International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 05 | May 2019 www.irjet.net p-ISSN: 2395-0072

ANOMALY DETECTION SYSTEM IN CCTV DERIVED VIDEOS

PROF NANDHINI N[1], BARATH KUMAR M R[2], LALIT SHARMA[3], ANKIT GUPTA[4]

[1]*Prof NANDHINI N, Dept. of CSE, Sapthagiri College of Engineering, Karnataka, INDIA*
[2]*BARATH KUMAR M R, Dept. of CSE, Sapthagiri College of Engineering, Karnataka, INDIA*
[3]*LALIT SHARMA, Dept. of CSE, Sapthagiri College of Engineering, Karnataka, INDIA*
[4]*ANKIT GUPTA, Dept. of CSE, Sapthagiri College of Engineering, Karnataka, INDIA*

ABSTRACT - ***People in the present world each and everyone thinks about one thing that's security. Providing this security depends on various things. One of them is to install surveillance cameras where it helps in preventing and alerting the people. Analysis of the information captured using these cameras can play effective roles in event prediction, online monitoring applications including anomalies. Nowadays, various Artificial Intelligence techniques have been used to detect anomalies, amongst them convolutional neural networks using deep learning techniques improved the detection accuracy significantly. The goal is to propose a new method based on deep learning techniques for anomaly detection in video surveillance cameras with higher accuracy.***

Keywords — CNN, Deep Learning, Image classification ALGORITHMS, surveillance cameras.

1. INTRODUCTION

In this project we are trying to develop a system for detecting anomalies in CCTV derived videos. As in today's world everyone wants to live in a secure Environment, and we believe that education without giving back to the society is meaningless. So through this project we are trying to improve society's security by some level. In our project we are using deep learning techniques to detect anomalies. The deep learning falls under machine learning. All machine learning tasks are classified into two broad categories first is supervised learning (requires labels) and other is unsupervised learning (does not require labels). We are using the CNN algorithm which is a completely supervised learning method. Anomaly activities will be different in different scenarios, for example the movement of vehicle on a pedestrian pathway will be unusual, and the movement of a person on foot on a highway will be unusual. The information or input to be read by the system in our system will be in high dimensions (large size).Detection in videos is

More difficult than other data, since it involves many methods and also requires video processing as well. One of the best approaches for processing this information is using advanced machine learning techniques such as deep learning. The main purpose or aspect behind deep learning is the Feature extraction. It means extracting data from the given CCTV derived videos. The architecture of this method has two main phases. The first one is the Train network and the second one is detection classifier. The train network deals with the feature extraction step and the detection classifier takes up the decision of whether there is an anomaly activity or not by taking the final decision.

The processing of surveillance cameras information in crowded scenes poses serious challenges and difficulties. If this process is online, the complexity will even increase. One of the best approaches for processing this information and consequently achieving the goal-oriented pattern is the use of advanced machine learning techniques such as deep learning approaches. The advantage of these types of processes, which usually have a high dimensional data, can be traced back to the existence of an end-to-end system. The main contribution of this paper is the use of deep learning techniques in all phases of anomaly detection. One of the main purpose of using deep learning is to extract information from high dimension data. In other sections of this paper, at first, an introduction is given and at section 2.

2. REALATED WORKS

There has been a lot of research in developing an artificial intelligent system which detects anomalies in the environment. We have referred some papers which have the similar idea and methodology. These papers have helped us in gaining at least some knowledge about our project. The survey papers don't exactly have the same procedures but they do have a common goal and that is detecting any anomaly activity happening in the area of surveillance. Here are the lists of survey papers we have referred in implementing our project.

A. Improved Anomaly Detection in Surveillance Videos based on A deep Learning method

This paper was published in the year 2017 by Ali Khaleghi. This paper introduces an anomaly detection method based on deep learning techniques. The architecture of this method has two main phases which are called train network and detection classifier. The first phase aims for feature extraction and is consisted of five components with a deep structure. The aim of the second phase is detection.

This phase is consisted of five deep neural network classifiers and reconstruction network. Each component in detection phase produces a detected class and a score. At last, by these detection classes and scores, the ensemble classifier performs the final detection and announces it.

The main contribution of this paper is the use of deep learning techniques in all phases of anomaly detection. One of the best approaches for processing this information and consequently achieving the goal-oriented pattern is the use of advanced machine learning techniques such as deep learning.

B. Artificial Neural Network based Anomaly Detection System

This paper was published in the year 2015 by Marjan Bahrololum. The basic idea of this paper was to detect any intrusion/anomaly activity in the CCTV installed areas. The anomaly can be different for different scenarios. In this paper he introduced a method of using a combination of both Neural Network (NN) and Decision Trees (DT's). Neural Network is a computer system which is modeled on human brain and nervous system.

Decision Tree is a graph that branching method which helps in achieving every possible outcome of a decision. In this method the Decision Trees were useful in detecting known attacks, whereas the Neural Networks were useful in detecting unknown attacks.

Advantages

- Artificial neural networks are a uniquely power tool in multiple class classification.
- The first advantage in the utilization of a neural network in the detection of the network intrusion would be the flexibility that the network would provide.

C. Support Vector Machine based Anomaly Detection

This paper was published in the year 2016 by Latifur Khan. The goal of this paper was to detect anomaly particularly when contracting with large datasets. In this method the use of DGSOT (Dynamically growing self-organizing Tree) was implemented. At starting the datasets were small and then according to learning of the machine the dataset was increased as more and more data (training examples) were stored in them.

This method of using the super vector machine is a completely supervised method of learning. In this method the set of examples are represented as points in the space, and according to the gaps present between the groups the examples are grouped with similar points.

This process of grouping points in different groups is known as clustering. The new training examples are set into the space and according to the gap in which they fall; they are classified in that particular group. This method was useful as the machines were able to detect the new types of anomalies happening in the environments. So summarized intro of this system would be that the events/activities recorded by the cameras were stored and then they were distributed among certain groups and finally they were being detected by the system.

D. K-means algorithm based Anomaly Detection System

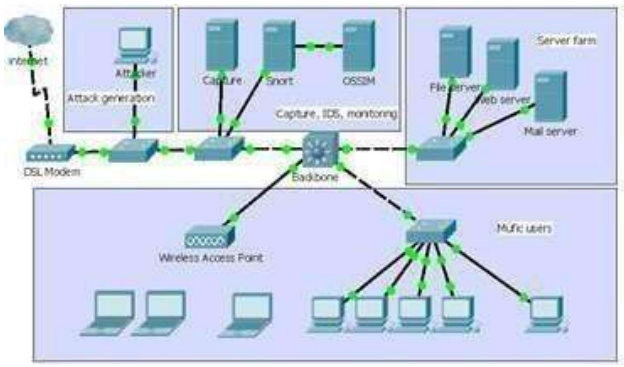
This paper was published in the year 2013 by a Chinese researcher named Yu Guan. He was the first person who proposed a method to detect the Anomalies by using the K-means algorithm. In K- means algorithm method the different numbers of observations (dataset) are divided in a number of clusters.

The clusters are represented by K- clusters. Each cluster has its own features based on which the data points are put in a group. Each cluster has a centroid which is basic element of the clusters; this centroid has some features (valuable information) which separate that particular cluster from other clusters created.

The K-means algorithm is a completely unsupervised method of learning. It means that the input data which is provided to the system does not need any label for getting the output. As soon as the input is given to the system, it is analyzed by the system by some algorithms and the feature (property) of that data is compared with the centroids of the clusters which are already present, whichever cluster has similar features as of the given input becomes the cluster (group) for that particular input. After this a detection algorithm is used to detect the unusual activity happening and the authorized user is informed.

E. Hybrid Anomaly-based Detection Approach

This paper was published by Shekhar R. Gaddam in the year 2017. In this method he proposed a method to detect anomaly activities happening in the environment by combining the K-means algorithm and the ID3 (Iterative Dichotomiser 3) Decision Tree. The process of decision making from a dataset uses the ID3 decision tree. The ID3 decision tree begins at a root node S (the base category), and as the algorithm reaches the next step from the first step, it iterates from one variable to another. The algorithm selects the unused attribute in the tree and calculates the entropy H(S) or information gain IG(S). It selects the attribute which has the smallest entropy or largest information gain. Then S is partitioned or divided from the selected attribute to create subsets of the dataset. In the proposed paper the K-means algorithm and ID3 decision tree were combined together to get the results. The K-means algorithm first divided the training cases into K-clusters, and on each cluster an ID3 Decision Tree was built which classified the clusters into two categories. They were "Normal instances" and "Abnormal instances". By studying the subgroups inside each cluster, the ID3 Decision Tree purified the decision boundaries. This method of combining different techniques to achieve the goal was a really effective method.



F. Multilevel hierarchical Kohonen Net (K-Map) based Anomaly Detection.

This paper was published in the year 2016 by Mrs. Susheela T. Kohonen's networks are one of basic types of self-organizing neural networks. The ability to self-organize provides new possibilities - adaptation to formerly unknown input data. It seems to be the most natural way of learning, which is used in our brains, where no patterns are defined. Every level of the hierarchical map was modeled as a straightforward winner-take-all K-Map. The computational effectiveness is the major advantage of this multilevel hierarchical K-Map. As there were different levels of computations so the tasks were divided among different levels and hence it was efficient in making the decisions. In this proposed method of anomaly detection the main advantage was the small size of the network. The concept of self-organizing maps was used in this method. A SOM (self-organized map) is basically a method which reduces the size of high dimensional data, so that computation can take place more efficiently. The categorization method was used in this type of anomaly detection. The dimension of the input data was also chosen accordingly.

G. Soft Computing approach for Anomaly detection.

This paper was introduced by Adal Nadjara Toosi and Mohsen Kahani in the year 2010.This method used the ANFIS (Adaptive neuro fuzzy inference system) network. This network is a kind of artificial network that applies logical rules to knowledge base to deduce new information. It works on both neural networks and fuzzy logics also. Fuzzy logic is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1. It is employed to handle the concept of partial truth, where the truth value may range between completely false. The ANFIS network is really good in generating fuzzy rules without human interactions. This was the main idea behind this method of anomaly detection.

H. TCM-KNN Based Network Anomaly Detection

This method was introduced by Yang Li and Li Guo in the year 2015. TCM-KNN stands for Transductive Confidence Machines for K- Nearest Neighbors. It was successfully identified anomalies with elevated detection rate, low false positives under the condition of utilizing much less chosen data as well as selected features for training in association managed intrusion detection techniques. The recommended method was more tough and successful than the state-of-the-art intrusion detection methods which were explained by a chain of experimental results on the familiar KDD Cup 1999 data set. In this method also classification of data points was done, and use of nearest neighbor algorithm was used.

Advantages

- Transduction can offer measures of reliability to individual points, and uses very broad assumptions except for the well-known id assumption (the training as well as new (unlabeled) points are independently and identically.

I. Anomaly Based DDoS Attack Detection

There are a lot of other methods which were used in developing a system to detect any anomaly activity happening in the environment. In the year 2015, Iwan Syarif has illustrated the compensation of utilizing the variance detection approach over the mishandling detection technique in detecting unknown network intrusions or attacks. When applied to anomaly detection it also examined the presentation of different grouping algorithms. We have five different clustering algorithms:

k-Means, improved k-Means, k-Medoids, EM clustering and distance-based outlier detection algorithms were utilized. Their testing showed that mishandling detection techniques, which executed four dissimilar classifiers (naïve Bayes, rule induction, decision tree and nearest neighbor) unsuccessful to detect network traffic, which enclosed a large number of unknown interferences.

In the same year in 2015, an intangible model for identifying and mitigating Distributed Denial-of-Service (DDoS) attacks and its incomplete achievement has been offered by Sajal Bhatia. To identify DDoS attacks, to distinguish them from like looking FEs, and to bring about source IP based mitigation strategies upon attack identification an assembly of network traffic and MIB server load data analysis was used in the mold. The testing and presentation assessment of the suggested model was performed by means of artificial network traffic, intimately on behalf of real-world DDoS attacks and FE traffic, a produced using a software-based traffic generator developed. Prasanta Gogoi has suggested an actual dataset to modernize this critical inadequacy. A test bed has been set up by them to begin network traffic of both attack as well as standard nature by means of attack tools. The network traffic in sachet and flow format was incarcerated by them. To produce a featured dataset the incarcerated traffic was sorted out and preprocessed. For investigate purpose the dataset was made accessible. They have High-level study of the KDD Cup 1999 and NSL-KDD datasets which are offered by them.

3. METHODOLOGY

The proposed method of this paper is based on deep learning techniques for detecting anomalies in video. Two main components are considered for this method. The first component is the extraction and learning of the feature and the second component is the detection of anomalies. Apart from these two components, there is a pre-processing step which is related to background estimation and removal. Like all machine learning approaches, this method also has two main train phase and test phase.

The first component is the extraction and learning of the feature and the second component is the detection of anomalies. Apart from these two components, there is a pre-processing step which is related to background estimation and removal. Like all machine learning approaches, this method also has two main train phase and test phase. In train phase, features are trained by train parts of dataset which contains only normal frames, and trained model in test phase is used by other parts of dataset which contain abnormal frames. Figure 3 illustrates the overall framework of the proposed method.

As can be seen in the figure, learning features are of four main types. For some types, feature extraction processes are performed on single frames, and others are based on patch frames in order to reduce cost and training time. The first feature is appearance which is related to object detection in each frame; and by comparing each frame with previous and next frames the detection score is generated. The second feature is density which is about density of objects in each frame; the final score is generated based on frames comparison and average speed. The third feature is motion which is based on the flow of objects between patch frames and it generates optical flow and a sequence of video then used for another score on anomaly. The last feature is scene which is based on patch.frames and reconstructing a scene from learned model.

A. *Pre-Processing*

The first step before starting extracting and learning features is to estimate and remove the background. The background is indeed different for different scenarios as there are various methods for its removal. For instance, the background might include empty spaces or street borders. In this method, the background estimation is based on most occurrence of frequency (MOF) between video frame patches [9]. For the background estimation steps at first, a histogram is generated for each frame of the video which is based on pixels and their location in the image. Then the histogram of the frames in each patch is compared with each other, and the maximum values per patch are identified as background and are thus grayed. Removing the background will reduce the cost of the computing and the processing time. This step is considered as a part of train network.

B. *Feature Extraction and Learning Component*

In addition to background estimation, train network has four main components. The deep network for extracting appearance feature uses a stacked denoising auto-encoder (SDAE) with 6 encode layer and the same structure of decode layer [17, 23]. Each frame is convolving to network with 1*1 window size and it includes stride and padding. All frames normalize in binary mode. This SDAE has 6 encode layers and 6 same structure in decode layer which is deeper than the existing methods. The output of this step is detected objects which are called appearance representation. This output is used in detecting phase and also is utilized as an input to density estimation component in order to increase the accuracy of estimation.

Density Estimation [25] is carried out by convolutional neural network with 8 * 8. Windows filter. This network is shown in Figure 4. The output of this component is feature map and the loss function is computed based on square error. In the estimation of the density, the sectors associated with the background are considered zero.

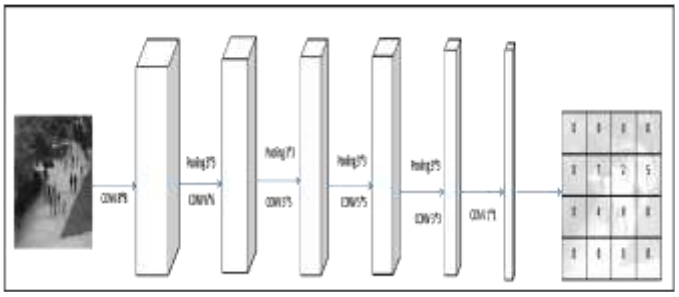


Fig. 4. The structure of density estimation.

C. *Detection Component*

In the detection component, learned features which are generated in train network are given to a classifier with two classes of normal and abnormal. Features are given as individual and combined feature to these networks. Reconstruction error and appearance features are given to network as a combined feature since the appearance feature or object detection with a reconstruction error can be a strong feature for the detection of anomalies. The lower reconstruction error for the corresponding frame will make the detection more accurate.

Two other combination features are Motion Feature and density map. These are two complementary features and the direction of motion must be equal to the transfer of density direction.

The classifiers used in this method are simple deep classifiers which used the softmax function. As can be seen in Figure 4, five classifiers with the same structure are used in the detection step. There are 5 hidden layers in these networks in order to reduce the computing cost overhead. The last layer of these networks is fully connected. Each of these classifiers finally detects anomaly or normal situation and produces a score for the percentage of anomalies presence. This score ranges between [0 – 1].

The last component is final decision-making (ensemble) which determines the final detection result. This classifier is a simple linear classifier that declares the final result based on the percentage of votes and the score of other classifiers. The structure of this component is defined in a way that if four out of six classifiers vote for anomalies, the detection is declared as anomaly and the score is announced as the average of other classifier scores.

D. *Text messaging using the gateway API.*

After a Anomaly is been detected by the system we have trained it will notify the authorized number mentioned in API. The gateway provide with API key after the purchasing of the API. Then the gateway tells the network operator that to send a text to the registered number. Later the user gets a text saying there was anomaly activity detected from surveillance camera 2 or how many surveillance cameras that were installed.

The last component is final decision-making (ensemble) which determines the final detection result. This classifier is a simple linear classifier that declares the final result based on the percentage of votes and the score of other classifiers. The structure of this component is defined in a way that if four out of six classifiers vote for anomalies, the detection is declared as anomaly and the score is announced as the average of other classifier scores.

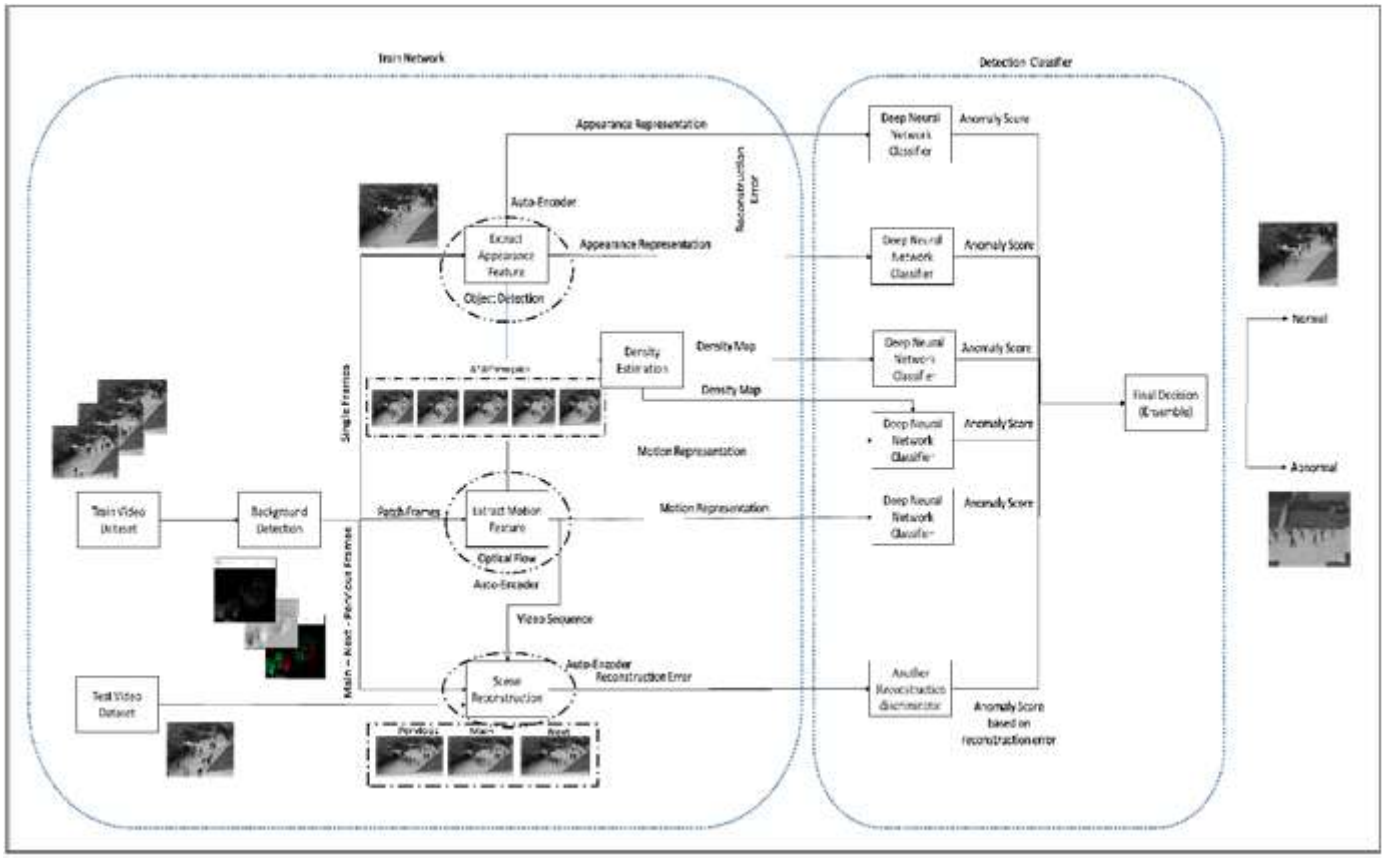


Fig. 3. The overall framework of proposed method

The density estimation and appearance representation is generated based on single frame analysis.

The third component is motion feature extractor [17, 23]. It performs a feature extraction based on the direction of moving objects in the scene of video patches. This deep network also has a similar structure to appearance feature extractor but it is based on frames patches. After entering the patch frame into the network, computing optical flow will be done based on comparison of frames in a patch. The output of this step is Motion Representation which is used for future detection.

The last component is Scene Reconstruction which is based on reconstruction network [26]. The structure of this reconstruction network is based on convolutional Auto- Encoder with the same CNN generator and discriminaton networks. Generator part regenerate the scene which has 10 layers to reconstruct frames based on the previous and the next

frame in same patch and the discriminator compares the generated scene with original one in order to compute the reconstruction error. It should be mentioned that discriminator part has the same structure as that of the generator. A high reconstruction error during test indicates anomalies. The reconstruction error in train network is low and this will be a measure for detecting anomalies.

At the end of the training step, a set of learned and combined features is created in order to achieve anomaly detection.

4) EXPERIMENTAL RESUALTS

To evaluate the proposed method and compare it with other available methods, public UCSD dataset is used which is one of the most famous dataset related to the anomaly detection. This dataset is related to the pedestrian walkaway surveillance camera. Any objects other than people are identified as anomaly, such as bicycle or car. This dataset has ped1 and ped2 parts that are related to cameras with a different angle. Both parts have test and train data [23].

In this section, by evaluating the proposed method on this dataset, it will be shown that the proposed method indicates a significant improvement on the existing methods.

A. Evaluation Configuration

At first, the train part of ped1 is given to network in order to train the network. The trained network produces the necessary outputs. These outputs are given to detection classifier for anomaly detection. The test part of ped1 has also evaluated this method and its outcomes.

Then, the results are compared with the output of other existing methods that are implemented and simulated in a completely similar situation.

All evaluations are conducted under similar conditions and with three parallel systems with processor Intel® Core™ i7 – 7700HQ and a graphics processor NVIDIA GeForce GTX 1050 . Each learning operation takes more than 24 hours.

Figure 5 is an example of anomaly detection in the data that is related to the extraction of features. The first-row images are the original frame. The second-row images are detected objects and the third-row shows optical flow. The last images are decision-making on the anomaly or normal situation.

B. Basic Methods and Evaluation Metric

The proposed method is compared with the following methods in quite similar situations. The evaluation results show the improvement in the proposed method. Also one-class SVM has been previously used in order to evaluate the deep neural network classifier accuracy whose comparison with the proposed method indicates an increase of 15% to 20% in the accuracy [22].

- Multi-Column Convolutional Neural Network (MCCNN) because it used CNN technique and the output of the network is density map which has similarities to the proposed method [25].
- Learning Deep Representations of Appearance and Motion (LDRAM) which use appearance and motion feature for anomaly detection [17, 23].
- Deep learning-based anomaly detection (DLAD) system which uses deep learning classifier [9].
- Deep Generative which is used for Auto-Encoder reconstruction [6].

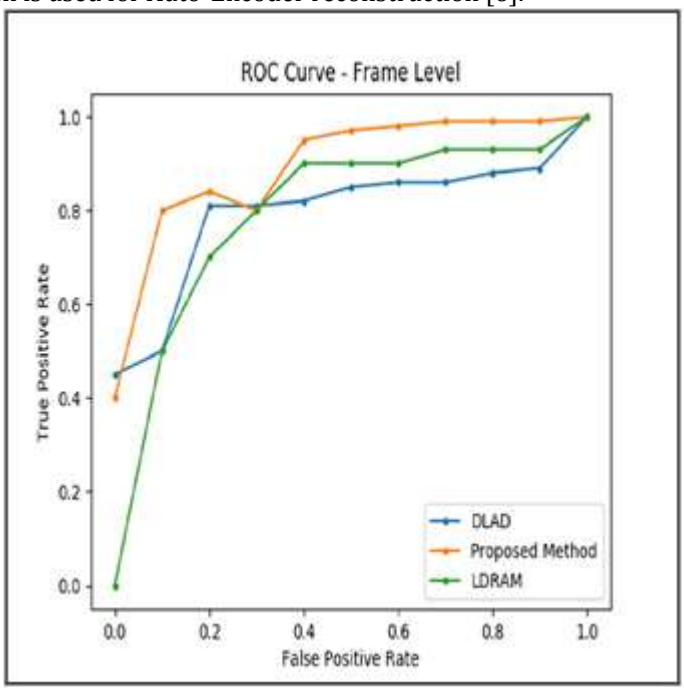
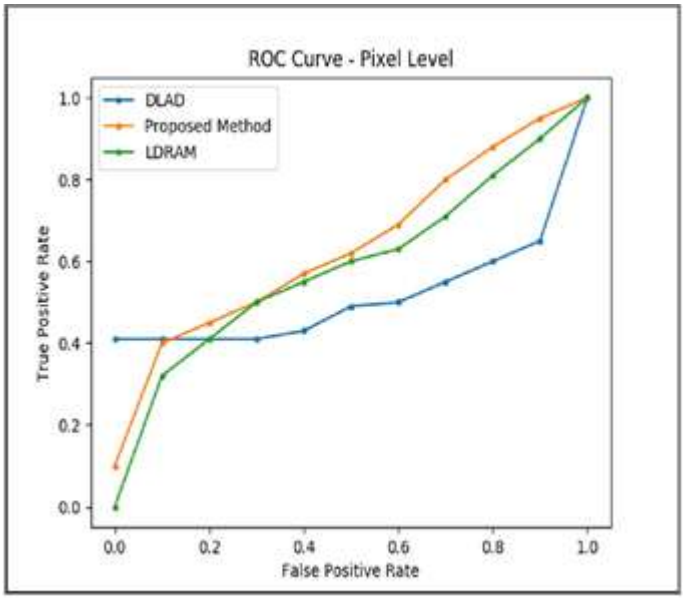


Fig. 6. Evaluation results based on ROC metric in frame level

The graph plotted is based on the accuracy of anomaly detected.

C. Evaluation Results

Evaluation is done on ped1 part of UCSD dataset. The four method are also implemented and MAE method to plot the graph and to evaluate the graph.



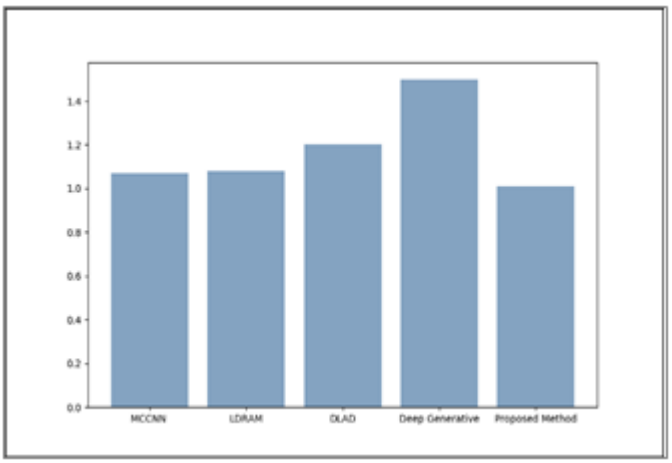


Fig. 7 Evaluation matrix graph using ROC.

5) CONCLUSION AND FUTURE WORK

In this paper, a new deep learning based for anomaly detection of video surveillance cameras is introduced. One advantage of this method is the use of deep learning techniques in all train and detection components. The two main components of this method are evaluated based on some metrics and with UCSD dataset which is the most famous anomaly detection dataset. Another benefit of this method is the isolation of train network phase. So it can use as a pre-train network in similar works.

For further improvement, it is possible to add a component which can add descriptions to each detection classifier or to the last one; or it is possible to add a component in the detection phase which can localize the anomaly accurately.

ACKNOWLEDGEMENT

We, would like to thank our Principal Dr. K L SHIVABASAPPA, for giving us an opportunity to do this Project.
We also would to like thank our parents for the moral support.

REFERENCES

[1] Ali Khaleghi "Improved anomaly detection in surveillance videos using A deep learning method" Department of Computer and Information Technology Engineering Qazvin Branch, Islamic Azad University Qazvin, Iran, 2017.

[2] Marjan Bahrololum "Neural network based Anomaly detection", Department of computer Science and Engineering, 2017.

[3] Yang Li and Li Guo "TCM-KNN based Anomaly Detection" TCM-KNN means Transductive Confidence Machines in the year 2015.

[4] Iwan Syarif "Other methods for detection of Anomaly" has illustrated the compensation of utilizing the variance detection in the year 2015.

[5] Shean Chong, Yong Haur Tay, Yong, "Modeling Representation of Videos for Anomaly Detection using Deep Learning: A Review", arXiv:1505.00523v1, 2015.

[6] Shekhar R. Gaddam "Hybrid Technique based Anomaly Detection". By combining the K-means algorithm and the ID3 (Iterative Dichotomiser 3) Decision Tree, 2015.

[7] Latifur Khan. "Support Vector Machine based Anomaly Detection", 2017.

[8] Siqi Wanga, E.Z., Jianping Yin, "Video anomaly detection and localization by local motion based joint video representation and OCELM", Neurocomputing, 2017.

[9] Hung Vu, Tu Dinh Nguyen, Anthony Travers, Svetha Venkatesh and Dinh Phung, "Anthony Travers, Energy-Based Localized Anomaly Detection in Video Surveillance", Springer International Publishing AG, 2017