International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

Semen Morphological and Motilty Image Analysis

Dr. V S Abbiramy[1], Dr. Ani Brown Mary N[2], Joel Sundar[3]
[1,2]*Asst Professor, Department of Computer Science, Sarah Tucker College*
[3]*Guidewire Developer at Zurich Insurance, Mexico*

***Abstract*— Most of the medical research laboratory accomplish a simple valuation of the percentage of normal shape semen in the specimen to evaluate the semen morphology. As trained technicians are needed to perform a "strict" semen analysis, CASA system-based image analysis is nowadays utilized for the valuation of semen concentration, morphology and motility features. The morphological image analysis of spermatozoa involves the following steps to be carried out in a sequence: Image preprocessing, Extraction and Detection of individual Spermatozoon, Image Segmentation and Feature Extraction. The pre-processing techniques comprise image smoothing and filtering, brightness and contrast enhancement. The individual spermatozoon from the pre-processed micrograph is then extracted. The extracted individual spermatozoon is then segmented into head, midpiece and nuclei using the Marker-Controlled Watershed Segmentation algorithm and Clustering Based Color Image Segmentation algorithm. Martin"s Global Consistency Error (GCE), William Rands Random Index (RI) and Variation of Information (VOI) were calculated for each individual region (head, nucleus and midpiece) for the segmented image. In watershed segmentation algorithm, the RI values were found to be higher and VOI, GCE values were lower for the segmented images. Thus, by using the segmentation metrics, it was clearly verified that watershed segmentation algorithm is well suited for segmenting the micrograph spermatozoa images. Dimensionality reduction techniques like Singular value decomposition (SVD), Principal Component Analysis (PCA) and Factor Analysis (FA) were applied. The performance of Singular value decomposition, Principal Component Analysis using the Covariance method and Correlation method and Factor Analysis were compared in terms of the number of mechanisms to be mined. It was detected that Covariance created PCA was initiated to be more effective in dropping the dimensionality of the feature space.**

***Keywords*—Spermatozoon, Segmentation, Semen analysis, Morphology, Oocyte, Acrosome, Spermatogenesis**

I. INTRODUCTION

The existing confidence is that semen morphological valuation must be utilized mostly as a productiveness device for analyzing the man's ability to conceive children. The measurement of the percentage of spermatozoa consuming an 'ideal' morphology by means of so-called severe technique.

It is the technique to mention in the newest publication of the World Health Organization (WHO) laboratory labour-intensive for semen investigation [1].

Males with a fault in semen development incline to have difficulties with semen geomorphology and might then be at jeopardy for disappointment of their semen to pollinate their spouse's eggs [2]. Also, geomorphology is an analyst of accomplishment in pollinating oocyte during in vitro pollination. Up to 10 per cent of all spermatozoa have visible faults and as such are deprived in terms of pollinating an oocyte [3].

Semen are tiny beings which look like tiny tadpoles swimming about at a desperate leap [4]. A strong humanoid semen is about 40 to 250 μm lengthy and is composed of neck, head, midpiece and tail. The semen skull, consists of an extended haploid nucleus, surrounded by a nuclear membrane and portion outside the nuclear membrane is called the acrosome followed by semen neck and then the semen midpiece. Finally, the semen has an elongated tail in order to boost the skull of the semen, that transfers all the DNA data, to the egg [11].

Evaluation of Semen morphological is done in two different ways: Furthermost medical laboratories accomplish a basic valuation of the proportion of semen in the uttered specimen that look to have regular shape. Nevertheless, an enlarged number of unusually shaped semen with 'non-strict' principles may specify a fertilization delinquent, but to be confident this test should possibly be recurrent in a laboratory that achieves firm criteria semen investigation [5]. Only dedicated andrology laboratories have skilled technicians who can investigate the semen affording to these "strict criteria". The semen is discoloured and observed under lubricant absorption (1000 x) for regular size and shape of the head, mid piece, tail or for any other irregularities. Here, a slight fault in any class rates the semen as irregular. Consequently, comparatively insufficient semen is evaluated as "normal" or perfect (near-perfect) through the firm morphology examination, as associated to the "estimated crude morphology" completed in a consistent semen examination [6].

Numerous dissimilar shapes or procedures of human semen have been recognized and categorized. These forms drop into one three main classes, namely normal forms, abnormal heads and abnormal tails [12].

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

II. Morphological Analysis Of Spermatozoon

The morphological image analysis of spermatozoa involves the following stages to be carried out in a sequence: Image preprocessing, Extraction and Detection of distinct Spermatozoon, Image Segmentation and Feature Extraction [16] as shown in the Figure 1.

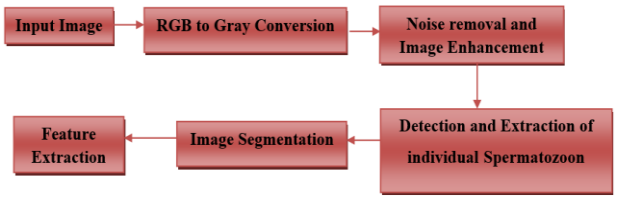

Figure 1 Architectural model of morphological image analysis

The first step is the image preprocessing phase which includes the conversion of RGB image to a gray scale image and then image noises are detached utilizing median filter. The second step is the recognition and abstraction of individual spermatozoon which contains the extraction of semen objects from images by means of Sobel edge detection process. The third step segments the spermatozoon into several region of interest, namely semen head, midpiece and tail. The fourth step comprises the removal of Morphological features, motility tracking features and feature dimensionality reduction techniques are applied [20]. Figure 2 shows the different stages in morphological image analysis. Each stage in morphological image processing is described in the Figure 1.

III. Image Preprocessing

An The impartial of the image pre-processing techniques is to alter pixel values in the image so that it is more appropriate for succeeding investigation than the creative image. The pre-processing techniques include image smoothing and filtering, brightness and contrast enhancement.

Micrograph images are frequently tainted by casual variations in intensity values named noises. The noise in micrograph is found to be Gaussian in nature. Gaussian noise is a set of values reserved from a zero mean Gaussian distribution which are additional to each pixel value. Gaussian filters are a class of linear smoothing filters with the weights selected affording to the shape of a Gaussian function. For image processing, the two-dimensional zero mean discrete Gaussian function as shown in Eq. (1).

$$g[i,j] = e^{\frac{-(i^2+j^2)}{2\sigma^2}} \qquad (1)$$

This Eq. (1) is used as a smoothing filter with σ=0.5 and kernel size of 3 × 3.

Unfortunately, image smoothing also hazes all sharp boundaries that accept significant data about the image. To enhance the edge contrast of an image, unsharp masking is performed. The knowledge of unsharp masking is to detract a mounted unsharp version of the image from the creative.

The spermatozoa micrograph images are poor in contrast and the intensity peak values of background and foreground are not widely separated. Power law transformation is utilized to improve the intensity of the individual pixels in the images. The power-law transformation is usually defined in Eq. (2).

$$s = cr^{\gamma} \qquad (2)$$

where s and r are the gray levels of the pixels in the yield and the input images, correspondingly and c is a positive constant. The power law transformation is applied with c=1 and γ=.5

IV. Extraction And Detection Of Individual Spermatozoon

The main purpose of this stage is to extract individual spermatozoon from the original RGB color image. The microscopic image is analysed and the presentation of the image is examined. If an image is in RGB method, it is converted to a gray scale image. The algorithm used to detect and extract individual spermatozoon is outlined in Listing 1 and the steps are discussed below:

TABLE I
Listing 1 Algorithm to detect and extract individual spermatozoon

Step 1: Edges of semen objects are extracted by applying Sobel edge detection method.
Step 2: Removal of artifacts present in the microscopic image.
Step 3: Image smoothening to decrease the number of linked components by applying convolution.
Step 4: Calculation of number of linked components in a binary image.
Step 5: Extraction of any individual spermatozoon from an original image.

Step 1: Edges of semen objects are extracted by applying Sobel edge detection technique. Edge detection is the greatest mutual method for detecting expressive cutoffs in intensity values. Such discontinuities are perceived by using first and second-order derivatives.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

First-order derivatives in an image are computed using the gradient. Second-order derivatives are obtained using the Laplacian. The gradient of an image f(x,y) at location (x,y) is well-defined as the vector in Eq. (3).

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (3)$$

Sobel edge detector usages the mask to rough numerally the first derivatives Gx and Gy. The gradient at the center point in an area is calculated as follows by the Sobel detector is shown in Eq. (4) and (5).

$$g = \left[G_x^2 + G_y^2\right]^{1/2} \quad (4)$$

$$g = \{[(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)]^2 + [(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)]^2\}^{1/2} \quad (5)$$

A pixel at position (x, y) is an edge pixel if its two-dimensional first order derivative is better than a quantified threshold ie $(g \geq T)$. A set of such points that are linked to a predefined criterion of connectedness is an edge. The basic steps followed in edge detection using Sobel algorithm are

1. Read the image.
2. Convert the image to double.
3. Use the mask $F1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$ for x direction and $F2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ for y direction and obtain the gradient of the image.
4. Find the magnitude of the vector.
5. The border pixels are not considered and edge detection process starts from the pixel (2, 2).
6. Threshold the image.
7. Display the logical image.

The result of edge detection using Sobel operator is explained. It is found that under noisy conditions, Sobel operator exhibit better performance.

Removal of artifacts present in the microscopic image. An initial microscopic examination of the sample reveals the existence of cells other than spermatozoa, e.g. epithelial cells, and isolated semen heads or tails. All connected components that have fewer pixels can be removed from the binary image by applying the size filter which are not useful. Every area under T pixels in size are detached by altering the value of their pixels to zero. However, it is very hard to discover a correct value for T. If T is small, little noise will persist. If T is large, important data will be gone.
Step 3: The image is then smoothened to decrease the number of linked components utilizing convolution. A convolution is performed by multiplying a pixel and its adjacent pixels value by a matrix named kernel. A kernel is a slight matrix of numbers that is applied in image convolutions. The size of a kernel is arbitrary but 3 × 3 is frequently utilized. The convolution formula is provided by Eq. (6).

$$V = \left| \frac{\sum_{i=1}^{q} \left| \sum_{j=1}^{q} f_{ij} d_{ij} \right|}{F} \right| \quad (6)$$

where:

f_ij = the factor of a convolution kernel at position i,j

d_ij = the data information of the pixel that corresponds to f_ij

q = the dimension of the kernel is assigned a square kernel (if q=5, the kernel is 5×5)

F = moreover the sum of the coefficients of the kernel or 1 if the sum of the coefficients is 0

V = the result pixel information

Smoothening the image before finding the number of connected components would improve the extraction of individual spermatozoon.

The number of combined parts in a binary image is estimated. Let, S denotes a subsection of pixels in an image, two pixels p and q are called to be combined in S if there occurs a path among them containing completely of pixels in S. For at all pixel p in S, the set of pixels that are associated to it in S is named a associated part of S. The associated part method forms two passes over the image: the first pass is to allocate temporary labels and record equivalences and the second pass is to substitute each temporary label by the lowest label of its equivalence class.

Pass 1: Scan the image pixels starting from left to right and from top to bottom. For all pixel P of value 1 (an object pixel), test top and left neighbors (4-neighbor metric).

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

- If 2 of the neighbors equal 0: Allocate a novel mark to P.
- If 1 of the neighbors equals 1: Allocate the neighbor's mark to P.
- If 2 of the neighbors equal 1: Allocate the left neighbor's mark to P.

Pass 2: Divide all marks in to equivalence classes. Exchange each mark with the number of its equivalence class. Thus, it is possible to assign different labels to pixels belonging to different regions.

Step 5: Any individual spermatozoon can be extracted from the original RGB color image. Once the number of linked components is known, any connected components can be extracted utilizing the label. The extracted spermatozoon is warehoused in an array list.

V. Image Segmentation

Fereidoon et al have proposed a multi-step method for semen segmentation in microscopic images. Segmentation of spermatozoa was done using the morphological operations. Guihai et al have proposed an image segmentation technique for spermatozoa that combines canny operator with two-dimensional extreme entropy.

The important determination of the segmentation phase is to subdivide a spermatozoon into numerous constituent shares, namely head, mid-piece and tail. Segmentation is performed by applying the Marker-Controlled Watershed Segmentation algorithm and Clustering Based Color Image Segmentation algorithm. The input for these segmentation algorithms were the images of individual spermatozoa given by the detection and extraction method.

5.1 Marker-Controlled Watershed Segmentation

The Marker-Controlled Watershed Segmentation algorithm used to segment spermatozoa is outlined below:

Step 1: Read the micrograph spermatozoon image as shown in the Figure 2(a) and convert it to grayscale image. The converted gray scale image is shown in Figure 2(b)

Step 2: Gradient magnitude is used as the segmentation function. The gradient magnitude is computed using the Sobel edge masks. It is found to be that the gradient is high at the borders of the objects and low (frequently) confidential the objects. The Gradient magnitude image is shown in the Figure 2(c).

Step 3: Internal markers that are inside respectively of the objects of interest is computed. Morphological techniques like "opening-by-reconstruction" and "closing-by-reconstruction" were used to find the connected blobs of pixels that are private of the foreground objects.

Step 4: External markers that are confined inside the background is computed. These are pixels that are not portion of somewhat object. The watershed ridge lines image is obtained by calculating the watershed transform of the distance transform of the image acquired in the previous step and the resulting image is shown in the Figure 2(d).

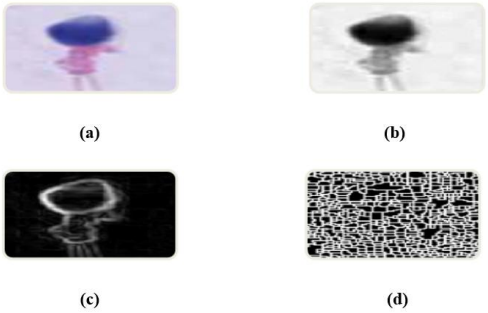

Figure 2 Marker-Controlled Watershed Segmentation: (a) Spermatozoon image (b) Gray scale image (c) Gradient magnitude image (d) Ridge lines image

Step 5: Internal and External markers is used to adjust the gradient image by a technique named minima imposition. It alters a gray-scale image so that regional minima happen only in marked positions.

Step 6: Finally, the watershed transform of the marker-modified gradient image is estimated.

VI. Clustering Based Color Image Segmentation

Clustering techniques categorizes the pixels with similar features into one cluster, thus making various clusters rendering to coherence among pixels in a cluster. A clustering method such as K-means clustering is used to form clusters [19].

The basic objective is to segment colors in an automated fashion using the K-means clustering algorithm. K-means is applied to cluster the objects into three clusters using the Euclidean distance metric. The entire process can be summarized in following steps:

Step 1: First, an image is engaged as an input.

Step 2: The input image which is in the form of RGB color space is transformed to a Lab color space. A Lab color space is a color-opponent space with dimension L for lightness and a and b for the color-opponent dimensions. The variance among two colors is restrained by applying the Euclidean distance metric.

Step 3: Following similar data points, i.e. the points which have like color, are combined together using K-means clustering technique.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

The data points with smallest Euclidean distance are gathered together to form the clusters.

Step 4: Then label all pixel in the image by means of the outcomes from K-means. Meant for all object in the input image, K-means provides an index matching to a cluster. Use this cluster index to label every pixel in the image.

Step 5: Segmented images are created by segmenting the original image by color. The objects in original image can be parted by color and this would outcome in segmented images [15].

VII. Performance Evaluation Of Segmentation Method

The existing evaluation methods for image segmentation can be classified as ground-truth based and non-ground-truth based [7, 8]. In non-ground-truth based techniques, the experiential golly procedures are projected to encounter the heuristic criteria in the necessary segmentations. Then the score is estimated based on these criteria to expect the quality of segmentation.

The ground-truth based techniques measure the change between the segmentation result and the human labeled ground truths [14]. The recognized Rand Index and its variants are of this kind. Three various parameters are utilized to estimate the performance of image segmentation approaches. The parameters applied for evaluation are

i. Global Consistency Error (GCE)
ii. Random Index (RI)
iii. Variation of Information (VOI)

The data obtained in the above methods were exposed to standard statistical analysis viz., Test of significance to compare the efficiency of color-based segmentation and watershed segmentation algorithms [13].Captions with table numbers must be placed before their associated tables, as shown in Table 1.

i) Global Consistency Error (GCE)

David (2002) has proposed several error procedures to quantify the constancy between image segmentations of divergent granularities and applied them to compare the results of normalized-cut methods to a database of physically segmented images.

Let S and S' be two segmentations of an image X={x_1,x_2,…,x_N} containing of N pixels. For a given pixel xi, contemplate the segments that contain xi in S and S'. These sets of pixels are denoted as C(S,x_i) and C(S^',x_i) respectively. Global Consistency Error (GCE) is defined in Eq. (7).

$$GCE(S,S') = \frac{1}{N}\ min\{\sum_i LRE(S,S',x_i), \sum_i LRE(S',S,x_i)\} \quad (7)$$

For GCE, a value of 0 designates no error and a value of 1 designates maximum deviation among the two segmentations being associated.

ii) Random Index (RI)

William (1971) has projected a similarity method that transformed the problem of comparing two partitions with probably divergent number of classes into a problem of computing pair wise label relationships.

Consider two valid label assignments S and S' with conforming labels {l_i } and {l_i^' } of N points X={x_1,x_2,…,x_N}. The rand index R can be calculated as the ratio of the number of pairs of points having the well-matched label association in S and S'.

$$R(S,S') = \frac{1}{\binom{N}{2}}\sum_{i,j,i\neq j}[I(l_i = l_j \wedge l_i' = l_j') + I(l_i \neq l_j \wedge l_i' \neq l_j')] \quad (8)$$

where I represents the identity function and the denominator is the number of likely unique pairs among N data points in Eq. (8).

This provides a measure of similarity with value ranging from 0 when the two segmentations have not one likeness to 1 when the segmentations are equal.

iii) Variation of Information (VOI)

It estimates the sum of data loss and information gain among the two clustering. The VOI metric is nonnegative, with lesser values representing greater similarity. It is based on relationship between a point and its cluster. It practices mutual data metric and entropy to rough the distance between two clustering across the lattice of possible clustering. More precisely, it measures the amount of information that is lost or gained in changing among one clustering to another.

The variation of data is a measure of the distance between two clusters. Suppose we have two clusters X and Y where X={X_1,X_2…X_k}, p_i=|X_i |/n, n=∑_k▒|X_i |. Then the variation of information between two clusters is:

$$VI(X;Y) = H(X) + H(Y) - 2I(X,Y) \quad (9)$$

where, H(X) is entropy of X and I (X,Y) is mutual information between X and Y.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

iv) Feature Extraction

Subsequent to image segmentation, different parts of the spermatozoa are identified. In additional analyze these parts, it is essential to make a numerical representation frequently in the form of a vector. Each element of the vector defines a various feature. A vector representation provides the application of classification algorithms and the extraction of important statistical data [17].

Standard statistical procedures to describe the features of the image were used in [1]. Statistical techniques examine the spatial distribution of gray values, by estimating local features at each point in the image, and originating a set of statistics from the distributions of the local features. Depending on the number of pixels defining the local feature statistical methods can be additional classified into first-order (one pixel), second-order (two pixels) and higher-order (three or more pixels) statistics.

Morphological features and the semen motility grades of individual spermatozoon are extracted as in [9].

v) Morphological Features

Morphological feature extraction is executed over the image extracted from the preceding phase. These features are applied to classify the Spermatozoa either as normal or abnormal. The following parameters are measured:

Head Area: It is the number of pixels controlled in the segmented head area.

Perimeter: It is the number of pixels in the boundary of the Spermatozoa.

Head Length: Head length is estimated with the major axis in pixels. Major axis is well-defined as a line that comprises the center of mass point and has a slope equal to the line well-defined by the highest value pixel in the Euclidean distance transform.

Head width: Head width is estimated with the minor axis in pixels. Minor axis is well-defined as the line perpendicular to the major axis.

Eccentricity: The ratio of the head length to the head width is termed as the eccentricity.

Head Orientation: It is the angle between the major axis of the head and the major axis of the midpiece.

Midpiece length: It is the major axis of the midpiece and it is measured in pixels.

Mid-piece width: It is the minor axis of the midpiece and it is measured in pixels.

Midpiece Orientation: It is the angle between the neck and tail to the major axis of the head in degrees.

Nucleus area: In this parameter the amount of area occupied by the nucleus is computed.

vi) Semen Tracking Features

The motility of the spermatozoa was established using videos in avi format with a minimum of 50 frames. Each frame has a resolution of 240 ×320 and its frame rate is 25fps. For each image, a fixed region of 220 × 300 pixels comprising the goal was composed and packed.

The semen's present in each frame were segmented and each of the frames was processed to determine the number of spermatozoa. The extracted frames from the video were processed sequentially to trail the motion of individual semen's from frame to frame. By conniving the distance trekked by each semen over a large dataset of frames for a certain period of time, the motility associated with each semen were determined.

The automated semen movement measurements also known as kinematics were determined for individual semen's in the video sequence by analyzing the semen tracks. These tracks Analysis provides the objective information on average speed, straight line velocity and mean speed of progressive and non-progressive spermatozoa.

The straightness of the swimming course for an individual spermatozoon was accessed unswervingly from the region by computing Total Distance (TD) of a track. Spermatozoa swimming along straight trajectories have a velocity v > 20. Spermatozoa with a velocity of less than 20 but greater than 4 may be swimming in a circle or their progression may be unpredictable. Spermatozoa with a velocity of less than 4 but greater than 0.4 move their tail but fail to progress forward. Spermatozoa with a velocity of less than 0.4 are immotile semen's that lack movement.

vii) Dimensionality Reduction Methods

Feature extraction converts the information in the high-dimensional space to a intergalactic of fewer dimensions. Feature extraction technique provides feature vectors by projecting parameter vectors onto a feature space through a linear transformation T_(p×m),p≥m:

$$y=T^{T} x \qquad (10)$$

where y is feature vector and x is parameter vector. Feature extraction technique nevertheless shows significant returns over feature selection technique [10, 18].

Principal Component Analysis (PCA), Singular Value Decomposition (SVD) and Factor Analysis (FA) are used as feature dimensionality reduction techniques. The presentation of PCA SVD and FA in feature dimensionality reduction is also examined for the extracted features.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

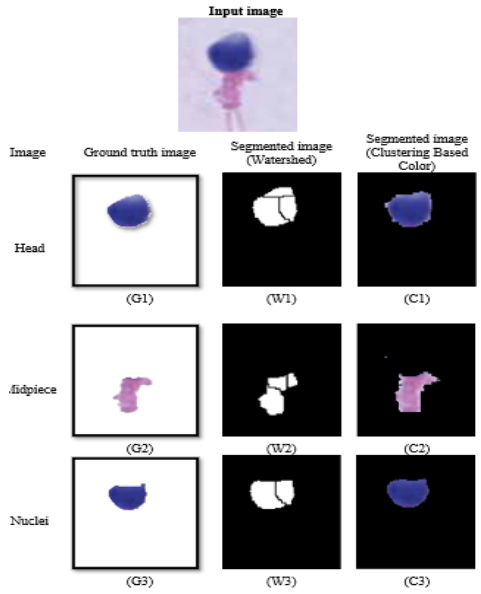

Figure 3 Ground truth, Watershed Segmentation and Clustering Based Color Image Segmentation of three image categories.

VIII. Results And Discussion

For the sample input image shown in Figure 3 (G1), (G2) and (G3) serves as ground truth images for head, midpiece and nuclei. The three distinct region of the spermatozoon, namely the head, midpiece and nuclei are segmented using marker controlled watershed segmentation and the results obtained are shown in Figure 3 (W1), (W2) and (W3). The segmented head, midpiece and nuclei region of the spermatozoon obtained using clustering-based color image segmentation is shown in Figure 3 (C1), (C2) and (C3).

Martin's Global Consistency Error (GCE), William Rand's Random Index (RI) and Variation of Information (VOI) are calculated for each individual region (head, nucleus and midpiece) segmented using marker-controlled watershed segmentation and clustering-based color image segmentation methods. The sample performance measures for both these segmentation methods on the segmented head, midpiece and nuclei are presented in the Table 1.

Evaluation of marker-controlled watershed segmentation and clustering-based color image segmentation according to the three metrics discussed is summarized in Table 2.

Applying the segmentation metrics, it can be visibly understood that watershed segmentation technique is well suitable for the micrograph spermatozoa images. The RI value should be developed for an image and VOI, GCE values must be lesser for an image.

Table 3 shows the values of morphological parameters for randomly selected spermatozoon images which is segmented using the marker-controlled watershed segmentation algorithm. The input data set as shown in the Table 4 is formed by linking the morphological parameters of spermatozoon and motility statistics. This data set has fifteen features, so every sample is a fifteen-dimensional vector.

The Eigen values, proportion and cumulative proportion of variance described by the principal components for the SVD, Co variation based PCA, Correlation based PCA and Factor Analysis are shown in the Table 5. Depending on the correlation matrix, eigen values are estimated.

In the case of N independent variables, there are N eigen values. For the specified data set, there are 15 eigen values. The variance column tells how much of the information from the original variables can be explained by each factor or component. The amount of total variance enlightened by the ith principal component is purely the ratio between the ith eigen value and the summation of all eigen values. The cumulative proportion column tells how much of the information from the original variables is explained by all the extracted factors. The cumulative proportion of variance is estimated by totaling the recent and the earlier proportion of variance.

Table 1 Performance measures for the segmentation outcomes

Segmented Image	Watershed Segmentation			Color Based Segmentation		
	RI	GCE	VOI	RI	GCE	VOI
Image1 Head	0.9480	0.0180	0.8045	0.6969	0.1390	1.8955
Image1 Midpiece	0.9418	0.0141	0.6432	0.9518	0.0507	0.4728
Image1 Nuclei	0.9582	0.0230	0.6005	0.9705	0.0363	0.3430
Image 2 Head	0.9286	0.0400	0.9434	0.9108	0.0895	0.8774
Image 2 Midpiece	0.9636	0.0116	0.3575	0.8040	0.0659	1.3824
Image 2 Nuclei	0.9614	0.0158	0.5063	0.9672	0.0270	0.3106

IJRDET

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

Table 2 Summarized results of Segmentation for quality metrics

Segmented Image (N=5)	Color Based Segmentation			Watershed Segmentation		
	RI	GCE	VOI	RI	GCE	VOI
Head*	0.890 [a] ± 0.049 *(12.25)*	0.074 [a] ± 0.019 *(56.16)*	1.067 [a] ± 0.10 *(20.94)*	0.924 [a] ± 0.008 *(1.92)*	0.032 [b] ± 0.004 *(27.09)*	0.874 [a] ± 0.264 *(67.62)*
Mid piece**	0.936 [a] ± 0.012 *(2.98)*	0.043 [a] ± 0.003 *(16.30)*	0.403 [a] ± 0.060 *(33.34)*	0.956 [a] ± 0.005 *(1.12)*	0.014 [b] ± 0.004 *(69.39)*	0.537 [a] ± 0.092 *(38.28)*
Nuclei	0.954 [a] ± 0.008 *(1.965)*	0.043 [a] ± 0.010 *(51.38)*	0.665 [a] ± 0.062 *(21.01)*	0.940 [a] ± 0.011 *(2.64)*	0.035 [a] ± 0.010 *(64.52)*	0.454 [b] ± 0.064 *(31.65)*

*Values bearing different superscript in a parameter differ significantly statistically- * (P<0.05), * (P<0.01) levels. Values in the parentheses indicates the Co efficient of Variation (CV) expressed in pe cent*

Table 3 Morphological Parameters of the human spermatozoon images

Images (jpg)	Head Properties						Midpiece Properties			Nucleus Area
	Area	Perimeter	Head Length	Head Width	Eccentricity	Orientation	Length	Width	Orien Tation	
Abnormal-1	797	117.36	34.33	30.30	0.47	18.70	1.15	1.15	0.00	580
Abnormal-2	575	109.46	36.13	20.88	0.82	-87.08	97.49	31.11	-70.33	409
Abnormal-3	619	129.15	42.17	20.09	0.88	-85.79	10.35	3.51	5.17	325
Abnormal-4.	936	127.01	36.46	33.59	0.39	-4.79	16.40	4.64	-85.15	467
Abnormal-5.	627	131.50	39.79	22.18	0.83	85.65	62.55	38.15	-88.49	471
Normal-1	671	114.33	34.66	25.14	0.69	-87.82	2.31	1.15	0.00	405
Normal-2	541	193.34	31.86	30.59	0.28	-78.38	91.24	47.34	74.84	134
Normal-3	702	113.60	33.08	28.00	0.53	88.99	3.46	1.15	90.00	538
Normal-4	446	102.87	30.20	25.89	0.52	-80.48	58.87	39.28	86.19	267
Normal-5	575	108.77	30.94	25.50	0.57	57.68	3.46	1.15	0.00	370

A line graph is designed with the values chosen from the Table 5 and it looks like what is publicized in the Figure 4. The x-axis or the horizontal axis depicts the principal component and the y-axis or the vertical axis depicts the cumulative proportion of variance

Table 4 Sample input dataset

Area	Perimeter	HeadLen	Head Width	Eccentricity	Mid Len	Tail Len	Orientation	Equiv Diam	Mean Dist	Mean Velocity	A	B	C	D
36.00	46.70	9.90	5.32	0.84	15.62	67.57	78.55	6.77	147.75	49.25	1.0	0.0	0.0	0.0
158.0	191.23	23.62	8.66	0.93	19.11	66.23	81.96	13.91	121.50	40.50	1.0	0.0	0.0	0.0
1.00	3.63	1.15	1.15	0.96	10.23	44.69	0.00	1.13	0.00	0.00	0.0	0.0	0.0	1.0
37.00	25.45	10.13	5.34	0.85	19.12	35.54	88.93	6.87	17.07	5.00	0.0	1.0	0.0	0.0
10.00	11.85	4.26	3.21	0.66	14.00	74.31	90.00	3.57	66.63	22.00	1.0	0.0	0.0	0.0
1.00	3.63	1.15	1.15	0.00	15.07	44.16	0.00	1.13	20.01	6.67	0.0	1.0	0.0	0.0
7.00	13.45	5.77	1.81	0.95	15.35	38.77	45.00	2.99	17.07	5.70	0.0	1.0	0.0	0.0
7.00	10.17	3.88	2.43	0.78	15.35	42.27	45.00	2.99	66.63	22.21	1.0	0.0	0.0	0.0
7.68	53.93	18.75	5.06	0.95	28.13	68.76	-89.91	81.05	53.81	17.94	0.0	1.0	0.0	0.0
4.57	35.91	1.66	1.07	0.96	2.49	45.56	89.76	87.23	58.47	19.49	0.0	1.0	0.0	0.0
6.41	48.76	8.46	5.26	0.96	12.70	32.45	-89.81	84.26	36.67	12.22	0.0	1.0	0.0	0.0
5.67	42.71	9.57	5.91	0.96	14.35	65.47	-89.96	85.36	30.91	10.30	0.0	1.0	0.0	0.0
7.19	52.16	2.47	1.03	0.96	3.70	45.26	89.87	84.45	7.68	2.56	0.0	0.0	1.0	0.0
8.16	63.25	8.86	5.61	0.96	13.29	35.89	-89.99	82.86	110.00	36.67	1.0	0.0	0.0	0.0
6.59	49.90	9.93	5.82	0.96	14.89	45.34	-89.69	83.58	46.07	15.35	0.0	1.0	0.0	0.0
3.97	7.69	7.66	4.12	0.96	11.49	30.98	89.38	89.86	11.96	3.99	0.0	0.0	1.0	0.0

Table 5 Eigen values and the proportion of variation explained by the principal components

Component	SVD			PCA						Factor Analysis		
				Covariation based PCA			Correlation based PCA					
	Eigen Values	Prop.	Cum. Prop.	Eigen Values	Prop.	Cum. Prop.	Eigen Values	Prop.	Cum. Prop.	Eigen Values	Prop.	Cum. Prop.
0	5.55	0.03	0.03	6.94	0.50	0.50	5.55	0.37	0.37	5.55	0.37	0.37
1	2.65	0.04	0.07	4.26	0.31	0.81	2.65	0.18	0.55	2.65	0.18	0.55
2	2.07	0.05	0.12	1.42	0.10	0.91	2.07	0.14	0.68	2.07	0.14	0.68
3	1.28	0.06	0.18	1.01	0.07	0.98	1.28	0.09	0.77	1.28	0.09	0.77
4	1.20	0.07	0.25	0.15	0.01	1.00	1.20	0.08	0.85	1.20	0.08	0.85
5	0.73	0.07	0.32	0.05	0.00	1.00	0.73	0.05	0.90	0.73	0.05	0.90
6	0.66	0.07	0.39	0.01	0.00	1.00	0.66	0.04	0.94	0.66	0.04	0.94
7	0.50	0.07	0.46	0.00	0.00	1.00	0.50	0.03	0.98	0.50	0.03	0.98
8	0.17	0.08	0.54	0.00	0.00	1.00	0.17	0.01	0.99	0.17	0.01	0.99
9	0.14	0.08	0.62	0.00	0.00	1.00	0.14	0.01	1.00	0.14	0.01	1.00
10	0.05	0.08	0.69	0.00	0.00	1.00	0.05	0.00	1.00	0.05	0.00	1.00
11	0.00	0.08	0.77	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
12	0.00	0.08	0.85	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
13	0.00	0.08	0.92	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
14	0.00	0.08	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00

Prop. – Proportion , Cum. Prop. – Cummulative Proportion

Table 6 represents the index of the module and its correspondent proportion of variance influence in terms of percentage for the three methods.

Table 6 Principal component base

Index	SVD	PCA		Factor Analysis
	Proportion of variance	Proportion of variance (Covariance)	Proportion of variance (Correlation)	Proportion of variance
	Per cent			
0	2.840	50.14	36.981	37
1	4.198	30.77	17.669	18
2	5.255	10.25	13.770	14
3	5.913	7.28	8.559	9
4	6.528	1.11	8.019	8
5	6.900	0.34	4.835	5
6	7.237	0.1	4.389	4
7	7.494	0.01	3.347	3
8	7.580	0	1.122	1
9	7.654	0	0.963	1
10	7.679	0	0.319	0
11	7.680	0	0.023	0
12	7.681	0	0.005	0
13	7.681	0	0.000	0
14	7.681	0	0.000	0

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

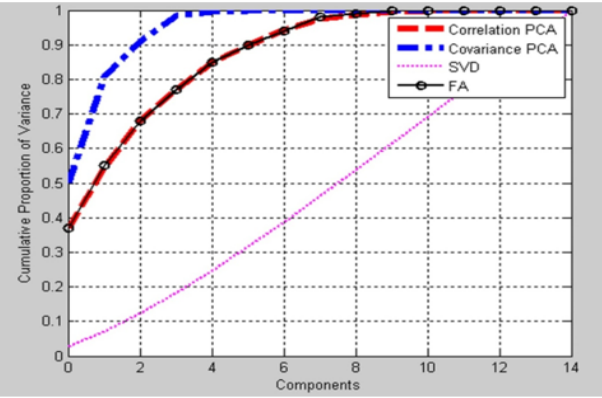

Figure 4 Imagining of component distribution for SVD, Covariation based PCA, Correlation based PCA and Factor Analysis

A previous measure has been established i.e. g[m=L]≥90% to choose the number of principal components that will elucidate the maximal amount of variance. In the case of SVD, the first 14 components explain 92.3% of proportion of variance and their extracted components are shown in the Table 7. In the case of Covariation constructed PCA, the first three components are chosen since it subsidizes to 91.16% of proportion of variance and their extraction is also shown in the Table 8. In the case of Correlation based PCA, the first six components are chosen since it donates to 89.83% of proportion of variance and their extraction is also shown in the Table 8.

In the case of Factor Analysis, the first six factors together accounted for the 91% of the total variance and the extracted factors are shown in Table 9. Principal axis factoring is used as the method of extraction where linear combinations of the observed variables are formed. Oblique rotation is used as the rotation method which allows the resulting factors to be correlated.

Table 9 depicts the assessment on the number of principal components obtained depending on different principles like

- *Kaiser Criterion:* The eigen value is associated to 1 and only workings with eigen values higher than 1 are engaged.
- *Percentage of variance criterion:* The number of workings to be calculated is resulted by the total explicated percentage of variance.
- *Scree test:* The eigen values are strategized in contradiction of the principal components in a simple line plot as shown in Fig 5.1. It suggests finding the place where the smooth decrease of eigen values appears to level off to the right of the plot.

Table 7 Projection of principal components for SVD

SVD												
Components												
1	2	3	4	5	6	7	8	9	10	11	12	13
-0.34	-0.06	-0.10	0.43	-0.16	0.29	0.06	-0.16	-0.07	0.02	-0.06	0.58	-0.46
-0.34	0.11	-0.22	0.21	-0.05	0.27	0.18	-0.36	0.04	0.39	0.10	-0.61	0.04
-0.35	0.28	-0.06	0.25	-0.06	-0.15	-0.08	0.05	0.24	0.04	-0.16	0.30	0.72
-0.34	0.27	-0.08	0.06	0.02	-0.02	-0.24	-0.01	-0.51	-0.65	0.06	-0.23	-0.01
-0.07	0.20	0.44	0.07	0.43	0.04	0.09	0.68	-0.10	0.22	0.19	0.03	-0.12
-0.22	0.23	0.36	0.21	-0.06	-0.43	-0.37	0.20	0.29	0.17	-0.14	-0.23	-0.42
-0.25	-0.07	0.17	0.01	0.03	-0.55	0.74	-0.03	-0.20	-0.02	0.07	0.03	-0.04
0.00	-0.45	-0.18	0.27	-0.38	0.08	0.17	0.44	0.13	-0.26	-0.39	-0.29	0.05
0.06	0.39	-0.41	-0.35	0.03	-0.10	0.16	-0.19	0.14	-0.08	-0.64	0.03	-0.21
-0.37	-0.14	0.00	-0.34	0.07	0.11	0.03	0.04	0.37	-0.19	0.17	0.02	-0.06
-0.37	-0.14	0.00	-0.34	0.07	0.11	0.03	0.04	0.37	-0.19	0.17	0.02	-0.06
-0.30	-0.35	0.02	-0.26	0.10	-0.05	-0.23	0.02	-0.35	0.31	-0.30	0.03	0.07
0.12	0.44	0.30	-0.02	-0.25	0.32	0.28	0.19	0.09	-0.09	0.09	-0.01	-0.02
0.15	-0.08	-0.53	0.10	-0.27	-0.42	-0.15	-0.17	0.17	-0.09	0.40	0.03	-0.07
0.13	-0.15	0.07	0.41	0.69	0.02	0.06	-0.21	0.25	-0.28	-0.14	-0.08	0.00

Table 8 Estimation of principal components for Covariation and correlation based PCA

Covariance Based PCA			Correlation Based PCA					
Components			Components					
1	2	3	1	2	3	4	5	6
-0.22	-0.47	-0.06	3.21	-2.02	-0.34	1.09	-0.22	0.32
-0.07	-0.64	-0.32	6.31	-0.42	0.99	-2.08	0.52	-0.78
0.00	-0.08	-0.02	-2.75	-1.45	-0.57	-1.96	-3.12	-0.04
0.00	0.03	0.00	-0.41	0.62	-0.79	-1.52	1.04	-0.60
0.00	0.00	0.00	0.72	-2.37	-1.04	0.58	0.22	1.15
0.01	-0.03	0.07	-2.20	-0.74	-2.87	0.13	1.80	-0.20
-0.02	-0.10	0.09	-1.59	0.16	-0.86	-0.65	0.45	-0.57
-0.92	0.21	-0.26	0.08	-1.93	-0.68	0.75	-0.14	0.05
0.28	-0.04	-0.83	1.45	2.84	-1.06	-0.24	-0.18	1.72
-0.10	-0.52	0.34	-1.57	-0.31	0.85	1.10	0.28	-1.53
-0.03	-0.17	0.11	-0.65	2.11	0.22	0.53	-0.33	-0.92
0.00	0.00	0.01	-0.01	2.18	-0.25	0.34	-0.37	0.58
0.00	0.00	0.00	-2.41	-0.95	2.98	-0.21	0.80	0.56
0.00	0.00	-0.01	1.74	0.13	0.77	2.03	-1.24	-0.47
0.00	0.00	0.00	-0.01	2.20	-0.03	0.51	-0.35	-0.28

Table 9 Evaluation on number of principal components obtained

Methods	Number of Dimensions in Input Data set	Number of Components Extracted		
		Kaiser criterion	*Percentage of variance>90%*	*Scree Test*
SVD	15	5	14	14
Covariance Based PCA	15	4	3	4
Correlation Based PCA	15	5	6	8
Factor Analysis	15	5	6	8

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 05, May 2026)

Table 9 Spermatozoa Features versus Rotated Component Matrix

Spermatozoa Features	Rotated Component Matrix					
	1	2	3	4	5	6
Area	0.80	-0.10	0.15	0.50	-0.18	-0.27
Perimeter	0.79	0.16	0.32	0.23	0.00	-0.26
Head Length	0.83	0.47	0.11	0.28	-0.03	0.15
Head Width	0.78	0.41	0.13	0.06	0.06	0.08
Eccentricity	0.14	0.23	0.48	0.06	0.34	-0.03
Midpiece Length	0.53	0.41	-0.50	0.21	-0.09	0.49
Tail Length	0.51	-0.07	-0.17	0.00	0.01	0.05
Orientation	-0.01	-0.66	0.19	0.25	-0.41	-0.08
Equiv Diameter	-0.14	0.60	0.65	-0.37	0.16	-0.01
Mean Distance	0.87	-0.24	-0.01	-0.38	0.08	-0.07
Mean Velocity	0.87	-0.24	-0.01	-0.38	0.08	-0.07
A	0.71	-0.56	-0.06	-0.27	0.07	0.13
B	-0.28	0.76	-0.38	-0.07	-0.28	-0.33
C	-0.36	-0.15	0.78	0.14	-0.24	0.33
D	-0.31	-0.25	-0.17	0.47	0.74	-0.05

IX. CONCLUSION

A significant constraint is evaluated in the semen investigation is the complete morphology. This is obtained with the help of the semen and its motility. The calculation of semen morphology for moreover normal spermatozoa or for semen defects is comparatively widespread and it is also very much important. The various steps that is very much applicable in the investigation of human spermatozoon morphology were discoursed. An uncharacteristic semen will not be intelligent to pollinate the egg. With the help of morphologic valuation deliberated in this work might be supportive to detect the pregnancy results in couples. This research will be very helpful for couples who are undergoing Intra Uterine Insemination (IUI) for In Vitro Fertilization (IVF).

REFERENCES

[1] World Health Organization, (2021), WHO laboratory manual for the examination and processing of human semen, Sixth edition,

[2] Pacific Fertility Center (2014), San Francisco, accessed 18 November 2014, http://www.pacificfertilitycenter.com/treatment-care/male-infertility-treatment.

[3] Sadler, T.W., (2012) Langman's medical embryology, Wolters Kluwer Health, Lippincott William & Wilkins, 12th Edition, Philadelphia.

[4] Malpani Infertility Clinic (2014), Semen Analysis - Everything about sperm & infertility, Mumbai, accessed 18 November 2014, http://www.drmalpani.com/articles/sperm.

[5] Andrology Institute of America (2009), ZDL, Inc. & Dr. Panayiotis Zavos, accessed 18 November 2014, http://www.aia-zavos.com/ strict_sperm_ morphology. html.

[6] Laurence, A. and Jacobs (2014), Screening for Male Infertility, Einstein Medical, accessed 18 November 2014, http://www.infertilitydoc.net/infertility-other-treatments/male-infertility.

[7] Fereidoon, N.R., Mohammad, H.M. and Vahid, R.N. (2005) "A Multi Steps Algorithm for Sperm Segmentation in Microscopic Image" in Proceedings of World Academy of Science, Engineering and Technology, Vol. 7, pp. 419-421.

[8] Guihai, L., Wenming, H. and Song, L. (2010) "2-D Maximum Entropy Spermatozoa Image Segmentation Based on Canny Operator" in International Conference on Intelligent Computing and Integrated Systems (ICISS), pp. 243 – 246.

[9] Abbiramy V.S. and Shanthi V., "Spermatozoa Detection, Counting and Tracking in Video Streams to Detect Asthenozoospermia", IEEE Sponsored International Conference on Signal and Image Processing (ICSIP), RMD Engineering College, Chennai, 15-17th December 2010.

[10] Paliwal, K.K. (1992) "Dimensionality Reduction of the Enhanced Feature Set for the HMM-Based Speech Recognizer", Digital Signal Processing, Vol. 2, No. 3, pp. 157-173.

[11] Richard, E.J. (1991) Human Reproductive Biology, United Kingdom Edition, Academic Press Limited, (1991), pp. 171-172.

[12] Brooks, A.K. and Bobby, W.W. (1990) Handbook of the Laboratory Diagnosis and Treatment of Infertility, CRC Press, Florida.

[13] George, W.S. and William, G.C. (1967) Statistical Methods, Iowa State University Press, 6th Edition.

[14] Spermatozoa Segmentation and Morphological Parameter Analysis Based Detection of Teratozoospermia, International Journal of Computer Applications, Vol. 3, No. 7, Article 4, pp. 19–23, ISSN 0975-8887.

[15] Abbiramy V.S. and Shanthi V., "A PCA Based Feature Extraction Approach for the Qualitative Assessment of Human Spermatozoa", International Journal of Computer Science and Information Security, Vol. 9, No. 1, (2011), Paper ID: 31121050, pp. 121-126, ISSN: 1947-5500.

[16] NAB Mary, D Dharma., "Coral reef image classification employing improved LDP for feature extraction", Journal of Visual Communication and Image …, 2017.

[17] NAB Mary, D Dharma., "Coral reef image/video classification employing novel octa-angled pattern for triangular sub region and pulse coupled convolutional neural network (PCCNN)", Multimedia Tools and Applications, 2018.

[18] NAB Mary, R Singh, S Athisayamani., "Banana leaf diseased image classification using novel HEAP auto encoder (HAE) deep learning", Multimedia Tools and Applications, 2020.

[19] N Ani Brown Mary, D Dejey, "Classification of coral reef submarine images and videos using a novel Z with tilted Z local binary pattern (Z⊕ TZLBP)", Wireless Personal Communications, 2018.

[20] N Ani Brown Mary, Dejey Dharma., "A novel framework for real-time diseased coral reef image classification", Multimedia Tools and Applications, 2019.