Movement Direction Estimation on Video using Optical Flow Analysis on Multiple Frames

Achmad Solichin

Faculty of Information Technology
Universitas Budi Luhur
Jakarta, Indonesia

Agus Harjoko[*], Agfianto Eko Putra[#]

Department of Computer Sciences and Electronics
Universitas Gadjah Mada
Jogjakarta, Indonesia

Abstract—**This study proposed a model for determining the movement direction of the object based on the optical flow features. To increase the speed of computational time, optical flow features derived into a Histograms of Oriented Optical Flow (HOOF). We extracted them locally on the grid with a certain size. Moreover, to determine the movement direction we also analyzed multiple frames at once. Based on the experiment results, showing that the value of accuracy, precision, and recall of the movement detection is good, amounting to 93% for accuracy, 73.07% for precision and 84.25% for recall. Furthermore, the results of testing using the best parameter shows the value of accuracy of 98.1%, 35.6% precision, 41.2% recall, and direction detection error rate (DDER) 25,28%. The results of this study are expected to provide benefits in video analysis studies such as the riots detection and abnormal movement in public places.**

Keywords—Video analysis; movement direction; optical flow; Histograms of Oriented Optical Flow (HOOF); multiple frames

I. INTRODUCTION

As a country of varying ethnicity, language and culture, Indonesia has a high potential for rioting. The National Disaster Management Authority (BNPB) noted that since 1998 to 2014 there have been 113 social conflicts [1]. The number of victims due to the conflict or the riots is also not small. According to BNPB's data, the number of dead is 6,022 people: injured 4,123 people, lost 476 people, and displaced 60,777 people. A large number of casualties in the event of social riots can be suppressed if it can be detected early. Therefore, continuous research is needed to detect riots using computers technology.

According to Mustofa [2], social riots involving many people is part of collective behavior. The collective behavior means a study that focuses on the patterns and sequences of events occurring in problematic situations [3]. There are three main characteristics of collective behavior that are spontaneity, volatility, and transitory. Every event of social riots occurs spontaneously, meaning that they are not predictable or engineered events. Individuals involved may initially be law-abiding individuals and do not like violence. But in problematic situations when individuals are involved in a collective behavior such as social riots, they suddenly commit acts of destruction and so on. The next characteristic of the riot is volatility, meaning that it is easy to change in a short time. In situations of riot, the people involved in it have changed behavior in such a short time as suddenly running or

screaming. While the third characteristic is transitory, that is a riot in general quickly subsided.

Based on the characteristics of the riots, the movement direction of an object is one of the most important in the process of riot detection. If it can be detected and known well, then the pattern of movement of the object will be known and analyzed. The main purpose of this research is to build a model that can determine the movement direction of objects in a video using computer vision technology. Computer vision is a set of methods for capturing, processing, analyzing and understanding images or videos. Basically, computer vision tries to imitate how humans capture, process, analyze and understand the environment.

In research on video analysis such as the analysis of social riot, the movement direction of objects within the video is very important. Movement direction can be used for the purpose of movement analysis, detection, and recognition. In a study by Martínez et al. [4], the movement direction is used as a descriptor to classify human activity within the video. Meanwhile, the histogram of the movement direction can also be used to analyze crowd behavior [5]. In the study, the histogram of movement direction is used as an indicator of movement speed. Speed changes are used to detect abnormal movements in the video.

Various methods and features used in many research related to the analysis of motion direction of objects on video. Optical flow is one feature that is quite widely used, such as in [6]. In addition to using the optical flow feature, other studies also add a mixture of Gaussian (MOG) feature to detect the direction and speed of moving objects in the video [6]. Although the results show that the direction and velocity of the object can be well-known, research has not been tested on video with the crowded situation and the diverse object. In addition, they used a low frame rate of density that is only 3 frames per second.

In a study by Benabbas et al. [7], the optical flow features are extracted globally from each video frame using the Kanade-Lucas-Tomasi method [8], [9]. The optical flow features are grouped into blocks of a certain size and each block is normalized using directional maps and von Mises distribution [10]. Movement directions are used to detect events such as walking, running, normal movement, evacuation, parting, assembling and spreading. Movement directions can also be used for recognizing human in a video

[11]. The research that conducted by Benabbas has been doing the division of frames into smaller blocks and analyzed the movement direction between the blocks. However, movement direction analysis has not been done for multiple frames at once.

Meanwhile, in a study by Colque et al. [12], the optical flow features and the direction vector become the descriptors of the object movement patterns in the video. The simple adjoining neighbor analysis method is applied to identify unusual movement patterns. The proposed model requires a training process that requires considerable computation time. In addition, the direction of optical flow is used only 4 directions, so it is not enough to represent the direction of the movement of an object.

Table I presents various studies related to the determination of movement direction of objects on video over the last 10 years. There are various features and methods used to determine the movement direction. There are 77% of researchers use optical flow features to get the direction of object movement because the optical flow has several advantages. One of its advantages is that analysis can be done directly on the pixels of successive frames, so it does not require the detection process first. However, some researchers also try to detect the object first before extracting the optical flow. Flexibility and ease in the extraction and analysis process make optical flow become popular especially to analyze movement in the video [13].

Another weakness of some research that aims to detect the movement direction of objects in the video is to require the object detection process first. With the object detection process will indeed increase the accuracy, but requires a longer detection time. In addition, object detection is not very effectively applied to video that contains large amounts of objects, such as in crowded videos. Therefore, we need a simpler model that no segmentation or object detection first.

In order to solve some problems in research to determine the movement direction of objects, in this study we proposed a new model of determining the movement direction of objects. In the previous studies, the movement direction of objects requires the process of segmentation and object detection in advance, so it requires considerable computing time, especially if applied to the crowded video. In this research, we developed a model for determining the movement direction of objects in the crowded video without involving segmentation and object detection process, using optical flow features that are derived to HOOF, dividing the frame into a certain size of grids, and the HOOF accumulation analysis for multiple frames. This research produces a model of determining the movement direction of objects on video that faster than previous methods.

TABLE I. STUDIES RELATED TO THE MOVEMENT DIRECTION OF AN OBJECT ON VIDEO

| #	Paper	Feature/ Method	Direction	Description
1	[14], [15]	Optical flow, von Mises distribution, and K-means	Circular directions	The extraction of motion pattern used to recognize object behavior on video.
2	[16]	Optical flow, HOOF	8 directions	The calculation of object direction is done on the silhouette of the object; the object is divided into 8 region
3	[17]	Optical flow, HOOF	12 directions	Histogram of direction used for segmentation of the crowded object on video.
4	[4]	Optical flow, HOOF	32 directions	Movement direction used as a descriptor to classify the movement pattern.
5	[18]	Optical flow and Mixture of Gaussian (MoG)	Based on optical flow directions	Video input only 3 fps have not been tested on higher video density.
6	[19]	Edge and KLT tracker	Based on optical flow directions	Objects are segmented based on the edge features and the KLT method, the movement direction is determined based on optical flow features.
7	[20]	Optical flow, Kalman filter	8 directions	Object directions were detected using the proposed method.
8	[12], [21], [22]	Optical flow, Simple Nearest Neighbor Search	4 directions	HOFM descriptor was implemented to detect abnormal events on a video.
9	[23]	Optical flow, Grid-based HOOF	12 directions	HOOF features were extracted locally on a grid and used to determine the movement direction of the object on every grid.
10	[24]	Optical flow, SLAM	-	The video shot motion analysis based on LK optical flow method to solve the problem of film video shot motion.
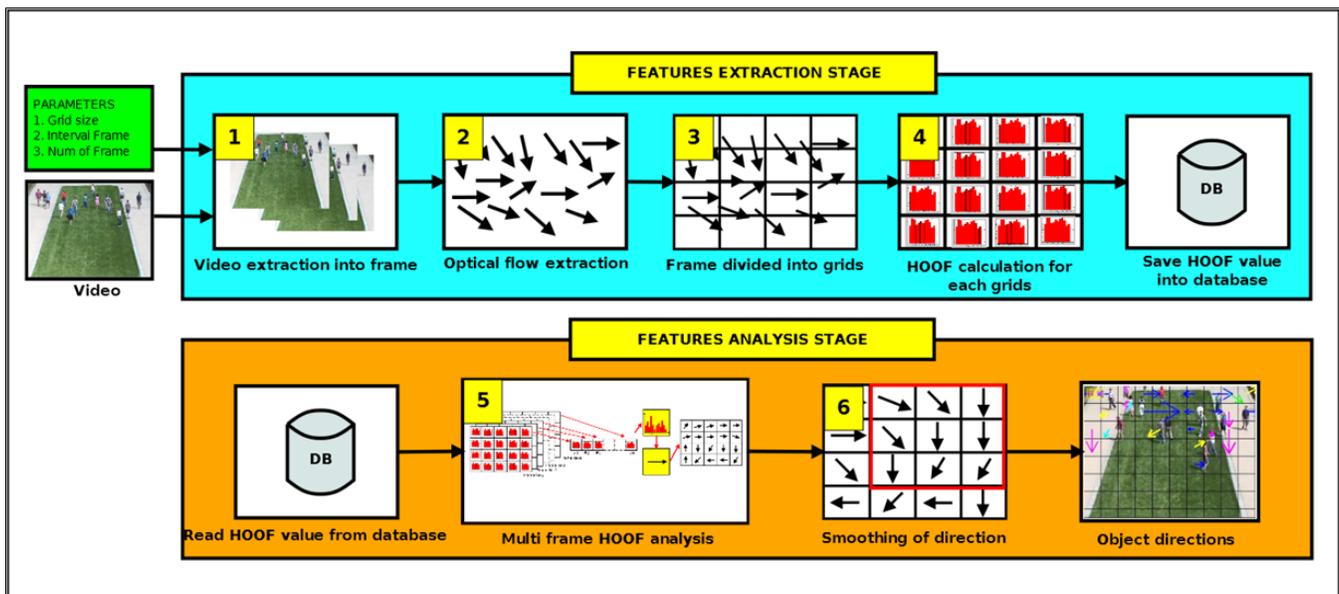
Fig. 1. The proposed model.

II. PROPOSED METHOD

This study aims to develop a model to determine the movement direction of objects in a crowd video using optical flow features that have computation time faster than previous methods while maintaining its accuracy value. To produce a model with faster computation time, then in this study:

- Derive the optical flow features into the Histogram of Oriented Optical Flow (HOOF), so it will reduce the dimensions of the features. To reduce it, we used simple statistical methods that are expected not to require a large computation time.

- Eliminate the process of segmentation and object detection because the process requires large computational time. The video constraint used in this study is a crowd video consisting of many human objects, so the process of segmentation and object detection is not effective to implement.

- Divide the frame into grids. It will keep a good accuracy value, and also accelerates the computing process because the accumulation of HOOF values is done only on the grid.

- Determination of the movement direction is not only done by analyzing HOOF on a single frame only but also done on multiple frames at once by accumulating the value of HOOF. It aims to maintain a good accuracy value.

We develop a video analysis model to determine the movement direction of objects consisting of two main processes: feature extraction and feature analysis. Fig. 1 presents the complete proposed model. This model is the main contribution of this research. Generally, the model has video input with 320x240 size, frame rate 25-30 fps and AVI (audio

video interleave) format. The model also accepts several parameters that affect the process at each stage. The main parameters of the proposed model are as follows:

- The size of the grid, denoted by N. The value of N determines the size of the grid that is used to divide the area of a frame in the HOOF feature extraction process.

- Interval frame, denoted with Iframe. Value Iframe affect HOOF features analysis process that has been stored in the database. The value of 1 means interval frame analysis performed for each frame and the interval frame of 2 means that the next frame will be ignored.

- The number of frames (P). It determines the number of frames that are included in the analysis process.

The feature extraction process begins by dividing the video into frames and extracting optical flow features for each frame. The process of feature extraction is a process to take the necessary features of an object. In this study, optical flow values were obtained using the classic Horn-Schunck [25] method by applying the convolution kernel [26], [27]. In this research, the optical flow features extraction process is performed for each frame of the input video. The optical flow value of each frame is stored in a variable for processing at a later stage.

The optical flow value reflects the movement of each point of the video frame. The greater value of the optical flow (r), the more significant the movement occurs. Under these conditions, then in this study applied a threshold value to eliminate the less significant movement of each video frame. Based on the results of program testing, in this study used the threshold value of 0.1. If the optical flow value (r) is less than the threshold, then the optical flow value is changed to 0.
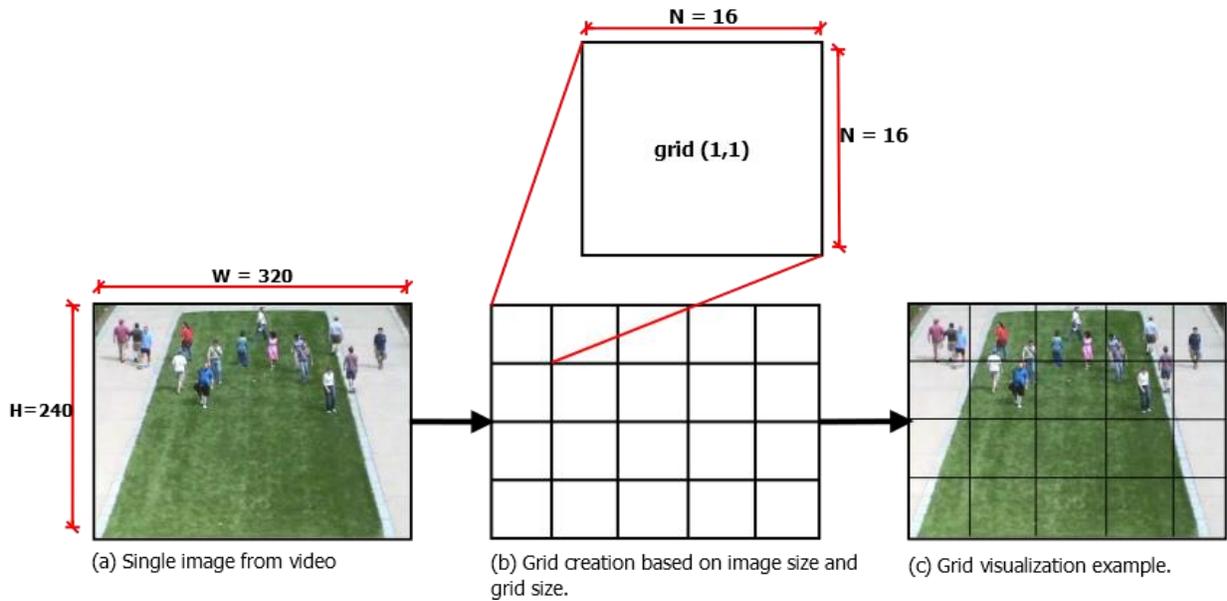
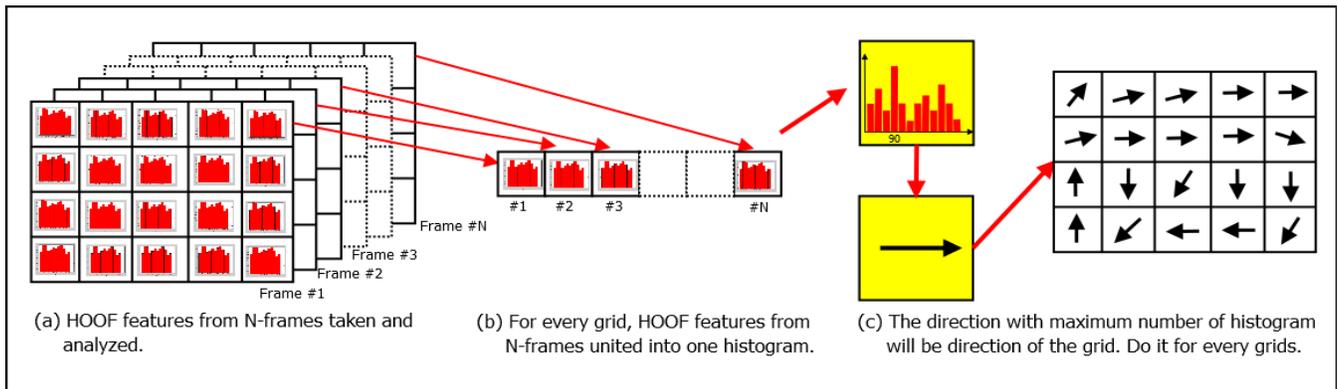Fig. 2. The frame is divided into a number of grids.

(a) Single image from video

(b) Grid creation based on image size and grid size.

(c) Grid visualization example.

(a) HOOF features from N-frames taken and analyzed.

(b) For every grid, HOOF features from N-frames united into one histogram.

(c) The direction with maximum number of histogram will be direction of the grid. Do it for every grids.

Fig. 3. Multiple frames of HOOF feature analysis process for determining the movement direction.

The frame is divided into a number of grids with a certain size (see Fig. 2). The division of frames into grids is based on model parameters as described in the previous sections. Grid size determination is highly dependent on the video dataset used. The grid size that is too large or small can cause movement direction information to be lost or distorted by another movement. In this study, we tested several grid sizes, namely 8x8, 16x16, and 32x32 pixels. With an N x N grid on an image of width W and height H, a grid of m x n will be obtained. The value of m is obtained by dividing the width of the image (W) by the grid size (N). The value of n is obtained by dividing the image height (H) with the grid size (N). If the division yields a fractional value, then the result is rounded up. The box coordinates on Grid (m, n) are (x, y, w, h). The value of x is obtained by equation $x = ((m-1) * N) + 1$ and y with equation $y = ((n-1) * N) + 1$. For w and h values are grid size (N).

Optical flow extraction produces values in complex numbers. Optical flow value consists of two components, the horizontal (u) and vertical (v). The direction of the optical flow (denoted by θ) at a point (x, y) can be calculated based on the

values u and v using (1). The value of θ is an angle between 0 and 360 degrees.

$$\theta_{x,y} = tan^{-1} \frac{v_{x,y}}{u_{x,v}} \qquad (1)$$

After the optical flow direction of each point is obtained, then it is normalized into 12 directions of movement. The normalized direction is denoted by θ_b and has a possible value of θ_b = {0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330}. The calculation of optical flow normalization at point (x, y) is done with (2). The twelve directions of movement are codified into numbers 1, 2, 3, ..., 12. We also add 13 as a code meaning no movement direction. The absence of a movement direction needs to be codified as it allows at a point no movement at all.

To calculate the accumulated HOOF values for each grid, we use the equation $hoof(b) = \sum \theta_h(x,v); (x,v) \in W$ with hoof (b) is HOOF value of direction code b, which b ∈ {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13}, θ_b (x,y) is an angle of b at point (x, y) such that point (x, y) is all point that located at area W. The process of histogram calculation of the optical flow direction is done by summing all the points that have a certain optical flow angle direction. The HOOF value calculation process is

performed for each grid. Thus the HOOF value on the grid (m, n) is the histogram obtained based on the optical flow direction of the grid (m, n) only.

$$\theta_b = \begin{cases} 0 & if & 0 \leq \theta_{x,y} < 15 \text{ or } 345 \leq \theta_{x,y} < 360 \\ 30 & if & 15 \leq \theta_{x,y} < 45 \\ 60 & if & 45 \leq \theta_{x,y} < 75 \\ 90 & if & 75 \leq \theta_{x,y} < 105 \\ 120 & if & 105 \leq \theta_{x,y} < 135 \\ 150 & if & 135 \leq \theta_{x,y} < 165 \\ 180 & if & 165 \leq \theta_{x,y} < 195 \\ 210 & if & 195 \leq \theta_{x,y} < 225 \\ 240 & if & 225 \leq \theta_{x,y} < 255 \\ 270 & if & 255 \leq \theta_{x,y} < 285 \\ 300 & if & 285 \leq \theta_{x,y} < 315 \\ 330 & if & 315 \leq \theta_{x,y} < 345 \end{cases} \quad (2)$$

The next process is analyzing the HOOF features to get the movement direction. It performed against a HOOF features database that coming up from the feature extraction process at an earlier stage. The analysis conducted on several frames at a time on the same grid position as shown in Fig. 3 Step (a) and (b). The number of frames being analyzed is determined based on the parameters of the model (denoted by P). For example, if you want to analyze the movement of the grid (1,1), the entire HOOF value of the grid (1,1) for each frame that is observed is taken and analyzed. HOOF value of the entire grid coupled into a single histogram. After the combined histogram is obtained, then the direction of movement is determined based on the highest histogram values. The process combines the histogram is done for each grid such that the entire grid of a series of frames can be determined the movement direction of the object as shown in Fig. 3 Step (c).

We also applied thresholding process to determine a grid will be analyzed its HOOF value or not. For grids that contain movement, it will be processed, while that does not contain movement does not need to be processed. Thus the computation process becomes faster. Based on the results of experiments that have been done, the best method is the max-min method. The threshold value (T) is obtained by adding the largest HOOF value with the smallest HOOF value, then divided by value 2.

The movement direction in a grid is obtained by accumulating the entire HOOF value of the entire grid analyzed. The HOOF grid values of the selected frames will be accumulated and recalculated to obtain the HOOF grid values that are accumulated from multiple frames at once. In other words, the analysis is not only done for a single frame but done for many frames (multi-frame). The process of accumulating HOOF values between grids is done using (3).

$$hoof_b(m,n) = \sum_{frame=1}^{P} hoof_b(m,n) \quad (3)$$

with $hoof_b$ (m, n) is the histogram of b of some frames on the grid (m, n). P notation is the number of frames analyzed.

III. RESULTS AND DISCUSSIONS

In this section, we explained the experiment results and discussed them.

A. The Experiment Scenario

The experiment on the whole system produced aims to determine the ability of the system of movement direction analysis on the video. It was performed using UMN public dataset [28]. In general, a series of tests are performed to find out:

- The value of success rate (SR) of the proposed model in order to determine the best parameter. The success rate of the detection process of a tested frame is calculated based on the number of grids detected correctly compared to the number of all grids detected by the system. Furthermore, the success rate of the overall tested data is averaged so as to produce the success value of the whole model. The evaluation of the proposed model is validated visually by the expert.

- The values of the accuracy, precision, and recall of the movement direction model using the best parameter. The experiment aims to determine the performance of the proposed model. Testing of accuracy, precision and recall value is performed using the best parameters based on test results. In addition to the accuracy, precision and recall value, it is proposed a measure of the performance of the classification model of the movement direction of the object on the video. It calls the direction detection error rate (DDER). The value of DDER indicates the error rate of the system classifies the movement direction. The error rate is based on how far the direction is detected by the system with the actual direction. For example, if the real direction of the object movement is the 3 o'clock direction, while the system detects it as the clock direction 9, then it has the highest error rate because the direction is opposite. While if the clock direction 3, classified as the direction of the clock 3 also then the error rate is 0 (smallest).

- The comparison of the speed of the HOOF features extraction process from the proposed model using a method that uses segmentation and object detection.
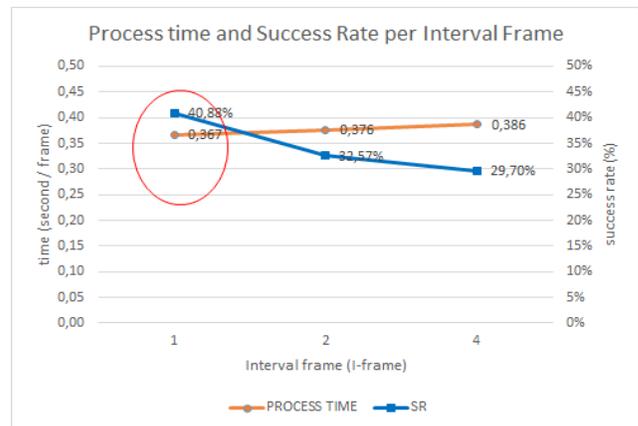

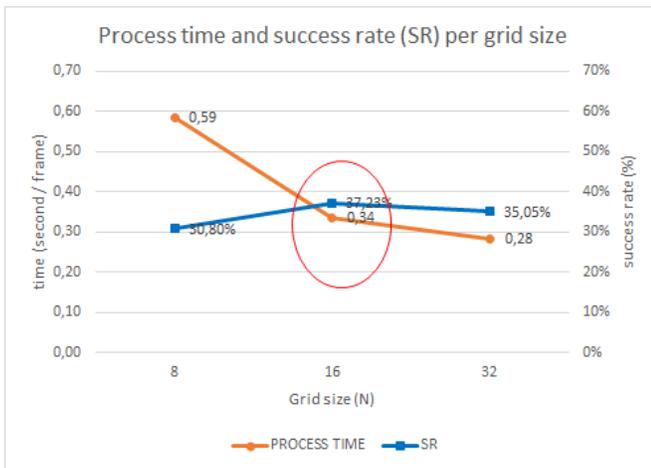Fig. 4. Process time and success rate (SR) per interval frame.

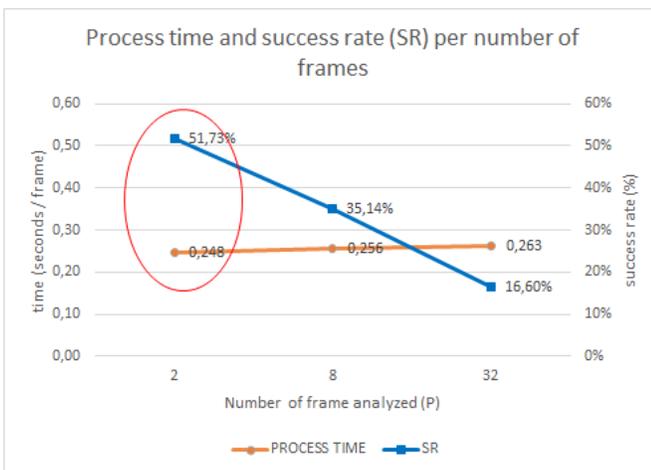Fig. 5. Process time and success rate (SR) per grid size.

Fig. 6. Process tima and success rate (SR) per number of frames.

B. The Testing of Success Rate (SR) of the Movement Direction Detection

The success rate test of the proposed model is performed for each combination of the grid size (N), frame interval (I_{frame}) and the number of frames analyzed (P). It turns out that the resulting SR value has a very wide variation. This happens because the value of SR depends on the value of the parameters used. The diversity of the success rate generated can be seen statistically by calculating the standard deviation value, the mean and the smallest and largest value. The standard deviation from the test result of 23.67% indicates that the data distribution is very spread and varied. The smallest and largest value range is also very wide, i.e., between 0% and 100%.

Based on the test results can also be seen that the average success rate of the proposed model in detecting the movement direction of objects in the video is 34.36%. The result is still below 50% and still needs to be improved in the future. Based on the test results, it can also be concluded that the best success rate is achieved on the test parameter with grid size 16, frame interval 1, and the number of frames 2. To strengthen the conclusion, the best parameter analysis based on two measures is the success rate (SR) and the process speed for each test parameter separately.

Fig. 4 presents the graph of process speed and success rate for each grid size (N). From the graph, it is seen that the best success rate value is achieved on grid size 16, while for the process speed is achieved by the grid size 32. As explained in the previous section that the grid size greatly affects the success rate of the process of determining the movement direction of the object and the optimal value depends on the size of the object in the video. While the speed of the process will be smaller as the size of the grid is used. From the graph, the optimal cutting point is reached at grid size 16, so the best grid size parameter is 16x16.

Meanwhile, the process speed and the success rate of movement direction detection for each frame interval is shown in Fig. 5. Based on the graph, it is seen that the process time increases the value of the larger frame interval. The best process speed is achieved at the frame interval parameter of 1, with a velocity of 0.367 seconds / frame. In line with that, when viewed from the value of success rate, then the greatest success rate value is also achieved with the parameter of the frame interval of 1. The best success rate value is 40.88%. Thus it can be concluded that the best frame interval parameters recommended for use are 1. Next in Fig. 6 presents the process time and success rate for each parameter of the number of frames analyzed (P value). Based on the graph, it is clear that the larger the number of frames that are analyzed then from the side of the process time greater and from the side of success rate decreases. Therefore, it can be concluded that the optimal frame number parameter is worth 2 frames.

C. The Testing of the Accuracy, Precision, and Recall

The testing of accuracy, precision, and recall of the proposed model is done using the best test parameters are grid size 16, frame interval 1, and a number of frames analyzed 2. The accuracy, precision and recall value is calculated by confusion matrix method. The movement direction for each grid is classified into 13 class labels consisting of sets {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}. The class 0 label indicates that there is no movement on a grid.
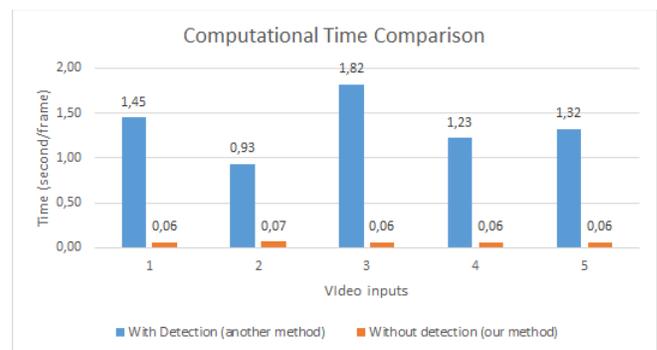

Fig. 7. Computational time comparison of the optical flow and HOOF extraction process.

Based on the test results, the accuracy of the model is quite high. The average accuracy for the entire class is 98.1%. It shows that the system is able to classify the movement direction correctly. Thus the system performance is in a good category. The results appear slightly different at the precision value. The precision value describes the number of correctly classified directions divided by the total grid having direction.

The average precision value of the whole class is only 35.6%. This indicates that the precision of the proposed model still needs improvement. Similarly, the value of recall that shows the level of system effectiveness in recognizing each class label is still quite low, that is equal to 41.2%.

In addition to the accuracy, precision, and recall of the proposed model, this study also calculated the DDER (direction detection error rate). The DDER value indicates the level of movement detection error from the proposed model. The higher the DDER value, the greater the error detection rate and the lower the proposed model's performance. Based on model performance test results, then we calculated DDER value. The resulting DDER value is 25.28%. Based on these values can be concluded that when the system detects the movement direction, the error rate in determining the direction of movement is 25.28%.

D. The Proposed Model Speed Comparison

In this research, we proposed a method of determining the movement direction of objects in a crowd video that does not involve the process of segmentation and object detection first, using the optical flow features that are derived into HOOF, dividing the frame into a number of grids, and HOOF accumulation analysis for multiple frames at once. The main purpose of this research is to produce a method or model of movement direction detection that has faster computation time than the previous method while maintaining its accuracy value.

To know the performance of the resulting model, especially in terms of time compared to other methods that require the process of segmentation and object detection, testing is done by comparing the proposed model with the previous method. Nevertheless, researchers have difficulty in finding similar methods that have the same research objective in determining the movement direction of objects. Therefore, in this study, we only compare computational time for the process of extracting HOOF features. The computation time of the process of determining the direction of movement of the object is not comparable because of the variety of techniques used by other researchers with unequal output results.

Fig. 7 presents computational time comparisons in the extraction process of optical flow and HOOF features. The comparison results show that for the computational time of optical flow features, eliminating the segmentation and object detection process saves computation time of 22 times more efficient. Meanwhile, for the computational time of feature extraction HOOF method without segmentation is slower than the method with segmentation, which is about 42% slower. It can be understood because by doing the detection and segmentation of the object first, the number of areas that the HOOF value is calculated will be less or limited only in the segmented area only.

IV. CONCLUSION

Based on the results of testing and discussion in the previous section some conclusions can be drawn as follows:

1) The video analysis model for determining the movement direction of objects on a video based on optical flow features that derived into Histogram of Oriented Optical Flow (HOOF) proved to be used to detect the movement direction of objects in the video.

2) The test also shows that the success rate of the model is influenced by three main parameters: grid size, frame interval and a number of frames analyzed. The success rate of the model is better on testing with 16x16 grid size, frame interval of 1, and a number of frames analyzed by 2 frames.

3) The testing of the model using the best parameter resulting in an average accuracy of 98.1%, 35.6% precision, 41.2% recall and a direction detection error rate of 25.28%.

4) The proposed model of this study has computation time of HOOF feature extraction which is faster than the similar method that requires segmentation and object detection first, with an average computation time of optical flow extraction of 0.06 seconds / frame and HOOF extraction of 0,29 seconds / frame. The results are 22 times faster than another method that implements segmentation and detection first.

In the future, our approach that proposed in this study will be developed with better optical flow estimation methods such as Deep Neural Network method [29], promising better computational time. Our method also needs to be implemented with another video dataset.

REFERENCES

[1] BNPB, "Data dan Informasi Bencana Indonesia," Badan Nasional Penanggulangan Bencana, 2015. [Online]. Available: http://dibi.bnpb.go.id. [Accessed: 07-Feb-2018].

[2] M. Mustofa, "Memahami kerusuhan sosial, suatu kendala menuju masyarakat madani," J. Kriminologi Indones., vol. 1, no. I, pp. 10–19, 2000.

[3] K. Lang and G. Lang, "Collective Behavior," International Encyclopaedia of the Social Sciences. Crowell Collier and Macmillan, Inc, 1968.

[4] F. Martínez, A. Manzanera, and E. Romero, "A Motion Descriptor Based on Statistics of Optical Flow Orientations for Action Classification in Video-Surveillance," Commun. Comput. Inf. Sci., vol. 346, no. 2, pp. 267–274, 2012.

[5] H. M. Dee and A. Caplier, "Crowd Behaviour Analysis Using Histograms of Motion Direction," in 2010 IEEE 17th International Conference on Image Processing, 2010, pp. 1545–1548.

[6] S. Fazli and H. Fathi, "Video Image Sequence Super Resolution using Optical Flow Motion Estimation," Int. J. Adv. Stud. Comput. Sci. Eng., vol. 4, no. 8, pp. 22–26, 2015.

[7] Y. Benabbas, S. Amir, A. Lablack, and C. Djeraba, "Human Action Recognition Using Direction and Magnitude Models of Motion," in International Conference on Computer Vision Theory and Applications (VISAPP) 2011, 2011, pp. 277–285.

[8] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," in Proc. 7th International Conference on Artificial Intelligence (IJCAI) 1981, 1981, pp. 121–130.

[9] C. Tomasi, "Detection and Tracking of Point Features," 1991.

[10] N. Ihaddadene and C. Djeraba, "Real-time crowd motion analysis," in 2008 19th International Conference on Pattern Recognition, 2008, pp. 1–4.

[11] Y. Benabbas, "Human behavior analysis from videos using optical flow," no. 33. UNIVERSITE DES SCIENCES ET TECHNOLGIES DE LILLE, pp. 1–49, 2011.

[12] R. V. H. M. Colque, C. A. C. Junior, and W. R. Schwartz, "Histograms of Optical Flow Orientation and Magnitude to Detect Anomalous Events in Videos," in Conference on Graphics, Patterns and Images (SIBGRAPI 2015), 2015, pp. 1–9.

[13] T. Li, H. Chang, M. Wang, B. Ni, and R. Hong, "Crowded Scene Analysis : A Survey," IEEE Trans. CIRCUITS Syst. VIDEO Technol., vol. 25, no. 3, pp. 367–386, 2015.

[14] M. Singh, a. Basu, and M. K. Mandal, "Human Activity Recognition Based on Silhouette Directionality," IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 9, pp. 1280–1292, Sep. 2008.

[15] Y. Benabbas, N. Ihaddadene, and C. Djeraba, "Motion pattern extraction and event detection for automatic visual surveillance," EURASIP J. Image Video Process., vol. 2011, pp. 1–15, 2011.

[16] K. Lertniphonphan, S. Aramvith, and T. H. Chalidabhongse, "Human action recognition using direction histograms of optical flow," 2011 11th International Symposium on Communications & Information Technologies (ISCIT). Ieee, pp. 574–579, Oct-2011.

[17] S. Cui, N. Li, and Z. Liu, "Multi-directional crowded objects segmentation based on optical flow histogram," in Proceedings - 4th International Congress on Image and Signal Processing, CISP 2011, 2011, vol. 1, pp. 552–555.

[18] O. Smirg, Z. Smekal, M. K. Dutta, and B. Kakani, "Automatic Detection of the Direction and Speed of Moving Objects in the Video," in Sixth International Conference on Contemporary Computing (IC3), 2013, pp. 86–90.

[19] S. D. Khan, "Estimating speeds and directions of pedestrians in real-time videos: A solution to road-safety problem," in CEUR Workshop Proceedings, 2013, vol. 1122, pp. 1–14.

[20] A. S. Rao, J. Gubbi, S. Marusic, A. Maher, and M. Palaniswami, "Determination of object directions using optical flow for crowd monitoring," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8034 LNCS, no. PART 2, pp. 613–622, 2013.

[21] R. V. H. M. Colque, C. Caetano, M. T. L. de Andrade, and W. R. Schwartz, "Histograms of Optical Flow Orientation and Magnitude and Entropy to Detect Anomalous Events in Videos," IEEE Trans. Circuits Syst. Video Technol., vol. 27, no. 3, pp. 673–682, Mar. 2017.

[22] C. A. Caetano, V. H. C. D. Melo, J. A. dos Santos, and W. R. Schwartz, "Activity Recognition Based on a Magnitude-Orientation Stream Network," in 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2017, pp. 47–54.

[23] A. Solichin, A. Harjoko, and A. E. Putra, "Grid-based Histogram of Oriented Optical Flow for Analyzing Movements on Video Data," in 2015 International Conference on Data and Software Engineering, 2015, pp. 114–119.

[24] P. Wang, C. Chen, C. Dong, H. Xu, and F. Tian, "The analysis method of video camera's motion based on optical flow and slam," in 2016 International Conference on Audio, Language and Image Processing (ICALIP), 2016, pp. 62–66.

[25] B. K. P. Horn and B. G. Schunck, "Determining optical flow," Artif. Intell., vol. 17, no. 1–3, pp. 185–203, Aug. 1981.

[26] J. L. Barron, D. J. Fleet, S. S. Beauchemin, and T. A. Burkitt, "Performance of Optical Flow Techniques," in 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition., 1992, pp. 236–242.

[27] S. S. S. Beauchemin and J. L. L. Barron, "The Computation of Optical Flow," ACM Comput. Surv., vol. 27, no. 3, pp. 433–467, 1995.

[28] UMN, "Unusual Crowd Activity Dataset of University of Minnesota, Department of Computer Science and Engineering," UMN, 2006. [Online]. Available: http://mha.cs.umn.edu/movies/crowd- activity-all.avi. [Accessed: 10-Aug-2017].

[29] X. Zhang, D. Ma, X. Ouyang, S. Jiang, L. Gan, and G. Agam, "Layered Optical Flow Estimation Using a Deep Neural Network with a Soft Mask," IJCAI, vol. May, no. 2018, 2018.