Using Resources Competition and Memory Cell Development to Select the Best GMM for Background Subtraction

Wafa Nebili, University 8 Mai 1945 Guelma, Guelma, Algeria

Brahim Farou, University 8 Mai 1945 Guelma, Guelma, Algeria

https://orcid.org/0000-0002-1609-6006

Hamid Seridi, University 8 Mai 1945 Guelma, Guelma, Algeria

ABSTRACT

Background subtraction is an essential step in the process of monitoring videos. Several works have proposed models to differentiate the background pixels from the foreground pixels. Mixtures of Gaussian (GMM) are among the most popular models for a such problem. However, the use of a fixed number of Gaussians influence on their results quality. This article proposes an improvement of the GMM based on the use of the artificial immune recognition system (AIRS) to generate and introduce new Gaussians instead of using a fixed number of Gaussians. The proposed approach exploits the robustness of the mutation function in the generation phase of the new ARBs to create new Gaussians. These Gaussians are then filtered into the resource competition phase in order to keep only ones that best represent the background. The system tested on Wallflower and UCSD datasets has proven its effectiveness against other state-of-art methods.

KEYWORDS

Background Pixel, Background Subtraction, Foreground Pixel, Gaussian Model, Moving Objects, Pixels Classification, Static Objects, Video Surveillance

INTRODUCTION

Moving objects segmentation from scenes that are captured with a stationary/non-stationary camera is one of the most difficult and interesting activities in computer vision (Brutzer et al., 2011; Lim and Keles, 2018b). Subtracting the background requires a powerful method that ensures a good separation between the background and the foreground.

In the literature, there are several methods for detecting moving objects without knowing any prior information about them (Toyama et al., 1999). Generally, all these methods share the following steps (Bouwmans, 2012):

Background Initialization

In this step, a primary background model is constructed and learned by a set of frames that have no moving objects. There are many ways which can be designed this model like (statistical, fuzzy, neuro-inspired, etc.).

DOI: 10.4018/IJSITA.2019040102

Foreground Detection

After initializing the background model, each frame is compared with the background model to define the foreground.

Background Maintenance

During this step, all settings of the background model are updated to pick up any novel changes in the background within a video.

GMM is one of the most popular methods that has achieved considerable success in detecting changes in videos. However, this method has failed in problems related to: lighting changes and hidden areas. Several studies showed that the number of Gaussians in GMM influence on the results quality. The contribution of this work is to manage dynamically the number of Gaussians based on the AIRS algorithm instead of fixing them a priori by the user. This paper proposes to generate a set of new Gaussians using two different strategies: the first one (Random generation) uses the AIRS to improve the system decision while the second one (Directed generation) uses the AIRS to improve the GMM learning phase.

Random Generation

Firstly, the system stars with a learning phase using the GMM algorithm. During the classification stage, the AIRS generates several Gaussian models using Memory cell identification and ARB generation process for all pixels regardless their nature. These models are filtered according to the resource competition and memory cell development process of the AIRS algorithm to select only the best models. Once the AIRS algorithm is finished, the GMM method is used to decide the pixels nature.

Directed Generation

It begins with the same first step as a random generation method and consists to apply Memory cell identification and ARB generation process only for pixels representing the background. Indeed, the system used the GMM algorithm to filter background from foreground pixels before the mutation process to reduce the time consumed to generate new models and to improve accuracy since the mutation process is based only on pixels representing the background.

To cover all sections, the paper is organized as follows: Section 2 provides an overview of literature works related to background subtraction in which we proposed a taxonomy. Section 3 and 4 present a definition of methods used (GMM, AIRS). Section 5 is dedicated to our contribution in which we present two propositions. Some experiments on Wallflower and UCSD datasets are discussed in section 6. Section 7 concludes the paper.

RELATED WORKS

Subtracting the background from videos remains a crucial problem due to the background variations. Several studies have been proposed to improve the quality of background subtraction results. These studies can be divided into two groups: the first group is focused on selecting a good feature (color, texture, edge), while the other try to choose the best algorithm for video changes detection. Among the approaches that are interested in selecting the right features:

St-Charles et al. (2015b) proposed a new universal pixel-level segmentation method based on the selection of spatiotemporal binary features and colors to detect video changes. Authors in (Wang et al., 2018) exposed a type of multi-view learning based on the use of heterogeneous features such as: brightness variation, chromaticity and texture variation to define background and foreground pixels. In (Allebosch et al., 2015), authors proposed a model that combines RGB color space and edge descriptors to classify the pixels.

Recently, many works focused on the development of parametric models for the background subtraction. Among this works: Wren et al. (1997) proposed a unimodal representation of the background based on the use of a single Gaussian probability distribution (SG) to describe pixel variations. This model is simple, very fast and computationally inexpensive, but it is sensitive to fast pixel variations. Indeed, only one Gaussian cannot memorize the old states of the pixel. This requires migration to more robust and multi-modal approach. Friedman and Russell (1997) proposed the first model, which describes the variance of the recent values of each pixel by a mixture of the Gaussians. In this model, the Expectation Maximization (EM) algorithm is used to initialize and estimate the parameters of each Gaussian. However, Stauffer and Grimson (1999) proposed a standardization of this model with efficient update equations.

Several studies have been proposed to improve GMM accuracy in complex scenes. Among this works, there are which interested in hybrid models such as GMM and Block matching (Farou et al., 2017), boosted Gaussian Mixture Model (Martins et al., 2017) which combines chromatic features by a hysteresis to classify pixels, GMM with a spatio-temporal distribution (Xia et al., 2016), GMM and k-means (Charoenpong et al., 2010).

There are also methods that focused on non-parametric background model. In (Elgammal et al., 2000) authors estimated the probability density function of the recent N values of each pixel with a kernel estimator (KDE). Jianzhao et al. (2017) proposed an improvement of this method with the LUT method. Authors of (Haritaoglu et al., 2000) used the notion of a visual dictionary words to model the pixels of the background. An identical approach (St-Charles et al., 2015a) modeled the background with a visual dictionary words, but it adds an automatic mechanism for updating inner parameters. Authors of (Krungkaew and Kusakunniran, 2016) presented a combination between the visual dictionary words and the lab colorimetric space (light, color channels) for the background subtraction in dynamic scenes.

Some research works have introduced the fuzzy concept to develop more efficient and robust methods for modeling the background, such as (El Baf et al., 2008; Sigari et al., 2008; Bouwmans and El Baf, 2010; El Baf et al., 2009; Zhao et al., 2012).

There are many approaches that used sub-spatial learning for the background subtraction. One of these methods used the principal component analysis (PCA) (Oliver et al., 2000) to create and to learn the background model. Works presented in (Javed et al., 2018; Hunziker et al., 2018; Shen et al., 2016) proposed to improve the performance of the PCA for the background subtraction. Authors in (Tsai and Lai, 2009) defined another type of sub-spatial learning which use the independent component analysis (ICA) in the video change detection. Another work of (Bucak and Gunsel, 2007) presented a decomposition of video content by an incremental non-negative matrix factorization (SL-INMF).

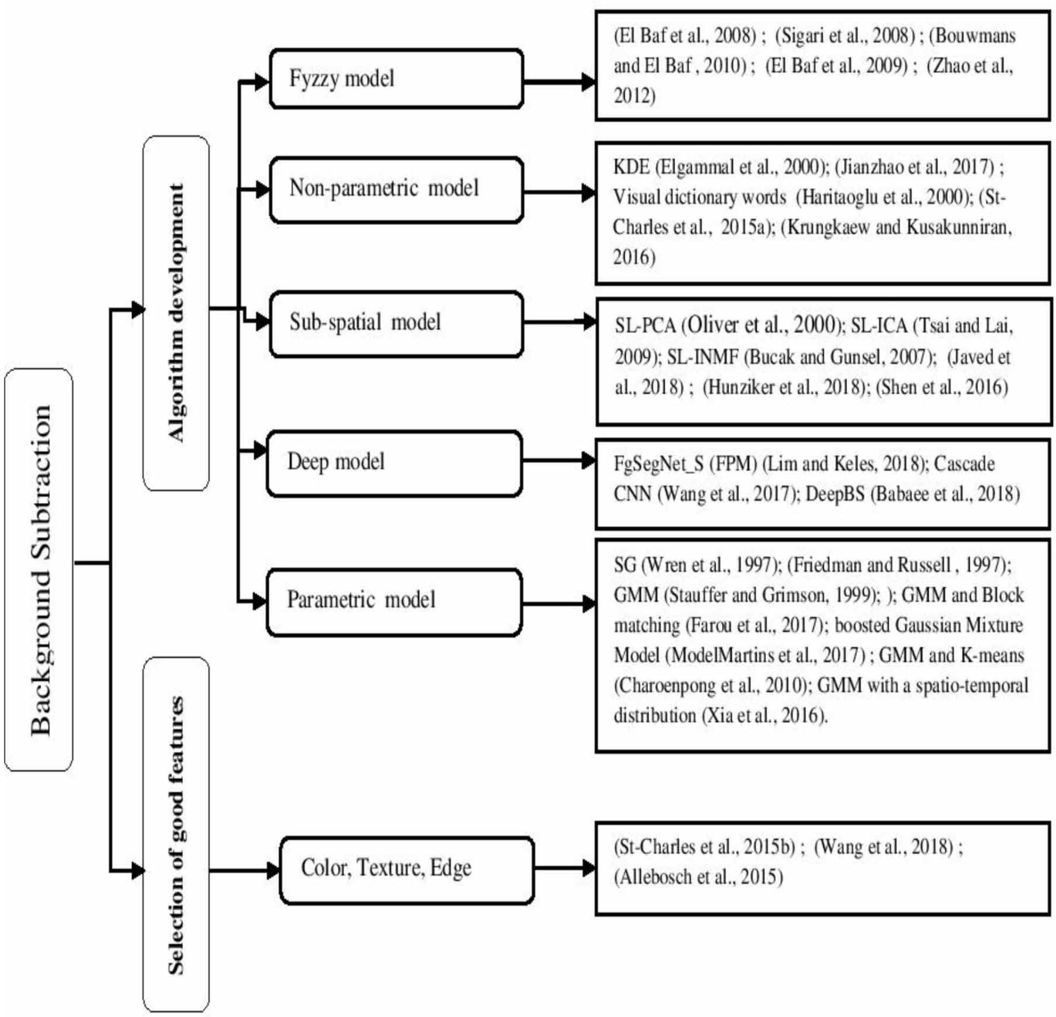
Recently, several methods used deep learning for subtracting the background, among this method: FgSegNet_S (FPM) (Lim and Keles, 2018a), Cascade CNN (Wang et al., 2017), DeepBS (Babaee et al., 2018). However, deep leaning methods require a large number of simples and needs more time for training. In addition, any change in the background leads to restart the learning of the model, which is impossible in a real monitoring application because of the high cost in time and power required for this operation (figure 1).

GAUSSIAN MIXTURE MODEL (GMM)

Gaussian mixture model (GMM) proposed by Stauffer and Grimson (1999) assumes that the observation of a given pixel value will be in one of K Gaussian distributions at one time.

The concept of the method is simple, for each new pixel, their intensity in the HSV color space is compared with the existing Gaussian probability distributions (figure 2). A high probability to these Gaussian mixtures indicates that the pixel represents the background. The probability of observing the value of the current pixel is:

Figure 1. Proposed taxonomy for Background subtraction



$$\mathrm{P}(P_t) = \sum_{k=1}^{K} w_{k,t}\,\eta(\mu_{k,t}, \Sigma_{k,t}, P_t) \qquad (1)$$

Where:

K is the number of Gaussians, $w_{k,t}$ the weight of the la kth Gaussian at time t, $\Sigma_{k,t}$ and $\mu_{k,t}$ respectively the covariance matrix and the average of the kth Gaussian at time t and η is the Gaussian probability density function.

After initializing the Gaussian parameters ($w_{k,t}$, $\mu_{k,t}$, $\Sigma_{k,t}$), each new pixel value is tested with these K Gaussian distributions to find the Gaussian that correspondent to it using the following equation: $\frac{|P_t - \mu_k|}{\sigma_k} < 2.5$ (2)

If the Equation (2) is satisfied, the parameters of the matched Gaussian are updated using the following equations:

$$w_{k,t} = (1-\alpha)w_{k,t} + \alpha \tag{3}$$

$$\mu_{k,t} = (1-\varphi_k)\mu_{k,t-1} + \varphi_k P_t \tag{4}$$

$$\sigma^2_{k,t} = \left(1-\varphi_k\right)\sigma^2_{k,t-1} + \varphi_k \left(P_t - \mu_{k,t}\right)^T \tag{5}$$

$$\varphi_k = \alpha\eta\left(P_t \mid \mu_k, \sigma_k\right) \tag{6}$$

Only the weights are updated for the other unmatched distributions (see Equation (7)).

$$w_{k,t} = (1-\alpha)w_{k,t-1} \tag{7}$$

After this step, all weights are normalized to ensure that their sum always equals 1.

If the match is not made for all the K Gaussians, the pixel is classified as foreground pixel and the least probable Gaussian will be replaced by a new distribution whose parameters are defined with:

$$\sigma^2_k = \sigma^2_k \tag{8}$$

$$w_k = \varphi_k \tag{9}$$

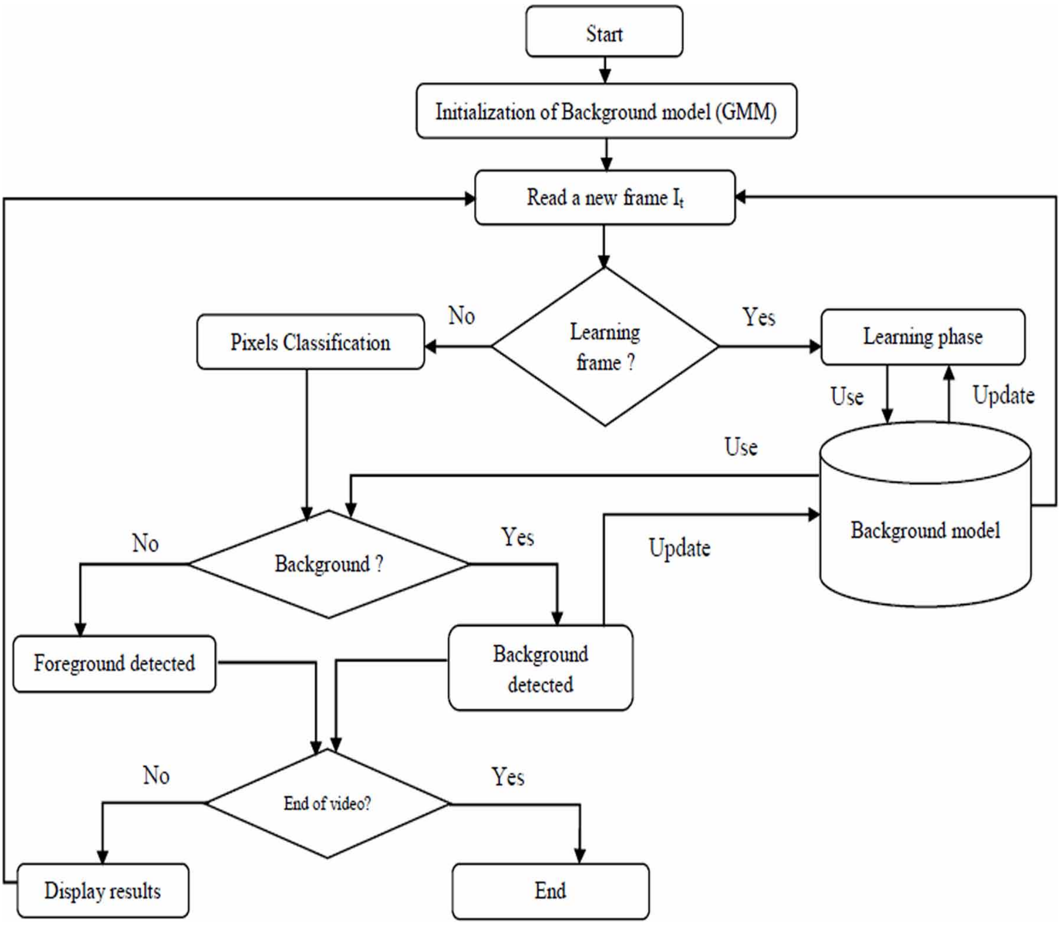
$$\mu_k = P_t \tag{10}$$

To distinguish the foreground pixels from the background pixels, the distributions will be ordred according to the value of $w_{k,t} / \sigma_{k,t}$. The first β distributions that verified Equation (11) are selected to represent the background.

$$\beta = \operatorname{argmin}\left(\sum_{k=1}^{K} w_{k,t} > B\right) \tag{11}$$

Where B is the minimum part of the data corresponding to the background.

Figure 2. Flow chart of GMM method



ARTIFICIAL IMMUNE RECOGNITION SYSTEM (AIRS)

In recent years there has been considerable interest in exploiting the bio-inspired approaches in computer applications. Among the approaches that have achieved great success in the optimization problem and machine learning is the artificial immune recognition system (AIRS). The first supervised artificial immune system was proposed by Watkins (2001), this model based on antigen-antibody representation, which measures a degree of "closeness" or similarity between training data (antigens) and cell B. The algorithm takes an antigen as an input and produces a set of memory cells. Memory cells represent the model that can be used in the classification stage. AIRS is described on four phases:

Phase One (Initialization)

It is considered as a data pre-processing stage, since all data will be normalized such that the Euclidean Distance (affinity) between two feature vectors is always in [0, 1]. After normalized the data, the affinity threshold (AT) will be calculated to use later in the decision of which memory cells will be replaced. At the end of this stage memory cell set MC and ARB set are initialized by some antigens selected randomly from AG set.

$$AT = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} Affinity(ag_i, ag_j)}{\frac{n(n-1)}{2}} \tag{12}$$

Once the initialization step is finished, each antigen in the learning set through the following phases.

Phase Two (Memory Cell Identification and ARB Generation)

The first step in this stage is to determine the memory cell (mc_{match}) the most stimulated with the current antigen *ag* and which has the same class of the latter. AIRS generates several new clones by mutating the feature vector mc_{match} according to a probability rate. The new clones placed with the clones previously generated in the AB set.

$$mc_{match} = \text{argmax}_{mc \in MC_{ag.c}} Stimulation(ag, mc) \tag{13}$$

With:

$$Stimulation(ag, mc) = 1 - Affinity(ag, mc) \tag{14}$$

Phase Three (Competition for Resources and Development of a Candidate Memory Cell)

AIRS has a mechanism to organize the survival of individuals within the ARB population (AB), this mechanism removes the bad quality of ARBs using the cumulative resource division technique according to the nature of the antigen class. ARBs of the same class of *ag* have half of the cumulative resources, the other half will be divided on the rest of the ARBs. Each class has a maximum resource allocation, if the allocation of resources of a class exceeds their maximum allocation, the excess resources will be removed from the least stimulated ARBs. Each *ab* has a number of resources less than or equal to 0 will be removed from the AB set. The remaining ARBs are given another opportunity to generate new clones. The new clones will be also filtered through the competition for resources process. These two steps will be repeated until the average stimulation of each class exceeds the stimulation threshold.

Once the stopping criterion is reached, the AIRS selects only one *ab* that is most stimulated with the current antigen *ag* as a candidate memory cell ($mc_{candidate}$).

Phase Four (Memory Cell Introduction)

This is the last phase in the process of training a single antigen. If the stimulation of $mc_{candidate}$ with the antigen *ag* greater than that of the mc_{match}. $mc_{candidate}$ will be introduced as a new memory cell (*mc*). If the affinity between $mc_{candidate}$ and mc_{match} is less than AT × ATS, $mc_{candidate}$ replaces mc_{match} in MC set.

PROPOSITIONS

The in-depth study made on Gaussian mixtures shows the important role of the number of Gaussians in describing the pixel variations. Indeed the use of a fixed number of Gaussians can influence results quality, because a small number of Gaussians can reduces the historic of the pixel, which causes a problem in the dynamic background. However, if we use a very large number of Gaussians without updating the Gaussian numbers we cannot correctly classify stationary objects that become moving objects. Following this principle propose two novel strategies to introduce new Gaussians based on the AIRS algorithm in order to be as faithful as possible to the background model (figure 3). Indeed, these ideas allows to the passage from a static model where the number of Gaussians is fixed empirically for all pixels towards a model dynamic and adaptive according to the environment and the background complexity.

Proposition One (Random Generation)

The first proposition consists to introduce to the GMM model of any pixel (background pixel or foreground pixel) a set of Gaussian representing the background using the AIRS algorithm. Then we decide the nature of this pixel according to the GMM method.

Firstly, we create a set of Gaussian (g_i) representing the background for the pixel P_t (at time t) that vitrifies:

$$Background\_set = \left\{ g_i, \frac{|P_t - \mu_i|}{\sigma_i} < 2.5 \right\} \tag{15}$$

Such that each Gaussian g_i is represented by: the pixel value P_i, the average μ_i, the variance σ_i and the weight w_k. After creating the background model, we choose the Gaussian (mc_{match}) that has the closest distance to the value of the current pixel.

$$mc_{match} = \operatorname{argmin}_{g_i \in Background\_set} \left(\frac{|P_t - \mu_i|}{\sigma_i} \right) \tag{16}$$

mc_{match} is mutated in the ARBs generation phase with a mutation function described below. At the end of this phase, a set of Gaussians (clones) are created.

The number of clones is calculated by the following equation:

$$NumClones = Clonal\_rate \times Hyper\_mutation\_rate \times Affinity\left(P_t, mc_{match}\right) \tag{17}$$

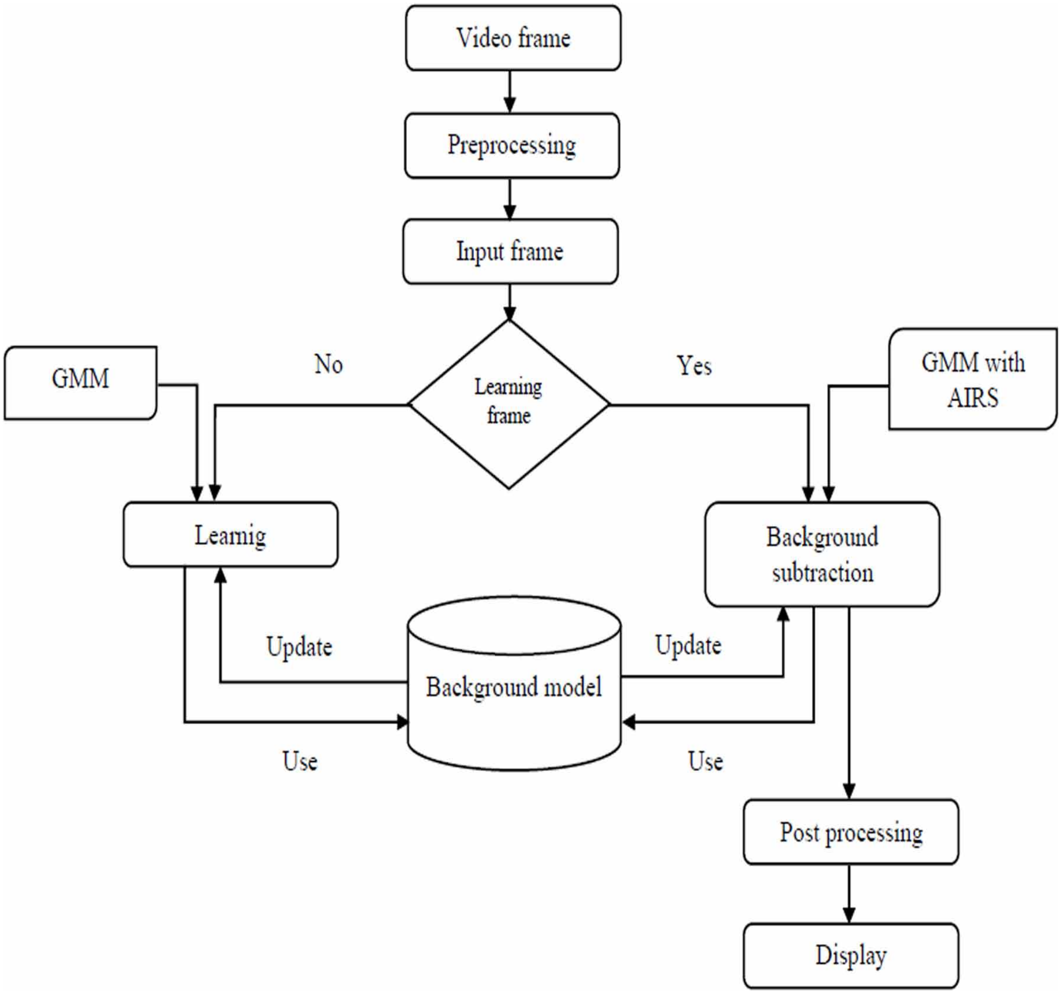
Clonal_Rate and Hyper_mutation_rate are two integer values chosen by the user.

Typically, the Affinity is calculated between the current pixel P_t and a Gaussian, in the Equation (18) it calculated between P_t and the Gaussian mc_{match} (figure 4):

$$Affinity\left(P_t, mc_{match}\right) = \frac{\left|P_t - \mu_{mc_{match}}\right|}{\sigma_{mc_{match}}} \tag{18}$$

$$AB = \left\{ g_{clone_1}, g_{clone_2}, \dots, g_{NumClones} \right\} \tag{19}$$

Figure 3. Global architecture of the proposed system



Mutation_rate is a value between 0 and 1.

New clones must be filtered through competition for resources process, keeping only the best and the correct Gaussians (*ab*) that verified the following condition.

$$\frac{|P_{ab} - \mu_{ab}|}{\sigma_{ab}} < 2.5 \tag{20}$$

The remaining *ab* are selected as a set of memory cell candidate ($MC_{candidate}$) instead to choose only the best *ab* as a memory cell candidate like that indicate in original AIRS.

$$MC_{candidate} = \left\{ ab, \frac{|P_{ab} - \mu_{ab}|}{\sigma_{ab}} < 2.5 \right\} \tag{21}$$

Figure 4. Pseudo-code for Mutation Function used

```
Mutation function (a Gaussian g (pixel value P, mean μ, variance σ), b : Boolean value): g

Begin
 For each g_i ∈ g, i ∈ [2,3] do
     change ← random value between 0 and 1
     changeto ← random value between 0 and 254
     If change < Mutation_rate then
         g_i ← changeto × random value + 1
         b ← true
     End If
  End For
  If b = true then
     change ← random value between 0 and 1
     changeto ← random value between 0 and 255
     If change < Mutation_rate then
         g_1 ← changeto × normalization_value
     End If
  End If
End
```

The last step of the AIRS algorithm is to introduce the memory cells *mc* from the previous set $MC_{candidate}$. This operation consists of choosing the most representative Gaussians among the new Gaussians and adding them to the background model according to the following condition:

$$Affinity\left(P_t, mc_{candidate}\right) < Affinity\left(P_t, mc_{match}\right) \tag{22}$$

If the previous condition is verified, we compare the mean Affinity of mc_{match} and $mc_{candidate}$ with the affinity threshold AT multiplied by the scalar affinity threshold ATS:

$$\frac{Affinity\left(P_t, mc_{candidate}\right) + Affinity\left(P_t, mc_{candidate}\right)}{2} < AT \times \mathrm{AT}S \tag{23}$$

Where AT is the average distance of all background models generated in the learning phase.

$$AT = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} Avrage\_Affinity(P_{i,j}, G_{i,j})}{n \times m} \tag{24}$$

$G_{i,j}$ is a set of Gaussians for the pixel $P_{i,j}$.

If the Equation (23) is satisfied, mc_{match} will be removed from MC set.

After these steps, GMM method is applied on the background model of this pixel to determine whether the pixel belongs to the background or foreground.

Proposition Two (Directed Generation)

This proposition is based on the same principle of proposition one, but we have led the production of the new Gaussians that are generated by the AIRS algorithm only for pixels representing the background instead to precede for all pixels whatever their nature (background or foreground). This mechanism started by classifying the pixels using GMM method (see section two). After this step, if the pixel belongs to the background, the memory cell mc_{match}, that satisfies the Equation (25), is chosen from the first β distribution Gaussians.

$$mc_{match} = \operatorname{argmin}_{g_i \in \beta\_GMM} \left(\frac{| P_t - \mu_i |}{\sigma_i} < 2.5 \right) \tag{25}$$

The mc_{match} is muted to create a *NumClones* novel Gaussians according to the mutation function algorithm. The affinity is calculated between P_t and the Gaussian mc_{match} according to the Equation (18).

The AB set (see Equation (19)) contains all the new Gaussians that are created in the mutation step. During the competition for resources process the *ab* are filtered and only AB set which satisfies Equation (20) is kept. To minimize the used memory space in the memory cell introduction step, only one memory cell ($mc_{candidate}$) is introduced. This latter represents the average of Gaussians of the AB set.

$$mc_{candidate} = Avrage(AB) \tag{26}$$

If the Equation (22) is satisfied, the system introduced the Gaussian $mc_{candidate}$ as a new background model and compare the mean Affinity of mc_{match} and $mc_{candidate}$ with $AT \times ATS$ according to the Equation (23) and Equation (24). The mc_{match} that verifies Equation (23) is removed from the set of background models.

EXPERIMENTAL VALIDATION AND DISCUSSION

The system presented in this paper is implemented in Python on a computer with an Intel Core i7 and 8 GB memory capacity. This section presents the results of our methods on the datasets Wallflower (Toyama et al., 1999) and UCSD (Mahadevan and Vasconcelos, 2009).

After several empirical tests, the learning rate, the minimum part of the data corresponding to the background, Clonal_rate, Hyper_mutation_rate, Mutation_rate, ATS) are respectively fixed to 0.01, 0.3, 10, 2, 0.1, 0.2.

Qualitative Approach

The following table shows the observable results of the propositions based on Wallflower and UCSD datasets (table 1 and 2). Obtained results are compared with images of the ground truth and with some method cited in the state of the art. However, the qualitative result does not describe objectively the system performance (figure 5). For this purpose, quantitative tests are necessary to bravely demonstrate the robustness of the approaches.

Quantitative Approach

In the quantitative approach, the authors have chosen as metrics: Recall, precision, F-measure to evaluate the performance of the propositions. Indeed these three metrics are the popular and the most used in the pattern recognition and information extraction with binary classification (Makhoul et al., 1999; Liang et al., 2015) (Table 3).

- Recall (Re): $\dfrac{TP}{TP+FN}$
- Precision: $\dfrac{TP}{TP+FP}$
- F-measure: $\dfrac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$

Where:

- **True positive (TP):** The result is positive (255), while the ground truth is also positive (255).
- **False positive (FP):** The result is positive (255), but the ground truth is negative (0).
- **False negative (FN):** The result is negative (0), but the ground truth is positive (255).

Discussion

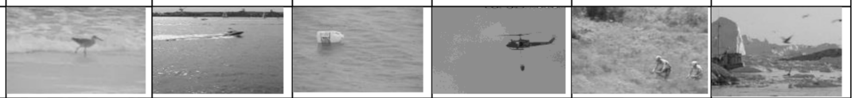
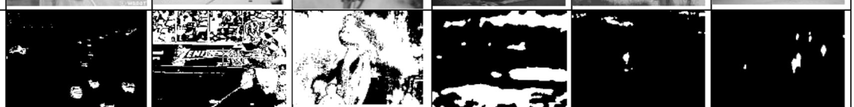
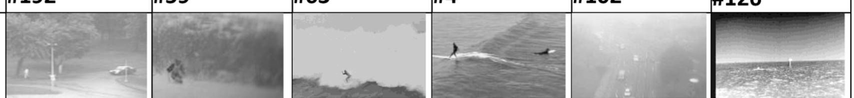
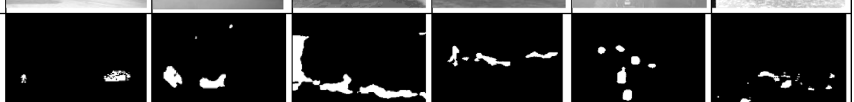
This section discuss the obtained results of the two propositions on Wallflower and UCSD datasets, these results are compared with the most referenced state of art methods in the modeling of the background and with the ground truth (s 4 and Figures 6-8).

Proposition one achieved good results in moved object, Foreground Aperture, Camouflage, Bootstarp, Waving trees videos, but they have some false detection. However, this proposition failed to detect objects in scenes that have a large change in illumination like Light Switch and Time of Day videos. For UCSD dataset, the random generation realized generally acceptable results, indeed, all moving objects are detected, but this proposition has achieved poor detection quality in scenes that have a progressive pixel variation. This due to the random generation of Gaussians.

The second proposition achieved good detection rate for Wallflower dataset compared to the other methods cited in the state of the art, it occupies the third place with a total error of 5180. Indeed, the directed generation has reduced the drawbacks of the first proposition when scenes have a large change in illumination, this is due to the mechanism that generates new Gaussian models only for pixels that represent the background. The obtained results on UCSD clearly show that this proposition has obtained good results compared to the proposition one, since it has detected all moving objects with some false negative in Hockey and Jump videos, this is due to the nature of the videos which does not contain a sufficient number of samples for learning the system.

Noted that these propositions can achieve more efficient results by adding other features. Indeed, the authors used H component of HSV color space as a feature. This choice was based on the capacity of this space compared to the RGB space since it allows to channel the light into a single component (V). Although variations related to light are reduced, only one feature remains insufficient for background modeling. This study used only the H component, since the objective in this work is to propose a new method of background subtraction and not selecting the good discriminator features.

Table 1. Qualitative results on UCSD dataset

Video Frame	Birds #63	Boats #25	Bottle #11	Chopper #51	Cyclists #23	Flock #171
Test Images						
Proposition 1 Random Generation						
Proposition 2 Directed Generation						
Video Name	Freeway #41	Hockey #46	Jump #68	Landing #32	Ocean #151	Peds #100
Test Images						
Proposition 1 Random Generation						
Proposition 2 Directed Generation						
Video Name	Rain #192	Skiing #59	Surf #63	Surfers #4	Traffic #102	Zodiac #126
Test Images						
Proposition 1 Random Generation						
Proposition 2 Directed Generation						

CONCLUSION

This paper proposed two new mechanisms, which allows to reduce the drawbacks of GMM for background subtraction. The idea is to introduce new Gaussian models randomly for all pixels or only for pixels that represent background using Artificial Immune Recognition System. This allows to move from a static to dynamic approach that can easily adapt the model to nature of the environment. Obtained results on several videos from a public benchmark (Wallflower, UCSD) showed the effectiveness of these propositions with small variations in the background. However, the random generation of Gaussians is sensitive when the scene contains high illumination. This is due to the nature of the method that create background models for all pixels (background, foreground pixels).

Based on the promising results of this work, further work is recommended to test the proposed system on another public datasets as CDnet 2014 and to increase the number of features used to have additional discriminative power instead to use only H value of HSV color space.

Table 2. Comparison of qualitative results with well-known background subtraction methods on Wallflower dataset

	MO	TD	LS	WT	Ca	Bo	FA
Test Images							
Ground Truth							
SG (Wren et al., 1997)							
MOG (Stauffer and Grimson, 1999)							
KDE (Elgammal et al., 2000)							
SL-ICA (Tsai and Lai, 2009)							
SL-INMF (Bucak and Gunsel, 2007)							
SL-IRT (Li et al., 2008)							
GMM and Block Matching (Farou et al., 2017)							
MOG with MRF (Schindler and Wang, 2006)							
Proposition 1 Random Generation							
Proposition 2 Directed Generation							

Table 3. Comparison of quantitative results with well-known background subtraction methods on the Wallflower dataset.

Algorithm	Error Type	MO	TD	LS	WT	Ca	Bo	FA	Total	Total Errors
SG	FN	0	949	1857	3110	4101	2215	3464	15696	35133
	FP	0	535	15123	357	2040	92	1290	19437	
MOG	FN	0	1008	1633	1323	398	1874	2442	8678	27053
	FP	0	20	14169	341	3098	217	530	18375	
KDE	FN	0	1298	760	170	238	1755	2413	6634	26450
	FP	0	125	14153	589	3392	933	624	19816	
SL-PCA	FN	0	879	962	1027	350	304	2441	5963	17677
	FP	1065	16	362	2057	1548	6129	537	10649	
SL-ICA	FN	0	1199	1557	3372	3054	2560	2721	14463	15308
	FP	0	0	210	148	43	16	428	845	
SL-INMF	FN	0	724	1593	3317	6626	1401	3412	17073	19098
	FP	0	481	303	652	234	190	165	2025	
SL-IRT	FN	0	1282	2822	4525	1491	1734	2438	14292	17053
	FP	0	159	389	7	114	2080	12	2761	
GMM and Block Matching	FN	0	64	88	367	184	187	89	979	3920
	FP	0	90	605	783	645	457	361	2941	
MOG with MRF	FN	0	47	204	15	16	1060	34	1376	3808
	FP	0	402	546	311	467	102	604	2432	
Proposition 1 Random Generation	FN	0	348	527	3	18	542	786	2224	37293
	FP	0	15686	15678	2000	460	430	815	35069	
Proposition 2 Directed Generation	FN	0	758	1087	60	140	709	227	2981	5180
	FP	0	202	712	173	185	424	503	2199	

Figure 5. Total errors on the Wallflower dataset

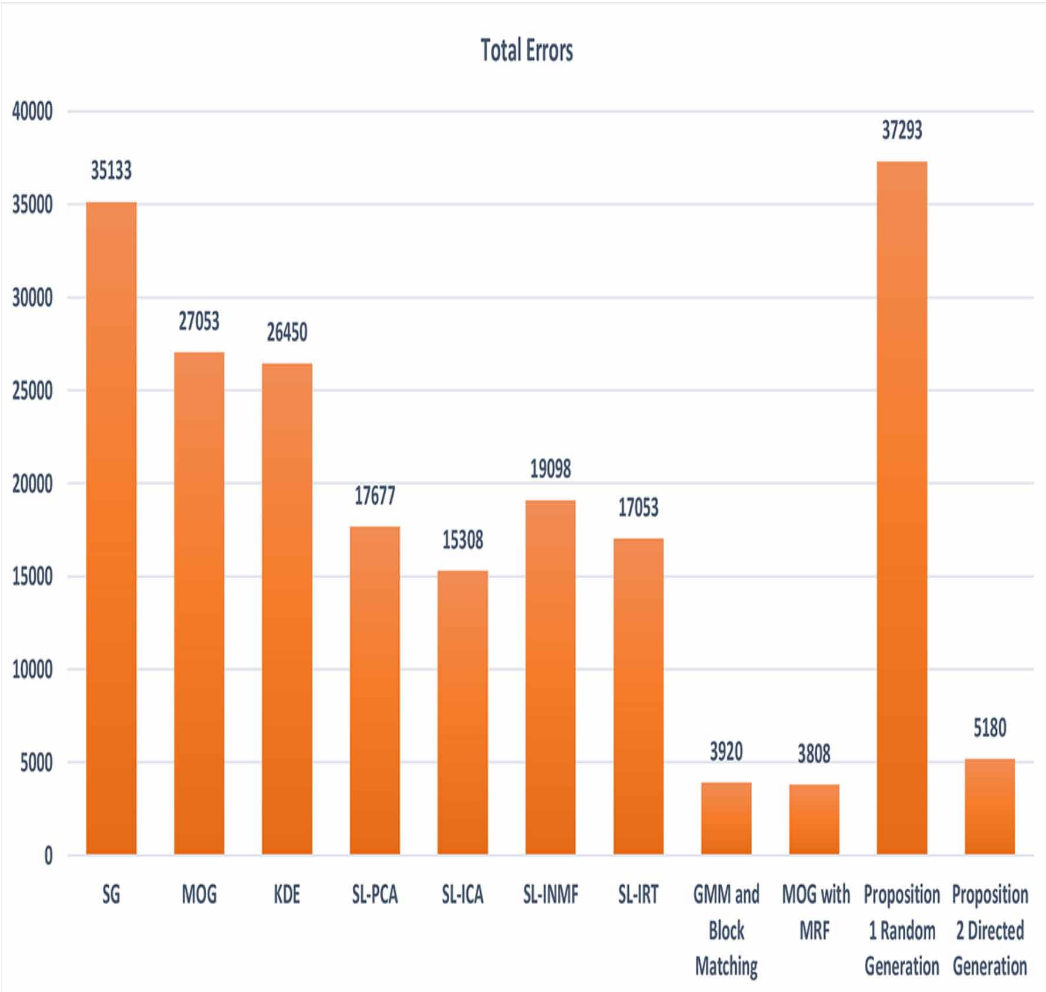


Table 4. Comparison in term of Recall, Precision and F-measure with well-known background subtraction methods on the Wallflower dataset

Algorithm	Error Type	MO	TD	LS	WT	Ca	Bo	FA
SG	Recall	1.000	0.949	0.545	0.835	0.761	0.884	0.807
	Precision	1.000	0.971	0.128	0.978	0.865	0.995	0.918
	F-measure	1.000	0.960	0.207	0.901	0.810	0.936	0.859
MOG	Recall	1.000	0.932	0.849	0.991	0.985	0.904	0.870
	Precision	1.000	0.993	0.232	0.969	0.821	0.947	0.963
	F-measure	1.000	0.962	0.365	0.980	0.896	0.925	0.914
SL-PCA	Recall	1.000	0.954	0.949	0.940	0.980	0.977	0.869
	Precision	0.945	0.999	0.980	0.887	0.918	0.676	0.968
	F-measure	0.971	0.976	0.964	0.913	0.948	0.799	0.916
SL-ICA	Recall	1.000	0.938	0.918	0.823	0.841	0.867	0.855
	Precision	1.000	1.000	0.988	0.991	0.997	0.999	0.974
	F-measure	1.000	0.968	0.952	0.899	0.912	0.928	0.911
SL-INMF	Recall	1.000	0.961	0.916	0.821	0.651	0.926	0.821
	Precision	1.000	0.974	0.983	0.959	0.981	0.989	0.990
	F-measure	1.000	0.968	0.948	0.885	0.782	0.957	0.897
SL-IRT	Recall	1.000	0.933	0.850	0.764	0.922	0.899	0.873
	Precision	1.000	0.991	0.976	1.000	0.994	0.881	0.999
	F-measure	1.000	0.961	0.909	0.866	0.956	0.890	0.932
GMM and Block Matching	Recall	1.000	0.997	0.995	0.980	0.990	0.990	0.995
	Precision	1.000	0.995	0.968	0.958	0.966	0.976	0.981
	F-measure	1.000	0.996	0.982	0.969	0.978	0.983	0.988
MOG with MRF	Recall	1.000	0.997	0.989	0.999	0.999	0.944	0.998
	Precision	1.000	0.979	0.971	0.984	0.976	0.994	0.968
	F-measure	1.000	0.988	0.980	0.991	0.987	0.969	0.983
Proposition 1 Random Generation	Recall	1.000	0.732	0.823	0.999	0.998	0.784	0.841
	Precision	1.000	0.057	0.135	0.746	0.957	0.821	0.836
	F-measure	1.000	0.106	0.232	0.854	0.977	0.802	0.838
Proposition 2 Directed Generation	Recall	1.000	0.416	0.635	0.990	0.986	0.717	0.954
	Precision	1.000	0.728	0.726	0.974	0.982	0.807	0.904
	F-measure	1.000	0.530	0.678	0.982	0.984	0.758	0.928

Figure 6. Recall results on the Wallflower dataset

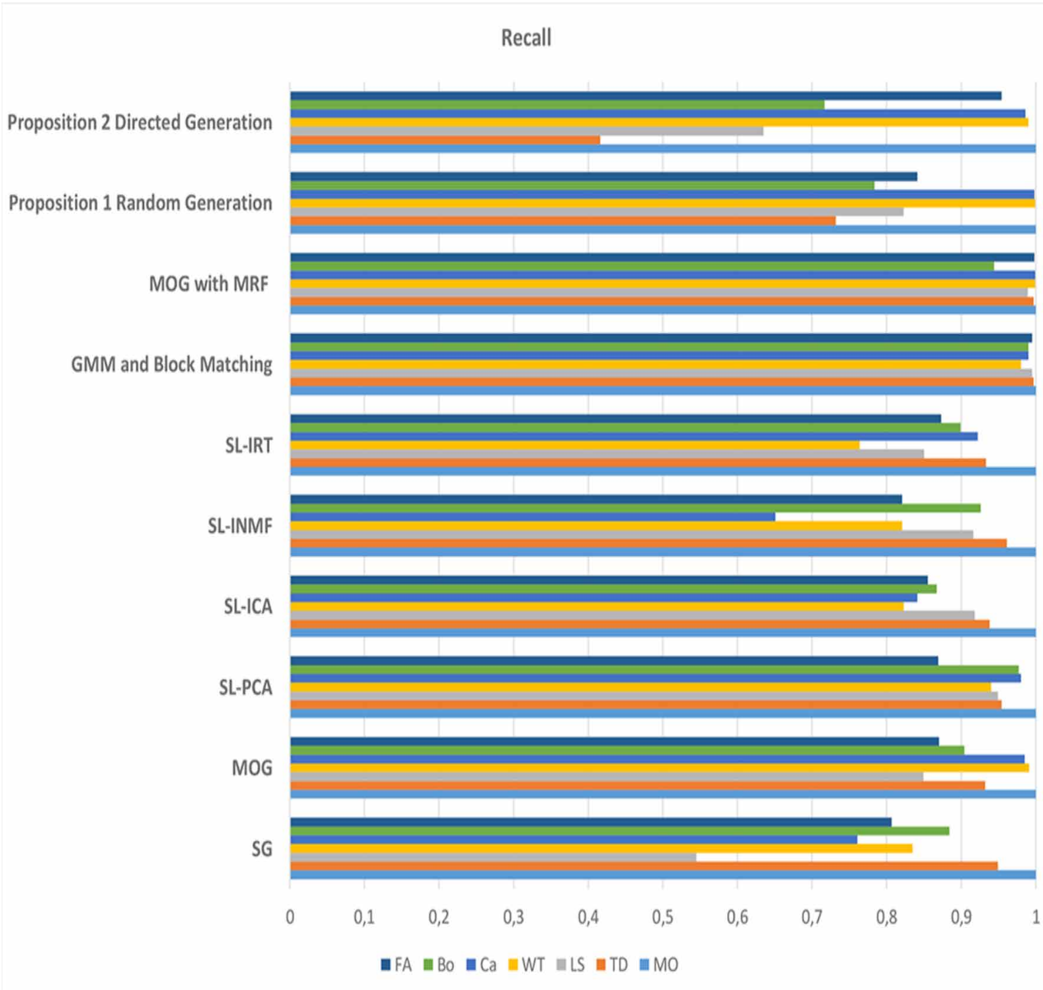


Figure 7. Precision results on the Wallflower dataset

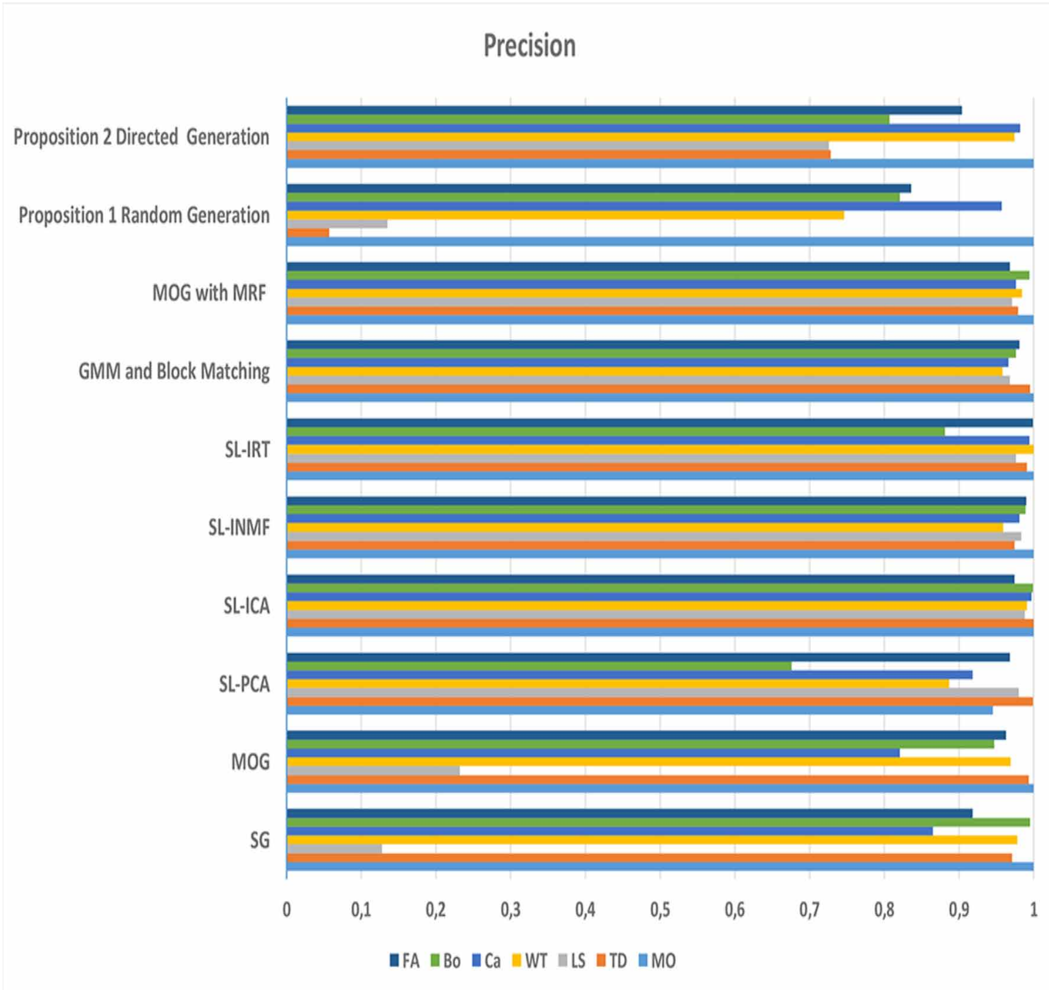
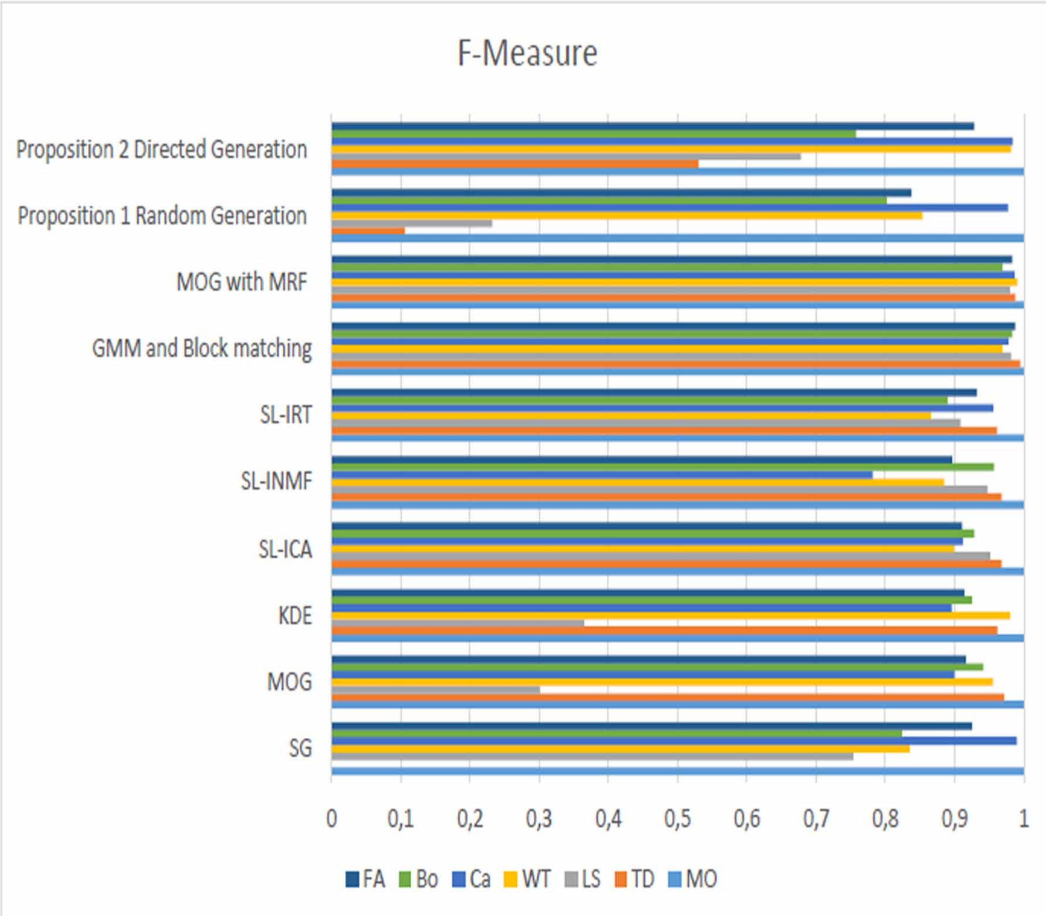


Figure 8. F**-measure results of well-known background subtraction methods on the Wallflower dataset.**



REFERENCES

Allebosch, G., Van Hamme, D., Deboeverie, F., Veelaert, P., & Philips, W. (2015). C-efic: color and edge based foreground background segmentation with interior classification. *Proceedings of the International Joint Conference on Computer Vision*, *Imaging and Computer Graphics* (pp. 433-454). Springer.

Babaee, M., Dinh, D. T., & Rigoll, G. (2018). A deep convolutional neural network for video sequence background subtraction. *Pattern Recognition*, *76*, 635–649. doi:10.1016/j.patcog.2017.09.040

Bouwmans, T. (2012), Background subtraction for visual surveillance: A fuzzy approach, Handbook on soft computing for video surveillance, 5: 103-138.

Bouwmans, T., & El Baf, F. (2010). Modeling of dynamic backgrounds by type-2 fuzzy gaussians mixture models. *MASAUM Journal of of Basic and Applied Sciences*, *1*(2), 265–276.

Brutzer, S., Höferlin, B., & Heidemann, G. (2011). Evaluation of background subtraction techniques for video surveillance. Proceedings of the CVPR 2011 (pp. 1937-1944). IEEE. doi:10.1109/CVPR.2011.5995508

Bucak, S. S., & Gunsel, B. (2007). Video content representation by incremental non-negative matrix factorization. *Proceedings of the 2007 IEEE International Conference on Image Processing* (Vol. 2, pp. 113). IEEE. doi:10.1109/ICIP.2007.4379105

Charoenpong, T., Supasuteekul, A., & Nuthong, C. (2010). Adaptive background modeling from an image sequence by using k-means clustering. Proceedings of the 2010 ECTI International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology ECTI-CON2010 (pp. 880-883). IEEE.

El Baf, F., Bouwmans, T., & Vachon, B. (2008). Fuzzy integral for moving object detection. *Proceedings of the IEEE International Conference on Fuzzy Systems* (pp. 1729-1736). IEEE. doi:10.1109/FUZZY.2008.4630604

El Baf, F., Bouwmans, T., & Vachon, B. (2009). Fuzzy statistical modeling of dynamic backgrounds for moving object detection in infrared videos. *Proceedings of the Computer Vision and Pattern Recognition Workshops* (pp. 60-65). IEEE. doi:10.1109/CVPRW.2009.5204109

Elgammal, A., Harwood, D., & Davis, L. (2000). Non-parametric model for background subtraction. *Proceedings of the European conference on computer vision* (pp. 751-767). Springer.

Farou, B., Kouahla, M. N., Seridi, H., & Akdag, H. (2017). Efficient local monitoring approach for the task of background subtraction. *Engineering Applications of Artificial Intelligence*, *64*, 1–12. doi:10.1016/j.engappai.2017.05.013

Friedman, N., & Russell, S. (1997). Image segmentation in video sequences: A probabilistic approach. *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence* (pp. 175-181). Morgan Kaufmann Publishers Inc.

Haritaoglu, I., Harwood, D., & Davis, L. S. (2000). W/sup 4: Real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(8), 809–830. doi:10.1109/34.868683

Hunziker, S., Quanz, S. P., Amara, A., & Meyer, M. R. (2018). Pcabased approach for subtracting thermal background emission in high-contrast imaging data. *Astronomy & Astrophysics*, *611*, A23. doi:10.1051/0004-6361/201731428

Javed, S., Narayanamurthy, P., Bouwmans, T., & Vaswani, N. (2018). Robust pca and robust subspace tracking: A comparative evaluation. Proceedings of the 2018 IEEE Statistical Signal Processing Workshop (SSP) (pp. 836-840). IEEE. doi:10.1109/SSP.2018.8450718

Jianzhao, C., Victor, O. C., Gilbert, O. M., & Changtao, W. (2017). A fast background subtraction method using kernel density estimation for people counting. *Proceedings of the 2017 9th International Conference on Modelling, Identification and Control (ICMIC)* (pp. 133-138). IEEE. doi:10.1109/ICMIC.2017.8321626

Krungkaew, R., & Kusakunniran, W. (2016). Foreground segmentation in a video by using a novel dynamic codebook. *Proceedings of the 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* (pp. 1-6). IEEE. doi:10.1109/ECTICon.2016.7561253

Liang, D., Hashimoto, M., Iwata, K., & Zhao, X. (2015). Co-occurrence probability-based pixel pairs background model for robust object detection in dynamic scenes. *Pattern Recognition*, *48*(4), 1374–1390. doi:10.1016/j.patcog.2014.10.020

Lim, L. A., & Keles, H. Y. (2018a). Foreground segmentation using convolutional neural networks for multiscale feature encoding. *Pattern Recognition Letters*, *112*, 256–262. doi:10.1016/j.patrec.2018.08.002

Lim, L. A., & Keles, H. Y. (2018b). Learning multi-scale features for foreground segmentation.

Mahadevan, V., & Vasconcelos, N. (2009). Spatiotemporal saliency in dynamic scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(1), 171–177. doi:10.1109/TPAMI.2009.112 PMID:19926907

Makhoul, J., Kubala, F., Schwartz, R., & Weischedel, R. et al.. (1999). Performance measures for information extraction. *Proceedings of the DARPA broadcast news workshop*, pages 249-252. Herndon, VA.

Martins, I., Carvalho, P., Corte-Real, L., & Alba-Castro, J. L. (2017). Bmog: boosted gaussian mixture model with controlled complexity. *Proceedings of the Iberian Conference on Pattern Recognition and Image Analysis* (pp. 50-57). Springer. doi:10.1007/978-3-319-58838-4_6

Oliver, N. M., Rosario, B., & Pentland, A. P. (2000). A bayesian computer vision system for modeling human interactions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(8), 831–843. doi:10.1109/34.868684

Shen, Y., Hu, W., Yang, M., Liu, J., Wei, B., Lucey, S., & Chou, C. T. (2016). Real-time and robust compressive background subtraction for embedded camera networks. *IEEE Transactions on Mobile Computing*, *15*(2), 406–418. doi:10.1109/TMC.2015.2418775

Sigari, M. H., Mozayani, N., & Pourreza, H. (2008). Fuzzy running average and fuzzy background subtraction: Concepts and application. *International Journal of Computer Science and Network Security*, *8*(2), 138–143.

St-Charles, P.-L., Bilodeau, G.-A., & Bergevin, R. (2015a). A self-adjusting approach to change detection based on background word consensus. *Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision* (pp. 990-997). IEEE. doi:10.1109/WACV.2015.137

St-Charles, P.-L., Bilodeau, G.-A., & Bergevin, R. (2015b). Subsense: A universal change detection method with local adaptive sensitivity. *IEEE Transactions on Image Processing*, *24*(1), 359–373. doi:10.1109/TIP.2014.2378053 PMID:25494507

Stauffer, C., & Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. *Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Vol. 2, pp. 246-252). IEEE. doi:10.1109/CVPR.1999.784637

Toyama, K., Krumm, J., Brumitt, B., & Meyers, B. (1999). Wallflower: Principles and practice of background maintenance. *Proceedings of the Seventh IEEE International Conference on Computer Vision* (Vol. 1, pp. 255-261). IEEE.

Tsai, D.-M., & Lai, S.-C. (2009). Independent component analysis-based background subtraction for indoor surveillance. *IEEE Transactions on Image Processing*, *18*(1), 158–167. doi:10.1109/TIP.2008.2007558 PMID:19095527

Wang, K., Gou, C., & Wang, F.-Y. (2018). m{4}cd: A robust change detection method for intelligent visual surveillance. *IEEE Access*, *6*, 15505–15520. doi:10.1109/ACCESS.2018.2812880

Wang, Y., Luo, Z., & Jodoin, P.-M. (2017). Interactive deep learning method for segmenting moving objects. *Pattern Recognition Letters*, *96*, 66–75. doi:10.1016/j.patrec.2016.09.014

Watkins, A., & Boggess, L. (2002). A new classifier based on resource limited artificial immune systems. *Proceedings of the 2002 Congress on Evolutionary Computation CEC'02* (Vol. 2, pp. 1546-1551). IEEE.

Watkins, A. B. (2001). *AIRS: A resource limited artificial immune classifier*. PhD thesis, Mississippi State University Mississippi.

Wren, C. R., Azarbayejani, A., Darrell, T., & Pentland, A. P. (1997). Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(7), 780–785. doi:10.1109/34.598236

Xia, H., Song, S., & He, L. (2016). A modified gaussian mixture background model via spatiotemporal distribution with shadow detection. *Signal, Image and Video Processing*, *10*(2), 343–350. doi:10.1007/s11760-014-0747-z

Yu, J., Zhou, X., & Qian, F. (2010). Object kinematic model: A novel approach of adaptive background mixture models for video segmentation. *Proceedings of the 2010 8th World Congress on Intelligent Control and Automation* (pp. 6225-6228). IEEE.

Zhao, Z., Bouwmans, T., Zhang, X., & Fang, Y. (2012). A fuzzy background modeling approach for motion detection in dynamic backgrounds. In *Multimedia and signal processing* (pp. 177–185). Springer. doi:10.1007/978-3-642-35286-7_23

Wafa Nebili received her Bachelor's degree with very honors in 2013, from the University of 08 Mai 1945 Guelma, Algeria, and the Master's degree from the University of 08 Mai 1945 Guelma, Algeria in 2015, both in Computer Science with very honors. Currently, she is PhD student in computer science at University of 08 Mai 1945, Guelma, Algeria. She is member in the Laboratory of Science and Information Technologies and Communication "LabSTIC," http://labstic.univ-guelma.dz/fr. She is interested in research on video surveillance, data mining, Big Data, and artificial intelligence.

Brahim FAROU received State Engineer degree in computer science systems from National School of Computer Science (Algiers, Algeria) in 2006 and the Magister degree in the sciences and technologies of information and communication from Guelma University in 2009. He received his DSc in Computer Science with distinction in 2016 and the HDR degree with distinction in 2018 from the University of Annaba, Algeria. He is currently associate professor in the computer science department, Guelma University and GADM team member at LabSTIC laboratory. He also occupied many administrative positions, he was deputy head of department responsible for teaching from 2010 to 2012 and he is currently deputy head of the department responsible for post-graduation. His research interests include color constancy, video mining, object recognition, object tracking, surveillance systems and computer vision.

Hamid Seridi received his Bachelor's degree with honours in 1981, from the University of Annaba, Algeria, and the Master's degree from the Polytechnic Institute of New-York, USA in 1984, both in Electrical Engineering. He received his PhD in Computer Science with distinction in 2001 from the University of Reims, Champagne Ardenne, France. He was Vice Dean of the Post-Graduation, Scientific Research and External Relations in the University of Guelma. Currently he is Professor and Director of Laboratory of Science and Information Technologies and Communication "LabSTIC," http://labstic.univ-guelma.dz/fr. He is also Chairman of the Scientific Council of the Faculty of Mathematics and Computing and Material Sciences. He is an expert member at the national committee for evaluation and accreditation national projects research. His research interests include approximate knowledge management, pattern recognition and artificial intelligence, data mining, video mining, machine learning, and cryptography.