International Journal of
HRM and Organizational Behavior

www.ijhrmob.com

editor@ijhrmob.com

International Journal of
HRM and Organizational Behavior

ISSN 2454 - 5015

MCS-YOLO: AN AUTOMATED DRIVING ROAD ENVIRONMENT RECOGNITION SYSTEM WITH MULTISCALE OBJECT DETECTION

A DHANASEKHAR REDDY[1], C CHAITHANYA[2], K YATHEENDRA[3], P GOPICHAND[4]
[1]Assistant Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email: dhanasekhar918@gmail.com
[2]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email: chaithucherry1430@gmail.com
[3]Associate Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email: k.yatheendra84@gmail.com
[4]Assistant Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email:pgopichand1@gmail.com

Abstract: Improving object identification algorithms' accuracy and speed is a critical issue in autonomous driving technology. Our MCS-YOLO procedure adds a direction consideration module to the backbone to further develop feature map spatial direction and cross-channel collection. Moreover, a multiscale little object detection structure builds aversion to thick minuscule articles, and CNNs utilize the Swin Transformer construction to focus on relevant spatial data. The BDD100K autonomous driving dataset shows that MCS-YOLO beats YOLOv5s in mean normal accuracy and recall rates. Curiously, our innovation detects 55 casings each second continuously driving circumstances. Further testing with YoloV5x6 shows a 0.798% mean typical accuracy increment. This exploration gives a solid and viable strategy for further developing autonomous driving object recognition, encouraging canny transportation frameworks.

Index terms *- Coordinate attention mechanisms, autonomous driving, road environmental object detection, swin transformer, YOLOv5.*

1. INTRODUCTION

In the 21st hundred years, new vehicle enlistments and authorized drivers have expanded as vehicles become more normal around the world. In any case, this quick development in engine vehicles has caused car crashes, blockage, and natural issues. Autonomous driving technology further develops wellbeing and course arranging [1], [2].

The ecological perception system, which rapidly and precisely recognizes street objects, supports autonomous driving. Decision systems need this information to enhance course arranging [3]. The early improvement of independent driving utilized costly single or multi-sensor combination moves toward that expected manual vehicle boundary change and human intercession. Nonetheless, progresses in profound picking up, detecting, and innovation have made computer vision (CV) and natural language processing (NLP) more proficient.

Girshick et al's. R-CNN model superior acknowledgment [5]. Later progressions like He et al's. Spatial Pyramid Pooling (SPP) [6], Fast R-CNN [7], and Faster R-CNN [8] with a Region Proposal Network (RPN] upgraded detection accuracy and handling effectiveness. Great detection and segmentation are added utilizing Cover R-CNN [9].

These advances exhibit the groundbreaking capability of DL-based object identification calculations for continuous, precise, and conservative independent vehicle ecological detecting.[47]

The You Only Look Once (YOLO) and Single Shot MultiBox Detector (SSD) calculations use regression approaches for object order and bounding box prediction. The YOLO calculation inputs the total picture and regresses bounding box area and class in the result layer. Industry involves YOLO and SSD calculations for faster constant detection than R-CNN. A Transformer-based convolutional neural network for thick vision applications was utilized by Liu et al.

The Swin Transformer [20], [21] shows its power for classification, detection, and division. ConvNext [22] trains convolutional neural networks utilizing Swin Transformer's enhancement approach. ConvNext outflanks Swin Transformer in deduction and accuracy with similar Failures. Chen et al. [23] formulated a DW-YOLO procedure that increments network profundity and expansiveness to perceive vehicle objects. Zhou et al. [24] recommended a lightweight MobileYOLO method that limits boundaries and paces detection. Wang et al. [25] involved MobileNet on YOLOv4 for driving circumstances and got 35 FPS detection. A superior SA-YOLOv3 detector by Tian et al. [26] balances speed and accuracy. Gupta et al. [27] utilized detection and segmentation to further develop self-driving vehicle versatile way of behaving by identifying street climate objects. Wang et al. [28] present an autonomous driving detection network for foggy circumstances that improves object identification accuracy and speed. Li et al. [29] fostered a Res-YOLO network model that limited missed detections and upgraded vehicle object detection accuracy.

2. LITERATURE SURVEY

This paper surveys the creators' momentum research on protected and strong autonomous driving in metropolitan settings with unforeseen traffic. The review incorporates constant innovations for climate detecting, restriction, arranging, and control to fabricate a completely utilitarian vehicle stage. [1] The creators' work on Junior, Stanford's 2007 DARPA Metropolitan Test passage, is extended to cover more practical driving conditions. The creators depict three unsupervised techniques that consequently align a 64-pillar turning LIDAR with preferred exactness over hand perceptions. Online confinement with centimeter-level accuracy requires high-goal ecological guides. Obstruction following, bicycle, person on foot, and vehicle detection, and traffic light detection are conceivable with further developed insight and recognizable proof calculations [6,29,39]. In light of approaching information, a progressive arranging framework makes many potential directions each second to improve the vehicle's course. An updated regulator upgrades choke, brake, and controlling to augment solace and limit direction error. These calculations function admirably in each climate, day or night. Junior has driven independently for many kilometers in different genuine settings on account of these advancements.

This study examined the quick advances in ML, computer vision, AI, and autonomous vehicles [2]. The creators give a definite outline to assist experienced scientists and novices with staying aware of this quick extending point. This book gives an exhaustive presentation of independent vehicle PC vision challenges, datasets, and approaches, in contrast to earlier examinations. The review covers authentic writing and current advances in autonomous driving fields such distinguishing proof, recreation, movement gauges, following, scene translation, and start to finish learning. The creators use benchmarking datasets including KITTI, Saying, and Cityscapes to assess algorithmic execution. Open worries and proceeding with research difficulties make the review applicable to autonomous vehicle advances [2,4,27]. The creators give a devoted site to smooth route among subjects and approaches, giving setting and data to further develop openness and correct missing references. This exhaustive appraisal helps scholastics, professionals, and rookies understand the advancing climate of computer vision in autonomous vehicles.

The impending send off of autonomous vehicles and the need to give security, reliability, and an agreeable client experience for general acknowledgment. As client solace in driving styles goes from lively to quiet, the creators propose a gaining from show methodology to tailor autonomous vehicle conduct. [4] Clients may physically drive the vehicle to represent their ideal driving style, staying away from the burdensome and blunder inclined errand of physically tweaking speed increase profiles, distances to different vehicles, and path change speed. An expense capability models the driving style, and element based converse support learning tracks down the model boundaries that best fit it. The model successfully figures vehicle directions in autonomous mode in the wake of getting the hang of, permitting it to repeat and adjust to various driving styles. It can learn cost works and impersonate driving ways of behaving utilizing genuine driver information, demonstrating its convenience. This client driven procedure works on the independent vehicle's responsiveness to individual inclinations and coordinates autonomous innovation into changed client encounters.[48]

The PASCAL VOC dataset's stale thing detection execution is tended to with a creative, basic, and adaptable recognition approach that significantly works on mean average precision. The technique [5] accomplishes a staggering 53.3% Guide, surpassing the past high by practically 30%. The technique consolidates two principal experiences: the utilization of high-limit Convolutional Neural Networks (CNNs) to handle base up area proposition for exact item restriction and division, and the adequacy of supervised pre-training for an assistant assignment followed by space explicit tweaking, particularly in circumstances with restricted labeled training information. R-CNN (Regions with CNN features) utilizes these bits of knowledge to support execution. R-CNN beats OverFeat, a sliding-window detector in view of a comparative CNN engineering, on the 200-class ILSVRC2013 detection dataset [5,7,8,17,18]. R-CNN's outcome demonstrates the way that area proposals can build CNN execution, conquer past cutoff points, and further develop object detection.

Existing deep convolutional neural networks (CNNs) that need fixed-size input pictures lose acknowledgment accuracy for pictures or sub-pictures of various sizes. The technique utilizes "spatial pyramid pooling" in SPP-net, another organization structure [6]. This plan produces a fixed-length portrayal free of picture size or scale, making it more versatile. CNN-based picture order is improved by SPP-net's item disfigurement obstruction. The article shows that SPP-net upgrades CNN design accuracy on ImageNet 2012. SPP-net produces cutting edge characterization scores on Pascal VOC 2007 and Caltech101 datasets utilizing a solitary full-picture portrayal and no calibrating. In object detection, SPP-net paces feature map processing and beats R-CNN [39]. In the ImageNet Large Scale Visual Recognition

Challenge (ILSVRC) 2014, SPP-net methodologies put #2 in object recognition and #3 in picture classification among 38 groups. The book depicts significant cutthroat upgrades that exhibit SPP-net's flexibility and proficiency in visual recognizable proof undertakings.

3. METHODOLOGY

i) Proposed Work:

Our recommended framework involves another MCS-YOLO calculation for autonomous driving object detection and recognition with a direction consideration module, multiscale tiny object detection structure, and Swin Transformer. This strategy attempts to further develop object detection speed and accuracy. Following broad removal tests and correlation preliminaries on the BDD100K dataset [41], our MCS-YOLO calculation performed best with a 53.6% mean average accuracy (mAP). Our recommended framework utilizes Yolov5x6 and YoloV8 modern ways to deal with further develop detection. These strategies try to build the mAP above 60% for dependable and powerful object identification. The calculations Faster RCNN, AD-Faster RCNN, YoloV3, YoloV3-tiny, YoloV4, YoloV5s, YoloV5s Further developed Form, Yolo V7-tiny, Yolo V5x6, Yolo V8, and MCS YoloV5s assess detection abilities in assorted settings [12,13,14,15,23,24]. This diverse technique streamlines the autonomous driving natural insight framework to further develop object recognition for protected and reliable vehicle route.

ii) System Architecture:

The unequivocally evolved system architecture flawlessly processes information from contribution to image processing utilizing progressed data augmentation methods. Model creation utilizes a large number of modern models, including YoloV5s, YoloV5s Improved Version, MCS YoloV5, Yolo V5x6, YoloV4, YoloV3, YoloV3-little, Yolo V7, Yolo V8, Faster RCNN, and AD-Faster RCNN [12,13,14,15,23,24]. Model precision, recall, and mean average precision (mAP) are assessed [40]. Utilizing these actions, the best object detection model is picked. This total design smoothes out and works on autonomous driving by areas of strength for offering exact discernment in various street settings. A versatile model set assists autonomous vehicle innovation with improving by permitting the framework to prevail in different settings.[50]

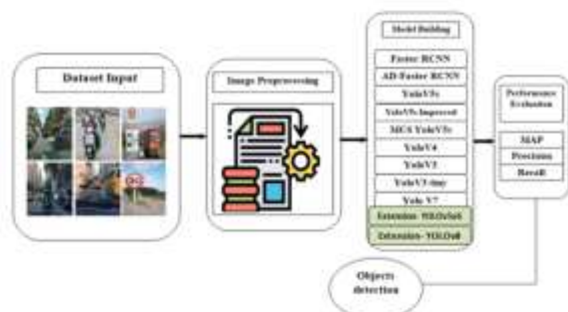


Fig 1 System Architecture

iii) Dataset collection:

The BDD100K dataset [41], known for its accuracy and fulfillment, is utilized to assess the MCS-YOLO calculation in autonomous driving discernment. This legitimate public dataset incorporates eleven objective classifications and different climate, driving, and season of day settings from genuine situations. The 100,000-picture assortment contains brilliant, cloudy, blustery, cold, and hazy settings. 20,000 unlabeled photographs were erased to work on model approval, and the excess information was parceled 8:1:1 for preparing, approval, and testing sets. The preparation set has 64,800 pictures, the approval set 7,200, and the test set 8,000. Object focus focuses are moved in the focal picture locale, guaranteeing a uniform dispersion of items and an enormous portrayal of little focuses in the dataset, giving a strong groundwork to surveying the MCS-YOLO calculation's autonomous driving discernment execution.[52]

Fig 2 Dataset images

iv) Image Processing:

Autonomous driving systems use image processing to recognize objects in different levels. Advancing the information picture for examination and change starts with mass item transformation. Following this, the calculation's objective classifications are determined by characterizing object classes. Bounding boxes are additionally characterized to show where things ought to be in the image. Changing over handled information into a NumPy exhibit is fundamental for mathematical calculation and investigation.

Stacking a pre-prepared model with huge datasets follows. This includes getting to the pre-prepared model's organization layers, which incorporate learnt elements and boundaries for compelling object identification. Extraction of result layers gives last forecasts and helps object recognition and classification.

Annexing the image and explanation record in the image processing pipeline guarantees total information for examination. Switching BGR over completely to RGB changes the variety space, and a cover features significant qualities. A last resize streamlines the picture for handling and examination. This total picture handling procedure lays the preparation for strong and accurate object recognition in autonomous driving systems' dynamic setting, further developing street wellbeing and navigation.

v) Data Augmentation:

Data augmentation is fundamental for creating assorted areas of strength for and datasets for ML models, particularly in image processing and PC vision. The first dataset is improved by randomizing, pivoting, and twisting the picture.

Picture fluctuation is made by randomizing brilliance, difference, and variety immersion. This stochastic strategy works on model speculation to new information and different conditions.

Changing the picture's direction by degrees is called pivot. This expansion technique helps the model to recognize objects from different points, duplicating true conditions.

Scaling, shearing, and flipping change the image. These contortions look like certifiable item look and direction, advancing the dataset.

These information increase techniques extend the preparation dataset, assisting the model with procuring vigorous elements and examples. This upgrades the model's speculation and execution on various and troublesome test conditions. Information expansion decreases overfitting, work on model execution, and further develop ML model constancy, quite in autonomous driving picture acknowledgment.

vi) Algorithms:

YoloV5 (You Only Look Once) distinguishes protests rapidly and accurately. It frameworks an image and predicts bounding boxes and class probabilities for every cell. **YoloV5s**, the more modest variation, balances execution and productivity.

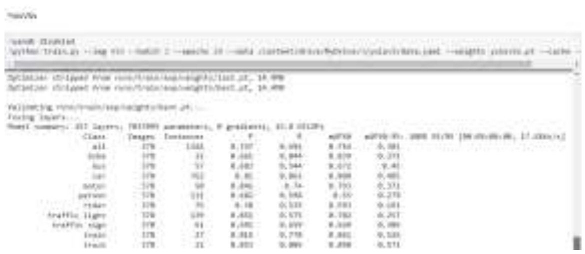


Fig 3 YOLOV5s

YoloV5s Improved Version: This contains engineering, training, and hyperparameter tuning enhancements over YoloV5s. Upgrades try to further develop object detection accuracy and effectiveness.

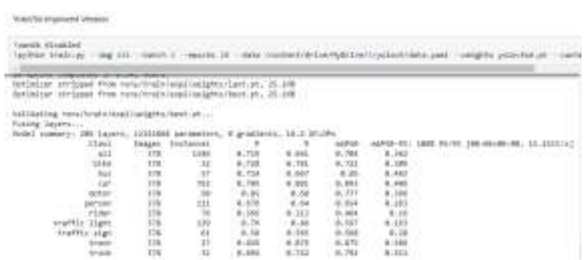


Fig 4 YOLOV5s improved version

This exploration presented **MCS YoloV5s** with a direction consideration module for spatial and cross-channel data conglomeration. Moreover, its multiscale little object detection structure further develops responsiveness and thick little article acknowledgment. Swin Transformer reconciliation further develops the organization's logical spatial data accentuation [40].

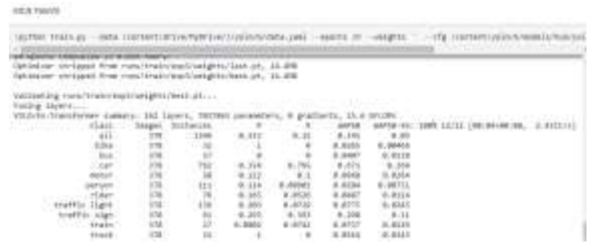


Fig 5 MCS YOLOV5s

YoloV4 progresses the Yolo series with speed and accuracy. It further develops object detection with CSPDarknet53 as a backbone, PANet, and SAM block.

Fig 6 YOLOV4

YoloV3, a more established rendition of the series, has a three-stage recognizing method. Bounding boxes at various scales are anticipated utilizing a Darknet-53 backbone. YoloV3 adjusts object detecting speed and accuracy.

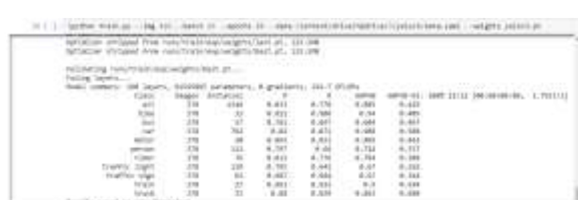


Fig 7 YOLOV3

ISSN 2454 - 5015

YoloV3-tiny is a lightweight rendition for faster surmising on asset obliged frameworks. It's quick yet less precise, making it fitting for constant applications.

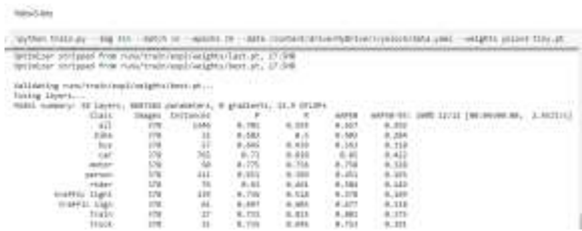


Fig 8 YOLOV3-tiny

YOLOv7, a better adaptation, joins components from YOLOv4, Scaled YOLOv4, and YOLO-R. The E-Spirit further develops learning, and Compound Model Scaling allows you to modify width, profundity, and goal autonomously. With its speed, adaptability, and accuracy progressively object identification, YOLOv7 addresses the undertaking's issues.

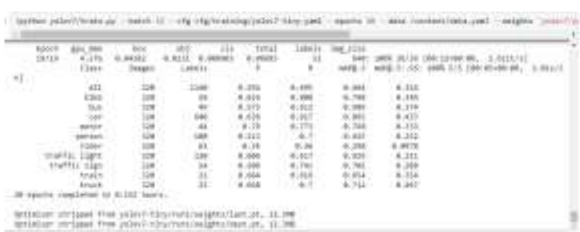


Fig 9 YOLOV7

Faster RCNN is a two-stage object detection framework. It arranges Region Proposal Network (RPN) - proposed districts of interest.[54]

Fig 10 Faster RCNN

AD-FRCNN further develops object detection execution by adding a unique region proposal network, a visual consideration plot for feature age, and a versatile unique training module [42].

Fig 11 AD-FasterRCNN

Yolo V5x6, a quick and exact type of the YOLO object detection model is enhanced for this venture. Its matrix based bounding box and class likelihood prediction gives it multiple times the handling limit. Fast deduction and exact object identification are fundamental for autonomous driving innovation in fluctuated street conditions, and this registering increment is fundamental for fulfilling project necessities.

ISSN 2454 - 5015

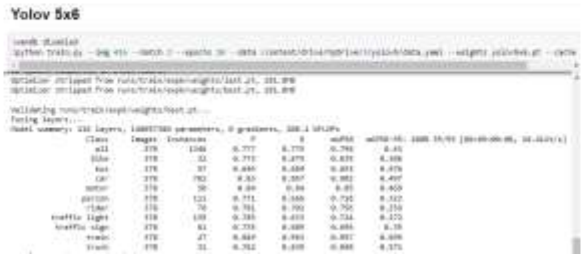


Fig 12 YOLOV5x6

YOLOv8, The YOLO series' top entertainer distinguishes many articles all the while by gridding pictures and assessing bounding boxes and class probabilities. It upholds Object Detection, Instance Segmentation, and Image Classification with an easy to use Programming interface and high accuracy and speed. New engineering with C2f modules and a sans anchor head further develops proficiency and flexibility. For this undertaking, YOLOv8 was picked for powerful, continuous object recognition.

Fig 13 YOLOV8

4. EXPERIMENTAL RESULTS

Precision: Precision estimates the level of positive cases or tests precisely sorted. Precision is determined utilizing the recipe:

$$\text{Precision} = \text{True positives}/\ (\text{True positives} + \text{False positives}) = TP/(TP + FP)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

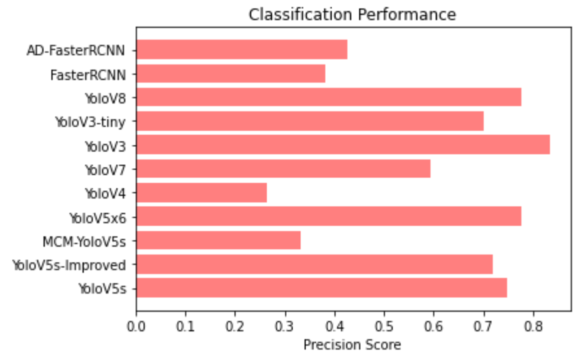


Fig 14 Precision comparison graph

Recall: Machine learning recall assesses a model's ability to perceive all significant examples of a class. It shows a model's culmination in catching occasions of a class by contrasting accurately anticipated positive perceptions with complete positives.

$$Recall = \frac{TP}{TP + FN}$$

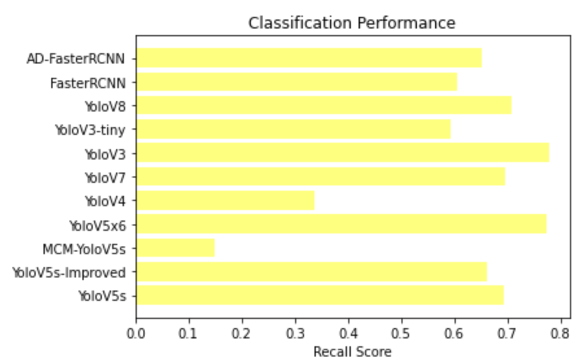


Fig 15 Recall comparison graph

mAP: Positioning quality measure Mean Average Precision (MAP). Number of significant ideas and rundown position are thought of. MAP at K is the math mean of Average Precision (AP) at K across all clients or inquiries.[56]

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{ the AP of class } k$$

$$n = \text{ the number of classes}$$

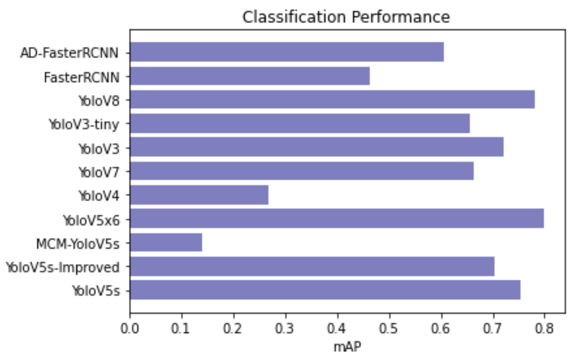


Fig 16 mAP comparison graph

	ML Model	Precision	Recall	mAP
0	YoloV5s	0.747	0.694	0.754
1	YoloV5s-Improved	0.719	0.661	0.704
2	MCM-YoloV5s	0.333	0.150	0.141
3	**YoloV5x6**	**0.777**	**0.775**	**0.798**
4	YoloV4	0.265	0.338	0.269
5	YoloV7	0.594	0.695	0.664
6	YoloV3	0.833	0.778	0.720
7	YoloV3-tiny	0.701	0.593	0.657
8	**YoloV8**	**0.776**	**0.708**	**0.782**
9	FasterRCNN	0.382	0.606	0.463
10	AD-FasterRCNN	0.427	0.653	0.605

Fig 17 Performance Evaluation table

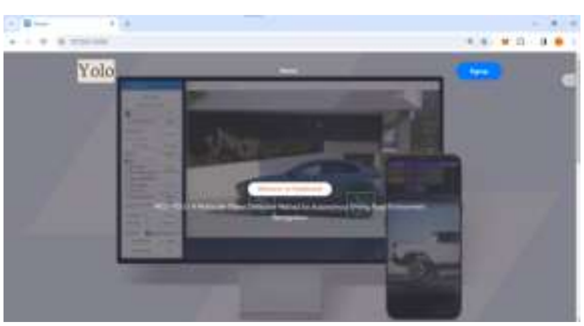


Fig 18 Home page

Fig 19 Registration page

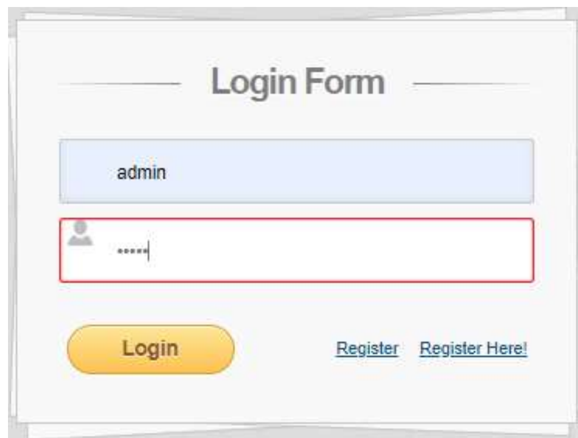


Fig 20 Login page

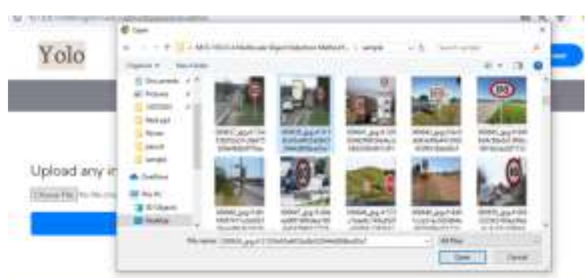
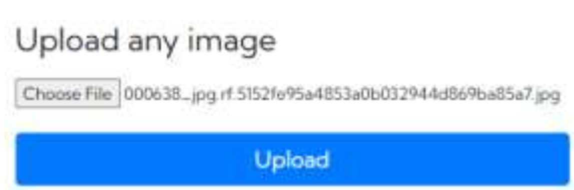


Fig 21 Input image folder



ISSN 2454 - 5015

Fig 22 Upload input image

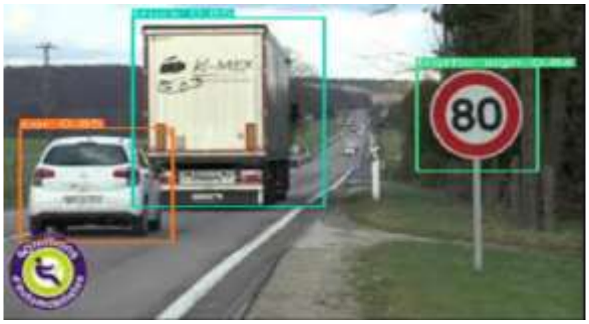


Fig 23 Predict result for given input

5. CONCLUSION

At long last, we show the MCS-YOLO calculation's prevalence in autonomous driving object identification. The technique upgrades ID accuracy and speed with a direction consideration module, multiscale tiny object detection structure, and Swin Transformer. Removal studies and BDD100K dataset correlations [41] show its better presentation over existing strategies. Future work includes utilizing MCS-YOLO to Multiple Object Tracking (MOT) to guarantee its adaptability and strength in autonomous driving situations. This drive tends to the critical need to work on autonomous driving wellbeing in the midst of rising mishaps and blockage. We work on autonomous portability by advancing ecological discernment with DL calculations like YOLOv5s, an overhauled adaptation, and MCS-YOLOv5s [25,46]. Our mechanical brightness is shown through benchmark correlations, modern model investigation, and connection with Flask and SQLite for client testing. Our drive further develops wellbeing, effectiveness, contamination, and autonomous driving advances, helping clients and networks.

6. FUTURE SCOPE

Future undertakings include adding radar and LiDAR sensors for natural comprehension and object detection. To streamline constant handling for dynamic situations, utilize improved equipment speed increase, equal handling, and model pressure. Decentralizing handling, lessening dormancy, and further developing adaptability are objectives of consistent edge figuring joining, particularly in asset compelled or time-touchy conditions. Investigating and coordinating new calculations and structures guarantees adaptability to autonomous driving innovation challenges [42].

REFERENCES

[1] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling, and S. Thrun, ‘‘Towards fully autonomous driving: Systems and algorithms,’’ in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2011, pp. 163–168.

[2] J. Janai, F. Güney, A. Behl, and A. Geiger, ‘‘Computer vision for autonomous vehicles: Problems, datasets and state of the art,’’ Found. Trends Comput. Graph. Vis., vol. 12, no. 1–3, pp. 1–308, 2020.

[3] Y. Wang, ‘‘Overview on key technology of perceptual system on selfdriving vehicles,’’ Auto Electr. Parts., vol. 2016, no. 12, pp. 12–16, Dec. 2016.

[4] M. Kuderer, S. Gulati, and W. Burgard, ‘‘Learning driving styles for autonomous vehicles from demonstration,’’ in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), May 2015, pp. 2641–2646.

[5] R. Girshick, J. Donahue, T. Darrell, and J. Malik, ‘‘Rich feature hierarchies for accurate object detection and semantic segmentation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 580–587.

[6] K. He, X. Zhang, J. Sun, and S. Ren, ‘‘Spatial pyramid pooling in deep convolutional networks for visual recognition,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904–1916, Jan. 2015.

[7] R. Girshick, ‘‘Fast R-CNN,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 1440–1448.

[8] S. Ren, K. He, R. Girshick, and J. Sun, ‘‘Faster R-CNN: Towards realtime object detection with region proposal networks,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[9] K. He, G. Gkioxari, P. Dollár, and R. B. Girshick, ‘‘Mask R-CNN,’’ in IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 2980–2988.

[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ‘‘You only look once: Unified, real-time object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 779–788.

[11] J. Redmon and A. Farhadi, ‘‘YOLO9000: Better, Faster, stronger,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Honolulu, HI, USA, Jul. 2017, pp. 6517–6525.

[12] J. Redmon and A. Farhadi, ‘‘YOLOv3: An incremental improvement,’’ 2018, arXiv:1804.02767.

[13] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, ‘‘YOLOv4: Optimal speed and accuracy of object detection,’’ 2020, arXiv:2004.10934.

[14] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, ‘‘YOLOX: Exceeding Yolo series in 2021,’’ 2021, arXiv:2107.08430.

[15] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W. Nie, Y. Li, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei, and X. Wei, ‘‘YOLOv6: A single-stage object detection framework for industrial applications,’’ 2022, arXiv:2209.02976.

[16] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, ‘‘YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,’’ 2022, arXiv:2207.02696.

[17] W. Liu, ‘‘SSD: Single shot multibox detector,’’ in Proc. Eur. Conf. Comput. Vis. Amsterdam, The Netherlands, 2016, pp. 21–37.

[18] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, ‘‘DSSD : Deconvolutional single shot detector,’’ 2017, arXiv:1701.06659.

[19] Z. Li and F. Zhou, ‘‘FSSD: Feature fusion single shot multibox detector,’’ 2017, arXiv:1712.00960.

[20] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, ‘‘Swin transformer: Hierarchical vision transformer using shifted Windows,’’ in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2021, pp. 9992–10002.

[21] Z. Liu, H. Hu, Y. Lin, Z. Yao, Z. Xie, Y. Wei, J. Ning, Y. Cao, Z. Zhang, L. Dong, F. Wei, and B. Guo, ‘‘Swin transformer v2: Scaling up capacity and resolution,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2022, pp. 12009–12019.

[22] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, ‘‘A ConvNet for the 2020s,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2022, pp. 11976–11986.

[23] Y. Chen et al., ‘‘DW-YOLO: An efficient object detector for drones and self-driving vehicles,’’ Arabian J. Sci. Eng., vol. 48, pp. 1–10, May 2022.

[24] Y. Zhou, S. Wen, D. Wang, J. Meng, J. Mu, and R. Irampaye, ‘‘MobileYOLO: Real-time object detection algorithm in autonomous driving scenarios,’’ Sensors, vol. 22, no. 9, p. 3349, Apr. 2022.

[25] H. Wang and W. Zang, ‘‘Research on object detection method in driving scenario based on improved YOLOv4,’’ in Proc. IEEE 6th Inf. Technol. Mechatronics Eng. Conf. (ITOEC), Mar. 2022, pp. 1751–1754.

[26] D. Tian, ‘‘SA-YOLOv3: An efficient and accurate object detector using self-attention mechanism for autonomous driving,’’ IEEE Trans. Intell. Transp. Syst., vol. 23, no. 5, pp. 4099–4110, May 2022.

[27] A. Gupta, K. Illanko, and X. Fernando, ‘‘Object detection for connected and autonomous vehicles using CNN with attention mechanism,’’ in Proc. IEEE 95th Veh. Technol. Conf., (VTC-Spring), Jun. 2022, pp. 1–6.

[28] H. Wang, Y. Xu, Y. He, Y. Cai, L. Chen, Y. Li, M. A. Sotelo, and Z. Li, ‘‘YOLOv5-fog: A multiobjective visual detection algorithm for fog driving scenes based on improved YOLOv5,’’ IEEE Trans. Instrum. Meas., vol. 71, pp. 1–12, 2022.

[29] Y. Li, J. Wang, J. Huang, and Y. Li, ‘‘Research on deep learning automatic vehicle recognition algorithm based on RES-YOLO model,’’ Sensors, vol. 22, no. 10, p. 3783, May 2022.

[30] J. Hu, L. Shen, and G. Sun, ‘‘Squeeze-and-excitation networks,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 7132–7141.

[31] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, ‘‘CBAM: Convolutional block attention module,’’ in Proc. Eur. Conf. Comput. Vis. (ECCV), Sep. 2018, pp. 3–19.

[32] Y. Liu, Z. Shao, and N. Hoffmann, ‘‘Global attention mechanism: Retain information to enhance channel-spatial interactions,’’ 2021, arXiv:2112.05561.

[33] X. Pan, C. Ge, R. Lu, S. Song, G. Chen, Z. Huang, and G. Huang, ‘‘On the integration of self-attention and convolution,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2022, pp. 815–825.

[34] Q. Hou, D. Zhou, and J. Feng, ‘‘Coordinate attention for efficient mobile network design,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 13708–13717.

[35] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, ‘‘Feature pyramid networks for object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 936–944.

[36] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, ‘‘Path aggregation network for instance segmentation,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 8759–8768.

[37] G. Ghiasi, T.-Y. Lin, and Q. V. Le, ''NAS-FPN: Learning scalable feature pyramid architecture for object detection,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 7036–7045.

[38] S. Qiao, L.-C. Chen, and A. Yuille, ''DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 10213–10224.

[39] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, ''An image is worth 16×16 words: Transformers for image recognition at scale,'' 2020, arXiv:2010.11929.

[40] J. Park, S. Woo, J. Lee, and I. Kweon, ''BAM: Bottleneck attention module,'' presented at the 29th Brit. Mach. Vis. Conf., Newcastle, U.K., Sep. 2018.

[41] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, ''BDD100K: A diverse driving dataset for heterogeneous multitask learning,'' 2018, arXiv:1805.04687.

[42] Y. Zhou, S. Wen, D. Wang, J. Mu, and I. Richard, ''Object detection in autonomous driving scenarios based on an improved faster-RCNN,'' Appl. Sci., vol. 11, no. 24, Dec. 2021, Art. no. 11630.

[43] L. Li and X. Li, ''H-YoLov3: High performance object detection applied to assisted driving,'' in Proc. Asia Conf. Algorithms, Comput. Mach. Learn. (CACML), Mar. 2022, pp. 462–467.

[44] F. Yang, X. Zhang, S. Zhang, C. Li, and H. Hu, ''Design of real-time vehicle detection based on YOLOv4,'' in Proc. Int. Conf. Control, Autom. Inf. Sci. (ICCAIS), Oct. 2021, pp. 824–829.

[45] Y. Li and F. Yu, ''CDMY: A lightweight object detection model based on coordinate attention,'' in Proc. IEEE 10th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC), Jun. 2022, pp. 1258–1263.

[46] Q. Luo, J. Wang, M. Gao, Z. He, Y. Yang, and H. Zhou, ''Multiple mechanisms to strengthen the ability of YOLOv5s for real-time identification of vehicle type,'' Electronics, vol. 11, no. 16, p. 2586, Aug. 2022.

[47] G.Viswanath, "Hybrid encryption framework for securing big data storage in multi-cloud environment", Evolutionary intelligence, vol.14, 2021, pp.691-698.

[48] Viswanath Gudditi, "Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage", Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol.12, 2021, pp.545-552.

[49] Viswanath Gudditi, "A Smart Recommendation System for Medicine using Intelligent NLP Techniques", 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), 2022, pp.1081-1084.

[50] G.Viswanath, "Enhancing power unbiased cooperative media access control protocol in manets", International Journal of Engineering Inventions, 2014, vol.4, pp.8-12.

[51] Viswanath G, "A Hybrid Particle Swarm Optimization and C4.5 for Network Intrusion Detection and Prevention System", 2024, International Journal of Computing, DOI: https://doi.org/10.47839/ijc.23.1.3442, vol.23, 2024, pp.109-115.

ISSN 2454 - 5015

[52] G.Viswanath, "A Real Time online Food Ording application based DJANGO Restfull Framework", Juni Khyat, vol.13, 2023, pp.154-162.

[53] Gudditi Viswanath, "Distributed Utility-Based Energy Efficient Cooperative Medium Access Control in MANETS", 2014, International Journal of Engineering Inventions, vol.4, pp.08-12.

[54] G.Viswanath," A Real-Time Video Based Vehicle Classification, Detection And Counting System", 2023, Industrial Engineering Journal, vol.52, pp.474-480.

[55] G.Viswanath, "A Real- Time Case Scenario Based On Url Phishing Detection Through Login Urls ", 2023, Material Science Technology, vol.22, pp.103-108.

[56] Manmohan Singh,Susheel Kumar Tiwari, G. Swapna, Kirti Verma, Vikas Prasad, Vinod Patidar, Dharmendra Sharma and Hemant Mewada, "A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification" published in Journal of Computer Science, Available at: https://pdfs.semanticscholar.org/69ac/f07f2e756b79181e4f1e75f9e0f275a56b8e.pdf