Surveillance Video Fire Detection by using Wavelet and Support Vector Machine

1st Yusron Hanan Zain Vidi Imtinan
Faculty of Informatics
Telkom University
Bandung, Indonesia
yusron@students.telkomuniversity.ac.id

2nd Febryanti Sthevanie
Faculty of Informatics
Telkom University
Bandung, Indonesia
sthevanie@telkomuniversity.ac.id

3rd Kurniawan Nur Ramadhani
Faculty of Informatics
Telkom University
Bandung, Indonesia
kurniawannr@telkomuniversity.ac.id

Abstract
The occurrence of fire is quite often in many countries all over the world. Fire causes harm to human life and human's property. Most of the time, every building potentially to get fire accidentally [1]. That is why fire detection systems have an important role in raising the alarm if a fire occurs. The approach methods in this study will be conducted in four critical steps. First step is Gaussian Mixture Model based background subtraction. Second step is color segmentation to select the candidate regions by using CIE L∗a∗b∗ color space. Third step is extract the candidate regions features in terms of distinguishing between actual fire and fire like objects by using wavelet analysis. Then, fourth step is classifying the candidate regions features to either actual fire or non-fire by using Support Vector Machine (SVM). The result shows the average accuracy is reached 75.463%.

Keywords: surveillance video, fire detection, image processing, CIE L∗a∗b∗ color space, SVM, wavelet

I. INTRODUCTION

A. Background

The occurrence of fire is quite often in many countries all over the world. Fire causes harm to human life and human's property. Most of the time, every building potentially to get fire accidentally [1]. That is why fire detection systems have an important role in raising the alarm if a fire occurs.

There are two ways to detect the occurrence of fire. First by using conventional fire detection, sensor detection, but it has a fatal flaw where they will only work when a certain condition or coverage area has been reached [2]. To solve sensor problems, second option is chosen. Second option is by having surveillance camera being installed, then the advancement of image processing techniques will involve on each frame of the video. By using surveillance camera, it will not have limitation like sensors, it will analyze into each pixel of each frame and able to cover large areas with great results, image processing methods have a faster reaction time than existing sensor systems [3]. In addition, videos can provide more detailed information, such as the location of the fire, size, growth rate, etc. [4].

Several research works have been made in the fire detection system, especially in a way to detect the occurrence of fire in video format and surveillance camera. In [5], the method only based on the chrominance components of the YCbCr color model. The method works fast. However, the proposed method is too simple and only based on YCbCr color model. In [6], it applied Gaussian Mixture Model (GMM) and multi-color feaFtures. The method showed it can achieve average rate of detection around 96%. However, it suffers from high false positive rate. The number of false positive frames is 29% when it tested to video of dancing man with fire-colored shirt [6]. It shows the weakness of this method is only able to detect the fire-color and fire-shape, not the fire-texture. In [7], to extract the features, it use wavelet analysis. The method uses wavelet analysis because it able to extract the candidate region features into their color and texture. The method produces an average accuracy around 97.4% without false alarms. In [8], this survey paper conclude that the fire pixel using CIE L*a*b* color space model is fast and efficient method for fire detection rather than the other color space model. CIE L*a*b* color space model able to reach the recognition rate of 99.88%. Although in [7], there are many classification techniques, the author is considering Support Vector Machine (SVM), because fewer hyper parameters are enough, also the number of training data are less required to achieve high accuracy and require less grid searching to obtain a reasonably accurate model.

The methods in this study will be conducted in four critical steps: (1) Gaussian Mixture Model based background subtraction; (2) color segmentation to select the candidate regions by using CIE L∗a∗b∗ color space; (3) extract the candidate regions features in terms of distinguishing between actual fire and fire like objects by using wavelet analysis; (4) classifying the candidate regions features to either actual fire or non-fire by using Support Vector Machine (SVM)

B. The Problem Statement

The problem statement in this study are:

1. How to implement of detect fire objects on surveillance video by using Wavelet and Support Vector Machine?
2. How is the performance of the fire detection system on surveillance video with the parameters used on Wavelet and Support Vector Machine?

The scope and limitation of this study are:
1. Video data input is offline.
2. The resolution of each video is different.

C. The Objective

The following are the objectives to be achieved in this study:
1. Build a system that can detect fire objects on surveillance video by using Wavelet and Support Vector Machine.
2. Knowing the performance of the fire detection system built with the Wavelet and Support Vector Machine.

This study contains five sections. In Section 2, related work according to the methods are discussed. Then, it will be followed by system design in Section 3. Section 4 provides test results and evaluations of the method. Finally, the author conclude this study in Section 5.

II. LITERATURE REVIEW

A. Related Works

Fire detection systems based on image processing, especially in a way to detect the occurrence of fire in video format and surveillance camera are available in the previous research. In [5], the method only based on the chrominance components of the YCbCr color model. In this research, it has four steps to be able detect fire region in forest by using drone’s camera. First, it will compute the difference between Cr and Cb to be able differentiate fire region in video containing fire. Second, the difference is enhanced by calculate the square of it and normalize the result in range 0 to 255. Then third, binarize it uses automatic threshold to get segmentation of fire region from non-fire region. And last step, to get original fire region, fire region in binary format will be located using connected component analysis and the region is mapped to the original image frames. The method works fast. However, the proposed method is too simple and only based on YCbCr color model. In [6], it applied Gaussian Mixture Model (GMM) and multi-color features. This method has three steps to do video fire detection. First step is motion detection by using GMM. GMM is used to remove background noise and locate moving things in the foreground. Second step is multi-color detection. In this step, to get possible fire areas, it combines RGB, HIS, and YUV color spaces. It will make each frame to be analyzed in RGB, HIS, and YUV rules. Then, third step is by combined the result from first step and second step to get the accurate fire areas. The method showed it can achieve average rate of detection around 96%. However, it suffers from high false positive rate. The number of false positive frames is 29% when it tested to video of dancing man with fire-colored shirt [6]. It shows the weakness of this method is only able to detect the fire-color and fire-shape, not the fire-texture. In [7], to extract the features, it uses wavelet analysis. The method uses wavelet analysis because it able to extract the candidate region features into their color and texture. The method produces an average accuracy around 97.4% without false alarms. In [8], this survey paper concludes that the fire pixel using CIE L*a*b* color space model is fast and efficient method for fire detection rather than the other color space model. CIE L*a*b* color space model able to reach the recognition rate of 99.88%. Although in [7], there are many classification techniques, the author is considering Support Vector Machine (SVM), because fewer hyper parameters are enough, also the number of training data are less required to achieve high accuracy and require less grid searching to obtain a reasonably accurate model.

B. Wavelet

Wavelet is a function in mathematics that can help transform the original image into a frequency, then be divided into sub-band images with different frequency components [9]. Analysis on wavelet is transformed by decomposing the signal into different frequency components. Then, each of these frequency components can be analyzed according to the scale of resolution or level of decomposition. From [10], in image processing, the type of wavelet analysis that can be implement is Discrete Wavelet Transform (DWT). DWT itself support the data that represented by a finite number of values [9], that is why DWT is important and useful in image processing. DWT are efficient and flexible sub-band to decompose the signal. DWT provides information about the frequency and time of a signal that works in multiresolution.

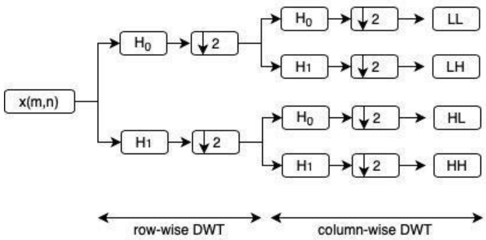


Figure 1. 2D Discrete Wavelet Transform [11]

The transform of DWT is started from vertical transform and horizontal transform. Then, the result will be divided into four blocks who have a same value. In Figure 1 shows 2D DWT. If the decompose process is continue, it will create another sub band with a new level, in Figure 2 shows the example of decompose process that create another new sub band with a new level.

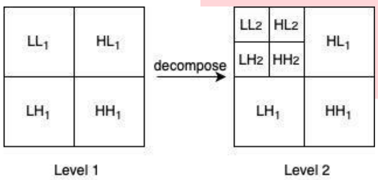


Figure 2. Decompose process create a new sub band [11]

C. Support Vector Machine

Support Vector Machine (SVM) is a machine learning system that uses a hypothetical space in the form of linear functions in a feature space with high dimensions, trained with a learning algorithm based on optimization theory by implementing biased learning derived from statistical learning theory [13]. SVM is a relative technique new compared to other techniques but has better performance in various application fields such as image processing, handwriting recognition, text classification, etc.

The concept of SVM can be explained simply as an effort to find the best hyperplane that functions as a separator of two classes in the input space [13]. Figure 3 shows the SVM illustration. It has several patterns that are members of two classes, namely +1 and -1. The problem of classification can be translated by trying to find the lines or in SVM it is called hyperplane that separates the two class. It uses linear separating hyperplane to create a classifier with a maximal margin [13].

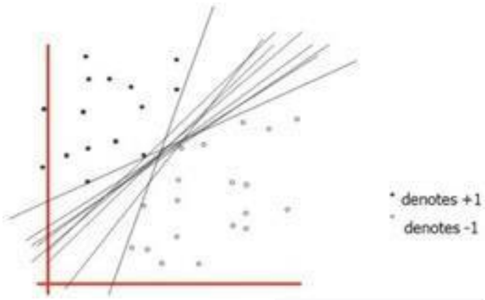


Figure 3. SVM illustration [14]

III. SYSTEM DESIGN

The system in this study is a system that can detect the occurrence of fire contained in the surveillance video data set using Wavelet and Support Vector Machine. In this stage, two processes are obtained, the training and testing process. The training process is a system to build a model, in terms to obtain system parameter values that will be implemented in the test process. Then, in the test process, the model from training process will be tested by unseen parameter values.

A. Data Collection

The dataset in this study is collected from various sources data, from the internet and self-made. For training, the number of dataset is 63 with 26 fire video and 37 non fire video. For testing, it tested to 11 fire video. The self-made dataset is took by using CCTV Outdoor IP Camera 1080p. In Table 1, it is the detailed of the dataset.

Table 1. Scenario of the obtain dataset.

Fire/Non Fire	Scenario	Source
Fire	Took in evening from various angle. In balcony, few moving object other than fire. Environment have low light.	Self-made
Fire	Took in daylight from various angle. In balcony, few moving object other than fire. Environment have enough light.	Self-made and Internet
Fire	Took in afternoon from various angle. . In balcony, few moving object other than fire. Environment have low light.	Self-made
Fire	Took in indoor from various angle. There is no moving object other than fire. Environment have low light.	Self-made and Internet
Fire	Took in outdoor from various time (Day/Night). There are many moving objects other than fire.	Self-made
Non Fire	Took in outdoor from various angle. There are many moving objects. Environment is under the sun, have enough light.	Self-made
Non Fire	Took in outdoor from various angle. There are many moving objects. Environment is night, not have enough light.	Self-made
Non Fire	Took in indoor. There are many moving objects. Environment is have enough light.	Internet

Surveillance Camera is used in many place and environment, that is why the used datasets contains various scenario. From no moving objects rather than fire, few moving objects rather than fire, and many moving objects other than fire. Also, it took in various time, light environment, and angle. Thus, the training model able to recognize the vary type of fire.

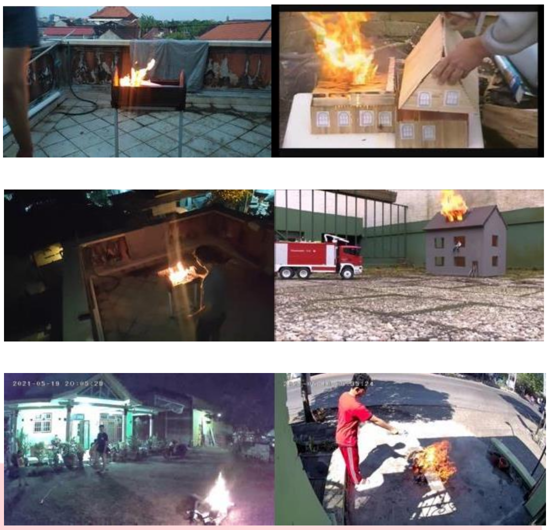
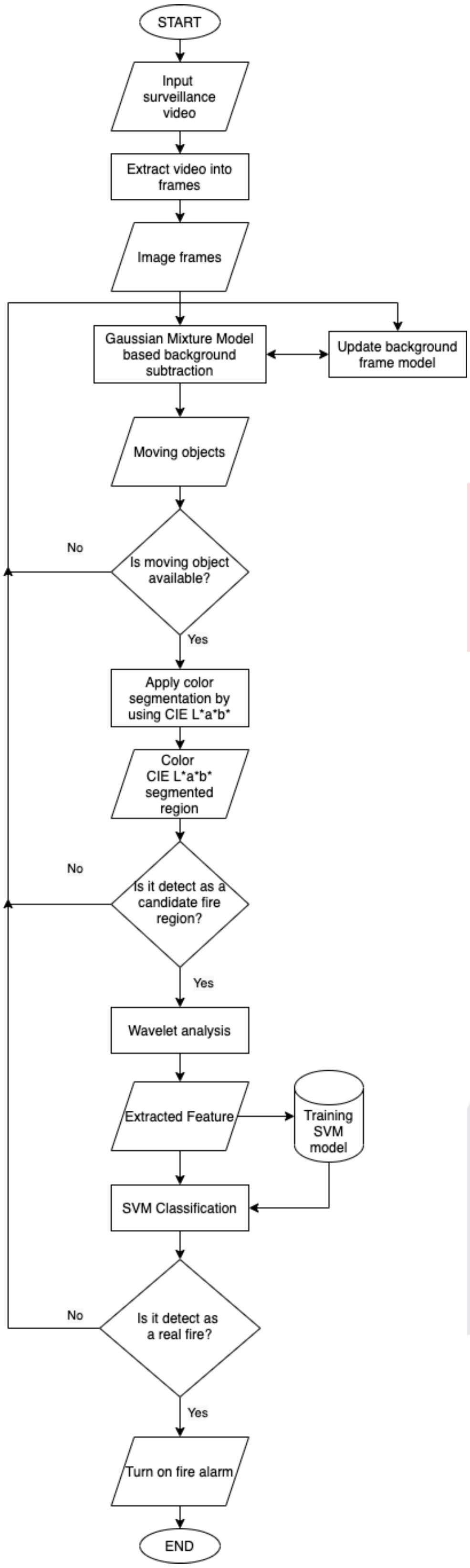


Figure 4. Sample datasets of different scenarios from the collected data



B. Training Process

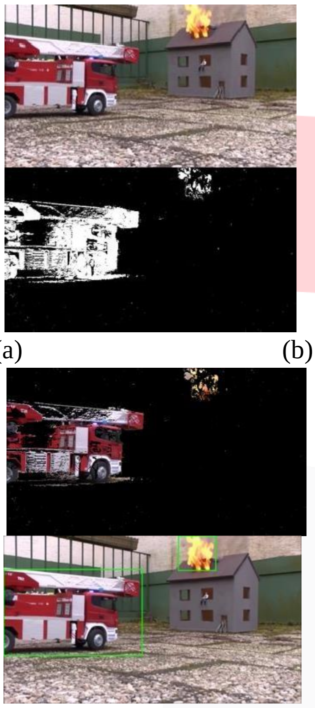
Figure 5. Flowchart method of training process

1. Gaussian Mixture Model based background subtraction:

Since the dataset in this study is in video format, thus it is needed to differentiate between moving objects and non-moving objects.

$$\boldsymbol{R_{motion}(i,\ j,\ n) = gmm(I(i,\ j,\ n))} \qquad \text{Equation 1 [6]}$$

Where *I* is the image frame, *R* is the foreground moving region from the GMM algorithm, and *gmm()* is the operation function of the Gaussian Mixture Model algorithm.



(a) (b) (c) (d)

Figure 6. An original frame containing moving object (a), the binary images produced from GMM background subtraction, the frame in RGB from masking original frame and binary images (c), and the original frame with bounding box on moving objects (d)

2. Select candidate fire regions:

   In this step, after get list of moving objects from GMM based background subtraction, will use CIE L∗a∗b∗ color space to select the candidate fire regions. First, RGB to CIE L*a*b* conversion is needed. In this study, the conversion is performed by using Equation 2:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$L^* = \begin{cases} 116 * \left(\frac{Y}{Y_n}\right) - 16, & \left(\frac{Y}{Y_n}\right) > 0.008856 \\ 903.3 * \left(\frac{Y}{Y_n}\right), & Otherwise \end{cases}$$

$$a^* = 500 * \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \qquad \text{Equation 2 [3]}$$

$$b^* = 200 * \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right)$$

$$f(t) = \begin{cases} t^{1/3}, & t > 0.008856 \\ 7.787 * t + \frac{16}{116}, & Otherwise \end{cases}$$

where X_n, Y_n, and Z_n are the color white values. The range of RGB, L*, a*, and b* are [0, 255], [0, 100], [–110, 110], and [–110, 110] respectively. Then as (L^*m, a^*m, b^*m) is average channel, in this study the average channel calculation use Equation 3:

$$L^*_m = \frac{1}{N} \sum_x \sum_y L^*(x, y)$$

$$a^*_m = \frac{1}{N}\sum_x \sum_y a^*(x,y) \qquad \text{Equation 3 [3]}$$
$$b^*_m = \frac{1}{N}\sum_x \sum_y b^*(x,y)$$

N is the total number of image pixel. Then, to produce candidate regions, four rules are obtained. The rules are as in Equation 4, 5, 6, and 7. Those rules will calculate the fire region with brightest area and the brightest itself is close to red. The results from each rule will be binary images.

$$R1(x,y) = \begin{cases} 1, & L^*(x,y) \geq L*m \\ 0, & Otherwise \end{cases} \qquad \text{Equation 4 [3]}$$

$$R2(x,y) = \begin{cases} 1, & a^*(x,y) \geq a*m \\ 0, & Otherwise \end{cases} \qquad \text{Equation 5 [3]}$$

$$R3(x,y) = \begin{cases} 1, & b^*(x,y) \geq b*m \\ 0, & Otherwise \end{cases} \qquad \text{Equation 6 [3]}$$

$$R4(x,y) = \begin{cases} 1, & b^*(x,y) \geq a^*(x,y) \\ 0, & Otherwise \end{cases} \qquad \text{Equation 7 [3]}$$

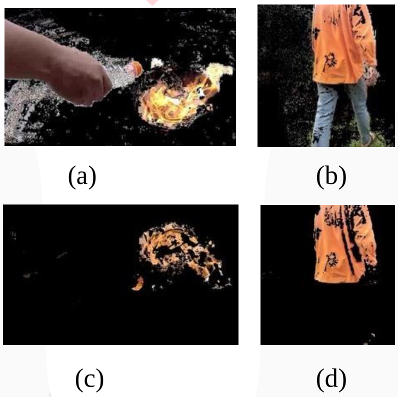


Figure 7. Applying the rules on moving object from GMM background subtraction. (a),(b) are objects before rules applied and (c),(d) are after rules applied respectively.

3. Extract the features:
Extract the features from candidate fire regions. In this study, wavelet analysis will be used to extract the features of fire. Wavelet analysis is a good image-processing method that can be used to distinguish between real fire regions and fire-like regions. To implement it, it shows in Equation 8, will use 2D wavelet filter to analyze the red color and will use spatial wavelet to compute the energy of each pixel values.

$$E(x,y) = (HL\,(x,y)^2 + LH(x,y)^2 + HH(x,y)^2) \qquad \text{[3]Equation 8}$$

From Equation 8, *E(x, y)* is the spatial wavelet computation for each pixel, while for wavelet sub-images, *HL* is low-high, *LH* is high-low, and *HH* is high-high. The spatial wavelet computation for each block in Equation 9. It adds the specific energy from every pixel in the block.

$$\boldsymbol{E_{block}} = \frac{\boldsymbol{1}}{\boldsymbol{N_{block}}} \sum_{x,y} \boldsymbol{E(x, y)} \quad \text{Equation 9 [3]}$$

N_{block} is the total number of pixels in the block and E_{block} is stored for the next classification step.

4. Classifying the candidate regions features by using Support Vector Machine (SVM).

In this step, SVM is used to classify the regions of interest to either fire or non-fire. The applied classification function is in the Equation 3:

$$\boldsymbol{f(x) = sign(\textstyle\sum_{i=0}^{l-1} w_i . k(x, x_i) + b)} \quad \text{Equation 10 [3]}$$

sign() is to get the class either +1 class or –1 class. +1 class belongs to fire and -1 class belongs to non-fire. While *k(), wi, xi,* and *i* are the kernel function, output weights of the kernel, the support vectors, and support vectors number respectively. SVM has several kernel, some of them are:

1. Linear kernel

$$k(\boldsymbol{x_1}, \boldsymbol{x_2}) = \boldsymbol{x_1} . \ \boldsymbol{x_2} \quad \text{Equation 11 [14]}$$

2. RBF kernel

$$k(\boldsymbol{x_1}, \boldsymbol{x_2}) = exp\left(-\frac{\|\boldsymbol{x_1} - \boldsymbol{x_2}\|^2}{2\sigma^2}\right) \quad \text{Equation 12 [14]}$$

3. Polynomial kernel

$$k(\boldsymbol{x_1}, \boldsymbol{x_2}) = (\boldsymbol{x_1} . \boldsymbol{x_2})^d \quad \text{Equation 13 [14]}$$

(x, y) is the input feature in vectors type, while σ is the parameter for controlling the width of the effective basis function and *d* is indicate as polynomial degree. Then, to build SVM model, dataset consisting of actual fire and non-fire is used.

C. Testing Process

In this stage, the model from training process will be tested by unseen data surveillance video. In Figure 5 and Figure 8 shows that the different between training and testing process is in wavelet analysis. In the training process, extracted feature is an input, while in testing process is not.

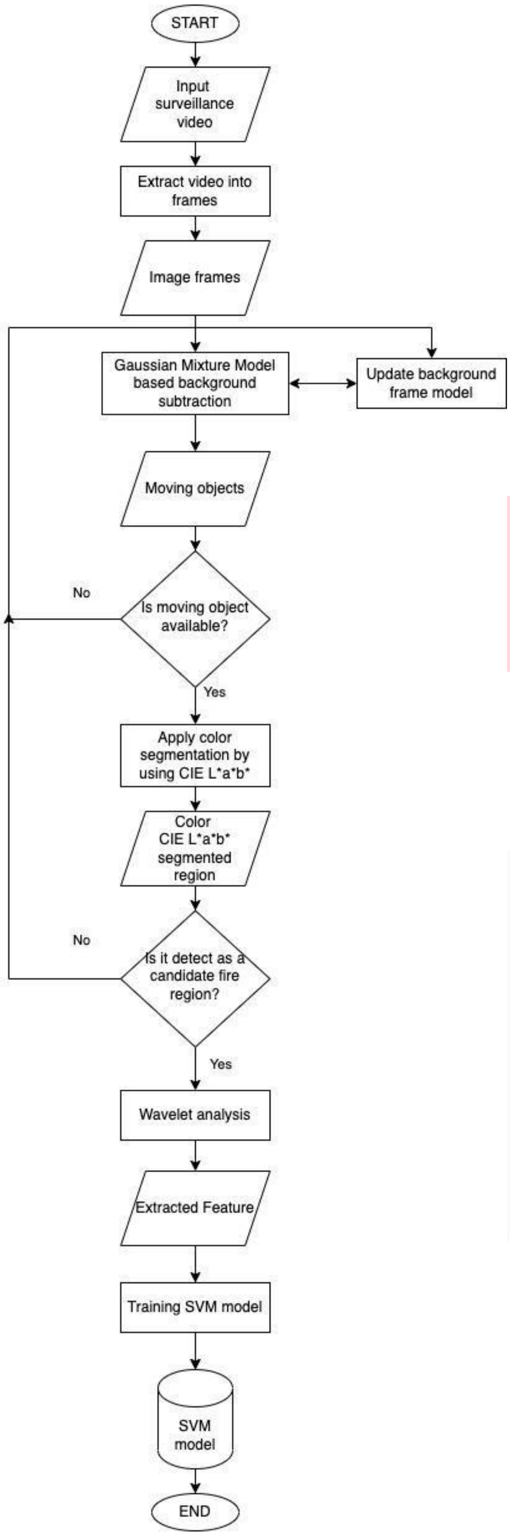


Figure 8. Flowchart method of testing process

Then, the result is evaluated by using hypothetical test, which are:

1. Accuracy
   In Equation 14, it calculates the accuracy of the built model with the test dataset.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad \text{Equation 14 [15]}$$

Where the detail of *TP*, *TN*, *FP*, and *FN* is in Figure 9.

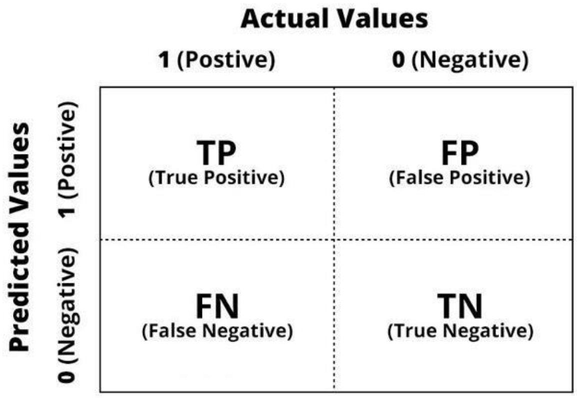


Figure 9. Confusion matrix [15]

IV. EVALUATION

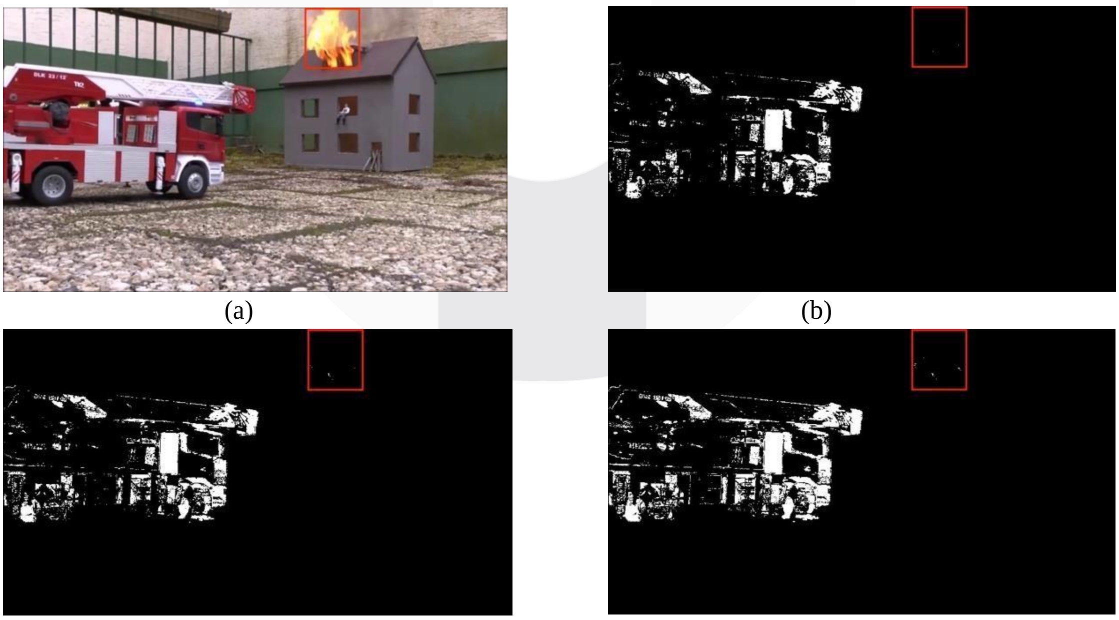
A. Test Result

The test result have several scenarios. In each scenario, the author changed the parameters in terms of getting optimal solution for this study.

Test Result Analysis

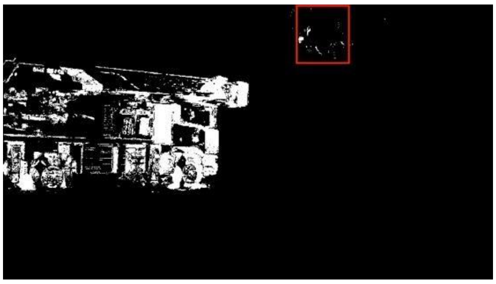
1. Scenario-1: Test scenario on threshold of GMM background subtraction

   In scenario-1, it tested several threshold value on GMM background subtraction. The threshold is determine how well a pixel is defined in background model. The high value of the threshold will effect the way of background model detect new pixels that extremely different from previous background model. The tested threshold value are 0, 100, 200, 300, and 400. It shows in Figure 10, the threshold with value 0 able to obtain the fire as moving object. Thus, threshold 0 is used in the system.

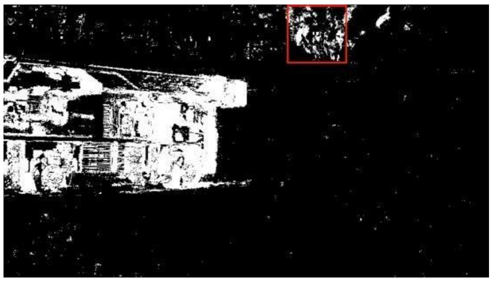


(a) (b)

(c) (d)



(e)



(f)

Figure 10. (a), (b), (c), (d), (e), and (f) are original frame, sample result of 400, 300, 200, 100, and 0 respectively. Red rectange is indicating the location of the fire.

2. Scenario-2: Test scenario on combination of mother wavelet and SVM kernel

.

In scenario-2, to find the best combination of several mother wavelet and several SVM kernel. The tested wavelet are biorthogonal, daubechies, and symlets. While for SVM kernel, the tested kernel are linear, polynomial, and gaussian

Table 2. Test result Scenario 2

Wavelet	SVM Kernel	Average Accuracy
Biorthogonal	Linear	70.313%
	Polynomial	70.473%
	RBF	55.687%
Daubechies	Linear	71.549%
	Polynomial	73.105%
	RBF	59.418%
Symlets	Linear	71.801%
	Polynomial	75.463%
	RBF	63.836%

The best combination is Symlets as mother wavelet and Polynomial as SVM kernel with average accuracy is 75.463%

B. Overall System Analysis

It use the best parameter from several test scenario. In Table 3, it shows the details of evaluation for each testing dataset. The measurement of the system evaluation is based on the objects

Table 3. Evaluation for each testing dataset

Video	TP	FP	TN	FN	Accuracy
Video_01	482	50	636	137	85.670%

Video_02	334	6	276	84	87.142%
Video_03	7	23	47	10	62.068%
Video_04	23	1	7	13	68.181%
Video_05	82	48	59	11	70.5%
Video_06	73	0	64	22	86.163%
Video_07	252	20	3	70	73.913%
Video_08	154	256	55	21	43.004%
Video_09	97	117	209	34	66.958%
Video_10	130	28	130	57	75.362%
Video_11	94	58	83	6	73.443%
Average Accuracy	1728	607	1569	465	75.463%

In the Table 3, the result shows the lowest accuracy is Video_08 with 43.004%. Video_08 took in outdoor with many moving objects with fire color, it cause the high number in false positive. While the highest accuracy is Video_02 with 87.142%. Video_02 environment is ideal with not many moving objects around and it took during the day. The decrease of accuracy is occurs because the non-fire with fire-color is detected as fire as in Figure 11.

(a) (b)

Figure 11. Sample of incorrect in detecting fire object (a) and sample of correct detecting fire object (b)

V. CONCLUSION

This study shows it is able to implemented wavelet and SVM for detecting fire in surveillance video. From an experimental point of view, it can be concluded that system have average of accuracy is 75.463%. The used parameters in the system for GMM is threshold 0, which shows better performance in getting moving object compare with threshold 100, 200, 300, and 400. While for wavelet, the selected mother wavelet is symlets, with Polynomial as SVM kernel. The selected combination is compare among combination of symlets, biorthogonal, and daubechies as mother wavelet and linear, polynomial, and RBF. The decrease of accuracy in the system is occurs because the non-fire with fire-color is detected as fire.

Future work concerns on use more datasets from different video resolution and fire color and shape to train the model in terms of getting better result, specifically in reduce the false positive, and can be used in many environment and deeper analysis in evaluating detecting fire based on moving objects.

BIBLIOGRAPHY

[1] D. S. Riyadi and S. Aisyah, “Vision Based Flame Detection System For Surveillance Camera,” Proc. 2018 Int. Conf. Appl. Eng. ICAE 2018, pp. 1–6, 2018, doi: 10.1109/INCAE.2018.8579405.

[2] K. Poobalan and S.-C. Liew, “Fire Detection Algorithm Using Image,” no. December, pp. 12–13, 2015, [Online]. Available: http://worldconferences.net.

[3] M. Adam, I. Mahmoud, and H. Ren, “Forest Fire Detection and Identification Using Image Processing and SVM,” vol. 15, no. 1, pp. 159–168, 2019.

[4] L. Shi, F. Long, Y. Zhan, and C. Lin, “Video-based fire detection with spatio-temporal SURF and color features,” Proc. World Congr. Intell. Control Autom., vol. 2016-Septe, pp. 258–262, 2016, doi: 10.1109/WCICA.2016.7578269.

[5] Y. K. Singh and D. Deb, “Detection of fire regions from a video image frames in YCbCr color model,” Int. J. Recent Technol. Eng., vol. 8, no. 3, pp. 6082–6086, 2019, doi: 10.35940/ijrte.C5637.098319.

[6] X. Han, L. Xiao, J. S. Jin, M. W. Wei, and J. Lei, “Video fire detection based on Gaussian Mixture Model and multi-color features,” Signal, Image Video Process., 2017, doi: 10.1007/s11760-017-1102-y.

[7] D. K. Appana, R. Islam, S. A. Khan, and J. M. Kim, “A video-based smoke detection using smoke flow pattern and spatial-temporal energy analyses for alarm systems,” Inf. Sci. (Ny)., vol. 418–419, pp. 91–101, 2017, doi: 10.1016/j.ins.2017.08.001.

[8] S. Wilson, S. P. Varghese, G. A. Nikhil, I. Manolekshmi, and P. G. Raji, “A Comprehensive Study on Fire Detection,” Proc. IEEE Conf. Emerg. Devices Smart Syst. ICEDSS 2018, no. March, pp. 242–246, 2018, doi: 10.1109/ICEDSS.2018.8544329.

[9] L. Debnath and F. A. Shah, Lecture Notes on Wavelet Transforms, 1st ed. Birkhäuser, Cham, 2017.

[10] C. L. Wang, R. H. Hwang, T. S. Chen, and H. Y. Lee, “Detecting and restoring system of tampered image based on discrete wavelet transformation and block truncation coding,” Proc. - Int. Conf. Adv. Inf. Netw. Appl. AINA, vol. 2, no. 1, pp. 79–82, 2005, doi: 10.1109/AINA.2005.160.

[11] M. Kociolek, A. Materka, M. Strzelecki, and P. Szczypinski, “Discrete Wavelet Transform – Derived Features for Digital Image Texture Analysis,” Int. Conf. Signals Electron. Syst., no. September, pp. 163–168, 2001.

[12] A. Hyvärinen, J. Hurri, and P. O. Hoyer, Natural Image Statistics-A Probabilistic Approach to Early Computational Vision., vol. 39. 2009.

[13] V. Kecman, “Support Vector Machines – An Introduction,” no. May, pp. 1–47, 2005, doi: 10.1007/10984697_1.

[14] V. Jakkula, “Tutorial on Support Vector Machine (SVM),” Sch. EECS, Washingt. State Univ., pp. 1–13, 2011, [Online]. Available: http://www.ccs.neu.edu/course/cs5100f11/resources/jakkula.pdf.

[15] K. M. Ting, “Precision and Recall,” in Encyclopedia of Machine Learning, C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2010, p. 781.