# Anomalous Detection in Noisy Image Frames using Cooperative Median Filtering and KNN

Dorcas. Esan, Pius. A Owolawi and Chunling Tu

***Abstract*—The widespread of surveillance cameras has consequently led to a significant increment in the global surveillance video market over the years. However, as the number of surveillance systems is increasing, the rate of crime threats is also increasing. Despite the advancements in surveillance systems, there are still challenges in the detection of anomalous behavioural patterns from noisy image frames, which significantly affect the accuracy of intelligent surveillance systems. This paper employs the median filtering technique and the k-nearest neighbor (KNN) classifier for the detection of behavioural patterns. Median filtering is used because of its ability to preserve the image edges while removing the noise; the statistical property of KNN is used to obtain vector distribution from images during the detection stage. Different analyses were conducted using the publicly available dataset repository that has been used by many researchers in the field of computer vision in the detection of anomalous behaviour. The results are compared with existing - state-of-anomalous detection models and the results obtained show that the proposed system outperforms the others mentioned by 85.15% and an F1-score of 0.54, equal error rate of 0.15, precision value of 0.90 and recall value of 0.40.**

***Index Terms*—Surveillance, False alarm, k-nearest neighbour, Noisy image frames, Median filtering.**

## I. Introduction

Detection of anomaly behavioural means identifying behaviors that do not conform to the normal behaviour [1]. With the increase in security threats around the world, the safety and wellbeing of people in the community raise a huge challenge [2]. Governments and many organizations have focused on installing surveillance cameras in different public locations in order to protect people's lives and valuable property [3]. Recent studies conducted by Jon Cropley, investigated and reported the statistics of the amount spent on video surveillance system from 2015 to 2019[4].

Manuscript received July 24, 2020; revised July 1, 2021.

This work was supported by the Department of Computer System Engineering. Anomalous Detection in Noisy Image Frames using Cooperative Median Filtering and KNN.

Dorcas. O. Esan is a Master Student at Tshwane University of Technology, Department of Computer Systems Engineering, Soshanguve Campus, South Africa. (e-mail: oladayojadesola10@gmail.com, Phone Number: +27677476441).

Pius. A. Owolawi is the Head of Department of Computer Systems Engineering, Tshwane University of technology, Soshanguve Campus, South Africa. (e-mail: OwolawiPA@tut.ac.za).

C. Tu is the Head of Research of Computer Systems Engineering Department and a Senior Lecturer in Computer Systems Engineering Department, Tshwane University of Technology, Soshanguve Campus, South Africa. (e-mail: duc@tut.ac.za).

The report revealed that a total of $ 14.9 billion was spent in 2015, approximately $15.1bilion was spent in 2016, $16.1bilion was spent in the year 2017, $18.2 billion was spent in 2018 and the worldwide market revenue will reach $19.9 billion in 2019 according to the IHS Markit [5]. However, with the recent spread of surveillance cameras there is still a significant increase in anomalous behaviour which often result in crimes.

The term anomalous behaviour can be defined as a behavioural pattern that does not conform to normal and regular patterns [6]. This anomalous activity is detected by installing multiple cameras at different locations. These cameras are often connected to monitors in the control room, which are constantly watched by security experts to detect any malicious activity [7]. This implies that the traditional surveillance system works with the active participation of human intervention, which consequently results in errors or omissions due to human intervention. Thus, the detection efficiency of traditional surveillance relies solely on the operator's ability to detect anomalies [8].

Traditional surveillance system detects anomalous behavioural activities with involvements of human participation who retrieved the archived video tape and analyzed the data after the occurrence of the crime [9, 10]. However, manual analyzing this huge dataset by security officers can be tiring and overwhelming [11]. However, in most cases these dataset are affected by environmental noise, which makes accurate analysis of the crime to be difficult [12], consequently leads to the loss of important information that could be used for the detection of anomalous activities by security experts [13]. To detect anomalous behavioural patterns in noisy image frames, this study utilizes the combination of median filtering and k-nearest neighbour classifier. The median filtering is used to remove unwanted noise from the image and the kNN uses the statistical properties to detection activities as either normal or anomalous.

## II. Related Works and Theoretical Background

### A. Related Works

The authors in [14] used background subtraction method to control environmental conditional changes that created noise in image frames I moving objects. The method achieved adequate and better results than other existing methods by reducing the amount of noise caused by the processing of pixels. Although the authors obtained a satisfactory result using the approach with an accuracy of 81%. The approach only worked in a static environment and its performance in dynamic environments was not studied and evaluated.

Also, the authors in [15] proposed anomalous behaviour detection in video surveillance for the classifier normal and anomalous behavioural patterns. The gaussian mixture model (GMM) was employed in analyzing the underlying data distributions. The result obtained indicated that the approach can detect anomalous behaviour correctly with 84% accuracy, and an F1-score of 0.5926 along with 8.2386 roots mean square error (RMSE). The limitation of this approach is that the approach fails in a crowded environment.

The work done in [16] presents the detection of anomalies in surveillance videos. The approach used a support vector machine and convolutional auto-encoder (CAE). CAE is used to capture the image structure and SVM is used to train to learn the normal patterns. Each sample of a testing set is classified as normal or anomalous based on the decision function learned in the trained phase. The proposed approach achieved a better result in the classification of anomalies with acceptable performance compared with other popular detection techniques with 79% accuracy. Although, the approach has low detection accuracy in a complex scene.

The contribution of Kundu et al in reference [17] introduced human detection speed to reduce the error rate in a controlled environment. The approach used the center of gravity of the image skeleton. The proposed algorithm was simple, effective, and efficient and it produced an encouraging result with an accuracy of 84%. However, the noisy environment generates numerous different images, and the algorithm was not tested on humans, which limited its performance in specific cases. In [18], automatic event detection for anomaly image analysis to confirm good image quality was introduced. The approach used a statistical technique for event detection and a Kalman filter to reduce noise that gives a false alarm. The experimental results showed that the method achieved an accuracy of 83%. However, the system has low detection accuracy in a complex environment. The summary of comparison for various anomalous detection techniques in terms of the specific technique, data used, the problem addressed, and accuracy are highlighted in Table I.

TABLE I
SUMMARY OF VARIOUS ANOMALY DETECTION TECHNIQUES

| Authors | Problem Addressed | Specific Techniques | Dataset used | Acc. (%) |
|---|---|---|---|---|
| [14] | Anomalous detection in a dynamic environment | Block-based adaptive threshold parameter. | CD net 2014 dataset | 81 |
| [15] | Anomalous detection in video surveillance to classify normal from anomalous. | Mixture Model | CAVIAR "INRIA" Dataset | 84 |
| [16] | Detection of anomalies in surveillance videos. | Support Vector Machine and Convolutional Autoencoder | Avenue dataset and UMN crime dataset | 79 |
| [17] | Human detection speed to reduce the error rate in a controlled environment. | Centre of gravity of the image skeleton | Private CCTV dataset | 84 |
| [18] | Anomalous detection in the noisy surveillance image frames. | Kalman filter | University of Houston Camera Tampering Detection dataset (UHCTD) | 83 |

With all these classical approaches used in literature, one can see from Table 1 that none address the issue of reducing the noise level in the image frame to improve the detection accuracy (i.e., maximizing the true positive rate while minimizing the false error rates). In this study, cooperative median filtering and k-NN is used to address the issue of noise in image frames and detection accuracy due to the statistical property exhibited by proposed model.

The subsequent section discusses the background theory of the new techniques. These techniques will be employed in this study.

### *B. Theoretical Technique*

The brief theorical background used for the implementation of the new technique on noisy image frames for the detection of anomalies behaviour.

### *C. Modeling of Noise*

The noise is introduced to the surveillance camera and thus gives false detection errors. There are two ways a model described the noise in digital images. These include (i) impulse noise and the (ii) additive Gaussian model. The impulse noise model is described as the noise that replaces a pixel value using a random value and this can be computed as in (1).

$$x_{ij}^{n} = \begin{cases} n_{ij}\,P \\ x_{ij}^{0}(1-P) \end{cases} \tag{1}$$

where $x_{ij}^{n}$ represent the noisy pixel$(i,j)$, $x_{ij}^{0}$represents the original pixel and $n_{ij}$is noise added to the image which is the salt-and-pepper noise, $P$ is the probability that is equivalent to the noise level in the image. This model in equation (1) described the uniform noise. The additive Gaussian noise is computed as in (2).

$$x_{ij}^{n} = x_{ij}^{0} + n_{ij} \tag{2}$$

The model in (2) describes the salt-and-pepper noise. However, since both noises are seen on both surveillance camera image frames, the combination approach of both noises is mitigated and used in this research.

### *D. Median Filtering Techniques*

The median filter is a non-linear signal-processing technique that is based on statistics [19]. The image pixel with the noise is substituted by middle value of the neighbour (mask) pixel and sorted in order of their grey values as in (3).

$$I'(x',y') = median\{g'(x'+i),(y'+j), i,j \in w'\} \tag{3}$$

where $I'(x',y')$ is the image median output and $g'(x',y')$ is the input image, and $w'$represents 2-D image mask. The mathematical analysis of median filtering is relatively difficult due to the non-linear properties of the model exhibits in dealing with noisy image. Thus, the noisy image variance is computed as in (4).

$$\sigma_{med}^{2} = \frac{1}{4nf^{2}\bar{n}} \approx \frac{\pi\sigma_{i}^{2}}{2n+\pi-2} \tag{4}$$

where σ is the image noise variance, n is the image filtered mask, and f(n) represents noisy image density function. The average filter image variance is computed as in (5).

$$\sigma_0^2 = \frac{\sigma_i^2}{n} \tag{5}$$

The effects of median filtering on image depends on the image mask size and noise distribution in the image as in (2) and (3) respectively. One of the advantages of median filtering is that it is an efficient filter and simple to implement.

### *E. k-Nearest Neighbor Technique*

The k-NN classifier is a supervised machine learning technique that trained data using posterior probability distribution function as in (6).

$$k - NN = P_{post}(H_i'|X_b') \tag{6}$$

where $H_i$ is the hypothesis combined with the observed X [20].

Then the probability density function $X'_b$ is computed as in (7).

$$P(X_b') = \frac{k_b}{N_b V_b} \tag{7}$$

where $X'_b$ represents the observation data connected to the $H'_b$, $N_b$ represents the number of behavioural patterns that are related to the hypothesized data which is represented as in (8).

$$N_b = N_1 + N_2 + \cdots + N_c \tag{8}$$

The represents the data volume around the observation data $X'_b$ containing hypothesized k[i] patterns and is calculated as in (9).

$$k = H_i'(k[1] + k[2] + \cdots . + k[c]) \tag{9}$$

Having obtained the $k$ in equation (9), the Bayes theorem is applied on equation (6) to give equation as in (10).

$$P((H_i'|X_b') = \frac{P(X_b'|H_i')P(H_i')}{P(X_b')} \tag{10}$$

Equation (10) established the fact that for any given observation $X_b'$, there should be decision to maximize any associated posterior probability. In k-NN technique, the value of k, which is the number of patterns that are included in the dataset during training is obtained around the observed data point using Euclidean distance to measure the minimum distance between a test query sample and a set of training data samples that stores in the database as given in (11).

$$d(x, y) = \sum_{i=1}^{N} \sqrt{x_i^2 - y_i^2} \tag{11}$$

To determine the value $k$, the distances in the entire space are sorted in ascending order and the detection of the query sample is done using the majority vote of the k-NN. The k-NN classifier is used in this research owing to the statistical property it exhibits that can be used to obtain vector distribution from images in the detection stage [21, 22].

### *F. Common Red Flag Indicators used in the Detection of Anomalous Behavioral Patterns*

The This section describes common anomalous indicators identified and used in the implementation of the proposed detection security system [23]. The common red flag indicators of anomalous behavioural patterns used in this research implementation are shown in Table II.

TABLE II
ANOMALOUS DETECTION INDICATORS

| S/N | Indicators | Meaning |
|---|---|---|
| 1 | Pocket Picking | In a scenario that involves multiple people, where two persons bump into each other, the third person takes the wallet from the victim and delivers the wallet to the fourth person. |
| *2* | Snatching | This involves a situation where a person grabs another person properties and runs away. |
| *3* | Pursuing | This is the situation where someone starts running or following another person with intention of knowing about his/her movements. |
| *4* | Loitering | This occurs when an individual is seen in certain surveillance areas and staying too long in a place without any obvious movement [24]. |
| *5* | Running | This occurs when an individual is running where other people usually walk. |
| 6 | Theft | This is another red flag indicator, and it often occurs when the perpetrator goes up and circles the object, then retreats before making a move [25] |
| 7 | Cycling or moving of vehicles | This can be another indicator that occurs when an individual is riding a bicycle or driving a vehicle in a vehicle or cyclists restricted areas such as pedestrian walkways [24]. |

These red-flag indicators in Table II are used for the implementation of anomalous behavioural pattern detection on image frames affected by environmental noise using the new technique. The next section discusses the methodologies used for the implementation of the cooperative median filtering and k-NN model.

## III. METHODOLOGY

This section explains the step-by-step methods used for the development of the new technique for the detection of anomalous behavioural patterns in the crowded environment.

### *A. Image Acquisition Subsystem*

The University of California San Diego (UCSD) dataset repository is utilized in this experiment. These datasets contain videos from different static cameras overlooking a pedestrian walkway each frame has 238 × 158 pixels, and the crowd density is sometimes high to the point of causing severe occlusions. The data consist of different anomalous activities combined with normal activities. The types of anomalies present are “cycling”, “skater”, “cart”,

"wheelchair", "walk across", and "other". The result is fed into another subsystem for further processing. The experimental implementation of all other subsystems is done in the Python environment using some Python libraries

### B. Image Pre-processing Subsystems

The acquired image data set is passed into the image pre-processing stage. At this stage, the image is enhanced, and features are extracted for further processing. The next sections explain the components of the image pre-processing stages further.

### C. Background Subtraction

In this research, the image is modelled from the sequence of image frames using the median modelling approach to extract the foreground image pixels from the background. This is done by estimation of the background image at the time I (x,y,t). The estimated background, which is the previous image frames, is then subtracted from the input image I(x,y,t-1). Then the background subtraction is computed as in (12).

$$B(x,y,t) = \left(I_{(t-1)}(x_{t-1},y_{t-1}) - I_t(x_t,y_t)\right) > Thr \quad (12)$$

where Thr is the selected threshold value, and this value is chosen dynamically to adapt to changes in the environment for each coming image frame. The background image is updated as in (13).

$$I_{t+1} = \begin{cases} I_t(x_t,y_t) > B(x,y,t), & foreground \\ otherwise & \\ I_t(x_t,y_t) < B(x,y,t), & background \end{cases} \quad (13)$$

### D. Noise Removal

After the extraction of the foreground in the image, there is still some discrepancy in the image that is noticed in the foreground, such as a motion region as well as noise. Hence, this noise is removed from the detected foreground by using a median filter as in (3).

### E. Morphological Operation Subsystem

Because of the composition of the environment, isolating the noise comes along with the composite of the environment, therefore further processing is done by morphological processing (dilation) to remove isolated spots on the detected foreground image and interference from the image while still preserving the shape of human motion, as in (14).

$$X \Rightarrow A \oplus B = \{Z \in B | \rightleftarrows (B^s)_z \cap \neq \varphi\} \quad (14)$$

where $A$ is a set structuring element $B$ that reflects the origin of the image structure, which is shifted by Z. $B^s$ denotes the symmetric of *B*.

This process helps to remove the broken edges in the image. The output of dilated image is fed into detection subsystem for further processing.

### F. Detection Subsystem

Here, the anomalous behavioural detection is done by modelling the temporal context of the behavioural pattern using past observations of the behaviour and the anomalous interaction temporal neighbour set. The *K* spatial nearest neighbour of the behavioural pattern of the image vector at a time (t) are searched from its spatial neighbour vector at a time (t+1) and the distance between the vectors is computed using the Euclidean distance as in (11). The k distance values are sorted in ascending order. Having obtained the distance between the image vector at time $X_{(t)}$ and time $X_{(t+1)}$, the mean is computed as in (15).

$$\mu = \frac{\sum_{j=0}^{i=0} X_{(t)} - X_{(t+1)}}{N} \quad (15)$$

For the detection, the difference between the image vector at a time (t) and the image vector at a time (t+1) is summed up to give the mean value, which is set as an assigned threshold value in this study. Thus, the image is considered anomalous or normal if it satisfies the condition as in (16).

$$d \approx X_t - X_{t+1} = \begin{cases} anomalous & d < \mu \\ normal & otherwise \end{cases} \quad (16)$$

Equation (16) implies that if the similarity distance between the image vector at time (t) and time (t+1) is greater than the assigned threshold value, then such an image is considered anomalous, otherwise, it is considered normal. The pseudo-code for the proposed technique for the detection of an anomalous behavioural pattern in an image frame is explained in the subsequent section.

### G. Algorithmic Analysis

The acquired image frame at a time (t) and its corresponding image frame (t+1) is divided into blocks *M *N*. The background subtraction subsystem is done on the partitioned image frames to extract the image foreground regions. The extracted foregrounds are fed into the median filtering to remove any unwanted environmental noise in the image frames. The two enhanced image frames are further passed into the morphological (dilation) operation, where the whole and broken edges in the images are removed.

The dilated images are fed into the image entropy to obtain the vectors from the images at times (t) and (t+1) respectively. The extracted vectors are passed into the classifier (k-NN). At the k-NN stage, the similarity distance between the vector of an image at a time (t) and the image at a time (t+1) is computed using Euclidean distances. The calculated Euclidean distances between the two frames are sorted in non-decreasing order. The $k$-value is taken from the short-listed distances. These distances between all feature vectors are computed and their mean values are computed. This mean value is used as a specified threshold value for the detection of anomalous behavioural patterns in image frames. That means that if the width between the image vector at a time (t) and the neighboring vector at a time (t+1) is greater than or equal to the mean value (μ) the image frame is considered anomalous, otherwise it is considered normal. This is explained in algorithm 1.

**Algorithm 1:** Anomalous detection using proposed model.
Input image *I*, $D= \{(x_1, c_1) \ldots, (x_n, c_n)\}$,
K=number of nearest neighbors
Output: detected class
1. divide an image $X_{(T)}$ and $X_{(t+1)}$ into N×N sliding windows.
2. for each slide k
3. compute the variance $var^k D$ and $var^k T$
4. endfor
5. calculate the mean $\mu_T$ of the image $X_T = T_p = t_r \times \mu_T$
6. for each slide k in $X_T$
7. while $(r(k) \leq T_P)$
8. $T_{final} = T_{global} + T_{batch} / 2$
9. if $(|T_{final} - T_{global}| \leq 1)$
10. consider pixel as background
11.else
12. temp = $T_{final}$
13||median filtering
14. allocate =outputPixelValue [image width* window height]
15. edgex = width / 2
16. edgey = height / 2
17. for x from edge x to image width – edgex **do**
18. for $y$ from edgy to image width - edgey **do**
19. I = 0
20 for fx from 0 to width **do**
21. for fy from 0 to height **do**
22. window[i]=inputPixelValue $[x + f_x - edgex] [y + f_y - edgey]$
23. i=i+1
24. sort entries in a window[ ]
25. return
26 outputpixelvalues[$x$][$y$] = window [width×height /2]
27. endfor
28. endfor
29. ||classification of image
30. if A← {}
31. for I = 1:m **do**
32. computed $(X_{(t)}, X_{(t+1)})$
33. endfor
34. $A \leftarrow A \cup \{x, c_1, D\}$
35. endfor
36. sort in order
37. $C_y \leftarrow$ themostfrequentclass
38. || Detection stage
39. compute the mean (μ) value for the most frequent class
40. if $X_{(t)} - X_{(t+1)} \geq \mu$ considered image as anomalous.
41. otherwise, considered as normal.
42. endfor
43. endif
43. endif
44. return

Algorithm 1. Pseudocode for proposed technique for detection of anomalies in the image frame

### H. Performance Evaluation Mechanism

In this section, the performance evaluation of the new technique is studied through qualitative and quantitative measure. During the quantitative inspection, the quality of the noisy pixel value is compared with those of the filtered image frames. The following evaluation techniques were used as quantitative scoring scheme MSE, RMSE and confusion matrix. These are further explained in subsequent sections.

*[i] Mean Squared Error (MSE):* This is the squared of original image intensity minus the squared of filtered image intensity. This is mathematically computed as in (17).

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad (17)$$

where $y_i$ is the noisy image and $\hat{y}_i$ the filtered image of size *M*N*.

*[ii] Root Means Square (RMSE):* RMSE is defined as the square root of mean square error this is computed as in (18).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N}} \quad (18)$$

where $y_i$ is the noisy image and $\hat{y}_i$ the filtered image of size *N*. The reduced score for both MSE and RMSE for the noise levels are expected.

*[iii] Cross-validation*: The Cross-validation of 90% of dataset for training and 10% for testing is used for performance evaluation of the proposed technique.

*[iv] Confusion Matrix:* This is a tabular representation of behavioural pattern instances that are correctly detected and those that are wrongly detected, and this is illustrated as in Table III.

TABLE III
CONFUSION MATRIX FOR THE DETECTION OF ANOMALIES IN A NOISY IMAGE

| Detected | Normal | Anomalous |
|---|---|---|
| Normal | *P* | *Q* |
| Anomalous | *R* | *S* |

where P represents true positive (TP), Q represents true negative (TN), R represents false positive (FP), and S is false negative (FN). These terms are explained in (19) - (23).

*[v] Precision ($P_r$):* This can be defined as true positive (P) divided by addition of true positive (P) and false positive (R) detected by the model [26]. This is computed as in (19).

$$P_r = \frac{P}{P + R} \quad (19)$$

*[vi] Recall* ($R_e$)*:* This is the true positive (P) divided by addition of the true positive (P) and false negative (Q), which can be calculated as in (20).

$$R_e = \frac{P}{P + Q} \quad (20)$$

*[vii] F1-Score:* This is multiplication of Recall ($R_e$) and Precision ($P_r$) divided by addition of recall and precision with all multiplied by two [26]. This is illustrated as in (21).

$$f1 - score = 2 \times \frac{P_r \times R_e}{P_r \times R_e} \quad (21)$$

*[viii] Equal Error Rate (ERR):* This is the sum of false negative (Q) and false positive (R) divided by total sum of all number of behavioural patterns used by the model, as in (22).

$$ERR = \frac{Q + R}{Q + R + P + S} \quad (22)$$

*[ix] Accuracy:* This is the addition of true positive (P) and true negative (S) divided by sum of all behavioural patterns used by the model, which is calculated as in (23).

$$Acc = \frac{P+S}{P+S+Q+R} \quad (23)$$

*[x]Area Under the Curve (AUC):* This is calculated for the Sensitivity- Specificity curve and detection rate.

*[xi] Receiver Operating Characteristic (ROC) Curve:* This is the plot of graph of behavioural patterns that are normal which model correctly detected as normal (TPR) against behavioural patterns that are false and correctly detected by the model as false (FPR) with different thresholds.
The proposed model performance on detection of anomalous in crowded environment is evaluated using the performance metrics in (19) - (23) respectively.

## IV. ANALYSIS OF RESULT

Here the results analysis of the new technique on anomalous behavioural detection is discussed. This is shown in sub-sections A and B.

### *A. Results and Analysis 1: Quantitative Evaluation of Proposed Techniques on Anomalous Detection as in Cycling*

The objective of this section is to evaluate the performance of the new technique on noisy image frames. Fig 1(a) contains a 10% noise level on a normal behavioural pattern image. One can see that the noise introduced in Fig 1(a) is light compared to Fig 1(e) with a 50% noise level on anomalous behavioural pattern such as cycling as shown in Fig 1.

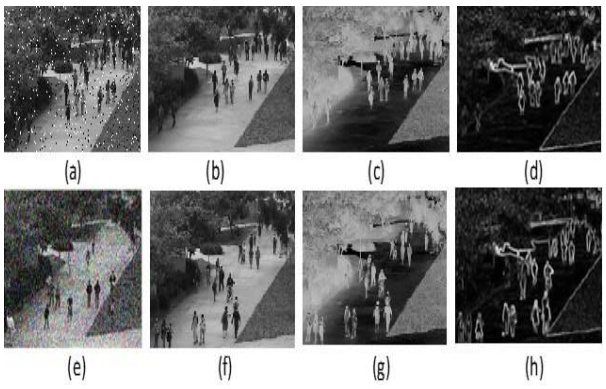


Fig 1. Quantitative evaluation of image; (a)-(d) are the light noisy ground truth image with the enhanced image using median filtering.

Figs 1(a) and 1(e) show the normal and anomalous image affected by environmental noise; Figs with the enhanced image using median filtering; Figs 1(e)-(f) show heavily noisy anomalous behaviour Figs 1(b) and 2(f) represent the output of background subtraction; Figs 1(c) and 1(g) show the result of the enhanced noisy image median filtering technique and finally Figs 1(d) and 1(h) show the result of the morphological operation. From Table IV, one can observe the performance of new technique on the noise level which ranges between 10% and 50%.

TABLE IV
COMPARING THE NOISY IMAGE AND THE FILTERED IMAGE BASED ON THE NOISE LEVELS

| Noisy Image | | | Median Filtered | |
|---|---|---|---|---|
| Noise Level (%) | MSE | RMSE | MSE | RMSE |
| 10 | 14.17 | 31.92 | 8.09 | 20.2 |
| 20 | 18.39 | 35.45 | 12.29 | 26.29 |
| 30 | 23.09 | 41.21 | 16.94 | 28.94 |
| 40 | 26.61 | 46.61 | 22.24 | 30.24 |
| 50 | 30.88 | 48.88 | 26.31 | 32.31 |

Having compared the noisy ground truth (original) with filtered image frames, the result of MSE and RMSE are shown in Figs 2 and 3, respectively.

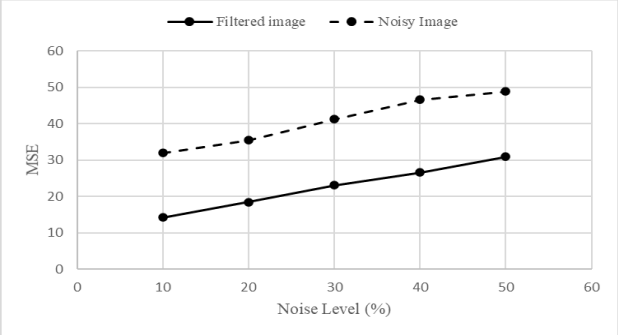


Fig 2. MSE at various levels for the filtered and Noisy Image

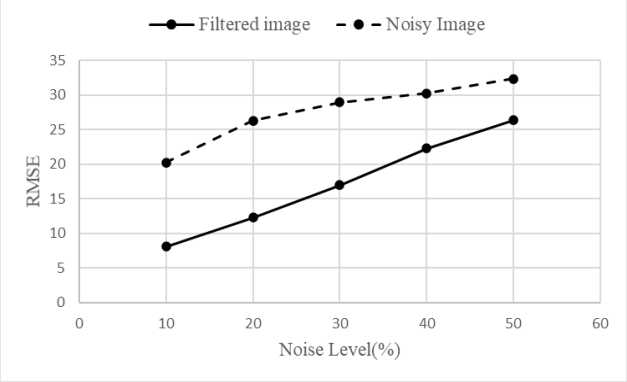


Fig 3. RMSE at various levels for the filtered and Noisy Image.

From Figs 2 and 3 one can see that the trend of the median filter result minimizes the noise level significantly compared to the noisy image. This serves as evidence for the lower scoring results are better for the MSE and RMSE. To further affirm the performance of new technique, a cross-validation technique of 90 % for the training dataset and the remaining 10% for the testing was conducted. The result is shown in Table V.

TABLE V
CONFUSION MATRIX

| Detected | Normal | Anomalous |
|---|---|---|
| Normal | TP = 115 | FN = 285 |
| Anomalous | FP = 20 | TN = 1280 |

From Table V, one can see that the FP is reducing while the FN is increasing, and the true positive rate value from the confusion table gives 0.82, the F1-score that measures model accuracy that combines precision and recall gives 0.4. Finally, an accuracy of 82% was achieved by the proposed model. Furthermore, precision, recall, F1-score, AUC, equal error rate, and accuracy are computed to validate the

performance of the new model. The results of this are shown in Table VI.

TABLE VI
DETECTION OF ANOMALOUS WITH OTHER PERFORMANCE METRICS

| Metrics | Performance (%) |
|---|---|
| F1 Score | 0.4 |
| ERR | 0.17 |
| Precision | 0.85 |
| Recall | 0.28 |
| AUC | 0.728 |
| Accuracy | 82 |

Table VI represents different performance evaluation metrics used on the model for detection of anomalous on selected image frames from the dataset, it is found that the accuracy of the model is 82%, recall or sensitivity which indicates the correctly detected portion of actual anomalous behavioural patterns of 0.28 and precision is 0.85 was achieved. The F1-score of 0.4, AUC of 0.728, and EER of 0.17 were obtained from the new technique.

To interpret the result presented in Table VI for the anomalous detection system results, the corresponding ROC curves which is the trade-off graph of the true positive rate against the false-positive rates with varied thresholds for the new model is shown in Fig 4.

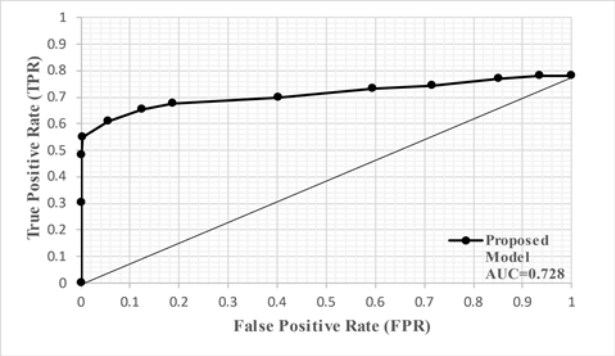


Fig 4. ROC curve for anomalous behaviour such as cycling using proposed technique.

From Fig 4, it can be observed that the new model produced high true positive rates with low false-positive rates. This implies the higher the true positive values the lower the false alarms, demonstrating the effectiveness of cooperative median filtering and KNN for anomalous detection on noisy image frames.

### *B. Results and Analysis 2: Quantitative Evaluation of Proposed Technique on Anomalous Detection of Moving Vehicle*

Here the performance evaluation of the new technique on noisy image frames is presented. Fig56(a) contains a 10% noise level on a normal behavioural pattern image. One can see that the noise introduced in Fig 5(a) is light compared to Fig 5(e) with a 50% noise level. Quantitative analysis of the proposed technique on a noisy normal behavioural pattern and anomalous behavioural pattern such as moving vehicle is shown in Fig 5.

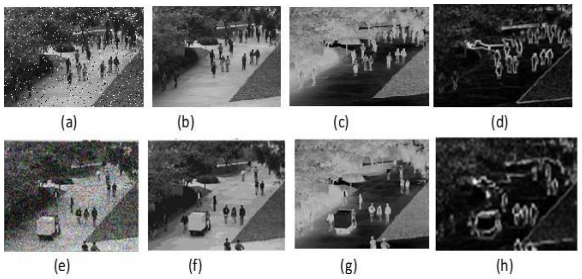


Fig 5. Quantitative evaluation of image; (a)-(d) are the noisy ground truth image with the enhanced image using median filtering; (e)-(f) reflect noisy anomalous behaviour with the enhanced image using median filtering.

Figs 5(a) and 5(e) show the normal and anomalous image affected by environmental noise; Figs 5(b) and 5(f) represent the output of background subtraction; Figs 6(c) and 5(g) show the result of the enhanced noisy image median filtering technique and finally Figs 5(d) and 5(h) show the result of the morphological operation.

From the result in Table VII, one can specifically access the qualitative performance of proposed approach for the noise level which ranges between 10% and 50%.

TABLE VII
COMPARING THE NOISY IMAGE AND THE FILTERED IMAGE BASED ON THE NOISE LEVELS.

| Noisy Image | | | Median Filtered | |
|---|---|---|---|---|
| Noise Level (%) | MSE | RMSE | MSE | RMSE |
| 10 | 15.17 | 32.52 | 9.11 | 25.24 |
| 20 | 19.3 | 38.45 | 13.69 | 29.28 |
| 30 | 23.11 | 43.22 | 18.84 | 34.9 |
| 40 | 27.71 | 46.91 | 24.27 | 38.14 |
| 50 | 29.88 | 49.82 | 26.71 | 40.11 |

Having compared the noisy ground truth (original) with filtered image frames, the result of MSE and RMSE are shown in Figs 6 and 7, respectively.

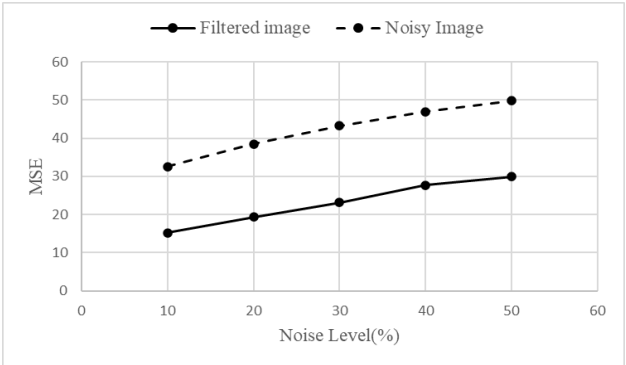


Fig 6. MSE at various levels for the filtered and Noisy Image.

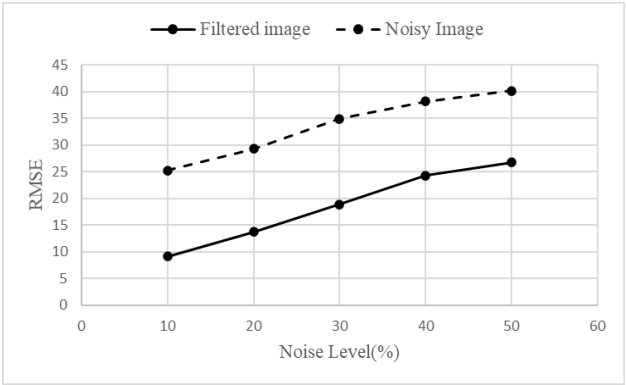


Fig 7. RMSE at various levels for the filtered and Noisy Image.

From Figs 6 and 7, one can see that the trend of the median filter result minimizes the noise level significantly compared to the noisy image. This serves as evidence for the lower scoring results are better for the MSE and RMSE.

To further affirm the performance of proposed technique, the cross-validation which utilizes training data of 90% and testing data of 10%. The result of experiment conducted is computed in Table VIII.

TABLE VIII
CONFUSION MATRIX

| Detected | Normal | Anomalous |
|---|---|---|
| Normal | TP = 203 | FN =12 |
| Anomalous | FP = 197 | TN = 1288 |

From Table VIII, the F1-score of 0.66 and accuracy of 88.3% was achieved by the proposed model. To further validate the performance of the proposed model, other performance evaluation metrics were used such as precision, recall, F1-score, AUC, equal error rate and accuracy. The results are shown in Table IX.

TABLE IX
ANOMALOUS DETECTION WITH OTHER PERFORMANCE METRICS

| Metrics | Performance (%) |
|---|---|
| F1 Score | 0.66 |
| ERR | 0.12 |
| Precision | 0.94 |
| Recall | 0.51 |
| AUC | 0.814 |
| Accuracy | 88.3 |

Table IX represents the different performance evaluation metrics used on the model for detection of anomalous image frames from the dataset, it is found that the accuracy of the model is 88.3%, recall which indicates the correctly detected portion of actual anomalous behavioural patterns of 0.51 and precision is 0.94 was achieved. The F1-score of the model is found to be 0.66 while the AUC of 0.814 and EER of 0.12 are produced by the model. To interpret the result presented in Table IX for the anomalous detection system results, the corresponding ROC curves which is the trade-off graph of the true positive rate against the false-positive rates with varied thresholds for the proposed model is shown in Fig 9.

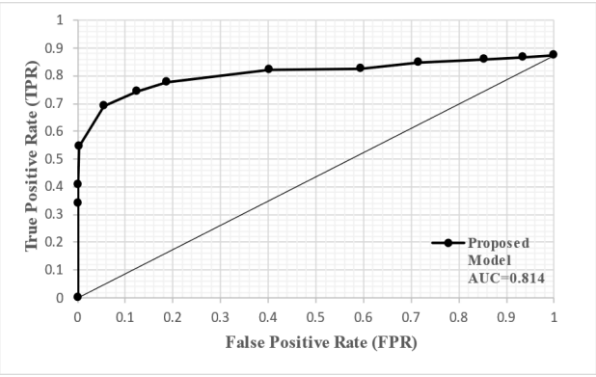


Fig 8. ROC curve for anomalous behaviour on the image with a moving vehicle using proposed technique.

From Fig 8, one can see that the high TPR was produced with low FPR, proving that the proposed approach yields a high correct detection with low false alarms. Fig 8 also demonstrates the performance of new model on the detection of an anomalous behavioural pattern on image frames affected by environmental noise. This shows that the output of the proposed technique can help security analysts to locate behavioural patterns quickly where possible anomalies may have occurred. Security personnel no longer need to search through all behavioural patterns looking for anomalies. The average accuracies of the two analyses are given in Table X.

TABLE X
PERFORMANCE OF THE PROPOSED APPROACH IN TWO ANALYSES

| Analysis | Accuracy (%) | F1-score | ERR | Precision | Recall |
|---|---|---|---|---|---|
| Analysis 1 | 82 | 0.4 | 0.17 | 0.85 | 0.28 |
| Analysis 2 | 88.3 | 0.66 | 0.12 | 0.94 | 0.51 |
| Average Accuracy | 85.15 | 0.54 | 0.15 | 0.90 | 0.40 |

Table X shows that the overall performance accuracy of the proposed technique in the two analyses was 85.15, overall equal error rate of 0.15, the precision of 0.90, recall value of 0.40 and the overall value of the F1-score was 0.54. This result is a clear indication that the new approach can be used effectively for the detection of anomalous behavioural patterns in a noisy environment.

### *C. Comparison of New Model with Other Popular Existing Models Using Accuracies and F1-Scores*

Also, the comparison of performance accuracy of the proposed method with the other three models from the experiment conducted on the UCSD dataset is as shown in Fig 9.

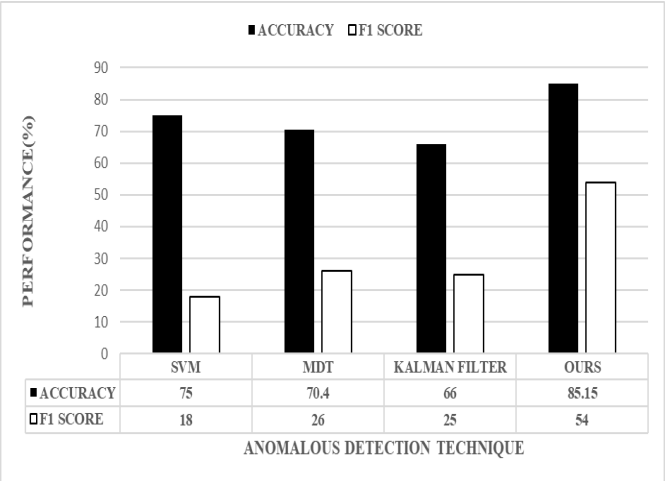


Fig 9. Graph Showing the Performance Comparison of the proposed model with other popular detection models.

From Fig 9, the proposed model gives a performance accuracy of 85.15% with an F1-score of 0.54% on the UCSD dataset for detection of anomalous behavioural patterns, the SVM technique is the runner-up with correct accuracy of 75% and F1-score18%, the Kalman filter gives correct accuracy of 66% and F1-score of 0.25% and the MDT produced correct accuracy of 70.4% and F1-score of 0.26%. From this graph, it obvious that the proposed model outperformed the three conventional contender models in the UCSD dataset for the detection of anomalous behavioural patterns. This is due to its capabilities in the training of the data with optimal parameter values of the model.

*D. Benchmarking Proposed Technique with Other Existing Techniques in Literatures*

The objective is to benchmark some favorable advantages of new technique against popularly used anomalous behavioural detection techniques on UCSD datasets to affirm the performance of the new model. Table XI presents a comparison of existing techniques with the proposed technique in terms of the methods, year, F1-score, the dataset used, and accuracy obtained from each method.

The literature cited in Table XI presents different methodologies for the detection of anomalous behavioural patterns compared with the new method. Research in [27]used a social force flow approach for detection of anomalous behavioural patterns in image frames. This is done by placing grid on the image. The authors performed the detection process on the image by utilizing bag of word approach. Although, the accuracy of the approach under the influence of environmental noise was not investigated.

Also, the mixture of dynamic texture-Spatial (MDT-Spat) as presented in [28] to detect anomalous activities in image frames and to optimize the detection accuracy of the anomalous model.

TABLE XI
PERFORMANCE COMPARISON OF THE PROPOSED MODEL WITH OTHER EXISTING TECHNIQUES

| Model | Year | F1 Scores | Dataset used | Accuracy (%) |
|---|---|---|---|---|
| Social Force Flow[27]. | 2009 | - | UCSD (ped 1) | 78.9 |
| Mixture of Dynamic Texture-Spatial (MDT-Spat)[28] | 2013 | - | UCSD (ped 1) | 71.3 |
| Non parametric [29] | 2011 | - | UCSD (ped1) | 66 |
| Mixture of Optical Flow model [30] | 2018 | - | UCSD (ped 1) | 77 |
| deep-learning [31] | 2018 | | UCSD (ped 1) | 78.1 |
| Proposed model | 2020 | 0.54 | UCSD (ped 1) | 85.15 |

These approaches have shown relatively high accuracy but have not been used for the detection of anomalous behaviour in crowded environments such as university environment. Research in [29] proposed an approach for the detection of an anomaly in real-time using Spatio-temporal features that capture scene dynamic statistics together with appearance. Detection of an anomaly in real-time was performed with an unsupervised approach using nonparametric modelling, evaluating directly multi-scale local descriptor statistics. The result in Table XI shows that the proposed technique has an accuracy which is superior to other popular anomalous detection techniques on publicly available anomalous behavioural pattern dataset (UCSD ped1). The result of this new technique is a piece of evidence that the technique can be used effectively by security personnel in the detection of anomalous behavioural patterns in a crowded environment.

## V Conclusion

This work has demonstrated the proof of concept about the ability of proposed technique. The new model is investigated as a means of detecting an anomalous behavioural pattern in image frames in the dynamic environment due to their non-parametric properties. In this method, the k-NN algorithm classified the image feature vectors as either normal or anomalous. Qualitative and quantitative experimental analysis was conducted on a cyclist and a moving vehicle on the UCSD pedestrian 1 dataset using the proposed technique for anomalous detection. The median filtering technique was utilized in removing image noise, as shown in Figs 2 and 6, while k-NN, which is a non-parametric classifier technique, was used for classification.

The average detection performance of proposed technique is shown in Table X with accuracy of 85.15, F1-score of 0.54, EER of 0.15, precision of 0.90 and recall of 0.40 were achieved. However, when using proposed model, one can observe from the results that the anomalous regions in the noisy image frames were identified when enhanced with median filtering and the detection classification was done correctly with the k-NN classifier. The outcome of this research can be a useful application to assist the security personnel to prevent crime in a dynamic environment. However, the current detection system for suspicious behaviour can still be improved upon by using more data-mining algorithms to predict the link between behavioural patterns in multivariate time series datasets in crowded environments.

## References


[1] A. A. Sodeman and M. P. Ross, "A Review of Abnormality Detection in Automated Surveillance," *IEEE Trans. Syst.Man Cybern*, no. 1, pp. 257–1272, 2012.

[2] P. Lv, S. Liu, M. Xu, and B. Zhou, "Abnormal Event Detection and Location for Dense Crowds Using Repulsive Forces and Sparse Reconstruction " *IEEE Transactions on Image Processing*, 2016.

[3] R. Collins, A. Lipton, and T. Kanade, "Introduction to the special section on video surveillance," *IEEE Transactions on PAMI*, pp. 745–746, 2015.

[4] J. Cropley, "Global Professional Video Surveillance Equipment Market Set For Third Year of Near Double-Digit Growth In 2019."

[5] I. M. s. V. S. I. Service, "Global Revenue Forecast for Professional Video Surveillance Equipment Market," 2019.

[6] E. Varghese, J. Mulerikkal, and A. Mathew, "Video Anomaly Detection in Confined Areas," *International Confrence on Advances in Computing and Communications,ICACC*, pp. 22-24, 2017.

[7] D. O. Esan, P. A. Owolawi, and C.Tu, "Anomalous Detection System in Crowded Environment using Deep Learning," *2020 International Conference on Computational Science and Computational Intelligence*, pp. 29-35, 2020.

[8] V. A.Kotkar and V.Sucharita, "A Comparative Analysis Of Machine Learning-Based Anomaly Detection Techniques In Video Surveillance," *Journal Of Engineering And Applied Sciences*, pp. 9376-9381, 2017.

[9] S. V. Rajenderan and K. F. Thang, "Real-Time Detection of Suspicious Human Movement " *international Conference on Electrical Electronics Computer Engineering and their Applications*, 2014.

[10] D. O. Esan, P. A. Owolawi, and C. Tu, "Detection of Anomalous Behavioural Patterns In University Environment Using CNN-LSTM," in *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, South A, 2020, pp. 1-8, doi: 10.23919/FUSION45008.2020.9190406.

[11] K. C. Baumgartner, S. Ferrari, and C. G. Salfati, "Bayesian Network Modelling of Offender Behaviour for Criminal," *Master of Science, Department of Mechanical Engineering and Material science, Duke University*, 2005.

[12] T. Zhang, Y. Y. Tang, Z. Shang, and X. Liu, "Face Recognition Under Varying Illumination Using Gradientfaces," *IEEE Transactions On Image Processing*, vol. 18, pp. 2599 – 2606, 2009

[13] M.Sankari and C.Meera, "Estimating of Dynamic Background and Object Detection in Noisy Visual Surveillance " *International Journal of Advanced Computer Science and Applications(IJASCSA)*, vol. 2, pp. 77-83, 2011.

[14] M. F. Savas, H. Demirel, and B. Erkal, "Moving Object Detection Using An Adaptive Background Subtraction Method Based On Block-Based Structure In A Dynamic Scene," *IEEE*, 2018.

[15] R. R. Sillito and R. B. Fisher, "Semi-Supervised Learning For Anomalous Trajectory Detection," 2016.

[16] M. Giuhoski, N. Marcelo, R. Aquino, M. Ribeiro, A. E. Lazzaretti, and H. S. Lopes, "Detection Of Video Anomalies Using Convolutional Autoencoder And One-Class Support Vector Machine," presented at the Brasillero Of Intelligence Computational 2017.

[17] M. Kundu, D. Sengupta, and J. G. Destidar, "Tracking Direction Of Human Movement An-Efficient Implementation Using Skeleton," *IEEE*, 2019.

[18] C. Wang, X. Wu, N. Li, and Y.-L. Chen, "Abnormal Detection Based On Gait Analysis," *Proceedings Of The 10th World Congress On Intelligent Control And Automation*, 2012.

[19] Y. Zhu and C. Huang, "An Improved Median Filtering Algorithm for Image Noise Reduction," presented at the 2012 International Conference on SOlid State Device and Materials Science, 2012.

[20] R. Gil-Pita and X. Yao, "Evolving Edited K-Nearest Nneighbor Classifiers," *International Journal of Neural Systems*, vol. 18, p. 6, 2008.

[21] Z. Al-asady and A. Al-amery, "Human Action Recognition using a Corners and Blob Detector with different Classification Methods," *International Confrence on Sustainable Engineering Techniques, ICSET*, 2019.

[22] S.-R. Ke, H. L. U. Thuc, Y.-J. Lee, J.-N. Hwang, J.-H. Yoo, and K.-H. Choi, "A Review on Video-Based Human Activity Recognition," *Computers*, vol. 2, pp. 88-131, 2013.

[23] S. Chaudhary, M. A. Khan, and C. Bhatnagar, "Multiple Anomalous Activity Detection in videos " *6th International Conference on Smart Computing and Communication*, pp. 336-345, 2018.

[24] P.Kamala, R. RS, and Manjula.P, "Automated Intelligent Surveillance using Human Behaviour Analysis in Shopping Malls," *International Journal of Computer Science and Information Technologies*, vol. 6, no. 5, pp. 4392-4396, 2015.

[25] N. B. Gadhe, B. K. Lande, and B. B. Meshram, "Intelligent System for Detecting, Modeling and Classification of Human Behaviour using Image Processing, Machine Vision and OpenCV," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 1, pp. 590-599, 2012.

[26] P. Chouhan and V. Richhariya, "Anomaly Detection in Network using Genetic Algorithm and Support Vector Machine " *international Journal Computer science and Information technologies* vol. 6, pp. 4429-4433, 2015.

[27] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behaviour detection using social force model," *IEEE Xplore*, 2009.

[28] A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz, "Robust Real-Time Unusual Event Detection using Multiple Fixed-Location Monitors," *IEEE Transactions on Image Processing*, 2008.

[29] M. Bertini, A. D. Bimbo, and L. Seidenari, "Multi-scale and real-time non-parametric approach for anomaly detection and localization," *Computer Vision and Image Understanding*, vol. 116, no. 3, pp. 320-329, 2011.

[30] K. G. Gunale and P. Mukherji, "Deep Learning with a Spatiotemporal Descriptor of Appearance and Motion Estimation for Video Anomaly Detection," *Journal of Imaging*, vol. 4, no. 79, pp. 1-17, 2018.

[31] F. Turchini, L. Seidenari, T. Uriccgio, and A. D. Bimbo, "Deep-Learning Based Surveillance System for Open Critical Areas," *Inventions*, vol. 3, no. 69, pp. 1-13, 2018.