International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 3, May-Jun-2022, ISSN (Online): 2395-566X

Object Detection Deep Learning Using Yolo, Darknet

Dharun Kumar, Premnath. S
Dept. of Information Technology
Dr. Mahalingam College of Engineering and Technology,
Pollachi.

Abstract- Picture characterisation stands out enough in the field of computer vision to be noticed. Over the last few years, there has been a lot of research done on picture characterisation using traditional AI and deep learning techniques. Deep learning-based techniques have yielded astounding results thus far. Despite the fact that various deep learning-based methods have demonstrated excellent picture sorting performance, deep learning methods are unable to separate all significant data from the image due to a variety of difficulties. As a result, characterization precision was significantly reduced. The goal of the current study is to improve image classification performance by combining deep features extracted using the popular YOLO deep convolutional neural network. From the experiment, we achieved an accuracy of 94.51 percent.

Keywords- Deep Learning, Picture Characterisation, Computer Vision, Picture Sorting, Deep Features, Yolo Algorithm, Deep CNN.

I. INTRODUCTION

The fundamental examination subject in the field of Deep Learning and AI is picture characterisation. The goal of image classification is to accurately distinguish an object in a photograph. Local, global, or both highlights may be used for image classification. Deep learning is a subset of artificial intelligence. In the deep learning approach to characterise the pictures with comparable items, both component extraction and order are done sequentially.

There is no requirement for specialists to physically perform both errands, as in traditional AI. Deep learning makes use of a number of different neural layers to process a large amount of data. It's known as a Deep Convolutional Neural Network (DeepCNN) when used this way.

Nonetheless, due to the low resolution of the image, insufficient extraction of local and/or global features, geometric variation, and other factors, determining the correct class of a given item is extremely difficult. A managed learning approach is used in the paper; with highlight extraction performed using a pre-trained deep learning model (YOLO model). The YOLO model has recently demonstrated excellent image classification performance. In any case, the trial results demonstrate that even this model is insufficient for precise image grouping. As a

result, a model combination is insufficient for accurate image classification. As a result, there's a fusion of deep features (pre-trained YOLO). For the concentrate, a standard information dividing methodology is used, with 70% of the pictures from each class considered in the preparation dataset and the remaining 30% in the testing dataset.

The proposed combination highlight extraction framework achieves 94.51 percent recognition precision, outperforming many experts' methodologies. Precision (93.70 percent), Recall (93.73 percent), F1 score (93.22 percent), Area Under Curve (96.79 percent), False Positive Rate (0.15 percent), Root Mean Square Error (20.05 percent), and Average CPU Time are all presented as results of the proposed approach in the paper (0.39 min).

II. CHALLENGES

Deep Convolutional Neural Network (CNN) has uncovered massive results in computer vision in recent years. At the same time, the CNN model is proving difficult to implement for the analysts. Using a deep brain organisation, the proposed framework is used to determine the issues that have arisen for the picture arrangement task. The most difficult task is to create an organisational model. CNN has a lot of layers, which means there are a lot of boundaries to pick up during the preparation stage.

Planning a CNN model from scratch necessitates a number of resources for execution, such as a large memory limit, a fast processor, a large dataset, high power consumption, and so on. YOLOv1 is a model for locating a single stage object. Object recognition is described as a recurrence problem involving spatially isolated bouncing boxes and associated class probabilities.

In a single assessment, a single brain network predicts jumping boxes and class probabilities directly from full pictures. Because the entire recognition pipeline is managed by a single company, it tends to be streamlined from beginning to end. Directly on the performance of identification. [5] [6]

© 2022 IJSRET

International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 3, May-Jun-2022, ISSN (Online): 2395-566X

To anticipate each jumping box, the organisation uses highlights from the entire picture. It also predicts all jumping boxes for a picture across all classes at the same time. This means that the organisation thinks about the whole picture and each of the articles in it all over the world.

1. YOLOv3:
Darknet-52 is used in YOLOv3, which is 1.5 times faster than ResNet101. The portrayed precision and speed between Darknet spines also involves no compromise, as it is still as precise as ResNet-152 while being twice as fast. Furthermore, without the need for model retraining, you can easily compromise between speed and accuracy by changing the model's size. [8] [14]

2. YOLOv4:
YOLOv4 is a one-stage object discovery model that improves with the addition of a few secret stashes and modules mentioned in the text. The parts section beneath it hides the stunts and modules used.

3. Darknet:
Darknet is a neural network framework that is open source. It's a quick and extremely precise method (precision for specially prepared models is dependent on data preparation, ages, and clump size, and a few other components) system for continuous article recognition (also works with images)[4]. Because it is written in C and CUDA, the main explanation is brief. It's a convolutional brain network that serves as the foundation for the YOLO object recognition method. The use of leftover associations, as well as more layers, is improvements over its predecessor Darknet-19. [2] [8]

III. GRAPH

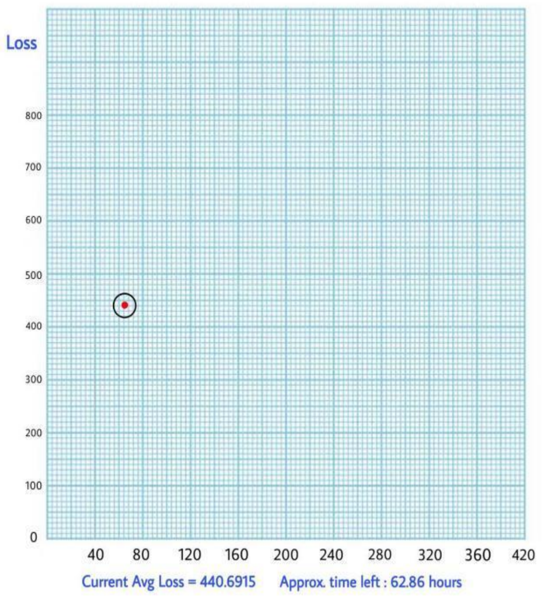

Fig 1. YOLO v1

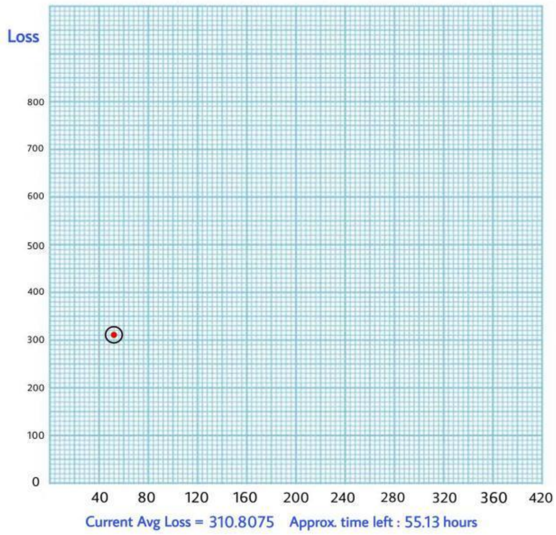

Fig 2. YOLO V3.

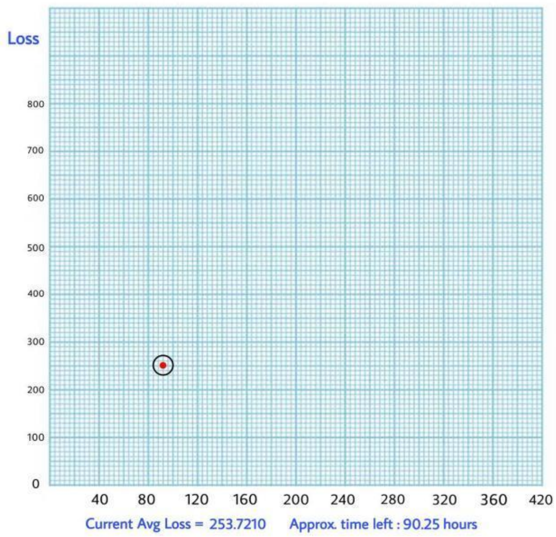

Fig 3. YOLO V4.

From these graphs given above, Current average loss gets decreased and approximate time left gets increased from YOLO version 1 to YOLO version 4.

IV. CONCLUSION

Different component extraction procedures are investigated, including a deep learning model (YOLO) and various handmade element extraction techniques. The results show that the group technique for extracting different component extraction procedures performs better than a single element extraction strategy.

The results show that using a well-known technique called YOLO to extract features is still insufficient for image classification. The analysis shows that the proposed technique is extremely powerful and consistently

© 2022 IJSRET

outperforms other scientists' strategies. The paper also mentioned various difficulties that occurred in the task of arranging pictures. This article will aid various specialists in researching other consolidated approaches for image classification, including the use of various recent deep learning models.

REFERENCES

[1] Garg A, Tandon N, Varde A (2020) I am guessing you can't recognize this: generating adversarial images for object detection using spatial commonsense (student abstract). Proc AAAI Conf Artif Intell 34(10):13789–13790. https://doi.org/10.1609/aaai.v34i10.7166

[2] Harris C, Stephens M (1988) A combined corner and edge detector. In: Proceedings of the Fourth Alvey vision conference, pp 147–151 Karthikeyan D, Varde AS, Wang W (2020) Transfer learning for decision support in Covid-19 detection from a few images in big data. IEEE Int Conf Big Data (big Data) 2020:4873–4881. https://doi.org/10.1109/ igData50022.2020.9377886

[3] Kataoka H, Iwata K, Satoh Y (2015) Feature evaluation of deep convolutional neural networks for object recognition and detection. https://arxiv.org/abs/1509.07627

[4] Srivastava S, Mukherjee P, Lall B, Jaiswal K (2017) Object classification using an ensemble of local and deep features. In: 2017 ninth international conference on advances in pattern recognition (ICAPR), pp 1–6. IEEE.

[5] Talaat A, Yousri D, Ewees A, Al-qaness MAA, Damasevicius R, Elaziz MEA (2020) COVID-19 image classification using deep features and fractional-order marine predators' algorithm. Sci Rep 10(1):15364. https://doi.org/10.1038/s41598-020-71294-2.

[6] Varde A, Rundensteiner E, Javidi G, Sheybani E, Liang J (2007) Learn ing the relative importance of features in image data. In: 2007 IEEE 23rd international conference on data engineering workshop, pp 237–244. https://doi.org/10.1109/ICDEW.2007.4400998

[7] Yadav SS, Jadhav SM (2019) Deep convolutional neural network-based medical image classification for disease diagnosis. J Big Data 6:113. https://doi.org/10.1186/s40537-019-0276-2

[8] Kleinberg EM (1996) An overtraining-resistant stochastic modeling method for pattern recognition. Ann Stat 24(6):2319–2349.

[9] Kumar V, Recupero DR, Riboni D, Helaoui R (2021) Ensembling classical machine learning and deep learning approaches for morbidity identification from clinical notes. IEEE Access 9:7107–7126. https://doi.org/10.1109/ACCESS.2020.3043221

[10] Liu L, Xie C, Wang R, Yang P, Sudirman S, Zhang J, Li R, Wang F (2020) Deep learning-based automatic multi-class wild pest monitoring approach using hybrid global and local activated features. IEEE Trans Ind Inf 17(11):7589–7598.

[11] Lowe DG (2004) Distinctive image features from scale-invariant key- points. Int J Comput vis 60(2):91–110.

[12] Mahmood A, Bennamoun M, An S, Sohel F (2017) Resfeats: residual network-based features for image classification. In: 2017 IEEE International conference on image processing (ICIP). https://doi. org/10.1109/ ic ip.2017.8296551.

[13] Mingyuan X, Wang Y (2019) Research on image classification model based on deep convolution neural network. EURASIP J Image Video Process 2019(1):1–11.

[14] Pandey A, Puri M, Varde A (2018) Object detection with neural models, deep learning, and common sense to aid smart mobility. In: 2018 IEEE 30th international conference on tools with artificial intelligence (ICTAI), pp 859–863.

© 2022 IJSRET