sustainability

Article

Vehicle Detection and Classification via YOLOv8 and Deep Belief Network over Aerial Image Sequences

Naif Al Mudawi [1], Asifa Mehmood Qureshi [2], Maha Abdelhaq [3,*], Abdullah Alshahrani [4], Abdulwahab Alazeb [1], Mohammed Alonazi [5,*] and Asaad Algarni [6]

1 Department of Computer Science, College of Computer Science and Information System, Najran University, Najran 55461, Saudi Arabia; naalmudawi@nu.edu.sa (N.A.M.); afalazeb@nu.edu.sa (A.A.)
2 Department of Creative Technologies, Air University, E-9, Islamabad 44000, Pakistan; asifa.mahmood19@gmail.com
3 Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
4 Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah 23218, Saudi Arabia; asalshahrani2@uj.edu.sa
5 Department of Information Systems, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 16273, Saudi Arabia
6 Department of Computer Sciences, Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia; asaad.algarni@nbu.edu.sa
* Correspondence: msabdelhaq@pnu.edu.sa (M.A.); mn.alonazi@psau.edu.sa (M.A.)

Abstract: Vehicle detection and classification are the most significant and challenging activities of an intelligent traffic monitoring system. Traditional methods are highly computationally expensive and also impose restrictions when the mode of data collection changes. This research proposes a new approach for vehicle detection and classification over aerial image sequences. The proposed model consists of five stages. All of the images are preprocessed in the first stage to reduce noise and raise the brightness level. The foreground items are then extracted from these images using segmentation. The segmented images are then passed onto the YOLOv8 algorithm to detect and locate vehicles in each image. The feature extraction phase is then applied to the detected vehicles. The extracted feature involves Scale Invariant Feature Transform (SIFT), Oriented FAST and Rotated BRIEF (ORB), and KAZE features. For classification, we used the Deep Belief Network (DBN) classifier. Based on classification, the experimental results across the three datasets produced better outcomes; the proposed model attained an accuracy of 95.6% over Vehicle Detection in Aerial Imagery (VEDAI) and 94.6% over Vehicle Aerial Imagery from a Drone (VAID) dataset, respectively. To compare our model with the other standard techniques, we have also drawn a comparative analysis with the latest techniques in the research.

Keywords: YOLOv5; vehicle detection; classification; segmentation; DBN

Citation: Al Mudawi, N.; Qureshi, A.M.; Abdelhaq, M.; Alshahrani, A.; Alazeb, A.; Alonazi, M.; Algarni, A. Vehicle Detection and Classification via YOLOv8 and Deep Belief Network over Aerial Image Sequences. *Sustainability* **2023**, *15*, 14597. https://doi.org/10.3390/su151914597

Academic Editors: Ahmad Hassanat, Sami Mnasri and Ahmad S. Tarawneh

Received: 13 September 2023
Revised: 29 September 2023
Accepted: 5 October 2023
Published: 8 October 2023

1. Introduction

In recent years, vehicle detection and classification has been an emerging research area due to its various applications in intelligent traffic management systems. Road Traffic management applications include congestion detection, categorizing the various vehicle types, recognizing doubtful vehicles on the road, and parking management system [1]. All these systems mainly depend on vehicle identification, which has become a significant and crucial issue in aerial imagery [2]. In conventional systems, vehicle detection was primarily conducted by estimating motion in the image pixels [3–6]. However, the methods are not efficient enough in remote sensing data because motion is also detected in pixels other than the targeted objects [7]. Recently, researchers have proposed many improved techniques, which include object segmentation [8], silhouette extraction [9], feature extraction, and classification [10], to enhance the object detection capabilities of a system [11–16].

Aerial images provide a better and broader view, thus providing significant information about the sensed environment [17]. These images are used in numerous applications, such as deforestation detection [18], agriculture field monitoring [19], and disaster management systems [20]. The aerial traffic data is also collected to do traffic analysis to efficiently use the road network, forecast forthcoming transportation requirements, and improve traveler protection [21].

In our proposed model, we have used aerial images to recognize and classify vehicles. In our model, the aerial videos are first converted into image frames. These frames are pre-processed for noise removal and brightness enhancement using defogging and gamma correction techniques, respectively [22–25]. Then, the images are segmented to reduce the background complexity using Fuzzy C Mean segmentation. To detect vehicles in each extracted frame, YOLOv8 is employed, which can detect small objects effectively. In the end, all the detected vehicles are subjected to SIFT, ORB, and KAZE feature extraction to classify them into multiple vehicle classes. For classification, we used the Deep Belief Network, which is a simple classifier that uses neural networks, thus providing better classification accuracy. Our accuracy has proven to be a result of an efficient model design. The following is our system's primary contribution:

- Our model combines the pre-processing methodologies with the segmentation technique to prepare images before passing them to the detection phase to reduce model complexity.
- We used the newest YOLOv8, which has improved architecture to enhance vehicle detection in segmented images as it can effectively detect objects of varying sizes.
- To classify vehicles, multiple features, including SIFT, ORB, and KAZE features, are extracted. Combining scale and rotation invariant, 2D and fast and robust local feature vectors are effective in classifying vehicles in aerial images.
- The proposed system uses a deep learning-based DBN classifier to achieve higher classification accuracy.

The following is a list of the remaining sections of this article. Related work analysis of the current approaches is included in Section 2. The suggested system's architecture is presented in Section 3. The experimental portion with a system performance evaluation is shown in Section 4. Section 5 presents the system's conclusion and the direction of future efforts.

2. Related Work

In this section, we presented the most relevant and popular systems designed for vehicle detection and classifications. Table 1 presents the details of the different models proposed by the researchers in the literature.

Table 1. Related Work for Vehicle Detection and Classification.

Authors	Methodology
Arinaldi et al. [26]	The paper implements two different methodologies for vehicle detection and classification. The first method uses a Mixture of Gaussian (MoG), combined with a Support Vector Machine Classifier (SVM) classifier. The other method only uses faster Recurrent Convolutional Neural Network (RCNN). However, there was still a large number of vehicles that were left undetected.

Table 1. *Cont.*

Authors	Methodology
Aqel et al. [27]	This study uses the background subtraction method to detect moving autos. To lower the occurrences of false positives, morphological corrections are performed. In the end, the classification is accomplished using the invariant Charlier moments. The method uses conventional image processing techniques, that limits its applicability to diverse traffic scenarios. Also, the background subtraction method will eliminate the cars which are not in motion, thus reducing the true positives.
Sarikan et al. [28]	The model uses a K-nearest neighbor classifier to automatically detect and classify vehicles. For feature extraction, windows and hollow areas of the vehicles are constructed to classify it as a motorcycle or car. The model is not applicable for broader views and dense traffic conditions.
Tan et al. [29]	The authors presented a method to classify vehicles using a Convolutional Neural Network (CNN). It uses an aerial image dataset. The proposed model firstly determines whether the area contains any vehicle or not by evaluating motion changes, feature matching and heat maps. Then, the classification is conducted using the classification layer of inception-v3 and AlexNet.
Hamzenejadi et al. [30]	This paper presents real-time vehicle detection solution based on Yolov5. The existing model is improved by adding attention mechanism and a new concept of ghost convolution. The experimental results prove the efficiency of the YOLO model in object detection models.
Ozturk et al. [31]	In this paper, a vehicle detection method has been presented. The vehicles are detected via miniature CNN architecture combined with morphological corrections. The model requires intensive post-processing to achieve good results. Also, the accuracy is not consistent on other datasets.
Roopa Chandrika et al. [32]	A model for vehicle recognition and classification has been presented. The model incorporates adaptive background subtraction along with binary label segmentation to locate vehicles. The approach is not suitable for stationary car detection or during traffic jam conditions.
Kumar et al. [33]	A new approach that uses You Only Look Once (YOLO) with Long Short-Term Memory (LSTM) to detect and classify vehicles. To reduce the model complexity, the images are segmented into binary labels in the pre-processing stage. The detected vehicles are also counted by counting the bounding boxes and classified into lightweight and heavy-weight vehicles.
Zhang et al. [34]	The paper proposes a method that uses an improved YOLOv3 algorithm to detect vehicles. The pre-trained YOLO network is trained with a new structure to improve the accuracy of the detection method. However, YOLOv3 is one of the oldest versions. The detection results can be improved by using the newest architectures.

Even though extensive research has been completed in the field of automated traffic monitoring systems, there is still room for improvement. The detection of vehicles in aerial images specifically in intensive traffic conditions requires efficient and specialized architectures to obtain good results. Machine learning methods are not good enough to differentiate between objects that have motion in their pixels [35,36]. Therefore, YOLOv8 is the newest and most effective object detector based on convolution layers [37,38]. Moreover, combining different feature sets to classify vehicles can contribute to reducing classification errors.

3. Proposed System Methodology

The proposed architecture identifies vehicles in the images and classifies them into multiple vehicle classes. Primarily, the videos are first converted into frames. Pre-processing procedures are applied to the images, i.e., defogging for noise reduction, and then Gamma correction is used to modify the intensity of the images for improved detection. On the filtered images, FCM segmentation is applied to separate the foreground and background objects [39–41]. The detection is performed using the YOLOv8 algorithm. After vehicle detection, SIFT, ORB, and KAZE features were extracted [42–44]. On this feature vector, the DBN classifier was trained to classify each detected vehicle into its corresponding class. The proposed system design is shown in Figure 1.

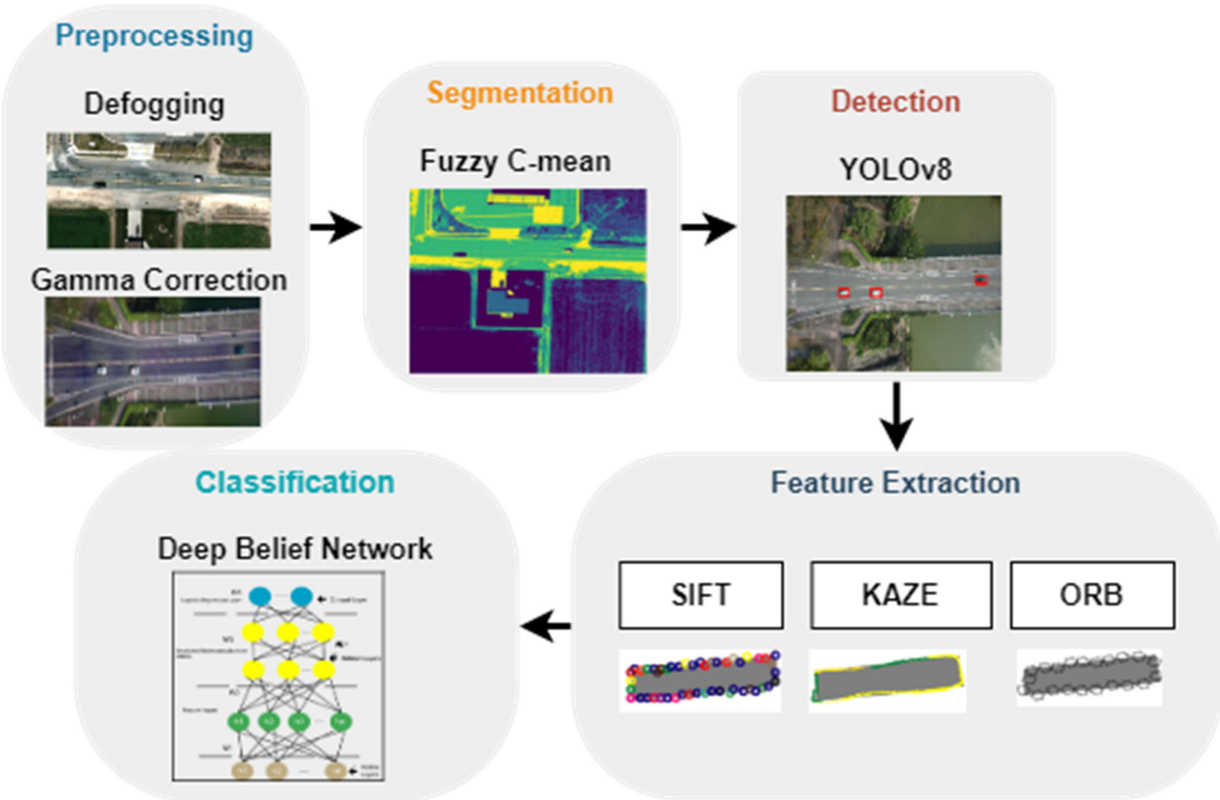


Figure 1. Proposed architecture for vehicle detection and classification.

3.1. Images Pre-Processing

Noise reduction is required in the obtained image to remove additional pixel information, since the extra pixels make detection more difficult [45–47]. Only the most appropriate filter that incorporates defogging techniques is applied to the specific noise for good results [48]. The defogging method determines the amount of noise present in each pixel of the image, then eliminates it as follows:

$$G(x) = X(x)Y(x) + Z(1 - p(x)) \tag{1}$$

where x specifies the location of the pixel, Z is the fog density, and $Y(x)$ is the transmission map. Figure 2 shows the defogged images.

(**a**) (**b**)

Figure 2. Defogging results over the VEDAI and VAID datasets (**a**) original Images (**b**) defogged images.

In the next step, Gamma correction [49,50] is used to alter the denoised image's intensity since the region of interest can be detected most effectively when the brightness is high [51]. The power-law for gamma correction is given as:

$$V_o = TV_I^{\gamma} \tag{2}$$

where T is a constant that is typically equal to 1, V_I is the input's non-negative values with power γ whose range can be between 0 and 1. V_o represents the resultant image [52–55]. Figure 3 displays the denoised, intensity-adjusted image with the plot. The gamma-corrected images are given in Figure 3.

Figure 3. Pre-processed image using gamma correction over the VEDAI and VAID datasets.

3.2. Fuzzy C-Mean Segmentation

In this section, the foreground objects are separated from the background to reduce the complexity of the images. For this purpose, we used the FCM segmentation technique, that groups the image pixels into one or more clusters [56]. In FCM segmentation, the pixels which belong to more than one cluster are known as fuzzy logic [57,58]. While grouping the pixels, the objective function is optimized during numerous iterations of the process [59,60]. The clustering centers and membership degrees have been regularly changed during the iterations [61]. A finite collection of N elements Q = q_1, q_2, . . ., q is divided into a set of M clusters via the FCM method. Each element of the vector w_j, where j = 1, 2, . . ., N, has n dimensions [62,63]. We define a technique to divide Q into M clusters using the cluster centers c_1, $c_2 \ldots, c_m$ in the centroid set c [64]. In the FCM technique, h is a representative matrix that shows each element's participation in each cluster [65,66]. It can be well-defined as:

$$h(j,y),\ 1 \leq j \leq N; 1 \leq y \leq M \tag{3}$$

where the membership value of the element q_j with cluster center c_y is represented by $h(j,y)$. We are more certain that the element q_j belongs to the y cluster if the value of $h(j,y)$ is higher [67,68]. Moreover, when calculating the performance index L_f, the weighted sum of the distance between the components of the relevant fuzzy cluster and the cluster center is calculated [69,70].

$$L_f = (h,c) = \sum_{i=1}^{v} \sum_{a=1}^{y} h_{ia}^{t} \| \ q_i - c_a \ \|^2,\ 1 \ < \ t \ < \infty \tag{4}$$

where c_a is the ath cluster center, q_i is the ith pixel, v is the cluster number, y is the number of pixels, and t is the blur exponent [71–75]. The following formula is used to update the membership function:

$$h_{ia}^{t} = \frac{1}{\sum_{h=1}^{m} \left(\frac{dis_{ia}^{2}}{dis_{ha}^{2}} \right)^{\frac{2}{t-1}}} \tag{5}$$

where the distance between the cluster centroid c_a and the pixel q_i is supplied by dis_{ia}^{2}, and the membership matrix is represented by h_{ia}^{t}, which ranges (0, 1). The point of cluster centroid is calculated as follows:

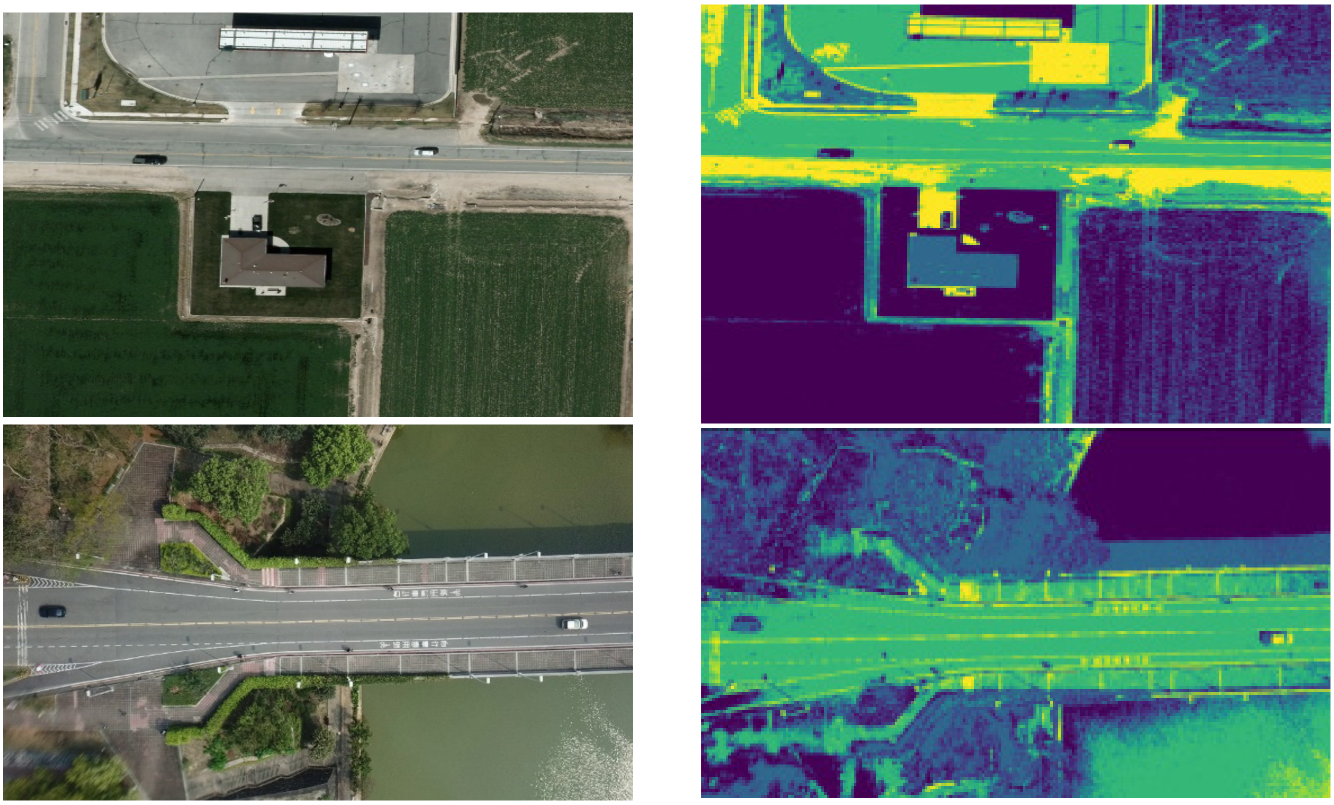
$$c_a = \frac{\sum_{j=1}^{N} h_{ij}^{t} q_j}{\sum_{j=1}^{N} h_{ij}^{t}} \tag{6}$$

When a pixel gets close to the cluster center to which it belongs, it receives a high membership value, and vice versa. The result of FCM segmentation is seen in Figure 4.

3.3. Vehicle Detection via YOLOv8

For vehicle detection, we used the YOLOv8 algorithm. YOLOv8 is an efficient single-shot detector that can be used for detection, segmentation, and classification tasks [76]. Furthermore, it requires fewer parameters for training [77–80]. Based on the CSP concept, the C2f module replaces the C3 module, whereas the YOLOv8 backbone is mostly the same as the YOLOv5 backbone [81,82]. The C2f combines C3 and ELAN to create the C2f module, building on the ELAN concept from YOLOv7, so that YOLOv8 might continue to be portable while obtaining more comprehensive gradient flow information [83]. The SPPF module was still utilized at the end of the backbone, and three Maxpools of size 5 $\times$ 5 were sequentially applied before each layer was concatenated to ensure the precision of objects of varying scales while also maintaining a low weight [84].

The feature fusion approach still employed by YOLOv8 in the neck section is PAN-FPN, which improves the fusion and usage of feature layer data at numerous scales. The neck module is made up by combining the final decoupled head structure, numerous C2f modules, and two upsamplings [85–88]. The final component of the neck in YOLOv8 was constructed using the same concept as the head in YOLOx. It increased accuracy by combining confidence and regression boxes. Moreover, it is an anchor-free model which can directly detect the object's center. In order to expedite Non-Maximum Suppression (NMS), a challenging post-processing step that sorts through potential detections following inference, anchor-free detection lowers the number of box predictions. The detected vehicles using the YOLOv8 are given in Figure 5.



(**a**) (**b**)

Figure 4. Semantic Segmentation using FCM over VEDAI and VAID dataset (**a**) original dataset image (**b**) segmented image.

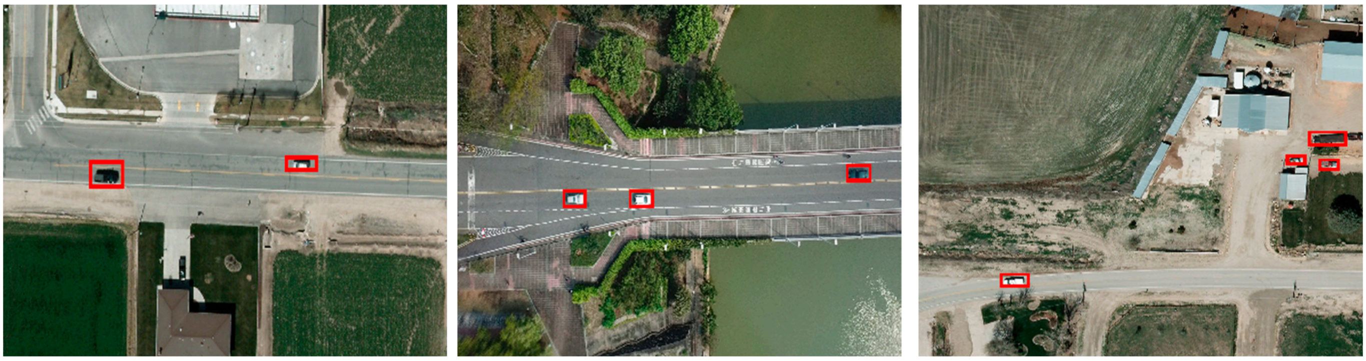


Figure 5. Vehicle Detection marked with red boxes via the YOLOv8 algorithm.

3.4. Feature Extraction

This section describes a method for extracting various features. The feature set comprises three different features: SIFT, KAZE, and ORB.

3.4.1. SIFT Features

We used the Scale Invariant Feature Transform (SIFT) technique to obtain important features [89–91]. SIFT reduces an image's information to a set of points that can be used to identify recurrent patterns in other pictures [92]. Scale and rotation invariant features are retrieved using SIFT [93–95]. Figure 6 shows the steps of SIFT feature extraction.

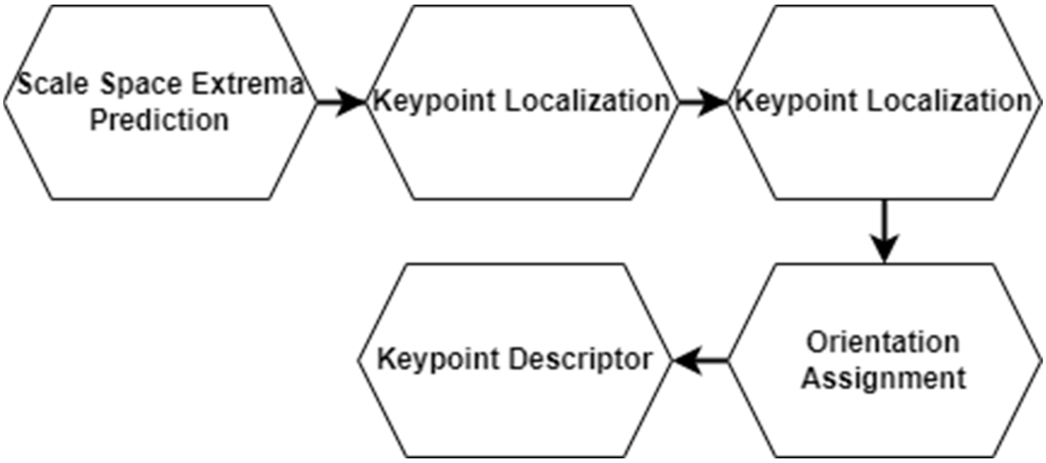


Figure 6. Steps for SIFT feature extraction.

The extracted SIFT features are given in Figure 7.

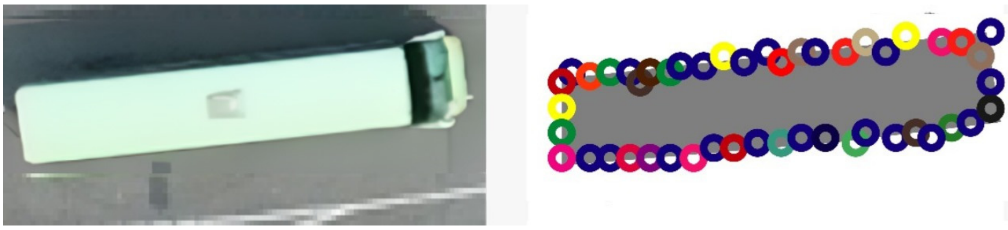


Figure 7. SIFT feature extraction.

3.4.2. KAZE Features

In order to extract KAZE features, a Gaussian kernel is convolved with an input image [96]. The convolved image is used to construct an image gradient histogram, and computer code is used to calculate the contrast parameters [97]. Values for the contrast parameter and evolution time are used to calculate the nonlinear scale space as follows:

$$t^{j+1} = \left(I - f_j + 1 - f_j \sum\nolimits_{i=1}^{m} B_i\left(t^j\right)\right)^{-1} t^j \tag{7}$$

to determine the response of the scale normalized determinant of the Hessian at various levels to identify interesting locations, we use the formula:

$$F_{Hess} = \sigma^2\left(t_{xx}t_{yy} - t_{xy}^2\right) \tag{8}$$

The second-order cross-derivative is presented as t_{xy}, the second-order horizontal derivative as t_{xx}, and the vertical derivative is given as t_{yy}. The extracted KAZE features are shown in Figure 8.

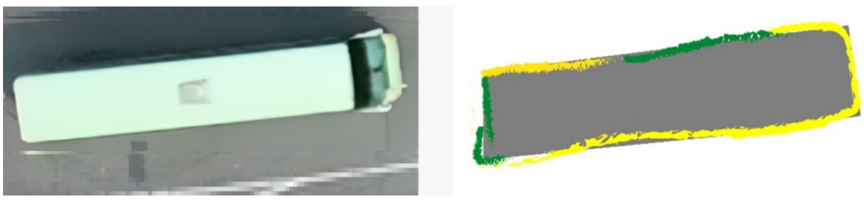


Figure 8. KAZE feature extraction.

3.4.3. ORB Features

The Oriented FAST and Rotated BRIEF (ORB) is an efficient feature extractor. To identify key points, it uses the FAST (Features from Accelerated Segment Test) keypoint detector [98–100]. It makes more complex use of the BRIEF (Binary Robust Independent Elementary Features) description. Additionally, it is dimensionally and rotationally invariant [101]. The patch moment is obtained as follows:

$$m_{uv} = \sum x^u y^v l(j,k) \tag{9}$$

where *u* and *v* represent the intensities of the picture pixels at the *j* and *k* locations. Moreover, the mass center is calculated by using the following formula.

$$W = \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \tag{10}$$

The patch orientation is obtained by:

$$\theta = atan(m_{01}, m_{10}) \tag{11}$$

The final extracted feature is seen in Figure 9.

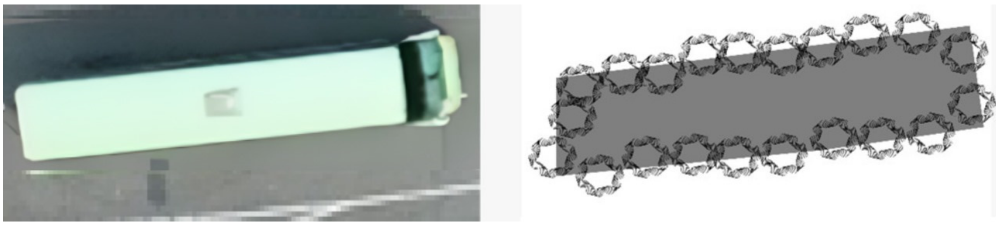


Figure 9. ORB feature extraction.

3.5. Classification via DBN

A Deep Belief Network (DBN) classifier is being used to classify vehicles. A deep neural network serves as a DBN's building block, which is composed of layers of latent variables connected only between the layers as a whole and not between the units within each layer [102]. For the creation of DBN, Restricted Boltzmann Machines (RBN) act as the fundamental building blocks [103]. A layer of RBN's visible and hidden units combine to form a two-layer structure [104]. The collective energy arrangement of the two units is calculated as:

$$\begin{aligned} Enr(MN, WM, \theta) &= -\textstyle\sum_{i=1}^{D} rM_i v_i - \sum_{j=1}^{F} a_{Hj} h_j - \sum_{i=1}^{D} \sum_{j=1}^{F} s_j MN_i WM_j \\ &=> \ -r^T MN - a^T WM - MN^T MWM \end{aligned} \tag{12}$$

where $\theta = \{rM_i,\ aW_j,\ se_{ij}\}$, a_{Hj} and rM_i stand for the bias conditions of the visible and hidden components, respectively. The hidden *j* and visible component i are given different weights by se_{ij}. The following determines the combined unit's configuration:

$$Pr(MN, WM, \theta) = \frac{1}{PC(\theta)} exp(-Enr(MN, WM, \theta)) \tag{13}$$

$$QC(\theta) = \sum_{MN} \sum_{WM} Enr(MN, WM, \theta) \tag{14}$$

where $QC(\theta)$ denotes a regularisation constant. In the network, the energy function acts as a probability distribution, and Equation (12) can be used to modify the training vector. It is not recommended to use the RBN's hidden layers alone to extract features from the data [105–109]. The output of the RBN from layer one serves as the input for layer two, and layer two's output serves as the input for layer three. A hierarchical approach to DBN, which is created by the hierarchical layer-by-layer RBN structure, is more effective in extracting characteristics from the dataset [110–112]. The DBN architecture is displayed in Figure 10. Also, Algorithm 1 shows the steps in classification via DBN.

Algorithm 1: Classification via DBN

Input: $I : \mathrm{I} = \{\mathrm{i}_1, \mathrm{i}_2, \ldots\ldots, \mathrm{i}_\mathrm{n}\}$; ***image fames***
Output: $\mathrm{C} = (n_0, n_1, \ldots\ldots, n_N)$: the classification;
$D \leftarrow$ []: Vehicle Detections
$F \leftarrow$ []: Feature Vector
Method:
Video = VideoReader ('videopath')
img_frame = read (video)
for k = 1 to size (img_frame)
    resize_img = imresize (img_frame$_\mathrm{k}$, 768 $\times$ 768)
    seg_img = FCM (resize_img)
        $D \leftarrow$ YOLOv8 (seg_img)
        for s = 1 to size D
            F $\leftarrow$ SIFT ($\mathrm{D_s}$)
            F $\leftarrow$ KAZE ($\mathrm{D_s}$)
            F $\leftarrow$ ORB ($\mathrm{D_s}$)
            veh-class = DBN (F)
        end for
        return *veh-class*
return img_frame

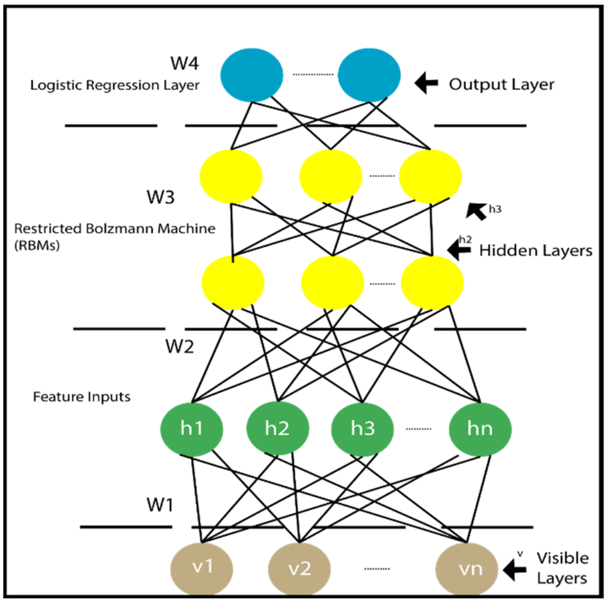


Figure 10. The detailed architecture of DBN classifier.

4. Experimental Setup and Evaluation

Experiments were conducted on a computer with the specs Intel Core i5-7200U 2.30 GHz processor, 8 GB RAM, and x64-based Windows 10. Results were obtained using Google Colab. The system examined the proposed architecture's performance on three benchmark datasets called: VEDAI, and VAID datasets. To evaluate the dependability of our suggested system, the k-fold cross-validation is applied on all three datasets. This section describes the dataset, details the trials, and compares the system to other state-of-the-art technologies.

4.1. Dataset Description

4.1.1. VEDAI Dataset

The VEDAI [113] is a public dataset for vehicle detection in aerial imagery. It was proposed in 2015. The collection aids researchers in locating cars in aerial photographs. The dataset contains miniature automobiles with a variety of properties, including variable lighting conditions, shadows, and obstructed objects. In this dataset, vehicles are classified into nine separate categories: “car”, “truck”, “pick-up”, “plane”, “tractor”, “boat”, “camping car”, “van”, and the “other” category. The average number of cars is 5.5, and they take up around 0.7% of the total number of pixels in each photograph. It also includes a common technique for replicating and contrasting the findings of other studies. Figure 11 shows some of the images from the VEDAI dataset.

Figure 11. Sample images frame from the VEDAI dataset.

4.1.2. VAID Dataset

The VAID dataset [114] included 6000 vehicle photos that were divided into seven categories, including minibus, truck, cement truck, sedan, pickup, bus, trailer, and truck. These images were taken by a drone in various lighting situations. The drone was positioned between 90 m and 95 m above the ground. Images taken at 23.98 frames per second have a resolution of 2720 × 1530. Ten locations in southern Taiwan’s traffic and road conditions are included in the dataset. The traffic images show an urban setting, a suburban city, and a university campus. Figure 12 shows sample photos from the VAID dataset.

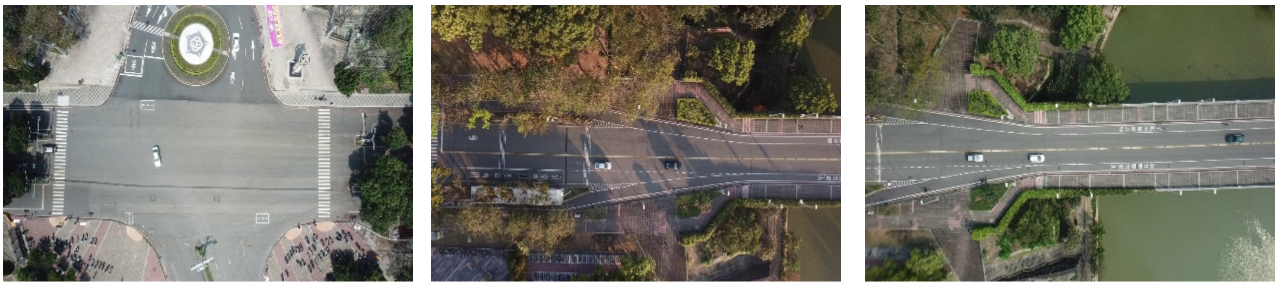


Figure 12. Sample images frame from the VAID dataset.

4.2. Performance Metric and Experimental Outcome

The studies demonstrated the efficiency of the proposed system after we analyzed its performance across the two datasets. Figures 13 and 14 represent the classification accuracies of both the datasets. Tables 2 and 3 demonstrated the vehicle detection accuracies, precision, recall, and F1-score. Tables 4 and 5 illustrates the confusion matrices for the VEDAI and VAID dataset achieving an accuracy of 95.6% and 94.6%, respectively. The experiments were repeated to assess the effectiveness of the findings. The comparison of our system with other widely used research models is shown in Table 6.

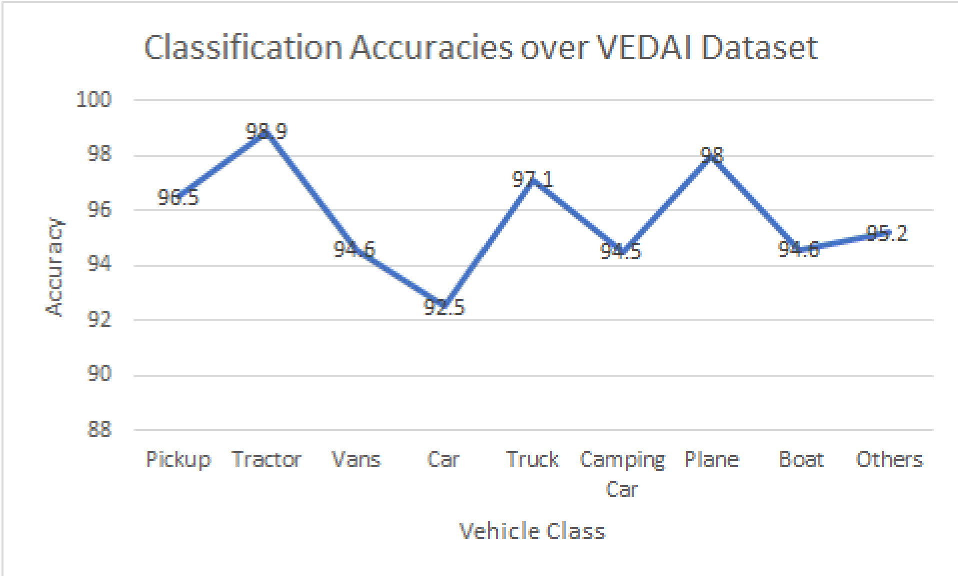


Figure 13. Vehicle Classification accuracies over the VEDAI Dataset.

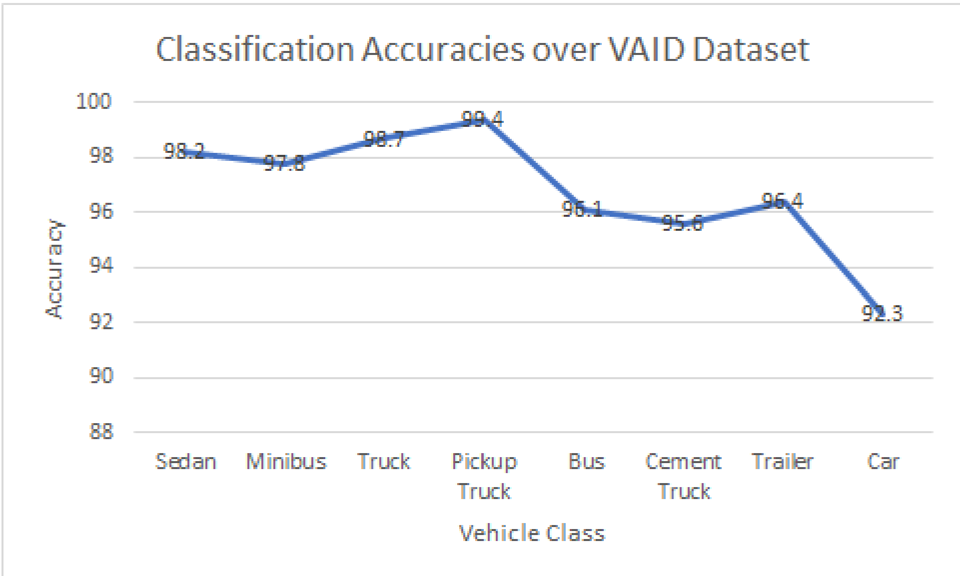


Figure 14. Vehicle Classification accuracies over the VAID Dataset.

Table 2. Overall accuracy, precision, recall, and F1-score for vehicle detection over the VEDAI dataset.

Vehicle Class	Precision	Recall	F1-Score
Pickup	0.985	0.967	0.975
Tractor	0.991	0.987	0.988
Vans	0.941	0.958	0.949
s	0.907	0.910	0.908
Truck	0.934	0.971	0.952
Camping Car	0.956	0.945	0.950
Plane	0.977	0.936	0.956
Boat	0.965	0.971	0.968
Others	0.962	0.934	0.947
Mean	**0.957**	**0.953**	**0.955**

Table 3. Overall accuracy, precision, recall, and F1-score for vehicle detection over the VAID dataset.

Vehicle Class	Precision	Recall	F1-Score
Sedan	0.963	0.974	0.968
Minibus	0.986	0.965	0.975
Truck	0.975	0.989	0.982
PickupTruck	0.988	0.946	0.967
Bus	0.941	0.978	0.959
Cement Truck	0.944	0.912	0.927
Trailer	0.973	0.956	0.964
Car	0.945	0.901	0.922
Mean	**0.964**	**0.953**	**0.958**

Table 4. Confusion matrix for vehicle classification by proposed approach on the VEDAI dataset.

Vehicle Class	Pickup	Tractor	Vans	Car	Truck	Camping Car	Plane	Boat	Others
Pickup	0.98	0	0	0	0	0	0	0	0
Tractor	0.02	0.97	0	0	0	0.01	0	0	0
Vans	0	0.01	0.95	0.02	0	0.02	0	0	0
Car	0	0	0.04	0.93	0	0.02	0	0	0.01
Truck	0	0.03	0	0	0.97	0	0	0	0
Camping Car	0.02	0	0.03	0.02	0.01	0.92	0	0	0
Plane	0	0	0	0	0	0	0.96	0	0.04
Boat	0	0	0	0	0	0	0.01	0.95	0.04
Others	0	0	0	0	0	0	0.01	0.02	0.97
Mean = 95.6%									

Highlights show the score for correct classification for each class.

Table 5. Confusion matrix for vehicle classification by proposed approach on the VAID dataset.

Vehicle Class	Sedan	Minibus	Truck	Pickup Truck	Bus	Cement Truck	Trailer	Car
Sedan	0.98	0.01	0.01	0	0	0	0	0
Minibus	0	0.95	0.02	0	0.03	0	0	0
Truck	0	0.01	0.99	0	0	0	0	0
Pickup Truck	0	0.01	0	0.96	0.02	0	0.01	0
Bus	0.01	0.02	0	0	0.97	0	0	0
Cement Truck	0.01	0	0	0	0	0.99	0	0
Trailer	0.01	0	0	0.01	0	0.01	0.98	0
Car	0.03	0.01	0.01	0	0	0	0.02	0.93
Mean = 94.6%								

Highlights show the score for correct classification for each class.

Table 6. Comparison of the proposed method with conventional systems over VEDAI and VAID Datasets.

Methods	VEDAI	VAID
Wang et al. [115]	93.96	-
Mandal et al. [116]	51.95	-
Terrail et al. [117]	83.50	-
Wang et al. [118]	91.27	-
Lin et al. [114]	-	89.3
Rafique et al. [1]	92.2	-
Hou et al. [119]	75.54	-
Our proposed Model	**95.6**	**94.6**

The results and comparison with other models show that our model performs well in detection and classification of the vehicles in aerial images. Additionally, YOLOv8 is an efficient algorithm in detecting objects of different sizes and appearance. Moreover, the classification accuracy can be improved further by extracting more useful features that are based on the texture and shape of the objects.

5. Conclusions

This study proposes an innovative method for identifying and categorizing vehicles in aerial image sequences. The model preprocesses the aerial images for noise removal before the detection phase. To reduce the complexity, all the images are segmented using the FCM segmentation technique. The vehicle detection task is accomplished using the YOLOv8 algorithm. All the detected vehicles are subjected to SIFT, KAZE, and ORB feature extraction. The extracted feature is then used to train the DBN classifier to classify vehicles into their corresponding classes. The proposed technique has produced promising results over both datasets. The accuracy attained over the VEDAI dataset is 95.6%, and on VAID it was 94.6%.

The proposed system needs to be trained with more vehicle classes. Also, more features can be added to improve the classification accuracy of the vehicles. In the future, to increase the efficiency of our system and make it a standard for all traffic environments, we intend to add more features and reliable algorithms.

Author Contributions: Conceptualization: A.M.Q., N.A.M. and A.A. (Asaad Algarni); methodology: A.M.Q. and M.A. (Mohammed Alonazi); software: A.M.Q. and M.A. (Maha Abdelhaq); validation: N.A.M., M.A. (Mohammed Alonazi) and A.A. (Abdulwahab Alazeb); formal analysis: A.A. (Abdullah Alshahrani) and N.A.M.; resources: N.A.M., A.A. (Asaad Algarni), M.A. (Maha Abdelhaq) and A.A. (Abdulwahab Alazeb); writing—review and editing: N.A.M. and A.M.Q.; funding acquisition: N.A.M., M.A. (Maha Abdelhaq), A.A. (Asaad Algarni), A.A. (Abdulwahab Alazeb) and A.A. (Abdullah Alshahrani). All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R97), Riyadh, Saudi Arabia. This research was supported by the Deanship of Scientific Research at Najran University, under the Research Group Funding program grant code (NU/RG/SERC/12/40). This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are thankful to Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R97), Riyadh, Saudi Arabia. The authors are thankful to Prince Sattam bin Abdulaziz University for supporting this study by project number (PSAU/2023/R/1444).

Conflicts of Interest: The authors declare that they have no conflict of interest to report regarding the present study.

References

1. Rafique, A.A.; Al-Rasheed, A.; Ksibi, A.; Ayadi, M.; Jalal, A.; Alnowaiser, K.; Meshref, H.; Shorfuzzaman, M.; Gochoo, M.; Park, J. Smart Traffic Monitoring Through Pyramid Pooling Vehicle Detection and Filter-Based Tracking on Aerial Images. *IEEE Access* **2023**, *11*, 2993–3007. [CrossRef]
2. Qureshi, A.M.; Jalal, A. Vehicle Detection and Tracking Using Kalman Filter Over Aerial Images. In Proceedings of the 2023 4th International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, 20–22 February 2023; pp. 1–6. [CrossRef]
3. Yang, M.; Wang, Y.; Wang, C.; Liang, Y.; Yang, S.; Wang, L.; Wang, S. Digital Twin-Driven Industrialization Development of Underwater Gliders. *IEEE Trans. Ind. Inform.* **2023**, *19*, 9680–9690. [CrossRef]
4. Liu, L.; Zhang, S.; Zhang, L.; Pan, G.; Yu, J. Multi-UUV Maneuvering Counter-Game for Dynamic Target Scenario Based on Fractional-Order Recurrent Neural Network. *IEEE Trans. Cybern.* **2022**, *53*, 4015–4028. [CrossRef]

5. Zhou, D.; Sheng, M.; Li, J.; Han, Z. Aerospace Integrated Networks Innovation for Empowering 6G: A Survey and Future Challenges. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 975–1019. [CrossRef]
6. Jiang, S.; Zhao, C.; Zhu, Y.; Wang, C.; Du, Y. A Practical and Economical Ultra-Wideband Base Station Placement Approach for Indoor Autonomous Driving Systems. *J. Adv. Transp.* **2022**, *2022*, 3815306. [CrossRef]
7. Schreuder, M.; Hoogendoorn, S.P.; Van Zulyen, H.J.; Gorte, B.; Vosselman, G. Traffic Data Collection from Aerial Imagery. In Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems, Shanghai, China, 12–15 October 2003; Volume 1, pp. 779–784. [CrossRef]
8. Ahmed, A.; Jalal, A.; Rafique, A.A. Salient Segmentation Based Object Detection and Recognition Using Hybrid Genetic Transform. In Proceedings of the 2019 International Conference on Applied and Engineering Mathematics (ICAEM), Taxila, Pakistan, 27–29 August 2019; pp. 203–208.
9. Farooq, A.; Jalal, A.; Kamal, S. Dense RGB-D Map-Based Human Tracking and Activity Recognition Using Skin Joints Features and Self-Organizing Map. *KSII Trans. Internet Inf. Syst.* **2015**, *9*, 1856–1869. [CrossRef]
10. Hsieh, J.W.; Yu, S.H.; Chen, Y.S.; Hu, W.F. Automatic Traffic Surveillance System for Vehicle Tracking and Classification. *IEEE Intell. Transp. Syst. Mag.* **2006**, *7*, 175–187. [CrossRef]
11. Bai, X.; Huang, M.; Xu, M.; Liu, J. Reconfiguration Optimization of Relative Motion Between Elliptical Orbits Using Lyapunov-Floquet Transformation. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *59*, 923–936. [CrossRef]
12. Min, H.; Fang, Y.; Wu, X.; Lei, X.; Chen, S.; Teixeira, R.; Zhu, B.; Zhao, X.; Xu, Z. A Fault Diagnosis Framework for Autonomous Vehicles with Sensor Self-Diagnosis. *Expert Syst. Appl.* **2023**, *224*, 120002. [CrossRef]
13. Zhang, X.; Wen, S.; Yan, L.; Feng, J.; Xia, Y. A Hybrid-Convolution Spatial–Temporal Recurrent Network For Traffic Flow Prediction. *Comput. J.* **2022**, *10*, bxac171. [CrossRef]
14. Li, B.; Zhou, X.; Ning, Z.; Guan, X.; Yiu, K.F.C. Dynamic Event-Triggered Security Control for Networked Control Systems with Cyber-Attacks: A Model Predictive Control Approach. *Inf. Sci.* **2022**, *612*, 384–398. [CrossRef]
15. Xu, J.; Park, H.; Guo, K.; Zhang, X. The Alleviation of Perceptual Blindness during Driving in Urban Areas Guided by Saccades Recommendation. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 16386–16396. [CrossRef]
16. Xu, J.; Guo, K.; Zhang, X.; Sun, P.Z.H. Left Gaze Bias between LHT and RHT: A Recommendation Strategy to Mitigate Human Errors in Left- and Right-Hand Driving. *IEEE Trans. Intell. Veh.* **2023**, *1*, 1–12. [CrossRef]
17. Qureshi, A.M.; Butt, A.H.; Jalal, A. Highway Traffic Surveillance Over UAV Dataset via Blob Detection and Histogram of Gradient. In Proceedings of the 2023 4th International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, 20–22 February 2023; pp. 1–5. [CrossRef]
18. Torres, D.L.; Turnes, J.N.; Vega, P.J.S.; Feitosa, R.Q.; Silva, D.E.; Marcato Junior, J.; Almeida, C. Deforestation Detection with Fully Convolutional Networks in the Amazon Forest from Landsat-8 and Sentinel-2 Images. *Remote Sens.* **2021**, *13*, 5084. [CrossRef]
19. Chen, P.C.; Chiang, Y.C.; Weng, P.Y. Imaging Using Unmanned Aerial Vehicles for Agriculture Land Use Classification. *Agriculture* **2020**, *10*, 416. [CrossRef]
20. Munawar, H.S.; Ullah, F.; Qayyum, S.; Khan, S.I.; Mojtahedi, M. UAVs in Disaster Management: Application of Integrated Aerial Imagery and Convolutional Neural Network for Flood Detection. *Sustainability* **2021**, *13*, 7547. [CrossRef]
21. Zhang, J.; Ye, G.; Tu, Z.; Qin, Y.; Qin, Q.; Zhang, J.; Liu, J. A Spatial Attentive and Temporal Dilated (SATD) GCN for Skeleton-Based Action Recognition. *CAAI Trans. Intell. Technol.* **2022**, *7*, 46–55. [CrossRef]
22. Ma, X.; Dong, Z.; Quan, W.; Dong, Y.; Tan, Y. Real-Time Assessment of Asphalt Pavement Moduli and Traffic Loads Using Monitoring Data from Built-in Sensors: Optimal Sensor Placement and Identification Algorithm. *Mech. Syst. Signal Process.* **2023**, *187*, 109930. [CrossRef]
23. Chen, J.; Wang, Q.; Peng, W.; Xu, H.; Li, X.; Xu, W. Disparity-Based Multiscale Fusion Network for Transportation Detection. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 18855–18863. [CrossRef]
24. Chen, J.; Xu, M.; Xu, W.; Li, D.; Peng, W.; Xu, H. A Flow Feedback Traffic Prediction Based on Visual Quantified Features. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 10067–10075. [CrossRef]
25. Zheng, Y.; Zhang, Y.; Qian, L.; Zhang, X.; Diao, S.; Liu, X.; Cao, J.; Huang, H. A Lightweight Ship Target Detection Model Based on Improved YOLOv5s Algorithm. *PLoS ONE* **2023**, *18*, e0283932. [CrossRef]
26. Arinaldi, A.; Pradana, J.A.; Gurusinga, A.A. Detection and Classification of Vehicles for Traffic Video Analytics. *Procedia Comput. Sci.* **2018**, *144*, 259–268. [CrossRef]
27. Aqel, S.; Hmimid, A.; Sabri, M.A.; Aarab, A. Road Traffic: Vehicle Detection and Classification. In Proceedings of the 2017 Intelligent Systems and Computer Vision (ISCV), Venice, Italy, 17–19 April 2017. [CrossRef]
28. Sarikan, S.S.; Ozbayoglu, A.M.; Zilci, O. Automated Vehicle Classification with Image Processing and Computational Intelligence. *Procedia Comput. Sci.* **2017**, *114*, 515–522. [CrossRef]
29. Tan, Y.; Xu, Y.; Das, S.; Chaudhry, A. Vehicle Detection and Classification in Aerial Imagery. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 86–90. [CrossRef]
30. Hamzenejadi, M.H.; Mohseni, H. Fine-Tuned YOLOv5 for Real-Time Vehicle Detection in UAV Imagery: Architectural Improvements and Performance Boost. *Expert Syst. Appl.* **2023**, *231*, 120845. [CrossRef]
31. Ozturk, M.; Cavus, E. Vehicle Detection in Aerial Imaginary Using a Miniature CNN Architecture. In Proceedings of the 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Kocaeli, Turkey, 25–27 August 2021. [CrossRef]

32. Roopa Chandrika, R.; Gowri Ganesh, N.S.; Mummoorthy, A.; Karthick Raghunath, K.M. Vehicle Detection and Classification Using Image Processing. In Proceedings of the 2019 International Conference on Emerging Trends in Science and Engineering (ICESE), Hyderabad, India, 18–19 September 2019. [CrossRef]
33. Kumar, S.; Jain, A.; Rani, S.; Alshazly, H.; Idris, S.A.; Bourouis, S. Deep Neural Network Based Vehicle Detection and Classification of Aerial Images. *Intell. Autom. Soft Comput.* **2022**, *34*, 119–131. [CrossRef]
34. Zhang, X.; Zhu, X. Vehicle Detection in the Aerial Infrared Images via an Improved Yolov3 Network. In Proceedings of the 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), Wuxi, China, 19–21 July 2019; pp. 372–376. [CrossRef]
35. Javid, A.; Nekoui, M.A. Adaptive Control of Time-Delayed Bilateral Teleoperation Systems with Uncertain Kinematic and Dynamics. *Cogent Eng.* **2018**, *6*, 1631604. [CrossRef]
36. Lu, S.; Ding, Y.; Liu, M.; Yin, Z.; Yin, L.; Zheng, W. Multiscale Feature Extraction and Fusion of Image and Text in VQA. *Int. J. Comput. Intell. Syst.* **2023**, *16*, 54. [CrossRef]
37. Cheng, B.; Zhu, D.; Zhao, S.; Chen, J. Situation-Aware IoT Service Coordination Using the Event-Driven SOA Paradigm. *IEEE Trans. Netw. Serv. Manag.* **2016**, *13*, 349–361. [CrossRef]
38. Shen, Y.; Ding, N.; Zheng, H.T.; Li, Y.; Yang, M. Modeling Relation Paths for Knowledge Graph Completion. *IEEE Trans. Knowl. Data Eng.* **2021**, *33*, 3607–3617. [CrossRef]
39. Zhou, X.; Zhang, L. SA-FPN: An Effective Feature Pyramid Network for Crowded Human Detection. *Appl. Intell.* **2022**, *52*, 12556–12568. [CrossRef]
40. Zhao, Z.; Xu, G.; Zhang, N.; Zhang, Q. Performance Analysis of the Hybrid Satellite-Terrestrial Relay Network with Opportunistic Scheduling over Generalized Fading Channels. *IEEE Trans. Veh. Technol.* **2022**, *71*, 2914–2924. [CrossRef]
41. Chen, J.; Wang, Q.; Cheng, H.H.; Peng, W.; Xu, W. A Review of Vision-Based Traffic Semantic Understanding in ITSs. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 19954–19979. [CrossRef]
42. Hou, X.; Zhang, L.; Su, Y.; Gao, G.; Liu, Y.; Na, Z.; Xu, Q.Z.; Ding, T.; Xiao, L.; Li, L.; et al. A Space Crawling Robotic Bio-Paw (SCRBP) Enabled by Triboelectric Sensors for Surface Identification. *Nano Energy* **2023**, *105*, 108013. [CrossRef]
43. Yu, J.; Shi, Z.; Dong, X.; Li, Q.; Lv, J.; Ren, Z. Impact Time Consensus Cooperative Guidance Against the Maneuvering Target: Theory and Experiment. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *59*, 4590–4603. [CrossRef]
44. Fang, Y.; Min, H.; Wu, X.; Wang, W.; Zhao, X.; Mao, G. On-Ramp Merging Strategies of Connected and Automated Vehicles Considering Communication Delay. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 15298–15312. [CrossRef]
45. Balasamy, K.; Shamia, D. Feature Extraction-Based Medical Image Watermarking Using Fuzzy-Based Median Filter. *IETE J. Res.* **2021**, *69*, 83–91. [CrossRef]
46. Somvanshi, S.S.; Kunwar, P.; Tomar, S.; Singh, M. Comparative Statistical Analysis of the Quality of Image Enhancement Techniques. *Int. J. Image Data Fusion* **2018**, *9*, 131–151. [CrossRef]
47. Zaman Khan, R. Hand Gesture Recognition: A Literature Review. *Int. J. Artif. Intell. Appl.* **2012**, *3*, 161–174. [CrossRef]
48. Liu, W.; Zhou, F.; Lu, T.; Duan, J.; Qiu, G. Image Defogging Quality Assessment: Real-World Database and Method. *IEEE Trans. Image Process.* **2021**, *30*, 176–190. [CrossRef] [PubMed]
49. Kong, X.; Chen, Q.; Gu, G.; Ren, K.; Qian, W.; Liu, Z. Particle Filter-Based Vehicle Tracking via HOG Features after Image Stabilisation in Intelligent Drive System. *IET Intell. Transp. Syst.* **2019**, *13*, 942–949. [CrossRef]
50. Xu, G.; Su, J.; Pan, H.; Zhang, Z.; Gong, H. An Image Enhancement Method Based on Gamma Correction. In Proceedings of the 2009 Second International Symposium on Computational Intelligence and Design, Washington, DC, USA, 12–14 December 2009; Volume 1, pp. 60–63. [CrossRef]
51. Veluchamy, M.; Subramani, B. Image Contrast and Color Enhancement Using Adaptive Gamma Correction and Histogram Equalization. *Optik* **2019**, *183*, 329–337. [CrossRef]
52. Liu, H.; Yuan, H.; Liu, Q.; Hou, J.; Zeng, H.; Kwong, S. A Hybrid Compression Framework for Color Attributes of Static 3D Point Clouds. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 1564–1577. [CrossRef]
53. Luo, J.; Wang, G.; Li, G.; Pesce, G. Transport Infrastructure Connectivity and Conflict Resolution: A Machine Learning Analysis. *Neural Comput. Appl.* **2022**, *34*, 6585–6601. [CrossRef]
54. Liu, Q.; Yuan, H.; Hamzaoui, R.; Su, H.; Hou, J.; Yang, H. Reduced Reference Perceptual Quality Model with Application to Rate Control for Video-Based Point Cloud Compression. *IEEE Trans. Image Process.* **2021**, *30*, 6623–6636. [CrossRef] [PubMed]
55. Yang, B.; Wang, J.; Clark, R.; Hu, Q.; Wang, S.; Markham, A.; Trigoni, N. Learning Object Bounding Boxes for 3D Instance Segmentation on Point Clouds. *arXiv* **2019**, arXiv:1906.01140.
56. Rafique, A.A.; Gochoo, M.; Jalal, A.; Kim, K. Maximum Entropy Scaled Super Pixels Segmentation for Multi-Object Detection and Scene Recognition via Deep Belief Network. *Multimed. Tools Appl.* **2022**, *82*, 13401–13430. [CrossRef]
57. Li, J.; Han, L.; Zhang, C.; Li, Q.; Liu, Z. Spherical Convolution Empowered Viewport Prediction in 360 Video Multicast with Limited FoV Feedback. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *19*, 1–23. [CrossRef]
58. Liang, X.; Huang, Z.; Yang, S.; Qiu, L. Device-Free Motion & Trajectory Detection via RFID. *ACM Trans. Embed. Comput. Syst.* **2018**, *17*, 1–27. [CrossRef]
59. Jalal, A.; Ahmed, A.; Rafique, A.A.; Kim, K. Scene Semantic Recognition Based on Modified Fuzzy C-Mean and Maximum Entropy Using Object-to-Object Relations. *IEEE Access* **2021**, *9*, 27758–27772. [CrossRef]
60. Miao, J.; Zhou, X.; Huang, T.Z. Local Segmentation of Images Using an Improved Fuzzy C-Means Clustering Algorithm Based on Self-Adaptive Dictionary Learning. *Appl. Soft Comput.* **2020**, *91*, 106200. [CrossRef]

61. Jun, M.; Yuanyuan, L.; Huahua, L.; You, M. Single-Image Dehazing Based on Two-Stream Convolutional Neural Network. *J. Artif. Intell. Technol.* **2022**, *2*, 100–110. [CrossRef]
62. Yu, H.; Wu, Z.; Wang, S.; Wang, Y.; Ma, X. Spatiotemporal Recurrent Convolutional Networks for Traffic Prediction in Transportation Networks. *Sensors* **2017**, *17*, 1501. [CrossRef] [PubMed]
63. Pan, S.; Xu, M.; Zhu, S.; Lin, M.; Li, G. A Low-Profile Programmable Beam Scanning Array Antenna. In Proceedings of the 2021 International Conference on Microwave and Millimeter Wave Technology (ICMMT), Nanjing, China, 23–26 May 2021. [CrossRef]
64. Zong, C.; Wan, Z. Container Ship Cell Guide Accuracy Check Technology Based on Improved 3d Point Cloud Instance Segmentation. *Brodogradnja* **2022**, *73*, 23–35. [CrossRef]
65. Han, Y.; Wang, B.; Guan, T.; Tian, D.; Yang, G.; Wei, W.; Tang, H.; Chuah, J.H. Research on Road Environmental Sense Method of Intelligent Vehicle Based on Tracking Check. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 1261–1275. [CrossRef]
66. Cao, B.; Zhang, W.; Wang, X.; Zhao, J.; Gu, Y.; Zhang, Y. A Memetic Algorithm Based on Two_Arch2 for Multi-Depot Heterogeneous-Vehicle Capacitated Arc Routing Problem. *Swarm Evol. Comput.* **2021**, *63*, 100864. [CrossRef]
67. Dai, X.; Xiao, Z.; Jiang, H.; Chen, H.; Min, G.; Dustdar, S.; Cao, J. A Learning-Based Approach for Vehicle-to-Vehicle Computation Offloading. *IEEE Internet Things J.* **2023**, *10*, 7244–7258. [CrossRef]
68. Xiao, Z.; Fang, H.; Jiang, H.; Bai, J.; Havyarimana, V.; Chen, H.; Jiao, L. Understanding Private Car Aggregation Effect via Spatio-Temporal Analysis of Trajectory Data. *IEEE Trans. Cybern.* **2023**, *53*, 2346–2357. [CrossRef] [PubMed]
69. Mi, C.; Huang, S.; Zhang, Y.; Zhang, Z.; Postolache, O. Design and Implementation of 3-D Measurement Method for Container Handling Target. *J. Mar. Sci. Eng.* **2022**, *10*, 1961. [CrossRef]
70. Jiang, H.; Chen, S.; Xiao, Z.; Hu, J.; Liu, J.; Dustdar, S. Pa-Count: Passenger Counting in Vehicles Using Wi-Fi Signals. *IEEE Trans. Mob. Comput.* **2023**, *1*, 1–14. [CrossRef]
71. Ding, Y.; Zhang, W.; Zhou, X.; Liao, Q.; Luo, Q.; Ni, L.M. FraudTrip: Taxi Fraudulent Trip Detection from Corresponding Trajectories. *IEEE Internet Things J.* **2021**, *8*, 12505–12517. [CrossRef]
72. Tian, H.; Pei, J.; Huang, J.; Li, X.; Wang, J.; Zhou, B.; Qin, Y.; Wang, L. Garlic and Winter Wheat Identification Based on Active and Passive Satellite Imagery and the Google Earth Engine in Northern China. *Remote Sens.* **2020**, *12*, 3539. [CrossRef]
73. Yang, M.; Wang, H.; Hu, K.; Yin, G.; Wei, Z. IA-Net: An Inception-Attention-Module-Based Network for Classifying Underwater Images From Others. *IEEE J. Ocean. Eng.* **2022**, *47*, 704–717. [CrossRef]
74. Shi, Y.; Hu, J.; Wu, Y.; Ghosh, B.K. Intermittent Output Tracking Control of Heterogeneous Multi-Agent Systems over Wide-Area Clustered Communication Networks. *Nonlinear Anal. Hybrid Syst.* **2023**, *50*, 101387. [CrossRef]
75. Lu, S.; Liu, M.; Yin, L.; Yin, Z.; Liu, X.; Zheng, W. The Multi-Modal Fusion in Visual Question Answering: A Review of Attention Mechanisms. *PeerJ Comput. Sci.* **2023**, *9*, e1400. [CrossRef] [PubMed]
76. Lou, H.; Duan, X.; Guo, J.; Liu, H.; Gu, J.; Bi, L.; Chen, H. DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor. *Electronics* **2023**, *12*, 2323. [CrossRef]
77. Zhang, X.; Fang, S.; Shen, Y.; Yuan, X.; Lu, Z. Hierarchical Velocity Optimization for Connected Automated Vehicles with Cellular Vehicle-to-Everything Communication at Continuous Signalized Intersections. *IEEE Trans. Intell. Transp. Syst.* **2023**, *1*, 1–12. [CrossRef]
78. Tang, J.; Ren, Y.; Liu, S. Real-Time Robot Localization, Vision, and Speech Recognition on Nvidia Jetson TX1. *arXiv* **2017**, arXiv:1705.10945.
79. Guo, F.; Zhou, W.; Lu, Q.; Zhang, C. Path Extension Similarity Link Prediction Method Based on Matrix Algebra in Directed Networks. *Comput. Commun.* **2022**, *187*, 83–92. [CrossRef]
80. Wang, B.; Zhang, Y.; Zhang, W. A Composite Adaptive Fault-Tolerant Attitude Control for a Quadrotor UAV with Multiple Uncertainties. *J. Syst. Sci. Complex.* **2022**, *35*, 81–104. [CrossRef]
81. Ahmad, F. Deep Image Retrieval Using Artificial Neural Network Interpolation and Indexing Based on Similarity Measurement. *CAAI Trans. Intell. Technol.* **2022**, *7*, 200–218. [CrossRef]
82. Hassan, F.S.; Gutub, A. Improving Data Hiding within Colour Images Using Hue Component of HSV Colour Space. *CAAI Trans. Intell. Technol.* **2022**, *7*, 56–68. [CrossRef]
83. Dong, Y.; Guo, W.; Zha, F.; Liu, Y.; Chen, C.; Sun, L. A Vision-Based Two-Stage Framework for Inferring Physical Properties of the Terrain. *Appl. Sci.* **2020**, *10*, 6473. [CrossRef]
84. Bawankule, R.; Gaikwad, V.; Kulkarni, I.; Kulkarni, S.; Jadhav, A.; Ranjan, N. Visual Detection of Waste Using YOLOv8. In Proceedings of the 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 14–16 June 2023; pp. 869–873. [CrossRef]
85. Wen, C.; Huang, Y.; Davidson, T.N. Efficient Transceiver Design for MIMO Dual-Function Radar-Communication Systems. *IEEE Trans. Signal Process.* **2023**, *71*, 1786–1801. [CrossRef]
86. Wen, C.; Huang, Y.; Zheng, L.; Liu, W.; Davidson, T.N. Transmit Waveform Design for Dual-Function Radar-Communication Systems via Hybrid Linear-Nonlinear Precoding. *IEEE Trans. Signal Process.* **2023**, *71*, 2130–2145. [CrossRef]
87. Ning, Z.; Wang, T.; Zhang, K. Dynamic Event-Triggered Security Control and Fault Detection for Nonlinear Systems with Quantization and Deception Attack. *Inf. Sci.* **2022**, *594*, 43–59. [CrossRef]
88. Yu, S.; Zhao, C.; Song, L.; Li, Y.; Du, Y. Understanding Traffic Bottlenecks of Long Freeway Tunnels Based on a Novel Location-Dependent Lighting-Related Car-Following Model. *Tunn. Undergr. Sp. Technol.* **2023**, *136*, 105098. [CrossRef]

89. Peng, J.; Wang, N.; El-Latif, A.A.A.; Li, Q.; Niu, X. Finger-Vein Verification Using Gabor Filter and SIFT Feature Matching. In Proceedings of the 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Piraeus/Athens, Greece, 18–20 July 2022; pp. 45–48. [CrossRef]
90. Hua, Y.; Lin, J.; Lin, C. An Improved SIFT Feature Matching Algorithm. In Proceedings of the 2010 8th World Congress on Intelligent Control and Automation, Jinan, China, 7–9 July 2010; pp. 6109–6113. [CrossRef]
91. Yawen, T.; Jinxu, G. Research on Vehicle Detection Technology Based on SIFT Feature. In Proceedings of the 2018 8th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 15–17 June 2018; pp. 274–278. [CrossRef]
92. Xiaohui, H.; Qiuhua, K.; Qianhua, C.; Yun, X.; Weixing, Z.; Li, Y. A Coherent Pattern Mining Algorithm Based on All Contiguous Column Bicluster. *J. Artif. Intell. Technol.* **2022**, *2*, 80–92. [CrossRef]
93. Alhwarin, F.; Wang, C.J.; Ristic-Durrant, D.; Gräser, A. Improved SIFT-Features Matching for Object Recognition. In Proceedings of the Visions of Computer Science—BCS International Academic Conference (VOCS), London, UK, 22–24 September 2008. [CrossRef]
94. Battiato, S.; Gallo, G.; Puglisi, G.; Scellato, S. SIFT Features Tracking for Video Stabilization. In Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP 2007), Modena, Italy, 10–14 September 2007; pp. 825–830. [CrossRef]
95. Mu, K.; Hui, F.; Zhao, X. Multiple Vehicle Detection and Tracking in Highway Traffic Surveillance Video Based on Sift Feature Matching. *J. Inf. Process. Syst.* **2016**, *12*, 183–195. [CrossRef]
96. Alcantarilla, P.F.; Bartoli, A.; Davison, A.J. KAZE Features. *Lect. Notes Comput. Sci.* **2012**, *7577*, 214–227. [CrossRef]
97. Sharma, T.; Jain, A.; Verma, N.K.; Vasikarla, S. Object Counting Using KAZE Features under Different Lighting Conditions for Inventory Management. In Proceedings of the 2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 15–17 October 2019. [CrossRef]
98. Dai, X.; Xiao, Z.; Jiang, H.; Lui, J.C.S. UAV-Assisted Task Offloading in Vehicular Edge Computing Networks. *IEEE Trans. Mob. Comput.* **2023**, *1*, 1–18. [CrossRef]
99. Zhang, Z.; Guo, D.; Zhou, S.; Zhang, J.; Lin, Y. Flight Trajectory Prediction Enabled by Time-Frequency Wavelet Transform. *Nat. Commun.* **2023**, *14*, 5258. [CrossRef] [PubMed]
100. Zhang, C.; Xiao, P.; Zhao, Z.T.; Liu, Z.; Yu, J.; Hu, X.Y.; Chu, H.B.; Xu, J.J.; Liu, M.Y.; Zou, Q.; et al. A Wearable Localized Surface Plasmons Antenna Sensor for Communication and Sweat Sensing. *IEEE Sens. J.* **2023**, *23*, 11591–11599. [CrossRef]
101. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An Efficient Alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571. [CrossRef]
102. Salakhutdinov, R.; Murray, I. On the Quantitative Analysis of Deep Belief Networks. In Proceedings of the 25th international conference on Machine learning, New York, NY, USA, 5–9 July 2008; pp. 872–879. [CrossRef]
103. Zheng, M.; Zhi, K.; Zeng, J.; Tian, C.; You, L. A Hybrid CNN for Image Denoising. *J. Artif. Intell. Technol.* **2022**, *2*, 93–99. [CrossRef]
104. Li, C.; Wang, Y.; Zhang, X.; Gao, H.; Yang, Y.; Wang, J. Deep Belief Network for Spectral–Spatial Classification of Hyperspectral Remote Sensor Data. *Sensors* **2019**, *19*, 204. [CrossRef]
105. Qi, M.; Cui, S.; Chang, X.; Xu, Y.; Meng, H.; Wang, Y.; Yin, T. Multi-Region Nonuniform Brightness Correction Algorithm Based on L-Channel Gamma Transform. *Secur. Commun. Netw.* **2022**, *2022*, 2675950. [CrossRef]
106. Liu, A.A.; Zhai, Y.; Xu, N.; Nie, W.; Li, W.; Zhang, Y. Region-Aware Image Captioning via Interaction Learning. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 3685–3696. [CrossRef]
107. Li, Q.K.; Lin, H.; Tan, X.; Du, S. H∞Consensus for Multiagent-Based Supply Chain Systems under Switching Topology and Uncertain Demands. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *50*, 4905–4918. [CrossRef]
108. Ma, K.; Li, Z.; Liu, P.; Yang, J.; Geng, Y.; Yang, B.; Guan, X. Reliability-Constrained Throughput Optimization of Industrial Wireless Sensor Networks with Energy Harvesting Relay. *IEEE Internet Things J.* **2021**, *8*, 13343–13354. [CrossRef]
109. Yao, Y.; Shu, F.; Li, Z.; Cheng, X.; Wu, L. Secure Transmission Scheme Based on Joint Radar and Communication in Mobile Vehicular Networks. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 10027–10037. [CrossRef]
110. Xu, J.; Guo, K.; Sun, P.Z.H. Driving Performance Under Violations of Traffic Rules: Novice vs. Experienced Drivers. *IEEE Trans. Intell. Veh.* **2022**, *7*, 908–917. [CrossRef]
111. Xu, J.; Pan, S.; Sun, P.Z.H.; Park, S.H.; Guo, K. Human-Factors-in-Driving-Loop: Driver Identification and Verification via a Deep Learning Approach Using Psychological Behavioral Data. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 3383–3394. [CrossRef]
112. Xu, J.; Park, S.H.; Zhang, X.; Hu, J. The Improvement of Road Driving Safety Guided by Visual Inattentional Blindness. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 4972–4981. [CrossRef]
113. Razakarivony, S.; Jurie, F. Vehicle Detection in Aerial Imagery: A Small Target Detection Benchmark. *J. Vis. Commun. Image Represent.* **2016**, *34*, 187–203. [CrossRef]
114. Lin, H.Y.; Tu, K.C.; Li, C.Y. VAID: An Aerial Image Dataset for Vehicle Detection and Classification. *IEEE Access* **2020**, *8*, 212209–212219. [CrossRef]
115. Wang, B.; Xu, B. A Feature Fusion Deep-Projection Convolution Neural Network for Vehicle Detection in Aerial Images. *PLoS ONE* **2021**, *16*, e0250782. [CrossRef]
116. Mandal, M.; Shah, M.; Meena, P.; Devi, S.; Vipparthi, S.K. AVDNet: A Small-Sized Vehicle Detection Network for Aerial Visual Data. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 494–498. [CrossRef]
117. du Terrail, J.O.; Jurie, F. Faster RER-CNN: Application to the Detection of Vehicles in Aerial Images. *arXiv* **2018**, arXiv:1809.07628.

118. Wang, B.; Gu, Y. An Improved FBPN-Based Detection Network for Vehicles in Aerial Images. *Sensors* **2020**, *20*, 4709. [CrossRef]
119. Hou, S.; Fan, L.; Zhang, F.; Liu, B. An Improved Lightweight YOLOv5 for Remote Sensing Images. In Proceedings of the 32nd International Conference on Artificial Neural Networks, Heraklion, Greece, 26–29 September 2023; pp. 77–89. [CrossRef]