International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 06 | June 2022 www.irjet.net p-ISSN: 2395-0072

Automatism System Using Faster R-CNN and SVM

Thrishul P [1], Sindhu Manvitha M V [2], Sivani M [3], Amarnadha Reddy K[4]

[1,2,3,4] *Student, Dept. of CSE, Madanapalle Institute of Technology & Science, Madanapalle, A.P., INDIA.*

Abstract

In worldwide car sales nearly 3000 million registered vehicles in the US, China, India etc., According to various reports, this figure continues to rise year after year. Needless to say, the busiest towns and cities are always clogged with automobiles. As a result, traffic and parking problems are almost constant in many regions of the world. Parking has been a major issue as the number of automobiles has increased. While parking in the large commercial areas, everyone is facing the problem in finding the parking space and in some complex situations, it leads to blocking of vehicles. The use of recent improvements in computer vision to solve this problem has a number of benefits. In this work, we will describe a model that makes use of already installed surveillance cameras. The objects are identified using a faster R-CNN that use the Region Proposal Network. This model detects items in the quickest possible time and may be used to any actual setting. SVM is used for classification of the objects. By this driver are assisted and the vacant places are managed accordingly. Our proposed solution was created for realistic settings, taking into account various closures, light changes, and environmental variables.

Keywords *- **Regional Proposal, Object detection, Classification, SVM (Support Vector Machine), Faster R-CNN (Faster Region Convolutional Neural Network), Image processing***

1. INTRODUCTION

Each year, parking lots spend a significant amount of money. Many vehicles like Car, Bus etc., park administration is expensive and complicated in many circumstances, especially in large regions such as airports, commercial areas, and other such locations. The cost and time required to solve this problem utilising computer vision rather than intrusive sensors such as induction loops or other weight-in-motion sensors are decreased, and the system functions efficiently. It assists drivers to park cars in lower occupancy areas in faster route.

The prior systems were complex and relied on picture segmentation or machine learning (SVMs, NN). However, due to advancements in object detecting algorithms, accurate detection is now achievable. Because most parking lots include surveillance cameras, the solution in many cases is simply to process the data collected by the cameras.

In this paper, we perform object detection using regional proposal technique to identify the objects present in image. The output of model contains different colours of bounding boxes representing the free space and occupied place so that every individual park his/her vehicle without delay.

A deep neural network is made up of several nonlinear process layers that use simple, parallel-operating parts that are powered by biological nerve systems. This system consists of an input layer, many hidden layers, and an output layer. The layers are linked together by nodes, or neurons, with each hidden layer using the previous layer's output as its input. Let's say we've got a collection of pictures wherever every image contains one in all four totally different classes of objects, and that we need the deep learning network to mechanically acknowledge that object is in every image. we tend to label the pictures so as to own coaching information for the network. The network will then begin to comprehend the object's specific options and associate them with the correct class using this coaching knowledge. Every layer of the network takes data from the previous layer, modifies it, and then sends it to the next layer. From layer to layer, the network will improve the quality and detail of what it learns. Notice that the network learns directly from the data - we don't have any influence over what options are being learned.

A deep neural network is made up of several nonlinear process layers that use simple, parallel-operating parts that are powered by biological nerve systems. This system consists of an input layer, many hidden layers, and an output layer. The

layers are connected by nodes, or neurons, with each buried layer using the previous layer's output as its input. One of the most extensively used deep learning approaches is the convolutional neural network (CNN, or ConvNet). A CNN has an input layer, an output layer, and numerous hidden layers in between, just like other neural networks.

The data is processed in one of three ways by these layers: convolution, pooling, or corrected linear measure (ReLU). Convolution uses a succession of convolutional filters to process the input images, each of which activates bound alternatives in the images. Pooling simplifies the output and minimises the number of parameters that the network must learn about corrected linear measure (ReLU), which allows for faster and easier coaching by mapping negative values to zero and keeping positive values. These three operations repeat itself across tens or hundreds of layers, with each layer learning to recognise entirely new options. The architecture of a CNN is critical when it comes to feature detection. The next-to-last layer could be a absolutely connected layer (FC) that outputs a vector of K dimensions wherever K is that the range of categories that the network are going to be able to predict. This vector contains the chances for every category of any image being classified.

The classification output is produced by the final layer of the CNN design, which employs a softmax operation. If you're unaccustomed deep learning, a fast and simple thanks to start is to use associate degree existing network, like GoogLeNet, a CNN trained on quite 1,000,000 pictures. GoogLeNet is most ordinarily used for image classification. It will classify pictures into a thousand completely different classes, as well as keyboards, laptop mice, pencils, and alternative workplace instrumentation, likewise as numerous breeds of dogs, cats, horses, and alternative animals. coaching a deep learning model will take hours, days, or weeks, reckoning on the scale of the information and also the quantity of process power you've got on the market. choosing a procedure resource could be a essential thought after you created your advancement.

2. LITERATURE SURVEY

HaihuiXie et al. a new detection method based on a CNN is proposed, which recognises and locates automobiles. The CNN is trained for object detection using a preset sliding window. After that, a distribution matrix is created to calculate the vehicle density and process so that the vehicles' positions may be accurately determined. In the suggested approach, a CNN was merged with a distribution matrix [1]. Yucheng Guo el al. described an intelligent parking system having multiple layers. We were able to better detect parking spaces using CNN. The position and direction of the objects present are used to evaluate the path. Dynamic Programming is used to find the shortest path. Because of the problems of DP, the interference is eliminated and the path is identified easily in [2]. Imen Masmoudi et al. presented a system for intelligent parking lot management and free parking spot detection. It is a camera-based system, which provides the real time identification of number of vacant places and assist path to drivers. To remove the real world challenges and work effectively, the ABS algorithm combined with the SURF algorithm in [3]. Hilal Al-Kharudi et al. described system captures and processes the image which gets as input from the external cameras. The camera is used because it captures image in which many objects are detected. Through this, the particular car parks which are vacant can be identified. This system has been developed in both software and hardware platform from[4].T.Fabian et al. describes a vision-based system for parking lot occupancy detection that is unsupervised. The system is simpler, with only a few frames every minute. This system is divided into three stages. The image is initially pre-processed. Shadows and other obtrusions are eliminated. Later, the parking spots for individual vehicle are evaluated and at last the best spot is mapped to each vehicle. Because of image processing, the system works accurately in [5]. Junzhao Liu et al. developed a method for locating vacant parking places based on images. Edge detection, colour density, and the foreground/background pixel ratio are all determined and provides the information whether the place is occupied or not in low complexity in [6]. Ching-Chun Huang et al. proposed a system which works in day-and-night. The issues like lighting fluctuations, shadow effects and more are present in day time. Identifying objects at night is difficult. A plane-based method is utilised to solve these difficulties. It utilizes a structural 3-D model with many planar surfaces. By developing Bayesian hierarchical framework, we can identify the objects in both day and night in [7]. Katy Bulmer et al. proposed a system for detection of vacant spots. It combines two types of algorithms. The static one using histogram and edge detection. The dynamic using blob analysis. By this, the system is more accurate and works in different conditions in [8]. Ching-Chun Huang et al. described a system which detects the unoccupied places instead of identifying cars. It considers a parking lot to be a system with a variety of surfaces. The interface of vacant space is done by surface-based hierarchical architecture. Each patch is retrieved using HOG technique in [9]. Chen et al. proposed a system which first tracks the vehicles present based on colour, position, and motion. Later the parking vacant spaces are identified. Then, by using various weights, a hybrid technique is

created for assisting individuals to their respective parking places in optimal time in [10]. The management of these car parks is very expensive and in many cases complex, especially in the case of those that have many places such as airports or large commercial areas. When compared to utilising intrusive sensors like induction loops or other weight-in-motion sensors, using computer vision to tackle this problem offers significant advantages. Surveillance cameras are readily available in most car parking lots, so in many cases the solution is only to adequately process the information available from the already existing cameras, or complete the deployment by adding some cameras to have a full coverage that allows the system to operate. The systems that have been developed so far are primarily based on picture segmentation or machine learning (SVMs, NN). over spot patches, but due to the evolution in the last years of object detection algorithms, it is possible to use the detections of these algorithms for the proper operation of automatic parking management systems. For increasing the coverage area, we have to increase the cameras count. Doesn't work efficiently in all the weather conditions

3. PROPOSED SYSTEM

The technology is built in such a way that it may be used with current parking lot cameras. The system faces many complicated situations such as blocking of vehicles, climatic changes, which reduces the performance. With such a variable background, it is hard to extract the background and also complex to detect and label the vacant spots and occupied spots. The main advantage is that the system works efficiently in all the weather conditions.

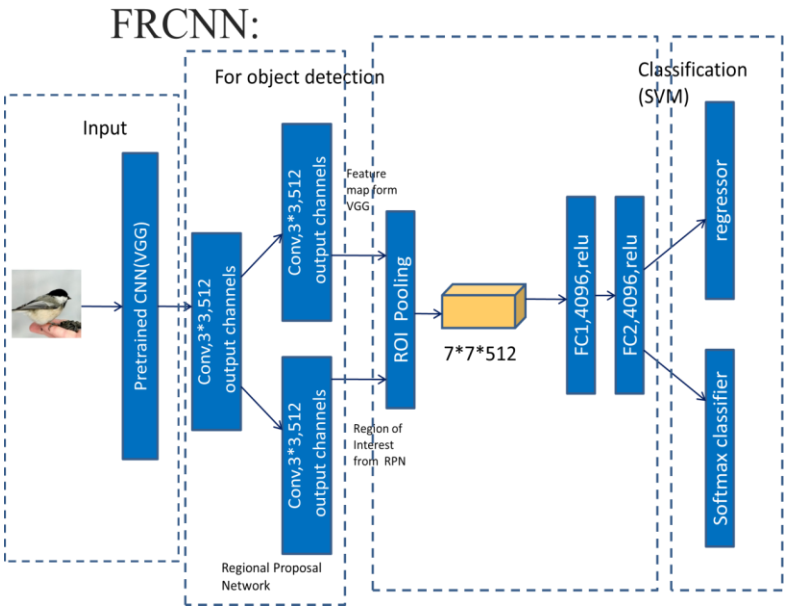

Fig. 1: Faster R-CNN

For object detection we can use sliding window approach as the entire image is divided into different sized boxes and locate object using brute-force method. But it requires more time and very complex. Hence, we use regional proposals. The Faster R-CNN model is used to identify objects in optimum time.

4. SYSTEM IMPLEMENTATION

The goal of our project, "AUTOMATIC VACANT PARKING PLACES MANAGEMENT SYSTEM," is to assist a person with parking. It identifies the objects present in the video which wegot from surveillance cameras and process the information. It deals with generating captions for a given image. The semantic meaning in the image is captured and converted into a natural language. The capturing mechanism involves a tedious task that collaborate both image processing and computer vision.The most accelerated technologies of this era are deep learning and machine learning. Artificial intelligence is now compared to the human brain, and it does great work than people in some fields. New research in this area occurs every day and this field is

growing very quicklybecause we now have enough computational power to do this. Multi-layered neural networks are used in deep learning, which is a branch of machine learning. Deep learning networks are often enhanced by increasing the amount of data used to train them.

A. Regional Proposals

The R-CNN system attempts to solve the challenge of finding objects in a picture (object detection). What is one to do in this situation? you may begin with a window approach. once exploitation this technique, you simply think again the entire image with totally different sized rectangles and appearance at those smaller pictures in a very brute-force-method. the matter is you may have an enormous variety of smaller pictures to seem at. To our luck, alternative sensible folks developed algorithms to neatly select those supposed region proposals. To change this concept:

Region proposals square measure simply smaller components of the first image, that we predict might contain the objects we have a tendency to square measure looking for. Region Proposals square measure simply smaller components of the first image, that we predict the item is gift. Fig. one represents the design of the quicker R-CNN. The image is provided as input to a CNN that provides feature map as output. There square measure differing types of CNNs like VGG, ResNet50 and additional. the item detection is performed within the RPN (Regional Proposal Network) and classified.

B. Region Proposal network

The Fast R-CNN model extends the R-CNN model. By learning a series of box offsets, a box regression layer is added to improve the position of the object in the image. To pool CNN features for each region suggestion, a ROI pooling layer is introduced to the network. The CNN feature map is fed into the RPN as an input. A convolutional network that performs two operations at the same time is known as an RPN. Objects that are present in the image are recognised, and a bounding box is formed around them. Anchor boxes are used to identify items in these networks. The image's detected objects are of various sizes.

In the next step, we have a tendency to take every region proposal and Using a Convolutional Neural Network, a feature vector representing this image in a much smaller dimension (CNN). Well, this is often one elementary issue with this R-CNN system. You can't train the entire system all at once (This are going to be solved by the quick R-CNN system). Rather, you may need to train each part multiple times. This indicates that AlexNet has already been trained on a classification task. They removed the last softmax layer during the coaching. The final layer is a 4096-dimensional one that is completely connected. this suggests that our options area unit 4096 dimensional.

Another necessary factor to stay in mind is that the input to the AlexNet is often an equivalent (227, 227, 3). The image proposals have totally different shapes tho'. several of them area unit smaller or larger than the desired size. As a result, we'll start sizing up each region proposal.

C. ROI Pooling layer

The RPN detects the items, although they are of different sizes. This layer is responsible for converting different sized objects which are detected into a particular size and shape. All the objects are reshaped into same shape so that the classification in the next stage can be performed.

D. SVM Classifier

The final step is to categorise the feature vectors. The categorization task is carried out based on the similarity of the objects. Each object's class is determined and classified. Hence the Support Vector Machine is used. Based on the classification the different colours of bounding boxes are drawn and provided as output. We produce feature vectors from the image proposals. currently we want to classify those feature vectors. we would like to observe what category of object those feature vectors represent. We have a tendency to apply AN SVM classification for this. we've one SVM for every object those categories and that we use all. this suggests that for one feature vector we've n outputs, wherever n is that the variety of various objects we would like to observe. The output could be a confidence score. however, assured are we have a tendency to that this

explicit feature vector represents this class? The Support Vector Machine, or SVM, is a well-known supervised learning technique that is used for both classification and regression problems. However, it is largely utilised in Machine Learning for Classification difficulties. The purpose of the SVM algorithmic rule is to create the most effective line or call boundary for categorising n-dimensional houses so that we can easily place additional data into the correct class in the future. A hyperplane is the name given to this best call boundary. SVM selects acute points/vectors that aid in the formation of the hyperplane. Support vectors are the extreme examples, which is why the algorithmic rule is called a Support Vector Machine. Consider the diagram below, which shows how a call boundary or hyperplane is used to classify two completely different classes.

SVM can be of two types: Linear SVM is a classifier for linearly separable data, which means a dataset can be classified into two classes using a single straight line, and the classifier is called Linear SVM.

Non-linear SVM: Non-linear SVM is a classifier that is used for non-linearly separated data. This means that if a dataset cannot be classified using a straight line, it is non-linear data, and the classifier used is termed Non-linearSVM.

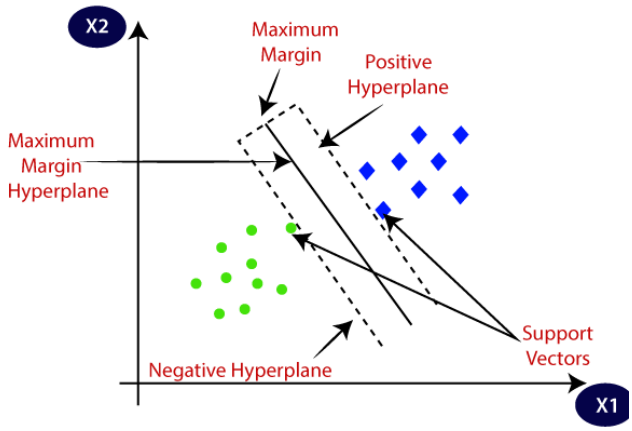

Fig 3: SVM 1Algorithm

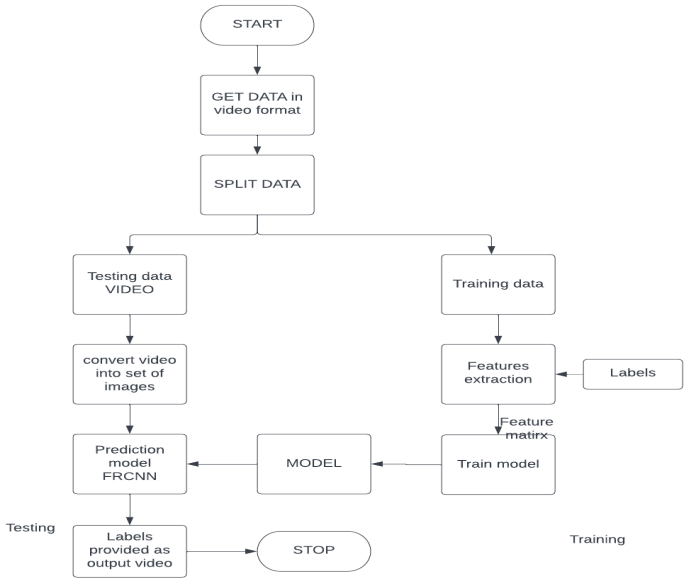

Fig. 3: Flow Chart

Fig. 2 represents the flow chart in which the model is developed at the time of training phase. The model has been efficiently trained to deliver accurate outputs. Later the video provided as input to model. The video is divided into pictures, with each image being processed separately. The objects are identified and a bounding box is drawn around the objects to indicate whether the parking spot is occupied or vacant. Bounding boxes come in a variety of colours to represent various things. Based on the colour, of the box, we can identify the free space and park vehicle. Our model also assists vehicles for not blocking in parking area.

3.1 System Design

The raw data that is processed to produce output is referred to as input in an information system. Developers must consider input devices such as PC, MICR, OMR, and others while designing input. As a result, the quality of the system's input dictates the system's output. The qualities of well-designed input forms and displays are as follows. It should effectively perform a specific purpose, such as storing, recording, and retrieving data.

The most important task for every system is to design output. Developers define the types of outputs that are required, as well as the necessary output controls and report layout prototypes, during output design. It should prioritise the user's attention, consistency, and ease of use. All of these goals are achieved by understanding basic design concepts relating to how end users react to various parts of forms and screens.

The most important task for every system is to design output. Developers define the types of outputs that are required, as well as the necessary output controls and report layout prototypes, during output design.

The live video from cameras which are present in the parking areas is taken as input to our model. Objects are identified using the Regional Proposal approach and then sorted into distinct groups. The input video is divided into images. We are going to use FRCNN (Faster Regional Convolution Neural Network) model to get the output effectively in less time. The objects in the image are identified with the help of Regional Proposal Network which is a part of FRCNN. The output of RPN is a feature vector which is provided as input to the ROI pooling layer to make all the identified objects into same size. Later with the help of SVM classifier, the cars are detected and provided as output to assist people for finding their parking slots.

The following are the goals of output design: To create an output style that fulfils the intended goal while avoiding the creation of undesired output. To create an output style that fits the needs of the most important users. To produce a sufficient amount of output. Form the result into a suitable format and deliver it to the right individual. To make the output out there on time for creating sensible selections. For producing region recommendations, the R-CNN and fast R-CNN models use the Selective Search rule, which takes a long time to identify an item. Every proposition is fed into a CNN that has been pre-trained to classify it. They planned a network called the region proposal network (RPN) in FRCNN to efficiently produce region proposals.

The region proposals square measure currently generated employing a network that would be trained and customized in step with the detection task. as a result of the proposals square measure generated employing a network, this may be trained end-to-end to be custom-built on the detection task. As a result, it generates more region ideas than generic tactics such as Selective Search and Edge Boxes. In the FR-CNN detection network, the RPN analyses the image victimisation similar convolutional layers. Thus, the RPN doesn't take longer to supply the proposals compared to alternative algorithms like Selective Search. Because the convolutional layers are shared, the RPN and hence the fast R-CNN can be integrated into a network. Thus, coaching is finished one time with efficiency. The output feature map from the convolutional layer shared with the FR-CNN is used by the RPN. An oblong window of size nxn is supported, and it travels through the feature map. for every window, many candidate's regions proposals square measure generated. These proposals aren't the ultimate proposals as they're going to be filtered supported their "objectness score" (explained below).

5. RESULTS AND ANALYSIS

The video from the surveillance cameras is fed into the system as input. The model processes the information and if the car is present in a spot, then a blue bounding box is drawn around it. If the space is vacant, then a green bounding box is drawn. The output is provided in a video format which contains the bounding boxes of different colours as shown below. We

can provide information to drivers about occupied and unoccupied places so that they move to vacant places without and delay in time. The data set is taken from the open source which is PK lot dataset which contains more than 12000 images of parking lots from the kaggle platform. 8000 images are used for training.

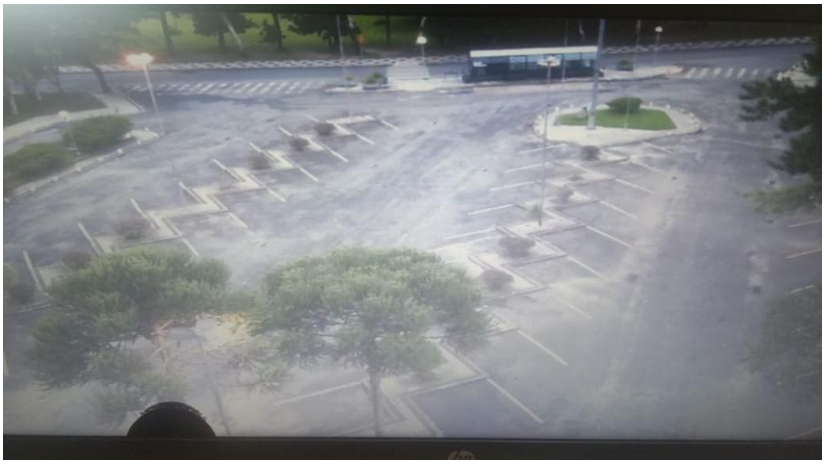

Fig. 3: The empty parking slot as input

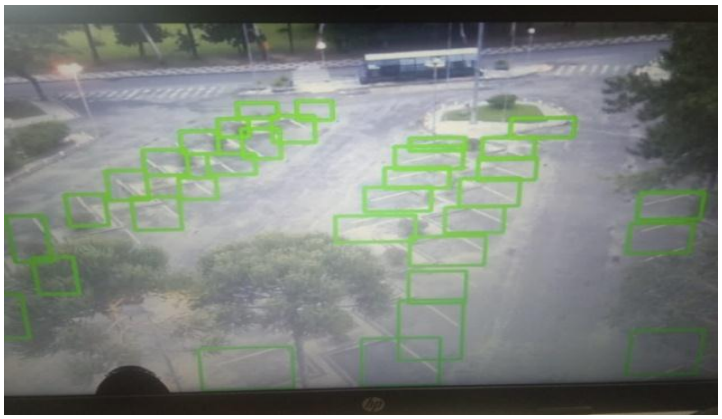

Fig. 4: Bounding boxes drawn in empty places as output

As the entire parking slot is empty, the bounding box of green colour is drawn and provided as output as shown in Fig. 4. If there are any vehicles present in the parking lots, then the blue colour boxes are drawn around the vehicles.

Fig. 5: Parking lot containing vehicles

In the Fig. 5 there are many cars parked and some of the places were vacant. If we use this image as an input, the result is as seen in Fig. 6

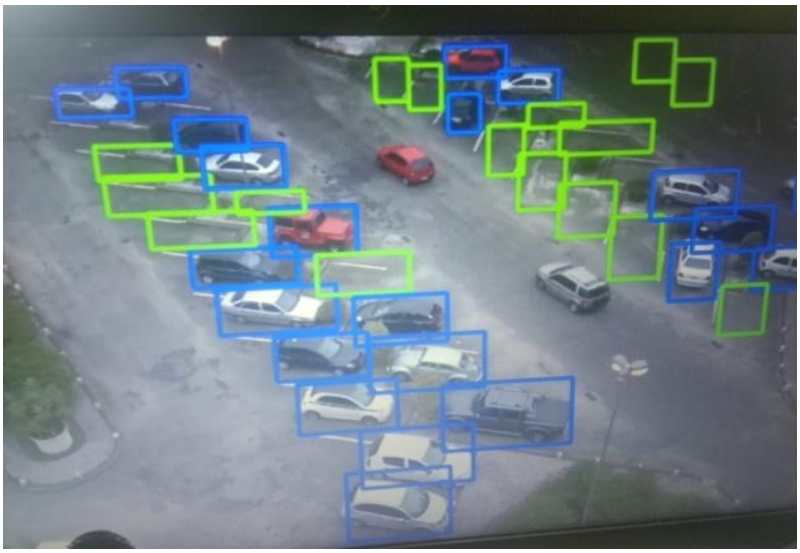

Fig. 6: Output of different colours bounding boxes

6. CONCLUSION AND FUTURE WORK

A system for managing vacant parking spaces using vehicle detection and the regional proposal approach is presented in this study. More complex conditions, such as climate changes (cloudy, rain, snow, and more) and car obstruction, are faced by the system. The management of cars in parking lots was made more efficient and cost-effective by using this strategy.

Based of further investigation on the object detection techniques in future, the solution is much optimised. It is also essential to select images from different parking lots datasets and thus helps us to enhance accuracy.

REFERENCES:

[1] Xie, H., Wu, Q., Chen, B., Chen, Y., & Hong, S. (2015, August). Vehicle detection in open parks using a convolutional neural network. In 2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA) (pp. 927-930). IEEE.

[2] Guo, Y., & Shi, H. (2021). Automatic parking system based on improved neural network algorithm and intelligent image analysis. Computational Intelligence and Neuroscience, 2021.

[3] Masmoudi, I., Wali, A., Jamoussi, A., & Alimi, A. M. (2014, January). Vision based system for vacant parking lot detection: Vpld. In 2014 International Conference on Computer Vision Theory and Applications (VISAPP) (Vol. 2, pp. 526-533). IEEE

[4] Al-Kharusi, H., & Al-Bahadly, I. (2014). Intelligent parking management system based on image processing. World Journal of Engineering and Technology, 2014.

[5] Fabian, T. (2008, June). An algorithm for parking lot occupation detection. In 2008 7th Computer Information Systems and Industrial Management Applications (pp. 165-170). IEEE.

[6] Liu, J., Mohandes, M., & Deriche, M. (2013, December). A multi-classifier image based vacant parking detection system. In 2013 IEEE 20th International Conference on Electronics, Circuits, and Systems (ICECS) (pp. 933-936). IEEE.

[7] Huang, C. C., Tai, Y. S., & Wang, S. J. (2013). Vacant parking space detection based on plane-based Bayesian hierarchical framework. IEEE Transactions on Circuits and Systems for Video Technology, 23(9), 1598-1610.

[8] Blumer, K., Halaseh, H. R., Ahsan, M. U., Dong, H., & Mavridis, N. (2012, November). Cost-effective single-camera multi-car parking monitoring and vacancy detection towards real-world parking statistics and real-time reporting. In International Conference on Neural Information Processing (pp. 506-515). Springer, Berlin, Heidelberg.

[9] Huang, C. C., Dai, Y. S., & Wang, S. J. (2012, November). A surface-based vacant space detection for an intelligent parking lot. In 2012 12th International Conference on ITS Telecommunications (pp. 284-288). IEEE.

[10] Chen, L. C., Hsieh, J. W., Lai, W. R., Wu, C. X., & Chen, S. Y. (2010, October). Vision-based vehicle surveillance and parking lot management using multiple cameras. In 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (pp. 631-634). IEEE.