25th International Conference on Pattern Recognition - ICPR 2020

DETECTING ANOMALIES FROM **VIDEO-SEQUENCES: A NOVEL DESCRIPTOR**

GIULIA ORRÙ, DAVIDE GHIANI, MAURA PINTOR, GIAN LUCA MARCIALIS, FABIO ROLI Department Of Electrical And Electronic Engineering, University Of Cagliari, Italy





CROWD ANALYSIS









A NOVEL DESCRIPTOR FOR ANOMALOUS EVENTS DETECTION IN HIGH-DENSITY CROWDS

- in the scenario
- Inspired by the concept of one-dimensional local binary pattern



Chatlani, N., & Soraghan, J. J. (2010). Local binary patterns for 1-D signal processing. 95-99. Paper presented at 18th European Signal Processing Conference (EUSIPCO-2010), Aalborg, Denmark.

based on the hypothesis that anomalous events happen when multiple group formation events and group breaking-up events suddenly appear









LOW LEVEL FEATURES: GROUP COUNTS

We compare 4 methods for counting the groups:

- Manual counting as ground truth (MC)
- Clustering of optical flow (COF)
- OpenCV Cascade detector (CD)
- Blob detector (BD) [our implementation]

G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.

B. K. Horn and B. G. Schunck, "Determining optical flow," Artificial intelligence, vol. 17, no. 1-3, pp. 185–203,1981.









HIGH LEVEL FEATURES: TRINARY CODES









EXPERIMENTAL PROTOCOL

- Motion Emotion data set: 31 video sequences of around 44000 frames in total.
- (panic, fight, congestion, obstacle and neutral).

Metrics:

$$precision = \frac{TP}{TP + FP} \qquad F1_{score} = 2 * \frac{pre}{pre}$$

Descriptor parameters have been set using a grid search, which maximize the F1score:

- in a supervised way F1score maximised on all videos
- left out

The videos contain both normal and abnormal behavior, labeled frame-by-frame as 5 classes

ecision * recall $recall = \frac{TP}{TP + FN}$ ecision + recall

• with a Leave-one-out cross validation F1score maximised on N-1 videos and test on the video





EXPERIMENTAL PROTOCOL

- Experiment 1: set of videos recorded by the camera in frontal position with respect to the scene.
- Experiment 2: set of videos recorded by the camera in lateral position with respect to the scene.



Frontal view, slope $< = 5^{\circ}$



Lateral view, slope $> 5^{\circ}$





OVERALL RESULTS

	Supervised			Leave-one-out		
	Precision	Recall	F1	Precision	Recall	F1
MC	88.89%	94.12%	91.43%	79.31%	71.87%	75.41%
COF	71.11%	88.89%	79.01%	52.50%	60.00%	56.00%
CD	75.00%	91.67%	82.50%	73.17%	83.33%	77.92%
BD	70.45%	86.11%	77.50%	56.52%	74.29%	64.20%

 Different position of the camera leads to very different performances: better performance on lateral videos





ANALYSIS OF INDIVIDUAL CASES: STRUCTURED CROWD



VIDEO 009





ANALYSIS OF INDIVIDUAL CASES: UNSTRUCTURED CROWD









VIDEO 023

Frame number



70

70



CONCLUSIONS

WHITE BOX DESCRIPTOR

CONTROL OF STRUCTURED CROWDS

PIPELINE WITHOUT DELAY

THANKS FOR THE ATTENTION!

RESEARCH (MIUR) WITHIN THE PRIN2017 - BULLYBUSTER - A FRAMEWORK FOR BULLYING AND CYBERBULLYING ACTION DETECTION BY COMPUTER VISION AND ARTIFICIAL INTELLIGENCE METHODS AND ALGORITHMS (CUP: F74I19000370001)





THIS WORK IS SUPPORTED BY THE ITALIAN MINISTRY OF EDUCATION, UNIVERSITY AND

giulia.orru@unica.it



