IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

# An Efficient Learning with Sparse Regularized Kernel for One-Class Classification

**S. Bhuvanesh[1], Arun Mathew[2], P.M. Ashfaq Ahmed[3], A. Muhammed Rafi[4]**

Assistant Professor, Department of Computer Science and Engineering, Dhaanish Ahmed Institute of Technology, Coimbatore, India[1]

Student, Department of Computer Science and Engineering, Dhaanish Ahmed Institute of Technology, Coimbatore, India[2,3,4]

**Abstract**: Speedy abnormal event detection meets the growing demand to process an enormous number of surveillance videos. Although real-time processing is a key criterion to a practically employable system given continuously captured videos, most sparse code methods cannot be performed fast enough. A novel anomaly detection framework with transferred deep Convolutional Neural Network (CNN) is propose. The presents an online learning with sparse regularized kernel based one-class Extreme Learning Machine (ELM) classifier and is referred as “sparse-OC-ELM”. The baseline kernel hyperplane model considers whole data in a single chunk with regularized ELM approach for offline learning in case of One-Class Classification (OCC). The regularized kernel ELM based online learning and consistency-based model selection has been employed to select learning algorithm parameters. The online RK-OC-ELM has been evaluated on standard benchmark datasets as well as on artificial datasets and the results are compared with existing state-of-the art one class classifiers. The proposed method achieves high detection rates on benchmark datasets at a speed of 140-150 frames per second on average when computing on an ordinary desktop PC using MATLAB. The experimental results evaluate the proposed method on two publicly available video surveillance datasets, showing competitive performance with respect to state of the art approaches.

**Keywords**: Abnormal Event Detection, Outlier Detection, Video Data Stream, Sparse Learning, Dynamic Detection.

## I. INTRODUCTION

A fundamental challenge in intelligent video surveillance is to automatically detect abnormal events in long video streams. This problem has attracted considerable attentions from both academia and industry in recent years. Video anomaly detection is also important as it is related to other interesting topics in computer vision, such as dominant behaviour detection, visual saliency and interestingness prediction. A typical approach to tackle the anomaly detection task is to learn a model which describes normal activities in the video scene and then discovers unusual events by examining patterns which distinctly diverge from the model. However, the complexity of scenes and the deceptive nature of abnormal behaviours make anomaly detection still a very challenging task.

Outliers often arise due to human carelessness, faults in systems, natural deviation in dataset, fraud etc. However, it is important to differentiate between the applications of outlier detection e.g. if there is a clerical error in the data of a, the entry clerk should be notified and the data should be corrected. However, in industry when there are faults in a machine and the machine is damaged or in a safety critical system like an intrusion monitoring system or a fraud detection system, an alarm must be sounded to notify the system administrators about the problem. There are three approaches to the problem of outlier detection.

Most outlier detection techniques treat objects with K attributes as points in $\Re$K space and these techniques can be divided into three main categories. The first approach is distance-based methods, which distinguish potential outliers from others based on the number of objects in the neighborhood. Distribution-based approach deals with statistical methods that are based on the probabilistic data model. A probabilistic model can be either a priori given or automatically constructed using given data. If the object does not suit the probabilistic model, it is considered to be an outlier. Third, density-based approach detects local outliers based on the local density of an object’s neighborhood. These methods use different density estimation strategy. A low local density on the observation is an indication of a possible outlier.

Density-based methods have been developed for finding outliers in a spatial data. These methods can be grouped into two categories called multi-dimensional metric space-based methods and graph-based methods. In the first category, the definition of spatial neighborhood is based on Euclidean distance, while in graph-based spatial outlier detections the definition is based on graph connectivity. Whereas distribution-based methods consider just the statistical distribution of attribute values, ignoring the spatial relationships among items, density-based approach consider both attribute values and spatial relationship. Among previous works, several anomaly detection approaches are based on analysing individual moving objects in the scene. Tracking is usually an initial step for this class of methods. By using

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

accurate tracking algorithms, trajectory extraction can be carried out to further perform trajectory clustering analysis or design representative features to model typical activities and subsequently discover anomalies. In trajectories which are spatially close and have similar motion patterns are identified and used for detecting unusual events. The a “shape activity” model to describe moving objects and detect anomalies. However, as tracking performance significantly degrades in the presence of several occluded targets, tracking-based methods are not suitable for analysing complex and crowded scenes. In Distance-based methods outlier is defined as an object that is at least dmin distance away from k percentage of objects in the dataset. The problem is then finding appropriate dmin and k such that outliers would be correctly detected with a small number of false detections.

## II. Related Work

Credit card fraud falls broadly into two categories: behavioural fraud and application fraud. Application fraud occurs when individuals obtain new credit cards from issuing companies using false personal information and then spend as much as possible in a short space of time. However, most credit card fraud is behavioural and occurs when details of legitimate cards have been obtained fraudulently and sales are made on a 'Cardholder Not Present' basis.

Intrusion detection corresponds to a suite of techniques that are used to identify attacks against computers and network infrastructures. Anomaly detection is a key element of intrusion detection in which perturbations of normal behaviour suggest the presence of intentionally or unintentionally induced attacks, faults, defects, etc. This paper focuses on a detailed comparative study of several anomaly detection schemes for identifying different network intrusions. Several existing supervised and unsupervised anomaly detection schemes and their variations are evaluated on the DARPA 1998 data set of network connections as well as on real network data using existing standard evaluation techniques as well as using several specific metrics that are appropriate when detecting attacks that involve a large number of connections. During a clinical trial of a new treatment, a large number of variables are measured to monitor the safety of the treatment. It is important to detect outlying observations which may indicate that something abnormal is happening. To do this effectively, techniques are needed for finding multivariate outliers.

Outliers, or commonly referred to as exceptional cases, exist in many real-world databases. Detection of such outliers is important for many applications and has attracted much attention from the data mining research community recently. However, most existing methods are designed for mining outliers from a single dataset without considering the class labels of data objects. Anomaly detection is an important problem that has been researched within diverse research areas and application domains. Many anomaly detection techniques have been specifically developed for certain application domains, while others are more generic. This survey tries to provide a structured and comprehensive overview of the research on anomaly detection. We have grouped existing techniques into different categories based on the underlying approach adopted by each technique. For each category we have identified key assumptions, which are used by the techniques to differentiate between normal and anomalous behaviour. finding outliers (exceptions) in large, multidimensional datasets. The identification of outliers can lead to the discovery of truly unexpected knowledge in areas such as electronic commerce, credit card fraud, and even the analysis of performance statistics of professional athletes. Existing methods that we have seen for finding outliers in large datasets can only deal efficiently with two dimensions/attributes of a dataset. Here, we study the notion of DB- (Distance- Based) outliers. While we provide formal and empirical evidence showing the usefulness of DB-outliers. In addition to developing relatively straightforward solutions to finding such outliers based on the classical nested-loop join and index join algorithms, we develop a highly efficient partition-based algorithm for mining outliers. This algorithm first partitions the input data set into disjoint subsets, and then prunes entire partitions as soon as it is determined that they cannot contain outliers. This results in substantial savings in computation.

This degree is called the local outlier factor (LOF) of an object. It is local in that the degree depends on how isolated the object is with respect to the surrounding neighborhood. We give a detailed formal analysis showing that LOF enjoys many desirable properties. Using real world datasets, we demonstrate that LOF can be used to find outliers which appear to be meaningful but can otherwise not be identified with existing approaches. Finally, a careful performance evaluation of our algorithm confirms we show that our approach of finding local outliers can be practical. Outlier detection has recently become an important problem in many industrial and financial applications. In this paper, a novel feature bagging approach for detecting outliers in very large, high dimensional and noisy databases is proposed. It combines results from multiple outlier detection algorithms that are applied using different set of features. Every outlier detection algorithm uses a small subset of features that are randomly selected from the original feature set. As a result, each outlier detector identifies different outliers, and thus assigns to all data records outlier scores that correspond to their probability of being outliers.

The existing method for evaluating outlier-ness, which we call the Local Correlation Integral (LOCI). As with the best previous methods, LOCI is highly effective for detecting outliers and groups of outliers (a.k.a. micro-clusters). In addition, it offers the following advantages and novelties: (a) It provides an automatic, data-dictated cut-off to determine whether a point is an outlier in contrast, previous methods force users to pick cut-offs, without any hints as to what cut-off value is best for a given dataset. (b) It can provide a LOCI plot for each point; this plot summarizes a wealth of information about the data in the vicinity of the point, determining clusters, micro-clusters, their diameters

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

and their inter-cluster distances. The existing approach utilizes the Alternating Direction Method of Multipliers (ADMM) to recover simultaneously the sparse representations and the outlier's components for the entire collection. This approach provides a unified solution both for jointly sparse and independently sparse data vectors. We demonstrate the usefulness of the proposed approach for irregular heartbeats detection in Electrocardiogram (ECG) as well as for specular reflectance and shadows removal from natural images.

The K-SVD algorithm is a highly effective method of training over complete dictionaries for sparse signal representation. In this report we discuss an efficient implementation of this algorithm, which both accelerates it and reduces its memory consumption. The Incremental Coding Length (ICL) of a feature is a measure of its entropy gain. Given a dictionary, the ICL computation does not involve any parameter, is computationally efficient and has been used for saliency detection in images with impressive results. sparse reconstruction cost (SRC) over the normal dictionary to measure the normalness of the testing sample. By introducing the prior weight of each basis during sparse reconstruction, the proposed SRC is more robust compared to other outlier detection criteria.

## III. PROPOSED APPROACH

The proposes a novel pseudo data generation-based automatic parameter selection method, which is named minimal spanning tree (MST)-GEN, for sparse kernel-based OCELM. MST-GEN makes two contributions. First of all, with information embedded in n-round MST, MST-GEN generates a controllable number of high-quality pseudo outliers at proper locations by edge pattern detection (EPD) and a novel repelling process.

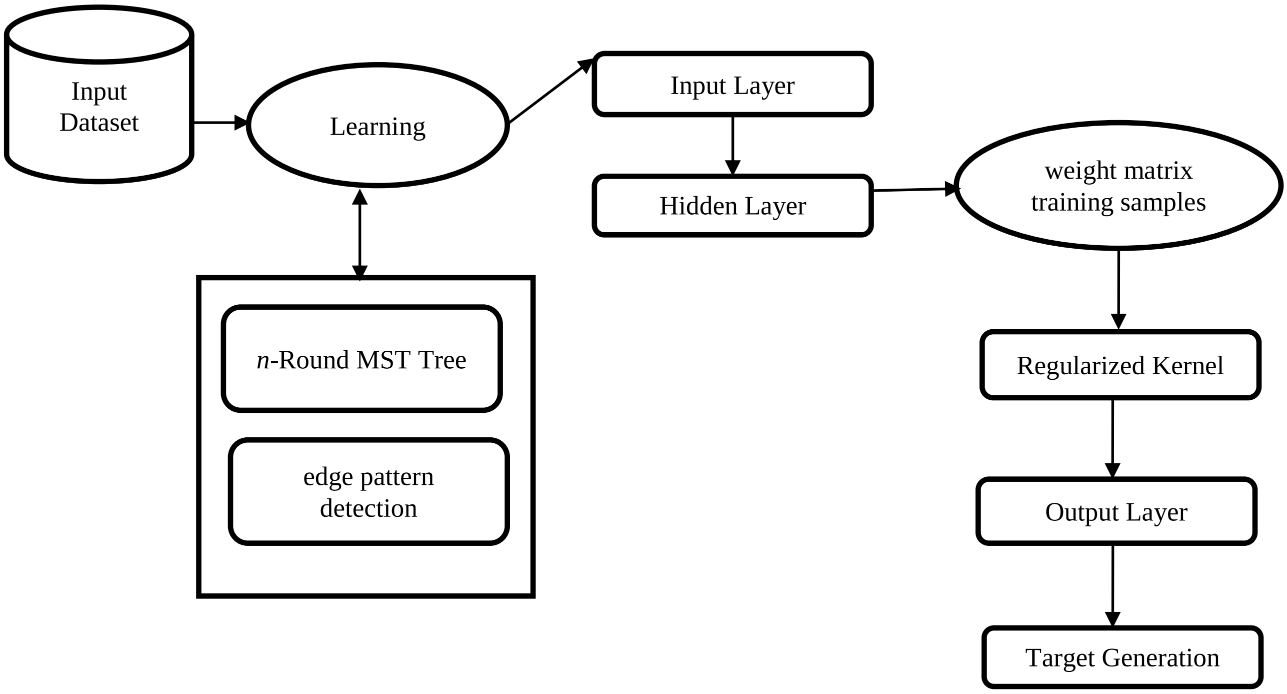


Fig. 1 Proposed architecture diagram

The proposed pseudo outlier generation can readily address the following two challenging problems that previous outlier generation methods cannot handle: where the generated outliers should be and how many outliers to be generated. Second, MST-GEN for the first time introduces a novel way to efficiently generate pseudo target set by n-round MST for OCC model validation.

A. one-class extreme learning machine

Numerous research attempts have been made on parameters election of OCC due to its vital importance and indispensability. OCELM is a simple variant of original ELM. For a ELM, weights of connections between the input layer and the hidden layer are randomly generated and do not need to be tuned in subsequent training, while the output weights between the hidden layer and the output layer are determined by analytically solving a least square optimization problem instead of iterative back-propagation.

$$\beta = H^T \left(\frac{I}{C} + HH^T\right)^{-1} T \quad (1)$$

where C, T, and I are the regularization coefficient, target output, and identity matrix, respectively. The prediction of a new sample x is given by.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

B. N-Round MST

The key of OCC is to model the data of the target set, which inspires us to introduce n-round MST as the foundation of the proposed approach. If data points in the target set X are viewed as vertexes of a complete undirected graph G, the MST T = $f_{MST}$(V, E) of graph G can serveas an efficient nonparametric graph representation of X [$f_{MST}$(V, E) returns the MST of a graph with vertex set V and edge set E]. The pointed out, T provides information on the target set in terms of two aspects. First, it reveals the underlying data structure and distribution of the target set X. Second, each edge in T's edge set ET reflects how target data can be linearly transformed into their neighbouring data through the edges. Thus, with edge set ET, MST T can be viewed as an enhanced representation of the target set X.

The n-round MST algorithm builds one MST at each iteration with those unselected edges after previous iterations. The function Remove (Vcur, Ecur) removes any isolated vertex in current vertex set Vcur with current edge set Ecur. n-round MST contains more edge information than a single MST. For example, as to the banana dataset, two-round MST and three-round MST contain more useful edges along the banana boundary than the single MST, while our experiments also verify that n-round MST typically yields better performance than single MST. As to the value of n, experiments show that 1 < n ⩽ 3 yields the best performance. However, a larger n does not lead to further performance improvement, as n-round MST will gradually approach the original complete graph when n is very large, which erases data structure and distribution information.

C. Sparse Regularized Kernel

The two objectives contradict each other in a sense. Reducing *K* could increase reconstruction errors. It is not optimal to fix *K* as well, as content may vary among videos. This problem is addressed in our system with a maximum representation strategy.

$$\forall_j \in (1,\dots,n\}, t_j = \sum_{i=1}^{K} \{\gamma_j^i \|x_j - s_i\beta_j^i\|_2^2 - \lambda\} \le 0,$$

$$s.t \sum_{i=1}^{K} \gamma_j^i = 1, \gamma_j^i = \{0,1\} \quad (2)$$

It automatically finds K while not wildly increasing the reconstruction error t. In fact, error t for each training feature is upper bounded in our method. obtain a set of combinations with a small K by setting a reconstruction error upper bound *λ* uniformly for all elements in *S*. If the reconstruction error for each feature is smaller than *λ*, the coding result is with good quality. So, we update function (2).

D. Edge Pattern Detection

We define "edge patterns "as those target data lying at the boundary of the target set. Our first step is to locate those edge patterns. Proposed an EPD method by local data statistics and geometry, which is efficient and easy to implement. Specifically, in the *ith* pass, given the leftover training data *Xc* ⊆ *X* that cannot be represented by previous combinations *{***S**1, . . ., **S***i*−1*}*, we compute **S***i* to bound most data in *Xc*. Our objective function becomes

$$\boldsymbol{min}_{s_i,\gamma,\beta} \sum_{j\in\Omega_c}^{\boldsymbol{K}} \{\gamma_j^i \|x_j - s_i\beta_j^i\|_2^2 - \lambda\} \quad (3)$$

$$s.t \sum_{i=1}^{K} \gamma_j^i = 1, \gamma_j^i = \{0,1\} \quad (4)$$

where Ω*c* is the index set for *Xc*. It is easy to prove that this cost satisfies condition (3) and the resulting **S***i* can represent most data. Specifically, if **x***j* –**S***i**β**i j* 22 −*λ* ≥ 0, setting *γi j* = 0 yields a smaller value compared to setting *γi j* = 1. Contrarily, *γi j* should be 1 if **x***j* –**S***i**β**i* −*λ* <0, complying with condition (3).

In each pass *i*, we solve the function in Eq. (4) by dividing it into two steps to iteratively update *{***S***i s*, ***β****}* and ***γ*** using the following procedure.

E. Quadratic Function Update

*{***S***i s*, ***β****}* With fixed ***γ***, Eq. (4) becomes a quadratic function

$$L(\beta, S_i) = \sum_{j\in\Omega_c} \{\gamma_j^i \|x_j - s_i\beta_j^i\|_2^2 (5)$$

Following the traditional procedure, we optimize ***β*** while fixing **S***i* for all *γi j* = 0 and then optimize **S***i* using block coordinate descent. These two steps alternate. The closed-form solution for ***β*** is

$$\beta_j^i = (S_i^T S_i)^{-1} S_i^T X_j \quad (6)$$

$S_i^T$ finds its solution as

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

$$S_i = \prod [s_i - \delta_t \Delta_{s_i} L(\beta, s_i] \quad (7)$$

where $\delta t$ is set to $1E{-}4$ and $\prod$ denotes projecting the basis to a unit column. Block-coordinate descent can converge to a global optimum due to its convexity. Therefore, the total energy for $L(\boldsymbol{\beta}, \mathbf{S}i)$ decreases in each iteration, guaranteeing convergence.

$\gamma$ with the $\{\mathbf{S}i, \boldsymbol{\beta}\}$ output, for each $\mathbf{x}j$, the objective function becomes

$$\boldsymbol{min}_{\gamma_j^i} \{\gamma_j^i \left\| x_j - s_i \beta_j^i \right\|_2^2 - \lambda \gamma_j^i \quad (8)$$
$$s.t\ \gamma_j^i = 0 \ or\ 1$$

$\gamma_j^i$ has a closed-form solution

$$\gamma_j^i = \begin{cases} 1 \quad if \quad \left\| x_j - s_i \beta_j^i \right\|_2^2 < \lambda \\ \quad 0 \quad otherwise \end{cases} \quad (9)$$

The proposed algorithm is controlled by $\lambda$, the upper bound of reconstruction errors. Reducing it could lead to a larger $K$. Our approach is expressive because all training normal event patterns are represented with controllable reconstruction errors under condition.

## IV. EXPERIMENTAL RESULTS

In this method, size of $\mathbf{S}i \in \mathrm{R}p\times s$ controls the sparsity level. We experimentally set $s = 0.1\times p$ where $p$ is the data dimension. $\lambda$ in Eq. (4) is the error upper bound, set to 0.04 in experiments. Given the input video, we resize each frame to 3 scales with 20×20, 30×40, and 120×160 pixels respectively and uniformly partition each layer to a set of non-overlapping 10 × 10 patches, leading to 208 sub-regions for each frame in total. For each frame, we compute an abnormal indicator $V$ by summing the number of cubes in each scale with weights. It is defined as $V = n\ i{=}1\ 2n{-}ivi$, where $vi$ is the number of abnormal cubes in scale $i$. The top scale is with index 1 while the bottom one is with $n$. All experiments are conducted using MATLAB.

For both the cases, true positive rate and false positive rate are calculated as follows:

$$TPR = \frac{\text{number of true positive frames}}{\text{number of positive frames}} \quad (10)$$
$$FPR = \frac{\text{number of false positive frames}}{\text{number of negative frames}} \quad (11$$

TPR and FPR is calculated for different threshold values. Then, ROC curve is drawn as the TPR versus FPR. Finally, the performance is summarized using the equal error rate (EER) which is the ratio of misclassified frames at which FPR is equal to 1-TPR in the ROC curve, for both frame-level and pixel-level criteria. A low EER value indicates a better performance.

TABLE I COMPARE OUTLIER SCORE IN THE INPUT VIDEO

| **Algorithm** | **Video Frame** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| MST-GEN | 0.04 | 0.05 | 0.09 | 0.14 | 0.08 | 0.05 | 0.07 | 0.04 | 0.1 |
| Sparse-OC-ELM | 0.05 | 0.07 | 0.1 | 0.18 | 0.1 | 0.07 | 0.09 | 0.05 | 0.15 |

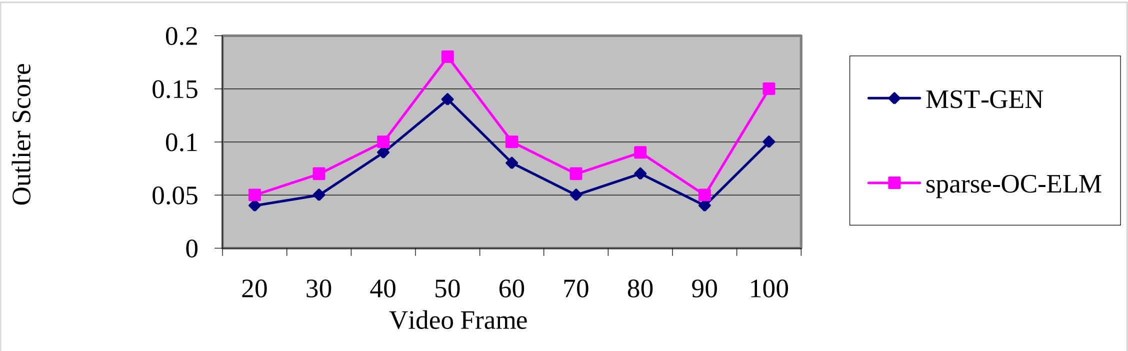


Fig. 2 Compare different frame sequence with outlier score

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

TABLE II COMPARE FPR AND TPR IN THE INPUT VIDEO

| Algorithm | FPR | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.9 | 1 |
| MST-GEN | 0.24 | 0.37 | 0.46 | 0.57 | 0.67 | 0.71 | 0.78 | 0.84 | 0.88 |
| Sparse-OC-ELM | 0.34 | 0.56 | 0.68 | 0.74 | 0.79 | 0.86 | 0.93 | 0.95 | 1 |

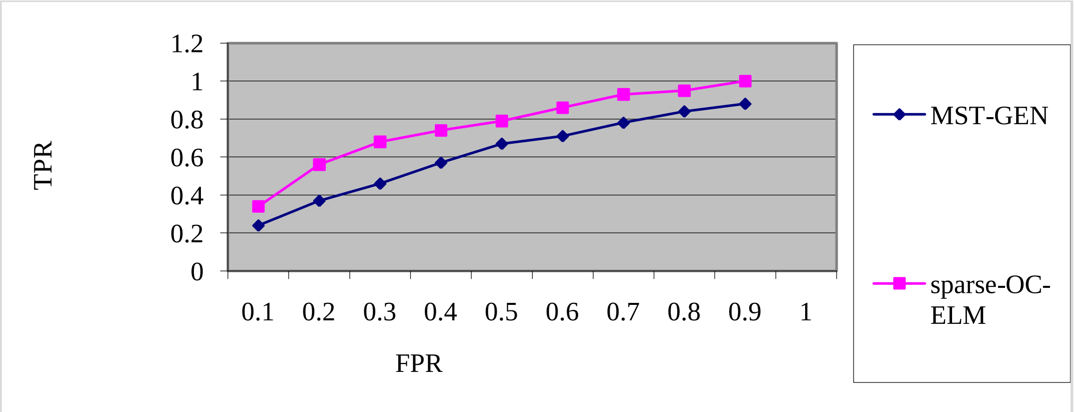


Fig. 2 Compare true positive rate with video frame sequence

## V. CONCLUSION

The presents an online learning with sparse regularized kernel based one-class Extreme Learning Machine (ELM) classifier and is referred as "sparse-OC-ELM". The baseline kernel hyperplane model considers whole data in a single chunk with regularized ELM approach for offline learning in case of One-Class Classification (OCC). The regularized kernel ELM based online learning and consistency-based model selection has been employed to select learning algorithm parameters. The online RK-OC-ELM has been evaluated on standard benchmark datasets as well as on artificial datasets and the results are compared with existing state-of-the art one class classifiers.

## REFERENCES

[1] R. J. Bolton and D. J. Hand, "Unsupervised profiling methods for fraud detection," in Proc.Credit Scoring Credit Control VII, 2001, pp. 235–255.
[2] A. Lazarevic, L. Ert€oz, V. Kumar, A. Ozgur, and J. Srivastava, "A comparative study of anomaly detection schemes in network intrusion detection," in Proc. 3rd SIAM Int. Conf. Data Mining, 2003, pp. 25–36.
[3] K. I. Penny and I. T. Jolliffe, "A comparison of multivariate outlier detection methods for clinical laboratory safety data," J. Roy. Statistical Soc.: Series D (The Statistician), vol. 50, no. 3, pp. 295–307, 2001.
[4] Z. He, X. Xu, J. Z. Huang, and S. Deng, "Mining class outliers: concepts, algorithms and applications in CRM," Expert Syst. Appl., vol. 27, no. 4, pp. 681–697, 2004.
[5] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Comput. Surveys, vol. 41, no. 3, p. 15, 2009
[6] P. J. Rousseeuw and M. Hubert, "Robust statistics for outlier detection," Wiley Interdisciplinary Rev.: Data Mining Knowl. Discovery, vol. 1, no. 1, pp. 73–79, 2011.
[7] E. M. Knox and R. T. Ng, "Algorithms for mining distance-based outliers in large datasets," in Proc. Int. Conf. Very Large Data Bases, 1998, pp. 392–403.
[8] S. Ramaswamy, R. Rastogi, and K. Shim, "Efficient algorithms for mining outliers from large data sets," ACM SIGMOD Rec., vol. 29, no. 2, pp. 427–438, 2000.
[9] M. M. Breunig, H. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," ACM SIGMOD Rec., vol. 29, no. 2, pp. 93–104, 2000.
[10] S. Chawla and A. Gionis, "k-means–: A unified approach to clustering and outlier detection," in Proc. SIAM Int. Conf. Data Mining, 2013, pp. 189–197.
[11] A. Lazarevic and V. Kumar, "Feature bagging for outlier detection," in Proc. 11th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2005, pp. 157–166.
[12] A. Zimek, E. Schubert, and H. Kriegel, "A survey on unsupervised outlier detection in high-dimensional numerical data," Statistical Anal. Data Mining, vol. 5, no. 5, pp. 363–387, 2012.
[13] S. Papadimitriou, H. Kitagawa, P. B. Gibbons, and C. Faloutsos, "LOCI: Fast outlier detection using the local correlation integral," in Proc. 19th Int. Conf. Data Eng., 2003, pp. 315–326.
[14] A. Adler, M. Elad, Y. Hel-Or, and E. Rivlin, "Sparse coding with anomaly detection," in Proc. IEEE Int. Workshop Mach. Learn. Signal Process., 2013, pp. 1–6.
[15] B. Zhao, L. Fei-Fei, and E. P. Xing, "Online detection of unusual events in videos via dynamic sparse coding," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2011, pp. 3313–3320.
[16] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," CS Technion, p. 40, 2008.
[17] D. Pokrajac, A. Lazarevic, and L. J. Latecki, "Incremental local outlier detection for data streams," in Proc. IEEE Symp. Comput. Intell. Data Mining, 2007, pp. 504–515.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

[18] F. Angiulli and F. Fassetti, "Distance-based outlier queries in data streams: the novel task and algorithms," J. Data Mining Knowl. Discovery, vol. 20, no. 2, pp. 290–324, 2010.
[19] D. Yang, E. A. Rundensteiner, and M. O. Ward, "Neighbor-based pattern detection for windows over streaming data," in Proc. 12th Int. Conf. Extending Database Technol.: Adv. Database Technol., 2009, pp. 529–540.
[20] M. Kontaki, A. Gounaris, A. N. Papadopoulos, K. Tsichlas, and Y. Manolopoulos, "Continuous monitoring of distance-based outliers over data streams," in Proc. IEEE 27th Int. Conf. Data Eng., 2011, pp. 135–146.
[21] J. K. Dutta and B. Banerjee, "Online detection of abnormal events using incremental coding length," in Proc. 29th AAAI Conf. Artif. Intell., 2015, pp. 3755–3761.
[22] B. Banerjee and J. K. Dutta, "SELP: A general-purpose framework for learning the norms from saliencies in spatiotemporal data," Neurocomputing, vol. 138, pp. 41–60, 2014.
[23] A. Borji and L. Itti, "Exploiting local and global patch rarities for saliency detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recog. 2012, pp. 478–485.
[24] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254–1259, Nov. 1998.
[25] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 1, pp. 18–32, Jan. 2014.
[26] Y. Cong, J. Yuan, and J. Liu, "Abnormal event detection in crowded scenes using sparse representation," Pattern Recog., vol. 46, no. 7, pp. 1851–1864, 2013.
[27] J. Kim and K. Grauman, "Observe locally, infer globally: A spacetime MRF for detecting abnormal activities with incremental updates," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2009, pp. 2921–2928.
[28] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2009, pp. 935–942.
[29] A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz, "Robust real time unusual event detection using multiple fixed-location monitors," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 3, pp. 555–560, Mar. 2008.