Received: 17 June 2021 | Revised: 13 August 2021 | Accepted: 13 August 2021
DOI: 10.37917/ijeee.17.2.10 Vol. 17| Issue 2| December 2021

Open Access

*Iraqi Journal for Electrical and Electronic Engineering*
*Review Article*

# A Survey on Segmentation Techniques for Image Processing

**Wala'a N. Jasim *[1], Rana Jassim Mohammed [2]**
[1] Department of Pharmacognosy, College of Pharmacy, University of Basrah, Iraq.
[2] Department of Computer Science, College of Education for Pure Sciences, University of Basrah, Iraq.

**Correspondence**
*Wala'a N. Jasim
Department of Pharmacognosy,
College of Pharmacy, University of Basrah, Iraq.
Email: walaa.jasim79@gmail.com

**Abstract**
***The segmentation methods for image processing are studied in the presented work. Image segmentation can be defined as a vital step in digital image processing. Also, it is used in various applications including object co-segmentation, recognition tasks, medical imaging, content based image retrieval, object detection, machine vision and video surveillance. A lot of approaches were created for image segmentation. In addition, the main goal of segmentation is to facilitate and alter the image representation into something which is more important and simply to be analyzed. The approaches of image segmentation are splitting the images into a few parts on the basis of image's features including texture, color, pixel intensity value and so on. With regard to the presented study, many approaches of image segmentation are reviewed and discussed. The techniques of segmentation might be categorized into six classes: First, thresholding segmentation techniques such as global thresholding (iterative thresholding, minimum error thresholding, otsu's, optimal thresholding, histogram concave analysis and entropy based thresholding), local thresholding (Sauvola's approach, T.R Singh's approach, Niblack's approaches, Bernsen's approach Bruckstein's and Yanowitz method and Local Adaptive Automatic Binarization) and dynamic thresholding. Second, edge-based segmentation techniques such as gray-histogram technique, gradient based approach (laplacian of gaussian, differential coefficient approach, canny approach, prewitt approach, Roberts approach and sobel approach). Thirdly, region based segmentation approaches including Region growing techniques (seeded region growing (SRG), statistical region growing, unseeded region growing (UsRG)), also merging and region splitting approaches. Fourthly, clustering approaches, including soft clustering (fuzzy C-means clustering (FCM)) and hard clustering (K-means clustering). Fifth, deep neural network techniques such as convolution neural network, recurrent neural networks (RNNs), encoder-decoder and Auto encoder models and support vector machine. Finally, hybrid techniques such as evolutionary approaches, fuzzy logic and swarm intelligent (PSO and ABC techniques) and discusses the pros and cons of each method.***

**KEYWORDS: Classification, segmentation, digital image processing, thresholding, image segmentation.**

## I. INTRODUCTION

ONE of the most challenging and important processes related to image processing is the image segmentation. The approaches of image segmentation were used for partitioning images into meaningful parts. In image processing, the major applications are machine vision, content based image retrieval, object detection, medical imaging, video surveillance, recognition tasks and object co-segmentation. In the process of segmentation, the major goal is getting more information from the image's region of interest and allow the object scene's annotation [1]. Image segmentation is challenging in image processing, also it is an active research area which allow the extraction of image objects. Segmentation is considered as an unsupervised learning. The techniques of image segmentation might be categorized into 3 types on the basis of the image properties: edge-based segmentation approaches, region-based segmentation approaches and thresholding segmentation approaches [2], along with deep neural network approaches, hybrid approaches and clustering approaches.

The presented survey is covering the latest image segmentation literature and discussing various segmentation methods proposed until 2020. We review six approaches these were used for segmentation image. The most important techniques used in segment images presented in this paper are: first, thresholding segmentation techniques such as global thresholding (iterative thresholding**,** minimum error thresholding, otsu's, optimal thresholding, histogram concave analysis and entropy based thresholding), local thresholding (Sauvola's approach, T.R Singh's approach, Niblack's approaches, Bernsen's approach Bruckstein's and Yanowitz method and Local Adaptive Automatic Binarization) and dynamic thresholding. Second, edge-based segmentation techniques such as gray-histogram technique, gradient based approach (laplacian of gaussian, differential coefficient approach, canny approach, prewitt approach, Roberts approach and sobel approach). Thirdly, region based segmentation approaches including Region growing techniques (seeded region growing (SRG), statistical region growing, unseeded region growing (UsRG)), also merging and region splitting approaches. Fourthly, clustering approaches, including soft clustering (fuzzy C-means clustering (FCM)) and hard clustering (K-means clustering). Fifth, deep neural network techniques such as convolution neural network, recurrent neural networks (RNNs), encoder-decoder and Auto encoder models and support vector machine. Finally, hybrid techniques such as evolutionary approaches, fuzzy logic and swarm intelligent (PSO and ABC techniques)

## II. THE CLASSIFICATION OF IMAGE SEGMENTATION TECHNIQUES

In literature, there are many utilized methods, each one of them is different from the other in terms of the utilized approach via for segmentation. The CLASSIFICATION of image segmentation methods has been illustrated in Fig. (1). [3].

### *1. Thresholding Segmentation Approaches*

One of the significant image segmentation techniques is thresholding. It is a simply methods utilized for image SEGMENTATION, also it is presenting a clear and fast method to extract objects problem from background. The techniques of thresholding are extracting objects from background through setting a sensible gray threshold (*T*) for the image pixels [2]. The thresholding techniques are on the basis of images features. They are selecting adequate thresholds *Tn* for splitting the pixels of an image into classes as well as subtracting the foreground from background. With regard to single threshold *T*, "object point" was provided as any point *(x, y)* that *f (x,y) >T* , and "background point" was provided as any point (x, y) *that f (x, y)< T* [4 ]; indicating that thresholding is making binary images from grey-level ones via converting every pixel of the image below certain threshold into zero, while every pixel over the certain threshold will be one. In Figure 2, the intensity histogram is matching the image *f (x, y)* consisting of object's lighting on dark background, the pixels related to the object and background have intensity that is grouped with 2 dominant modes [5].

The techniques of thresholding were categorized into global, dynamic and local thresholding methods.

#### 1.1 Global Thresholding

In the case when *T* is on the basis of gray level values, such technique of thresholding will be referred to as "*global thresholding"* [4]. In the case when choosing an adequate single threshold value *T* to input image, then the pixels of the image will be divided into 2 classes (foreground and background), the global thresholding is indicated in the Equation (1) :

$$g(x,y) = \begin{cases} 1, if\ f(x,y) \geq T_{global} \\ 0, iff(x,y) < T_{global} \end{cases} \quad (1)$$

In which:

$g(x,y)$ represents the binary image, $f(x,y)$ represents the original image, T represents the threshold value [5]. There were a few global thresholding approaches, including minimum error thresholding, iterative thresholding, histogram concave analysis, Otsu, optimal thresholding, entropy-based thresholding, and MoM keeping thresholding are presented [4].

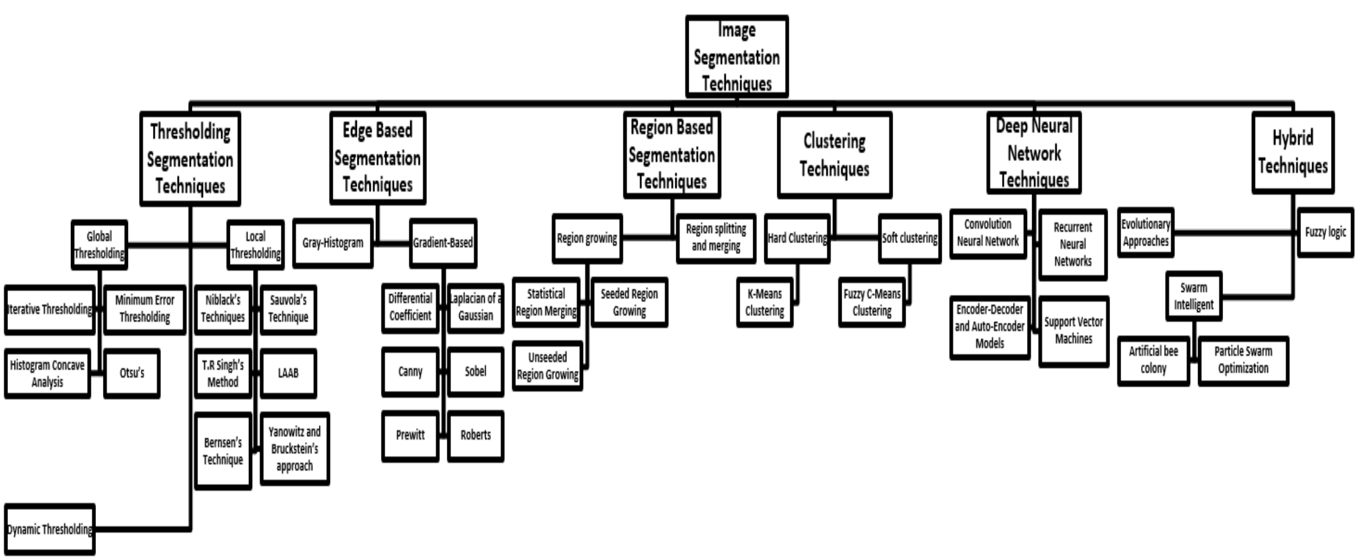
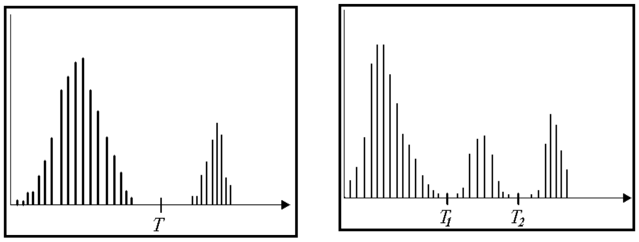



Fig. 1: The Classification of Image Segmentation Methods




(A)Single (B)Multiple

Fig.2: Thresholding. [4][6]

*A) Iterative Thresholding*

This is considered as a global thresholding approach. It is specified as fairly simple approach which doesn't need a great deal of knowledge regarding the image, also it showed good performance against image noise [7]. In as study conducted by [8], procedure automatically computed the intensity threshold *T* that is indicated via Woods and Gonzalez in (2002) [6]. Below, is the summarization of the technique:

1- Making initial estimation related to *T*, for instance, the middle intensity.

Segmenting the histogram of the image into group *G1* of intensities < *T* and a group *G2* of intensities ≥ *T* and computing the mean of intensity values $\mu 1$ as well as $\mu 2$ of pixels *P(I)* in the 2 groups in the following way:

$$\mu = \sum_{I=0}^{N-1} I\, P(I) \qquad (2)$$

$$P(I) = n1 / N \qquad (3)$$

In which *N* represents the total number of pixels, while *n1* represents the number pixels with intensity *I*.

2- Computing a new threshold

$$T = 0.5(\mu 1 + \mu 2) \qquad (4)$$

3- Repeating steps 2 & 3 till the success value difference of *T* is not more than a pre-defined limit.

$$|T - Tn| < 0.1 \qquad (5)$$

*B) Minimum Error Thresholding*

Illingworth and Kittler (1986) suggested a minimum error thresholding algorithm which is minimizing the classification error probability via fitting error expression. In addition, the gray-level histogram in this approach is indicated as an estimate related to probability density function *p(g)* of mixture population consisting of the gray-level of background as well as the object pixels [9][10]. Typically, it is suggested that the 2 mixture components $p(g|i)$ were normally distributed with $\mu_i$, standard deviation $\sigma_i$ and priori probability $p_i$ that is:

$$p(g) = \sum_{i=1}^{2} pi\, p(g|i) \qquad (6)$$

In which,

$$(g|i) = \frac{1}{\sqrt{2\pi\sigma i}}\, exp\left(-\frac{g-\mu i}{2\,\sigma_i^2}\right) \qquad (7)$$

The threshold value might be selected via quadratic equation

$$\frac{(g-\mu 1)^2}{\sigma_1^2} + \log_e \sigma_1^2 - 2\log_e P_1 = \frac{(g-\mu 2)^2}{\sigma_2^2} + \log_e \sigma_2^2 - 2\log_e P_2 \quad (8)$$

The parameters $\mu_i$, $\sigma_i^2$ as well as $p_i$ ($i$=1,2) regarding the density of the mixture *p(g)* associator with the image to be thresholded aren't typically known. To overcome the complexity of assessing the unknown parameters, in a study conducted by [10] Illingworth and Kittler provided a criterion function *J(t)*, given as follows:

$$J(t) = 1+2\{P1(t)\log e\sigma 1(t) + P2(t)\log e\, \sigma 2(t)\} - 2\{P1(t)\log e P1(t) + P2(t)\log e\, P2(t)\} \qquad (9)$$

In which,

$$P_1(t) = \sum_{g=0}^{t} h(g) \quad , \quad P_2(t) = \sum_{g=t+1}^{t-1} h(g) \qquad (10)$$

$$\mu_1(t) = \frac{\{\sum_{g=0}^{t} h(g)g\}}{P_1(t)} \quad , \quad \mu_2(t) = \frac{\{\sum_{g=t+1}^{t-1} h(g)g\}}{P_2(t)} \qquad (11)$$

$$\sigma_1^2(t) = \frac{\{\sum_{g=0}^{t} (g-\mu_1(t))2\, h(g)\}}{P_1(t)} \qquad (12)$$

and

$$\sigma_1^2(t) = \frac{\{\sum_{g=t+1}^{t-1} (g-\mu_2(t))2\, h(g)\}}{P2(t)} \qquad (13)$$

The optimize threshold is get through minimizing *J(t)* via finding:

*t* = Arg Min i ∈ G J(t)*

*B) Histogram Concave Analysis*

A study conducted by [11] indicated that the concavity points related to maximum "depth" were excellent candidates for the thresholds. Also, such points might be effectively located via designing the histogram's convex hull. Assuming the given histogram H is specified over set of gray-levels *K*,…., *L* , with bar heights *h(H)*,…, *h(L)*, in which *h(K)* as well as *h(L)* were non-zero, while all h's were non-negative. *H* represents a 2D region bounded on left, bottom, and right by the straight lines *(K,0 )(K, h(K)),(K, 0),(L,0)* , and *(L,0 ),(L,h(L))*, consecutively Fig. (3).

To find the concavities of *H* convex hull was developed; this was the minimum convex polygon *H* contains *H*, and concavities were identified via taking the set-theoretic difference $\overline{H}$ - *H*. (A concavity might be specified as a connected component related to $\overline{H}$ - *H*.) It has been indicated that the 3 straight lines which bound *H* on left, right, and bottom were the sides of *H*; therefore, the part of *H* specified via the histogram bars' top was required to construct. An easy algorithm to construct H via a left to right scan of H was followed [12].

1- Beginning at *(K, h(K))*, calculate the slopes $\theta_i$ of line segment *(K, h(K)) (i, h(i))*, $K+1 \le i \le L$ , in which $-90^{\circ} < \theta_i < 90^{\circ}$
2- Let $\theta_{k1}$ is the largest of the slopes and $K_1$ is the right most point having such slope.
3- *(K,h(K))(K1,h(K1))* is a side of the convex hull shows Fig. (4).
4- The process is repeated with $K_1$ replacing K, which means, calculating the segments' slopes $(K_1, h(K_1))(i\ h(i)),\ K_1+1 \le i \le L$
5- Existing the largest slope $\theta_{K2}$ and the right most point $K_2$ having such slope, yielding $(K_1, h(K_1))(K_2,h(K_2))$ as a side of the convex hull, and so on, until access *L*.

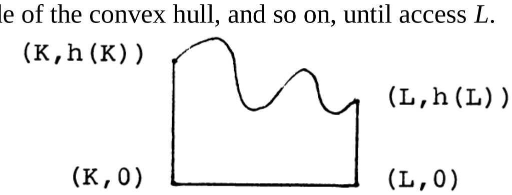



Fig. 3: Histogram specified as 2D region

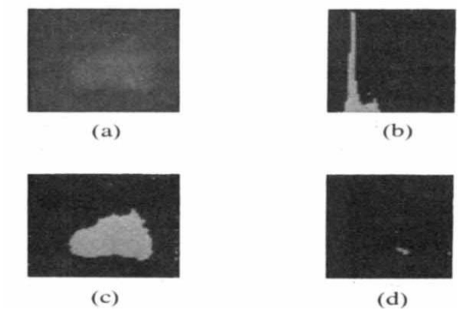



Fig.4: Test image. (a) Image (b) Histogram (c) & (d) Thresholded images

*D) Otsu's*

Otsu's can be defined as one of the unsupervised and non-parametric approaches related to automatic threshold selection for image segmentation. The approach is used for outdoing the iterative thresholding drawbacks, and estimating the mean following each one of the steps. Otsu's has the aim to find the optimal value in terms of global threshold [13].

Otsu's include all possible threshold values and computes the pixel levels in each side of the threshold value. Threshold value split the foreground or background of pixels. This algorithm categorizes the image into two categories of pixels such as within category and between category variance [14-15].

The formulation of this technique will present by the following equations, Assuming an image's pixels are represented in L gray levels [1, 2, ,L], then the number of pixels at level i was indicated via $n_i$, while the total amount of the pixels is indicated through $N = n_1 + n_2 + + n_L$. with a view to simplify the discussion, gray level histogram was normalized and specified as a distribution of the probabilities:

$$p_i = n_i / N,\ p_i \geq 0,\ \sum_{i=1}^{L} pi = 1 \qquad (14)$$

The pixels are divided to 2 categories, which are $C_0$ & $C_1$ (objects and background, or vice-versa) through the threshold at level k; $C_0$ represents the pixels with levels [1, , k], and $C_1$ represents the pixels with levels [k 1, , L]. In addition, the class probabilities occurrence as well as the mean levels of class, consecutively, are provided as follows:

$$\omega_0 = P_r(C_0) = \sum_{i=1}^{k} p_i = \omega(k) \qquad (15)$$

$$\omega_1 = P_r(C_1) = \sum_{i=k+1}^{L} p_i = 1 - \omega(k) \qquad (16)$$

and

$$\mu_0 = \sum_{i=1}^{k} iP_r(i|C_0) = \sum_{i=1}^{k} \frac{ip_i}{\omega_0} = \mu(k)/\omega(k) \qquad (17)$$

$$\mu_1 = \sum_{i=k+1}^{L} iP_r(i|C_1) = \sum_{i=k+1}^{L} \frac{ip_i}{\omega_1} = \frac{\mu_T - \mu(k)}{1-\omega(k)} \qquad (18)$$

Where

$$\omega(k) = \sum_{i=1}^{k} p_i \qquad (19)$$

and

$$\mu(k) = \sum_{i=1}^{k} ip_i \qquad (20)$$

First- and zeroth-order cumulative moments related to histogram up to k-th level, consecutively, and

$$\mu_T = \mu(L) = \sum_{i=1}^{L} ip_i \qquad (21)$$

The total average level regarding the original image, the next relation for any choice of k might be simply verified.

$$W_0\mu_0 + W1\,\mu_1 = \mu_T \quad ,\ w_0 + w_1 = 1 \qquad (22)$$

The class variances are given by

$$\sigma_0^2 = \sum_{i=1}^{k} (i - \mu_0)^2\ \Pr(i|\, c_0) = \sum_{i=1}^{k} (i - \mu_2)^2\ pi\ /w_0 \qquad (23)$$

$$\sigma_1^2 = \sum_{i=k+1}^{k} (i - \mu_1)^2\ \Pr(i|\, c_1) = \sum_{i=k+1}^{L} (i - \mu_2)^2\ pi\ /w_1 \qquad (24)$$

With a view to evaluation the "finedness" of threshold (at level k), the next discriminant criterion measure is provided (or measures of class separability) utilized in the discriminant analyses.

$$\lambda = \sigma_B^2 / \sigma_W^2,\ K = \sigma_T^2 / \sigma_W^2,\ \eta = \sigma_B^2/\sigma_T^2 \qquad (25)$$

In which,

$$\sigma_W^2 = \omega_0\sigma_0^2 + \omega_1\sigma_1^2$$

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 \qquad (26)$$

$$= \omega_0\omega_1(\mu_1 - \mu_0)^2$$

$$\sigma_T^2 = \sum_{i=1}^{L} (i - \mu_T)^2\, p_i \qquad (27)$$

Otsu's technique is simple implementation but the major disadvantage of Otsu's technique of threshold chosen is that it is assuming that the histogram was bimodal. Also, Otsu's is considered to be effective in the case when 2 classes were of different sizes and with changeable illumination and not producing excellent results in the case when the histogram of the image has over 2 peaks [13].

*E) Optimal Thresholding*

Pun (1980, 1981) suggested an optimum criterion with regard to image thresholding. The criterion has been corrected as well as enhanced via Kapur et al. (1985). The researchers revised and enhanced the algorithm of Pun's through taking into account 2 probability distributions for background and objects and maximized the image entropy for obtaining optimal threshold [17]. In addition, the optimal thresholding is specified as a histogram of the image which will be weighted sum of at least 2 probability densities. After that, the threshold was set as the gray-level that outcomes in minimum number of pixels mis-classified bad split, indicating that the background pixels are split as vice versa and foreground.

Fig. (5) shows a noisy of image and Fig. (6) shows the optimal threshold as well as conventional threshold salient on its histogram. Fig. (7) shows thresholding at the optimal threshold outputted a better outcome compared to the thresholding at the conventional threshold [16] .

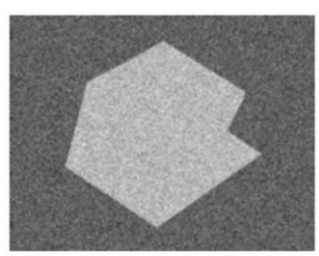


Fig. 5: A noisy of image

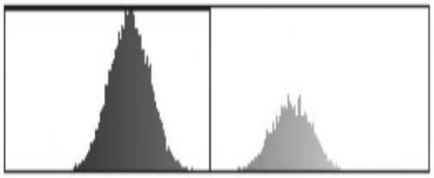


Fig. 6: The optimal threshold and conventional threshold salient

Fig. 7: Thresholding at the optimal threshold

$$\mu(k) = \sum_{i=k+1}^{L-1} p_i \qquad (28)$$

Where L is the number of gray levels (e.g. 256, for an 8-bit image) . By defintion

$$\omega(k) + \mu(k) = \sum_{i=0}^{L-1} p_i = 1 \qquad (29)$$

$K$ is found to maximize the difference between $\omega(k)$ and $(k)$ . which may be performed through initially defining the image average as :

$$\mu_T = \sum_{i=0}^{L-1} i p_i \qquad (30)$$

then finding the value of $K$ that maximizes

$$\frac{(\mu T\omega(k) - \mu(k))^2}{\omega(k)\mu(k)} \qquad (31)$$

That efficiently maximizes between class variances or minimizes within class variance. $K$ was selected for maximizing the disengagement regarding 2 classes "foreground" and "background", or alternatively minimizing their overlap. In addition, the optimum threshold which it finds is stable, based as it is on the integration of the grey level histogram (i.e. the global characteristics) instead of its differentiation (i.e. local property like the valley).

Optimal thresholding might be iteratively conducted via isodata (iterative self organizing data analysis method), such process is explained as follows:

1- Thresholding the image with the use of the mean regarding 2 peaks or mean pixel value,

2- Calculating the mean value related to pixels below such threshold $\mu 1$, as well as the mean regarding pixels above such threshold $\mu 2$.

3- Thresholding the image at new threshold $T_i = (\mu_1 + \mu_1)/2$.

4- Repeating the steps (2, 3, and 4) till $T_i - T_i\text{-}1 \le \Delta$, in which the change $\Delta$, might be specified in various ways, either through evaluating the relative changes in the threshold value or through the percentage of pixels which change sides (background to foreground or the other way around) between the iterations.

### F) Entropy-Based Thresholding

The entropy that is utilized as a fuzziness measures is considered to be in analogy with information theory entropy, yet with insignificant differences in definition. A study conducted by Fan et al. (1996), suggested a rapid entropic approach for automatically obtaining the global threshold through decreasing the computation complexity. A study conducted by Portes de Albuquerque et al. (2004) suggested an entropic thresholding approach which has been customized from the non-extensive Tsallis entropy concept. A study conducted by Xiao et al. (2008) suggested an entropic thresholding approach on a basis of the grey level spatial correlation (GLSC) histogram. The researchers revised as well as extended the algorithm of Kapur et al.'s (Kapur et al. 1985) [17].

The presented study is reviewing the entropy thresholding suggested via Xiao et al. (2008). The GLSC histogram was estimated in the following way. Assuming g(x,y) is the number of pixels to small windows containing $N$ x $N$ pixels:

$$g(x,y) = \sum_{i=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{-\frac{N-1}{2}}^{\frac{N-1}{2}} v(|f(x+i, y+i) - f(x,y)| \le \zeta) \qquad (32)$$

The pixel gray values, g(x,y) and f(x,y) were utilized for creating GLSC histogram:

$$h(k,m) = Pr\,ob(f(x,y) = k \text{ and } g(x,y) = m), \qquad (33)$$

where $k \in G$, & $m \in \{1,2,\ldots, N \times N\}$

The probability function p(k,m) is

*p(k,m) = hˆ(k,m).*

The distributions of the probability that are related to background and object were acquired as follows:

$$G_{obj} = \left\{\frac{P(0,1)}{P_{obj}}, \cdots, \frac{P(0,N\times N)}{P_{obj}}, \cdots, \frac{P(0,1)}{P_{obj}}, \cdots, \frac{P(TH,N\times N)}{P_{obj}}\right\}, \qquad (34)$$

$$P_{obj} = \sum_{k=0}^{TH} \sum_{m=1}^{N\times N} P(k,m), \qquad (35)$$

and

$$G_{bg} = \left\{\begin{matrix} \frac{P(TH+1.1)}{P_{bg}}, \cdots, \frac{P(TH+1,N\times N)}{P_{bg}}, \cdots \\ , \frac{P(255.1)}{P_{bg}}, \cdots, \frac{P(255.1,N\times N)}{P_{bg}} \end{matrix}\right\}, \qquad (36)$$

$$P_{bg} = \sum_{k=TH-1}^{255} \sum_{m=1}^{N\times N} P(k,m), \qquad (37)$$

The entropy calculation related to the elements in GLSC histogram were weighted via non-linear function linked to m and N acquired via:

$$\omega(m,N) = \frac{1+e^{-9m/N\times N}}{1-e^{-9m/N\times N}} \qquad (38)$$

Therefore, the entropies related to objects as well as background distributions are as follows:

$$H_{obj}(TH,N) = -\sum_{k=0}^{TH} \sum_{m=1}^{N\times N} \frac{P(k,m)}{P_{obj}} \ln\left[\frac{P(k,m)}{P_{obj}}\right] \omega(m,N), \qquad (39)$$

$$H_{bg}(TH,N) = -\sum_{k=TH-1}^{255} \sum_{m=1}^{N\times N} \frac{P(k,m)}{P_{bg}} \ln\left[\frac{P(k,m)}{P_{bg}}\right] \omega(m,N), \qquad (40)$$

### 1.2 Local Thresholding

The local threshold method needs to select multiple segmentation thresholds and divides the image into multiple target regions and backgrounds by multiple thresholds [18] . In the case when threshold T is on the basis of the point's gray value and certain point's local property including the average gray value regarding the neighborhood which is centered on the point (x,y), the thresholding is referred to as "local thresholdin" [4]. The approach is specifying the values of threshold locally with regard to each one of the image regions as indicated in Equation(41) [19]. In addition, a single threshold won't work-well in the case when having uneven illumination because of the shadows or because of the illumination direction. The major approach is portioning the image into m x m sub-images and after that choosing the threshold.

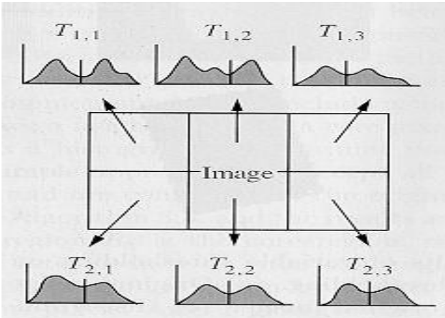



Fig.8: Sub images of an image [13]

$$g(x,y) = \begin{cases} 1, \text{if } f(x,y) \geq \text{T local} \\ 0, if\ f(x,y) \geq \text{T local} \end{cases} \quad (41)$$

With regard to such condition, multi values regarding local thresholding might be utilized and needed for many image areas, such process is referred to as multi-level thresholding, Equation (42) is explained [19].

$$g(x,y) = \begin{cases} a, & if\ f(x,y) > T2 \\ b, & if\ T1 < f(x,y) \leq T2 \\ c, & if\ f(x,y) \leq T1 \end{cases} \quad (42)$$

The histogram of the image is specified via 3 controlling modes (the example of 2 light objects' types in dark background). If *T1 < f(x ,y)*
*≤ T2*, then might be classified a point *(x,y)* as single object class, in the case when *f(x, y) > T2* is classified to another object class, and in the case when *f(x, y) ≤ T1* it will be classified to background as indicated in Fig. (2).

A lot of aspects are affecting the suitability histogram to choose the threshold [5][19]:

- Image noise.
- Related size of background and objects.
- Illumination's regularity.
- Reflectance regularity.

*A) Niblack's Techniques*

Niblack thresholds were specified as local thresholding approaches which were of high importance for images in which the background isn't uniform, particularly for text recognition. Fig.(9) is showing many methods of segmentation, such as Niblack's threshold.

With regard to this technique, the local threshold value *T(x, y)* at *(x, y)* was evaluated in window of size *w* x *w as follows:*

$$T(x, y) = m(x, y) + k * \delta(x, y) \quad (43)$$

In which m(x, y) as well as δ(x, y) were the local mean and SD regarding the pixels in local window and k represents the bias. Set as k = -0.20 and local window size is w=15. Standard deviation δ(x, y) and Local mean m(x, y) are adapting the threshold value on the basis of the contrasts in pixel's local neighborhood. Also, bias k is controlling the adaptation level changing the threshold value. Here, the pixel = (pixel > mean + k * standard deviation) object : background [20].

This approach provides good results, in which the intensity variation was different along the location of pixel, yet the drawbacks is that the threshold value is on the basis of the choice ok the bias value[21].

*B) Sauvola's Technique*

This technique, which has been proposed by Sauvola's, Sauvola's thresholds is also method of local thresholding, which can be beneficial for the images in which background is uneven, in particular for the recognition of text as Niblack threshold. A threshold T(x, y) in the Sauvola's method is calculated with the use of standard deviation δ(x, y) and mean m(x, y) of the intensity values of the pixels in a window that has w × w dimensions, and centered around pixel at (x, y) and it can be represented by:

$$T(x,y) = m(x,y) \left[1 + k\frac{\delta(x,y)}{R} - 1\right] \quad (44)$$

The advantages of Sauvola's Technique is that it provides sufficient results where there's a change between the background and the foreground variation over an image, however, the drawbacks lie in the fact that the value of the threshold is dependent on the bias value and the maximal value of the standard deviation (SD) for an image [21] .

R represents the maximal SD value (R=128 for the gray-scale documents), and k represents a parameter that would take positive values ranging between [0.20 and 0.50]. Where a pixel = (pixel > mean × (1 + k × (SD / r - 1))) ? object : background [20] .

In the case where there is a high level of the contrast in a certain image area, δ(x, y) ~ R that yields T(x, y) ~ m(x, y). which represents same result like in the Niblack's approach. None-the-less, difference is resulted in the case where the local neighborhood's contrast is rather low. In such case, threshold T(x, y) decreases less than the average value, which results in the successful removal of rather dark background areas [22]. For the purpose of computing the value of the threshold T(x, y), the local SD and mean must be calculated for every one of the pixels its computation complexity is O ($n^2$ x $w^2$) in naïve manner for an n x n image [23]. Which indicates the fact that its computation complexity depends upon the size of the window. T. Singh has suggested window size independent thersholding approach utilizing integral summation image as prior procedure. Fig. (9) shows various segmentation methods , including Sauvola's technique .

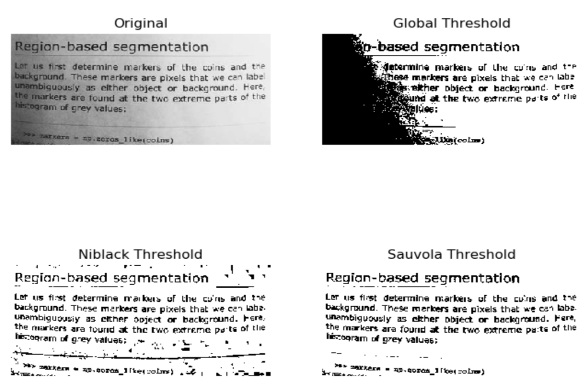



Fig. 9: Various Segmentation Methods

*C) T.R Singh's Method*

For the purpose of the minimization the local thresholding calculation's computational time, Singh proposed a sufficient manner for the determination of the local threshold with the use of the integral summation image as one of the prior processes for the determination of the local summation [20]. It only utilizes the local mean and it's quite convenient using integral summation while other methods such as Niblack and Sauvola's. Approaches aren't quite convenient due to the use of both standard deviation and local mean. As a result of the utilization of the integral summation image, this approach is the local window size independent. This approach is represented by:

$$T(x,y) = m(x,y)\ [1 + k\frac{\partial(x,y)}{1-\partial(x,y)} - 1] \qquad (45)$$

Where $\partial(x,y)= I(x, y) - m(x, y)$ represents local mean and is bias that may be controlling the degree of the adaptation varying value of the threshold. k value can take a significant impact on the determination of the value of the threshold. The lower k value results in increasing the value of the threshold and the other way around [20].

*D) LAAB (Local Adaptive Automatic Binarization)*

The automatic binarization, which has been suggested by Singh, can be represented a procedure for the transformation of grey scale image I(x, y) to binary image b(x, y) in an automatic manner with no use of any threshold value T(x, y) through the adaptation of pixels in the environment of the local region. It's an automated binarization with the local adaptations [24]. The local adaptations are performed in a local w x w window using local mean m(x, y) of the values of the pixel intensities of the pixels in local area. Automated binarization has been designed based on the following equation:

$$b(x,y) = \frac{|1-2v| - |1-2v|}{2|1-2v|} \qquad (46)$$

Where $v = \frac{k(1+\partial)}{(1-\partial)}$

*k* represents the value of a bias such that 0.50< k < 0.60

$\partial$ = {g(x, y) - m(x, y)}{1-m(x, y)} & g(x, y) represents original pixel at (x,y).

Bias is responsible for the control of the pixel adaptation level in local area at the transformation time to b(x, y) binarized image. An increase in k value, results in increasing the background and decreases the foreground area and the other way around [20].

*E) Bernsen's Technique*

This approach, which has been suggested by Bernsen, is an approach of local binarization, using the value of the local contrast for the determination of the value of the local threshold, for every one of the pixels (x, y) is computed from the equation:

$$T(x,y)= \frac{Imax+Imin}{2}$$

Where I max and I min stand for the maximal and minimal value of the gray level in a window of w x w size, which has been centred respectively at (x, y) [25]. However, the assignment of the threshold has been based upon the value of the local contrast, which is why, it may be represented as: if the value of $(I\ max – I\ min) > L$ i.e. in the case where, grey scale image is non-uniform:

Then

$$T(x,y)= \frac{Imax+Imin}{2}$$

Otherwise

$T(x, y) = GT$ //(in this case, the value of the threshold is computed from the method of the global thresholding.)

Where L represents a threshold of the contrast and GT represents a value of the global threshold.

*F) Yanowitz and Bruckstein's approach*

Bruckstein and Yanowitz suggested the use of gray-level values at areas that have high gradients as known data for the interpolation of the threshold surface of the texture characteristics of the image [24].The fundamental operations of that this approach can be summarized below:

1.Smoothing images through the use of the average filters.
2. Deriving the magnitude of the gradient.
3. Applying an algorithm of thinning for the purpose of finding the boundary points of the object.
4. Sampling the gray-levels in smoothed image at boundary points. Those can be defined as support points for the interpolations in the step5.
5. Finding threshold surface T(x, y) which equals the values of the image at support points and is satisfying Equation (47) with the use of the Southwell's approach of the successive over relaxation

$$\frac{\partial^2 p(x,y)}{\partial x^2} + \frac{\partial^2 p(x,y)}{\partial y^2} \qquad (47)$$

6. Utilizing obtained T(x, y), segmenting the image.
7. Applying an approach of the post-processing for the validation of segmented image.

1.3 Dynamic Thresholding

Local adaptive thresholding is done by analyzing local image features to determine a threshold value for each pixel; such a method is sometime called dynamic threshold [26-27].

In the case where there is a number of the objects in an image residing in a variety of the grey level areas, then that image has to be divided by a variety of the dynamic threshold values (T1, T2, …Tn), which are dependent upon p(x, y), f (x, y), and spatial coordinate values x & y. Generally, the methods ofthe dynamic thresholds include Watershed, thresholding image, interpolatory thresholding, etc. [4].

One dynamic thresholding method (Morrin / 1974) uses an edge gradient to bias the thresholding value from the local average gray-levels to produce a sharper image. The current pixel is stset to be black if the current scan value *x* is less than a threshold value *T(x)* , where

$$T(x) = A + G + b\ ,$$

Where

*A* is the average value of the surrounding pixels
*G* is the gradient biasing for enhancement
*b* is a fixed negative constant to bias background toward white .

A desirable gradient operation in dynamic thresholding is sensitive to character edges but insensitive to random noise and halftone backgrounds. A mini-max corner –sum gradient in a (3 × 3 ) matrix and the clustering of scan pixels improves rejection of random noise in a document ( Wong / 1978 ) . In a dynamic – thresholding algorithm developed for OCR applications (White and Rohrer / 1983) , two first order – difference equations are used to update the background

levels in the both the horizontal and vertical directions . The thresholding scheme is similar to ( Morrin/1974), except that gradient enhancement is not used.
We summarizes the benefits and the drawbacks of the image segmentation threshold based method which proposed in the literature.

TABLE 1
SUMMARY OF THE BENEFITS AND THE DRAWBACKS OF THE IMAGE SEGMENTATION THRESHOLDING TECHNIQUES [18] [28]

| Advantages | Disadvantages |
| --- | --- |
| • the calculations are simple<br>• Computationally inexpensive.<br>• Requires no prior knowledge about the image.<br>• May be utilized in the real time applications.<br>• the speed of the operation is faster<br>• In particular, when the background and the target have high contrast, the effect of the segmentation may be produced. | • For images with flat and broad valleys or with no peaks, it does not operate sufficiently.<br>• There is a difficulty in obtaining precise results for the problems of the image segmentation in which there aren't any considerable differences in the gray scale or a large gray scale values' overlapping in that image.<br>• Disregards the image spatial information, can't ensure the fact that the areas that have been segmented are contiguous.<br>• Due to the fact that it can only consider the image's gray information with no consideration of the image spatial information, it has a sensitivity to the gray-scale unevenness and noise, which leads to usually combining it with other approaches.<br>• Threshold selection is crucial, as the wrong choices can result in over or under segmentation.<br>• High sensitivity to the noise. |

## 2. Edge Based Segmentation Techniques

Image segmentation can be defined as an approach that is utilized to separate image from background and read the content. The edge has been considered as an important and significant image feature. The edge may be characterized as the boundary that separates 2 different image areas. The operation of the Edge detection mean the procedure of the identification and localization of the sharp discontinuities that may be present in the image [29].
The methods of the edge detection are well advanced image processing methods on their own. Edge based approaches of the segmentation have been based upon the abrupt intensity change in the image due to the fact that the single value of intensity provides no sufficient data on the edges. The methods of the edge detection are responsible for the location of edges, in which either the 1$^{st}$ intensity derivative is higher than a specific threshold or the 2$^{nd}$ derivation has 0 crossing. In the edge based approaches of the segmentation, initially, edges will be detected, after that, they are connected together for the purpose of forming object boundaries for the segmentation of the needed regions. There are 2 fundamental approaches of the edge-based segmentations, which are: the grey-histogram approach and the gradient-based approach [3][4]. There are 3 most widely utilized Gradient based approaches, which are: the approach of the differential coefficient, Canny method, and Laplace of Gaussian (LoG) [4] .

### 2.1 Gray-Histogram Method

Edge detection quality is greatly dependent upon the fitness of thresholdT However; it is actually complicated searching for maximal and minimal grey values, due to the fact that the grey histogram is uneven for the noise impact. In such case, it is possible to approximatively replace the object and background curves by 2 conic Gauss curves, whose intersection is the histogram valley. The value of the threshold T represents the grey value of point at this valley.

### 2.2 Gradient-Based Approach

The Gradient can be defined as the 1$^{st}$ image f (x, y) derivative. In the case where the grey value variation near the edge is sufficiently intense and there isn't much noise in the image, the Gradient-based approach operated sufficiently, and the result of the segmentation is adaptive to the gradient oritentation [4]. The gradient-based segmentation principle can be described as associating the object of interest's boundaries with the ensity of the gradientint crests observed in an image. For the purpose of obtaining meaningful segmentation of the images, a number of the conditions have to be met. In the case of the low intrinsic resolution of the imaging device in comparison to voxel size, the transitions between the areas of the variety of the image activities appear blurry. Which is why, the peaks of the intensity of the gradient aren't sharp, which is why, they are harder to identify. Another limitation has been associated with the fact that the noise can get amplified in the image of the gradient-intensity, in comparison to initial image. As first approximation, the noise and resolution effects may be shown below. At first, an 'ideal' unknown image (which is free of any noise and has an infinitely high level of the resolution) undergoes blurring through the convolution of that image with the considered imaging device's point spread function. Second, this blurred image will be corrupted through the use of the statistical noise. The really obtained images may be considered to be resulting from that 2-step model. Which has suggested the fact that ideal images may be recovered through the use of the reverse model that has been abovementioned, in the first approximation [30]. There have been 3 most widely utilized Gradient-based approaches utilized, which are: LoG, Canny method and differential coefficient technique. Amongst those, the Canny approach has been defined as most representative method [4] . In addition to Sobel method , Prewitt method and Roberts method.

*A) Differential Coefficient Technique*

Utilizing the partial differential equation (PDE)-based approaches and solving PDE equations through the use of a numerical scheme, an image may be segmented. The most significant approaches of the image segmentation that have been based upon PDEs, are: fast marching techniques, parametric techniques, and level set techniques.

- *Parametric techniques*

Lagrangian methods have been based upon the parameterization of contours based on a certain strategy of the sampling and after that, evolving every one of the elements based upon the internal terms and the image. Those approaches are known for their speed and efficiency, none-the-less, the original formulation, which is "purely parametric" (based on Terzopoulos and Kass in 1987 and has been referred to as the "snakes"), it has received a considerable share of the criticism, in general, as a result of its drawbacks, in regards to the sampling strategy choice, the internal geometrical characteristics of a curve, changes of the topology (splitting and merging of the curve), addressing issues in some higher dimensions, and so on. In the present days, the sufficient "discretized" formulations were advanced for addressing those issues at the same time as preserving the high level of the efficiency. In the two cases, the minimization of the energy is performed in general with the use of a steepest-gradient descent, while the derivatives are calculated with the use of, for example, the finite differences [31].

- *Level set techniques*

This approach has been has been suggested first by Osher and Sethian in 1988 for tracking the moving interfaces and spread over a variety of the domains of imaging in late 90s.[32-33] It may be utilized for the efficient addressing of the problems of the curve/surface/etc. propagations in implicit way. The main concept is representing the evolving contour with the use of a signed function, where its 0 level is corresponding to actual contour. After that, based on the contour motion equation, there is a possibility for easily deriving an equivalent flow for the implicit surface which, in the case of being applied to 0-level will be reflecting the contour propagation. This approach may be utilized due to a wide range of the benefits: it is parameter free, implicit, intrinsic, presents a direct method for the estimation of the geometric characteristics of evolving structures, and has the ability of changing the topology. In addition to that, they may be utilized for defining a framework of optimization as it has been suggested in 1996 by Zhao, Osher and Merriman. Which is why, there is a possibility for concluding the fact that it's quite a convenient model for addressing a wide range of the medical image analysis and computer vision applications. In addition to that, researches about a variety of the level set data structures resulted in a highly sufficient implementations of that approach [31].

- *Fast marching techniques*

This approach is utilized in segmentation of images, and was enhanced (which results in permitting a positive as well as a negative propagation of the speed) in a method that has been referred to as generalized fast marching approach [33]. This approach can be considered as a numerical approach to solve the problems of the boundary value of Eikonal formula [31][33] .

$$|\nabla T(x)| = F(x) = 1 \tag{48}$$

Usually, this types of the problems describes the closed curve evolution as a time T function with speed F(x) in normal direction at a point x on curve. The function of the speed is identified, and time where a contour crosses a point x results from solving the formula. The algorithm is like the Dijkstra's algorithm and utilizes the fact that the information can only flow outwards from the region of the seeding. That issue has been defined as one of the special cases of the level set approaches. There are some more general algorithms, however, they but are usually more time consuming [31]. The Extensions to the non-flat (i.e. triangulated) domains solving:

$$F\ x\ |\nabla_s\ T\ x| = 1 \tag{49}$$

for the surface S ,and x ∈ S was introduced by Ron Kimmel and Sethian. The fast marching approach is only suitable for the problem of the boundary value. The main concept that underlies this approach is that the problem of the boundary value $|\nabla T|\ F = 1$ includes a front which is always contracting or expanding.

*B)Laplacian of a Gaussian(LoG)*

The Laplacian-of-Gaussian (LoG) utilizes a Gauss filter for blurring the image and a Laplacian for enhancing the edges. It is seldom used on its own for edge detection because of its sensitivity to noise. Edge localisation is done by finding zero crossings. Also known as Marr & Hildreth edge detector [34].

The LoG algorithm is explained by the following steps :

1- Convolving the image with a 2-D Gaussian function
2- Computing the Laplacian of the convolved image L
3- Identify edge pixels as the ones for which there is a 0-crossing in L.

A radially-symmetric 2D Gaussian:

$$G_0(x,y) = e^{\left(\frac{-r^2}{2\sigma^2}\right)} \tag{50}$$

Where

$$r2 = x2 + y2 \tag{51}$$

The Laplacian of this is:

$$G(x,y) = \left(\frac{r^2 - \sigma^2}{\sigma^4}\right) e^{\left(\frac{-r^2}{2\sigma^2}\right)} \tag{52}$$

This function has a minimum at its origin, but it is usual to invert the filter, so that it has a maximum at its origin. It is a classic "Mexican hat" shape which shows in Fig. (10).

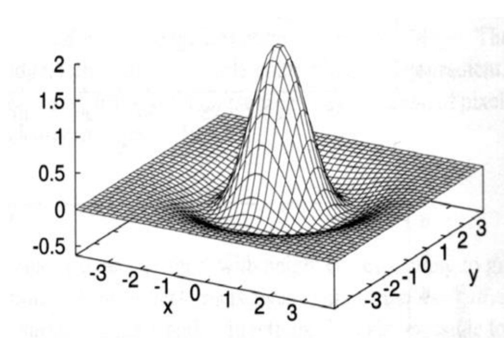



Fig. 10: Classic Mexican hat

The value of σ is responsible for the determination of the filter width and controlling the smoothing amount that results from the Gaussian component.

- σ tunes the filter to detect edges at different scales
- Should have a half-width of at least 3σ.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

5×5 Laplacian of Gaussian kernel. The LoG can be approximated by convolving with a kernel that is the difference of two Gaussian kernels with substantially different σ's. Which is known as the Difference-of-Gaussian (DoG).

*C) Canny technique*

In [35] Canny methods was presented by JOHN CANNY in 1986 to edge detection. Canny is an efficient and effective method of edge detection which yields a significant reduction of the requirements of memory, reduction in the latency and increase in the throughput without any loss in the efficiency of the edge detection [36]. Detecting the image edges can be helpful for segmenting the image, compressing data, as well as playing a role in some significant operations like the image modernization [29].

The canny edge detector can be characterized most strictly defined operator and it's commonly utilized amongst the approaches of the edge detection, the popularity of this detector results from the fact that it's the best approach to find edges with good localization, good detection, as well as a single response to the edge. Canny can determine the edges through the implementation of a process of optimization and suggested an approximation to optimum detector as the gradient magnitude's maxima of the Gaussian-smoothed images [29].

There are five separate steps for detecting the edges by the Canny edge detection method :

**Step 1 Smoothing:** this process is carried out with the use of the canny detector, which aims at filtering the noise out in original image. Which is carried out through the conversion of input image to grayscale through the adjustment of the brightness and contrast, thus, the image is blurred out for the purpose of removing noise. In general, Gaussian filter is utilized to remove noise.

**Step2 Finding the Gradients:** the pixels of the edge are the ones in which there is a sudden variation in the values of the gray level. Those pixels can be characterized through the calculation of image gradient, which can be defined as a unit vector, pointing in towards the maximal change of the intensity. The gradient's horizontal and vertical components are initially calculated and after that, the gradient direction and magnitude are calculated.

**Step3 Non-Maxima Suppression:** fundamentally, the thinning of the edge is carried out in the no-maxima suppression. Throughout this procedure, based on the magnitudes of the gradient, thick image edges are converted into nearly sharp and thin edges that may be utilized additionally for the purposes of recognition. In that step, images are scanned throughout the direction of the edge and any pixel values which are not considered as an edge will be discarded, and that will produce a thin line in the resultant image.

**Step4 Double Thresholding :** Two values of a threshold are determined in the method of canny edge detection, T1, which represents the High Threshold, and T2, which represents the Low one. Pixels that have gray scale level values that are greater than T1 can be considered as strong pixels of the edge, and edge area is the result. Pixels that have grey scale level values < T2 can be considered as weak pixels of the edge, and result is the non-edge area. In the case where pixels have grey scale values that range from T1 to T2, result will be dependent upon neighboring pixels.

**Step5 Edge Tracking through the Hysteresis:** Edges which don't connect to very strong (i.e. strong) edge are eliminated in the ultimate output image. The strong edges can be also referred to as "Certain Edges" and are included in final edge image. The edges which aren't strong ones, however, are connected to strong edges have been included in output image.

The Canny approach includes the steps below, which are carried out in a sequential manner:

1- The image is low pass filtered by the use of a Gaussian Mask.
2- Calculating the vertical and horizontal gradients at every one of the pixel locations.
3- Calculating the magnitudes of the gradient at every one of the pixel locations.
4- Calculating a lower and higher threshold that has been based upon the gradient histogram of the whole image.
5- Suppression of the Non Maximal Strong (NMS) edges.
6- Carrying out hysteresis thresholding for the purpose of determining the map of the edge.

The Benefits of the Canny Edge Detection method are:

• Sufficient localization.
• Enhanced signal/noise ratio.
• Low spurious responses (Single-edge responses).
• carrying out methods of image processing on the reconfigurable hardware results in minimizing the time-to-market costs.
• More sufficient detection in the conditions of the noise.
• Simplified verification and debugging.
• Provides the ability for the rapid prototyping of the complicated methods.

Drawbacks of the Canny Edge Detector are:

• Complicated and time-consuming calculations, as canny edge detection methods includes lengthy steps of preprocessing and post-processing.
• It's difficult giving a general threshold which works efficiently on every image.
• False zero crossing.
• High memory requirements, resulting in high degree of latency.

*C) Sobel method*

The Sobel operator can be defined as an approach that uses the average value of the upper and lower, right and left neighboring pixels, and extremum value is obtained at the edge. Due to the weighted average, the operator only has the ability of getting the information of the edge, as well as the capability of the anti-noise, it is a normal operator of edge detection. Sx & Sy values of the Sobel operator may be computed according to the template of the convolution.

| | | |
|---|---|---|
| -1 | -2 | -1 |
| 0 | 0 | 0 |
| 1 | 2 | 1 |

(a) horizontal direction

| | | |
|---|---|---|
| -1 | 0 | 1 |
| -2 | 0 | 2 |
| -1 | 0 | 1 |

(b) vertical direction

Fig. 11: Templates of Sobel operator [37]

Based on the information, of the detection, an image may be considered as curved surface, the edge is where there is the highest change. The information of the edge includes 2 aspects: one of them is the edge direction, the isolated edges are insignificant, the solution is stitching the corresponding edge to line of the edge. A different one is the specific location, like the pixel coordinates [38]. Based on the differential knowledge, the information of the position on surface can be judged with the use of the derivation. The detection of the edges based upon the 1st order differential operator considers the image as a surface, and after that, it applies the gradient operator on the image f(x,y), at last vector field form.

$$\nabla f(x,y) = \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \tag{53}$$

The field of the vector presents 2 information aspects: one of them is all local gradient strength.

$$\|\nabla f(x,y)\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \tag{54}$$

The other one, is the local gradient orientation.

$$\nabla f(x,y) = arctan\left(\frac{\partial f}{\partial x} / \frac{\partial f}{\partial y}\right) \tag{55}$$

In the particular edge detection's experimental scheme, the individuals are usually concerned with the way of finding out ||f(x,y)||, which is the way for getting the information of the direction. And concerning the differential operator, the difference or other similar technology is typically utilized for the discrete, ultimately evolved to a form of the template process; it is quite simple. The generalized type of the use of the Sobel operator for the calculation of the value of the ||f(x,y)|| can be represented based on the equation below:

$$|T_1\, f(i,j)| \quad = \sqrt{(|S_x \ |2 + |S_y \ |2\,)} \tag{56}$$

Sx and Sy are as follows:

$S_x = f(i+1,j+1) + 2f(i+1,j) + f(i+1,j-1) - f(i-1,j+1) - 2f(i-1,j) - f(i-1,j-1)$ (57)

$S_y = f(i+1,j+1) + 2f(i,j+1) + f(i-1,j+1) - f(i-1,j-1) - 2f(i,j-1) - f(i+1,j-1)$ (58)

(a) (b)

Fig. 12: (a) original image , (b) image that has been detected using Sobel

Through the comparison of Fig. (12), the Sobel operator increases the central pixel's coefficient of weight in template floor, which is why, it has the ability of the detection of edges without any great mutations of the gradient. The experiments have shown: the conventional Sobel operator which is produced by the feature of the theory that the grey function has the ability of getting the maximum of the part at edge have the ability of achieving edge detection with the use of the grey weighted differences' summation results from 4 directions' neighboring pixels. Such algorithm type can be easily realized, and it has the ability of providing the precise information of the edge direction. It possesses a smoothing impact on the noise and is capable of the anti-noise, particularly in the case of being utilized in the large-size images. None-the-less, as can be seen from the simulation results, there are false edges, the image edge location precision is not good. In the case of being utilized in the large-size images, the computation amount is large, and edge isn't obvious; the calculations' amount increases with the increase on the direction [37].

*D) Prewitt method*

Prewitt can be defined as a discrete operator of differentiation, which computes a function of the approximation of the image intensity gradient at every one of the points in an image [39]. Prewitt operator result can correspond the norm gradient vector of that vector. Prewitt edge detector has been considered as one of the suitable ways for the estimation of the edge orientation and magnitude. Even though the edge detection of the differential gradient requires quite a lengthy computation for the estimation of the orientations from x & y-direction magnitudes, the edge detection of the compass gets orientation in a direct manner from kernel with maximal degree of the response. Prewitt operator has been limited to eight possibilities of the directions, none-the-less, the experience has shown that the most direct estimations of the orientation aren't considerably more precise. Such gradient-based detector of the edges is calculated in a 3x3 neighborhood for 8 orientations as can be seen from the Fig. (13). Each one of the 8 convolutional masks are computed [40].

Prewitt operator utilizes those equations as Sobel operators, except the constant c = 1. Which is why: it has to be noted that in contrast to Sobel operator, it doesn't concentrate upon the pixels which are nearer to the masks' center. Prewitt

operator can measure 2 elements. The vertical component of the edge is computed by the kernel Gx and horizontal component of the edge is computed by the kernel Gy.
|Gx| + |Gy| indicate the gradient intensity in current pixel[40]. One of the kernels simply represents the other, which is rotated by

| | | |
|---|---|---|
| -1 | 0 | +1 |
| -1 | 0 | +1 |
| -1 | 0 | +1 |

G(x)

| | | |
|---|---|---|
| +1 | +1 | +1 |
| 0 | 0 | 0 |
| -1 | -1 | -1 |

G(y)

Fig. 13: Masks that are utilized by the Prewitt operator [39].

Those may be combined for finding the gradient's absolute magnitude at every one of the points and Eqs. [39] :

$$|\nabla f| = \sqrt{G2\,X} + \sqrt{G2\,Y} \qquad (59)$$

$$|\nabla f| = |\sqrt{G2\,X}| + |\sqrt{G2\,Y}| \qquad (60)$$

The edge orientation angle (relative to the grid of the pixel) gives rise to spatial gradient has been represented as:

$$Angle\ of\ \nabla f = tan^{-1}\,(Gy/Gx) \qquad (61)$$

Computing image gradient has been based upon obtaining partial derivations of $\partial f / \partial x$ & $\partial f/\partial y$ at each one of the pixel locations. The 3x3 have been shown in the Fig.(14) are gray levels in an image neighborhood.
Those derivations may be carried out for a whole image with the use of 3x3 masks that have been illustrated in Fig. (14) with convolution procedure.

$$Gx= (z7+z8+z9)-(z1+z2+z3) \qquad (62)$$

$$Gy= (z3+z6+z9)-(z1+z4+z7) \qquad (63)$$

A slide change in those 2 equations utilizes a weight of two in central coefficient. A value of weight of two is utilized for the achievement of a level of smoothing through granting a higher level of significance to central point Fig. (13), which is referred to as prewitt operator, is utilized for the implementation of those 2 equations.

| | | |
|---|---|---|
| Z1 | Z2 | Z3 |
| Z4 | Z5 | Z6 |
| Z7 | Z8 | Z9 |

Fig. 14: A 3x3 image region [39].

The strengths of Prewitt operator is simple, detects edge and their orientation but the weakness of it is inaccurate and sensitive to noise [39].

*E) Roberts method*

The Roberts Cross operator carries out a quick to calculate, simple, 2D spatial gradient measure on image. Which is why, it can highlight the high spatial frequency areas that are usually corresponding to the edges. In the most common utilization, the operator input is a gray-scale image, like the output. The values of the Pixels at every one of the points in output are calculated absolute spatial gradient magnitude of input image at this point [40].
Roberts cross operator provides a simple gradient magnitude approximation:

$$G[f[i,j]]=|f[i,j]-f[i+1,j+1]|+|f[i+1,j]-f[i,j+1]| \qquad (64)$$

Utilizing the convolutional masks, the equation will be:

$$G[f[i,j]] = |G_x| + |G_y| \qquad (65)$$

Where Gx & Gy can be computed with the use of the masks below:

| | |
|---|---|
| 1 | 0 |
| 0 | -1 |

Gx

| | |
|---|---|
| 0 | -1 |
| 1 | 0 |

Gy

Fig. 15: Masks that have been utilized by Roberts Operator [41]

Fig. 16 shows the original image and it's the changing when used Sobel, Roberts, Prewitt, LOG and Canny methods.

(A)
(B)
(C)
(D)
(E)
(F)
(G)

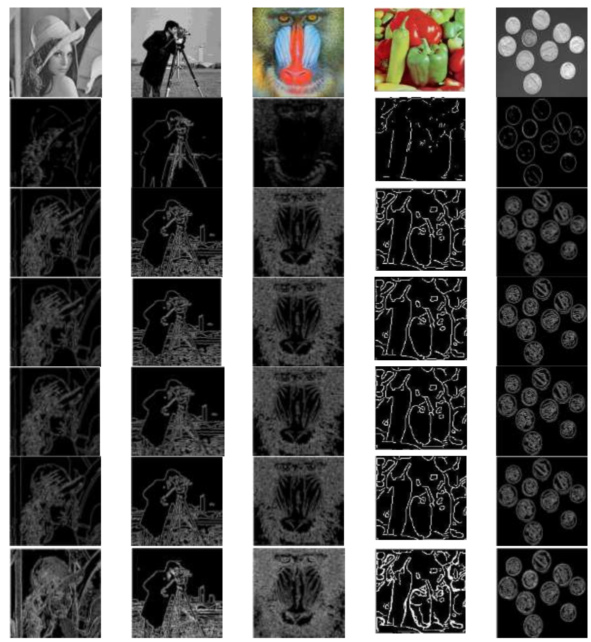


Fig. 16: (A) original images; (B): detected edges with the use of the Sobel approach; (C): detected edges with the use of the Prewitt approach; (D): obtained edges with the use of the Roberts approach; (E): obtained edges with the use of the LOG approaches; (F): obtained edges with the use of the Canny approach; (G): the obtained edges with the use of the suggested approach in [36].

## 3. Region Based Segmentation Techniques

Region based approaches of the segmentation are approaches which are utilized for the segmentation of images to a variety of the areas that have similar properties [3]. In the region based areas of segmentation have been produced through the association or the dissociation of the neighboring pixels. It operates according to the homogeneity principle, through the consideration of the fact that the neighbor pixels within an area possess similar properties and not similar to pixels in the other areas. Every one of the pixels is compared to the neighbor pixel for the check of the similarity, like the color, gray level, shape and texture. In the case where results are

positive then this specific pixel has been added to the pixel for the purpose of growing the area.
In the case where the complete image has been represented as area R, then for the segmentation compose it to n of the disjoint areas S1, S2, S3,……., Sn so that

$$S_i = S, \quad S_i \cap S_j = \emptyset, \ if\ i \neq j$$
$$Prop(S_i) = True, \quad f\ i = 1,2,3,\ldots,n \qquad (66)$$
$$Prop(S_i \cup S_j) = False, \quad f\ i = 1,2,3,\ldots,n$$

Where *prop*($S_i$ ) has been characterized according to the feature values through the area R. Those areas can be disjoint, connected and homogeneous [28]. There exist 2 main approaches that have been based upon that approach: the methods of region splitting and region growing and approaches of the merging [3][28].

3.1 Region growing methods
In this approach, the pixels within an area have been labeled by a distinctive lable that differs from other area labels. That approach may categorized additionally as SRG (i.e. Seeded Region Growing) and UsRG (i.e. Unseeded Region Growing), the latter is a semi-automatic approach, and the UsRG is a fully automated approach [28].
The approach of the region growing can be defined as a simple region-based approach of the image segmentation [42]. It has been categorized as well, as a pixel-based approach of the segmentation of the image, due to the fact that it includes selecting the points of the initial seeds. That segmentation method can examine the neighbor pixels of the initial “points of the seed” and specifies whether or not the neighborhoods of the pixel have to be added to that area. This procedure is iterated on, in a similar way as the general algorithms of the data clustering. The main disadvantage of the histogram-based detection of the region is the fact that the histograms do not provide any spatial knowledge (only grey level distributions). The approaches of the region growing usually provide highly good segmentation processes, corresponding well to observed edges [43].

3.1.1 Statistical Region Merging (SRM)
SRM can be defined as a method which is utilized for the segmentation of the images. This approach is utilized for the evaluation of values in regional span and combined according to the criteria of merging, which results in shorter list. Some of the beneficial examples create a set of the generations in one of the populations, or in the image processing, combining several neighbor pixels that have been based upon the shades falling in a certain value of the threshold [44].

3.1.2 Seeded Region Growing (SRG)
SRG is It is proposed by R. Adam [45]. It is a semi-automatic approach of merge type [2][28]. The SRG has been known for being rapid and robust, and it’s free from the parameters of the tuning. This procedure begins with the selection of a seed pixel in an image. The suitable seed option is highly necessary in this approach, due to the fact that it’s focused on the general quality of the segmentation.
The main SRG steps can be summarized below[28]:

1- Selecting the pixel of the seed in the image for the purpose of starting the process of the segmentation.
2- Deciding the criterions for region growing.
3- Including the pixel in the area in the case where it’s eight – connected to a minimum of 1 pixel in the area.
4- Each region is labeled, following the testing of every pixel for the allocation.
5- Merging the areas in the case where 2 separate areas have been assigned the identical labels.

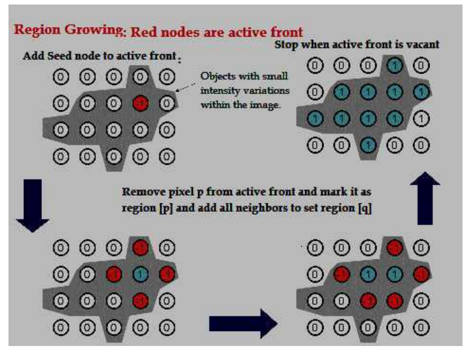



Fig. 17: .SRG [28]

3.1.3 Unseeded Region Growing
This approach UsRG can be described as a fully automated approach, it has been based upon the similarities of the pixel in an area. This approach has flexibility, it is fully automated and it is dependent upon the parameters of the tuning [28].
The main UsRG method steps can be summarized as:
1- Initializing the process of the segmentation with the region *S1* that contains 1 pixel and then produces *S1, S2, …, Sn* areas following the completion.
2- For the allocation of the pixel, test pixel difference measure with the average statistics value are implemented.
3- Allocating a pixel to a certain area, such as *Si*, in the case where the value of the difference is < a specific threshold value; in the opposite case, this pixel is allocated to a new area *Sj*.
4- The abovementioned steps are repeated for the rest of the pixels.

3.2 Region splitting and merging methods

Those methods have been presented by B. Penetal [45], they operate based on the quad-tree with a fundamental goal of distinguishing image homogeneity, where the whole image can be viewed as a complete area and, after that, this image is divided to 4 different quadrant areas that have been based upon specific predetermined criterions [28] . Fig (18) shows this approach.
The main steps of this approach can be summarized as [28]:
1- Defining the condition of the homogeneity.
2- Creating a pyramid data structure for the image.
3- Forming a quad-tree with numbers for the levels and forming a node fragment number.
4- Repeating those steps to the point where there aren’t any more possible operations of merging nor splitting.

TABLE 2
SUMMARY OF THE BENEFITS AND THE DRAWBACKS OF THE IMAGE SEGMENTATION REGION GROWING TECHNIQUES [42][43]

| • Advantages | • Disadvantages |
|---|---|
| • Those approaches have the ability of correctly separating areas with identical characteristics.<br>• Those approaches have the ability of providing original images that have obvious edges with sufficient results of the segmentation.<br>• The idea is simply a limited amount of the seed points are required for the representation of a characteristics are desired, after that, the region will be grown. | • Those methods are computationally expensive.<br>• They are local approaches without any global views of the problem.<br>• Noise- Sensitive.<br>• Unless an image had undergone the application of a thresholding function, a continual path of the points that are associated with the color can exist, which results in connecting any 2 image points. |

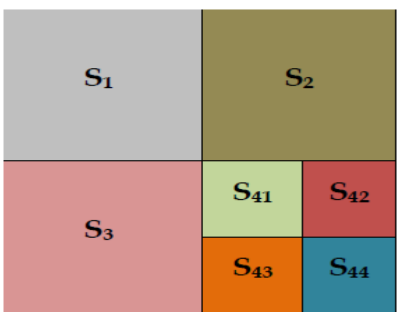


Fig. 18: The approach of the Region Splitting and Merging method [28]

## 4. Clustering Techniques

The cluster based approaches can be defined as the methods that are utilized for the segmentation of an image to clusters that have pixels with the same properties. Clustering is utilized for dividing data elements to clusters in a way that the same cluster elements are more like one another compared to the others. There exist 2 main clustering method classes, and those are: partition-based and hierarchical methods. The latter type is based upon the tree concept. Partitioning based approaches utilize the approaches of optimization in an iterative manner for the minimization of an objective function. Besides those 2 approaches there is a variety of the algorithms for clustering. Clustering can be classified to 2 main types [3][46-47].

4.1 Hard Clustering

This types can be defined as a simple method of clustering, which operates by dividing an image to a group of the clusters, in a way that 1 pixel may only be part of one single cluster, i.e. it may be considered that every one of the pixels may be part of only 1 cluster. Those approaches utilize the functions of the membership that have either 0 or 1 values. In other words, one certain pixel may either be part of a certain cluster or otherwise. One of the hard clustering based approach examples is the k-means clustering based approach, which has been referred to as the HCM. [3].

*K-Means Clustering*

K-Means is a very simple unsupervised algorithm of learning, which has the ability of solving the common problems of the clustering. K-Means is good technique which used for segmenting image. In that approach, initially, centers are calculated, and after that, every one of the pixels is given to its closest centre. It puts an emphasis upon the maximization of intracluster similarity as well as the minimization of the intercluster equality.

The process operates through an easy and simple approach for the classification of a certain dataset by a specific amount of the clusters (assuming a k number of the clusters) fixed apriori. The fundamental concept is defining k of the centroids, 1 for every one of the clusters. Those centroids must be allocated in cunning manner due to the fact that different locations cause different results. Which is why, the optimal option is placing them as far away from one another as possible. The following step will be taking every one of the points that belong to a specified dataset and associate it with the closest one of the centroids. In the case where none of the points is pending, the initial step is done and early grouping will be performed. Here, the necessity for the recalculation of k new cluster centroid values that result from the preceding step. Following the calculation of those k new centroid values, a new binding must be performed between identical points of the dataset and the new nearest centroid. A loop was produced. Due to that loop, it can be noticed that those k centroids alter their positions gradually, to the point where there aren't any more changes occurring, i.e. the centroids don't move any more. In the end, this approach has the aim to minimize an objective function; in the present case a function of the summation squared error. The objective function will be represented as:

$$J= \sum_{j=1}^{k} \sum_{i=}^{n} ||x_i^{(j)} - c_j||^2 \qquad (67)$$

$\| x_i(j) - c_j \|^2$ represents a specified measure of the distance between the data point xi (j) and the center of the cluster Cj, indicates the distances of n data points to their corresponding centers of the cluster. This approach includes the steps below [47] :

General steps in K-mean algorithm are:

1- Placing K points in a space which is represented by the clustered objects. Those points are the centroids of the initial group.

2- Assigning every one of the objects to a group which includes the nearest one of the centroids.

3- After assigning all the objects, the K centroids' positions will be recalculated.

4- The Step2 and Step3 are repeated to the point where centroids do not change anymore. Which results in the production of the separation of objects to groups, and the metric will be minimized and computed from it.

4.2 Soft clustering

This clustering type can be considered as the most natural clustering due to the fact that in the real life, the precise division is impossible as a result of the presence of the noise. Which is why, the methods of the soft clustering are most beneficial for the segmentation of images where the division isn't strict. One of the examples of this method is the fuzzy c-means approach of clustering[3].

*Fuzzy C-Means Clustering (FCM)*

In this method, the pixels are divided to clusters according to the partial memberships, in other words, one of the pixels may be part of several clusters and such belonging degree is characterized by the values of membership. This approach has higher flexibility compared to other approaches [46].

The FCM performs the partitioning of a group of the n objects x {x , x ,..., xn} in $R^d$ dimensional space to c(1 < c < n) fuzzy clusters with y = { $y_1$, $y_2$ , $y_3$ ,..., $y_c$} centroids or centers of the clusters [48]. FCM of the objects can be identified by fuzzy matrix μ that has c columns and n rows, where n represents the amount of the data objects and c represents the amount of the clusters. μij represents an element which is in i-th row and j-th column, μ , stands for the level of the membership or association function of i-$^{th}$ object with j-$^{th}$ cluster. FCM algorithm's objective function is minimizing the equation below [46].

$$J_m = \sum_{j=1}^{c} \sum_{i=1}^{n} u_{ij}^{m} \; d_{ij} \tag{68}$$

Where

$$d_{ij} = \| x_{i-} y_j \| \tag{69}$$

The $y_j$ represents $j^{th}$ cluster's centroid, which can be calculated from the equation below:

$$y_j = \frac{\sum_{i=1}^{n} u_{ij}^{m} \; x_i}{\sum_{i=1}^{n} u_{ij}^{m}} \tag{70}$$

FCM algorithm is one of the iterative algorithms and may be represented by the following steps [46]:

1- Select m(m > 1); initializing the values of the function of membership $\mu_{ij}$ ,
 i = 1,2,...,n; j = 1,2,...,c.

2- Computing the centers of the clusters $y_j$ , j = 1,2,..., ,c . based on eq.70

3- Computing the Euclidian distance $d_{ij}$ , i=1,2…,n

TABLE 3
SUMMARY OF THE BENEFITS AND THE DRAWBACKS OF THE IMAGE SEGMENTATION TECHNIQUES CLUSTERING TECHNIQUES [3]

| Advantages | Disadvantages |
|---|---|
| • It utilizes the partial membership, which is thus, more beneficial for solving the real tasks. | • The determination of the function of membership is a complex process. |

**5. Deep Neural Network Techniques**

5.1 Convolution Neural Network

CNNs have been identified as amongst the most beneficial and commonly utilized architectures in deep learning community, particularly for the tasks that are related to the computer vision. CNNs have been suggested initially by Fukushima in one of his seminal papers on the "Neocognitron" [49], according to the visual cortex's hierarchical receptive field model.

The Convolutional Neural Network (CNN) may be utilized to learn the ways for segmenting the images. The CNNs perform the extraction of the features in a direct way from the pixel images that have minimum pre-processing.

The typical CNN usually includes 3 layer type: a) convolution layers, in which a filter (or a kernel) of the weights undergoes convolution for the purpose of extracting the features; b) non-linear layers, applying a function of activation on the feature maps (typically element-by-element) for the purpose of enabling to model nonlinear functions by networks; c) pooling layers that substitute a small neighbourhood of the feature map by a statistical information (max, mean, and others.) around that the neighbourhood and decrease the spatial resolution. The layer units are connected locally; which means that every one of the units can receive weighted input values from a small neighbourhood, which is referred to as receptive field, of the units in the preceding layer. Through the stacking of the layers for forming pyramids with multiple resolutions, the layers of the higher-levels learn the features from considerably wider receptive fields. The fundamental CNN's computational strong point is the fact that each receptive field in the layer shares weight with the others, which results in a considerably fewer parameters compared to the fully-connected NNs. [50].

CNNs can be characterized as directed graph. Nodes are corresponding to image pixels. Edges are corresponding to the filters. CNNs are a multi-layer perceptron (MLP) that have been particularly designed for the recognition of 2-D shapes that have a high invariance level to translations, skewing, scaling, or any other distortion forms. The architecture has been shown in Fig (19) [51].

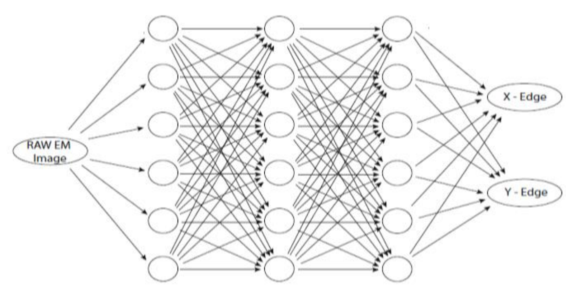



Fig 19: Convolutional neural network [51]

Using the weight sharing can allow the parallel implementation of the CNN. Reducing the number of the free parameters can be accomplished by using the process of the weight sharing. As a result, the machine learning capacity is decreased, which results consequently in improving the generalization ability of the machine. Adjusting the network's free parameters are done through the use of back propagation's stochastic mode. CNN learning has twice the advantages. With a prior information about images, it has the ability for learning the complicated, non-linear, high dimensional mapping. In addition to that, it has the ability for learning the bias levels and syntactic weight values [51].

### 5.2 Recurrent Neural Networks (RNNs)

The RNNs have been commonly utilized for the processing of the sequential data, like the text, speech, time-series, and videos, in which the data at certain position/ time is dependent upon data that have been previously encountered [50]. At every one of the time-stamps, the model can collect inputs from current time $X_i$ and hidden state from an earlier process, $h_{i-1}$, and output a new hidden state and a needed value Fig. (20).

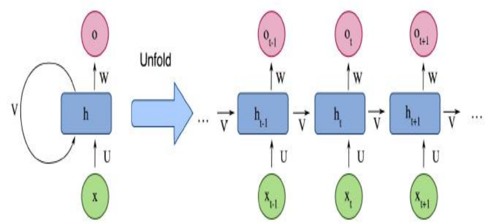



Fig. 20: Simple RNN Architecture [50]

The RNNs are usually problematic in the case of the long sequences, due to the fact that they are not capable of capturing the long-term dependency levels in a wide range of the real-world implementations (even though they do not show any theoretical limitations in that matter) and they are usually suffering from the problems of exploding or vanishing of the gradients. None-the-less, an RNN type, which has been referred to as the Long Short Term Memory (LSTM) has been modeled for avoiding those problems. The architecture of the LSTM, which has been illustrated in Fig.(21), comprises 3 gates (which are input, forget and output gates) that are responsible for the regulation of the information flow to and out of a memory cell that is responsible for storing the values over random intervals of time.

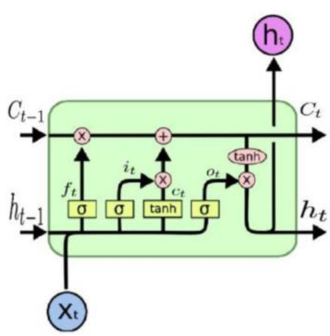



Fig. 21: Typical LSTM module Architecture [50].

The correlation between the input, hidden states, and various gates can be represented as:

$$
\begin{gathered}
f_t = \sigma\,(w^{(f)xt} + U^{(f)} + b^{(f)}\,, \\
i_t = \sigma\,(w^{(i)xt} + U^{(f)ht-1} + b^{(i)}\,, \\
o_t = \sigma\,(w^{(o)xt} + U^{(o)ht-1} + b^{(f)}\,, \\
c_t = f_t \odot c_{t-1} + i_t \odot tanh\,(w^{(c)}\,xt + U^{(c)}\,h_{t-1} + b^{(c)} \\
h_t = o_t \odot \tanh(c_t)
\end{gathered}
\quad (71)
$$

where $xt \in \mathrm{R}^d$ represents input at the time-step $t$, and $d$ represents the dimension of the feature for every one of the words, $\sigma$ represents element-wise sigmoid function (for the mapping of values in the range of [0 , 1]), $\odot$ represents element-wise product, and ct represents memory cell which has been designed for the purpose of lowering the risks of the gradient's exploding/vanishing (as a result, providing the ability to learn of dependencies over the larger time periods, which is feasible with the conventional RNNs). A forget gate, *ft*, has been aimed for resetting memory cell. *ot* and *it* respectively represent the output and input gates, and in fact regulate the memory cell's output and input [50].

### 5.3 Encoder-Decoder and Auto-Encoder Models

The models of the Encoder-Decoder can be defined as a group of the models that learn the mapping of the data-points from a domain of the inputs to a domain of the outputs through the use of a 2-stage network: an encoder, which is characterized with a function of encoding $z = f(x)$, results in compressing input to a latent-space representation; the decoder, $y = g(z)$, has the aim of predicting the value of the output from the representation of the latent space. In this case, the latent representation in fact indicates a representation of the features (i.e. vectors) that has the ability of capturing the input's underlying semantic information, which has been beneficial to predict the output. Those models have gained extreme popularity in the image-to-image problems of the translation, in addition to the NLP sequence models. Fig. (22) depicts a diagram of the simple model of encoder-decoder. Those models are typically trained through the minimization of the loss of reconstruction $L(y,\hat{}y)$, measuring the difference values between the output $y$ of the ground-truth and the successive reconstruction $\hat{}y$. in this case, the output may refer to an improved image version (like in the super-resolution or de-blurring of the image), or a map of segmentation [50].

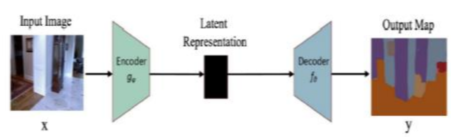



Fig.22: The simple encoder-decoder model architecture [50]

The auto-encoders can be defined as special encoder-decoder model case, where the output is the same as the input. Numerous auto-encoder variations were suggested. One of the most common ones is SDAE (i.e. stacked denoising auto-encoder), stacking numerous auto-encoders and utilizing those for the purposes of the image de-noising. One more common type is VAE (i.e. the variational auto-encoder), imposing an earlier distribution on latent representations. The VAEs have the ability of generating realistic samples from a certain distribution of the data. A different type is the adversarial auto-encoders, introducing an adversarial loss on latent representation for the purpose of encouraging them for approximating a prior distribution[50].

### 5.4 Support Vector Machines

SVMs have a highly important impact on the approaches of image segmentation, particularly, in the processing of the medical images. There are numerous modified approaches that have been based upon the SVMs such as modified SVM approaches in medical image segmentation. There are numerous modified approaches of the SVMs that have been presented during the past 5 years, particularly the modified SVMs in the segmentation of the medical images [52].

In this review we will display how to used modified SVM approaches were capable of achieving good effects on the methods of image segmentation or the threshold SVM medical image segmentation. We will present an overview of the modified approaches of the SVM from the past decade, like the Least squares SVM (LS-SVM) as well as the genetic algorithm-SVM (GA-SVM) .

- Least Squares Vector Machine (LS-SVM) :

Yang[53] porposed a sufficient color image segmentation method that has been based upon the classification of the pixels by using the LS-SVMs. The pixel-level color features as well as the texture features have been utilized as LS-SVM model (i.e. classifier) input, and LS-SVM classifier has been trained through the selection of training samples with the Arimoto method of entropy thresholding. The LS-SVM is an approach which is utilized for improving the SVM through the modification of its function or parameters. It makes no changes to SVM's concepts and theory. It has the ability for improving segmentation accuracy. None-the-less, the drawback of that approach is that it still unable of achieving the precise segmentation. LS-SVM is an SVM approach which has been altered through changing the SVM theory [53].

- Genetic Algorithm-Support Vector Machine (GA-SVM):

Zhang [54] had researched the GA-SVM for the determination of SVM hyper-parameters (c, g).

Those modified approaches of the SVMs methods have the objective to obtain sufficient segmentation of image, searching for the parameters for determining SVM hyper-parameters (c, g) for the purpose of achieving the automated training of the data selection as well as the automated extraction of the features. GA-SVM is a combination of two methods, which are; GA and SVMs. The approach of combining 2 methods results in the improvement of SVM's theory and principle. Its benefit lies in the fact that it is insensitive to the noise, which results in the fact that it does not have the ability for accurately segmenting the images [52]. GA-SVM is one of the updated methods of the SVMs that has been based upon the combinations with the genetic algorithms. Such SVM approach is modified with the use of other approaches [54].

## 6. Hybrid Techniques

In general the hybrid approaches can be defined as combination of 2 computational approaches or more, providing more benefits compared to the single approaches and enhance the analysis of the data. The hybrid method of the segmentation performs the integration of the region-based and the boundary-based approaches of segmentation amplifying the strength at the same time as reducing the weaknesses of the two methods.

We'll show some hybrid methods like Evolutionary Approaches , Fuzzy logic and Swarm Intelligent that are used for segmentation.

### 6.1 Evolutionary Approaches

The evolutionary algorithms are utilized in a wide range of the applications of engineering for optimizing problems which are usually hard to solve with the use of the traditional approaches. One of those problems is the segmentation of image. This task is utilized for the extraction of objects (contours) from the images for the purpose of creating sensible image representation [55].

This type of computation is a part of the Bio-inspired computations and AI field, providing approaches that are utilized for the optimization of the problems of the continuous space [55]. It's one of the suitable tools for the flexible and effective segmentation of the time series [56].

### 6.2 Fuzzy logic

In general Fuzzy logic can be defined as one of the approaches for the computations, based upon the "truth degrees" instead of usual "true or false" (i.e. 1 or 0), which has been referred to as the Boolean logic, which modern computers have been based upon. It can be helpful in seeing the fuzzy logic as one of the ways for the reasoning that in fact, works and the Boolean or binary logic is simply a special case of that.

Fuzzy logic theory (FL) has been an increasing use for solving the problems in image processing such as segmentation problems. In [57] software system developed a by using combination of fuzzy and neural network techniques to implement an image segmentation algorithm. So fuzzy logic theory with neural network techniques can be combined and produced hybrid techniques used for segmentation.

On the other hand fuzzy logic with Genetic algorithms (GAs) have a significant impact on solving numerous problems in the image processing and pattern recognition. In [58] an FL and GAs hybrid approach has been proposed for the purpose of being utilized for segmentation and extraction of the features from the intensity and range images.

### 6.3 Swarm Intelligent (SI)

Swarm intelligence can be defined as a collective behavior of the self-organized, decentralized systems, artificial or natural. In the present day, SI-based approaches are utilized in numerous applications and presented good efficiency [59] . In [60] two swarm intelligent approaches for image thresholding are display these are artificial bee colony method and PSO method have been utilized for the purpose of dealing with it.

*A. Artificial bee colony (ABC) approach*

The ABC algorithm includes 3 fundamental elements, which are: nectar-amounts, food source positions, and several honey-bee classes[60-61].In [62] ABC is used with MRF (Markov Random Fields) in for the purpose of obtaining the best segmentation for brain image segmentation. ABC is robustness, particularly in the discrete multi-variable optimization tasks. So that is used with MRF for the purpose of improving the quality of the image segmentation [62]. Research has proven that APPS is good for Image Segmentation [60]. In [63] an (ABC) optimization algorithm has been proposed, which has been utilized as one of the clustering techniques for the segmentation of liver in CT images.

*B. Particle Swarm Optimization (PSO) approach*

PSO is one of the most powerful approaches of the optimization, inspired by the social behavior animals that live or move in a swarm like fish schooling or bird flocking. In [64] the POS has been utilized to segment the medical images and solve problems occurring in the case of the medical image segmentation.

The PSO is one of the new classes of the metaheuristics that have been presented by Kennedy and Eberhart in 1995 [65]. This approach has been studied by several other studies [66-70]. This algorithm has been inspired by the social behavior of the animals that move in swarms as schooling fish or flocking birds. The swarm efficiency is higher compared to the summation of the efficiency of its parts.

This approach of optimization has been based upon collaborating between the individuals. A swarm individual only knows the speed and the position of the nearest neighbors. Every one of the particles adjusts its behavior according to its experience as well as its neighbors' experiences for building a problem solution. By the use of the simple rules of displacement (in the space of the solution), particles may converge in a gradual manner toward the problem solution.

Formally, every one of the particles i has a location $x_i(t)$ at time t in the possible solution space changing at the time $t+1$ by the velocity $v_i(t)$, which is affected by the optimal position $y_i(t)$ that is visited by itself (i.e. its own experience) and zi t represents the optimal position that is visited by every particle (which can be referred to as the global optimal). The positions have been measured with a fitness function *f* that is dependent upon the problem of the optimization and K represents the dimension of the space.

$$xi(t)=(xi1,xi2,\ldots, xij ,\ldots , xik ) \quad (72)$$
$$vi(t)=(vi1,vi2,\ldots, vij ,\ldots , vik ) \quad (73)$$
$$yi(t)=(yi1,yi2,\ldots, yij ,\ldots , yik ) \quad (74)$$
$$zi(t)=(zi1,zi2,\ldots, zij ,\ldots , zik ) \quad (75)$$

$y_i$ is updated time based upon the following equation:

$$y_i(t+1)=\left(\begin{cases} y_{i(t)} \text{ if } f(x_i\ (t+1)\ge f(y_i\ (t)) \\ x_{i(t+1)} \text{ if } f(x_i\ (t+1)\ge f(y_i\ (t)) \end{cases}\right. \quad (76)$$

The best position that has been visited by every particle until the time t, z t will be computed at time t with the equation below:

$$Z(t) \in (y1(t),y2(t),\ldots,ys(t) = \min\{f(y1(t)),f(y2(t)),\ldots,f(ys(t))\} \quad (77)$$

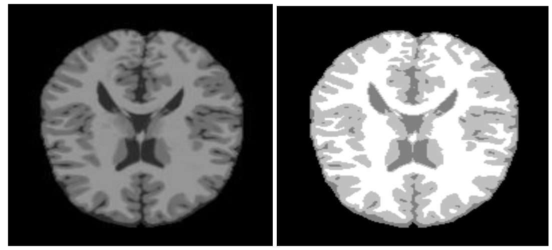
where s represents the number of particles (the swarm size). The velocity $vi\ t = (vi1(t), vi2(t), \ldots , vij\ (t), \ldots , viK(t))$ of the particle $i$ at time $t$ is updated with the use if the equation below:

$$vij\ t+1 = w*vij\ t + c1*r1j*\ yij\ t - xij\ t + c2*r2j*\ zj\ t - xij\ t \quad (78)$$

$w$ has been referred to as inertia weight; $c1$ & $c2$ represent the constants of acceleration; $r1j$ & $r2j$ represent the random variables in the fange from 0 to 1; velocity $vij$ is dependent upon the value of the V$max$ for ensuring the convergence. The location $xi$ of particle $i$ is updated with the use of the equation below:

$$Xi(t+1) = xi(t) + vi\ (t+1) \quad (79)$$

Fig. (23) shows MRI scan and segmentation image by using POS.



(a) (b)

Fig. 23: (a) MRI scan (b) PSO

## III. Conclusion

In this survey, we carried out a detailed study of state-of-the-art of a number of the algorithms that were suggested earlier for the improvement of the segmentation techniques for image and proposed a hierarchical taxonomy for classifying these techniques. We surveyed different approaches, which were classified into six categories which are: thresholding segmentation approaches, edge based segmentation approaches, region based segmentation approaches, clustering approaches, deep neural network approaches and hybrid approaches. Each of these approaches employ to segment image. We discussed the sub classification of each six approaches separately and their advantages and disadvantages.

This paper has also reviewed the study on a variety of the research methods that have been implemented for the segmentation of images and a variety of the research issues in that study area. This research has the aim of providing a simple guide for researchers who have conducted their researches on image segmentation.

## Conflict of Interest

The authors have no conflict of relevant interest to this article.

## References

[1] R. B. Fisher, T. P. Breckon, K. Dawson-Howe, A. Fitzgibbon, C. Robertson, E. Trucco, and C. K. Williams, "Dictionary of computer vision and image processing," NIST Special Publication, vol. 7, no. 4, 2013, pp. 22.

[2] D. Venkateshwar Rao, S. Patil, N. Anne Babu and V. Muthukumar , "Implementation and evaluation of image processing algorithms on reconfigurable architecture using C-based hardware descriptive languages," International Journal of Theoretical and Applied Computer Sciences, vol. 1, no.1 2006, 9-34.

[3] D. Kaur and Y. Kaur, "Various image segmentation techniques: A Review," International Journal of Computer

Science and Mobile Computing, vol. 3, no. 5, 2014, pp. 809-814.
[4] W.-X. Kang, Q.-Q. Yang, and R.-P. Liang, "The comparative research on image segmentation algorithms," [in Proceedings of the 1st International Workshop on Education Technology and Computer Science, IEEE, pp. 703-707], 2009.
[5] R. F. Yaser, "Multiobject extraction from color image, " PhD Thesis, College of Science, University of Al-Mustansirya, 2016.
[6] R. C. Gonzalez and R. E. Woods, "Digital image processing," Prentice Hall, Upper Saddle River, New Jersey 07458, 2002.
[7] Deepa M, "Wavelet and curvelet based thresholding techniques for image denoising," International Journal of Advanced Research in Computer Science and Electronics Engineering, vol. 1, no. 10, 2012, pp. 77-81.
[8] V. Zharkova, "Artificial intelligence in recognition and classification of astrophysical and medical images," vol. 46. Springer Science & Business Media, 2007.
[9] P. K. Sahoo, S. Soltani, and A. K. Wong, "A survey of thresholding techniques," Computer Vision, Graphics, and Image Processing, Elsevier, vol. 41, no. 2, 1988, pp. 233-260.
[10] F. Deravi and S. K. Pal, "Grey level thresholding using second-order statistics," Pattern Recognition Letters, Elsevier, vol. 1, no. 5-6, 1983, pp. 417-422.
[11] A. Rosenfeld and P. De La Torre, "Histogram concavity analysis as an aid in threshold selection," IEEE Transactions on Systems, Man, and Cybernetics, no. 2, 1983, pp. 231-235.
[12] D. Rutovitz, "An algorithm for in-line generation of a convex cover," Computer Graphics and Image Processing, Elsevier, vol. 4, no. 1, 1975, pp. 74-78.
[13] K. Bhargavi and S. Jyothi, "A survey on threshold based segmentation technique in image processing," International Journal of Innovative Research and Development, vol. 3, no. 12, 2014, pp. 234-239.
[14] T. Kalaiselvi, P. Nagaraja, and V. Indhu, "A comparative study on thresholding techniques for gray image binarization," International Journal of Advanced Research in Computer Science, vol. 8, no. 7, 2017, pp. 1168-1172.
[15] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, 1979, pp. 62-66.
[16] G. Dougherty, "Digital image processing for medical applications," Cambridge University Press, 2009.
[17] M. Athimethphat, "A review on global binarization algorithms for degraded document images," AU J.T, vol. 14, no. 3, 2011, pp. 188-195.
[18] I. Pitas, "Digital image processing algorithms and applications, " John Wiley and Sons, 2000. .
[19] W. N. Jasim, "A human activities recognition system for the healthcare services, " M Sc. Thesis, College of Science, University of Basrah, 2018.
[20] R. Firdousi and S. Parveen, "Local thresholding techniques in image binarization," International Journal of Engineering and Computer Science, vol. 3, no. 3, 2014, pp. 4062-4065.
[21] N. R. Pal, E. S. Corchado, L. T. Kóczy, and V. Kreinovich, "Advances in intelligent systems and computing," [in Proceedings of the 2nd International Conference on Data Engineering and Communication Technology, 2017.
[22] O. Imocha Singh, Tejmani Sinam, O. James and T.Romen Singh, "Local contrast and mean based thresholding technique in image binarization," International Journal of Computer Applications, vol. 51, no.6, 2012, pp. 0975-8887.
[23] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," Journal of Electronic Imaging, vol. 13, no. 1, 2004, pp. 146-165.
[24] G. Leedham, Y. Chen, K. Takru, J. H. N. Tan, and L. Mian, "Comparison of some thresholding algorithms for text/background segmentation in difficult document images," [in Proceedings of the 7th International Conference on Document Analysis and Recognition, Citeseer, pp. 859-864, 2003].
[25] N. Kaur and R. Kaur, "A review on various methods of image thresholding," International Journal on Computer Science and Engineering, vol. 3, no. 10, 2011, pp. 3441-3443.
[26] J. S. Weszka, "A survey of threshold selection techniques," Computer Graphics and Image Processing, Elsevier, vol. 7, no. 2, 1978, pp. 259-265.
[27] R. Kasturi, and M.M. Trivedi, "Image analysis applications," Marcel Dekker, Inc, 1990.
[28] M. Jogendra Kumar, Dr. GVS Raj Kumar and R. Vijay Kumar Reddy, "Review on image segmentation techniques," International Journal of Scientific Research Engineering & Technology, ISSN, 2014, pp.2278-0882.
[29] R. Kirti and A. Bhatnagar, "Image segmentation using canny edge detection technique," International Journal of Techno-Management Research, vol. 04, no.04, 2017, pp. 8-14.
[30] X. Geets, J. A. Lee, A. Bol, M. Lonneux, and V. Grégoire, "A gradient-based method for segmenting FDG-PET images: methodology and validation," European Journal of Nuclear Medicine and Molecular Imaging, Springer, vol. 34, no. 9, 2007, pp. 1427-1438.
[31] S. Tara, R. B. B. Reddy, G. Ramesh, and K. S. Sandeep, "Various image segmentation methods based on partial differential equation-a survey," International. Journal of Advances in Computer, Electrical & Electronics Eng., vol. 3, no. 1, 2014, pp. 183-186.
[32] D. Cremers, "Dynamical statistical shape priors for level set-based tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 8, 2006, pp. 1262-1273.
[33] D. MARR and E. HILDRET, "Theory of edge detection," [in Proceedings of the Royal Society of London. Series B. Biological Sciences, pp. 187-217, 1980].
[34] J. A. Sethian, "Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science," vol. 3: Cambridge University Press, 1999

[35] J. Canny, 1986, "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, 1986, pp. 679-698.

[36] W. Rong, Z. Li, W. Zhang and L. Sun, "An improved canny edge detection algorithm," [IEEE international conference on mechatronics and automation, pp. 577-582, 2014]

[37] Y. Yao, "Image segmentation based on sobel edge detection," [in Proceedings of the 5th International Conference on Advanced Materials and Computer Science, Atlantis Press, pp. 141-144, 2016].

[38] F. Truchetet, F. Nicolier, and O. Laligant, "Subpixel edge detection for dimensional control by artificial vision," Journal of Electronic Imaging, vol. 10, no. 1, 2001, pp. 234-239.

[39] Priyam, Diganta Dey, Shreya, and Dipanjan Polley, "Edge detection by using canny and prewitt," International Journal of Scientific & Engineering Research, vol. 7, no. 4, 2016, pp. 251-254.

[40] S. S. Al-Amri, N. Kalyankar, and S. Khamitkar, "Image segmentation by using edge detection," International Journal on Computer Science and Engineering, Citeseer, vol. 2, no. 3, 2010, pp. 804-807.

[41] G. Shrivakshan and C. Chandrasekar, "A comparison of various edge detection techniques used in image processing," International Journal of Computer Science Issues, Citeseer, vol. 9, no. 5, 2012, pp. 269-276.

[42] D. Oliva, M. Abd Elaziz, and S. Hinojosa, "Metaheuristic algorithms for image segmentation: theory and applications," Springer, 2019.

[43] S. Kamdi and R. Krishna, "Image segmentation and region growing algorithm," International Journal of Computer Technology and Electronics Engineering, vol. 2, no. 1, 2012, pp. 103-107.

[44] R. Adams and L. Bischof, "Seeded region growing," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 6, 1994, pp. 641-647.

[45] J. G. Rodriguez , "Advancements in computer vision and image processing," IGI Global, 2018.

[46] M. Yambal and H. Gupta, "Image segmentation using fuzzy C means clustering: A survey," International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, no. 7, 2013, pp. 2927-2929.

[47] V. K. Dehariya, S. K. Shrivastava, and R. Jain, "Clustering of image data set using k-means and fuzzy k-means algorithms," [in Proceedings of the International Conference on Computational Intelligence and Communication Networks, IEEE, 2010, pp. 386-391.]

[48] W.-C. Lin, E. C.-K. Tsao, and C.-T. Chen, "Constraint satisfaction neural networks for image segmentation," Pattern Recognition, Elsevier, vol. 25, no. 7, 1992, pp. 679-693.

[49] K. Fukushima, "Neocognitron A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by shift in position," Biological Cybernetics, Springer-Verlag, 1980, pp. 193-202.

[50] G. Wang, et al., "Interactive medical image segmentation using deep learning with image-specific fine tuning," Interactive medical image segmentation using deep learning with image-specific fine tuning. IEEE transactions on medical imaging, vol. 37, no.7, 2018, 1562-1573

[51] V. Govindan, "Convolutional neural network based segmentation," [in Proceedings of the International Conference on Information Processing, Springer, 2011, pp. 190-197].

[52] T. C. Zhang, J. Yang, J. P. Zhang, J. Zhang, "SVM methods in image segmentation," [in Proceedings of the 6th International Conference on Advanced Collaborative Networks, Systems and Applications, IARIA,, pp. 62-65, 2016].

[53] TH.-Y. Yang, X.-Y. Wang, Q.-Y. Wang, and X.-J. Zhang, "LS-SVM based image segmentation using color and texture information," Journal of Visual Communication and Image Representation, Elsevier, vol. 23, no. 7, 2012, pp. 1095-1112.

[54] Z. Zhang, J. Yang, Y. Wang, D. Dou, and W. Xia, "Ash content prediction of coarse coal by image analysis and GA-SVM," Powder technology, Elsevier, vol. 268, 2014, pp. 429-435.

[55] K. Mozdren, T. Burianek, J. Platos, and V. Snášel, "Evolutionary techniques for image segmentation," [in Proceedings of the 5th International Conference on Innovations in Bio-Inspired Computing and Applications IBICA, Springer, pp. 291-300, 2014].

[56] J. Yu, J. Yin, D. Zhou, and J. Zhang, "A pattern distance-based evolutionary approach to time series segmentation," [in Proceedings of the Intelligent Control and Automation, Springer, pp. 797-802, 2006].

[57] B. Karasulu and S. Balli , "Image segmentation using fuzzy logic, neural networks and genetic algorithms: survey and trends," Machine Graphics & Vision International Journal, vol. 19, no. 4, 2010, pp.367-409.

[58] M. Abdulghafour, "Image segmentation using fuzzy logic and genetic algorithms," Journal of WSCG, vol. 11, no. 1, 2003, pp. 1-8.

[59] I. Brajevic, M. Tuba, and M. Subotic, "Performance of the improved artificial bee colony algorithm on standard engineering constrained problems," [in Proceedings of the International Journal of Mathematics and Computers in Simulation, vol. 5, no. 2, 2011, pp. 135-143.

[60] M. Li and H. Duan, "Hybrid artificial bee colony and particle swarm optimization approach to protein secondary structure prediction," [In Proceedings of the 10th World Congress on Intelligent Control and Automation, IEEE, pp. 5040-5044, 2012.

[61] E. Cuevas, F. Sención-Echauri, D. Zaldivar, and M. Pérez, "Image segmentation using artificial bee colony optimization," in Handbook of Optimization, Springer, pp. 965-990, 2013.

[62] M. Bou-Imajjane and M. Sbihi, 2016, "Brain image segmentation using artificial bee colony optimization and markovian potts model," [in Proceedings of the 5th International Conference on Multimedia Computing and Systems, IEEE, pp. 141-147, 2016].

[63] A. Mostafa, A. Fouad, M. Abd Elfattah, A. E. Hassanien, H. Hefny, S. Y. Zhu, and G. Schaefer, "CT liver segmentation using artificial bee colony optimisation," [in Proceedings of the 19th International Conference on Knowledge Based and Intelligent

Information and Engineering Systems, Procedia Computer Science, Elsevier, vol. 60, pp. 1622-1630, 2015].

[64] S. Ait-Aoudia, E.-H. Guerrout, and R. Mahiou, “Medical image segmentation using particle swarm optimization,” [in Proceedings of the 18th International Conference on Information Visualization, IEEE, pp. 287-291, 2014].

[65] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” in Proceedings of the 6th International Symposium on Micro Machine and Human Science," IEEE, pp. 39-43.

[66] B. Birge, “PSOt-A particle swarm optimization toolbox for use with matlab,” [in Proceedings of the IEEE Swarm Intelligence Symposium, IEEE, pp. 182-186, 2003].

[67] A. P. Engelbrecht, “Fundamentals of computational swarm intelligence,” John Wiley & Sons, Inc, 2006.

[68] J, Kennedy, “Particle swarm optimization,” [in Proceedings of International Conference on the Encyclopedia of Machine Learning, Springer, pp. 760-766, 2010].

[69]Y. Shi, “Particle swarm optimization: developments, applications and resources,” [in Proceedings of the Congress on Evolutionary Computation, IEEE, pp. 81-86, 2001].

[70] D. Van der Merwe and A. P. Engelbrecht, “Data clustering using particle swarm optimization,” [in Proceedings of the Congress on Evolutionary Computation, IEEE, pp. 215-220, 2003.