

Efficient and Lightweight Traffic Sign Detection using YOLO V5 Algorithm

Papasahebgari Jakeer Hussain¹, Radhika N² and Remyakrishnan P³

¹⁻³Department of Computer Science and Engineering, Amrita School of Computing, Coimbatore, Amrita Vishwa Vidyapeetham, India

Email: cb.en.p2aie21019@cb.students.amrita.edu, {n_radhika, p_remyakrishnan}@cb.amrita.edu

Abstract—An essential component of many user applications today, as well as security surveillance systems, text recognition, and some diagnosis for diseases from few scans like CT or MRI, is object identification, a computer vision problem. Also if we can see one of the most important components like object detection which can help detect any object etc. The traffic sign recognition is also a main important task which have a crucial role in intelligent autonomous vehicles. In our work we have implemented different models of YOLO v5 (You only look once) algorithm for traffic sign detection. To train yolo algorithm for the traffic sign detection on German Traffic Sign detection benchmark which have multiple scenario of images like moving vehicle images, clear images, uncleared images, and different environmental conditional images all these will help in to detect the four different classes and improve the mAP of the model with the existing work and comparing results of different models of Yolo v5 algorithm with the existing work.

Index Terms— YOLO, GTSDb, Traffic Sign Detection.

I. INTRODUCTION

A self-driving vehicle referred as an autonomous vehicle has the ability to sense its surroundings and move safely with little bit or no human intervention. The fundamental goal of computer vision is to identify image. When we see the world with our own eyes, our brains are continually engaged in doing classification activities. Furthermore, for autonomous vehicles an accurate object detection and classification algorithms are crucial to enabling the vehicle to see the world similarly to humans. The objectives of our work is to create a model which will be able to detect traffic signs provides sufficient performance which can be used in real-time applications. Metrics like box_loss, objectness loss, and classification loss these are all the loss functions which the yolo has to evaluate the results. Where box_loss is an mean squared error, the confidence of object present is the obj_loss and classification loss is the cross entropy.

II. LITERATURE SURVEY

[1] In this paper authors started with the explanation of one stage and two stage detectors and explained different models of one stage and two stage detectors. The authors have discussed about the few most used datasets which will be used for detection and also the applications of the algorithms. Further latest trends in the detection algorithms have been explained.

- [2] In this paper authors used 2 models in which first model is trained to locate traffic signs into 4 categories and the model 2 has been trained to classify fragments located of images in to one of 43 classes in which both the models solve detection and classification issues. For detection authors have used Yolo v3 algorithm on the GTSTB which has total of 900 RGB images and have achieved 97.22mAP and model 2 i.e., CNN has been trained on GTSRB with a 66000 amount of RGB images.
- [3] In this paper authors started with the explanation of two stage and one stage detectors. Further explained the YOLO algorithm and its working with the history of the versions following to this authors have explained MobilenetSSD, SSD and Mobilenet. Further comparative analysis of the YOLO v3 and Mobilenet SSD and finally concluded YOLO has been provides better accuracy if we compare to the Mobilenet SSD which has been given more detection speed.
- [4] In this paper author has conferred literature review on the traffic sign identification by using techniques of machine learning, also the comparative study and reasoning of those techniques. Author proposed an solution which gives a voice alert message to the driver using CNN algorithm and have explained about the image preprocessing which will be done before classification which helps to reduce noise and improve precision of algorithm, convolution neural network architecture and traffic sign recognition.
- [5] In this paper author has presented the system for doing detection and recognition for traffic signs that has applied to the Spanish traffic signs dataset. Author followed 3 stages for detection and classification like segmentation, shape classification, and recognition and explained the working of all the 3 methods. After the segmentation support vector machine is used for classification which classify 4 different color signs with 2 or 3 different shapes in each color. Once the classification is done recognition is implemented by SVMs with Gaussian kernels for the training LIBSVM library for support vector machines has been used.
- [6] The mask RCNN has been selected in this study. Convolutional neural network (CNN) for doing the detection and recognition part of traffic signs has been used. Rather than using the Hough transform method, the scale invariant feature transform, local binary patterns, the authors employed CNN for broad feature extraction. They have also used the data augmentation method to defeat real-time issues with traffic sign appearance. Models like Faster RCNN and Mask R-CNN are checked using the Swedish traffic-sign dataset (STSD). Further they have created a new data set called DFG traffic-sign which help to have less inter-class variability and more intra-class variability. Average precision change been performed in Mask R-CNN method to increase the overall recall.
- [7] In this paper researchers came up with the challenges and successors for object detection using Yolo. Authors explained about the fundamental tasks involved in the object detection. Further explored the different two stage detectors like RCNN, Fast-RCNN and Faster-RCNN and explained the important applications for each of the models. Further explored the different versions of yolo and done a comparative analysis using darknet architecture. The different versions of YOLOs compared in this are Yolo v1, Yolo v2, Yolo v3, Yolo v4 and explained various applications in which each versions has been showed with performance metrics. Authors further concluded Yolo v3 is better detector, which is fast and accurate compared with different versions.
- [8] The researcher in this work has concentrated on the problems with cramped objects detection and developed an effective CNN model for cramped traffic sign detection, then comparing accuracy with R-CNN and Faster R-CNN model. Forward, backward, and loss functions are all thoroughly explained along with the CNN model. In order to down-sample this network all along the stage of feature extraction, authors have created max-pooling layers with a stride of 2 and expanded the figure of the convolutional kernels per each layer of convolution from the start. To do a optimization to this model further three strategies has been used like convolution factorization, fully connected transformation and redundant layer cropping. Dataset called Tsinghua Tencent has been used for doing the part of evaluation. The proposed methodology is not only for adequate but it also consume less memory of GPU and reduces the figure of computation.
- [9] In this paper authors come up with two algorithms which work fast for doing traffic sign detection as well as classification. For the detection part German Traffic Sign Detection Benchmark and chinese traffic sign dataset has been used and for classification german traffic sign recognition benchmark has been used. In this detection algorithm first do extracting traffic signs by using model called color probability and MSER region detector. Further SVM is used to do the classification which can refine false positives based on a odd color HOG features.
- [10] In this paper researchers came up with the challenges and open issues with the object apprehension in self driving cars. Further authors explained about expansion of the autonomous cars with their advantages also the drawbacks, levels of the autonomous vehicles have been briefed. Further discussed about the deep learning architectures for different object detection models and computer vision in autonomous vehicle.

III. DATASET

German Traffic Sign Detection Benchmark has been used for the detection process. This dataset consists of 900 RGB images. This dataset consists of images files and text files with their annotations. The dataset includes images with no traffic signs to be used for training. However, it was decided to eliminate such images. Those txt files will include all annotations of every bounding box of traffic signs in format of YOLO.

In this Dataset the traffic signs has been concerted into four different categories like prohibitory, mandatory, danger and other. The following signs of traffic under the category of prohibition are: speed restriction, no passing, no vehicle in both directions, and no trucks.

Priority at next intersection, hazard, animals, cycles crossing, pedestrain crossing, narrow road, bend, construction, slippery road and traffic signal are include in the danger category. The following traffic signs are included in the mandatory category: go straight, round about, go left or straight, go right, go left or straight, go left. The following traffic signs falls under the other category: no entry, halt, priority ends, give way, and restriction ends. This dataset contains different scenerios of weather conditions and different scenrio of images like moving, clear, blur etc.

IV. EXISTING WORK

Valentyn Sichkar, Sergey A. Kolyubin[2] describes the “Real time detection and classification of traffic sign based on YOLO version 3 algorithm”. In this paper authors used 2 models in which first model is trained to locate traffic signs into 4 categories and The model 2 is trained to classify located fragments of images into one of the 43 classes in which both the models solve detection issue. The darknet framework has been used by model 1. Every image will be having an annotations files in the yolo format where bounding boxes and class number will be reorded. Images are collected in the batches of 64 while training. Once the traffic sign is detected using the model 1 the fragment will be fed to the model 2. The model 2 have been trained on GTSRB in which the dataset is divided as 50000 images for training, 12000 images for validation and 4000 images for testing the model. This model 2 will have one convolutional layer with rectified linear unit activation function, 32 filters and one down sampling layer with 2*2 maximum factor. For detection authors have used Yolo v3 algorithm on the GTSTB which has total of 900 RGB images and have acheived 97.22mAP and model 2 i.e., CNN was trained on GTSRB with 66000 RGB images in which the result will be the one of the traffic sign out of all 43 classes.

V. PROPOSED WORK

In our proposed work the Yolo v5 algorithm has been trained for the traffic sign detection on German Traffic Sign detection benchmark helps to detect in four different classes and improve the mAP of the model with the existing work. In this for yolo v5 Pytorch framework has been used which is just an extension for yolo v3. The main objective is to design a YOLO algorithm that accurately detects traffic signs and further comparing our all different models of Yolo v5 algorithm like v5s, v5m, v5l, v5x with the existing work. In our work the stochastic graient descent optimizer has been used. Our yolo v5 model architcture consists of backbone, neck and head. We have used the cross stage partial networks as a backbone to extract the features from the input feed image. Further PANet has been used as a neck to get features and the model head has been used to perform the detection part which gives the generated final output vectors with the bounding boxes and probabilities as shown in the fig-2.

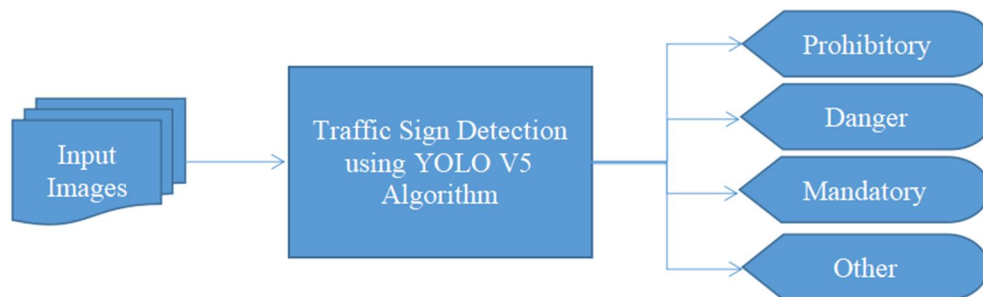


Figure 1: Flowchart of the Traffic Sign detection process using Yolo V5



Figure 2: Original Image and Traffic Sign Detected Image using Yolo v5

VI. RESULTS

TABLE I. TRAINING RESULTS OF YOLO V5S ALGORITHM

Epoch	box_loss	Obj_loss	Cls_loss	Precision	Recall	mAP
30	0.03304	0.00714	0.00498	96.8	89.3	94.6
60	0.02497	0.00578	0.00260	97.7	92.1	97.1
90	0.02096	0.00481	0.00208	98.0	95.2	98.0
120	0.01728	0.00418	0.0018	98.4	94.4	97.7
150	0.01318	0.00387	0.00115	99.1	95.4	97.5

The above Table 1 shows the training results of the yolo v5s algorithm in the different epochs range. As we can see the precision and recall has been increased with the epochs and achieved the highest as 99.1% and 95.4% with the 150th epoch. Where as the best mAP has been achieved as 98% in the 90th epoch.

TABLE II. TRAINING RESULTS OF YOLO V5M ALGORITHM

Epoch	box_loss	Obj_loss	Cls_loss	Precision	Recall	mAP
30	0.03187	0.00604	0.00451	97.0	92.2	97.0
60	0.02322	0.00512	0.00249	98.0	96.3	98.7
90	0.01949	0.00418	0.00201	99.4	95.5	97.5
120	0.015	0.00351	0.00177	99.2	93.7	97.4
150	0.01087	0.00314	0.00108	98.4	96.1	97.8

The above Table 2 shows the training results of the yolo v5m algorithm in the different epochs range. As we can see the the model has achieved the best precision results at 120th epoch i.e., 99.2% and the recall has got a good result which is 96.3% at 60th epoch. Where as the best mAP has been achieved as 98.7% in the 60th epoch.

TABLE III. TRAINING RESULTS OF YOLO V5L ALGORITHM

Epoch	box_loss	Obj_loss	Cls_loss	Precision	Recall	mAP
30	0.03149	0.00573	0.00387	96.2	95.9	97.6
60	0.02272	0.00468	0.00226	99.4	95.5	98.0
90	0.01785	0.00382	0.00175	98.9	96.8	97.9
120	0.01436	0.00319	0.00151	98.8	95.8	98.2
150	0.00996	0.00273	0.00102	98.5	97.9	98.1

The above Table 3 shows the training results of the yolo v5l algorithm in the different epochs range. As we can see the the model has acheived the best precison results at 60th epoch i.e., 99.4% and the recall has got a good result which is 97.9% at the 150th epoch. Where as the best mAP has been achieved as 98.2% in the 120th epoch.

TABLE IV. TRAINING RESULTS OF YOLO V5X ALGORITHM

Epoch	box_loss	Obj_loss	Cls_loss	Precision	Recall	mAP
30	0.03201	0.00570	0.00404	99.3	95.9	98.2
60	0.02258	0.00460	0.00210	99.2	97.6	98.8
90	0.01818	0.00368	0.00169	98.5	97.6	98.8
120	0.01384	0.003085	0.00159	99.4	96.4	98.3
150	0.00948	0.00267	0.00097	98.7	97.5	98.3

The above Table 4 shows the training results of the yolo v5m algorithm in the different epochs range. As we can see the the model has acheived the best precison results at 120th epoch i.e., 99.4% and the recall has got a good result which is 97.6% at the 90th epoch. Where as the best mAP has been achieved as 98.8% in the 90th epoch.

TABLE V. PRECISION AND MAP COMPARISON WITH THE EXISTING WORK[2]

Model	Precision				mAP
	Prohibitory	Danger	Mandatory	Other	
Yolo v3	100	100	89.2	90.9	97.22
Yolo v5s	99.2	100	98.5	99	98.0
Yolo v5m	99.2	98.8	96.2	100	98.7
Yolo v5l	98.3	99.8	96.3	100	98.2
Yolo v5x	99.7	100	95.8	99.1	98.8

The above Table 5 shows the precision of all 4 classes and mAP metrics comparison of different yolo v5 models with the existing work. As we can see that the mAP performance is better than the existing work for all the models of yolo v5 and the highest mAP achieved around 98.8% with yolo v5x model.

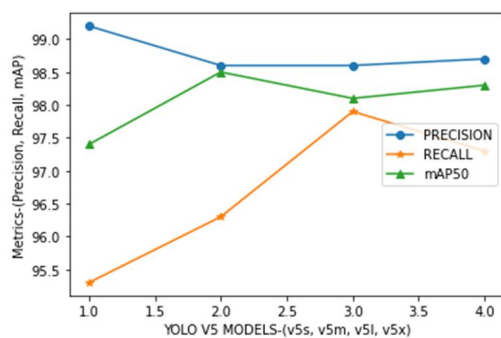


Figure 3: Metrics comparison of different Yolo v5 models

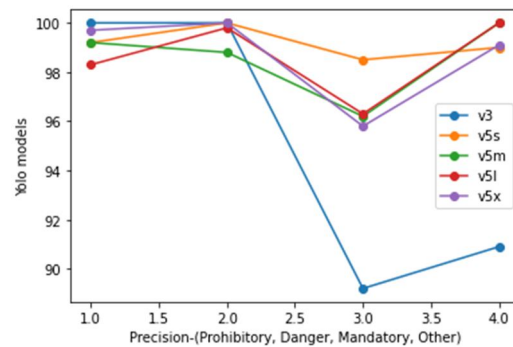


Figure 4: Precision comparison of different Yolo v5 models with existing work[2]

Figure 3 shows the different metrics comparison for different models of yolo v5. where the precision is represented in blue color, recall is represented using orange and mAP using green color.

The above Figure 4 represents the precision comparison of different yolo v5 models with the existing work[2]. In this the precision has been compared for all the four different classes.



Fig-5: Results of sample detected traffic sign image using Yolo v5

VII. CONCLUSION

The study gives the detection issue for different variety of traffic signs. As for doing the detection version 5 Yolo model has been trained on German Traffic sign Detection Benchmark to predict all the locations among the 4 categories. Our work has been showed the training of deep network with 4 categories to detect traffic signs achieves high mAP reaching 98.8 using yolo v5x. Further we have done the comparison of different yolo v5 models with the previous existing work. The main aim for our model is giving the improvement in mAP and get the better results than the previous existing work using the different yolo v5 models. As a future work we have been planned to do the segmentation and classification of the traffic signs which uses the real time images for training and testing our models.

REFERENCES

- [1] Jiao L, Zhang F, Liu F, Yang S, Li L, Feng Z, Qu R. "A survey of deep learning-based object detection". IEEE access, Volume 7, 2019, ISSN 2169-3536
- [2] Valentyn Sichkar, Sergey A. Kolyubin. "Real time detection and classification of traffic sign based on YOLO version 3 algorithm". Scientific and Technical Journal of Information Technologies Mechanics and Optics, Volume 20, 2020, ISSN 2226-1494.
- [3] Sabina N, Aneesa M.P, Haseena P.V. "Object Detection Using YOLO and Mobilenet SSD: A Comparative Study". International Journal of Engineering Research and Technology, Volume 11 Issue 26, June 2022, ISSN 2278-0181.
- [4] Vaibhavi Golgire. "Traffic sign Recognition using Machine Learning: A Review". International Journal of Engineering Research and Technology, Volume 10, May 2021, ISSN 2278-0181.
- [5] Saturnino Maldonado-Bascon, Sergio Lafuente-Arroyo, Pedro Gil-Jimenez, Hilario Gomez-Moreno, Francisco Lopez-Ferreras. "Road-Sign Detection and Recognition Based on Support Vector Machines". Volume 8, June 2007.
- [6] Domen Tabernik, Danijel Skocaj. "Deep Learning for Large-Scale Traffic-Sign Detection and Recognition". IEEE Transactions of Intelligent Transportation Systems, Volume 21, April 2020.
- [7] Diwan, T., Anirudh, G. & Tembhurne, J.V. Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimed Tools Appl* (2022). <https://doi.org/10.1007/s11042-022-13644-y>.
- [8] S. Song, Z. Que, J. Hou, S. Du and Y. Song, "An efficient convolutional neural network for small traffic sign detection", *Journal of Systems Architecture*, vol. 97, pp. 269- 277, 2019. Available: 10.1016/j.sysarc.2019.01.012.
- [9] Yi Yang, Hengling Luo, Haurang Xu, Fuchao Wu. "Towards Real Time Traffic Sign Detection and Classification" *IEEE Transactions of Intelligent Transportation Systems*, Volume 17, April 2016.
- [10] Abhishek Gupta, Algaan Anapalagam, Ling Guan, Ahmed Shaharyar Khwaja "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues", *Array*, Volume 10, 2021, 100057, ISSN 2590-0056.
- [11] Mehdi Masoudi, Hamdi Friji, Hakim Ghazzai, Yehia Massoud "A Reinforcement Learning Framework for Video Frame-Based Autonomous Car-Following", *IEEE Open Journal of Intelligent Transportation Systems*, Volume 2, 2021, ISSN 2687-7813.
- [12] Aleksej Avramovic, Davor Sluga, Domen Tabernik, Daniel Skocaj, Vladan Stojnic, Nejc Ilc. "Neural-Network-Based Traffic Sign Detection and Recognition in High-Definition Images using Region Focusing and Parallelization". *IEEE Access* Volume 8, ISSN 2169-3536.
- [13] Arnold E, Al-Jarrah OY, Dianati M, Fallah S, Oxtoby D, Mouzakitis A. "A survey on 3d object detection methods for autonomous driving applications". *IEEE Trans Intell Transport Syst*, Volume 20, 2019, ISSN 1524-9050.

- [14] Ersaa Khatab, Ahmed Onsy, Martin Warley, Ahmed Abouelfarag, "Vulnerable Object Detection for autonomous Driving: A Review", Volume 78, Integration, 2021.
- [15] K. N. R. Chebrolu and Dr. (Col.) Kumar P. N., "Deep Learning based Pedestrian Detection at all Light Conditions", in Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019, 2019, pp. 838-842.
- [16] H. Haritha and Dr. Senthil Kumar T., "A modified deep learning architecture for vehicle detection in traffic monitoring system", International Journal of Computers and Applications, pp. 1-10, 2019.
- [17] O. K. Sikha and Dr. Soman K. P., "Multi-resolution Dynamic Mode Decomposition-based Salient Region Detection in Noisy Images", Signal, Image and Video Processing, 2020.
- [18] K. Hena, Amudha J., and Aarthi, R., "A Dynamic Object Detection In Real-World Scenarios", Proceedings of International Conference on Computational Intelligence and Data Engineering (Lecture Notes on Data Engineering and Communications Technologies), vol. 28. Springer Singapore, Singapore, pp. 231-240., 2019.
- [19] S. A. Shreyas, Abishek, S., and Dr. Radhika N., "Performance analysis of linear grid stability using classifiers and advanced ensemble techniques", Journal of Advanced Research in Dynamical and Control Systems, vol. 11, no. 3, pp. 1574-1579, 2019.
- [20] Chandrakala, S., Deepak, K. and Revathy, G., 2022. Anomaly detection in surveillance videos: a thematic taxonomy of deep models, review and performance analysis. Artificial Intelligence Review, pp. 1-50. (IF-9.55)
- [21] "A hybrid architecture for Action Recognition in Videos using Deep Learning", V S Akella, C Vijayendra Sai, Kakarla Ajay Kumar Reddy, Priyanka Kumar presented at 3rd International Conference on Data Science and Applications (ICDSA 2022), Jadavpur University, Kolkata, India, March 26-27, 2022.
- [22] A. Ćorović, V. Ilić, S. Đurić, M. Marijan and B. Pavković, "The Real-Time Detection of Traffic Participants Using YOLO Algorithm," 2018 26th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2018, pp. 1-4, doi: 10.1109/TELFOR.2018.8611986.
- [23] Namman Mittal, Akarsh Vaidhya, Asst.Prof.Shreya Kapoor. "Object Detection and Classification using Yolo" International Journal of Scientific Research & Engineering Trends, Volume 5, Issue 2, Mar-Apr-2019, ISSN 2395-566X.[7] M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.