Deployment of YOLOv5 for Traffic Sign Detection and Recognition.

CS405 Machine Learning

Langchu Huang	Yutong Wu	Yuhang Huang	Tianye Wu
12213009	12213012	12213015	12211201

Content

Background and Significance

1.1 Traffic Sign Detection and Recognition (TSDR)

enables real-time recognition and understanding of road signs,
making it an essential technology for modern transportation systems

M. Swathi and K. V. Suresh, "Automatic traffic sign detection and recognition: A review," 2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET), Chennai, India, 2017, pp. 1-6,

1.2 Significance

- a critical component of Advanced Driving Assistance Systems (ADAS) and autonomous driving, help to improve driving safety, reduce traffic accidents, and enhance the driving experience
- help Intelligent Transportation Systems (ITS) monitor road conditions in real time, adjust traffic signals, and provide traffic information
- has immense commercial value

Market Size and CAGR Estimates (2024-2029)			
Market	**Estimated Size (2024)**	**Estimated Size (2029)**	**CAGR (2024-2029)**
ADAS	USD 49.65 billion	USD 107.47 billion	16.70%
Autonomous driving	USD 41.10 billion	USD 114.54 billion	22.75%
ITS	USD 33.38 billion	USD 46.36 billion	6.79%

Analysis of Current Research Status

2.1 Datasets & Benchmark

- Road-sign-detection Dataset:

  Contains 877 images across 4 classes for road sign detection.

- Lisa traffic sign dataset：

  Consists of images and video frames for traffic sign detection, with annotations for 47 types of traffic signs.

- CCTSDB (Chinese City Traffic Sign Database) Dataset:

  Contains more than 10,000 images with over 60 categories of Chinese traffic signs.

- TT100k Dataset:

  A large-scale traffic sign dataset containing 100,000 images.

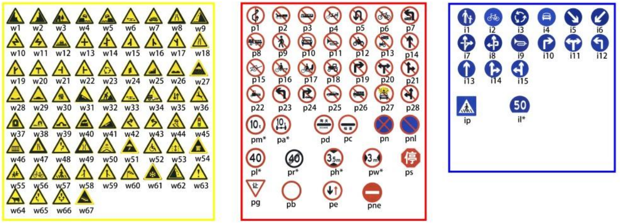
Comparison

different dataset with
different methods

Dataset	Methods	Prohibitive (AUC)	Danger (AUC)	Mandatory (AUC)	Time (s)
GTSDB	HOG+LDA [6]	70.33%	35.94%	12.01%	N/A
	Hough-like [6]	26.09%	30.41%	12.86%	N/A
	Viola-Jones [6]	90.81%	46.26%	44.87%	N/A
	HOG+LDA+SVM [89]	100%	99.91%	100%	3.533
	ChnFtrs [25]	100%	100%	96.98%	N/A
	HOG+SVM [67]	99.98%	98.72%	95.76%	3.032
	SVM+Shape [68]	100%	98.85%	92.00%	0.4-1
	SVM+CNN [69]	N/A	99.78%	97.62%	12-32
	SFC-tree [88]	100%	99.20%	98.57%	0.192 (3.19 GHz CPU)
	CNN [E-53]	99.89%	99.93%	99.16%	0.162 (Titan X GPU)
	ACF+SPC+LBP+AdaBoost [58]	100%	98.00%	97.57%	N/A
	AdaBoost+SVR [59]	100%	100%	99.87%	N/A
	AdaBoost+CNN+SVM [73]	99.45%	98.33%	96.50%	N/A
BTSD	ChnFtrs [25]	94.44%	97.40%	97.96%	1~3 (Intel Core i7 870 CPU, GTX 470 GPU
	AdaBoost+SVR [59]	93.45%	99.88%	97.78%	0.05~0.5 (Intel Core-i7 4770 CPU)
	AN+FRPN [72]	AP(%): 50.82%(Small), 88.05%(med), 96.82%(large)			0.128 (Tesla K20 GPU)
	Faster-RCNN in [72]	AP(%): 43.93%(Small), 97.8%(medium), 98.31%(large)			0.165 (Tesla K20 GPU)
TT100k	Fast R-CNN in [10]	Recall: 56%; Accuracy: 50% Curves can be found in [10]			N/A
	Multi-class Network [10]	Recall: 91%; Accuracy: 88% Curves can be found in [10]			N/A
	AN+FRPN [72]	AP(%): 49.81%(Small), 86.9%(med), 96.05%(large)			0.128 (Tesla K20 GPU)
	Faster-RCNN in [72]	AP(%): 31.22%(Small), 77.17%(med), 94.05%(large)			0.165 (Tesla K20 GPU)
LISA	ICF in [11]	87.32% (Diamond)	96.03% (Stop)	91.09% (NoTurn)	N/A
	ACF in [11]	98.98% (Diamond)	96.11% (Stop)	96.17% (NoTurn)	N/A

C. Liu, S. Li, F. Chang and Y. Wang, "Machine Vision Based Traffic Sign Detection Methods: Review, Analyses and Perspectives," in IEEE Access, vol. 7, pp. 86578-86596, 2019, doi: 10.1109/ACCESS.2019.2924947.

The characteristics of the TT100K dataset



- **High-resolution images**: Each image has a resolution of 2048x2048, providing rich details.
- **Diverse scenes**: The images are captured in various locations, lighting, and weather conditions, increasing the dataset's diversity.
- **Rich categories**: The dataset includes 221 categories of traffic signs, providing a wide range of samples for traffic sign recognition and classification.
- **Detailed annotations**: Each traffic sign comes with detailed annotation information, including category IDs and icons.

In the TT100K, there are some data issues in traffic sign datasets mainly caused by the following factors.

Illumination | Perspective | Blur | Occlusion

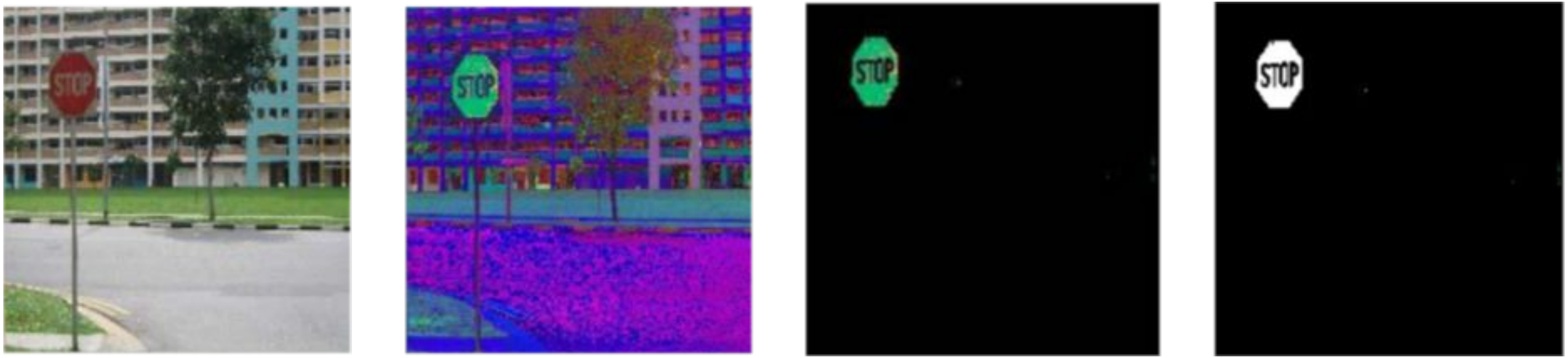
2.2 Models Architecture and Principle

- Traditional Method
  - color and shape analysis
  - Feature-Based Methods
  - Ensemble learning

- The novel deep learning-base method

color and shape analysis

Color segmentation process eliminates the unnecessary objects and hence it reduces the search area of the image or video frame.



RGB　　HSI　　Pixel of interest in HSI　　Binary

Swathi M , Suresh K V .Automatic traffic sign detection and recognition: A review[C]//International Conference on Algorithms.2017:1-6.DOI:10.1109/ICAMMAET.2017.8186650.

color and shape analysis

more color based detection methods

	Category	Paper	Year	Method	Detected colors
Color Based Detection Methods	RGB based thresholding	[2]	2010	Normalized RGB thresholding	Red, blue, yellow
		[30]	2010	Color Enhancement	Red, blue, yellow
		[31]	2015	Color Enhancement	Red, blue, yellow
	Hue and saturation thresholding	[2]	2010	Hue and saturation thresholding	Red, blue, yellow
		[33]	2004	LUTs based HS thresholding	Red, blue, yellow
	Thresholding on other spaces	[2]	2010	Ohta thresholding	Red, blue, yellow
		[34]	2015	Lab thresholding	Red, blue, yellow, green
	Chromatic/Achromatic Decomposition	[2]	2010	RGB, HIS, Ohta decomposition	white
		[34]	2015	RGB based achromatic segment	white
	Pixel classification	[2]	2010	SVM classification	Red, blue, yellow
		[36]	2012	Probabilistic neural networks	Red, blue, yellow

C. Liu, S. Li, F. Chang and Y. Wang, "Machine Vision Based Traffic Sign Detection Methods: Review, Analyses and Perspectives," in IEEE Access, vol. 7, pp. 86578-86596, 2019, doi: 10.1109/ACCESS.2019.2924947.

color and shape analysis

more shape based detection methods

	Category	Paper	Year	Method	Detected shapes
Shape Based Detection Methods	Shape detection	[38]	2015	Hough	Circle and triangle
		[39]	2008	Radial symmetry transform	Circle
		[86]	2004	Radial symmetry transform	Polygons
	Shape analysis and matching	[41]	2003	Complex shape models	Circle, polygons
		[42]	2008	Shape decomposition	Circle, square, triangle
	Fourier transformation	[26]	2011	Fourier descriptors	Circle, square, triangle
		[43]	2008	Fast Fourier Transformation	Circle, square, triangle
	Key points detection	[45]	2014	SIFT	Circle, square, triangle, octagon
		[15]	2014	Harris corner	Circle, triangle
		[46]	2014	Interest points clustering	Different shapes

C. Liu, S. Li, F. Chang and Y. Wang, "Machine Vision Based Traffic Sign Detection Methods: Review, Analyses and Perspectives," in IEEE Access, vol. 7, pp. 86578-86596, 2019, doi: 10.1109/ACCESS.2019.2924947.

The novel deep learning-base method

- The Convolutional Neural network (CNN) based detection methods learn features through convolutional network.
- You only look once (YOLO)

Zhang et al. utilized YOLOv2 to design their real-time traffic sign detection method.Liu et al.[5] use a YOLO CNN to classify traffic signs and MSRCR image augmentation during pre-processing.

Sharma and Kumar’s study provides YOLOv8 for traffic signal recognition in the advanced version that takes place in a real time environment for road safety improvement.

R-CNN

Region-based Convolutional Neural Network

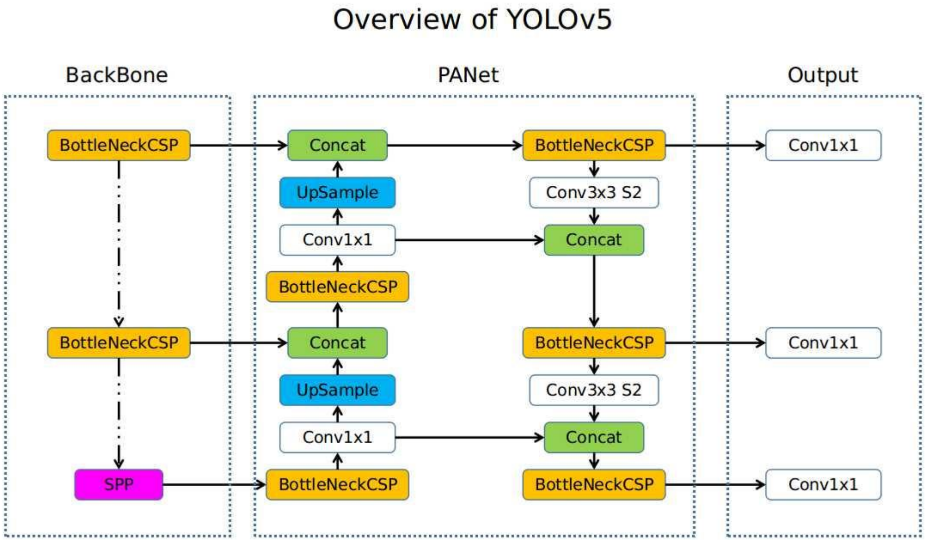
generating a set of region proposals by selective search, which groups similar pixels into regions based on color, texture, and other visual cues.

Merged the region by calculate the similarity of adjacent region such as….

Generate candidate region and repeat recursion

YOLOv5 architecture

- Backbone (Feature Extraction):
  - CSPDarknet
  - Extracts deep semantic features for object recognition.
- Neck (Feature Fusion):
  - PANet
  - Combines features from different scales to enhance the ability to detect multi-scale objects.
- Head (Detection and Output):
  - One-Stage
  - Transforms feature maps into specific detection results.



Small

YOLOv5s

14 MB$_{FP16}$
6.4 ms$_{V100}$
37.2 mAP$_{COCO}$

You Only Look **Once**

One-Stage Detection

- **Process**: Directly predicts object categories and bounding boxes in a single step.
- **Speed**: Fast, suitable for real-time applications.
- **Accuracy**: Slightly lower, struggles with small objects.
- **Examples**: YOLO, SSD, RetinaNet.

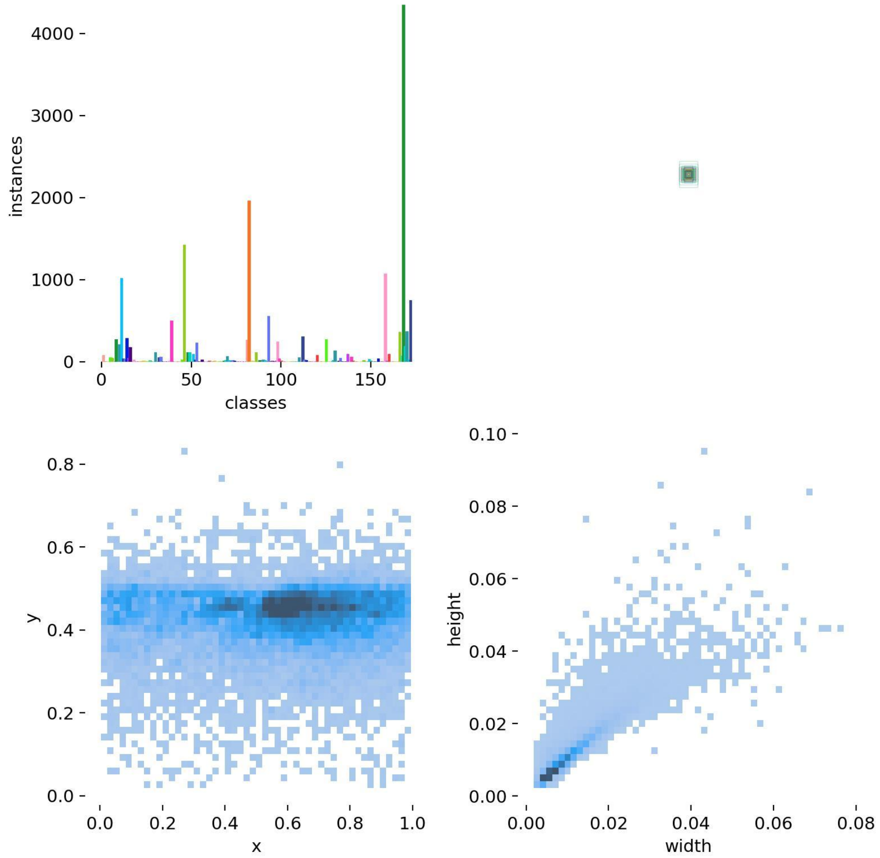
Two-Stage Detection

- **Process**: First generates region proposals, then refines classification and bounding boxes.
- **Speed**: Slower, computationally intensive.
- **Accuracy**: Higher, better for small objects and complex tasks.
- **Examples**: Faster R-CNN, Mask R-CNN, Cascade R-CNN.

You Only Look **Once**

One-Stage Detection

- **Process**: Directly predicts object categories and bounding boxes in a single step.
- **Speed**: Fast, suitable for real-time applications.
- **Accuracy**: Slightly lower, struggles with small objects.
- **Examples**: YOLO, SSD, RetinaNet.



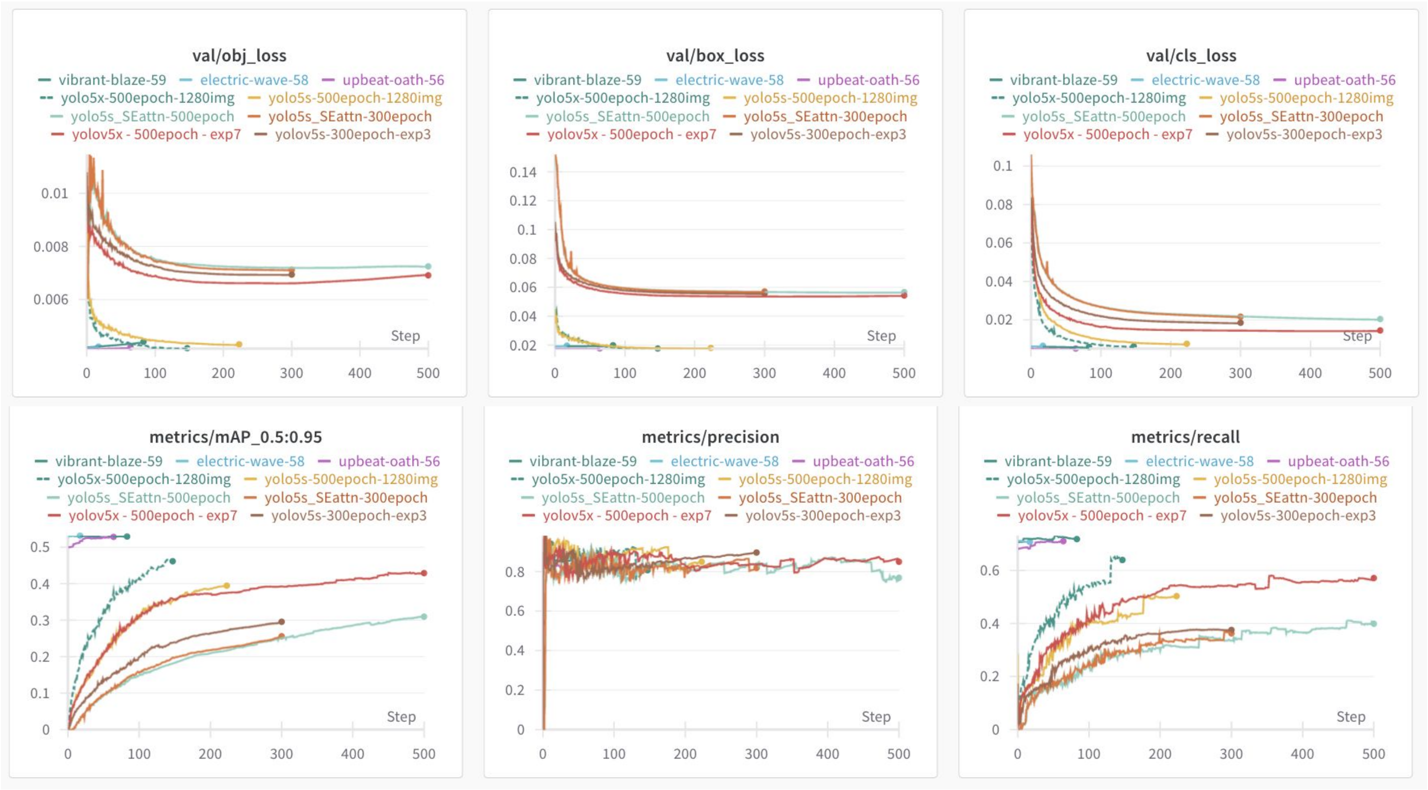
Contribution of This Study

Contribution

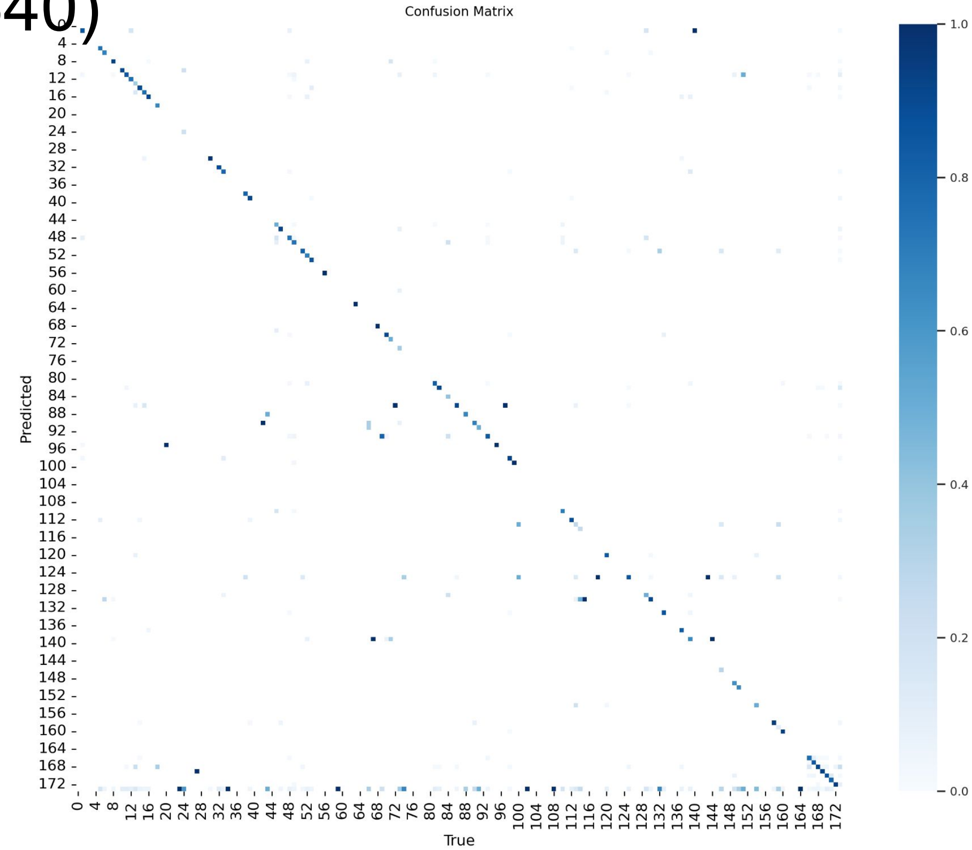
1. Applying Transfer Learning for variations of YOLOv5 models in TT100K datasets
2. Integrating Attention Module into YOLO to enhance the ability
3. Analysing and Comparing performances of different parameters sets
4. Deploying the Model in host and AI edge device in both online and offline way

- yolo5x-500epoch-1280img
- yolo5s-500epoch-1280img
- yolo5s_SEattn-500epoch
- yolo5s_SEattn-300epoch
- yolov5x - 500epoch - exp7
- yolov5x - 5epoch - exp6
- yolov5s - 300epoch - exp3
- yolov5s - 10epoch

Training Process

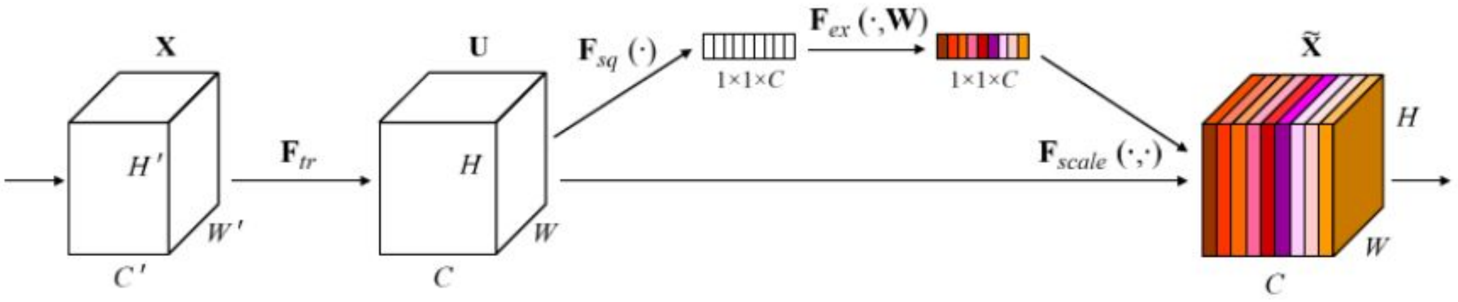


Training Result (yolov5x 640)



Achieved goals

Add the SE module



https://blog.csdn.net/qq_45981086/article/details/130869929

Performance Comparison

Metrics of YOLO Variations and Baseline Models (TT100K)

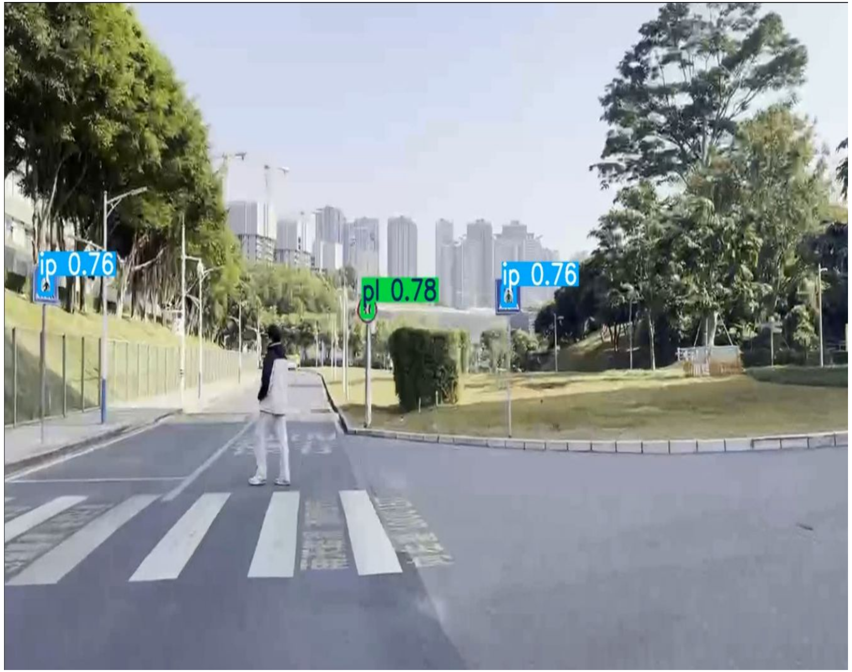
Variations	$mAP_{0.5:0.95}$	$mAP_{0.5}$	Precision	Recall
$yolov5s_{640}$	0.310	0.489	0.870	0.410
$yolov5x_{640}$	0.431	0.656	0.866	0.580
$yolov5s\text{-}seattn_{640}$	0.358	0.526	0.897	0.421
$yolov5s_{1280}$	0.462	0.671	0.723	0.510
$yolov5x_{1280}$	0.532	0.781	0.855	0.720
Fast R-CNN	N/A	N/A	0.50	0.56
Multi-class Network	N/A	N/A	0.88	0.91
AN+FRPN	0.4981	N/A	N/A	N/A
Faster-RCNN	0.3122	N/A	N/A	N/A

C. Liu, S. Li, F. Chang and Y. Wang, "Machine Vision Based Traffic Sign Detection Methods: Review, Analyses and Perspectives," in IEEE Access, vol. 7, pp. 86578-86596, 2019, doi: 10.1109/ACCESS.2019.2924947.

Research Effect Demonstration

Deploying - Offline

YOLOv5s - SEattn



YOLOv5s

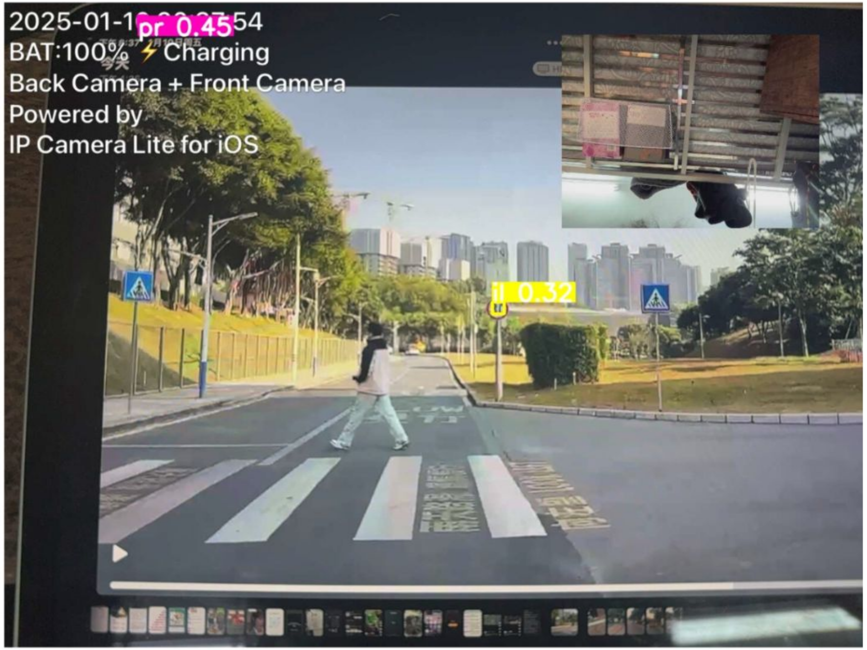
Deploying - Offline

YOLOv5s - SEattn

YOLOv5s

Deploying - Real Time

摄像头实时视频流



COMPARISION BETWEEN YOLOV5S AND YOLOV5S-SE

Model	Parameters	GFLOPs	Inference time
yolov5s	7476706	17.2	1.2ms
yolov5s-se	7613578	17.3	1.2ms

Future Work

- Advanced Attention Mechanisms:

Fuse self- attention with conventional attention mechanisms.

- Optimization for Resource-Constrained Devices:

Optimize the model for ultra-low-power devices (e.g., microcontrollers, IoT edge nodes) without sacrificing accuracy.

- Extension to Multi-Language Traffic Sign Recognition:

Adapt the model for multilingual traffic signs to support international traffic systems.

- Real-Time Multi-Task Learning:

Combine traffic sign detection with other tasks (e.g., lane detection, pedestrian recognition) in a unified framework.

Reference

[1] World Health Organization, "Global status report on road safety 2015," 2015.

[2] M. R. Research and Advisory, "ADAS market size & share analysis - growth trends & forecasts (2024-2029)," June 2024, retrieved December 3, 2024. [Online]. Available:

[3] "Autonomous vehicle market size & share analysis - growth trends & forecasts (2024-2029)," July 2023, retrieved December 3, 2024.

[4] "Smart transportation market size & share analysis - growth trends & forecasts (2024-2029)," June 2024, retrieved December 3, 2024.

[5] "Road signs dataset." [Online]. Available: https://makeml.app/datasets/road-signs

[6] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," in *International Joint Conference on Neural Networks*, no. 1288, 2013.

[7] R. Timofte, M. Mathias, R. Benenson, and L. Van Gool, "Traffic sign recognition - How far are we from the solution?" in *International Joint Conference on Neural Networks (IJCNN)*, August 2013.

[8] "LISA traffic sign dataset," 2012, the world’s largest dataset containing pictures of US traffic signs. [Online]. Available: https://vbn.aau.dk/en/datasets/lisa-traffic-sign-dataset

[9] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-sign detection and classification in the wild," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[10] J. Zhang, Y. Lv, J. Tao, F. Huang, and J. Zhang, "A robust real-time anchor-free traffic sign detector with one-level feature," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 2, pp. 1437-1451, 2024.

[11] M. Swathi and K. V. Suresh, "Automatic traffic sign detection and recognition: A review," in *International Conference on Algorithms*, 2017, pp. 1-6.

[12] C. Liu, S. Li, P. Chang, and Y. Wang, "Machine vision-based traffic sign detection methods: Review, analyses, and perspectives," *IEEE Access*, vol. 7, pp. 86578-86596, 2019.

[13] S. H. S. G. and C. M. Patil, "An approach towards efficient detection and recognition of traffic signs in videos using neural networks," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2016, pp. 456-460.

[14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. CVPR*, San Diego, CA, USA, Jun. 2005, pp. 886-893.

[15] S. Salti, A. Petrelli, P. Tombari, N. Fioraio, and L. D. Stefano, "Traffic sign detection via interest region extraction," *Pattern Recognition*, vol. 48, no. 4, pp. 1039-1049, Apr. 2015.

[16] F. Zaklouta and B. Stanciulescu, "Real-time traffic sign recognition in three stages," *Robotics and Autonomous Systems*, vol. 62, no. 1, pp. 16-24, 2014.

[17] Y.-L. Hou, X. Hao, and H. Chen, "A cognitively motivated method for classification of occluded traffic signs," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 2, pp. 205-262, Feb. 2017.

[18] Y. Wu, Y. Liu, J. Li, H. Liu, and X. Hu, "Traffic sign detection based on convolutional neural networks," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Dallas, TX, USA, Aug. 2013, pp. 1-7.

[19] R. Qian, B. Zhang, Y. Yue, Z. Wang, and P. Coenen, "Robust Chinese traffic sign detection and recognition with deep convolutional neural networks," in *Proceedings of the International Conference on Natural Computation*, Zhangjiajie, China, Aug. 2015, pp. 791-796.

[20] J. Zhang, M. Huang, X. Li, and X. Jin, "A real-time Chinese traffic sign detection algorithm based on modified yolov2," *Algorithms*, vol. 10, no. 4, p. 127, Nov. 2017.

[21] A. Liu, Y. Liu, and S. Kifah, "Deep convolutional neural network for enhancing traffic sign recognition developed on yolo v5," in *2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASIC)*, Bhubaneswar, India, 2024, pp. 1-6.

[22] L.-J. Wang, T. Suwattanapunicul, J. Thalauy, and P. Jansengrat, "Research on traffic sign detection and recognition system using deep ensemble learning," in *2024 6th International Conference on Computer Communication and the Internet (ICCCI)*, Tokyo, Japan, 2024, pp. 61-66.

[23] J. Sharma, D. Kumar, and A. Malhotra, "Advanced traffic signal recognition using real-time detection using yolo v8," in *2024 2nd World Conference on Communication & Computing (WCONP)*, Raipur, India, 2024, pp. 1-4.

[24] J. Arcos-Garcia, J. A. Alvarez-Garcia, and L. M. Soria-Morillo, "Deep neural network for traffic sign recognition systems: An analysis of spatial transformers and stochastic optimization methods," *Neural Networks*, vol. 99, pp. 158-165, 2018.

[25] B. Sanyal, R. K. Mohapatra, and R. Dash, "Traffic sign recognition: A survey," in *2020 International Conference on Artificial Intelligence and Signal Processing (AISP)*, 2020, pp. 1-6.