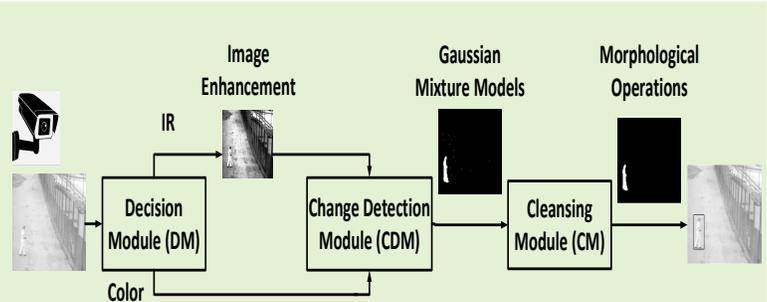


# Improved Change Detector using Dual-Camera Sensors for Intelligent Surveillance Systems

Ajmal Shahbaz , *Student member, IEEE* and Kang-Hyun Jo, *Senior member, IEEE*

**Abstract**—Unauthorized entrance in a prohibited area might create a security risk. An intelligent surveillance system should be able to mitigate such a problem by incorporating a sterile zone monitoring algorithm. The algorithm is challenged by a dual-camera sensors (color/IR), dynamic backgrounds, illumination changes, camouflaged, and static foreground objects, etc. This paper proposes an improved change detector (ICD) to mitigate the above-mentioned challenges. It employs a novel statistical decision criterion (SDC) based on skewness patterns. The SDC helps to differentiate time of day using the camera sensors (color/IR). The input frames are processed according to the time of day. For instance, IR input is image-enhanced to differentiate between camouflaged intruders from the background. Then input goes through Gaussian Mixture Models (GMM) based change detector to segment foreground (intruder). The foreground object is further cleansed using morphological operations for possible isolated noise and holes. The ICD was tested on three datasets and outperformed top-ranked change detection algorithms.



**Index Terms**—IR camera, camouflaged intruder, dual-camera sensors, intelligent surveillance systems.

## I. INTRODUCTION

THE camera-based surveillance systems have become an integral part of a smart city. They are widely applied for security purposes in public and private domains such as bus stops, airports, shopping malls, industrial complexes, or border monitoring. The task involves monitoring a certain area with the help of human input and requires a high level of concentration.

Intelligent Surveillance Systems (ISSs) are improving conventional surveillance systems. They allow autonomous detection of anomalies with minimal human intervention. Sterile zone monitoring (SZM) is an important task of the ISS, which enables the detection of any object entering the prohibited area. The task of SZM may seem trivial. But, it suffers due to IR camera, dynamic backgrounds, illumination changes, camouflaged, and static foreground objects, etc [1]–[6]. SZM has numerous applications depending on the user’s choice. It could be the fence of a prison or an industrial area with expensive equipment or a rooftop of a building abstaining

someone from committing suicide.

The task of SZM requires the change detection algorithm to segment out an intruder (foreground) from a scene (background). It is a well-documented pre-processing task in a multistage computer vision system [1]–[4]. Gaussian Mixture Models (GMM) based algorithms are classified as parametric algorithms [7]–[10]. GMM [7] can be considered as the most employed algorithms in the field of change detection [8]–[10]. They model the background by using a Gaussian fitted with mean and variance. Parametric based algorithms are efficient and showed good results towards the challenge of illumination changes. But, they are affected by the camouflage foreground object due to color/intensity features based background modeling [5]. Self-Balanced SENSitivity SEGmenter (SuBSENSE) [11] is classified as a non-parametric algorithm. It models background using a spatial pattern in a pixel region. Pixel-based Adaptive Word Consensus Segmenter (PAWCS) [12] employed an adaptive update mechanism in the SuBSENSE. Weight Sample Background Extractor (WeSamBE) [13] improved SuBSENSE by introducing a penalty weight strategy for a misclassification. Such algorithms are promising with high-end hardware [6].

Flux Tensor and Split Gaussian (FTSG) [14] is classified as a hybrid algorithm. It exploits flux tensor and Gaussian to model the background. Similarly, In Unity There Is Strength (IUTIS) [15] proposed a genetic algorithm that handpicks the best algorithm for the particular video sequence from a subset of algorithms. These algorithms showed promising results with a high computational cost. Hence, unsuitable for real-

Manuscript received June 26, 2020; accepted July 13, 2020. Date of publication Month, xx, 20xx; date of current version Month, xx, 20xx.

This work has been supported by the National Research Foundation of Korea (NRF) grant funded by the government (MSIT) (No. 2020R1A2C2008972). (Corresponding author: Ajmal Shahbaz)

Ajmal Shahbaz is with the School of Electrical and Computer Engineering, University of Ulsan, Ulsan, 44610 South Korea (e-mail: ajmal@islab.ulsan.ac.kr).

Kang-Hyun Jo are with the Department of Electrical and Computer Engineering, University of Ulsan, Ulsan, 44610 South Korea (e-mail: acejo@ulsan.ac.kr).

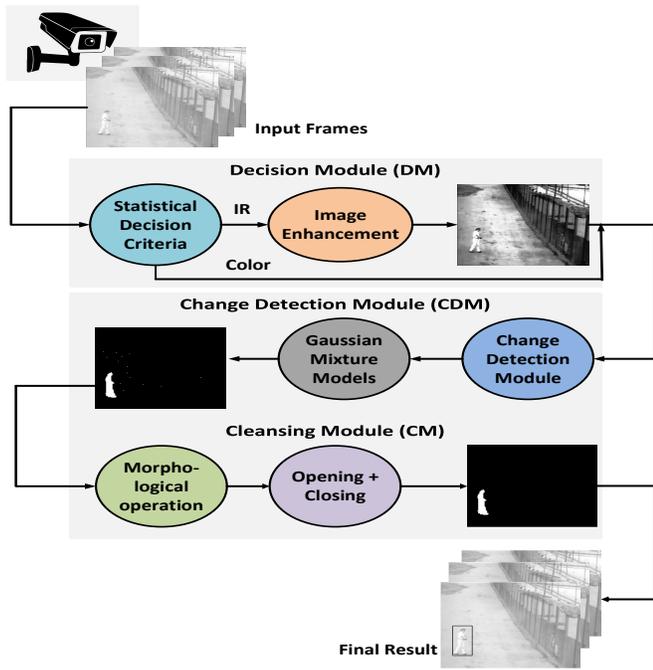


Fig. 1. System overview of proposed algorithm.

time applications. Subspace/low-rank based algorithms [16]–[18] are the popular choice to segment a foreground. Multi-Layer Robust Principal Component Analysis (ML-RPCA) [18] extracts low-rank information using multiple dimensional arrays. Such algorithms require high memory power to store input in the memory for modeling, thus, they are inefficient as well [18].

Deep learning based algorithms showed promising prospects with high computational complexity [19]–[23]. Braham *et al.* [20] proposed training 50% of video and testing remaining using the model [24]. DeepBS [21] trained a CNN model with for whole Change Detection dataset using 5% training frames [25]. CascadeCNN [22] trained multiscale input based CNN from a video sequence. Deep learning algorithms require expensive hardware and are cost-inefficient for real-time applications [26].

This paper proposes an improved change detector (ICD) incorporating a dual-camera sensors for intelligent surveillance systems. The word improved refers to the improvement built over the GMM such as statistical decision criterion and image enhancement scheme for the camouflaged foreground object. This work is an expanded version of the conference paper [27] and the contributions are as follows:

- A novel statistical decision criterion (SDC) helps to integrate a dual-camera sensors (color/IR) by detecting a switch between them. Also, restarting the change detection module.
- Cost-efficient image enhancement schema to differentiate the camouflaged intruder from the background.
- The ICD was incorporated with top-ranked change detectors such as SuBSENSE [11], PAWCS [12], WeSamBE [13], and ML-RPCA [18] to prove effectiveness and generalization.

- The proposed algorithm was tested on three datasets and extensive experiments were performed to select optimal parameter setting.
- Introducing a dataset of HD videos for sterile zone monitoring in an industrial setting using the dual-camera sensors (color/IR).

The rest of the manuscript is divided as: Section II describes the improved changed detector (ICD) in detail. Section III provides experimental analysis in support of the claims.

## II. PROPOSED ALGORITHM

Figure 1 shows a system overview of the proposed algorithm, which consists of three modules:

### A. Decision Module (DM)

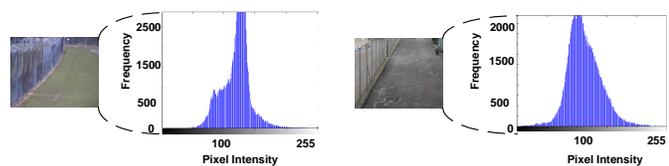
The current ISSs are equipped with a dual-camera sensors (color/IR). The color camera is operated during the day. As soon as the sunlight diminishes, it switches to the IR camera (night). While such cameras are economical, they come with drawbacks. Sudden camera switches may distort the change detection algorithm giving false alarms. Also, strong camouflage due to the IR sensor may result in false negatives, leading to system failure. Such a problem can be resolved with a manual reset of the system. However, it decreases the autonomy of the ISS [27].

The decision module does three crucial jobs. First, it detects a switch in the camera sensor due to the time of day or otherwise. Secondly, it restarts the change detection module again if there is a switch between the camera sensor. This helps to mitigate the problem of false positives/alarms due to a sudden switch. Lastly, it applies an image enhancement to the IR camera frame to tackle the problem of camouflage intruder, decreasing false negatives.

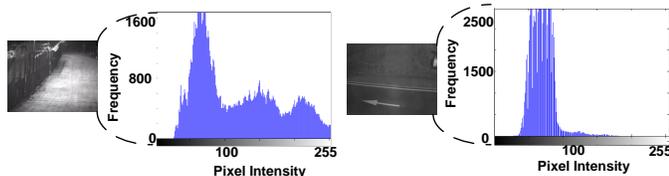
1) *Statistical Decision Criterion (SDC)*: A novel statistical decision criterion (SDC) is responsible for detecting the switch between color and IR sensor. The premise of SDC is inferred from the general characteristics of color and IR cameras [28]. The color camera gives rich detail of the scene, a more spread on intensity range owing to RGB color space. IR sensor gives pixel intensity as shades of gray, a much more congested intensity range. This leads to the assumption that the histograms of color and IR sensor should differ remarkably and follow third order image moments (skewness). Three skewness patterns can be defined using third order image moments (mean  $\mu$ , median  $m$ , and mode  $M$ ):

- 1) If  $\mu = m = M \rightarrow$  Symmetrical pattern
- 2) If  $\mu > m > M \rightarrow$  Left-skewed pattern
- 3) If  $\mu < m < M \rightarrow$  Right-skewed pattern

Figure 2 illustrates the three skewness patterns mentioned above. It can be seen that the color frame (day) follows an approximately symmetrical, while the IR frame (night) could either left or right skewed (Fig. 2b). It was assumed that mean  $\mu$ , median  $m$ , and mode  $M$  of the color frames would be approximately equal.  $|m - M|$  for the day time images varies from 5-7 (Fig. 2a). Similarly, mean  $\mu$ , median  $m$ , and mode  $M$  of the IR frames would be far apart.  $|m - M|$  varies



(a) Day frames (color) with symmetrical pattern. 1st frame ( $m=123, M=130, |m-M|=7$ ) and 2nd frame ( $m=110, M=105, |m-M|=5$ )



(b) Night frames (IR) with left or right skewed pattern. 1st frame ( $m=98, M=63, |m-M|=35$ ) and 2nd frame ( $m=74, M=44, |m-M|=30$ ).

Fig. 2. Illustration of Skewness patterns shown by the color and IR sensor, where x-axis is pixel intensity and y-axis is frequency respectively. It is evident that the color and IR frames exhibit patterns defined in II.A.1.

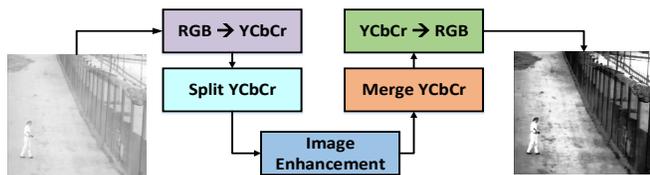


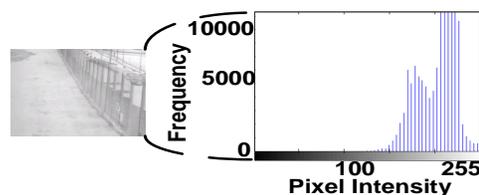
Fig. 3. Flow diagram of image enhancement on IR camera frames (II.A.2).

during night time from 30-35 (Fig. 2b). The difference is large and leads to the detection of the switch between the camera sensors. Based on this experiment, statistical decision criterion SDC was formulated as:

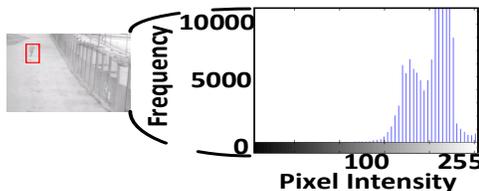
$$SDC = \begin{cases} IR, & |m - M| \geq T \\ Color, & else \end{cases} \quad (1)$$

where  $T$  is a statistical decision threshold. The third order image moments are averaged values over the three channels in the Eq. 1. If  $|m - M| \geq T$ , there is a switch from the color camera to IR. The decision module would restart the change detection module (CDM) again. Furthermore, image enhancement is incorporated to deal with a camouflage foreground object. If  $|m - M| < T$ , there is a switch from the IR to the color camera and CDM would be restarted again. Statistical decision threshold  $T$  is selected heuristically as explained in section III.B.

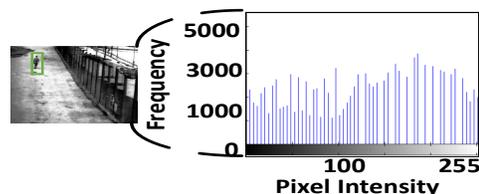
2) *Image Enhancement (IE)*: Input frames from the IR camera are enhanced using a cost-effective schema shown in Figure 3. IE being an intensity based operation, may result in drastic color imbalance for RGB color space. Hence, RGB input is converted into a purely intensity-based color space known as YCbCr. Later, YCbCr color space is split into respective channels (Y, Cb, and Cr). The process of image enhancement consists of calculating probability mass function (PMF) and cumulative density function (CDF) for a pixel intensity. PMF is the probability of a pixel intensity at frame-



(a) Input frame and its histogram



(b) Input frame with camouflage intruder and its histogram



(c) Equalized input frame with camouflage intruder and its histogram

Fig. 4. The effectiveness of image enhancement (IE), where x-axis is pixel intensity and y-axis is frequency respectively. It can be seen that Fig. 4a and Fig. 4b have a similar histogram with a camouflage intruder in the background. The enhanced IR frame reveals the camouflage intruder.

level while CDF is the summation of probabilities of the current pixel and previous pixels. The final values are mapped onto the histogram. Later the channels are merged and YCbCr is converted back to RGB color space. The equalized frame appears to have stretched the range of intensity with the camouflaged object more visible as shown in Fig. 3.

Figure 4 shows the image enhancement (IE) improving the contrast of IR input images, thereby, making the camouflaged intruder/object visible. It could be seen that there is not much difference in the histograms of the input image with the background (Fig. 4a) and the input image with the camouflaged intruder/object without IE (Fig. 4b). Also, the intensity range is congested and the histogram distribution appears to be quite similar (Fig. 4a and 4b).

IE improves the contrast of the image by expanding a pixel's intensity range by flattening the curve (Fig. 4c). The camouflage foreground object is distinguishable from the background. This helps the change detection module (CDM) to better detect the camouflaged intruder/object in the proposed algorithm. While it is arguable that any image enhancement can be applied. However, real-time performance can not be guaranteed. As compared to other histogram-based methods such as CLAHE (8.1 ms) [29] and FCCE (4.6 ms) [30], proposed IE (3.1 ms) gave better processing time.

TABLE I  
COMPUTATIONAL ANALYSIS IN FRAMES PER SECOND (*fps*)

Algorithm	Processing speed ( <i>fps</i> )
GMM	25-35
SuBSENSE	4
PAWCS	2
WeSamBE	2
ML-RPCA	0.5

### B. Change Detection Module (CDM)

The real-time performance of the change detector is a must for the ISS. Extensive experiments were performed on top-ranked change detectors on low-end hardware as shown in Table I. The GMM and its improvement gave good processing speed. While comparative algorithms are computationally inefficient. The ICD is integrated with GMM [9] due to real-time performance and good accuracy [5]-[7]. This section gives an overview of GMM. There are three steps: background modeling, foreground detection, and background model update [9]:

1) *Background Modeling*: Each pixel of initial frames without foreground is modeled using GMM [9]. The probability of observing a particular pixel  $X$  at time  $t$  is:

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \eta(X_t; \mu_{i,t}, \Sigma_{i,t}), \quad (2)$$

where  $\eta$  is probability density function.  $K$  usually 2-5 defines number of Gaussian needed to model the background. While  $\omega_{i,t}$ ,  $\mu_{i,t}$ , and  $\Sigma_{i,t}$  are estimate of weight, mean, and covariance of the  $i$ th Gaussian in the mixture at time  $t$ .

2) *Foreground Detection*: The decision criterion to label particular pixel at time  $t$  as background or foreground is:

$$|X_t - \mu_{i,t}| > \lambda \sigma_{i,t}, \quad (3)$$

where  $\lambda$  is a pixel labeling threshold equal to 2.5. Hence, A foreground pixel is the one located at  $> 2.5$  standard deviations  $\sigma$  away from the background component [7].

3) *Background Model Update*: The background model needs to be updated with new background and foreground values after foreground detection. The new background model  $M_t$  is updated using a weighting factor multiplied by pixel intensity in the current frame  $C_t$  and pixel intensity in the previous background model  $M_{t-1}$ :

$$M_t = \alpha C_t + (1 - \alpha) M_{t-1}, \quad (4)$$

where  $\alpha$  is a learning rate defines the weighting factor in the above equation. The  $\alpha$  can be set according to the background and application. A low  $\alpha$  value can deal with a slowly changing background environment. Similarly, a high  $\alpha$  value is better for a gradual changing background environment. It can be calculated as [25],

$$\alpha = \frac{1}{\tau \times f}, \quad (5)$$

where  $f$  is frames per second. Learning rate  $\alpha$  and time span  $\tau$  show an inverse relationship.  $\tau$  is set as per application. For example, for  $\tau$  of certain pixels more than 20 seconds

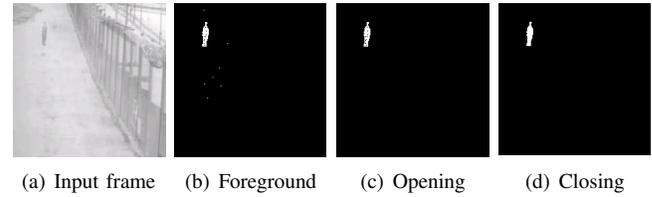


Fig. 5. The effectiveness of cleansing module (CM). CM helps to get rid of isolated noise by opening operation (c). It helps to fill holes in foreground object by using the closing operation (d).

with  $f = 25$  fps, substituting values gives  $\alpha = 0.002$ . Similarly, for  $\tau$  of certain pixels more than 30 seconds with  $f = 25$  fps, substituting values gives  $\alpha = 0.0013$ .

### C. Cleansing Module (CM)

The foreground mask obtained from CDM has isolated noise which may result in false positives (Fig. 5b). The cleansing module (CM) employs the morphological opening and closing. It has two jobs: First, getting rid of the isolated noise in the foreground mask by opening (Fig. 5c). It is the erosion followed by the dilation. The foreground mask obtained is multiplied pointwise with an opening kernel  $K_o$ . First, an erosion operation is performed which is similar to an AND operation in image processing. The foreground value (1) is only outputted when the kernel matches the foreground object in the kernel mask.

Secondly, CM improves the geometry of the foreground object by morphological closing (Fig. 5d). It is the dilation followed by erosion. The foreground mask obtained after the opening operation is multiplied pointwise with the closing kernel  $K_c$ . First, dilation operation is performed which is like an OR operation in image processing. The foreground value (1) is outputted even when the kernel matches partially with the foreground object in the kernel mask.

## III. EXPERIMENTAL RESULTS AND ANALYSIS

The top-ranked change detectors namely GMM [7], its improvements [8]-[9], SuBSENSE [11], PAWCS [12], WeSamBE [13], and ML-RPCA [16] are employed for comparative analysis. Deep learning methods are not included in the comparison as they require ground-truth information of foreground and background during training. Such consensus is established within the change detection research community.

### A. Datasets Description

The experiments were performed on three datasets:

1) *Imagery Library for Intelligent Detection Systems (i-LIDS) Dataset* [24]: The general scenario of the dataset is that an intruder tries to enter a prohibited area by cutting the fence. The intruder comes with the challenges of scale-variance, speed-variance, camouflage, and static foreground object. The night sequences are recorded using an IR camera which is challenging for change detectors. It comprises of 10 videos. There are 5 videos each for day and night. Each video consists of 1,000 frames.

TABLE II  
PARAMETER DEFINITION AND SETTING.

Definition	Symbol	Optimal Setting
Statistical Decision Threshold	$T$	20
Number of Gaussian	$K$	3
Pixel Labeling Threshold	$\lambda$	2.5
Learning Rate	$\alpha$	0.001
Opening Kernel	$K_o$	$3 \times 3$
Closing Kernel	$K_c$	$5 \times 5$

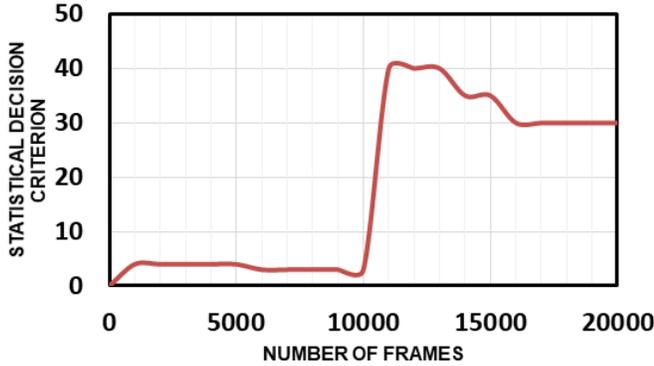


Fig. 6. Statistical decision criterion SDC, where x-axis is number of frames and y-axis is SDC  $|m - M|$  (II.1). First 10,000 frames are day (color) whereas later frames are night (IR). It can be seen that  $|m - M|$  suddenly jumps after day frames.

2) *Intelligent Systems Laboratory dataset for Industrial Sterile Zone Monitoring (ISL-ISZM) Dataset*: The dataset is constructed in an industrial setting according to the challenges outlined by the i-LIDS benchmark. ISL-ISZM dataset comprises of 15 videos. There are 10 videos for the day and 5 videos for the night. Each video varies from 1,000-2,300 frames. The dataset can be accessed at <https://drive.google.com/file/d/1qIRUPgAQeY42zeRlTg2oTqvCC9ImYxMG/view?usp=sharing>.

3) *Change Detection (CDNet) Dataset [25]*: It is considered as a realistic benchmark in the field of change detection with 11 challenging categories and 150,000 video frames. Each category has 4-6 videos. Experiments are performed on five categories such as baseline, dynamic backgrounds, bad weather, shadows, and thermal relevant to ISS. There are 25 videos in these categories with more than 80,000 video frames.

## B. Parameter Setting

The parameter definition and their optimal setting is shown in Table II. The optimal parameter values of  $T$ ,  $K$ , and  $\alpha$  were selected through extensive experiments. The parameters of GMM such as  $K$ ,  $\lambda$ , and  $\alpha$  are well defined and discussed in the literature [5]-[9].

The variation of statistical decision criterion SDC for color and IR frames is shown in Figure 6 to determine the threshold  $T$ . The analysis was performed on 20,000 frames arranged as the color (first 10,000) and IR sequences. The frames were chosen from i-LIDS and ISL-SZM datasets with six different background settings. It is evident that variation of  $|m - M|$  was small (i.e., 5-7) for color (day). The value increased above

TABLE III  
QUANTITATIVE ANALYSIS ON PERFORMANCE METRICS NAMELY RECALL  $R$ , PRECISION  $P$ , AND F-MEASURE  $F$ .

Algorithm	i-LIDS Dataset			ISL-SZM Dataset		
	$R$	$P$	$F$	$R$	$P$	$F$
GMM	0.20	0.20	0.20	0.00	0.00	0.00
SuBSENSE	0.60	0.60	0.60	0.20	0.20	0.20
PAWCS	0.20	0.20	0.20	0.20	0.20	0.20
WeSamBE	0.60	0.60	0.60	0.00	0.00	0.00
ML-RPCA	1.00	0.50	0.66	1.00	0.50	0.66
Proposed+GMM	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
Proposed+SuBSENSE	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
Proposed+PAWCS	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
Proposed+WeSamBE	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
Proposed+ML-RPCA	<b>1.00</b>	<b>0.70</b>	<b>0.82</b>	<b>1.00</b>	<b>0.70</b>	<b>0.82</b>

40 and varies between 35-45 for IR (night). Thus,  $T=20$  was selected as the optimal value which lies in the middle of variation of  $|m - M|$  between color and IR frames.

Pixel labeling threshold  $\lambda$  value is implicated from the 68-95-99.7 standard deviation  $\sigma$  rule from statistics [26]. The one, two, and three  $\sigma$  cover 68%, 95%, and 99.7% of data points (pixel values belonging to the background) in the Gaussian. Thus,  $2.5 \sigma$  covers roughly 99% of pixel values in a Gaussian. The small kernel size is used for opening  $K_o$  ( $3 \times 3$ ) and closing  $K_c$  ( $5 \times 5$ ) to avoid distortion in the foreground object.

The comparative methods were tested with the original setting provided by the respective authors. The proposed algorithm uses six parameters. It is still fewer than the SuBSENSE, PAWCS, and WeSamBE, which employ parameters  $\geq 10$ .

## C. Quantitative Analysis

1) *i-LIDS Dataset*: The performance metrics namely recall  $R$ , precision  $P$ , and F-measure  $F$  on night time sequences were used (Table III). The i-LIDS dataset defines a frame-level evaluation criterion. An intruder must be detected for at least 75% of a challenge (video) to be marked as a success [24]. Each video sequence contributes 20% of the average  $F$ . The GMM detected intruder in one sequence. SuBSENSE, PAWCS, and WeSamBE were able to detect and track an intruder in 3, 1, and 3 sequences respectively. ML-RPCA detected intruders in all the videos with false positives due to illumination changes and shadows.

2) *ISL-ISZM Dataset*: ISL-ISZM poses a severe camouflage effect with illumination changes and shadows. GMM based algorithms and WeSamBE failed to detect an intruder. SuBSENSE and PAWCS were able to detect an intruder in one sequence successfully. ML-RPCA detected all the intruders. But, it gave too many false positives decreasing overall performance. The proposed algorithm when combined with all the comparative algorithms improved their average  $F$  by 16-80% (Table III).

3) *Change Detection Dataset (CDNet)*: Table IV shows the quantitative analysis of the CDNet. CDNet defines a pixel-wise evaluation criterion to compare the foreground mask with ground-truth. The GMM was improved 5-8% in terms of recall  $R$ , precision  $P$ , and F-measure  $F$  by the proposed algorithm. The SuBSENSE was also improved by the 1-3% in  $P$  and  $F$ . The proposed algorithm improves the precision of GMM and SuBSENSE which is crucial for the ISS.

TABLE IV  
 QUANTITATIVE ANALYSIS ON THE CDNET AS AVERAGE VALUE OF THE METRICS.

Algorithm	Recall	Specificity	False Positive Rate	False Negative Rate	% Wrong Classifications	F-measure	Precision
GMM	0.7334	0.9928	0.0071	0.2660	1.9973	0.7164	0.7663
SuBSENSE	0.8616	0.9958	0.0041	0.1383	0.4855	0.8691	0.8895
PAWCS	0.8626	0.9961	0.0039	0.137	0.7195	0.8744	0.8979
WeSamBE	0.8302	0.9963	0.0036	0.1697	0.9861	0.8484	0.8945
Proposed+SuBSENSE	<b>0.8861</b>	<b>0.9966</b>	<b>0.0033</b>	<b>0.1138</b>	<b>0.7916</b>	<b>0.8988</b>	<b>0.9133</b>
Proposed+GMM	<b>0.7897</b>	<b>0.9946</b>	<b>0.0054</b>	<b>0.2123</b>	<b>1.4748</b>	<b>0.8028</b>	<b>0.8242</b>

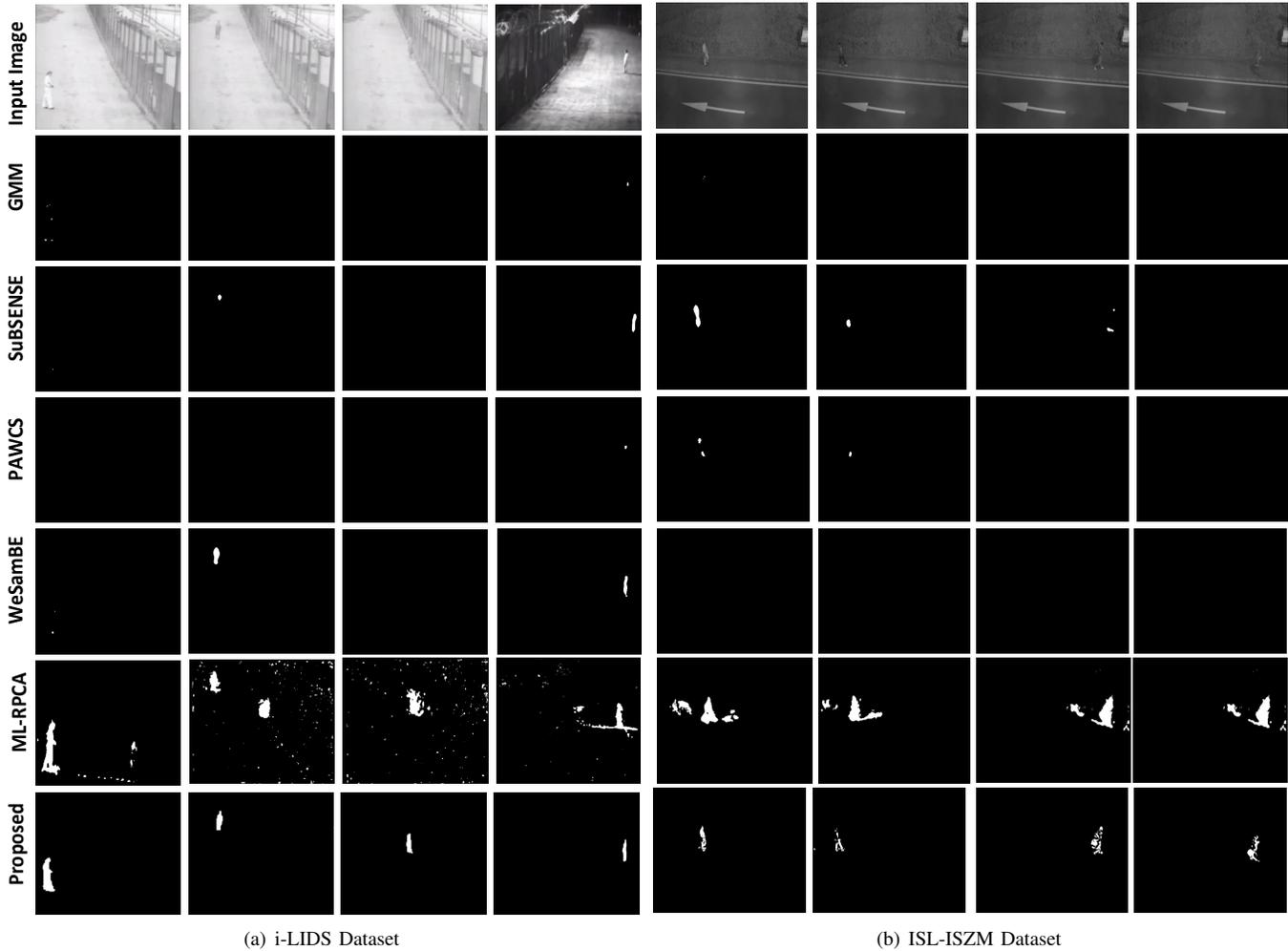


Fig. 7. Qualitative analysis of the proposed algorithm with the comparative methods. GMM and its improvement had the same performance and are shown in one row. Similarly, all of the comparative methods missed intruders except ML-RPCA. But, it gave false positives.

#### D. Qualitative Analysis

1) *i-LIDS Dataset*: Figure 7a shows the qualitative comparison of the i-LIDS dataset on the night sequences to validate the effectiveness of the proposed algorithm. GMM (2nd row) were only able to segment the distinct part of the intruder from the background. SuBSENSE was able to segment intruder partially or fully (3rd row). PAWCS was only able to segment intruder in one sequence (4th row). WeSamBE had similar performance to SuBSENSE (5th row). ML-RPCA was able to segment foreground objects in all the sequences but it gave too many false positives (6th row). The proposed algorithm (7th row) detected the precise geometry of intruders.

2) *ISL-ISZM Dataset*: Figure 7b shows the qualitative comparison of the ISL-SZM dataset. The dataset is more challenging as it poses a severe camouflage effect. Also, it has the challenge of multiple intruders. GMM (2nd row) and WeSamBE (5th row) failed all sequences. SuBSENSE, PAWCS, and WeSamBE which are top-ranked algorithms on CDNet failed to cope with challenges of the camouflaged intruder. SuBSENSE (3rd row) and PAWCS (4th row) were able to segment an intruder in one sequence only. ML-RPCA was able to segment foreground objects in all the sequences but it gave too many false positives (6th row). The proposed algorithm detected intruders with precise boundaries (7th row).



Fig. 8. Final detection results of the proposed algorithm.

3) *Final Detection Results*: The final detection results on the i-LIDS dataset, ISL-ISZM dataset, and CDnet dataset is shown in Figure 8. The quantitative analysis of Figure 8 is shown in Table V. The proposed algorithm was successfully able to detect all the intruders and track them. It is evident that the proposed algorithm was able to cope with multiple intruders/objects and dynamic backgrounds challenge as shown in Figure 8b and 8c.

### E. Computational Comparison

The ICD was implemented on Intel core i5 processor with 3.40 GHz and 8 GB RAM in the C++ programming language. The videos were resized to 640×480. The proposed algorithm integrated with GMM showed better processing speed (*fps*) than the other comparative methods (Table VI). It is evident from Table I and VI that the proposed algorithm does not constitute much processing time. The Proposed+GMM gave good processing speed. While comparative methods are computationally inefficient due to their background modeling on low-cost hardware.

Per-operation processing time in milliseconds (*ms*) is shown in Table VII. The operations from GMM such as the background modeling, foreground detection, and background model update accumulate to most of the processing time. The improvements accumulate to 5-16% of processing time only (1.2±3.1 ms). Due to the statistical decision criterion, ample processing time (3.1 ms) was also saved.

TABLE V

PERFORMANCE OF PROPOSED ALGORITHM ON VARIOUS CHALLENGES, WHERE *GT*, *TP*, *FP*, AND *F* ARE GROUNDTRUTH, TRUE POSITIVE, FALSE POSITIVE, AND F- MEASURE, RESPECTIVELY.

#	Challenges	<i>GT</i>	<i>TP</i>	<i>FP</i>	<i>F</i>
<b>i-LIDS Dataset</b>					
1	Illumination changes	1	1	0	1
2	Illumination changes	1	1	0	1
3	Illumination changes	1	1	0	1
4	Illumination changes	1	1	0	1
5	Illumination changes	1	1	0	1
6	IR, camouflaged intruder	1	1	0	1
7	IR, illumination changes	1	1	0	1
8	IR, illumination changes	1	1	0	1
9	IR, camouflaged intruder	1	1	0	1
10	IR, camouflaged intruder	1	1	0	1
<b>ISL-SZM Dataset</b>					
11	Illumination changes	1	1	0	1
12	Illumination changes	1	1	0	1
13	Illumination changes	1	1	0	1
14	Illumination changes	1	1	0	1
15	Multiple intruders	2	2	0	1
16	Illumination changes	1	1	0	1
17	Illumination changes	1	1	0	1
18	Illumination changes	1	1	0	1
19	Dynamic backgrounds	1	1	0	1
20	Dynamic backgrounds	1	1	0	1
21	IR, camouflaged intruder	1	1	0	1
22	IR, camouflaged intruder	1	1	0	1
23	IR, camouflaged intruder	1	1	0	1
24	IR, camouflaged intruder	1	1	0	1
25	IR, multiple intruders	2	2	0	1
<b>CDNet Dataset</b>					
26	Multiple objects, dynamic backgrounds	5	5	0	1
27	Multiple objects, dynamic backgrounds	5	5	0	1
28	Multiple objects, dynamic backgrounds	4	4	0	1
29	Multiple objects, dynamic backgrounds	3	3	0	1
30	Multiple objects, dynamic backgrounds	2	2	0	1

TABLE VI

COMPUTATIONAL ANALYSIS IN FRAMES PER SECOND (*fps*)

Algorithm	Processing Speed ( <i>fps</i> )
Proposed+GMM	30-33
Proposed+SuBSENSE	3.9
Proposed+PAWCS	2
Proposed+WeSamBE	2
Proposed+ML-RPCA	0.5

TABLE VII

PROCESSING TIME IN MILLI-SECONDS (*ms*)

Operation	Processing Time ( <i>ms</i> )
Statistical Decision Criterion	1.2
Image Enhancement	±3.1
Background Modeling	16.6
Foreground Detection	6.6
Background Model Update	5.6
Cleansing Module	0.3
Total	29.8±3.1

## IV. CONCLUSION

This paper presented an improved change detector for intelligent surveillance systems in general and sterile zone monitoring in particular. It overcomes the inherent drawback of change detection algorithms posed by the dual-camera sensors (color/IR). It was tested on three datasets and compared with top-ranked change detection algorithms. The proposed algorithm might be integrated with the deep learning based

intruder classification in the future.

## REFERENCES

- [1] S. Jan and Y. Chen, "Development of a new airport unusual-weather detection system with aircraft surveillance information," *IEEE Sensors Journal*, vol. 19, no. 20, pp. 9543–9551, 2019.
- [2] S. Huang, H. Liu, B. Chen, Z. Fang, T. Tan, and S. Kuo, "A gray relational analysis-based motion detection algorithm for real-world surveillance sensor deployment," *IEEE Sensors Journal*, vol. 19, no. 3, pp. 1019–1027, Feb 2019.
- [3] S. S. M. S. A. Narayanan, and V. Menon, "Maximizing camera coverage in multi-camera surveillance networks," *IEEE Sensors Journal*, pp. 1–1, 2020.
- [4] Wahyono, A. Filonenko, and K. Jo, "Unattended object identification for intelligent surveillance systems using sequence of dual background difference," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2247–2255, 2016.
- [5] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview," *Computer Science Review*, vol. 11–12, pp. 31 – 66, 2014.
- [6] A. Shahbaz, L. Kurnianggoro, Wahyono, and K.-H. Jo, *Recent Advances in the Field of Foreground Detection: An Overview*. Cham: Springer International Publishing, 2017, pp. 261–269.
- [7] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," in *Computer Vision and Pattern Recognition. IEEE Computer Society Conference on*, vol. 2, 1999, p. 252 Vol. 2.
- [8] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Pattern Recognition. ICPR. Proceedings of the 17th International Conference on*, vol. 2, 2004, pp. 28–31 Vol.2.
- [9] P. KaewTraKulPong and R. Bowden, *An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection*. Boston, MA: Springer US, 2002, pp. 135–144.
- [10] Y. Chen, J. Wang, and H. Lu, "Learning sharable models for robust background subtraction," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2015, pp. 1–6.
- [11] P. L. St-Charles, G. A. Bilodeau, and R. Bergevin, "Subsense: A universal change detection method with local adaptive sensitivity," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 359–373, 2015.
- [12] P. St-Charles, G. Bilodeau, and R. Bergevin, "Universal background subtraction using word consensus models," *IEEE Transactions on Image Processing*, vol. 25, no. 10, pp. 4768–4781, Oct 2016.
- [13] S. Jiang and X. Lu, "Wesambe: A weight-sample-based method for background subtraction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 9, pp. 2105–2115, Sept 2018.
- [14] R. Wang, F. Bunyak, G. Seetharaman, and K. Palaniappan, "Static and moving object detection using flux tensor with split gaussian models," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, June 2014, pp. 420–424.
- [15] S. Bianco, G. Ciocca, and R. Schettini, "Combination of video change detection algorithms by genetic programming," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 6, pp. 914–928, 2017.
- [16] J. Giraldo-Zuluaga, A. Gómez, A. Salazar, and A. Diaz-Pulido, "Camera-trap images segmentation using multi-layer robust principal component analysis," *CoRR*, vol. abs/1701.08180, 2017.
- [17] G. Liu and S. Yan, "Active subspace: Toward scalable low-rank learning," *Neural Computation*, vol. 24, no. 12, pp. 3371–3394, Dec 2012.
- [18] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 11:1–11:37, Jun. 2011.
- [19] T. Akilan, Q. M. J. Wu, and W. Zhang, "Video foreground extraction using multi-view receptive field and encoderdecoder dcnn for traffic and surveillance applications," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9478–9493, 2019.
- [20] T. Akilan, Q. J. Wu, A. Safaei, J. Huo, and Y. Yang, "A 3d cnn-lstm-based image-to-image foreground segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 959–971, 2020.
- [21] M. Babae, D. T. Dinh, and G. Rigoll, "A deep convolutional neural network for video sequence background subtraction," *Pattern Recogn.*, vol. 76, no. C, pp. 635–649, Apr. 2018.
- [22] Y. Wang, Z. Luo, and P.-M. Jodoin, "Interactive deep learning method for segmenting moving objects," *Pattern Recogn. Lett.*, vol. 96, no. C, pp. 66–75, Sep. 2017.
- [23] A. Cioppa, M. V. Droogenbroeck, and M. Braham, "Real-time semantic background subtraction," *CoRR*, vol. abs/1409.1556, 2020.
- [24] H. O. S. D. Branch, "Imagery library for intelligent detection systems (i-lids)," in *Crime and Security. The Institution of Engineering and Technology Conference on*, 2006, pp. 445–448.
- [25] Y. Wang, P. M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, "Cdnet 2014: An expanded change detection benchmark dataset," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 393–400.
- [26] A. Shahbaz, L. Kurnianggoro, J. Hariyono, and K. H. Jo, "Parameter analysis of probabilistic foreground detector for intelligent surveillance system," in *IECON- 42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016, pp. 5574–5578.
- [27] A. Shahbaz, D. Cceres Hernandez, A. Filonenko, J. Hariyono, and K. Jo, "Probabilistic foreground detector with camouflage detection for sterile zone monitoring," in *IEEE 25th International Symposium on Industrial Electronics (ISIE)*, June 2016, pp. 997–1001.
- [28] A. Filonenko, D. C. Hernandez, and K. Jo, "Fast smoke detection for video surveillance using cuda," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 725–733, Feb 2018.
- [29] H. Yue, J. Yang, X. Sun, F. Wu, and C. Hou, "Contrast enhancement based on intrinsic image decomposition," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 3981–3994, Aug 2017.
- [30] A. S. Parihar, O. P. Verma, and C. Khanna, "Fuzzy-contextual contrast enhancement," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1810–1819, April 2017.

**Ajmal Shahbaz** (S'16) Ajmal Shahbaz received the Bachelors of Electronics Engineering from Islamia University of Bahawalpur, Bahawalpur, Pakistan in 2014. He is working towards the Ph.D degree in Electrical Engineering, University of Ulsan, Ulsan, South Korea. His research interests include computer vision, image processing, and deep learning.



**Kang-Hyun Jo** (M'96–SM'16) Kang-Hyun Jo received the BS degree from Busan National University, Korea, 1989 and MS. and Ph.D. degrees in Computer Controlled Machinery from Osaka University, Japan, in 1993 and 1997, respectively. After a year of experience at ETRI as a postdoctoral research fellow, he joined the School of Electrical Engineering, University of Ulsan, Ulsan, Korea. He has served as a director or an AdCom member of Institute of Control, Robotics and Systems, The Society of Instrument and Control Engineers, and IEEE IES Technical Committee on Human Factors Chair, AdCom member, and the Secretary until 2019. Currently, he is serving as Faculty Dean of School of Electrical Engineering, University of Ulsan. He has also been involved in organizing many international conferences such as International Workshop on Frontiers of Computer Vision, International Conference on Intelligent Computation, International Conference on Industrial Technology, International Conference on Human System Interactions, and Annual Conference of the IEEE Industrial Electronics Society. At present, he is an Editorial Board Member for the International Journal of Control, Automation, and Systems and a Vice President for Membership of ICROS. His research interests include computer vision, robotics, autonomous vehicle, and ambient intelligence.

