An MRF based Abnormal Event Detection Approach using Motion and Appearance Features

Hajananth Nallaivarothayan, Clinton Fookes, Simon Denman, Sridha Sridharan
Image and Video Laboratory, Queensland University of Technology
GPO Box 2434, Brisbane 4001, Australia
{h.nallaivarothayan, c.fookes, s.denman, s.sridharan}@qut.edu.au

Abstract

Abnormal event detection has attracted a lot of attention in the computer vision research community during recent years due to the increased focus on automated surveillance systems to improve security in public places. Due to the scarcity of training data and the definition of an abnormality being dependent on context, abnormal event detection is generally formulated as a data-driven approach where activities are modeled in an unsupervised fashion during the training phase. In this work, we use a Gaussian mixture model (GMM) to cluster the activities during the training phase, and propose a Gaussian mixture model based Markov random field (GMM-MRF) to estimate the likelihood scores of new videos in the testing phase. Furthermore, we propose two new features: optical acceleration, and the histogram of optical flow gradients; to detect the presence of any abnormal objects and speed violations in the scene. We show that our proposed method outperforms other state of the art abnormal event detection algorithms on publicly available UCSD dataset.

1. Introduction

Abnormal event detection has attracted a lot of attention in the computer vision community in recent years due to the increased focus on automated surveillance systems to improve security in public places. As the technology evolves, embedded devices such as CCTV cameras become more affordable, enabling CCTV cameras to be more widely deployed. However, it is seldom that the CCTV footage is monitored in real time due to the scarcity and cost of human resources. This motivates the need for an automated abnormal event detection framework using computer vision technologies, to detect abnormal events in real time.

An abnormal event has no consistent definition, as it varies according to the context. In general, it is defined as an event which stands out from the normal behavior within a particular context. In this work, we consider the context of pedestrian walkways, where the normal behavior is people walking. Events which involve speed violations, the presence of abnormal objects and trespassing are considered to be abnormal. Due to occlusions in crowded scenes, bottom up approaches which involve object detection and tracking are unsuitable. Thus we adopt a top down approach, where low level features are extracted at the pixel level and encoded in to non-overlapping spatio-temporal cuboids.

In this work we use Gaussian mixture models to cluster the training data, which contains the normal data pertaining to the scene and context. Low level features extracted at a pixel level are summed for each and every non-overlapping spatio-temporal cuboid to create a feature vector to represent behavior. During the testing phase, to account for the spatial causality of the activities, we propose using a Gaussian mixture model based Markov random field to calculate the likelihoods. Performance of the proposed method is compared with that of the GMM alone. Furthermore, we propose two new motion and appearance based features: optical flow acceleration, and the histogram of optical flow gradients; to detect anomalies relating to speed and the presence of abnormal objects. These features are combined with raw optical flow to enhance the detection performance.

The remainder of this paper is structured as follows: Section 2 summarises related work in this field; Section 3 describes the GMM-MRF model proposed in our work; Section 4 describes the proposed motion and appearance features; Section 5 presents experimental results on the publicly available USCD database [7]; and Section 6 presents conclusions and directions for future work.

2. Related work

Abnormal event detection is a key problem within crowd surveillance, and it has been an active research topic for several years. The definition of an abnormality differs with the context, and the events which are considered as abnormal typically occur very rarely compared to normal events.

The above reasons lead to a scarcity of examples of abnormal events, making it difficult to train event specific models. Due to this scarcity of training data, the problem of abnormal event detection is typically formulated as a novelty detection task, where the system is trained on normal data in an unsupervised manner and events which do not fit the learned 'normal' model are detected.

Abnormal event detection can be divided into two major categories: top-down approaches, and bottom-up approaches. Top-down approaches [10], otherwise called trajectory analysis, involve object detection and tracking for feature extraction. This approach can be effective only in sparsely crowded scenarios, and will degrade in performance in densely crowded areas due to the challenges involved in tracking in crowded scenes.

In bottom up approaches low level features are extracted which represent the underlying scene characteristics and crowd behavior. As this approach involves feature extraction at a pixel level instead of at the object level, this can work well in densely crowded environments and will be robust to extensive clutter and dynamic occlusions. Due to this robustness, the majority of recent works adopt this approach for feature extraction. The feature extraction process plays a key role in the abnormal event detection framework, as the features contain the information which describes the events present in the scene. Because of the huge variety of contexts, each with unique characteristics, features have to be sufficiently descriptive to capture the unique properties of the normal behavior, such that the any outliers can be distinguished from the normal behavior model effectively. Furthermore, the features must also be invariant and robust to variations such as brightness changes, occlusions, presence of clutter, etc. Features such as optical flow [10] and spatio-temporal gradients which contain motion rich information were proposed to detect anomalies related to speed. Features which contain shape, size and texture related information have been used to detect the presence of any abnormal objects [11]. Other texture [13] and shape related features [5] which also contain motion information, have also been proposed and can be used to detect both speed related anomalies, and the presence of abnormal objects.

Ryan et al [13] used optical flow vectors in their work to model the motion information, and they proposed the textures of optical flow feature to capture the smoothness of the optical flow field across a region enabling the detection of objects such as bicycles and vans. This approach was extended in [9] by incorporating additional texture features extracted through Gabor wavelets. Reddy et al [11] used foreground texture features extracted through Gabor filters, a size feature which is calculated based on the number of foreground pixels, and optical flow vectors to represent the motion characteristics, to detect the presence of abnormal objects. The histogram of optical flow, which contains both motion and shape information, was used by Kim et al [5] and Adam at al [1]. Kratz et al [6] used the distribution of spatio-temporal gradients to model the local spatio-temporal motion patterns.

Mahadevan [7] proposed a detection framework which uses mixture of dynamic textures to represent both motion and appearance features. Andrade et al [2] calculated a dense optical flow field for each pixel, followed by dimensionality reduction using PCA. Though the above feature representations contain motion and appearance information, they dont extract the information specifically belonging to human motion (i.e. leg movements), where optical flow varies temporally and spatially. Extracting this information alongside other motion and appearance related information will enhance the performance of detection systems.

Extracted features which contain information about the scene have to be modeled to learn the patterns related to the normal activities. Popular learning models include GMM [13], LDA [10], Support Vector Machine (SVM) [10], Mixture of Probabilistic PCA [5], and Gaussian kernel based density estimation [11]. There have also been models proposed to account for the temporal causality of the activities present in the scene. HMM and Multi observation HMM [15] are two examples of such models. Coupled HMMs [6] and Semi-2D HMMs [9] are proposed to account for both the spatial and temporal causality of the activities, whereas Kim et al [5] proposed an MRF based labeling of abnormal and normal events to ensure the spatial consistency of the labeling. Most of these state of the art models do not completely model both the spatial and temporal causality, while the MRF based approach proposed by Kim et al [5] has high data requirements due to the use of location specific model parameters.

3. Proposed GMM-MRF model

People inside a crowd tend to show similar behavior, regardless of their self behavior which varies with every individual. The degree of behavior similarity between the individuals in the crowd rises with the density of the crowd. Therefore there is a necessity to take this similarity into account during the classification of abnormal activities. We propose a Gaussian mixture model based Markov random field (GMM-MRF) to detect the abnormal activities during the classification process in our detection framework. A two dimensional Markov random field is created by connecting the adjacent cuboids in vertical and horizontal directions in a four neighbor fashion, as shown in Figure 1b.

Initially, during the training process we divide the video clips which contain the normal activities in to non-overlapping spatio-temporal cuboids. Here, the spatio-temporal cuboids are created around a centre pixel with fixed dimensions. Centre pixels are chosen so that the created spatio-temporal cuboids cover the entire scene, and

avoiding redundant overlapping. Low level features are extracted for every single pixel and are then summed to extract the feature vector of each spatio-temporal cuboid.

Feature vectors are created for each and every spatio-temporal cuboid, and are then modeled using a GMM with the number of mixtures in the model determined through the Bayesian information criterion [14]. After determining the parameters of the GMM, the parameters of the MRF framework are estimated based on the co-occurrence statistics of the spatio-temporal cuboids extracted from the training data. The co-occurrence statistics contain information about the co-existence of each and every state of the cuboids in adjacent spatial locations. Here the states refer to the mixtures obtained during the GMM training process, and every cuboid in the plane is considered to be a node of the particular 2D-MRF network.

We estimate the co-occurrence probabilities for our two-dimensional MRF model in four spatial directions. We only consider the immediate neighbors of the cuboid, i.e. those locations adjacent in the up, down, left and right directions. After calculating the individual co-occurrence probabilities for each instance in the space-time domain, we estimate the global co-occurrence probabilities for each of the four directions by summing up the statistics obtained for the individual instances. A schematic diagram of our MRF model is shown in Figure 1. The co-occurrence probability of two states, q and r, at horizontally adjacent spatial locations, (x, y) and $(x + 1, y)$, is given by:

$$P(q, r)_{(x,y)->(x+1,y)} = P(q|O_{(x,y)})P(r|O_{(x+1,y)}), \quad (1)$$

where the posterior probabilities of the states q and r at locations (x, y) and $(x + 1, y)$ given the respective observations ($O_{(x,y)}$ and $O_{(x+1,y)}$), are determined as follows,

$$P(q|O_{(x,y)}) = \frac{\omega_q N(O_{(x,y)}|\mu_q, \Sigma q)}{\sum_s \omega_s N(O_{(x,y)}|\mu_s, \Sigma s)}, \quad (2)$$

$$P(r|O_{(x+1,y)}) = \frac{\omega_r N(O_{(x+1,y)}|\mu_r, \Sigma r)}{\sum_s \omega_s N(O_{(x+1,y)}|\mu_s, \Sigma s)}, \quad (3)$$

where $N(.)$ denotes the normal distribution and s is the number of mixtures in the GMM.

Hence, the global co-occurrence probability in the horizontal-right direction is given by

$$G_{right}(q, r) = \sum_{\forall x, \forall y, \forall t} P(q, r)_{(x,y)->(x+1,y)}, \quad (4)$$

where t is the total number of cuboids created from the training video in the temporal direction.

Due to the symmetrical nature of this technique, the global co-occurrence probability in the right direction is equal to that of the left direction. Similarly those of the up and down directions are equivalent to each other, and are given by the following equation,

$$G_{down}(q, r) = \sum_{\forall x, \forall y, \forall t} P(q, r)_{(x,y)->(x,y+1)}. \quad (5)$$

During the testing phase of the detection framework, sub-sequences of video frames are bundled using a sliding window, and are then used to create a two dimensional MRF network of spatio-temporal cuboids. The likelihood of each and every node is calculated using the the loopy-belief propagation (sum-product) algorithm.

Based on the minimum likelihood of the node among all the nodes in the MRF net, a frame is classified as either normal or abnormal. Here, the length of the bundle is the length of the spatio-temporal cuboid in the temporal dimension. When a new frame arrives for processing, it will be added to the bundle while the first frame will be removed to keep the length of the bundle constant.

This two dimensional Markov Random Field is defined by two functions: node evidence, and pair wise potential. Node evidence measures the likelihood of the observation, $O_{(x,y)}$, of a particular node at location (x, y) being generated by the state q, and is given by the equation

$$ne_{(x,y)}(q) = P(O_{(x,y)}|\mu_q, \sigma_q) = N(O_{(x,y)}|\mu_q, \sigma_q). \quad (6)$$

Pairwise potential on the other hand measures the co-existence of two states, q and r, in adjacent locations in the four specified directions. We use the global co-occurrence statistics which are obtained from the training data, and the likelihoods of the new observations belonging to the adjacent locations generated by the states, q and r. Pairwise potential in the rightward direction for the states q and r of the node at (x, y) is given by the Equation 7, where G (calculated by Equation 4) is the global co-occurrence statistic for the co-existence of a state q with a second state, r at the node in the rightward direction; and $N(.)$ denotes the normal distribution. Similarly, pairwise potential in the leftward, upward and downward directions are given by Equations 8, 9 and 10 respectively.

The joint probability function of all the observations and node variables of the Markov random field is given by the following equation,

$$p(Z, O) = \prod_{i,j} pp_{(z_i->z_j)} \prod_i ne_{z_i}, \quad (11)$$

where i, j represent the node index of the MRF net.

Then, the marginal distribution of a particular node Z_i at location (x, y) is obtained by summing up all the configurations of the node variables Z, except Z_i,

$$m(z_i) = \sum_{\forall z_1, \forall z_2, ..., \forall z_n \sim z_i} p(Z, O). \quad (12)$$

Finally the likelihood of the node Z_i at location (x, y) is given by the equation,

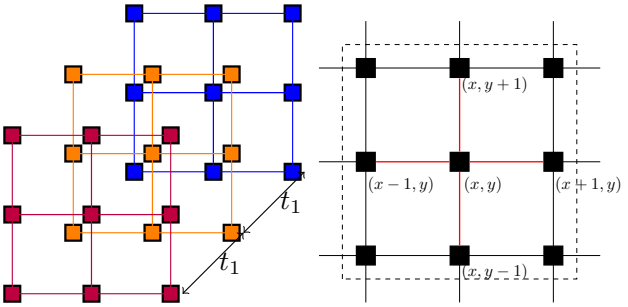
$$l(z_i) = \sum_{\forall z_i} m(z_i). \quad (13)$$

$$pp_{(x,y)->(x+1,y)}(q,r) = N(O_{(x,y)}|\mu_q,\sigma_q)P(O_{(x+1,y)}|\mu_r,\sigma_r)G_{right}(q,r) \quad (7)$$

$$pp_{(x,y)->(x-1,y)}(q,r) = N(O_{(x,y)}|\mu_q,\sigma_q)P(O_{(x-1,y)}|\mu_r,\sigma_r)G_{left}(q,r) \quad (8)$$

$$pp_{(x,y)->(x,y-1)}(q,r) = N(O_{(x,y)}|\mu_q,\sigma_q)P(O_{(x,y-1)}|\mu_r,\sigma_r)G_{up}(q,r) \quad (9)$$

$$pp_{(x,y)->(x,y+1)}(q,r) = N(O_{(x,y)}|\mu_q,\sigma_q)P(O_{(x,y+1)}|\mu_r,\sigma_r)G_{down}(q,r) \quad (10)$$



(a) multiple 2D MRF nets along the time axis

(b) Neighbors of the proposed 2D MRF Network

Figure 1: MRF Structure

4. Proposed Motion and Appearance Features

Features have to be descriptive enough to extract the important information about the normal behavior pertaining to the context in consideration. Here in this work we consider the context of pedestrian walkways, and anything which stands out from the normal behavior, such as high speed objects and the presence of abnormal objects, are considered to be anomalies. State of the art feature vectors do not completely describe the human motion behavior, which has its own unique characteristics such as variation in the motion features across the human body due to the non-linear nature and the repetitiveness of the limb movements. Here in this work we propose two features named optical acceleration and the histogram of optical flow gradients, to extract information about the temporal and spatial variations in the optical flow respectively.

4.1. Optical Acceleration

Optical flow features have been used as a very strong cue to detect speed related anomalies. In addition to using traditional optical flow features, we propose a new feature called optical flow acceleration, to extract information about the variation of optical flow temporally. It can be noticed that the optical flow across the human body varies over the time due to the nature of the limb movements, particularly due to the legs. This results in acceleration across the human body varying significantly in direction, but with a small magnitude. Furthermore, objects like vehicles and bicycles tend to have high acceleration due to the thrust applied to them, but the direction of their acceleration is predominately uniform due to their rigid motion.

To model both the acceleration direction and the magnitude, we create a two dimensional feature vector which contains both the horizontal and the vertical components of the acceleration. We first calculate the pixel level optical flow vectors for each and every frame using the method proposed by Black and Anandan [3], and then we create an optical flow magnitude image for every frame where the magnitude of the optical flow at position (x, y) is given by,

$$M(x,y) = \sqrt{U_x^2 + U_y^2}, \quad (14)$$

where U_x and U_y are the horizontal and vertical components of the optical flow vector respectively. Acceleration at each pixel is calculated from the time derivative of the optical flow magnitude image. We choose to use the same algorithm which has been used for the optical flow calculation instead of using temporal differencing of the magnitude images due to the robustness of Black and Anandan's [3] algorithm in calculating the flow vectors.

Finally, the acceleration information across a spatio-temporal patch, P, is directly incorporated using a summation of the acceleration vectors at the pixel level to create a two dimensional acceleration feature vector, $[A_x, A_y]$

4.2. Histogram of Optical Flow Gradients

Due to the non-rigid nature of the human motion caused by the limb movements, optical flow varies across the human body, thus creating a need to extract information about the spatial variations in the optical flow. At the same time, abnormal objects such as bicycles and vans have laminar and smooth flow across their surface. So additional information about the flow variation across an object's surface can be indicative of appearance variations, in addition to providing motion information. This can aid in the detection of abnormal objects such as slow moving bicycles which don't present significant motion variation to classify them as anomalies.

We propose a similar feature to the Motion Boundary Histogram used in Human Detection [4] to differentiate between humans and other abnormal objects in the scene. We term this feature, the histogram of optical flow gradients. We treat both the horizontal and vertical optical flow components as separate images and calculate the gradients of each image using Sobel operators. Four bins have been used each for the horizontal and vertical flow components, and a

soft binning based approach is used to calculate the weight assigned to two adjacent bins based on the distance to the bin centers. Furthermore, these soft votes are weighted based on the magnitude of the gradient.

Gradient components of the horizontal flow at a pixel location p are given by $U_{xx}(p) = \frac{d}{dx}(U_x)$ and $U_{xy}(p) = \frac{d}{dy}(U_x)$, where U_{xx} and U_{xy} are the horizontal and vertical gradient components of the horizontal flow.

Similarly, gradient components of the vertical flow are given by $U_{yx}(p) = \frac{d}{dx}(U_y)$ and $U_{yy}(p) = \frac{d}{dy}(U_y)$, where U_{yx} and U_{yy} are the horizontal and vertical gradient components of the vertical flow.

The orientation of the horizontal optical flow gradient at pixel location p is given by $\theta_h(p) = \arctan\frac{U_{xy}(p)}{U_{xx}(p)}$, and the orientation of the vertical flow, $\theta_v(p)$ can be calculated similarly.

After calculating the orientation at each pixel, weights are added to bins based on the calculated orientation. We don't consider the sign of the orientation like the techniques used in human detection, and any negative angle has π radians added to it to convert it to it's positive counterpart. We use four bins centered at the angles: $\frac{\pi}{8}$, $\frac{3\pi}{8}$, $\frac{5\pi}{8}$ and $\frac{7\pi}{8}$. Weights are added to the bins in a soft manner, such that for the gradient orientation at location p, $\theta_h(p)$, which lies between two adjacent bin centers θ_n and θ_{n+1} (i.e $\theta_n < \theta_h(p) < \theta_{n+1}$), the soft weight for the nth bin is given by

$$h^h_n(p) = \frac{\theta_{n+1} - \theta_h(p)}{\theta_{n+1} - \theta_n}\sqrt{U^2_{xx}(p) + U^2_{xy}(p)}, \qquad (15)$$

and the weight for the $n+1$th bin is given by

$$h^h_{n+1}(p) = \frac{\theta_h(p) - \theta_n}{\theta_{n+1} - \theta_n}\sqrt{U^2_{xx}(p) + U^2_{xy}(p)}. \qquad (16)$$

All other bins will be unchanged for the particular gradient. Histogram weights for the gradients of the vertical component of the optical flow are calculated in a similar manner.

Finally, the histogram information across a spatio-temporal patch, P, is directly incorporated using a summation of histogram values at the pixel level, the n^{th} bin of the horizontal optical flow gradient and vertical optical flow gradient are given by $H^h_n = \sum_{p\in P} h^h_n(p)$ and $H^v_n = \sum_{p\in P} h^v_n(p)$ respectively.

The two histograms for a patch are concatenated into a single eight dimensional feature vector,

$$[H^h_{\frac{\pi}{8}}, H^h_{\frac{3\pi}{8}}, H^h_{\frac{5\pi}{8}}, H^h_{\frac{7\pi}{8}}, H^v_{\frac{\pi}{8}}, H^v_{\frac{3\pi}{8}}, H^v_{\frac{5\pi}{8}}, H^v_{\frac{7\pi}{8}}].$$

5. Experimental results

We test our proposed approach by combining the proposed features (see Section 4) with the optical flow features $[U_x\ , U_y]$ from [13]. The final feature vector contains two features describing the horizontal and vertical optical flow fields, two optical acceleration features (horizontal and vertical directions) and eight histogram of optical flow gradient components, four components each for vertical and horizontal directions. Thus we have a twelve dimensional feature vector for each spatio-temporal cuboid. Feature vectors are normalized for perspective using the technique proposed in [8].

The above feature vector is tested with both the proposed GMM-MRF classifier and a GMM classifier on the publicly available UCSD dataset [7]. Dimensions of the spatio-temporal cuboids are chosen to be 11x11 spatially and 13 temporally for the ped2 dataset, while we use 15x15 spatially and 20 temporally for the ped1 dataset. Dimensions reported here are based on the best results among the several configurations tested during the experiments.

The UCSD video dataset contains bi-directional pedestrian traffic from two camera view points. Several video sequences (each of 200 frames duration) which contain normal pedestrian movements are used for training. The testing video sequences contain abnormalities, such as the presence of abnormal objects, anomalous pedestrian motions and spatial abnormalities, and are annotated with frame-level ground truth.

Table 1 shows the equal error rate and area under the curve obtained for both classifiers for the ped1 and ped2 dataset, and examples of detected anomalies are shown in Figure 2. It can be seen that for the ped1 dataset, using the MRF with the GMM gives better performance compared to using the GMM alone, whereas for the ped2 dataset, using the MRF with the GMM doesn't give the same improvement compared to using the GMM alone. This lack of performance in ped2 can be attributed to the lack of training data for the ped2 dataset, which causes incompleteness in estimating the co-occurrence statistics of the MRF network.

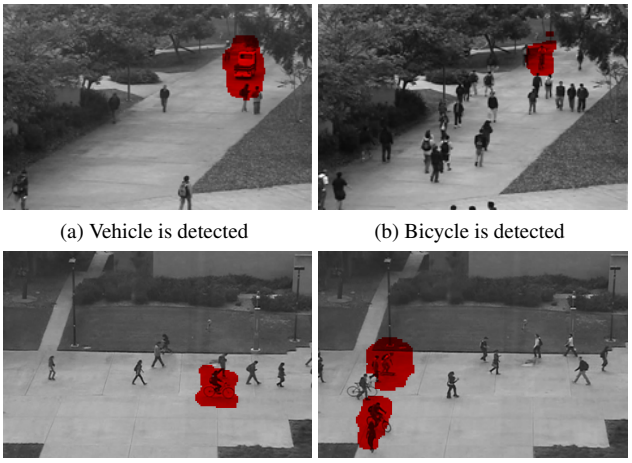
Furthermore, ranks of both the GMM and the GMM-MRF classifier scores are calculated and the addition of these ranks is used to create a combined classifier. The equal error rate and the area under the curve are reported for this classifier, and shown in Table 1. ROC curves for our proposed approach for both the ped1 and peds2 datasets are shown in Figure 3. This proposed approach outperforms all the state of the art approaches for the ped2 dataset, and narrowly beats the best performing state of the art approach [12] for the ped1 dataset.

6. Conclusion and Future Work

We have proposed a GMM based Markov random field approach for abnormal event detection, which accounts for the spatial causality of motion patterns in the scene, enhancing the performance of the GMM based abnormal event detection framework. Furthermore, we propose two new fea-

System	EER (Ped1)	AUC (Ped1)	EER (Ped2)	AUC (Ped2)
GMM	16.50%	0.903	5.44%	0.979
GMM-MRF	16.00%	0.891	7.80%	0.966
Combined	**14.90%**	**0.908**	**4.89%**	**0.979**
Roshtkhari et al [12]	15.0%	-	-	-
Ryan et al [13]	23.1%	0.838	12.7%	0.939
Reddy et al [11]	22.5%	-	20%	-
Mahadevan et al [7]	25%	-	25%	-

Table 1: Comparison of the proposed classifiers with state of the art on the UCSD datasets [7]. Equal error rate (EER) and area under the curve (AUC) are reported.



(a) Vehicle is detected (b) Bicycle is detected

(c) Slow moving bicycle is detected (d) Bicycle and skateboarder are detected.

Figure 2: Detected anomalies

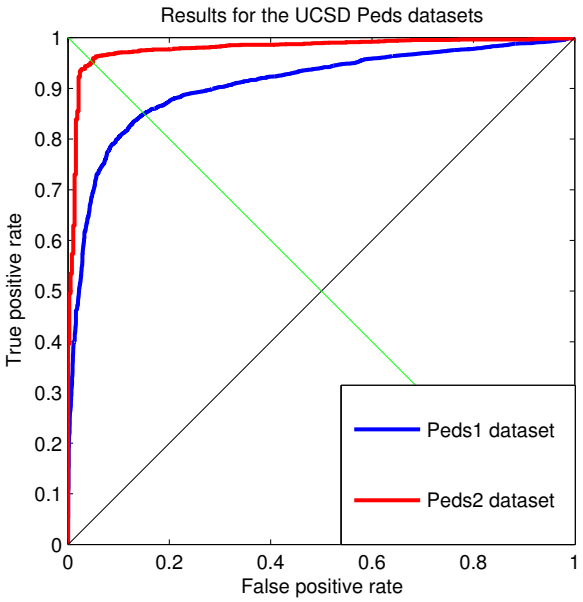


Figure 3: ROC curves of Ped1 and Ped2 datasets obtained by our proposed method.

tures, optical acceleration and histogram of optical flow gradients, to extract information about the temporal and spatial variations of the optical flow respectively. Both the traditional GMM and proposed GMM-MRF classifiers are combined by adding their ranks to create a new combined classifier, which clearly outperforms all state of the art approaches for the ped2 dataset with an EER of 4.89 %, while also slightly outperforming the best state of the art approach for the ped1 dataset [12] with an EER of 14.90 %. Future work will focus extending this evaluation to other databases, and experimenting with other features and learning models to detect anomalies related to different contexts.

References

[1] A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz. Robust real-time unusual event detection using multiple fixed-location monitors. In *IEEE Trans. Pattern Anal. Mach. Intell*, volume 30, pages 555–560, Mar 2008.

[2] E. Andrade, S. Blunsden, and R. Fisher. Modelling crowd scenes for event detection. In *ICPR*, volume 1, page 4, 2006.

[3] M. J. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. In *Comput. Vis. Image Underst*, 1996.

[4] N. Dalal, B. Triggs, and C. Schmid. Human detection using oriented histograms of flow and appearance. In *ECCV 2006*.

[5] J. Kim and K. Grauman. Observe locally, infer globally: A space-time mrf for detecting abnormal activities with incremental updates. In *CVPR*, pages 2921–2928, 2009.

[6] L. Kratz and K. Nishino. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In *CVPR*, pages 1446–1453, 2009.

[7] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos. Anomaly detection in crowded scenes. In *CVPR*, 2010.

[8] H. Nallaivarothayan, D. Ryan, S. Denman, S. Sridharan, and C. Fookes. An evaluation of different features and learning models for anomalous event detection. In *DICTA 2013*.

[9] H. Nallaivarothayan, D. Ryan, S. Denman, S. Sridharan, and C. Fookes. Anomalous event detection using a semi-two dimensional hidden markov model. In *DICTA*, pages 1 – 7, 2012.

[10] O. P. Popoola and K. Wang. Video-based abnormal human behavior recognitiona review. In *IEEE Trans. on System, Man, and Cybernetics, Part C*, 2011.

[11] V. Reddy, C. Sanderson, and B. C. Lovell. Improved anomaly detection in crowded scenes via cell-based analysis of foreground speed, size and texture. In *CVPRW*, 2011.

[12] M. J. Roshtkhari and M. D. Levine. Online dominant and anomalous behavior detection in videos. In *CVPR 2013*.

[13] D. Ryan, S. Denman, C. Fookes, and S. Sridharan. Textures of optical flow for real-time anomaly detection in crowds. In *AVSS*, pages 230–235., 2011.

[14] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6:461464, 1978.

[15] T. Xiang and S. Gong. Incremental and adaptive abnormal behaviour detection . In *CVIU,*, volume 111, pages 59–73,, 2008.