International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

Harmful Object Detection Using Deep Learning

R. Sravani[1], D. Venkata Shiva Reddy[2] B.E., MBA, M.Tech, (Ph.D.)
[1]Department Of Computer Science, Besant theosophical College, Madanapalli, Andhra Pradesh, India
[2]Assistant Professor, Head of Department of Computer Science, Besant Theosophical College, Madanapalli, Andhra Pradesh, India

Article Info
Volume 9, Issue 2
Page Number: 116-121

Publication Issue :
May-June-2021

Article History
Accepted : 10 June 2021
Published: 23 June 2021

ABSTRACT

Mobile networks and binary neural networks are the most commonly used techniques for modern deep learning models to perform a variety of tasks on embedded systems. In this paper, we develop a technique to identify an object considering the deep learning pre-trained model Mobile-Net for Single Shot Multi-Box Detector (SSD). This algorithm is used for real-time detection, and for webcam feed to detect the purpose webcam which detects the object in a video stream Therefore, a module to track the objects in the video stream is used. We combine the Mobile-Net and the SSD system to deploy the module in a rapid and effective, thorough learning method of object detection. Our research focuses on improving the precision of the SSD object detection procedure and on the importance of Mobile-Net's pre-trained deep learning model. This increases the accuracy of conduct identification at a speed necessary for real-time detection and the standards of indoor and outdoor day-to-day surveillance.

Keywords: - Video Processing, Image, Classification, Computer Vision

I. INTRODUCTION

Now, object identification is widely used in many areas such as video monitoring, foot show, slander detection, self-conducting vehicles, and the identification of appearances. Deep Learning field, the sub-discipline which is called Object Detection, includes an image such as a photo, video, or webcam feed. Deep learning has shown considerable application in research into computer vision. Deep convolution neural networks have various models in which the computer-view subjects are learned to essentially classify images, detect objects, recreate images, assess object poses, learn events, and so on. CNNs rely mainly on the fact that they can acquire substantial mid-scale image features rather than hand crafted low-level representations that are commonly found in special methods of image classification. Instead of The successful success of object recognition is the result. The key features that include the shape, scale, color, texture and other features define an item in this issue.

In order to identify this object, an image clearly shows the existence of the object and, moreover, the position of it is seen in the image. The identification of objects may thus be described as a way to find real-world object instances in images. Detection is closely connected with the classification as it is about saying a certain object's presence and placement in an image. Various objects, for instance car detection, football, houses, road markings and human faces, etc, can be found in a photograph. This will lead to more efficient monitoring and security mechanisms for the detection of moving objects from video by enhancing object detection accuracy, robustness and efficacy by deeper learning methods, including deep neural networks, regionally convolutionary neural networks, and profoundly convolutive neural networks.

This is especially important in order to trace security risks like intruders in a sensitive environment, to locate lost artefacts which may be abnormalities in a scene such as bombs or weapons and to track vehicles of theft. Moreover, smart visual cameras can be used for tracking animals' activity and behavior in protected areas, both for ethological and for the protection of our natural world, through inspiration from deep learning in object sensing. The use of profound study algorithms for target detection has also been an important tool for medical imaging and cancer cell detection in the human body Object detection is one of the machine vision activities benefitting in many articles in literature from the Deep Learning techniques.

This paper discusses the deep learning algorithms and methods used to detect objects in fixed and video files. It includes a comprehensive analysis of deep-technology and its uses in picture detection. Furthermore, the precision function and efficiency of deep neural networks in target recognition over conventional machine learning techniques has been clearly demonstrated. It also incorporates traditional deep learning methods for the investigation of images and objects, and provides the most impressive faming's of recent years; In recent years, many notable and groundbreaking approaches have been employed to improve detection accuracy of deep learning models and to resolve difficult problems faced by deep learning target detection algorithm training and testing processes. The alteration of the activation mechanism of CNNs, transfer learning and intelligent approaches for the combined collection and the optimization method of the proposed deep learning model is among these groundbreaking techniques.

The following document is organized: Section 2 presents the context in the identification of objects and traditional methods of machine learning applied to them.the common DL and its analytical methods for object detection and offers an overview of DL models for solving advanced object detection problems.

View of the machine. Several groundbreaking strategies have been developed for improving and optimizing profound learning models and solving some of the problems during training and research.

II. RELATED WORKS

BACKGROUND REVIEW

Object Detection:

Color also provides important knowledge about the environment. Objects in photographs are different in colour and can be differentiated from the background based on their colour. As a consequence, sometimes objects are separated in different colours from a scene. The technique is also used to perform this task directly classifying the pixels into 'objects and context.' An object will almost definitely be marked with or outlined by a series of colours or co-ordinate it. The backdrop may therefore be represented by way of its characteristic colours as the other picture values or in a similar way.

This technique is used in the process of filtering colours from the other elements and context of a given object. By the Direct Pixel Classification method Cyganek applied a 'Road Signs Detection' strategy. Classification techniques are good for reducing dimensionality and can be applied to rapid image preprocessing. Furthermore, features other than colour may be included.There are two pixel-based methods commonly used for the identification of road signs. These techniques include extracting samples manually from such photographs seen in live scenes of traffic . A large number of techniques for refining object detection specific to categories have previously been established. Traditionally the most popular and commonly accepted methods for pedestrian detection are the histogram of oriented gradients-based detectors using the multiscal sliding window system.There has obviously been a good detection quality for pedestrian detection using filtered channel features. In order to provide insights into the common errors of such approaches and explain the implications of training efficiency data, a detailed study was performed on the performance of cutting edge pedestrian detectors. These observations have been used for the analysis and refinement of sophisticated techniques like the filtered channel characteristics and the R-CNN detectors.One approach is to construct a fuzzy colour histogram classifier. A few thousand samples have been taken to establish their colour histograms for each of the distinctive colours that belong to a category of traffic signs. It is easy to apply and requires less computing time, since this approach is useful. However, the flush method also reveals a significant proportion of false positive and therefore low precision.

The Support Vector Machines is a pixel-based method (SVM). Vapnik , under the Structural Risk Minimisation (SRM) process, submitted support vector machines. SVMs are extremely generalised and have considerable potential for regularisation.

They are known as massive binary classifiers.One of the basic low vision functions of the machine is the detection of regular figures, namely, arcs, circles, polygoms, etc. This numbers can be represented by mathematics parametrically. Hough invented the widely esteemed procedure for the identification of forms as a means of voting for line recognition. The system was eventually expanded to Ballard's random detection of figures. The technique is computationally costly when typical forms are detected. The use of the tensor therefore significantly improved the identification of regular types.Another approach to identifying features from the images is used in figure recognition. On the basis of their defining points, objects can be decided. This approach is based on the dynamically advancing sparse image coding region. The description concerns the identification of distinguishing points which are the most intrinsic to an object, the possible geometrical transition, noise and other disturbances of the object's view. HOG, and SIFT are among the well-established point descriptors with their many variants, including OpponentS1FT and PCA-SIFT

Different investigators compared the efficiency of sparse descriptors, for instance Mikolajczyk, and Schmid, which carried out the SIFT procedure in many cases, and submitted a paper on the Gradient Position and Orientation Histograms. But these methods are disadvantageous since only small-scale information of any statistic are taken into account. Functions which can collect mid-level data such as edge intersections or high-level representations such as object parts are more difficult to create.

This is evident even when distinctive image characteristics are extracted from invariant key points

III. METHODOLOGY

SSDs are two-way: a backbone and an SSD head. In general, the spinal model is a network of pre-trained image classifications. We normally use a Mobile-Net network here which trains more than a million images which are removed from the associated ranking layer. The effect is described as groups of cases and artefacts linked to the dimensional position where the closure of the layer of activation is kept.

SSD Head This waist has only one or more fixed layers.

Some Parameters in SSD

Grating/Grid Cell

Detecting objects means estimating the class and location of an element directly in this area. For example (see Figure 3) we use the 4x4 grid. Individually, the grating focuses on creating the space and the shape of the space that suits it.

Anchor Box

Grid cells with separate anchors or prefixes are uniquely assigned to SSD. These anchor boxes will monitor any form and size in each grid cell. The cats are similar to the different anchor boxes, one high anchor box, while the other box is broader and thus of different sizes. These anchor boxes with a wealth of intersection through an object complete the class and its place

Zoom Level

No anchor boxes must be as large as the cells in the grid. It is used with grid cells to identify to what degree the anchor box independently requires posterity up or down.

Aspect Ratio

Any shaped artefacts are broader and some with different grades are longer. The SSD frame makes the anchor box aspect ratio to it. The proportion range is used to characterise the variable aspect ratios of anchor box connections

Receptive Field

As a region seen by a specific CNN, the appropriate field input area is segregated. Zeiler and Rob[21] used the characteristics and the actuate to represent them at the corresponding position as a backline mix. The properties of the different layers, because of the compromised activity, indicate the different area in the picture. Place the lowest layer (5x5) and then a console in which a single green pixel reflects 3x3 area with a single input layer (3x3) is the central layer (bottom layer).

The convolution is then added to the middle layer(green), with the upper red layer (2x2), in which each attribute corresponds to the 7x7 region of the picture.Related map features have a similar area in which the similar patterns are identified in different locations in return. A Convolutionary Network local level is thus created.

IV. DEEP LEARNING TECHNIQUES

Deep learning methods:

Different methods were developed to enhance the identification of objects in the field of deep learning. Profound knowledge Belief Networks include a two-layer, bipartisan undirected graphic model with full communication among the visible units and individual units of the basic hidden layer, the restricted band Boltzmann machine (RBM). Networks A Stack Auto-Encoder (SAE) is a two-layer neural network stack that is programmed to remodel its own inputs.

It is a model which records an automatic fountain of regulation, without the need for any algorithm to support it. The aim of CNNs is to dctcnnine filters using a data-driven approach in order to extract input description functionality. Deep learning seeks to meet the problems facing functional approaches, such as mid- and high-level knowledge learning. This is achieved naturally by the learning of visual characteristics directly from input, both in supervised and uncontrolled techniques.

Many profound learning methods are usually grouped in three categories for target detection. In the first group, the functional education is uncontrolled and only functions are extracted using the philosophy and the deep learning principle. These features are then provided with incredibly simple machine learning techniques, depending on the process, for tasks like grouping, identification or monitoring. The second is regulated education methods, which are used to jointly refine the extractor and the classifier units of the whole model by including significant mark data steps.The third classification is hybrid deep networks, which include the use of génerative functional learning models to improve formation of deep neural networks and the preparation of other deeply supervised approaches through regularising optimization strategies more efficiently.

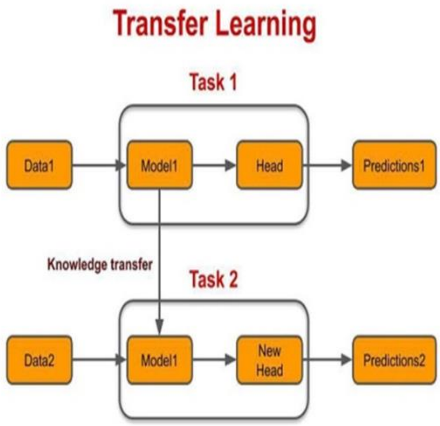
By using Mobile-Net and the SSD detector for object detection, a high-precision object detection process is achieved, this will push the processor speed to 14 fps and make the operation effective for those who can process often just 6 fps. In this scheme, objects such as a vehicle, cycles, bottles, chairs can be detected inside the dataset. The dataset can be extended by using photos killed by adding an infinite number of objects.

For experiments with the SSD algorithm, we used Ubuntu 18.04.2, OpenCV 3.4.2 and Python. The aim of this research is to create an autonomous system that allows the community to make the system engaging and attractive by recognising objects and scenes. This work would specifically be used to classify the item in the external world with best features for potential work.

Detection Contains Three Steps

1-Using OpenCV's deep neural network (DNN) module to load a pre-trained object detection network. 2-Set of input to the network and compute the forward pass for the input. Storing the result as Detections. 3-Then, loop through the Detection and determine what and where the objects are in the images.

the location in which the individual, the site and the precise detection are in the frame of a given object. A personal picture and then the mark that is already present in the pre-trained model for the specific detection. A mark is named the individual and the label is precise . Via these three measures, these three steps retrained and used a loop and drew a boundscape around this certain entity.



Files for Model

These files are the setup of our pretrained versions, and the second is the weights. In fact, the model is how neurons in a neural network are organised.

V. EXPERIMENTAL RESULTS

The detection algorithm for this item is up to 14 fps, which can provide good results in poor quality cameras of any fps. We consider a webcam of 6 fps in this case. In our tests, indoor and outdoore feed video frames are demonstrated by the SSD algorithm by webcam, but artefacts vary from two consecutive frames. The video and algorithms taken by the webcam transform the size of a single frame to 300 to 300. By building the bounded box around the detected object, SSD can detect frame by frame artefacts with the precision of a class marker.

The findings obtained from this treatment in the home-made video

The SSD is able, with the help of a higher per cent default cases, to generate several bond boxes for various groups with a different degree of trust . Based on the frame gap, this suggested single-shot multi-box sensing system. The efficiency of the proposed approach was analysed by Frames. Determining the accuracy and strength of the proposed approach is tested by foggy weather conditions.

VI. CONCLUSION

The MobileNet and SSD detectives for object detection have been used to obtain a highly precise object detection technique that can shift a processing speed to 14 fps, and make it effective for any camera that can process at 6 fps only once. In its data collection this machine is able to identify objects such as a motorcycle, bicycle, bottle and chair. By inserting a number of things via pictures,

The aim of this research is to create an autonomous system that allows the community to make the system engaging and attractive by recognising objects and scenes. This work would specifically be used to classify the item in the external world with best features for potential work. known as "deep learning technology," a dataset can be extended

VII. REFERENCES

[1]. Hong, Y. C., Chung, Y. S, An Enhanced HybridMobileNet 2018 9th International Conference onAwareness Science and Technology (iCAST)

[2]. Shraddha, M., Supriya, M. Moving object detection and tracking using convolutional neural networks IEEE Xplore ISBN:978-1-5386-2842-3

[3]. Ojala. T., Matti. P., Topi, M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis & Machine Intelligence, no. 7, pp. 971–987, 2002

[4]. J. Sun and Y. Fang, "Cross-Domain Data Sharing in Distributed Electronic Health Record Systems," in IEEE Transactions on Parallel and Distributed Systems, vol. 21, no. 6, pp. 754-764, June 2010.

[5]. Hu. M.K., Visual pattern recognition by moment invariants, IRE transactions on information theory, vol. 8, no. 2, pp. 179–187, 1962.

[6]. Khotanzad. A., Yaw. H. H. H. Invariant image recognition by Zernike moments. IEEE Transactions on pattern analysis and machine intelligence, vol. 12, no. 5, pp. 489–497, 1990.

[7]. Huang. J, Kumar S. R., Mitra. M, Zhu. W. J., Zabih. R. Image indexing using color correlograms. incvpr. IEEE, 1997, p. 762.

[8]. J. Han, W. Susilo, Y. Mu, J. Zhou and M. H. A. Au, "Improving Privacy and Security in Decentralized Ciphertext-Policy Attribute-Based Encryption"

[9]. Dalal, N., Triggs, B. Histograms of oriented gradients for human detection. An International conference on computer vision & Pattern Recognition (CVPR'05), vol. 1. IEEE Computer Society, 2005, pp. 886–893

[10].Han. F., Shan, Y., Cekander R., Sawhney, H.S., Kumar R, A two-stage approach to people and vehicle detection with hog-based SVM in Performance Metrics for Intelligent Systems 2006 Workshop, 2006, pp. 133–140.

[11].D.Venkata Siva Reddy, Deep Intelligent Prediction Network: A Novel Deep Learning based Predictions model on SpatiotemporalCharacteristicsandLocation based Services for Big data driven Intelligent IJTEE 8(11),5,0.