TSDR: Traffic-Sign Detection and Recognition

——Research Proposal in CS405 Machine Learning

proposers: Langchu Huang,Yutong Wu
Yuhang Huang, Tianye Wu

Contents

1. Background and significance
2. Analysis of Current Research Status
3. Completed Work
4. Research Plan and Expected Results
5. Potential Challenges and Solutions

Background and significance

1.1 TSDR: Traffic Sign Detection and Recognition

- used to identify and interpret various road signs and signals by computer vision.
- a critical component of Intelligent Transportation Systems (ITS), Advanced Driving Assistance Systems (ADAS) and autonomous driving.

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

- reduce the likelihood of accidents by **providing drivers with essential information** such as road conditions, speed limits, potential hazards and so on

  over 1.2 million people die across the globe annually due to road accidents.

- as **a key component** in ADAS and autonomous driving
- help ITS **monitor road conditions** in real time, adjust traffic signals, and provide traffic information

1.3 Motivations

- its related technologies advance

  like computer vision, deep learning & neural networks

- 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.

- CCTSDB (Chinese City Traffic Sign Database) Dataset:

  A large-scale dataset with 5183 images of 43 distinct traffic sign types from real-world traffic scenarios in Germany.

- 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.

2.2 Models Architecture and Principle

- color and shape analysis
- Feature-Based Methods
- Ensemble learning
- The novel deep learning-base method

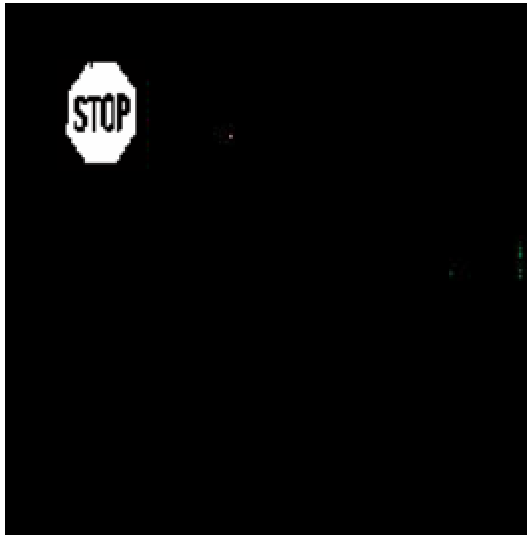
color and shape analysis

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

RGB representation

HSI representation

Pixels of interest in HSI



Binary image

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

As for the shape analysis, it can eliminate the issue in color based detection is the ambient illumination variation

Color Space	Detection Rate
RGB	88.75%
HSV	95%
HSI	91.35%

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:

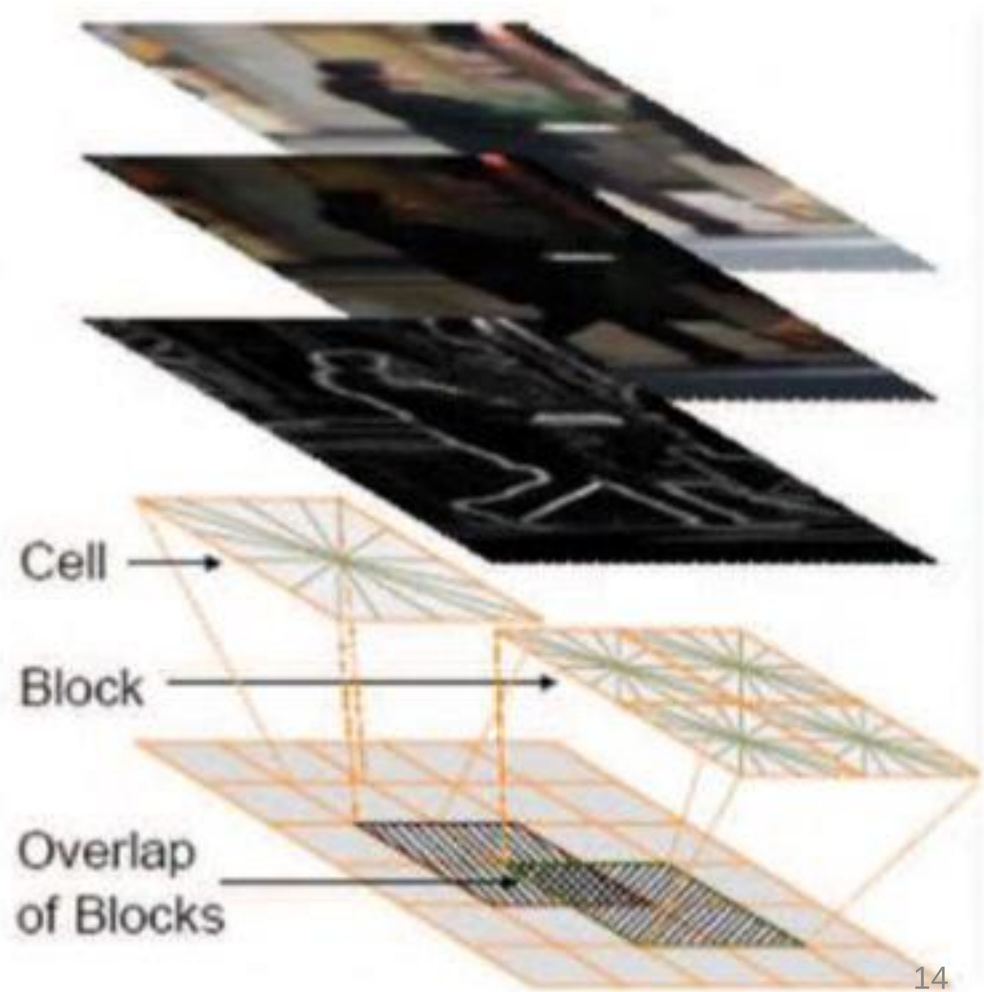
color and shape analysis

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

Feature-Based Methods

This structure utilizes HOG-like features to express the objects and treats the object detection problem as an SVM classification problem, in which each candidate is classified into objects or backgrounds.



HOG descriptor

N. Dalal and B. Triggs, ‘‘Histograms of oriented gradients for human detection,’’ in*Proc. CVPR*, San Diego, CA, USA, Jun. 2005, pp. 886–893.

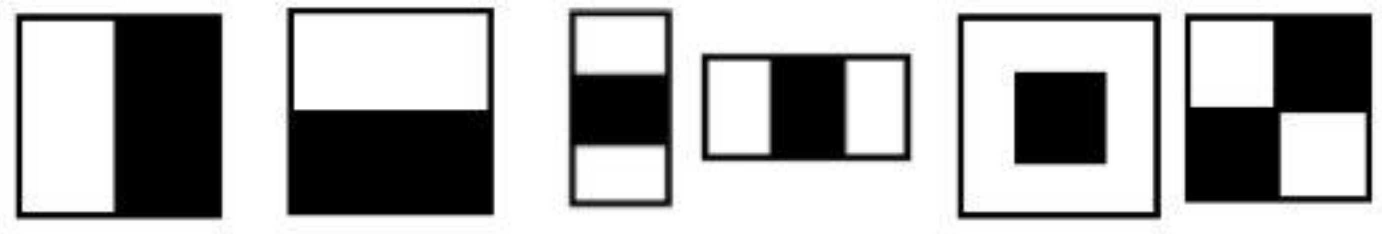
Ensemble learning

Viola and Jones' AdaBoost and cascade based detection structure (VJ) has been proved very efficient in some object detection problems. The selection of features is crucial for AdaBoost based TSDR detectors.

- Haar-like feature
- Dissociated dipoles feature
- Binary Pattern (MB-LBP) feature

Ensemble learning

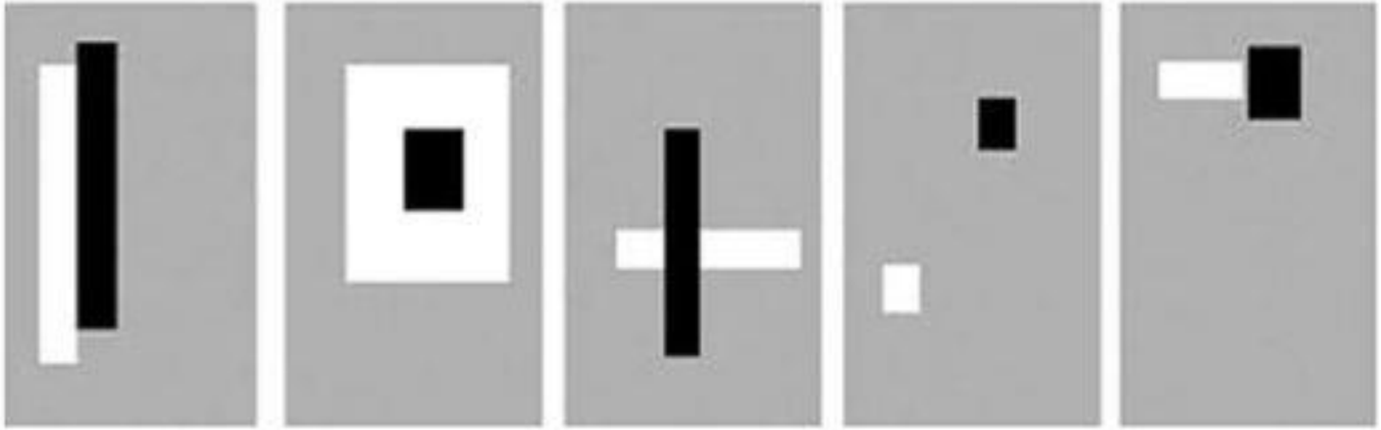
The Haar-like feature can express the gray level difference of traffic signs.



X. Baró, S. Escalera, J. Vitri, O. Pujol, and P. Radeva, ‘‘Traffic sign recognition using evolutionary Adaboost detection and forest-ECOC clas- sification,’’ IEEE Trans. Intell. Transp. Syst., vol. 10, no. 1, pp. 113–126 Mar. 2009.

Ensemble learning

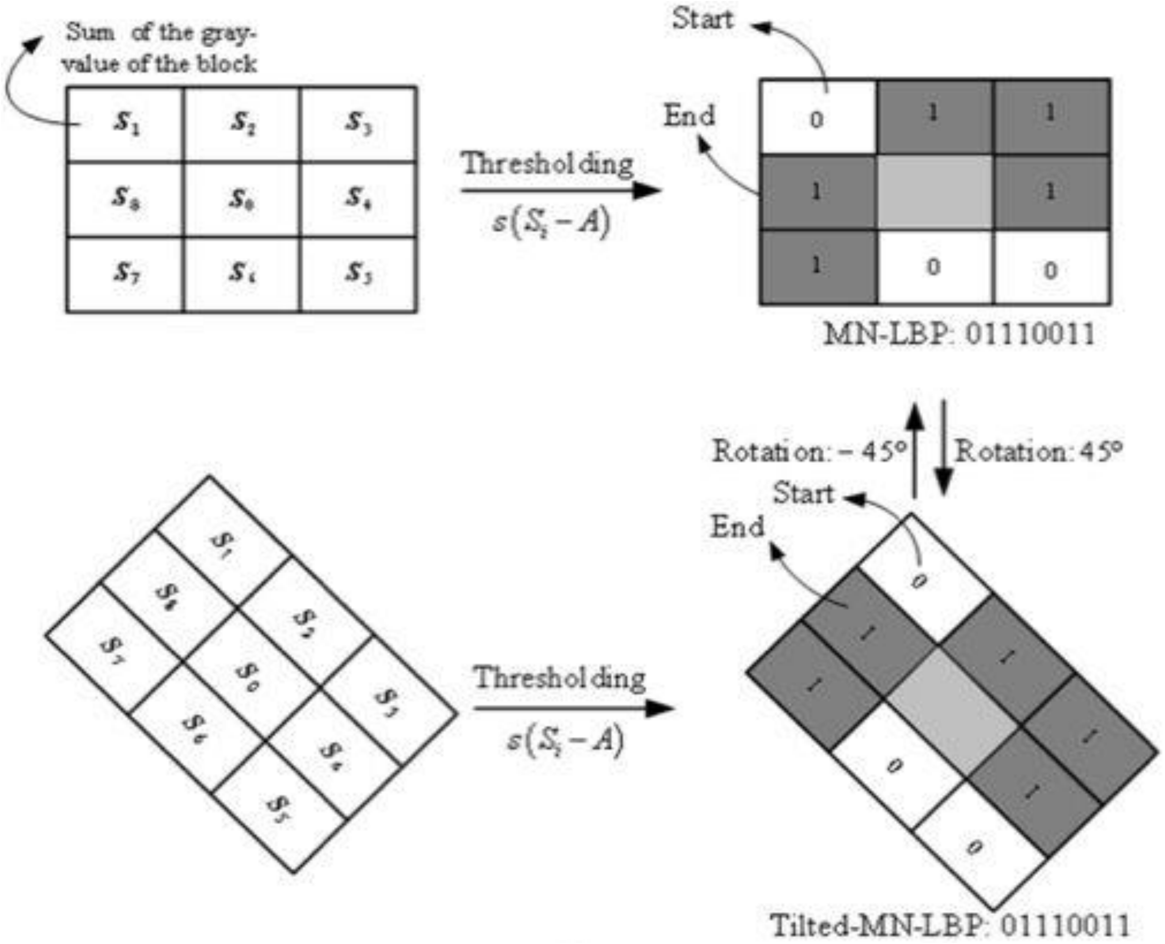
Dissociated dipoles feature is a more general rectangular feature. Using uncon- nected two dipoles, the dissociated dipoles feature can pro- duce more features to express traffic signs.



X. Baró, S. Escalera, J. Vitri, O. Pujol, and P. Radeva, ʻ ʻTraffic sign recognition using evolutionary Adaboost detection and forest-ECOC clas- sification,ʼ ʼ IEEE Trans. Intell. Transp. Syst., vol. 10, no. 1, pp. 113–126,Mar. 2009.

Ensemble learning

Multi-block normalization LBP (MN-LBP) features can express different types of features.



L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, ′′Face detection basedon multi-block LBP representation,′′ inProc. Int. Conf. Biometrics, Seoul,South Korea, 2007, pp. 11–18.

Database Size	Feature Extractor	Classifier Used	Recognition Rate (in %)	Remarks
	RGB colour segmentation			
	Proposed algorithm for circular signs	SVM	97.04	Novel idea
3 Signs	1NN each for matching shape and colour		>95	No proper database Not applicable in real time
	Circle by Hough Triangle by Ramer-Douglus-Rucar	SURF & FLAN	85	
	Traffic sign detection	LDA		Also uses motion information.
48 signs Polish,Japanese	Equiangular polygons	Gaussian	85.3	Novel idea

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:

The novel deep learning-base method

The Convolutional Neural network (CNN) based detection methods learn features through convolutional network.

In recent years, with the development of deep learning, many different deep neural network structures have appeared and made breakthrough in different detection areas.

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.

Comparison

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.

Reproduction of Current Paper

Zhang, J., Huang, M., Jin, X., & Li, X. (2017). A real-time Chinese traffic sign detection algorithm based on modified YOLOv2. Algorithms, 10(4), 127

Sanyal, B., Mohapatra, R. K., & Dash, R. (2020, January). Traffic sign recognition: A survey. In 2020 International Conference on Artificial Intelligence and Signal Processing (AISP) (pp. 1-6). IEEE.

Completed Work

3.1 Model Selection

- **Deep learning revolution in TSDR:**

  Traditional machine learning: SVM, KNN, HOG ---> CNN.

- **Advantages of YOLO:**

  Balance of speed and accuracy

  Transfer learning

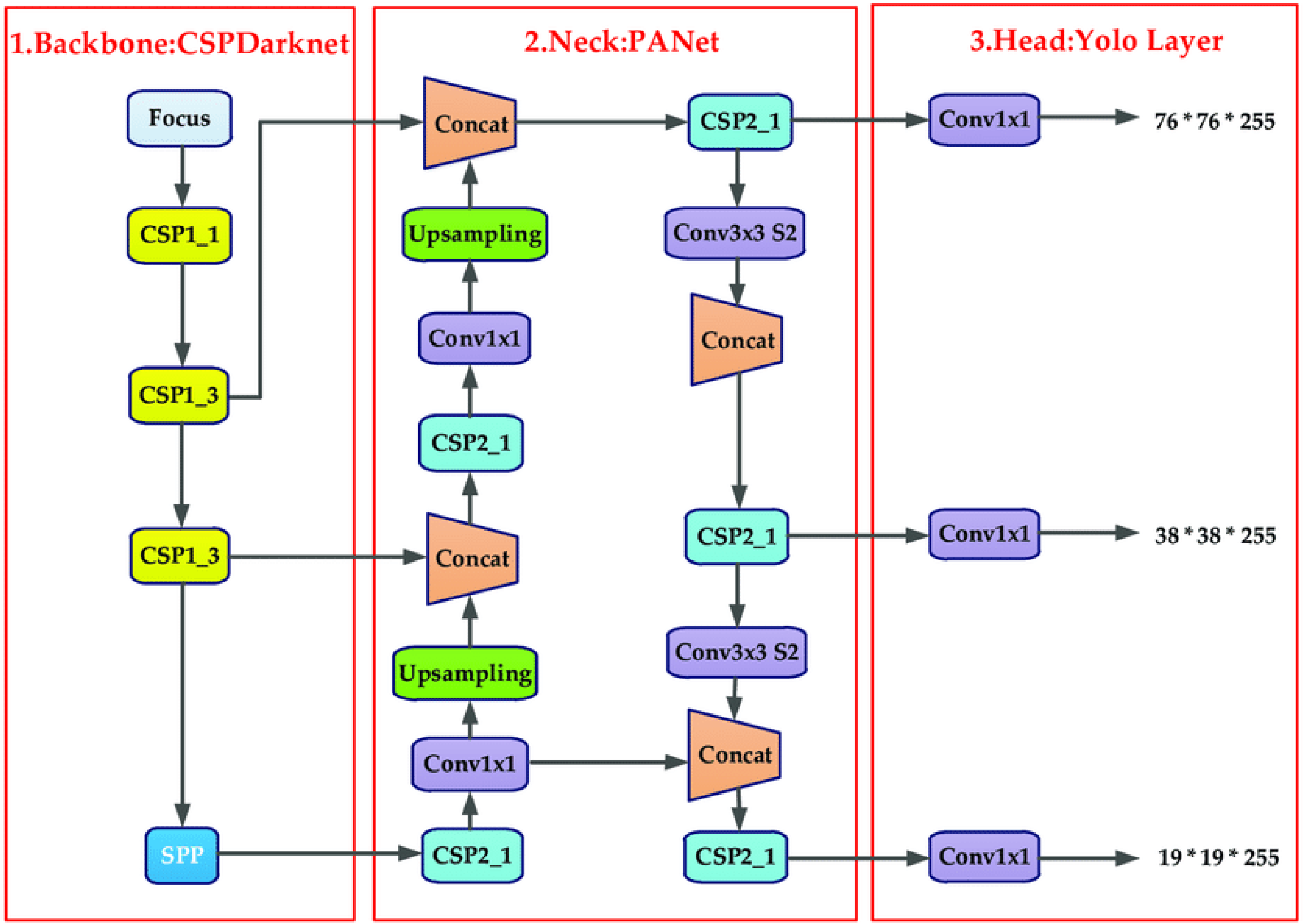
  Flexibility for deployment

3.2 Model Principle

Architecture:

Traditional machine learning:

SVM, KNN, HOG ---> CNN.



https://www.researchgate.net/figure/The-network-architecture-of-YOLOv5-1-Backbone-CSPDarknet-for-feature-extraction-2_fig1_358553872

3.2 Model Principle

- **Backbone**: The backbone network extracts features from the input image. YOLOv5 uses CSPDarknet53 as its backbone, a variant of Darknet53 that introduces Cross-Stage Partial Networks (CSP) to enhance feature propagation and reduce the computational load.
- **Neck:** The neck in YOLOv5 uses the PANet (Path Aggregation Network) to combine feature maps from different layers, allowing the model to leverage both low-level and high-level features. This helps the model make more accurate predictions, especially for small or distant objects.
- **Head**: The head of the network is responsible for making the final predictions (bounding box, class probability, and object confidence) from the features extracted by the backbone and aggregated by the neck.
- s m l x x6

3.3 Experiment result

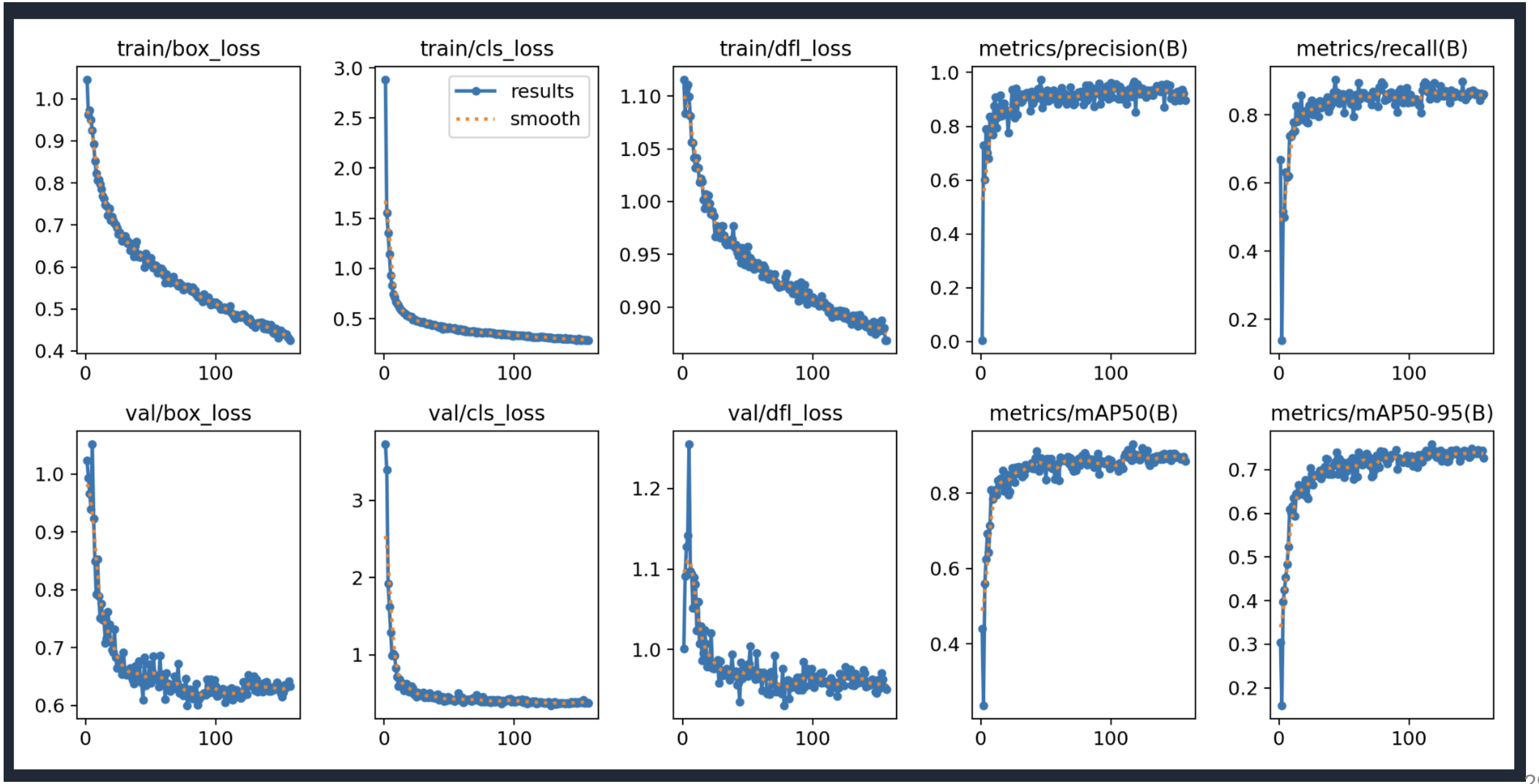
- **Dataset:**

Road Sign Detection dataset, 877 images of 4 distinct classes (https://www.kaggle.com/datasets/andrewmvd/road-sign-detection/data)

- **Result:**

mAP	93.1%
Precision	95.5%
Recall	87.6%

3.3 Experiment result



3.3 Experiment result

Research Plan and Expected Results

4.1 Time schedule

2024/10/24 week 7	**Project selection: TSDR problem （Traffic Sign Detection and Recognition）**
2024/11/11 week 8	Midterm Preparation
2024/11/7 week 9	Reproduce provided reference paper indivizually. Read field background information and the latest paper overview. Collect dataset, test different methods and models.
2024/11/14 week 10	Make comparision from previous work. Choose yolov5 as the base model of the project.

2024/11/21 week 11	**Train yolov5's branch models on collected dataset.** **Analyse the training results.**
2024/11/28 week 12	**Do Research Proposal and PPT.** Summarize previous work and make plan for the following weeks.
2024/12/5 week 13	Clean the tt100k dataset and transform it to YOLO format and train it on the YOLOv5 models. Compare experiment results with previous paper and work, try to do optimization and add some plug-and-play modules to the chosen model. Analyse the difference and disscuss further work.
2024/12/12 week 14	Training final optimized model and implement it on the device. Test it in real life and analyse the result. Disscuss the further improvment in terms of real life implementation.
2024/12/19 week 15	Write final report and prepare for the presentation

4.2 Expected results

Use YOLO to do regression and use Sill-net to do light optimization and do reclassification to achieve better accuracy compared with present results.
Add attention mudule(time, space, frequency...) to YOLO, in order to increase the accuracy of classification.
The expected performance is to be able to process real-world traffic sign images or video streams by 30 fps on a mobile phone device.

Potential Challenges and Solutions

5.1 Data Challenge

- Illumination Issues
- Perspective Issues
- Motion Blur
- Partial Occlusion


Illumination

Perspective

Blur


Occlusion

5.2 Architecture Challenges

- Dataset Limitations
- Model Complexity
- Overfitting Risk
- Real-Time Performance
- Integration with Real-World Systems
- Dynamic Environment Adaptability

References

[1] World Health Organization, “Global status report on road safety 2015,” 2015. [Online]. Available: https://www.who.int/publications-detail/global-statusreport-on-road-safety-2015

[2] Mordor Intelligence Research & Advisory. (2024, June).ADAS Market Size & Share Analysis - Growth Trends & Forecasts (2024 - 2029). Mordor Intelligence. Retrieved December 3, 2024, from

https://www.mordorintelligence.com/industry-reports/advanced-driver-assistance-systems-market

[3] Mordor Intelligence Research & Advisory. (2023, July). Autonomous Vehicle Market Size & Share Analysis - Growth Trends & Forecasts (2024 - 2029). Mordor Intelligence. Retrieved December 3, 2024, from

https://www.mordorintelligence.com/industry-reports/autonomous-driverless-cars-market-potential-estimation

[4] Mordor Intelligence Research & Advisory.(2024, June).Smart Transportation Market Size & Share Analysis - Growth Trends & Forecasts (2024 - 2029).Mordor Intelligence.Retrieved December 3, 2024, from

https://www.mordorintelligence.com/industry-reports/smart-transportation-market

[5] 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

References

[6] 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.

[7] H.s.g, Supreeth., & Patil, C. M. (2016). An approach towards efficient detection and recognition of traffic signs in videos using neural networks.*2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 456-459.

[8] 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.

[9] N. Dalal and B. Triggs, ‘‘Histograms of oriented gradients for human detection,’’ inProc. CVPR, San Diego, CA, USA, Jun. 2005, pp. 886–893.

[10] S. Salti, A. Petrelli, F. Tombari, N. Fioraio, and L. D. Stefano, ‘‘Traffic sign detection via interest region extraction,’’Pattern Recognit., vol. 48, no. 4,pp. 1039–1049, Apr. 2015.

References

[11] F. Zaklouta and B. Stanciulescu, ‘‘Real-time traffic sign recognition in three stages,’’Robot. Auto. Syst., vol. 62, no. 1, pp. 16–24, 2014.

[12] P.ViolaandM.J.Jones,‘‘Robustreal-timefacedetection,’’Int.J.Comput. Vis., vol. 57, no. 2, pp. 137–154, 2004.

[13] X. Baró, S. Escalera, J. Vitri, O. Pujol, and P. Radeva, ‘‘Traffic sign recognition using evolutionary Adaboost detection and forest-ECOC clas- sification,’’IEEE Trans. Intell. Transp. Syst., vol. 10, no. 1, pp. 113–126,Mar. 2009.

[14] C. Liu, F. Chang, and Z. Chen, ‘‘Rapid multiclass traffic sign detection in high-resolution images,’’IEEE Trans. Intell. Transp. Syst., vol. 15, no. 6, pp. 2394–2403, Dec. 2014.

[15] J. Sharma, D. Kumar and A. Malhotra, "Advanced Traffic Signal Recognition Using Real-Time Detection Using YOLO v8," 2024 2nd World Conference on Communication & Computing (WCONF), RAIPUR, India, 2024, pp. 1-4, doi: 10.1109/WCONF61366.2024.10692265.

References

[16] Joseph Redmon, et al. “You only look once: Unified, real-time object detection.” CVPR 2016.

[17] 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 (ASSIC), Bhubaneswar, India, 2024, pp. 1–6.

[18] J. Sharma, D. Kumar, and A. Malhotra, “Advanced trafficsignal recognition using real-time detection using yolo v8,” in 2024 2nd World Conference on Communication & Computing(WCONF), Raipur, India, 2024, pp. 1–4.