International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 06 | June 2019 www.irjet.net p-ISSN: 2395-0072

# APPLICATION OF MCNN IN OBJECT DETECTION

**Kritika Yadav[1], Faseeh Ahmad[2]**

*[1]M. Tech Student, Department of Electronics & Communication Engineering, Goel Institute of Technology & Management, Lucknow*
*[2]Assistant Professor, Department of Electronics & Communication Engineering, Goel Institute of Technology & Management, Lucknow*

---

***ABSTRACT:****When we're shown an image, our brain instantly recognizes the objects contained in it. On the other hand, it takes a lot of time and training data for a machine to identify these objects. But with the recent advances in hardware and deep learning, this computer vision field has become a whole lot easier and more intuitive. Our work is to focus on achieving more accuracy rate on object detection in videos. The main objective of moving object detection is to take a video sequence from a fixed or moving camera and output a binary mask representing moving objects for each frame.*

*In recent years videos are widely adopted to monitor the security sensitive areas includes Highway, borders, banks and various public places except the development in computing power infrastructure of high-speed network large capacity storage device multi-sensor videos system. The important part of a machine to interact with human in an easy manner is the capability of Machines to identify the object and further identify the activities in the environment. In this research work MCNNs method is used for object detection on road which has a highest data accuracy.*

**Keywords:** ***Neural Network, CNN, MCNN, Object detection, Camera***

## 1. INTRODUCTION

Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance.

As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner. As a technological discipline, computer vision seeks to apply its theories and models for the construction of computer vision systems.

Digital Image Processing is the use of algorithms to carry out processing of digital image in the field of digital signal processing, digital image processing has many advantages over analog image processing. It allows a several type and range of algorithms to be applied as input data and avoids the build-up of noise and signal distortion during processing problems. Since images are created in two dimensional or more, digital image processing is modeled in the form of multidimensional systems.

Face detection is a technology being used in a variety of applications that recognize human faces in digital images.

Face detection can be defined as a specific case of object-class detection (OCD). In object-class detection, the main task is to find out the locations and sizes of each object in an image related to a given class. Face detection system recognize the faces, which include upper torsos, pedestrians, and cars etc. Face-detection algorithms focus on the frontal human faces detection. It is analogous to image detection in which the image of a person is matched bit by bit. Image matches with the image stores in database. Any facial feature changes in the database will invalidate the matching process.

A reliable face-detection approach based on the genetic algorithm and the eigen-face technique, At first, the possible human eye regions are detected by testing all the valley regions in the gray-level image. Then the genetic algorithm is used to generate all the possible face regions which include the eyebrows, the iris, the nostril and the mouth corners.

Each possible face candidate is normalized to reduce both the lightning effect, which is caused by uneven illumination; and the shirring effect, which is due to head movement. The fitness value of each candidate is measured based on its projection on the eigen-faces. After a number of iterations, all the face candidates with a high fitness value are selected for further verification. At this stage, the face symmetry is measured and the existence of the different facial features is verified for each face candidate.

## 2. APPROACHES

At first video is recorded using camera. Recorded video is used as an input for the image processing. Real time recording device is used to create movie in .mp4 format. Then python with Tensorflow is used for image generation. Yolo data segmentation method with MCNN is used and compared with CNN. Different pooling methods are also used and including BOW-Color with Max-Pooling, BOW-Color with Sum-Pooling, HOG-BOW-Gray with Max-Pooling and a comparison graph for accuracy is created.

Following major steps are involved in the image detection process:

1. We take video and extract frames from videos.
2. The Image will be divided into different reasons.
3. We will then consider each region as a separate image.
4. Pass all these regions of images to the CNN and classify them into various classes.
5. Once we have divided each region into its corresponding class we can combine all the Seasons to get the original image with the detect object.
6. Extracting 2000 regions for each image based on selected search.
7. Features are extracted using CNN for each reason of image, suppose we have K images then the number of CNN features will be K*2,000.
8. Multiscale CNN method is used for feature extraction.

## 3. RESULTS

Yolo data segmentation method with MCNN is used and compared with CNN. And with other pooling methods including BOW-Color with Max-Pooling, BOW-Color with Sum-Pooling, HOG-BOW-Gray with Max-Pooling.

Fig.1 show that the accuracy level for pedestrian detection is more than 95%. The probability of injury reduces if accuracy is over 95% as the sensor can send a warning to vehicle driver in the form of beep.

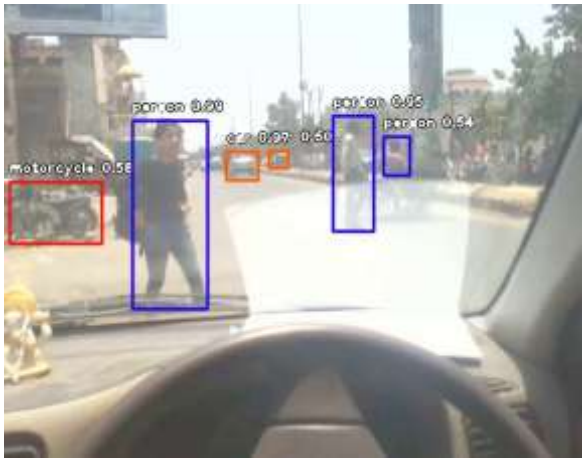


**Fig 1. Pedestrian Detection-1-MCNNs**

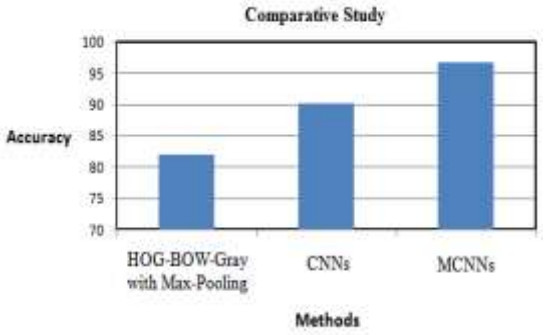


**Fig2.MCNNs Vs HOG-BOW-Gray with Max Pooling & CNNs**

Fig.2 shows that the MCNNs method is the best among other methods as accuracy is more than 95%.

## 4. CONCLUSION

Hand detection in still images plays an important role in many hand-related vision problems, for example, hand tracking, gesture analysis, human action recognition and human-machine interaction, and sign language recognition. Although hand detection has been extensively studied for decades, it is still a challenging task with many problems to be tackled. The contributing factors for this complexity include heavy occlusion, low resolution, varying illumination conditions, different hand gestures, and the complex interactions between hands and objects or other hands. The MCNNs is very useful in AI especially in medical field.

In the current study MCNNs method is used. The detected image has over 90% accuracy. The method can be modified to get more frames per second (fps).

## REFERENCES

[1] Viola P., "Feature-Based Recognition of Objects", AAAI Technical Report FS-93-04, 1993.

[2] LeCun Y., Bengio Y. and Hinton G., "Deep Learning" NATURE, Vol 521, 2015
[3] Tian B., Li L.， Qu Y.， Yan L.," Video Object Detection for Tractability with Deep
Learning Method ", Fifth International Conference on Advanced Cloud and Big Data, 2017

[4] Liu L., Ouyang W., Wang X., Fieguth P.,Chen J., Liu X., Pietik M., "Deep Learning for Generic Object Detection: A Survey" Research Gate, Sep 2018

[5] Dheekonda R. S. R., Panda S. K., Khan MD. N.,Hasan M. A., Anwar S., "
Object Detection from a Vehicle using Deep Learning Network and Future Integration with Multi-Sensor Fusion Algorithm," SAE, 2017.

[6] Tripathi S., Belongie S., Hwang Y., and Nguyen T.," Detecting Temporally Consistent Objects in Videos through Object Class Label Propagation", IEEE, 2016

[7] Roselo P., and Kochenderfer M J., "Multi-Agent Reinforcement Learning for Multi-Object Tracking," AAMAS, July 2018, pp 10-15.

[8] Lahamy H. and Litchi D., "REAL-TIME HAND GESTURE RECOGNITION USING RANGE CAMERAS ", ResearchGate, Jan 2009

[9] Srnivas B., Shivaranjani V., Udaykumar M., and Reddy N. A., " Moving Object Detection for Real-Time Traffic Surveillance using Genetic Algorithm, International Journal of Engineering Trends and Technology (IJETT), Vol. 49, No. 6, July 2017.

[10] Belhani H. and Guezouli L.,"Automatic detection of moving objects in video Surveillance".IEEE, Global Summit on Computer & Information Technology (GSCIT), 2016

**BIOGRAPHIES**

I have completed my B.tech in (Electronics and Engineering) from BBDNITM, Lucknow affiliated to AKTU. I am pursuing my M. tech in electronics & communication engineering from Goel Institute of Technology and Management.

Mr. Faseeh Ahmad is currently working as Head of Department (Electronics & Communication Engineering) at Goel Institute of Technology & Management, Lucknow. He has approximately 10 years of academic experience in his field. After pursuing his B.Tech degree in Electronics & Communication stream, he further enhanced his skills by acquiring PG Diploma in Embedded Systems Design from a premier institution C-DAC, Pune. Further he also pursued a PG Diploma in Wireless Telecommunication and worked as RF/DT engineer in telecom sector also. Afterwards he pursued his M.Tech degree in Electronics & Communication Engineering to serve in academics field in a more skilled & competent way.