**Design and implementing an algorithmic framework with stacking/bagging of ensemble learning to support Genetic Algorithm to solve Object Detection problems.**

**19AI798 – Dissertation**

Submitted by

CB.EN.P2AID20011 ALLAPARTHI SREE ROJA RANI

in partial fulfillment of the requirements for the degree of

master of technology

in

ARTIFICIAL INTELLIGENCE



Department of computer science and engineering

amrita school of engineering

amrita vishwa vidyapeetham

coimbatore – 641 112

DEC 2021

amrita vishwa vidyapeetham

amrita school of engineering, coimbatore

department of computer science and engineering



bonafide CERTIFICATE

This is to certify that the **19AI798 – Dissertation** entitled **Design and implementing an algorithmic framework with stacking/bagging of ensemble learning to support Genetic Algorithm to solve Object Detection problems** submitted by **Allaparthi Sree Roja Rani(CB.EN.P2AID20011)** in partial fulfilment of the requirements for the award of Degree of **Master of Technology in Artificial Intelligence** is a bonafide record of the work carried out under my guidance and supervision at Amrita School of Engineering, Coimbatore.

ALLAPARTHI SREE ROJA RANI

**CB.EN.P2AID20011**

Dr. G. Jeyakumar Dr. (Col) P. N. Kumar

Thesis Advisor Chairperson

Vice Chairperson & Associate Professor Department of CSE

Department of CSE

Evaluated on:

INTERNAL EXAMINER EXTERNAL EXAMINER

**ABSTRACT**

The project work mainly focuses on integrating machine learning approaches with evolutionary algorithms. This work includes the ensemble learning approach of machine learning (ML) and genetic algorithm (GA) of evolutionary algorithm (EA). The objective of this project work is to investigate utilizing the support of machine learning approaches to optimization techniques. This investigation proposes to design and implement an algorithmic framework to solve object detection problems using convolutional neural networks, ensemble learning, and genetic algorithm. The initial step is to perform stacking (ensemble learning) for multiple YOLOV3(You Only Look Once) models. Then, to perform object detection using a genetic algorithm. Finally, integrates object detection using stacking with multiple YOLOv3 models and object detection using a genetic algorithm. The Object Detection problem is an application of computer vision and basically detects all classes in an image. The efficiency of the proposed framework will be tested on object detection problems using relevant performance metrics. The YOLOv3 models like tiny, SPP and 320 weights are implemented and performed ensemble on these models. These models predicted objects correctly with IoU > 0.5 and mAP is greater than 80. And, future work includes object detection using genetic algorithm and then integrating yolo ensemble model with genetic algorithm object detection model.

Keywords: machine learning, evolutionary algorithms, stacking, bagging, ensemble learning, genetic algorithm, optimization, object detection, convolutional neural networks, and YOLOV3(You Only Look Once).

ACKNOWLEDGEMENT

I wish to record my deep sense of gratitude and profound thanks to my research supervisor Dr. G Jeyakumar, Vice-Chairperson & Associate Professor, Department of CSE, Amrita Vishwa Vidyapeetham, Coimbatore, for his / her keen interest, inspiring guidance, constant encouragement with my work during all stages, to bring this thesis into fruition.

I am extremely indebted to **Dr.**(Col) P. N. Kumar, Chairperson, Department of CSE, Amrita Vishwa Vidyapeetham, Coimbatore and panel members for their valuable suggestions and support during the course of my research work.

I also thank the faculty and non-teaching staff members of the Department of CSE, Amrita Vishwa Vidyapeetham, Coimbatore, for their valuable support throughout the course of my research work.

ALLAPARTHI SREE ROJA RANI **TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTER** | **TITLE** | **PAGE NO** |
|  | **ABSTRACT** | **4** |
|  | **LIST OF TABLES** | **7** |
|  | **LIST OF FIGURES** | **8** |
|  | **LIST OF ABBREVIATIONS** | **9** |
| 1 | **INTRODUCTION** | 10-13 |
| 1.1 | Problem Definition | 10-12 |
| 1.2 | Motivation / Challenges | 12-13 |
| 2 | **LITERATURE SURVEY** | 13-17 |
| 2.1 | Survey | 13-17 |
| 2.2 | Findings | 13-17 |
| 3 | **PROPOSED WORK** | 18-20 |
| 3.1 | Architecture of the System | 18-20 |
| 3.2 | Data sets for the study and Platform | 20 |
| 4 | RESULTS AND INFERENCES | 21-24 |
|  | Results and discussion | 21-24 |
| 5 | CONCLUSION AND FUTURE ENHANCEMENTS | 25 |
| **6** | **REFERENCES** | **26-27** |
| **7** | **LIST OF PUBLICATIONS ( IF APPLICABLE)** | - |

LIST OF TABLES

|  |  |  |
| --- | --- | --- |
| TABLE NO | TITLE | PAGE NO |
| 1 | Summary of papers on Ensemble Learning | 17 |
| 2 | Summary of papers on Evolutionary Algorithms | 17 |

LIST OF FIGURES

|  |  |  |
| --- | --- | --- |
| FIGURE NO | TITLE | PAGE NO. |
| 1 | BLOCK DIAGRAM OF PROPOSED SYSTEM | 18 |
| 2 | YOLO ARCHITECTURE | 19 |
| 3 | RESULTS | 21-24 |
|  |  |  |
|  |  |  |

LIST OFABBREVIATIONS

|  |  |
| --- | --- |
| ML | Machine Learning |
| EL | Evolutionary Learning |
| EA | Evolutionary Algorithm |
| GA | Genetic Algorithm |
| EL | Ensemble Learning |
| YOLO | You Only Look Once |

Introduction

* 1. **Problem Definition**

This research work mainly focuses on Evolutionary Learning. Evolutionary Learning (EL) is the integration of machine learning (ML) with evolutionary algorithms (EA). In simpler terms, the process of applying optimization to machine learning. Machine learning is the type of learning which makes computers learn without being explicitly programmed. And the types are supervised machine learning, unsupervised machine learning, reinforcement learning, and ensemble learning. Supervised learning trains the input data and gives desired output in such a way that a function is mapped. Unsupervised learning only trains the input data and has no desired output. Reinforcement learning has a sequence of states and actions to achieve maximum rewards. This research work mainly focuses on ensemble learning, where predictions are based on combined results of individual models.

Evolutionary algorithms (EA) are a heuristic-based approach to solve problems that cannot be easily solved in polynomial time such as classically NP-Hard problems. EA needs a target function that is to be optimized. EA methods can apply to any complex problems like discontinuous, non-differentiable, and noisy target functions. EA is also known as Evolutionary Computation (EC). There are four types of EAs: Genetic Algorithm (GA), Genetic Programming (GP), Evolution Strategies (ES), and Evolutionary Programming (EP). A genetic algorithm is the process of natural selection where individuals are selected for reproduction in order to produce offspring for the next generation. Genetic Programming is a branch of GA and has a tree-shaped representation of the solution. Evolution Strategies are mainly used for combinatorial optimization, constrained optimization, and multi-objective optimization. Evolutionary Programming is used for finite state machines to study the development of intelligence and problem-solving.

1. **ENSEMBLE LEARNING**

Ensemble learning is a type of learning where predictions are based on combined results of individual models. This gives better performance and reduces variance and bias. The result is based on maximum or average voting. Types are Stacking, Blending, Bagging, and Bootstrap. Stacking is the method that uses predictions from multiple models (SVM, KNN, Decision Tree) to build a new model and this is used for making predictions on the test set. Blending is the special case of Stacking that used validation set in place of test set to make predictions. Examples for Stacking are Stacked Models, Super Ensemble. Bagging follows bootstrapping. Bootstrapping is a sampling technique in which the dataset is divided into several subsets with replacement. The size of subset should be same as the size of original set. Bagging uses these subsets to get combined results of the voting. All subset models run in parallel and are independent of each other. Bagging reduces variance. Examples are Random Forest, Extra Tree, Bagged Decision Trees.

Boosting is an ensemble model where each subsequent model attempts to correct the errors of the previous model. The succeeding models depend on 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. Types are: Adaboost (Adaptive Boosting) assigns weights to the observations which are incorrectly predicted and the next model works to predict these values correctly, Gradient Boosting (GBM) uses regression trees which is used as a base learner, and each subsequent tree is built on previous errors, Extreme Gradient Boosting (XGBoost) also known as Regularized Boosting and works 10 times faster than other boosting techniques. Also uses regularization techniques to reduce overfitting and improve overall performance, Light GBM performs better than all the other algorithms when the dataset is large and follows a leaf-wise approach, Cat GBM mainly handles categorical variables and does not require any data preprocessing. This research work focuses on Stacking/Bagging algorithms of Ensemble learning.

1. **GENETIC ALGORITHM**

Genetic Algorithm is the process of natural selection where individuals are selected for reproduction in order to produce offspring to the next generation. GA has a fixed length and is represented in the form of integer vectors. GA has some terminology like Gene (a single element), Chromosome (entire row), Population (set of individuals). Genes are represented with binary values (1s or 0s). Also, GA uses genetic operators like mutation and crossover. The process of altering the value of a gene is called a mutation.

For Example: [ 1,0,0,1] 🡪 [1,1,0,1] – second element is mutated.

To generate new offspring from two-parent chromosomes with selection is called crossover.

For Example:

Parent members Crossover

Chromosome 1: 0 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 🡨 Selection

Chromosome 2: 1 0 0 1 1 1 0 0 1 0 0 1 1 1 1 1 🡨 Offspring

Here, parent chromosome 2 is replaced with the selection which yields offspring.

Genetic Algorithm can be solved in five steps: Initial Population, Fitness Function, Selection, Crossover, and Mutation. Fitness function determines how to fit an individual and gives some fitness score to each individual.

In Selection, individuals are selected based on fitness scores for the objective of the problem like maximum or minimum one.

1. **EVOLUTIONARY LEARNING**

Evolutionary learning (EL) is the integration of machine learning with evolutionary algorithms. First, machine learning should be applied for a problem and that should be given as population to perform the genetic algorithm. EL gives optimized results with good performance. This learning can be applied to any problem which uses both learning and optimization. One of the applications is: Dynamic Pricing or Price Optimization: It is the concept of offering goods at different prices which varies according to customer’s demand. For instance, the cost of a ticket is based on the type of ticket like business class, general class. This application gives a good profit by rising in seasonal pricing.

One of the papers is Evolutionary-based Federated Ensemble learning on Face Recognition [10]: Federated learning is collaborative machine learning without centralized training data. This paper works on ensemble learning with CNN and evolutionary methods like crossover, mutation, and selection.

* 1. MOTIVATION AND CHALLENGES

Object Detection is an application of Computer Vision with Deep Learning approaches used for self-driving cars, face recognition applications. This research work is to detect required objects in given images using Ensemble learning approaches which combines the results of multiple CNN to achieve good performance. Then, using genetic algorithm to optimize the results. This research work is mainly to give optimized results with good performance.

In Object Detection problems, there are some challenges. They are:

1. Multiple Scales: In many applications of object detection, objects in images and videos will be in different sizes and different views. The solution for this is using Anchor boxes.
2. Execution time: For Real-time processing, detecting object takes more execution time. The solution for this is to use pre-trained models.
3. Low Data: If the data is low, the model cannot be trained well and may predict the object misclassifically.

literature survey

In modern times, forest fire detection [1] is a challenging task due to the many shapes, textures, and colors of fires. This work is implemented using the ensemble learning method. Two individual models Yolov5 and EfficientDet are integrated to detect the fire in this problem. And, EfficientNet is responsible for learning global information to avoid false positives. Final predictions are based on the decision of three learners. Images were collected from multiple public fire datasets: BowFire, FD-dataset, ForestryImages, VisiFire and integrated all above into one which contains 10581 images with both fire and non-fire images. Non-maximum suppression is applied after integration of Yolov5 and EfficientDet and the model EfficientNet gives classified results. These results are evaluated based on metrics like frame accuracy (FA) and false positive rate (FPR). Finally, the ensemble model performed better on ground fires, trunk fires and canopy fires.

The graph coloring problem [2] can also be solved using an evolutionary algorithm. This is a constraint satisfaction problem (CSP). The main aim of this problem is there is an undirected graph with vertices and edges, to color each node such that no two adjacent nodes have the same color. The accuracy can be verified by measuring success rate and computational complexity- will be measured by average number of fitness evaluations in successful runs. Here, grouping genetic algorithms is used, two parts like object part and group part are included. When a chromosome is selected for mutation, a number of elements in group part are deleted. When a group is deleted, the color will be temporarily uncolored.

This project work [6] focuses on detecting moving objects from traffic video surveillance. Continuous video tracking leads to some issues like more number of moving items suddenly, alter in light conditions, size variety, and darkness conditions. Firstly, extract necessary video frames and apply genetic vehicle detection algorithm. Then, generate population by randomization method, Calculate the value of fitness and choose chromosomes randomly and then apply crossover and do validation of chromosomes. This proposed system is quick, easy, efficient. This system detects images, background in a video and gives bound boxes with a detected class label.

The paper on Stacking Ensemble Learning [7] for short-term electricity consumption forecasting is used to tackle the short-term load forecasting problem. The more efficient usage of power reduces cost. Here, multiple models were used like Evolutionary Algorithm, ANN, Random Forest and Gradient Boosted Tree. Ensemble learning is performed between these models. Gradient Boosted Tree is used as meta regressive learner.

This paper [9] works on deep learning-based image analysis for road damage detection. To solve this object detection problem, three ensemble learning approaches were implemented:

1. Ensemble Model Approach (EM) – which uses multiple trained models for prediction
2. 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.

1. Hybrid Approach (EM+EP) – which uses an ensemble model from EM for generating predictions for the images generated by the TTA procedure.

The image dataset is taken from IEEE Big Data 2020 Global Road Damage Detection Challenge 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, u-YOLO approaches. In these, u-YOLO achieved the highest F1 scores and this model is used for further evaluation by considering hyperparameters on augmented data.

Federated Learning [10] is basically collaborative machine learning without centralized training data. Object detection is sub-task of Face recognition. This system proposes a model trained with ensemble learning and evolutionary methods. First, the population is initialized and performs genetic operators like mutation, crossover, selection. From this, population gives inference to ensemble learning (to combine individual models). 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 this project work [11] is Real-time adaptation of genetic algorithms. This system detects and estimates distance and orientation to key elements on a football field. This depends on the similarity between captured images and information obtained with the filtering process. In this process, the population will be initialized randomly and calculate the value of fitness. Information from previous populations gives the similarity between consecutive images taken by camera. This technique yield results in short execution time by reducing the number of iterations and individuals. This system gives good performance and efficiency.

The objective of this paper [12] is to review the video segmentation and moving object detection methods to categorize them into classes. Video segmentation is done using genetic dynamic saliency map (GDSM) and background subtraction (also known as foreground detection). These techniques are used for identifying the moving object and the maximum distance moved by the object in the given group of frames. Steps for this proposed technique: Video Frames, Preprocessing, Background Modeling, Foreground Detection, Data Validation, and Foreground Masks. Finally, this system detects the object by removing background subtraction.

The paperwork on COVID-19 detection [13] in chest radiographs are classified based on ensemble learning. Here, multiple CNNs are used to perform ensemble and get good performance in Computer Vision tasks. This study has many limitations like random noise during training, artificial images, deep learning behavior to support decision-making, inter-reader variability affecting learning and evaluation. The steps in overcoming these limitations are pre-training using ensemble learning and then statistical analysis at learning stages.

The state-of-the-art research works in the area chosen for this project work are discussed here. The work presented in [14] mainly focuses on implementing object detection using stacked YOLO v3(You Only Look Once) by finding bounding boxes. The model is evaluated using COCO dataset and Intersection over Union (IoU) is the suitable metric for evaluating results. The class of object is classified correctly if the IoU of the box is greater than 0.5. Non-Maximal Suppression (NMS) is used to keep the best bounding box in all boxes. Yolov3 is the fastest algorithm for real-time object detection which uses Darknet-53 architecture and has 53 convolutional 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 paper on solving multi-objective Travelling Salesman Problem [15] using Evolutionary Algorithm: The main aim of this problem is optimizing the order of cities. To minimize the total distance covered, each city should visit once and should get back to starting city. Uses genetic algorithm based on crossover and mutation. Each chromosome is represented by set of cities and encodes a complete solution. Single point crossover is used to create the offspring.

**Table No 1:** Summary of Papers on Ensemble Learning:

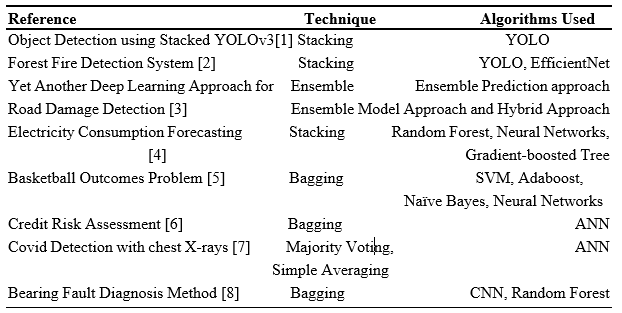
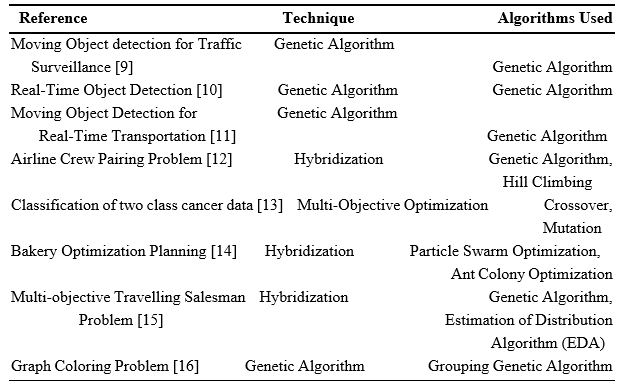


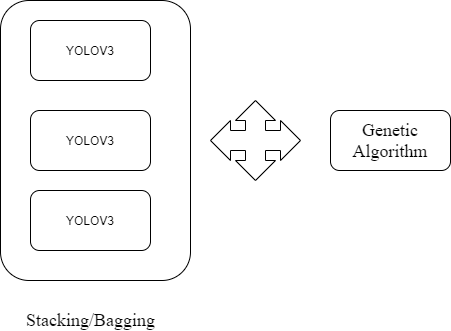
Table No 2: Summary of Papers on Evolutionary Algorithms:



**CHAPTER 3**

proposed work

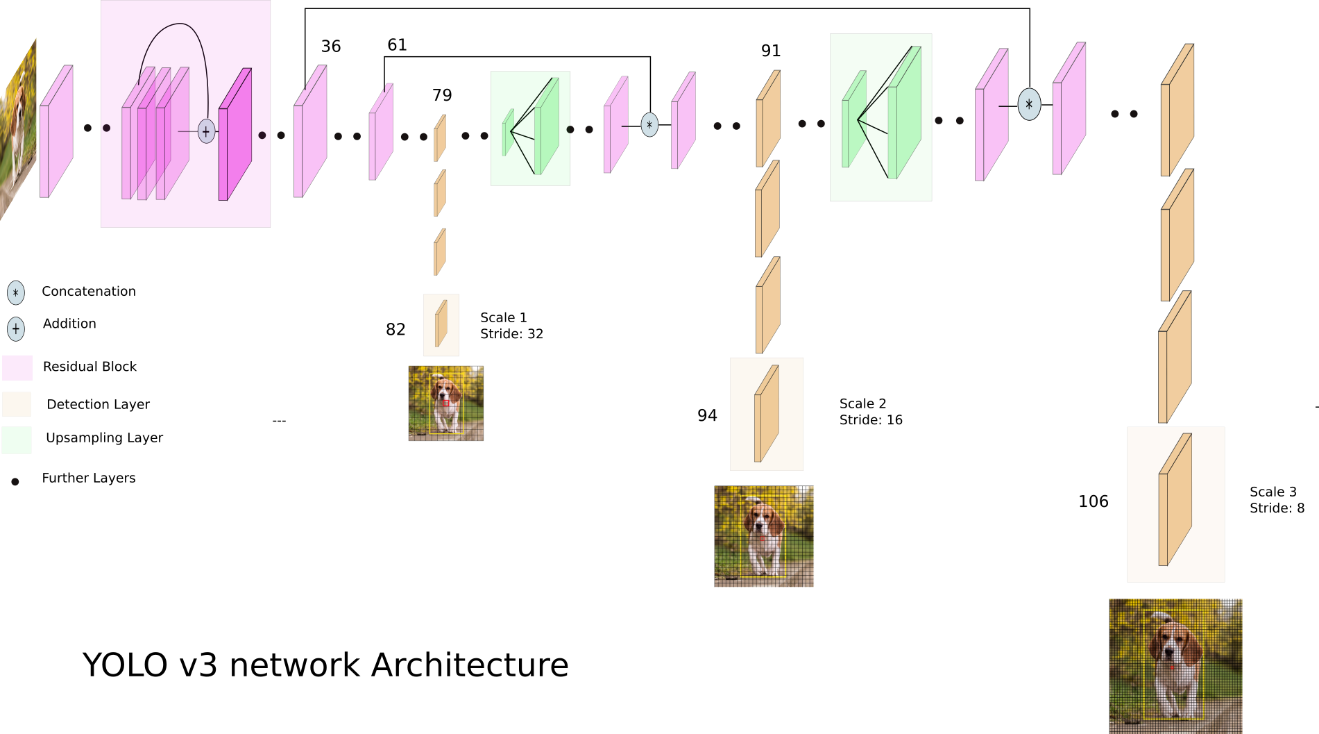
The Block Diagram of Proposed System:



Workflow of the Project:

**Module 1** – Object Detection using ensemble learning: Applying ensemble learning for multiple YOLOV3 to solve object detection problems. This can be implemented using Keras.

YOLO is an object detection algorithm that identifies objects in images and videos. YOLOV3 uses a variant of Darknet which has 53 layers trained on ImageNet. Another 53 layers are stacked to detect the object. A 106 fully connected convolutional layers architecture is used for yolov3.



Here, three YOLOV3 models are stacked. They are: YOLOV3-320 weights, YOLOV3-Tiny, YOLOV3- SPP.

YoloV3 uses downsampling(stride=2) in convolutional layers.

YoloV3 Tiny uses downsampling(stride=2) in Max-pooling layers.

YoloV3 SPP uses downsampling(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.

**Module 2** - Object Detection using Genetic Algorithm: Applying genetic algorithm for object detection, which includes initialization of population, calculate the value of fitness, then genetic operators like crossover and mutation can be performed. Finally, selection will be done which gives optimal results.

The genetic algorithm is as follows:

START

Generate the initial population

Compute fitness function

REPEAT

Selection

Crossover

Mutation

Compute fitness

UNTIL population has converged

STOP

**Module 3 –** Integrating Module 2 and Module 3:

Integrating Object Detection using Ensemble Learning with Object Detection using Genetic Algorithm.

**3.2 DATASET**

COCO (Common Objects in Context) Dataset:

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 taken from the link: <https://cocodataset.org/>

YOLO models are already trained on COCO dataset and gave a good performance. By using this pre-trained model, there is a chance to reduce execution time. This can solve a challenge. These pre-trained models are taken from: <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.

Chapter 4

## RESULTS AND INFERENCES

**Evaluation Metrics:**

The object detection problem has some specific evaluation metrics. They are Intersection over Union (IoU), mean Average Precision(mAP).

Intersection over Union (IoU): It is the ration of measure of area of intersection to area of union in bounding boxes.

If IoU is greater than 0.5, the class is predicted correctly.

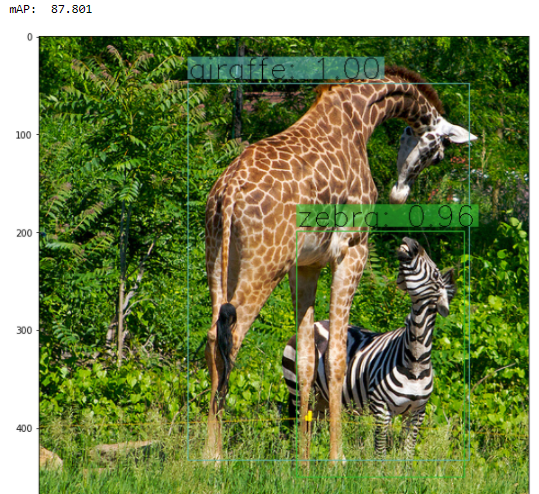
If IoU is less than 0.5, the class may not be predicted correctly.

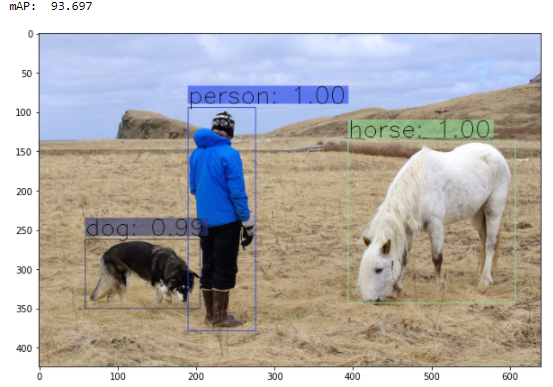
Mean Average Precision(mAP): It is calculated by taking the mean AP over all classes or overall IoU thresholds.

**RESULTS:**

1. YOLOV3-320 Weights:

Here, the objects in the images are classified correctly with IoU more than 0.5.

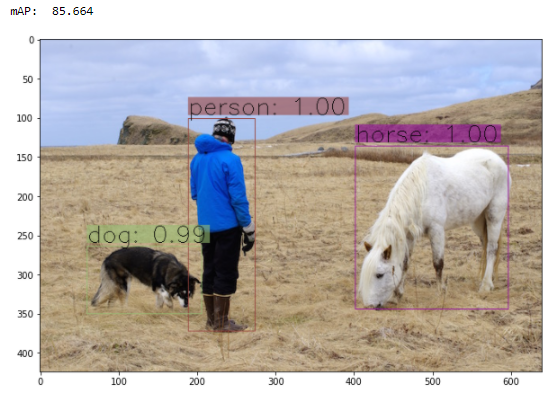


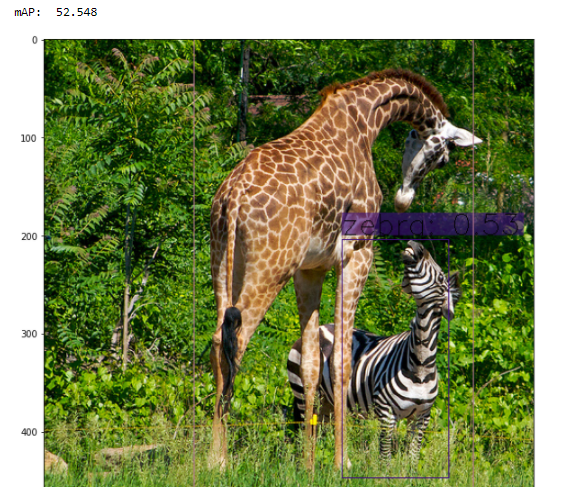


Each image is classified with bounded boxes and predicted class, IoU score is displayed above the bounded box. All the objects in the images are classified correctly.

This model performed good in all three models with good IoU and mAP Score.

1. YOLOV3 SPP:





Here, the model had predicted objects with good IoU and mAP. But it couldn’t predict some objects like giraffe and horse in the above images.

1. YOLOV3 Tiny:



A giraffe and a zebra in a zoo exhibit

Description automatically generatedA picture containing text, grass, outdoor, mammal

Description automatically generated

In this model, some objects are predicted differently from the above two models.

1. YOLOV3 Ensemble:



A picture containing text, mammal

Description automatically generated

As Ensemble is a combination of individual models, this model is predicted based on three predicted models.

This model has predicted objects correctly with good IoU and mAP greater than 70.

Chapter 5

**CONCLUSION AND FUTURE ENHANCEMENTS**

This project presented a work on ensemble learning, genetic algorithm and evolutionary learning. Many computer vision problems, NP-Hard problems are solved with good performance. Ensemble model gives good performance by combining the results. The metrics like Intersection over Union and mean Average precisions are evaluated, got the good score performance in each model. All the objects are predicted correctly. While in all three models, yolov3 320 weights model had performed better. Also, this research work has solved one of the challenges by using trained yolo models.

Next step includes object detection using a genetic algorithm. For EAs, fitness function must be evaluated correctly based on the objective of the problem. My future work involves investigating the support of ensemble learning with genetic algorithm on object detection problem.

Chapter 6

REFERENCES

1. 1, R. X. (2021). A Forest Fire Detection System Based On Ensemble Learning. Multidisciplinary Digital Publishing Institute.
2. A.E. Eiben, J. V. (1998). Graph Coloring With Adaptive Evolutionary Algorithms. Journal Of Heuristics.
3. Cai, W. (2019). A Hybrid Ensemble Learning Framework For Basketball Outcomes Prediction. Sciencedirect.
4. Deb, K. (2003). Reliable Classification Of Two-Class Cancer Data Using Evolutionary Algorithms. Sciencedirect.
5. Deveci, M. (2018). Evolutionary Algorithms For Solving The Airline Crew Pairing Problem. Sciencedirect.
6. Dey, J. (2016). Moving Object Detection Using Genetic Algorithm For Traffic Surveillance. International Conference On Electrical, Electronics, And Optimization Techniques (Iceeot) - 2016.
7. Divina, F. (2018). Stacking Ensemble Learning For Short-Term Electricity Consumption Forecasting. Multidisciplinary Digital Publishing Institute.
8. Hecker, F. T. (2013). A Case Study On Using Evolutionary Algorithms To Optimize Bakery Production Planning. Sciencedirect.
9. Hegde, V. (2020). Yet Another Deep Learning Approach For Road Damage Detection Using Ensemble Learning. Ieee International Conference On Big Data (Big Data).
10. Li, L. (2021). Evolutionary-Based Federated Ensemble Learning On Face Recognition. Imcec.
11. Mart´Inez-G´Omez, J. (2010). Using Genetic Algorithms For Real-Time Object Detection. Springer.
12. R. Aruna Jyothi, K. R. (2019). Moving Object Detection Using The Genetic Algorithm For Real Times Transportation. International Journal Of Engineering And Advanced Technology (Ijeat).
13. Rajaraman, S. (2020). Analyzing Inter-Reader Variability Affecting Deep Ensemble Learning For Covid-19 Detection In Chest Radiographs. Plos One.
14. Sai Shilpa Padmanabula, R. C. (2020). Object Detection Using Stacked Yolov3. International Information And Engineering Technology Association.
15. Shim, V. A. (2011). Evolutionary Algorithms For Solving Multi-Objective Travelling Salesman Problem. Springer.
16. Xu, G. (2019). Bearing Fault Diagnosis Method Based On Deep Convolutional Neural Network And Random Forest Ensemble Learning. Multidisciplinary Digital Publishing Institute.
17. Yu, L. (2008). Credit Risk Assessment With A Multistage Neural Network Ensemble Learning Approach. Sciencedirect.