I.INTRODUCTION

Object Detection System is related with Intelligent visual monitoring which is able to perform automatic video analysis. Intelligent visual monitoring would be very helpful when implemented in security systems. The Intelligent visual monitoring can be implemented to monitor ordinary public places or can also be used in monitoring more complex areas. This system ignores useless information and keeps it’s focus towards the necessary data for better processing. Intelligent visual monitoring can detect, classify, recognize objects without the involvement of humans. This technology when implemented in Security systems can be used to trigger an alarm in order to alert the security personnel to take the necessary action before any great damage is caused. Intelligent visual monitoring system can be used for continuous real-time monitoring and to respond to abnormal targets or behaviors accurately, which is very essential for places that requires high security. With a good network and processing capability, this system can perform high speed analysis on a continuous stream of video data.

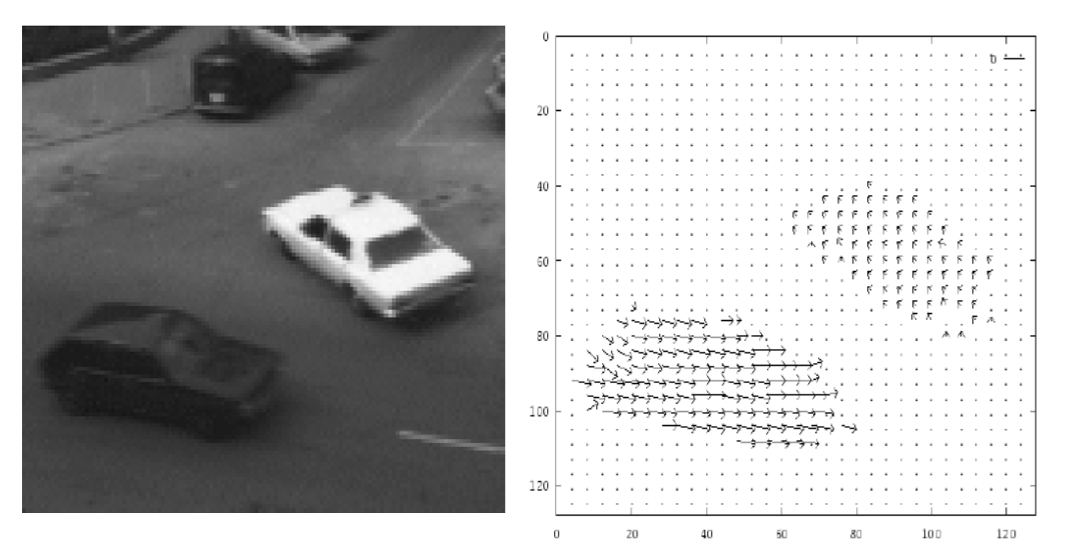
1.BASIC CONCEPT

Day by day, thefts and attacks on people are being increased rapidly in public places. People are feeling more insecure to be present in public places like banks, malls, etc. So, in order to avoid much damage or prevent thefts and attacks on people in public places and instead of taking action on criminals after any damage, preventing large scale criminal activity in public places before it even happens will be a huge step in ensuring safety of the people. So, an Intelligent visual monitoring system using object detection can be implemented in surveillance cameras for better monitoring of public places. Through this system, abnormal or illegal objects and illegal behaviors based on movement speed and other tracking details of a person can be detected and can be further analyzed. An alarm can be triggered if the system detects any unusual objects or behaviors which allows the security personnel to take effective measures in time before any huge damage occurs. Therefore this paper proposes two moving target tracking methods which are widely used in the intelligent visual monitoring systems.

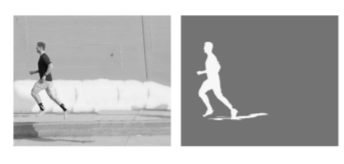
2.DESIGN

**Optical flow vector method**

After recognizing the moving objects in a given video sequence, the Hue Saturation Value Space Shadow detection is applied to remove the influence the shadow part.



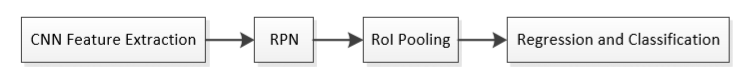
**Fig.no 1**

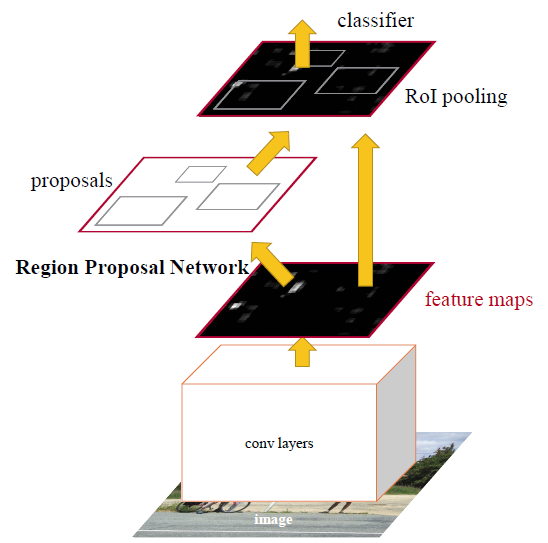
**RCNN method**

* **The R-CNN object detection algorithm usually contains three parts:**

1. **Features extraction**
2. **Choice of detection window**
3. **Design of classifier**

****

**The process of RCNN method:**

****

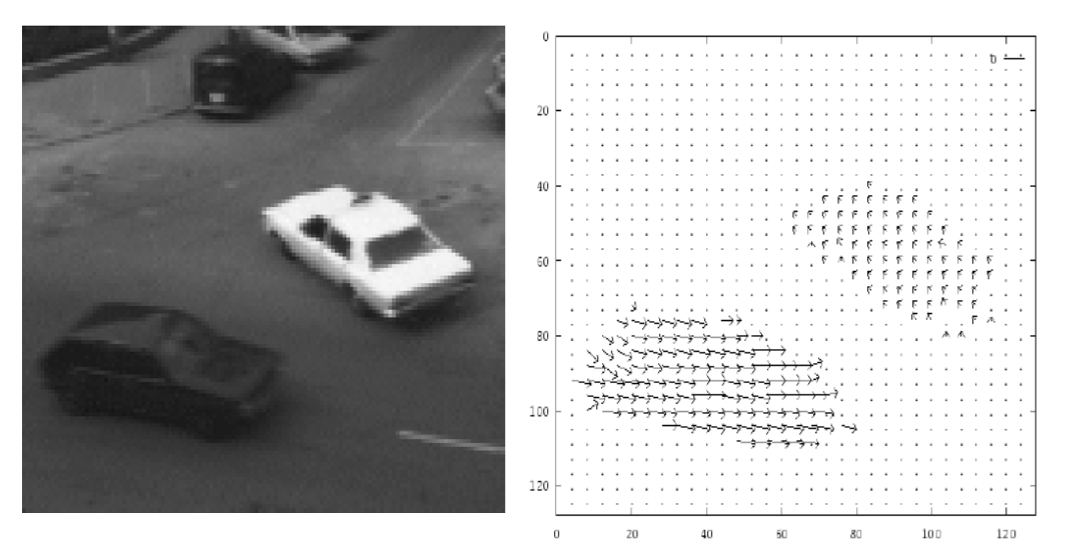
**Fig no 3**

3.Methodology

Two widely used methodologies, Optical flow vector with HSV Space shadow detection and RCNN for intelligent visual monitoring have been included in this paper.

**Optical flow vector method:**

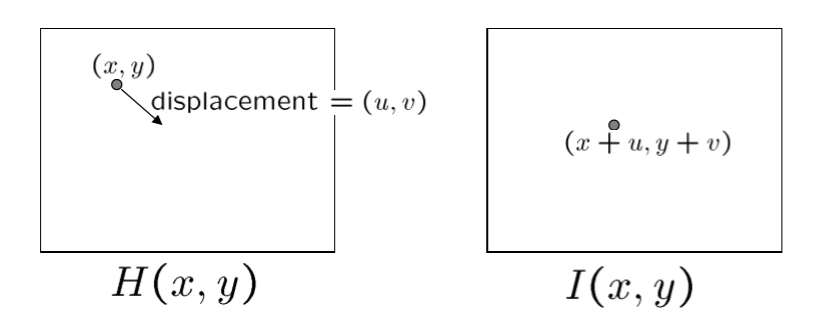
One of the main objectives of Optical flow method is to identify moving targets in given video image sequence. The optical flow method checks for changes in the gray scale values of the pixels of the image. The image can be tracked by identifying the change in the gray scale values of the pixels. The method gives the gray scale change trend of each pixel of the image. The above picture shows the change in gray scale values of the pixels, which are indicated by optical flow vectors. By that representation, the moving objects in a video sequence can be differentiated from the static background.



The above image represents the change in the gray scale values of the pixels as the object which is a car moves in the given video sequence. The movement is represented by the optical flow vectors which represents the change in the gray scale values.

By that representation the moving objects from a static background can be differentiated.

The below picture represents the change in the grayscale value of a pixel.

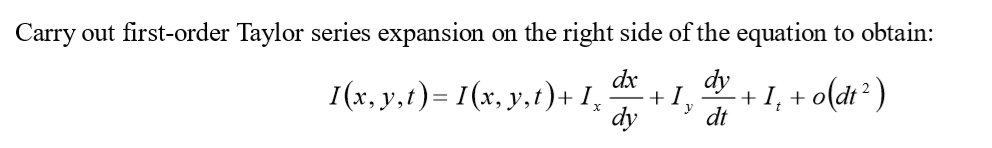


The grayscale values of a particular pixel are represented by H(x,y) and the new gray scale values can be obtained as the object changes it’s position with displacement (u,v) as I(x,y) where the new gray scale values are (x+u,y+v).

For Suppose the grayscale value of the pixel of image (x,y) is I(x,y,t) at time t , then the new grayscale values after any tiny movement can be given by,

Ix,Iy,It are the partial derivatives of the grayscale value of the pixel in the x, y and t directions, and according to the brightness constant of its tiny movement in a very short time, we can get:

**I (x, y, t) = I (x + dx, y + dy, t + dt)**



The generalized equation can be given by:

**Ix U + Iy V + It = 0**

**where u and v are components of optical flow vector.**

The optical flow vectors u and v can be calculated. Moving target tracking,



The optical flow method can directly calculate the motion information of any pixel in the image of video sequence. Without knowing any kind of information of the scene, the moving target can be detected easily. But the optical flow method for moving target detection algorithm is complex, computational and anti-build performance is poor, it is not easy to be real-time and high precision monitoring system in the use of morphological processing (technique for modifying the pixels of an image).

HSV Space Shadow Detection

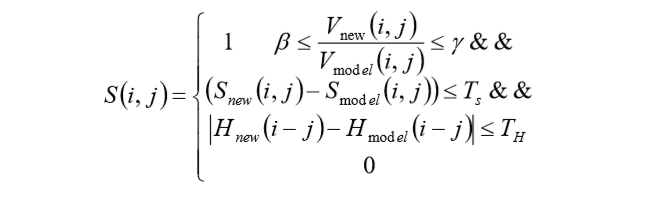
After the moving object is detected from the region, the pixels from the detected region can again be divided into two parts:

1. Original Part
2. Shadow Part

The shadow is formed by the occlusion of the object to the light source in the scene. Shadow part plays a vital role in processing of the detected motion region. The shadow part will have a bad influence on the processing of the detected motion region, so, the shadow part needs to be eliminated from the detected motion region. HSV space is directly corresponding to the brightness, color and saturation of the three elements of human eye color visual characteristics, and the three components are independent of each other.

Hue, the difference between the appearances of a color, used to share dividends, orange, yellow, green, blue and other color characteristics, not affected by the color of light, light and shade. (Lightness, Brightness, Lightness Value) - Color shade of gray degree, given all the same properties. The degree of Saturation (Saturation) - colors should be bright pale, namely the same hue of the color in the mixed proportion of white, mixed unless otherwise stipulated in the color white, we generally referred to as the pure color, the more the lower the proportion of white color is bright, on the other hand, will become pale. In HSV color space model, the actual visual experience conclusion, we can come to the conclusion that under the condition of certain Brightness, the same object in the shadow area and not in the shadow area Hue (Hue) is approximate consistent, the shadow in the region mainly make change colors of light and shade, namely

Let H new, S new and V new represent the three components of chromaticity(H), color (S) and gray scale (V) of a pixel point in the moving region to be detected, and H model, S model and V model represent the three components of chromaticity, color and gray scale of the corresponding background pixel point in the region. The specific algorithm is as follows:



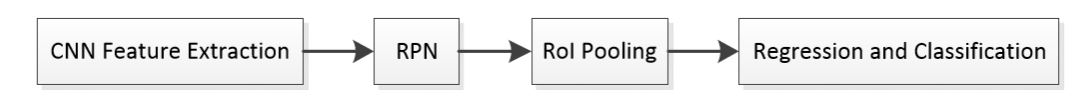
Where, T s and T h respectively represent the threshold value of color and chromaticity component. Since the V value of the shaded point is always less than the V value of the corresponding point that is not shaded, Gamma is less than 1, while beta takes into account the current light intensity. Generally, the stronger the light is, the smaller the beta value wills be. For S, the shadow usually has a relatively low value, and the difference between the shadow and the background model is often negative. Finally, considering that the difference between the background and the shadow is relatively small, the difference of H value is considered.

**RCNN Method**

R-CNN method can also be used for object detection. The detection algorithm contains design of features (derived values from built data), detection window, design of classifier, CNN, RPN (Region Proposal Networks), Soft max Classification Processing.

* The R-CNN object detection algorithm usually contains three parts:

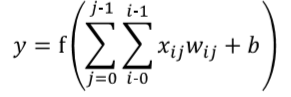
1. Features extraction
2. Choice of detection window
3. Design of classifier



CNN Feature Extraction:

The image feature extraction is done by using filters (convolution kernels) in RCNN method. Different types of filters extract different features. The activation function that is generally used is ReLU (rectified linear unit) which is most commonly used activation function in neural networks.

The function is given by:



where y is the output

f represents the activation function

w represents parameters of filter

x is input vector

b is bias term (used to increase capacity of the network)

RPN:

This stage takes the output as the feature map generated by CNN extraction and fixes some windows in the original design. Using this window scale, the RPN will be able to learn where are the objects inside the window. Only the approximate place needs to be found, because that the precise positioning position and size can be accomplished by following works. The anchors can be fixed in three aspects: fixed scale changes (three scales), fixed length and width ratio changes (three ratio), fixed sampling method.

After getting the proposal on feature map, convolution calculation can be shared in front of the network. The result of this network is that each point of the convolution layer has an output about the k anchor boxes, including whether it is an object or not, adjusting the corresponding position of the box. The RPN's overall Loss function can be defined as:



where the i denotes the i-th anchor, pi\* = 1 when the anchor is positive, and pi\*= 0 when the anchor is negative. Ti\* represents a ground true box coordinate associated with the positive sample anchor.

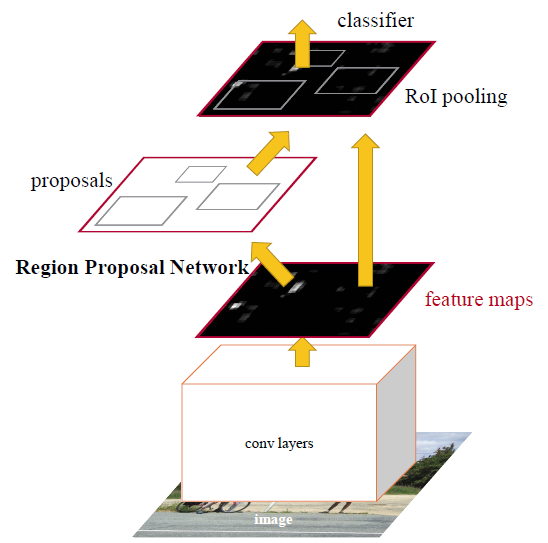
ROI Pooling:

For normal CNN, when the network is trained, the size of input must be a fixed value. So, ROI Pooling is introduced to solve this. By that, the training time of the model is reduced to a great extent.

Classification and regression:

This is the final stage in this method in which the final probability vector is calculated based on the regional proposal feature map. Bounding box regression is used to obtain the accurate position of the target. Bounding-box regression is a popular technique to refine the prediction vector in recent object detection approaches.

The below image represents the different processes involved:



4.RESULTS

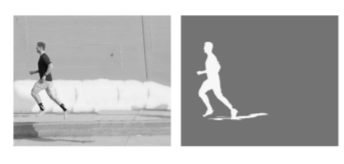
**Optical flow vector method:**

After the detection of the motion region, HSV Space shadow detection is applied on the sequence of frames to remove the influence of the shadow. After applying the shadow detection, the center of mass of the moving body will fall on the body itself thereby helping in better estimate of the trajectory.

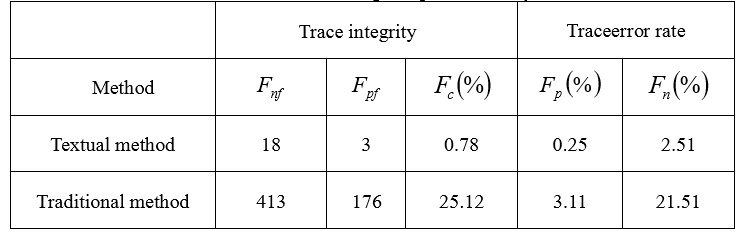
Detected motion region before applying the HSV Space shadow detection,



Detected motion region after applying the HSV Space shadow region, as we can see from the below image, the center of mass of the body will fall on the body itself and the influence of the shadow to estimate the trajectory of the target is removed which leads to better processing and results.



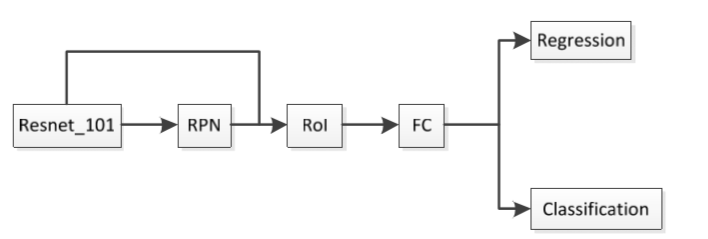
Video tracking performance is analyzed. The tracking method is compared with the tracking performance index and tracking time on the input video. Video used in this paper has a total of 2413 frames, and the number of frames covered by the target is 300 frames. It can be seen from the below that the performance of the tracking method proposed in this paper is greatly improved compared with the traditional method.

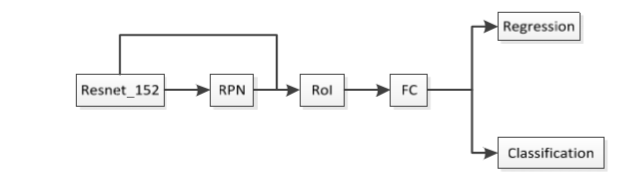


We can see from the above table that performance is better after applying the HSV Space shadow detection.

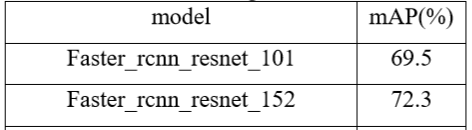
**RCNN method:**

The object detection data set is trained with different object detection models by using different types of classification, regression and pooling techniques, and the results of object detection under different models are obtained. The structure is shown in the following figures:





Here RPN represents the area to generate the network, CONV represents convolution layer, FC represents the full connection layer, and Position sensitive RoI represents the pooling operation.



The below table represents the mean average precision percentages of different models of their performance when the same data set is used. The performance can be increased by repeated experiments by fine tuning the parameters.

5.CONCLUSIONS

**Optical Flow Vector:**

The commonly used normal optical flow vector method is studied, the principle is discussed, and through repeated experiments the disadvantages of the influence of the presence of the shadow part is recognized and the influence of the shadow part is removed using HSV Space shadow detection. By that an improved motion region detection method is proposed. The algorithm principles of HSV spatial shadow detection and is introduced. After the detection of moving regions, the accuracy of detection was improved by shadow detection removal and moving targets were obtained more clearly.

**RCNN:**

A small daily object detection data set is taken and the different models (created by using different parameters) are trained with this data set. The models shown good accuracy in detection and more better models can be obtained by optimization of the models and by fine tuning the parameters.

6.REFERENCES

1. Rui Liang, Hui Zhi, M. M. Kamruzzaman “Methods of Moving Target Detection and Behavior Recognition in Intelligent Vision Monitoring” Acta Microscopica 2019.
2. Sheng Ding, Kun Zhao “Research on Daily Objects Detection Based on Deep Neural Network”.
3. Junjie Huang, Wei Zou, Zheng Zhu, Jiagang Zhu “Optical Flow based Real-time Moving object Detection in Unconstrained scenes.
4. Li, L., Xia, Z., Hadid, A., Jiang, X., Zhang, H., & Feng, X. (2019) “Replayed Video Attack Detection Based on Motion Blur Analysis”, IEEE Transactions on Information Forensics and Security.
5. Kaiming He, Georgia Gkioxari, Piotr Dollar, et al. Mask R-CNN[C] //IEEE Conference on Computer Vision and Pattern Recognition.2018.
6. Lee, S. Y., Kim, S. Y., Lee, J. H., Park, T. K., & Kim, J. T. (2019) “An Unmanned Aerial Vehicle System for Detecting and Tracking Moving Objects”, Advanced Science Letters.
7. Z. Deng, S. Hao, S. Zhou, J. Zhao, and H. Zou, ‘‘Toward fast and accurate vehicle detection in aerial images using coupled region-based convolutional neural networks,’’ IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., Aug. 2017
8. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// Computer Vision and Pattern Recognition. IEEE, 2016.
9. Kwak S, Cho M, Laptev I, et al. Unsupervised object discovery and tracking in video collections [C]//IEEE International Conference on Computer Vision, 2015.
10. Wang, Q., & Lu, P. (2019) “Research on Application of Artificial Intelligence in Computer Network Technology”, International Journal of Pattern Recognition and Artificial Intelligence, 33(05) pp. 1959015.