International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 6 Issue 2, January-February 2022 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Helmet Detection Using Yolo -V3 Technique

Vijay Khare, Pratiroop Puri, Kratik Chaurasia

Jaypee Institute of Information Technology, Noida, Uttar Pradesh, India

ABSTRACT

Bike mishaps have been quickly enlarged over time in many countries. A helmet or a protecting cap is the main safety equipment of motorcycle riders, and two wheelers, however many driver's abandonment wearing helmets. The main outcome of wearing a helmet is to protect the head of a person travelling on two-wheelers just in case of a major or minor accident or fall from a running bike. Because of different social and monetary elements, this sort of vehicle is turning out to be progressively famous. The head protector is the fundamental security gear of motorcyclists; however, anyway numerous drivers don't utilize it. This paper proposes a framework for identification of motorcyclists without helmets. For this, we have applied the round Hough change and the Histogram of Oriented Gradients descriptor to remove the picture credits. Then the YOLO v3 was utilized and acquired outcomes. The system has given an average recognition accuracy of 75% that is satisfactory.

KEYWORDS: *YOLO v3, CNN, CHT, Darknet-53*

How to cite this paper: Vijay Khare | Pratiroop Puri | Kratik Chaurasia "Helmet Detection Using Yolo -V3 Technique" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-6 | Issue-2, February 2022, pp.986-989, URL: www.ijtsrd.com/papers/ijtsrd49341.pdf

IJTSRD49341

I. INTRODUCTION

Recently, because of the constant advancement of profound learning research, Girshick [1] and Ren [2] have separately proposed Fast Regional Convolution Neural Network (Fast R-CNN) which have worked on the exactness and realtime identification speed, yet there is a sure hole between them. In 2015, Redmon J [3] proposed YOLO object identification calculation, which adjusted the precision and recognition speed. In 2016, Redmon made YOLO-V2 [4] and YOLO-V3 [5] discovery calculations through progress. YOLO-V2 zeroed in on little article identification, expanded the mean exactness of mAP (mean normal accuracy) by 2%. The most recent YOLO-V3 further fortified multi-mark characterization and organization design, considering both exactness and identification speed, which has great discovery impact in development furthermore, different fields. In any case, there are still a few lacks in discovery exactness in the current circumstance. In this paper, in light of the YOLO-V3 discovery calculation, bunch calculation is utilized to anticipate the objective casing of the cap, and afterward the precision is advanced through the blend of profundity lingering organization and multi-scale location preparing in the preparation interaction[6].

II. Methodology

Three phases are merged in the proposed vision-based framework. The first step is to create a suitable dataset for training our model, as there is none available off the shelf. Then does the data pre-processing, which is divided into three parts: data acquisition, data enhancement, and data annotation. The photographs acquired had high resolution, various angles, and different backgrounds to create a more realistic scenario. The collected images are expended through augmentation techniques in terms of scaling, dropping, and changing brightness to increase the diversity and richness of the experimental dataset. The next step after image augmentation is image annotation, which involves creating a boundary box surrounding the objects and its label that is helmet or no helmet. Following the augmentation and annotation, a dataset of 2480 images was generated, with 80% of the images being randomly selected for the training dataset and the remaining 20% for the testing collection. During the training process, the image size will be reset, and the batch size will be fixed based on the memory constraints of the GPU. We will use the optimizer in the training, with the learning rate set to 0.001 and the other parameters remaining the same as they were in the YOLO model. The testing process will follow, during which a wide variety of images will be run through the proposed solution and results will be registered.

A. Data Set

The dataset consists of 2480 images of two classes

1. with helmet
2. without helmet

B. Feature Extraction

For training our custom object detection model, we will need a lot of images of objects which we're going to train nearly a few thousand. Number of images is directly proportional to accurate precision. We first perform feature extraction to determine the distribution and mathematical characteristics of the dataset; then we build YOLOv3 on our pre-processed data for training to build our model to detect helmets on the camera.

Based on the features of the dataset, we can obtain relevant information that will provide better support in building neural network training. For feature extraction, Calculation of the proportion of each target in the original image, calculate the average length of the target, calculate the average width, calculate the average area, calculate the average percentage of the target.Figure 1 shows the control flow diagram of Helmet Detection while capturing live feed. First of all, there will be a background subtraction from the extracted frame. The next stage would be that whether the output of first stage consists of a bike or not. If not then the process would be ended otherwise it will go to third stage i.e. the "Helmet Detection Module", Output of which would obviously our main concern.

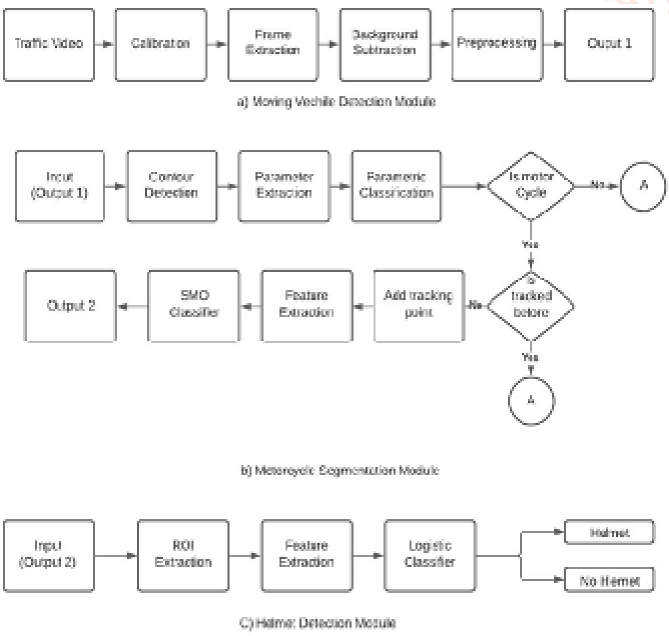

Fig 1: The proposed end to end Helmet Detection System

Fig. 2 and Fig.3 respectively show the structure of Optimization and workflow of the approach.

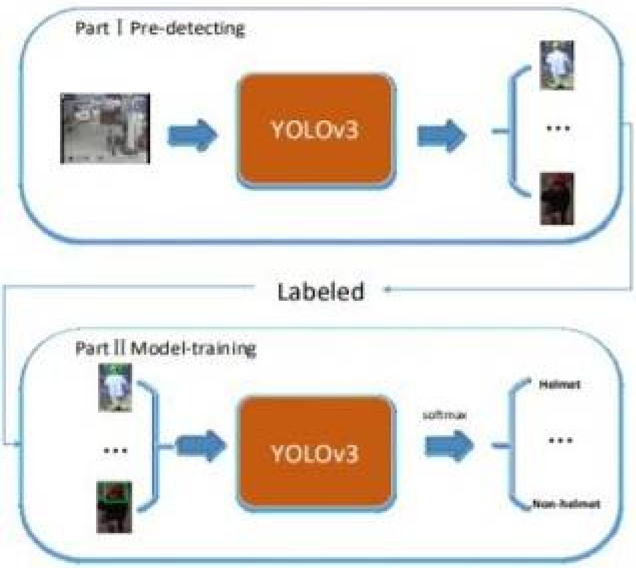

Fig 2 The structure of optimization method

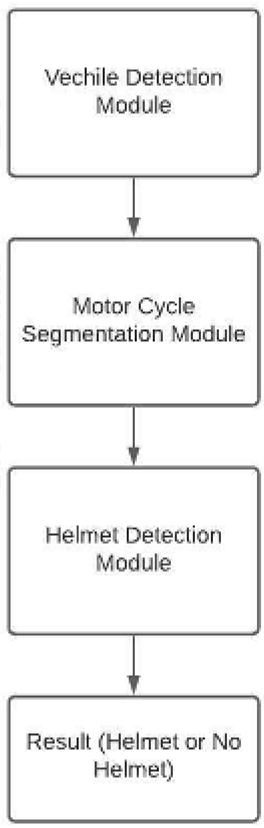

Fig 3: Workflow Diagram

III. Algorithms

In YOLO v3, the new association darknet-53 is used for incorporate extraction. There are 53 convolutional layers and 5 most limit pooling layers in the association structure. To keep away from over-fitting, a bundle normalization framework and a dropout movement are introduced after each convolutional layer. YOLO v3 improves target distinguishing proof accuracy by using a multi-scale incorporate blend computation to appraise the position and characterization on a multi-scale feature map. As far as possible encases YOLO v3, estimation bunches are used as before boxes. The k-infers approach is used to perform dimensional bundling on the goal encloses the dataset, achieving 9 priori boxes of various sizes that are reliably passed on among incorporate outlines of various scales. More unobtrusive concluded boxes are utilized for incorporate charts with a more noteworthy scale. Finally, the pack local area will be used to do security defensive cap wear recognizable.

A. Architecture of YOLO v3

- YOLO v3uses a variant of Darknet CNN architecture of Darknet has **53 layer network trained on Imagenet.**
- In YOLO v3, the detection is done by applying 3x3 and 1 x 1 detection kernels on feature maps of three different sizes at three different places in the network.
- The output is a list of bounding boxes along with the recognized classes. Each bounding box is represented by 6 numbers **(pc, bx, by, bh, bw, c)**.
- Finally, we do the IoU (Intersection over Union) and Non-Max Suppression to avoid selecting overlapping boxes.
- YOLO v3uses**binary cross-entropy** for calculating the **classification loss** for each label while **object confidence and class predictions are predicted through logistic regression.**

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
	Convolutional	32	1 × 1	
1×	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
	Convolutional	64	1 × 1	
2×	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
	Convolutional	128	1 × 1	
8×	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
	Convolutional	256	1 × 1	
8×	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
	Convolutional	512	1 × 1	
4×	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

B. Hyper-parameters used

- **Class_threshold**- Defines probability threshold for the predicted object.
- **Non-Max suppression Threshold -** It helps overcome the problem of detecting an object multiple times in an image. It does this by taking boxes with maximum probability and suppressing the close-by boxes with non-max probabilities (less than the predefined threshold).
- **input_height&input_shape** - Image size to input.

C. Training and Optimization

The preparation information is parted into 8:2 with 8 sections for training and 2 sections for testing. As the camera situations are mind boggling and diverse camcorders have various goals, the full association layer is eliminated in YOLO v3, so the prepared model can be taken care of pictures of various scales. So, we focus harder on the most proficient method to distinguish the far off, little and unclear targets better. In the exploratory interaction, we tracked down that the YOLOv3 model has a decent reaction to the ID of "individual". Hence, right off the bat, the specialists in the video are distinguished and caught by utilizing the YOLO v3 model, and afterward certain and negative examples are made. Because of the low goal of the video, which is fluffy and difficult to recognize. Part of the information from the positive examples are haphazardly separated and fluffy handling is to recreate the little impact.

IV. Result and Discussion

The preparation information is parted into 8:2 with 8 sections for preparing and 2 sections for testing. As the camera situations are perplexing and diverse camcorders have various goals, the full association layer is taken out in YOLO v3, with the goal that the prepared model can be feeded pictures of various scales. So we focus harder on the best way to recognize the inaccessible, little and ambiguous targets better. In the exploratory interaction, we tracked down that the YOLOv3 model has a decent reaction to the distinguishing proof of "individual". Hence, first and foremost, the laborers in the video are distinguished and captured by utilizing the YOLO v3 model, and afterward certain and negative examples are made. Because of the low goal of the video, which is fluffy and difficult to distinguish. Part of the information from the positive examples are arbitrarily removed and fluffy preparing is to recreate the little and fluffy examples, to improve the location precision.

In this paper helmet detection using YOLOv3 has been implemented and Figures 4 and 5 show the implemented results as helmet detection with a accuracy of 75%and non- helmet detection with 47% accuracy.

Fig. 4 Helmet is detected with a probability of 75%

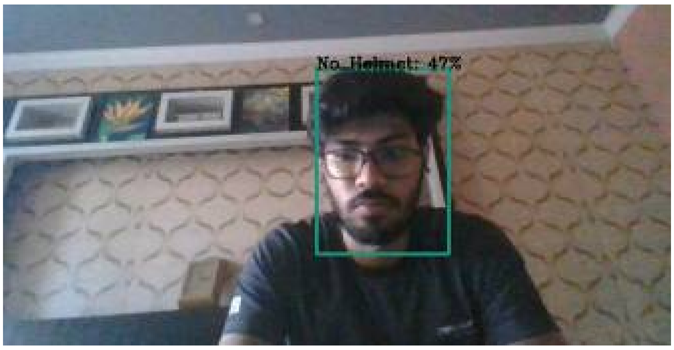

Fig. 5No Helmet with a probability of 47%

V. Conclusions

In this paper, we have used the YOLO v3 for identification of real time person with and without helmets. YOLO is suitable to detect the single object from the image, YOLO has a limitation that if there are multiple object in a single cell then YOLO is not suitable to all objects. Therefore if you know that your data set consist many small object in group then YOLO will unable to detect all the objects.

VI. Acknowledgment

We would like to place on record our deep sense of gratitude to Dept. of Electronics and Communication Engineering, Jaypee Institute of Information Technology, Noida, for their guidance, support, continuous encouragement and supervision throughout the course of the work.

References

[1] Girshick, R. (2015) Fast R-CNN[C]//IEEE International Conference on Computer Vision. IEEE, 2015:1440-1448.

[2] Ren, S., He, K., Girshick, R., et al. (2015) Faster R-CNN: towards real-time object detection with region proposal networks[C] //International Conference on Neural Information Processing Systems. MIT Press, 2015: 91-99.

[3] Redmon, J., Divvala, S., Girshick, R., et al. (2015) You Only Look Once: Unified, Real-Time Object Detection. J.2015: 779-788.

[4] Redmon, J., Farhadi. A., (2017) YOLO9000: Better, Faster, Stronger[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017:6517-6525.

[5] Sanjana, S., et al. "A review on various methodologies used for vehicle classification, helmet detection and number plate recognition." Evolutionary Intelligence (2020): 1-9.

[6] VinayakMudhomath | Prof. S. A. Angadi "Automatic Detection of Bike-Riders Without Helmet" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-2 | Issue-5, August 2018, pp.1972-1975, URL: https://www.ijtsrd.com/papers/ijtsrd18217.pdf