Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

An Investigational Study on Ensemble Learning Approaches to Solve Object Detection Problems in Computer Vision

Sree Roja Rani Allaparthi [1, a)] and Jeyakumar G [2, b)*]

[1,2] Department of Computer Science Engineering

Amrita School of Engineering, Coimbatore

Amrita Vishwa Vidyapeetham, India.

[a)]cb.en.p2aid20011@cb.students.amrita.edu,
*Corresponding Author [b)] g_jeyakumar@cb.amrita.edu

Article Info
Page Number: 399 – 412
Publication Issue:
Vol. 71 No. 3s (2022)

Abstract

Object Detection is a challenging task in computer vision, which is used to identify all or required objects in the given images or videos. The object detection tasks are widely used in many real-world image classifications and face recognition applications like self-driving cars and autonomous robots etc. Considering the challenges in object detection, this paper proposes to present a study on ensemble learning-based approaches to solving object detection problems. The proposed study uses the YOLO algorithmic model (You Only Look Once) to formulate the ensemble learning model with multiple YOLOV3 variants (YOLOV3-320-weights, YOLOV3-SPP and YOLOV3-Tiny). This ensemble learning model, formulated in this study (named YOLOV3-ensembled) is a combination of these algorithmic models. This study, initially, predicts the objects using the YOLO variants individually. Then the variants are combined to detect the objects. The experimental setup included the evaluation metrics IoU (Intersection over Union) and mAP (mean Average Precision). The comparative performance analysis of the ensemble model with other individual models is presented in this paper. It is observed from the results that the YOLOV3-320-weight model could predict the objects more accurately with good IoU scores and mAP scores.

Keywords: Object Detection, Ensemble Learning, Computer Vision, Image Classification, Face Recognition, and YOLOV3 (You Only Look Once).

Article History
Article Received: 22 April 2022
Revised: 10 May 2022
Accepted: 15 June 2022
Publication: 19 July 2022

I. INTRODUCTION

Machine learning is a computing paradigm that makes computers learn without being explicitly programmed. Its types are supervised machine learning, unsupervised machine learning, reinforcement learning, and ensemble learning. Supervised learning trains based on the input data to get the desired output, in such a way that a function is mapped. Unsupervised learning only trains based on the input data but has no desired output. Reinforcement learning has a sequence of states and actions to achieve maximum rewards. In ensemble learning the predictions are based on combined results of individual models.

The ensemble learning approaches give better performance and reduce variance and bias. The result is based on maximum or average voting. The types of ensemble learning are stacking, blending, bagging, and bootstrap. Stacking is the method that creates a new model from multiple individual predicted models viz., Random Forest, Decision Tree, and K-Nearest Neighbour. Blending is the special case of stacking that uses a validation set in place of a test set to make predictions. Bootstrapping is a technique in which the dataset is divided into several subsets with replacement. The size of the subset will be the same size as the dataset. Bagging uses these subsets to get combined results of the voting. All subset models will run at a time and so these models are independent of each other. Bagging reduces variance. Examples are Random Forest, Extra Tree, and Bagged Decision Trees. Boosting is an ensemble model where each successive or next model learns from the errors of the previous model. Each model increases the performance of the ensemble model. Boosting reduces both bias and variance. Also, uses a learning rate, a hyper-parameter. Its types are Adaboost (Adaptive Boosting) adds weights to the points which are incorrectly predicted and the successive model will predict these values correctly, Gradient Boosting (GBM) uses regression trees, a base learner, and each successive tree is built on previous tree errors, Extreme Gradient Boosting (XGBoost) also known as Regularized Boosting and works faster than other algorithms. Also uses regularization techniques to reduce overfitting and improve overall performance, Light GBM performs better than other algorithms when the data is large and follows a leaf-wise approach, Cat GBM mainly handles categorical variables(classification) in data and does not require any data pre-processing.

This paper is an attempt to implement and analyse the performance of ensemble learning approaches, on solving the challenging object detection problems of computer vision.

The rest of the paper is organized as follows – section 2 discusses the works in the literature related to the work presented in this paper, section 3 presents the proposed study, Section 4 details the experimental setup, results, and the discussions and finally, section 5 concludes the paper.

II. RELATED WORKS

This section presents the research works in the literature which are solving the object detection problems in real-world applications, using different computing methodologies including the ensemble learning approaches.

In computer vision applications forest fire detection is a challenging task because of the shapes, textures, and colors of the fire. The work presented in [1] is proposing a fire detection method using an ensemble learning method. In this method, the models Yolov5 and EfficientDet are combined. The EfficientNet is used to avoid false positives. The final predictions are based on the combined decisions of both models. Images required for this study were collected from multiple public fire datasets viz., BowFire, FD-dataset, ForestryImages, and VisiFire which includes 10581 images with both fire and non-fire images. Non-maximum suppression is applied after the integration of Yolov5 and EfficientDet. The proposed model was evaluated based on metrics like frame accuracy (FA) and false positive rate (FPR). This study observed that the ensemble model performed better in dark climates.

The work presented in [2] focuses on detecting moving objects from traffic video surveillance. Continuous video tracking leads to some issues like a greater number of moving items suddenly, alter in light conditions, size variety, and darkness conditions. This study extracts necessary video frames first and then applies a genetic vehicle detection algorithm. This system is able to detect the images and the background in the video and gives bound boxes with a detected class label. This proposed system is proved to be quick, easy, and efficient.

The work presented in [3] proposes a model trained with ensemble learning and evolutionary methods. A population is initialized and then the genetic operators - mutation, crossover, and selection were performed. This work uses Random Forest of Ensemble learning with multiple CNNs. Datasets, namely GFR-R and GFR-V, are taken from the NDWD platform, which contains all human faces over the globe. This federated ensemble learning model performs better than the federated averaging and federated file system.

The objective of the work presented in [4] is Real-time object detection using Genetic Algorithms. This system detects distance and key elements in a football match. This depends on the correlation between captured images and information gained in key elements. In this process, the population will be initialized randomly and the fitness is calculated. Information from previous populations gives the correlation between consecutively captured images. This technique yield results in short execution time by reducing the number of iterations and individuals.

The article [5] is about video segmentation and moving object detection to categorize defined classes. Video segmentation is performed using GDSM and foreground detection. These techniques are used for identifying the moving object and the max distance covered by an object in the group of boxes. Steps for this proposed technique are Video Frames, Pre-processing, Background Modelling, Foreground Detection, Data Validation, and Foreground Masks. Finally, this system detects the object by removing background subtraction.

The work presented in [6] is for COVID-19 detection from chest radiographs, by following a classification based on an ensemble learning approach. Here, multiple CNNs are used to formulate the ensemble framework. This study has many limitations like random noise during training, artificial images, deep learning behavior to take decisions, and the variability affecting the learning and evaluation. As suggested by the authors, the steps in overcoming these

limitations are pre-training using ensemble learning and then statistical analysis at learning stages.

The work presented in [7] focuses on implementing object detection using stacked YOLOv3 (You Only Look Once) by finding bounding boxes. The model is evaluated using the COCO dataset and the Intersection over Union (IoU) metric. The classes of an object are classified correctly if the IoU of the box is greater than 0.5. The Non-Maximum Suppression (NMS) is used to find the best bounding box. Yolov3 is the fastest algorithm for real-time object detection which uses the Darknet-53 network and has 53 CNN layers to recognize 80 classes. Two different types of threshold values were used to evaluate the model and both performed with a good accuracy score and detected all objects in an image at threshold 0.5.

The article [8] presents a work on deep learning-based image analysis for road damage detection. To solve this object detection problem, three ensemble learning approaches were implemented - Ensemble Model Approach (EM) – which uses multiple trained models for prediction, Ensemble Prediction Approach (EP) – which applies an ensemble of the predictions obtained from images generated by the test time augmentation procedure, Test Time Augmentation (TTA) - This applies several transformations like horizontal flipping, increasing image resolution to test an image and the Hybrid Approach (EM + EP) – which uses an ensemble model from EM for generating predictions for the images generated by the TTA procedure. For this study, the image dataset is taken from IEEE Big Data 2020 Global Road Damage Detection Challenge, which was collected from three countries: the Czech Republic, India, and Japan. Here, the training set consists of 21,041 images with some classes. Models are implemented using Faster-RCNN, YOLOv3, and u-YOLO. In these, u-YOLO achieved the highest F1 scores and this model is used for further evaluation by considering hyperparameters on augmented data.

The article [9] discusses an object detection model using the SSD MobileNet algorithm. For large datasets, the model is trained with high-performance machines. The single-shot detector-MobileNet (SSD) is used to predict multiple class objects in images and the CNN is used to train the model. The SSD has given good advantages of speed and performance.

The work presented in [10] focuses on object detection by combining CNN and Adaptive Color Prior Features. One of the challenges in this problem is insensitive to scale, light, and dark conditions. To improve the accuracy of predicted models, this paper has found an approach to model color priors. The initial step is to get the color of prior features of target samples by a cognitive-driven model. Then, these features are weighted accordingly and obtained the prior-based saliency image. These images are called features maps and merged with a convolutional neural network at the extraction stage. Here, this proposed system has experimented on Cascade R-CNN, SSD300, Libra R-CNN, and Retina Net.

The article [11] studies deep learning and GPU computing-based models for object detection. This study covered two-stage detectors like RCNN, Fast RCNN, R-FCN, FPN, Mask RCNN, and one-stage detectors like YOLO, YOLOv2, YOLOv3, SSD, RetinaNet, DSSD, and RefineDet. This study also discussed some applications like face detection, pedestrian detection, and object tracking.

The work presented in [12] proposed a Differential Evolution (DE) algorithm-based approach to solving object detection problems. This work focuses on selecting an optimal number of feature descriptors for object detection, to give a good accuracy score.

The work presented in [13] is to work on atomic clocks using an ensemble algorithm. The principle of atomic clocks is to work on frequent times generated by the atoms of the elements. This timescale is generated from atomic clock frequencies. Here, ensemble learning performs the combination of atomic clocks to yield an optimal clock. This clock is stable and optimal in frequency. The proposed system performed an artificial neural network (ANN) ensemble approach to observe the changes in clock behavior.

A detailed study on using deep learning approaches for object detection problems is presented in [14]. The study considered a smart surveillance system architecture for the experiment. In this comparative study of all deep learning algorithms for smart surveillance systems, the YOLO architecture and its versions were found to be efficient for object detection.

Considering all the above research works related to object detection in images, and the idea of ensemble-based learning approaches, this paper proposes to implement a YOLOV3 based ensemble framework for object detection. The performance of the ensembled model is compared with other models and the experimental results and the inferences are discussed in the next sections.

III. THE PROPOSED INVESTIGATION

The YOLO (You Only Look Once) is an object detection algorithm that identifies objects in images and videos. YOLOV3 uses a variant of Darknet which has 53 CNN layers. Another 53 layers are stacked to detect the object. Totally, 106 CNN layers architecture is used for YOLOV3. The proposed ensemble YOLOV3-ensembled model is the combined framework of the models YOLOV3-320-weights, YOLOV3-SPP and YOLOV3-TINY, following the stacking approach of an ensemble.

The YOLOV3-320-weights model uses down-sampling (stride=2) in convolutional layers, the YOLOV3-Tiny uses down-sampling (stride=2) in Max-pooling layers and the YOLOV3-SPP uses down-sampling (stride=2) in Convolutional layers and Max-pooling layers. These three models are combined using the ensemble boxes library to classify the object in an image. The block diagram of the ensembled model is depicted in Figure 1.

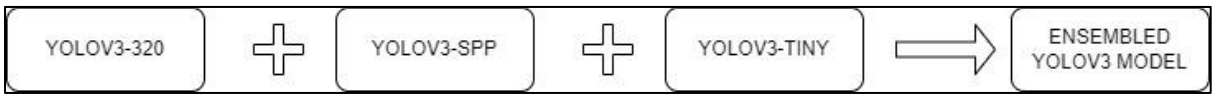


FIGURE 1. The block diagram of the ensembled model

IV THE EXPERIMENTAL RESULTS AND DISCUSSIONS

The experimental setup of this study used the COCO dataset. The COCO dataset is a large dataset that consists of 330K images and 80 object-defined classes used for object detection,

object recognition, and image segmentation problems. Many datasets can be created for required classes by using this coco dataset. This dataset is available at https://cocodataset.org/.

The YOLO models were already trained on the COCO dataset and have given a good performance. Using this pre-trained model is expected to reduce execution time and solve the challenges of object detection problems. These pre-trained models are available at https://pjreddie.com/darknet/yolo/. These trained models consist of weights and configuration files. All the layers in a network are optimized and mapped in weights and configuration files.

The evaluation metrics – Intersection of Union (IoU) and Mean Average Precision (mAP) - specific to the object detection problems are used in the experiments. The IoU is the ratio of the area of intersection to an area of a union in bounding boxes. If IoU is greater than 0.5, the class is predicted correctly. If IoU is less than 0.5, the class may not be predicted correctly. The mAP is calculated by considering average precision (AP) overall classes or overall IoU thresholds.

In this experimental study, first the YOLOV3-320-weights, YOLOV3-SPP, and YOOV3-TINY models are applied individually for the object detection problem, and then the ensembled model of these models (YOLOV3-ensembled) is applied. The output attained for the object detection problem and the evaluation metrics (IoU and mAP) measured are discussed next.

The images considered for our study are Image of giraffe and zebra, Image of dog, bicycle and truck, Image of dog, person, and horse, and Image of four horses. Henceforth, these four images are denoted as Image_gz, Image_dbt, Image_dph, and Image_4h, respectively.

On applying the YOLOV3-320-weights model the objects in the given images are classified correctly with IoU of more than 0.9. The sample images used and the results of object detection using YOLOV3-320-weights are depicted in Figure 2. Each image is classified with bounding boxes and predicted class. In the Figure, IoU score is displayed above the bounded box and the mAP score is displayed on the top of the images. All the objects in the images are classified correctly. The YOLOV3-320-Weights model performed well compared to the YOLOV3-SPP and YOLOV3-TINY models with good IoU and mAP Scores. Table 1 summarizes the performance of the YOLOV3-320-weights model.

TABLE 1. Performance summary of YOLOV3-320-weights model.

Sno	Image	mAP	IoU
1	Image_gz	87.80	Giraffe – 1.00 Zebra – 0.96
2	Image_dbt	90.85	Dog – 1.00 Bicycle – 0.99 Truck – 0.94
3	Image_dph	93.70	Dog – 0.99 Person – 1.00 Horse – 1.00
4	Image_4h	85.99	Horse1 – 0.99

Horse2 – 0.88
Horse3 – 0.96
Horse4 – 0.99

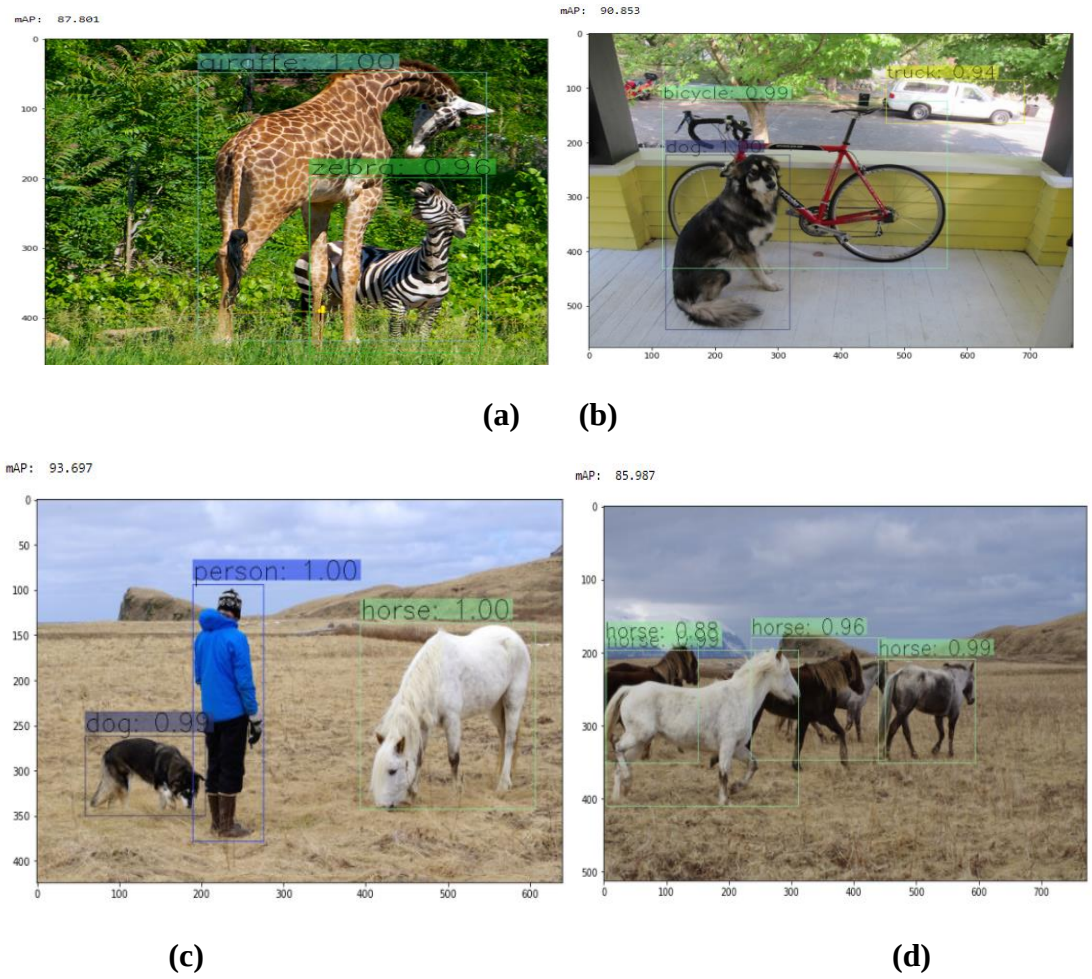
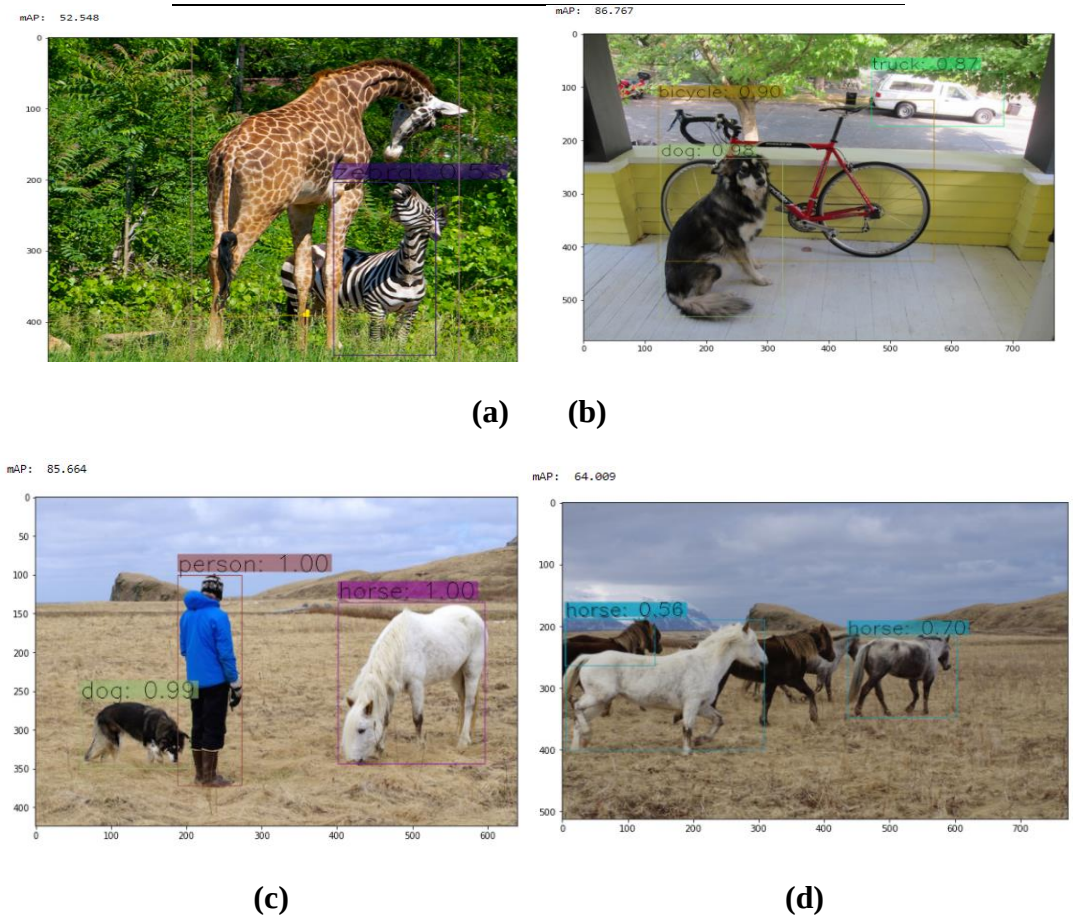


FIGURE 2. The output of YOLOV3 320 Weights model for (a) Image_gz (b) Image_dbt (c) Image_dph and (d) Image_4h.

Next, the YOLOV3-SPP model is used for object detection problems for the same set of images Image_gz, Image_dbt, Image_dph, and Image_4h. It is observed during the experiments that the YOLOV3-SPP model could predict the objects with good IoU and mAP. However, it failed to detect some objects. It could not detect the giraffe object in the image Image_gf and the horse objects in the image Image_4h. The performance summary of the YOLOV3-SPP model is presented in Table 2 and the images and the corresponding object detection outputs are depicted in Figure 3. On comparing the YOLOV3-SPP model with the YOLOV3-320-weights model, the YOLOV3-SPP performed poorly providing lower mAP and IoU scores.

TABLE 2. Performance summary of YOLOV3-SPP model.

Sno	Image	mAP	IoU
1	Image_gz	52.55	Giraffe – Not detected. Zebra – 0.53 Dog – 0.98
2	Image_dbt	86.77	Bicycle – 0.90 Truck – 0.87 Dog – 0.99
3	Image_dph	85.66	Person – 1.00 Horse – 1.00
4	Image_4h	64.01	Horse1 – Not Detected Horse2 – 0.56 Horse3 – Not Detected Horse4 – 0.70



(a) **(b)**

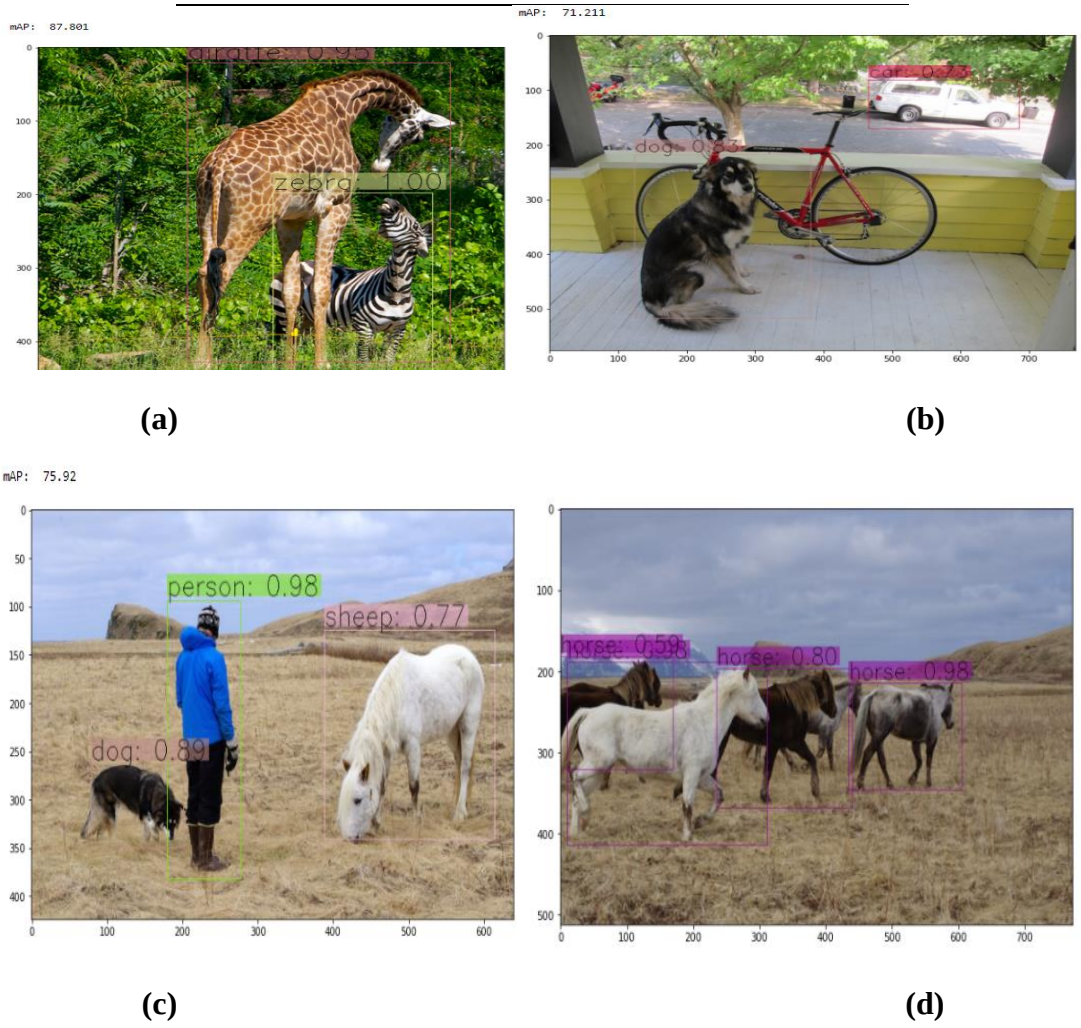
(c) **(d)**

FIGURE 3. The output of YOLOV3-SPP model for (a) Image_gz (b) Image_dbt (c) Image_dph and (d) Image_4h.

Now the third model YOLOV3-Tiny has experimented with the object detection problem on the same set of sample images (Image_gz, Image_dbt, Image_dph, and Image_4h.). The images and the object detection outputs are visualized in Figure 4, and the mAP and IoU scores of this model are presented in Table 3.

Table 3. Performance summary of YOLOV3-Tiny model**.**

Sno	Image	mAP	IoU
1	Image_gz	87.80	Giraffe – 0.95 Zebra – 1.00 Dog – 0.83
2	Image_dbt	71.21	Bicycle – Not Detected. Truck – 0.73, but detected as car. Dog – 0.89
3	Image_dph	75.92	Person – 0.98 Horse – 0.77, but detected as sheep. Horse1 – 0.58
4	Image_4h	64.01	Horse2 – 0.59 Horse3 – 0.80 Horse4 – 0.98



(a) **(b)**

(c) **(d)**

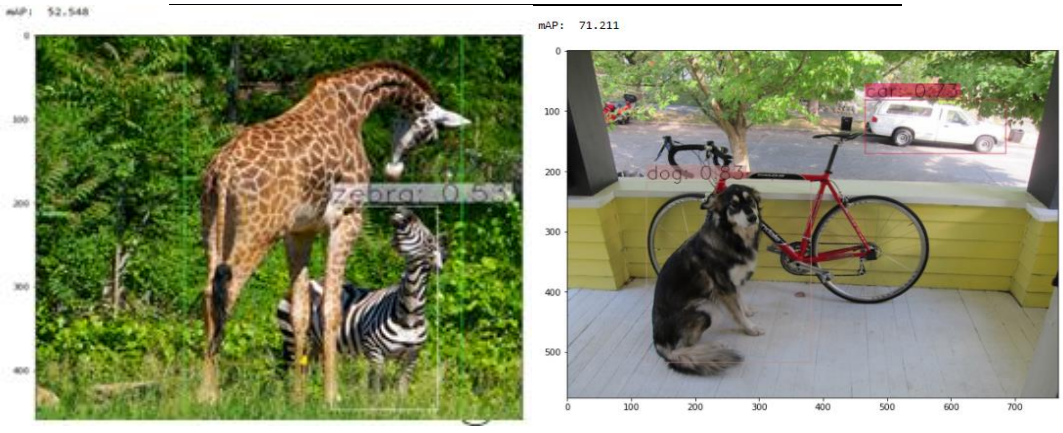
FIGURE 4. The output of YOLOV3-Tiny model for (a) Image_gz (b) Image_dbt (c) Image_dph and (d) Image_4h.

It is observed from the experimental results that the YOLOV3-Tiny failed in detecting some of the objects correctly. It detected the truck as the car in the image Image_dbt, and the horse as sheep in the image Image_dph. As well as it could not detect the bicycle object in the image Image_dbt.

Finally, the ensembled YOLOV3-ensembled model is implemented for the object detection problems on the images Image_gz, Image_dbt, Image_dph, and Image_4h. The YOLOV3-ensembled model is the combination of the individual three models studied above. The evaluation score metrics of the ensembled model are presented in Table 4, and the outputs of the object detection process are visualized in Figure 5.

TABLE 4. Performance summary of YOLOV3-ensembled model.

Sno	Image	mAP	IoU
1	Image_gz	52.55	Giraffe – Not Detected. Zebra – 0.53 Dog – 0.83
2	Image_dbt	71.21	Bicycle – Not Detected. Truck – 0.73, but detected as car.
3	Image_dph	75.92	Dog – 0.89 Person – 0.98 Horse – 0.77, but detected as sheep.
4	Image_4h	64.01	Horse1 – 0.56 Horse2 – Not detected. Horse3 – Not detected. Horse4 – 0.70



(a) (b)

(c) (d)

FIGURE 5. The output of YOLOV3-ensembled model for (a) Image_gz (b) Image_dbt (c) Image_dph and (d) Image_4h.

The comparative performance analysis of the models considered (YOLOV3-320-weights, YOLOV3-SPP, YOLOV3-Tiny, and YOLOV3-ensembled) in this study was carried out comparing their mAP and IoU scores.

The mAP comparison is presented in Table 5 and is visualized in Figure 6. It is observed from the results that the YOLOV3-320-weights model has provided a higher mAP score compared to other models in all the images. In Image_gz, the YOLOV3-Tiny model also performed similarly to YOLOV3-320.

TABLE 5. Comparison of mAP scores.

Sno	Image	Comparison of mAPs			
		YOLOV3-320	**YOLOV3-SPP**	**YOLOV3-Tiny**	**YOLOV3-Ensembled**
1	Image_gz	**87.8**	52.55	**87.8**	52.55
2	Image_dbt	**90.85**	86.77	71.21	71.21
3	Image_dph	**93.7**	85.66	75.92	75.92
4	Image_4h	**85.99**	64.01	64.01	64.01

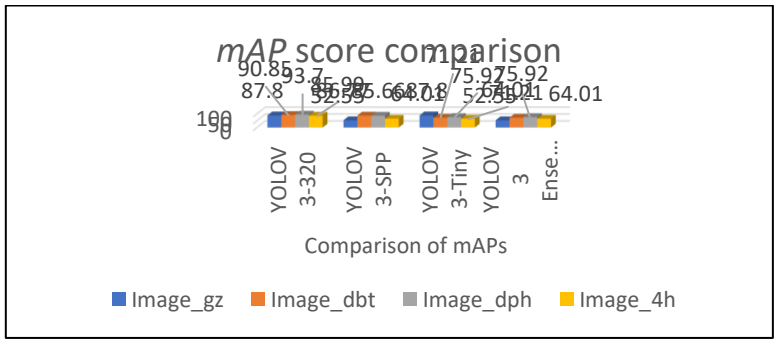


FIGURE 6. The comparison of mAP scores of the models.

The comparison of the models based on the IoU score is presented in Table 6 and is visualized in Figure 7. It is observed from the results in Table 6 that the YOLOV3-320-weights model could detect all the objects in the given images successfully with higher IoU scores. The YOLOV3-SPP model could not detect the Giraffe, Horse1, and Horse2 objects in the images Image_gz and Image_4h. The YOLOV3-Tiny model failed in detecting the Bicycle object in the image Image_dbt, as well as detecting the Truck object in Image_dbt as Car and the Horse object in Image_dbh as Sheep. The ensemble model, YOLOV3-ensembled, failed to detect fours objects – the Giraffe object in Image_gz, the Bicycle object in Image_dbt, and the Horse2, Horse3 objects in the image Image_4h. It is also observed that the ensemble model showed the combined performance of YOLOV3-SPP and YOLOV3-Tiny in many cases. Also, in all the cases where the YOLOV3-SPP, YOLOV3-Tiny, and the YOLOV3- ensembled, models detected the objects correctly their IoU score is lesser than the corresponding IoU score of the YOLOV3-320 model.

The above-said observations can be reiterated by referring the Figure 8, also. In this figure, the IoU values for the 'Not Detected' cases are assigned as 0 (zero), and for the cases of the wrong detection of objects, it is assigned as '-1'.

Table 6. The comparison of IoU scores

Sno	Image	Object to be detected	YOLOV3-320	YOLOV3-SPP	YOLOV3-Tiny	YOLOV3-Ensembled
1	Image_gz	Giraffe	1	Not detected.	0.95	Not Detected.
		Zebra	0.96	0.53	1	0.53
2	Image_dbt	Dog	1	0.98	0.83	0.83
		Bicycle	0.99	0.9	Not Detected.	Not Detected.
		Truck	0.94	0.87	0.73, but detected as car.	0.73, but detected as car.
3	Image_dph	Dog	0.99	0.99	0.89	0.89
		Person	1	1	0.98	0.98
		Horse	1	1	0.77, but detected as sheep.	0.77, but detected as sheep.
4	Image_4h	Horse1	0.99	Not Detected	0.58	0.56
		Horse2	0.88	0.56	0.59	Not detected.
		Horse3	0.96	Not Detected	0.8	Not detected.
		Horse4	0.99	0.7	0.98	0.7

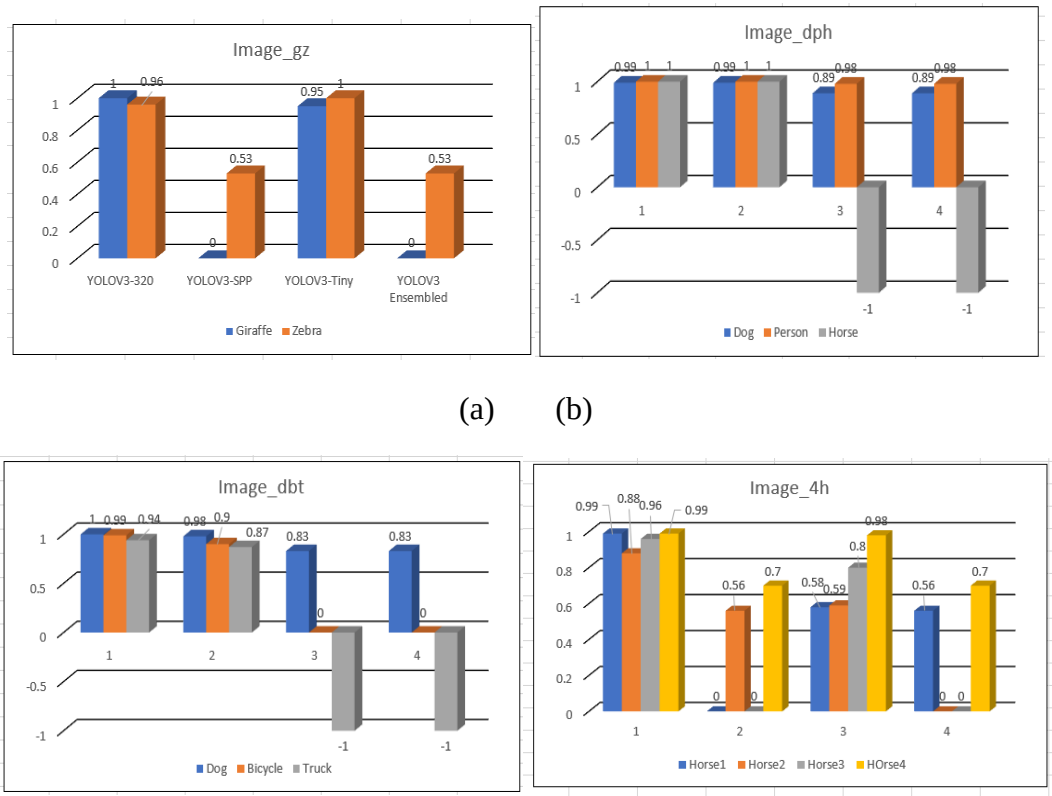


FIGURE 7. The comparison of IoU scores for (a) Image_gz (b) Image_dbt (c) Image_dph and (d) Image_4h.

V. CONCLUSIONS

This paper presented a study on implementing and comparing an ensemble learning model with the constituent individual models on solving object detection problems of computer vision. The YOLOV3 object detection algorithm and its variants YOLOV3-320-weights, YOLOV3-SPP, and YOLOV3-Tiny are considered for this study. An ensemble model, name YOLOV3-ensembled, is formulated. The experimental study covered implementing these four algorithmic models on solving object detection problems. The performances of these models were compared based on the evaluation metrics Intersection of Union (IoU) and Mean Average Precision (MAP). The comparative study revealed that the individual model YOLOV3-320-weights could outperform other models with accurate detection of all the objects in the images considered for the experiment.

It is also noted that the ensembled model could not show any advantages over other individual models and performed as similar to YOLOV3-SPP and YOLOV3-Tinny models. This necessitates a detailed investigation on ensembled learning approaches, which will be taken as future work of this study by the authors.

REFERENCES

1. Xu R, Lin H, Lu K, Cao L, Liu Y. "A Forest Fire Detection System Based on Ensemble Learning", Forests. Vol. 2, Issue 2., pp. 217, 2021.

2. J. Dey and N. Praveen, "Moving object detection using genetic algorithm for traffic surveillance," In the Proceedings of 2016International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 2289-2293, 2016.
3. L. Li, M. Li, F. Qin and W. Zeng, "Evolutionary-based Federated Ensemble Learning on Face Recognition," In the proceedings of the IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), pp. 815-819, 2021.
4. Martínez-Gómez J., Gámez J.A., García-Varea I and Matellán V. "Using Genetic Algorithms for Real-Time Object Detection", In the proceedings of RoboCup 2009: Robot Soccer World Cup XIII, Lecture Notes in Computer Science, Springer, Vol. 5949, 2010.
5. R. Aruna Jyothi, K. Ramesh Babu, Srinivas Bachu, "Moving Object Detection Using the Genetic Algorithm for Real Times Transportation", International Journal of Engineering and Advanced Technology, Vol. 8., No. 6, 2019.
6. Rajaraman S, Sornapudi S, Alderson PO, Folio LR and Antani SK, "Analyzing inter-reader variability affecting deep ensemble learning for COVID-19 detection in chest radiographs", PLoS One, Vol. 15., No. 11, 2020.
7. Padmanabula, S.S., Puvvada, R.C., Sistla, V and Kolli, V.K.K. (2020), "Object detection using stacked YOLOv3", Ingénierie des Systèmes d'Information, Vol. 25., No. 5., pp. 691-697., 2020.
8. V. Hegde, D. Trivedi, A. Alfarrarjeh, A. Deepak, S. Ho Kim and C. Shahabi, "Yet Another Deep Learning Approach for Road Damage Detection using Ensemble Learning," In the proceedings of 2020 IEEE International Conference on Big Data (Big Data), pp. 5553-5558., 2020.
9. Sanjay Kumar K.K.R., Subramani G., Thangavel S and Parameswaran L, "A Mobile-Based Framework for Detecting Objects Using SSD-MobileNet in Indoor Environment.", Advances in Intelligent Systems and Computing, Vol., 1167, pp. 65-76., 2021.
10. Gu P, Lan X and Li S, "Object Detection Combining CNN and Adaptive Color Prior Features", Sensors, Vol. 21., No. 8, pp. 2796., 2021.
11. Pal, S.K., Pramanik, A., Maiti, J and Pabitra Mitra, "Deep learning in multi-object detection and tracking: state of the art", Applied Intelligence, Vol. 51., pp. 6400–6429., 2021.
12. Sree Katamneni Vinaya and Jeyakumar, G., "An Evolutionary Computing Approach to Solve Object Identification Problem in Image Processing Applications", Journal of Computational and Theoretical Nanoscience, Vol. 17., No. 1., pp. 439-444., 2020.
13. Rajathilagam B., S. N., Maharana, S., Subramanya Ganesh, T., and Krishnamoorthy, S, "An Artificial Neural Network Model for Timescale Atomic Clock Ensemble Algorithm", . MAPAN- Journal of Metrology Society of India, Vol. 35. No. 4., pp. 547 – 554., 2020.
14. U. Subbiah, D. K. Kumar, S. K. Thangavel and L. Parameswaran, "An Extensive Study and Comparison of the Various Approaches to Object Detection using Deep Learning," In the proceedings of Third International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 1018-1030., 2020.