Iranian Journal of Medical Physics
Vol. 12, No. 3, Summer 2015, 209-222
Received: April 15, 2015; Accepted: September 21, 2015

Original Article

A Novel Spot-Enhancement Anisotropic Diffusion Method for the Improvement of Segmentation in Two-dimensional Gel Electrophoresis Images, Based on the Watershed Transform Algorithm

Sina Shamekhi[1], Mohammad Hossein Miran Beygi[1]*, Bahareh Azarian[2], Ali Gooya[1]

Abstract

Introduction
Two-dimensional gel electrophoresis (2DGE) is a powerful technique in proteomics for protein separation. In this technique, spot segmentation is an essential stage, which can be challenging due to problems such as overlapping spots, streaks, artifacts and noise. Watershed transform is one of the common methods for image segmentation. Nevertheless, in 2DGE image segmentation, the noise and artifacts of images cause over-segmentation in the watershed algorithm.
Materials and Methods
In this study, we proposed a novel spot-enhancement anisotropic diffusion (SEAD) method, based on multi-scale second-order derivatives and eigensystemto enhance the spots and remove noise and artifacts. The proposed SEAD algorithm was plugged to a watershed transform in order to improve the performance of watershed segmentation algorithm.
Results
The performance of the proposed SEAD method was evaluated on synthetic and real 2DGE images. The proposed algorithm was compared with other segmentation methodsin terms of different criteria including efficiency, precision and true positive rate. The performance of the methods were evaluated in the presence of noise and the results were evaluated by t-test. According to the count of detected spots, precision and efficiency of the proposed method were 0.82 and 0.67 respectively. The precision and efficiency values of the comparative methods were as follows: 0.65 and 0.42 for MCW algorithm, 0.40 and 0.37 for BWT method, 0.74 and 0.53 for the method proposed by Kostopoulou and 0.76 and 0.55 for the method proposed by Mylona.
Conclusion
The comparison of the proposed method with four other conventional methods revealed the superiority and effectiveness of the proposed SEAD method.

Keywords: Diffusion, Noise, Segmentation, Spot, Two-dimensional gel electrophoresis

1- Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran
* Corresponding author: Tel: +982182883370; Fax: +982182884325; E-mail: miranbmh@modares.ac.ir
2- Protein-Chemistry Laboratory, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

1. Introduction

Proteomics is a research area exploiting the temporal dynamics of proteins, expressed in a given biological compartment at a given time. Proteomics was first introduced in the mid-1990s to characterize and analyze an entire set of proteins in a cell line, tissue, organism or other biological samples [1-4].

As an established scientific branch, proteomics governs various methods and instruments such as two-dimensional gel electrophoresis (2DGE) and mass spectrometry. Generally, 2DGE is an important, practical and unrivalled method for protein separation, which is widely used in the search for disease biomarkers, particularly for cancer and neurological conditions [5]. Electrophoresis was introduced in the early 1970s and was soon widely applied by many researchers [6]. The initial reports on 2DGE, which combined isoelectric focusing and sodium dodecyl sulfate-polyacrylamide gel electrophoresis, were published in 1974 and 1975 [7, 8].

Overall, 2DGEprovides a complete map of proteins within the primary sample. The proteins are separated due to differences in their isoelectric points (pI) in the first dimension and subsequently, their molecular weight (MW) in the second dimension. These separated proteins appear as spots in the resulting gel, as well as the digital scanned image (Figure 1). Each provided image contains a few hundreds to several thousands of protein spots [9, 10].

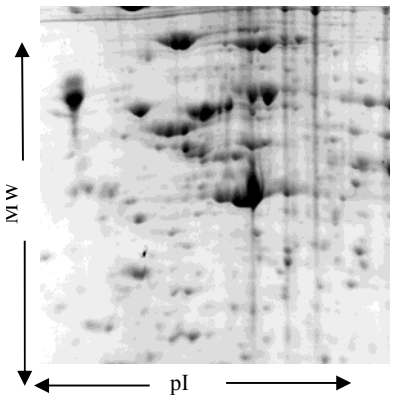


Figure 1. A real 2DGE image indicating isoelectric points (first dimension) and molecular weight (second dimension)

Based on the literature, 2DGE images have various limitations such as streaks, cracks, fingerprints, dust, staining variations and overlapping spots [11, 12]. These limitations can affect the reliability of analysis [13]. Moreover, these problems have been categorized by Dowsey as artifacts, co-migration, intensity inhomogeneity and geometric distortion [14]. These artifacts must be filtered and the overlapping spots must be correctly segmented.

Considering the abovementioned issues, the processing of 2DGE images is a complex and time-consuming procedure. Among the processing techniques of 2DGE images, spot segmentation encounters serious drawbacks, particularly in cases such as overlapping spots, streaks, staining problems and fingerprints, which lead to inaccurate segmentation. Therefore, development of spot detection and segmentation algorithms for accurate segmentation is of pivotal importance.

Several methods including parametric models, thresholding, edge detection, scale-space blob detection and mathematical morphology have been proposed to overcome problems associated with the segmentation of 2DGE images. Parametric models are based on some common characteristics of the shape of protein spots, which can be captured by a mathematical model [14]. Nevertheless, overlapping spots are the main problem in model-based techniques.

The thresholding method applies an increasing threshold to 2DGE images in order to split unified regions into multiple connected sub-regions [15, 16]. On the other hand, edge detection methods aim to identify large variations in image intensity, which are often associated with spot boundaries [17]. The overlapping spot segmentation is the shortcoming of both thresholding and edge detection methods. Also, these methods are highly sensitive to artifacts and noise [18, 19].

Mathematical morphology in image analysis is a broad field of research. H-dome and watershed transform are two widely used

techniques in this category [20-23]. In H-dome method, maximal structures are determined in grayscale-inverted 2DGE images. The overlapping neighboring spots are not regional maxima and H-dome cannot separate these overlapping spots[24].

Additionally, the watershed transform is a region-growing algorithm, which analyzes an image as a topographic surface. The advantage of this segmentation algorithm is that it produces closed, adjacent and accurate contours [25]. Nevertheless, the watershed transform facilitates the detection of false positive (FP) regions (over-segmentation) [24, 26].

There are two main approaches to reduce the problem of over-segmentation in watershed transformation [21, 25, 27, 28]. The first strategy is region merging in which the unwanted adjacent watersheds are merged and removed after segmentation. [29] Considering the large, non-linear variations in the intensity of pixels, the region-merging method often provides poor results [25]. The second approach to overcome over-segmentation is marker-controlled watershed (MCW) transform. A spot marker is a connected component in the image, which is imposed as regional minima on the image, while all other minima are suppressed. The drawback of these methods based on MCW transform is that the accuracy of segmentation depends on the number and location of the marker.

In 2DGE images containing anisotropic spots with varying intensities and noise, it is difficult to identify accurate spot markers and background regions. In other words, the current watershed-based methods are highly sensitive to noise and other artifacts in 2DGE images. Therefore, over-segmentation in watershed algorithm is still a challenging, unresolved problem. To improve the performance of watershed transform, we need a more developed and sophisticated image denoising algorithm for 2DGE images.

A technique for accomplishing image denoising is the scale-space denoising method, also known as non-linear anisotropic diffusion. This method was first introduced by Perona-Malik in 1990 for image denoising and edge detection where the amount of diffusion depends on the gradient of the image [30]. By applying Perona-Malik diffusion method, the image can be selectively smoothed, while preserving the edges.

In the Perona-Malik method, diffusion is inhibited in edges across the image. Therefore, the edges are not enhanced in this method. Moreover, Weickert introduced the concept of diffusion tensor and established a new anisotropic diffusion, which simultaneously enhances the edges, while smoothing the images [31]. This method has been recently developed and applied to some images [32], particularly medical images. The proposed method includes speckle suppression and edge detection in ultrasound images [33, 34] and vessel enhancing diffusion filters [35-37].

In the present study, we proposed a novel spot-enhancement anisotropic diffusion (SEAD) method for 2DGE images to improve marker selection in watershed segmentation algorithm. The proposed SEAD method is a multi-scale method, based on the second-order local Hessian matrix and its eigenvectors and eigenvalues.

In the present study, we proposed a novel spot-enhancement anisotropic diffusion (SEAD) method for 2DGE images to improve marker selection in watershed segmentation algorithm. The proposed SEAD method is a multi-scale method, based on the second-order local Hessian matrix and its eigenvectors and eigenvalues.

In the present study, the proposed SEAD watershed method is introduced in section 2. In section 3, the quantitative evaluation and results of SEAD application to synthetic and real images are presented. Moreover, the discussion is presented in section 4 and the conclusion is provided in section 5.

2. Materials and Methods

In this study, we proposed a novel anisotropic diffusion method, powered by a spot-enhancing filter (SEF) for improving 2DGE images. SEF is a second-order, local and structural image filter, which searches for

geometrical circular Gaussian structures across an image. Since the spots appear in different sizes, the scales of spots should be considered, as well. The proposed SEAD and SEF are based on the multi-scale eigenanalysis of the Hessian matrix of the image. We used the output of SEAD and SEF to improve MCW segmentation algorithm.

2.1. Watershed transform

In terminology of geosciences, watershed is defined as the geographical boundary of a region in a landscape where all water drains. The main idea is that water (i.e., in a river) or raindrops, falling on a surface, flow into the regional minima. The set of these minimum points constitutes a catchment basin and the borders between these catchments are watershed lines [23].

In the scope of image processing, watershed was first proposed by Beucher et al. as an effective segmentation approach[38]. The watershed transform has a wide range of applications, particularly in medical image processing for purposes such as cancer and tumor detection, brain MRI segmentation and tissue segmentation [39-42]. Nevertheless, this method suffers from over-segmentation in the presence of noise.

As mentioned in Section 1, 2DGE images have various artifacts and noise, which appear as various local minima across the image. These local minima lead the watershed towards over-segmentation. One proposed approach to overcome this problem is the MCW algorithm. The basic steps of a typical MCW are shown in Figure 2. In MCW algorithm, the markers of the spots and image background are initially generated. Spot markers are commonly detected by a regional minima detection algorithm, applied after using a smoothing filter. In the next step, the gradient of the raw image is calculated and used as the control surface in the watershed transform.

Some optional post-processing steps such as calculating and correcting the center and size of the spots have been applied in various studies [21, 23]. As mentioned before, the MCW algorithm is highly sensitive to the number and location of detected markers. The wide range of maximal intensity and size of protein spots and their anisotropic shape make the selection of a marker difficult in the presence of noise.

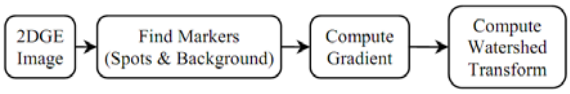


Figure 2. The basic steps of a typical marker controlled watershed (MCW) algorithm

In this study, we proposed a novel anisotropic diffusion method, which enhances spot regions across 2DGE images and suppresses noise and artifacts in these images. The proposed SEAD was plugged to a watershed transform in order to improve the performance of the watershed segmentation algorithm. The main steps of the proposed SEAD watershed algorithm are shown in Figure 3. The proposed SEAD algorithm had two main iterative sub-routines. The first algorithm was SEF and the second algorithm was a tensor-based anisotropic diffusion, proposed in this study.

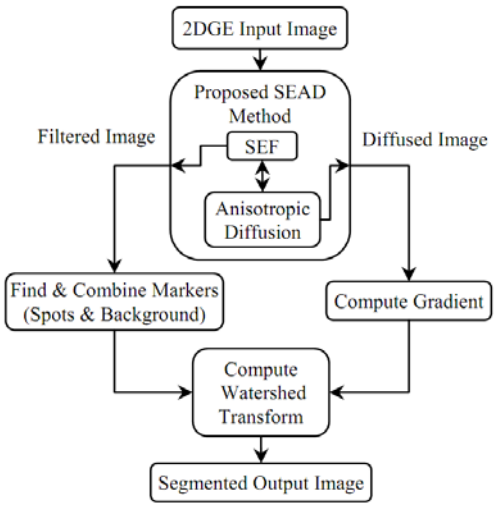


Figure 3. The proposed SEAD watershed algorithm

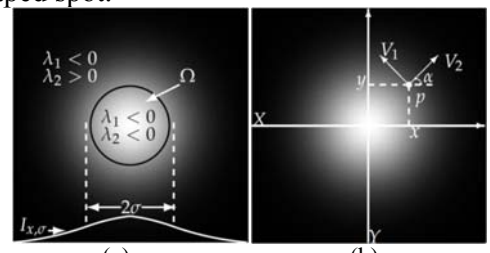
The proposed method was based on a multi-scale Hessian matrix and its eigenanalysis. Therefore, we calculated the eigenvalues and eigenvectors of the Hessian matrix of a Gaussian spot. Afterwards, we defined SEF and presented the SEAD diffusion algorithm.

2.2. Eigenanalysis of the Gaussian spot

The majority of parametric models of protein spots are commonly based on a two-dimensional Gaussian spot model [43, 44]. Therefore, we investigated the status of eigenvalues and eigenvectors of an ideal two-dimensional Gaussian spot. (a) (b) Figure 4(a) indicates the regional map of the states of eigenvalues in a second-order system on a Gaussian spot sample, located in the center of an image. Also, the directions of the eigenvectors are plotted in (a) (b) Figure 4(b). The equation of this Gaussian spot can be written as follows:

$$I_o = exp^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

where x and y are the coordinates of pixel p and σ is the space-scale of the Gaussian-shaped spot.



(a) (b)

Figure 4. (a) The regional map of the states of eigenvalues on a Gaussian spot. In the spot region, both eigenvalues (λ_1 and λ_2) are negative and on the background (non-spot) region, the signs of eigenvalues can take any other combinations; (b) the eigenvectors of the local Hessian at pixel *p*. As shown at pixel *p*, the eigenvectors V_1 and V_2 are orthogonal.

We also calculated the Hessian matrix of pixel p:

$$H = \frac{I_0}{\sigma^2}\begin{bmatrix} r^2-1 & rr' \\ rr' & r'^2-1 \end{bmatrix} \quad (2)$$

where $r = x/\sigma$ and $r' = y/\sigma$ are the normalized coordinate values. The eigenvalues of the Hessian matrix can be calculated as follows:

$$\lambda_1 = -\frac{I_0}{\sigma^2}, \quad \lambda_2 = \frac{I_0}{\sigma^2}(r^2+r'^2-1) \quad (3)$$

In this study, the eigenvalues were sorted in an ascending order $(\lambda_1 \le \lambda_2)$. Therefore, λ_1 is always negative and λ_2 can take negative, zero or positive values, depending on the pixel location:

$$\begin{aligned} r^2+r'^2<1 &\Rightarrow x^2+y^2<\sigma^2 \Rightarrow \lambda_2<0 \\ r^2+r'^2=1 &\Rightarrow x^2+y^2=\sigma^2 \Rightarrow \lambda_2=0 \\ r^2+r'^2>1 &\Rightarrow x^2+y^2>\sigma^2 \Rightarrow \lambda_2>0 \end{aligned} \quad (4)$$

Therefore, a border-line (Ω) is located where λ_2 changes. The relations between eigenvalues of the Hessian matrix and geometric structures are summarized in Table 1.

Table 1. Typical structures in 2DGE images and their corresponding eigenvalues

Patterns	λ_1	λ_2
Background	NL	NL
Bolb-like structure (dark)	H-	H-
Linear structure (dark)	L	H-
Saddle-like region	H+/-	H-/+

H ≡ high, L ≡ low, N ≡ Noisy and + /− ≡ Positive / Negative

We calculated the corresponding eigenvectors as follows:

$$\begin{aligned} V_1 &= c_1\frac{I_0}{\sigma^2}\begin{bmatrix} 1 \\ -r/r' \end{bmatrix} = c_{1'}\frac{I_0}{\sigma^2}\begin{bmatrix} r' \\ -r \end{bmatrix} \\ V_2 &= c_2\frac{I_0}{\sigma^2}\begin{bmatrix} 1 \\ r'/r \end{bmatrix} = c_{2'}\frac{I_0}{\sigma^2}\begin{bmatrix} r \\ r' \end{bmatrix} \end{aligned} \quad (5)$$

Where c_1, c_2, and $c_{2'}$ are constant values? The second eigenvector $V_2 = c''[r \quad r']^T$ is a radial vector and the first eigenvector V_1 is perpendicular to V_2 throughout the Gaussian spot, as presented in (a) (b) Figure 4(b). These eigenvectors were used to determine the directions needed for the proposed diffusion algorithm (Section 0).

2.3. SEF

We proposed a novel SEF to primarily detect the spot regions in 2DGE images. The proposed SEF is a multi-scale Hessian-based (second order) structural filter. To calculate the derivatives of image I, convolution with Gaussian derivatives was applied. The Hessian matrix (H) was a 2×2 matrix with two eigenvalues, i.e., λ_1 and λ_2. These eigenvalues were used to make three exponential structural measures and the final filter function. The final function of SEF consisted of exponential functions:

$$B_o(x)=\begin{cases} e^{(-\frac{R_{BN}^2}{2\alpha^2})}(1-e^{(-\frac{S^2}{2\beta^2})})e^{(-\frac{T}{\gamma})^2} & \lambda_1,\lambda_2<0 \\ 0 & o.w. \end{cases} \quad (6)$$

$$R_{BN}=\frac{\lambda_1-\lambda_2}{\lambda_1+\lambda_2} \quad (7)$$

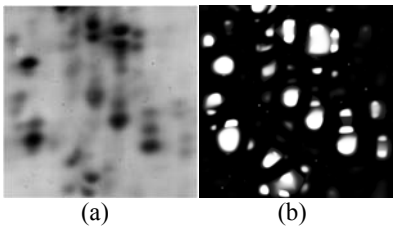
$$S=\sqrt{\lambda_1^2+\lambda_2^2} \quad (8)$$

$$T=exp(-(\frac{\lambda_2}{s})^2) \quad (9)$$

where R_{BN} is a proposed geometric ratio to measure local blobness, S differentiates between structures and background (noise), α, β and γ are the threshold parameters, which control the sensitivity of SEF to measure R_{BN} and S, respectively, λ_1 and λ_2 are the eigenvalues of the Hessian matrix of the image at each pixel, s is the scale of the pixel and T is a proposed measure to extract the internal border.

SEF is a multi-scale spot-enhancing filter. To specify the final values of SEF at each pixel, we computed the filter in a range of s scales and selected the maximum response over the range. The value of the proposed SEF was close to one on spot regions and close to zero on the background. A sub-image of the 2DGE input image and the corresponding output of the proposed SEF are demonstrated in Figure 5(a) and (b), respectively.

In this study, the filtered image was used as the input data to set the parameters of SEAD algorithm. As shown in Figure 3, the markers required in MCW algorithm were obtained from the SEF output.



(a) (b)

Figure 5. (a) The sub-image of the original 2DGE image, (b) the image filtered by the proposed blobness filter

2.4. SEAD algorithm

In this section, the proposed SEAD algorithm is explained. We established an anisotropic diffusion method, based on Hessian eigenanalysis, which enhances spot regions and reduces background noise in 2DGE images. We used the output scales obtained from the previous step (blobness filtering) to calculate the Hessians in correct scales. Afterwards, we implemented and run the anisotropic diffusion, which is an iterative and time-dependent process, based on Fick's law and continuity equation [30]. The diffusion equation can be written as follows:

$$I_t=\partial I/\partial t=div(g(|\nabla I(x,y,t)|)\nabla I)=g\Delta I+\nabla g.\nabla I \quad (10)$$

where I is the input image and div, ∇ and Δ are the divergence operator, gradient and Laplacian operator, respectively. We could replace g with a diffusion matrix D and rewrite the diffusion equation as follows [31]:

$$I_t=div(D\cdot\nabla I) \quad (11)$$

D was reconstructed based on the eigenvectors of Hessian matrix as follows [31]:

$$D=\begin{bmatrix} a & b \\ c & d \end{bmatrix}=[V_1 \quad V_2]\begin{bmatrix} f_1 & 0 \\ 0 & f_2 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (12)$$

$f_1(x,y,t)$ and $f_2(x,y,t)$ were applied as diffusion coefficients along V_1 and V_2 where $V_1=[V_{1x}\ V_{1y}]^T$ and $V_2=[V_{2x}\ V_{2y}]^T$ are the eigenvectors of the local hessian matrix. As shown in Figure 4(b), the eigenvectors had singular directions on the spot regions. The direction of V_1 is along the isolevels and V_2 is parallel to the gray-level gradient.

The objective here was to construct an anisotropic diffusion, which preserves blob-like structures (spot regions), while suppressing the background artifacts and noise. This can be achieved by defining the diffusion matrix D for spot regions. The diffusion moderately occurs in the direction of boundaries or parallel to isolevels (V_1), while the diffusion along V_2 is inhibited. Also, in

non-spot regions, diffusion should be increased in both directions to rapidly eliminate the noise and other background artifacts. The constraints suggested for the proposed SEAD are summarized in Table 2. It should be mentioned that ω and ς are positive large and small values, respectively .

Table 2. Diffusion constraints suggested for the proposed SEAD

Spot regions	$B_o(x) \to 1$	$f_1 \to 1$	$f_2 \to \varsigma$
Non-spot regions	$B_o(x) \to 0$	$f_1 \to \omega$	$f_2 \to \omega$

To formulate the diffusion constraints, we adopted the following diffusion coefficients:

$$f_1 = \omega e^{-B_o/\zeta} + 1$$
$$f_2 = \omega e^{-B_o/\zeta} + \varsigma \quad (13)$$

where B_o is the SEF function (6) and ζ is a threshold to control the sensitivity of functions to B_o .

As demonstrated in Figure 3, we used the diffused image (the output of SEAD) as the input of the final watershed transform. As previously mentioned, the markers required for MCW were obtained from the SEF output.

3. Results

3.1.Image acquisition and processing tools

Eight real 2DGE images and their corresponding segmentations (ground-truth images) were acquired in the Protein-Chemistry Laboratory of Biotechnology Research Center at Pasteur Institute (Tehran, Iran). Also, eight synthetic noise-free 2DGE images and the same synthetic images with common noise were obtained in this study. These synthetic images were used to evaluate the performance of methods, particularly in terms of noise effects.

To develop these synthetic images, synthetic two-dimensional Gaussian spots with common noise and artifacts, e.g., streaks, fingerprint, background noise and staining noise, were combined. Both real and synthetic images were obtained in a TIFF format with 24-bit grayscale depth (400 DPI). The size of real and synthetic images was 3181×3361 and 1000×1000 pixels, respectively.

The utilized real dataset contained approximately 5000 protein spots, while the synthetic dataset contained approximately 2000 spots. The ground-truth images of the synthetic 2DGE images were provided, while creating the synthetic image. The proposed algorithms were implemented using the open source Insight Toolkitand executed on a 3.2 GHz Intel workstation, running under Linux operating system.

3.2. Results of the proposed SEAD

We applied the proposed SEAD watershed algorithm on synthetic and real 2DGE images. The main advantage of synthetic images for the evaluation of the methods is that the effects of the presence or absence of noise can be investigated by these images. All parameters were set empirically. The sensitivity parameters α, β and γ were set at 5, 5 and 1, respectively, while the diffusion parameters ω and ς were set at 25 and 0.1, respectively.

3.2.1. SEAD on synthetic images

To evaluate the performance of the proposed anisotropic diffusion, we initially applied the method on a set of synthetic Gaussian spots, as depicted in Figure 6(a): one Gaussian spot, two spots, two closely-located spots, one noisy Gaussian spot, two noisy spots and two noisy closely-located spots. The diffused images of the original synthetic images, as shown in Figure 6(b), had clear spot regions and corrected backgrounds, compared to the corresponding original images [Figure 6(a)].

To further clarify the superiority of the proposed SEAD method on watershed segmentation algorithm, the segmentation results of the proposed method, the basic watershed transform (BWT) and MCW algorithms are shown in Figures 6 (c)-(e), respectively.

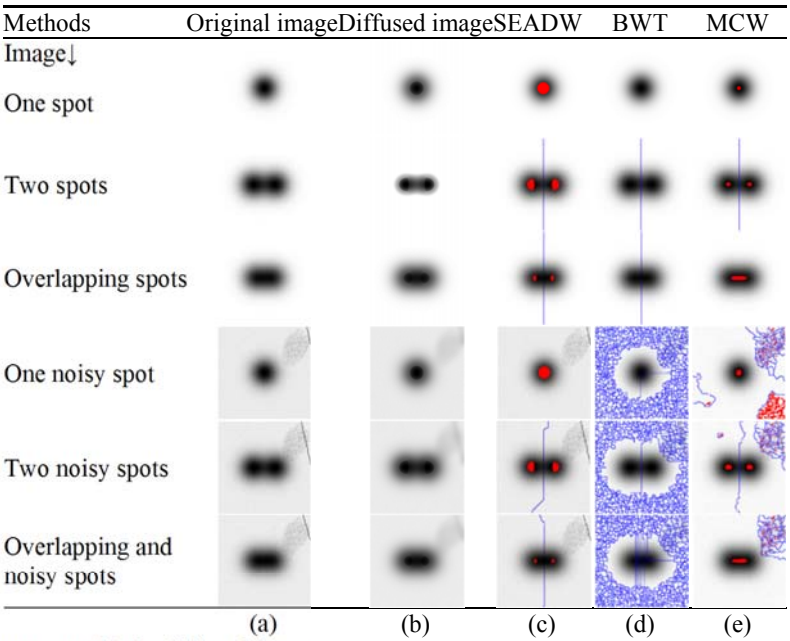


Figure 6. (a) Gaussian spots, (b) the diffused image obtained by the proposed SEAD method, (c) the segmented image created by the proposed SEAD watershed algorithm, (d) the segmented image created by BWT, and (e) the segmented image created by MCW algorithm. The proposed SEADW is superior to other watershed algorithms in segmentation of overlapping and noisy spots.

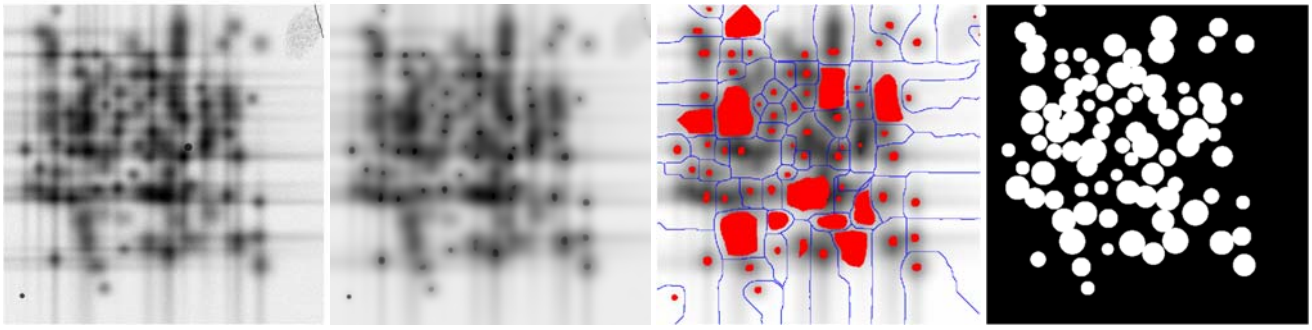


Figure 7. (a) A synthetic 2DGE image from our set of synthetic images, (b) the diffused image (output of the proposed SEAD), (c) the output of the SEAD watershed segmentation algorithm, (d) the ground-truth image

We applied the proposed SEAD watershed algorithm to a set of eight synthetic 2DGE images. A synthetic 2DGE image from this set and its diffused, segmented and ground-truth images are shown in Figure 7. The quantitative results of this evaluation are presented in Section 0.

3.2.2. SEAD on real 2DGE images

In the next step, we applied the proposed SEAD method to real 2DGE images. The final results of the diffused image, constructed by the proposed SEAD method, are demonstrated in Figure 8(a). The segmentation results of the SEADW algorithm are shown in Figure 8(b). To evaluate the segmented images, developed by the proposed method, the ground-truth images were required. In this study, the ground-truth images for each 2DGE image were provided by the expert biologists of Pasteur Institute of Iran. The ground-truth image for the 2DGE image in Figure 1 is depicted in 8(c). In the ground-truth images, the red borders indicate the spot areas.

For further clarification, the sub-images of the inverted real images are shown in Figure 9 (a). The corresponding inverted diffused images, obtained from the proposed SEAD diffusion algorithm, the segmented sub-images captured by the proposed SEAD watershed algorithm and the ground-truth images are shown in Figure 9 (b) to (d), respectively. As depicted in Figure 9 (b), the spot regions in Figure 9 (a) images are more recognizable than the original real sub-images. Also, the spots located within the streaks were isolated and the overlapping spots were successfully separated, as depicted in Figure 9 (c).

3.3.Evaluation of the performance of the proposed SEAD watershed method

We compared our proposed method with other available segmentation methods, implemented in this study. BWT was the first comparative method [38], followed by MCW transform, proposed in the literature [23]. Another comparison was based on Otsu's thresholding method [45], utilized by Kostopoulou and colleagues [16]. Furthermore, we implemented the method by Mylona et al.[18] and compared the segmentation results.

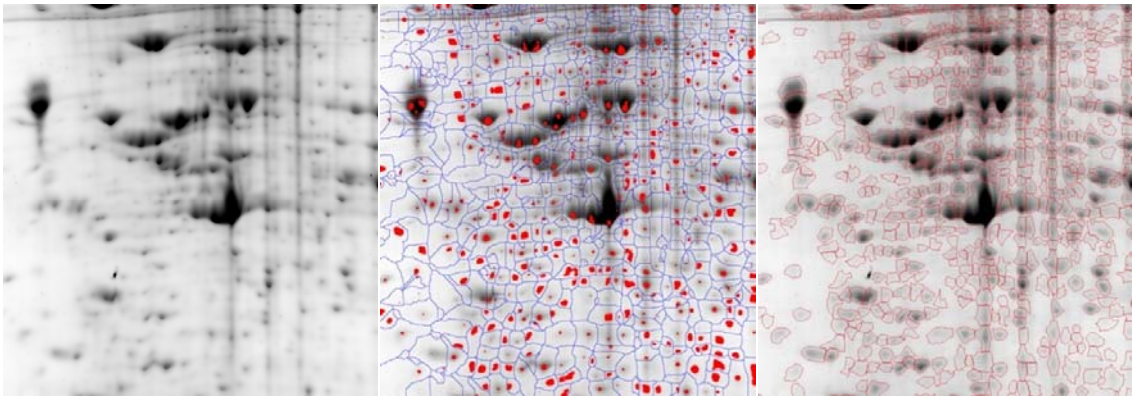
Some criteria are required to evaluate the performance of the proposed method in comparison with the above-mentioned methods. We used efficiency, precision and true positive rate (TPR) as the main criteria for comparing the number of detected spots by each method. These measures are defined as follows [25, 46, 47]:

$$Efficiency = \frac{TNDS - FP}{TNDS + FN} = \frac{TP}{TP + FP + FN} \quad (14)$$

$$Precision = \frac{TNDS - FP}{TNDS} = \frac{TP}{TP + FP} \quad (15)$$

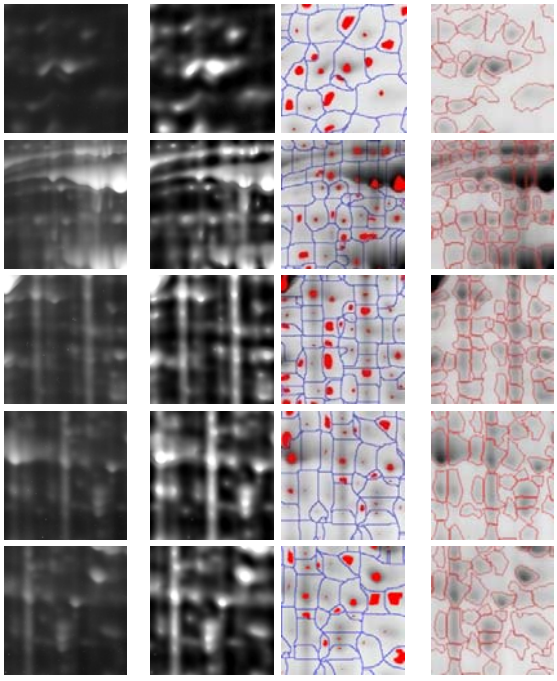
$$TPR = \frac{TP}{P} = \frac{TP}{(TP + FN)} \quad (16)$$

where TP is the true positive, FP is the false positive, FN is the false negative and $TNDS$ denotes the total number of detected spots.



(a) (b) (c)

Figure 6. (a) The diffused output of the proposed SEAD method on a real 2DGE image (shown in Figure 1), (b) the segmented image in this study, (c) the corresponding ground-truth image



(a) Sub-image (b) Diffused (c) Segmented (d) Ground-truth

Figure 7. Sub-figures and diffused images, (a) sub-images of the real 2DGE image, (b) the diffused image obtained from SEAD, (c) the segmented output created by the proposed SEAD watershed transform, and (d) the ground-truth images of the selected regions

3.3.1. Evaluation of the number of detected spots

The number of the detected spots by the proposed filter was evaluated and compared with other methods. To compute TP, FP and FN of spot count, we extracted he locations of the local maxima within the watersheds and compared the extracted locations with true locations of the spots, using precision (15) and efficiency criteria (14). True locations of the spots are available in the ground-truth database.

Generally, noise and artifacts lead the watershed transform towards over-segmentation. We investigated the effects of noise and artifacts by comparing the performance of the methods on two types of synthetic images: noisy images (NI) and images without noise (WN), which were developed by the authors. The mean values of precision and efficiency, based on the comparison, are shown in Table 3 and Figure 8; the standard deviations are plotted as error bars in Figure 8. The number of detected spots in real 2DGE images, calculated by different methods, is shown in and Table 4 and Figure 9. Additionally, the standard deviations of TPR, precision and efficiency are shown in Figure 9.

Table 3. Comparison of the mean values of precision (Pr.) and efficiency (Eff.)of the methods applied to synthetic noisy images (NI) and images without noise (WN)

	SEADW		MCW		BWT		Kostopoulou		Mylona	
I	WN	NI	WN	NI	WN	NI	WN	NI	WN	NI
Pr.	0.83	0.75	0.73	0.4	0.55	0.1	0.7	0.42	0.81	0.64
Eff.	0.71	0.64	0.63	0.37	0.51	0.09	0.63	0.39	0.7	0.57

(a) (b)

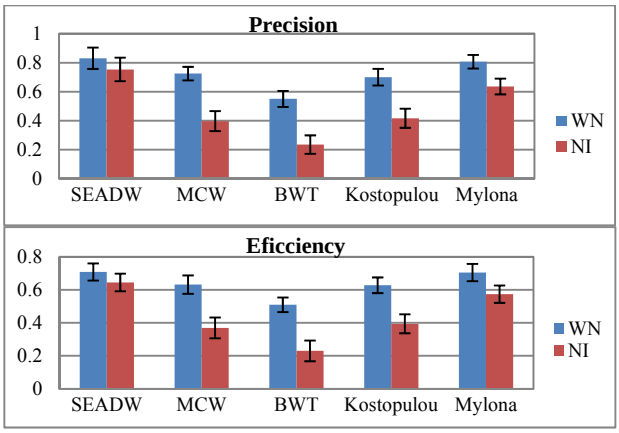


Figure 8. Comparison of the mean values of (a) precision (Pr.) and (b) efficiency (Eff.) of the methods applied to synthetic noisy images (NI) and images without noise (WN). BWT, MCW and Kostopoulou methods are highly sensitive to noise.

Table 4.The count of detected spots on real 2DGE images

I	Methods														
	SEADW			MCW(2006)			BWT			Kostopoulou (2014)			Mylona(2011)		
	TPR	Pr.	Eff.	TPR	Pr.	Eff.	TPR	Pr.	Eff.	TPR	Pr.	Eff.	TPR	Pr.	Eff.
1	0.83	0.86	0.73	0.50	0.71	0.42	0.82	0.46	0.41	0.60	0.81	0.52	0.68	0.77	0.57
2	0.78	0.86	0.69	0.51	0.70	0.42	0.86	0.44	0.41	0.57	0.74	0.47	0.65	0.79	0.58
3	0.8	0.84	0.70	0.63	0.62	0.45	0.90	0.38	0.36	0.67	0.71	0.53	0.64	0.79	0.55
4	0.69	0.73	0.55	0.67	0.56	0.44	0.81	0.32	0.30	0.71	0.71	0.55	0.64	0.72	0.51
5	0.83	0.84	0.72	0.50	0.74	0.42	0.84	0.47	0.43	0.61	0.81	0.54	0.66	0.82	0.58
6	0.78	0.86	0.70	0.49	0.71	0.41	0.86	0.42	0.40	0.56	0.78	0.49	0.69	0.82	0.60
7	0.79	0.81	0.67	0.64	0.57	0.44	0.93	0.35	0.34	0.80	0.61	0.53	0.71	0.67	0.52
8	0.71	0.74	0.56	0.57	0.57	0.40	0.79	0.35	0.32	0.73	0.76	0.60	0.69	0.70	0.53
M	0.78	0.82	0.67	0.56	0.65	0.42	0.85	0.40	0.37	0.66	0.74	0.53	0.67	0.76	0.55
±SD	±0.05	±0.05	±0.07	±0.07	±0.09	±0.02	±0.05	±0.0.5	±0.05	±0.08	±0.07	±0.04	±0.03	±0.06	±0.03

TPR ≡ True Positive Rate, Pr.≡ Precision, and Eff.≡ Efficiency

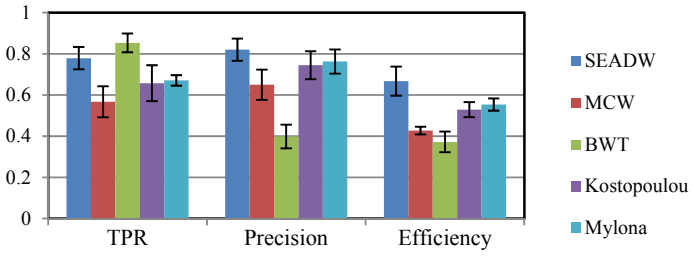


Figure 9. Comparison of TPR, precision and efficiency of the methods in real 2DGE images

Table 5. Paired t-test results pertaining to the number of detected spots using the proposed SEAD watershed and other methods

	MCW		BWT		Kostopoulou		Mylona	
	h	*P*	*H*	*p*	*h*	*P*	*h*	*p*
Precision	1	9.9×10−6	1	3.9×10−9	1	0.0186	1	0.0069
Efficiency	1	3.2×10−5	1	5.7×10−8	1	0.0053	1	5.0×10−4

We have further investigated the statistically significant differences between the results of proposed SEAD watershed and the other implemented methods by performing t-test evaluations [48].The results of t-test is presented in Table 5.

4. Discussion

Based on the findings, the proposed method had a superior performance on 2DGE images. The presence of noise caused over-segmentation in previous techniques, while the proposed SEAD watershed was proved to be a robust and reliable method in 2DGE images, particularly in noisy images. As indicated in Figure 6, the noisy overlapping spots were only segmented correctly to two separate regions by the proposed SEAD watershed algorithm. Also, the over-segmentation problem of BWT and MCW in the presence of noise was clearly noted.

Based on the results presented in Table 3 and Figure 10, the presence of noise reduces the efficiency and precision of all methods. Nevertheless, the presence of noise leads to the detection of many FPs in watershed-based methods (i.e., BMT and MCW) and considerably decreases their segmentation performance. However, considering the proposed SEAD algorithm, the segmentation performance of SEAD watershed was not sensitive to the presence of noise. Unlike the proposed method, the previous watershed-based methods and the method proposed by Kostopoulou were highly sensitive to the presence of noise.

Comparison between the number of detected spots by the proposed SEAD watershed method and previous techniques reveals that the precision and efficiency of this method are maximal, whereas the TPR of BWT is maximal. This is due to the large number of detected spot regions in BWT method. In fact, the large number of FPs reduces the efficiency and precision of BWT method. The methods proposed by Mylona and Kostopoulou had good precision and moderate efficiency on our database.

The results of t-test on these quantitative parameters demonstrated considerable differences between SEAD watershed transform the proposed and previous methods. In other words, the proposed method showed a considerably effective performance in comparison with previous methods. In this study, all the parameters of the proposed SEF and diffusion algorithm were set empirically. In order to improve the proposed method, the effects of variations in these parameters should be further investigated.

5. Conclusion

In this study, a novel SEAD method was proposed for spot enhancement and noise removal in 2DGE images. This method was based on the multi-scale Hessian analysis,which facilitates a more accurate detection of local structures and their properties. The proposed method was used to improve the segmentation performance of watershed transform. One important contribution of this study was the introduction of MCW segmentation algorithm, which was not sensitive to noise. Unlike the proposed SEAD watershed algorithm, noise and artifacts in 2DGE images caused over-segmentation in previous watershed algorithms. We used the SEAD output as a marker to extract the accurate locations of the spots. This approach reduced the number of detected FP spots and eliminated over-segmentation in watershed transform in 2DGE images. We compared our segmentation method with other published methods in terms of spot count and t-test results. We also evaluated our proposed method, using synthetic and real 2DGE images. Our quantitative evaluations showed that the proposed SEAD watershed segmentation method is a viable multi-scale tool, which can be employed for the segmentation of 2DGE images.

Acknowledgment

The authors would like to thank Dr. Behrouz Vaziri, the head of Protein-chemistry laboratory at Pasteur Institute of Iran for his support and guidance.

References

1. Anderson NG, Anderson NL. Twenty years of two-dimensional electrophoresis: past, present and future. Electrophoresis. 1996 Mar;17(3):443-53.
2. Graves PR, Haystead TA. Molecular biologist's guide to proteomics. Microbiology and Molecular Biology Reviews. 2002;66(1):39-63.
3. Yan JX. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. Electrophoresis. 1995;16:1090-4.
4. Wilkins MR, Sanchez J-C, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnology and genetic engineering reviews. 1996;13(1):19-50.
5. Issaq HJ, Veenstra TD. The role of electrophoresis in disease biomarker discovery. Electrophoresis. 2007;28(12):1980-8.
6. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. nature. 1970;227(5259):680-5.
7. Macgillivray AJ, Wood DR. The Heterogeneity of Mouse-Chromatin Nonhistone Proteins as Evidenced by Two-Dimensional Polyacrylamide-Gel Electrophoresis and Ion-Exchange Chromatography. European Journal of Biochemistry. 1974;41(1):181-90.
8. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. Journal of biological chemistry. 1975;250(10):4007-21.
9. Rabilloud T, Chevallet M, Luche S, Lelong C. Two-dimensional gel electrophoresis in proteomics: past, present and future. Journal of proteomics. 2010;73(11):2064-77.
10. Rabilloud T, Lelong C. Two-dimensional gel electrophoresis in proteomics: a tutorial. Journal of proteomics. 2011;74(10):1829-41.
11. Rye M, Fargestad EM. Preprocessing of electrophoretic images in 2-DE analysis. Chemometrics and Intelligent Laboratory Systems. 2012;117:70-9.
12. Nhek S, Tessema B, Indahl U, Martens H, Mosleth E. 2D electrophoresis image segmentation within a pixel-based framework. Chemometrics and Intelligent Laboratory Systems. 2015;141:33-46.
13. Yang G-Z. The role of bioinformatics in two-dimensional gel electrophoresis. Proteomics. 2003;3:1567-96.
14. Dowsey AW, English JA, Lisacek F, Morris JS, Yang GZ, Dunn MJ. Image analysis tools and emerging algorithms for expression proteomics. Proteomics. 2010;10(23):4226-57.
15. Tyson JJ, and Haralick RH. Computer analysis of two-dimensional gels by a general image processing system. Electrophoresis. 1986; 7(3):107-113.
16. Kostopoulou E, Zacharia E, Maroulis D. An effective approach for detection and segmentation of protein spots on 2-d gel images. Biomedical and Health Informatics, IEEE Journal of. 2014;18(1):67-76.
17. Lemkin PF, Lipkin LE. 2-D Electrophoresis gel data base analysis: Aspects of data structures and search strategies in GELLAB. Electrophoresis. 1983;4(1):71-81.
18. Mylona E, Savelonas M, Maroulis D, Kossida S. A Computer-Based Technique for Automated Spot Detection in Proteomics Images. Information Technology in Biomedicine, IEEE Transactions on. 2011;15(4):661-7.
19. Kim Y, Kim J, Won Y, In Y. Segmentation of protein spots in 2D gel electrophoresis images with watersheds using hierarchical threshold. Computer and Information Sciences-ISCIS 2003: Springer; 2003. p. 389-96.
20. Vincent L. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. Image Processing, IEEE Transactions on. 1993;2(2):176-201.
21. Tsai M-H, Hsu H-H, Cheng C-C. Watershed-based protein spot detection in 2DGE images. ICS 2006,Int'l Computer Symposium; 2006.
22. Tsai T-S, Chen T-S, Tang S-S, Chou M-C, Li S-Y, Wang T-S, editors. A new progressive image transmission scheme based on content feature and Gaussian model of 2D gel electrophoresis images. Advanced Communication Technology, 2006 ICACT 2006 The 8th International Conference; 2006: IEEE.
23. Beare R, Lehmann G. The watershed transform in ITK-discussion and new developments. The Insight Journal. 2006;(1):1-24.
24. Pedersen L. Analysis of two-dimensional electrophoresis gel images. Information and Mathematical Modeling, Revision. 2002;1.

25. Sengar RS, Upadhyay AK, Singh M, Gadre VM, editors. Segmentation of two dimensional electrophoresis gel image using the wavelet transform and the watershed transform. Communications (NCC), 2012 National Conference on; 2012: IEEE.
26. Horgan GW, Glasbey CA. Uses of digital image analysis in electrophoresis. Electrophoresis. 1995;16(1):298-305.
27. Hamarneh G, Li X. Watershed segmentation using prior shape and appearance knowledge. Image and Vision Computing. 2009;27(1):59-68.
28. Sun C-l, Wang X-m, editors. Spot segmentation and verification based on improve marker controlled watershed transform. Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on; 2010: IEEE.
29. Haris K, Efstratiadis SN, Maglaveras N, Katsaggelos AK. Hybrid image segmentation using watersheds and fast region merging. Image Processing, IEEE Transactions on. 1998;7(12):1684-99.
30. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 1990;12(7):629-39.
31. Weickert J. Anisotropic diffusion in image processing: Teubner Stuttgart; 1998.
32. Liu F, Liu J. Anisotropic diffusion for image denoising based on diffusion tensors. Journal of Visual Communication and Image Representation. 2012;23(3):516-21.
33. Yue Y, Croitoru MM, Bidani A, Zwischenberger JB, Clark JW. Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images. Medical Imaging, IEEE Transactions on. 2006;25(3):297-311.
34. Fu S, Ruan Q, Wang W, Li Y. Adaptive anisotropic diffusion for ultrasonic image denoising and edge enhancement. International Journal of Information Technology. 2006;2(4):284-92.
35. Manniesing R, Viergever MA, Niessen WJ. Vessel enhancing diffusion: A scale space representation of vessel structures. Medical image analysis. 2006;10(6):815-25.
36. Enquobahrie A, Ibanez L, Bullitt E, Aylward S. Vessel enhancing diffusion filter. The Insight Journal, 2007;(1):1-14.
37. Abdollahi B, El-Baz A, Amini A, editors. A multi-scale non-linear vessel enhancement technique. Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE; 2011: IEEE.
38. Beucher S, Lantuejoul C. Use of Watersheds in Contour Detection. International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation. Rennes. France. 1979.
39. Salman SD, Bahrani AA. Segmentation of tumor tissue in gray medical images using watershed transformation method. International Journal of Advancements in Computing Technology. 2010;2(4):123-127.
40. Mustaqeem A, Javed A, Fatima T. An efficient brain tumor detection algorithm using watershed & thresholding based segmentation. International Journal of Image, Graphics and Signal Processing (IJIGSP). 2012;4(10):34.
41. Balafar MA, Ramli AR, Saripan MI, Mashohor S. Review of brain MRI image segmentation methods. Artificial Intelligence Review. 2010;33(3):261-74.
42. Massoptier L, Casciaro S. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans. European radiology. 2008;18(8):1658-65.
43. Bettens E, Scheunders P, Sijbers J, Van Dyck D, Moens L. Automatic segmentation and modelling of two-dimensional electrophoresis gels. Image Processing, 1996. Proceedings., International Conference on.1996. p. 665-668.
44. Bettens E, Scheunders P, Van Dyck D, Moens L, Van Osta P. Computer analysis of two-dimensional electrophoresis gels: A new segmentation and modeling algorithm. Electrophoresis. 1997;18(5):792-8.
45. Otsu N. A threshold selection method from gray-level histograms. Automatica. 1975;11(285-296):23-7.
46. Comar SR, Malvezzi M, Pasquini R. Are the review criteria for automated complete blood counts of the International Society of Laboratory Hematology suitable for all hematology laboratories? Revista brasileira de hematologia e hemoterapia. 2014;36(3):219-25.
47. Fawcett T. An introduction to ROC analysis. Pattern recognition letters. 2006;27(8):861-74.
48. Kremelberg D. Practical Statistics: A Quick and Easy Guide to IBM® SPSS® Statistics, STATA, and Other Statistical Software: Sage Publications; 2010.