[1]Ranjitha P

[2] Saira Banu Atham

GHE-Ensemble: Enhanced Hybrid Image Enhancement Model for Night Vision Multi-Object detection in Autonomous Vehicle

Journal of Electrical Systems

Abstract: - Autonomous vehicle is the exponential growing topic in the research filed due its huge demand in the current industry. Where object detection plays the major role to take any major decision. Object detection in the normal vision yields the excellent performance due to deep learning models. But object detection in the night vision does not yield the good performance due to various challenges in it. In the proposed work created a night vision database called BDD-Darko where it contains night vision images with multi objects. Applied novel GHE-Ensemble model for image enhancement technique and trained using deep learning model Yolov5 for the multi-object detection in the night vision. Which resulted in 65.3 % accuracy. Proposed model is yielding better results for image enhancement technique for night vision than the previous existing model and detecting the multi object in the single frame with better performance. This paper proposes (1)A new database called as BDD-Darko which contains night vision image with multiple objects. (2) A Novel GHE-Ensemble model for image enhancement technique for night vision images

Keywords: Multi Object Detection, Image Enhancement, Night Vision Image, Gaussian Filtering, Adaptive Histogram Equalization, Pixel Normalization, Yolov5

I. INTRODUCTION

Autonomous vehicle is the type of smart vehicle where it has ability to sense the surrounding environment and take its own decision like when to move, stop or slow down the vehicle by detecting the objects in the nearby [1]. Therefore, object detection is the crucial task in the autonomous vehicle. There are superior works in the field of deep learning and computer vision to detect the object in the normal condition. But only few works carried out in the extreme condition-night vision. It is very hard to identify the objects from the naked eyes. Thermal sensors are used for this but it is very expensive which cannot used in all the scenario. There are huge challenges in the night vision due to poor quality of the image. Where image suitable enhancement technique plays crucial role in working in this condition.

The main aim of any image enhancement techniques is the improve the given image so that it sharpens the image, reduce the noise if it is available, normal distribution of the pixel throughout the image, resizing the image if the given image is small or it is too big. For the better understanding of image, analysis of the image and working on the image [2]. There are two types of image enhancement technique i.e. techniques in the spatial domain and techniques in the frequency domain. In the spatial domain, gray level transformation is used in the proposed work. In the frequency domain histogram, adaptive histogram equalization, Gaussian filtering and pixel normalization is used. Apart from Autonomous vehicle, image enhancement techniques can also be used in the medical image, color processing, pattern recognition, coding and decoding the images [3].

A. Challenges

Challenges in the night vision images is 1 Due to uneven or sometimes low intensity in the night vision images foreground and background treated as the same [4]. Which leads to segmentation issues.2. Many information like vehicle, road, traffic light, person will be lost due to poor lighting. 3 No accurate object is detected due to motion blurring, blurred images or partial occlusion [5]. 4 Night images does not receive the sufficient light due to which there will be more noise resulting in the failure of the multi object detection. 5 If the light is more due to street light or vehicle which leads to the high exposure level in the image or if light is less low exposure level of the image. Both leads to the loss of important information. If we try to maintain the exposure level of light to be normal then also it leads to the loss of the information [6]. Extracting or detecting object information in the night image is not possible without applying the most feasible pre-processing technique (image enhancement technique). After going through all this challenges and difficulty HEG-Ensemble model for this work is proposed as pre-processing technique.

[1] *Ranjitha P: Assistant Professor, School of Computer Science and Engineering and Information Science, Department of Computer Science and Engineering, Presidency University, Bangalore, India.

[2] Saira Banu Atham: Professor, School of Computer Science and Engineering and Information Science, Department of Computer Science and Engineering, Presidency University, Bangalore, India

B. Contribution of the Proposed Work

1. Created a new dataset named BDD-Darko from the available open source BDD100K video. BDD-Darko has 20K night images with multi-object in the road. Which helps other researcher to work on the night vision in the autonomous vehicle. After completing the research work this dataset will be deployed as open source. 2. All the images are pre-processed with the Novel GHE-Ensemble image enhancement technique for the better detection of the object in the night vision. Which is yielding the better results in detecting the objects in night vision than the base line method.

II. RELATED WORKS

This work aims to detect the pedestrian in the night image with the help of visible light. Where used RCCN for the object detection. Even through the proposed work is increasing the accuracy but not reliable in the detecting the multi object from the single image. This method resulted in the 57.59% of LAMR (low average miss rate) with fusion and without fusion in the feature extraction. [5] Used BDD open-source dataset with generative adversial network for the enhance feature extraction on the road. In the deep learning framework used ROI pooling method which resulted in 45% average precision in the confidence level 75. [8] Worked for the night images with and without applying image enhancing technique.

In image enhancement techniques they have applied like entropy, mean square error and pixel density. They achieved 58% accuracy for without image enhancement technique and 71% of accuracy in detecting the single object after applying the image enhancement techniques. [9] Worked on the sensor fusion camera and radar for object detection in the road condition. Where they focus on distance, angle of capturing images, environmental set up, converting 3D perspective to 2D perspective, 1D segmentation [10].

Worked on the degraded pixel in the image where there is huge loss of data due to heavy noise and applied robust RCNN with Yolov4 resulted in the mean accuracy precision of 61.85.[11] Few works focus on detecting the pedestrians and tracking their movements using the graph-based method. Where each object that is pedestrian is treated as the nodes in the graph. [12] For working with challenging environment applied deep learning neural network for generalization. To see whether object is present or not in the given image [13].

While working on the underwater images which is not receiving the proper light. Applied image dehazing and deblurring methods are used. Few other works deal with applying feature attention regional feature information and feature maps for dealing with the night images or low-level feature image. [14] Segmentation is the challenging task in the night vision images for that instance segmentation is used for boosting the object detection the challenging weather condition. [15] For working with night surveillance camera yolov4, yolov5, single shot detector (SSD), retina net are used as comparative study and their results are compared.

III. METHODOLOGY

Figure1 shows the overall working of the proposed model. As a first step video is taken and converted into the frames. After discarding few images due to completely blur or poor quality collected 25K images. Then collected images subjected to the GHE-ensemble model where each image or frame undergo four major process 1. Gray level transformation 2. Gaussian filtering 3. Histogram 4. Enhanced adaptive histogram equalization. Once the pre-processing is done for each image. Then dataset is ready for further training. Then it is trained using YoloV5 model. Once the knowledge base is ready it is subjected to validating and testing. Finally model performance measuring parameters like precision, recall, accuracy, mean average precision is calculated and compared with the baseline model.

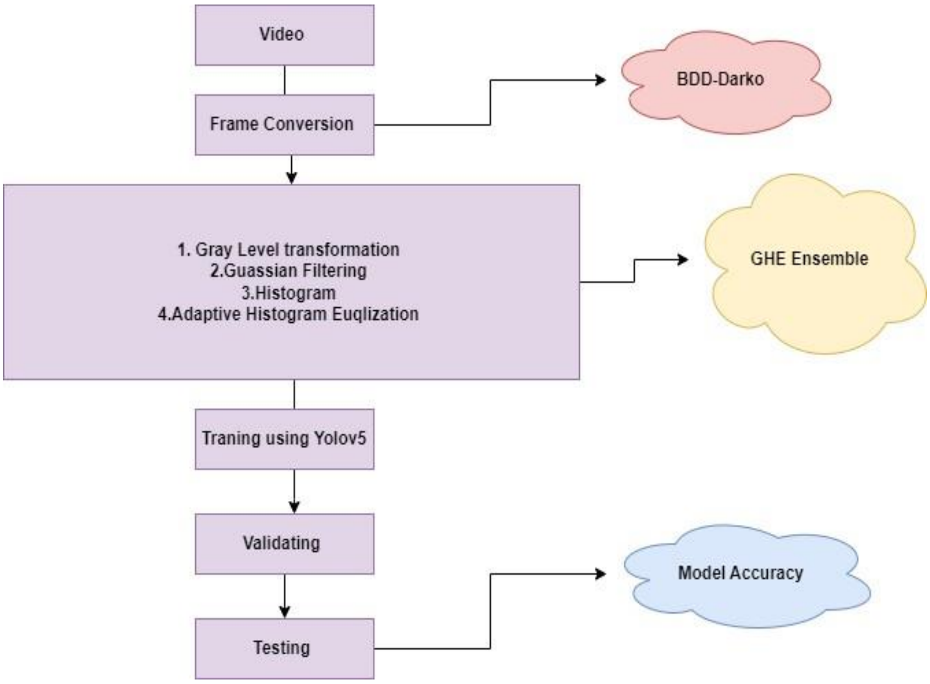

Figure 1: Working of the Proposed Model

A. Dataset Creation

Creating the suitable dataset which aligns with our research problem is the crucial role any research. Here to work with autonomous vehicle object detection the night vision initially various data sets are considered. They are KITTI, Caltech USA, ImageNet, Cityscapes, Microsoft COCO, Berkeley DeepDrive, Pascal VOC, Astyx Dataset HiRes2019, Waymo Open Dataset, NuScenes Dataset, Indian Driving dataset [16]. After the clear invesgition found Berkeley DeepDrive dataset which is suitable for the proposed work [1]. Berkeley DeepDrive dataset comprise of 100K videos of autonomous vehicle with various extreme weather condition [17]. From the huge dataset 200 videos are taken particularly for the night vision. 200 videos are converted into frames resulted in 20K images of night vision images. This 20K images of night vision is considered as dataset for the proposed work. Newley created dataset termed as BDD-Darkco. Figure 2 shows the sample image from the created BDD-Darko dataset and figure 3 shows the created dataset after image enhancement technique.

Figure 2: Sample image from the BDD-Darko dataset

Figure 3: Sample image from the BDD-Darko dataset after applying GHE Ensemble.

B. FRAME CONVERSION

Frames are converted into 640 Pixel and transformed coloured to the gray scale for the future analysis of the frame. used scikit-image 0.13.1 and OpenCV-python 3.4.0.12. If the frames are less than 640 pixels then it is resized using math function and then converted into the gray scale.

C. PIXEL NORMALIZATION

It is the one of normalizing technique in the image processing. Where it equally normalizes or distribute the pixel throughout the given input image. It takes the maximum pixel value available in the given image and divide all the other pixel value by it. Later it recomputes the pixel value by the equation 1. For the proposed model initially tried this method but more than this gaussian filter worked well for the model [18].

$$P(a,b) = \frac{(m(p(a,b) - fc\ )}{rc} \quad where\ fc = \frac{\sum_a \sum_b P(a,b)}{n} \qquad 1$$

$$rc = \sqrt{\sum_a \sum_b (p(a,b) - fc)^2} / n$$

where m is the scaling actor, P (a, b) is the image at the pixel value at a, b point, n is the total number of pixels.

D. GAUSSIAN FILTER

Filtering the image by deleting the low pass part of the image from the original image. In the simpler word adding high frequency content to the raw input image to increase the quality to the image. Intern it decrease the noise in the image [19]. There is two ways to find the gaussian filter i.e., one-dimension gaussian filter and two-dimension gaussian filter. Equation for two-dimension calculation [20].

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} \qquad 2$$

E. HISTOGRAM

It is the graphical representation of the given input image. Where there is two axis x and y in that x axis represents the variations and y axis represents the number of pixel available in that particular point [21]. Histogram has various application they are 1. Analysing the images 2. Increasing the brightness of the given image .3 Image equalization: equal distribution of the pixel throughout the image. 4 Thresholding the image with the customized value. 5. Any other type of transformation can also be done. Working of histogram is explained in the algorithm 1.

Algorithm 1: Working of Histogram

Step1: Find the maximum intensity and fix the gray levelling range which includes the maximum intensity in that.

Step 2: Find the number of occurrences of pixel in each gray level.

Step 3: Plot the graph based on the number of occurrences of the pixel and respective gray scale.

Step 4: Find the probability density function using the below equation.

$$M(K) = \frac{F(K)}{R} \qquad 3$$

Where M(K) is the probability density function F is the frequency of the gray scale, R is the highest value of the gray scale.

Step 5: Find the cumulative distribution function for each M(K) in the above step 4.

$$C(K) = F(K) = \sum_{j=0}^{K} S_R\,(R_j) \quad K = 0,1, \ldots.last\ element - 1 \qquad 4$$

Step 6: Multiply the C(K) with the maximum gray value.

Step 7: Apply the round off operation for each of the value obtained and finally plot the histogram graph again.

F. ADAPTIVE HISTOGRAM EQUALIZATION

It is the one of the most important image enhancement techniques in the histogram [22]. It computes the many histograms for the single given image. Each part of the image it is applying and it will redistribute the pixel value of the image. Which intern increases the contrasts of the given image. For the created dataset it is best suitable [23]. Here we modified the various clip value for the adaptive histogram equalization. Clip value of this method tells that how much maximum distribution of pixel can be done for the given any image. For the created dataset worked with 0.1, 0.2, 0.3, 0.4 ,0.5 Clip value. From that 0.2 yielded the best result for the proposed model. Advantages of using this technique is it computes many histograms in different section of the given image. It increases the contrast of the input image locally. It preserves the edges in the various part of the image [24].

G. WORKING OF YOLOV5 MODEL

Yolov5 is the one the most popular image processing or computer vision algorithm or model for detecting the objects in the given frame or video in the single shot. It is called as you only look once [26]. It is based on the deep learning framework. Where there are many versions in the algorithm. Depending on the type of dataset, training time, output required Yolov3, Yolov5, Yolov7 are used. For the proposed model yolov5 is used to detect the multi object in the night vision environment. In yolov5 also there is four different kinds available i.e 1. Small 2. Medium 3. Big 4. Extra-large. For the proposed work big is used. First yolov5 algorithm generates the various feature from the given input frame or image. Later this feature is subjected to the learning model and creates the bounding box for detection of object and predicting the suitable class label. In the proposed model the image or frame obtained from the GHE ensemble is passed to the YOLOv5 model. Where there total 6 classes i.e 1. Car 2. Person 3. Building 4. Road 5. Bus 6. Truck. Once the input image is sent to the yolov5 it will undergo major three process 1. Backbone 2. Neck 3. Head. Figure 4 shows the various task carried out by the three processes in the model [27] [28]. Figure 5 shows the overall working algorithm of the yolov5 algorithm. Where x is the bounding box, c is the predicted class label for the model, xw, xh is the width and hight of the bounding box, n is the number of labels in the class, x1 and y1 are the two coordinates in the bounding box [29].

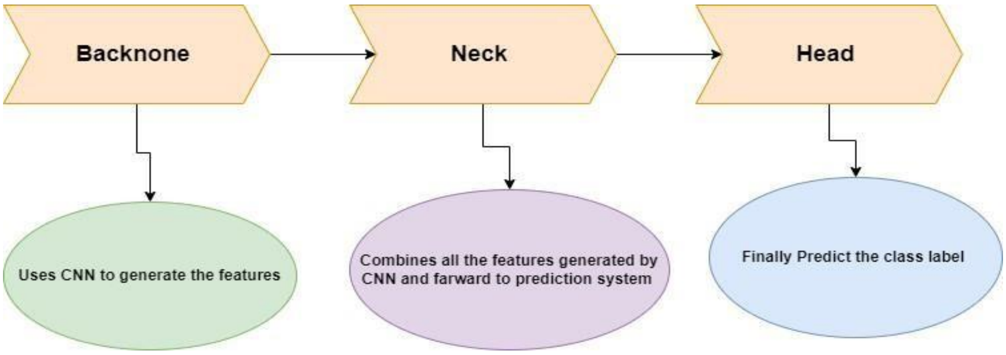

Figure 4: Working of three major steps in the Yolov5 model.

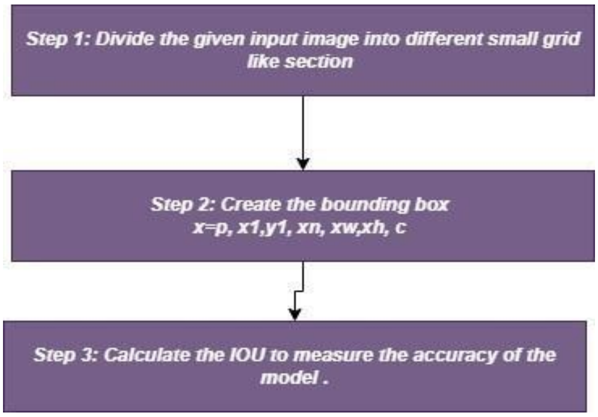

Figure 5: Working algorithm of the YOLOV5 model.

IV. RESULTS AND DISCUSSION

For the proposed model to run Jupiter and google colab is used. For plotting the graph MATLAB 2020 and Roboflow is used. For training, validating and testing 67% 23%, 10% Of dataset is used respectively. Figure 6 and 8 shows the sample image of the created dataset without applying any pre-processing technique. Figure 7 and 9 shows the image after applying the GHE ensemble image enhancement model. Figure 10 and 11 shows the histogram graph before applying any pre-processing technique and after applying the GHE ensemble model. Figure 11 clearly shows even distribution of pixel throughout the image. Table a show the comparison of MAP before applying GHE Ensemble model and after applying it. Where there is huge difference in the value. Figure 12 and 13 shows the comparison of MAP before and after applying GHE Ensemble model. The final accuracy and other performance measuring parameter is compared and discussed in the comparative study in the next section.

Figure 6: Sample dataset before image enhancement.

Figure 7: Sample dataset after applying image enhancement for the Figure 6.

Figure 8: Sample dataset before image enhancement.

Figure 9: Sample dataset after applying image enhancement for Figure 6.

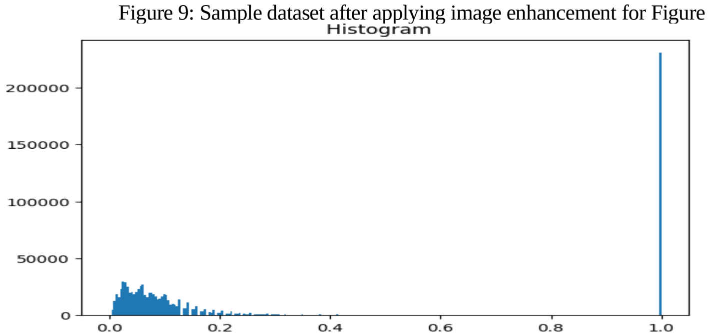

Figure 10: Graph of pixel before applying Histogram

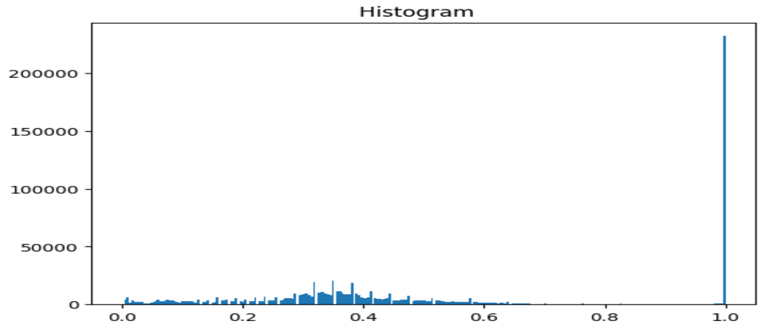

Figure 11: Graph of pixel after applying Histogram.

Table 1: Comparison of mean average precision of original data and GHE- Ensemble data.

Dataset	Number of images	Number of objects	MAP
Original training dataset	20000	79413	33.4
GHE- Ensemble trained data set	25000	90126	65.3

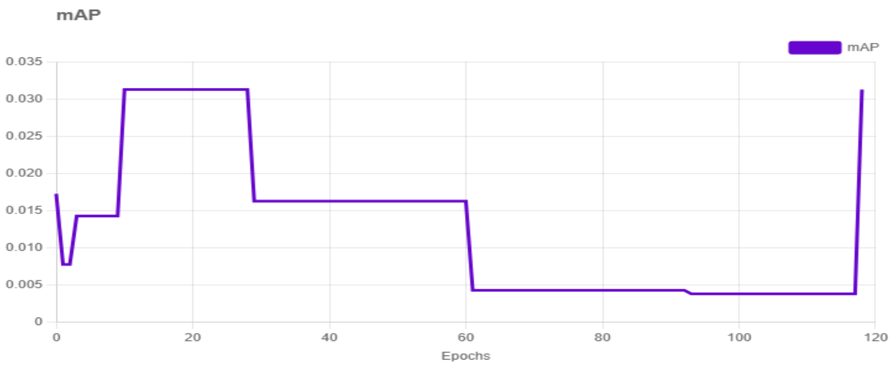

Figure 12: Shows the MAP value for created dataset before applying image enhancement technique

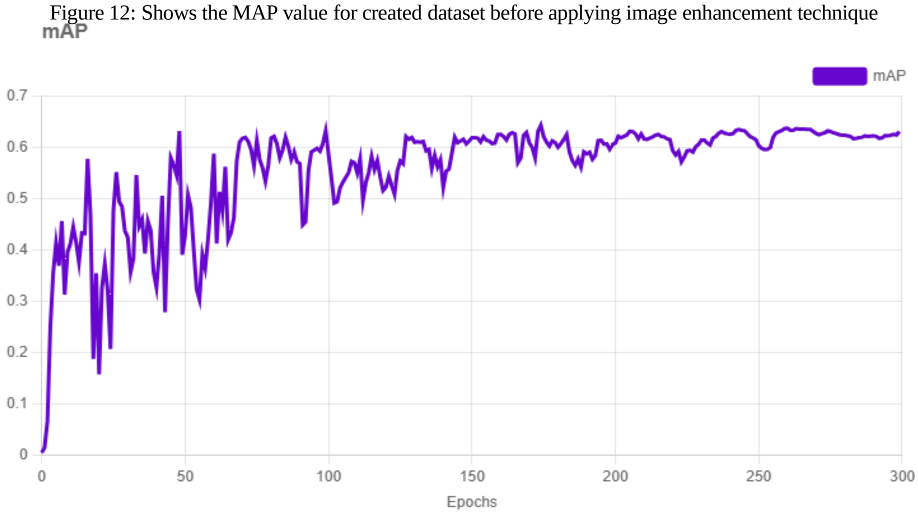

Figure 13: Shows the MAP value for created dataset after applying image enhancement technique -GHE Ensemble.

V. COMPARATIVE STUDY

Table 2 show the comparison of our model GHE- Ensemble with Yolov5 with other previously existing model at the threshold value of average precision at 25, 50, 75. Where proposed model is working better than the other baseline model at all the threshold value. Table 3 shows the comparison of the proposed model with other baseline model like Yolov5, retina net, Yolov4, faster RCNN with respect to MAP at 0.5% confidence level. The proposed model MAP is better than the previously existing model.

Table 2: Comparison of average precision at the different threshold value 25, 50, 75.

Algorithm	AP(25)	AP(50)	AP(75)
Faster R-CNN[9]	35.7	58.2	38.8
Cascade R-CNN[9]	39.3	61.9	41.7
Mask R-CNN[9]	32.3	56.1	34.5
Retinal Net[9]	29.9	53	30.2
SSN[9]	29.9	53	30.2
GHE-Yolov5(Proposed model)	45.9	65.3	55.9

Table 3: Comparison of MAP of various existing method at the confidence level 0.5%.

Method	Mean average precision (MAP) at 0.5%
Yolov5[13]	62.9
Retina net [13]	58.2
Yolov4[13]	59.8
Faster RCNN[13]	54.3
Proposed Model	65.3

A. EVALUATION METRICS

Precision: It is defined as ratio of the predicted sample in the ground truth to the total number of predicted samples in the given dataset. Where A is the true positive and B is the false positive [31]. Equation 5 shows the formula to calculate the precision of any given model.

$$Precision = \frac{A}{A + B} \qquad 5$$

Recall: It is the one of the measuring parameters of any given model. Here it measures the ratio of true positive to the true positive and true negative. Given any model as precision and recall value is high is considered as the good performance of the system. Always value lies between 0 to 1. If calculated in percentage then it will range from 0 to 100% [32]. A is true positive and B is true negative. Equation 6 gives the formula to calculate the recall of the given model.

$$Recall = \frac{A}{A + B} \qquad 6$$

Average precision: Equation 7 shows the way to calculate the average precision for the

proposed model. This parameter shows how accurate the model is predicting the class. where r is the recall, p is the precision, m is the threshold value [33].

$$Ap = \sum_{m} \frac{(r_m - r_{m-1})}{p_m} \qquad 7$$

Mean average precision: Equation 8
gives the formula to create the MAP where m is the threshold value, i is variable which is initialized to 1 and which goes upto m, APi is the average precision at the ith value [33].

$$MAP = \frac{1}{m} \sum_{i=1}^{m} AP_i \qquad 8$$

VI. CONCLUSION AND FUTURE SCOPE

For the proposed night vision created the database BDD-Darko from the standard BDD dataset. For the dataset applied GHE- Ensemble image enhancement technique. Where it resulted in equally distributed pixel value in the images. Then subjected to the Yolov5 training model and tested. Which resulted in far better performance than the baseline model or previous existing model [34]. Further it can be trained using hybrid deep learning model for the better improvement of the system [35].

REFERENCES

[1] P Ranjitha, Saira Banu, N Manohar. "Comparative Analysis of Autonomous Vehicle in Computer Vision" , 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC), 2022.

[2] https://medium.com/@gokcenazakyol/what-is-image-enhancement-image-processing-3-32a813087e0a.

[3] A. Łoza, D. R. Bull, P. R. Hill, and A. M. Achim, "Automatic contrast enhancement of low-light images based on local statistics of wavelet coefficients," Digital Signal Processing, vol. 23, no. 6, pp. 1856–1866, 2013.

[4] YUXUAN XIAO, AIWEN JIANG1,JIHUA YE ,MINGWEN WANG," Making of Night Vision: Object Detection under Low-illumination", 10.1109/ACCESS.2017.DOI.

[5] Liu, Y.; Qiu, T.; Wang, J.; Qi, W. A Nighttime Vehicle Detection Method with Attentive GAN for Accurate Classification and Regression. Entropy 2021, 23, 1490.

[6] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the exclusively dark dataset," Computer Vision and Image Understanding, vol. 178,pp. 30–42, 2019.

[7] Jong Hyun Kim, Ganbayar Batchuluun, Kang Ryoung Park, "Pedestrian detection based on faster R-CNN in nighttime by fusing deep convolutional features of successive images", Expert Systems With Applications 114 (2018) 15–3

[8] Aashish Bhandari, Aayush Kafle, Pranjal Dhakal, Prateek Raj Joshiand Dinesh Baniya Kshatri, "Image Enhancement and Object Recognition for Night Vision Surveillance", arXiv:2006.05787v1 [cs.CV] 10 Jun 2020.

[9] Kamil Kowol, Matthias Rottmann, Stefan Bracke, Hanno Gottschalk," YOdar: Uncertainty-based Sensor Fusion for Vehicle Detection with Camera and Radar Sensors", arXiv:2010.03320v1 [cs.CV] 7 Oct 2020.

[10] Nilantha Premakumara, Brian Jalaian , Niranjan Suri, and Hooman Samani, "Enhancing Object Detection Robustness: A Synthetic And Natural Perturbation Approach", arXiv:2304.10622v1 [cs.CV] 20 Apr 2023.

[11] Hrishikesh Vachhani, Thangarajah Akilan, Yash Devmurar, Nisharaff Shaik, Dhruvisha Patel , "Low-Light Pedestrian Detection In Visible And Infrared Image Feeds: Issues And Challenges", arXiv:2311.08557v1 [cs.CV] 14 Nov 2023.

[12] Ahmed, M.; Hashmi, K.A.; Pagani, A.; Liwicki, M.; Stricker, D.; Afzal, M.Z. Survey and Performance Analysis of Deep Learning Based Object Detection in Challenging Environments. Sensors 2021, 21, 5116.

[13] Wu, S.; Liu, Z.; Lu, H.Huang, Y. Shadow Hunter: Low-Illumination Object-Detection Algorithm. Appl. Sci. 2023, 13, 9261.https://doi.org/10.3390/app13169261.

[14] Hashmi, K.A.; Pagani, A.; Liwicki, M.; Stricker, D.; Afzal, M.Z. Exploiting Concepts of Instance Segmentation to Boost Detection in Challenging Environments. Sensors 2022, 22, 3703. https://doi.org/ 10.3390/s22103703

[15] Dipali Bhabad , Surabhi Kadam, Tejal Malode , Girija Shinde , Dipak Bage, "Object Detection for Night Vision using Deep Learning Algorithms", International Journal of Computer Trends and Technology Volume 71 Issue 2, 87-92, February 2023 ISSN: 2231 – 2803/https://doi.org/10.14445/22312803/IJCTT-V71I2P113

[16] https://analyticsindiamag.com/top-10-popular-datasets-for-autonomous-driving-projects/

[17] https://github.com/SysCV/bdd100k-models/tree/main/pose

[18] Parth Sane1 , Ravindra Agrawal "Pixel Normalization from Numeric Data as Input to Neural Networks", 978-1-5090-4442-9/17/$31.00c 2017 IEEE.

[19] https://www.southampton.ac.uk/~msn/book/new_demo/gaussian/#:~:text=The%20Gaussian%20Smoothing%20Operator%20performs,defines%20the% 20amount%20of%20blurring.

[20] https://blog.demofox.org/2022/02/26/image-sharpening-convolution-kernels/

[21] Zhu, H.. "Image Contrast Enhancement by Constrained Local Histogram Equalization" , Computer Vision and Image Understanding, 19990201.

[22] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. T. H. Romeny, and J. B. Zimmerman, "Adaptive histogram equalization and its variations," Computer Vision, Graphics, and Image Processing, vol. 39, p. 355–368, Sept. 1987.

[23] https://in.mathworks.com/help/visionhdl/ug/contrast-adaptive-histogram-equalization.html

[24] https://www.geeksforgeeks.org/adaptive-histogram-equalization-in-image-processing-using-matlab/

[25] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (Las Vegas, Nevada, USA), pp. 779–788, 2016.

[26] S. Nikkath Bushra, G. Shobana, K. Uma Maheswari, Nalini Subramanian. "Smart Video Survillance Based Weapon Identification Using Yolov5" , 2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC), 2022

[27] Markus Braun, Sebastian Krebs, Fabian Flohr, and Dariu M. Gavrila. \The EuroCity Persons Dataset: A Novel Benchmark for Object Detection". In: IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI) (2019).

[28] L. Tao, C. Zhu, J. Song, T. Lu, H. Jia, and X. Xie, "Low-light image enhancement using cnn and bright channel prior," in Proceedings of IEEE International Conference on Image Processing, (Beijing, China), pp. 3215–3219, 2017

[29] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele.\The Cityscapes Dataset for Semantic Urban Scene Understanding". In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2016..

[30] Yalun Wu, Yanfeng Gu, Yuanwan Chen, Xiaoshu Cui, Qiong Li, Yingxiao Xiang, Endong Tong, Jianhua Li, Zhen Han, Jiqiang Liu. "Camouflage Backdoor Attack against Pedestrian Detection" , Applied Sciences, 2023.

[31] https://www.analyticsvidhya.com/blog/2020/09/precision-recall-machine-learning/

[32] https://www.iguazio.com/glossary/recall/

[33] https://towardsdatascience.com/breaking-down-mean-average-precision-map-ae462f623a52

[34] Saira Banu Atham., R. P. . (2024). Enhanced Yolov5 Deep Learning Technique for Multi-Object Detection in Autonomous Vehicle in Extreme Weather Condition. International Journal of Intelligent Systems and Applications in Engineering, 12(3), 1162–1168. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/5395

[35] Ahamed, S. Aarif, and Chandrasekar Ravi. "Novel deep learning model for bitcoin price prediction by multiplicative LSTM with attention mechanism and technical indicators." International Journal of Engineering Systems Modelling and Simulation 13.2 (2022): 164-177.