The over-height vehicle detection using the computer vision method

T M Iqbal1*, F Arnia1, R Muharar1
1Department of Electrical and Computer Engineering, Univeritas Syiah Kuala, Aceh, Indonesia.
*E-mail: twk.m@mhs.unsyiah.ac.id

Abstract. Over-height vehicles are not only threats for the over-height vehicles themselves, but they also bring about a hazard to the tunnels they pass in, and to other drivers on the road way when a clash happens with a low clearance structure. Therefore, it is needed an over-height detection system which is affordable yet also reliable. Nowadays, there are the over-height detection system using laser and infrared. However, they are quite pricy. In this research a computer vision-based system is proposed to detect the height of vehicles and provide a warning for the over-height vehicles. The height determination was conducted using Gaussian Mixture Model (GMM) and blob detection. The GMM is used to detect the vehicle and the blob detection is applied to produce the vehicle coordinates which determines the direction of movement and the over-height vehicle. The over-height vehicles are detected by using 5 (five) variance total frames within 3 (three) conditions. The accuracy of the test proves that the method is reliable in determining the height of vehicles, achieving 100% accuracy of the detected vehicles. The significance of this work is the design of a vision-based method which can determine the height of the vehicles and is a low-cost alternative to the current costly laser and infrared detection systems.

1. Introduction
Statistics shows that the number of accidents occurred in Indonesia are relatively high. In 2019 there were 107,500 traffic accidents with the total victims of 23,530 [1]. The traffic accidents can be caused by several factors such as human errors, vehicle eligibility, and environmental factors. One of the examples of human errors is physical condition such as fatigue, while the vehicle eligibility factor is like the unoptimized condition of machine, brake, and wheels. Some of the environmental factors are climbing and downhill roads, as well tunnels and underpasses.

The accidents at the tunnel or underpass can be caused by the over-height vehicles. It causes vehicles striking the underpass. This condition brings about material losses, victims, underpass damage, and traffic lag/jam. To prevent those strikes, it is needed a system which can detect the over-height vehicle. The system can direct the vehicles passing the tunnels. Should the vehicles height is over the tunnels, the system will warn and alert drivers to stop or proceed to the exit way.

There are several pieces of previous research related to the over-height vehicles detection conducted by using various methods such as laser [2] [3] and optics [4] [5] [6] [7]. The study using laser method shows that the rate of error accuracy is 0.66 inch [6]. However, the laser or infrared methods are costly [5] [10]. Another study reveals that the infrared dual beam (Z-Pattern) method can detect the false over-height vehicles and alert the drivers [5]. This system is also relatively expensive as it can spend USD $150,000 – 200,000 (around Rp 2,9 billion) per installation and per direction. It includes equipment costs, excluding the maintenance fees. Meanwhile, the Sydney Harbor Tunnel using infrared method...
and pseudo-holographic system displaying “STOP” word at the wall can spend AUD $150.000 – 200.000 (around Rp 2 billion) and maintenance fee AUS $25.000 – 35.000 (around Rp 352 million). The optic method (prototype) installed at the Fairlop Station Bridge, East London, UK costs £2.810 (around Rp 50,9 million); it counts only for the vehicle detection equipment (camera, computer, hard disk, router, mounting, cable, and outdoor box), excluding the warning system [8].

Among those methods, optics is one of the systems which needs the lowest costs in the over-height vehicle detection [4] [9] [10]. The previous research using optics to detect the over-height vehicle reveals the varied rate of accuracy using various methods. The blob detection method and Kanade Lucas Tomase (KLT) shows the accuracy of over height vehicle up to 92,2% [10]. Meanwhile, the blob detection method and Histogram of Oriented Gradients (HOG) marks the accuracy of vehicle detection at 95,83% and the over height vehicle detection at 95,77% [7]. The KLT detection method generates 68,7% for vehicle detection accuracy and 83,3% for over-height vehicle detection accuracy [5]. Lastly, KLT and SURF detection method marks 83,3% accuracy for vehicle detection and 96,6% accuracy for over-height vehicle detection [6].

Those researches show that the optics method can generate a more varied result of accuracy while the laser method needs costly expenses for equipment and maintenance. Therefore, it is required another optics method which can detect the over-height vehicle more accurately with affordable cost. This study proposes the use of Gaussian Mixture Models (GMM) feature. The GMM feature is applied to split between the foreground and background [11] [12]. The detected foreground object is assumed as a vehicle. The vehicle is then measured its height.

The previous research reveals that GMM method can generate high rate accuracies in detecting vehicles. The result of Nurhadiyatna et al. study shows that the foreground detection using GMM features combined with the Hole Filling (HF) algorithm marks the accuracy rate up to 97,9% [13]. Meanwhile, in Indrabayu et al., study, the accuracy rate generated by the Gaussian Mixture Models (GMM) in detecting the vehicle achieves 97,2% [14]. The results explain that the Gaussian Mixture Models (GMM) can yield a high rate of accuracy in vehicle detections. It, then, becomes the researcher’s reason to choose the Gaussian Mixture Models (GMM) to be applied in detecting the over-height vehicles. It is expected that the system can detect the vehicles more accurately so that can warn and give alert precisely. Also, it is hoped that the system can cut the traffic accident and bridge strike numbers.

2. Literature Review
Gaussian Mixture Models (GMM) is the components of the Gaussian function forming a kind of density model from various threshold. This feature segregates between the foreground and background [11] [12] in the video. The segregation is produced by comparing an image from a frame with the next frame. If a density group moves, it will be classified as the foreground and the rest as background represented in a binary matrix. The detected foreground object is assumed as a vehicle. Later, this system will measure the height of the vehicle. Should the height of the vehicle is beyond the threshold, the system will trigger an alert in the form of voices, lights, and warning boards.

GMM is generated from the pixel color value based on the frame changes. The models is classified into two groups which are background and foreground [11]. Equation 1 is the equation for the Gaussian Mixture Model (GMM) [15]:

\[ P(X_t) = \sum_{i=1}^{K} \omega_{i,t} \ast \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \]  \hspace{1cm} (1)

With \( P(X_t) \) is the pixel probability of a frame. \( K \) is the sum of Gaussian Distribution, \( \omega_{i,t} \) is the weight estimation of Gaussian to \( i \) in \( t \) frame. Then, \( \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \) is the probability of density function of Gaussian Mixture Model to \( i \) in \( t \) frame, represented in Equation 2.

\[ \eta(X_t, \mu_{i,t}, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(X_t-\mu_{i})\Sigma^{-1}(X_t-\mu_{i})} \]  \hspace{1cm} (2)
where $\mu_{i,t}$ is the mean of Gaussian to $i$ in $t$ frame, and $\Sigma_{i,t}$ is the covariance matrix of Gaussian to $i$ in $t$ frame. The equation for covariance matrix is as Equation 3.

$$
\Sigma_{i,t} = \sigma_k^2 I
$$

(3)

Where $\sigma_k^2$ is the variance of Gaussian distribution, $I$ is the identity matrix.

GMM feature is a statistic model which is effective to segregate between the foreground and the background with minimal movement. It is carried out by comparing an image of a frame with the next frame. Should there a density group moves, it will be classified as foreground and the rest as background represented in the form of binary matrix.

The steps in classifying the background is presented in Picture 1. First is matching every distribution until it is found the appropriate distribution. The matching is performed by determining ranges of a pixel with deviation standard value 2.5. It can effectively classify a distribution. Should a pixel does not match with all existing distributions, a new distribution is made and replace the low weight value, mean, and high variance value. The renewal of the weight value to the frame changes (time) presented in Equation 4.

$$
\omega_{k,t} = (1 - \alpha)\omega_{k,t-1} + \alpha(M_{k,t})
$$

(4)

Where $\omega_{k,t}$ is the Gaussian distribution weight and $\alpha$ is learning rate of a model with value 0 if the model is matching and with 0 value if the model is not matching. After the weight value is updated, the weight value is normalized of all distribution to get value 1. The update of mean and variance score to frame (time) changes carried out with unmatched distribution presented in Equation 5 to 7.

$$
\mu_{i,t} = (1 - \rho)\mu_{i,t-1} + \rho X_t
$$

(5)

The variance value update to the frame (time) change is shown in the following equation:

$$
\sigma_k^2 = (1 - \rho)\sigma_k^2 + \rho(X_t - \mu_t)(X_t - \mu_t)^T
$$

(6)

Where:

$$
\rho = \alpha \eta_k(X_t|\mu_k, \sigma_k)
$$

(7)

And $\sigma_k^2$ is the Gaussian variance distribution value.

The next step is choosing the models of background distribution. The background determination based on the similarity value of the models in the distribution. Should the model has a large similarity value, the model is classified as background. In this step, the level of model similarity is sorted based on the weight value ($\omega$) or deviation standard value ($\sigma$). The bigger the similarity level of Gaussian distribution, the bigger the weight value ($\omega$) of Gaussian distribution and the smallest the deviation standard value ($\sigma$).

On the other hand, if the deviation standard value is smaller than the weight value, it is assumed as the moving object, or called as foreground. B distribution is determined as background model as presented in Equation 8.

$$
B = \arg\min_b \left( \sum_{k=1}^b \omega_k > T \right)
$$

(8)

Where $T$ is the threshold of minimum data assumed as background. Then, foreground is determined from the smallest weight value. The feature is used widely, such as for modelling a dynamic environment in real-video and in the online-based fashion [16]. In transportation sector, this feature is used to detect vehicles [11].

3. Research methodology

This research was carried out by recording dataset video in two different locations, Simpang Surabaya, Banda Aceh and Pango bridge, Aceh Besar. The Simpang Surabaya location has traffic characteristics with a concrete wall background. Meanwhile, at the Pango bridge has traffic characteristics with a background of small plants, grasses, tress, and two-ways street. The vehicle detection was carried out
by determining the length of time of object and movement direction in frame. This method aims to reduce errors in vehicle detections. In this study, there are five variance of frame i.e 3, 5, 7, 9, and 11. Meanwhile the over-height vehicle detection was performed in three conditions. First is without ROI and without vehicle movement direction detection. Second is with ROI and without vehicle movement direction detection. Lastly is with ROI and vehicle movement direction detection. The flowchart of the over-height vehicle detection is presented in Figure 1.

The flowchart begins with the video inputs. After that, the researcher determines the Region of Interest (ROI), as shown in Figure 2. The ROI is used to remove unnecessary objects. The ROI frame is used to detect foreground by using the GMM method. The next step is to remove noises. The noise removal is done by using the morphology filter. The filter cleans the noise by restructuring the morphological flat element of a pixel. The restructuring element is performed by using dilation and erosion operations. The form used in this filter is in square with 3 pixel width. The center element acts as the origin pixel determining the values of the surrounding pixels.

Next step is detecting the blob area. The output of the blob detection is the coordinate value of the object which is used as the basis in determining the movement direction and height. Then, the detection of the movement direction of the vehicle moving to the forward is determined by the coordinate y_now smaller from y_preceding. The reason is the coordinate O(0,0) is located at the left side of the top of the frame. The later step is counting the number of frames detecting the foreground object and the number of movement direction to ensure that the object is the vehicle. The counting of the consistency foreground_detected and direction_movement is bigger or equal than total_frames. The total_frames is the length time of the object in the frames and the variance of total_frames is 3, 5, 7, 9, and 11. The next phase is deciding the vehicle height by determining the coordinate object which is smaller than coordinate value of the threshold.

The threshold coordinate frame is determined from the result of the position point of camera calibration. The height of the camera position is at 2.5 meters and the direction is 90° and at 5 meters from the road side. On the other side of the road, the researcher places a 2.5 meters stick for calibrating the coordinate of real world to the pixel of frame. It is assumed that the threshold height is 2.5 meters. If x_foreground is smaller from x_threshold then the object is determined as the over-height object and the system will activate the trigger of warning system. Otherwise, should the height of the vehicle does not exceed the threshold, the system will ignore the vehicles. The camera calibration point and pixel coordinate is presented in Figure 3.

In Figure 3 the coordinate pixel x_threshold is 200 which is determined from the position of camera calibration point (P). The line OB-AC is the threshold line. The object A with the value coordinate x_object_A is 150 which is smaller than x_threshold. Then the object A is called as the over-height object. Meanwhile, the object B with the value coordinate x_object_B is 230 which is bigger than x_threshold. The object B is then determined as a not over-height object. The position of camera calibration point is preferred to be placed at the center of the frame to decrease the effect of lens radial distortion [17].
Start

Input video

Determine ROI

Detect the foreground using GMM

Remove the noise

Blob detection determine the foreground coordinate output

Detect the direction movement $y_{\text{now}} < y_{\text{preceeding}}$

$\text{Foreground\_detected} \geq \text{total\_frames}$ and $\text{direction\_movement} \geq \text{total\_frames}$

Yes

Threshold of camera calibration

Detect the over-height vehicle $x_{\text{foreground}} < x_{\text{threshold}}$

Yes

Activate the trigger alarm warning

Finish

No

Figure 1. The system simulation flowchart of the over-height vehicle detection
Figure 2. The Region of Interest (ROI) area

The display of simulation system shows a passing vehicle height is over the threshold line. Therefore, the system activates the warning of over-height vehicle. The system also displays the counter of five variance total frames detection, i.e. in 3, 5, 7, 9 and 11 frames. It also presents the conditions of over-height vehicle detection, i.e. with the ROI, and direction detection, or not. The output display of the simulation system is shown on Figure 4.

Figure 3. Camera calibration point and pixel coordinate
4. Result and discussion

This study used Gaussian Mixture Model and blob detection feature to detect the over height vehicles. The study was conducted at two locations, Simpang Surabaya and Pango bridge. The traffic documentation at the Simpang Surabaya was at the roadway from Simpang Surabaya to Lambaro with a concrete wall background. Meanwhile, the traffic documentation at the Pango bridge is at the roadway from Lambaro to Simpang Surabaya with a background of small plants, grasses, trees, and two ways street. The video was recorded for 7 minutes 54 second with 640x480 pixel resolution and frame rate 30 frames per second (fps) in cloudy conditions and light rains. The video testing data in both locations is represented in the Table 1 and Table 2

| Condition                        | Frame | Actual |
|----------------------------------|-------|--------|
| Without ROI and without direction detection | 12 9 8 8 6 |        |
| With ROI and without direction detection | 5 4 4 4 4 | 2      |
| With ROI and direction detection | 2 2 2 2 2 |        |

**Table 2. Video testing data at Pango bridge, Aceh Besar**

| Condition                        | Frame | Actual one way | Actual two ways |
|----------------------------------|-------|----------------|-----------------|
| Without ROI and without direction detection | 34 23 19 19 16 |            |                |
| With ROI and without direction detection | 15 10 8 8 7  | 2             | 6              |
| With ROI and direction detection | 4 2 2 1 1 |            |                |

The result of the data shown that the there is an effect of frame number on the accuracy rate. The object consistency detected in a particular number of frames sequentially revealed the rate of accuracy [6]. The over-height vehicle detection using 3 frames shown the lowest accuracy rate. It is because the 3 frame cannot reduce the detection error. The over-height vehicle detection using 5 frames shown a rise in accuracy rate [6]. The climb is caused by the increase number of frames which can reduce the
error in the over-height vehicle detection. The accuracy rise with the increase of frame number is presented in Figure 5. The number of frame which produced the smallest number of errors in the over-height vehicle detection is frames 11 and 7. However, the percentage of the accuracy average revealed that the optimal accuracy is at frames 5 and 7. Therefore, referring to the number of errors of the over-height vehicle detection and the accuracy rate, the optimal frame number is 7 frames. Figure 6 shows the traffic video accuracy enhancement graph at Simpang Surabaya while Figure 7 Presents the traffic video accuracy enhancement graph at Pango bridge.

The result revealed that the Gaussian Mixture Model (GMM) and blob detection can detect the over-height vehicle with the accuracy up to 100%. The background differences do not affect the system performance on conditions with ROI dan direction detection. It, thus, can be concluded that Gaussian Mixture Model (GMM) and blob detection can accurately detect the over-height vehicles. This study is further expected can be a reference for stakeholders in transportation sectors as an alternative to reduce the bridge strikes.

**Figure 5.** The effect of frame number in decreasing the error of the over height vehicle detection

**Figure 6.** The traffic video accuracy enhancement graph at Simpang Surabaya
5. Conclusion
The result of the study revealed that the GMM and blob detection can detect the over-height vehicles accurately up to 100%. The study was conducted within three conditions, without ROI and without direction detection, with ROI and without direction detection, and with ROI and direction detection. The accuracy rate significantly rises within condition with ROI and direction detection. This condition was applied at two locations with different backgrounds. One is with a concrete wall background and the other is with a background of grasses, small plants, trees, two ways street. The accuracy rate in both places achieves 100%.

The method to reduce the errors in detecting the over-height vehicles can be done by adding the number of frame. The best number of frame in detecting the over-height vehicles is 5 and 7 frame. Referring to the accuracy rate, frame 7 is marked as the most optimal number in detecting the over-height vehicles as it can obtain accuracy 100% with the lowest errors.

6. References

[1] Akbar N A, Amanda C G 2019 Available: https://nasional.republika.co.id/berita/q37i3n423/jumlah-kecelakaan-lalu-lintas-naik-3-persen-pada-2019. (in Bahasa Indonesia).

[2] Singhal A, Kanga C, Agrawal A K 2018 IET Intell. Transp. Syst. 12 689-695.

[3] Garcia F, Cerri P, Broggi A, Armingol J M, Escalera A D L 2009 Computer Aided Systems Theory - EUROCAST 2009, EUROCAST 2009, Las Palmas de Gran Canaria, Spain.

[4] Massoud M A 2013 World Congress on Engineering, London, U.K.

[5] Nguyen B, Brilakis I, Vela P A 2017 J. Comput. Civ. Eng. 31(5).

[6] Nguyen B, Brilakis I 2018 Adv. Eng. Inform. 38 67-80.

[7] Dai F, Park M W, Brilakis M S D I 2015 Automat. Constr. 50 29-39.

[8] Nguyen B 2017 Vision-Based Over-Height Vehicle Detection for Warning Drivers. Cambridge: University of Cambridge.

[9] Nguyen B, Brilakis I 2016 Transp. Res. Procedia. 14 3915 – 24.

[10] Khorramshahi V, Behrad A, Kanhere N K 2008 WASET, Open Science Index 15, Int. J. Transp. Veh.Eng. 2 681-685.

Figure 7. The traffic video accuracy enhancement graph at Pango bridge
[11] Alfina 2016 *Perancangan dan Pengujian Piranti Pemantauan Visual untuk Menentukan Volume Lalu Lintas*. Banda Aceh: Universitas Syiah Kuala. *(in Bahasa Indonesia).*

[12] Goyal K, Singhai J 2018 *Artif. Intell. Rev.* 50 241-259.

[13] Nurhadiyatna A, Jatmiko W, Hardjono B, Wibisono A, Sina I, Mursanto P 2013 *IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, UK.

[14] Indrabayu, Bakti R Y, Areni I S, Prayogi A A 2016 *Int. Conf. Comput. Intell. Cybern.*, Makassar, Indonesia.

[15] Stauffer C, Grimson W E L 1999 *CVPR*, Fort Colins, CO, USA.

[16] Najar F, Boururouis S, Al-Azawi R, Al-Badi A 2020 *Mixture Models and Applications*. Switzerland: Springer.

[17] Iqbal T M, Munadi K, Arnia F. 2011. *Seminar Nasional dan Expo Teknik Elektro 2011*, Banda Aceh. *(in Bahasa Indonesia).*