MAT JOURNALS
Journal of Image Processing and Artificial Intelligence
e-ISSN: 2581-3803
Volume 5 Issue 2

Survey on Object Detection Algorithms using Neural Networks

Vitthal Bohra[1], Atharva Lad[1], Parth Sagar[2]
[1]UG Students, [2]Professor
[1,2]Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, Maharashtra, India
Email: *vbohra0@gmail.com*
DOI: *http://doi.org/10.5281/zenodo.2667077*

Abstract

In day to day life, we encounter several different types of things like people, things, animals etc., but it becomes difficult for a system to identify them when processing them in images. For such identification purpose, we use Object Detection algorithms. These algorithms can be used for various uses like defence systems, security management units, and other fields like healthcare. In the following study, various Object Detection algorithms such as facial recognition, facial feature detection like skin, colour etc., and target detection have been performed and used to detect various types of objects with increased accuracy and efficiently in the fields like surveillance. Further, different difficulties and uses of Object Detection strategies are expounded. Pictures have numerous objects in complex foundation, how to distinguish these objects, the principle objects in that and comprehend the connection between them two that is, the primary objects and different objects are the significant focal point of this study. There are numerous ways which can be utilized for object detection and acknowledgment; however a large portion of them can't check the primary objects of the picture. In the accompanying study, we are playing out a review on the different picture capturing strategies for object detection.

Keywords: *Deep learning, image capturing, object detection.*

INTRODUCTION

Object detection and recognition are not difficult for humans. Humans can easily understand the main objects of the image and the meaning of the image, but the same is a very tedious task for a machine i.e., a computer. Usually, there are more objects in the complex background. It is difficult for a computer to identify the main objects and understand the relationship between the main objects and the other insignificant objects. Object detection mainly deals with identification of real-world objects such as people, animals and objects of suspense or threatening objects. Object detection algorithms use a wide range of image processing applications for extracting the objects desired portion.

The main purpose of this research is to contribute to the existing methodologies by giving a new perspective in which the process of Object detection can be improved upon. In the following survey paper different methodologies that are in use for object detection have been studied. The positives and the areas in which they lack also have been highlighted.

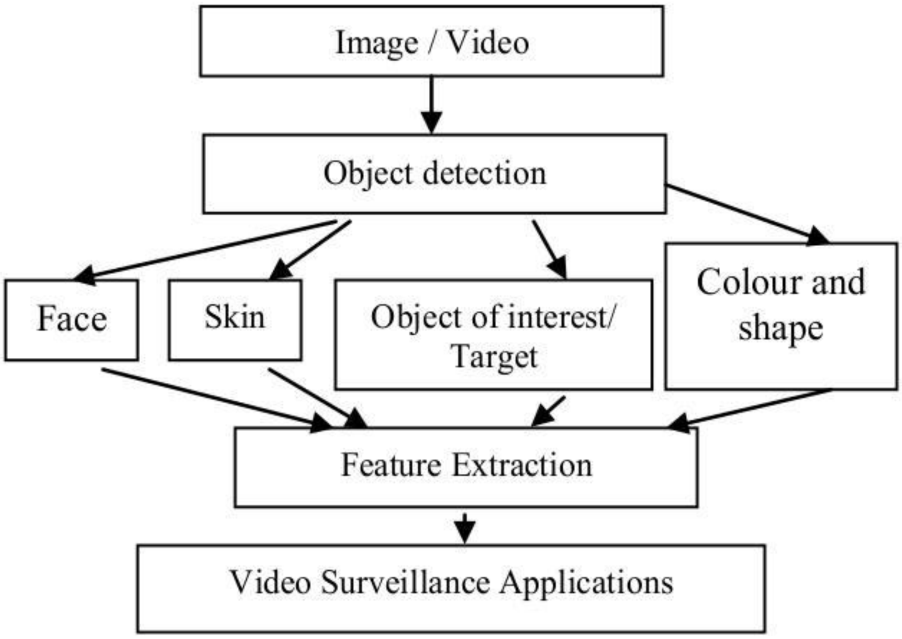

Figure 1: *Basic block diagram of object detection process.*

Fig. 1 shows the Basic block diagram of object detection process. Frames are extracted from image or video. Objects are detected based on user's desired choice such as face, skin, colour, target of interest. Further various features of object are extracted for video surveillance applications.

PROBLEM STATEMENT

Our problem is that the existing systems work using classifiers which are unable to perform detection at fast speed hence, making it really difficult to use the same in real time. This study proposes to use Neural Networks instead of classifiers in combination with processing using GPU so as to make it usable for real-time detection making it more practical for use in several applications like face detection, defence and machine learning.

DESIGN OF IMAGE MAIN OBJECTS SCORING SYSTEM

There are many different objects in complex images and it is difficult to distinguish these objects i.e., to identify if it is a main object or other object. For example, two images in Fig. 2, we can easily discern the main object of the left image is the girl, and the other object is the red road sign. So, the score of the girl is higher than the red road sign. The main object in the right image is the elderly; the other object is the hat, so the score for the elderly is higher than the hat. Usually, the main objects of the image have two main features. First, the ratio of the objects volume to the image. For example, the volume of the girl and the elderly in Fig. 2 is larger than the other objects, so they are the main objects. The second is the rarity of all the objects in the image. If there is a person and a panda in the image, the panda can be considered as the main object of the image. So for the main objects in the image this section will propose a set of criteria for scoring.

Figure 2: *An example of image scoring system.*

FUNDAMENTALS OF OBJECT DETECTION

Object detection is a method of identifying a main object in a frame with respect to other background objects. The ideal object could be an individual, a creature or some other object or point of interest.

Foreground Object: A foreground object is that different object which can be identified distinctively from the unimportant objects. It could be done so by observing it movement in relation to the other objects or in general by its appearance as it will mostly change with every frame.

Background Object: Stationary objects in a frame which are part of the background are called background objects.

In the next diagram that is Fig. 3, we can see the basic structure of the background subtraction process. This process basically used to differentiate between the main object and other objects and is performed frame by frame. Next step involves the background update process in which the current frame is compared with the previous frame to detect the object. We consider various objects and parameters like living, non-living things, animals, furniture, chairs, tables, human beings, multiple objects.

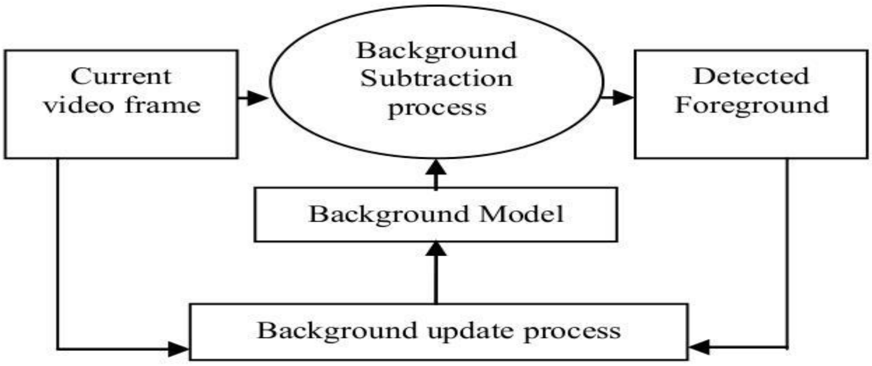

Figure 3: *Block diagram of background subtraction process.*

The following diagram i.e., Fig. 4 explains the relation between two of the most common color spaces i.e., the YCbCr colour space model and RGB colour space model. The outer cube is used to show the YCbCr colour space and the inner one for RGB. For converting the color space from normal RGB to YCbCrcolor space the formula used is as follows:

Y=0.299R + 0.587G + 0.114B Cr = R - Y Cb = B - Y

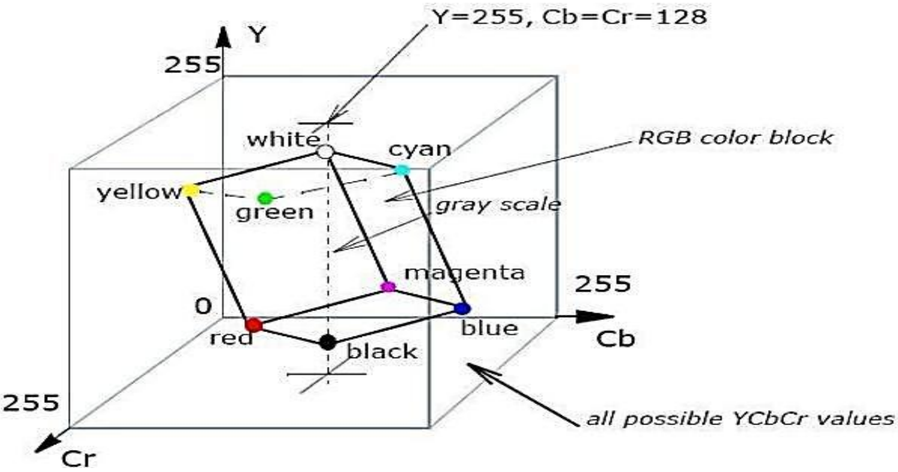

Figure 4: *The RGB and YCbCr model.*

OBJECT DETECTION ALGORITHMS

This section helps us to identify the available algorithms that are used for object detection. An object can be categorized based on detection of various features like face, skin and colour.

Face Detection

Facial recognition or face detection is a method used to identify individual human faces. Viola Jones algorithm is one such method that can be used to detect human faces by determining specific features like eyes, nose and mouth. Hence, it can also be used to identify various parts using the same method.

Skin Detection

Skin detection as we can understand is a method used to detect the existence of skin. It is basically done by identifying the skin as pixels like that of an image. These pixels are represented by 1 and 0 where 1 signifies presence and the non-skin areas are specified by 0 respectively. Some of the major colour spaces that are used in Skin detection are as follows:

1. RGB
2. Orthogonal
3. Perpetual
4. Perpetually uniform

Skin detection is performed using the YCbCr Model. The translation of RGB into YCbCr colour space mainly involves separation of luminance from chrominance. Therefore, the model does not change with variation of illumination. Here, Y represents Luminance and Cb and Cr indicate the Chromi-nance parameters.

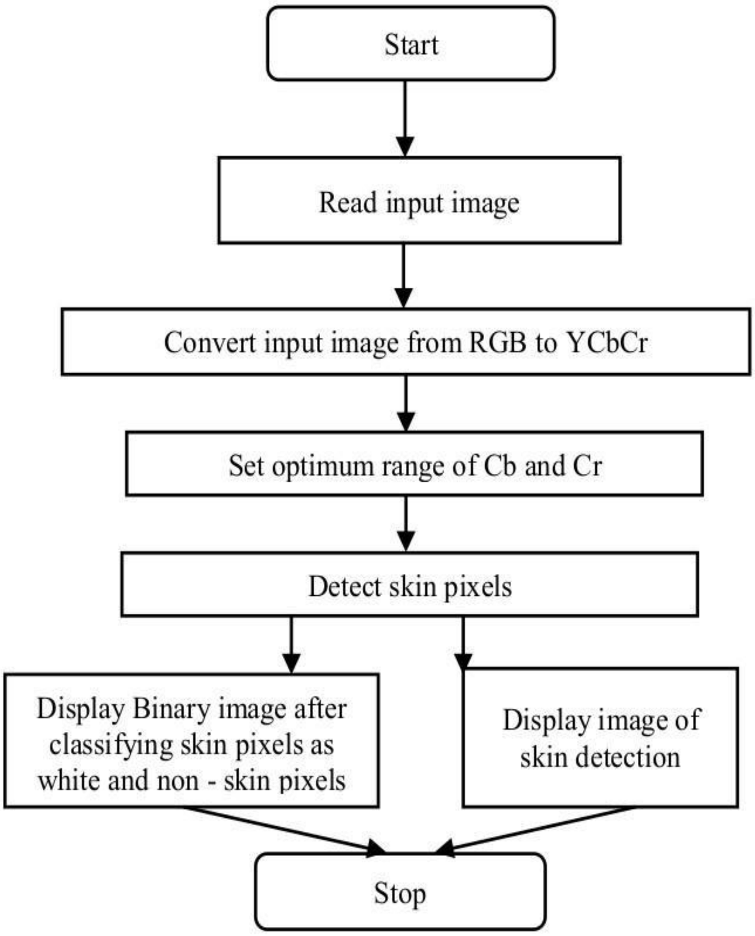

Figure 5: *Flow chart for skin detection.*

Target Detection

Target detection algorithm is basically used for main object detection, what that means is we can detect target object using this method. As we saw before, background subtraction is one of the most used method for performing target detection.

Background subtraction from a stationary camera: The image is captured from a stationary camera. The posterior probability is computed using Bayes rule.

$$p(\beta / xt) = (p(xt/\beta) . p(\beta))/ p(xt) \quad (4)$$
$$p(F / xt) = (p(xt/F) . p(F))/ p(xt) \quad (5)$$

Where (F/xt) and p(/xt) indicate probabilities of observing xt in the foreground and background models, respectively. p(F) and p() are probabilities of the pixels belonging to the foreground and background respectively. If equal priorities are assumed, then expression can be denoted by the likelihood ratio.

$$p(xt/\beta)/ \ p(xt/F) \quad (6)$$

Colour Detection

Fig. 5 shows the flow chart of skin detection. First step is to read the image and convert RGB to YCbCr model. Second step is adjusting the values of Cb and Cr to detect skin pixels accurately. In the third step, detection of skin pixels is done. Finally, the output image i.e., the image of skin detection is displayed, in this image the skin pixels are represented by 1 and non-skin pixels are represented by 0 (binary format).

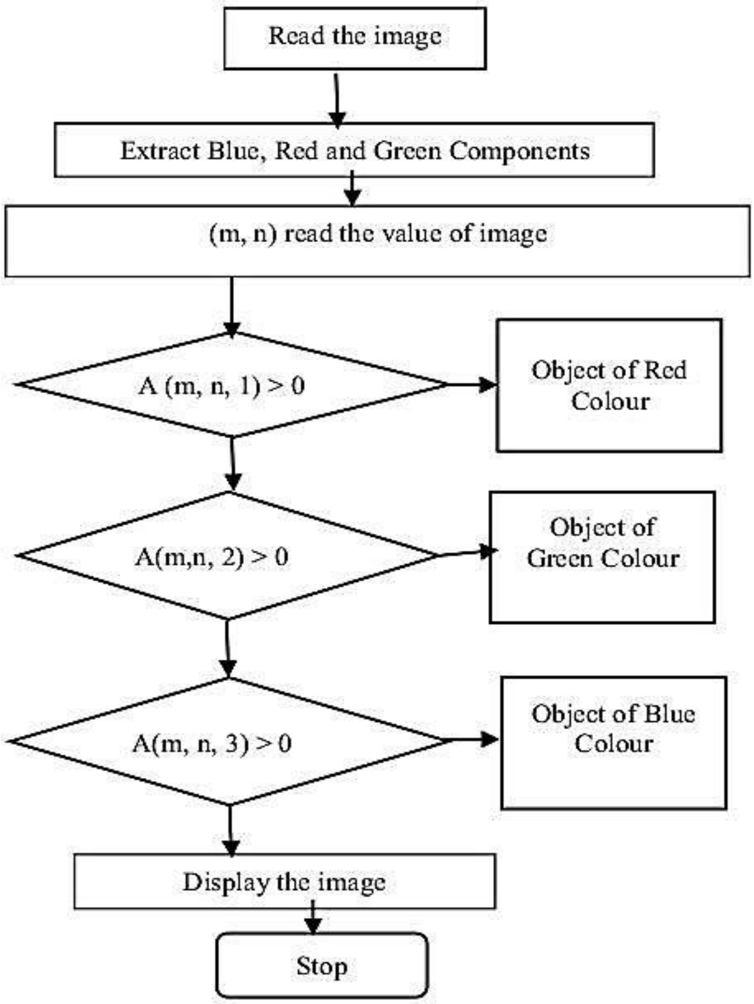

Figure 6: *Flowchart for colour detection.*

The above Fig. 6 shows the flow of the colour detection algorithm. In the first step, we accept and process the image. In the next step, the Gray scale image is converted into a binary image by comparing the value of each pixel with a preset threshold value. If the threshold value is less than thevalue of the pixel then, it is assigned 1 or else it is given a value of 0. If the value of pixel is assigned as 1 then the colour is black or else if the value is 0 then it is assigned white colour. Colour is a very important parameter that is used in object recognition. The lighting on the objects which differs across the scene is a trivial segment to be taken into consideration when selecting colour models. Impact of robustness with the structure of an object, blockages and messy images also play an important role.

OBJECT DETECTION ALGORITHMS CHALLENGES AND APPLICATIONS

This section elaborates challenges and applications of vari-ous object detection algorithms.

Challenges

The challenges or the problems identified in each method namely the skin detection algorithm, the colour detection algorithm, W4, background subtraction, mean shift algorithm, Kalman filtering, and LIBS have been highlighted ahead. Like when using LIBS, the model fails when dynamic objects are present in the background and hence gives inaccurate results. If there are small changes in the background like the waving of leaves or any subtle changes that may occur in the background. In colour detection algorithm, the algorithms in use are useful just for primary colours and detects them accurately but fails for new colours like it will detect cyan as blue. In W4, only people in upright position can be detected using the cardboard model. If people are in different poses, or are crawling and climbing, it becomes challenging. In Behavior Subtraction, detection of spatial anomalies like U-turns is challenging in this method. If it is necessary in detecting outliers, only the ones that are spatially localized and temporal can be detected. Behavior camouflage takes place especially when there is a foreground object during background activity. In Kalman Filtering, MSA and GMM when the is obstruction they fail to detect multiple images and at the same time if there are various objects in the image that can be identified the existing algorithms used for skin detection fail resulting in being useless to detect skin.

Other than the above challenges from these, some of the basic issues faced are like, if there is any change in lighting on the other objects they could be wrongly detected as main or foreground object. Similar main object and other objects can also result in the effect of camouflaging which makes it difficult to differentiate between objects creating confusion of camouflage. Whereas some algorithms dont have the ability to differentiate between light and detect a shadow. Detection gets tougher in situations where the background is dynamic. And even creates problem when there are a lot of objects in a single frame that overlap each other. Therefore, it makes it difficult to classify a fixed main object and the background due to continuous change.

Applications

Object detection finds scope in various areas such as defense and border security, medical image processing, video surveillance, astronomy and other security related applications.

CONCLUSION

The conclusion is that the various object detection algorithms such as skin detection, colour detection, face detection and target detection are simulated using

MATLAB 2017b with an accuracy of approximately 95 percent. Parameters like the accuracy of detection of object, the YCbCr colour model in skin detection, threshold value in target detection using RGB Euclidean threshold T have been tested and then used so as to result in a better efficiency of the algorithms majorly for video surveillance applications. Further, a single algorithm maybe designed by considering various detection parameters such as colour, face, skin and target of interest to meet video surveillance applications.

ACKNOWLEDGMENT
The authors would like to thank and express profound gratitude to Prof.ParthSagar for his valuable support, encouragement, supervision and useful suggestions throughout this project work. His moral support and continuous guidance enabled to complete the work successfully. His regular suggestions made the work easy and proficient. Last but not the least, the authors are thankful and indebted to all those who helped directly or indirectly in completion of this work.

REFERENCES

1. LiyanYu,Xianqiao Chen, Sansan Zhou (2018), "Research of Image Main Objects Detection Algorithm Based on Deep Learning", *3rd IEEE International Conference on Image, Vision and Computing*
2. ApoorvaRaghunandan, Mohana, PakalaRaghav, H. V. Ravish Aradhya (April 3-5, 2018), "Object Detection Algorithms for Video Surveillance Applications", *International Conference on Communication and Signal Processing,*
3. K. K. Hati, P. K. Sa, B. Majhi (Aug. 2013), "Intensity Range Based Background Subtraction for Effective Object Detection", *IEEE Signal Processing Letters,* Volume 20, Issue 8, pp. 759−762.
4. I. Haritaoglu, D. Harwood, L. S. Davis (1998), "W4: Who? When? Where? What? A real time system for detecting and tracking people", *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara*, pp. 222−227.
5. D. H. Santosh, P. G. K. Mohan (2014), "Multiple objects tracking using Extended Kalman Filter, GMM and Mean Shift Algorithm - A comparative study", *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram,* pp. 1484−1488.
6. J. C. Nascimento, J. S. Marques (Aug. 2006), "Performance evaluation of object detection algorithms for video surveillance", *IEEE Transactions on Multimedia,* Volume 8, Issue 4, pp. 761−774.
7. H. Fradi, J. L. Dugelay (2012), "Robust foreground segmentation using improved Gaussian Mixture Model and optical flow", *2012 International Conference on Informatics, Electronics Vision (ICIEV), Dhaka*, pp. 248−253.
8. R. Girshick, J. Donahue, T. Darrell, et al (2014), "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation" *Computer Vision and Pattern Recognition. IEEE,* pp. 580−587.
9. N. Dalal, B. Triggs (2005), "Histograms of oriented gradients for human de-tection," *Computer Vision and Pattern Recognition, CVPR 2005. IEEE Computer Society Conference on. IEEE,* pp. 886−893.
10. P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan (Sep. 2010), "Object detection with discriminatively trained part based models", *IEEE Trans Pattern Analysis and Machine Intelligence,* pp. 1627−1645.
11. R. Girshick (2015), "Fast R-CNN", *IEEE Internation Conference on*

Computer Visio. IEEE Computer Socity, pp. 1440−1448.
12. S. Ren, K. He, R. Girshick, et al. (2017). "Faster R-CNN: Towards real-time object detection with region proposal networks", *IEEE Transactions on Pattern Analysis Machine Intelligence*, pp. 1137−1149.
13. J. Redmon, S. Divvala, R Girshick, et al. (June 27, 2016), "You only look once: Unified, real-time object detection", *Computer Vision and Pattern Recognition*, pp. 779−788.
14. J. Uijlings, K. van de Sande, T. Gevers (2013), "Selective search for object recognition", *International Journal of Computer Vision*, pp. 154−171.
15. K. Simonyan, A. Zisserman (Sep. 2014), "Very deep convolutional networks for large-scale image recognition", *Computer Vision and Pattern Recognition*.

Cite this article as: *Vitthal Bohra, Atharva Lad, & Parth Sagar. (2019). Survey on Object Detection Algorithms using Neural Networks. Journal of Image Processing and Artificial Intelligence, 5(2), 1–8. http://doi.org/10.5281/zenodo.2667077*

