JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Review of Real Time Object Classification Using CNN Algorithm

Prashant Kumar Tripathi, Ms. Namrata Sahayam

Department Of Electronics and Telecommunication Engineering, Jabalpur Engineering College M.P.

Abstract-

The field of artificial intelligence is built on object detection techniques. YOU ONLY LOOK ONCE (YOLO) algorithm and it's more evolved versions are briefly described in this research survey. This survey is all about YOLO and convolution neural networks (CNN) in the direction of real time object detection. YOLO does generalized object representation more effectively without precision losses than other object detection models. CNN architecture models have the ability to eliminate highlights and identify objects in any given image. When implemented appropriately, CNN models can address issues like deformity diagnosis, creating educational or instructive application, etc. This article reached at number of observations and perspective findings through the analysis. Also it provides support for the focused visual information and feature extraction in the financial and other industries, highlights the method of target detection and feature selection, and briefly describe the development process of yolo algorithm.

Keywords- Object Detection, YOLO, CNN.

1. INTRODUCTION

The primary goal of Object Detection is to detect all of the items in a given image. The main point in object detection is to teach the system how to detect the object by training the dataset and letting the system to detect the object on its own. In a real-time environment, the searching or recognizing procedure is extremely difficult. So far, no viable remedy has been discovered. Despite the fact that there is a lot of work being done, the approaches that have been created so far are inefficient, need a long training period, are not ideal for real-time application, and are not scalable to huge numbers of classes.

When the machine is seeking for a specific thing to detect, object detection is rather simple. Recognizing all of the things, on the other hand, necessitates the ability to distinguish one object from another, even if they are of the same type. Such a task is extremely tough for robots to solve if they are unaware of the numerous possibilities of objects.

When someone visits a hospital, a traffic control monitor, or a self-driving car, all they do is detect items for the optimum performance. In a hospital, for example, the doctor will be able to detect whether or not there is a virus in our body based on the scan report. As a result, this work decided to create a REAL TIME OBJECT DETECTION System. It can be difficult to choose between several Object detection approaches at times. This work assign a class name to an image in image classification, and build bounding boxes around objects in object localization. Detecting the item is difficult since combine both duties, including both the bounding box around each item and then combine them. Rest of the paper is organized as section II Review motivated.

2. LITERATURE REVIEW

Ruiyang Xia et al. [1] In this letter, present a real-time detector for FSOD called Bi-path Combination You Only Look Once (BC-YOLO) to maintain a high inferring speed and equivalent detection precision. BC-YOLO is a one-stage object detector that uses transfer learning and has a two-phase training method. It is made up of bi-path parallel detection branches that detect base and novel class objects separately and usually detect objects with a discriminator in the inferring step. Furthermore, this work present an Attentive Drop Block technique to make the detector focus on the whole details of objects rather than the small discriminative regions in order to improve model generalisation trained from few-shot objects.

Punam Sunil Raskar et al. [2] This research provides a novel method for detecting Copy-Move attacks in passively blind films. The quick and real-time object detector "You Only Look Once -Version 2": YOLO (V2) is used to implement the object-based forgery detection technique. With a confidence score of 0.99, the system is trained on the benchmark dataset movies for the automatic detection of forged objects within the film. The trained YOLO (V2) model correctly classifies and localizes forged and non-forged items in the given input video. The results and experimental analysis show that the suggested YOLO (V2) model performed admirably in identifying simple and complicated Copy-Move attacks such as scaling, rotation, and flipping. The performance is superior for object-based forgery detection in terms of speed and accuracy when compared to other state-of-the-art deep learning algorithms.

Xiaohong Han et al. [3] Object detection is critical in Automatic Driving Systems (ADS) and Driver Assistance Systems (DAS). However, present real-time detection methods for small vehicle objects have limited precision and poor performance. To address these challenges, this work offer Optimized You Only Look Once Version 2 (O-YOLO-v2), a unique real-time object detection model for micro vehicle objects based on the You Only Look Once Version 2 (YOLO-v2) deep learning framework.

Zuxiang Situ et al. [4] Object detection is essential in Automatic Driving Systems (ADS) and Driver Assistance Systems (DAS). Nonetheless, present real-time detection algorithms for microscopic vehicle objects suffer from low precision and poor performance. To address these challenges, this work offer Optimized You Only Look Once Version 2 (O-YOLO-v2), a unique real-time object detection model based on the You Only Look Once Version 2 (YOLO-v2) deep learning framework for micro vehicle objects. A new structure is implemented in the suggested model to increase the network's feature extraction ability by adding convolution layers at various spots. Meanwhile, adding residual modules solves the problem of gradient disappearance or dispersion produced by growing network depth.

Guanbo Wang et al. [5] Based on YOLO v4, this research proposes a real-time detection technique for CCTV autonomous weapons. This work propose the YOLO v4 backbone with Spatial Cross Stage Partial-ResNet (SCSP-ResNet) for the peculiarities of CCTV scenarios. Meanwhile, it is demonstrated that the receptive field improvement module can capture fine semantic aspects of high-dimensional tiny objects. To boost the model's perceptive power on the region of interest, the Fusion-PANet (F-PaNet) module was employed to fuse multi-scale information. Furthermore, combine synthetic and real-world datasets to examine the effects of synthetic datasets on detectors in depth. Our proposed detection model enhances mAP (mean Accuracy Precision) and inference time by 7.37% and 4.2%, respectively, according to the experimental data. The parameter of the model has been lowered by 0.349 BFLOP/s (billion floating point operations per second).

Jinjie Zhou et al. [6] In this research, propose a new object detection technique for infrared (IR) images called YOLO-CIR, which is based on YOLO and ConvNext. present an augmentation approach for infrared images and improve the pre-processing algorithm without bit-width compression to support high-bit-width infrared images. In addition, a multi-scale feature extraction network based on ConvNext was constructed to adapt low-resolution infrared images. Furthermore, the ConvNeXt block introduces the coordinate attention module to focus on targets and suppress the background, and the neck introduces a split attention module to improve feature fusion capabilities.

Shu-Jun Ji et al. [7] This study proposes MCS-YOLO v4, a small object detection technique based on YOLO v4 and Multi-scale Contextual information and Soft-CIOU loss function. To gather rich location information, MCS-YOLO v4 adds a scale detection to the existing three scales. MCS-YOLO v4 incorporates an enlarged field-of-perception block to improve the network's capacity to locate and classify objects. This block gets object contextual features and combines them with convolutional features to produce more robust and discriminative features.

Arunabha M. Roy et al. [8] Present WilDect-YOLO, an automated high-performance detection model based on deep learning (DL) for real-time endangered wildlife detection. In the model, this work incorporate a residual block in the CSPDarknet53 backbone for strong and discriminating deep spatial feature extraction, and combine DenseNet blocks to increase important feature information preservation.

Fengying Dang et al. [9] This study presents a new dataset (CottoWeedDet12) of weeds important to cotton production in the southern United States (U.S.); it includes 5648 images of 12 weed classes with a total of 9370 bounding box annotations, collected in cotton fields under natural light conditions and at various weed growth stages. On the dataset, a novel, comprehensive benchmark of 25 cutting-edge YOLO object detectors of seven versions, including YOLOv3, YOLOv4, Scaled-YOLOv4, YOLOR and YOLOv5, YOLOv6, and YOLOv7, has been built for weed detection.

Huayi Zhou et al. [10] this work provide a novel semi-supervised domain adaptive YOLO (SSDA-YOLO) technique for improving cross-domain detection performance by combining the compact one-stage stronger detector YOLOv5 with domain adaptation. In particular, modify the knowledge distillation framework with the Mean Teacher model to help the student model gain instance-level properties of the unlabeled target domain. Also use scene style transfer to generate fake images in other domains to compensate for image-level variations. To better align cross-domain predictions, an intuitive consistency loss is presented.

Zhiyang Zheng et al. [11] To solve the problem of missed detection and false detection caused by complicated environments in cow individual detection and tracking, a multi-object tracking method (YOLO-BYTE) is presented. The method enhances the YOLO v7 Backbone network feature extraction module by incorporating a Self-Attention and Convolution mixed module (ACmix) to account for the cows' uneven spatial distribution and target size fluctuation. Additionally, to reduce model complexity, an updated lightweight Spatial Pyramid Pooling Cross Stage Partial Connections (SPPCSPC-L) module is used to reduce the number of model parameters.

Chenchen Jiang et al. [12] This study offered a framework for UAV TIR object recognition in pictures and movies. The You Only Look Once (YOLO) models were created using Convolutional Neural Network (CNN) architecture to extract characteristics from ground-based TIR pictures and videos acquired by Forward-looking Infrared (FLIR) cameras. Finally, evaluation metrics determined the best successful algorithm, which was subsequently used to detect objects on TIR films captured by UAVs. In the validating task, the maximum mean average precision (mAP) of the human and car occurrences was 88.69%.

Qiwen Qiu et al. [13] The incorporation of You Only Look Once (YOLO) into an unmanned aerial vehicle (UAV) is presented in this paper to enable real-time crack detection on tiled sidewalks. YOLOv2 tiny, Darknet19-based YOLOv2, ResNet50-based YOLOv2, YOLOv3, and YOLOv4 tiny network designs are reframed and compared to improve detection accuracy and speed.

Thi-Thu-Huyen Vu et al. [14] In this research, present a method for developing a real-time packaging defect detection system based on deep learning techniques, with the goal of automatically detecting defective packed products in industrial package quality control. To be more specific, describe a real-time defect detection system that uses the YOLO (You Only Look Once) method to automatically classify product quality. The system can be connected into factories and manufacturing lines, assisting in the optimisation of efficiency and the reduction of operational costs.

Mehmet Emin Salman et al. [15] To find key regions (for the localization job) and grade the found regions (for the classification goal), the Yolo general-purpose object identification algorithm was used. Our prostate cancer dataset was used to retrain the system.

S.N.	Author name	Year	Method	Advantage	Disadvantage
1	Xia, R., Li, G., Huang, Z., Meng, H., & Pang, Y.	2023	Attentive DropBlock algorithm	Method can achieve a better tradeoff between speed and precision than state-of-the-art methods.	Low accuracy.
2	Situ, Z., Teng, S., Feng, W., Zhong, Q., Chen, G., Su, J., & Zhou, Q.	2023	CNNs, Resnet	Better sewer defect detection.	Low detection precision
3	] Zhou, J., Zhang, B., Yuan, X., Lian, C., Ji, L., Zhang, Q., & Yue, J.	2023	ConvNext Model	Enhance feature fusion ability.	Low bit-width compression
4	Ji, S. J., Ling, Q. H., & Han, F.	2023	MCS-YOLO v4 Model	Obtain more robust and discriminative features	Poor ability of network to locate
5	Roy, A. M., Bhaduri, J., Kumar, T., & Raj, K.	2023	WilDect-YOLO Model	Enhance receptive field representation And preserve fine-grain localized information.	Higher complex model

3. RESEARCH MOTIVATION

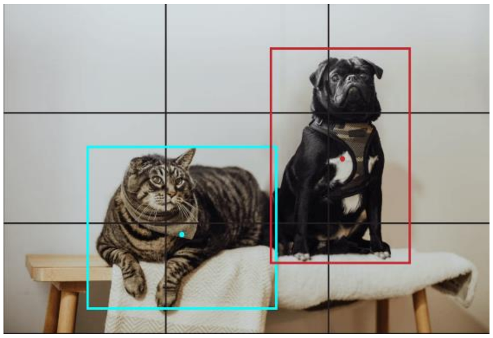
In the field of computer vision, advanced concepts such as neural networks and deep learning are gaining traction. In real time, the solution supplied by these strategies can be very adaptive and dependable. Traditionally, visually impaired persons navigate the outdoors using a white cane, which is of limited utility. To increase support, a sophisticated system is essential to assure safety and make the individual highly aware of his or her surroundings. Before can implement object detection and categories objects, must first understand the distinction between object detection and picture classification.

Image classification is the process of categorizing an image based on recognized features and patterns, whereas object detection is the process of determining the bounding box of coordinates precisely where a specific object is located in the image. In a picture, can recognize many objects of different classes. In brief, object detection can tell us not only what is in an image, but also where it is. There are various methods for detecting objects in images [2].

3.1 Yolo

You Only Look Once (YOLO) is one of the most popular model architectures and object detection algorithms. It uses one of the best neural network architectures to produce high accuracy and overall processing speed, which is the main reason for its popularity. If search Google for object detection algorithms, the first result will be related to the YOLO model.

YOLO algorithm aims to predict a class of an object and the bounding box that defines the object location on the input image.



YOLO stands for real-time object detection. It uses a single neural network to the entire image, separating it into sections and predicting bounding boxes and possibilities for each. These bounding boxes are weighted based on predicted probabilities. In a single evaluation, a single neural network predicts bounding boxes and class possibilities straight from entire images. Because the entire detection pipeline is a single network, detection performance may be optimized end-to-end.

3.2 YOLO algorithm is important

YOLO algorithm is important because of the following reasons:

- **Speed:** This algorithm improves the speed of detection because it can predict objects in real-time.
- **High accuracy:** YOLO is a predictive technique that provides accurate results with minimal background errors.
- **Learning capabilities:** The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection.

3.3 Design Disruption

Object detection is defined as determining where objects exist in a given image (object localization) and which category each object belongs to (object classification). As a result, the typical object detection model pipeline can be separated into three stages:

Informative region selection, feature extraction and classification.

Feature extraction.

Classification.

3.4 Convolutional neural network CNN

Object detection is a tool that recognizes semantic Objects in virtual photographs and films. Self-driving cars are undoubtedly one of its Real-time programs. Our task in this game is to find a couple of items in a photo. The most prevalent items found on this utility are the automobile, motorbike, and pedestrian. Object localization and should find many items in real-time structures for locating the items within the photograph. There are several approaches for item detection that can be classified. The first is Algorithms, which are mostly based on classifications. CNN and RNN fall into this category. Select the concerned Regions from the image and classify them using the Convolutional neural network community.

Instead, it anticipates the training and bounding containers of the entire image in a single run of the algorithm and detects a couple of gadgets using a single neural community. When compared to other classification algorithms, the Yolo Set of rules is faster. In real time, our algorithm technique produces 45 frames that are consistent with 2d. Although the Yolo algorithm commits localization errors, it predicts fewer false positives in the background [3]. Humans may easily come across and pick out devices of their surroundings without paying attention in their situations, regardless of what position they're in and whether or not they're the wrong way up, unique in shade or texture, partially obscured, and so on.

The same item identification and recognition on a computer requires a lot of processing to extract information about the objects and shapes in a picture. In computer vision, identifying an object in an image or video is referred to as object detection. The primary steps in object detection or future extraction, which is used for monitoring, cancer reduction, car identification, and underwater item detection, among other things. To accurately and correctly identify the object for specialized packages, various ways had been employed. However, these proposed methods entail inaccuracies and inefficiencies. Device learning and deep neural network techniques are more effective in addressing these object detection difficulties [4].

4. CONCLUSION AND FUTURE SCOPE

4.1 Conclusion-

This study provides an overview of several object detection, tracking, and recognition approaches, feature descriptors, and segmentation methods based on video frames and various tracking technologies. With new concepts, this approach was applied to improve object detection. In addition, the bibliography content includes tracking the item from video frames with theoretical explanation.

The bibliography content is the most significant contribution of research because it will lead to a new area of research. Now a days are seeing that object detections applications are becoming popular in most of the fields, so based on that perception we've developed the console based application which takes image as an input and returns the same image with detecting object names on the top of the bounding boxes drawn around the image. Because this is a custom trained data set, trained the dataset with Google Colab because this is supervised learning and labelling the data with the Labeling. So, in this case, we're using YOLO to get faster results, and we're using the most recent version.

4.2 Future use-

1. Create and simulate complex video sequences, then evaluate them using the same tracking technique. In the hypothetical case, occlusion is utilized for an object of the same colour for moving objects, or else larger occlusion with a longer occlusion time is used. Increasing the number of objects aids in determining the tracking algorithm's effectiveness and functioning.
2. Weight parameters must be added for each pixel's specific intensity level. If an intensity value is allocated as foreground in an image based on the current frame, there is a lower possibility that the foreground also has comparable pixel coordinates, therefore the BG weightage for the pixel is set to the minimal value rather than the starting value. By increasing weightage less than the original value, the old pixel value is removed with the least likelihood rather than the evolved scene.
3. The variance data of each channel should be improved using the distance calculation. This allows for a change in the quick scene via the Euclidean distance technique.

REFERENCE

[1] Xia, R., Li, G., Huang, Z., Meng, H., & Pang, Y. (2023). Bi-path combination YOLO for real-time few-shot object detection. *Pattern Recognition Letters*, *165*, 91-97.

[2] Raskar, P. S., & Shah, S. K. (2021). Real time object-based video forgery detection using YOLO (V2). *Forensic Science International*, *327*, 110979.

[3] Han, X., Chang, J., & Wang, K. (2021). Real-time object detection based on YOLO-v2 for tiny vehicle object. *Procedia Computer Science*, *183*, 61-72.

[4] Situ, Z., Teng, S., Feng, W., Zhong, Q., Chen, G., Su, J., & Zhou, Q. (2023). A transfer learning-based YOLO network for sewer defect detection in comparison to classic object detection methods. *Developments in the Built Environment*, 100191.

[5] Wang, G., Ding, H., Duan, M., Pu, Y., Yang, Z., & Li, H. (2022). Fighting against terrorism: A real-time CCTV autonomous weapons detection based on improved YOLO v4. *Digital Signal Processing*, *132*, 103790.

[6] Zhou, J., Zhang, B., Yuan, X., Lian, C., Ji, L., Zhang, Q., & Yue, J. (2023). YOLO-CIR: The network based on YOLO and ConvNeXt for infrared object detection. *Infrared Physics & Technology*, *131*, 104703.

[7] Ji, S. J., Ling, Q. H., & Han, F. (2023). An improved algorithm for small object detection based on YOLO v4 and multi-scale contextual information. *Computers and Electrical Engineering*, *105*, 108490.

[8] Roy, A. M., Bhaduri, J., Kumar, T., & Raj, K. (2023). WilDect-YOLO: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection. *Ecological Informatics*, *75*, 101919.

[9] Dang, F., Chen, D., Lu, Y., & Li, Z. (2023). YOLOWeeds: a novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems. *Computers and Electronics in Agriculture*, *205*, 107655.

[10] Zhou, H., Jiang, F., & Lu, H. (2023). SSDA-YOLO: Semi-supervised domain adaptive YOLO for cross-domain object detection. *Computer Vision and Image Understanding*, *229*, 103649.

[11] Zheng, Z., Li, J., & Qin, L. (2023). YOLO-BYTE: An efficient multi-object tracking algorithm for automatic monitoring of dairy cows. *Computers and Electronics in Agriculture*, *209*, 107857.

[12] Jiang, C., Ren, H., Ye, X., Zhu, J., Zeng, H., Nan, Y., ... & Huo, H. (2022). Object detection from UAV thermal infrared images and videos using YOLO models. *International Journal of Applied Earth Observation and Geoinformation*, *112*, 102912.

[13] Qiu, Q., & Lau, D. (2023). Real-time detection of cracks in tiled sidewalks using YOLO-based method applied to unmanned aerial vehicle (UAV) images. *Automation in Construction*, *147*, 104745.

[14] Pham, D. L., & Chang, T. W. (2023). A YOLO-based real-time packaging defect detection system. *Procedia Computer Science*, *217*, 886-894.

[15] Salman, M. E., Çakar, G. Ç., Azimjonov, J., Kösem, M., & Cedimoğlu, İ. H. (2022). Automated prostate cancer grading and diagnosis system using deep learning-based Yolo object detection algorithm. *Expert Systems with Applications*, *201*, 117148.