scales. It requires high computation time to be trained for car detection using GPU. Directly combining CNN with sliding windows has some difficulties to precisely localize objects [11]. To address this issue, Sppnet [12], region-based CNN (R-CNN) [13], and Fast R-CNN, have been proposed. However, the region proposal generation step consumes too much computation time. Ren et al. further improved Fast R-CNN [14] and developed Faster R-CNN [15], which achieves state-of-the-date object detection accuracy with real-time detection speed.

In this paper, a two-level ensemble deep transfer learning (named 2EDL) is introduced for multiple vehicle detection in UAV images. Inspired by the success of Faster R-CNN in object detection, we apply it for multiple vehicle detection in UAV images. The proposed 2EDL model contributes to the existing methods by introducing a two-level ensemble deep transfer learning model based on Faster R-CNN comprising three feature extractors and five classifiers, resulting a two-level ensemble of 15 base learners. In the proposed 2EDL model, three Faster R-CNN models (VGG16, ResNet50, and GoogLeNet) which have pre-trained on huge auxiliary data are trained on a UAV dataset using five learning models (KNN, SVM, MLP, C4.5 Decision Tree, and Naïve Bayes). The final output of the 2EDL is obtained via a majority vote ensemble of these 15 base learners, which effectively reduces the detection error with less sensitivity to noise through aggregation of different feature extraction models and classification methods.

The rest of this paper is organized as follows: Section 2 reviews the literature on vehicle detection. Section 3 provides the detail of Faster R-CNN for vehicle detection. The proposed 2EDL model for multiple vehicle detection is presented in Section 4. Simulation results are provided in Section 5, and finally, Section 6 concludes the paper with some future research directions.

2. Literature review

There are many object detection techniques in literature, which have been applied for vehicle detection in UAV images, e.g., Viola-Jones (VJ) object detection algorithm [16] and Discriminatively Trained Part Based Models (DPM) [17]. Generally, these techniques are less sensitive to image noise and complex scenarios, and thus, are more robust for vehicle detection. However, most of these techniques are very sensitive to object rotation. Moreover, some methods such as VJ suffer from high sensitivity to illumination changes.

There are different methods in literature, which have used machine learning for vehicle detection. In the case of UAV image, car detection introduces more challenges than other object detection problems, due to the extremely high resolution of the UAV images. Moranduzzo and Melagni [18] proposed a scale-invariant feature transform (SIFT) for detection of the interest points of cars, in which, support vector machine (SVM) is used to classify the extracted interest points into “car” and “no-car”, based on the SIFT model descriptor. Finally, the SIFT points belonging to the same car are merged together to represent a single car. Moranduzzo et al. [19] developed a method based on higher-order gradients and gaussian process regression for object detection, and applied it to the problem of car detection in UAV images in an urban environment. Xu et al. [20] combined the SVM classifier with VJ (SVM+VJ) and with histograms of gradient (SVM+HOG) for car detection in UAV street videos.

Recently, different deep learning techniques have been presented for vehicle detection. Perez et al. [21] introduced a traditional object detection method based on the sliding window technique utilizing CNN. The main problem of this method is the time-consuming procedure of the sliding window strategy to handle multiscale object detection. Chen et al. [8] proposed a CNN-based deep learning for vehicle detection in satellite images. To bypass the problem of extracting a huge number of regions in traditional CNNs, Girshick et al. [13] proposed a selective search to extract a subset of regions from the image, also called region proposals. Ammour et al. [22] presented a three-stage method for car detection, including generating candidate regions, feature extraction, and classification. At first, mean-shift algorithm [23] is used to segment images, and then, fine-tuned VGG16 model [24] is applied to extract region features. Finally, SVM is applied to classify the selected regions. The procedure of generating candidate regions is time-consuming. Moreover, different models should be trained for the three stages, which increases the complexity of this model. Bazi et al. [25] developed a convolutional support vector machine network model (CSVM) for the vehicle detection in UAV images. The CSVM is based on several alternating convolutional and reduction layers, which ended by a linear SVM classifier.

To overcome the time-consuming process of region proposal generation in the R-CNN techniques, Ren et al. [15] presented Faster R-CNN as an appealing solution. There are different techniques based on the Faster R-CNN, which have been presented for vehicle detection from low-altitude UAV images [26-35]. Inspired by the success of the Faster R-CNN in terms of both detection accuracy and speed, this paper also utilizes the Faster R-CNN technique to detect multiple vehicles from UAV imagery.

3. Vehicle detection using Faster R-CNN

Faster R-CNN consists of two units: Regional Proposal Network (RPN) and Fast R-CNN detector (FR-CNN), as seen in Fig. 1. The RPN is a fully convolutional network for generating region proposals (candidate regions for object) with a wide range of scales and aspect ratios, which would be fed into the second unit. Proposals are rectangular regions, which may or may not contain target objects. The FR-CNN is used to refine the proposals. The FR-CNN runs through the CNN only once for the entire input image and then refines object proposals.

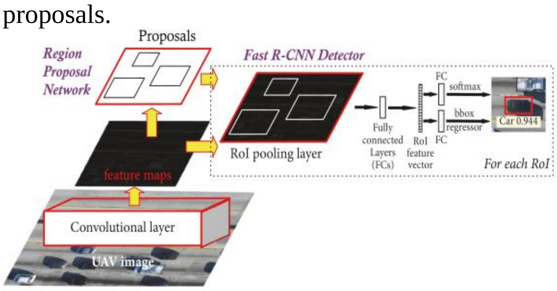


Fig. 1. Architecture of object detection by Faster R-CNN.

A) Region proposal network

The RPN takes a UAV image as input, and outputs a set of bounding boxes, each with an object probability score for each class. In this paper, VGG16, ResNet50, and GoogLeNet, are used as the Faster-RCNN convolutional backends. The RPN utilizes sliding windows over the convolutional feature map output by the last convolutional layer, to generate rectangular region proposals for each position. As many proposals highly overlap with each other, non-maximum suppression (NMS) is applied to merge proposals that have high intersection-over-union (IoU). After NMS, the remaining proposals are ranked based on the object probability score, and only the top N proposals are used for detection.

B) FR-CNN detector

The FR-CNN takes multiple regions of interest (RoIs) as input. For each RoI, RoI pooling layer from the convolutional layer extracts a fixed-length feature vector. Each feature vector is fed into a sequence of fully connected layers. The final outputs of the detector through softmax and bounding-box regressor layers include softmax probabilities which estimate over K object classes (5 in this paper), and the related bounding-box values.

C) Faster R-CNN training

For training RPNs, each proposal is assigned a K+1 class label which indicates whether the proposal is an object from K vehicle types or not (background). Since both RPN and FR-CNN networks share the same convolutional layers, they can be trained jointly to learn a unified network at the same time. The learning process follows four steps to train the Faster R-CNN model [4]:

– Train the RPN network as described above.
– Train the detector network using region proposals generated by the RPN trained at the first step.
– Initializing the RPN training by the detector, but only train the specific layers of the RPN.
– Train the detector network model using new region proposals of the RPN.

4. Proposed 2EDL for multiple vehicle detection

Generally, each deep learning and classification model has its own advantages in detecting different objects, and consequently, the different models have different performances in terms of the FPR-FNR trade-off. As a result, an ensemble of different learning models may result better performance by aggregation of the results of the different base learners [36].

A) Transfer learning model

Since the training process in the Faster R-CNN requires a large number of data samples and consumes huge computation time, transfer learning [37] is used to construct the vehicle detection model in this paper. The transfer learning is able to take less time to build a higher-precision model, i.e., transfer learning begins not from scratch, but from previous models that have addressed various problems.

In the proposed 2EDL model, the convolutional base of the existing Faster R-CNN model is used as the feature extractor and add a new classifier to the top of the redefined model. For example, the pretrained VGG16 has 1000 different class labels as the output, however, our newly model corresponding to five specific object classes (car, van, truck, bus, and trailer) as the output. The parameters of the new classifier can be trained using sample data with the labeled UAV images. Framework of the vehicle detection using a single transfer learning can be seen in Fig. 2, which is composed of the convolutional base of a pre-trained R-CNN as the feature extractor, and a newly defined classifier in the top-level as the output.

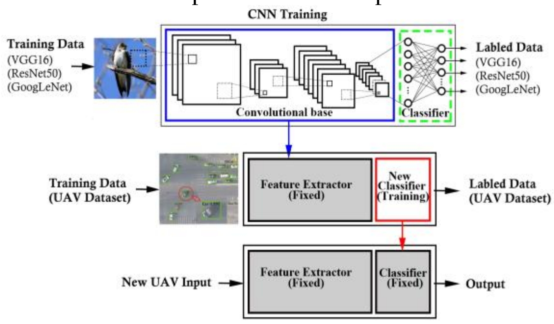


Fig. 2. Framework of transfer learning for a single base learner.

To obtain a transfer-learning model, three steps should be performed, as:

- Several models pretrained on the ImageNet are selected (VGG16, ResNet50, or GoogLeNet).
- The convolution base of these models is taken out in turn as the feature extractor, and the newly defined classifier layer (KNN, SVM, MLP, C4.5 DT, or Bayes) is connected with it to construct a new Faster R-CNN model.
- The new model is trained with the new dataset, i.e., sample UAV image dataset. During the training process, the weights of the convolutional base are frozen, and only the parameters of the added classifier layer are trained.

B) Ensemble learning model

Vehicle detection in the 2EDL model is done using an ensemble deep transfer learning in two levels, as shown in Fig. 3. The first level includes 3 pre-trained CNN-based feature extractors (VGG16, ResNet50, and GoogLeNet), while the second level comprises 5 classifiers (KNN, SVM, MLP, C4.5 DT, and Bayes), which results 3×5=15 base learners to be obtained. Each base learner is trained separately, and the trained learners are used to aggregate their results for vehicle detection in each region proposal for any new UAV image in the test dataset. The final object score of the region proposal *k* can be obtained by majority vote of the output class of the 15 base learners.

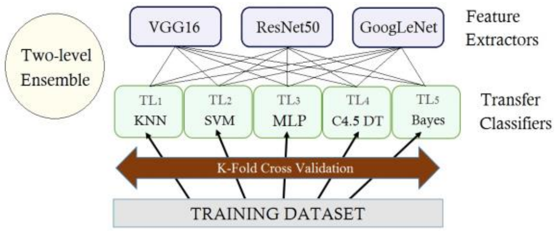


Fig. 3. Two-level ensemble structure of the 2EDL model.

At the training stage, N_1^{Train} cropped target objects from five types of vehicles (car, van, truck, bus, and trailer) in different orientations, as well as N_0^{Train} class of "background" are extracted from the different UAV images. These cropped images include label 0 for background (no-vehicle), and labels 1 to 5 for the different object classes including car, van, truck, bus, and trailer, respectively. The cropped images are used as the same for the transfer learning of all base learners within 3×5 ensemble structure of the 2EDL. After training of the 2EDL model, the trained model can be used for the real-time multiple vehicle detection in new unseen UAV images.

5. Simulation results

A) UAV Dataset: AU-AIR

To evaluate the performance of the proposed 2EDL model, it is applied for vehicle detection in a recently published UAV data, named AU-AIR [37]. It is a multi-modal aerial dataset captured by a UAV. Having visual data, object annotations, and flight data (time, GPS, altitude, IMU data, velocities), the AU-AIR meets vision and robotics for UAVs. It is accessible in https://bozcani.github.io/auairdataset.

B) Simulation Settings

As mentioned above, three pre-trained CNN are used as feature extractors of the proposed vehicle detection model. These feature extractors include VGG16, ResNet50, and GoogLeNet. Moreover, five classifiers including KNN, SVM, MLP, C4.5 DT, and Naïve Bayes, are employed as the transfer learning classifiers. To train each base transfer learning model of the 2EDL, a subset comprising N1Train=1000 cropped objects from five classes and N0Train=2000 background is used, which results 3000 training samples in total. After transfer learning of 15 base learners on the training dataset in an offline scheme, the final tuned 2EDL model is used for the online vehicle detection in new unseen UAV images. To evaluate the performance of the 2EDL model, it is tested on 300 new UAV images including 1252 vehicle objects.

Performance measures are considered as accuracy, precision, and recall. Precision is the ability of the model to not label negative samples as positive. On the other hand, recall is the ability of the model to find positive samples (i.e., vehicle objects). The less FPR (FNR), the more precision (recall). Accuracy, precision and recall for class c can be expressed as:

$$\text{Accuracy}_c = \frac{TP_c}{TP_c + FP_c + FN_c} \tag{1}$$

$$\text{Precision}_c = \frac{TP_c}{TP_c + FP_c} \tag{2}$$

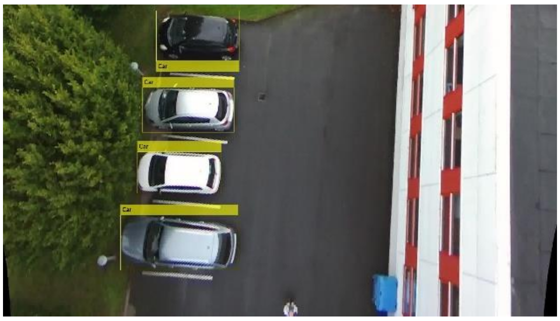
$$\text{Recall}_c = \frac{TP_c}{TP_c + FN_c} \tag{3}$$

where TP_c is the number of positive objects of vehicle class c correctly identified, FP_c is the number of regions incorrectly identified as vehicle class c, and FN_c is the number of vehicles class c which have not identified as object class c.

C) Simulation Results

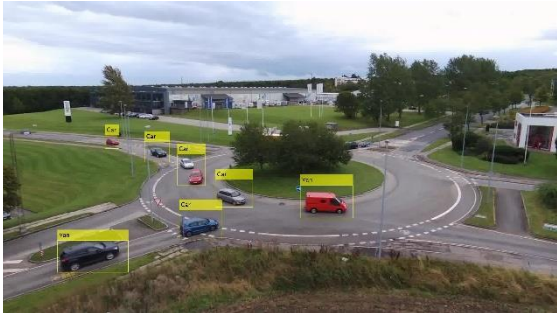
As mentioned above, the proposed 2EDL model was trained on 3000-cropped objects (1000 vehicle objects from the five object classes, and 2000 background objects). The trained model was tested on 300 new UAV images containing 1252 different vehicle objects in five classes. The results of the 2EDL for some test images are shown in Fig. 4. In Fig. 4 (a), totally 5 vehicles have been identified with label “car”. In Fig. 4 (b), 10 cars, a bus, a trailer, and a truck have been detected. In Fig. 4 (c), 5 cars and 2 vans have been identified. Finally, in Fig. 4 (d), 3 cars and a truck have been detected. The results in Fig. 4 demonstrate the effectiveness of the 2EDL to detect multiple objects. The model can detect even semi-distorted vehicles, e.g., the car at the right-side of the Fig. 4 (b).

The proposed 2EDL model has different error types for some UAV images. Generally, the closer the UAV to the target object, as in Fig. 4 (a), the more accuracy. For example, in Fig. 4 (c) and (d), the vehicles at the top left of the image were ignored, which led to false negative. An influential parameter is the field of view of the UAV and the orientation of the UAV against the object, which highly affects the detection accuracy. As an example, in Fig. 4 (b), the model has detected some cars parked at the top left of the image. However, these cars have not been detected in Figs. 4 (c) and (d), because of the less height of the UAV camera against the ground. Another type of error is the merging error, in which, two or more vehicles may be merged together and detected as a single object. An example to this type of error can be seen in Fig. 5. In this figure, there are two merging error, in which, two cars were merged into a single object “car”, a pair at the center of the image, and a pair at the top left of the image.



(a)

(b)



(c)

(d)

Fig. 4. Sample results of the 2EDL model on some UAV images.

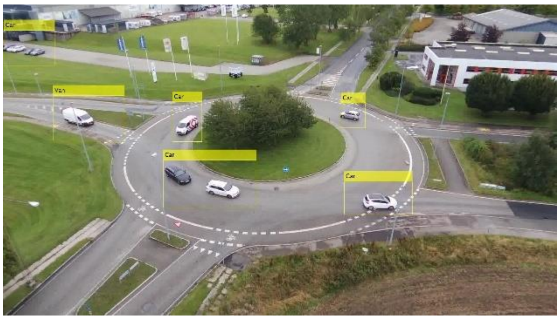


Fig. 5. Merging error.

D) *Comparison with single Faster R-CNN models*

In this section, the performance measures of the proposed ensemble deep transfer learning model (2EDL) on test dataset are compared with those of achieved by single base learners. Comparison of the 2EDL with different pre-trained CNN models (considering MLP as classifier) can be summarized as Table 1. Moreover, comparison of the 2EDL with different transfer classifiers (considering VGG16 as feature extractor) are provided in Table 2. The obtained results clearly show the effectiveness of the proposed eansemble model to improve the total detection accuracy, precision, and recall.

E) *Comparison with single Faster R-CNN models*

In this section, the performance of the 2EDL is compared with a classical method (Viola-Jones) [16], a machine learning technique (HOG+SVM) [20], and a deep learning model (SW-CNN) [22]. Comparison of the 2EDL with these methods is summarized in Table 3, and graphically illustrated in Fig. 9. According to the obtained results, the proposed 2EDL model outperforms the existing techniques, in terms of the total accuracy, precision, and recall.

6. Conclusion

In this paper, a two-level ensemble deep learning (2EDL) has been presented for the multiple vehicle detection in UAV images. The 2EDL is based on the Faster R-CNN object detection technique, in which, a set of region proposals are extracted using the RPN, and then, CNN is used to classify the region proposals into five vehicle classes (car, van, truck, bus, and trailer), or no object (background). In the proposed transfer-learning model, three pre-trained CNNs (VGG16, ResNet50, and GoogLeNet) have been used as feature extractors, combined with five classifiers (KNN, SVM, MLP, C4.5 Decision Tree, and Naïve Bayes). Simulation results on a recently published UAV dataset (AU-AIR) have demonstrated the superiority of the 2EDL model against the existing methods in terms of accuracy, precision, and recall. In the proposed 2EDL model, the different base learners are aggregated via a majority vote ensemble. As a future work, weighted averaging ensemble can be used, in which, the weights of the different base learners are optimized via a metaheuristic algorithm, e.g., genetic algorithm.

Table.1.
Comparison of 2EDL with Different Single CNNs (Considering MLP as Classifier)

Detection Method	*Accuracy*	*Precision*	*Recall*
VGG16	80.4	91.2	83.7
ResNet50	78.6	92.5	80.9
GoogLeNet	79	90.8	82.7
2EDL (Proposed)	85.3	95.6	88.8

Table.2.
Comparison of 2EDL with Different Classifiers (Considering VGG16 as Feature Extractor).

Detection Method	*Accuracy*	*Precision*	*Recall*
KNN	77.5	87.9	81.7
SVM	82.8	92.7	84.9
MLP	80.4	91.2	83.7
C4.5 Decision Tree	79.5	89.8	83.9
Naïve Bayes	77.2	90.2	79.4
2EDL (Proposed)	85.3	95.6	88.8

Table.3.
Comparison of 2EDL with Existing Techniques.

Detection Method	*Accuracy*	*Precision*	*Recall*
Viola-Jones [16]	72.7	87.3	81.4
HOG+SVM [20]	74.4	88.4	82.4
SW-CNN [22]	81.5	92.2	83.9
2EDL (Proposed)	85.3	95.6	88.8

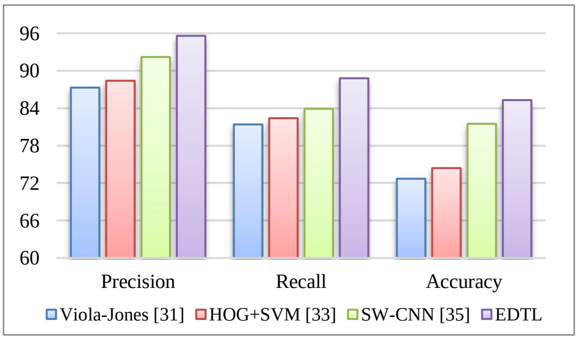


Fig. 6. Comparison of the 2EDL with the existing techniques.

References

[1] A. Angel, M. Hickman, P. Mirchandani, and D. Chandnani, "Methods of analyzing traffic imagery collected from Aerial platforms," IEEE Transactions on Intelligent Transportation Systems, vol. 4, no. 2, pp. 99–107, 2003.

[2] Shokouhifar, M., & Jalali, A. (2017). Optimized sugeno fuzzy clustering algorithm for wireless sensor networks. Engineering applications of artificial intelligence, 60, 16-25.

[3] J. Leitloff, D. Rosenbaum, F. Kurz, O.Meynberg, andP. Reinartz, "An operational system for estimating road traffic information from aerial images," Remote Sensing, vol. 6, no. 11, pp. 11315–11341, 2014.

[4] Xu, Y., Yu, G., Wang, Y., Wu, X., & Ma, Y. (2017). Car detection from low-altitude UAV imagery with the faster R-CNN. Journal of Advanced Transportation, 2017.

[5] Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS J. Photogramm. Remote Sens. 2014, 92, 79–97.

[6] G. Salvo, L. Caruso, and A. Scordo, "Urban traffic analysis through an UAV". Procedia-Social and Behavioral Sciences, 111,1083-1091, 2014.

[7] Huang, G. L., He, J., Xu, Z., & Huang, G. (2020). A combination model based on transfer learning for waste classification. Concurrency and Computation: Practice and Experience, e5751.

[8] El Mostafa Hambi, F. B. A New Online Plagiarism Detection System based on Deep Learning. International Journal of Advanced Computer Sciences and Applications, 11(9), 470-478.

[9] Shustanov, A., & Yakimov, P. (2017). CNN design for real-time traffic sign recognition. Procedia engineering, 201, 718-725.

[10] X. Chen, S. Xiang, C.-L. Liu, and C.-H. Pan, "Vehicle detection in satellite images by hybrid deep convolutional neural networks," IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 10, pp. 1797–1801, 2014.

[11] P. Sermanet, K. Kavukcuoglu, S. Chintala, and Y. Lecun, "Pedestrian detection with unsupervisedmulti-stage feature learning," in Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13), pp. 3626–3633, IEEE, June 2013.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1904–1916, 2015.

[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587, Columbus, Oh, USA, 2014.

[14] R. Girshick, "Fast R-CNN," in Proceedings of the 15th IEEE International Conference on Computer Vision (ICCV '15), pp. 1440– 1448, December 2015.

[15] Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: towards real-time object detection with region proposal networks. IEEE transactions on pattern analysis and machine intelligence, 39(6), 1137-1149.

[16] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 511–518, December 2001.

[17] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, pp. 1627–1645, 2010.

[18] Zhao, T.; Nevatia, R. Car detection in low resolution aerial image. In Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV 2001), Washington, WA, USA, 7–14 July 2001.

[19] Moranduzzo, T.; Melgani, F. Automatic Car Counting Method for Unmanned Aerial Vehicle Images. IEEE Trans. Geosci. Remote Sens. 2014, 52, 1635–1647.

[20] Xu, Y.; Yu, G.; Wang, Y.; Wu, X.; Ma, Y. A Hybrid Vehicle Detection Method Based on Viola-Jones and HOG + SVM from UAV Images. Sensors 2016, 16, 1325.

[21] A. Perez, P. Chamoso, V. Parra, and A. J. Sanchez, "Ground vehicle detection through aerial images taken by a UAV," in Proceedings of the 17th International Conference on Information Fusion, (FUSION '14), July 2014.

[22] N. Ammour, H. Alhichri, Y. Bazi, B. Benjdira, N. Alajlan, and M. Zuair, "Deep learning approach for car detection in UAV imagery," Remote Sensing, vol. 9, no. 4, p. 312, 2017.

[23] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 603–619, 2002.

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Computer Science, 2014.

[25] Bazi, Y., & Melgani, F. (2018). Convolutional SVM networks for object detection in UAV imagery. Ieee transactions on geoscience and remote sensing, 56(6), 3107-3118.

[26] Uto, K.; Seki, H.; Saito, G.; Kosugi, Y. Characterization of Rice Paddies by a UAV-Mounted Miniature Hyperspectral Sensor System. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2013, 6, 851–860.

[27] Bazi, Y., & Melgani, F. (2018). Convolutional SVM networks for object detection in UAV imagery. Ieee transactions on geoscience and remote sensing, 56(6), 3107-3118.

[28] Li, Y., Melgani, F., & He, B. (2019, July). Fully Convolutional SVM for Car Detection In Uav Imagery. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium (pp. 2451-2454). IEEE.

[29] Wang, L., Liao, J., & Xu, C. (2019, February). Vehicle detection based on drone images with the improved Faster R-CNN. In Proceedings of the 2019 11th International Conference on Machine Learning and Computing (pp. 466-471).

[30] Espinosa, J. E., Velastin, S. A., & Branch, J. W. (2017, November). Vehicle detection using alex net and faster R-CNN deep learning models: a comparative study. In International Visual Informatics Conference (pp. 3-15). Springer, Cham.

[31] Mittal, U., Potnuru, R., & Chawla, P. (2020, June). Vehicle Detection and Classification using Improved Faster Region Based Convolution Neural Network. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 511-514). IEEE.

[32] Sommer, L., Schumann, A., Schuchert, T., & Beyerer, J. (2018, March). Multi feature deconvolutional faster r-cnn for precise vehicle detection in aerial imagery. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) (pp. 635-642). IEEE.

[33] Ji, H., Gao, Z., Mei, T., & Li, Y. (2019). Improved faster R-CNN with multiscale feature fusion and homography augmentation for vehicle detection in remote sensing images. IEEE Geoscience and Remote Sensing Letters, 16(11), 1761-1765.

[34] Ammar, A., Koubaa, A., Ahmed, M., & Saad, A. (2019). Aerial images processing for car detection using convolutional neural networks: Comparison between faster r-cnn and yolov3. arXiv preprint arXiv:1910.07234.

[35] Bai, T., Pang, Y., Wang, J., Han, K., Luo, J., Wang, H., ... & Zhang, H. (2020). An Optimized Faster R-CNN Method Based on DRNet and RoI Align for Building Detection in Remote Sensing Images. Remote Sensing, 12(5), 762.

[36] Shair, E. F., Jamaluddin, N. A., & Abdullah, A. R. (2020). Finger Movement Discrimination of EMG Signals Towards Improved Prosthetic Control using TFD. International Journal of Advanced Computer Sciences and Applications, 11(9), 244-251.

[37] Bozcan, I., & Kayacan, E. (2020). AU-AIR: A Multi-modal Unmanned Aerial Vehicle Dataset for Low Altitude Traffic Surveillance. arXiv preprint arXiv:2001.11737.