Shiliang PU, Wei ZHAO, Weijie CHEN, Shicai YANG, Di XIE, Yunhe PAN, 2021. Unsupervised object detection with scene-adaptive concept learning. *Frontiers of Information Technology & Electronic Engineering*, 22(5):638-651. https://doi.org/10.1631/FITEE.2000567

Front Inform Technol Electron Eng

Unsupervised object detection with scene-adaptive concept learning

Key words: Visual knowledge; Unsupervised video object detection; Scene-adaptive learning

Corresponding author: Di XIE
E-mail: xiedi@hikvision.com
ORCID: https://orcid.org/0000-0001-8065-5901

Motivation

1. The mainstream detection methods still cannot generally be used across different kinds of scenes and cannot autonomously evolve along with the accumulated unlabeled scene data.

2. For changing scenes, new manual annotations and new models are expensive and time-consuming.

3. The concepts and prototypes of object groups can be used to weaken the influence of the scene change.

Front Inform Technol Electron Eng

Main idea

1. The classes and object appearances are similar in the original and new scene data, which can weaken the influence of the scene change.

2. Positive samples of the same class have similar prototypes.

3. Graph neural network and group information propagation can be used to mine similar object prototypes.

Front Inform Technol Electron Eng

Method

1. A scene-adaptive evolution unsupervised video object detection method is proposed which consists of a prototype dictionary generation (PDG) strategy and a graph-based group information propagation (G-GIP) strategy.

2. PDG is introduced to weaken the influence of the scene change on object detection.

3. G-GIP creates the prototype category, which can mine the positive samples with pseudo labels from each object group. Then, the new model is fine-tuned by positive samples.

Front Inform Technol Electron Eng

Method (Cont'd)

Framework of the proposed method

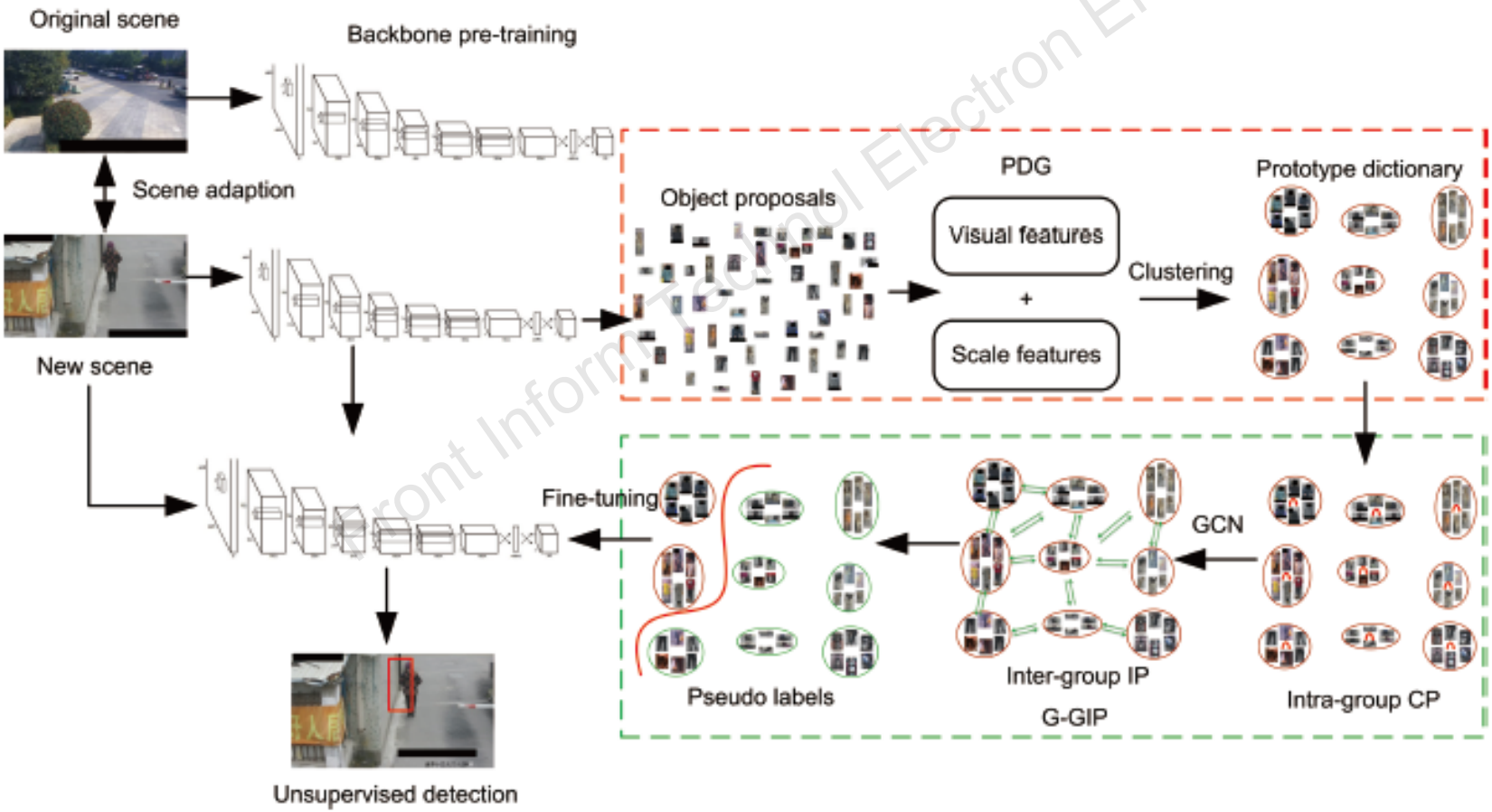


Fig. 1 Framework of the proposed method

Front Inform Technol Electron Eng

Major results

1. Monitoring test data results of our model and base model

Front Inform Technol Electron Eng

Table 6 RF0.1 of all video surveillance datasets

Model	RF0.1 (%)						Mean (%)
	J-road	S-path	S-gate	SW-gate	Residential-M	Residential-K	
YOLO-v3	68.08	63.48	40.36	77.67	74.59	94.93	69.85
G-GIP	73.64	72.17	50.07	82.63	85.64	96.46	76.80

Table 7 AP of all video surveillance datasets

Model	AP (%)						Mean (%)
	J-road	S-path	S-gate	SW-gate	Residential-M	Residential-K	
YOLO-v3	95.20	83.73	60.60	91.75	84.82	96.17	85.38
G-GIP	96.31	87.37	66.02	92.60	90.41	97.29	88.33

Major results (Cont'd)

2. Test results of our model and related methods on public datasets

Table 8 AP of the KITTI-to-Cityscapes (K2C) and Sim10k-to-Cityscapes (S2C)

Model	AP (%)	
	K2C	S2C
Khodabandeh et al. (2019)'s	43.00	42.60
Zhu XG et al. (2019)'s	42.50	43.00
Shen et al. (2019)'s	41.90	42.60
He and Zhang (2019)'s	41.00	41.10
Faster R-CNN	35.76	33.10
G-GIP(F)	42.75	37.81
YOLO-v3	47.90	53.70
G-GIP(Y)	51.50	59.40

Front Inform Technol Electron Eng

Conclusions

1. The proposed scene-adaptive evolution unsupervised video object detection algorithm can adapt to and evolve with changing scenes. The experimental results showed the effectiveness of the proposed method.

2. The PDG strategy can be used to build a dictionary of object prototypes to represent the visual knowledge of object groups.

3. The G-GIP strategy can be used to create the object prototype category for mining positive samples from the unlabeled new scene dataset.

Front Inform Technol Electron Eng



Shiliang PU received his PhD degrees from University of Rouen and Zhejiang University in 2005 and 2007, respectively. He is currently the chief research scientist at Hikvision and president of Hikvision Research Institute. His research interests include AI, machine perception, and robotics.



Di XIE received his BS degree in software engineer and his PhD degree in computer science from Zhejiang University, in 2007 and 2012, respectively. He is currently a senior research scientist and a principal research manager at Hikvision Research Institute. His research interests include computer vision, detection & recognition, and neural network optimization.

Front Inform Technol Electron Eng