Deep Residual Neural Network for Efficient Traffic Sign Detection

Hanlin Cai, Zheng Li, Jiaqi Hu, Wei Hong Lim,
Sew Sun Tiang, Mastaneh Mokayef, Chin Hong Wong

Fuzhou University, China

Feb 11st, 2023

About the Presenter

Hanlin Cai 蔡汉霖

- A junior student majoring in Robotics and Intelligent Devices at Fuzhou University.
- I am extremely fortunate to be advised by Chin Hong Wong and Zhezhuang Xu.
- Currently, I am interested in Machine Learning and its applications on the Internet of Things (IoT).
- hanlin.cai@ieee.org
- `https://caihanlin.com`

1. Research Overview
2. Proposed Methodology
3. Experimental Results
4. Conclusion and Plan

1. Research Overview
   - Background
   - Literature Review
   - Gaps in existing works
2. Proposed Methodology
3. Experimental Results
4. Conclusion and Plan

1. Research Overview
   - Background
   - Literature Review
   - Gaps in existing works
2. Proposed Methodology
3. Experimental Results
4. Conclusion and Plan

Background

Neural Network for Traffic Sign Detection.

- **Traffic Sign Detection System (TSDS)**
  1. Traffic Sign Detection (TSD)
  2. Traffic Sign Recognition (TSR)
  3. Traffic Sign Classification (TSC)
- **Famous Neural Network for Image Processing**
  1. VGGNet[1]
  2. GoogLeNet[2]
  3. ResNet[3] (SOTA)
- **Significance of Related Works**
  1. Increase the efficiency and safety of urban transportation
  2. Bring the development of self-driving cars that are able to detect traffic.

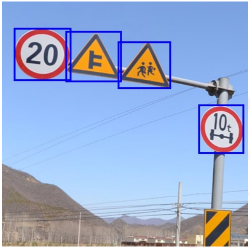


Figure 1: Traffic Sign Example

$$\text{Problem} \Rightarrow \begin{cases} \text{TSD} \xrightarrow{\text{YOLO}} \text{Search for the Signs} \\ \text{TSR} \xrightarrow{\text{RNN}} \text{Understand the Signs} \\ \text{TSC} \xrightarrow{\text{RNN}} \text{Classify the Signs} \end{cases} \xrightarrow{\text{How to get}} \text{Optimization}$$

1. Research Overview
   - Background
   - Literature Review
   - Gaps in existing works
2. Proposed Methodology
3. Experimental Results
4. Conclusion and Plan

Literature Review

$$\text{Review} \Rightarrow \begin{cases} \text{Traditional Method} \xrightarrow{\text{2003-2009}} \text{Color, Features} \\ \text{Machine Learning} \xrightarrow{\text{2010-2015}} \text{Classifier, NN} \\ \text{Deep Learning} \xrightarrow{\text{2016-2022}} \text{CNN, RNN Model} \end{cases}$$

Table 1 A summary of related literature works

Techniques	Descriptions
Colour Segmentation[12,13]	Easily affected by daylight conditions.
Texture Features[16,17]	Highly depending on the quality of the images.
SVM Classifier[18]	Good classification accuracy, but low speed.
NN Models[19]	High accuracy, but a large resource is required.
LeNet-5 CNN[4]	Utilizing Gabor Based Kernel, high accuracy
CNN+RPN[5]	Very high real-time detection speed.
R-CCN+RPN[20]	Very high accuracy, close to 99.9%.
Mask R-CNN[21]	Based on highly challenging datasets.

1. Research Overview
   - Background
   - Literature Review
   - Gaps in existing works
2. Proposed Methodology
3. Experimental Results
4. Conclusion and Plan

Research Gaps and Contributions

Gaps in existing knowledge

- Large-scale categories datasets for training and testing
- Systematic evaluation method for configuration optimization

Contributions of this works

- The model training and testing are based on large-scale datasets (Tsinghua-Tencent 100k).
- A new systematic analytic hierarchy process (AHP Method) for model evaluation has been proposed.

Overview of Our Works

Deep Residual Neural Network for Efficient Traffic Sign Detection*

- **Traffic Sign Detection System**
  1. Traffic Sign Detection (TSD)
  2. Traffic Sign Recognition (TSR)
- **Analytic Hierarchy Process Method**
  1. Accuracy (Performance Metrics)
  2. Stability
  3. Response
  4. System Capability

$$\text{Our works} \Rightarrow \begin{cases} \text{TSD} \xrightarrow{\text{YOLO}} \text{Search Sign} \\ \text{TSR} \xrightarrow{\text{RNN}} \text{Understand} \end{cases} \xrightarrow{\text{AHP method}} \text{Optimization}$$

Traffic Sign Detection System (TSDS)

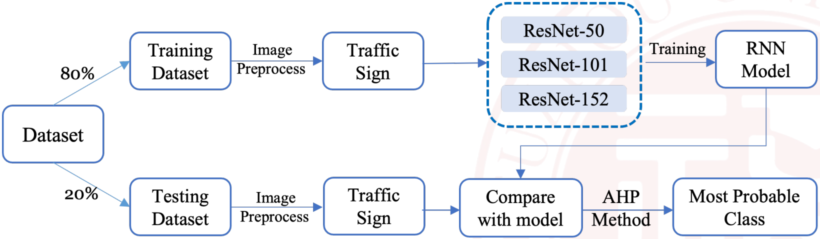


Figure 2: Flowchart of the proposed system

$$\text{Our works} \Rightarrow \begin{cases} \text{TSD} \xrightarrow{\text{YOLO}} \text{Search Sign} \\ \text{TSR} \xrightarrow{\text{RNN}} \text{Understand} \end{cases} \xrightarrow{\text{AHP method}} \text{Optimization}$$

Step 1: Pre-processing

Image Pre-processing

- Edge Detection
  1. Noise processing and gradient computation
  2. Non-maximum suppression
  3. Double threshold detection
- Corrosion Expansion
- After that, a traffic sign without background noise is obtained.

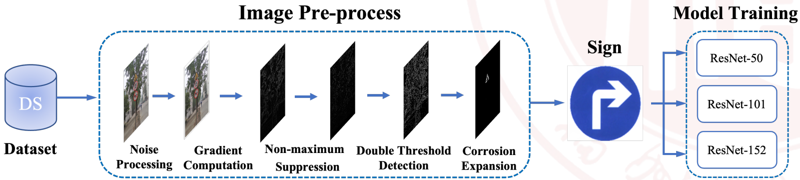


Figure 3: Flowchart of the pre-process

Proposed ResNet Models

- Three ResNet models with different architecture
  1. ResNet 50
  2. ResNet 101
  3. ResNet 152
- Two main types of blocks used in ResNet
  1. Identity Block: Input activation $\xleftrightarrow{\text{same dimension}}$ Onput activation
  2. Conv Block: Input activation $\xleftrightarrow{\text{not match}}$ Onput activation

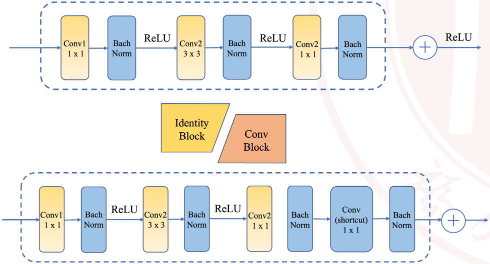
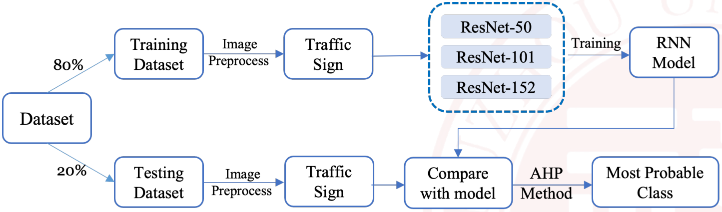
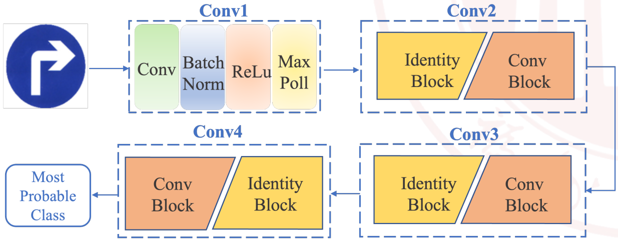


Figure 4: Structures of ResNet blocks

Model Training and Testing





1 Research Overview

2 Proposed Methodology
- Overview of Our Works
- Step 1: Image Pre-processing
- Step 2: Residual Neural Network
- Step 3: Analytic Hierarchy Process

3 Experimental Results

4 Conclusion and Plan

AHP Method

Analytic Hierarchy Process Method

- Maturity level evaluation model
  1. Accuracy (AC)
  2. Stability (SB)
  3. Response Time (RT)
  4. System capability (SC)
- Weight Determination
  1. Indicator judgment matrix
  2. Weight of each indicator

Table 2. Notations of the AHP method

Symbol	Description	Unit
Score	Total score	1
RT	Response time score	1
SB	Stability score	1
AC	Accuracy score	1
SC	System capability score	1
PS	Process Speed	Piture/ms
CPR	Computing Requirement	FLOPS

Step 3-1: Maturity level evaluation model

- Four performance indicators
  1. Accuracy: Correct detection / Total datasets
  2. Stability: Realistic accuracy / Theoretical accuracy
  3. Response Time: Time of processing 1000 images
  4. System capability: Floating-point operations per second (FLOPS)

Table 3 Accuracy score table

Accuracy	***AC (Score)***
Less than 0.75	0
0.75-0.80	1
0.80-0.85	2
0.85-0.90	3
0.90-0.95	4
More than 0.95	5

Table 4 Stability score table

Stability	***SB (Score)***
Less than 75%	0
75%-80%	1
80%-85%	2
85%-90%	3
90%-95%	4
95%-100%	5

Table 5 Response time score table

Process Speed	***RT (Score)***
More than 5s	0
4-5s	1
3-4s	2
2-3s	3
1-2s	4
0-1s	5

Table 6 System capability score table

CPR	***SC (Score)***
More than 3.5G	0
2.0-3.5G	1
1.5-2.0G	2
1.0-1.5G	3
0.5-1.0G	4
0-0.5G	5

Step 3-2: Weight Determination

- Weight Determination
  1. Indicator judgment matrix
  2. Weight of each indicator

The judgment matrix of the indicators is constructed according to the nine-point scale (Table 7), to compare the four indicators in the scores pairs.

Table 7 Nine-point table

Scaling	Definition
1	Factor i is as important as factor j
3	Factor i is slightly more important than factor j
5	Factor i is significantly more important than factor j
7	Factor i is much more important than factor j
9	Factor i is extremely more important than factor j
2,4,6,8	The scale value of the importance of factor i over factor j is between the above two adjacent levels
Reciprocal of scaling value	Inverse comparison of factor i and factor j: $x_{ij} = 1/x_{ji}$

Step 3-2: Weight Determination (Proof)

The elements in the matrix should satisfy:

$$x_{ij} = \frac{1}{x_{ji}}, (i, j = 1, 2, 3, 4) \tag{1}$$

The weight vector can be obtained by the arithmetic mean below

$$\omega_{1i} = \frac{1}{n} \sum_{j=1}^{5} \frac{x_{ij}}{\sum_{k=1}^{5} x_{ki}} (i = 1, 2, 3, 4) \tag{2}$$

And the geometric mean method to find the weight vector is

$$\omega_{2i} = \frac{\left(\prod_{j=1}^{5} x_{ij}\right)^{\frac{1}{5}}}{\sum_{k=1}^{5} \left(\prod_{j=1}^{5} x_{ij}\right)^{\frac{1}{5}}}, (i = 1, 2, 3, 4) \tag{3}$$

Step 3-2: Weight Determination

- Weight of each indicator in AHP evaluation model
  1. Accuracy: 0.342
  2. Stability: 0.211
  3. Response Time: 0.275
  4. System capability: 0.172

The consistency test of the judgment matrix is less than 0.1, indicating that the weight data obtained are valid.

Table 8: Weight of indicators

Indicator	**Weight**
AC	0.342
SB	0.211
RT	0.275
SC	0.172

Methodology of Results Analysis

The experiment result of training and testing is usually assessed by metrics derived from the confusion matrix:

Table 9 Confusion matrix for performance evaluation

Input Image	Positive Predictive	Negative Predictive
Positive Sample	True Positive (TP)	False Nagatice (FN)
Negative Sample	False Positive (FP)	True Negative (TN)

1. TP: Number of positive samples correctly classified as positive.
2. TN: Number of negative samples correctly classified as negative.
3. FN: Number of positive samples incorrectly classified as negative.
4. FP: Number of negative samples incorrectly classified as positive.

Three evaluation metrics

In order to avoid a biased analysis, credible metrics namely False Alarm Rate and Un-Detection Rate metrics were used:

Table 10 Evaluation metrics and explanations

Evaluation Metrics	Corresponding Formula	
Accuracy	$\frac{1}{n}\sum_{1}^{n}\frac{TP+TN}{TP+TN+FP+FN}\times 100\%$	(8)
False Alarm Rate (FAR)	$\frac{1}{n}\sum_{1}^{n}\frac{FP}{FP+TN}\times 100\%$	(9)
Un-Detection Rate (UND)	$\frac{1}{n}\sum_{1}^{n}\frac{FN}{FN+TP}\times 100\%$	(10)

- Three evaluation metrics
  1. Accuracy
  2. False Alarm Rate
  3. Un-Detection Rate

1. Research Overview

2. Proposed Methodology

3. Experimental Results
   - Methodology of Results Analysis
   - Experimental Results Analysis

4. Conclusion and Plan

The experimental results were proposed by comparing the training and testing results of different models based on the same datasets:

Table 11 Experimental results of training and testing

Evaluation Metrics		VGG	GoogleNet	RNN
		Deep Learning Models		
Training (80%)	**Accuracy**	98.24%	98.89%	**99.03%**
	FAR	0.06%	0.03%	**0.01%**
	UND	0.87%	0.86%	**0.41%**
Testing (20%)	**Accuracy**	83.60%	96.62%	**98.01%**
	FAR	2.47%	0.19%	**0.09%**
	UND	56.73%	2.94%	**1.28%**

- Result of ResNet-50 (Training and Testing, Best)
  1. Accuracy: 99.03% and 98.01%
  2. False Alarm Rate: 0.01% and 0.09%
  3. Un-Detection Rate: 0.41% and 1.28%

By comparing the testing results of different parameters, this work also proposed the optimal parameters configurations.

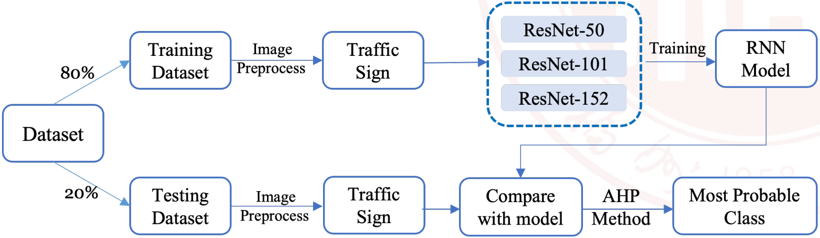
Table 12 Parameter list of the AHP method

Training Parameters	Parameter 1	Parameter 2	Parameter 3
Convolution Layers	**RNN 50**	RNN 101	RNN 152
Learning Rate	**Step**	Low	High
Split Strategy	**Classification Split**	Random Split	
Image Enhancement	**Brightness Enhancing**	**Image Scale**	Contrast Enhancing
Colour Processing	**HSV**	RGB	

- Optimal configurations:
  1. Model Layers: ResNet-50
  2. Learning Rate: Step
  3. Split Strategy: Classification Split
  4. Image Enhancement: Brightness Enhancing
  5. Colour Processing: HSV

Conclusion and Future Plan

- **Conclusion**
  1. Proposed ResNet-50 model got the highest accuracy of 99.03% and 98.01%.
  2. Optimal configurations were suggested through AHP method.
- **Future Plan**
  1. Training and testing based on more complex and large-scale datasets.
  2. Implement image classification capabilities.



Thanks for Listening !