Mixture gaussian V2 based microscopic movement detection of human spermatozoa

Ariyono Setiawan [a,1], I Gede Susrama Mas Diyasa [b,2,*], Moch Hatta [c,3], Eva Yulia Puspaningrum [b,3]

[a] Department of Air Transportation Management, Aviation Polytechnic of Surabaya, Indonesia
[b] Department of Informatics Engineering, Universitas Pembangunan Nasional "Veteran" Jatim, Indonesia
[c] Department of Computer Engineering, Universitas Maarif Hasyim Latif Sidoarjo, Indonesia
[1] rmaryo4u@gmail.com; [2] igsusrama.if@upnjatim.ac.id; [3] moch.hatta@dosen.umaha.ac.id; [4] evapuspaningrum.if@upnjatim.ac.id
* corresponding author

ARTICLE INFO

ABSTRACT

Healthy and superior sperm is the main requirement for a woman to get pregnant. To find out how the quality of sperm is needed several checks. One of them is a sperm analysis test to see the movement of sperm objects, the analysis is observed using a microscope and calculated manually. The first step in analyzing the scheme is detecting and separating sperm objects. This research is detecting and calculating sperm movements in video data. To detect moving sperm, the background processing of sperm video data is essential for the success of the next process. This research aims to apply and compare some background subtraction algorithms to detect and count moving sperm in microscopic videos of sperm fluid, so we get a background subtraction algorithm that is suitable for the case of sperm detection and sperm count. The research methodology begins with the acquisition of sperm video data. Then, preprocessing using a Gaussian filter, background subtraction, morphological operations that produce foreground masks, and compared with moving sperm ground truth images for validation of the detection results of each background subtraction algorithm. It also shows that the system has been able to detect and count moving sperm. The test results show that the MoG (Mixture of Gaussian) V2 (2 Dimension Variable) algorithm has an f-measure value of 0.9449 and has succeeded in extracting sperm shape close to its original form and is superior compared to other methods. To conclude, the sperm analysis process can be done automatically and efficiently in terms of time.

1. Introduction

Infertility can occur in both women and men. One of the factors that cause infertility in men is poor sperm quality, such as low sperm concentration. Medically, a normal sperm concentration is 20 million or more sperm/ml of semen [1]. Sperm test may determine the quality of sperm. An expert uses a microscope to perform the test. As a consequence, assessment is subjectively performed through observations by experts on a set of parameters.

Research on sperm analysis continues to be developed as in research conducted by several researchers. One of them conducts sperm analysis through classification based on the shape of the sperm head by using threshold segmentation and decision trees [2]. Sperm images are from the WHO standard book [1]. The researchers conducted a preprocessing (image adjusting process), segmentation process using the Otsu threshold method, and classification using a decision tree to distinguish between normal and abnormal spermatozoa heads [2]. Similarly, Susrama et al. [3] created the Automated Analysis of Sperm

Concentration Counters (A2SC2) system using the Otsu and morphological threshold segmentation processes. Other studies [4][5] discussed the analysis of sperm motility by dividing it into three modules: head detection, head tracking, and flagellum tracing. This unique framework aims at providing both the head trajectories and the flagella beat patterns to assess sperm motility quantitatively.

In order to obtain good tracking results, the Background subtraction process is needed. It may separate moving objects (foreground) and background in the video. The background subtraction process can detect sperm motility in the video. One example of sperm detection related to the background subtraction process was conducted by Hidayatullah *et al.* [6]. They proposed the Gaussian Mixture Model with Hole Filling Algorithm in the sperm detection process, compared to the other six methods for detecting motilities.

The sperm motility measurement in the video required a sperm detection method [7], such as the Adaptive Local Threshold (ALT) and Morphology (M). The method works by separating the background of the object to be detected, removing unwanted objects, and then detecting the ellipse, which is assumed to be the sperm head. The sperm detection methods using Adaptive Local Threshold and Morphology have a detection accuracy of up to 82%. At the same time, Ghasemian *et al.* [8] focuses on Sperm morphology analysis (SMA) using images of sperm cells, by detecting and analyzing various parts of human sperm. The first step of this research is to eliminate background noise. Then the system recognizes the sperm part of the head, tail and analyzes the size and shape of each part by classifying normal and abnormal. As previously explained, this study only detects images and is different from the research conducted by analyzing spermatozoa videos.

Maggavi *et al.* [9] and Keskenler *et al.* [10] implemented the first step of the syllable frame difference algorithm for background subtraction. There were some limitations in selecting the appropriate threshold values [8]. The precision of the output depends significantly on the threshold value selected, used non-linear diffusion filtering in time to eliminate this dependence. Li *et al.* [11] performed sperm detection based on its contours using the Gaussian Modeling algorithm to segment the foreground. Sperm counts as foreground since it continues to move on video frames in the research data. Furthermore, Khachane *et al.* [12] divided the background subtraction process based on the method used. The divisions are Basic methods, Fuzzy based method, Statistical method, Type-2 Fuzzy based methods, Statistical methods using color and texture features, Non-parametric methods, Methods based on eigenvalues and eigenvectors, and Neural and neuro-fuzzy methods.

Some of the researched data is image processing so that the process of eliminating noise in the background requires low computational time. In other studies, the processed data is spermatozoa video data or moving sperm data, but video data that is synthetic data and video data that are already available. Video data has a frame rate of 30 fps, and the magnification of video is 20x. The method used for the detection of moving objects uses statistical methods but only uses one or two background subtraction algorithms even though they have quite high accuracy results. In this research, sperm video data from several volunteer scanners were taken in real-time, with a frame rate of 50-60 fps, and data processing for background subtraction using 23 algorithms, both from basic methods and statistical methods. The goal is to get one best algorithm out of 23 algorithms tested. Thus this study will contribute to further research relating to the analysis of sperm infertility levels in determining the right algorithm for sperm detection and calculation.

2. Method

This research compared moving object detection methods in the case of human sperm detection. Fig. 1 shows the flowchart that consists of two parts. The first part is the process of sperm video production by testing 25 human sperm video data, and the second part is sperm video processing to enable the detection and calculation of human sperms.

There were four processes done in the second part. The first process was preprocessing using a Gaussian filter [13] on every frame read from the sperm video. The next was the process of background subtraction [14]. The result of this process was a binary image, which represented the area of moving

objects visible in the frame. A morphological operation consisting of opening operation and closing operation [15] followed the subtraction, eliminated noises, and perfected the shapes of extracted moving sperms. Subtraction algorithm used to validate the detection results of each background. The study compared the foreground mask results of the morphological operations with the ground truth image of moving sperms - the results of manual observation. Moreover, for visualization purposes, any blob area (white objects on the binary images) on the foreground mask would be given bounding boxes on the original frame. Simultaneously, the number of blob objects that were on the foreground mask would be calculated. The result determined that the system had been able to perform detection and calculation of moving sperm.
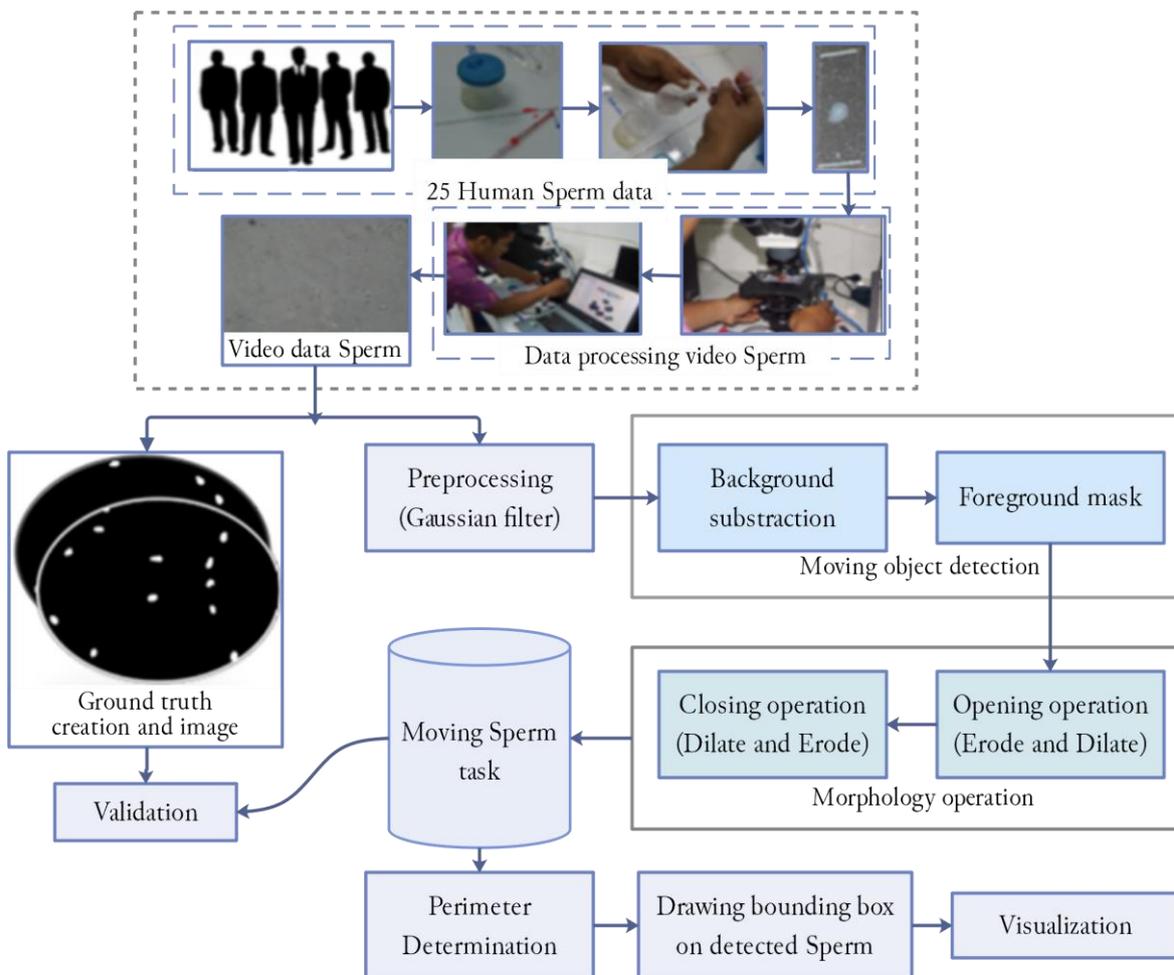

Fig. 1. Overview of the sperm detection and calculation system.

2.1. Materials

The sperm data used was data from one of the patients who performed fertility tests in the integrated laboratory of the Politeknik Kesehatan Surabaya, Indonesia. Following the WHO standard sperm test procedure [1], sperm was retrieved directly from the patient after ejaculation. Then, the sperm that has been liquefied for 10-20 minutes as well and dripped on object-glass, than was observed using a bright-field microscope connected with a Point Gray camera type FL3-U3-13S2C-CS. However, an essential early stage in sperm infertility research is the sperm detection phase or sperm object separation from images/videos obtained from observations on semen. The success rate in separating sperm objects from cement fluids has an essential role for further analysis of sperm objects, as shown in Fig. 2, not all sperm videos taken in real-time yield results cleared of background. Fig. 2(a)-(c) were taken by microscope using the Point Gray type FL3-U3-13S2C-CS camera with 100 X magnification. Fig. 2(a) is cleaner than pictures Fig. 2(c) and the picture Fig. 2(b), which are the results of videos unclean background.
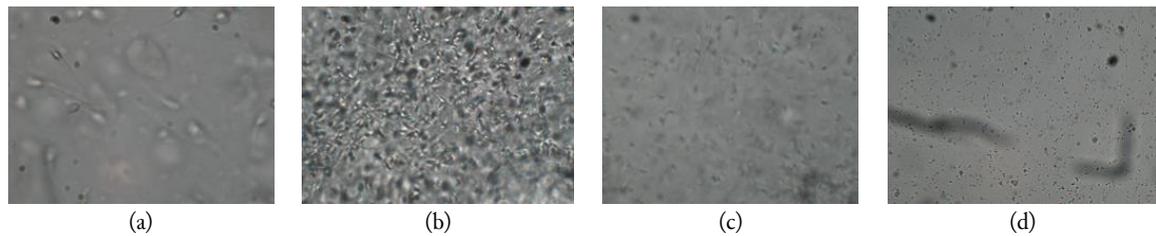
Fig. 2. Sperm shapes from the results of data collection using Point Gray Point type FL3-U3-13S2C-CS camera

2.2. Preprocessing (Gaussian filter)

Preprocessing is the original data processing before the sperm video frame data is processed. This process aims to remove noise, clarify data features, minimize/enlarge data size, and convert the original data to obtain data as needed. The preprocessing method used in this study was the Gaussian filter. The filter may increase the image frames of the video blur, yet decreasing the noise, and the detail on the image [16] would affect the results in the next process. Equation (1) expresses one dimension Gaussian function.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \tag{1}$$

where σ is the standard deviation of the distribution that assumed to have 0 mean value. When applied to images, 2D Gaussian distribution is required. Therefore, there were two 1D Gaussian distributions used, each for the x-axis and y-axis [17]. Equation (2) presents the 2D Gaussian distribution.

$$G(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2 y^2}{2\sigma^2}} \tag{2}$$

The amount of kernel used on the Gaussian filter in this research was 5x5, as in Fig. 3.

$$\frac{1}{273} \begin{array}{|c|c|c|c|c|} \hline 1 & 4 & 7 & 4 & 1 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 7 & 26 & 41 & 26 & 7 \\ \hline 4 & 16 & 26 & 16 & 4 \\ \hline 1 & 4 & 7 & 4 & 1 \\ \hline \end{array}$$

Fig. 3. Gaussian Kernel 5x5

2.3. Background subtraction (BS)

BS detected a foreground mask (a term for a binary image containing information about moving objects in a frame) in a video recording or a camera capture. This technique is often and commonly used in the field of image processing or computer vision. Calculating the difference between the video's current frame value and the background model obtained the foreground mask. According to Sheng-Yi et al. [18], the background subtraction process outline can be divided based on the method used. The divisions are Basic methods, Fuzzy based methods, Statistical methods, Non-parametrical methods, a method based on eigenvalues and eigenvectors, Neural and neuro-fuzzy methods. The ones used in this research were the Basic Method and the Statistical Method [19].

In a study of sperm infertility, the sperm that will eventually fertilize the egg is the sperm that keeps moving. Therefore, the background subtraction process is necessary for detecting moving sperm. In the case of sperm detection, the advantages of the background subtraction process are the data used has uni-modal characteristics, the distance between frames is short, and the effect of light changes is absent. On the other hand, the challenge faced is the presence of background objects that move and the existence of new objects is considered background. The input from the background subtraction process is preprocessed video frames, while the output is binary images that represent the objects (sperm) that move in the video.

2.3.1. Basic background subtraction models

The basic model performs mathematical calculations in conducting background image modeling, which calculates the value differences between frames, obtains the background model by calculating the average of the frame history, and calculates the variance of the frame history. The basic model algorithms used in this research were weighted moving means (WMM), weighted moving (WM), and variance frame difference (FD) [20]. The difference calculation between the background image and the current frame of the video obtained the pixel value, which was classified as foreground. We used a thresholding process to convert the foreground image into a binary image. As a result, the pixels representing the moving objects would have a value of 1, while the regions' background would have a value of 0. The primary method algorithms used in this research were WMM, WM, and variance FD. Frame Difference is expressed by (3) and (4) [21].

$$B = f_{t-1} \tag{3}$$

$$F = |f_t - B| \tag{4}$$

where the frame (*f*) means a moment before the current frame, background (*B*), and foreground (*F*).

Equation (5) shows the Weighted Moving Mean [22].

$$B = (f_t.w_1) + (f_{t-1}.w_2) + (f_{t-2}.w_3) \tag{5}$$

where frame (*f*) is the current frame and two previous frames *(frame_t, frame_{t-1}, and frame_{t-2})*, and *weight (w)*.

Weighted Moving Variance modeled the background image by counting the moving variance [22]. In calculating the average value (μ), this algorithm uses the formula of the weighted moving mean algorithm. Equation (6) used to model the background image (B) in the Weighted Moving Variance algorithm.

$$B = \sqrt{0.5} \sum_{i=1}^{t}(f_i - (f_i. w_1) + (f_{i-1}. w_1 + (f_{i-1}. w_2))^2 \tag{6}$$

Adaptive Background Learning involves subsequent sequential frames to get motion differences from the image and forecast the foreground and background areas [23]. Two classes are used for segmentation: class 1 as background and class 2 to note the foreground point. It aims to provide background images regularly based on steady motion and image segmentation.

2.3.2. Statistical background subtraction models

The statistical models appropriate in addressing the situation are camera automatic adjustment, moved background object, and inserted background object. This situation corresponds to the challenges present in detecting moving sperms. The background subtraction algorithms in this model statistically modeled every pixel in the frame to be then classified into the categories of foreground pixels background pixels. The Single Gaussian are represented as in (7), (8), and (9) [24].

$$\mu(x, y, t) = \sum_{i=1}^{t} p(x, y, i)/t \tag{7}$$

$$\sigma(x, y, t) = (\sum_{i=1}^{t}(p^2(x, y, i)/t) - \mu^2(x, y, t))^2 \tag{8}$$

$$|\mu(x, y, t) - p(x, y, t)| > c.\sigma(x, y, t) \tag{9}$$

where the mean value (μ) and standard deviation (σ) are in the YUV color range, $P(x, y, i)$ is the intensity value of the pixel in the position (x, y) at time i c is the value of a particular constant value.

The Gaussian Mixture (GMM) is a type of density model consisting of components of Gaussian functions [25]–[27]. For each pixel, $\{X_1, \ldots, X_t\}$, is modeled with a Gaussian distribution K mixture. As for taking the probability value in each pixel, it is obtained through (10), (11), and (12).

$$P(X_t) = \sum_{i=1}^{k} W_{i,t} \cdot \eta(x_t | \mu_{i,t}, \Sigma i, t) \tag{10}$$

$$\eta(x_t | \mu_{i,t}, \Sigma i, t) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma i,t|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_t - \mu_{i,t})^T \Sigma i,t^{-1}(x_t - \mu_{i,t})^T} \tag{11}$$

$$B = \arg_b \min\left(\sum_{j=1}^{b} w_j > T\right) \tag{12}$$

The mathematical model for selecting the first distributed background *(B)* is defined as in (12). K is the number of distributions, $\omega_{i,t}$ is the estimated weight of the Gaussian mixture at the time t, $\mu_{i,t}$ is the mean value of the Gaussian mixture at the time t, $\sum i, t$ is a covariance matrix to Gaussian-mixture at time t, η is the Probability Density Function, $|\sum i, t|$ is the determinant of covariance, the T power is a transpose matrix, the power of -1 is an inverse matrix, e is exponential, and π (phi) is a scalar image size X, as well as vector image X (RGB)

Kernel Density Estimation (KDE) is a non-parametric statistical approach to estimating the probability distribution function of a random variable if it is assumed the shape or distribution model of the random variable is unknown [28]. KDE is defined as in (13).

$$Pr(X_t) = \frac{1}{2}\sum_{i-1}^{n} K(X_t - X_i) \tag{13}$$

The foreground is detected using the following rules: If $Pr(X_t) < TH$, then the pixel includes the foreground. Also, pixels are included in the background. This algorithm is the same as GMM, able to adapt to the multimodal background, but it is not necessary to estimate the parameters of Gaussian. A mixture of Gaussians (MoG) [29][30] is a distribution that has a pdf (probability density function) of the kind with (14) and (15).

$$p(x) = \sum_{j=1}^{M} p(x | \mu_j, \Sigma_j) P_j \tag{14}$$

$$0 \le P_j \le 1 \text{ and } \sum_{j=1}^{M} P_j = 1 \tag{15}$$

where $p(x | \mu_j, \Sigma_j)$ is the pdf of a Gaussian with parameters μ_j, Σ_j, where the *mixture coefficients* P_j satisfy [30].

2.4. Morphology operation

After the background subtraction process was done, then the foreground mask was obtained, which is in the form of a binary image, representing the pixels moving. The resulting foreground image still had noises, and sometimes the moving sperm objects extracted were not intact; they were divided into two or more parts [31]. For this reason, this research applies morphological operation that performed in this study was the opening and closing operation.

The opening operation occurs morphological erosion operation followed by dilated morphology operation. The erosion process aimed to eliminate noise that appeared in the image foreground from the background subtraction process. Then, the dilation process was performed to restore the shape of the changed object after the first erosion process. The closing operation process is morphological dilation operation followed by erosion morphology operation. The dilation process may close small holes in image sperm and connect separate sperm shapes. The last in-depth erosion process in the closing operation refines the shapes of detected moving sperms. The elemental structure used in morphological operations performed has an ellipse shape with a kernel 5x5. The shape of the elemental structure can be seen in Fig. 4.

0	0	1	0	0
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
0	0	1	0	0

Fig. 4. The structure of the element used

2.5. Ground truth creation and image

A ground truth image in this research is an image that contains information about the real area of the moving sperm object in the frame of sperm data video. The ground truth images were obtained by manually observing the regions on video frames that have moving sperm objects.

For sperm movement verification, ten frames before and ten frames after the frame were observed, from which aground truth image would be drawn. For example, the ground truth image in the 50th frame was created by observing sperm movement from frame 40 to frame 60 of the video. The area where there was a moving sperm object was marked by assigning a 255 (white) pixel value, and an area that does not have a moving sperm object was characterized by giving a 0 (black) pixel value. In this manner, a ground truth image is formed, which will refer to the testing process of the results of the detection and calculation of sperm. Fig. 5 illustrates the ground truth image on frame 50.
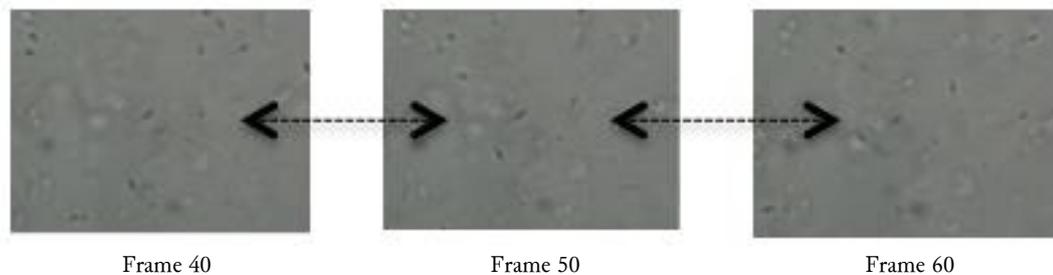

| Frame 40 | Frame 50 | Frame 60 |

Fig. 5. An illustration of ground truth image creation

2.6. Contour detection and validation

In the previous process, the morphological operation had successfully eliminated noise, which arose due to the foreground separation process from the background and refinement of the extracted sperm form, obtaining a binary image, which represents moving sperms. Any blob on the binary image is a representation of a moving sperm object. From the resulting image of the morphological process, the sperm object is detected based on its contour, so the contour shape, the contour area, and the midpoint of the sperm position on the frame can be known. Based on this information, any detected sperm will be given a bounding box and count in the original frame of the video, so it can be observed that the system succeeded in detecting sperm.

The validation process was conducted by comparing the results of moving sperms from each algorithm tried with ground truth images, which were the result of the manual observation of the location of sperm location moving in the video frame. The comparison result was then analyzed using the receiver-operating characteristic, so the level of validation of each algorithm tried can be known.

The detection result of each algorithm used will be compared with the ground truth image, obtaining three values, namely False Positive (FP), True Positive (TP), and False Negative (FN). True positive is for an existed sperm that was detected as existing. False Positive (FP) indicates an existing sperm that was undetected. False Negative (FN) for a sperm that did not exist, but was detected. From the above test results, the precision, recall, and f-measure values were calculated, with equations (16), (17) and (18) [32].

$$Precision = tp/(tp + fp) \qquad (16)$$

$$Recall = tp/(tp + fn) \tag{17}$$

$$f - Measure = 2((Precision * Recall)/(Precision + Recall)) \tag{18}$$

3. Results and Discussion

The testing results shown were the results of preprocessing, background subtraction, morphological operation, and sperm count detection and calculation. The testing compares sperm detection results and calculation of sperm with ground truth images of moving sperm objects. The comparison was then analyzed using receiver operating characteristics.

3.1. Preprocessing results

The preprocessing is a Gaussian filter with a kernel size of 5x5. The inputs of preprocessing were video frames of the observed semen liquid. Fig 6(a) presents an example of the frame of the sperm video data used while the result of preprocessing is in Fig. 6(b). The preprocessing eliminated white noises that appeared from the camera, blurring the images, and reducing the detail of the image.
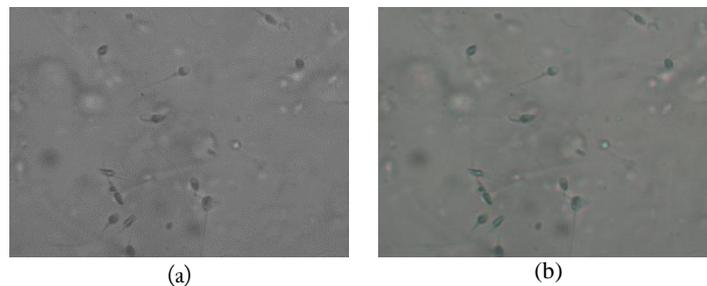

(a)   (b)

Fig. 6. An original frame from sperm video (a) and the frame after preprocessing (b).

3.2. Ground truth image results

The ground truth images were obtained by manually observing the regions in video frames that have moving sperm objects. These ground truth images would be used as a comparison of the results of the detection and calculation of sperms. The observation of ten frames before and 10 frames after the frame may ensure the movement, from which aground truth image would be drawn. For example, the ground truth image in the 50th frame was created by observing sperm movement from frame 40 to frame 60 of the video. An example of an original frame and its ground truth image in frame 60 and 90 can be seen in Fig. 7.
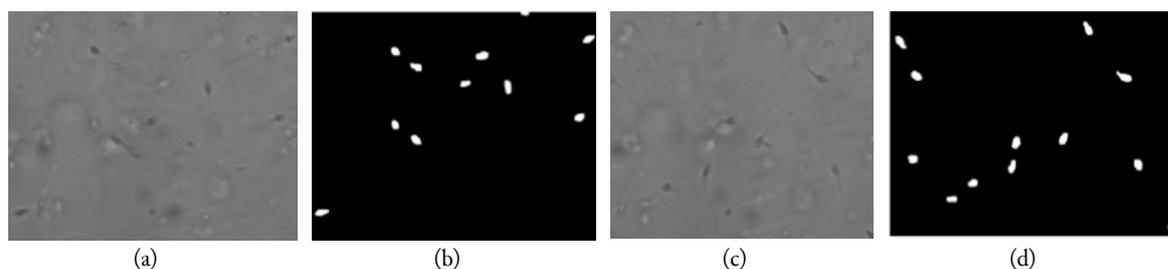

(a)   (b)   (c)   (d)

Fig. 7. Ground truth image results: (a) Frame 60 (b) Ground truth of frame 60 (c) Frame 90 (d) Ground truth of frame 90

An example of background subtraction result from Fig. 8 (Frame Difference) and Fig. 9 (Grimson GMM). The red box in the resulted image is an area where there are background objects which are moving in the video scene, and the yellow box is an example of an area where there is a detected moving sperm object. In the yellow box, the detected sperm looks divided into several parts, including noises around the detected sperm.

This also happens to all detected moving sperm, as seen in the foreground mask picture in Fig. 8 and Fig. 9. One detectable moving sperm object can appear divided into several parts. Between the sperm head and tail, the area of the sperm head was divided into two. Noises also appear in other areas on the

frame, not just around the moving sperm region. In the red box, it can be seen that a moving background object was also detected. The background object detected was also divided into sections like the moving sperm object. However, it was the background object movement that was detected.
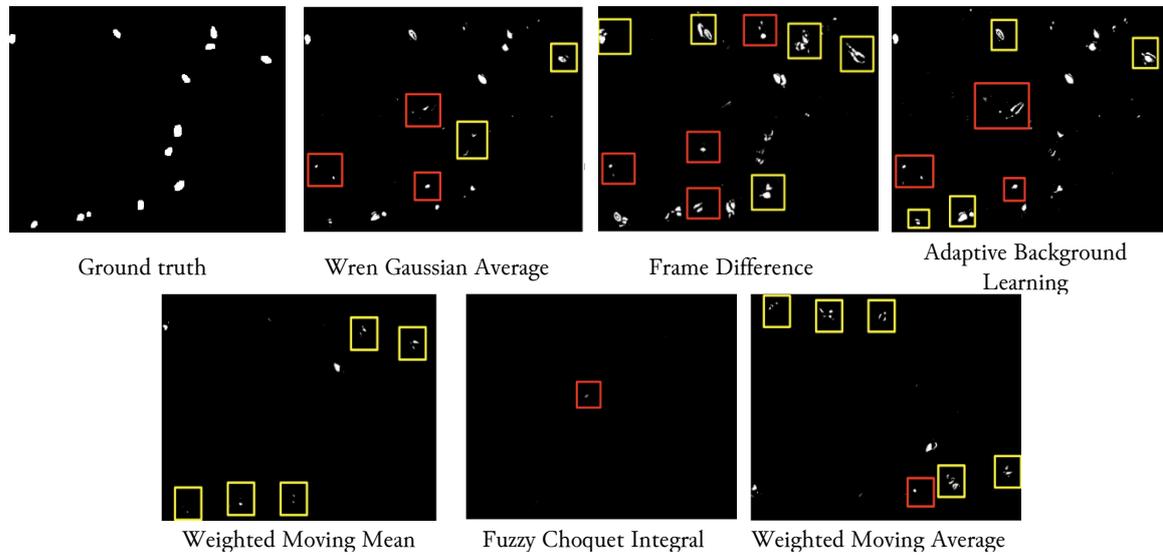

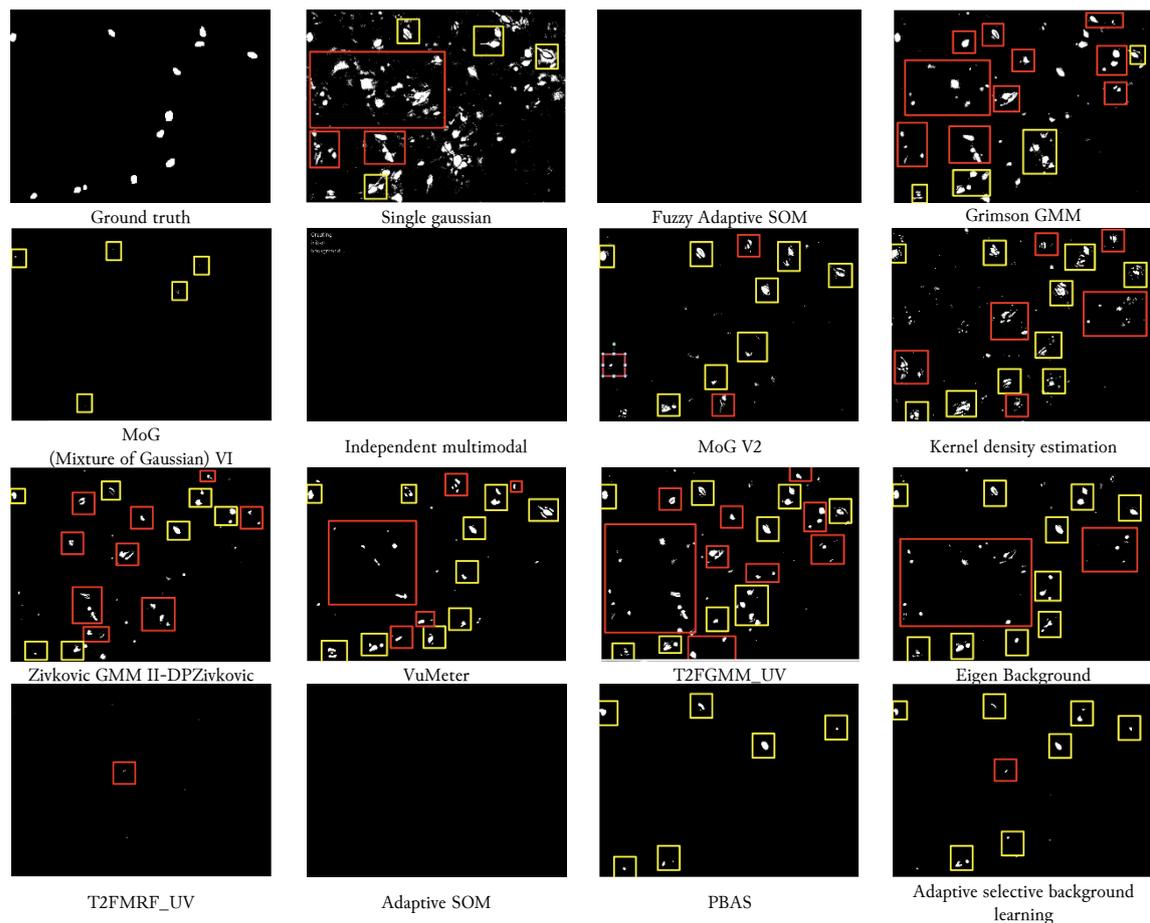Fig. 8. Testing Results of basic model algorithm

Fig. 9. Testing Results of statistic model algorithm.

3.3. Results of morphological operation

The Morphological Operation conducted consisted of an Opening Operation followed by a Closing Operation, or successively erosion-dilation-dilation-erosion. The opening operation aimed to eliminate

noises, which appeared on the foreground image of the background subtraction results, and returned the morphed object shape after the noise removal process (erosion). The Closing operation aimed to cover small holes in objects, connect separate sperm forms, and perfect the shape of the detected moving sperm.

The input of the morphological operation process was foreground mask resulting from background subtraction, where the image has noises. Sometimes, the moving sperm object is not intact divided and was into two parts or more. Moreover, results from the morphological operation were a foreground mask in which the noises had been removed. The detected sperm object is intact so that each BLOB (large binary object) that exists became a representation of a moving sperm object on the video frame. Any background subtraction algorithm produced varying foreground mask, and the results of each morphological operation of each background subtraction algorithm is explained in the next sections.

3.4. Sperm detection and calculation of the test

After the process of morphological operation, with an assumption of the foreground mask that was already formed, it did not have noise, and severed sperm objects had been reconnected. Any blob (Binary Large Object) that was on the foreground mask became a representation of a moving sperm object. For visualization purposes, each blob that was formed would be detected based on its contour so that its contour shape, total contours, and the location of the midpoint of each moving sperm object can be known. From this information, the bounding box and the sequence of sperm moves detected on the original frame from the video were given, so it can be observed that the system has succeeded in performing the detection and calculation of sperms. The details of the results of sperm detection and calculation and its test results can be seen in Table 1.

Table 1. Test Results of sperm detection and calculation and the validation process

Background Subtraction Models	Algorithm	Test Results of sperm detection and calculation			Results of the validation process		
		True Positive	*False Negative*	*False Positive*	*Precision*	*Recall*	*F-Measure*
Basic Model	Frame Difference	112	0	25	0.8175	1	0.8995
	Weighted Moving Mean	12	100	0	1	0.1071	0.1935
	Weighted Moving Average	23	89	0	1	0.2053	0.3407
	Adaptive Background Learning	**110**	**2**	**17**	**0.8661**	**0.9821**	**0.9205**
	Fuzzy Choquet Integral	0	112	0	0	0	0
	Wren Gaussian Average	74	38	0	1	0.6607	0.7956
Statistical Model	Single Gaussian	112	0	417	0.2117	1	0.3494
	Grimson GMM	112	0	47	0.7044	1	0.8265
	MOG V1	1	111	0	1	0.0089	0.0176
	MOG V2	**103**	**9**	**3**	**0.9716**	**0.9196**	**0.9449**
	Zivkovic GMM	106	6	37	0.7412	0.9464	0.8313
	T2FGMM_UV	109	3	97	0.5291	0.9723	0.6855
	T2FMRF_UV	52	60	2	0.9629	0.4642	0.6265
	PBAS	70	42	0	1	0.625	0.7692
	VuMeter	92	20	10	0.9019	0.8214	0.8598
	Eigen Background	112	0	53	0.6787	1	0.8086
	Adaptive SOM	99	13	44	0.6923	0.8839	0.7764
	Fuzzy Adaptive SOM	87	12	17	0.8365	0.8787	0.8571
	Adaptive Selective Background Learning	40	72	0	1	0.3571	0.5263
	Independent Multimodal	102	12	13	0.8869	0.8947	0.8908
	Kernel Density Estimation	112	0	73	0.6054	1	0.7542

The testing of detection and calculation of sperms would be done by manually comparing the results of detection and calculation of sperm with images ground truth of moving sperm objects, which was the result of manual observation. This comparison was made 10 times by taking the detection results on every 30 frames of the video, thus forming the sequence of frames: 30, 60, 90, 120, 150, 180, 210, 240,

270, and 300. The comparison results were then analyzed using ROC Analysis, obtaining the value: True positive (TP) for existing sperms being detected. False Positive (FP) for a real sperm that was undetected. False Negative (FN) for a sperm that did not exist, but was detected. After the values are obtained, then the f-measure, recall, and precision values of each algorithm used were calculated, so it can be determined which background subtraction algorithm is the most appropriate in detecting and calculating moving sperms in a video.

The Basic algorithm test model to detect and calculate sperm count includes the Frame Difference, Adaptive Background Learning, Weighted Moving Mean, Fuzzy Choquet, Integral, Weighted Moving Average, and Wren Gaussian Average algorithms. The test results of the six basic model algorithms, the Adaptive Background Learning algorithm has the highest accuracy, followed by the Frame Difference and Weighted Moving Average (Table 1). Meanwhile, Fuzzy Choquet Integral cannot be used for detection analysis and calculating sperm count because of its accuracy. The statistical models' test shows that the highest accuracy is achieved by the MoG V2 algorithm, while MoG V1 has the lowest accuracy. If both the best models, Adaptive Background Learning and MoG V2 algorithms, were compared, MoG V2 has the highest f-measure of 0.9449 compared to Adaptive Background Learning with 0.9205.

4. Conclusions

In this research, we have implemented a series of procedures for double-tracking spermatozoa motility based on the basic model and statistical model algorithms and created a complete integrated system from data collection to analysis. The MOG V2 algorithm in the background subtraction process was capable of detecting and performing calculations of moving sperm objects in the video. The algorithm is the resulting foreground with little noise, background object moving in the video being undetected as foreground objects, and sperm forms being extracted more perfectly. The test results of performing detection and calculation of sperm motility show that MOG V2 has the highest f-measure value of 0.9449 compared to the other tested background subtraction algorithms in this research. Another best performance algorithm is the Adaptive Background Learning algorithm that has an f-measure value of 0.9205. It shows that the difference value between the MOG V2 and Adaptive Background Learning is only 0.0244. It means that the MOG V2 and Adaptive Background Learning algorithms are promising for detecting and calculating sperm motility. Both algorithms have successfully overcome the challenges and advantages that exist in this case.

Declarations

Author contribution. All authors contributed equally as the main contributor to this paper. All authors read and approved the final paper.
Funding statement. The authors received no specific funding for this work.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] W. H. Organisation, *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction*, fifth edition, Cambridge university press, 2010, available at: Google Scholar.

[2] I. G. S. Masdiyasa, I. K. E. Purnama, and M. H. Purnomo, "Teratozoospermia Classification Based on the Shape of Sperm Head Using OTSU Threshold and Decision Tree," *MATEC Web Conf.*, vol. 58, p. 03012, May 2016, doi: 10.1051/matecconf/20165803012.

[3] I. G. Susrama, K. E. Purnama, and M. H. Purnomo, "Automated Analysis of Human Sperm Number and Concentration (Oligospermia) Using Otsu Threshold Method and Labelling," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 105, p. 012038, Jan. 2016, doi: 10.1088/1757-899X/105/1/012038.

[4] M. U. Daloglu and A. Ozcan, "Computational imaging of sperm locomotion," *Biol. Reprod.*, vol. 97, no. 2, pp. 182–188, Aug. 2017, doi: 10.1093/biolre/iox086.

[5] B. J. Walker, K. Ishimoto, and R. J. Wheeler, "Automated identification of flagella from videomicroscopy via the medial axis transform," *Sci. Rep.*, vol. 9, no. 1, p. 5015, Dec. 2019, doi: 10.1038/s41598-019-41459-9.

[6] P. Hidayatullah, T. L. E. R. Mengko, and R. Munir, "A Survey on Multisperm Tracking for Sperm Motility Measurement," *Int. J. Mach. Learn. Comput.*, vol. 7, no. 5, pp. 144–151, Oct. 2017, doi: 10.18178/ijmlc.2017.7.5.637.

[7] I. G. Susrama, Tobing L, I. D. G. H. Wisana, I. K. E. Purnama, and M. H. Purnomo, "Counting the Amount of Spermatozoa Active Perframe Video Using Morphology and Local Adaptive Threshold," in *Nusantara Science and Technology Proceedings*, 2018, pp. 118–126, doi: 10.11594/nstp.2018.0118.

[8] F. Ghasemian, S. A. Mirroshandel, S. Monji-Azad, M. Azarnia, and Z. Zahiri, "An efficient method for automatic morphological abnormality detection from human sperm images," *Comput. Methods Programs Biomed.*, vol. 122, no. 3, pp. 409–420, Dec. 2015, doi: 10.1016/j.cmpb.2015.08.013.

[9] R. R. Maggavi, S. A. Pujari, and V. C.N, "Motility Analysis with Morphology: Study Related to Human Sperm," *Procedia Comput. Sci.*, vol. 152, pp. 179–185, 2019, doi: 10.1016/j.procs.2019.05.041.

[10] M. F. Keskenler, A. Hasiloglu, G. T. Ozyer, B. Ozyer, and E. Simsek, "Sperm Detection and Analysis Using Feature Description Algorithms," in *2019 27th Signal Processing and Communications Applications Conference (SIU)*, 2019, pp. 1–4, doi: 10.1109/SIU.2019.8806287.

[11] Q. Li *et al.*, "Automatic human spermatozoa detection in microscopic video streams based on OpenCV," in *2012 5th International Conference on BioMedical Engineering and Informatics*, 2012, pp. 224–227, doi: 10.1109/BMEI.2012.6513003.

[12] M. Y. Khachane, R. R. Manza, and R. J. Ramteke, "Fuzzy rule based classification of human spermatozoa," in *2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO)*, 2015, pp. 1–5, doi: 10.1109/EESCO.2015.7253656.

[13] Kshema, M. J. George, and D. A. S. Dhas, "Preprocessing filters for mammogram images: A review," in *2017 Conference on Emerging Devices and Smart Systems (ICEDSS)*, 2017, pp. 1–7, doi: 10.1109/ICEDSS.2017.8073694.

[14] S. Kumar and J. Sen Yadav, "Background Subtraction Method for Object Detection and Tracking," 2017, pp. 1057–1063, doi: 10.1007/978-981-10-1708-7_125.

[15] A. Raid, W. Khedr, M. . El-dosuky, and M. Aoud, "Image Restoration Based on Morphological Operations," *Int. J. Comput. Sci. Eng. Inf. Technol.*, vol. 4, no. 3, pp. 9–21, Jun. 2014, doi: 10.5121/ijcseit.2014.4302.

[16] W. Piao, Y. Yuan, and H. Lin, "A Digital Image Denoising Algorithm Based on Gaussian Filtering and Bilateral Filtering," *ITM Web Conf.*, vol. 17, p. 01006, Feb. 2018, doi: 10.1051/itmconf/20181701006.

[17] L. Cadena, A. Zotin, and F. Cadena, "Enhancement of medical image using spatial optimized filters and OpenMP technology," in *Lecture Notes in Engineering and Computer Science*, 2018, available at: Google Scholar.

[18] S.-Y. Chiu, C.-C. Chiu, and S. Xu, "A Background Subtraction Algorithm in Complex Environments Based on Category Entropy Analysis," *Appl. Sci.*, vol. 8, no. 6, p. 885, May 2018, doi: 10.3390/app8060885.

[19] I. G. S. Masdiyasa, I. D. G. Hari Wisana, I. K. Eddy Purnama, and M. Hery Purnomo, "Modified Background Subtraction Statistic Models for Improvement Detection and Counting of Active Spermatozoa Motility," *Lontar Komput. J. Ilm. Teknol. Inf.*, p. 28, May 2018, doi: 10.24843/LKJITI.2018.v09.i01.p04.

[20] A. A. Moulay and A. Amine, "A Novel Background Subtraction Algorithm for Person Tracking Based on K-NN," in *Computer Science & Information Technology (CS & IT)*, 2017, pp. 125–136, doi: 10.5121/csit.2017.70113.

[21] I. G. S. Masdiyasa, I. K. E. Purnama, and M. H. Purnomo, "A new method to improve movement tracking of human sperms," *IAENG Int. J. Comput. Sci.*, vol. 45, no. 4, pp. 1–9, 2018, available at: Google Scholar.

[22] H. B. Basoeki, A. D. Wibawa, and I. K. E. Purnama, "Improving sperms detection and counting using single Gaussian background subtraction," in *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, 2016, pp. 295–299, doi: 10.1109/ISEMANTIC.2016.7873854.

[23] R. Azzam, M. S. Kemouche, N. Aouf, and M. Richardson, "Efficient visual object detection with spatially global Gaussian mixture models and uncertainties," *J. Vis. Commun. Image Represent.*, vol. 36, pp. 90–106, Apr. 2016, doi: 10.1016/j.jvcir.2015.11.009.

[24] A. N. Rumaksari, S. Sumpeno, and A. D. Wibawa, "Background subtraction using spatial mixture of Gaussian model with dynamic shadow filtering," in *2017 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2017, pp. 296–301, doi: 10.1109/ISITIA.2017.8124098.

[25] N. Kim, M. Heo, R. Fleysher, C. A. Branch, and M. L. Lipton, "A Gaussian Mixture Model Approach for Estimating and Comparing the Shapes of Distributions of Neuroimaging Data: Diffusion-Measured Aging Effects in Brain White Matter," *Front. Public Heal.*, vol. 2, Apr. 2014, doi: 10.3389/fpubh.2014.00032.

[26] S. Qi, T. Nie, Q. Li, Z. He, D. Xu, and Q. Chen, "A Sperm Cell Tracking Recognition and Classification Method," in *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2019, pp. 163–167, doi: 10.1109/IWSSIP.2019.8787312.

[27] G. Gao, H. Jiang Inc., J. C. Vink, C. Chen, Y. El Khamra Inc., and J. Ita Inc., "Gaussian Mixture Model Fitting Method For Uncertainty Quantification By Conditioning To Production Data," 2018, doi: 10.3997/2214-4609.201802279.

[28] Y. He, J. Jiang, D. Dai, and K. Fabrice, "An Incremental Kernel Density Estimator for Data Stream Computation," *Complexity*, vol. 2020, pp. 1–17, Feb. 2020, doi: 10.1155/2020/1803525.

[29] T. Trnovszký, P. Sýkora, and R. Hudec, "Comparison of Background Subtraction Methods on Near Infra-Red Spectrum Video Sequences," *Procedia Eng.*, vol. 192, pp. 887–892, 2017, doi: 10.1016/j.proeng.2017.06.153.

[30] M. A. Soeleman *et al.*, "An Improvement for Background Modelling using a Mixture of Gaussian and Region Growing in Moving Objects Detection," *J. Phys. Conf. Ser.*, vol. 1430, p. 012032, Jan. 2020, doi: 10.1088/1742-6596/1430/1/012032.

[31] A. Ajmal and I. M. Hussain, "Vehicle detection using morphological image processing technique," in *2010 International Conference on Multimedia Computing and Information Technology (MCIT)*, 2010, pp. 65–68, doi: 10.1109/MCIT.2010.5444851.

[32] I. G. S. M. Diyasa, E. Y. Puspaningrum, M. Hatta, and A. Setiawan, "New Method For Classification Of Spermatozoa Morphology Abnormalities Based On Macroscopic Video Of Human Semen," in *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2019, pp. 133–140, doi: 10.1109/ISEMANTIC.2019.8884348.