# Deep Anomaly Net: Detecting Moving Object Abnormal Activity Using Tensor Flow

Ebtesam Shadadi[1], Saahira Ahamed[2*], Latifah Alamer[3] and Mousa Khubrani[4]

[1]Department of Computer Science, College of Computer Science and Information Technology, Jazan University, Jazan, Kingdom of Saudi Arabia. ashedadi@jazanu.edu.sa

[2*]Department of Computer Science, College of Computer Science and Information Technology, Jazan University, Jazan, Kingdom of Saudi Arabia. sahamed@jazanu.edu.sa

[3]Department of Information Technology and Security, College of Computer Science and Information Technology, Jazan University, Jazan, Kingdom of Saudi Arabia. aalamer@jazanu.edu.sa

[4]Department of Computer Science, College of Computer Science and Information Technology, Jazan University, Jazan, Kingdom of Saudi Arabia. mmkhubrani@jazanu.edu.sa

Received: August 02, 2022; Accepted: September 30, 2022; Published: November 30, 2022

**Abstract**

Sparse secret writing, primarily based on abnormal detection, has shown promising performance, key features being feature learning, subtle illustrations, and vocabulary learning. Propose a replacement neural network for anomaly detection called AnomalyNet by deep feature learning, sparse representation, and dictionary learning in 3 collaborative neural processing units. In particular, to obtain higher functions, form the motion fusion block in the middle of the function transfer block to enjoy the benefits of eliminating background noise, motion capture, and eliminating information deficit. In addition, to deal with some of the shortcomings (such as non-adaptive updating) of existing sparse coding optimizers and to take advantage of the advantages of neural network (such as parallel computation), design a unique continuous neural network, which will be told as a thin illustration of a docent dictionary by proposing a consistent iterative rule of hard threshold (adaptive ISTA) and the reformulation of adaptive ISTA as a substitute for long-term memory (LSTM). As far as we know, this may be one of the first works to link the `1-solvers and LSTM and offer new insights into LSTM and model-based refinement (or so-called differential programming), but primarily in the form of detection-based sparse secret writing anomaly. In-depth, experiments show the progressive performance of our technique in the task of detecting abnormal events.

**Keywords:** Deep Learning, CNN, LSTM, Object Detection, Prediction.

## 1 Introduction

Increasing security demands have led to the widespread deployment of police investigation cameras. Surveillance video analysis is faced with a major challenge when it comes to police work abnormal

*Journal of Interner Services and Information Security (JISIS)*, volume: 12, number: 4 (November), pp. 116-125
DOI: *10.58346/JISIS.2022.I4.008*

*Corresponding author: Department of Computer Science, College of Computer Science and Information Technology, Jazan University, Jazan, Kingdom of Saudi Arabia.

events, which requires exhausting human efforts. An effortful task like this can be recast as an anomaly detection downside. Anomaly detection differs from normal classification in the following ways: 1) listing all possible negative (anomaly) samples is terribly difficult. The rarity of the samples makes it difficult to gather adequate negative samples. Among the most popular techniques to detect anomalies is to use videos of traditional events as coaching information to find out a model, and then police the anomalous events that don't match the learned model. [1] [2].

By combining wordbook learning and sparse representation, thin committal to writing has proven to be successful in anomaly detection. By learning a dictionary from a set of normal events, SCAD (sparse coding-based anomaly detection) discovers abnormal events that cannot be precisely recreated by normal events alone [3]. An overview and systematic comparison of the state of the art on crowd video analysis is presented in this paper. As a result of an increase in intelligent video police investigation algorithms capable of analyzing mechanically visual streams of terribly jammed and littered scenes, such as aerodrome concourses, railway stations, shopping malls, and also the like, the explanation for our review can even be found in this review. Computer vision analyzers have been focused on intelligent solutions to protect doubtless very crowded public areas since security became a priority [4, 5].

The aim of this paper is to propose a review article of existing literature concerning the automated analysis of complicated and jammed scenes. The literature is split into 2 broad categories: the mega scopic and also the microscopic modelling approach. The hassle is supposed to produce a indicator for all pc vision practitioners presently engaged on crowd analysis. We have a tendency to discuss deserves and weaknesses of assorted approaches for every topic and supply a recommendation on however existing ways may be improved. In the past decades, SCAD has utilized a variety of options. There is extensive use of bar graphs of oriented gradients (HOG) [6], 3D spatiotemporal gradients, and also bar graphs of orienting flows (HOF). Data-driven works are more favorable since they may result in better performance, whereas handcrafted ones have the main disadvantage of being handcrafted. Recent works have attempted to marry deep learning and anomaly detection in order to enhance the representative capability of neural networks. Convolutional filters are used in conjunction with a repeated neural network (RNN), for example. [7,8].

## 2 Related Works

As their ways adapt to long-variable discourse dynamics, the motion and the look are implicitly encoded. Despite their promising performance, these methods suffer from two limitations. Motions and appearances are encoded by the RNN and convolutional filters separately, which suggests that spatial-temporal relationships between motions and appearances are broken. As a result, performance is also inferior. As opposed to that, options are generally learned from scratch while not considering pre-trained models from relevant connected tasks [9, 10].

The performance of methods can be significantly improved by using transferable models, according to various studies. Existing 1-solvers that resemble unvarying hard-thresholding (ISTA) use a non-adaptive change approach by updating the parameters with a set learning rate on all dimensions. For sparse/big information, for example, the per-dimension updating theme is typically needed to save memory and price. However, this strategy might not be best in some cases and result in inferior performance. The 1-solvers resembling ISTA don't take into account the historical info once coming up with the updating rules [11]. Numerous studies in the optimization community have demonstrated that incorporating historical information enhances algorithm convergence. Predicting thin codes in inference is terribly computationally intensive. A sparse committal to writing has a time quality proportional to its scale dimension, still because of the input dimension for every information point. Sparse illustration and wordbook learning are conducted uniformly. All ancient "1-solvers" may obtain an honest sparse representation, but never produce an interesting dictionary since sparse representation and dictionary learning are separate processes [12].

As an alternative to the above limitations, we propose a feature learning network that consists of a motion fusion block and a transfer block. Video clips are compressed into one image using motion fusion while the orthogonal background is suppressed [13, 14]. In this way, the motion and look are often combined into a single image. The spatial-temporal options (i.e., appearance and motion) are extracted from the compressed pictures using the feature transfer block. Alternatively, utilize information from other related tasks/domains to enhance feature learning. Using SCAD's key technique, thin writing requires lexicon learning optimization through 1-regularization that is computationally inefficient, particularly in the context of video analysis. Though variety of ways is proposed, they need still suffered from limitations. To beat the limitations, propose a completely unique `1-solver termed adaptive ISTA, by introducing an adaptive momentum vectors to change per-parameter updates and encapsulate the historical information. Concomitant with benefits of the adaptive ISTA, the disadvantage is that the problem in the optimization of parameters. Our adaptive ISTA requires mechanically learning ds parameters, while ISTA relies solely on a single parameter, ds, and the dimension of thin codes. Adaptive ISTAs can be recast as continuous neural network units (RNNs), referred to as sparse long short-term memories (SLSTMs), which are a variant of long short term memories (LSTMs), in order to overcome the optimization issue and therefore these two limitations.

## 3 Proposed Methodology – Sparse Method

The adaptive momentum vectors work because the predicted SLSTM has input and output gates. In the new formulation, optimized lexicons and sparse codes correspond to the load and outputs of SLSTM. Build a neural network (called SC2Net) to implement thin codes in an end-to-end unsupervised manner and use SC2Net to detect anomalies. SLSTM solves four constraints using a continuous neural network, unlike conventional '1solvers such as ISTA. The predicted SLSTM performs sparse coding in a completely different way compared to the LSTM. Among existing RNN-based type-1solvers such as Learned ISTA (LISTA), the main variations can be found in the following aspects is given in fig. 1.

First, SLSTM achieves the optimization of the sparse codes of the LSTM block rather than the simple RNN block. Thus, SLSTM is ready to capture historical information that is useful for increasing the convergence associated with improving the performance of our model. In other words, LISTA still suffers from the same main 2 limitations as ISTA, while our SLSTM does not. Second, the SC2Net project will not depend on other demand optimizers. In contrast, LISTA needs to victimize the subtle codes that other 1-solvers such as ISTA provide because the supervisor. Such variations make it possible for our technique to function in an end-to-end manner. The supported above feature extraction network and optimization network, proposed a replacement model called AnomalyNet to observe the abnormal events in the videos shown in Fig. 2.

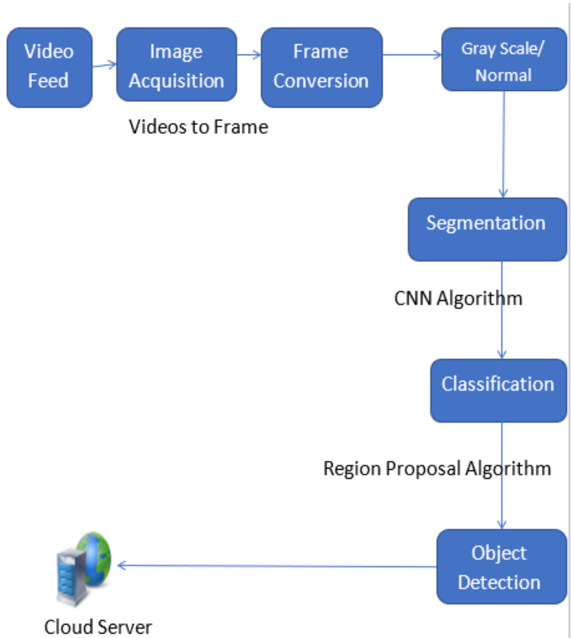



Figure 1: Object Detection and Store Procedure

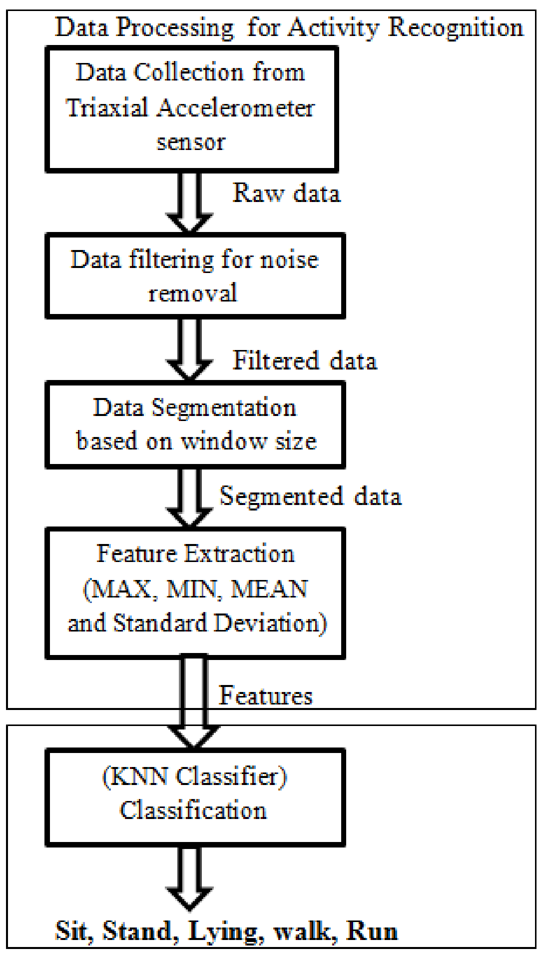



Figure 2: AnomalyNet – Data Processing and Recognition

Feature extraction and optimization networks are the key components of the designed Anomaly Net. Motion combining units and feature transfer units make up the feature extraction network. For fine coding, a proposed SLSTM is used to support the optimization network. Existing 1solvers, for example Iterative Rigid Threshold Estimation (ISTA), use a rigid and fast updating technique, updating parameters along each dimension with a non-adaptive change approach. There are some cases where this approach isn't optimal and leads to poor performance, for example, thin/large knowledge sometimes requires a subject to change each dimension to save memory and cost. '1 solvers like ISTA do not consider historical data when coming up with update rules. Tons of research in the optimization community has proven that incorporating historical information is useful for improving the convergence performance of algorithms. Predicting sparse codes in the output is very computationally expensive. For each data point, the time complexity of the sparse entry is proportional to the scale and size of the dictionary used, and because the input size.

Learning the vocabulary associated with subtle illustrations is done in an iterative way. In other words, all ancient solvers 1 could lead to an honest sparse representation, but could never produce an exciting dictionary, since sparse representation and dictionary learning are treated separately. Subtle illustrations related to the study of lexicons are carried out in a repetitive manner. In other words, all the ancient solvers 1 could lead to an honest sparse representation, but could never produce an exciting dictionary, since sparse representation and dictionary learning are treated separately.

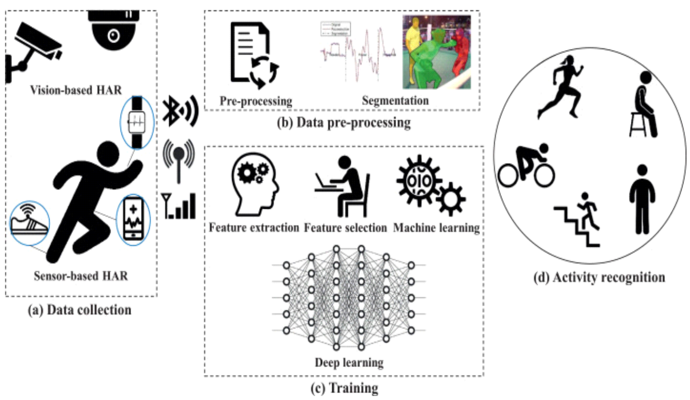



Figure 3: Vicinity Algorithm Process

Frames are analyzed using the grayscale image method, which improves the analysis process. This provides a clear visual image for the output. During this module, consider the built-in transfer of video recording. Use the inline code layer to override the image. The file uploads the video to make the stills frames for the result. Numerous image frames were collected from video feeds for the detection process by victimization and information segmentation. Metametric data is classified using the RPA algorithmic rule to obtain better data validation, which enriches the output through a detailed view.

Thus, the administrator will receive the most accurate information about video streams. By classifying the video channels, the administrator can get a higher performance of the victimization analysis of the algorithmic rule of the SVM classifier. Define proximity suggestion algorithms based on 3 criteria: repeatability, recall, and detection. The execution time of the algorithm was recorded. Repeatability is defined as the tendency of RPA to re-localize similar image content across different image propagation intervals. It predicts a situation with one thing in an image, and the same prediction will be repeated when presented with a changed image. By obtaining results in segmentation and classification, objects are detected with a higher victimization of analysis by the RPA algorithm. It is the entire output for the entire system from which the output to its core can be derived.

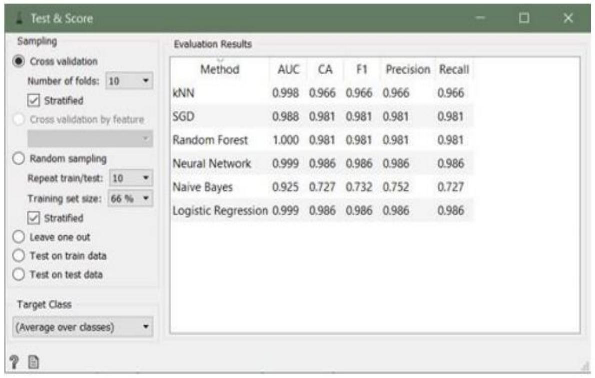



Figure 4: Data Selection Process - Tensorflow

## 4 Experimental Setup

Predictive modeling will be assembly exploitation classification, association and regression. Selection tree and association rule mining are used for tasks and categorization. Association rule mining applies to dependencies and correlations from completely different objects. Once the window is filled, the classification section can be started and the mean, minimum, maximum variance values of the information in the window are calculated and these values are compared one by one with the values in the compact training sets that were created during the previous processing step.

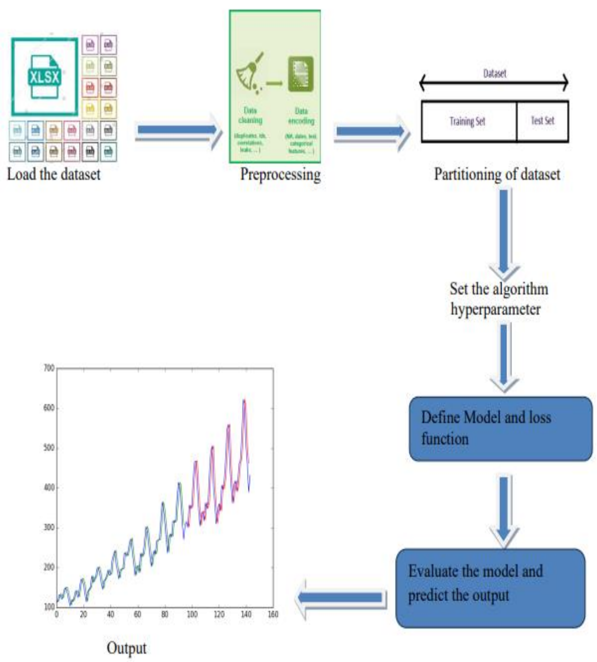



Figure 5: Result Processing Tensor Flow

The nearest test sample is selected from the training sets, and the option ends by looking at the final action list. This marks the information in the connected window because the activity that occupies the most information in the final set K. By using the standard deviation function, the activity closest to the quality deviation of this exact window is recognized as the recognized activity. The results of the options for each function will give you four shortcuts.

Mark the window as the activity for which we received the best vote and record the classification. The naïve Thomas Bayes classifier divides the data into completely different categories according to Thomas Bayes' theorem, while each predictor is independent of the other. It is assumed that the selected function in a particular category does not apply to the availability of alternative functions.

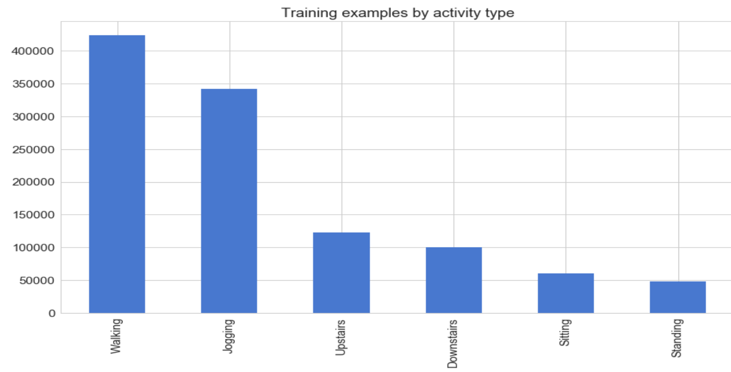



Figure 6: Training Set – Activity Selection

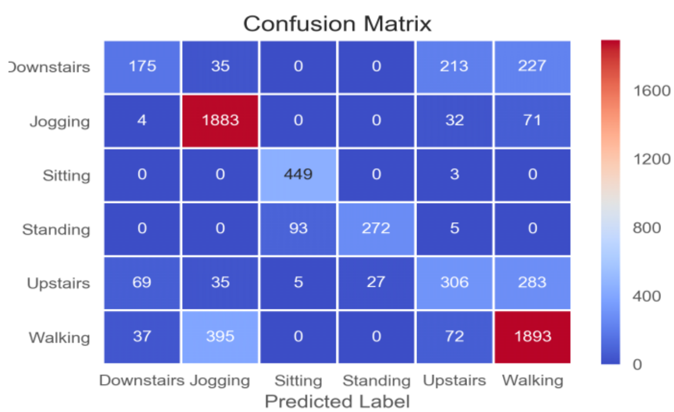



Figure 7: Confusion Matrix

This rule is kind of trendy because it can even destroy extremely advanced classification methods. Moreover, it is quite simple and you will build it in no time. Here is Bayes' theorem, which is the basis for this algorithm:

A. *Tensor Board*: Tensor Board is an interface used for record visualization, graphics and gross error detection mechanisms.

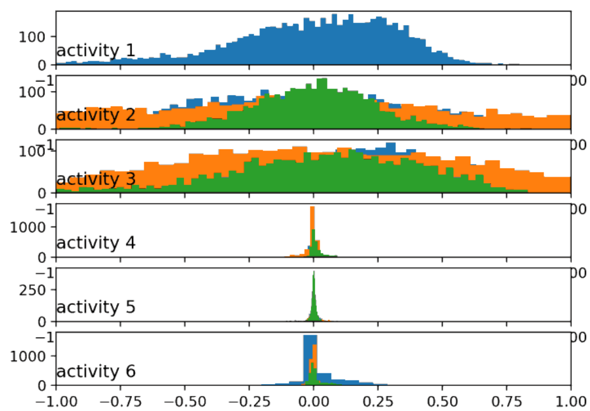
B. *Graphs*: A TensorFlow framework follows a schema and then runs, which refers to a static graph computation method. TensorFlow has a true computational graph visualization that is superior to other libraries.

C. *Enterprise-oriented*: TensorFlow is a framework with high overall performance that is superior to other frameworks available in the market. TensorFlow has a unique method that allows you to observe the learning process of neural modes and monitor many metrics. TensorFlow is a framework that has excellent network support.

D. *Customer Orientation*: The framework of TensorFlow provides Tensor Board, TensorFlow helps you perform graph subdivisions, which provides it with a higher hand, as you can start and get better discrete entries per part and find errors when using Tensor Board.

TensorFlow is extraordinarily parallel and built for a specific backend application. The TensorFlow software includes record and version parallelism so you can split a version into segments and run them in parallel. TensorFlow programs have faster compilation times with electric drive than some of the alternative frameworks like Theano etc.

*E. Delivery focus*: Libraries can be deployed on a range of hardware machines, from mobile equipment to complex computer systems.



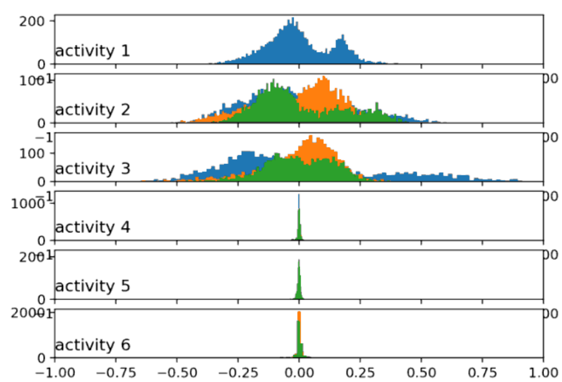





Figure 8: Final Result and Object Selection

# 5 Conclusion

The purpose of this document is to promote unified deep learning, based almost entirely on the detection of strange cases. There are three blocks in Anomaly Net that can be designed to detect anomalies in neural networks using three keys. The motion fusion unit maintains a temporal and spatial relationship between motion and view signals. It exploits the transferability of a neural community to extract discriminative features from unique tasks/domains via a feature switching block. The coding block is an LSTM that produces fast sparse coding that can experience fast output and continuous learning. Experimental results demonstrate the effectiveness of our technique in detecting anomalies during video surveillance and reconstructing photos. While development has been done in providing personalized additions to the dedicated pipeline, additional drawings are needed to similarly enhance and verify the accuracy of the mixed pipeline. In particular, the COCO dataset needs detection parameter settings for IoU Edgeboxes. Also, the CNN feature extractor needs to be improved on the under-deployment COCO dataset. A similar optimization can be achieved for CNNs by adopting the Fast-CNN framework. As an "in-the-wild" control case, a digital digital camera is primarily based entirely on infrastructure, the use of R-CNN modes proposed in this paper can be used to tune motors in a parking area. These records can be used to view parking zone occupancy situations in real time. Both assigned and unassigned R-CNN modes can be deployed and tested on a sequence of metrics along with latency and accuracy.

# References

[1] Lu, C., Shi, J., & Jia, J. (2013). Abnormal event detection at 150 fps in mat lab. *In Proceedings of the IEEE international conference on computer vision*, 2720-2727.

[2] Sabokrou, M., Fathy, M., Hoseini, M., & Klette, R. (2015). Real-time anomaly detection and localization in crowded scenes. *In Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 56-62.

[3] Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A.K., & Davis, L.S. (2016). Learning temporal regularity in video sequences. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 733-742.

[4] Zhao, B., Fei-Fei, L., & Xing, E.P. (2011). Online detection of unusual events in videos via dynamic sparse coding. *In IEEE CVPR 2011*, 3313-3320.

[5] Luo, W., Liu, W., & Gao, S. (2017). A revisit of sparse coding-based anomaly detection in stacked rnn framework. *In Proceedings of the IEEE international conference on computer vision*, 341-349.

[6] Ahamed, M.I., & Kumar, K.S. (2019). Modelling of electronic and optical properties of Cu2SnS3 quantum dots for optoelectronics applications. *Materials Science-Poland, 37,* 108-115.

[7] Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *In IEEE computer society conference on computer vision and pattern recognition (CVPR'05), 1*, 886-893.

[8] Kratz, L., & Nishino, K. (2009). Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. *In IEEE conference on computer vision and pattern recognition*, 1446-1453.

[9] Manikandan, S., Dhanalakshmi, P., Priya, S., & Teena, A.M.O. (2021). Intelligent and deep learning collaborative method for E-learning educational platform using TensorFlow. *Turkish Journal of Computer and Mathematics Education*, *12*(10), 2669-2676.

[10] Jiang, F., Yuan, J., Tsaftaris, S.A., & Katsaggelos, A.K. (2011). Anomalous video event detection using spatiotemporal context. *Computer Vision and Image Understanding*, *115*(3), 323-333.

[11] Zhang, D., Gatica-Perez, D., Bengio, S., & McCowan, I. (2005). Semi-supervised adapted hmms for unusual event detection. *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 1, 611-618.

[12] Kim, J., & Grauman, K. (2009). Observe locally, infer globally: a space-time MRF for detecting abnormal activities with incremental updates. *In IEEE conference on computer vision and pattern recognition*, 2921-2928.

[13] Medel, J.R., & Savakis, A. (2016). Anomaly detection in video using predictive convolutional long short-term memory networks. *arXiv preprint arXiv:1612.00390*.

[14] Babu, P.J.S., Padmanabhan, T.S., Ahamed, M.I., & Sivaranjani, A. (2021). Studies on copper indium selenide/Zinc sulphide semiconductor quantum dots for solar cell applications. *Chalcogenide Letters*, *18*(11), 701 – 715.

[15] Giorgi, G., Abbasi, W., & Saracino, A. (2022). Privacy-Preserving Analysis for Remote Video Anomaly Detection in Real Life Environments. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 13*(1), 112-136.