International Journal of Interdisciplinary Innovative Research &Development (IJIIRD)

ISSN: 2456-236X

Vol. 04 Issue 02 |2020

Household Object Detection using Neural Networks

Kundan Kumar[1,] Divya Kiran[2]

[1,2] *Master student, Department of MCA, Jain Deemed-to-be University, Bengaluru, Karnataka, India*

ABSTRACT

Household object detection is a new technology in computer which is connected to image processing and computer vision that deal with detecting object which is in house. The camera will detect all object which is store in kitchen, room etc. the object detection is low-end device algorithms to detect individuals in video or image. It us lose relationship with image and video analysis.

Keywords- object detection model, neural network, deep learning, python

1. INTRODUCTION

The topic of household object detection has been finding with in history many times, but each approach deals with a different method. The implementation of machine learning in this area uses a lot of mathematical calculations. Each image in the video feed needs to be further divided into pixels and analyzed in details. This project is equipped with image detection algorithms in order to make lives of elderly people easier. In place of manual calculation, different frameworks are used like Regression Based Framework, Region Proposal Based framework [1]. The model needs to be given the input. Further processing is done using deep neural networks. In this, there are more than one layer in the network helps increase the accuracy of the output. People at their old age face difficulty in visually recognizing object and they need a 24*7 human assistance. So, this system aims to give the remote in their hands. The remote will identifies the object and speaks out the name of object and old age people and eyeless people can easily get find the object. As the aerial of technology has made advancement, object detection is playing a very important role. The power of machine to recognize object just like a human does, can be used in a variety of domains. This project deals with one such domain, object detection using video and image for home assistance and also eye tracking system.

The Dalal-Triggs detector [2], which won the 2006 PASCAL object detection problem, used a standard gradient histogram filter (SOG) function to describe a group of objects [3]. This sensor uses a moving window technique, where a filter is added to any image or video location. We may think of the sensor as just a compiler that accepts an object, a position within the same image and a scale as feedback. The classifier decides whether at the specified location and scale and instance of the target group occurs or not. It has several uses in real life, for example the IoT[4].

2. HOUSEHOLD OBJECT DETECTION METHODS

Household object detection is a part of old age person which is detect all object in house. below are different methods given [5]: -

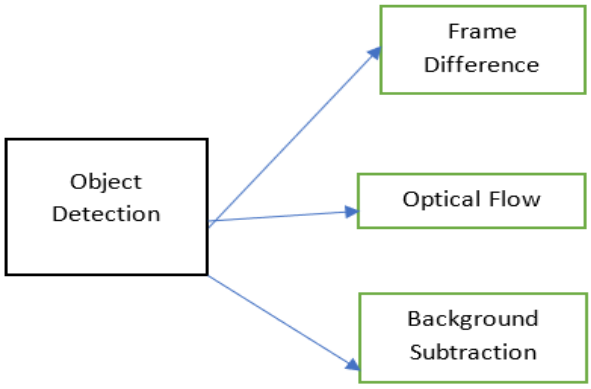

Fig. 1 Object detection methods

2.1 Frame Difference

Detection of moving object from a series of frames taken from a static camera is commonly achieved by means of frame gap. The frame differential approach is the common movement detection approach. This approach adopts variations dependent on pixels to locate the moving object[6-9].

2.2 Optical flow

On the image plane, optical flow is defined as the apparent movement of individual pixels. Sometimes it serves as a decent approximation to the actual physical motion projected onto the plane of the film. Many optical flow estimation methods assume that a pixel's color/intensity under the displacement from one video frame to the next is invariant [10].

2.3 Background Subtraction

Context subtraction is a common technique to separate the moving parts of a picture by segmenting it into the context and the foreground for an example, the object deducted from either the background series. The motion of person movement can be readily observed [11].

The background subtraction is divided into sub-methods. They all are given below: -

2.3.1 Recursive

This approach requires little use of a buffer to approximate the context. Then it uses an input element to algorithm update a standard context construct. It means objects from such a long ago will influence the image in the past. Relatively speaking, recursive methods require less computing space but a context error will last longer [12].

Recursive method approaches that are widely used include mixture of gaussian, Kalman filtering and conditional median filtering [13].

2.3.2 Non-recursive

This procedure uses a moving for techniques for estimating contact by saving a buffer from previous L video frames. It is estimating the history by evaluating the temporal variance of a pixel in the buffer [12].

3. LITERATURE REVIEW

A literature review or narrative review is a type of review article. A literature reviews is a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic.

3.1 Md Zahangir Alom1 et al, Chris Yakopcic1 et al, Stefan Westberg et al"Optimization Techniques for Artificial Neural Networks", IEEE 2007.

In such an un-supervised training area, this group suggested model, neural network would outperform human performance in tasks such as speech recognition, image recognition, predicting. We can render any standard model capable of solving multiple tasks in various application domains [14].

Advantage

Illustratively, the paper describes gradient descent, back-propagation, and Stochastic Gradient Descent principles for transmitting data to specific neural models.Methods of optimization include SGD, Ad grad, Ad Delta, RM Sprop, and Adam which help to evaluate the learning rate.Intuitions from this paper help to realize that a broad dataset makes for more accurate tests [14].

Disadvantage

Larger weights are not stored, and the parameter shifting results.These boundaries often limit proper back-propagation which eventually leads to reduced accuracy [14].

3.2 Karen Simonyan Andrew Zisserman, "Introduction to Localization in ConvNet", IEEE 2011.

This paper came up with substantially better reliable ConvNet architectures later on, which not only achieve region-of - the-art precision on ILSVRC identification and optimization tasks, but are also relevant to other object recognition datasets, where outstanding efficiency is achieved even if used as part of a fairly simple pipeline [15].

Advantage
The teaching scale is dynamically calculated by either single-or multi-teaching.The model itself will change the scale for the latter by controlling the jittering out of data.Fully linked network is implemented at the end, which improves the parameter by having a more robust output [16].

Disadvantage
As the fully connected layer has expanded parameters, the operation of the fully connected layer in certain situations allows output to be affected by the device threshold.

3.3 Discussion of Existing System

- Early work on object recognition was based on strategies for the matching of templates and basic partly based models.Methods were later adopted, based on statistics.This original popular family of object detectors, all focused on mathematical classifiers, laid the foundations for most study interms of preparation, measurement and classification techniques [17].
- Object recognition is a vital function for any device that communicates with humans; it is the most popular computer vision feature.Many external identification problems were studied, however.Most instances refer to objects in which humans often communicate, such as other person and body arts, such as ears, hands, and arms, as well as vehicles, such as cars, aircraft, and animals [18].

3.3.1 Disadvantage of existing system

- Most object recognition systems consider the same simple technique, generally known as sliding window: an exhaustive search is performed in order to identify the objects that appear in the image at various sizes and locations.This quest requires use of a classifier, the detector's central component, indicating whether or not a given image patch corresponds to the target.
- Since the classifier essentially operates at a specified scale and patch size, many iterations of the input image are created at various sizes, and the classifier is used to identify all potential patches of the same size, with each version of the image downscaled.
- They cannot accommodate well the case of two cases of the object being next to each other, and may not be sufficient to find the object.

4. PROPOSED SYSTEM

Household object detection system aims at achieving high accuracy in categorising the common household items for assistance in movement and other basic activities of visually impaired individuals. The system focuses on non-gpu deployment of detection algorithms for efficient use and easy access. The system at hand enables following.

- Real time identification of artefacts on video stream.
- Performance rating dependent on predefined degree of trust.
- Increase in processing speed by decreased ram intensive process reliance

4.1 Advantage of proposed system

- The developed algorithm works properly and is able to detect an item in households.
- Estimating the identification of items from individual is finished.
- Max target detection distance from camera mounted on computer was a max of 2 metres.
- Objects such as glass, table, human, books, dog, chair etc. are identified and the user is alerted by speech created that tells about the name of the recognized object.

5. ALGORITHM USED

The algorithm is a well-defined method that helps a computer to solve a problem. A series of unambiguous instructions is another way of defining the algorithm.Use the word 'unambiguous' is symbolic of no room for contextual interpretation.When you ask your machine to run the same algorithm, with the exact same outcome, it will do so precisely the same way. They are given below: -

5.1 Softmax Function

We use convex analysis and monotone operator theory results to obtain additional softmax function properties that are not yet addressed in the current literature.In particular we show that the softmax function is the log-sum-exp function's monotonous gradient map.By making use of this relation, we show that the inverse temperature parameter defines the Lipschitz and Softmax function co-obligation property [18].

ISSN: 2456-236X

Vol. 04 Issue 02 |2020

Softmax feature measures the distribution of the event's probability over various occurrences 'n.' In general terms, this equation would determine the probability of increasing target class for all other target groups. The estimated probabilities for deciding the target class for the specified inputs will be helpful later.

The principal benefit of utilizing Softmax is the spectrum of potential probabilities. The spectrum would be from 0 to 1, and the sum of all odds equals one. If the softmax method used for the model of multi-classification returns the probabilities of each class and the high likelihood of the goal class.

The formula computes the exponential (e-power) of the given input value and the of all the values in the inputs. Then the ratio of the exponential of the input value and the sum of exponential values is the output of the softmax function. The above graph in (Fig 3.1) is a graph which is used to show relation between input value and softmax score.

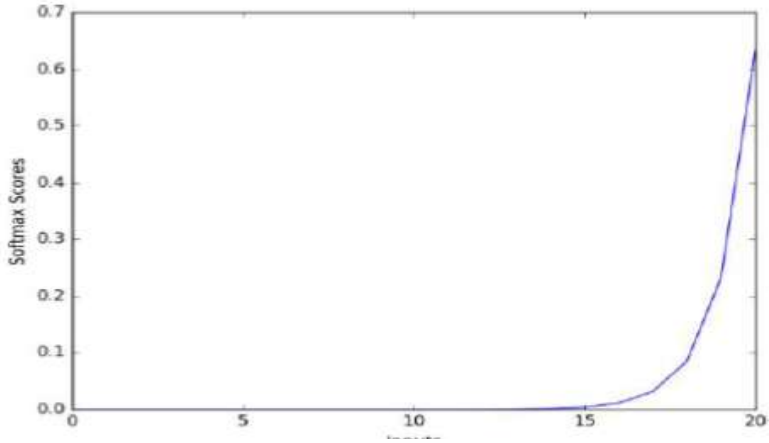

Fig.2. Graph of softmax function input and output

5.2 CNN Classifier

The co-evolutionary neural network (CNN) proposed by LeCun et al. is the most common type of deep neuralnetwork in use for machine vision issues[19-20]. One of the major initial attempts to use CNN for action recognition was by Baccouche et al. in[21]. A 3D coevolutionary neural network is trained in this work to allocate a vector of functions to a small number of consecutive frames. A recurrent neural network makes use of the spatio-temporal evolution of these characteristics for classification. We must create a CNN in this post, capable of classifying pictures.

A 3D coevolutionary neural network is trained in this work to allocate a vector of functions to a small number of consecutive frames. A recurrent neural network makes use of the spatio-temporal evolution of these characteristics for classification[22].

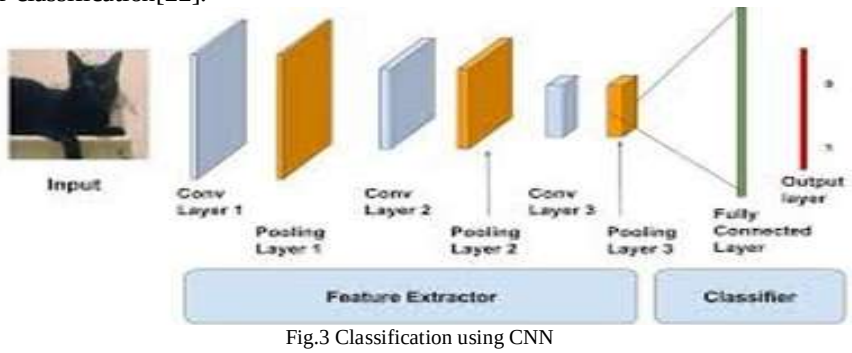

Fig.3 Classification using CNN

6. RESULT AND DISCUSSION

Color is an integral attribute that defines the image quality and, as seen in Figure 4, colors that appear in the photographs can be described successfully in image classification.The collection of the amount of the quantization rates in the color classification is an essential matter.

- At the other hand, the more precise multilevel color classification can be accomplished when color representations are merged using various quantization rates.The classification is obtained by merging separate base classifiers, using image histograms as their inputs at various stages of quantization.

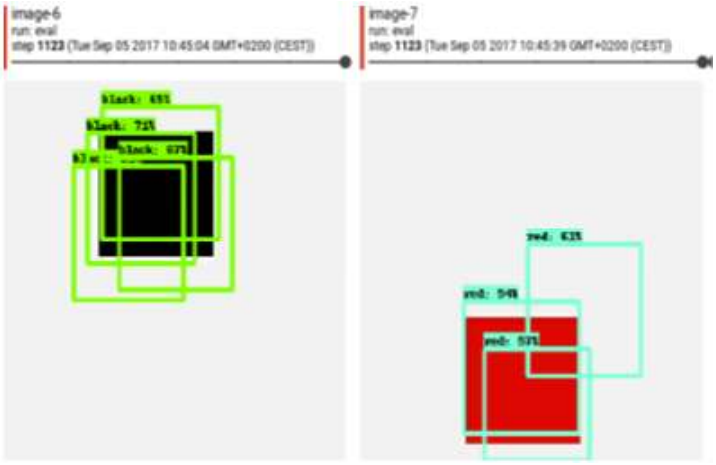

Fig 4Base classification for colour identification

- The mobile phone object class displayed in Figure 7.2 identifies a pattern as cell phones dependent on rectangular pixel density in the foreground.

Fig. 5 Mobile object classification

- Although with many interacting artifacts the algorithm can quickly distinguish the dominant groups with ease for these situations the most important aspect concentrated here is the distance between bounding box.The algorithm will differentiate between the class "chair" for the example seen above in Fig. 6

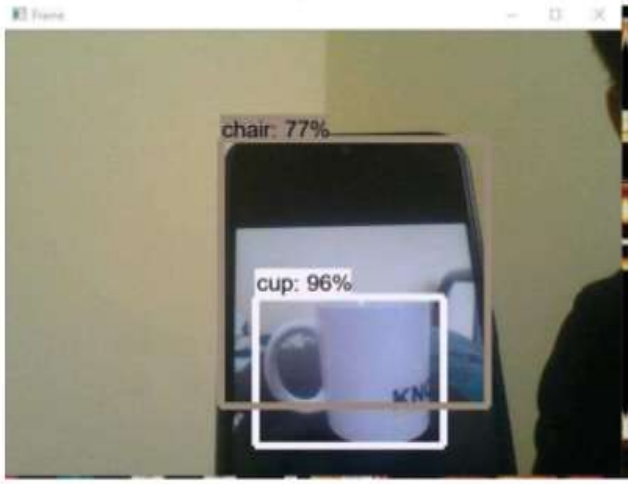

Fig 6. Cup object detection (box within box)

7. CONCULATION

Analyzing various approaches, it is deduced that it is much more practical and reliable to use deep learning rather than traditional machine learning strategies for target recognition.Implementation of machine learning involves a lot of mathematical equations which are boring to a computer program.Implementing profound coevolutionary neural networks cuts computations by a significant amount.Successfully the generated system recognizes basic objects such as container, table, human, bottle, device etc.Humans use all kinds of artifacts in their everyday lives.This initiative also reduces the expense of processing by using CNN as well as precision.

8. ACKNOWLEDGEMENT

I would like to express my profound gratitude to professor Dr. MN Nachappa and Prof Subarna Panda, for their patient, encouragement and valuable assessments of this research work. I appreciate his willingness to generously contribute time.

9. REFERENCES

[1] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," arXiv:1312.6229, 2013
[2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in IEEE Conference on Computer Vision and Pattern Recognition, 2005.
[3] pedro F. Felzenszwalb, Rose B. Girshick, David McAllester and Deva Ramanan, Object Detection with Discriminatively Trained part Based Models.
A. Behura and H. Mohapatra, "lot based smart city with vecular safety monitoring", EasyChair, Tech Rep, 2019.
[4] Himani S. Parekh, Darshak G. Thakore, Udesang K. Jaliya,"A Survey on Object Detection and Tracking Methods",2014.
[5] ZhengjieWang, YuanZhao, JifenZhang, YinjingGuo, Research on Motion Detection of Video Surveillance System, 3 rd International Congress on Image and SignalProcessing, vol.1, pp. 193-197, October 2010.
[6] K.Kavitha, A.Tejaswini, Background Detection and Subtraction for Image Sequences in Video,International Journal of Computer Science and Information Technologies, vol.3, Issue 5, pp. 5223-5226, 2012.
[7] M. H. Ali, Fadhan Hafiz, A. A Shafie, Motion Detection Techniques using Optical Flow, World Academy of Science, Engineering and Technology, vol. 32, pp 559-561, 2009.
[8] WeiShuigen, ChenZhen, DongHua, Motion Detection Based on Temporal Difference Method and Optical Flowfield, second International Symposium on Electronic Commerce and Security, pp 85-89, May 2009.
[9] Pavan Turaga,Ashok Veeraraghavan, in Advances in Computers, 2010.
[10] Pavan Turaga, Ashok Veeraraghavan, in Advances in Computers, 2010.
[11] S Cheung Sen-Ching and Chandrika Kamath. Robust techniques for background subtraction in urban traffic video. In Visual Communications and Image Processing 2004, volume 5308, pages 881–893. International Society for Optics and Photonics, 2004.
[12] Kirubaraj Ragland and P Tharcis. A survey on object detection, classification and tracking methods. Int. J. Eng. Res. Technol, 3:622–628, 2014.
[13] Md Zahangir Alom1 et al, Chris Yakopcic1 et al, Stefan Westberg et al "Optimization Techniques for Artificial Neural Networks ", IEEE 2007.
[14] Karen Simonyan Andrew Zisserman, "Introduction to Localization in ConvNet", IEEE 2011.
[15] C. Farabet, B. Martini, P. Akselrod, S. Talay, Y. LeCun, and E. Culurciello. Hardwareaccelerated convolutional neural networks for synthetic vision systems. In Proc. International Symposium on Circuits and Systems (ISCAS'10), 2010.
[16] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In IJCAI, 2011.
[17] Bolin Gao, Lacra Pavel, On the Properties of the Softmax Function with Application in Game Theory and Reinforcement Learning, 21 Aug 2018.
[18] Y. LeCun, K. Kavukcuoglu, C. Farabet, Convolutional networks and applications in vision, in: Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS), 2010, pp. 253–256. http://dx.doi.org/10.1109/ISCAS.2010. 5537907.
[19] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[20] M. Baccouche, F. Mamalet, C. Wolf, C. Garcia, A. Baskurt, Sequential deep learning for human action recognition, In: Proceedings of the Second International Conference on Human BehaviorUnterstanding, HBU'11, SpringerVerlag, Berlin, Heidelberg, 2011.
[21] S. Ji, W. Xu, M. Yang, K. Yu, 3d convolutional neural networks for human action recognition, IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 35 (1) (2013).