IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Novel Study for Identification of Vehicles Using Deep Learning Techniques

[1]K. Vishal
[1]Student, Department of Information Technology
G M R Institute of Technology, Rajam, AP

Abstract: The paper addresses the problem of detection of vehicles from aerial images using various deep learning techniques. Unmanned Aerial Vehicles (UAV's) are used to capture the images which gives us a clear view of all the vehicles. The main aim of this project is to detect all types of vehicles present in the image, so that it is useful for traffic analysis, free slots in parking system, counting vehicles, and many more. There are many deep learning algorithms implemented for the detection of vehicles. It includes the deep learning algorithms such as Convolutional Neural Network -Support Vector Machine (CNN-SVM), Faster Region – based Convolutional Neural Network (Faster R-CNN), You Look Only Once (YOLO) and Single Shot Detector (SSD). The objective of the study is to conduct a comparison between the approaches used by Faster R-CNN, YOLO, SSD and CNN-SVM algorithms. The differences and similarities of different techniques are highlighted in this paper and future challenges are also discussed.

Keywords: - **Vehicle Detection, Deep Learning, Unmanned Aerial Vehicles (UAV's), Convolutional Neural Network - Support Vector Machine (CNN-SVM), Faster Region – based Convolutional Neural Network (Faster R-CNN), You Look Only Once (YOLO) and Single Shot Detector (SSD).**

I. INTRODUCTION

Drones, also known as unmanned aerial vehicles, are quickly becoming a widely used and very effective technology for monitoring infrastructure and the environment. Particularly, there is a lot of interest in the employment of UAVs in the field of road traffic monitoring (RTM). In the above mentioned deployments, UAVs are in charge of looking for, gathering, and sending vehicle information from on-board video sensors in real time for traffic regulation purposes.. Since the introduction of convolution neural networks and other deep learning techniques, object identification and recognition have displayed a noticeable improvement in accuracy. This opens the door for the broad use of UAVs for data collection and analysis across a variety of technical domains. The accuracy of categorization and object identification has significantly increased as a result of developments in deep learning, particularly convolution neural network (CNN) applications. There were several suggested CNN architectures and algorithms, including YOLO and its variants, R-CNN and its variants, and R-CNN, which is a region-based CNN. To overcome the limitation of R-CNN, selective search of region, Faster R-CNN was introduced.

II. LITERATURE SURVEY

In paper [1], Dikbayir et.al discussed about two most popular algorithms Faster R-CNN and YOLO. The aim of Faster R-CNN (region based convolutional neural network) is to create a certain number of regions with a selective search method and search through the regions instead of searching through the whole picture and find the right object. The YOLO algorithm aims to offer a structure suitable for real-time processing by taking the picture completely convolutional rather than a regional-based approach. In this paper, firstly; Munich Vehicle Data Set was used, because it includes images of different vehicle types over 100m, and high resolution images obtained from the "Google Earth" application and DJI drone images were also included in the set to expand the data set.

In paper [2], Ammar, A. et.al assessed the performance of three cutting-edge CNN algorithms, including the most well-known region-based technique, Faster R-CNN, and the fastest detection algorithms, YOLOv3 and YOLOv4. Faster R-CNN is a deep convolutional network that simulates an end-to-end network to the user and is used for item detection. The network can accurately and rapidly predict the locations of different objects. YOLO, an acronym for You Only Look Once, processes the image using a Fully Convolutional Neural Network that predicts the bounding boxes and their corresponding class probabilities, based on the global context of the image. The model was successfully validated on two datasets (Stanford & PSU). Future work includes Extending our results to the newly released EfficienDet detector and to much larger datasets of aerial images.

In paper [3], Lu, J. et.al performed research and compared all the versions of YOLO. By combining three open-source aerial picture datasets, the method creates an aerial image dataset appropriate for YOLO training. The VEDAI (Vehicle Detection in Aerial Imagery) dataset is made by Sebastien Razakarivony and Frederic Jurie of University of Caen, whose original material is from public Utah AGRC database. COWC (Cars Overhead with Context) dataset was designed by T. Nathan Mundhenk and others of Lawrence Livermore National Laboratory, whose original materials are from six opensource websites. DOTA (Dataset for Object detection in Aerial images) is an aerial image dataset made by researchers from Wuhan university. Future work includes integration of more public aerial image datasets to increase the number and diversity of training samples, at the same time, optimize the YOLO algorithm to further improve the detection accuracy.

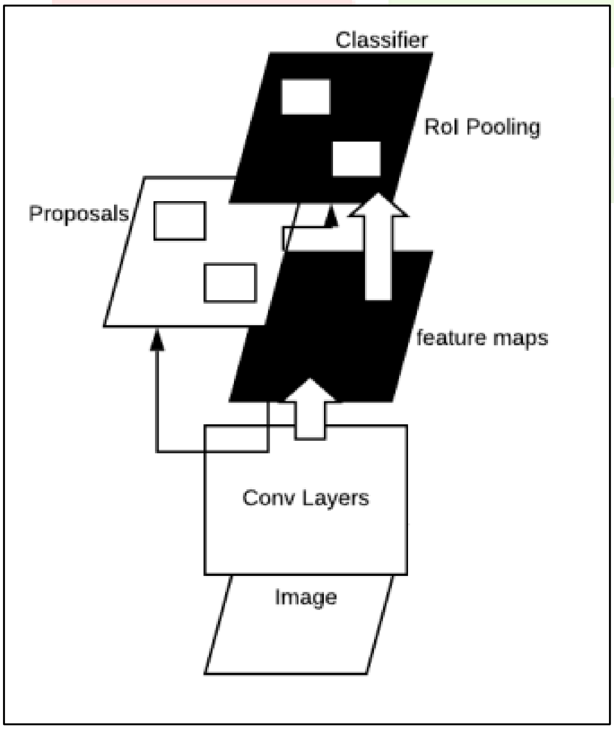
In paper [4], Liao et.al presented an experimental study to evaluate the performances of several state-of-the art deep learning-based detection approaches on vehicle detection from aerial imagery using deep learning techniques such Faster R-CNN, R-FCN, and SSD. As a benchmark, the VEDAI dataset is utilised. Two different aerial picture sizes are included in the VEDAI dataset: VEDAI 512 and VEDAI 1024 (1024 × 1024 and 512x 512 pixels, respectively). VEDAI dataset 1024 consists of 1250 instances in total, and 1164 images are remaining after we removed the scarce categories mentioned above.

In paper [5], Valappil et.al proposed a procedure which includes the Kanade-Lucas optical flow method for detecting moving objects, the construction of connected graphs to separate items, convolutional neural networks (CNN), and support vector machines (SVM) to determine the outcome of the classification. The classifier eliminates the possibility of any additional (moving) items being present and being identified as cars. On both fixed and moving aerial films, the approach is evaluated. The capacity of CNN to extract features and subsequently apply the effectiveness of SVM for binary classification. The future scope is Development of multi-class classification applied to deep learning approaches in situation where various categories of vehicles are being detected.

III. METHODOLOGY OF EXISTING TECHNOLOGY

Faster R-CNN:

Faster R-CNN is the most accurate version of region-based convolutional neural networks. Faster R-CNN is a two-stage network, region proposal network is used for generating the region proposals and these proposals are used for detection of objects. In Faster R-CNN the use of the selective search algorithm is discarded and is replaced with Region Proposal Network (RPN). RPN provides the probable regions in the image and proposes the ones most likely to contain objects. RPN has learnable parameters which make it more efficient than the previous versions of region-based convolutional neural networks. Faster R-CNN is very accurate and also fast in processing therefore can be extended to soft real-time applications.



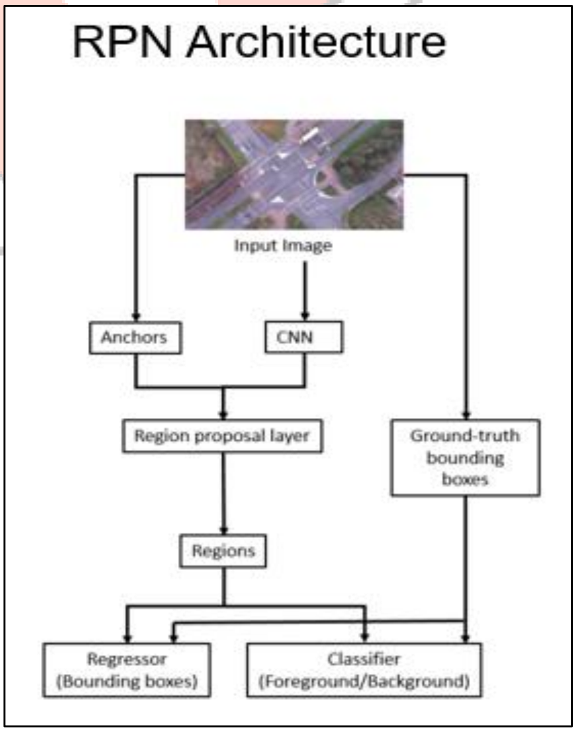


Fig: RPN Architectures

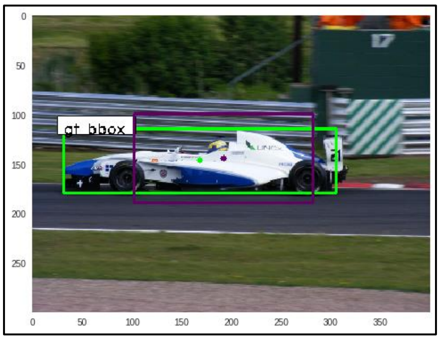


Fig: Bounding Boxes

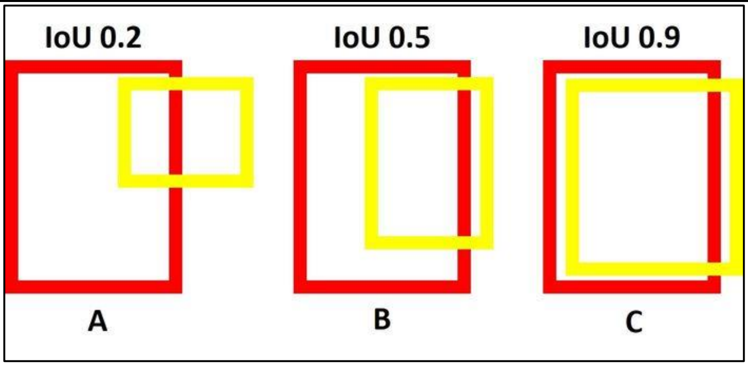


Fig: Intersection of Unions(IoU)

YOLO v3:

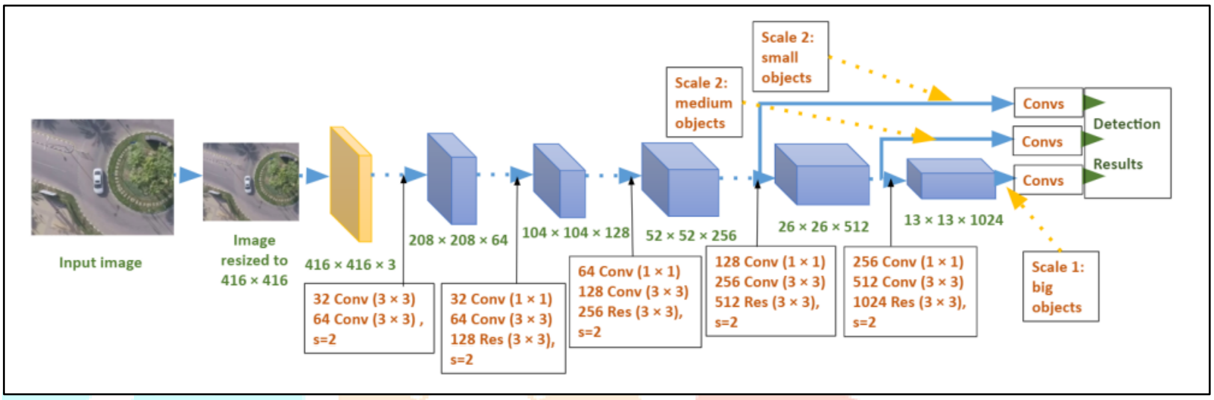


Fig: YOLOv3 Architecture

YOLO stands for You Only Look Once. YOLOv3 is a single stage detector that utilizes convolutional layers and anchors to perform the task of detection and localization. The algorithm divides the images into regions and predicts the bounding box probabilities for each region. The anchor boxes are used for localization of the object in the image. These bounding boxes are weighted by the prediction probabilities. A single CNN is used to simultaneously predict multiple bounding boxes and class predictions for the boxes, thereby reduces the number of parameters and reduces the process overhead. YOLOv3 is very fast at inference and can be used for real – time object detection.

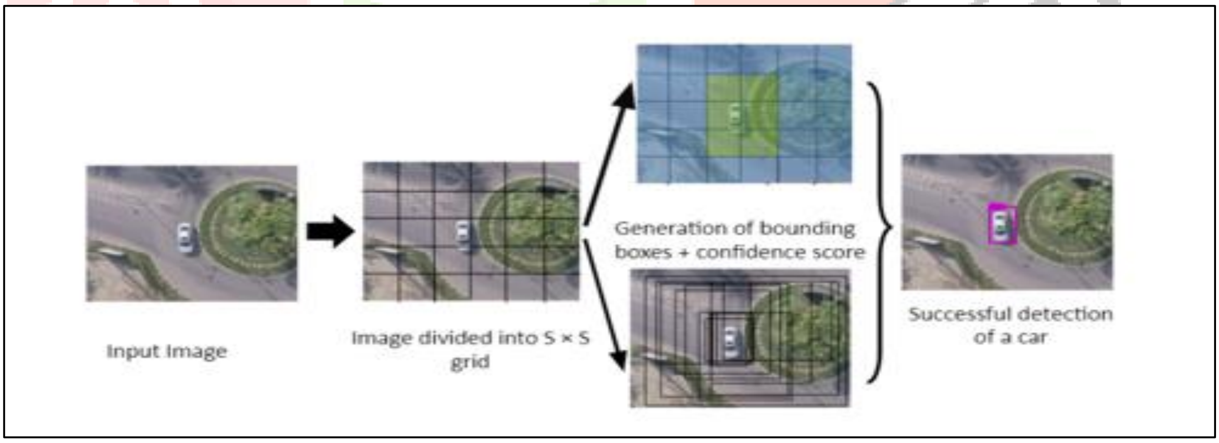


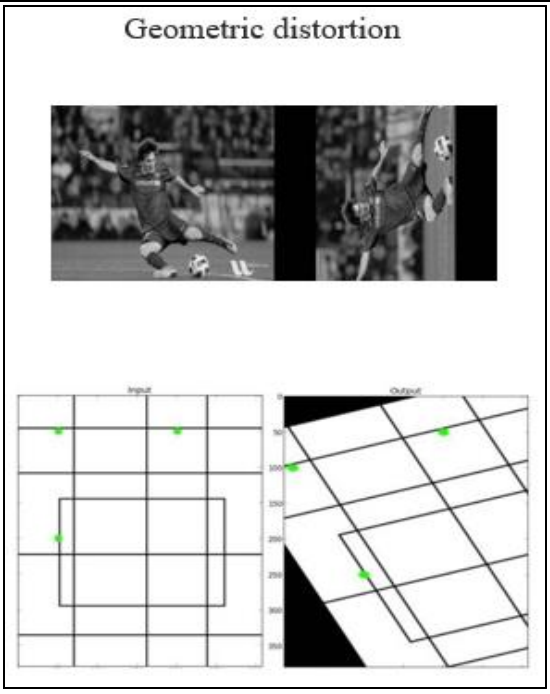
Fig: Successive stages of YOLOv3 applied on Car Detection

YOLO v4:

The goal of YOLOv4 is to make object detection operate effectively and without bugs on the hardware found on edge devices. They fall into two groups in terms of YOLOv4's technical advancements.

The Bag of Freebies (BoF), the first category, refers to enhancements that can be applied during training without influencing the inference time. This includes the CutMix and Mosaic data augmentation methods, the Geometric distortion, class label smoothing, the Complete IoU (CIoU) loss, the Photometric distortion, the Self Adversarial Training (SAT), the use of multiple anchors for a single ground truth, the cosine annealing scheduler, and the genetic algorithm-derived optimal hyper-parameters.

The second category, known as Bag of Specials (BoS), includes advancements that just marginally affect inference time while significantly increasing accuracy. It also comprises the Distance IoU Loss (DIoU) utilised as a factor in the Non-Maximum-Suppression (NMS) step, the Spatial Pyramid Pooling (SPP) block, the Spatial Attention Module (SAM) block, the Multi-input Weighted Residual Connection (MiWRC), and the Mish Activation Function.



SSD:

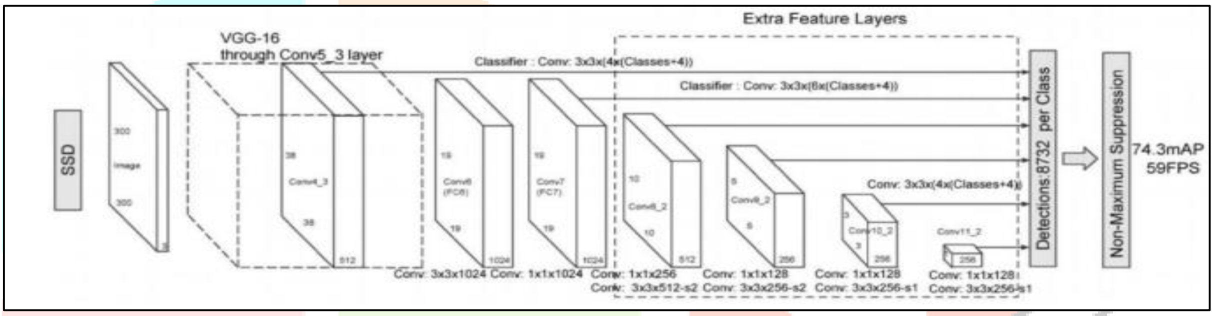


Fig: Architecture of SSD

SSD is a one-stage approach for object detection called Single Shot Multi-Box Detector (SSD). It was inspired by the anchors adopted in Multi-Box, RPN, and multi-scale representation that aims to deal with strong spatial constraints imposed on bounding box predictions to overcome the shortcoming of detecting small objects in groups. The algorithm extracts feature maps from different scales where they could use to recognize small and large objects well. The SSD method applies anchor boxes with varying ratios of aspects to discretize the space of bounding boxes. It uses VGG16 net as the backbone, and extra feature layers are added to the end of networks that are responsibly predicting boxes with different scales, aspect ratios, and confidences.

CNN-SVM:

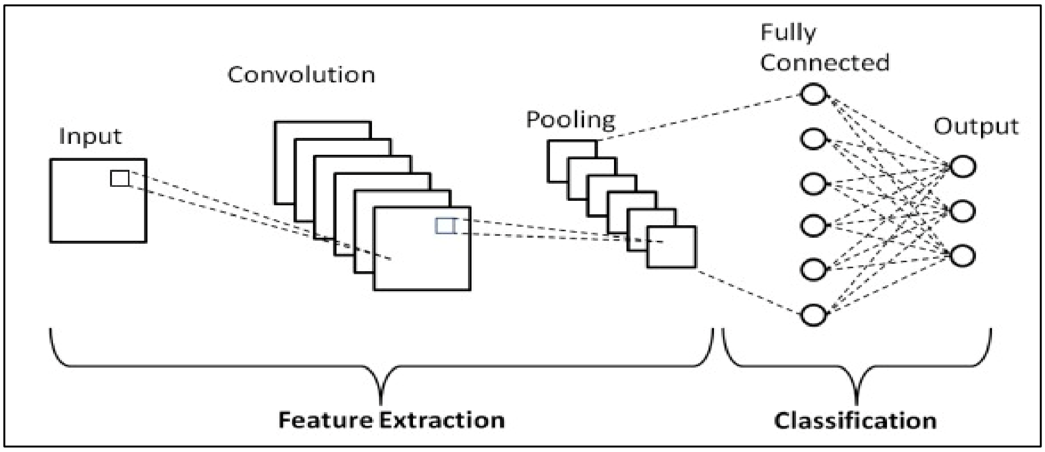


Fig: CNN Architecture

The methodology involves multiple steps, which are building motion vectors, developing connectivity graph, detection followed by classification. As a first step, optical flow techniques are used for motion estimation. The binary classification stage is completed with support vector machine, where extracted features from CNN layers form the input. The interest regions built during connectivity graph undergo final step in CNN-SVM for classification. This step does the feature engineering on its own, and fully exploits the capability of CNN to extract features and later use the efficiency of SVM for binary classification. During feature extraction, the data is searched for feature correlation, and later combination enables faster learning. It consists of various layers in the form of convolution and pooling stacks with linear and nonlinear operators joined end-to-end for solving the problem. The convolutional layer is significant as it contains small tunable filter applied across the whole image for extracting image features.

IV. RESULTS

The data in the table below shows the final outcome, which is utilised to train and test the predictive models. Apply deep learning techniques to the final dataset after it has been prepared. Finally, apply deep learning algorithms after creating the final dataset to obtain precision which is being displayed in Table1. Fig1 shows the graphical representations of the precisions.

S.No	Algorithm	Precision
1	Faster R-CNN	97.5%
2	YOLO v3	94.7%
3	YOLO v4	91.6%
4	SSD	96.49%
5	CNN-SVM	98%
6	CNN	97.8%

Table-1: - Different DL Algorithms with their Precision

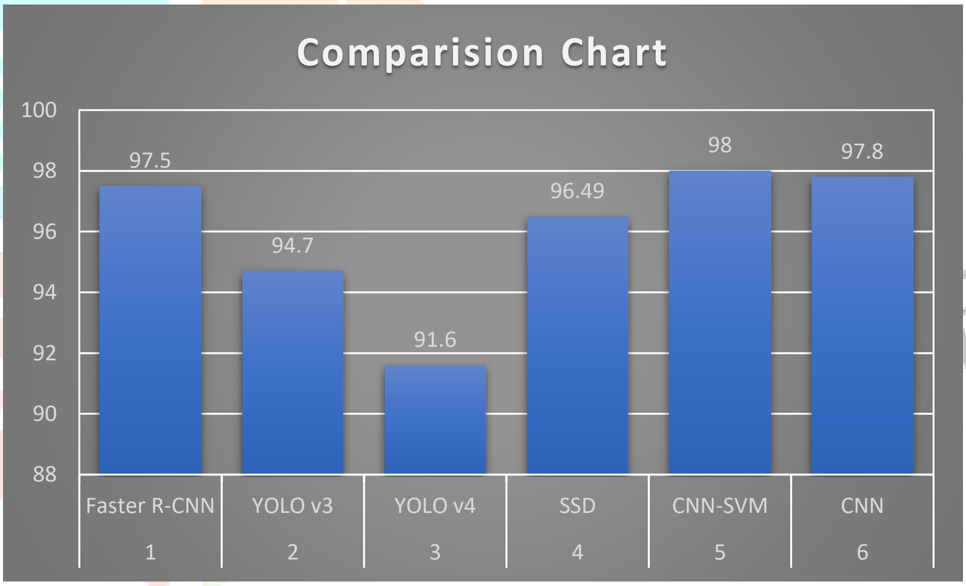


Fig-1: - Analysis of various DL algorithms

V. CONCLUSION

Based on the analysis performed in this paper, various Deep Learning algorithms are applied to the data to detect the vehicles from UAV's using the datasets previously available. Comparative analysis of the various algorithms used to include Faster R-CNN, YOLO v3, YOLO v4, SSD and CNN-SVM was carried out. After examining the methods for detection of vehicles, CNN-SVM outperforms the other algorithms in terms of performance and accuracy. Therefore, the Deep Learning model that was created using the CNN-SVM algorithm may be used efficiently to detect the vehicles from the UAV's accurately.

VI. REFERENCES

[1] H. S. DIKBAYIR and H. İbrahim BÜLBÜL, "Deep Learning Based Vehicle Detection From Aerial Images," 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA),2020, pp. 956-960, doi: 10.1109/ICMLA51294.2020.00155.

[2] Ammar, A.; Koubaa, A.; Ahmed, M.; Saad, A.; Benjdira, B. Vehicle Detection from Aerial Images Using Deep Learning: A Comparative Study. Electronics 2021, 10, 820. https://doi.org/ 10.3390/electronics10070820

[3] Lu, J. , Ma, C. , Li, L. , Xing, X. , Zhang, Y. , Wang, Z. and Xu, J. (2018) A Vehicle Detection Method for Aerial Image Based on YOLO. Journal of Computer and Communications, 6, 98-107. doi: 10.4236/jcc.2018.611009.

[4] Liao, Xuanzhi & Sahran, Shahnorbanun & as, Syaimak. (2020). An experimental study of vehicle detection on aerial imagery using deep learning-based detection approaches. Journal of Physics: Conference Series. 1550. 032005. 10.1088/1742-6596/1550/3/032005.

[5] Valappil, Najiya K. and Memon, Qurban A. 'CNN-SVM Based Vehicle Detection for UAV Platform'. 1 Jan. 2021 : 59 – 70.

[6] Radovic, M.; Adarkwa, O.; Wang, Q. Object Recognition in Aerial Images Using Convolutional Neural Networks. *J. Imaging* 2017, *3*, 21. https://doi.org/10.3390/jimaging3020021

[7] X. Li, X. Li and H. Pan, "Multi-Scale Vehicle Detection in High-Resolution Aerial Images With Context Information," in IEEE Access, vol. 8, pp. 208643-208657, 2020, doi: 10.1109/ACCESS.2020.3036075.

[8] H. -Y. Lin, K. -C. Tu and C. -Y. Li, "VAID: An Aerial Image Dataset for Vehicle Detection and Classification," in IEEE Access, vol. 8, pp. 212209-212219, 2020, doi: 10.1109/ACCESS.2020.3040290.

[9] Yuni Zeng, Qianwen Duan, Xiangru Chen, Dezhong Peng, Yao Mao, and Ke Yang. 2021. UAVData: A dataset for unmanned aerial vehicle detection. Soft Comput. 25, 7 (Apr 2021), 5385–5393. https://doi.org/10.1007/s00500-020-05537-9

[10] A. Bouguettaya, H. Zarzour, A. Kechida and A. M. Taberkit, "Vehicle Detection From UAV Imagery With Deep Learning: A Review," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2021.3080276.