Remote Sensing: (c) Applications

Detection of Collapsed Buildings in Post-Earthquake Remote Sensing Images Based on the Improved YOLOv3

Qiguang Xie

Kexin Wang

CONTENT

1

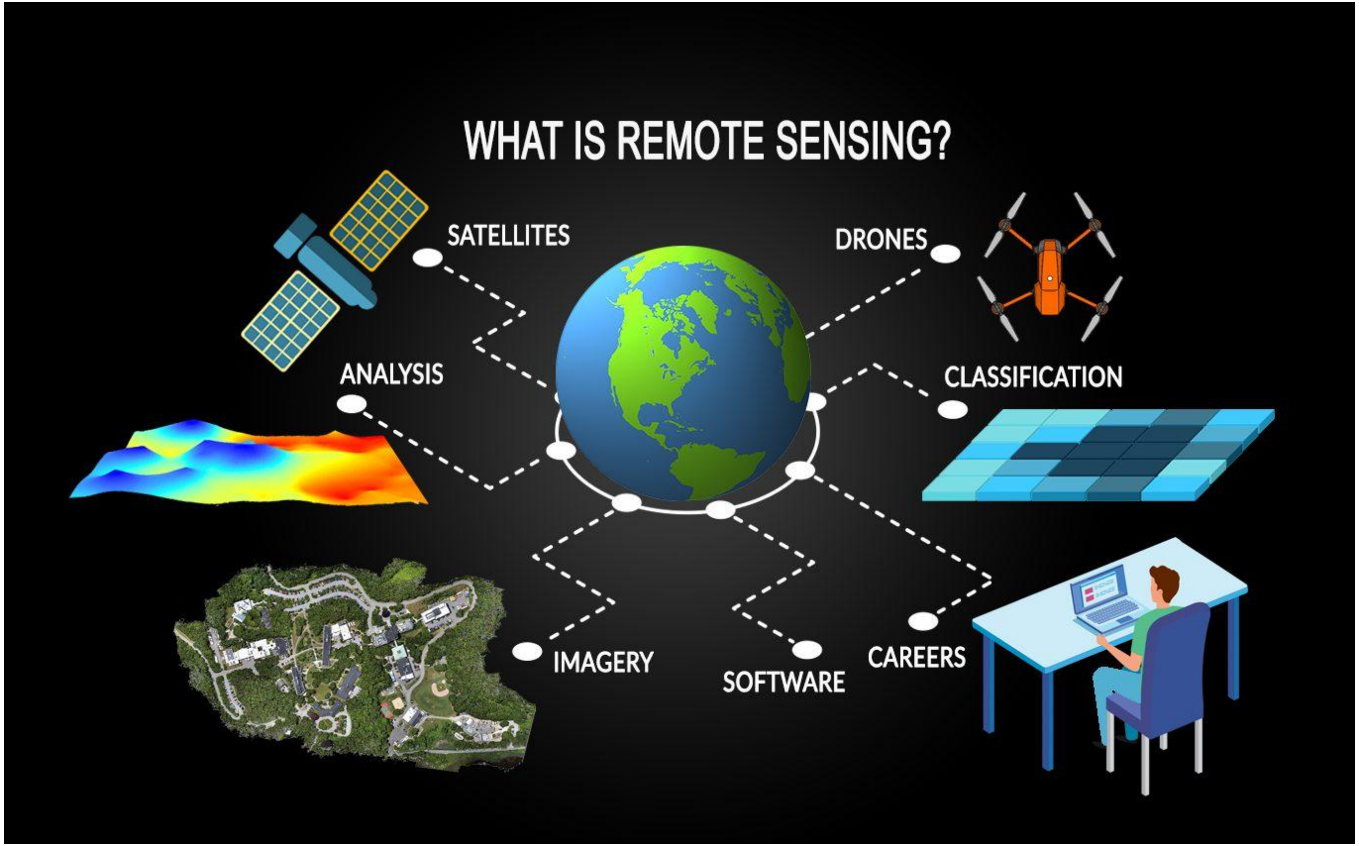
PART ONE

Introduction

Building damage after an earthquake

Remote Sensing

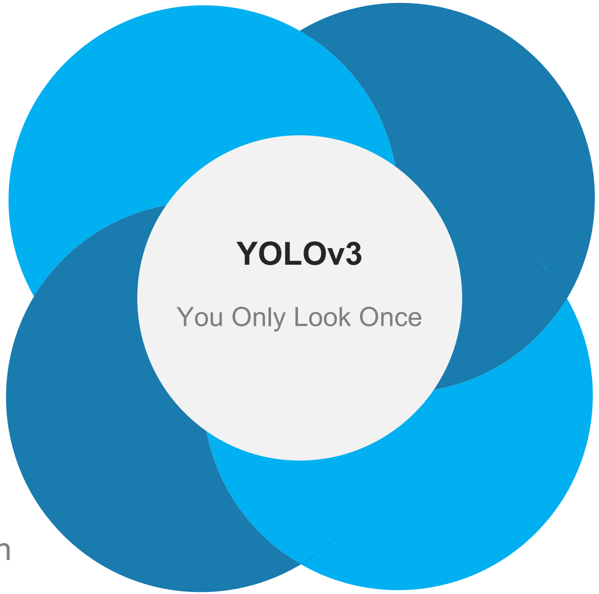
Remote sensing is the acquisition of information about an object or phenomenon without making physical contact with the object and thus in contrast to on-site observation, especially the Earth.



Previous Studies

01. Multi-temporal evaluation method

Mainly based on detecting changes to evaluate the information on building damage.



02. Single-temporal evaluation method

It has become an effective technical means that can be directly used to extract and evaluate information on building damage.

03. Region-based method

The combination of a region-based extractor and detection network(CNN) has become a classic method.

04. Regression-based method

It uses a single CNN to predict the boundary box and classify, and transform the object detection problem into a regression problem.

YOLOv3

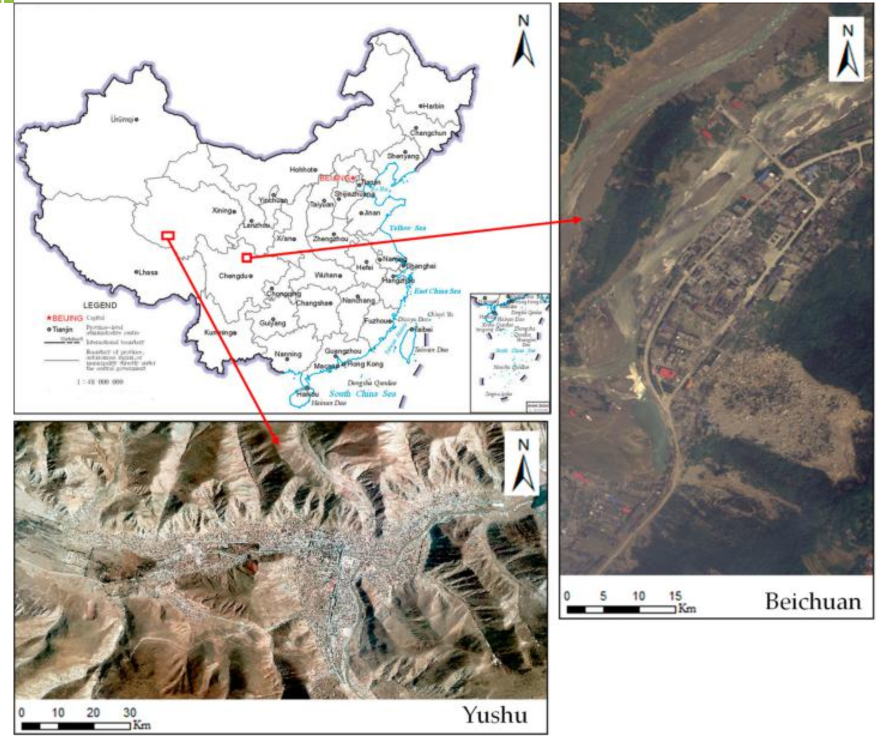
- YOLOv3, a CNN-based object detection method.
- The YOLO series of algorithms have a better generalization capability and faster detection speed than the R-CNN series of algorithms.
- Achieve higher efficiency and precision
- Improve a part of its network structure and loss function to improve the efficiency and accuracy of detection.

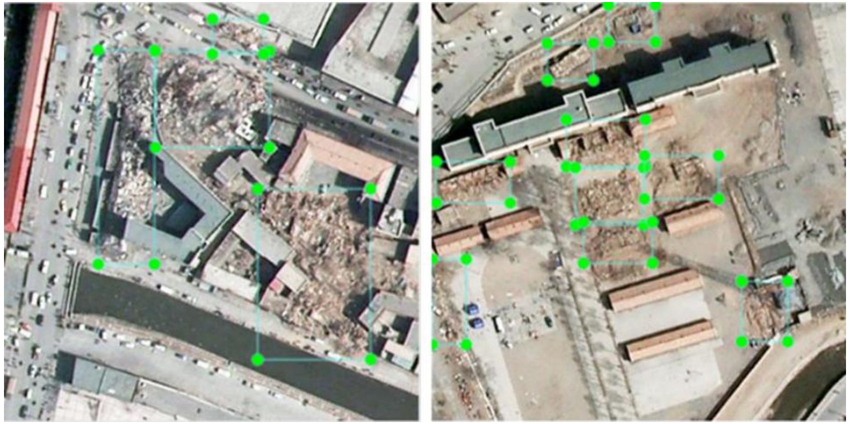
2

PART TWO

Materials and Methods

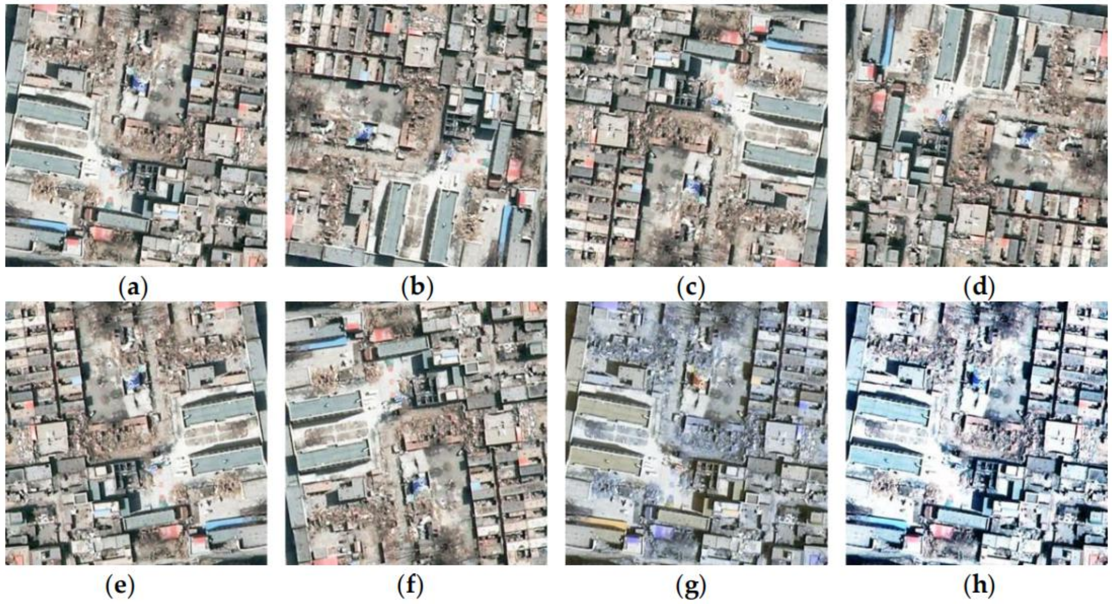
Remote Sensing Data Acquisition





The LabelImg software was used to label collapsed buildings in the image block in PASCAL VOC format.

Dataset Enhancement

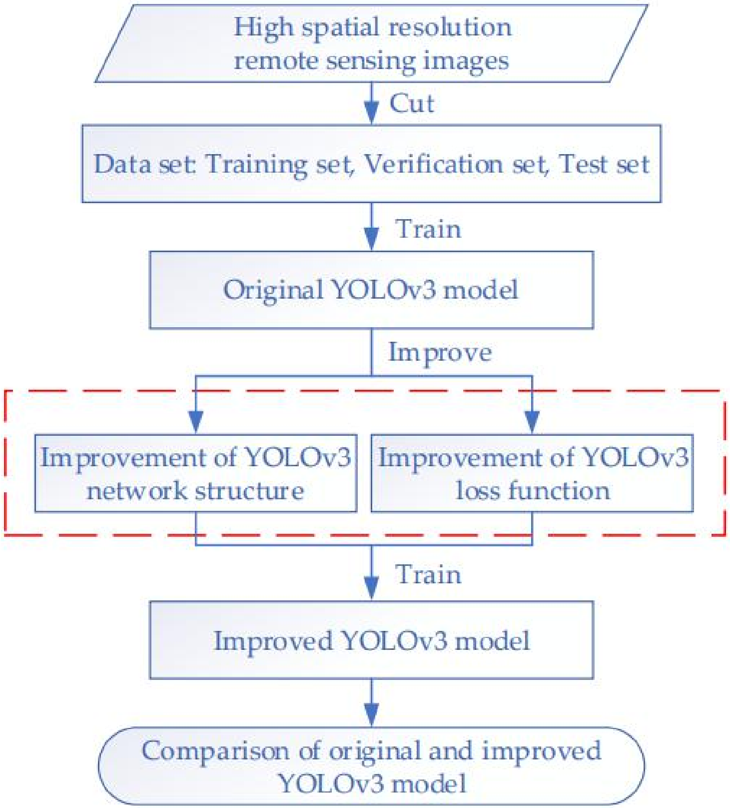


(a)original image
(b)90-degree rotation
(c)180-degree rotation
(d)270-degree rotation
(e)horizontal flip
(f)up-and-down flip
(g)color transformation
(h)image stretch

	Number of Sample Images	Number of Collapsed Buildings
Training set	1456	8751
Validation set	364	2516
Testing set	360	2234

Method Flow

YOLOv3 network:
too complex
too redundant



The red dotted line in the figure is the improvement and optimization of the network structure and loss function of the YOLOv3 model.

Improvement of YOLOv3

Darknet53

complicated and redundant

Shufflenet v2

a lightweight classification network

XY+WH

center coordinate, XY

Generalized intersection over union (GIoU)

replace the regression parameters for the distance loss of the prediction box

3

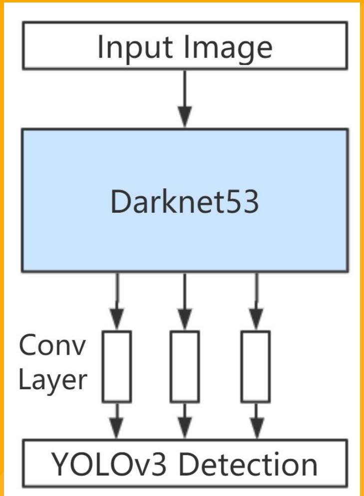
PART THREE

Experimental Settings

3 MODELS

YOLOv3

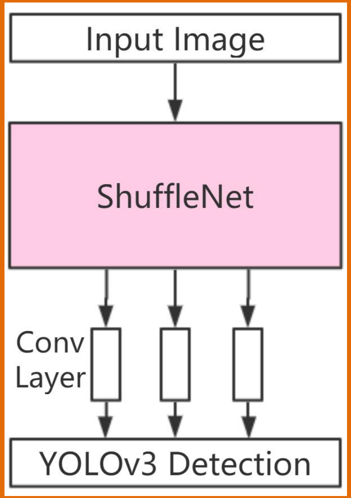
(model Y)



Coord Loss :XY+WH

YOLOv3-ShuffleNet

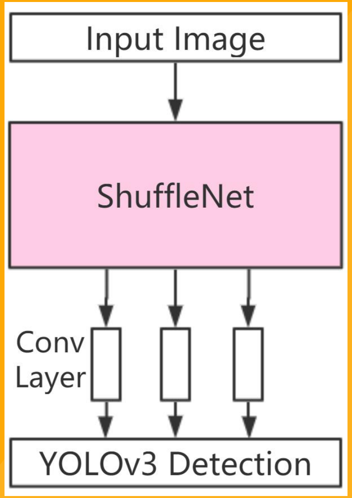
(model YS)



Coord Loss :XY+WH

Yolov3-S-GIoU

(model YSG)



Coord Loss: GIoU

Experimental Settings

Hardware Environment

RTX2080Ti graphics, Intel i7-8700k processor, 32 GB of memory.

Initial Settings

Optimizer: Adam, Image Batch Size:8, Initial Learning Rate:10^-3

Training

If the loss value did not ↓ after 20 epochs, the learning rate should be reduced by 0.1 fold (min learning rate:10^-6).

Performance Indicators

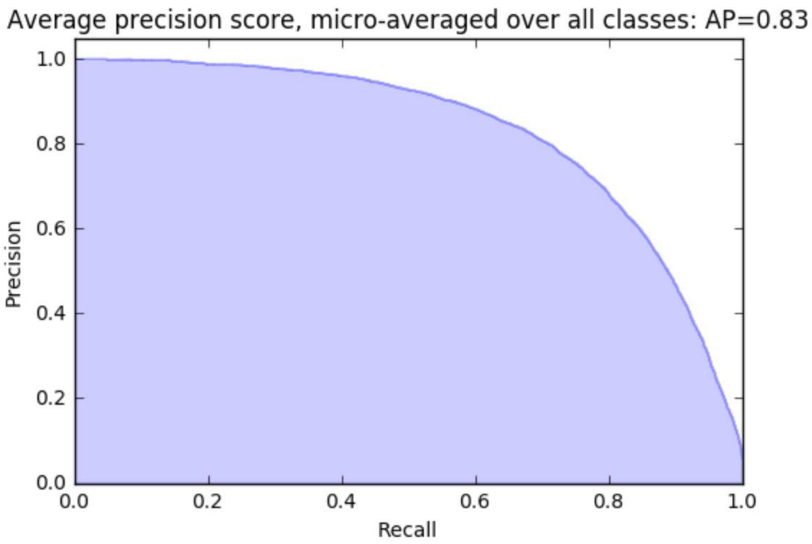
PR-Curve(Precision Recall Curve)

A curve that X-axis is **recall**, $R = \frac{TP}{FN+TP}$,

and Y-axis is **precision** $P = \frac{TP}{FP+TP}$

Table 2. Confusion matrix for predicted results and ground truth.

	Ground Truth	
	Collapsed Building	**Others**
Collapsed building	True Positive (TP)	False Positive (FP)
Others	False Negative (FN)	True Negative (TN)



Performance Indicators

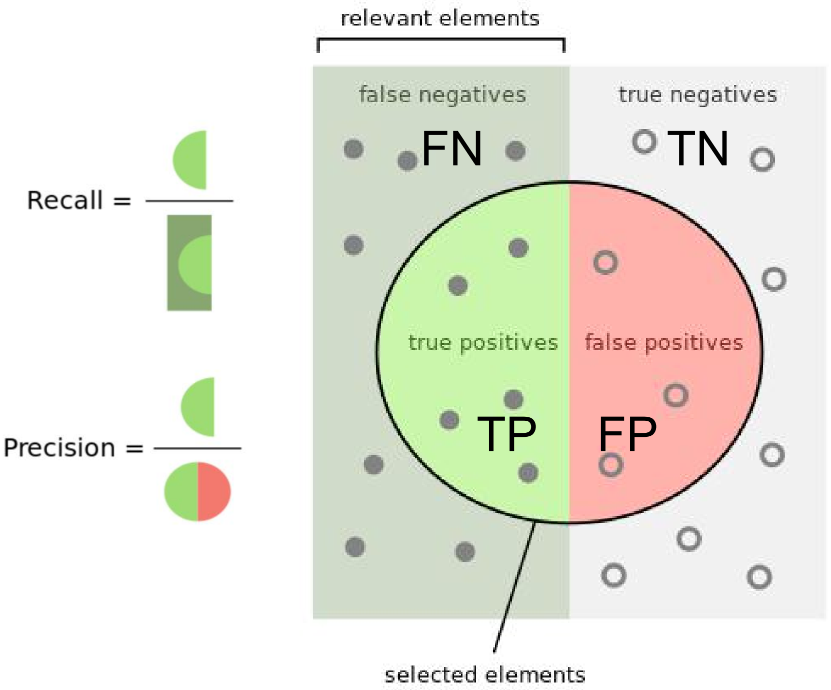
PR-Curve(Precision Recall Curve)

A curve that X-axis is **recall**, $R = \frac{TP}{FN+TP}$,

and Y-axis is **precision** $P = \frac{TP}{FP+TP}$

Table 2. Confusion matrix for predicted results and ground truth.

	Ground Truth	
	Collapsed Building	**Others**
Collapsed building	True Positive (TP)	False Positive (FP)
Others	False Negative (FN)	True Negative (TN)



Performance Indicators

PR-Curve(Precision Recall Curve)

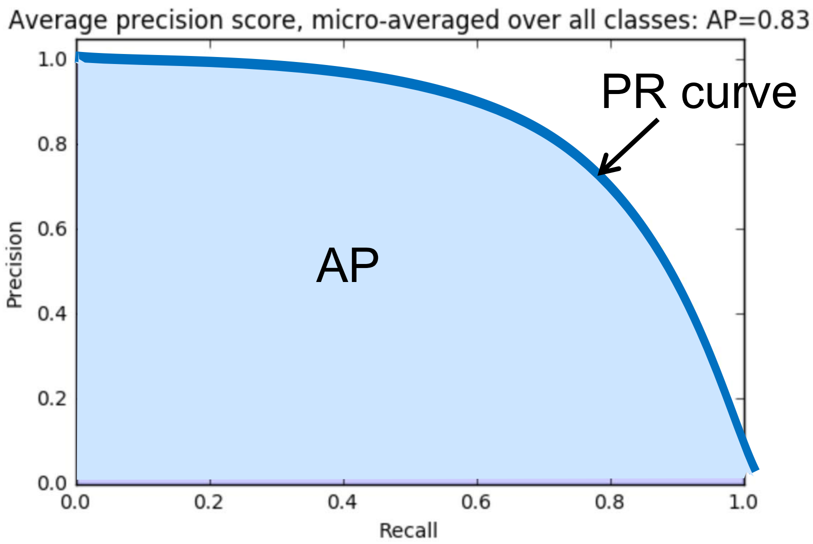
A curve that X-axis is **recall**, $R = \frac{TP}{FN+TP}$,

and Y-axis is **precision** $P = \frac{TP}{FP+TP}$

AP(Average Precision)

Average precision value within all the recall rate(from 0 to 1), which is also **the area under the PR curve**.

$$AP = \sum_{R=0}^{1} (R_{n+1} - R_n) \cdot P_{interp}(R_{n+1}),$$



Performance Indicators

PR-Curve(Precision Recall Curve)

A curve that X-axis is **recall**, $R = \frac{TP}{FN+TP}$,

and Y-axis is **precision** $P = \frac{TP}{FP+TP}$

AP(Average Precision)

Average precision value within all the recall rate(from 0 to 1), which is also **the area under the PR curve**.

$$AP = \sum_{R=0}^{1}(R_{n+1} - R_n) \cdot P_{interp}(R_{n+1}),$$

F1 score

the **harmonic mean** of precision and recall

$$F1 = \frac{2P \cdot R}{P+R}.$$

Speed Indicators

FPS(Frame per second)

the number of pictures processed per second (f/s)

FPS>=30 : real-time processing

This animated cartoon of a horse is displayed at 12 drawings per second,its FPS is 12

4

PART FOUR

Results

Loss Function

Loss Function

In YSG model,

1.violent jitter is shorter

2.loss value is mush lower

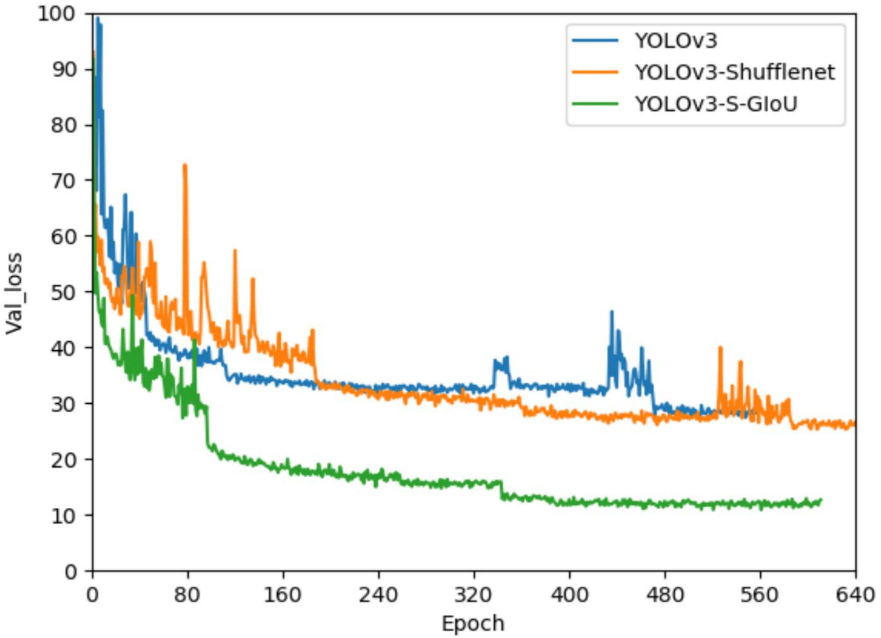


Figure 12. Loss curves on the verification set of three YOLOv3 models.

Quantitative Evaluation

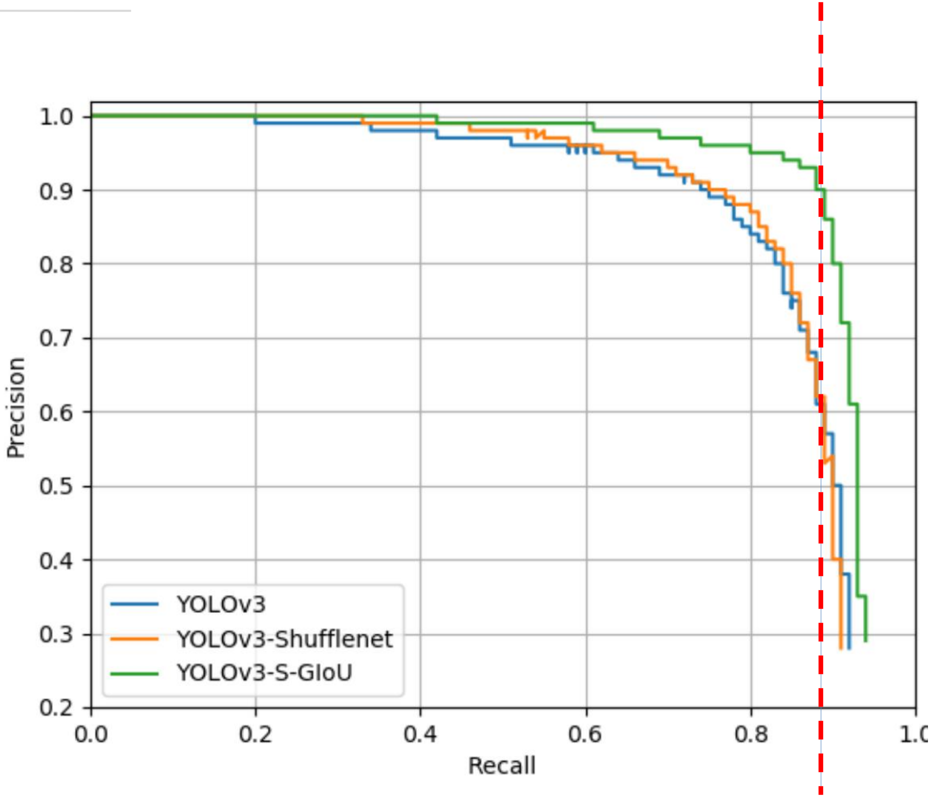
Table 3. Performance comparison between YOLOv3 and the two improved models.

	P (%)	R (%)	F1 (%)	AP (%)	FPS (f/s)	Parameter Size (M)
YOLOv3	88	78	82.7	85.84	23.95	241
YOLOv3-ShuffleNet	87	81	83.89	85.98	29.16	146
YOLOv3-S-GIoU	93	88	90.43	90.89	29.23	146

P-R Curve

PR-Curve

When the recall rate was approximately 0.88, the precision of model Y and YS declined to only about 0.6 but the precision for YSG remained at about 0.93

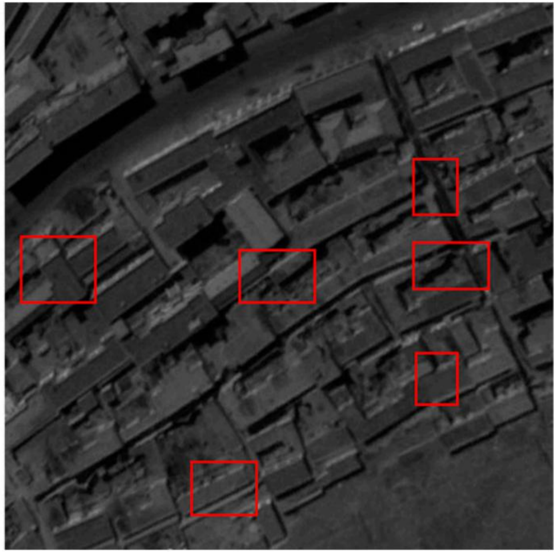


Detections

Model Y

Model Y-S-G

Before Earthquake



Robustness

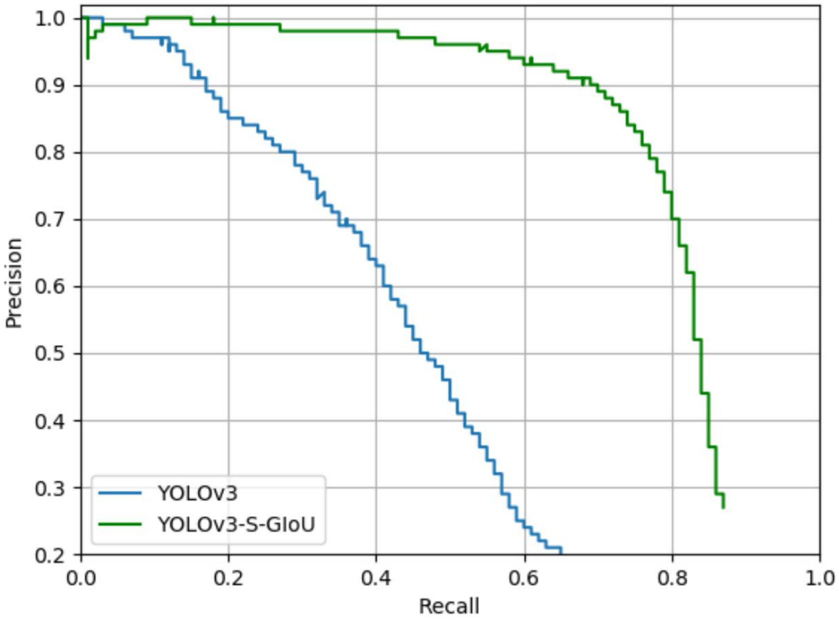
Original Image

Add Gaussian noise

Add salt-pepper noise

	P (%)	R (%)	F1 (%)	AP (%)
YOLOv3	63	41	49.67	44.3
YOLOv3-S-GIoU	86	74	79.55	79.8

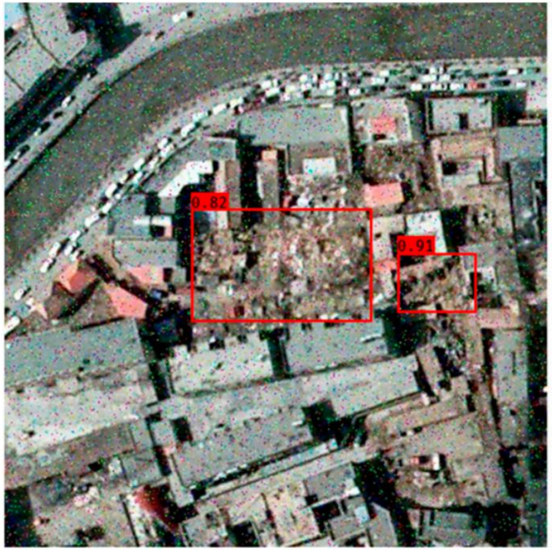
Robustness



	P (%)	R (%)	F1 (%)	AP (%)
YOLOv3	63	41	49.67	44.3
YOLOv3-S-GIoU	86	74	79.55	79.8

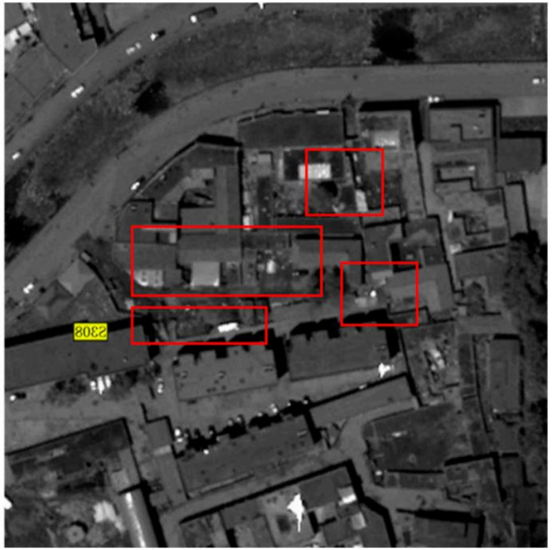
Robustness

Model Y



Model Y-S-G

Before Earthquake



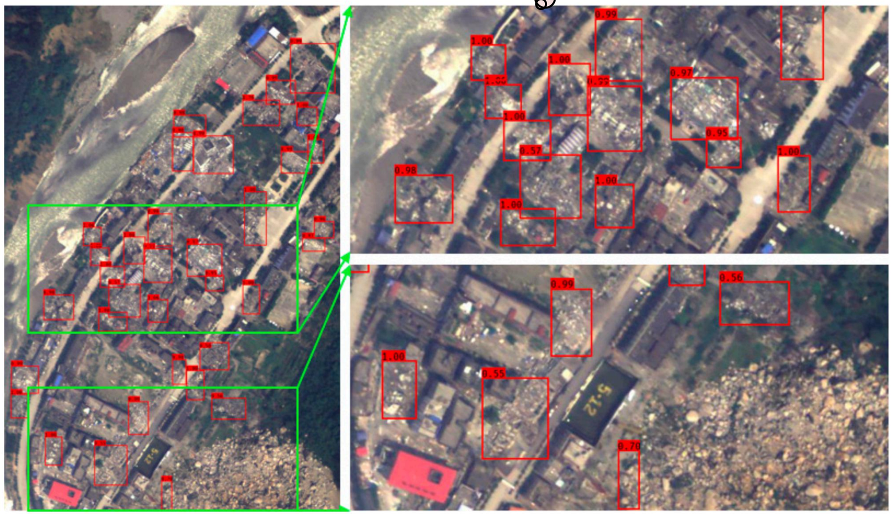
	P (%)	R (%)	F1 (%)	AP (%)
YOLOv3	63	41	49.67	44.3
YOLOv3-S-GIoU	86	74	79.55	79.8

Conclusion

Future works?

Conclusion

- The experimental results show that the improved YOLOv3 model (YSG) had sufficient robustness and a certain anti-noise ability.
- The test set reached 29.23 f/s, the average precision reached 90.89% and a significant reduction in the number of parameters, i.e., only 146 MB.



QUESTIONS?

THANKS FOR YOUR WATCHING