TÉCNICO LISBOA

Principles of Biosignals and Biomedical Imaging

Bioengineering Department (DBE)

MATLAB Project - Image Processing of Dermoscopic Images [EN]

Authors:

Armando Gonçalves (100290) armandogoncalves@tecnico.ulisboa.pt
Patrícia Marques (100357) patriciasmarques@tecnico.ulisboa.pt

2022/2023 – 2º Semester, P3

Contents

1 Background

Skin cancer is one of the most common type of cancer in the world. It occurs when cells in the skin mutate and grow uncontrollably, forming a tumor [11]. These mutations are often caused by exposure to ultraviolet (UV) radiation from the sun or other sources, such as tanning beds. Other risk factors include having fair skin, a history of sunburns, a weakened immune system, and exposure to certain chemicals or radiation. There are several different types of skin cancer, including basal cell carcinoma, squamous cell carcinoma, and melanoma [5].

(a) Basal cell carcinoma [2] **(b)** Squamous cell carcinoma [1] **(c)** Melanoma [4]

Figure 1: Most common types of skin cancer.

Melanoma is the deadliest form of skin cancer, and it is responsible for the majority of skin cancer-related deaths. Unlike other types of skin cancer, which typically develop in sun-exposed areas of the body, melanoma can occur anywhere on the skin [9]. It can also develop in areas that are not exposed to the sun, such as the soles of the feet and the palms of the hands.

One of the challenges of identifying melanoma is distinguishing it from keratosis-like lesions, which are benign growths that can look very similar to melanomas [3]. The **ABCDE rule** is a commonly used method for identifying melanomas.

1.1 ABCDE rule

ABCDE stands for
Asymmetry
Border irregularity
Color variation
Diameter greater than 6 millimeters
Evolution (changes in size, shape, or color over time) [6]

By using this rule, individuals can be more confident in identifying potentially dangerous skin growths and seek medical attention if necessary.

2 State-of-the-Art

Dermoscopy is a non-invasive diagnostic technique that involves the use of a handheld device called a dermatoscope to examine skin lesions for signs of melanoma. It works by illuminating the skin with polarized light and magnifying the area of interest. This allows doctors to see

the structure of skin lesions and identify features associated with melanoma, such as irregular borders, uneven color distribution, and abnormal blood vessels.

To enhance the accuracy of dermoscopy, various computer-aided diagnosis (CAD) systems have been developed in recent years. These systems use machine learning algorithms to analyze dermoscopic images and provide diagnostic assistance to dermatologists.

CAD systems typically rely on deep learning models that are trained on large datasets of dermoscopic images with annotations provided by expert dermatologists. The models learn to identify features associated with melanoma and can make predictions about the likelihood of malignancy for a given lesion [8].

There are various techniques used in deep learning models for melanoma detection, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models that combine both [7]. These models have shown promising results in clinical studies and are being integrated into commercial dermatoscopes to assist in the early detection of melanoma.

3 Methods

In this study, we implement a classic image processing pipeline to classify dermoscopy images into two classes: benign (keratosis-like lesions) and malignant (melanomas). The pipeline consists of the following steps:

3.1 Dataset

We use MATLAB's `imageDatastore` and `readall` functions to import all the .jpg images from our working directory. We then convert the images to grayscale and invert them.

```
% Define the folder path where the images are stored
folderPath = 'Data';

% Create an imageDatastore object to read all the .jpg images from the
    folder
imds = imageDatastore(folderPath, 'FileExtensions', '.jpg');

% Read all the images and convert them to grayscale and invert them
imds = readall(imds);
imds = cellfun(@(x) imcomplement(rgb2gray(x)), imds, 'UniformOutput', false)
    ;

% 'UniformOutput' is set to false to tell cellfun every output of the cell
% may not be uniform in size or shape (can have different data types)

% Display all inverted images in grayscale
figure;
for i = 1:numel(imds)
    subplot(4, 3, i);
    imshow(imds{i});
end
```

3.1.1 Output:

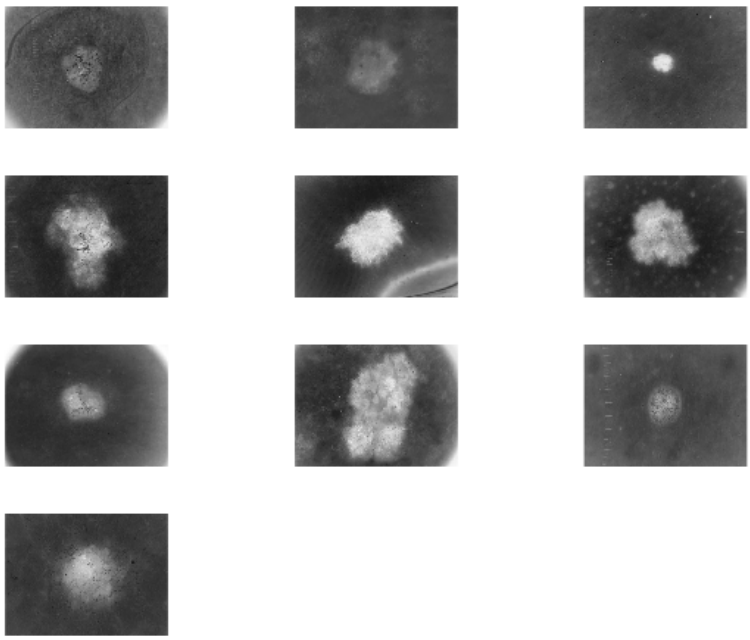


Figure 2: Filtered Images

3.2 Denoising

We use a Gaussian filter to remove noise from the images.

```
% Define the standard deviation and filter size of the Gaussian filter
sigma = 1;
filterSize = 5;

% Apply a Gaussian filter to each image in the cell array
imds_filtered = cellfun(@(x) imgaussfilt(x, sigma, 'FilterSize', filterSize)
    , imds, 'UniformOutput', false);

% Display all the denoised images in a loop
figure;
for i = 1:numel(imds_filtered)
    subplot(4, 3, i);
    imshow(imds{i});
end
```

3.2.1 Output:

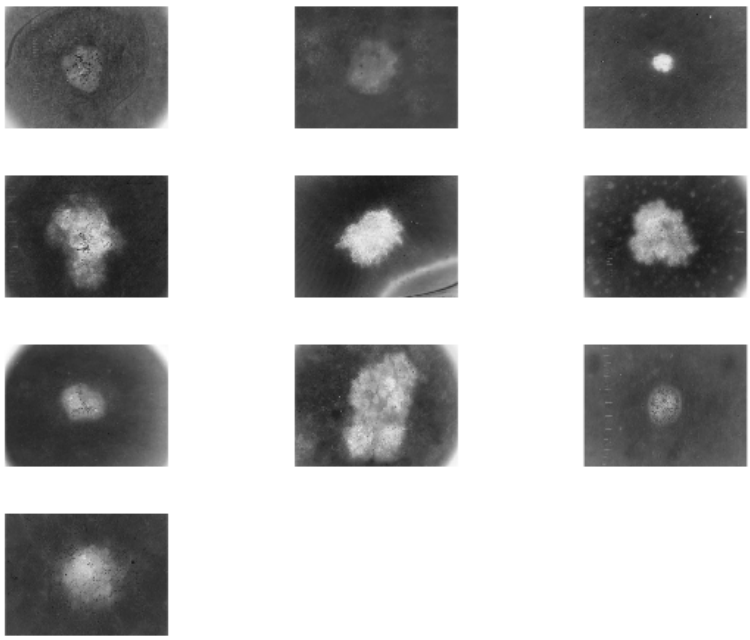


Figure 3: Filtered Images

3.3 Thresholding

We use Otsu's method to obtain binary masks that separate the background (skin, black) from the foreground (lesion, white).

```
% Apply Otsu's method to each image to obtain threshold values
thresholds = cellfun(@(x) graythresh(x), imds_filtered);

% Apply the obtained threshold values to each image to obtain binary masks
imds_binary = cellfun(@(x, t) imbinarize(x, t), imds_filtered, num2cell(
    thresholds), 'UniformOutput', false);

%imshow(imds_binary{1})
% Display all the denoised images
figure;
for i = 1:numel(imds_binary)
    subplot(4, 3, i);
    imshow(imds_binary{i});
end
```

3.3.1 Output:

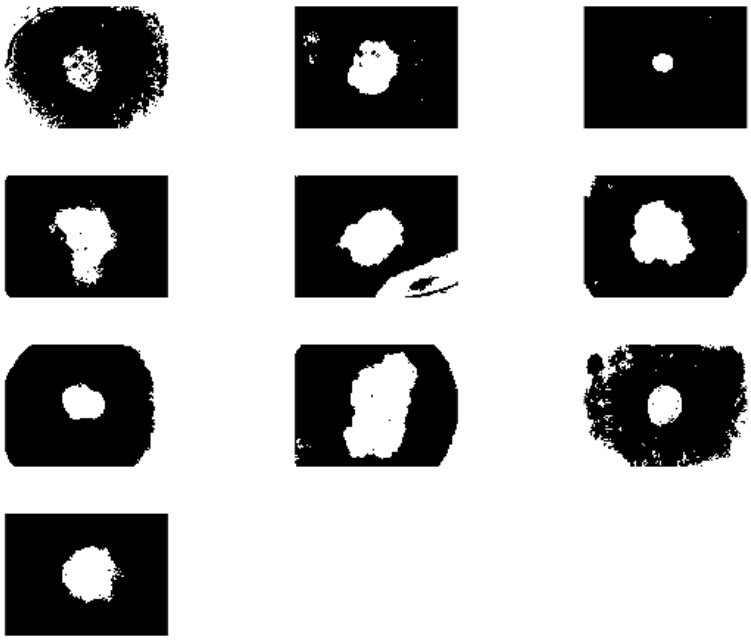


Figure 4: Thresholded Images

3.4 Morphological Operations

We clean the pre-processed masks by performing binary dilation, border clearing, small object removal and hole-filling within the lesion masks if needed.

```
% Apply morphological operations to each binary mask
% disk-shaped structuring element with radius 3
se = strel('disk', 3);
for i = 1:numel(imds_binary)
    % Perform binary dilation
    imds_binary{i} = imdilate(imds_binary{i}, se);

    % Perform border clearing
    imds_binary{i} = imclearborder(imds_binary{i});

    % Perform small object removal (<100 pixels)
    imds_binary{i} = bwareaopen(imds_binary{i}, 100);

    % Perform hole-filling
    imds_binary{i} = imfill(imds_binary{i}, 'holes');

    % Keep only largest connected component
```

```
    imds_binary{i} = bwareafilt(imds_binary{i}, 1);
end

% Show the cleaned and refined binary masks
figure;
for i = 1:numel(imds_binary)
    subplot(4, 3, i);
    imshow(imds_binary{i});
end
```

3.4.1 Output:

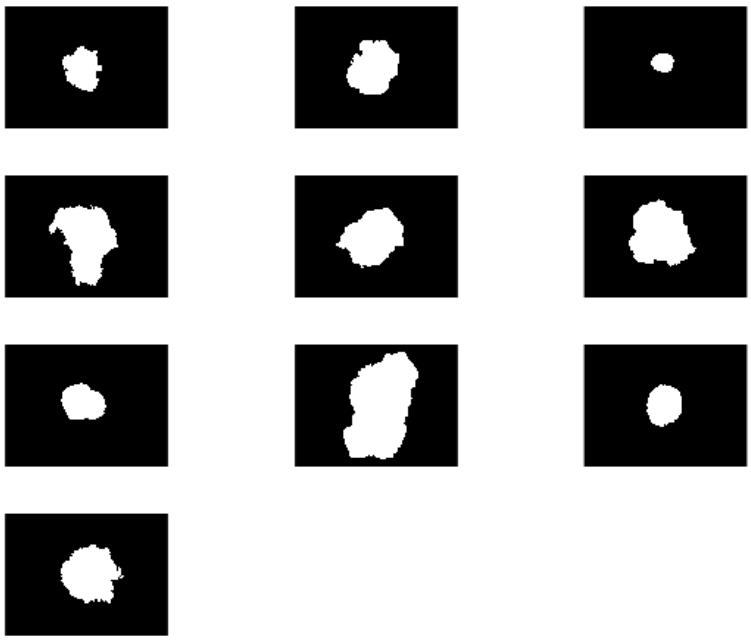


Figure 5: Masked Images

3.5 Feature Extraction

We multiply the grayscaled and inverted images with the final binary masks obtained from the previous step. We then compute the standard deviation of the intensity distribution of each lesion and use `regionprops` to obtain the circularity of each mask.

```
% Loop through each image
for i = 1:numel(imds)

    % Multiply with binary mask
```

```
    imds_masked = double(cell2mat(imds(i))) .* double(cell2mat(imds_binary(i
        )));

    % Compute standard deviation of intensity distribution
    std_intensity(i) = std(double(imds_masked(:)));

    % Use regionprops to obtain circularity of mask
    stats = regionprops('table', imds_binary{i}, 'Circularity');
    circularity(i) = stats.Circularity;
end

% Concatenate the features into a matrix and display it
features = [std_intensity', circularity']
```

3.5.1 Output:

Image	Std deviation	Circularity
"image0.jpg"	24.9175	0.4417
"image1.jpg"	32.7166	0.6175
"image2.jpg"	25.5942	0.9227
"image3.jpg"	55.5259	0.3276
"image4.jpg"	64.3821	0.6384
"image5.jpg"	56.5298	0.6349
"image6.jpg"	32.6795	0.7812
"image7.jpg"	75.1156	0.5627
"image8.jpg"	27.2339	0.8406
"image9.jpg"	47.7586	0.3981

(See note[1]).

3.6 Data Visualization and Analysis

We perform k-means clustering to automatically separate the two classes (benign / malignant lesions). We use `scatter` to plot both features simultaneously in the 2D feature plane along with the corresponding estimated class label.

```
% Set the number of clusters (2 for benign/malignant)
k = 2;

% Perform k-means clustering on the features
[idx, centroids] = kmeans(features, k);

% Plot the results
figure;
```

[1]This is not exactly the output, but to be more explicit about the result, the data was presented as shown in the table.

```
scatter(features(idx==1,1), features(idx==1,2), 40, 'filled', 'DisplayName',
    'Benign','MarkerFaceColor','#77AC30');
hold on;
scatter(features(idx==2,1), features(idx==2,2), 40, 'filled', 'DisplayName',
    'Malignant','MarkerFaceColor','r');
scatter(centroids(:,1), centroids(:,2), 150, 'x', 'Linewidth', 4, '
    DisplayName','Centroids');
legend();
title('Malignant vs. Benign Lesions')
xlabel('Standard Deviation');
ylabel('Circularity');
```

3.6.1 Output:

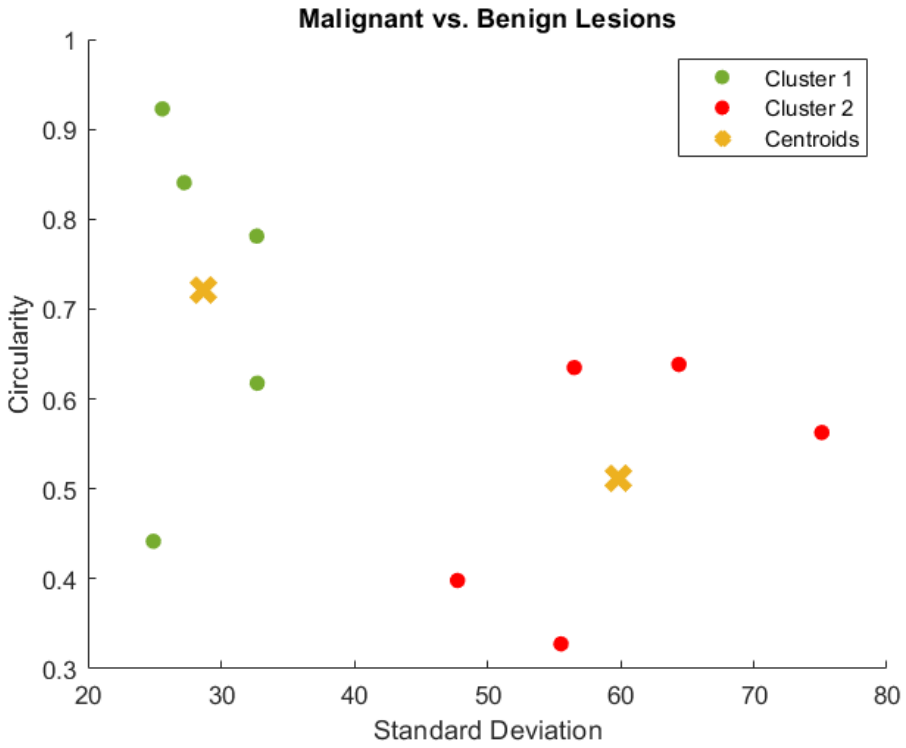


Figure 6: Cluster's plot

4 Results

Based on the results of k-means clustering, the two clusters correspond to benign and malignant lesions. The cluster with a lower standard deviation and higher circularity values (Cluster 1) is most likely to correspond to benign lesions (keratosis-like lesions) while the cluster with a higher standard deviation and lower circularity values (Cluster 2) is most likely to correspond to malignant lesions (melanomas).

This conclusion is based on the following observations:

- Circular lesions, which tend to have higher circularity values, are more likely to be benign (keratosis-like lesions).
- Lesions with a larger spread of color values, which tend to have higher standard deviation values, are more likely to be malignant (melanomas).
- The ABCDE rule 1.1 suggests that malignant lesions tend to have asymmetry, irregular borders, uneven color, larger diameter, and evolving features [10]. Cluster 2, which has lower circularity and higher standard deviation values, is more likely to correspond to malignant lesions based on these criteria.

Therefore, the final plot will be:

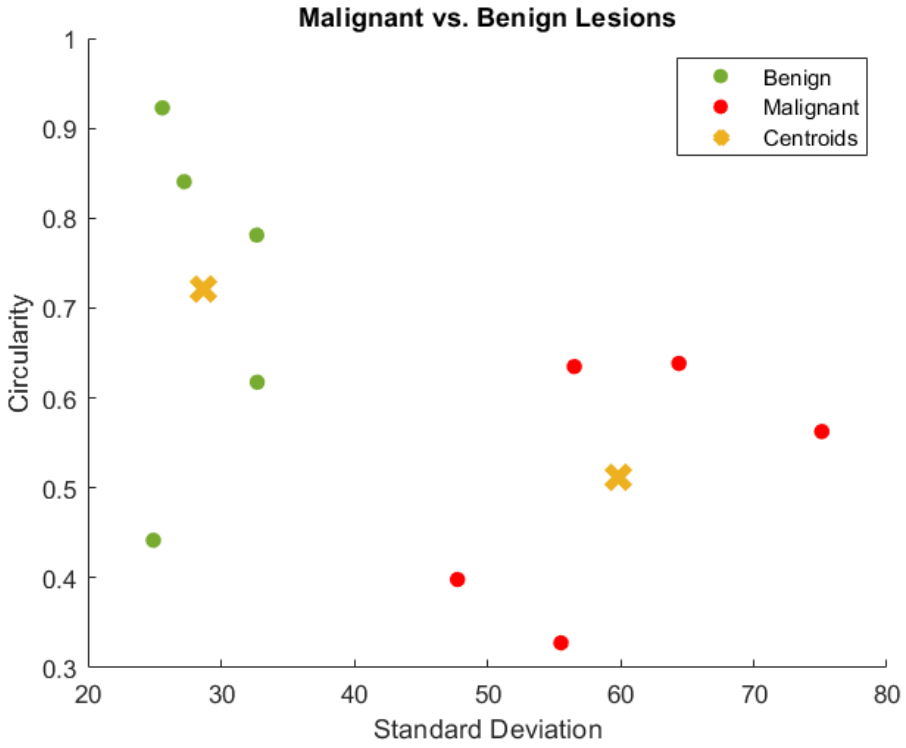
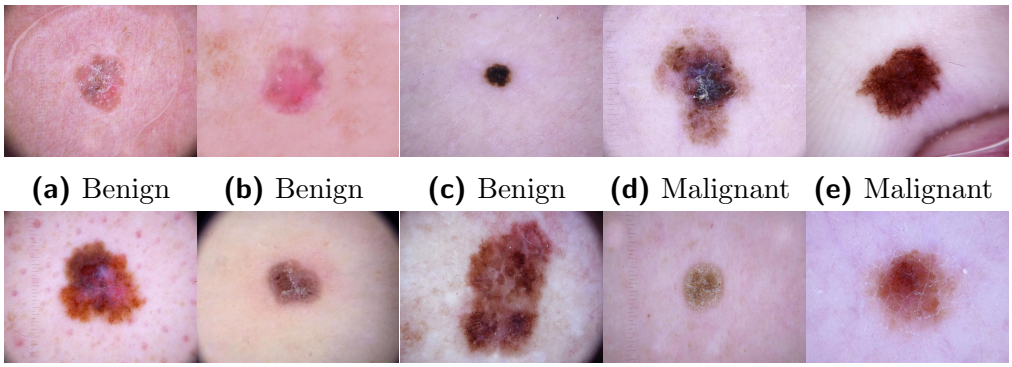


Figure 7: Final result

Together with the table obtained from step 3.5.1, we can finally classify the images with a high degree of confidence, looking at graph 4, as **"benign"** or **"malignant"**:



(a) Benign **(b)** Benign **(c)** Benign **(d)** Malignant **(e)** Malignant

(f) Malignant **(g)** Benign **(h)** Malignant **(i)** Benign **(j)** Malignant

Figure 8: Classification of images in "benign" or "malignant".

5 Final notes

In conclusion, we have used MATLAB to analyze a set of skin lesion images to classify them as benign or malignant based on their standard deviation and circularity.

We first plotted the data points on a scatter plot and identified two distinct clusters, which we further analyzed using the ABCDE rule. Our analysis showed that the cluster with a higher standard deviation and lower circularity corresponds to malignant lesions, while the cluster with lower standard deviation and higher circularity corresponds to benign lesions.

This project demonstrates the potential of using MATLAB for image analysis and classification tasks. By leveraging the power of MATLAB's built-in functions and tools, we were able to quickly and accurately analyze a large dataset of skin lesion images and make meaningful classifications.

Future work could include applying more advanced machine learning techniques to improve the accuracy of our classifications, as well as analyzing additional features of the images beyond standard deviation and circularity. Overall, this project highlights the importance of computer-aided analysis in the field of dermatology and shows the potential for using MATLAB as a tool for image analysis and classification.

References

[1] [Accessed: 13-Apr-2023]. 2020. URL: https://ballaratskincancer.com.au/squamous-cell-carcinoma-scc/.

[2] [Accessed: 11-Apr-2023]. 2022. URL: https://skinsight.com/skin-conditions/basal-cell-carcinoma-bcc/.

[3] American Academy of Dermatology Association. *Actinic Keratosis*. https://www.aad.org/public/diseases/scaly-skin/actinic-keratosis. [Accessed: 13-Apr-2023]. n.d.

[4] Contribuidores dos. *Melanoma*. [Accessed: 13-Apr-2023]. 2006. URL: https://pt.wikipedia.org/wiki/Melanoma.

[5] Skin Cancer Foundation. *Melanoma*. https://www.skincancer.org/skin-cancer-information/melanoma/. [Accessed: 12-Apr-2023]. 2022.

[6] Skin Cancer Foundation. *The ABCDEs of Melanoma*. https://www.skincancer.org/skin-cancer-information/melanoma/the-abcdes/. [Accessed: 13-Apr-2023]. n.d.

[7] James W. Harrison and Steven R. Feldman. "Artificial Intelligence, Skin Cancer, and the Melanoma Epidemic: Pressing the Refresh Button". In: *JAMA Dermatology* 154.9 (2018), pp. 991–992. DOI: 10.1001/jamadermatol.2018.1768.

[8] Alon Scope and Ashfaq A. Marghoob. "Computer-Assisted Analysis of Pigmented Skin Lesions". In: *JAMA Dermatology* 153.3 (2017), pp. 286–287. DOI: 10.1001/jamadermatol.2016.5602.

[9] American Cancer Society. *Melanoma Skin Cancer*. https://www.cancer.org/cancer/melanoma-skin-cancer.html. [Accessed: 13-Apr-2023]. 2022.

[10] American Cancer Society. *Signs and Symptoms of Melanoma Skin Cancer*. https://www.cancer.org/cancer/melanoma-skin-cancer/detection-diagnosis-staging/signs-and-symptoms.html. [Accessed: 12-Apr-2023]. 2022.

[11] American Cancer Society. *What is Skin Cancer?* https://www.cancer.org/cancer/skin-cancer/about/what-is-skin-cancer.html. [Accessed: 11-Apr-2023]. 2022.