CANDID: Robust Change Dynamics and Deterministic Update Policy for Dynamic Background Subtraction

24th IEEE International Conference on Pattern Recognition (ICPR), Beijing, China, pp. 2468-2473, 2018

By: Murari Mandal

CANDID

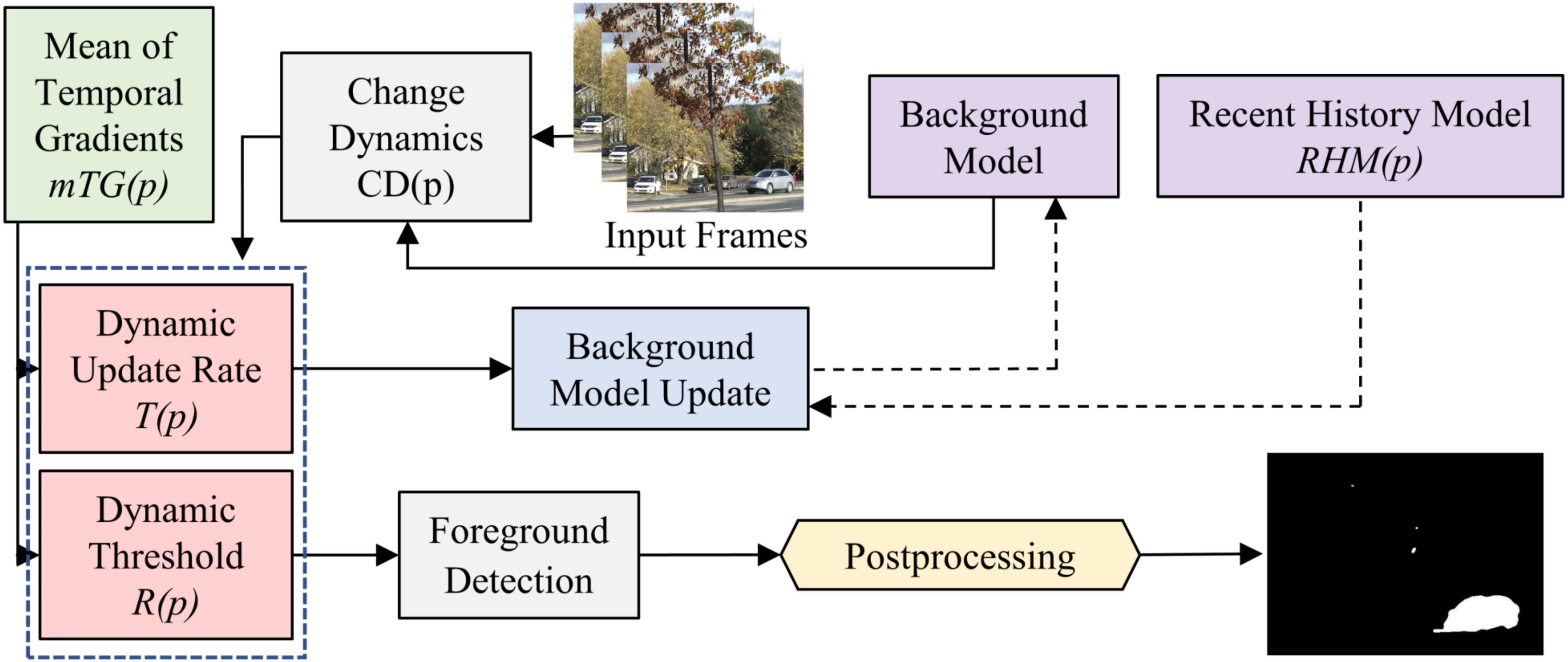


Fig. 1. Block diagram of the proposed method. In this context, *mTG(p)* represents the pixel-level average subtracted values, *CD(p)* is the change dynamics computed at every incoming frame, *RHM(p)* is the recent history model, *R(p)* contains the dynamic threshold for foreground decision making and *T(p)* is the update rate for the background pixels. Both *R(p)* and *T(p)* are controlled by *mTG(p)* and *CD(p)* respectively

State-of-the-art

State-of-the-art	Properties
ViBe [1] (2011)	• The authors proposed three important background model update strategies: random sample replacement, memoryless update policy, spatial diffusion via background sample propagation. • They further used a constant threshold and static update rate for foreground detection and background model maintenance
PBAS [2] (2012)	• The authors introduced dynamic controllers to update the per-pixel decision thresholds and learning rates.
SuBSENSE [3, 4] (2015)	• The SuBSENSE computes the pixel-level spatiotemporal feature descriptor LBSP [3], color channel intensity and incorporates the adaptive feedback information to perform background subtraction. • The adaptive feedback mechanism continuously monitors the model fidelity and segmentation entropy to update the decision thresholds, learning rates and background samples.

Technical Comparison

Method	ViBe	SuBSENSE	PBAS	PM
Initial Parameter Selection	Manual	Manual	Manual	Semi-automated (Adaptive Parameter Initialization)
Background Model Update	Non-Deterministic	Non-Deterministic	Non-Deterministic	Deterministic
Sensitivity to dynamic background changes	Sensitive	Non-Sensitive	Non-Sensitive	Non-sensitive
Dynamic Update Rate	No	Yes	Yes	Yes

Technical Comparison

Method	ViBe	SuBSENSE	PBAS	PM
Dynamic Distance Threshold	NO	YES	YES	YES
Color Images	NO	YES	NO	NO
Local Descriptor	NO	YES (LBSP)	NO	NO
Neighborhood pixel BM update	YES	YES	YES	YES

Proposed Method: CANDID

- All these methods use the random sample update policy for background model maintenance.

- According to Charles et al. [4], updating the samples randomly ensures the presence of long-term and short-term history of background representation in the background model.

- However, this approach gives **equal importance** to all the **background samples** and thereby, both **relevant** and **irrelevant samples** have equal probability to be updated.

CANDID

- This leads to insufficient or improper update of the background samples which is a common reason for unsatisfactory results in sample-based approaches.

CANDID

- Motivated by the preceding considerations, in this paper, we propose a new background subtraction technique which employs a **deterministic model update policy based on the observation of recent pixel history behavior**.

- Moreover, in order to **minimize dependence** on **manual parameter tuning** for different visual scenarios, we designed an **adaptive parameter initialization and maintenance scheme**.

CANDID

- The CANDID adaptively initializes the pixel-level distance threshold and update rate.
- These parameters are updated by computing the change dynamics at a location.
- Further, the background model is maintained by formulating a deterministic update policy.
- The performance of the proposed method is evaluated over various challenging scenarios such as **dynamic background** and **extreme weather conditions.**

Mean of Temporal Gradient (mTG)

Let I_k be a video frame of size $P \times Q$ at time k in a video stream $\{I_k\}_{k=1}^{V}$. The pixel coordinates of image I_k is represented as $I_k(a,b)$ $\forall a \in [1,P], \forall b \in [1,Q]$ and V is the length of the video. Then the mTG can be computed using Eq. (1) as below,

$$mTG(a,b) = \frac{1}{(F_n - 1)} \sum_{k=2}^{F_n} | I_k(a,b) - I_{k-1}(a,b) | \qquad (1)$$

where F_n is the total number initial frames selected for parameter initialization.

Mean of Temporal Gradients

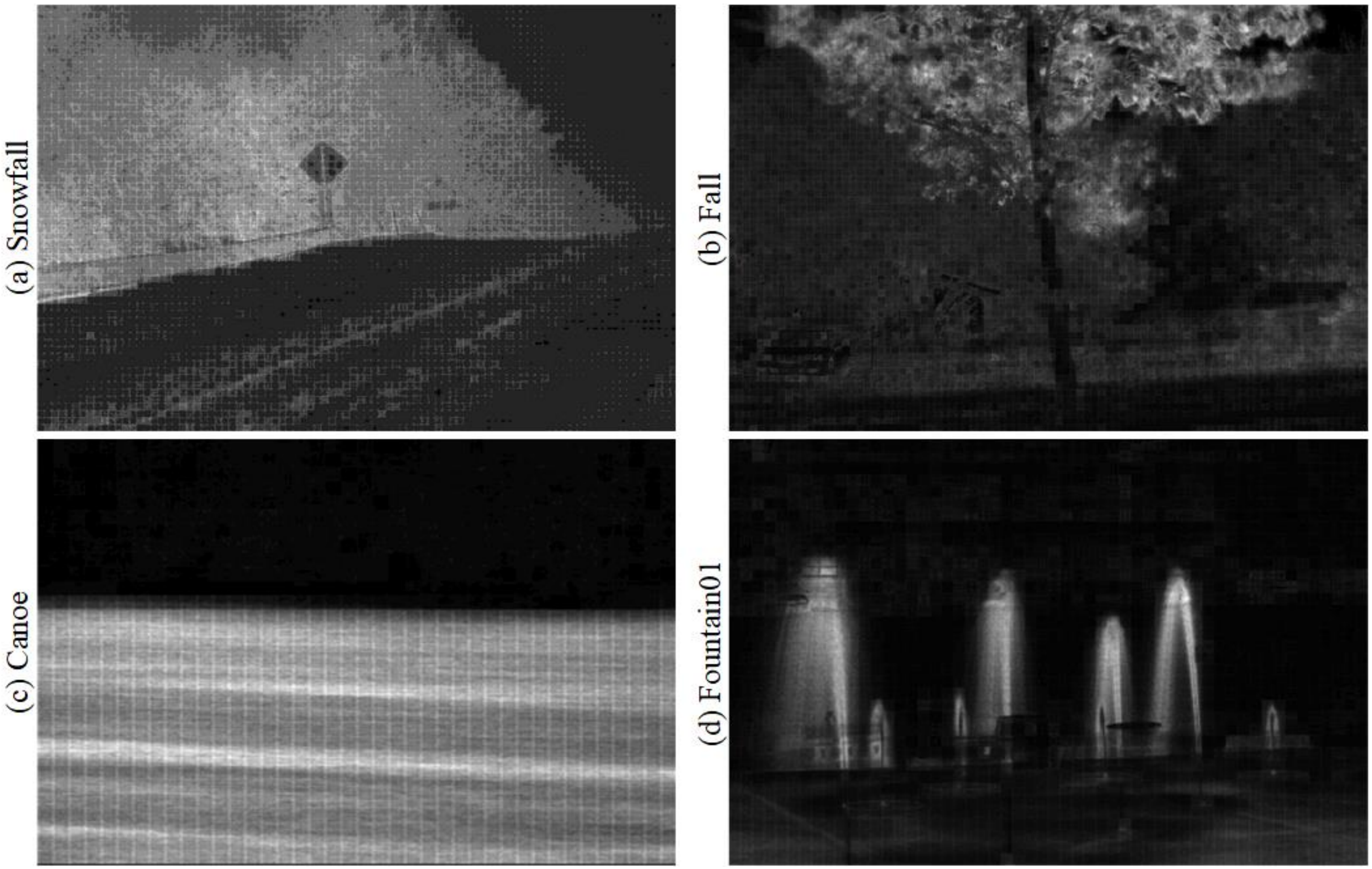


Fig. 2. The mean of temporal gradients (mTG) after initialization using F_n initial frames for (a) Snowfall, (b) Fall, (c) Canoe and (d) Fountain01 videos.

Adaptive Parameter Initialization

The initial parameters are adaptively initialized using Eq. (2) and Eq. (3).

$$R_0(a,b) = mTG(a,b) + \alpha \tag{2}$$

$$T_0(a,b) = \beta \times \frac{1}{(1 + mTG(a,b))^2} \tag{3}$$

where $R_0(a,b)$ is the adaptive distance threshold, $T_0(a,b)$ is the adaptive update rate, α and β are the offset parameters.

Background Model

After parameter initialization, the next task is to initialize the background model (*BM*) as defined in Eq. (4).

$$BM(a,b) = \{s_i(a,b)\}_{i=1}^{N} \tag{4}$$

where s_i is the background sample at index position i and N is the total number of background samples.

The recent history model (*RHM)* is computed in Eq. (5).

$$RHM(a,b) = \{I_i(a,b)\}_{i=1}^{H} \tag{5}$$

where H is the number of recent pixel history samples.

Change Dynamics (CD)

The CD at frame k can be computed as follows,

$$CD_k(a,b) = \frac{1}{2\times L^2} MP_k(a,b) \times [med(\{FS_{k,i}(a,b)\}_{i=1}^{N/2}) + med(\{FS_{k,i}(a,b)\}_{i=N/2+1}^{N})] \quad (6)$$

where $L = 255$ and $med(\cdot)$ is the median of the set. The mean of distances $MP_k(\cdot)$ and the sorted background model distances $FS_k(a,b)$ are computed using [Eq. (7) - Eq. (9)].

$$MP_k(a,b) = 1/N \sum_{i=1}^{N} \{DB_k(a,b)\}_i \quad (7)$$

$$FS_k(a,b) = sort(DB_k(a,b)) \quad (8)$$

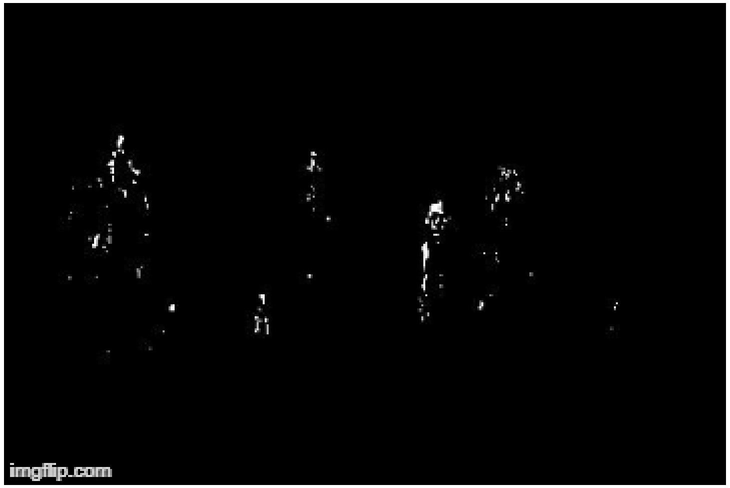
$$DB_k(a,b) = \{|\, I_k(a,b) - BM_i(a,b)\,|\}_{i=1}^{N} \quad (9)$$

Vision Intelligence Lab

Change Dynamics (CD)

Input “fountain01” video

Change Dynamics (CD)



Ground truth

Final results from CANDID

Fig. 3. The change dynamics and final detection results for input video “fountain01” with ground truth from frame no. 450 to 776

Dynamic Distance Threshold

The distance threshold $R_k(a,b)$ at frame k is computed using Eq. (10)

$$R_k(a,b) = \begin{cases} R_0(a,b) + \gamma, & if\ CD_k(a,b) > \xi \\ R_0(a,b), & otherwise \end{cases} \qquad (10)$$

where the γ is the offset parameter to adjust the $R_k(a,b)$ and ξ is the degree of change in the *CD*.

Foreground Detection

The foreground detection is performed as presented in Eq. (11).

$$F_k(a,b) = \begin{cases} 0, & if\ X_k(a,b) < C_{\min} \\ 1, & otherwise \end{cases} \tag{11}$$

where $C_{\min}$ is the minimum number of matches required to label a pixel as background.

In our experiments, we set the parameter $C_{\min} = 2$. The $X_k(a,b)$ can be computed through Eq. (12).

$$X_k(a,b) = \#\{DB_{k,i}(a,b) < R_k(a,b), \forall N\}_i \tag{12}$$

Dynamic Update Rate

The update rate $T_k(a,b)$ is computed using Eq. (13).

$$T_k(a,b) = \begin{cases} 2, & \text{if } CD_k(a,b) > \xi \\ K, & \text{otherwise} \end{cases} \tag{13}$$

where $K = 1/CD_k(a,b)$ and the value of K is bound within the interval $[2, 200]$.

Deterministic BM Update Policy

The recent history model $RHM_k(a,b)$ is utilized to determine the update procedure for the background model $BM_k(a,b)$. For this purpose, we first compute the $RDist_k(a,b)$ using Eq. (14).

$$RDist_k(a,b) = mean(RHM_k(a,b)) - I_k(a,b) \qquad (14)$$

If $RDist_k(a,b) > 0$ && $F_k(a,b) == 0$, then the background sample having minimum distance is replaced by the current pixel value. If $RDist_k(a,b) < 0$ && $F_k(a,b) == 0$, then the background sample with maximum distance is replaced by the current pixel value.

CANDID: Algorithm

Let I_k be a frame of size $P \times Q$ in a video stream $\{I_k\}_{k=1}^{V}$ where V is the length of the video. The pixel coordinates of image I_k is represented as $I_k(a,b) \forall (a \in [1,P], d \in [1,Q])$. Then the proposed method CANDID can be represented through Algorithm 1.

Algorithm 1 CANDID

$\alpha = 10, \beta = 50, \gamma = 10, N = 30, \xi = 0.1, F_n = 300$

for $i = 1 : F_n$

  compute the mean of temporal gradients $mTG(a,b)$

end

// parameter initialization

$R_0(a,b) = mTG(a,b) + \alpha$

$$T_0(a,b) = \beta \times \frac{1}{(1 + mTG(a,b))^2}$$

For $i = F_n + 1 : V$

  if $i <= N + F_n$

    initialize background model $BM(a,b)$

    initialize recent history model $RHM(a,b)$

  else

    //background sample distances with the current pixel

    $DB_k(a,b) = \{|I_k(a,b) - BM_u(a,b)|\}_{u=1}^{N}$

    compute change dynamics $CD_k(a,b)$

    update the dynamic threshold $R_k(a,b)$

    foreground segmentation $F_k(a,b)$

    update the dynamic update rate $T_k(a,b)$

    // update background model

    $RDist_k(a,b) = mean(RHM_k(a,b)) - I_k(a,b)$

    if $RDist_k(a,b) > 0 \, \&\& F_k(a,b) == 0$

      update the farthest background sample using $I_k(a,b)$ with $1/T_{k-1}(a,b)$ probability

      update the neighborhood pixel's farthest background sample using $I_k(a,b)$ with $1/T_{k-1}(a,b)$ probability

    else if $RDist_k(a,b) <= 0 \, \&\& F_k(a,b) == 0$

      update the nearest background sample using $I_k(a,b)$ with $1/T_{k-1}(a,b)$ probability

      update the neighborhood pixel's nearest background sample using $I_k(a,b)$ with $1/T_{k-1}(a,b)$ probability

    end

    update recent history model with a first in first out approach

  end

end

Performance Metrics

- Pixel-based Evaluation Metrics

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F\text{-}score = \frac{2*Precision*Recall}{Precision+Recall}$$

$$Specificity(Sp) = \frac{FN}{(FN+FP)}$$

$$PWC = \frac{100*(FN+FP)}{(TP+FN+FP+TN)}$$

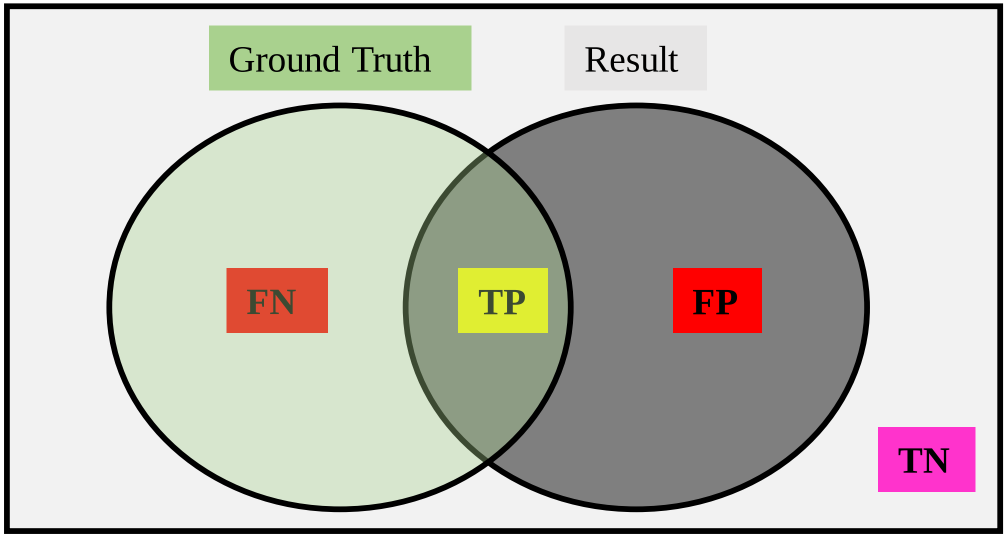


Fig. 4. Pixel based performance evaluation

Quantitative Results

TABLE I

EVALUATION RESULTS OF THE PROPOSED METHOD ON THE DYNAMIC BACKGROUND AND BAD WEATHER VIDEOS FROM CDNET 2014 DATASET

Video	Pr	Re	FM	Sp	PWC
Blizzard	0.93	0.80	0.87	1.00	0.31
Skating	0.96	0.88	0.92	1.00	0.80
Snowfall	0.75	0.77	0.78	1.00	0.39
wetSnow	0.83	0.82	0.83	1.00	0.45
Boats	0.93	0.50	0.66	1.00	0.34
Canoe	0.99	0.81	0.91	1.00	0.70
Fall	0.67	0.97	0.81	0.99	0.91
fountain01	0.38	0.87	0.55	1.00	0.13
fountain02	0.96	0.86	0.92	1.00	0.04
Overpass	0.97	0.84	0.92	1.00	0.25
Avg.	**0.84**	**0.81**	**0.82**	**1.00**	**0.43**

Comparative Performance

TABLE II

COMPARATIVE CHANGE DETECTION PERFORMANCE OF THE PROPOSED METHOD AND EXISTING STATE-OF-THE-ART METHODS BASED ON FM AND PWC OVER THE DYNAMIC BACKGROUND AND BAD WEATHER VIDEOS FROM CDNET 2014 DATABASE

Methods	Metric	blizzard	skating	snowFall	wetSnow	boats	canoe	fall	fount01	fount02	overpass	Avg.
GMM_Grim	FM	0.83	0.86	0.76	0.61	0.73	0.88	0.44	0.08	0.80	0.87	0.69
[10]	PWC	0.36	1.22	0.37	0.98	0.35	0.82	4.05	1.60	0.09	0.33	1.02
GMM_Zivk	FM	0.80	0.84	0.74	0.56	0.26	0.64	0.32	0.05	0.58	0.67	0.55
[12]	PWC	0.40	1.34	0.36	1.02	1.91	3.02	5.61	1.84	0.23	0.98	1.67
KDE	FM	0.54	0.80	0.41	0.12	0.03	0.18	0.08	0.01	0.19	0.24	0.26
[14]	PWC	0.81	2.01	1.24	12.32	35.14	33.19	35.45	13.88	1.72	8.33	14.41
VIBE	FM	0.53	0.71	0.66	0.55	0.22	0.75	0.42	0.09	0.65	0.68	0.53
[1]	PWC	0.75	2.95	0.42	0.91	1.61	1.80	3.30	0.76	0.14	0.82	1.35
PBAS	FM	0.82	0.89	0.73	0.72	0.21	0.40	0.89	0.59	0.90	0.66	0.68
[2]	PWC	0.38	1.03	0.37	0.61	0.56	2.67	0.40	0.10	0.04	0.70	0.69
LOBSTER	FM	0.47	0.78	0.65	0.53	0.58	0.93	0.25	0.16	0.83	0.70	0.59
[4]	PWC	0.81	2.08	0.42	0.89	0.37	0.49	8.90	0.67	0.07	0.99	1.57
SuBSENSE	FM	0.85	0.89	0.88	0.80	0.69	0.79	0.87	0.75	0.94	0.86	0.83
[3]	PWC	0.32	0.95	0.19	0.46	0.31	1.22	0.47	0.05	0.02	0.35	0.43
UBSS1	FM	0.87	0.92	0.85	0.56	0.90	0.93	0.57	0.52	0.92	0.90	0.80
[21]	PWC	0.29	0.72	0.25	1.53	0.11	0.44	2.01	0.06	0.03	0.25	0.57
UBSS2	FM	0.87	0.92	0.85	0.48	0.90	0.93	0.57	0.52	0.92	0.90	0.79
[20]	PWC	0.29	0.72	0.25	2.05	0.11	0.44	2.01	0.06	0.03	0.25	0.62
Spectral-360	FM	0.78	0.92	0.76	0.65	0.69	0.88	0.90	0.47	0.92	0.81	0.78
[24]	PWC	0.43	0.75	0.34	0.94	0.30	0.78	0.37	0.17	0.03	0.46	0.46
IUTIS-1	FM	0.67	0.71	0.76	0.55	0.32	0.41	0.18	0.04	0.74	0.83	0.52
[25]	PWC	0.59	3.44	0.33	1.32	2.02	9.90	14.83	3.39	0.14	0.48	3.64
IUTIS-2	FM	0.63	0.89	0.76	0.73	0.59	0.71	0.30	0.07	0.89	0.88	0.65
[25]	PWC	0.64	0.95	0.33	0.59	0.48	1.96	7.26	1.98	0.05	0.30	1.45
RMoG	FM	0.61	0.79	0.77	0.64	0.83	0.94	0.67	0.20	0.87	0.90	0.72
[13]	PWC	0.66	1.73	0.33	0.77	0.21	0.44	1.23	0.36	0.06	0.25	0.60
SC-SOBS	FM	0.59	0.89	0.65	0.51	0.90	0.95	0.28	0.12	0.89	0.88	0.67
[21]	PWC	0.68	1.04	0.43	1.23	0.13	0.34	8.35	0.93	0.05	0.34	1.35
BingWang	FM	0.73	0.89	0.78	0.67	0.85	0.93	0.63	0.77	0.93	0.95	0.81
[17]	PWC	0.50	0.97	0.30	0.78	0.19	0.53	1.97	0.04	0.03	0.14	0.55
CP3	FM	0.68	0.90	0.74	0.80	0.54	0.91	0.63	0.17	0.64	0.77	0.68
[15]	PWC	1.00	1.00	0.43	0.53	0.87	0.61	1.18	0.61	0.16	0.52	0.69
AAPSA	FM	0.84	0.85	0.78	0.69	0.76	0.89	0.75	0.44	0.36	0.82	0.72
[22]	PWC	0.33	1.31	0.34	0.64	0.24	0.72	0.79	0.11	0.71	0.41	0.56
EFiC	FM	0.73	0.92	0.86	0.57	0.36	0.36	0.72	0.23	0.91	0.88	0.66
[7]	PWC	0.50	0.74	0.21	1.58	0.53	2.88	1.26	0.47	0.04	0.32	0.85
C-EFiC	FM	0.76	0.90	0.87	0.62	0.37	0.34	0.56	0.27	0.93	0.90	0.65
[8]	PWC	0.45	0.94	0.20	1.20	0.50	2.91	2.43	0.37	0.03	0.26	0.93
Graphcut	FM	0.90	0.92	0.88	0.87	0.57	0.12	0.72	0.08	0.91	0.84	0.68
[23]	PWC	0.24	0.76	0.20	0.31	0.52	52.01	1.20	1.10	0.04	0.40	5.68
MultSpat	FM	0.71	0.62	0.71	0.57	0.48	0.89	0.41	0.14	0.82	0.84	0.62
[18]	PWC	0.52	4.78	0.37	0.97	0.70	0.83	4.27	0.51	0.08	0.43	1.35
PM	**FM**	**0.87**	**0.92**	**0.78**	**0.83**	**0.67**	**0.91**	**0.81**	**0.55**	**0.92**	**0.92**	**0.82**
	PWC	**0.31**	**0.80**	**0.39**	**0.45**	**0.34**	**0.70**	**0.91**	**0.13**	**0.04**	**0.25**	**0.43**

Qualitative Results

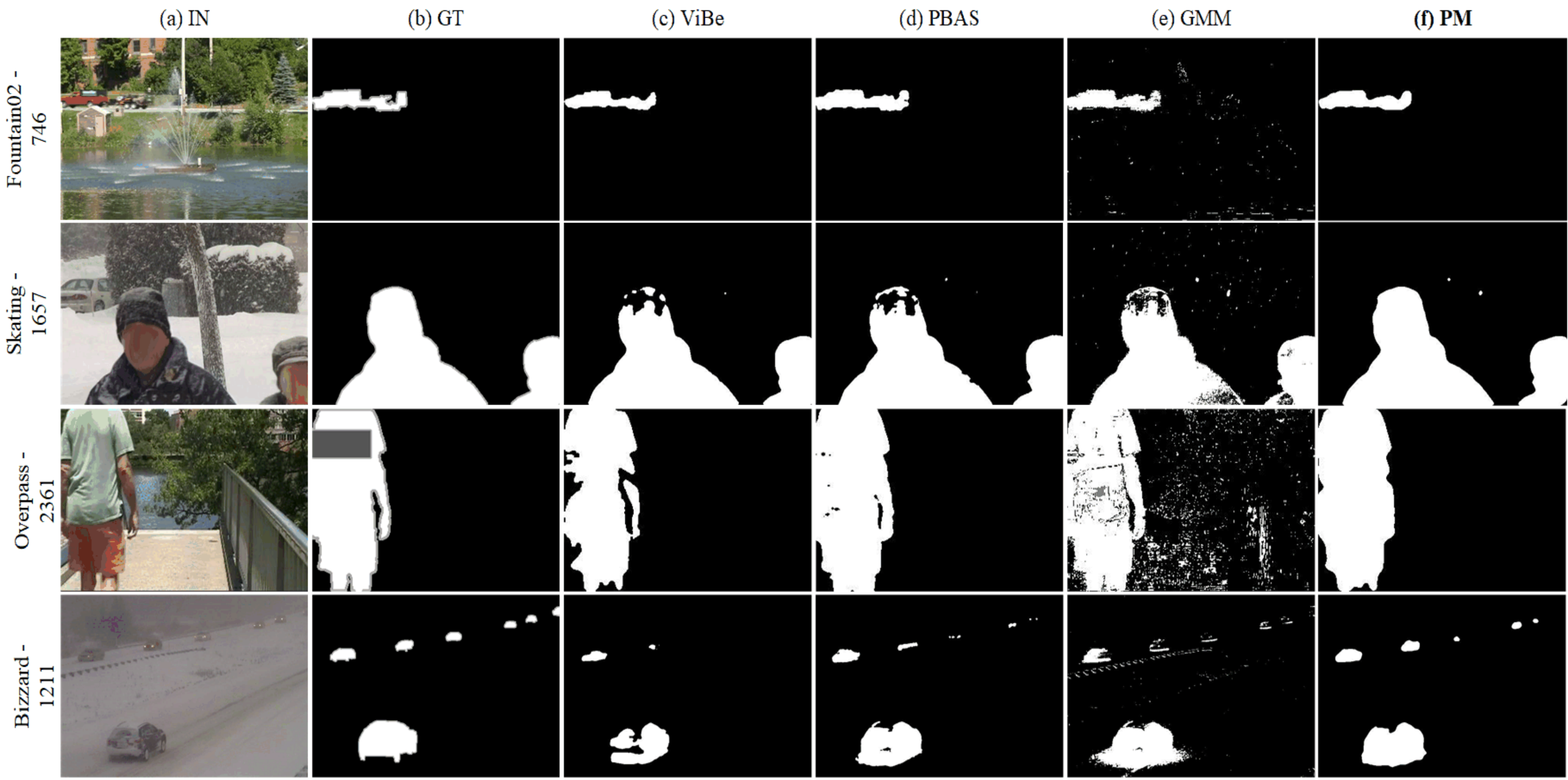


Fig. 5. Detection results of the proposed method and other state-of-the-art methods

References

[1] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1709–1724, Jun. 2011.

[2] M. Hofmann, P. Tiefenbacher, and G. Rigoll. Background segmentation with feedback: The pixel-based adaptive segmenter. In *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, pages 38–43, Jun. 2012.

[3] St. Charles, P. Luc, and G. A. Bilodeau. Improving background subtraction using local binary similarity patterns. In *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, pages 509-515, 2014.

[4] P. L. St-Charles, G. A. Bilodeau and R. Bergevin, "SuBSENSE: A universal change detection method with local adaptive sensitivity," IEEE Trans. Image Process., vol. 24, no. 1, pp. 359-373, 2015.