2014 22nd International Conference on Pattern Recognition

Motion Interaction Field for Accident Detection in Traffic Surveillance Video

Kimin Yun*, Hawook Jeong*, Kwang Moo Yi†, Soo Wan Kim* and Jin Young Choi*
* Perception and Intelligence Lab, ASRI
Department of Electrical and Computer Engineering
Seoul National University, South Korea.
$\{ykmwww, hwjeong, soowankim, jychoi\}@snu.ac.kr$
† Computer Vision Laboratory
School of Computer and Communication Sciences (IC)
Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Switzerland.
$kwang.yi@epfl.ch$

***Abstract*—This paper presents a novel method for modeling of interaction among multiple moving objects to detect traffic accidents. The proposed method to model object interactions is motivated by the motion of water waves responding to moving objects on water surface. The shape of the water surface is modeled in a field form using Gaussian kernels, which is referred to as the Motion Interaction Field (MIF). By utilizing the symmetric properties of the MIF, we detect and localize traffic accidents without solving complex vehicle tracking problems. Experimental results show that our method outperforms the existing works in detecting and localizing traffic accidents.**

I. Introduction

In these days, a large number of cameras are being installed for traffic monitoring purposes since the demand for accident detection and analysis is being increased. Since recent traffic surveillance systems mostly rely on human observation, it is difficult to monitor a large number of camera scenes at the same time and recognize abnormal events without missing. In order to overcome this limitation, much efforts have been tried to develop an automatic detection method via computer vision and pattern recognition techniques, but the level of current technique is still limited to be applied in actual environments. The existing methods for traffic accident detection have been developed via three approaches: modeling of traffic flow patterns, analysis of vehicle activities, and modeling of vehicle interactions.

For modeling of traffic flow patterns, they use a learning scheme to model the typical patterns of traffic flow according to traffic rules such as go-straight, U-turn, turn-right, and so on [1]–[3]. The typical traffic patterns are recognized as normal events and outliers are treated as abnormal traffic events. This approach is valid only when the normal pattern appears at a fixed region repeatedly, so it could not detect infrequent event such as collision. The analysis of vehicle activities [4]–[6] is done by using motion features extracted from moving vehicles after detecting and tracking of vehicles. The traffic accident can be detected by analyzing pattern changes such as distance between two vehicles, acceleration and direction of a vehicle. The technical level of this approach is not sufficient to be applied to actual environments because the detection and tracking performance is not satisfactory in crowded traffic scenes. The approach of interaction modeling has been inspired by sociological concepts for abnormal event detection where the social force model [7] and the intelligent driver model [8] have been used to model the interaction among vehicles and detect accidents. These methods do not show sufficient performance because they use only speed change information. In addition, they require an offline learning process with a large number of training data.

In this paper, we propose a new field-based method for modeling interactions among multiple moving objects to effectively detect and localize traffic accident without vehicle tracking and complex learning process. The proposed interaction model is inspired by the movement of water surface when multiple objects are moving on water. When an object moves on water, it pushes water molecules and creates waves where the water surface rises up at the front of the object and falls down at the backside of the object. This natural phenomenon is modeled in a field form using Gaussian kernels depending on both the speed and the direction of each moving object. We refer to this model as the Motion Interaction Field (MIF). In addition, we develop a criterion to detect traffic accidents from observing MIF characteristics. Through experiments, we show our method outperforms the state-of-the-art in detection and localization of traffic accident in video streams.

II. The Proposed Method

The overall scheme of the proposed method is depicted in Fig. 1. The motion information (speed and direction) of an object is obtained from recent optical flow algorithm[1] [9] based on [10], and the motion information is used to generate MIF for a scene. A kernel function is generated for each optical flow instead of object since the segmentation of an object is not always available in a crowded scene or requires much computation. The MIF for a scene is built by superposing the multiple kernel functions generated for each optical flow. The MIF depicts an overall view of interactions

[1]The code is available in http://people.csail.mit.edu/celiu/OpticalFlow/

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.528

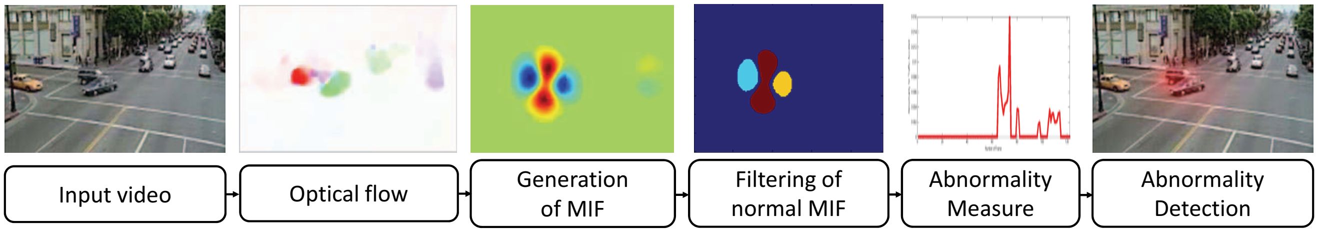


Fig. 1. The overall framework of the proposed method. *Best viewed in color.*

among moving objects within a scene. From the MIF, we can easily grasp the interaction status in a scene, *e.g.,* interaction intensity, extent of interaction, direction of interaction, and so on. To detect abnormality, we use a measure calculated from the shape of MIF, *e.g.,* maximum value(intensity) and symmetry representing direction and extent of interaction. The localization of accident region is visualized by blending the input video and MIF value.

A. Motion Interaction Field

The MIF illustrated as a field form in Fig. 2(c) is a new aspect distinctive from the existing approaches where the interaction is described by the particle's velocity variation with the advection method(*e.g.,* social force model [7] and intelligent driver model [8]) in Fig. 2(a) or by the network of related particles(*e.g.,* interaction energy potential [11]) in Fig. 2(b). The proposed MIF is designed to mimic the change of water surfaces when objects are moving on the water. When two objects become close to each other in a water, the height of the water surface between the two objects rises up. On the other hand, when one follows the other in the same direction, water surface level between the objects does not change. The former case corresponds to a car collision and the latter case represents the normal traffic. Hence the MIF mimicking the water shaping can be expected to be utilized for detection of abnormal interaction such as car collision.

To construct MIF, we design a kernel function to represent the intensity of interaction arising from an optical flow in a moving object. The function value should be proportional to the speed of object because the object with high speed has larger interaction force than the low speed objects. However, only the speed cannot provide complete information on the abnormality in diverse situations. For example, the collision of two objects with low speeds is more abnormal than a single moving object with high speed. For this reason, we also consider the direction of the objects to measure the interaction in the perspective of abnormality. That is, our model is designed so that the interaction intensity between two objects moving in opposite direction becomes larger than that between two objects moving in the same direction. For example, two objects moving to each other as shown in Fig. 3(c), (d) bring a higher interaction value than two objects going in the same direction as shown in Fig. 3(a), (b).

To make the kernel function which can describe the characteristics mentioned in the above, we use two Gaussian

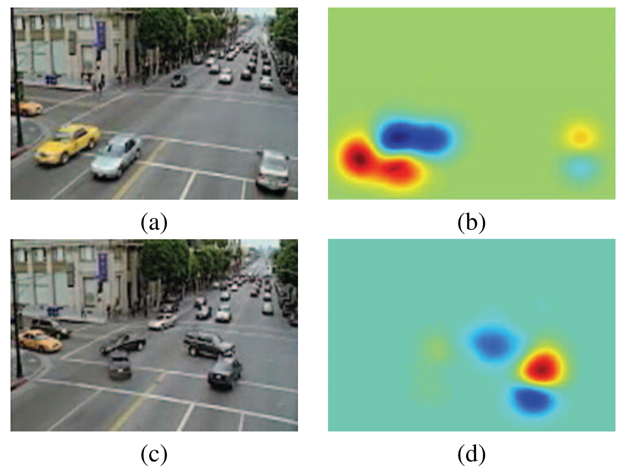


Fig. 3. Normal interaction and abnormal interaction. (a)-(b) Normal interaction and its MIF (c)-(d) Abnormal interaction and its MIF.

functions parameterized with direction and speed of a motion. With image positions (x_i, y_i) and the corresponding velocity (v_{x_i}, v_{y_i}), the kernel $K(x, y; x_i, y_i)$ is calculated by subtraction of two Gaussians with different center positions: one is $(x_i + v_{x_i}, y_i + v_{y_i})$ for forward direction and the other is $(x_i - v_{x_i}, y_i - v_{y_i})$ for backward direction as follows,

$$\begin{aligned} K(x, y; x_i, y_i) = k(x, y; x_i + v_{x_i}, y_i + v_{y_i}) \\ - k(x, y; x_i - v_{x_i}, y_i - v_{y_i}), \end{aligned} \tag{1}$$

where $k(x, y; x_c, y_c)$ is the general 2D Gaussian distribution with its center (x_c, y_c) as

$$k(x, y; x_c, y_c) = \exp\left(-\left(\frac{(x - x_c)^2}{2\sigma_x^2} + \frac{(y - y_c)^2}{2\sigma_y^2}\right)\right). \tag{2}$$

Then, we apply this kernel to all moving pixels. If we denote the MIF as $F(x, y)$ then,

$$F(x, y) = \sum_{x_i, y_i} K(x, y; x_i, y_i). \tag{3}$$

B. Filtering of Normal MIF

The proposed MIF shows different properties depending on normal/abnormal events in a traffic scene. In the previous section II-A, our kernel is designed to make the field have a symmetric structure if traffic situation is normal. Hence the

(a) (b) (c)

Fig. 2. The types for description of interaction. (a) Particle motions in Social force [7] (b) Network of neighbors in Energy potential [11] (c) Motion interaction field in our method. *Best viewed in color.*

MIF in a normal region could be filtered out by checking the symmetry of MIF. The Fig. 3 (b) shows symmetric interaction (normal case), whereas (d) shows non-symmetric interaction(abnormal case). This symmetry could be recognized by checking the one-to-one correspondence between the set of positive regions (red regions) and the set of negative regions (blue regions) of MIF as shown in Fig. 3. When two cars move side by side in parallel, the resulting field might be merged as shown in Fig. 3 (c), but still shows symmetric property. When multiple cars move one by one in series with the same direction, the field still remains in symmetry because the field between cascading cars may be cancelled out by overlapping of positive and negative kernels. In the abnormal traffic situation like accident, however, the symmetry property is broken, which can be utilized as a criterion to detect occurrence of the abnormal interaction as shown in Fig. 3(b).

The filtering procedure does a key role in our algorithm to improve performance in view of anomaly detection accuracy and anomaly localization ability. In the existing algorithms for detecting abnormal traffic situation, the abnormality is detected by using the holistic feature such as histogram on the entire area. In this case, the coverage of accident region takes a small portion of scene, so the influence of the normal regions is dominant in making the holistic feature. Hence this kind of holistic feature may not be discriminative in detecting local accident, while the filtering of normal MIF in our algorithm can focus on the abnormal region and enhance the detection performance.

Fig. 4 is a graphical explanation to remove the regions showing symmetry. First, the interaction field is divided into positive and negative regions by simple segmentation through thresholding on the field magnitude. To check the one-to-one correspondence between the set of positive regions and the set of negative regions, each region chooses the most similar shaped region with opposite sign(the detailed matching procedure will be described in the next paragraph). In Fig. 4, for example, the pair of region 2 and 5 is a one-to-one match and so the region 2 and 5 are regarded as normal. Although the pair region 1 and 3 is matched to each other, this pair is not a one-to-one match because region 4 is also matched to region 1. According to the matching result, the filtered MIF F_s is obtained as shown in the fourth column of the Fig. 4.

To drive the matching procedure, we define the similarity between the i-th positive region and the j-th negative region, where the similarity uses the shape and the degree of neighbors. If the two regions are not in neighbors, the similarity is set to zero. In order to find the neighboring regions, each region is fitted to the circle which covers the region. If these circles are overlapped, they are defined as neighboring regions. Fig. 4(b) shows this procedure and the mathematical expression is defined as following

$$S(i,j) = \begin{cases} S_{shape}(i,j) & \text{if } \|c_i - c_j\|_2 < r_i + r_j \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

where c_i, r_i and c_j, r_j are the center positions and radii of fitted circle of region i and j, respectively.

The shape similarity $S_{shape}(i,j)$ is defined as differences of Hu's image moments [12]. The shape similarity is defined as

$$S_{shape}(i,j) = \sum_{k=1,\ldots,7} \left| m_k^i - m_k^j \right|, \tag{5}$$

$$m_k^i = \mathrm{sign}(h_k^i) \cdot \log h_k^i, \tag{6}$$

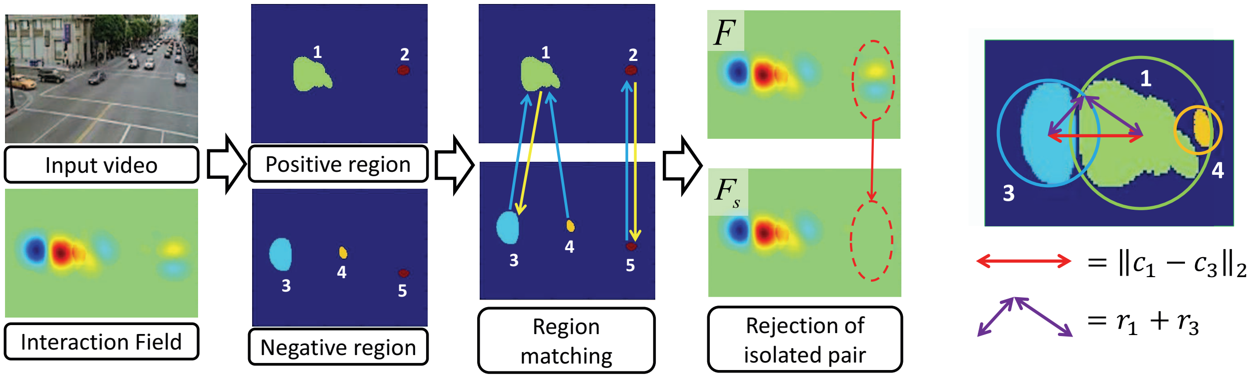
$$m_k^j = \mathrm{sign}(h_k^j) \cdot \log h_k^j, \tag{7}$$

and h_k^i, h_k^j are the Hu's image moments of region i and j, respectively. Image moment is a weighted average of the pixel intensities which includes shape properties of the image such as area, centroid and degree of rotation. Among various image moments, Hu's set of moments are most frequently used as a shape descriptor, because Hu's set of moments is invariant under translation, rotation, scale [12]. If the scene has projective distortion, we cannot use the simple shape feature like region area because the magnitude of optical flow is different along the position. On the contrary, Hu's set of moments has invariant properties in describing the shape and it is robust against projective distortion. With this similarity measure, we find regions which obey the one-to-one match rule and construct F_s by removing the one-to-one match pair from MIF F.

C. Abnormality measure using temporal information

1) Temporal Propagation: Using the F_s the degree of abnormality (abnormal MIF intensity) M_F is defined as

$$M_F = \max_{x,y} F_s(x,y). \tag{8}$$



(a) (b)

Fig. 4. (a) Overall procedure to check the local symmetry of MIF. First we divide the interaction field into positive and negative regions. Based on the similarity matrix by equation (4), each region chooses the most similar region with opposite sign. Then, the pair of regions which holds symmetry are filtered out and F_s is obtained as a result. (b) The method to find the neighboring regions. First, each region is fitted to circle which covers the area. Then, the neighboring relation is decided by checking if two circles are overlapped. *Best viewed in color.*

This measure does not contain temporal information because F_s is obtained from consecutive two frames. Since there is no motion of car after accident, this measure can detect only a start time of the accident. To complement this weak point, the dynamic motion of the MIF intensity is modeled by auto regression model along the time axis, which reflects the propagation property of water surface. Mathematical expression is given as

$$M_{Abn}^t = \begin{cases} M_F^t + \alpha M_{Abn}^{t-1} & \text{if } M_{Abn}^{t-1} > threshold \\ M_F^t & \text{otherwise} \end{cases} \quad (9)$$

where M_F^t indicates M_F at t-frame, α is the decaying parameter representing the degree of propagation and the range of α is 0 to 1. Each frame is then decided as either normal or abnormal through a threshold on M_{Abn}^t. This equation is defined as following

$$L^t = \begin{cases} Abnormal & \text{if } M_{Abn}^t > threshold \\ Normal & \text{otherwise} \end{cases} \quad (10)$$

where L^t is the final label of abnormality at t-frame.

2) Temporal Constraint in Region Matching: The number of positive and negative regions might be changed suddenly due to inaccurate optical flow, which occurs unexpectedly by using sparse optical flow for real time purpose. In this case, we regard the obtained MIF is not credible and we maintain the previous decision on abnormality without further processing of the region matching. The effectiveness of this constraint is evaluated through the ROC curve in the experiments.

III. Experiments

To validate the effectiveness of our method, we compared our method against the Intelligence Driver Model(IDM) [8], which is an interaction based approach without vehicle tracking like our method. For comparison, we used the car accident dataset [8] containing 8 videos with sudden accidents in traffic. Quantitative evaluation and comparison with IDM are presented in terms of ROC curve(*i.e.,* true positive ratio versus false positive ratio) by varying the threshold in (10). We used fixed parameters in all experiments of the proposed algorithm: $\sigma_x = \sigma_y = 10$, $\alpha = 0.8$.

Fig. 5 shows the accident detection result of the proposed method. Columns (a) and (c) of Fig. 5 are sample frames of normal situations from each accident videos whereas columns (b) and (d) are the accident detection results of the proposed method. As shown in Fig. 5, the local region of accident has been depicted by using the intensity of MIF decided as abnormal, which helps us understand the degree of abnormality and localization of abnormal accident.

Fig. 6 shows the accident localization results by the IDM and our method. The IDM method uses a watershed segmentation algorithm on gradients of IDM's abnormality value. Therefore, several normal regions are erroneously detected as abnormal region as shown in Fig. 6(a). On the contrary, our method does not use a segmentation method and just blend the input frame with the MIF. This simple blending is sufficient to localize the accident region because MIF has continuos value and reflects the abnormal degree.

In Fig. 7(a), it shows a false positive example of the IDM method, which is actually a normal event. However, the false positive case in our method was the case of high interaction forecasting accident to be happen even if it was not accident as shown in Fig. 7(b).

Fig. 8 shows the ROC curves for three methods : Intelligent Driver Model, the proposed method without temporal constraint in Section II-C2, and the proposed method. The average area under ROC(AUC) is given in Table I. As shown in Fig. 8 and Table I, we can see our method is superior to IDM. As mentioned in Section II-C2, it is shown that the temporal constraint in region matching reduces the false positives and improves the performance as shown in ROC curves in Fig. 8.

IV. Conclusion

Motivated by the wave of water, we developed a novel method to depict an interaction among objects and detect

(a) (b) (c) (d)

Fig. 5. The detection result of car accident dataset. Column (a) and (c) are the sample frames in normal situation of each video. Column (b) and (d) are the detection results of proposed method when traffic accidents occur. *Best viewed in color.*

(a) (b)

Fig. 6. The accident localization of each algorithm (a) Intelligent Driver Model (b) Proposed Method. *Best viewed in color.*

(a) (b)

Fig. 7. False positive example. (a) Intellient Driver Model (b) Proposed method. *Best viewed in color.*

Method	Area Under ROC
Intelligent Driver Model [8]	0.6270
Proposed Method without temporal constraint	0.8722
Proposed Method	0.8950

TABLE I
THE PERFORMANCE COMPARISON USING AVERAGE AREA UNDER ROC

accident events in traffic videos. It has been shown the value and the shape of the proposed motion interaction field could well describe the key aspects of the interaction among moving objects. Unlike the existing methods, our method does not use learning algorithm and can be applied to specific applications such as detection and localization of the car accidents in traffic videos. For this purpose, our method shows outperforming performance compared to the existing methods. For the future work, an extended version of our method could be tried to general applications beyond car accidents.

ACKNOWLEDGMENT

This work has been supported by Brain Korea 21 Program for Leading Universities & Students, and the Ministry of Science, ICT and Future Planning in Republic of Korea.

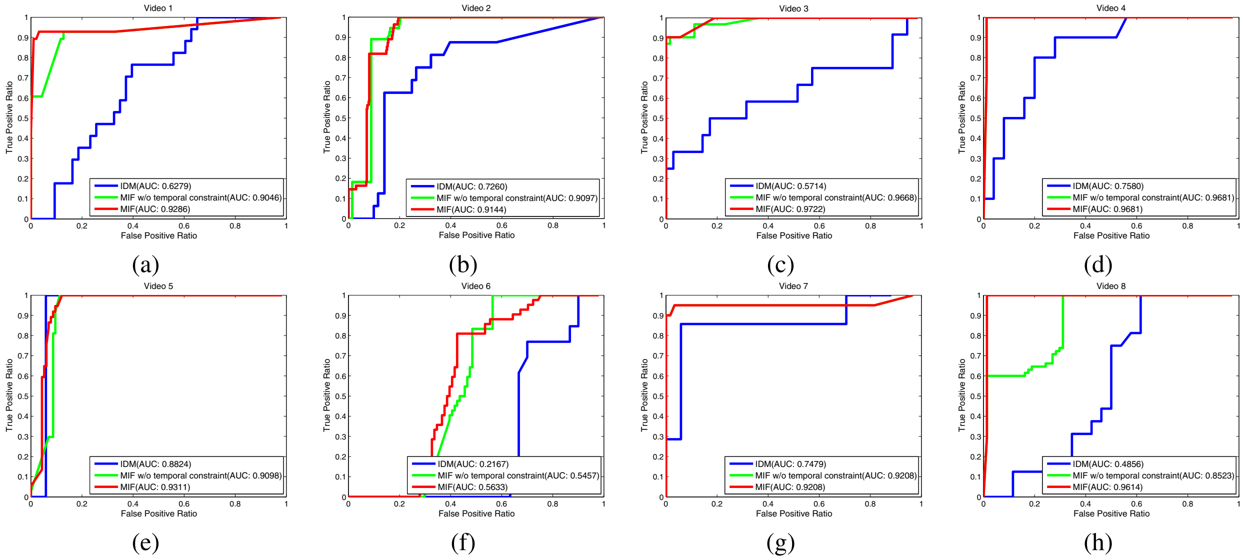


Fig. 8. The ROC curve result of each video. Blue plot is the result of IDM, Green plot is the result of proposed method without temporal constraint in Section II-C2 and Red plot is the result of proposed method. *Best viewed in color.*

REFERENCES

[1] J. Varadarajan, R. Emonet, and J. Odobez, "Bridging the Past, Present and Future: Modeling Scene Activities From Event Relationships and Global Rules," in *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2012, pp. 2096–2103.

[2] T. Hospedales, S. Gong, and T. Xiang, "A Markov Clustering Topic Model for mining behaviour in video," in *International Conference on Computer Vision (ICCV)*, 2009.

[3] W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, and S. Maybank, "A system for learning statistical motion patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1450–1464, 2006.

[4] S. Sadeky, A. Al-Hamadiy, B. Michaelisy, and U. Sayed, "Real-Time Automatic Traffic Accident Recognition Using HFG," in *International Conference on Pattern Recognition (ICPR)*. IEEE, 2010, pp. 3348–3351.

[5] Y.-K. Ki, "Accident Detection System using Image Processing and MDR," *International Journal of Computer Science and Network Security(IJCSNS)*, vol. 7, no. 3, pp. 35–39, 2007.

[6] D. Zeng, J. Xu, and G. Xu, "Data fusion for traffic incident detector using DS evidence theory with probabilistic SVMs," *Journal of Computers*, 2008.

[7] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Computer Vision and Pattern Recognition (CVPR)*, 2009.

[8] W. Sultani and J. Choi, "Abnormal traffic detection using intelligent driver model," in *International Conference on Pattern Recognition (ICPR)*, 2010.

[9] C. Liu, "Beyond pixels: exploring new representations and applications for motion analysis," Doctoral Thesis, Massachusetts Institute of Technology, May 2009.

[10] A. Bruhn, J. Weickert, and C. Schnorr, "Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 211–231, 2005.

[11] X. Cui, Q. Liu, M. Gao, and D. N. Metaxas, "Abnormal detection using interaction energy potentials," in *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2011, pp. 3161–3167.

[12] M. K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Info. Theory*, vol. IT-8, pp. 179–187, 1962.