International Journal of Innovations in Engineering and Technology (IJIET)
http://dx.doi.org/10.21172/ijiet.173.01

Automatic Detection of Helmet and Non-Helmet Motorcycle

Anjana George, Paul Augustine
Department of Computer Science and Engineering
Rajagiri School of Engineering and technology, Kakanadkochi,India

Abstract—Nowadays, one of the major reasons for the accident is the motorcyclists who don't wear a helmet during driving. According to the law, every motorcyclist needs to wear a helmet while riding the motorcycle. Many of them are denying this and using their vehicle without safety equipment. In this paper, we propose an approach for automatic detection of helmet and non helmet motorcyclists using surveillance videos in real-time. We propose a programmed method utilizing a profound learning framework dependent on one of the profound learning strategies, a Darknet-53 based Convolutional Neural Network which is called YOLOv3

Index Terms—Darknet-53, CNN, YOLOv3, Helmet, Motorcycle

I. INTRODUCTION

Motorcycles are widely used as a means of transportation in many countries. The major advantages are their low prices and low operation cost in comparison with other vehicles. In countries like India, Brazil, Thailand, the majority of the population uses motorcycles for the daily commute. In India, as of 31 March 2015, there were 154.3 million registered motorcycles. In most of these countries, wearing a helmet for motorcyclists is mandatory by law. Also, considering the safety of people using motorcycles, wearing a helmet is paramount.The MAIDS (Motorcycle Accidents In-Depth Study) report, which become achieved in five European countries (France, Italy, Netherlands, Spain, and Germany), 7% of motorcycle accidents reported, the helmet could have prevented or reduced the head injury sustained by the motorcyclist.[19]

Currently, in practice, Traffic Police are entrusted with the task of ensuring that motorcycle riders wear a helmet. But, this method of monitoring motorcyclists is inefficient due to insufficient police force and limitations of human senses. Also, all major cities use CCTV surveillance based methods. But, those require human assistance and are not automated. Here we are utilizing a profound learning method called YOLOv3. When compared with different forms of YOLO, YOLOv3 utilizes a variation of Darknet and posses high accuracy so it would be progressively ideal for our project.

Here we are utilizing a profound learning method called YOLOv3. When compared with different forms of YOLO, YOLOv3 utilizes a variation of Darknet and posses high accuracy, so it would be progressively ideal for our project.

II. LITERATURE REVIEW

Video content analysis is one of the trending subjects in video image processing. There are many techniques and methods used for content analysis from hardware-based signal processing to convolutional neural network-based deep learning techniques like smoke and fire detection for a live video using neural networks [4]. Analyzing and making a short description of the video or even movies[3], and also providing captions for video audio signal[2] is also a major achievement in video content analysis. In any case, the CNN algorithms utilize only rule-based models and feature vector for classification[4]. These features are hard to define and depend on the kind of object observed. And the result leads to low detection rate.CNN is used for identifying specified objects in videos and it performs great in the classification of the objects. This network can perform classification of objects and extraction in similar architecture.

The detection of a cyclist wearing or not wearing a helmet can be detected using CNN. There has been an improvement in deep learning models and one such model used is YOLOv2. YOLO v2 is utilized at two unique stages so as to improve the helmet identification correctness. This model is trained in COCO data set and can detect all classes in it. The edited pictures of different people are used as information to the second stage which was

prepared on the data set of helmeted pictures[1]. To recognize the drinking condition of a driver by utilizing a camera, so as to block drinking and driving.

The technique identifies whether a driver has drunk liquor or not by utilizing face pictures and their parts caught by a camera. In examinations, a three-layered neural network is utilized to identify drinking and decide the portions of a face helpful for location. Through investigations, the precision of the technique is appeared and is contrasted and that by the decision tree 4.5 [10]. or detecting the unusual activity in a large video set an unsupervised deep learning technique can be used. Here it divides the video into equivalent length fragments and groups the removed highlights into models, from which a prototype segment co-occurrence matrix is processed. It looks for a correspondence relationship among models and video fragments that fulfills the transitive conclusion constraint[6].

Human activity recognition was used to analyze different actions, posture and sign of a human being[7] using deep learning techniques. It is nowadays used exponentially surveillance systems and other monitoring systems. There are still studies going on to help increase the efficiency, speed, and accuracy of the current system. A system exists that will convert human action video into different images and then combine that image to create a binary motion image[8] which can be used to detect human activity. Another major area where video content analysis is used for detecting unusual activity in surveillance[6].

Here live feed is analyzed and the system will alert the user in case of anomalies, cases like this require faster detection and accuracy. C3D architecture which uses eight convolutional layers[11] is used to process a video and to tell the story on the video. This is clearly a good approach to video processing to extract its information. But the more a faster approach is needed.YOLOv3 is a reasonable detector. It is fast and accurate. It's not as great on the COCO average AP between .5 and .95IOU metric. However, it is excellent on the previous detection metric of .5 IOU[9].

In Sequential Deep Learning for Human Action Recognition model, it's a 2 step neural-based deep model in which the first part of the model is the extension of ConvNets[12] to the 3D case where the features are automatically learned. LSTM recurrent architecture is used. Then, in the 2ND part, these learned features are trained on a recurrent neural network model to classify the complete sequence. Video can be converted description and this helps in guiding visually impaired people through the movie or video that they are going through[13]. The description can be audio and video. This is one application of video to the description. Based on the literature survey, in this paper, we finally conclude using the YOLOv3 model for the detection of cigarette smoking in video frames, as it posses a high accuracy rate compared to other models and acceptable speed.

III. DESIGN

A. *Architecture*

The system architecture (Figure 1) is the model that defines the entire framework of the system's behavior and structure. Here the input video from the user is split into frames then each frame is filtered one by one to remove noise. After that, each frame is passed to a trained neural network model for detecting the activity, if any objectionable action is recognized then the corresponding disclaimer is added to the video. This is repeated for the entire video frames. At last entire video is saved as an output file.

B. *Data flow*

The video is preprocessed and send to activity recognition, where the object detection is done if any utilizing neural system and related alerts are affixed to the video. (Figure 2)

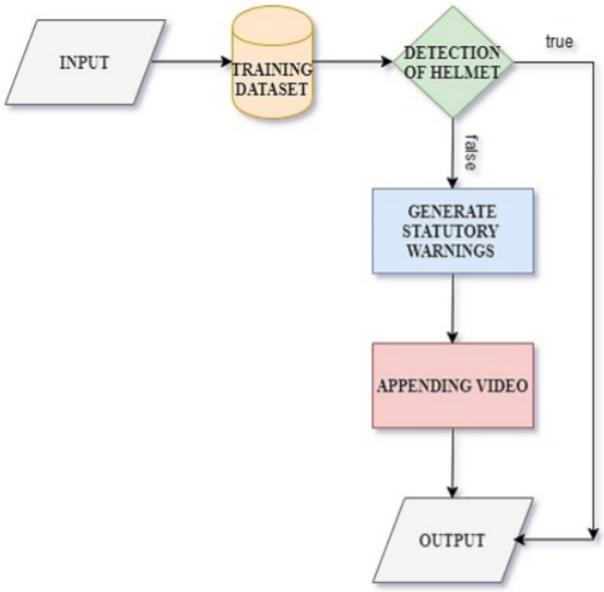

Fig. 1. Architecture

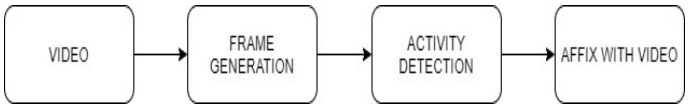

Fig. 2. Design

IV. IMPLEMENTATION

A. *Input video file to image conversion*

In spite of the fact that the venture depends on a video, the information is given as the casings of the video as the edges are handled to distinguish the unwanted substance in the video. The thought is to release 2 outlines for every second and then feed it to the neural system for handling.

B. *Object detection*

This is the principal part of the work. We are utilizing a convolutional neural system. For this purpose You only look once (YOLO) is a profound learning method and a constant object detection framework. On a Pascal Titan X, it forms pictures at 30 FPS and has anmAP of 57.9 rates on COCO test-dev. YOLOv3 is the most recent and it is incredibly quick and accurate.

Before YOLO was introduced, the detection systems apply the model to an image at multiple locations, then the scaling of the digital image is done.High scoring regions of the images are considered as detection. It is more than 1000x faster than R-CNN and 100x faster than Fast R-CNN.

In YOLO, it is absolutely an alternate methodology. The image goes through the convolutional neural system in a solitary pass. At that point, this system will separate the image into regions and predicts bounding boxes and probabilities for each region as shown in above Figure 3.

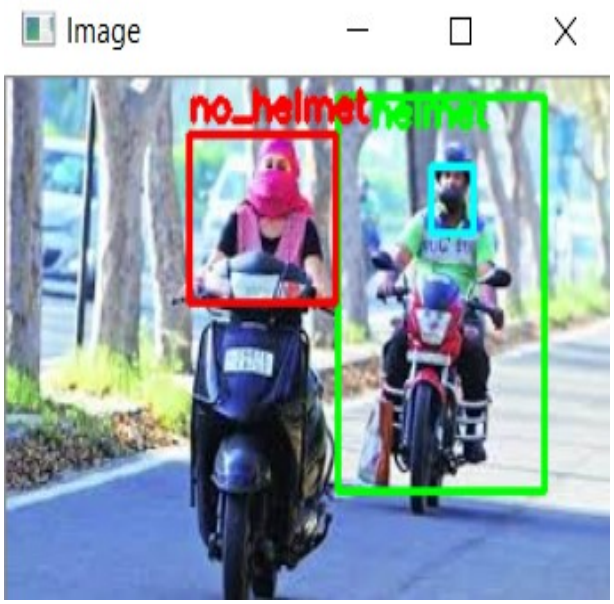

Fig. 3. Object Detection

The best thing about YOLO is that it gives a certainty score for each jumping boxes and these scores with a likelihood-appropriation over the class will give a superior result.YOLO separates the picture into 13x13 lattices. YOLO is basic. Give an information picture (for instance, re-sized to 416x416pixels). At that point, it experiences the convolutional neural system in a solitary pass and turns out the opposite end as a13x13x125 tensor depicting the bounding boxes for the lattice cells. At that point, locate the last score for bouncing boxes and set a limit. Dispose of the ones that don't fit.

V. CONCLUSION

This project proposing a design of automatic detection of the helmet and non helmet motorcyclists using Deep Learning. Normally this is done manually by screening the entire video file and manually inserting the warning. Detection of undesirable content in video scenes will take up time when done manually.

So, in this project, we have automated this entire process. The method will ensure that it will reduce both complexity and time. The automated system will recognize motorcyclists who drive without wearing a helmet.

VI. FUTURE SCOPE

This project is possible to computerize this procedure of including statutory warnings in videos or TV arrangement. These days the movie editor needs to manually check the whole video clips to detect the motorcyclists who don't wear a helmet. After scanning all the scenes, they put enough time for inserting statutory warnings and signs in videos. Statutory admonitions are shown in public interest declaring that the motion picture/Televisions indicates does not encourage smoking, drinking liquor, and so on. For example, if there is a person driving a motorcycle without a helmet scene in a film, at that point the undertaking will identify it and will demonstrate a related statutory cautioning as a subtitle.

REFERENCES

[1] Jimit Mistry, Aashish K. Misraa, Meenu Agarwal, Ayushi Vyas, Vishal M. Chudasama, Kishor P. Upla "An Automatic Detection of Helmeted and Non-helmeted Motorcyclist with License Plate Extraction using Convolutional Neural Network *International Conference on image processing theory, Tools and Application (IPTA)* 2018

[2] Subhashini Venugopalan1, Marcus Rohrbach, Jeff Donahue, Raymond Mooney, Trevor Darrell, Kate Saenko "Sequence to Sequence Video to Text *https://arxiv.org/abs/1505.00487*

[3] Anna Rohrbach, Marcus Rohrbach, BerntSchiele "The Long-Short Story of Movie Description *https://arxiv.org/abs/1506.01698*

[4] SebastienFrizzi, RabebKaabi, MoezBouchouicha, Jean-Marc Ginoux, Eric Moreau, FarhatFnaiech "Convolutional Neural Network for Video Fire and Smoke Detection IECON 2016 - *42nd Annual Conference of the IEEE Industrial Electronics Society*

[5] C. Vishnu, Dinesh Singh, C. Krishna Mohan and SobhanBabu "Detection of Motorcyclists without Helmet in Videos using Convolutional Neural Network *International Joint Conference on Neural Networks* 2017

[6] N Hua Zhong , Jianbo Shi , M. Visontai Detecting unusual activity in video *, Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2004

[7] Jagwinder Kaur Dhillon,Chandni,Alok Kumar Singh Kushwaha Human action recognition based on deep learning approach *fourth international conference on image information processing* 2017

[8] TusharDobhal, VivswanShitole Human activity recognition using binary motion image and deep learning *Second international symposium on computer vision and the internet* 2015

[9] Joseph Redmon, Ali Farhadi YOLOv3: An Incremental Improvement *University of Washington* 2018

[10] Kenichi Takahashi, Kana Hiramatsu, and MamiTetsuishi Experiments on Detection of Drinking from Face Images Using Neural Networks *Second International Conference on Soft Computing and Machine Intelligence* 2015

[11] HuijuanXu, Boyang Li, VasiliRamanishka, Leonid Sigal, Kate Saenko Joint Event Detection and Description in Continuous Video Streams *The IEEE and CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 2018

[12] MoezBaccouche ,Christophe Garcia Sequential Deep Learning for Human Action Recognition *Conference paper* 2011

[13] Anna Rohrbach, Marcus Rohrbach, NiketTandon, BerntSchiele A Dataset for Movie Description *The IEEE and CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 2015

[14] NarongBoonsirisumpun, WichaiPuarungroj, and PhonratichiWairotchanaphuttha Detector for Bikers with no Helmet using Deep Learning *International Computer Science and Engineering Conference (ICSEC)* 2019

[15] SumayyaFathima , UdayiniChandana "Bike Authentication by Helmet using Faster RCNN *International Journal of Research* 2019

[16] Dany Joy , Manu John† , Irene Abraham‡ , ElizebethShiju§ and Anjali A¶ Automatic generation of statutory warnings using deep learning *International Conference on Computing Methodologies and Communication (ICCMC)* 2019

[17] Geethapriya. S, N. Duraimurugan, S.P. Chokkalingam Real-Time Object Detection with Yolo *International Journal of Engineering and Advanced Technology (IJEAT)* 2019

[18] Reagan L. Galvez, Argel A. Bandala,Elmer P. Dadios Object Detection Using Convolutional Neural Networks *International Conference on Computer Science and Electronics Engineering*2018

[19] YogirajKulkarni,ShubhangiBodkhe,AmitKamthe, ArchanaPatil Automatic Number Plate Recognition for Motorcyclists Riding Without Helmet *IEEE International Conference on Current Trends toward Converging Technologies* 2018