Background subtraction for night videos

Hongpeng Pan[1,*], Guofeng Zhu[1,*], Chengbin Peng[1,2] and Qing Xiao[3]

[1] College of Information Science and Engineering, Ningbo University, Ningbo, China
[2] Ningbo Institute of Industrial Technology, Chinese Academy of Sciences, Ningbo, China, Ningbo, Zhejiang, China
[3] Electrical Engineering and Computer Science, Leibniz University Hannover, Hanover, Germany
[*] These authors contributed equally to this work.

ABSTRACT

Motion analysis is important in video surveillance systems and background subtraction is useful for moving object detection in such systems. However, most of the existing background subtraction methods do not work well for surveillance systems in the evening because objects are usually dark and reflected light is usually strong. To resolve these issues, we propose a framework that utilizes a Weber contrast descriptor, a texture feature extractor, and a light detection unit, to extract the features of foreground objects. We propose a local pattern enhancement method. For the light detection unit, our method utilizes the finding that lighted areas in the evening usually have a low saturation in hue-saturation-value and hue-saturation-lightness color spaces. Finally, we update the background model and the foreground objects in the framework. This approach is able to improve foreground object detection in night videos, which do not need a large data set for pre-training.

Subjects Algorithms and Analysis of Algorithms, Computer Vision, Data Mining and Machine Learning
Keywords Background Subtraction, Night Videos

Submitted 13 December 2020
Accepted 21 May 2021
Published 10 June 2021

Corresponding authors
Chengbin Peng,
pengchengbin@nbu.edu.cn
Qing Xiao,
qing.xiao@stud.uni-hannover.de

Academic editor
Qichun Zhang

Additional Information and Declarations can be found on page 11

DOI 10.7717/peerj-cs.592

INTRODUCTION

Background subtraction aims to identify moving objects from current video frames with the knowledge of a background model (*Sobral & Vacavant, 2014*). It is a very useful image preprocessing tool in many applications. For example, in video surveillance, background subtraction can improve object tracking and recognition (*Kim & Jung, 2017*). General background subtraction models consist of three parts: background initialization builds the initial background model from a few frames at the beginning. Foreground detection extracts moving objects from the current frame by comparing with the background model, while background maintenance updates the background model (*Bouwmans, 2012*).

Background subtraction has been of interest for to researchers for decades, and most detection algorithms have used the pixel-based approaches. Some researchers consider background pixel values at each location of video frames follow a Gaussian distribution, and others propose the use of the median value of each location as the corresponding background pixel (*Piccardi, 2004*). Some researchers take the pixel color frequency into account using weightless neural networks in an unsupervised mode (*De Gregorio & Giordano, 2017*). For video surveillance systems, Gaussian mixture models have been used to cluster foreground and background pixels, respectively (*Goyal & Singhai, 2017*).

Van Droogenbroeck & Paquot (2012) use a number of post-processing methods, such as eliminating small foreground blobs, to improve the performance. Their work is based on classifying foreground and background pixels and updating the background models accordingly. Probabilistic approaches (Ren, Zhang & Zhang, 2019) along with principle component analysis has also been used (Umer et al., 2021).

To improve the performance, researchers also propose to use descriptors. Some researchers transform the pixels from RGB space into other color spaces to separate color intensity from other color information (Balcilar, Amasyali & Sonmez, 2014; Martins et al., 2017). Another effective descriptor for background subtraction is texture-based local binary pattern (LBP) (Heikkila & Pietikainen, 2006). This descriptor combines neighboring pixels rather than single color information, and can better represent the local information of objects. Local binary similarity pattern (LBSP) is an improved LBP that uses larger patterns. St-Charles, Bilodeau & Bergevin (2015) used spatial information and temporal difference as an LBSP descriptor. Tan & Triggs (2010) developed local ternary patterns (LTP) which categorizes pixels into three threshold values.

Monitoring at night is also important for surveillance systems. Some researchers propose to use contrast analysis to capture local change over time to detect potential objects and then use spatial nearest neighbors to suppress false alarms (Huang et al., 2008). Some others use support vector machines and a combination of Kalman filter for object tracking (Fengliang Xu, Xia Liu & Fujimura, 2005). With a high performance background subtraction algorithm, such object detection algorithms can better focus on the moving objects.

However, most of the background subtraction these approaches are not designed for surveillance systems at night, because dark objects and reflection lights can significantly impact the performance of segmentation algorithms. In this paper, we propose a framework that utilizes multiple feature extractors to improve performance. Our contributions are:

1. We propose an unsupervised framework for background segmentation.
2. We propose a local pattern enhancement method for texture feature extraction in the evening.
3. We propose a combination of feature extractors to effectively obtain foreground objects in night videos.

RELATED WORKS

Most state-of-art algorithms such as SUBSENSE (St-Charles, Bilodeau & Bergevin, 2015), WeSamBE (Jiang & Lu, 2018) and C-EFIC (Allebosch et al., 2015) use the pixel-wise RGB descriptors, which are effective for describing the change of objects. Such algorithms have a similar ability to consider the absolute difference of color intensity between the current frame and the background model only. Unfortunately, the RGB descriptor is sensitive and may mistakenly classify the foreground. For example, strong lighting in night videos can become false positive foreground objects.

Researchers have proposed local patterns, including Local Binary Pattern(LBP) (Heikkila & Pietikainen, 2006), Local Ternary Pattern(LTP) (Tan & Triggs, 2010) and Local Binary

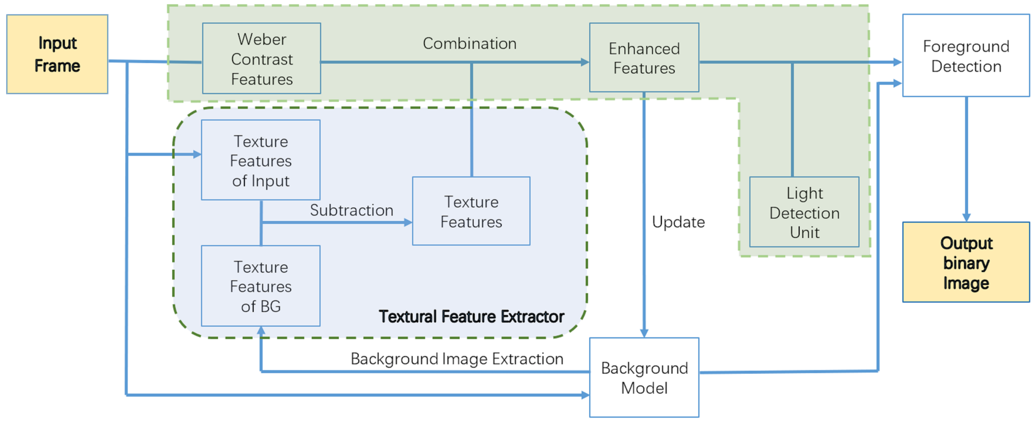

Figure 1 Framework of WCLPE.

Full-size DOI: 10.7717/peerjcs.592/fig-1

Similarity Pattern (LBSP) (*Bilodeau, Jodoin & Saunier, 2013*) to obtain local textural information. One common characteristic of those patterns is that the local difference information is given as ± 1 or 0 according to the threshold T_h.

Many model updating mechanisms have been proposed to date. *Wang & Suter (2007)* introduced the idea that the samples will be replaced according to its lifespan. Another effective solution (*Barnich & Van Droogenbroeck, 2011*) is using a stochastic strategy to update background samples. Both of these strategies are lacking evidence to decide whether current selected images should be replaced. An alternative solution was introduced by *Jiang & Lu (2018)* depends on the weight of each image.

We propose a framework that utilizes and develops multiple feature extractors and can significantly improve the performance of background subtraction in night videos.

METHOD

The framework of our proposed approach is shown in Fig. 1 and can be divided into four parts, namely: the Weber contrast descriptor, an enhanced local feature extractor, a light detection unit, and a background model.

Firstly, we define a background model. Video frames are captured and some of them are stored as "samples" in the pixel-level model, and these samples as a whole form our background model. We store a maximum of twenty-five samples in our background model. Secondly, we extract Weber features and texture features of input images, and compare with images in the background models, to obtain an enhanced representation of texture features through background subtraction. Finally, after eliminating the influence of illumination through the light unit detection, we separate the foreground objects through the foreground detection and output the results.

Weber contrast descriptor

As an example, in Fig. 2, area (3) contains a dim foreground object which has a small color deviation from the background samples, and the threshold to distinguish it from

Figure 2 Winter Street #001225 (1), (2), (3): dim foreground object. (4), (5): strong lighted area.
Full-size DOI: 10.7717/peerjcs.592/fig-2

background should be small. In area (2), the pixel changes in the foreground object is much greater, so the threshold should be larger.

To address this problem, we propose a more effective pixel-wise descriptor named the Weber contrast descriptor, which is a term borrowed from the Weber-Fechner law (*Fechner, 1860*; *Fechner, 1966*). This law indicates that visual systems have different sensitivity to changes in lightness when lightening is different (*Fechner, 1860*; *Fechner, 1966*). For example, in dim areas, our visual perception system can catch subtle change but in bright areas, it is more difficult to perceive same changes. With this idea, we define our Weber contrast descriptor in a simplified version as follows:

$$W = \frac{\Delta I}{I}, \tag{1}$$

where ΔI is the actual change, namely, the deviation of intensity between background and current frame, and I is intensity of current frame. Equation (1) shows that objects in lighted area have relatively low Weber contrast values while objects in dim areas have relatively high values. Therefore, the detection of dim foreground objects is more effective, and the impact of lighting is not too great with the improved descriptor. The Weber contrast features are shown in Fig. 3B.

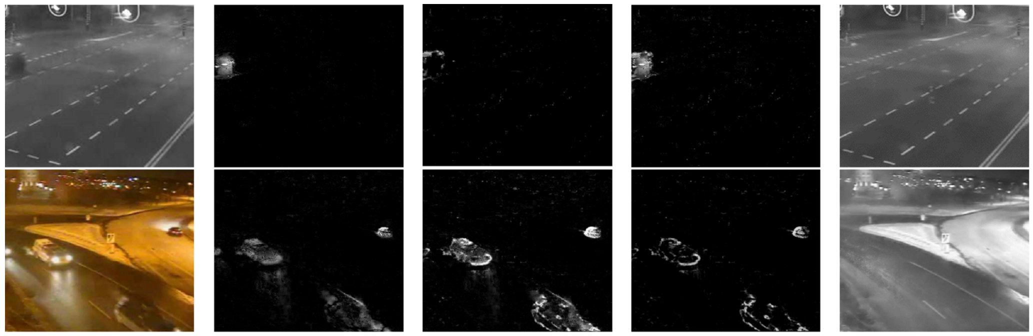

Figure 3 **Feature maps.** Input images: streetCornerAtNight-#002834 and winterStreet-#001225 (*Wang et al., 2014*).

Full-size DOI: 10.7717/peerjcs.592/fig-3

Textural features

Local pattern enhancement

The Weber contrast descriptor requires the elimination of strong lighting influences as well as another light detection method. Hence, we propose another light detection mechanism for further elimination of light.

A lighted area gradually fades in background with a smooth decrease of intensity. On the contrary, the foreground region of interests (FROI) like cars and pedestrians, have obvious silhouettes in common, which means that we can distinguish FROI and lights from changing areas using this edge features.

We seek to obtain more local textual information to better describe the silhouette of FROI. Inspired by local binary patterns, we design a local pattern that includes more detailed information, as shown in Fig. 4 to resolve the problem. With this pattern, we consider the differences between all the marginal pixels and the center pixel, and then summarize the differences. We use the summation as a new descriptor, to indicate local textural features at pixel x,y. Mathematically, it is defined as $\sum_{i=0}^{8s-1}|L_i - C_{x,y}|$, where s is the stride of pattern we have used in Fig. 4. In order to satisfy the condition to detect dim objects without strong lighting, the calculation of $A(x,y)$ is modified as:

$$A(x,y) = \sum_{i=0}^{8s-1} |L_i - C_{x,y}| * \frac{I_d}{max(C_{x,y}, T_c)}, \tag{2}$$

where I_d is intensity degree, meaning that when the brightness is less than I_d, the textural features should be enhanced. Otherwise, in high color intensity areas, the textural features will be suppressed by the multiplication of a small coefficient. T_c corresponds to threshold of camouflage. In low intensity areas, $A(x,y)$ is very large because of the additional term, so we set the threshold T_c at 75 by default to avoid such large value. We need to normalize $A(x,y)$ to combine the textural features with the Weber contrast features, which is shown in the following equation:

$$A_{norm}(x,y) = \frac{A(x,y)}{255}, \tag{3}$$

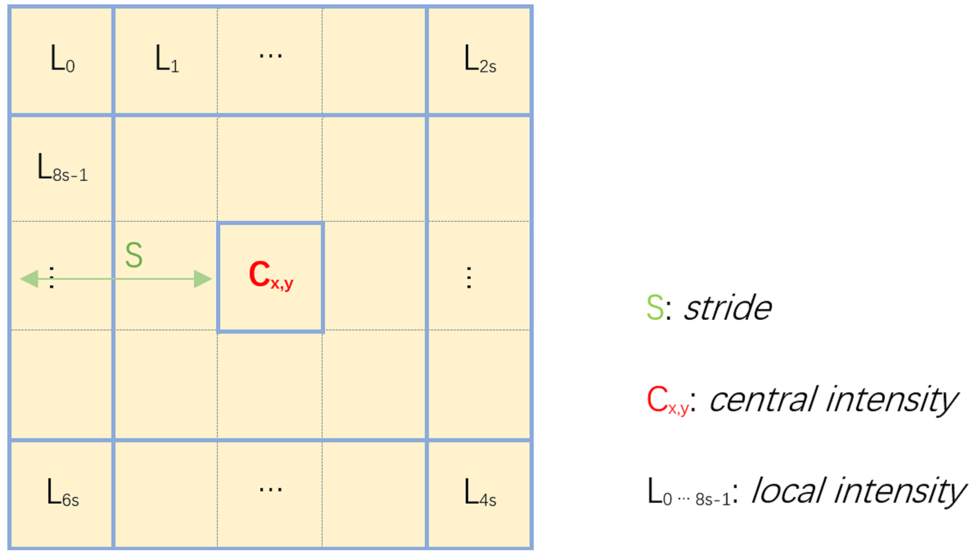

Figure 4 **Local pattern enhancement.**

Full-size DOI: 10.7717/peerjcs.592/fig-4

where s is stride with default value two. This is the Local Pattern Enhancement (LPE) approach.

Textural feature extraction

The local binary features based on LBP or LBSP can be stored in one or two bytes. Similarly, the LTP pattern can be separated into upper LBP and lower LBP meaning that each image could include color information and textural features together without utilizing a large amount of memory. In order to maintain more information in the feature, we consider extracting the background image from background model, and extracting the local textural features from the background image and the current frame. Samples with the most similar are stored in the background model. To better represent the background, each pixel of the background image was taken from the average value of the corresponding pixels from N background samples ($N = 30$), as shown shown in Fig. 3E. Finally, to reduce the noise, the calculation of the final utilized textural feature A_{final} is based on the textural feature difference of the current frame and the background model, as below:

$$A_{final}(x,y) = min\{\sum_{i=-1}^{1}\sum_{j=-1}^{1} max[0, A_{norm}^{cur}(x,y) - A_{norm}^{bg}(x+i,y+j)]\} \quad (4)$$

where A_{norm}^{cur} and A_{norm}^{bg} are the normalized amplitudes of the current frame and the background image, while x and y are the positions in the images. This rule assumes that the background image is smoother than the current frame. The final utilized textural feature maps are shown in Fig. 3C.

Light detection unit

We introduce detection of the light in night videos by combining an intrinsic attribute of color and saturation. There are many alternative representations of the RGB color model

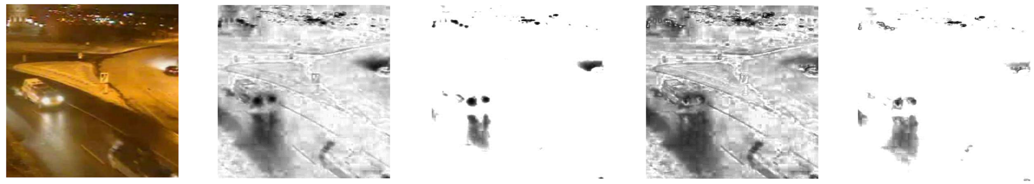

Figure 5 Saturation maps. Saturation maps. Input image: winterStreet-#001225 (*Wang et al., 2014*).
Full-size DOI: 10.7717/peerjcs.592/fig-5

according to different applications, HSV (Hue-Saturation-Value) or HSL (Hue-Saturation-Lightness) (*Joblove & Greenberg, 1978*). Color space is one of such representations. The lighted areas in night videos are similar to natural light, which has a low saturation. Thus, the lighted areas can be separated from other areas in HSV/HSL color space by using a specific threshold on saturation.

In our case we only consider the terms saturation (S) and value (V). The definition of saturation in HSV and HSL is only slightly different in that the colors with maximal saturation locate at lightness 0.5 in HSL, while they locate at value 1 in HSV. Light sources typically have low saturation but high value in HSV, such as headlights belong to foreground objects. Coincidentally, in HSL they have a relative high saturation just like other non-lighting objects. Thus, HSL is more suitable for our approach. Figure 5 shows the saturation maps calculated in HSV and HSL color space.

White or almost-gray objects have low saturation and the direct use of saturation will cause more false negatives. According to the Dark Channel Prior (*He, Sun & Tang, 2011*), in normal objects, there is at least one low intensity channel, but in lighted objects, characteristic are more like sky patches. Therefore, we can improve the saturation S_{cor} as follows:

$$S_{cor}(x,y) = \begin{cases} 1 & \text{if } \dfrac{maxRGB(x,y) - minRGB(x,y)}{N_{all}} < T_d \\ \dfrac{S(x,y)}{minRGB(x,y)} & \text{otherwise} \end{cases} \tag{5}$$

where S represents saturation in HSL, N_{all} is the number of pixels of an input image, and T_d is an upper boundary, $minRGB(x,y)$ and $maxRGB(x,y)$ are the minimum value and the maximum value of three channels (R, G, and B) for pixel position (x,y). In HSL biconical space, the closer the color vector is to the axis, the smaller saturation it has, and the saturation corresponding to the gray color is zero. Therefore, the light detection unit works only when the RGB information is not close to the gray value. When images taken at night are very close to gray images, the saturation method is not useful and we set $S_{cor}(x,y) = 1$ in this case.

Background match and classification

We have introduced the necessary features in our approach. The matching and background classification strategy are also important and must be combined to work effectively. One of the most important combined features in Fig. 1 is the effective feature which was calculated by following equation:

$$F_{effective} = W * S_{cor} + A_{final}^2, \tag{6}$$

where W, S_{cor} and A_{final} are the matrices corresponding to Weber contrast, corrected saturation and final utilized textural features respectively. The value of each pixel is at the corresponding position in the matrix. In fact, the lighted areas masked by light detection units belong to other motion objects and were not classified as FROI. Thus, their pixels should not be updated into the background model, and the model update must be separated from the classification. Here the additional enhanced features $F_{enhanced}$ for model update was defined as:

$$F_{enhanced} = W + A_{final}^2. \tag{7}$$

The process of classification is similar to SuBSENSE (St-Charles, Bilodeau & Bergevin, 2015), this process is presented in Eq. (8),

$$S_{mask} = \begin{cases} 1 & \text{if } \#\{F_{effective,i} < R, \forall i\} < \#min \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$M_{mask} = \begin{cases} 1 & \text{if } \#\{F_{enhanced,i} < R, \forall i\} < \#min \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where S_{mask} is a segment mask for background segmentation while M_{mask} is a motion mask for model update. R is the matching threshold. The symbol $\#\{.\}$ means the number of true elements. $\#min$ is defined as the minimum number of matches, and the parameter $\# min = 2$ is also used.

Background Model Update

To obtain a dynamic background model, once the pixel of the current detected frame is close to the background samples, this pixel should be updated into the background model. At the same time, the image with the farthest distribution in the background model should be replaced. We set a weight for each image to decide which image should be replaced or not. The similar images are allocated by higher weights, and vice versa. The definition of the weight for samples is as following equations:

$$S_{mask} = \begin{cases} W_i + 2 & \text{if } |I_{i,s} - I_{cur}| < T_{lower} \\ W_i + 1 & \text{if } |I_{i,s} - I_{cur}| < T_{upper} \\ W_i - 1 & \text{otherwise} \end{cases} \tag{10}$$

where $I_{i,s}$ is intensity of the ith image, while I_{cur} is the intensity of current frame. T_{upper} corresponds to threshold R_{color} in Jiang & Lu (2018) whose value is 23, while T_{lower} was given as 10 to increase the weight of highly similar images. Such updating strategy can ensure the high correlation of inner images. The lowest weight of image was replaced.

Additionally, to eliminate ghost objects, we adopt the policy of random updating neighboring pixels with time subsampling (St-Charles, Bilodeau & Bergevin, 2015). It should be noted that the updating policy of update is only suitable for neighboring pixels, and update of the current pixel is triggered when $M_{mask} = 1$.

Table 1 Results of all seven measures.

Videos	Recall	Specificity	fpr	fnr	pbc	Precision	F-measure
tramStation	0.7691	0.9955	0.0045	0.2308	0.9026	0.7737	0.7714
fluidHighway	0.6101	0.9893	0.0107	0.3899	1.7169	0.5000	0.5496
streetCornerAtNight	0.8904	0.9964	0.0036	0.1096	0.4117	0.5402	0.6724
winterStreet	0.6753	0.9906	0.0094	0.3247	1.8713	0.6874	0.6813
busyBoulvard	0.4177	0.9929	0.0071	0.5823	2.7403	0.6828	0.5183
bridgeEntry	0.6692	0.9964	0.0036	0.3308	0.8208	0.7299	0.6982

RESULT

Our approach has been evaluated based on CDnet2014 (*Wang et al., 2014*) to compare it with other algorithms and obtain an objective evaluation. This dataset incorporates 11 categories with a variety of scenes, including challenging weather, shadow, and night videos. We used night videos, including six different videos, namely tramStation, fluidHighway, brigeEntry, busyBoulvard, streetCornerAtnight, and winterStreet.

Parameter Initialization

We initialized a few parameters, and the threshold of the Weber contrast was one of the most important. The Weber contrast descriptor is highly sensitive in dark areas, while in some scenes such dark objects belong to FROI but in other scenes are noises such as shadows. Thus, a proper strategy to balance it is to use the global threshold rather than local threshold for parameter R, because the global threshold considers mainly changing areas. However, the global threshold means that each image has a separate threshold, and it is unfair to compare with other algorithms which are without parameter tuning. Thus, local threshold has been adopted in our approach by choose R as follows:

$$R = C_{coe} - \frac{I_{s\_norm}}{2} \tag{11}$$

where I_{s_norm} is the normalized intensity of samples, and C_{coe} is correction coefficient. In the evaluation, we use $C_{coe} = 0.51$ in tramStation and fluidHighway and $C_{coe} = 0.41$ in streetCornerAtNight, winterStreet, busyBoulvard and bridgeEntry. R is typically bound within the range of [0.12, 0.4]. It means that with the decrease of color intensity, the local threshold R will be increased. Using Eq. (11) we can suppress the high sensitivity in dark areas.

Evaluation

We evaluated our approach using the seven measures and their results are shown in Table 1. The F-measure plays an important role in the evaluation of overall performance. There are results of comparisons with other state-of-art algorithms as shown in Table 2 using series of video frames with ground-truth information. Our approach shows a competitive performance for night videos with dim objects. To illustrate the advantage of our approach, images representing typical conditions such as lighting and dim objects are shown in Fig. 6. For some frames, our approach does not perform as good as other approaches, which is mainly due to the lack in shadow detections. Our approach is implemented in python. This

Table 2 F-measure comparison with other state-of-art algorithms[1].

Algorithms	TS	FH	SC	WS	BB	BE
our approach	0.7714	**0.5496**	**0.6724**	**0.6813**	**0.5183**	**0.6982**
SUBsense	**0.7764**	0.3964	0.6036	0.4516	0.4251	0.3166
C-EFIC	0.7648	0.5480	0.6450	0.6348	0.4729	0.6183
EFIC	0.7621	0.5441	0.6705	0.6077	0.4182	0.5980
WeSamBE	0.7696	0.4432	0.6212	0.5211	0.4406	0.4101

Notes.
[1]The results is based on used ground truth frames, there are:TS:tramStation(#1210 –#1310), FH:fluidHighway(#415 –#655),SC:streetCornerAtNight(#800 –#2999), WS:winterStreet(#900 –#1339), BB:busyBoulvard(#730 –#1744), BE:bridgeEntry(#1000 –#1749).

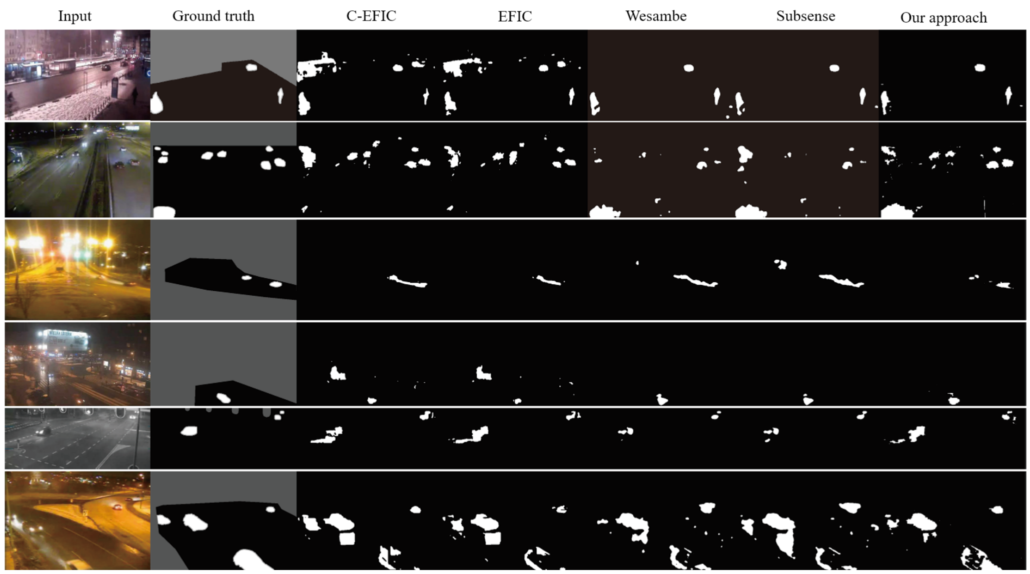

Figure 6 **Result comparison with other algorithms.** Input images from top to bottom are tramStation-#001131, fluidHighway-#000445, bridgeEntry-#001430, busyBoulvard-#001230, streetCornerAtNight-#002834, and winterStreet-#001225 (*Wang et al., 2014*).

Full-size DOI: 10.7717/peerjcs.592/fig-6

program has been optimized through *just in time*(*jit*) based on numba, running with Intel core i5 at 1.6GHz and the processing speed is three frames per second.

CONCLUSION

We proposed a new framework by integrating and improving a number of feature extractors to allow background subtraction. Our framework can enhance foreground object representation and reduce the impact of light reflections for videos in the evening. Our results justifies our approach. In the future, the detection performance can be further improved using temporal information, such as the intensity difference of two near frames.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the National Natural Science Foundation of China (NO. 61802372), the Natural Science Foundation of Zhejiang Province (NO. LGG20F020011), the Ningbo Science and Technology Innovation Project (NO. 2018B10080, 2019B10035), and the Qianjiang Talent Plan (NO. QJD1702031). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Natural Science Foundation of China: 61802372.
Natural Science Foundation of Zhejiang Province: LGG20F020011.
Ningbo Science and Technology Innovation Project: 2018B10080, 2019B10035.
Qianjiang Talent Plan: QJD1702031.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Hongpeng Pan, Guofeng Zhu, Chengbin Peng and Qing Xiao conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The CDnet2014 (2014 Dataset) are available at http://changedetection.net/. Python scripts are available in the Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj-cs.592#supplemental-information.

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. In: *International joint conference on computer vision, imaging and computer graphics*. Cham: Springer, 433–454.

Balcilar M, Amasyali MF, Sonmez AC. 2014. Moving object detection using Lab2000HL color space with spatial and temporal smoothing. *Applied Mathematics & Information Sciences* **8(4)**:1755–1766.

Barnich O, Van Droogenbroeck M. 2011. ViBe: a universal background subtraction algorithm for video sequences. *IEEE Transactions on Image processing* **20(6)**:1709–1724 DOI 10.1109/TIP.2010.2101613.

Bilodeau G-A, Jodoin J-P, Saunier N. 2013. Change detection in feature space using local binary similarity patterns. In: *Computer and Robot Vision (CRV), 2013 international conference on*. Piscataway: IEEE, 106–112.

Bouwmans T. 2012. Background subtraction for visual surveillance: a fuzzy approach. *Handbook on Soft Computing for Video Surveillance* **5**:103–138.

De Gregorio M, Giordano M. 2017. Background estimation by weightless neural networks. *Pattern Recognition Letters* **96**:55–65 DOI 10.1016/j.patrec.2017.05.029.

Fechner G. 1860. Elemente der Psychophysik (Elements of Psychophysics). Adler HE, trans. New York: Holt, Rinehart and Winston Inc.(1860/1966).

Fechner GT. 1966. Elements of psychophysics. In: *Obra publica da originaimente em 1860*. Nova York: Holt, Rinehart and Winston.

Fengliang X, Xia L, Fujimura K. 2005. Pedestrian detection and tracking with night vision. *IEEE Transactions on Intelligent Transportation Systems* **6(1)**:63–71 DOI 10.1109/TITS.2004.838222.

Goyal K, Singhai J. 2017. Review of background subtraction methods using Gaussian mixture model for video surveillance systems. *Artificial Intelligence Review* **50**:1–19.

He K, Sun J, Tang X. 2011. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33(12)**:2341–2353 DOI 10.1109/TPAMI.2010.168.

Heikkila M, Pietikainen M. 2006. A texture-based method for modeling the background and detecting moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28(4)**:657–662 DOI 10.1109/TPAMI.2006.68.

Huang K, Wang L, Tan T, Maybank S. 2008. A real-time object detecting and tracking system for outdoor night surveillance. *Pattern Recognition* **41(1)**:432–444 DOI 10.1016/j.patcog.2007.05.017.

Jiang S, Lu X. 2018. WeSamBE: a weight-sample-based method for background subtraction. *IEEE Transactions on Circuits and Systems for Video Technology* **28(9)**:2105–2115 DOI 10.1109/TCSVT.2017.2711659.

Joblove GH, Greenberg D. 1978. Color spaces for computer graphics. In: *ACM siggraph computer graphics, volume 12*. ACM, 20–25.

Kim W, Jung C. 2017. Illumination-invariant background subtraction: comparative review, models, and prospects. *IEEE Access* **5**:8369–8384 DOI 10.1109/ACCESS.2017.2699227.

Martins I, Carvalho P, Corte-Real L, Alba-Castro JL. 2017. BMOG: boosted Gaussian mixture model with controlled complexity. In: *Iberian conference on pattern recognition and image analysis*. Springer, 50–57.

Piccardi M. 2004. Background subtraction techniques: a review. In: *Systems, man and cybernetics, 2004 IEEE international conference on, volume 4*. Piscataway: IEEE, 3099–3104.

Ren M, Zhang Q, Zhang J. 2019. An introductory survey of probability density function control. *Systems Science & Control Engineering* **7(1)**:158–170 DOI 10.1080/21642583.2019.1588804.

Sobral A, Vacavant A. 2014. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Computer Vision and Image Understanding* **122**:4–21 DOI 10.1016/j.cviu.2013.12.005.

St-Charles P-L, Bilodeau G-A, Bergevin R. 2015. 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.

Tan X, Triggs B. 2010. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing* **19(6)**:1635–1650 DOI 10.1109/TIP.2010.2042645.

Umer S, Dawood H, Yousaf MH, Dawood H, Ahmad H. 2021. Efficient foreground object segmentation from video by Probability Weighted Moments. *Optik* **229**:166251 DOI 10.1016/j.ijleo.2020.166251.

Van Droogenbroeck M, Paquot O. 2012. Background subtraction: experiments and improvements for ViBe. In: *Computer vision and pattern recognition workshops (CVPRW), 2012 IEEE computer society conference on*. Piscataway: IEEE, 32–37.

Wang H, Suter D. 2007. A consensus-based method for tracking: Modelling background scenario and foreground appearance. *Pattern Recognition* **40(3)**:1091–1105 DOI 10.1016/j.patcog.2006.05.024.

Wang Y, Jodoin P-M, Porikli F, Konrad J, Benezeth Y, Ishwar P. 2014. CDnet 2014: An expanded change detection benchmark dataset. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. Piscataway: IEEE, 387–394.