# Unidirectional-Edge Detection based Background Subtraction method for Real-time Object Detection in Restricted Environments


1st Whui Kim
*Environment ICT Research Section*
*Electronics and Telecommunications Research Institute (ETRI)*
Daejeon, Korea
khsunkh@etri.re.kr

2nd Ju Derk Park
*Environment ICT Research Section*
*Electronics and Telecommunications Research Institute (ETRI)*
Daejeon, Korea
zdpark@etri.re.kr

3rd Jae Hong Ryu
*Environment ICT Research Section*
*Electronics and Telecommunications Research Institute (ETRI)*
Daejeon, Korea
jhryu@etri.re.kr

4th Byeong Cheol Choi
*Environment ICT Research Section*
*Electronics and Telecommunications Research Institute (ETRI)*
Daejeon, Korea
bcchoi@etri.re.kr

5th Chang Won Lee
*Environment ICT Research Section*
*Electronics and Telecommunications Research Institute (ETRI)*
Daejeon, Korea
changwonlee@etri.re.kr



***Abstract*—Although many studies focus on the deep learning algorithms, traditional image processing and machine learning technologies are still being developed and used as an auxiliary means. For examples, there is background subtraction based object localization. This can reduce the number of deep learning model inference. The most famous background method for subtraction is background subtraction using the GMM-derived MOG, KNN, and MOG2 algorithms. However, these algorithms still use a non-trivial of computing resources on lightweight single board computers. To alleviate this problem, we propose a unidirectional edge detection based background subtraction algorithm in restricted environments. In terms of processing time, proposed algorithm outperformed others. Although the processing time improved significantly, the precision ($77.8927\%$) was only about $1\%$ lower than the best method. These improvements will enable the video surveillance system to be implemented on lightweight single board computer, such as NVIDIA Jetson boards.**



***Index Terms*—background subtraction, mixture of gaussian, k-nearest neighbor, edge detection, object detection**


## I. Introduction

The vision industry is growing every year along with the semiconductor and communication industries. Advances in semiconductor process technology have enabled large-scale database construction and real-time processing. The development of next-generation communication technology has facilitated data collection from multiple agents. Accordingly, deep-learning based vision technology is being actively researched and applied to various centralized monitoring applications for smart cities. [1]

However, deep learning-based artificial intelligence technologies are not actively utilized in limited environments such as edge devices. In specific industries, such as vision surveillance, traditional image processing and machine learning techniques are still used as pre-processing tools for deep-learning base techniques. Representatively, there is object localization based on background subtraction. [2], [3] This can reduce resource usage for deep learning-based model inference.

Most popular methods are background subtraction based on MOG, KNN, and MOG2 derived from Gaussian Mixture Model (GMM). [4]–[6] They used a mixture of Gaussian distributions, k-nearest neighbor for better estimation, and Gaussian mixture probability density to improve performance, respectively. GMM derivative methods utilize Gaussian distribution information from historical images as background information, making it difficult to adapt to rapid changes. R. Javadzadeh et al proposed adaptive background subtraction algorithm using edge detection. [7] However, due to the large number of morphological operations, R. Javadzadeh's algorithm suffers from an increase in processing time. To overcome these problems, we propose a real-time background subtraction algorithm by reducing morphological operations and several edge detection processes.

## II. Proposed Method

GMM-driven Methods make background model from color or gray images. When based on a color or gray image, the characteristics of background model can be highly dependent on external environment such as illumination. In image processing, edge features can be used to avoid feature changes due to illumination over time. To obtain robust features, R. Javadzadeh et al applied edges to the background model.

In this paper, we propose Unidirectional-Edge Detection based Background Subtraction (UDBS), which is optimized version of R. Javadzadeh's method in consideration of highway characteristics. The UDBS method has three steps. First, a background edge model is created only once, depending on the established environment. Next, it performs preprocessing such as edge detection, image subtraction, and morphological operations. Afterwards, we localize objects in the preprocessed background subtraction image.

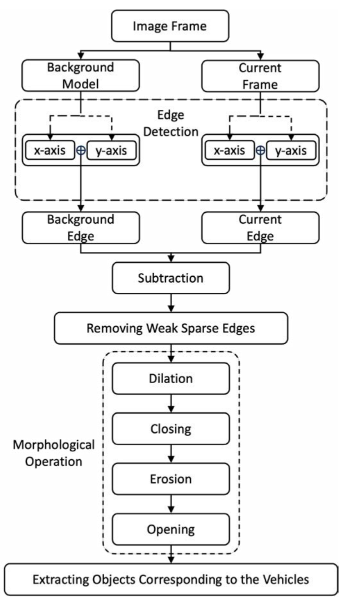


Fig. 1. R. Javadzadeh, et al's method

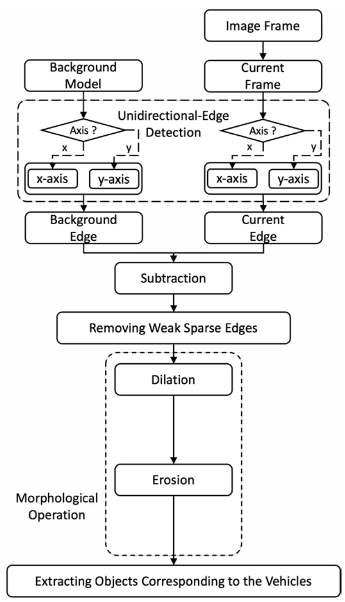


Fig. 2. The proposed method

## A. Preparing Background Edge Model

The UDBS algorithm is based on the premise that the camera is installed so that the vehicle in the video moves horizontally/vertically. Whenever the camera changes or moves, the background edge model is initially created only once. The edges are extracted from an empty highway image using the Prewitt filter along an axis that determined the directions of vehicle movement direction in the image.

## B. background subtraction

The edge of current image is extracted in the same procedure as preparing the background edge model. Next, the procedure for background subtraction and weak sparse edge removal is identical to that proposed by R. Javadzadeh. In the morphological operating procedure, we use only dilation and erosion with large square structural element. Appropriate square structural element enable dilation and erosion to replace closing and opening, in this work.

## C. Localizing Objects

The unidirectional edge detection highlights the front and rear sides of the vehicle more than the other. Depending on the gradient direction of the Frewitt filter used, one side is particularly prominent. In our experiments, it is the bottom side of the vehicle in the image as shown in Fig. 3. The result of finding object contours is the blue boxes in the Fig. 4. Considering that a typical object classification algorithm uses square input, we determine the final bounding box, such as the red boxes in Fig. 4. These bounding boxes are squares that contain the results of finding object contours.

Fig. 3. A sample of background subtraction

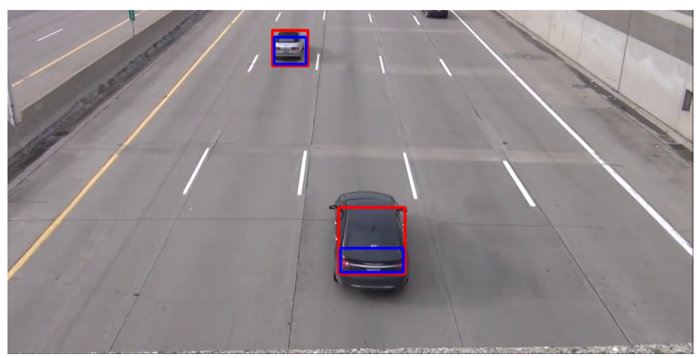

Fig. 4. A sample of object localization

# III. Experiment

## A. Dataset and Environment

We use the Highway CCTV dataset on kaggle. The data set provides images that contain only road surfaces on highways. Each image contains an average of 5.64 vehicle. Experiments are conducted on NVIDIA Jetson Xavier NX Developer Kit (8GB).

## B. Evaluation Metrics

The evaluation method follows mean Average Precision (mAP), which is one of common metrics in the field of object detection. [8] mAP is mean of AP values over classes. AP (5) calculates the average precision ($P_n$) for recall ($R_n$) over 0 to 1. Precision (3) and recall (4) are calculated with true positive ($TP_i$) and false positive ($FP_i$) determined by

Intersection over Union ($IoU_i$) and threshold ($IoU_{threshold}$). IoU (1) measures the overlap between the detected bounding box ($Rect_i$) and ground truth bounding box($Rect_j$). Since the purpose of this paper is dynamic/static object localization, the number of classes is one. Then mAP and AP are the same.

$$IoU_i = max(\frac{Area(Rect_i \cap Rect_j)}{Area(Rect_i \cup Rect_j)}) \quad (1)$$

$$TP_i, FP_i = \begin{cases} 1, 0 & \text{if } IoU_i \geq IoU_{threhold} \\ 0, 1 & \text{otherwise} \end{cases} \quad (2)$$

$$P_n = \frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n} TP_i + \sum_{i=1}^{n} FP_i} \quad (3)$$

$$R_n = \frac{\sum_{i=1}^{n} TP_i}{N} \quad (4)$$

$$\text{AP} = \sum_{n} (R_n - R_{n-1}) P_n \quad (5)$$

### C. Results

As shown in Table I, GMM-based and R. Javadzadeh et al's algorithm have trade-offs of precision and processing time. MOG2 and KNN reduce processing time by about $70\%$ compared to MOG, although their accuracy is reduced by about $5\%$. R. Javadzadeh et al's algorithm, motif of this paper, did not have meaningful improvement over GMM-based methods in terms of precision and processing time. The proposed algorithm showed an precision of $77.89\%$. This is slightly lower than the MOG algorithm, which showed the highest precision. However, it outperformed others in terms of processing time.

TABLE I
THE MEAN AVERAGE PRECISION AND PROCESSING TIME

| | **MOG** | **MOG2** | **KNN** | **R. Javadzadeh** | **UDBS** |
|---|---|---|---|---|---|
| **mAP** | **78.5016** | 73.9998 | 74.3482 | 74.9521 | **77.8927** |
| **ms** | 66.480 | **18.6469** | 19.789 | 34.913 | **11.115** |

Figure 5 show a sample of localized objects. The MOG Method showed localization results (blue) that included some of the object shadow. The localized bounding boxes (green) of the UDBS algorithm was a subset of ground truths (red).

## IV. CONCLUSIONS

This paper presents a unidirectional edge detection based background subtraction algorithm. Our proposed method is superior to other method in terms of processing time since we focus on procedure optimization. This enables the video surveillance system to be implemented on lightweight single board computer, such as NVIDIA Jetson boards. Processing time of our method (11.115 milliseconds) is $40\%$ to $83\%$ less than other methods. Although the processing time improved significantly, the precision ($77.8927\%$) was only about $0.6\%$ lower than the best method.

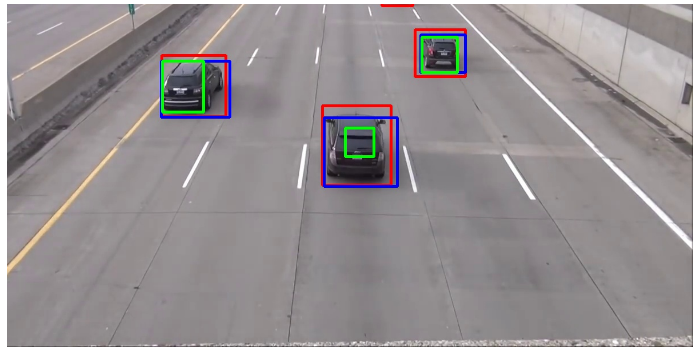

Fig. 5. A sample of localizating objects

In future work, we'd like to add some procedures for practical use to the proposed algorithm. For example, we will add a classifier to categorize objects to vehicle, dropouts, and etc. The hierarchical structure consisting of classifiers and proposed method is expected to show higher precision than the experiments of this study. We will evaluate our method with other challenge datasets, such as Background Models Challenge (BMC).


## ACKNOWLEDGMENT

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 22AMDP-C160501-02).


## REFERENCES


[1] H. Shin, K. I. Na, J. Chang, and T. Uhm, "Multimodal layer surveillance map based on anomaly detection using multi-agents for smart city security," ETRI Journal, vol. 44, issue 2, 2022, pp. 183–193.

[2] P. Alipour, and A. Shahbahrami, "An adaptive background subtraction approach based on frame differences in video surveillance," In 2022 International Conference on Machine Vision and Image Processing, 2022, pp. 1–5.

[3] K. Yun, Y. Kwon, S. Oh, J. Moon, and J. Park, "Vision-based garbage dumping action detection for real-world surveillance platform," ETRI Journal, vol. 41, issue 4, 2019, pp. 494–505.

[4] P. KaewTraKulPong, and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," Video-based surveillance systems: Computer vision and distributed processing, 2002, pp. 135–144.

[5] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," Proceedings of the 17th International Conference on Pattern Recognition, pp.28–31, August 2004.

[6] Z. Zivkovic, and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," Pattern recognition letters, vol. 27, issue 7, 2006, pp. 773–780.

[7] R. Javadzadeh, E. Banihashemi, and J. Hamidzadeh, "Fast vehicle detection and counting using background subtraction technique and prewitt edge detection," International Journal of Computer Science and Telecommunications, vol. 6, issue 10, 2015, pp. 8–12.

[8] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge," International Journal of Computer Vision, vol.88, issue 2, 2010, pp.303–338.