DENSITY BASED TRAFFIC CONTROLLER WITH CAMERAS

S.V. Lakshmi Kumari Assistant Professor, Department of ECE, Narayana Engineering College (Autonomous), Gudur, AP,524101
M. Sruthi Leela, K. Mounika, M. Jayarami Reddy, G. Koteswar UG Student, Department of ECE, Narayana Engineering College(Autonomous), Gudur, AP,524101.
Email I'd:[1]svlkumari.57@gmail.com,[2]sruthi15091999@gmail.com

Abstract : *Nowadays traffic congestion is a severe burning problem so this paper came with a proposal to introduce modified methods in controlling the traffic, mainly using modern technology (RCNN detector and SVM classifier) and also using signal lights (Red, Green, Yellow) through image processing. It is a system to control the traffic by measuring the real-time vehicle density using a convolution neural network with digital image processing is proposed. Besides that, the complete technique from image acquisition to the number of vehicles detection and finally green signal allotment using four sample images of different traffic conditions are illustrated with proper schematics and the final results are verified.*

Keywords— Image processing, Image acquisition, Convolution Neural Network, RCNN detector, SVM classifier

1. INTRODUCTION

Traffic congestion is when vehicles travel slower because there's an excessive amount of traffic on roads. This makes trip times longer, and increases queuing. This is also known as a traffic jam. Congestion may result from a decrease in capacity, for instance, accidents on the road or roads being closed. Bad road layouts can also restrict capacity. Increased traffic, for instance by many cars leaving a stadium at an equivalent time, also can cause congestion.

The first electric traffic light was developed by Lester Wire in 1912. The T-shaped design for traffic signals was invented by Garrett Morgan in 1923. The first fully electric traffic signal system was installed on 5 th August 1914.

Sitting in traffic, especially in a busy city can feel like a twisted version of the kids game "red light, green light ".While many traffic signals still run on timers that might cause delays in the future of traffic control at intersections.

Intermetropolitan area studies suggest that traffic congestion reduces regional competitiveness and redistributes economic activity by slowing growth in county gross output or slowing metropolitan area employment growth [1].

A. IMAGE PROCESSING

Image may be a two-dimensional candlepower function f (x, y), where x and y are spatial coordinates and therefore the amplitude f at any pair of coordinates (x, y) is named the intensity or gray level. When x, y, and f are discrete quantities the image is digital. 'f' can be a vector and represent a color image, e.g., using the RGB model, or in general a multispectral image. A video signal is similarly expressed as a sequence of frames f (x, y, and t). The gray-level image will be represented by a matrix with 8-bit integer values, in the range [0=black, 255=white]. Digital Image Processing concerns the transformation of a picture to a digital format and its processing by a computer or by dedicated hardware. Digital image processing allows one to enhance image features of interest while attenuating details irrelevant to a given application, and then extract useful information about the scene from the enhanced image. This introduction may be a practical guide to the challenges, and therefore the hardware and algorithms wont to meet them. Images are produced by a variety of physical devices, including still and video cameras, x-ray devices, electron microscope, radar, and ultrasound, and see for a variety of purposes, including entertainment, medical, business (e.g. documents), industrial, military, civil (e.g. traffic), security and scientific. The goal in each case is for an observer, human or machine, to extract useful information from the scene being imaged. The raw image is not directly suitable for extracting the information conveyed by that image and must be processed in some way. Such processing is called Image Enhancement;

processing by an observer to extract information is called Image Analysis. Enhancement and analysis are distinguished by their output, images versus scene information, and by the challenges faced and methods employed. Image enhancement has been done by chemical, optical, and electronic means, while analysis has been done mostly by humans and electronically. Digital image processing may be a subset of the electronic domain wherein the image is converted to an array of small integers, called pixels, representing a physical quantity like scene radiance, stored during a digital memory, and processed by computer or other digital hardware. Digital image processing, either as an enhancement for human observers or performing autonomous analysis, offers advantages in cost, speed, and adaptability, and with the rapidly falling price and rising performance of personal computers, it's become the dominant method in use.

2. EXISTING METHOD

In the existing method, we use the edge detection technique. The edge detection technique is imperative to extract the specified traffic information from the CCTV footage. It is often wont to isolate the specified information from the remainder of the image. There are several edge detection techniques available. They have distinct characteristics in terms of noise reduction, accuracy, etc. Among them, Prewitt [2], canny [3], Sobel [3], Roberts, and LOG are the most accredited operators. Canny edge detector depicts higher accuracy in detection of an object, PSNR (Peak Signal to Noise Ratio), MSE(Mean Square Error), and execution time compared with Sobel, Roberts, Prewitt, Zero crossing, and LOG [4-6]. In this paper, a system during which density of traffic is measured by comparing the captured image with real-time traffic information against the image of the empty road as a reference image is proposed.

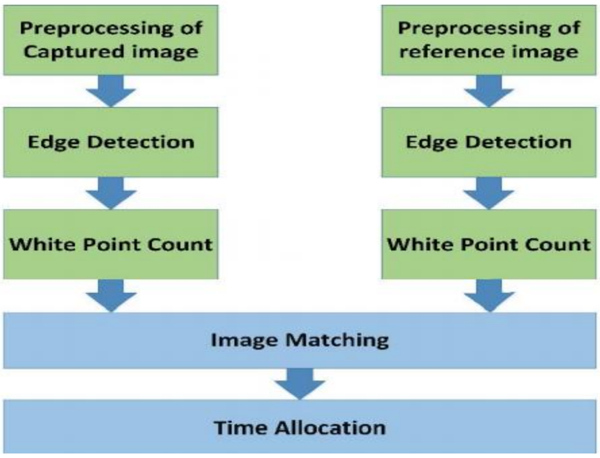

Fig. 1. Block Diagram of Existing Method

The matching is achieved by comparing the number of white points between two images. The entire image processing before edge detection i.e. image acquisition, image resizing, RGB to gray conversion is explained. Canny edge detection operations are depicted. The canny edge detector operator is chosen due to its greater overall performance. Percentage matching for various sample images and traffic time allocation for them are demonstrated. The content of this paper completely serves the aim of demonstrating the restrictions of current control techniques and therefore the solution of these limitations with detailed explanation. Image matching by comparing detected edges is an approach to identify the vehicle density with propitious accuracy. As far as we all know, matching images by comparing detected edges has not been used before for smart control applications.

3. PROPOSED METHOD

The proposed method consists of convolution neural networks which can be used to detect vehicle detection. To bypass the matter of choosing an enormous number of regions, Ross Girshick et al. proposed a way where we use selective search to extract just 2000 regions from the image and they are called region proposals. Therefore, now, rather than trying to classify an enormous number of regions, you'll just work with 2000 regions. These 2000 region proposals are generated using a selective search algorithm.

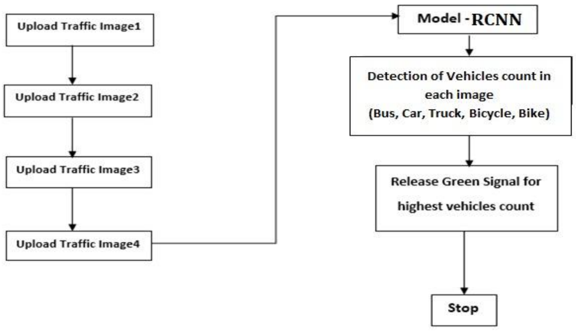

Fig. 2. Block Diagram of Proposed Method

To know more about the selective search algorithm. These 2000 candidate region proposals are warped as a square and fed into a convolutional neural network which gives a 4096-dimensional vector as output. The CNN acts as a feature extractor and the output dense layer consists of the features extracted from the image and the extracted features are fed into an SVM classifier to classify the presence of the thing within that candidate region proposal. In addition to predicting the presence of an object within the region proposals, the algorithm also predicts four values which are offset values to increase the precision of the bounding box. For example, given a neighborhood proposal, the algorithm would have predicted the presence of an individual but the face of that person within that region proposal could've been cut in half. Therefore, the offset value helps in adjusting the bounding box of the region proposal.

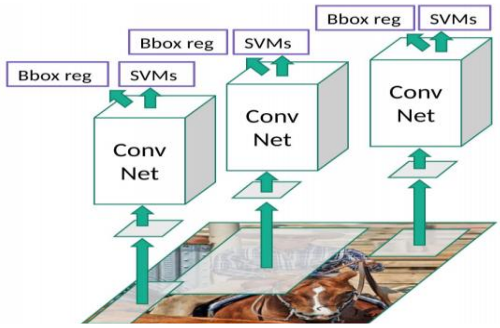

Fig. 3. Architecture for RCNN

When the paper "Rich feature hierarchies for accurate object detection and semantic segmentation" came out of UC Berkely in 2014 nobody could have predicted its impact. After 5 years it now has nearly 9000 citations. In this paper, the authors introduced a fundamental concept for all modern object detection networks: Combining region proposals with CNN's. They called this method R-CNN.

This will be the first entry in a 3-part series covering R-CNN, Fast R-CNN, and Faster R-CNN. You will get to fully understand the intuition behind the concepts here to know the subsequent articles. Remember, knowing the basics well is far more important than "half-knowing" modern state-of-the-art approaches.

A. R-CNN System:

The problem the R-CNN system tries to unravel it's to locate objects in a picture (object detection). What do you do to solve this? You could start with a sliding window approach. When using this method, you only re-evaluate the entire image with different sized rectangles and appearance at those smaller images during a brute-force method. The problem is you'll have an enormous number of smaller images to seem at. To our luck, other smart people developed algorithms to smartly choose those so-called region proposals.

1) Region proposals

There are different region proposal algorithms we will choose between. These are "normal" algorithms that employment out of the box. We don't have to train them or anything. You can choose any method you wish and it might work either way. This will create nearly 2000 different regions we'll need to check out. This seems like an enormous number, but it's still very small compared to the brute-force window approach.

2) CNN

In the next step, we take each region proposal, and that we will create a feature vector representing this image during a much smaller dimension employing a Convolutional Neural Network (CNN).

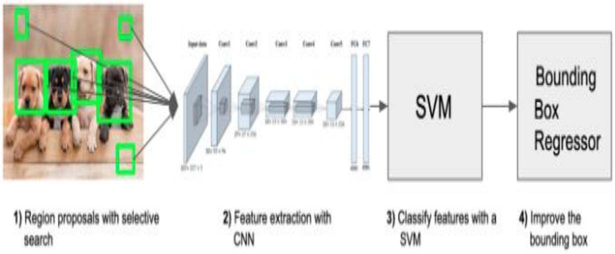

Fig. 4. The RCNN System

They use the AlexNet as a feature extractor. Don't forget it's 2014 and AlexNet remains quite state-of-the-art. Well, this is one of the fundamental issues with the R-CNN system. You can't train the entire system in one go (This is going to be solved by the fast R-CNN system). Rather, you'll get to train every part independently. The AlexNet was trained before the classification task. After the training, the last softmax layer was removed which is the fully connected 4096-dimensional one. This means that our features are 4096 dimensional. Another important thing to stay in mind is that the input to the AlexNet is usually equivalent (227, 227, 3). The image proposals have different shapes though. Many of them are smaller or many of them are larger than the specified size. So we'll get to resize every region proposals.

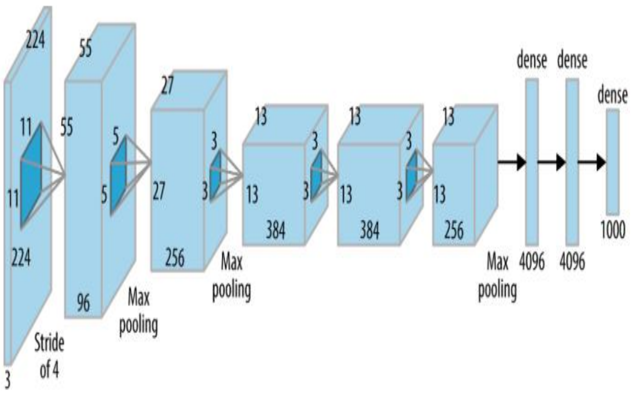

Fig. 5. Region Proposals

A. SVM

We created feature vectors from the image proposals. Now we'll get to classify those feature vectors. We want to identify what class of object those feature vectors represent. For this, we use an SVM classificatory. We have one SVM for every object class and that we use all of them. This means that for one feature vector we've n outputs, where n is that the number of various objects we would like to detect. The output is a confidence score.The thing that confused me once I read this paper for the primary time was how we trained those different SVM's. Well, we train them using feature vectors created by the AlexNet. That means, that we've to attend until we fully trained the CNN before we will train the SVM. The training is not parallelizable. Because we all know when training what feature vector represented which class we will easily train the various SVM's during a supervised-learning way.

1) Bounding Box Regressor

I want to say the Bounding Box Regressor at the top because it's not a fundamental building block of the R-CNN System. It's a great idea though and the authors found that it improves the average precision by 3%. When you are training the Bounding Box Regressor your input is that the center, width, and height in pixels of the region proposal and therefore the label is that the ground truth bounding box.

B. Layers of CNN:

1) Image Input Layer:

Create an image input layer using the image input layer. An image input layer inputs images to a network and applies data normalization. Specify the image size using the input Size argument. The size of a picture corresponds to the peak, width, and therefore the number of color channels of that image. For example, for a grayscale image, the amount of channels is 1, and for a color image, it's 3.

2) Convolutional layer:

A two dimensional convolutional layer applies to slide convolutional filters to the input. Create a 2-D convolutional layer using convolution2dLayer. The convolutional layer consists of various components. Filters and Stride a convolutional layer consists of neurons that connect to sub-regions of the input images or the outputs of the previous layer. The layer learns the features localized by these regions while scanning through a picture. When creating a layer using the convolution2dLayer function, you can specify the size of these regions using the filter Size input argument.

3) Batch normalization layer:

Batch normalization has the benefits of helping to make a network output more stable predictions, reduce overfitting through regularization, and speeds up training by an order of magnitude. Batch normalization is the process of carrying normalization within the scope activation layer of the current batch, subtracting the mean of the batch's activations and dividing by the standard deviation of the batch's activations. Create a batch normalization layer using the batch Normalization Layer. A batch normalization layer normalizes each input channel across a mini-batch. To speed up the training of convolutional neural networks and reduce the sensitivity to network initialization, use batch normalization layers between convolutional layers and nonlinearities, like ReLU layers.

The layer first normalizes the activations of every channel by subtracting the mini-batch mean and dividing by the mini-batch variance. Then, the layer shifts the input by a learnable offset β and scales it by a learnable multiplier γ. β and γ are the learnable parameters which updated during network training. Batch normalization layers normalize the activations and gradients propagating through a neural network, making network training a neater optimization problem. To take full advantage of this fact, you'll try increasing the training rate. Since the optimization problem is simpler, the parameter updates are often larger and therefore the network can learn faster. You can also reduce the L2 and dropout regularization. With batch normalization layers, the activations of a selected image during training depend upon which images happen to seem within the same mini-batch. To take full advantage of this regularizing effect, we can shuffle the training data before every training epoch. To specify how often to shuffle the data during training, use the 'Shuffle' name-value pair argument of training Options.

4) Max and Average Pooling Layers:

A max-pooling layer performs down-sampling by dividing the input into rectangular pooling regions and computing the utmost of every region. Create a max-pooling layer using maxPooling2dLayer. An average pooling layer performs down-sampling by dividing the input into rectangular pooling regions and computing the typical values of every region. Create an average pooling layer using averagePooling2dLayer. Pooling layers follow the convolutional layers for down-sampling, hence, reducing the number of connections to the subsequent layers. They do not perform any learning themselves but reduce the number of parameters to be learned within the following layers. They also help reduce overfitting.

A max-pooling layer returns the utmost values of rectangular regions of its input. the dimensions of the oblong regions are decided by the pool Size argument of the max Pooling Layer. for instance, if pool Size equals [2, 3], then the layer returns the utmost value in regions of height 2 and width 3. a mean pooling layer outputs the typical values of rectangular regions of its input. the dimensions of the oblong regions are decided by the pool Size argument of the typical Pooling Layer. for instance, if pool Size is [2, 3], then the layer returns the typical value of regions of height 2 and width 3.Pooling layers scan through the input horizontally and vertically in step sizes you'll specify using the 'Stride' name-value pair argument. If the pool size is smaller than or adequate to the stride, then the pooling regions don't overlap. For non-overlapping regions (Pool Size and Stride are equal), if the input to the pooling layer is n-by-n, and therefore the pooling region size is h-by-h, then the pooling layer down-samples the regions by h [6]. That is, the output of the max or average pooling layer for one channel of a convolutional layer is n/h-by-n/h. For overlapping regions, the output of a pooling layer is (Input Size – Pool Size + 2*Padding)/Stride + 1.

4. RESULTS

The input image is taken from the high resolution cameras which are placed at the junction. The RCNN detector is used such that it detects the image. The RCNN is a combination of Convolution Neural Network and SVM. The SVM classifier is used to classify the image and classifies the foreground images and the background images.

Fig. 6. Input Image

By using the CNN the bounding boxes are applied to the classified image and the accurate readings are taken. An empty RCNN network was created such that it was applied to all the images and detects the vehicles in the lane. Using the object annotation the bounding boxes are applied to the images .

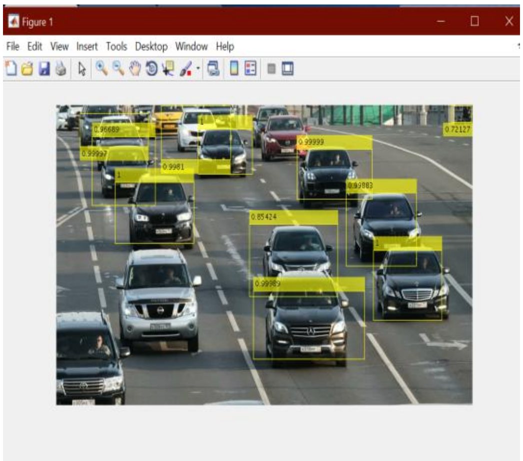

Fig. 7. Bounding Box Implementation

The final results are verified and are displayed in the command window in MATLAB.

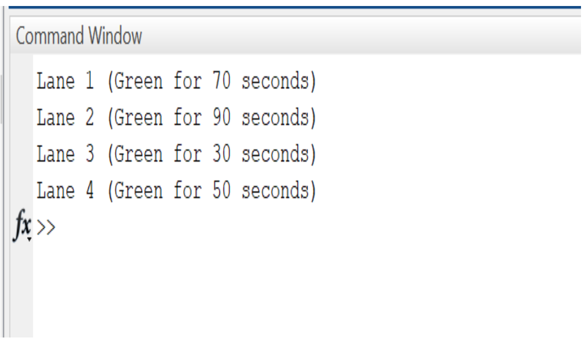

Fig. 8. Output

5. CONCLUSION

In this paper, a traffic control system with image processing as an instrument for measuring the density has been proposed. For this purpose, four sample images of different traffic scenarios have been attained. Upon completion of the RCNN detection then images are bounded by the bounding boxes for calculation of vehicles. Using this time allocation is carried out for each image. The time allocation has been illustrated for the four samples using the MATLAB programming language and the final results are verified.

ACKNOWLEDGEMENT

We would wish to thank our head of the department Dr. P. Sreenivasulu and my guide Ms. S.V. Lakshmi Kumari for motivating us in doing such kind of real-time projects and we wish to thank our friends who helped directly and indirectly in the completion of our project.

REFERENCES

1. M. Sweet, “Traffic Congestion’s Economic Impacts: Evidence from US Metropolitan Regions,” Urban Studies, vol. 51, no. 10, pp. 2088–2110, Oct. 2013.
2. Weibin Rong, Zhanjing Li and Wei Zhang, “An improved canny edge detection algorithm,”in IEEE International Conference on Mechatronics and Automation, August 2014.
3. Wenshuo Gao, Xiaoguang Zhang and Lei Yang, “An improved Sobel edge detection,” in3rd International Conference on Computer Science and Information Technology, July 2010.
4. G.T. Shrivakshan and Dr.C. Chandrasekar, “A Comparison of various Edge Detection Techniques used in Image Processing,” International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012.
5. Sarankumar, R., Sumathi, K., Sravanthi, G. S., Venkateswarlu, B., & Rajasekar, P. (2021). Reducing Complexity in PTS Scheme using Optimization Techniques to reduce PAPR in OFDM Systems. IOP Conference Series: Materials Science and Engineering, 1074(1), 012032. https://doi.org/10.1088/1757-899x/1074/1/012032
6. Pinaki Pratim Acharjya, Ritaban Das and Dibyendu Ghoshal, “Study and Comparison of Different Edge Detectors for Image Segmentation,” Global Journal of Computer Science and Technology Graphics& Vision, Volume 12 Issue 13 Version 1.0 Year 2012.
7. M. Venkatesan, C. S. Subash Kumar, M. Sathiyanathan (2020) A Step by Step Design of PV System: A Research Approach for Engineering Graduates (Vol. 16. No-1,), 30-35
8. Jiaqiao He, Liya Hou and Weiyi Zhang, “A Kind of Fast Image Edge Detection Algorithm Based on Dynamic Threshold Value,” Sensors & Transducers, Vol. 24, Special Issue, August 2013, pp. 179.
9. Penchalaiah P, Rajasekar P et al, "An Efficient Multi-User Hierarchical Authenticated Encryption Using Simultaneous Congruence for Highly Secure Data", International Journal of Future Generation Communication and Networking (WoS), ISSN: 2233-7857(Print); 2207-9645(Online), NADIA, (2020), Vol. 13, No. 2, pp. 1-10.
10. A.M.Mahaboob Basha, M.Rajaiah, Penchalaiah P et al., “Machine Learning-Structural Equation Modeling Algorithm: The Moderating role of Loyalty on Customer Retention towards Online Shopping”, International Journal of Emerging Trends in Engineering Research, ISSN: 23473983 (SCOPUS), Volume-8, No. 5, Page 1578–1585,May, 2020.