
M. Srikar¹, K. Malathi²

¹Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Tamil Nadu, India, PinCode: 602105
²Project Guide, Corresponding Author, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Tamil Nadu, India, PinCode: 602105

Abstract

Aim: The objective of the research is to increase the precision of object detection using novel image classification using machine learning algorithms. Materials and Methods: The categorising is performed by adopting a sample size of n = 10 in You Only Look Once (YOLO) and sample size n = 10 in Deformable Part Model (DPM) algorithms with a sample size = 10 and the G-Power analysis was carried out with 80% and confidence interval 95%. Results and Discussion: The experiment outcomes shows that the You Only Look Once (YOLO) has a high accuracy of 90.78% in comparison with the 83.47%. A statistically significant difference exists between the research groups with p=0.001 (2 tailed) (p<0.05). Conclusion: Detection of items and entities with high accuracy using machine learning algorithms shows that the You Only Look Once (YOLO) generates higher accuracy than the Deformable Part Model (DPM) algorithm.

Keywords: Object Identification, Novel Image Classification, Deformable Part Model, Bounding Boxes, Non-Maximal Suppression (NMS), Convolutional Neural Network, You Only Look Once.

https://doi.org/10.47750/pnr.2022.13.S04.204

INTRODUCTION

Today’s modern and quick changing world computer vision plays a major role in order to recognize items with greater speed and accuracy. The goal of this work is to increase the accuracy of object detection in computer vision through image classification by using a You Only Look Once (YOLO) algorithm (I and Ankith 2021). Computer vision’s major process is object detection, it is quite a complex problem for a system (Arulprakash and Aruldoss 2021). Computer vision is so essential in today’s environment, with autonomous vehicles, traffic detections, tracking items and many security purposes (Jiang et al. 2018). The approach used is image classification which helps the system to recognize and classify entities in a picture, and is used in medical imaging, detection through satellites, and traffic cameras. The main theme of the experiment is to produce an object identification model to improve precision for faster and more accurate recognition of items in a picture (Novotny and Matas 2015).

Predicting items using object detection algorithms over past years and several surveys have been published in the last years, over 13,900 articles from google scholar, 658 journals IEEE Xplore digital library 571 research articles from science direct. Among all the articles and journals, the most cited paper is (He et al. 2015). The work produced by (Redmon and Farhadi 2017) is very accurate and with upgraded speed compared to the Deformable Part Models algorithm (Tang 2016). In this paper the YOLOv3 algorithm uses a totally different technique. Input image will undergo through a single neural network layer. This layer produces bounding boxes with coordinate values over the item that is suspected in the input picture. Each box is predicted with certain proposition probabilities (Adarsh, Rathi, and Kumar 2020). It additionally determines with one network analysis unlike R-
CNN that needs multiple iteration through each layer. This compels it to be extraordinarily quick, quicker than R-CNN, Fast R-CNN and DPM algorithm (Kitano, Takiguchi, and Ariki 2015).

Our institution is passionate about high quality evidence based research and has excelled in various fields (Parakh et al. 2020; Pham et al. 2021; Perumal, Antony, and Muthuramalingam 2021; Sathiyamoorthi et al. 2021; Devarajan et al. 2021; Dhanraj and Rajeshkumar 2021; Uganya, Radhika, and Vijayaraj 2021; Tesfaye Jule et al. 2021; Nandhini, Ezhillarasan, and Rajeshkumar 2020; Kamath et al. 2020). This approach which was used earlier has much less accuracy on identifying items. It is essential to hit upon and detect the item in very milliseconds to save you problems. For example, self-riding vehicles want to detect the items in a fragment of seconds and analyse the image to move forward, in any other case there might be many consequences. The research gap found to be the low accuracy and constant speed to identify the object is lacking in the previous model. Different attributes proposed in the previous system have become less efficient as the environment changes. Order to collect the techniques and strategies on these studies usually stands far higher than the DPM algorithm. It additionally takes plenty of extra time to render all of the images to teach, the version in comparison to YOLOv3 to Deformable Part Models algorithm. The objective of the research work is to increase the precision of object classification using the You Only Look Once (YOLO) algorithm over a Deformable Part Models (DPM) algorithm.

MATERIALS AND METHODS

The experiment work was carried out in the Image Processing Lab, Department of Computer Science and Engineering, Saveetha School of Engineering, SIMATS. Basically it is considered with two groups of classifiers namely YOLOv3 and DPM algorithm which is used to detect objects in the image with various image datasets and labels. Group 1 is the YOLOv3 with the sample size of 10 and Group 2 is the DPM with sample size of 10 and it was used to compare for more accuracy score values for choosing the best algorithm to detect objects correctly. A Sample size of 20 different values has been calculated with the G power having 80%, 95% confidence interval, alpha as 0.05 and beta as 0.2 values and standard deviation for YOLOv3 = 0.32554 and DPM = 0.45353. The dataset used in the model for prediction process has been collected from the Kaggle website. The dataset can be found as the Fruit Recognition dataset created by the owner (Chris Gorgolewski). The dataset contains different label images around 660 files of apple class labels and equally other class labels named Class A and Class B respectively. The images were collected in various kinds of conditions to obtain accurate results even under different low light conditions. The data is further processed for training and testing datasets and has been split up in the ratio of 70% and 30% respectively. Adapting Google Collaboration as the software which will provide a wide Range of GPUs. The Gpu used is the NVIDIA Tesla K80. The configurations of the system used are Intel i5, 5th Gen CPU@2.8GHZ, 8 GB RAM, and 64-bit OS.

YOLO Algorithm

You Only Look Once (YOLO) is a revolutionary object recognition system. YOLOv3 is extremely fast and accurate. And by differing the dataset size, the process efficiency and precision can be managed. The previous models stated systems with classifiers or localizers to detect images. It’s one of the traditional methods. These types of models discover the pinnacle point of coordinates and sum up in the input image and classify the item. In this system the approach was quite peculiar to the previous systems. By feeding the input picture to the single convolutional neural network. Then the model breaks down the input picture into multiple portions using novel image classification. And Non-Maximal Suppression (NMS) predicts the bound boxes score for every divided portion. This approach is quite different and efficient than the classifier based models. It considers the input and passes it through the model single time and forms bounding boxes which will determine the outcome using Non-Maximal Suppression (NMS). It additionally determines with one network analysis in contrast to systems like Region Neural Networks that need 1000s for one image. This compels it to be extraordinarily quicker than R-CNN, Fast R-Convolutional Neural Network and DPM algorithm.

Pseudocode

Step 1: import tensorflow and numpy
Step 2: Make ready the image dataset to fed the model
Step 3: from PIL import Image, ImageDraw, ImageFont
Step 4: Read image → readImage()
Step 5: predict coordinates → new_array(size(NoOfCells,NoOfCells,NoOfClasses))
Step 6: Estimate bounding boxes array → new_array(size(NoOfCells,NoOfCells,))
Step 7: best_bounding_box → [0 if predictions_bounding_box_array[i,j,0, 4] > else 1]
Step 8: predicted_class → index_of_max_value(predictiction_class_array[i,j])
Step 9: print final_predictions
Deformable Part Models (DPM) Algorithm

The Deformable part model is another object detection model. Although DPM offers the benefit of managing huge appearance variances for difficult datasets, Pascal VOC takes more than 10 seconds per picture. The object detecting speed is the main issue in DPM, because it’s as important as the accuracy to detect an object. The DPM approach is quite peculiar from the YOLO approach here Non-Maximal Suppression (NMS) is not used for scoring bounding boxes which reduce the actual potential of the algorithm, DPM related approaches that promote a single category convolutional neural network in object identification proposed to convert star-structure to cascade-to-fine approach based on that model, at low resolution can prune a lot of hypotheses with low computational cost (Yan et al. 2014). The system is developed on deformable models which define items using common part models and symmetrical restraints on part locations. And by reducing object identification to classification with independent and dependent variables by performing Novel Image classification. As a result, a generic framework for feature selection classifier training using variables has been developed. A huge training dataset helps discriminative training. (P. Felzenszwalb et al. 2013). While deformable component models can identify large differences in texture, they are frequently insufficient to describe a complex object category. (P. F. Felzenszwalb et al. 2010).

Pseudocode

Step 1: import cv2, numpy
Step 2: import required image classifier.
Step 3: From skimage import io
Step 4: ready the feature classifier skimage.feature
Step 5: From skimage import data, color, exposure, transform
Step 6: train the classifier train_clf(self,P,Q,R, count kernel→'linear'):
Step 7: collect feature coordinates → [rect[A]:rect[B]+rect[C],rect[0]:rect[1]+rect[1]]=0
Step 8: test the image data→ cv2.imread('test_val/images?test'+str(count)+'.png,0)
Step 9: f1 = 2*precision * recall/(precision + recall)
Step 10: print final output value.

Statistical Analysis

IBM SPSS (version 26) statistical tool is used for analysis. 10 iterations were done with a sample size of 10 for each of the algorithms and predicted accuracy was noted for analysis. And the sample size was calculated using G power of 80% for each group and confidence interval of 95% for the two groups. The value obtained from the iterations of a total 20 samples 10 iterations from each of the algorithms and conducted Independent Sample T-test using SPSS statistical tool. The dependent data sets are ImageNet, PASCAL VOC 2007 (Thakkar et al. 2020). The independent values are AlexNet, VGGNet. The fragmented analysis has been done with independent variables labelled images, bounding box coordinates and dependent variables are accuracy, duration, feature graded object graph, val_loss and val_accuracy to find the objects with more accuracy and speed.

RESULTS

The model has trained through more than 642 files on specific labels. Group statistics of You Only Look Once (YOLO) by Deformable Part Models (DPM) by grouping with iterations sample size of 10 is represented in Table-1, mean = 90.7899 Standard Deviation = 1.07686, Standard Error Mean = 0.34053.

In Table 2, The statistical analysis of two independent groups defines that the You Only Look Once (YOLO) achieves higher accuracy mean 90.78% compared to Deformable Part Models (DPM) with accuracy 83.47%.

In Table 3, The Significant value= 0.649, Mean Difference= 7.31490 and confidence interval = (6.23666 - 8.39314) of You Only Look Once (YOLO) and Deformable Part Models (DPM) based Object detection algorithms are tabulated in Table 3. The significance value is p=0.001 (p<0.05) with an independent sample T-Test. Images, labels and tested image datasets independent variables.

Figure 1, displays the simple bar graph comparison of mean accuracy on YOLO model and DPM model.

DISCUSSION

This work on identifying items in pictures, subsequently termed object detection, is very essential in many industries in order to comprehend different scenarios by computer systems (Pathak, Pandey, and Rautaray 2018).
Object recognition approaches are basically categorised into two types, through classifiers (R-CNN) and regressor (YOLO). The mean accuracy 90.78% using YOLO and 83.47% using DPM. And the statistical 2-tailed significant difference in accuracy for two algorithms is 0.001 (p<0.05). YOLO, modern approach towards object identification. Earlier, it was performed by classifiers. But, this model considers object identification as a regression issue, it evaluates the photo only once and produces the object proposition coordinates with help of Non-Maximal Suppression (NMS). As the network flow is a single pipeline network, it’s easy to optimise based on the performance (Redmon et al. 2016).

This model proposed architecture is immensely quick and can process every image with more than 44 fps. And by considering the same network but a decreased version will achieve more than 150 frames per second for each image with help of Novel Image classification (Singh et al. 2021). In general the yolo method is very simple to execute, by passing an image as input that will feed into the single pipeline neural network will produce bounding boxes as output. Then the input picture is split into an S x S grid of cells, so each grid cell will predict the object in the provided image. Likely each object falls into certain cells accordingly. The single pipeline network of yolo contains convolutional neural networks and max pooling layers and Non-Maximal Suppression (NMS) then concluded with other fully connected layers.

Object detection is an important capability for computers and robotic futuristic systems. Despite modern advancement in robotic observation in recent years from consumer electronic devices to include driving technical assistance in the modern vehicles. Yet it’s far from reaching human level capability, specifically in an open world environment. The scope of object detection is inevitable, but the use of it in this current world is very less compared to what it can do (Verschae and Ruiz-del-Solar 2015). As in automatic guided robots and in stand-alone systems, which are getting to become more widely available deployed (quadcopters, drones and early services robots), the need for an item identification system is essential in this modern world. Finally, one must consider the need for advanced improvements in object identification, novel image classification and making machines capable of learning in the open world in real time independently (Lu, Zhang, and Xie 2020).

**CONCLUSION**

The objective of the research is to increase the precision of object detection using novel image classification using machine learning algorithms, the YOLO algorithm is notably better than the DPM algorithm. The novel image classification technique and Non-Maximal Suppression (NMS) is quite an enhancement. The YOLO algorithm is a better model as it obtained greater accuracy 90.78% than the DPM algorithm 83.47%

**DECLARATION**

**Conflicts of Interest**

No conflicts of interest in this manuscript.

**Authors Contribution**

Author MS was involved in data collection, data analysis, algorithm framing, implementation and manuscript writing. Author KM was involved in the designing the workflow, guidance and Critical review of manuscript.

**Acknowledgements**

The authors would like to extend their appreciation to Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure and constant help to carry out this work successfully.

**Funding**

We would like to express our gratitude to the following organisations for giving financial support that allowed us to finish the research successfully.

1. Qbec Infosol, Chennai.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

**REFERENCES**


Table 1. Accuracy values of YOLO and DPM models. The YOLO obtained accuracy of 90.78% compared to DPM having 83.47 %. The significant 2-tailed value for the two groups is P = 0.001 (p<0.05).

Table 2. Group Statistics of You Only Look Once(YOLO) by grouping the iterations with sample size 10, Mean = 90.7899, Standard Deviation = 1.07686. Descriptive Statistics of the mean and standard deviation of two groups with each sample size of 10 using T-Test.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std.Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>YOLO</td>
<td>10</td>
<td>90.7899</td>
<td>1.07686</td>
<td>0.34053</td>
</tr>
<tr>
<td>DPM</td>
<td>10</td>
<td>83.475</td>
<td>1.21422</td>
<td>0.38397</td>
</tr>
</tbody>
</table>

Table 2. Independent Sample T-Test is applied for the dataset fixing confidence as 95% and level of significance as P = 0.001 (p<0.05).

<table>
<thead>
<tr>
<th></th>
<th>Levene’s test for equality of variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>0.214</td>
<td>0.649</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>14.253</td>
<td>17.747</td>
</tr>
</tbody>
</table>
Fig. 1. Comparison of regional proposal network based YOLO in terms of mean accuracy. As the graph shows the mean accuracy of the YOLO is greater than the DPM. Graphical representation of the bar graph is plotted using group id as X-axis YOLO vs DPM, Y-axis displaying the error bars with mean accuracy of detection +/- 1 SD.