Anomalous Trajectory Detection in Surveillance Systems Using Surrounding Information

Doan Trung Nghia, Sunwoong Kim, and Hyuk-Jae Lee

*Abstract*— Concurrently detected and annotated abnormal events can contribute significant impacts on surveillance systems. While considering the specific domain of pedestrian trajectories, there are two main contributions in this paper. First, as introduced in a great amount of work of the trajectory-based anomaly detection literature, only the information of the pedestrian paths such as direction and speed is considered. Being differed from previous work, this paper proposes a framework that deals with an additional type of trajectory-based anomalies. This abnormal events happen when a person enters the prohibited regions. Those restricted regions are constructed by an online learning algorithm that uses surrounding information including detected pedestrians and background scenes. Second, a simple data-boosting technique is introduced as an aid for the lack of training data, such problem particularly challenges all previous work owing to the significantly low frequency of abnormal events. This technique only requires normal trajectories and fundamental knowledge of scenes to surge the number of training data for both normal and abnormal trajectories. With the increased number of training data, the conventional abnormal trajectory classifier is able to achieve better prediction accuracy without falling in the overfiting problem caused by complex learning models. Finally, the proposed framework, which annotates tracks entering prohibited areas, and a conventional abnormal trajectory detector using the data-boosting technique are integrated to form the united detector. Such detector faces with different types of anomalous trajectories in a hierarchical order. The experimental results show that all proposed detectors are able to effectively detect anomalous trajectories in the test phase.

*Index Terms*—Pedestrian detection, pedestrian tracking, anomalous trajectory, superpixel classification, trajectory features, Neural Network (NN).

# INTRODUCTION

C

urrently, an increasing number of cameras used for surveillance encounters a new challenge in terms of data analysis due to the fact that storing the information captured by those cameras is not the only pivotal task. As many cameras are installed in public places, manual monitoring and validation of the human behaviour in those systems are often beyond the ability of a human observer. In reality, it is hard for a camera observer to rapidly react to anomalous trajectories that captured by a surveillance system because the observer has to manage a great amount of information from many cameras at the same time. Consequently, an automatic analysis of video-based data as a support tool for monitoring personnel is required. For instance, potentially abnormal trajectories which are detected and annotated by such systems can contribute a significant impacts on monitoring surveillance cameras as only suspicious cases are warned and time to react is obviously shortened. In addition, cameras installed inside vehicles can be a good field for such applications where computer vision solutions assist drivers in case of early warning for abnormal movements of pedestrians like crossing the road while traffic lights are still green.

The definition of abnormal trajectories is understood in two ways. First, if there exists one or more regions which are prohibited for the access of humans in a video frame, any trajectories move through such areas are defined as abnormal trajectories. Secondly, a track is also assumed as abnormal one if it is considerably different from the dominant trajectories in directional and/or speed aspects. For instance, the dominant trajectories are the set of tracks that are parallel to the main walking path while the rare trajectories are the set of tracks in which people cross the main walking path. To deal with those abnormal trajectories, two frameworks which make use of pedestrian and surrounding information are introduced in this paper.

The rest of this paper is organized as follows. Section II presents the related work in this literature whereas Section III explains the proposed frameworks in detail. Sections IV and V give experimental results and conclusions of this paper, respectively.

# Related work

Analysis anomalies in terms of object trajectory have been studied for decades with a great deal of works. Owing to the fact that unusual behaviours are rarely occurred and strongly depended on specific contexts, many approaches employ prior knowledge of scenes where moving directions of objects are already determined. However, there are other works that successfully overcome the existed problems by building adaptive learning systems which are able to be applied to various scenes. Extensive surveys of related research that investigate and make use of trajectory information for video surveillance systems are shown in [1]-[3].

Hu *et al.* in [4] is one of the first works that addresses the problem of detecting abnormal object trajectories by applying an unsupervised learning algorithm. A fast and accurate *fuzzy* K-means algorithm is the core part of the tracking algorithm proposed in this work. Besides, a framework that automatically learns the motion patterns by using the spatial and temporal information is introduced and each motion pattern is then represented as a chain of Gaussian distributions. This framework not only annotates the unusual events by analysing the statistical data of motion patterns, it is also able to predict behaviours for early warning.

Jung *et al.* in [5] proposed a framework that includes four-dimensional (4-D) histograms and trajectory clustering. At the initial stage, sample trajectories are grouped into main clusters based on Mixture of Gaussian (MoG) followed by a process of removing outliers. In each refined cluster, the position and velocity of each tracked object are arranged to form the 4-D histogram that represents the local of trajectories within each cluster. In the test step, the 4-D histogram having the same format as that built in training phase is created for each new trajectory. With the newly calculated histogram, the coherence of the test track is to be evaluated with those in the training step in order to detect the abnormal events.

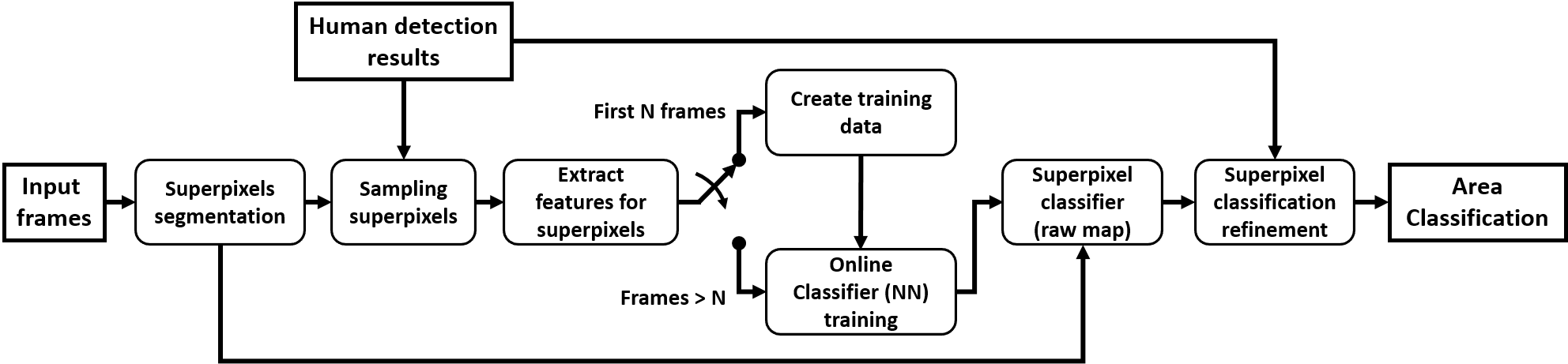


Fig. 1. The flow of the proposed framework to classify human-like and non-human-like areas

A method for online learning and sequential anomaly detection in trajectories is introduced by R. Laxhammar *et al.* in [6], in which the parameter-light algorithm called Sequential Hausdorff Nearest-Neighbor Conformal Anomaly Detector (SHNN-CAD) is investigated. This algorithm does not suffer from the difficulties of tuning of many parameters, over-fitting and poorly calibrated alarm rates during supervised online-learning process.

Having a slightly different approach while modelling the trajectories, H. Jeong *et al.* in [7] formulates a track regarding the temporal-spatial property of the overall path as opposed to the conventional similarity measures proposed in almost all other related papers. To represent a trajectory, position information is extracted by quantizing the location of trajectory sub-parts into cells where the size of the cell is 10×10 pixels. Every point in the original track is assigned to a word (a cell) and a complete track is represented by a vocabulary which is the combination of many words. Temporal model of a trajectory is then formulated using Hidden Markov Model (HMM) for detecting anomalous events.

To the best of our knowledge, for both clustering-based and non-clustering-based anomalous approaches, all related work in this literature only considers the trajectories with the perspectives of direction and/or speed while additional cues such as background information is not used effectively.

# Proposed Frameworks

The first framework tries to give a warning for a person who enters prohibited areas for a number of consecutive frames. The motivation idea of the first framework is that cameras used for surveillance systems often capture scenes where the regions indicated for human to appear have considerably different appearance than the remaining regions, which are less likely for human to walk through. Therefore, if those distinguished regions are correctly separated, annotations for human entering such restricted areas are useful helps for camera observers. Notice that to yield these kinds of abnormal event annotations, none of the directional properties of pedestrian trajectories is taken into account since the only required task is to compare the positions of the detected human with that of the restricted areas; therefore, fast annotations are possible to achieve. In the second framework, the directional and speed properties of trajectory are considered. As a traditional approach, a supervised learning model is used to train the classifier with the negative and positive trajectory samples. Notice that those samples are created by a data-boosting technique where only normal trajectories and basic prior knowledge of the scenes are involved. Furthermore, by assuming the average speed of human moving paths and its deviations determine the range of slow, normal and high speeds, the speed of a given track in test phase is to be compared with the defined speed thresholds in order to categorize that track into the groups of slow, normal, and fast moving tracks. Note that inputs of the proposed frameworks are the trajectories derived from human tracking algorithms, which are proposed in [8]-[9] where they use the human detection algorithms in [10]-[11]. Ultimately, a hierarchical abnormal trajectory detector that makes use of the two above frameworks is introduced as a new approach for real-life applications where it effectively deals with separate aspects of the anomalous trajectory-based events.

## Anomalous trajectory detection in prohibited areas

Fig. 1 illustrates the flow of the abnormal trajectory detector that tries to give annotations when a person enter the prohibited regions. Those restricted regions are either defined by an explicit manner such as hand labelling or a learning algorithm to automatically describe the restricted areas. In the proposed framework, the second approach is chosen since it can be adopted to many circumstances of real applications.

The main idea of the proposed framework is to make use of the existing human detection results in order to build a map which represents two kinds of areas. One of them indicates the regions in which human can appear, while the other points out the regions in which human are less likely to appear or the prohibited areas. To separate those regions, an image is initially segmented into sub-regions, particularly the choice of segmentation methods could affect strongly to the overall framework. Considering the simple grid segmentation, by which a given video frame is split into multiple sub-regions having the shapes of small squares. It is obvious that the boundaries of objects that exist in the image frame could be broken if a sparse grid segmentation is used, hence, the information contained in those sub-regions could end up having different appearance properties. On the other hand, using the dense grid-segmentation occupies more computational demands which is also not suitable for real time applications. Thanks to superpixels [12]-[13], such segmentation method splits the whole frame into smaller parts while preserving the boundaries among objects appear in the frame. Besides, all pixels inside a superpixel often contain similar appearance since they are grouped based on the small gradient differences. It turns out another advantage for using superpixel segmentation where spatial redundancy is often existed. Therefore, applying feature engineering to construct the important features for a superpixel could remarkably reduce the complexity of the learning model without affecting the classification accuracy. Human detection results such as human bounding boxes are used to indicate the labels of sampling superpixels by considering the relative location between a human bounding box and surrounding superpixels. The superpixels locating around the footage areas of a human bounding box are labelled as human-like area while some random superpixels locating distantly from the human bounding boxes are assumed as non-human-like area. Fig. 2 (a) and (b) give illustrations about the above superpixel sampling step, where the yellow rectangles are the human bounding boxes. The lighter superpixels surrounding the human bounding boxes are labelled as human-like area superpixels whereas the darker superpixel randomly located far from human bounding boxes are marked as non-human-like area superpixels. Once positive (human-like-area) and negative (non-human-like-area) superpixel samples are completely taken from the segmented frame, each superpixel is then gone through the features extraction step before an online learning algorithm is taken in to account. As location and appearance information of a superpixel are pivotal cues, the feature vector of the superpixel is defined as a set of *px*, *py*, *cR*, *cG*, *cB*, and *cGrey*. The parameters {*px, py*} are the average horizontal and vertical positions of all pixels inside a sample superpixel in the image frame correspondingly, which aims at describing the global location information of a superpixel. On the other hand, {*cR, cG, cB, cGrey*} define the appearance characteristics of a superpixel by exploiting the average values of 4 colour channels, Red, Green, Blue and Grey, of all pixels.

|  |
| --- |
| D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\rawClassification.PNG  (a) |
| D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\RefineSuperpixel.PNG  (b) |

Fig. 3. Human(grey) and non-human(black)-like area maps of the PETS09-S2-L1-VIEW-001 dataset. (a) raw map (b) refining process by using the moving window approach

|  |
| --- |
| D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\SP_Samples3.PNG  (a) |
| D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\SP_CurrentSence_Samples.PNG  (b) |

Fig. 2. Results in the superpixel segmentation step. (a) PETS09-S2-L1-VIEW-001 (b) PETS09-S2-L1-VIEW-004

At the first N-frames of the video sequence, all superpixels features and their labels are stored to make sure that the number of training data should be large enough prior to the training process. This preparation step is illustrated as the Create training data step in Fig. 1. In this step, it is necessary to guarantee that people are assumed to appear only in the allowed areas. When the number of frames exceeds the pre-determined value of N, the online NN classifier is trained by those prepared features and labels. The superpixel NN model consists of 2 hidden layers with the number of neurons are 10 and 6 for the first and the second hidden layers, respectively. Note that from the (N+1) frame, the processes of segmenting a new frame into superpixels and computing their features with another step that removes old superpixels feature vectors are required for the online training classifier. In this paper, the value of N is set to 200, which gives acceptably empirical results for the test datasets. Note that online classifier is able to deal with the changes in terms of illuminations of the objects such as light and weather conditions. Obviously, this aspect is a pros of the proposed framework when it is applied for real applications.

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |

Fig. 5. Examples of normalized trajectory generation of the TownCentre dataset. (a) a normal track (b) additional normal tracks (c) base lines of rare tracks (d) a normal trajectory (red color) and abnormal trajectories created from the normal one (other colors)

By using the already trained classifier, every segmented superpixel of the original frame is then categorized into two groups which are non-human-like-area superpixels and human-like-area superpixels. This very first classification results in a raw map where miss-classified superpixels can be existed. Fig. 3 shows the superpixel classification results in which human-like and non-human-like superpixels are represented by grey and black regions, respectively. The respective grey and black regions represent the non-prohibited and prohibited areas for human. Fig. 3 (a) shows a raw superpixel classification of a given frame where all superpixels in that frame are classified by the trained classifier one by one. As clearly seen from Fig. 3 (a), there exists wrong classified superpixels where some non-human-like superpixels are treated as the opposite kind. To handle that problem, a moving window approach is proposed, of which main purpose is to detect the wrong classification superpixels. The refinement step using the proposed window approach is illustrated in Fig. 3 (b). In this step, two different windows with the same centre recursively scan over the whole image, the content in the inner window is then compared with that of its outer window. It is obvious that the human like area cannot be completely covered by the prohibited area since a person cannot suddenly appear in any area of the image without his/her trajectories. It leads to the fact that such area is miss-classified and it is to be corrected as the non-human-like area by changing the label of all wrong detected human superpixel to non-human area superpixel. Although there exists a case that the refinment process makes a side effect when it also removes correct human-like area superpixels, the experimental results show that this negative aspect still dose not strongly affect the accuracy of the current framework.

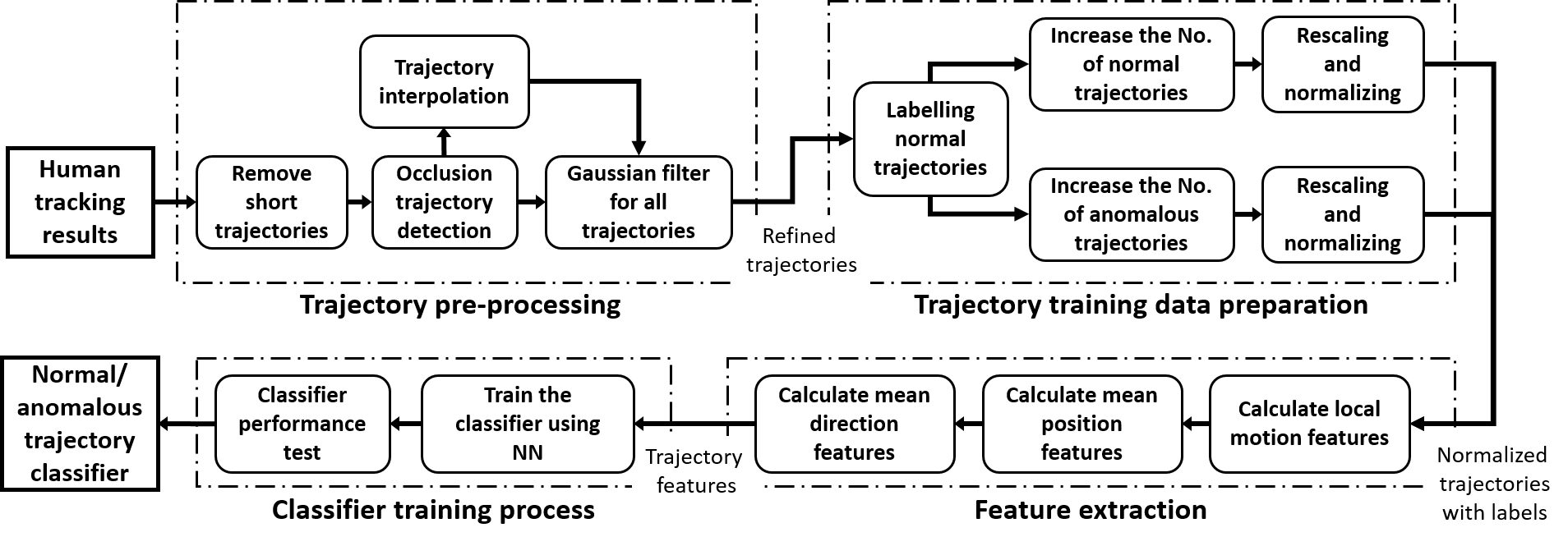


Fig. 4. The flow of the proposed framework to detect an anomalous pedestrian trajectory in which the direction is only considered

It turns out another problem when the online classifier is made use of, not all the superpixels surround human bounding boxes are the human-like superpixels. Therefore, only the “safe” superpixels are chosen to train the classifier. The safe superpixels are the one that belong to the humans appearing in non-prohibited areas, which is determined by the current classifier. Once the human-like area map is completed, the detection process turns out to be fairly easy. If the footage area of any person mostly contains the black area for consecutive frames, then this person is assumed to enter the prohibited area.

## Anomalous trajectory detection in terms of directional and speed

The second type of abnormal trajectory detection is to differentiate among dominants trajectories and the rare ones where the unusual trajectories contain significantly different expressions regarding their moving directions and speeds. The whole process of the proposed framework is shown in Fig. 4 where only the directional aspect is taken into account. Initially, given the raw human trajectories of a human tracking algorithm, those raw tracks enter the pre-processing step that eliminates error tracking data as well as refines tracks affected by noises and/or occlusions. All short trajectories are removed since their information is not sufficient for further steps, hence only the ones having the length greater than or equal to M-frames are kept. In this paper, the value of M is set to 50, which achieves good results. Occlusions often exist in crowded pedestrian scenes, which results in dis-continuous trajectories. Especially, trajectories with discrete expression lead to more difficulties in the steps that re-sample or extract features of trajectories. This negative impact of the occlusion effects is resolved by using both a linear and an n-tap interpolation techniques with weighted values used to combine the results of those two methods. The non-discrete tracks are then smoothed using a Gaussian filter to remove the noises caused by the tracking algorithm itself before entering the trajectory training data preparation step.

|  |
| --- |
| **FOR** i=1:maxTrackIndex  t=Tracks[i];  isInProhibitAreas =  **Abnormal\_Track\_Detector\_1**(t);  **If** isInProhibitAreas  Abnormal\_Track\_Notation(t);  **ELSE**  [isAbnDirection, isAbnSpeed] =  **Abnormal\_Track\_Detector\_2**(t);  **If** isAbnDirection  Abnormal\_Track\_Notation(t);  **END IF**  **If** isAbnSpeed  Abnormal\_Speed\_Notation(t);  **END IF**  **END IF**  **END FOR** |

Fig. 6. The pseudo code of the hierarchical abnormal trajectory detector. The Abnormal\_Track\_Detector\_1 is the detector trained by the first proposed framework introduced in Section III.A whereas the Abnormal\_Track\_Detector\_2 is that of the second framework explained in Section III.B

In the trajectory training data preparation step, positive (normal) and negative (abnormal) samples are created. The proposed framework is different from others in the point that only normal trajectories and basic prior knowledge of the given scenes are needed to build the training data for both abnormal and normal samples. The motivated idea is that in anomalous trajectory detection applications, unusual trajectories are occurred at significantly lower frequency in comparison with that of the normal ones. It leads to the fact that a greater training dataset for trajectories in general and rare trajectories in this particular issue have to be built to achieve high accuracy of the detection algorithm. First, based on the labelled normal tracks, a new set of normal tracks which are slightly rotated from the original ones is created. Fig. 5 (a) shows a normal trajectory; whereas Fig. 5 (b) depicts additional trajectories those are formed by rotating the track in Fig. 5 (a) using small rotate angles. Note that the rotate angles are constrained to completely preserve the general direction of the normal ones. Rescaling (up-scaling and down-scaling) the tracks plays such a pivotal role to ensure that the classifier is persistent with a wide variety of pedestrian speeds in real applications. Besides, instead of considering the complete path of a person, the smaller parts of the whole track are split and used to extract track features. The reason is that an abnormal track often contains a small part which is assumed to be abnormal whereas the remaining part of it is still normal; therefore it is more accurate to evaluate the track by considering the sub-parts of it only, not the complete one. Normalizing step is then performed to subsample the whole track into shorter tracks with the same length and the sub-tracks partly overlap each other. In addition, there is a little difference when the same measure is applied for abnormal tracks to rise the number of anomalous training data. In case of anomalous trajectories, the prior knowledge is required to create such samples for the training process. For instance, in the test sequence TownCentre, the dominant trajectories are the ones which pass alongside the pavements whereas the rare trajectories occur when people cross the main road. Then, the abnormal trajectory is created as follows. Firstly, the base lines that represent the directions of rare tracks are manually created as shown in Fig. 5 (c). The black lines represent the crossing movements alongside the roads. Second, for each base line of the rare tracks, additional abnormal tracks are created by rotating the normal track around that base line. Taking the normal track as a seed track for the creation of rare trajectories preserves the physical properties in terms of curve, fluctuation ratio, speed, etc. Fig. 5 (d) shows the abnormal tracks created from the normal one by using the above technique. At the end of this step, for each newly created anomalous track, several smaller parts are split from the complete track and all sub-tracks are normalized in the same manner as explained for the normal tracks.

|  |  |
| --- | --- |
| (a) | D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\SP_CurrentSence.PNG  (b) |
| (c) | D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\SP_CurrentSence_HMMap_refined.PNG  (d) |

Fig. 7. Results of the superpixel classification refinement step. (a) original frame of the PETS09-S2-L1-VIEW-001 dataset (b) original frame of the PETS09-S2-L1-VIEW-004 dataset (c) refined human-like area maps of (a) (d) refined human-like area maps of (b)

The next feature extraction step is applied for the normalized tracks that already contain either normal or abnormal labels. Owing to the fact that each track is represented by a group of consecutive positions; therefore, the global positions, the global directions as well as the local motions (between each pair of successive points) of a trajectory are elemental to represent important properties of that track. Specifically, a normalized track *t* with the length of *n* is the set of *n* successive points in a two-dimensional plane which is represented as:

 (1)

|  |  |
| --- | --- |
| D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\epr1.PNG  (a) | D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\epr2.PNG  (b) |
| D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\ProhibitedTrackNotation.PNG  (c) | D:\[2]DATA\[RESEARCH]\[CAPP]\[3]VideoMining\[1]Reports\Pictures\ProhibitedTrackNotation2.PNG  (d) |

Fig. 8. Prohibited trajectory annotations of the PETS09-S2-L1-VIEW-001 dataset. (a, b, c) two people with indices 3 and 4 are detected (d) one person with the index 4 is detected

A complete feature vector of *tn* comprises of five elements {*Dx*, *Dy*, *Px*, *Py*, *α*}, which are expressed as follows.

 (2)

In (2), *Dx* and *Dy* give the average values of the horizontal and vertical local motions between all consecutive positions in the normalized trajectory, respectively. On the other hand, *Px* and *Py* illustrate the respective global horizontal and vertical positions of the whole trajectory as written in (3).

 (3)

In (3), the parameters *Xmax* and *Ymax* represent the positions having the maximum values in horizontal and vertical directions of the frame, respectively. Ultimately, the last element of the trajectory feature vector is denoted as *α*, which expresses the average direction of the normalized track.

 (4)

Next, the trajectory feature vectors are transferred to the classifier training process step. In this step, a NN is trained by using all trajectory features as well as their labels followed by a classifier evaluation process using the test data. Note that the same NN configuration and training method as explained in Section III.A are used for the second framework. The only difference is the number of neurons in the input layer, which is 6 for the trajectory feature data.

When speeds are taken into consideration, pedestrians who move with high and slow speeds are also assumed as containing some abnormal movements. At the training process, after converting centroid positions of pedestrians in image plane to world plane by using a given camera model of the scene, all human trajectories are represented in world coordinate system. A trajectory speed is defined as fast, low, and medium as presented in [5]. In case the current trajectory is classified as low or high speed, an annotation is shown on top of the corresponding human bounding box to illustrate such anomalous trajectories.

## Hierarchical abnormal trajectory detector

The two proposed frameworks deal with different types of trajectory-based anomalous events. Regarding the main characteristics of each introduced framework, it is possible to combine them together in order to construct a united and consistent detector that effectively utilizes the advantages of both frameworks. Specifically, the first type of abnormal trajectory detection is fairly simple where one of the most complicated computational processes is segmenting a frame into superpixel which is required only once for the first background scene. The remaining online training classifier used to build the human-like and non-human-like map as depicted in Fig. 1 is not taken into account as the frequency of every new frame. Instead this process is scheduled at every 50 frames regarding the current configuration of the implementation program. Furthermore, the classifier training process in general can be either pipelined or paralleled with the detection process. This assumption holds for both the two proposed frameworks, which is also a key aspect for designing a real-time system in surveillance related applications. Using the human-like-area map to evaluate whether a trajectory is anomalous or not only requires simple comparisons. Besides, it is clear that there is no need to consider a track in the prohibited area is normal or abnormal in terms of directional and speed aspects or not. Therefore, a trajectory is firstly classified in this manner to accelerate the computing time of the whole system since all processes needed for the second framework are early terminated if the examining track lays inside the restricted areas. It turns out that the second proposed framework becomes involved when the position of a trajectory is detected as locating inside the human-like areas. This exclusive detector is able to detect rare trajectories by considering the directional and speed perspectives,

Fig. 6 illustrates the main idea of the combined detector. Given a human-like-area map, the hierarchical abnormal-trajectory detector examines all trajectory t existed in the current frame, the Abnormal\_Track\_Detector\_1 is brought out to tackle the simple cases where a track’s path moves through the prohibited area. If the current track t is found in such restricted areas, a notation for the person whose track is t is turned on and the process is finished for the current track t. Otherwise, the second classifier is picked to evaluate the normality of t in terms of speed and direction that is denoted as Abnormal\_Track\_Detector\_2 in Fig. 6. Finally, with respects to each unusual property of t, the notifications for abnormalities in directional aspect and/or speed aspect is noticed for the corresponding person. If t is a normal track, the two detectors are tested for t before the whole process is applied to the next track in the current frame.

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |

Fig. 10. The results of creating sample rare trajectories and superpixel classification map. (a) a background scene (b) normal trajectories (c) base lines of rare trajectories (d) the corresponding human-like and non-human-like areas of the background scene in (a)

# Experimental Results

All proposed frameworks are implemented in Matlab version R2015b in a Windows operating system. Two datasets, PETS09-S2-L1-VIEW-001 and PETS09-S2-L1-VIEW-004 [24], are used for the first proposed framework in which a person enters the restricted area. On the other hand, TownCentre [25] dataset is evaluated for the second proposed framework where rare tracks have different directions and moving speeds as compared to the dominant tracks. Another dataset named SNUCafe1 is recorded to evaluate the effectiveness of the hierarchical detector as explained in Section III.C.

Fig. 7 shows the experimental results of the first proposed framework. Fig. 7 (a) and (b) illustrate the original input frames whereas Fig. 7 (c) and (d) depict the refined human-like area map which are the output of the online NN classifier followed by a refinement process where the proposed moving window approach presented in Section III. A is involved. In Fig. 7 (c) and (d), the roads in which almost all people of the datasets appear correspond to the grey part of the human-like area map. In contrast, the grass areas are classified as black parts where humans are not supposed to appear. Note that human-like area map is the result of an online learning method, hence it is adaptively changed according to illumination conditions.

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| (e) | (f) |

Figure 9. Rare trajectories detection in terms of speed and direction of the TownCentre dataset. (a, b) human crossing the road with normal speed but anomalous directions (c, d) normal moving direction but with anomalous speed (e, f) anomalous trajectories in both speed and directional aspects

The annotations for persons who enter restricted areas are based on the human-like area maps as shown in Fig. 8. Particularly, if the track of that person lays in the black areas during the determined number of frames then this person is assumed to enter the prohibited areas. The notifications for such trajectories are denoted by the large red bounding boxes of the humans with indices 3 and 4. When the person containing the index 3 moves out of the grass area (prohibited area) the notification is turned off as shown in Fig. 8 (d).

In case of the second proposed framework in which the classifier trained by a NN is evaluated at the last step depicted in Fig. 4, the results show that 98.2% of the number of the trajectories in test data is correctly classified. In case of labelled normal trajectories, the accuracy of the trained classifier is 98.8% where that number for the labelled anomalous tracks is still significant high at 96.6%. Fig. 9 shows results in which the second proposed framework is applied to the TownCentre test sequence. If a person has the abnormal trajectory in case of direction, a large red bounding box is used to notify such trajectory. On the contrary, if a person moves with either low or high speed, the speed classification sign (“Fast” or “Slow”) is placed on top of his/her bounding box. Fig. 9 (a) and (b) capture the abnormal trajectories when people are crossing the road and all people in those two frames are moving with normal speeds. In contrast, Fig. 9 (c) and (d) only depict the anomalous movements with unusual speeds but normal moving directions. As shown in those Fig., two persons are both moving with high speed as compared the average speed calculated in the training phase. Finally, Fig. 9 (e) and (f) depict the anomalous trajectories where their speeds and directions are both abnormal. The man in Fig. 9 (e) stays still in the same position, which results in a low moving speed and an anomalous moving direction. The woman whose track crosses the road in Fig. 9 (f) has an unusual moving direction since the dominant tracks are often spread parallel to the road. Besides, the speed of this woman is also considered as high speed with the notification assigned on top of her bounding box.

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |

Fig. 10. The results of creating sample rare trajectories and superpixel classification map. (a) a background scene (b) normal trajectories (c) base lines of rare trajectories (d) the corresponding human-like and non-human-like areas of the background scene in (a)

Fig. 10 and Fig. 11 depict the experimental results of the hierarchical detector which is built based on the two proposed frameworks. Fig. 10 (a) shows a background scene whereas the majority of pedestrian trajectories, which contains the directions that are relatively parallel to the main moving path, is depicted in Fig. 10 (b). Base on that assumption, rare track base lines are created in the directions that are relatively perpendicular to the main moving path as shown in Fig. 10 (c). Note that the base lines do not need to exactly form with the main moving path with a specific angle since a significant number of rare tracks will be created by rotating a normal track within a range of rotation angles alongside the base lines. Fig. 10 (d) illustrates a human-like-area map constructed during the online learning process of the first framework. The main moving path in Fig. 10 (a) coressponds to the grey area of Fig. 10 (d). By contrast, other areas in Fig. 10 (a) where humans are not detected such as the soil area locating in the bottom left corner of the background scene result in the black areas in Fig. 10 (d).

Fig. 11 gives the notifications for the detected abnormal trajectories of the hierarchical detector. If a trajectory appears in the prohibitted area, then a large red bounding box is used to represent the warning sign. In this situation, owing to the fact that the man is standing inside the soil area which corresponds to the non-human-like areas based on the created map as depicted in Fig. 10 (d); therefore the notations are given as shown in Fig. 11 (a) and (b). On the other hand, if a trajectory locates in the allowed moving areas and it is classified as an unusual track in terms of direction, a large bounding box in orange is used to notice such events. Those examples are described in Figs 11 (c) and (d) when a man crosses the main moving path where others tracks behave as normal ones. Finally, regarding the trajectory speeds, slow or fast notification is placed on top of the human bounding box to signal for those anomalous events as shown in Fig. 11 (e) and (f) in which a man is moving with low speed whereas the other is running, respectively.

# Conclusions

Two frameworks that detect anomalous trajectories are introduced in this paper where the notifications are employed to indicate the tracks for humans entering restricted areas and containing unusual directions and speeds. In case of a superpixel in the first framework, the information of the global positions and the average values of R, G, B and grey colour channels are used to form the features. In the later framework, a complete track given by the pedestrian tracking algorithms is initially refined in the pre-process step where noise and/or occlusion effects are mitigated. Especially, a proposed data-boosting technique is carried out to significantly increase the number of training data for both anomalous and normal trajectories. Such method takes the refined tracks of the previous block then creates more training data using fundamental knowledge of the background scenes. The newly generated tracks are then re-scaled to deal with speed variance issues in real applications before being normalized into shorter parts. These normalized tracks are represented in terms of track features such that a complete feature vector includes the information of global position, local motion and direction. Besides, two trajectory speed thresholds are calculated from that of the training data. The speed of a test track is compared with those thresholds to indicate whether this track contains an anomalous (fast or slow) moving speed or not. The first and the second frameworks make use of NN-based classifiers. Finally, the experimental results are conducted in PETS09-S2-L1-VIEW-001, PETS09-S2-L1-VIEW-004, TownCentre and a self-recorded SNUCafe1 datasets to evaluate the effectiveness of the proposed abnormal trajectory detectors. It is observed that the empirical results achieve high accuracy when almost all of the anomalous trajectories are captured and given notifications as explained in Section IV.

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| (e) | (f) |

Figure 11. Rare trajectory detection results of the hierarchical detector conducted in the SNUCafe1 dataset. (a, b) anomalous trajectories detected in prohibited areas (c, d) anomalous trajectories in terms of directions (e, f) anomalous trajectories in terms of speed

References

1. Morris, Brendan Tran, and Mohan Manubhai Trivedi. "A survey of vision-based trajectory learning and analysis for surveillance." *Circuits and Systems for Video Technology, IEEE Transactions on* 18.8 (2008): 1114-1127.
2. Remagnino, Paolo, et al. "Novel concepts and challenges for the next generation of video surveillance systems." *Machine Vision and Applications* 18.3 (2007): 135-137.
3. Shah, Mubarak. "Understanding human behaviour from motion imagery." *Machine Vision and Applications* 14.4 (2003): 210-214.
4. Hu, Weiming, et al. "A system for learning statistical motion patterns." Pattern Analysis and Machine Intelligence, *IEEE Transactions* *on* 28.9 (2006): 1450-1464.
5. Jung, Cláudio Rosito, Luciano Hennemann, and Soraia Raupp Musse. "Event detection using trajectory clustering and 4-D histograms." *Circuits and Systems for Video Technology, IEEE Transactions* *on* 18.11 (2008): 1565-1575.
6. Laxhammar, Rikard, and Goran Falkman. "Online learning and sequential anomaly detection in trajectories." *Pattern Analysis and Machine Intelligence, IEEE Transactions* *on* 36.6 (2014): 1158-1173.
7. Jeong, Hawook, Hyung Jin Chang, and Jin Young Choi. "Modeling of moving object trajectory by spatio-temporal learning for abnormal behaviour detection." *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference* *on*. IEEE, 2011.
8. Führ, Gustavo, and Cláudio Rosito Jung. "Combining patch matching and detection for robust pedestrian tracking in monocular calibrated cameras." *Pattern Recognition Letters* 39 (2014): 11-20.
9. Milan, Anton, et al. "Joint tracking and segmentation of multiple targets." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.
10. Dollár, Piotr, et al. "Fast feature pyramids for object detection." *Pattern Analysis and Machine Intelligence, IEEE Transactions* *on* 36.8 (2014): 1532-1545.
11. Felzenszwalb, Pedro F., et al. "Object detection with discriminatively trained part-based models." *Pattern Analysis and Machine Intelligence, IEEE Transactions* *on* 32.9 (2010): 1627-1645.
12. Ren, Xiaofeng, and Jitendra Malik. "Learning a classification model for segmentation." *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference* *on*. IEEE, 2003.
13. Mori, Greg, et al. "Recovering human body configurations: Combining segmentation and recognition." *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference* *on.* Vol. 2. IEEE, 2004.
14. Dollár, Piotr, et al. "Integral channel features." (2009): 91-1.
15. Dollár, Piotr, Serge Belongie, and Pietro Perona. "The Fastest Pedestrian Detector in the West." *BMVC*. Vol. 2. No. 3. 2010.
16. Dollár, Piotr, Ron Appel, and Wolf Kienzle. "Crosstalk cascades for frame-rate pedestrian detection." *Computer Vision–ECCV 2012*. Springer Berlin Heidelberg, 2012. 645-659.
17. Andriyenko, Anton, Konrad Schindler, and Stefan Roth. "Discrete-continuous optimization for multi-target tracking."Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.
18. Milan, Anton, Konrad Schindler, and Stefan Roth. "Detection-and trajectory-level exclusion in multiple object tracking." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2013.
19. Bae, G. T., S. Y. Kwak, and H. R. Byun. "Motion pattern analysis using partial trajectories for abnormal movement detection in crowded scenes." *Electronics Letters* 49.3 (2013): 186-187.
20. Hou, A-Lin, et al. "Abnormal behaviour recognition based on trajectory feature and regional optical flow." *Image and Graphics (ICIG), 2013 Seventh International Conference on.* IEEE, 2013.
21. Piciarelli, Claudio, Christian Micheloni, and Gian Luca Foresti. "Trajectory-based anomalous event detection." *Circuits and Systems for Video Technology, IEEE Transactions on* 18.11 (2008): 1544-1554.
22. Li, Ce, et al. "Abnormal behaviour detection via sparse reconstruction analysis of trajectory." *Image and Graphics (ICIG), 2011 Sixth International Conference on.* IEEE, 2011.
23. Jiang, Fan, Ying Wu, and Aggelos K. Katsaggelos. "Abnormal event detection based on trajectory clustering by 2-depth greedy search." *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on.* IEEE, 2008.
24. Ferryman, J., and Anna Ellis. "PETS2010: Dataset and challenge."Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on. IEEE, 2010.
25. Benfold, Ben, and Ian Reid. "Stable multi-target tracking in real-time surveillance video." Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011.
26. Abu-Mostafa, Yaser S., Malik Magdon-Ismail, and Hsuan-Tien Lin. Learning from data. Berlin, Germany: AMLBook, 2012.