A hybrid skin lesions segmentation approach based on image processing methods

Hanae Moussaoui 1,*, Nabil El Akkad 1, Mohamed Benslimane 2

1Engineering, Systems and Applications Laboratory at ENSA of Fez, Sidi Mohamed Ben Abdellah University of fez, Morocco
2Innovative Technologies Laboratory at EST of Fez, Sidi Mohamed Ben Abdellah University of fez, Morocco

Abstract Presently image segmentation remains the most crucial stage in the image processing system. The main idea of image segmentation is to partition or divide a random image into several partitions depending on the problem to solve. In this paper, we will be presenting a new method of skin cancer detection based on Otsu’s thresholding algorithm and marker-controlled watershed method. This hybridization process is first of all started by segmenting the input image using fuzzy c-means algorithm which is a clustering method that gives the possibility to a pixel to belong to one or more clusters. After that, we will apply multi-Otsu which is a thresholding algorithm that separates the pixels of an image into a variety of classes depending on the intensity of the gray levels. The next step of this proposed method is the marker-controlled watershed algorithm that divides the image into homogenous areas or regions by using edge-detection concepts including mathematical morphology. The proposed technique was applied and experienced using several images of different types of skin cancer that were collected and gathered from the web and also from the Kaggle dataset. To assess the worth of the achieved results, we used several evaluation metrics like dice coefficient, sensitivity, specificity as well as Jaccard similarity that all have shown good and satisfactory results.

Keywords Image segmentation, skin cancer, fuzzy c-means, multi-Otsu, thresholding, marker-controlled watershed

DOI: 10.19139/soic-2310-5070-1549

1. Introduction

Skin cancer occurs when cells in the epidermis have turned into cancerous cells. Different types of skin lesions are determined by which cells in the epidermis turn cancerous [1, 2]. Among the skin cancer types, we can cite the nonmelanoma and the melanoma types. The nonmelanoma will include basal cell carcinoma, squamous cell carcinoma including actinic keratosis which can turn into a squamous cell carcinoma. Furthermore, the melanoma type remains one of the most highly aggressive types of skin cancer of all the types listed before. Non-melanoma skin cancers happen on their extremities, face and any place we get a lot of chronic sun damage. Chronic means that the more sun exposure we get, the higher the risk will be. Melanoma happens also in other areas where the sun doesn’t shine as constantly, like on the legs. Even where the sun often never shines because melanoma can happen in the GI tract, as well as uncommon melanoma that’s also in the eye. One of the most important signs we should be aware [4] of are spots that don’t look like any place of the body which is called the Ugly Duckling Rule. The other characteristic we want to look for is changing, so an area that’s changing more than the rest of the body. We all have spots and lumps that are relatively stable, but cancer generally isn’t. So, it’ll continue to evolve and change. Everybody at least is on the list of being at risk of skin cancer, and early detection means that you’ll have
a smaller scar while the surgical repair, or sometimes we can treat very early non-melanoma [5] skin cancers with non-surgical techniques such as creams or another kind of ablative techniques. The best ways to protect ourselves especially in the case of having fair skin, when the risk for UV damage is high is by avoiding the sun between 10 and 2 PM [3], wearing sun-protective clothing, applying sunscreens appropriately. Medical imaging [6] is a fascinating area, the challenges are incredible and so are the rewards. When we solve problems in medical imaging, we're helping humanity, and that’s a great reward. Medical imaging [7] is very different than many other areas of image processing. One of the differences is that very often we start the investigation with a problem presented to us such as brain tumor, leukemia, and skin cancer detection [8, 9] and segmentation. Image processing can be used in many different types of computer vision, particularly image segmentation [10, 11, 12, 13], 3D reconstruction [14, 15], the auto-calibration camera [16, 17, 18, 19, 20]. Image segmentation is the mechanism that gives the capacity to move apart objects from the background to make these objects easier to analyze later. Image segmentation helps to identify individual structures within the image, like in our case, finding regions of the skin that contain tumors. In other words, we’ll classify each pixel in the skin image as a tumor or not tumor. To segment an image, we can choose over many approaches, depending on the nature of the problem to solve. Among these approaches, we can list pixel-based [21] methods like thresholding and clustering. The main idea of these methods is to group pixels in the input image based on color similarities or the intensity of pixels. The second example is the method based on edges [22, 23] when the algorithm’s main objective is to identify the change of state of the intensity between pixels. Region-based algorithms are the third type of image segmentation when the purpose is to create larger and homogeneous areas by comparing every pixel with its neighbor. The last type of image segmentation is the hybrid approach when we use to combine two already existing methods to give birth to a new one with better results and performances. In this article, we’ll be presenting a novel technique for detecting and segmenting skin tumors, based on the hybridization between fuzzy c-means algorithm and marker watershed method using Otsu’s algorithm. The work will be organized as follows; we’ll start by listing the previous and different methods that have been proposed in the area of skin tumor detection. The third part will be consecrated to discuss our proposed method, followed by the part when we’ll compare the obtained results with the two methods that have been used using several evaluations metrics such as Shannon entropy, Jaccard Index, Specificity, Sensitivity and dice coefficient that all have shown satisfactory results.

2. Related works

R. Suganya [24] has proposed a method of skin cancer detection using a Median filter for hair removal, k-means clustering, and SVM classification. This method achieved 95.4 in sensitivity, 89.3 in specificity, and an accuracy of 96.8.

Chiem, A. Al-Jumaily, and R. N. Khushaba [25] have proposed a skin cancer detection method. For this, they first applied median filter and contrast enhancement, followed by the thresholding technique. For classification, they chose a backpropagation neural network and support vector machine.


M. W. Rashad and M. Takruri [27] have suggested a technique based on SVM classification by first applying Wiener and Median filters, and then k-means clustering and SVM classifier. This method gains a sensitivity rate of 80.7, a specificity rate of 62.5, and an accuracy rate of 75.31.

Smaoui. N, and Derbel. N [28] have proposed a study on skin cancer detection. For this and in the preprocessing stage, they used the median filter followed by applying morphological closing and contrast enhancement. Apropos feature extraction and classification, they opted for the ABCD rule. This process had an accuracy of 90.
3. The proposed method

In this unit, we will be talking about our proposed process for skin cancer detection and segmentation. For this, the first task was to gather images from the web and datasets such as the Kaggle dataset. The second step is the preprocessing phase when we apply the median filter. It’s a non-linear filter that is used for reducing the noise in the original image. In the next step, we opted for fuzzy $c$-means clustering which is a soft clustering with the ability to give to a pixel the capacity to be in one or more clusters at the same time. The resulting image will be after that going through the next phase of the proposed method. The next step is presented by Multi Otsu’s algorithm that looks at the histogram of the image, including the pixel values and the property to segment the object that we want. After that, we’ll use the marker watershed algorithm for the segmentation. The following schema shows in detail the sequence of the different stages in the proposed method.

![Diagram of the proposed method flowchart]

Figure 1 the proposed method flowchart
3.1. Median filter

Before applying any changes to any input image, the first thing we need to do is to preprocess. That means removing any imperfections, such as pepper and salt noise which are represented by small dark and white dots. The median filter [29] is a non-linear operation that helps during this stage by blurring the image. Something very important about this filter is that it preserves edges and sharp lines. This is an advantage in medical imaging when we often look for defining the edges of tumors for example. The main idea of the median filter is, assuming we’ve got an input image, we’ll choose a neighborhood of three by three. Then we open up the whole matrix horizontally and arrange it into ascending order. The median will be the value lying in the center. We then use the median filter for all pixels in the image using the sliding window.

\[
y(n) = \text{med}x(i)
\]

where

\[
i = n, n - 1, \ldots, n - M
\]

3.2. Fuzzy c-means

Fuzzy c-means [28] is the most widespread algorithm for clustering. It’s a soft clustering process, which is considered as an extension of the hard clustering k-means. In fuzzy c-means, a pixel in an image can be in one or more clusters at the same time, unlike k-means. This method is especially used for analyses that are based on the range between several input data points. The clusters are then formed depending on the separation between data points and the cluster-centers that are made for each cluster.

Fuzzy c-means works by following the steps below:

**Step 1:** arbitrary set the membership function employing the equation written down:

\[
\sum_{j=1}^{c} u_j(x_i) = 1
\]

where

\[
i = 1, 2, \ldots, k
\]

**Step 2:** compute the centroid value by employing the following expression:

\[
C_j = \frac{\sum_j [u_j(x_i)]^m x_i}{\sum_j [u_j(x_i)]^m}
\]

**Step 3:** compute the variance between data points and the centroid by applying the Euclidean distance:

\[
D_i = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2}
\]

**Step 4:** amend the membership matrix using the following function:

\[
u_j(x_i) = \left[\frac{1}{\sum_k \left[\frac{1}{d_{ki}}\right]^{m-1}}\right]^{-\frac{1}{m-1}}
\]

where \(m\) represents the fuzzification parameter which is always a value between 1.25 and 2.

**Step 5:** redo this mechanism from step 2 until the centroids stop changing.
3.3. Multi Otsu’s algorithm

Multi Otsu’s [30] process is an algorithm to automatically get the optimal threshold intensity. This thresholding method is looking at the histogram of the given image, the pixel values, and the property to obtain segments. In this case, we won’t be looking at the edges, but at the regions and inside the segments that we want to segment out the object. Multi-Otsu’s method works by searching for the threshold intensity that optimally separates the image into two various classes; the foreground and the background. It does this by maximizing a metric called the between-class variance. This variance is given by the equation below:

\[ \sigma^2_B = w_B w_f (\mu_b - \mu_f)^2 \]  

where \( w_B \) and \( w_f \) are called the weights and they are on a par with the number of pixels in the foreground or in the background divided by the total number of pixels in the image. \( \mu_b \) and \( \mu_f \) present the mean intensity of the background or the foreground pixels.

3.4. Marker watershed algorithm

The watershed algorithm [31] takes an image as an input. The algorithm would treat the intensity value of each pixel in the image as its height. So, the brighter the pixel, the higher it is. The important thing to know is that each object in the image must have low intensity. Picking again in terms of height means that each object must be a basin. The ridgelines are used to separate the objects, essentially to mark the edges.

The steps for executing the watershed transform are as follows:
**Step 1:** Read the given image and transform it into a grayscale

\[ G(I, j, k) = 0.30 \times R + 0.59 \times G + 0.11 \times B \]  

(7)

Where 'G' represents the grayscale image.

**Step 2:** Apply the gradient magnitude being the segmentation function [32].

**Step 3:** Define every foreground object by using some of the morphological techniques such as, erosion, closing, dilation and opening[33, 34, 35].

**Step 4:** Calculate the original maxima.

**Step 5:** Place the foreground-marker image over the original image.

**Step 6:** Calculate background-markers by applying the thresholding function.

**Step 7:** Apply at this stage the watershed algorithm to the adjusted gradient image[36, 37].
4. Results and discussion

In this part of the paper, we will discuss the proposed process which is made of three phases. The first thing we had to do is to gather the medical images showing skin lesions. For this, we had to collect about 300 images to test our proposed algorithm. The second stage is to go through the preprocessing phase. At this stage, the main purpose is to clarify the image, which means to remove noises so that the results will be more efficient. Here, we’ve chosen the median filter over the other filters because it has good results for pepper and salt enhancement. In addition to that, it doesn’t blur edges which is very interesting and advantageous. After that, we opted for fuzzy c-means segmentation which is a soft clustering that as said before, gives the chance for a pixel to be in one or more clusters at the same time. This resulting image will be then used as an input image to the next method that is the marker watershed algorithm. The purpose behind choosing the marker watershed algorithm is to finally mark the boundaries of the detected skin lesion. The table below shows the obtained result of the proposed process.

Table. 1 The obtained results

<table>
<thead>
<tr>
<th>Original image</th>
<th>Fuzzy c-means</th>
<th>Multi Otsu</th>
<th>First result</th>
<th>Marker Watershed</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Original Image" /></td>
<td><img src="image2.png" alt="Fuzzy c-means" /></td>
<td><img src="image3.png" alt="Multi Otsu" /></td>
<td><img src="image4.png" alt="First result" /></td>
<td><img src="image5.png" alt="Marker Watershed" /></td>
</tr>
<tr>
<td><img src="image6.png" alt="Original Image" /></td>
<td><img src="image7.png" alt="Fuzzy c-means" /></td>
<td><img src="image8.png" alt="Multi Otsu" /></td>
<td><img src="image9.png" alt="First result" /></td>
<td><img src="image10.png" alt="Marker Watershed" /></td>
</tr>
<tr>
<td><img src="image11.png" alt="Original Image" /></td>
<td><img src="image12.png" alt="Fuzzy c-means" /></td>
<td><img src="image13.png" alt="Multi Otsu" /></td>
<td><img src="image14.png" alt="First result" /></td>
<td><img src="image15.png" alt="Marker Watershed" /></td>
</tr>
<tr>
<td><img src="image16.png" alt="Original Image" /></td>
<td><img src="image17.png" alt="Fuzzy c-means" /></td>
<td><img src="image18.png" alt="Multi Otsu" /></td>
<td><img src="image19.png" alt="First result" /></td>
<td><img src="image20.png" alt="Marker Watershed" /></td>
</tr>
</tbody>
</table>
To assess the obtained result of the submitted method, we opted for several evaluation metrics like dice coefficient, Jaccard index, Sensitivity as well as Specificity. All the obtained results showed very satisfactory results of the given method compared to the two already existing algorithms, which are Fuzzy c means as well as the Watershed algorithm. Which means a good detection of skin lesions. The six tables below show the result of the six images shown at the top, using the four different measurements. As we can see, the proposed method achieved better results compared to fuzzy c-means besides marker watershed.

### Table 2: The obtained results for image 1

<table>
<thead>
<tr>
<th></th>
<th>Dice coefficient</th>
<th>Jaccard index</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proposed method</td>
<td>0.89</td>
<td>0.91</td>
<td>0.82</td>
<td>0.87</td>
</tr>
<tr>
<td>FCM</td>
<td>0.6</td>
<td>0.43</td>
<td>0.9</td>
<td>0.1</td>
</tr>
<tr>
<td>Marker watershed</td>
<td>0.59</td>
<td>0.42</td>
<td>0.97</td>
<td>0.04</td>
</tr>
</tbody>
</table>

### Table 3: The obtained results for image 2

<table>
<thead>
<tr>
<th></th>
<th>Dice coefficient</th>
<th>Jaccard index</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proposed method</td>
<td>0.86</td>
<td>0.86</td>
<td>0.79</td>
<td>0.72</td>
</tr>
<tr>
<td>FCM</td>
<td>0.7</td>
<td>0.54</td>
<td>0.9</td>
<td>0.1</td>
</tr>
<tr>
<td>Marker watershed</td>
<td>0.69</td>
<td>0.53</td>
<td>0.98</td>
<td>0.04</td>
</tr>
</tbody>
</table>

### Table 4: The obtained results for image 3

<table>
<thead>
<tr>
<th></th>
<th>Dice coefficient</th>
<th>Jaccard index</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proposed method</td>
<td>0.76</td>
<td>0.95</td>
<td>0.88</td>
<td>0.99</td>
</tr>
<tr>
<td>FCM</td>
<td>0.14</td>
<td>0.07</td>
<td>0.99</td>
<td>0.1</td>
</tr>
<tr>
<td>Marker watershed</td>
<td>0.13</td>
<td>0.07</td>
<td>0.98</td>
<td>0.79</td>
</tr>
</tbody>
</table>

### Table 5: The obtained results for image 4

<table>
<thead>
<tr>
<th></th>
<th>Dice coefficient</th>
<th>Jaccard index</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proposed method</td>
<td>0.82</td>
<td>0.95</td>
<td>0.89</td>
<td>0.98</td>
</tr>
<tr>
<td>FCM</td>
<td>0.2</td>
<td>0.11</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>Marker watershed</td>
<td>0.19</td>
<td>0.11</td>
<td>0.94</td>
<td>0.22</td>
</tr>
</tbody>
</table>

### Table 6: The obtained results for image 5

<table>
<thead>
<tr>
<th></th>
<th>Dice coefficient</th>
<th>Jaccard index</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proposed method</td>
<td>0.8</td>
<td>0.93</td>
<td>0.77</td>
<td>0.95</td>
</tr>
<tr>
<td>FCM</td>
<td>0.3</td>
<td>0.17</td>
<td>0.8</td>
<td>0.97</td>
</tr>
<tr>
<td>Marker watershed</td>
<td>0.3</td>
<td>0.17</td>
<td>0.9</td>
<td>0.81</td>
</tr>
</tbody>
</table>

### Table 7: The obtained results for image 6

<table>
<thead>
<tr>
<th></th>
<th>Dice coefficient</th>
<th>Jaccard index</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proposed method</td>
<td>0.93</td>
<td>0.97</td>
<td>0.95</td>
<td>0.99</td>
</tr>
<tr>
<td>FCM</td>
<td>0.3</td>
<td>0.18</td>
<td>0.9</td>
<td>0.1</td>
</tr>
<tr>
<td>Marker watershed</td>
<td>0.29</td>
<td>0.17</td>
<td>0.95</td>
<td>0.17</td>
</tr>
</tbody>
</table>
Dice coefficient:

\[
DC = \frac{2 \cdot |GT \cap CS|}{|GT| + |CS|}
\]  

(8)

Where CS is the lesion segmentation and GT presents the ground truth.

Sensitivity:

\[
Sensitivity = \frac{TP}{TP + TN}
\]

(9)

Where the main objective for sensitivity is computing if a pixel is belonging to the lesion detected.

Specificity:

\[
Specificity = \frac{TP}{FP + TN}
\]

(10)

Where the main objective for specificity is computing if a pixel doesn’t belong to the skin lesion.

Jaccard similarity:

\[
J(img1, img2) = \frac{img1 \cap img2}{img1 \cup img2}
\]

(11)

The Jaccard similarity measure compute the similarity between two images.

5. Conclusion

In this manuscript, we’ve been discussing the proposed process for skin lesion detection. For this, we used a hybridization between the two already existing methods which are fuzzy c-means as well as marker watershed algorithm. Also, we used Multi Otsu’s thresholding. The proposed method was evaluated using about 300 gathered and collected images. The obtained results show very satisfactory results compared to the two other methods.

REFERENCES

A Hybrid Skin Lesions Segmentation Approach Based on Image Processing Methods


