



## MULTIPLE OBJECT DETECTION IN IMAGES USING CNN AND DPM

**G.RAMAMOohana RAO**, Assistant Professor, Dept.of AI & IT, DVR & Dr.HS MIC college of Technology, A.P., India.

**P.NAGA PAVANI**, MCA, DCA, DVR & Dr.HS MIC College of Technology, A.P., India.

**Abstract**— Object detection is an essential task in computer vision and image processing. It has many applications in various domains like medical diagnosis, civil military, video surveillance, security, etc. With the development of intelligent device and social media, the bulk of data on Internet has grown with high speed. There are so many important aspect in image processing, object detection is one of the international demanded research field. Multiple object detection is an important concept in object detection. In object detection extracting the features and handling the occlusion are two most important components. A Convolution Neural Network (CNN) has achieved great success in extracting the region based features which may used for extracting multiple regions from the images and Deformable Part Based Model (DPM) improve the ability for handling the occlusion. Occlusion handling is nothing but when multiple objects are near to each other that time some objects are not detected so this drawback will be handled by DPM. Existing method not performing well in the aspect of detecting multiple objects. In this paper CNN and DPM are to be integrated to detect multiple objects. By combining these two models we are able to notice every single object with high accuracy.

**Keywords**—CNN, Multiple object detection.

### 1. INTRODUCTION

Object detection used as an integral part, such as semantic segmentation, instance segmentation, pose estimation, suspicious activity detection, etc. The first stage in the pipeline is to detect an object. The survey begins with significant highlights of deep learning for object detection. It provides a comprehensive study on object representation; Convolution Neural Network (CNN) and different Deep Convolution Neural Network architecture. It presents a concise review of renowned datasets and definitive measurement metrics, forming the primitive baseline to evaluate the detection framework. Then studies in



detail on detection framework one-stage and two-stage detectors and evaluates each framework with standard datasets listing its vital significance. The study also explores different issues of object detection like multi-scale, intra-class variations, generalization & security. Now a days the deep learning has achieved good results in number of research areas and companions by the continual improvement of convolution neural networks. Object detection is an important application in convolution neural network. The CNN has build valuable progress in object detection. In the actual application it is very challenging task to use computer technology to detect the objects. Complex background, noise disturbance, occlusion, low resolution and other factors will seriously affect the object detection performance. Feature extraction is very difficult in object detection if size of the object is very small than alternative object then this types of objects are excluded throughout the detection method. The R-CNN model is deep learning model for object detection. Here we are going to apply the Region based model for detecting multiple objects but in R-CNN Selective Search Algorithm (SSA) not performing well when there are several objects which is close to one another. Although the deformable part based model will be able to find every single object that is partially occluded. In this paper we take the benefits of R-CNN and DPM and we are going to develop a new system framework that integrates R-CNN with DPM. The DPM can almost cover the entire object with high resolution that covers smaller parts of the object. DPM can rectify the capability for handling the occlusion and R-CNN will not only used for extracting the features but also handling the incorrect bounding boxes that may generated by DPM. The contribution of this letter is as follows. First to propose a framework which can merge the R-CNN and DPM for object detection. The proposals which are achieved by DPM are cleaned by using filter based Density Subgraph Algorithm (DSD). In this system framework for detecting each of objects and guess all the available proposals of objects initially apply the DPM. After this the proposals that are generated by SSA and proposals generated by DPM are one by one send to the CNN model for extracting features. These features are used to detect object. Using PASCAL VOC 2007 dataset we calculate the performance of our framework.

## 2. LITERATURE REVIEW

**Feed forward neural network training based interactive shopping for blind:** Today shopping markets pay attention towards customer needs and services. Unfortunately the blind and vision impaired person are still incapable to access these environments without reliance. Assistive technology is trying to sway the living style of the blind by introducing support systems for routinely actions like reading, writing, walking,



Web surfing, and shopping. However, still the blind have to count on others for personal accessories shopping. To overcome this problem, we designed and developed a feed forward talking accessories selector. Our system is trained using feed forward techniques with a feature level block-based multi-focus image fusion method to provide suggestions regarding product selection, fitness, and color combination, for instance in dress and jewelry. The evaluation of our system takes into account specialists opinions, such that statistical analysis shows similarity between both.

**Smartphone based obstacle detection and classification system for assisting visually impaired people:**In this paper we introduce a real-time obstacle detection and classification system designed to assist visually impaired people to navigate safely, in indoor and outdoor environments, by handling a smartphone device. We start by selecting a set of interest points extracted from an image grid and tracked using the multiscale Lucas - Kanade algorithm. Then, we estimate the camera and background motion through a set of homographic transforms. Other types of movements are identified using an agglomerative clustering technique. Obstacles are marked as urgent or normal based on their distance to the subject and the associated motion vector orientation. Following, the detected obstacles are fed/sent to an object classifier. We incorporate HOG descriptor into the Bag of Visual Words (BoVW) retrieval framework and demonstrate how this combination may be used for obstacle classification in video streams. The experimental results demonstrate that our approach is effective in image sequences with significant camera motion and achieves high accuracy rates, while being computational efficient

### 3. PROPOSED WORK

Existing paper of an object detection system that will supported merging of multiscale DPM. In the PASCAL object detection this system represents the deep object categories and achieves the progressive decisions. Although the deformable part based models have become quite common, on the difficult benchmark their value had not been demonstrated like PASCAL datasets. The discriminative training system is depends on new methods for partially labeled information. For data-mining examples the author combines margin sensitive approach called SVM.

Drawbacks

Less accuracy

Only Single object detection



## Proposed Work

RCNN got major success for regions based object detection but it has occlusion problem where object near to one and other get ignored and this problem can be solved by DPM (Deformable Part Based Model) technique which can hold information about all objects and then passed that information to CNN for near object detection. In propose paper author combining CNN with DPM to detect multi nearer objects.

Advantages

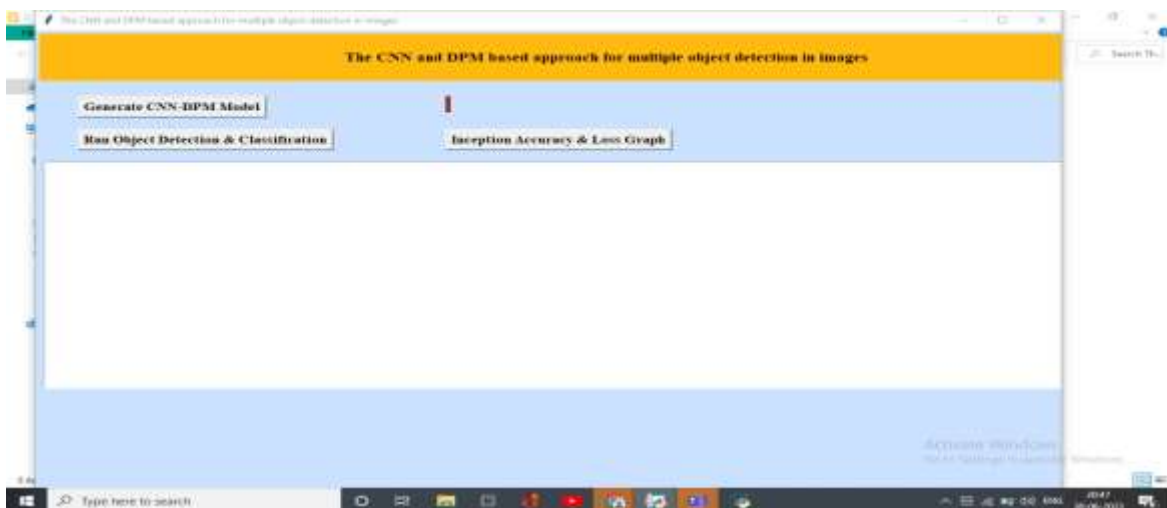
Better Prediction

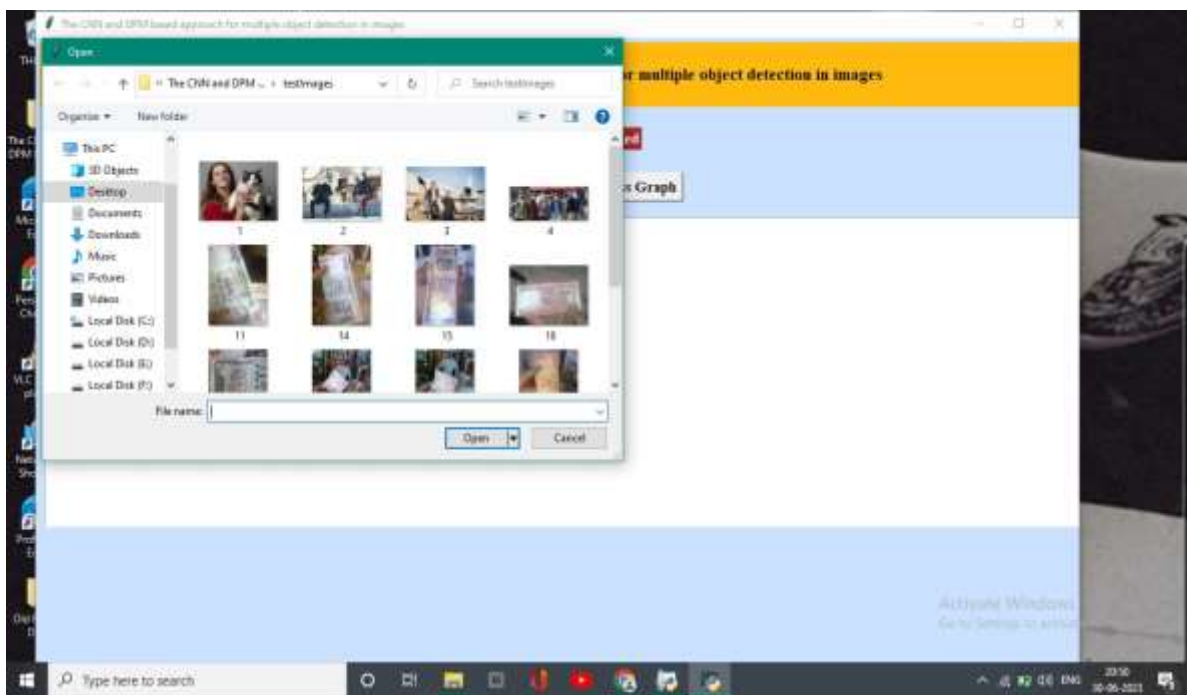
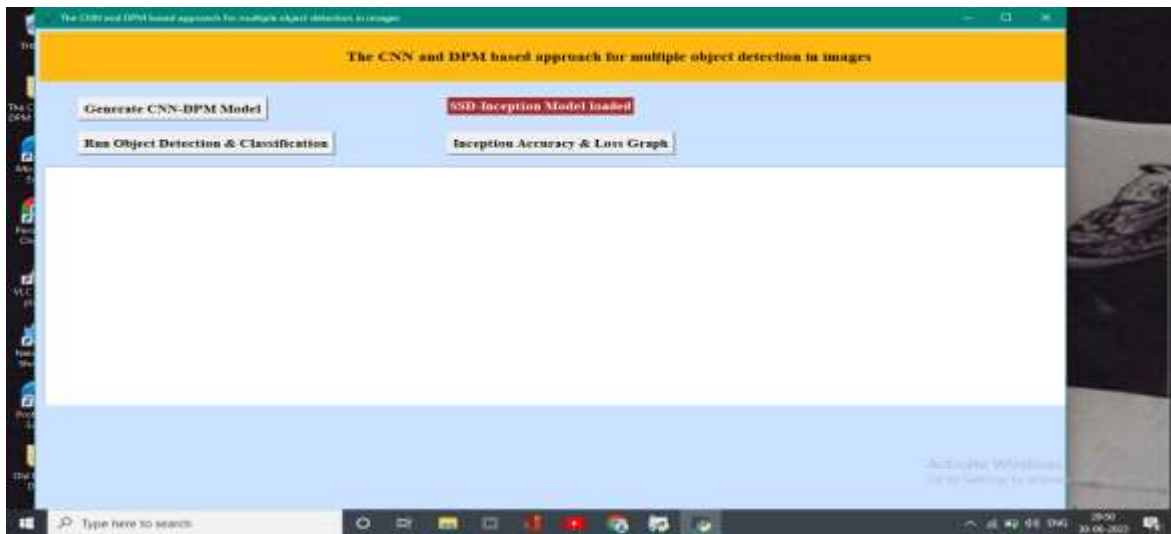
Multiple Object detection

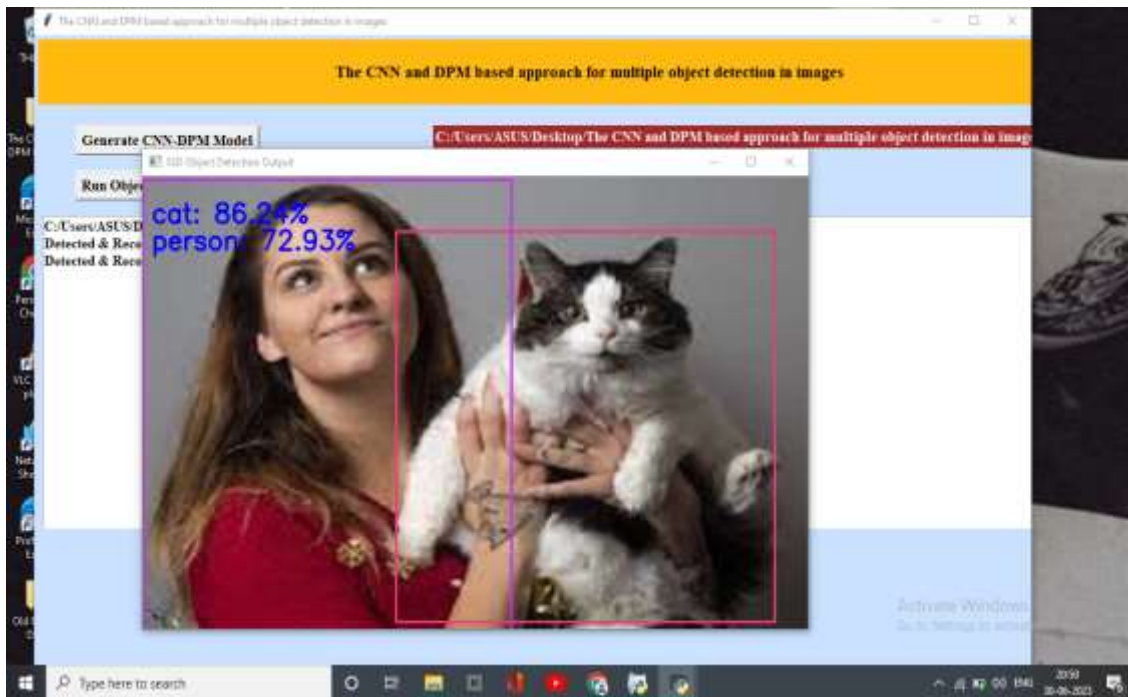
## 4. IMPLEMENTATION

- 1) Upload Pascal-VOC Dataset: using this module we will upload dataset to application
- 2) Generate & Load CNN-DPM model: using this module we will read all images from dataset and then train CNN-DPM model
- 3) Run Multi-object Detection: using this module we will upload test image and then application will detect multi objects and then surround it with bounding boxes.

### 4.1 RESULTS:







## 5. CONCLUSION

In this paper, we propose an integrate DPM and R-CNN techniques we get better result for multiple object detection. If we are using the DPM and R-CNN separately then it will not give more accuracy. In this framework we can use the DPM which can be used to generate the proposal that contains the part of the object for handling the occlusion and R-CNN is used for region based feature extraction. The proposals that are achieved from DPM and the proposals that are achieved from the SSA will be one by one send to the CNN model for extracting the features. When the CNN model is done with their process the DSD filter will clean the proposals generated from DPM. The proposals that are detected for every single object are outputted as an individual object.



## REFERENCES

- [1] F. Jabeen, A. Muhammad, and A. M. Enriquez, "Feed forward neural network training based interactive shopping for blind," in 2015 12th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). IEEE, 2015, pp. 1– 6.
- [2] R. Tapu, B. Mocanu, A. Bursuc, and T. Zaharia, "A smartphonebased obstacle detection and classification system for assisting visually impaired people," in Proceedings of the IEEE International Conference on Computer Vision Workshops, 2013, pp. 444– 451.
- [3] R. Tapu, B. Mocanu, and E. Tapu, "A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments," in 2014 11th international symposium on electronics and telecommunications (ISETC). IEEE, 2014, pp. 1– 4.
- [4] J. Monteiro, J. P. Aires, R. Granada, R. C. Barros, and F. Meneguzzi, "Virtual guide dog: An application to support visually-impaired people through deep convolutional neural networks," in 2017 International Joint Conference on Neural Networks (IJCNN). IEEE, 2017, pp. 2267– 2274.
- [5] R. Kumar and S. Meher, "A novel method for visually impaired using object recognition," in 2015 International Conference on Communications and Signal Processing (ICCSP). IEEE, 2015, pp. 0772– 0776.
- [6] R. Tapu, B. Mocanu, and T. Zaharia, "Seeing without sight-an automatic cognition system dedicated to blind and visually impaired people," in Proceedings of the IEEE International Conference on Computer Vision Workshops, 2017, pp. 1452– 1459.
- [7] B.-S. Lin, C.-C. Lee, and P.-Y. Chiang, "Simple smartphone-based guiding system for visually impaired people," *Sensors*, vol. 17, no. 6, p. 1371, 2017.
- [8] K. Potdar, C. D. Pai, and S. Akolkar, "A convolutional neural network based live object recognition system as blind aid," arXiv preprint arXiv:1811.10399, 2018.
- [9] J.-C. Ying, C.-Y. Li, G.-W. Wu, J.-X. Li, W.-J. Chen, and D.-L. Yang, "A deep learning approach to sensory navigation device for blind guidance," in 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS). IEEE, 2018, pp. 1195– 1200.



- [10] H. Alhichri, Y. Bazi, N. Alajlan, and B. Bin Jdira, “ Helping the visually impaired see via image multi-labeling based on squeezeNetCNN,” *Applied Sciences*, vol. 9, no. 21, p. 4656, 2019.
- [11] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “ Ssd: Single shot multibox detector,” in *European conference on computer vision*. Springer, 2016, pp. 21– 37.
- [12] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “ The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results,” <http://www.pascalnetwork.org/challenges/VOC/voc2007/workshop/index.html>.
- [13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “ Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818– 2826.
- [14] G. B. Huang, M. Mattar, H. Lee, and E. Learned-Miller, “ Learning to align from scratch,” in *NIPS*, 2012.
- [15] N. A. Semary, S. M. Fadl, M. S. Essa, and A. F. Gad, “ Currency recognition system for visually impaired: Egyptian banknote as a study case,” in *2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA)*. IEEE, 2019.