



UNIUNEA EUROPEANĂ



IOSUD – “DUNĂREA DE JOS” UNIVERSITY OF GALAȚI

Doctoral School of Mechanical and Industrial Engineering

GALATIENSIS

Project co-financed from the European Social Fund Operational Human Capital 2014-2020

PhD THESIS

ABSTRACT

MODERN TOOLS FOR SKIN MELANOMA DETECTION – SelfChecker application for the analysis of nevi and melanomas using digital images

PhD. student,
Eng. Felicia - Anișoara MICHÎȘ (DAMIAN)
Scientific Coordinator,
Prof. PhD. Eng. Phys. Luminița MORARU

Work completed within the project "Programme for increasing performance and innovation in advanced doctoral and

Postdoctoral research of excellence- PROINVENT"

Contract No: 62487/03.06.2022 POCU/993/6/13 - SMIS Code: 153299

Series I 4: Industrial Engineering No. 89

**GALAȚI
2023**



Parteneri:





UNIUNEA EUROPEANĂ



"Programme for increasing performance and innovation in advanced doctoral and Postdoctoral research of excellence- PROINVENT"

IOSUD – "DUNĂREA DE JOS" UNIVERSITY OF GALAȚI
Doctoral School of Mechanical and Industrial Engineering



PhD THESIS

ABSTRACT

MODERN TOOLS FOR SKIN MELANOMA DETECTION –

SelfChecker application for the analysis of nevi and melanomas using digital images

PhD. student,
Eng. Felicia - Anișoara MICHÎȘ (DAMIAN)

President,

Prof. PhD. Eng. Cătălin FETECĂU
"Dunarea de Jos" University of Galati

Scientific Coordinator,

Prof. PhD. Eng. Phys. Luminița MORARU
"Dunarea de Jos" University of Galati

Scientific referrers

Prof. PhD. Eng. Neculai Eugen SEGHEDEIN
"Gheorghe Asachi" Technical University of Iasi
Prof. PhD. Eng. Cristian Doicin
Polytechnic University of Bucharest
Prof. PhD. Eng. Cătălina ITICESCU
"Dunarea de Jos" University of Galati

Series I 4: Industrial Engineering No. 89

GALAȚI
2023

Lider:



Parteneri:



The PhD thesis series publicly presented in UDJG since October 1, 2013 are:

Major field ENGINEERING SCIENCES

- Series I 1: **Biotechnologies**
- Series I 2: **Computers and Information Technology**
- Series I 3: **Electrical Engineering**
- Series I 4: **Industrial Engineering**
- Series I 5: **Materials Engineering**
- Series I 6: **Mechanical Engineering**
- Series I 7: **Food Engineering**
- Series I 8: **Systems Engineering**
- Series I 9: **Engineering and management in agriculture and rural development**

Major field SOCIAL SCIENCES

- Series E 1: **Economics**
- Series E 2: **Management**
- Series E 3: **Marketing**
- SSEF series: **Sports science and physical education**
- SJ series: **Law**

Major field HUMANITIES

- Series U 1: **Philology - English**
- Series U 2: **Philology - Romanian**
- Series U 3: **History**
- Series U 4: **Philology - French**

Major field MATHEMATICS AND NATURAL SCIENCES

- Series C : **Chemistry**

Major field BIOMEDICAL SCIENCES

- Series M : **Medicine**
- Series F : **Pharmacy**

FOREWORD

With the completion of this stage in my life, I would like to thank those who have guided and supported me in the completion of this work. Finalizing the PhD thesis required going through a PhD program and ensuring a coherent scientific pathway that involves the accomplishment and endorsement of research reports, further study and dissemination of results through the publication of the scientific papers.

The progress made is based on collective effort rather than an individual achievement. So I express my gratitude to Prof. Dr. Phys. Engineer Luminița Moraru, the scientific coordinator of my PhD thesis, who has managed to form around an academic team with a professional vision and humanism for training the future researchers and professors.

I express my gratitude to Prof. Dr. Phys. Engineer Luminița Moraru, whom I thank for her trust, her advice and encouragement, and for the rigor she showed during the PhD program. I would also like to thank Prof. Simona Moldovanu for her patience, professionalism and personal support that contributed to the completion of this thesis.

I am honored to have met an exceptional team of teachers in the Faculty of Science and Environment, who received me with great warmth and understanding. I thank Prof. Dr. Camelia Frigioiu and Prof. Dr. Mihaela Carmen Baroni (posthumously), for involving me in the teaching concerns of the Department for Mathematics and Computer Science, and I thank them for their support throughout this research.

I would like to thank my colleague Engineer Lenuța Pană who supported me during the years of doctoral studies. I express my gratitude to my family, without whose support, I would not have been able to complete this thesis and I would not have found the strength to overcome all the obstacles encountered during my doctoral years.

Galați, 2023

Eng. Felicia-Anișoara MICHÎȘ (DAMIAN)

This summary contains a synthesis of the most eloquent experimental results obtained from theoretical and practical research.

The chapter headings, sub-chapters, figures, tables and bibliographical references are identical to those of the PhD thesis.

CONTENT	Pg. thesis	Pg. abstract
Foreword	III	III
Content (in Romanian)	V	
Content	IX	V
Summary	XIII	
Abstract	XV	
List of figures	XVII	
List of tabs	XIX	
Abbreviations	XXI	
Introduction	XXIII	VIII
1. Motivation for choosing the research topic	XXIV	IX
2. The importance of the research	XXIV	IX
3. Thesis objectives	XXV	IX
4. Dissemination of research results	XXV	X
5. Structure of the scientific paper	XXV	
Introduction (in English)	XXVII	
Chapter 1		
State of the art on methods for microstructural and morphological characterisation of digital dermoscopic images	1	1
1.1 Statistic of melanoma evolution	1	
1.2 Data sets	2	
1.3 Image acquisition	3	
1.4 Characteristics of dermoscopic images	4	
1.4.1 Methods for skin lesions examination	4	
1.4.2 Image features	6	
1.5 Image pre-processing	8	
1.6 Image segmentation	10	
1.7 Classification of skin lesions	11	
1.7.1 Artificial intelligence tools	11	
1.7.2 Mobile phone applications	12	
1.8 Partial conclusions	15	
Chapter 2		
Mathematical methods and models used in the analysis of digital dermoscopic images	17	2
2.1 Methods for processing digital dermoscopic images	17	2
2.1.1 Geometric features	17	
2.1.2 Colour features	19	2
2.1.3 Colour space selection	19	
2.1.4 Quantifying colour components by evaluating the relative colour Histogram	22	
2.1.5 Colours component extraction colour channels	22	
2.1.6 Quantified colours using the statistical parameters and colour Channels	23	
2.1.7 Frequency domain image analysis	23	
2.1.8 Hu moments	24	
2.1.9 Haar waves	25	
2.1.10 Discrete Wavelet Transform (DWT)	26	

2.1.11 Contrast Sensitive Function (CSF)	26	
2.1.12 Noise Visibility Function (NVF)	27	
2.1.13 First order features	28	
2.2 Segmentation methods	28	2
2.2.1 Segmentation based on morphological operators	28	2
2.2.2 Histogram-based segmentation	29	2
2.2.3 Otsu method	30	
2.3 Clustering algorithms	31	3
2.3.1 K-means method	32	3
2.3.2 Fuzzy C-means method	33	3
2.4 Methods for validating the proposed models	34	3
2.4.1 Receiver Operating Characteristics curven (ROC)	34	3
2.4.2 Cross validation method (CV)	35	3
2.4.3 Dispersion Index (ID)	35	4
2.5 Tools for skin lesion classification	35	4
2.5.1 Artificial Neural Networks (ANN)	35	4
2.5.2 Radial basis function neural network classifier (RBFNN)	36	5
2.5.3 The kNN classifier with cross-validation	37	5
2.5.4 Random Forest (RF)	38	6
2.6 Partial conclusions	38	
Chapter 3		
Personal contributions on the extraction and selection of relevant features specific to skin lesions	39	7
3.1 Analysis of skin lesions with significant features	40	7
3.2 Study of skin lesion asymmetry using artificial neural networks	44	9
3.3 Comparison of distance metrics for skin lesion differentiation	49	12
3.4 Selection of relevant features from non-dermatoscopic images for nevi vs. melanomas classification	52	14
3.5 Influence of colour space on skin lesion classification using statistical features	56	
3.6 Invisible watermarking algorithm applied to decomposed dermatoscopic images with discrete wavelet transform	60	
3.7 Partial conclusions	63	
Chapter 4		
Personal contributions on the classification of skin lesions based on colour clusters	65	18
4.1 Analysis of skin lesions with algorithm for selecting geometric features	65	
4.2 Colour histogram analysis for skin lesion differentiation	68	18
4.3 Colour cluster classification with the kNN algorithm	71	
4.4 Fractal dimension and statistical features of the colour cluster for machine learning classification	74	20
4.5 Skin lesion classification using the Random Forest (RF) algorithm	81	26
4.6 Partial conclusions	84	

Chapter 5		
Personal contributions on the implementation of the "Skin Lesions" expert system	85	29
5.1 Requirements for an application	85	29
5.2 The proposed technique	89	30
5.3 Description of the application.....	90	30
5.4 Original contributions	94	34
5.5 Partial onclusions.....	98	
Chapter 6		
General conclusions and future research directions	99	37
List of published works	101	39
References	107	44
Appendices	127	

Introduction

Clinical and statistical research has confirmed that melanoma is the most dangerous form of skin cancer, with a rapid evolution which causes a high mortality rate in the population worldwide. Due to the increased incidence of melanoma, new approaches need to be developed to accurately and correctly differentiate skin lesions.

The topic proposed by this PhD thesis is one of real interest as SelfChecker applications for skin lesion analysis offer a modern technological approach to help people with suspicious skin lesions decide whether to seek further medical attention.

As a result of the evolution in mobile phone technology and the development of appropriate digital dermatoscopic image processing techniques, there has been a significant increase in interest in recent years in the development of an automated expert system capable of assessing skin lesions. This automated system may be an important step in the treatment of melanoma based on early detection of the affected area. The final diagnosis must be made by the medical specialist, working with specialists from technical fields, so that the expert system “**Skin Lesions**” developed is as effective as possible and early detection of melanoma is achieved in a shorter time span.

The development of high-performance expert systems is closely linked to research results in the field of digital dermoscopic image processing and analysis. Thus, once the digital dermoscopic image is acquired, this must be processed and analysed in real time, as speed of diagnosis in melanoma is sometimes vital.

The present PhD thesis aims to develop an expert system “**Skin Lesions**” for the analysis of skin lesions using digital images. In order to achieve the aim of the PhD thesis, the following steps have been taken:

- literature review on the use of methods for automatic information processing and interpretation of digital dermatoscopic images;
- acquisition of skin lesion images using a mobile device able to generate a digital image at its output;
- the selection and description of features used to investigate useful information contained in the images;
- pre-processing of digital images based on enhancement algorithms, removal of artefacts and extraction of areas of interest based on linear, non-linear, wavelet or Fourier transform filtering algorithms;
- image segmentation;
- description, recognition, selection and classification operations of object features extracted in the segmentation process;
- analysis of experimental data and interpretation of the listed operations performed with different quality metrics or advanced statistical methods;
- implementation of innovative diagnostic methods based on colour clusters;
- securing digital images;
- implementation of the expert system “**Skin Lesions**” as a health care mobile phone application that identifies whether a skin lesion may be a cause for concern and displays an alert.

In this work, we reviewed the literature on digital dermoscopic image processing and evaluation and developed original processing methods that were used to create the Skin Lesions expert system.

1. Motivation for choosing the research topic

As mobile devices are used by billions of people all over the world and these are equipped with high resolution image sensors and powerful processors it is possible to design based on them new expert systems. The functions of mobile devices combined with advanced digital image processing algorithms can provide solutions for monitoring health and tracking the evolution of certain medical parameters.

The expert system “**Skin Lesions**” developed uses proprietary algorithms, which can also be used for educational purposes as an e-Learning opportunity for students. I believe that the scientific results obtained are benchmarks, for future research in the medical and industrial fields. At the end of the 5 years of study we have implemented an expert system for an objective evaluation of skin lesion features.

2. The importance of the research

Based on the topic of my PhD thesis, I proposed a series of original algorithms whose efficiency in differentiating skin lesions has been experimentally validated. For this, during the achievement of this scientific thesis, I followed step by step the entire process concerning the processing of digital dermatoscopic images. I have also taken into account that the functions of a mobile phone combined with advanced algorithms of digital image processing offer solutions for monitoring and tracking the evolution of certain parameters of skin lesions. Thus, efficient algorithms were developed and implemented in the proposed expert system:

- analysis of melanomas and nevi using geometric features and first-order features;
- the influence of color space on the classification of skin lesions using their features that can provide useful information about their structure;
- selecting relevant features from images for classification of nevi vs. melanomas;
- histogram analysis to investigate and classify skin lesions;
- classification of skin lesions using the k-NN algorithm and based on 2D fractal dimensions;
- determining the lesion contours;
- developing and implementing the expert system “**Skin Lesions**”.

To process the digital images and to obtain the experimental results we used the Matlab 2017a environment with the libraries: Graphical User Interface, Image Processing Toolbox, Wavelets, nntool. For the analysis of characteristic parameters we used the statistical software package MedCalc designed for statistical applications in medical laboratory. For designing the expert system „Skin Lesions” we chose the Android platform, this being a very popular operating system.

By publishing articles and presenting the results obtained during the PhD research at scientific conferences, we validated the experimental data obtained.

3. Thesis objectives

The aim of the research is to develop a patient-oriented SelfChecker application for skin lesion analysis from digital images using mobile devices.

The overall objective of the thesis is to develop an expert system „Skin Lesions” for automatic analysis of images acquired with a mobile device that ultimately, displays an alert if a skin lesion may be a cause for concern.

In accordance with the general objective of the PhD thesis "MODERN TOOLS FOR SKIN MELANOMA DETECTION - SelfChecker application for the analysis of nevi and melanomas using digital images" and the scientific reports submitted during the training period, the following specific objectives were proposed and achieved:

- improvement in melanoma detection and classification techniques;
- the use of mobile devices in medical systems and the development of computer-aided diagnosis (CAD);
- improvement of image processing methods for melanoma feature extraction.

4. Dissemination of research results

The research results of the PhD traineeship have been materialised in 13 scientific papers (indexed in ISI Web of Knowledge (Clarivate) and Scopus, IEEE Xplore and BDI/B+ databases indexed in EBSCO database). Among these, 2 papers were published in ISI refereed journals, 3 papers were published in ISI Proceedings volumes, 3 papers were published in international conference volumes and 3 papers were published in BDI refereed journals.

The novelty of this PhD thesis lies in the evolutionary, processual approach, based on the latest information, theoretical foundations, principles and methods supporting the process of skin lesion analysis. The methods proposed in chapters 3, 4 and 5 can be appreciated as new methods of differentiation of skin lesions, as well as the expert system materialized by the application „Skin Lesions". The thesis problem has been structured in such a way that the scientific novelty obtained from the research process and the practical value of the work are highlighted in several elements presented in the final chapter.

Chapter 1

Current status of research on methods for microstructural and morphological characterization of digital dermoscopic images

The aim of this chapter: to review the scientific literature on methods for microstructural and morphological characterization of digital dermoscopic images.

Starting from the study of algorithms reported in various scientific papers, the contribution made has resulted in the development of one's own algorithms. They can quickly and efficiently differentiate skin lesions and have been implemented in the expert system "**Skin Lesions**".

In short the summary of this chapter consists in the theoretical presentation of the following:

- An analysis of statistics on the evolution of melanomas underlines that melanoma is one of the most dangerous skin cancers in terms of death rate. The probability of death increases when the skin lesion is diagnosed too late. For the development of the proposed expert system an important step is the recognition of skin lesions with specific datasets available online and free of charge, such as MedNode, ISIC archive, Ph2, 7-POINT. There are also paid datasets such as "Dermofit", the interactive dermoscopy atlas. In the scientific literature, there are adopted screening methods that include reliable and standardized techniques which improve the early detection rate of melanoma. These methods support the process of self-examination of skin lesions: structural analysis [36], the ABCDE method [37], the 7-point checklist [38], or the Menzies [39], CASH [40], Chaos and Clues [41] and BLINCK [42] methods.
- For the implementation of the proposed expert system we chose the ABCD method because it is compatible with the computer implementation of, the SelfChecker application in the Matlab programming environment. In this way, the features of skin lesions can be evaluated which help us to develop support systems able to signal the severity of the situation and prompt us to see a specialist. The acronym ABCD refers to four parameters: asymmetry, contour irregularity (border), color variation and skin lesion diameter greater than 6 mm. These parameters provide a simple means of assessing skin lesions.
- The research reviewed in the field of image processing shows that the field is still in need of innovative ideas, both in terms of image processing software and hardware solutions that acquire digital dermoscopic images without defects, artefacts or noise.
- Mobile phone apps for early melanoma diagnosis are graphically friendly and easy to use to support people with skin lesions in deciding whether or not to seek further medical attention: DermLite app [116], Skin Cancer App - MySkinPal - Map your skin moles [118], MoleScope app [119], ApreSkin you app [120], Derma Analytics app, Miiskin app [122], "APD Skin Monitoring" app [123], SkinVision app [124], „SkinMD" app [127], „eSkin" app [112], etc.
- The Otsu image thresholding method [85] is based on image variance maximization and it's a commonly used method for skin lesion segmentation. In all fixed thresholding techniques, a cutoff value is selected based on which the image is converted from gray scale to binary.
- The classification of skin lesions when implementing the expert system „Skin Lesions" is done using the total dermoscopy score - TDS.

Chapter 2

Mathematical methods and models used in the analysis of digital dermoscopic images

Image analysis can be performed with mathematical methods and models that are designed to simplify images from complex structures into simple elements. Mathematical methods and models correlate objects of interest in images with regular or irregular geometric shapes.

The synthesis of this chapter is made up of a theoretical presentation of the principles and mathematical descriptions of certain methods and models for processing digital dermoscopic images, descriptions of segmentation methods based on contours, histograms and regions, clustering methods and data validation methods.

2.1 Digital dermoscopic image processing methods

Features successfully used in highlighting information from medical images are as it follows:

- geometric features such as asymmetry, compactness, circularity, eccentricity;
- color features;
- first order features such as skewness, kurtosis.

2.2 Segmentation methods

Since an important step in image analysis is to highlight certain objects or regions of interest (ROI) in the image, methods are used to segment the areas of interest by determining the boundaries of the objects under analysis. This can be used:

2.2.1 Segmentation based on morphological operators

The principle of morphological operators used for contour extraction is based on measuring the differences between the extreme values (minimum and maximum) of the neighbourhood of the current point; if the difference between these values is large enough it means that the current point is a contour point, being in a transition zone of pixel values.

2.2.2 Histogram-based segmentation

Scrolling the histogram pixel by pixel requires control of the grey levels so that the histogram function, which is a probability function, will satisfy the condition [162]:

$$\sum_{i=0}^{L-1} h(i) = 1 \quad (2.62)$$

where L is the number of grey levels of an image.

The color histogram is a method of describing the color content of an image and quantifies the number of occurrences of each color in an image [163]. For a given RGB image, let L be the intensity levels in the range [0, 1, 2,..., L-1]. The probability distribution can be defined as:

$$p_i^C = \frac{h_i^C}{N}, \sum_{i=0}^{L-1} p_i^C = 1 \quad (2.63)$$

where i is a specific intensity level in the interval $\{0 \leq i \leq L-1\}$ for the color component,

$C = \{R, G, B\}$, N is the total number of pixels in the image and h_i^C is the number of pixels for intensity level i in the component C [164].

2.3 Clustering algorithms

Image clustering uses classification algorithms that group similar pixels in the image into clusters [167]. Clustering pixels into clusters is based on the principle of minimizing intra-cluster similarity and maximizing inter-class differences [168].

2.3.1 K-means method

The K-means method [172] involves grouping objects in a dataset into multiple clusters, each cluster containing a set of objects from a particular category.

2.3.2 Fuzzy C-means method

The Fuzzy C-means method [173] is a clustering technique that allows an item in the dataset to belong, with a certain degree of membership, to one or more clusters. In non-fuzzy analysis, or hard clustering, the information is divided into clusters where each entity belongs to a single cluster [174]. The association of an object with a cluster is done using membership degrees:

2.4 Methods for validating the proposed models

Validation of the proposed method is a mandatory step in image analysis. It consists of verifying the accuracy of the method and can be performed using different methods:

2.4.1 ROC curve (Receiver Operating Features)

The ROC curve [175] is a two-dimensional curve plotting the sensitivity (on the Ox axis) vs. false positive rate (Oy axis) and specificity, respectively, where: Sensitivity = $TP/(TP+FN)$, Specificity = $TN/(TN+FP)$, where TP - Positive prediction, positive response, TN - Negative prediction, negative response, FP - Positive prediction, negative response, FN - Negative prediction, positive response.

The ROC curve is a useful test that checks with the area under the curve (AUC) how effective a feature is (the larger the area under the curve the higher the accuracy of the test, an area of 1 represents a perfect feature and if the $AUC < 0.6$ the feature becomes irrelevant) [176]. Another statistical interpretation of the ROC curve, that helps to determine the maximum cut-off point, is the Youden Index (J) which is defined as [177]: $J = \text{maximum}(\text{Sensitivity} + \text{Specificity} - 1)$. Lower J values indicate that the feature under analysis has no relevance for the proposed study [178].

2.4.2 Cross validation (CV) method

The data evaluation algorithm using the cross-validation method assumes the existence of n data sets structured as follows, $n - 1$ sets are used for training and the n th set for testing. The algorithm is repeated k times, (usually $k = 10$). At each step, the effect of adding or removing a relevant item for classification is evaluated. When this operation produces no further improvement in the results then the algorithm stops [179].

2.4.3 Dispersion index (ID)

This method quantifies whether a data set is clustered or dispersed compared to a standard statistical model:

If $ID = 0$ for the data there is no dispersion, if $ID \in (0,1)$ the data is under-dispersed and if $ID > 1$ the data is dispersed.

The dispersion index is calculated according to the formula [180]:

$$ID = \frac{\sigma^2}{\mu} \quad (2.89)$$

2.5 Tools for skin lesion classification

CAD classification systems require dedicated image processing algorithms to provide mathematical descriptions of suspected skin lesions.

2.5.1 Artificial Neural Networks (ANNs)

ANNs consist of a set of linked nodes that collaborate to solve problems. ANNs consist of an input level, an output level and a hidden level. The input nodes retrieve information in numerical form. The information is presented as activation values, where each node is assigned a number, so the higher the number, the higher the activation. In the case of a feed-forward network, information is forwarded throughout the network [182]. The activation value is passed from node to node, but weighted in a certain way. The activation values are reviewed at each node. The activation flows through the network, through hidden layers, until it reaches the output nodes. The difference between the predicted value and the actual value (error) will be propagated back throughout the network. The transfer function converts input signals into output signals, the sigmoid transfer function is used [183]. Each training network is a pair of the form $\{\bar{X}, \bar{Y}\}$, where X is a vector containing the input data values and Y is a vector of target values (desired values). To evaluate the performance of the proposed method, the following metrics are calculated from the confusion matrix [184].

$$\text{Sensitivity (SE)} = \frac{TP}{TP + FN} \quad (2.90)$$

$$\text{Specificity (SP)} = \frac{TN}{TN + FP} \quad (2.91)$$

$$\text{Accuracy (ACC)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.92)$$

$$\text{Precision (P)} = \frac{TP}{TP + FP} \quad (2.93)$$

$$\text{Dice coefficient (D)} = \frac{2TP}{2TP + FP + FN} \quad (2.94)$$

$$\text{Jaccard index (JAC)} = \frac{TP}{TP + FN + FP} \quad (2.95)$$

where TP is true positive, FP false positive, TN false true negative and FN false negative.

The simplest and yet most commonly used error function whose purpose is to estimate the performance of neural networks is the mean squared error (MSE) [185].

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.96)$$

where n is the number of variables, Y is the vector of variables.

2.5.2 Radial basis function neural network (RBFNN)

The RBF neural network is a three-layer feed forward neural network architecture. Each hidden node puts into action a nonlinear activation function, which is a radially symmetric function. The radial basis function is centered on a vector in the feature space and its response varies monotonically with distance from the center point. RBFNN can be formulated as the minimization of the MSE function. A critical point for the implementation of an RBF network is to establish the RBF centers and determine the weights. Once the RBF centers and nonlinearities in the hidden layer are determined, the weights are calculated based on linear regression of the hidden layer outputs to the desired outputs or target values. The RBFNN classifier is trained and validated using cross-validation to strengthen the capabilities of the predictive model [188]. This approach separates datasets into training and testing groups and avoids overfitting.

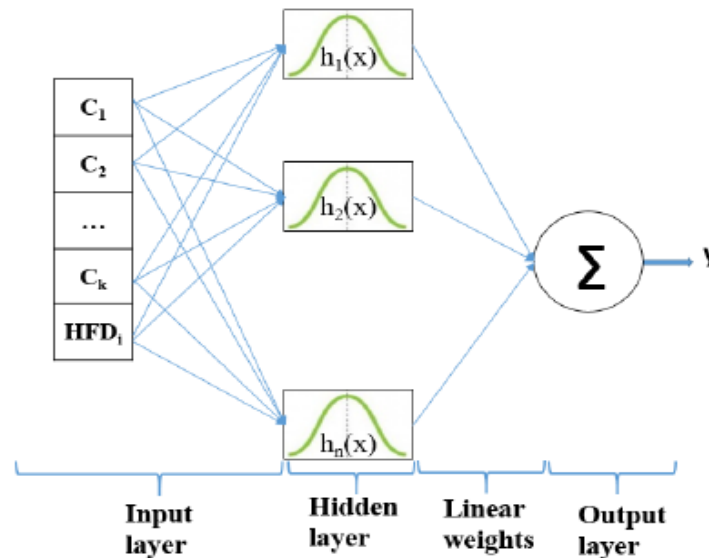


Figure 2.4 Structure of an RBFNN classifier [187].

2.5.3 The kNN -cross-validation

The kNN assumes that samples from each cluster are predominantly surrounded by samples from the same cluster [189]. The training stage preserves the features and class label of the training samples. Then, kNN classifies new instances based on a similarity operation. Euclidean distance is the most commonly used similarity measure. When only a small dataset is available, the cross-validation technique is an alternative solution to the augmentation technique. In addition, the sampling cost is low. kNN-CV with a 5-fold cross-validation algorithm uses a subset of the data for validation, but still uses all the data for the test phase. The input data is divided into five subsets. The algorithm is trained on four subsets and tested on one subset of data. Then this process is iteratively repeated until each subset of data becomes a test subset and all data is evaluated.

The training time complexity is $O(d \times n \times \log n)$, where the number of samples in the training dataset is denoted generically by n and the dimensionality of the data by d . The

complexity is average. For 5-fold kNN cross-validation, we repeat the calculations and increase the time complexity. This relatively higher computation time could be interpreted as a limitation of the kNN-CV algorithm. However, this drawback is compensated by the reduction of the overfitting influence. Moreover, the correct estimation of the test error is its main advantage. It also provides a good level of accuracy for the model [189].

2.5.4 Random Forest (RF)

RF is an algorithm that creates decision trees on datasets and then obtains the prediction for each of them and finally selects the best solution [190]. Breiman [191] mathematically described RF as a classifier based on a family of classifiers $h(x|\Theta_1), \dots, h(x|\Theta_k)$ based on a classification tree with parameters Θ_k chosen randomly from a random vector model Θ . For the final ranking $f(x)$ (which combines the classifiers $\{h_k(x)\}$), each tree votes for the most popular class X , on entry and the class with the most votes wins. Specific data $D = \{(x_i, y_i)\}_{i=1}^n$ we train a family of classifiers $h_k(x)$. Each classifier $h_k(x) \equiv h(x|\Theta_k)$ is in our case a predictor n , $y = \pm 1$ = the outcome associated with the input x .

The Random Forest model is used in machine learning problems in order to determine the importance of input features.

Chapter 3

Personal contributions on the extraction and selection of relevant features specific to skin lesions

The objective of this chapter focused on the idea of highlighting those relevant features extracted from skin lesion images that have increased ability to differentiate and classify them by implementing specific algorithms. In order to achieve this goal, the results provided by new processing methods were presented and discussed.

3.1 Analysis of skin lesions with significant features

We conducted a study where we used a set of features such as asymmetry, skewness, kurtosis, compactness, circularity, eccentricity [193]. The mathematical approaches of the methods and features are presented in Chapter 2. The study contains two important steps:

- image segmentation with a classical binarization method;
- extraction of first-order and geometric features.

For this study, image segmentation was performed with the Otsu segmentation method.

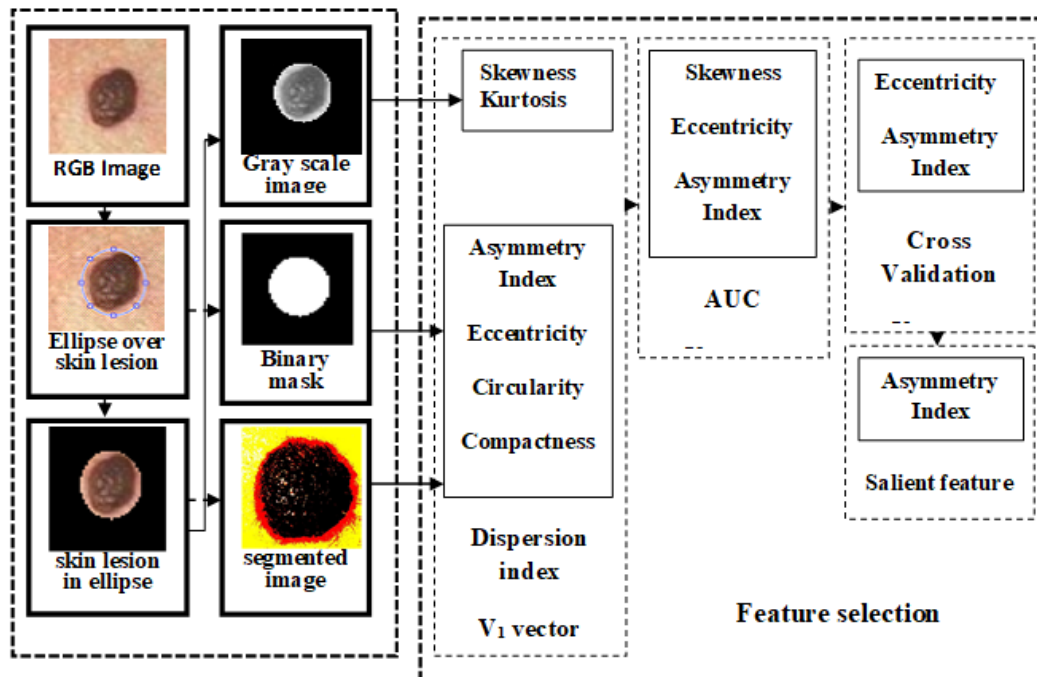


Figura 3.5 Overview of the study

Initially, the feature vector $V1 = \{AI, S, K, COMP, CIRC, E\}$ was formed (Table 3.1). For each feature the ID was calculated. Based on the values obtained for the ID and in accordance with the statistical thresholds set by it, those features with a high degree of dispersion ($ID > 1$) were removed from the vector $V1$ and those with a lower degree of dispersion $ID < 1$ were kept. A second vector of ID-validated features was obtained and denoted by $V2 = \{AI, S, E\}$ (Table 3.2). The stability of the features forming the vector $V2$, was determined using the area under the curve (AUC). The result of the analysis allows the construction of a third vector $V3 = \{IA, E\}$ consisting of features for which $AUC > 0.6$.

Personal contributions on the extraction and selection of relevant features specific to skin lesions

Tabelul 3.1 ID for selected features in V1.

The smaller average values of the ID select the relevant features

Original image	AI	S	K	Comp	Circ	E
Nevi	0.0109	0.178	5.461	24.947	15.559	0.079
Melanoma	0.091	0.307	8.467	4.308	19.148	0.080

Tabelul 3.2 Average AUC and J for features in V2

AUC/ skin lesion	AI	S	E
AUC/nev	0.994	0.547	0.625
J/nev	0.969	0.3582	0.615
AUC/ melanom	0.949	0.532	0.761
J/ melanom	0.897	0.3831	0.666

AUC values indicate that only features IA and E meet the AUC > 0.6 stability condition. Feature S is excluded. The thresholds that were associated with the best Youden indices are shown in Table 3.3 and Figure 3.6. In the proposed study, J was used to interpret the features in V2. As only IA and E are stable, we analyzed the J indices correlated with them.

From the data analysis we observe J = 0.969 for AI and J = 0.615 for E (for nevus), J = 0.889 for AI and J = 0.666 for E (for melanoma). As both values are closer to 1 than to 0 we resort to validating them by the K-fold cross-validation (CV) method.

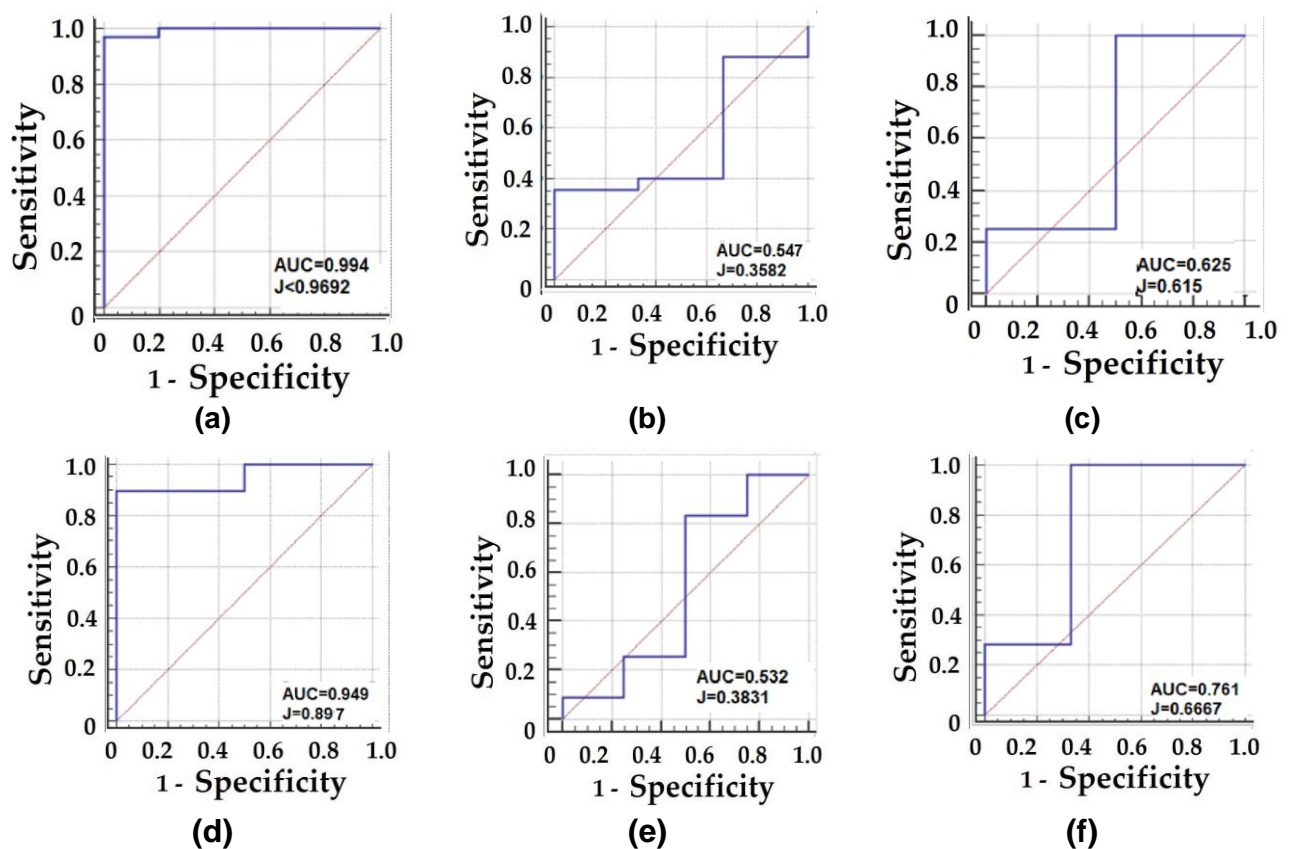


Figure 3.6 ROC curves and J index corresponding to the selected features for non-melanoma (top line) and melanoma (bottom line) skin lesion.

First column - AI; Second column - S; third column - E

Personal contributions on the extraction and selection of relevant features specific to skin lesions

CV estimates the misclassification percentage of the feature values in V3 after 10 iterations of the algorithm on the two data sets, training and test. Figure 3.7 shows the evolution of CV-errors as a function of the number of repetitions. We find that the CV-error is between 0.89 and 2.3 for feature E and 0.68 and 3 for feature IA.

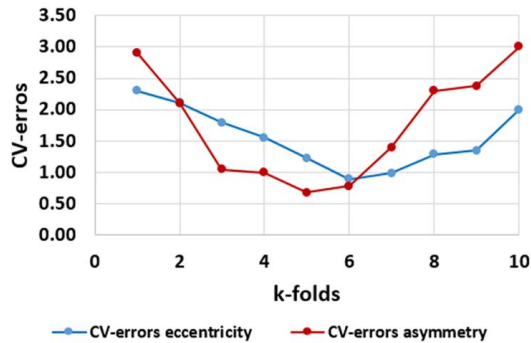


Figure 3.7 Evoluția eroare CV pentru E și AI.

Finally, the AI feature is selected as the most relevant in differentiating skin lesions, with a minimum of 0.68. Furthermore, the AI feature has $J = 0.89$ for nevi and $J = 0.96$ for melanoma.

3.2 Study of skin lesion asymmetry using artificial neural networks

Since asymmetry is the most relevant feature in differentiating of skin lesions, the proposed study analyses the asymmetry using two methods: determining asymmetry by projection of the image on the principal axes (GAF) and from histogram projections (AHP). Dermoscopic images from two databases, i.e., MED-NODE and PH2 were used. The RGB images were binarized with the Otsu method (Figure 3.8). The output of the algorithms are the asymmetry values by the two methods.

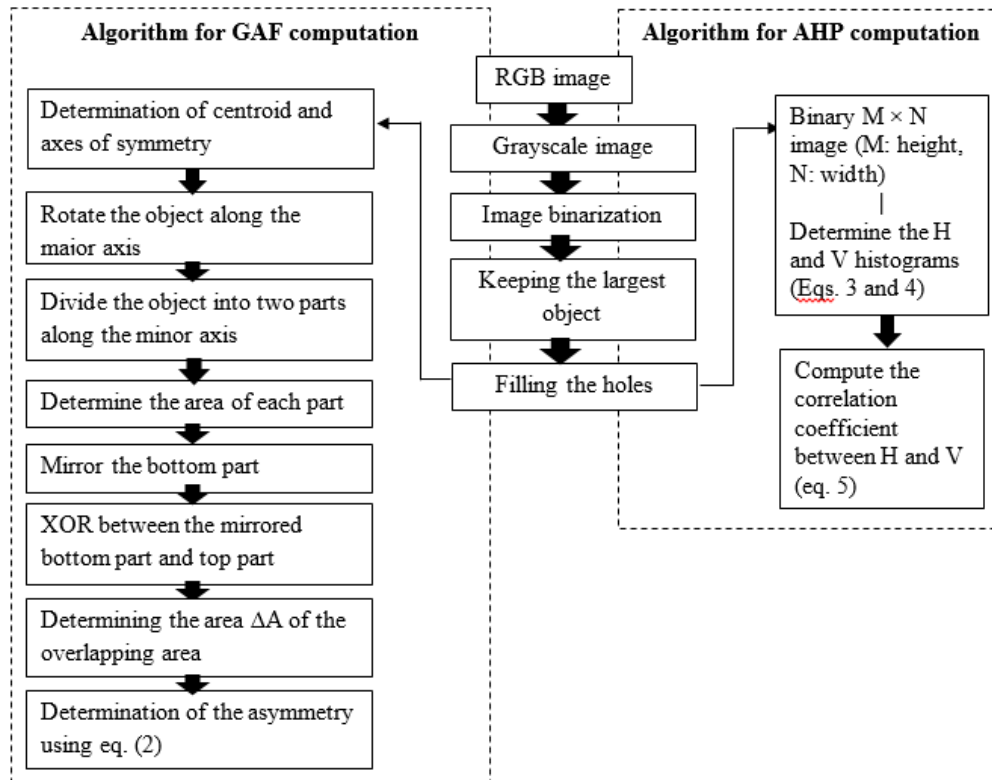


Figure 3.8 Algorithms for asymmetry calculation.

The asymmetry values, calculated using both methods, were divided into 70% training data, 15%, test data and 15% validation data. The network type selected for training was feed-forward backpropagation. The Levenberg-Marquardt back propagation (LMBP) algorithm is used for training the network. The training function updates the comparison values according to the Levenberg-Marquardt optimization. The learning function used was Learngdm. After creating the network, the next step is to train the network. A two-layer ANN was developed (Figure 3.11), a hidden layer and an output layer for bidirectional transmission and 10 neurons in the hidden layer using the scaled conjugate gradient procedure for training. As input the data, were used the GAF asymmetry values calculated for nevi and melanomas from the MED-NODE database. As output data, the AHP asymmetry calculated for nevi from MED-NODE were used.

ANN was created in Matlab with the purpose of performing the following operations:

- import input and output data into the ANN app from the Matlab environment;
- create the network;
- determine percentages of data used for training, validation and testing;
- calculate R and MSE statistical values.

Tabelul 3.3 Features of the constructed network

Feature	Name
Training algorithm	LMBP
Learning function	Learngdm
Square mean error, regression coefficient	MSE, R
Number of layers	2
Number of neurons on the hidden layer	8, 10, 12, 14

Tabelul 3.4 Input data and output data

Training options	Network training		
	Input data Lesion type / database / algorithm		Output data Lesion type / database / algorithm
V1	Nevi/MED-NODE/ GAF	V1	Nevi /MED-NODE/ GAF
V2	Nevi /MED-NODE/ GAF	V2	Naevus/ MED-NODE/ GAF
V3	Nevi / PH2/ GAF	V3	Nevi / PH2/ GAF
V4	Nevi / PH2/ GAF	V4	Nevi /PH2/ GAF

After training, the data were validated with the different network architectures obtained by changing the number of neurons in the hidden layer to the values 8, 10, 12 and 14. Table 3.5 shows the correct classification score (R and MSE) according to the number of neurons in the hidden layer. The training of the network stops automatically when an increase in the mean squared error (MSE) of the validation samples is obtained. The mean squared error (MSE) is the mean squared difference between the outputs and the desired results. The lower values of the MSE are the ones that are targeted. Regression analysis R is performed to measure the

correlation between outputs and desired outcomes. R close to 1 indicates a good correlation. Table 3.5 shows that the proposed training options V1 and V3 are robust for our purposes. Model V1 has lower MSE for 8 hidden neurons and is considered the optimal predictive model for the MED-NODE database. Model V3 has lower MSE for a structure with 12 hidden neurons and is also considered an optimal predictive model for the PH2 database. Models V2 and V4 are removed from our analysis.

Table 3.5 Training phase: MSE and R values and number of neurons in the hidden layer

Baza de date	Model	Număr de neuroni	MSE	R
MED-NODE	V1	8	0.003	0.89
		10	0.005	0.84
		12	0.004	0.86
		14	0.025	0.27
	V2	8	0.008	0.61
		10	0.006	0.79
		12	0.005	0.48
		14	0.004	0.36
PH2	V3	8	0.198	0.38
		10	0.022	0.62
		12	0.021	0.83
		14	0.037	0.69
	V4	8	0.030	0.52
		10	0.085	0.46
		12	0.069	0.34
		14	0.042	0.43

The ANN performance as a function of training, validation and test data is shown in Figure 3.12. This shows the interpretation of the mean squared error (MSE) as a function of the number of epochs, for the two variants for which the MSE is minimum i.e. V1 and V3. The MSE decreases after several training epochs, but may start to increase on the validation dataset as the network starts to match validation data with training data.

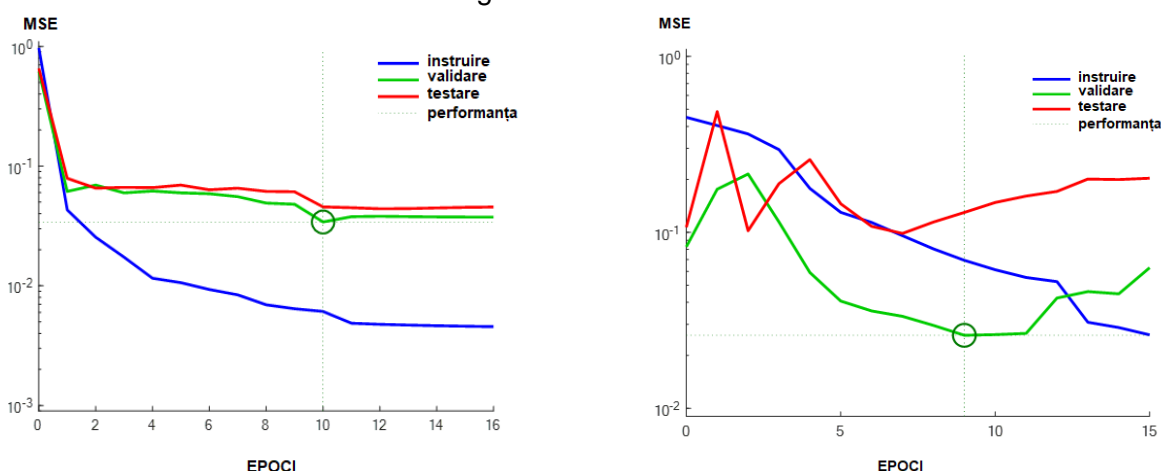


Figure 3.12 Best ANN model performance as a function of variation of MSE values according to the number of epochs (iterations) (a) V1; (b) V3.

The best performance was obtained at epoch 10 for V1 and at epoch 9 for V3. The MSE values are 0.003 (RMSE = 0.055) for V1 and 0.021 for V3 (RMSE = 0.445).

Finally, the V1 model was considered the optimal model for which the accuracy of the method is determined by the lowest MSE and the highest R.

3.3 Comparison of distance metrics for skin lesion differentiation

Since asymmetry and diameter can be calculated with good accuracy [181] we conducted a study on the effectiveness of implementing new distance metrics (Euclidean distance, quasi-Euclidean distance, city-block distance and chessboard distance) for skin lesion assessment [195]. The proposed algorithm is shown in Figure 3.13

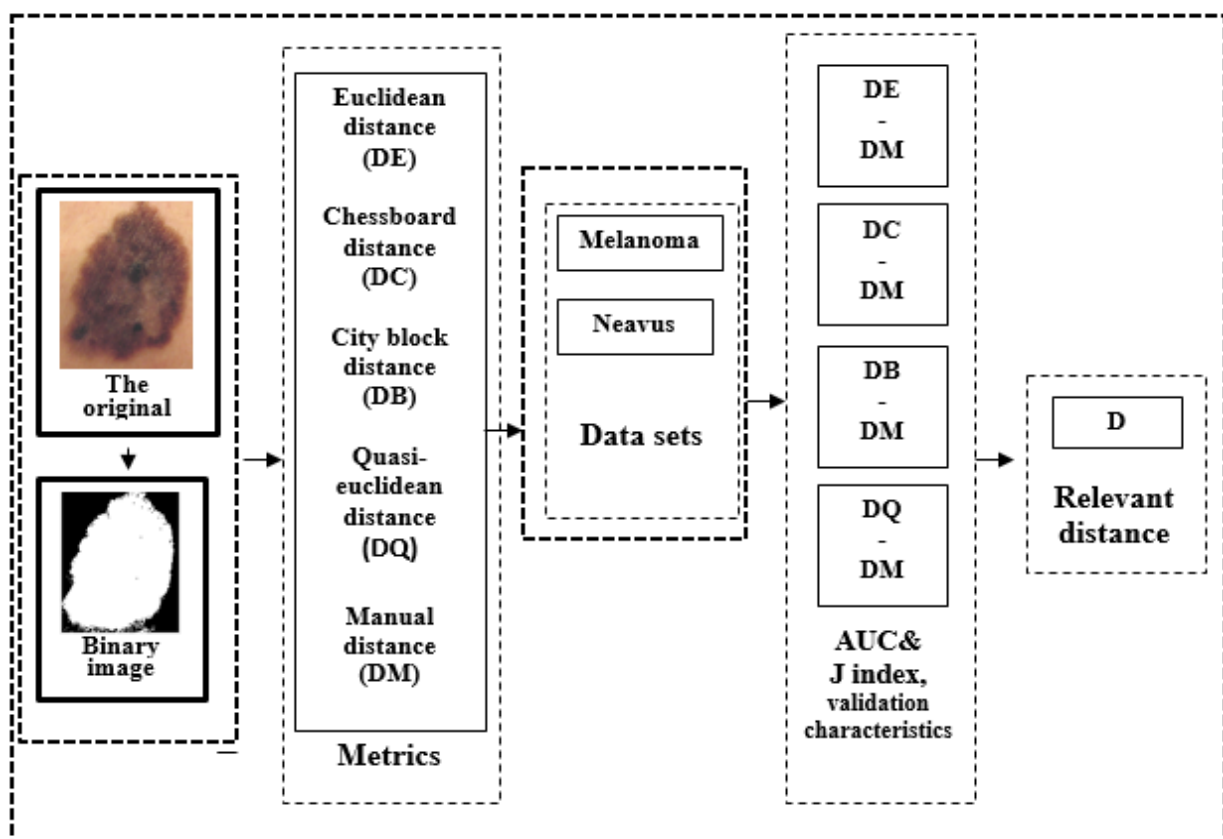


Figure 3.13 Overview of the study

The geometric metrics are compared with the manual distance using the area under the curve (AUC) and the thresholds that were associated with the best Youden index values. The cross-comparison process was used for the DE-DM, DC-DM, DB-DM, DQ-DM pairs. ROC curves, AUC values and threshold values are shown in Figures 3.14 and 3.15.

The stability of the proposed metrics is determined using the following threshold values: $AUC > 0.6$ and Youden index $J \rightarrow 1$. All investigated distance values used to measure skin lesion diameter had AUC values greater than 0.9. The best AUC values were determined for the chessboard distance (CD) metric, respectively 0.979 for nevi and 0.982 for melanoma. The Youden J index is 0.8426 for the nevi class and 0.875 for the melanoma class indicating an excellent relevance of this value for the stated objective.

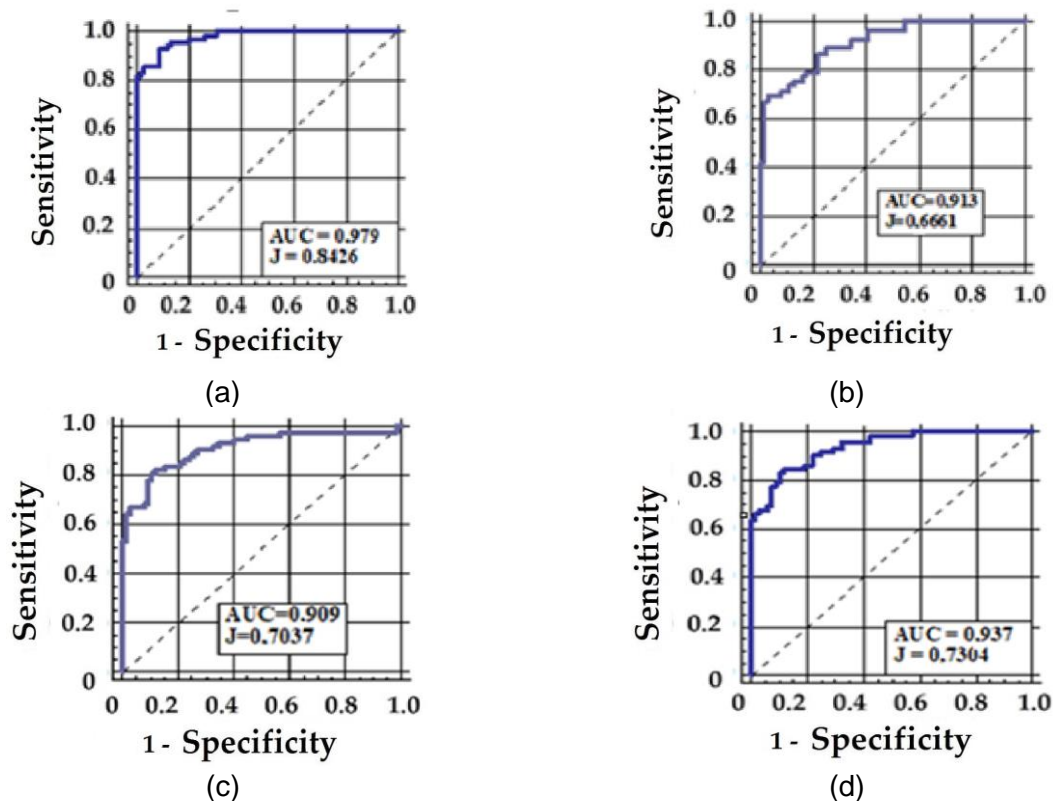


Figure 3.14 ROC curves for nevi.
(a) DC-DM; (b) DB-DM; (c) DE-DM; (d) DQ-DM.

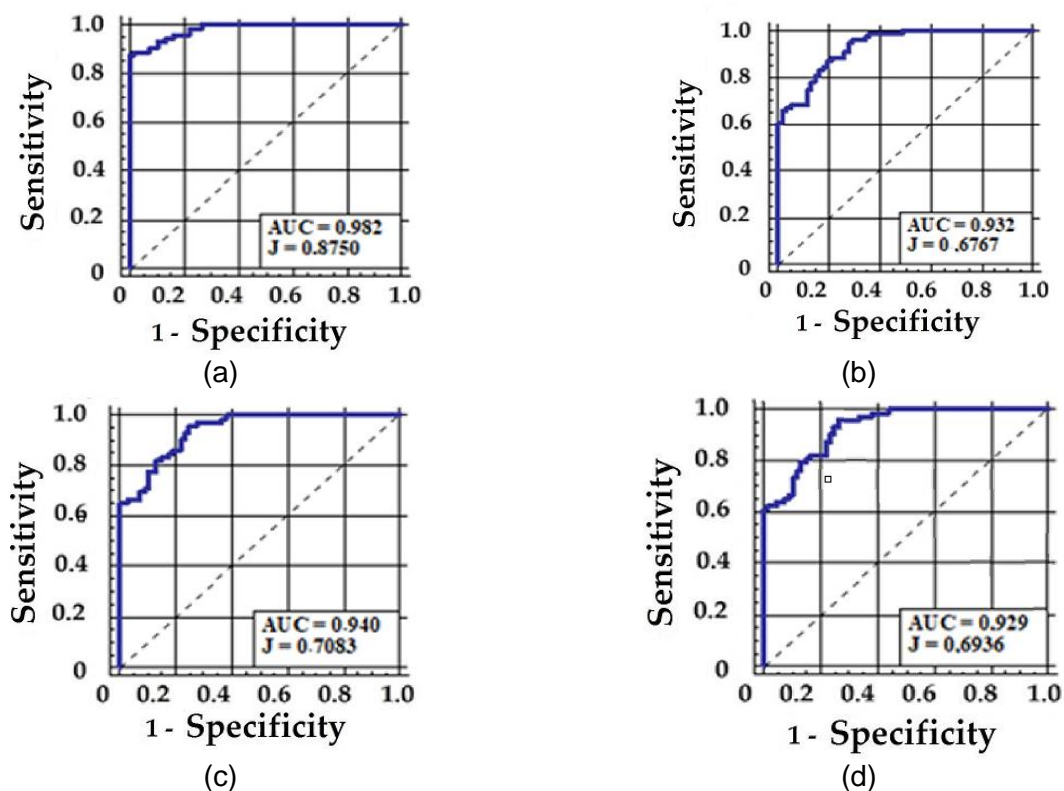


Figure 3.15 ROC curves for melanoma.
(a) DC-DM; (b) DB-DM; (c) DE-DM; (d) DQ-DM.

Lower AUC values were obtained for Euclidean distance (DE) with a value of 0.909 for the nevi class, and for quasi-Euclidean distance (DQ) with a value of 0.929 for the melanoma class, respectively. Similarly, lower values for J, were obtained for city block (DB) distance with 0.66661 for nevi class and 0.66767 for melanoma class.

The accuracy, sensitivity and specificity are shown in Table 3.6.

Table 3.6 Sensitivity, specificity and accuracy of the proposed distance metrics: DC-DM, DB-DM, DE-DM and DQ-DM for nevus and melanoma.

		DC-DM	DB-DM	DE-DM	DQ -DM
Sensitivity (%)	nevi	92.96	69.74	80.82	83.33
	melanoma	87.50	87.34	95.83	92.83
Specificity (%)	nevi	91.30	96.87	89.55	89.71
	melanom	100.00	80.33	75.00	73.53
Accuracy (%)	nevi	92.13	83.30	85.18	86.52
	melanoma	93.75	83.83	85.41	84.68

According to the data in Table 3.6 in terms of sensitivity and specificity, the chessboard distance (CD), a balanced accuracy is observed for both nevi (92.13) and melanoma (93.75). In contrast, the city block distance (DB) has sensitivity values of 69.74 / 87.34 against specificity of 96.87 / 80.33 and accuracy values of 83.30 / 83.83 respectively. Similarly, values for Euclidean distance (ED) show sensitivity values of 80.82 / 95.83 against specificity of 89.55 / 75.00 and accuracy of 85.18 / 85.41, and quasi-Euclidean (QD) values show sensitivity of 83.33 / 92.83 against specificity of 89.71 / 73.53 and 86.52 / 84.68 respectively.

The results show excellent performance values for DC, with AUC of 0.997 for nevi and 0.982 for melanomas, respectively. Also, J is 0.8426 for nevus class and 0.875 for melanoma class.

3.4 Selection of relevant features from non-dermatoscopic images for classification of nevi vs. melanomas

Trying to break out of the ABCD rule patterns, I conducted a study targeting eight features, namely asymmetry index (AI), eccentricity (E), circularity (CIRC), normalized amplitude (FFT), color distribution asymmetry (Q), quadrant asymmetry index (λ), and invariant Hu moments (F6, F7). These features are dedicated to the assessment of asymmetry based on the evaluation of the shape and colours of skin lesions. We aimed to determine their potential in skin lesion differentiation [196]. We focused on the assessment of lesion shape and colour, which are expressed by different mathematical approaches presented in Chapter 2. Seventy melanomas and 100 nevi are investigated. To evaluate the performance of the proposed feature selection procedure, two datasets were used: D1 containing original images and D2 containing images pre-processed for noise removal and non-uniform illumination reduction. The images are segmented, followed by the extraction of the proposed features. The ROC curve and area under the curve (AUC) are used to ascertain the relevance of the selected features and their ability to differentiate skin lesions. The steps of the proposed method are shown in Figure 3.16.

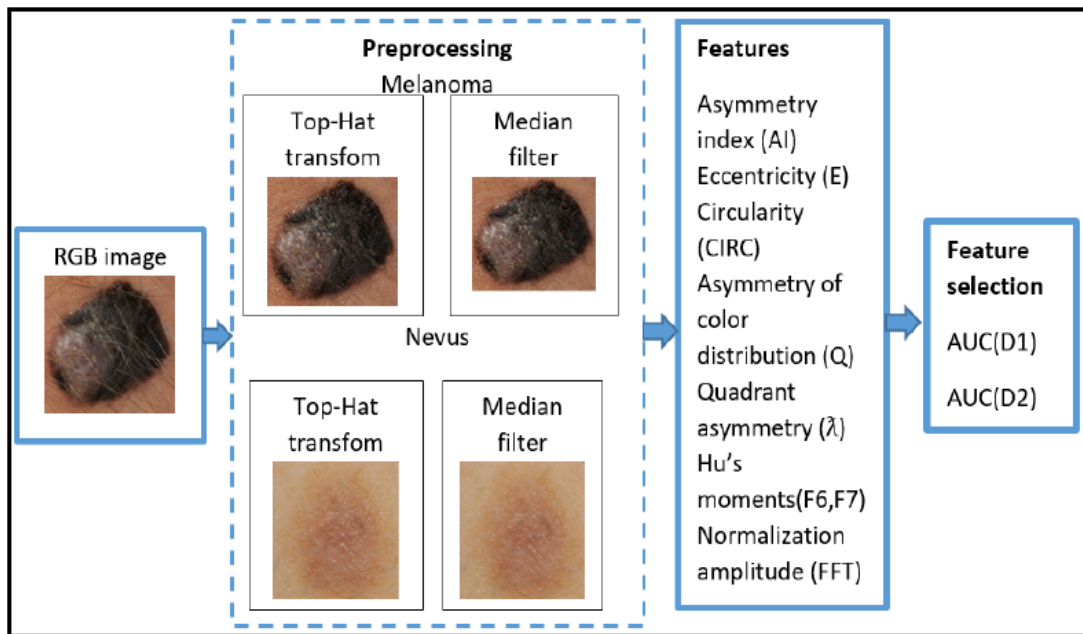


Figure 3.16 Proposed method for feature selection.

An important step to achieve high performance in skin lesion differentiation is the elimination of noise and artefacts due to the existence of hair. A median filter is used to enhance the image with minimal degradation of the original image. To correct for non-uniform illumination, an algorithm using the Top-Hat transform was implemented. Figure 3.17 shows examples of images processed for noise removal and non-uniform illumination correction.

Figure 3.18 displays the results of the skin lesion segmentation process. The Otsu method optimally converts a grey level image into a binary image by setting threshold values to reduce the overlap of the class distribution.

The ROC and AUC curves are illustrated in Figure 3.19. Table 3.7 showing the mean values of the performance measures for D1 containing original images and D2 containing images pre-processed for noise removal and non-uniform illumination reduction.

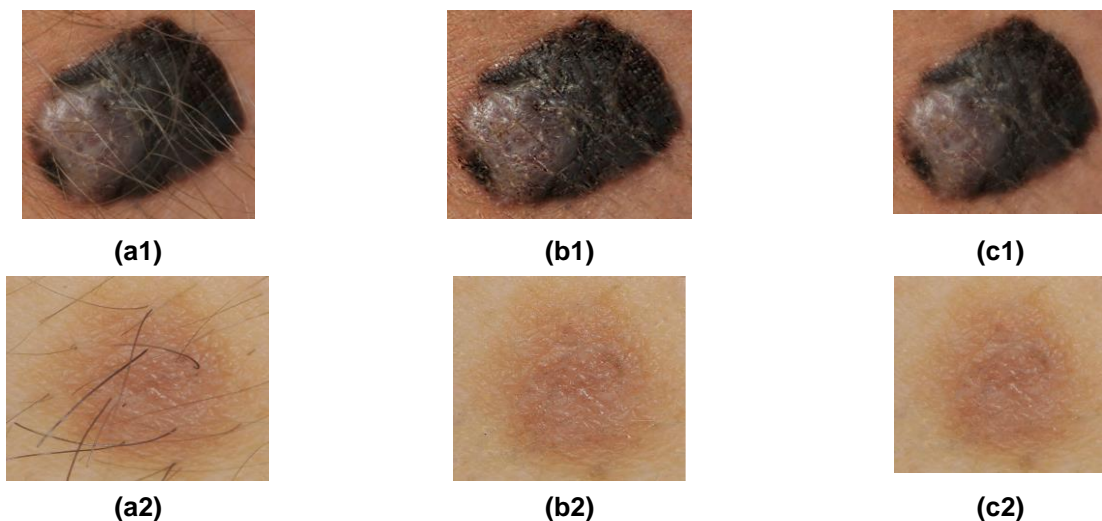


Figure 3.17 Preprocessing of a melanoma image (first row) and nevus image (second row): (a1,a2) Original image; (b1,b2) illumination equalization results obtained by employing classical top-hat transform; (c1,c2) the image given after applying the median filter.

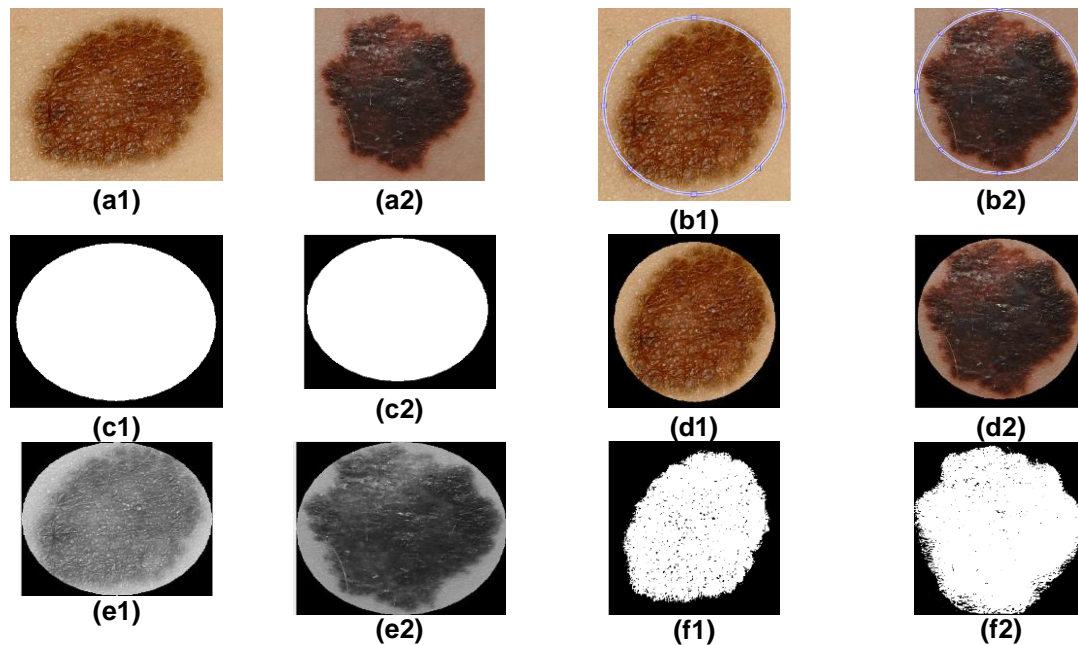


Figure 3.18 Skin lesion segmentation, with index 1 for a nevus and 2 for melanoma; (a1, a2) original RGB image; (b1, b2) ellipse circumscribing the lesion; (c1, c2) elliptical mask; (d1, d2) elliptical mask enclosing the skin lesion; (e1, e2) elliptical mask in grayscale containing the lesion; and (f1, f2) mask associated with the skin lesion (segmented image).

For the images in set D2, an average accuracy of the melanoma vs. nevi classification is observed (Table 3.7). Eccentricity (E) and normalized amplitude (FFT) are relevant features for differentiating skin lesions. Thus eccentricity (E) and normalized amplitude (FFT) for data set D2 indicate a sensitivity of 0.94 / 0.80, specificity of 0.63 / 0.51 and accuracy of 0.79 / 0.66 and for data set D1 eccentricity (E) and normalized amplitude (FFT) indicate a sensitivity of 0.90 / 0.53, specificity of 0.66 / 0.76 and accuracy of 0.78 / 0.65. The moment invariant Hu F6 is important for nevus analysis as the specificity/sensitivity is 0.82/0.72 with accuracy of 0.772. The invariant Hu moment F7 follows the same trend, but the accuracy is 0.67 (D2).

Table 3.7 Sensitivity, specificity and accuracy of analysed features

	Features							
	AI	E	CIRC	TFF	Q	λ	F6	F7
Sensitivity (D1)	0.98	0.90	0.81	0.53	0.61	0.67	0.68	0.76
Sensitivity (D2)	0.84	0.94	0.81	0.80	0.63	0.65	0.72	0.62
Specificity (D1)	0.64	0.66	0.32	0.76	0.57	0.46	0.80	0.51
Specificity (D2)	0.87	0.63	0.35	0.51	0.58	0.53	0.82	0.71
Accuracy (D1)	0.82	0.78	0.57	0.65	0.59	0.57	0.74	0.64
Accuracy (D2)	0.85	0.79	0.58	0.66	0.61	0.59	0.77	0.67

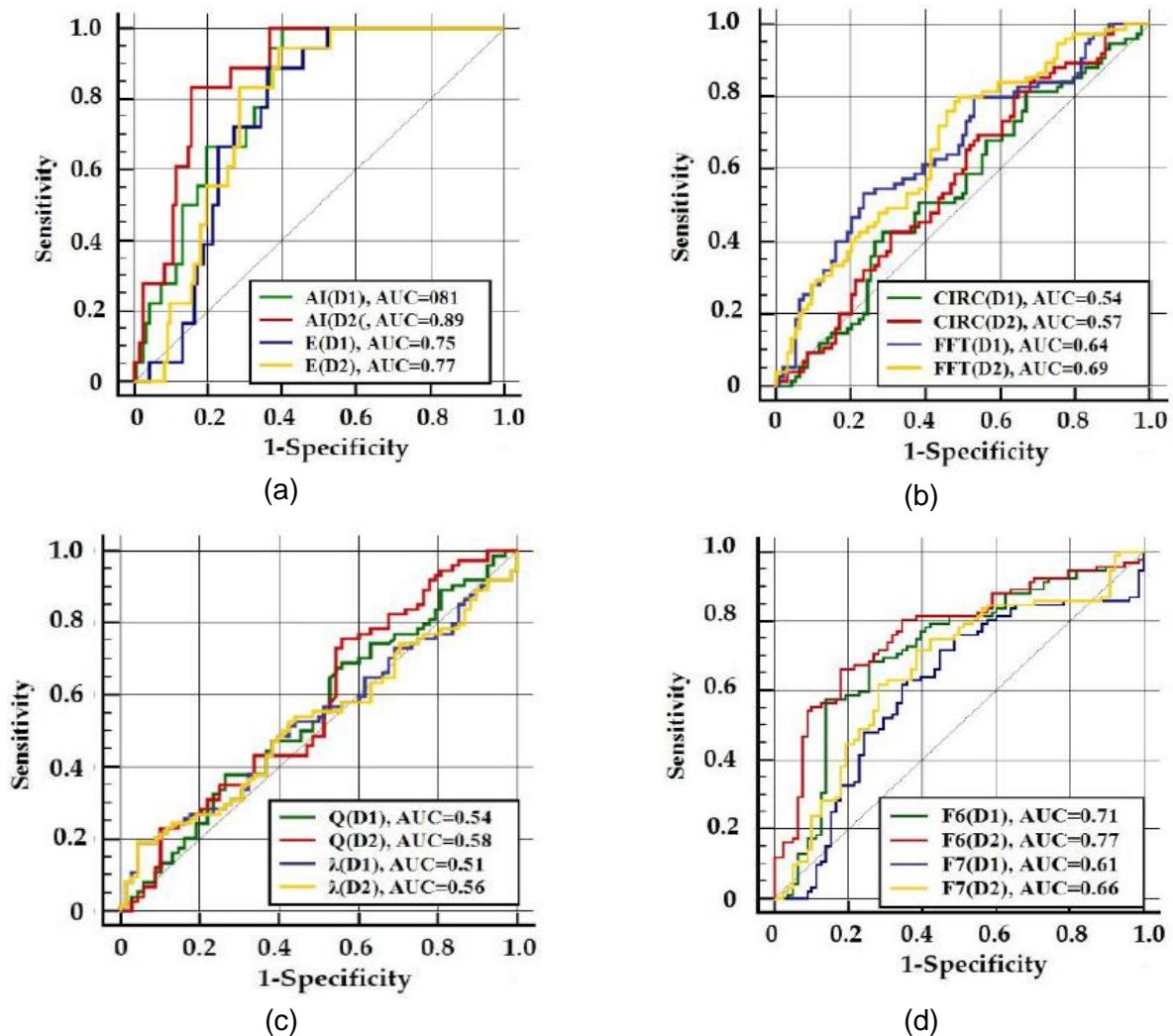


Figure 3.19 Receiver operating characteristic (ROC) curves and area under the curve (AUC) for discriminating melanoma from nevi in the D1 and D2 datasets: (a) AI (asymmetry index) and E (eccentricity) features; (b) circularity (CIRC) and FFT normalization amplitude features; (c) asymmetry of color distribution (Q) and quadrant asymmetry (λ) features; (d) F6 and F7 Hu's invariant moments.

The asymmetry of color distribution (Q) and the quadrant asymmetry (λ) showed no significant contribution assessed by sensitivity, specificity and accuracy. According to these results, asymmetry index (AI) and eccentricity (E) together with Hu F6 invariant moment, are quite competent to differentiate skin lesions. Also, normalized amplitude (FFT) shows great potential in classifying skin lesions. In Figure 3.19 we observe the following: the area under the curve (AUC) values for AI (asymmetry index) is 0.81 for D1 and 0.89 for D2 showing a very good separation between the two classes (melanoma and nevus).

Similarly, for feature E (eccentricity), the AUC values are 0.75 (D1) and 0.77 (D2) indicating good separation between the two classes. Circularity (CIRC), the asymmetry of color distribution (Q) and the quadrant asymmetry (λ) have areas under the ROC curve between values 0.51 and 0.58 indicating a modest separation of the two classes. The normalized amplitude (FFT) has AUC 0.64 for D1 and 0.69 for D2 and shows a moderate separation of the two classes. The invariant Hu moments F6 and F7 have AUC 0.71 for D1 and 0.77 for D2 and 0.61 for D1 and 0.66 for D2, respectively. F6 has a good separation of the two classes, while F7 has an average separation of the two classes.

Chapter 4

Personal contributions on the classification of skin lesions based on color clusters

The aim of this chapter is to bring to the forefront new advanced segmentation methods based on clustering and active contours capable of extracting skin lesion structures. The features extracted from the areas of interest can be analysed and classified with statistical methods so that the sharing of classes containing nevi and melanomas is as obvious as possible. The objective of this chapter focused on understanding the influence of morphological operators in differentiating skin lesions.

In this chapter the synthesis of scientific research is reported through studies that allow the selection of the relevant color distribution from skin lesion images using color histogram analysis, kNN algorithm, area fractal dimension and color group statistical features using a set of machine learning techniques, Random Forest (RF) algorithm.

4.2 Color histogram analysis for skin lesion differentiation

We proposed a new solution for detecting skin lesion boundaries using the color histogram as the main tool in evaluating the color distribution of red, green and blue channels [200]. For each monochrome channel, a global threshold value is calculated using the Otsu method. Pixel clustering of objects and image background is performed taking into account intra- and inter-class variations. The selected thresholds are clearly marked in Figure 4.4 and Figure 4.5.

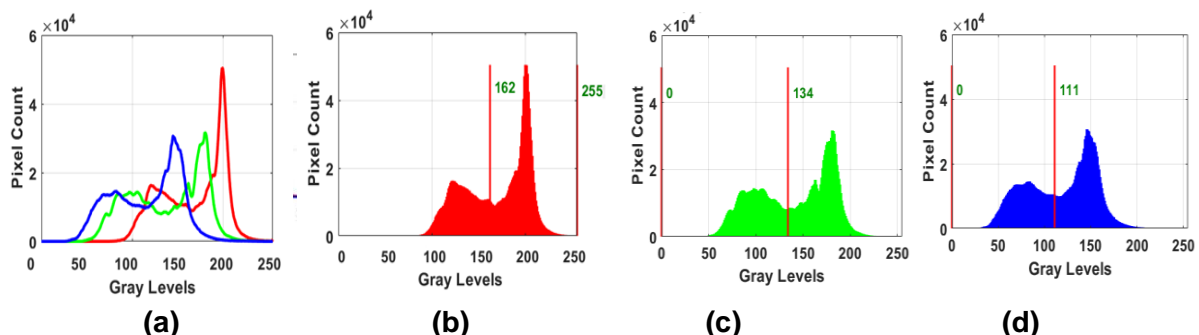


Figure 4.4 Histograms, R, G and B monochrome channels and threshold values for melanoma

(a) Histogram of RGB channels. (b) Red channel; (c) Green channel; (d) Blue channel.

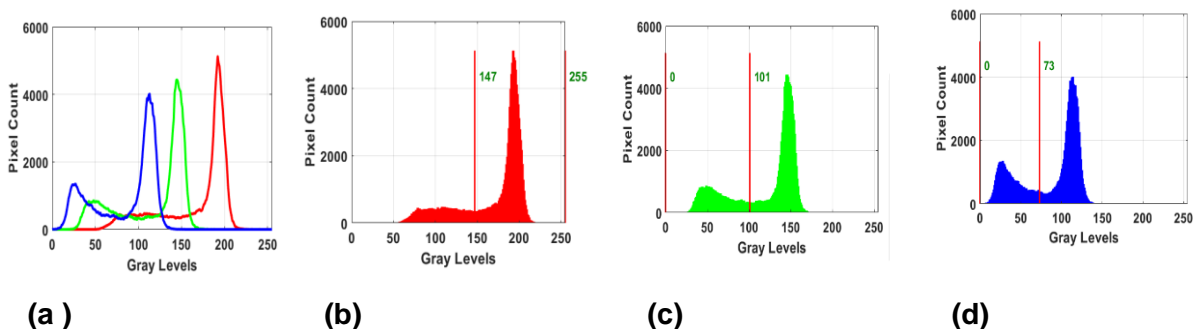


Figure 4.5 Histograms, R, G and B monochrome channels and threshold values for nevi

(a) Histogram of RGB channels. (b) Red channel; (c) Green channel; (d) Blue channel.

The data displayed in color histograms are analyzed using the spread of histogram (HS) [61] as the ratio of interquartile range (IQR) to monochromatic pixel amplitude (R). To determine the similarity between features we used the K-means algorithm. We assumed that once clusters are formed, there are no redundant features in our model.

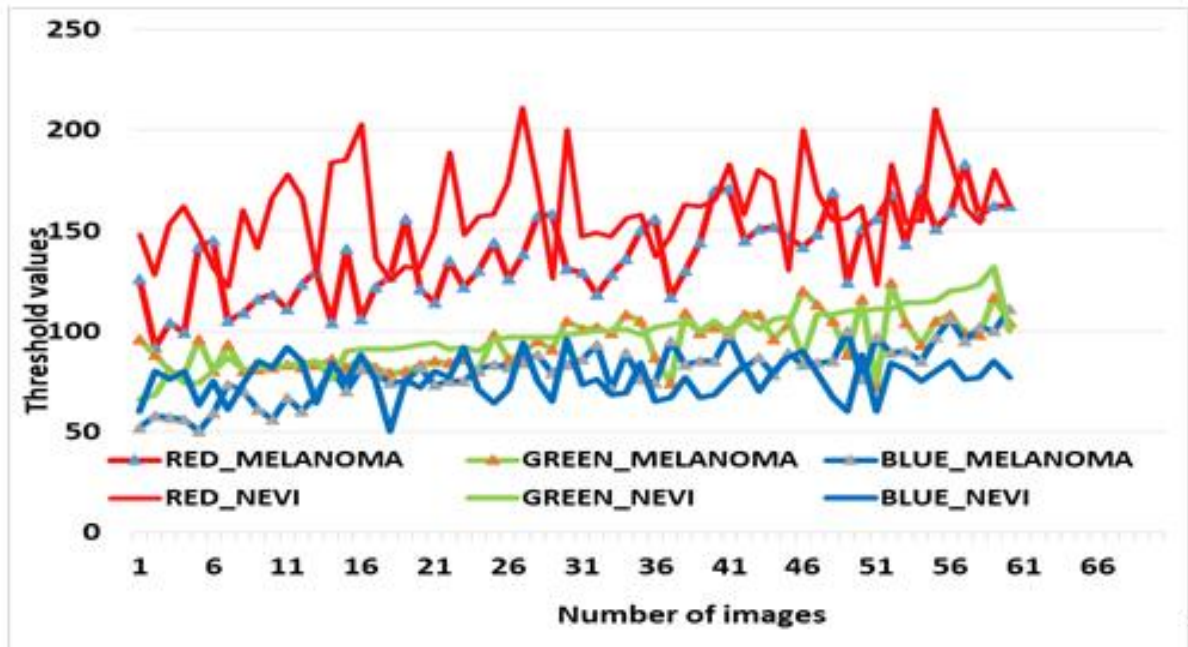


Figure 4.7 The threshold values for R, G and B monochromatic channels for melanoma and nevi

It can be seen that for channel G and B the threshold values are evenly distributed so they cannot provide meaningful information. For the R channel the threshold values are higher and the data distribution is better separated between classes so it could be easily analysed. These threshold values for each monochrome channel create a new database for the cluster.

To determine the quality and validity of the cluster, the following were calculated: the silhouette of an object, the average silhouette of a cluster, and the overall average silhouette for the whole set of skin lesion images (Figure 4.9). A decision on the importance of the monochrome channel in skin lesion classification is made by correlating HS with the relative weight of threshold values for each monochrome channel (Table 4.2).

Table 4.2 HS value for melanoma and nevi

	Melanoma	Nevi
HS	0,143178	0,100046

A proper allocation of objects in the cluster is assessed by the silhouette thickness.

The number of clusters was chosen a priori for 2 because there are two cluster classes: melanoma and nevus. So the clusters are distinct, non-uniform but also negative for G and B channels and for R channel the clusters are distinct, uniform and positive.

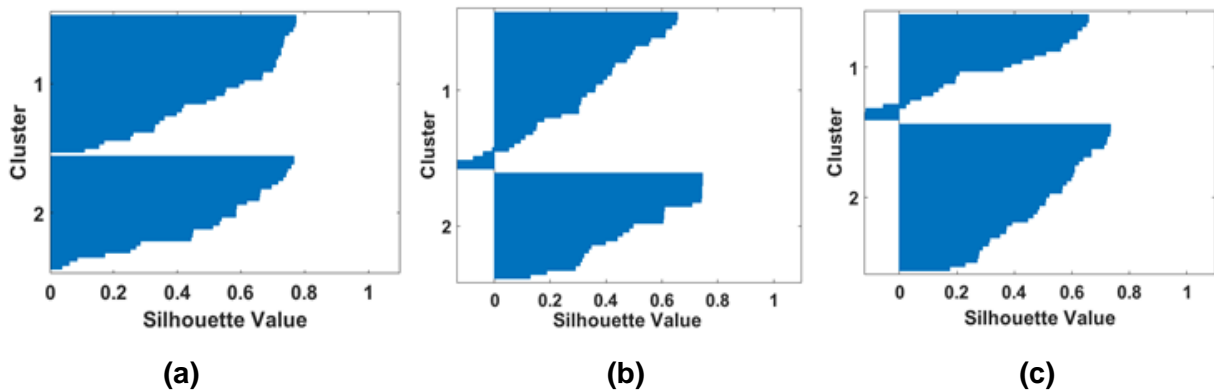


Figure 4.9 Silhouette analysis for data clustering, $n = 2$.
 (a) Red; (b) Green; (c) Blue channel. 1 for melanoma class; 2 for nevi class.

Tabelul 4.3 The values SC for the monochromatic channels

Skin Lesion	Red	Green	Blue
Melanoma	0.91	0.62	0.41
Nevi	0.90	0.61	0.39

The red channel proves to be the most suitable for this task as it provides good silhouette width uniformity, no negative silhouette coefficient and SC values close to 1.

4.4 Classification of skin lesions based on surface fractal sizes and statistical features of the colour cluster using a set of machine learning techniques

The objective of the study is to recognize and classify skin lesions by implementing a novel combination of features between the 2D Higuchi's surface fractal dimension and relevant color area features and two classifiers as an auxiliary diagnostic method for melanoma recognition. The experimental base consisted of 248 nevi images and 407 melanoma images belonging to public databases: 7-Point (68 nevi and 297 dermoscopic melanoma images), Med-Node (100 nevi 40 melanoma non-dermoscopic images), PH2 (80 nevi 40 melanoma dermoscopic images).

To achieve the proposed objective, the following directions were followed:

- Determination of the average representative percentage of colour areas of skin lesions for each data set considered.
- Proposing a descriptor for investigating skin surface fractal dimensions for channels in RGB colour images, i.e. 2D Higuchi fractal dimension as a quantitative objective.
- Implementation of two distinct machine learning classifiers, namely a kNN-CV algorithm and a RBFNN approach as a nonlinear classifier, to generate the prediction.
- Dynamic data partitioning is performed using the 5-fold cross-validation (CV) method.

These machine learning classifiers belong to different classification paradigms.

An example of image preprocessing for noise reduction, hair removal and image segmentation is shown in Figure 4.13. DullRazor software was used for hair removal and image segmentation based on colour thresholding.

Color features are selected using a clustering method that allows selection of the relevant color distribution in melanocytic lesion images and calculation of the average percentage of color areas in the lesion area of nevi and melanoma. For this, we analyzed the color histogram of each

skin lesion associated image. Thus, we considered twenty-three color groups, denoted as cl1, . . . , cl23, which are characterized by the largest differences between the minimum and maximum intensity in each R, G and B channel, as given in [203]. For each image, the percentage of each colour group is calculated as the ratio of the number of pixels in the lesion belonging to the specified colour group to the total pixels of the lesion.

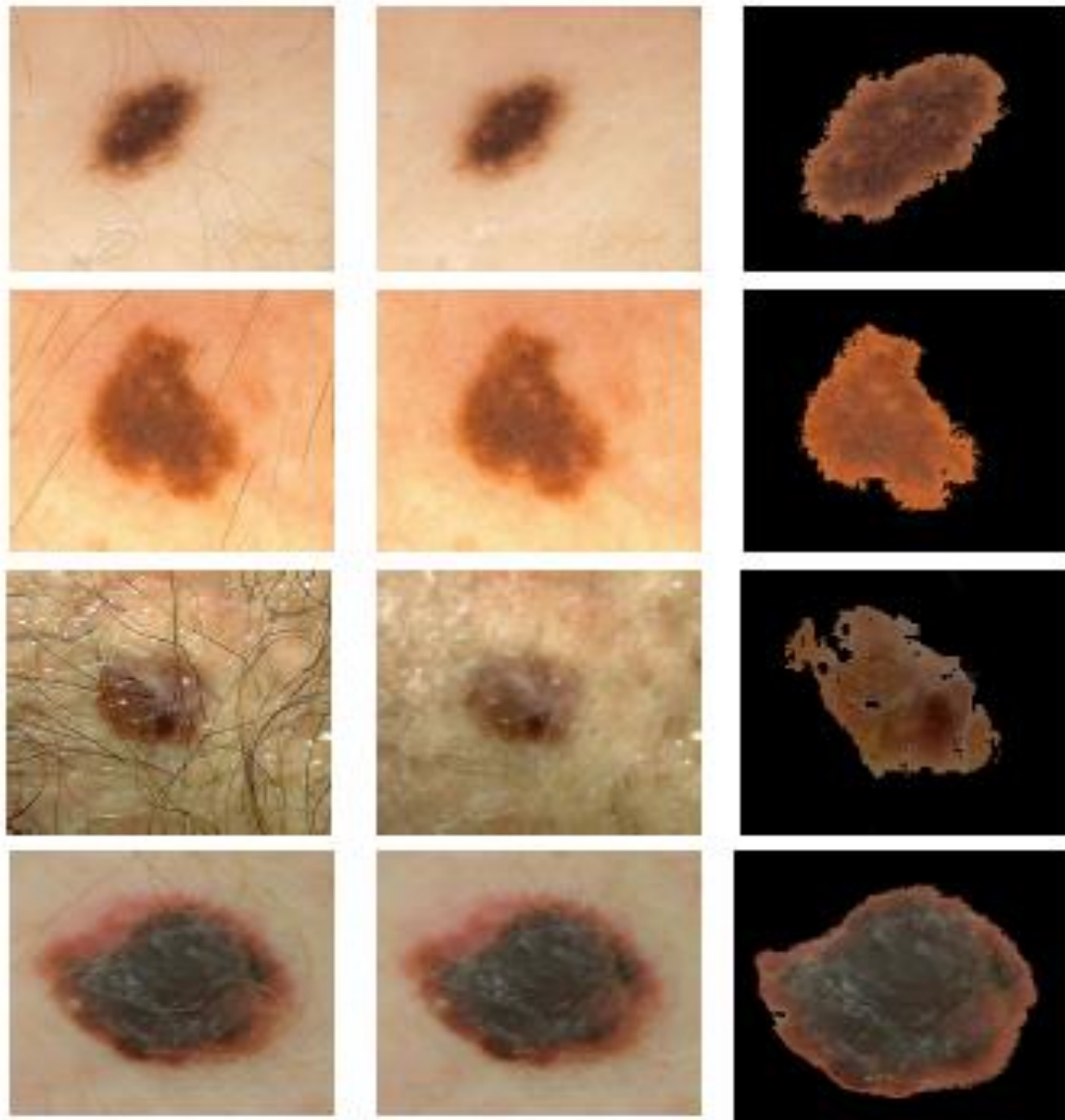


Figure 4.13 Illustration of preprocessing for hair removal and segmentation.

Rows 1 and 2: nevi image (PH2 database).

Rows 3 and 4: melanoma image (7-point database in row 3 and Med-Node database in row 4).

First column: the original image. Second column: image after hair removal.

Third column: segmented image.

These mean percentages of colour areas represent statistically relevant features for differentiating skin lesions. A vector with many features is constructed. To eliminate features without predictive information, a t-test is used. A p-value < 0.05 is considered statistically significant, i.e. the selected samples are statistically significantly different from each other. The selected features are universally accepted and used to differentiate skin lesions. An example of the colour content measurement and clustering method to select the relevant colour distribution is shown in Figure 4.14.

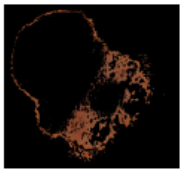
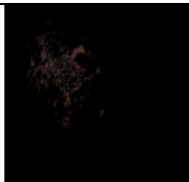
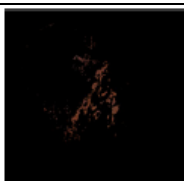
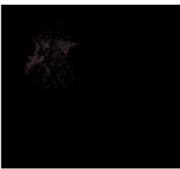



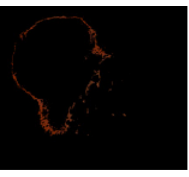


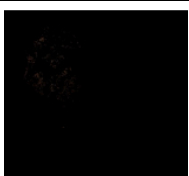
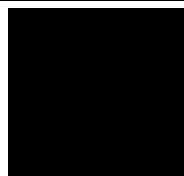

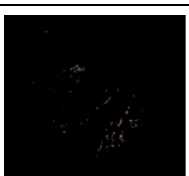
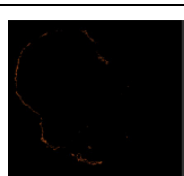
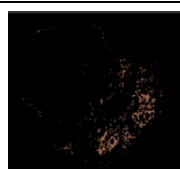
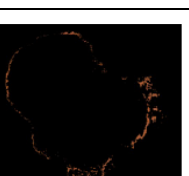
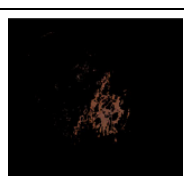



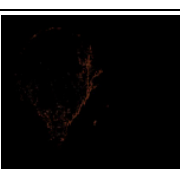
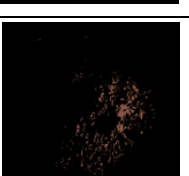

cl1 R _{min} 128 G _{min} 64 B _{min} 32 R _{max} 159 G _{max} 95 B _{max} 63 	cl9 R _{min} 64 G _{min} 32 B _{min} 32 R _{max} 95 G _{max} 63 B _{max} 63 	cl17 R _{min} 96 G _{min} 32 B _{min} 32 R _{max} 127 G _{max} 63 B _{max} 63 
cl2 R _{min} 32 G _{min} 0 B _{min} 0 R _{max} 63 G _{max} 31 B _{max} 31 	cl10 R _{min} 160 G _{min} 128 B _{min} 96 R _{max} 191 G _{max} 159 B _{max} 127 	cl18 R _{min} 192 G _{min} 96 B _{min} 96 R _{max} 223 G _{max} 127 B _{max} 127 
cl3 R _{min} 32 G _{min} 32 B _{min} 32 R _{max} 63 G _{max} 63 B _{max} 63 	cl11 R _{min} 96 G _{min} 32 B _{min} 0 R _{max} 127 G _{max} 63 B _{max} 31 	cl19 R _{min} 224 G _{min} 160 B _{min} 128 R _{max} 255 G _{max} 191 B _{max} 159 
cl4 R _{min} 160 G _{min} 96 B _{min} 64 R _{max} 191 G _{max} 127 B _{max} 95 	cl12 R _{min} 32 G _{min} 32 B _{min} 0 R _{max} 63 G _{max} 63 B _{max} 31 	cl20 R _{min} 192 G _{min} 128 B _{min} 64 R _{max} 223 G _{max} 159 B _{max} 95 
cl5 R _{min} 96 G _{min} 64 B _{min} 64 R _{max} 127 G _{max} 95 B _{max} 95 	cl13 R _{min} 128 G _{min} 96 B _{min} 96 R _{max} 159 G _{max} 127 B _{max} 127 	cl21 R _{min} 128 G _{min} 64 B _{min} 0 R _{max} 159 G _{max} 95 B _{max} 31 
cl6 R _{min} 128 G _{min} 96 B _{min} 64 R _{max} 159 G _{max} 127 B _{max} 95 	cl14 R _{min} 160 G _{min} 64 B _{min} 32 R _{max} 191 G _{max} 95 B _{max} 63 	cl22 R _{min} 96 G _{min} 64 B _{min} 32 R _{max} 127 G _{max} 95 B _{max} 63 
cl7 R _{min} 160 G _{min} 96 B _{min} 32 R _{max} 191 G _{max} 127 B _{max} 63 	cl15 R _{min} 160 G _{min} 96 B _{min} 96 R _{max} 191 G _{max} 127 B _{max} 127 	cl23 R _{min} 64 G _{min} 64 B _{min} 64 R _{max} 95 G _{max} 95 B _{max} 95 
cl8 R _{min} 64 G _{min} 32 B _{min} 0 R _{max} 95 G _{max} 63 B _{max} 31 	cl16 R _{min} 128 G _{min} 64 B _{min} 64 R _{max} 159 G _{max} 95 B _{max} 95 	<p>The segmented image</p> 

Figura 4.14 Example of colour content measurement and clustering method to select the relevant colour distribution from skin lesion images.

The fractal surface size is calculated using a 2D generalisation of the HFD. It calculates the HFD of the R, G and B channel colours in RGB images associated with skin lesions. The 2D Higuchi

surface fractal descriptor was calculated for each image with the fractal scaling parameter ranging from $k = 1$ to $k = 8$ (Figure 4.15).

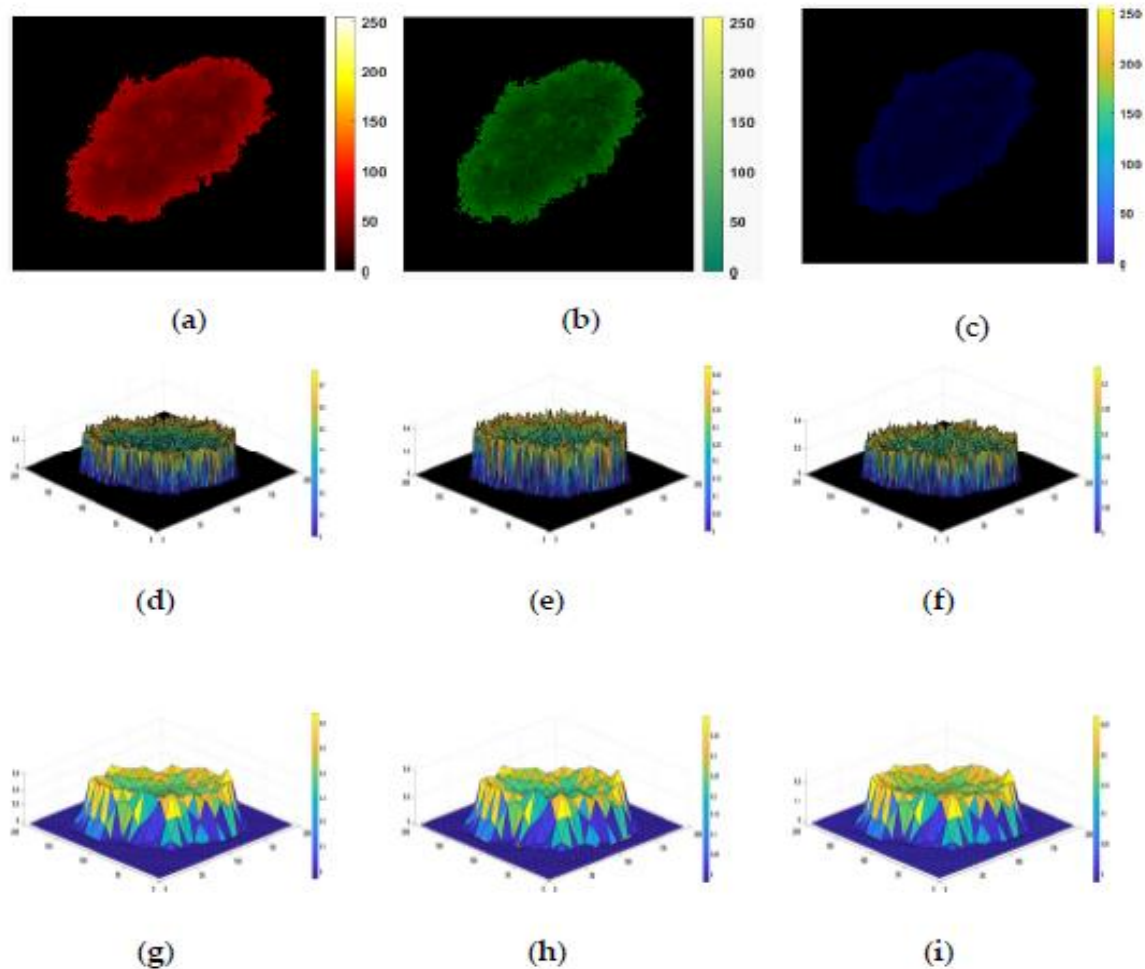


Figure 4.15 Example of HFD calculation.

(a - c) R, G and B color channels for a digital image; (d - i) An illustration of the tessellation patterns for each color channel. The triangular shapes of $k = 1$ are shown in (d - f); the fractal scaling parameter $k = 4$ is shown in (g, i); (d, g) are R - channel images; (e, h) are G - channel images; (f, i) are B-channel images.

The classification of both color and fractal features is performed separately using the kNN algorithm with a 5-fold cross-validation method. The same discriminative features are used for classification by the RBFNN approach as for a nonlinear classifier.

An analysis of the performance of these different classification methods is performed. The prediction performance of kNN based on the average percentage of feature descriptors of the colored areas differed significantly across color groups and datasets (Figure 4.16). The highest classification accuracies of 82.47% (clusters cl10 and cl15), 81.44% (cl23) and 80.41% (cl20) belong to dataset PH2. The second highest classification accuracy of 75.91% (cl5 cl7 and cl14) belongs to the Med-Node dataset. Clusters cl3, cl11 and cl13 from the 7-point dataset gave lower accuracy results. Color clusters cl8, cl16, cl17, cl19 and cl22 did not contain any relevant mean percentage of color area feature descriptors.

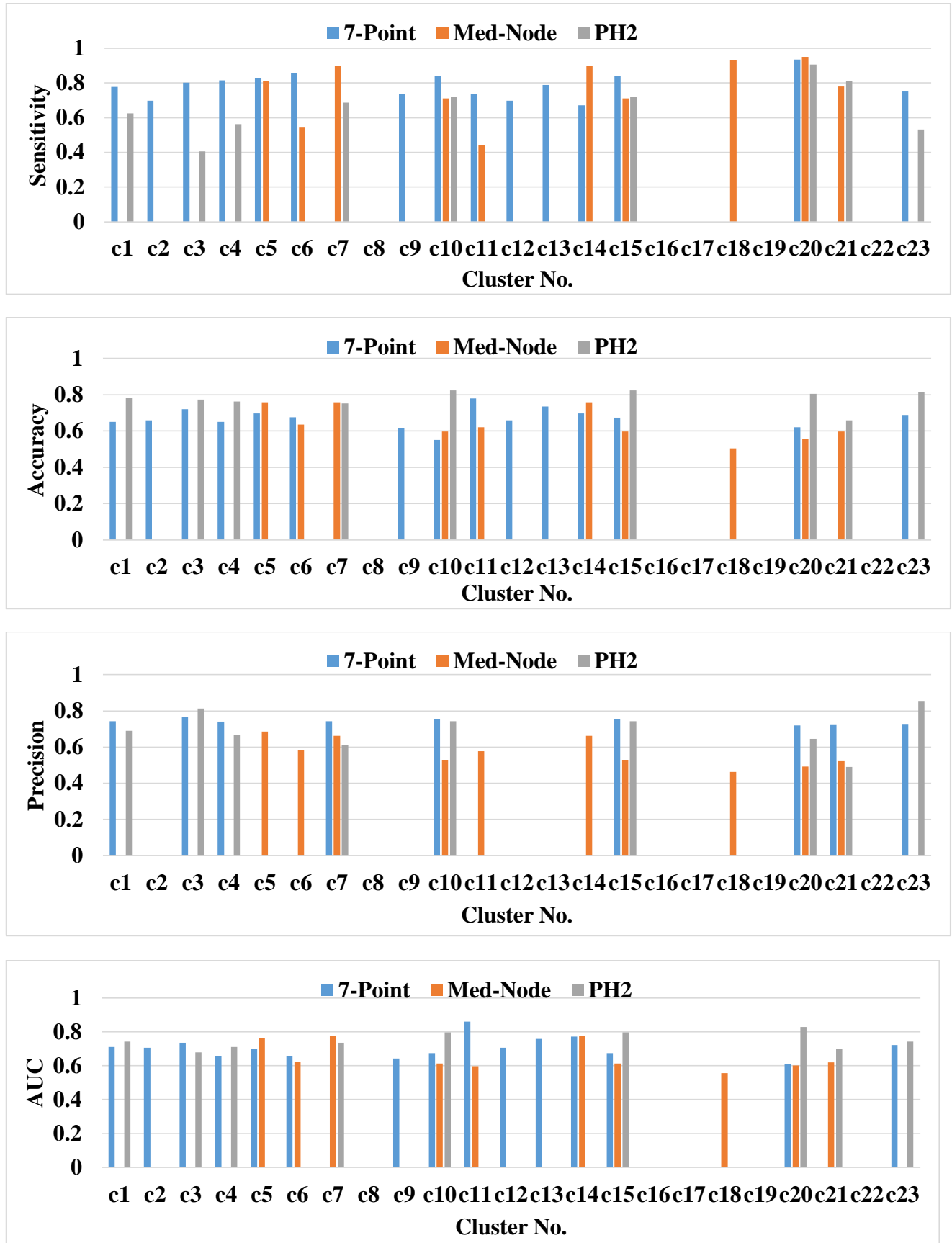


Figure 4.16 The prediction performance for 5-fold cross-validation and kNN classifier for different average percentage of colour areas/color cluster descriptors.

Classification performance based on the 2D Higuchi surface fractal descriptor is shown in Table 4.5. Classification results show that the 2D Higuchi surface fractal descriptor produces higher classification accuracy (79.38%) than the average percentage of color zone feature descriptors for the PH2 dataset.

Table 4.5. Accuracy for 5-fold cross-validation and kNN classifier for 2D Higuchi fractal surface descriptors

Dateset	Sensitivity (%)	Accuracy (%)	Precision (%)	AUC	Dice Scores
7-Point	80.77	71.43	73.26	0.6948	0.7683
Med-Node	30.19	64.23	57.14	0.6423	0.3951
PH2	83.33	79.38	62.50	0.8047	0.7143

The selected discriminant features were classified by a five-fold cross-validation and RBFNN approach. The input data were divided into five subsets; one subset was the test set and the other four subsets were for training. The Gaussian functions used h_1, h_2, \dots, h_g were as follows: $g = 15$ for the 7-Point data set; $g = 11$ for the Med-Node data set and $g = 10$ for the PH2 data set.

The investigation was dedicated to determine the best performance for the RBFNN classifier with different inputs: 50 neurons per hidden layer and two outputs. The number of hidden neurons varied from 0 to 50, and the new hidden nodes were automatically incorporated by the network. Experimental verification indicated that 50 neurons on the hidden layer provided the best classification performance, and the classifier obtained the global optimal solution characterized by lower MSE.

Our specified MSE target was 0.01. The diagnostic performance of the models was evaluated in terms of sensitivity, accuracy, precision, AUC, Dice scores and MSE provided by the RBFNN classifier. To highlight that both features can promote each other in the diagnostic process, the mean percentage of colour areas and HFD features were evaluated and compared together and separately, and the changes in performance are shown in Table 4.6.

Table 4.6 Performance of the RBFNN classifier in test experiments.

Dateset	RBFNN input	Sensitivity (%)	Accuracy (%)	Precision (%)	AUC	DICE Scores	No. of hidden neurons	MSE
7-POINT	Color cluster	97.77	95.12	94.32	0.9412	0.9603	50	0.1904
	Color cluster and HFD	98.01	95.42	94.44	0.9422	0.9630		0.0924
Med-Node	Color cluster	96.22	94.12	88.61	0.9550	0.9333	50	0.1789
	Color cluster and HFD	96.42	94.71	87.50	0.9588	0.9396		0.1662
PH2	Color cluster	1.00	94.17	85.03	0.9553	0.9195	50	0.1372
	Color cluster and HFD	1.00	94.88	85.62	0.9685	0.9211		0.1128

Table 4.6 shows that for all investigated datasets, when the neural network input consisted of both the average percentage of color areas/color cluster and 2D Higuchi surface fractal descriptors, the diagnostic performance was improved in terms of accuracy, AUC, Dice scores, and MSE. In addition, the proposed RBFNN was found to be more accurate and efficient than the kNN algorithm in recognizing and classifying skin cancer.

RBFNN errors were lower when the input data were color clusters and HFD descriptors, confirming our working hypothesis of using 2D Higuchi's fractal surface descriptors to improve classification performance was correct. In addition, skin lesion analysis was more accurate when the proposed RBFNN was used.

We compared the results obtained with machine learning methods as well as neural network methods (Table 4.7). The classification accuracy of kNN-CV with 2D Higuchi surface fractal features is comparable to that provided by other classifiers. It can be seen that the proposed RBFNN algorithm achieved a significant improvement in accuracy in all cases.

Table 4.7 The comparison of accuracy results of the proposed method with those of existing methods.

Autor	Acuratețe (%) și detalii
Nasiri et al.[204]	64% (for 1st test: kNN (300, 100) and spot features) 67% (2nd test: kNN (1346, 450) and spot features)
Kavitha et al. [205]	78.2 (kNN and GLCM features) 81.79% (Inception-ResNet-v2, ISIC 2016 dataset)
Al-masni et al. [206]	81.57% (ResNet-50, ISIC 2017 dataset) 89.29% (ResNet-50, ISIC 2018 dataset)
Seeja & Suresh [207]	79.26% (kNN, LBP and Edge histograms, HOG, Gabor filter)
Khan et al. [208]	94.50% (Neural Network/Feed Forward/sigmoid function/3 hidden layers, ISBI2016 dataset, 70:30 training and testing). 94.20% (Neural Network/Feed Forward/sigmoid function/3 hidden layers, ISBI2017 dataset, 70:30 training and testing).
Proposed kNN-CV	71.43% (7-Point dataset); 64.23% (Med-Node dataset) and 79.38% (PH2 dataset) for 2D Higuchi's surface fractal features
Proposed RBFNN	95.42% (7-Point dataset) 94.71% (Med-Node dataset) 94.88% (PH2 dataset)

4.5 Skin lesion differentiation using the Random Forest (RF) algorithm

Since the first warning signs for early detection and prevention of melanoma are colour and irregular edges a particular importance we have given to methods and techniques for investigating the colour of dermoscopic images based on clustering algorithms and characterizing the shapes of objects in an image based on Hu moments viewed as invariant moment features [209].

The objective of the research was to develop a novel algorithm for the differentiation of skin lesions based on color (C) and Hu moment (H) features using the Random Forest (RF) algorithm.

The experimental base used consisted of 68 nevi and 98 melanomas from the 7-Point database, 100 nevi and 70 melanomas from the Med-Node database and 80 nevi and 40 melanomas from the PH2 database. For all datasets, 70% of the total samples were taken in the

training set and the remaining 30% for the test set. The proposed methodology for identifying relevant features for classifying skin lesions is shown in Figure 4.17.

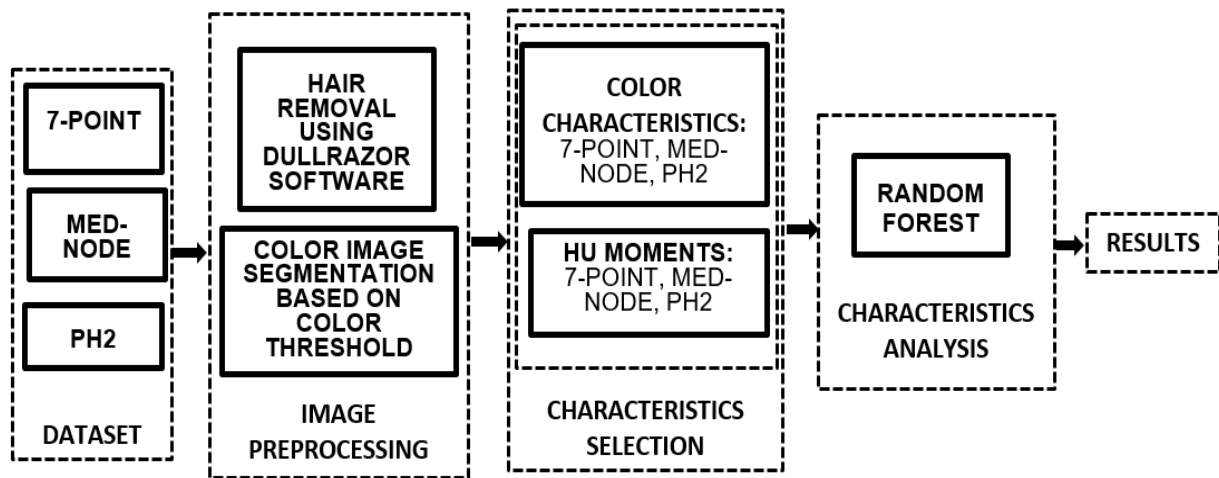


Figure 4.17 Overview of the methodology

For the extraction of colour features, the images were segmented based on the optimal colour threshold method applied on each RGB colour channel (red, green, blue). A binary mask was obtained for each image and the area of interest for the skin lesion was highlighted. Segmented images were also preprocessed using DullRazor hair removal software.

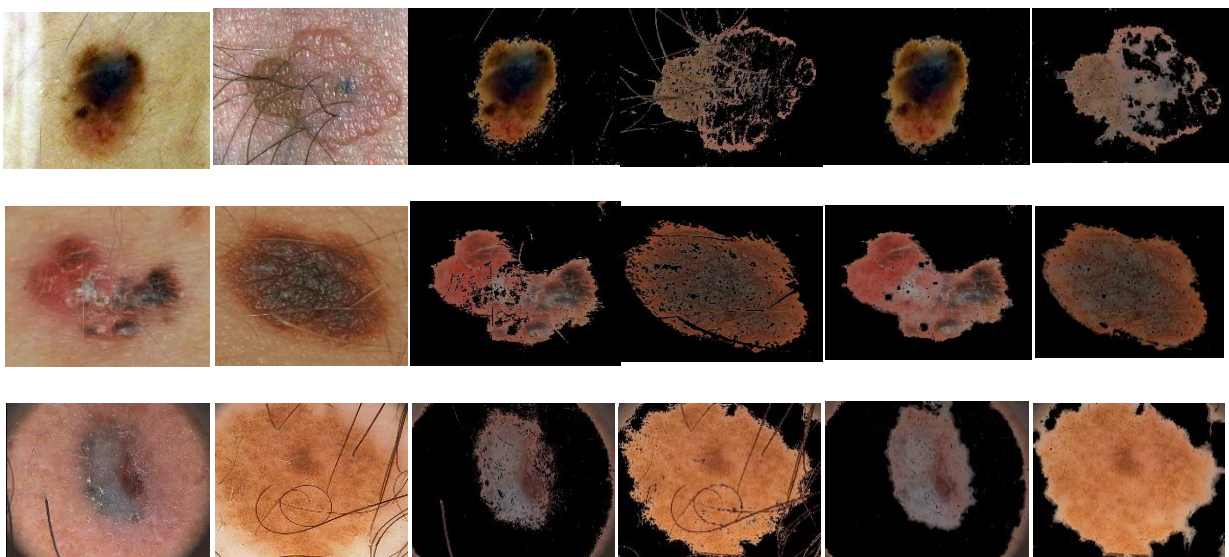


Figure 4.18 Examples of preprocessed images.

Row 1: images from the 7-Point database; Row 2: images from the MED-NODE database; Row 3: images from the PH2 database.

Figure 4.18 shows examples of preprocessed images. Each image is divided into color groups based on the method proposed by Seidenari et al. [203]. Using the minimum and maximum RGB intensity, 23 color features were extracted. The clustering method allows selection of the relevant color distribution, and calculates the average percentage of color areas in the area of the skin lesion [202].

We used the RF algorithm involving 30 input features for skin lesion images of three datasets and two output classes (nevi and melanomas). The obtained feature sets are fed into the random RF classifier to determine the relevant feature to classify skin lesions with the highest accuracy. As the RF algorithm for classification problems uses majority voting, and for regression problems it uses the average prediction of individual decision trees we will analyze those features for which the relevance exceeds the average (0.033). Figure 4.19 shows the calculated performance for the proposed features.

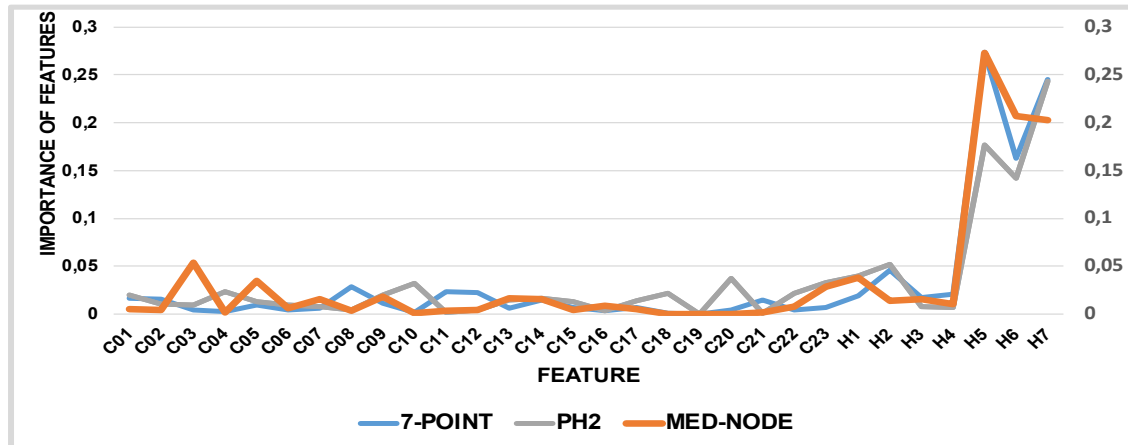


Figure 4.19 Feature accuracy.

Skin lesion differentiation accuracy was evaluated on a training set consisting of 30% skin lesion images and tested on 70% other skin lesion images for each dataset. Because the diagnostic performance may depend on the degree of image difficulty, we correlated the mathematical parameters of the skin lesion features and compared the accuracy for each database. The relevant features are:

- H5, H7, H6, H2, C08 for the 7 Point data set;
- H7, H5, H6, H2, H1 for the Med-Node dataset;
- H5, H6, H7, C03, H1 for PH2 dataset.

The impact of using the selected features for each data set is shown in table 4.8.

Table 4.8 Relevant features and performances of the RF classifier

Database	Accuracy	Sensitivity	Specificity
7-Point	94.0%	92.5 %	94.9%
Med-Node	98.1 %	99.0 %	97.1%
PH2	96.5 %	98.7 %	92.7%

The results obtained were compared with the results obtained in other studies. In the presented study, a higher accuracy was obtained in the classification of skin lesions for HU moments compared to color features, namely 94% for the 7-Point database, 98.1% for the Med-Node database and 96.5% for PH2 database. These features are dedicated to the evaluation of images based on the assessment of irregularity of edges and colors of skin lesions. Our approach to analyze the shape geometry and color of benign (nevi) and malignant (melanoma) lesions will be useful to detect lesions before they grow and become a case of melanoma.

Chapter 5

Personal contributions on the implementation of the "Skin Lesions" expert system

The aim of the **"Skin Lesions"** expert system is to provide the user with relevant information to enable them to draw their own conclusions based on the skin lesion analysis, with the clear specification that the system is not intended to replace doctors, but only to provide users with an additional motivation to seek medical assistance if necessary.

5.1 Requirements for an application

The expert system „Skin Lesions" has been developed into a mobile application (app) that can be downloaded and installed on Google Android mobile phone platforms.

The main requirements for the mobile app:

- to be able to capture images using the phone's camera;
- to be able to upload images from the phone's memory;
- the graphical user interface of the app must contain buttons to start the analysis processes;
- to display information messages.

An expert melanoma analysis system is built in four main steps.

The first stage is image acquisition which can be performed by different devices such as dermatoscope, spectroscope, standard digital camera or phone camera. The images acquired by these devices have particular features and different qualities that can significantly change the outcome of the analysis process.

The second stage involves detecting skin lesions and resizing the image.

The third stage calculates features A, B, C, D with own algorithms.

Finally, the fourth stage classifies the lesions according to the TDS score producing an estimate whether the lesion is benign or malignant (melanoma).

Own algorithms are used to calculate the TDS score, asymmetry features, edge/grain regularity, colour and diameter. Each of the features is then multiplied by a given weighting factor to obtain a total dermoscopic score (TDS). TDS values below 4.75 indicate a benign melanocytic lesion, values between 4.8 and 5.45 indicate a suspicious lesion, and values equal to or greater than 5.45 are highly suggestive of melanoma [210].

- For the development of the proposed expert system we used:
 - Java programming language development kit;
 - Android SDK (Software Development Kit);
 - Android Studio: an integrated development environment (IDE);
 - Android Virtual Device (delivered with the Android SDK), a device on which to run the applications;
 - A mobile phone with the Android operating system for which the "Skin Lesions" application has been developed.

The compiling and running on a device is done using the SDK. It is the component that creates the file that will be installed on the device and that has the .apk extension, merging all the resources in the application project.

We incorporated MATLAB codes for each feature using the ToApp Block from the Android Studio support package.

5.2 The proposed technique

Acquiring the digital image of the skin lesion is the first and main step in making the expert system "Skin Lesions". We use either images obtained using a commercially available digital camera or dermoscopy or images obtained with a mobile phone. After the image is acquired we extract discriminating features from the skin lesion. To assess asymmetry we used the geometric feature asymmetry (GAF) calculation algorithm presented in the study [194]. To assess contour (border) irregularity (B) we used the compactness index presented in the study [193]. Six different colors are considered to determine the color score: white, red, light brown, dark brown, blue-gray and black. White should only be considered if the area is lighter than the adjacent skin. The maximum color score is 6 and the minimum score is 0. For color detection we used the method proposed in study [200]. For diameter we used the Chessboard distance method presented in study [195].

For the classification of skin lesions I propose to calculate the total dermoscopy score (TDS) using formula 1.1 [211]:

After calculating the TDS, the diagnosis will be obtained according to the interpretation in Table 5.1.

Table 5.1 *Diagnosis of lesions according to TDS.*

Total dermoscopy score TDS	Interpretation
0 - 2.75	Benign lesion
2.76 - 4.75	Slightly suspicious lesion (benign)
4.76 - 5.45	Lesion suspected to be melanoma (melanoma)
>5.46	Lesion highly suspicious for melanoma (melanoma)

In the application the message displayed will be made according to the algorithm:

Proposed

algorithm:

```

if (TDS >= 0.00 && TDS<= 2.75)
    msgbox('We keep under observation')
elseif (TDS >= 2.76 && TDS<= 4.75)
    msgbox('Go to the doctor')
elseif (TDS >= 4.8 && TDS<= 5.45)
    msgbox('Suspicious Melanoma Cancer Detected')
% elseif (TDS >= 5.46)
else
    msgbox('Highly suspicious Melanoma Cancer Detected')
end.

```

5.3 Application description

Install the "Skin Lesions" application on your mobile phone:

➤ by clicking on the link:

https://drive.google.com/drive/folders/18ayYaLWY3hMrn12ghcuQ4g-30f_fWmoi?usp=share_link

- by scanning the QR code (figure 5.5):



Figure 5.5 QR code of the application "Skin Lesions"

To be able to install an app on a mobile device, you need to enable the possibility to run it, from Settings → System → Developer Options. This option must be enabled, as well as Debugging → Android Debugging (on some systems it may appear as USB Debugging).

After installing the application on your device, you will see the following icon in the phone's menu, which corresponds to the "Skin Lesions" application (figure 5.6):



Figure 5.6 The "Skin Lesions" application icon

In terms of architecture, the "Skin Lesions" expert system also contains a graphical user interface (GUI) (Figure 5.7) that facilitates the analysis of the digital image features of the skin lesion needed to calculate the final score (total dermoscopy score - TDS).

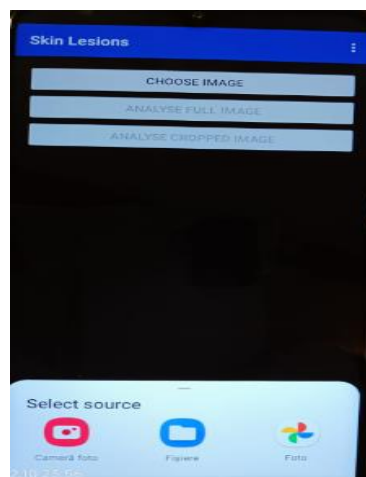


Figure 5.7 Graphical User Interface (GUI) - "Skin Lesions" application

Image selection can be done by clicking on the CHOOSE IMAGE button, which has the following options:

- camera button which will take an instant digital image (Figure 5.8);



Figure 5.8 Camera button

- File button which makes it possible to choose a digital image that is saved to the phone or to your own drive (Figure 5.9);



Figure 5.9 Fisier button

- Photo button which makes it possible to choose a digital image from the mobile phone, WhatsApp or Messenger (Figure 5.10).

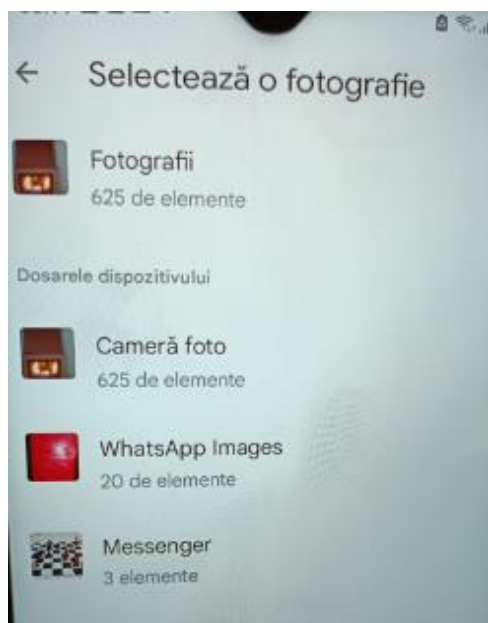


Figure 5.10 Foto button

After selecting the digital image of the skin lesion, the original image can be analysed or it can be analysed after performing digital image resizing by pinching (zooming in and out by removing two fingers or bringing two fingers together) (Figure 5.11).

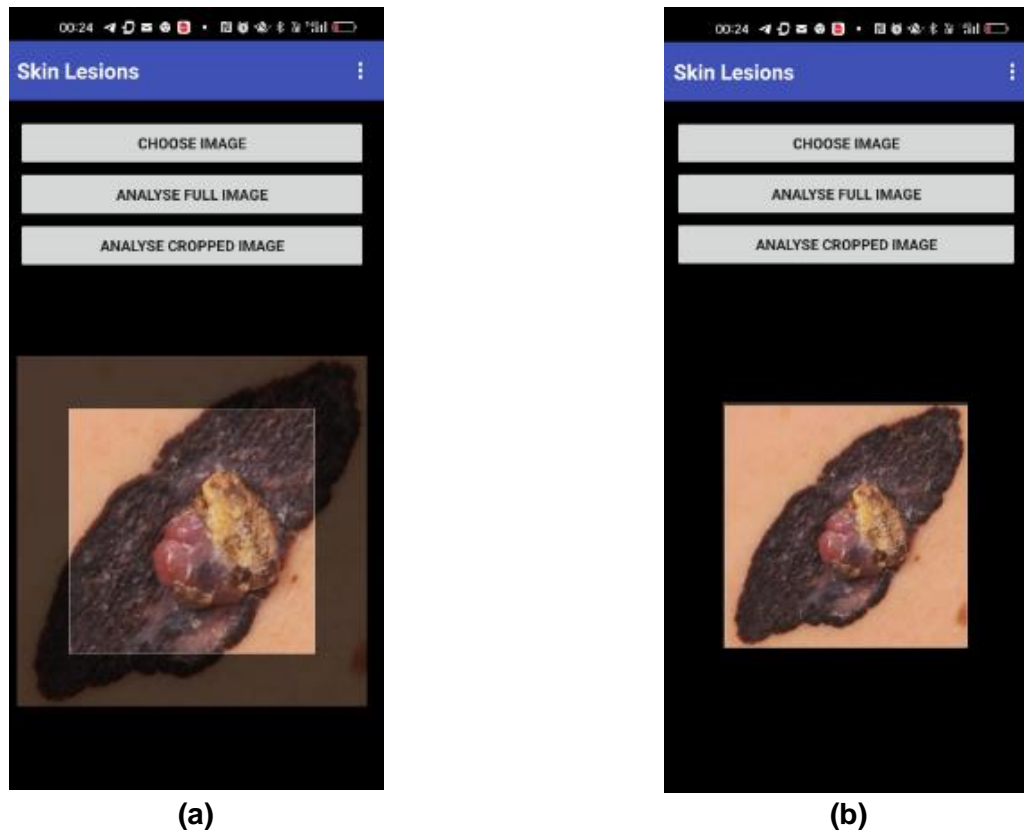


Figure 5.11 Graphical User Interface (GUI) window with selected digital image.
(a) Original image; (b) Resized image

După alegerea imaginii, aceasta poate fi analizată. În urma analizei și în funcție de valoarea TDS se pot obține următoarele rezultate (figura 5.12):

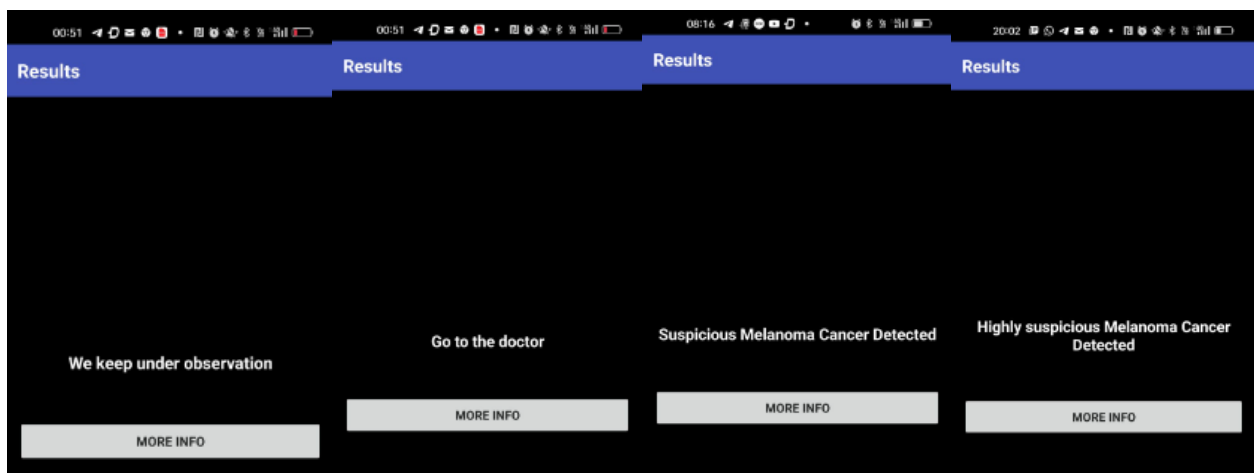


Figure 5.12 Graphical User Interface (GUI) window with messages obtained.
(a) Benign melanocytic lesion; (b) Slightly suspicious lesion; (c) Lesion suspected to be melanoma; (d) Lesion with a high degree of suspicion to be melanoma.

If you want to display the values of A, B, C, D and TDS score you can open the More info button (figure 5.13).

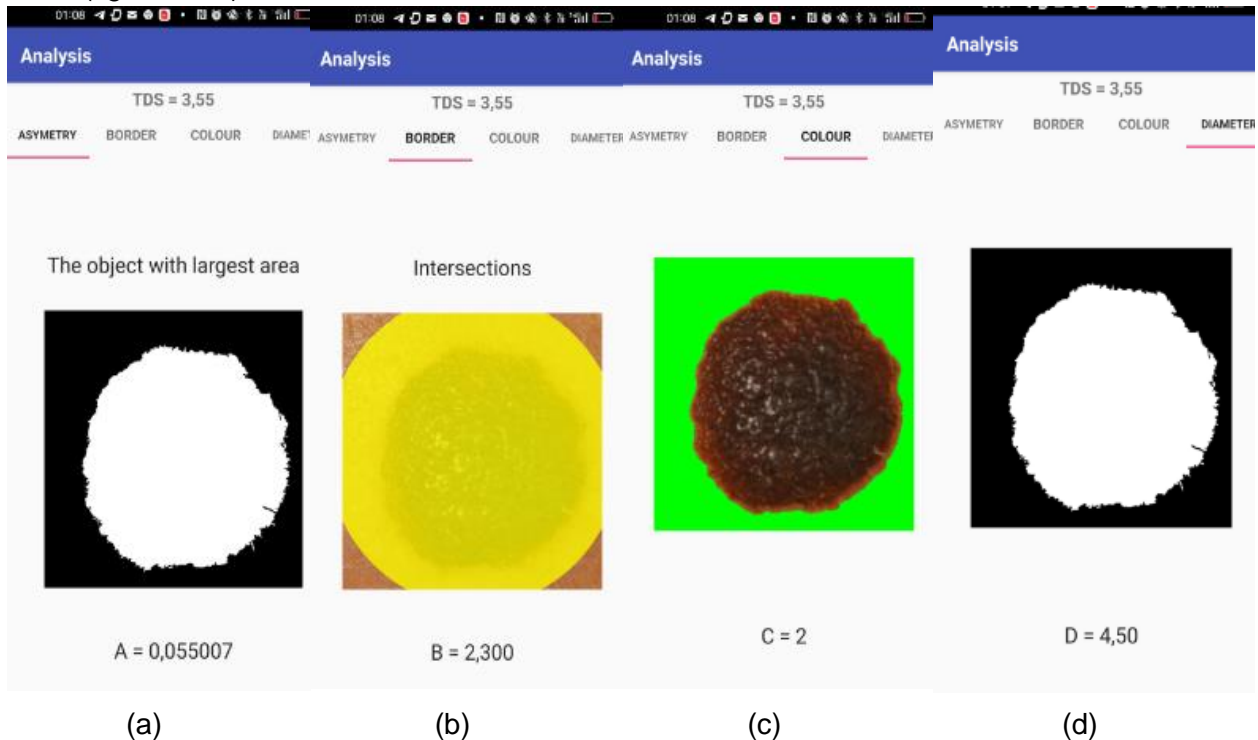


Figure 5.13 Graphical User Interface (GUI) window with the results.

(a) Asymmetry; (b) Border; (c) Color; (d) Diameter

These results show how efficiently melanoma detection is performed and how skin lesion classification is performed.

5.4 Original contributions

The original contributions to this chapter can be summarized as follows:

- presenting and explaining the stages of development of the "**Skin Lesions**" expert system;
- presentation of the methodology used for the development of the "**Skin Lesions**" expert systems.

To evaluate the performance of the proposed skin lesion differentiation model, 170 digital colour images (70 melanomas and 100 nevi) from the Digital Archive of the Department of Dermatology of the University Medical Centre Groningen (UMCG) were used. Table 5.2 shows the mean values of the performance measures. Performance was tested on the parameters sensitivity (SE), specificity (SP), accuracy (Acc) and precision (P). The formulas for calculating the mentioned evaluation values are presented in Chapter 2 (2.81 - 2.87). TP, TN, FP, FN represent true positive, true negative, false positive and false negative, respectively. According to the TDS values images in the melanoma class are considered as true positive (TP) if skin lesions are correctly detected, otherwise they are considered as false negative (FN). As for images in the nevus class they are considered as true negative (TN) if they are predicted as nevus, otherwise they are considered as false positive (FP). ROC and AUC curves are illustrated in Figure 5.14.

In the experiment, 80% of the images are used for training purposes and 20% are used for testing purposes.

Table 5.2 Values of performance measures..

TDS	TP	TN	FP	FN	SE (%)	SP (%)	ACC (%)	P (%)
MELANOMA	46	117	3	4	92.00	97.50	95.88	93.88
NEV	71	90	4	5	93.42	95.74	94.71	94.67

Table 5.2 shows an SE of 92% for melanoma versus 93.42% for nev, an SP of 97.50% for melanoma versus 95.74% for nev, an ACC of 95.88% for melanoma versus 94.71% for nev, a P of 93.8% for melanoma versus 94.67% for nev. In terms of SE, SP, ACC and P for the images analysed, it shows a balanced accuracy of correct classification for melanoma and nevus.

Figure 5.14 shows the AUC values from the ROC curves which have values between 0.826 and 0.938 indicating a good separation of the two classes: nevus and melanoma.

We estimated the values of the performance measures of the expert system „Skin Lesions" as the average between the two classes thus obtaining: for SE - 92.71%, SP - 96.62%, ACC - 95.29%, P - 94.27%.

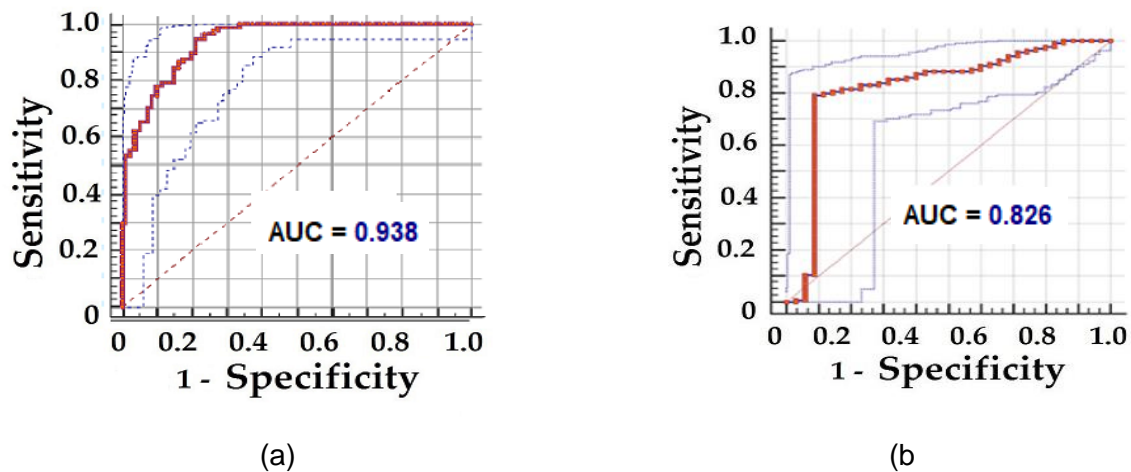


Figure 5.14 ROC curves and area under AUC corresponding to TDS values.
(a) for melanoma; (b) for nev

We compared the results obtained with the clinical risk assessment results of the integrated algorithms in the applications (Table 5.3). It can be seen that the proposed algorithm achieved a significant improvement in application accuracy.

Table 5.3 Comparison of the results of the proposed application with those of existing applications..

Apps	SE (%)	SP (%)
SkinVision [112]	73	83
Lübax [112]	90,4	91,5
SkinMD [212]	85,9	86
MelaFind [213]	82,5	52,4
FotoFinder Moleanalyzer Pro [213]	88,1	78,8
Verisante Aura [213]	21,4	86,2
Skin Lesions [propus]	92,71	96,62

The accuracy of skin lesion detection applications is measured by calculating sensitivity and specificity. Sensitivity quantifies the ability of the app to accurately diagnose the presence of melanoma, while specificity reflects the ability of the app to accurately diagnose the absence of melanoma.

This comparison demonstrates that the highest sensitivity and specificity of the tools were established with Skin Lesions, which could be a valuable tool to raise a red flag, but does not replace clinical decision making by the medical specialist. This app can help users be more mindful of their health care and improve communication between them and their doctors, but it is important that users do not allow the "app to replace doctor's advice and doctor's diagnosis".

Chapter 6

General conclusions and future research directions

The research undertaken shows that image processing is still in need of innovative ideas, both in terms of image processing software solutions and hardware solutions that acquire digital dermoscopic images without defects, artefacts or noise.

In this PhD thesis, I have developed a series of own algorithms that are implemented as software solutions well suited for the analysis and management of skin lesions to make it easier to identify and diagnose early stage of malignant melanomas. Therefore, I can state that I have successfully achieved the overall objective.

The studies cited in this thesis were taken after a careful literature search using several databases, namely:

- IEEE Xplore (<http://ieeexplore.ieee.org/>);
- Springer Link (<http://link.springer.com/>);
- Science Direct (<http://www.sciencedirect.com/>);
- Wiley Online Library (<http://onlinelibrary/>);
- wiley.com/, Web of Science (<http://webofscience.com/>);
- PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>);
- Google Scholar (<https://scholar.google.com/>).

The proposed methods and algorithms have been rigorously validated and accepted in the literature through scientific papers published in prestigious journals and communications at national and international conferences.

For the realization of the PhD thesis I have developed the theoretical part, the methodological part and the applied part which represent original contributions in the field of research aimed at the realization of an expert system for the evaluation of skin lesions. Thus:

➤ **In terms of research objectives:**

- in the introduction chapter, we identified the importance of developing an expert system capable of evaluating skin lesions based on a set of features correlated with the decision rules used by dermatological medical specialists and we have consequently established the research aim and objectives in the present thesis. We identified the possibility of realization and implementation of the expert system, "**Skin Lesions**" practicable for mobile phone users.

➤ **In terms of practical and applicative contributions:**

- In Chapter 3 we conducted studies on: analysis of skin lesions with significant features [193], study of skin lesion asymmetry using artificial neural networks [194], comparison of distance metrics for skin lesion differentiation [195], selection of relevant features from non-dermoscopic images for classification of nevi vs. melanomas [196], influence of color space on skin lesion classification using statistical features [197], invisible watermarking algorithm applied to dermoscopic images decomposed with discrete wavelet transform [195].

- In Chapter 4 we conducted studies on: skin lesion analysis with geometric feature selection algorithm [199], color histogram analysis for skin lesion differentiation [197], color cluster classification with kNN algorithm [198], skin lesion classification based on surface fractal dimensions and statistical color cluster features using a set of machine learning techniques [202], skin lesion differentiation using Random Forest (RF) algorithm [206].

- In Chapter 5 to ensure the applicability of the thesis we developed algorithms necessary for determining A, B, C, D features of skin lesions in order to implement the expert system "Skin Lesions".

It can be concluded that it is possible to use an expert system for self-assessment of skin lesions implemented on mobile phone using image analysis and machine learning, so that any

General conclusions and future research directions

person with a mobile phone can use it. This was the overall objective of the research thesis although melanoma image analysis and classification programs have been addressed and received solutions in several papers in the past, they are not publicly available, except for the MelApp and Skin Scan apps mentioned above, which are currently available for a fee.

Future research directions include:

- testing skin lesion classification algorithms on other digital dermoscopic images acquired with a mobile device. I believe that the major advantage of this approach is that it is cheap, fast, robust and easy to use, with very few constraints on the acquisition protocol of images containing melanomas;
- extending the training database would help future predictions given that in this thesis, the sweeps were performed on images collected from dedicated databases but using different acquisition techniques (camera settings, positioning, lighting conditions, etc.);
- adaptation of the expert system to multiple smartphone operating systems given that currently the application can only be used on mobile phones with the Android operating system.

List of published works

Papers published in ISI-listed journals

1. **Damian, F. A.**, Moldovanu, S., Dey, N., Ashour, A. S., Moraru, L., *Feature Selection of Non-Dermoscopic Skin Lesion Images for Nevus and Melanoma Classification*, MDPI, Journals Computation, Volume 8, 2020, Issue 2: 41, eISSN: 2079-3197, doi: 10.3390/computation8020041, <https://www.mdpi.com/2079-3197/8/2/41>, (IF=0, emerging) **11 citări.**
2. Moldovanu, S., **Damian Michis, F. A.**, Biswas, K. C., Culea-Florescu, A., Moraru, L., *Skin lesion classification based on surface fractal dimensions and statistical color cluster features using an ensemble of machine learning techniques*, Cancers, Volume 13, 2021, Issue 21: 5256, eISSN: 2072-6694, doi:10.3390/cancers13215256, <https://www.mdpi.com/2072-6694/13/21/5256>, (IF=6,639 Clarivate) **8 citări.**

Papers published in ISI Proceedings

1. Moldovanu, S., Stefanescu, D., Damian, A. F., *Watermarking skin lesion digital color images based on channels*, 24th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 2020, Pages 933-936, ISSN: 2372-1618 doi: 10.1109/ICSTCC50638.2020.9259631, <https://ieeexplore.ieee.org/document/9259631>
2. **Damian, F. A.**, Moldovanu, S., Moraru, L., *Skin Lesions Asymmetry Estimation Using Artificial Neural Networks*, 25th International Conference on System Theory, Control and Computing (ICSTCC), Iași, România, 2021, Pages 64-67, ISSN: 2372-1618, doi: 10.1109/ICSTCC52150.2021.9607133, <https://ieeexplore.ieee.org/document/9607133>
3. **Felicia Anișoara Damian**, Simona Moldovanu, Luminița Moraru, *Melanoma Detection using a Random Forest Algorithm*, The 10th IEEE International Conference on E-Health and Bioengineering - EHB 2022, Grigore T. Popa University of Medicine and Pharmacy, 17-19 November 2022, Iași, Romania, ISSN: 2575-5145, doi: 10.1109/EHB55594.2022.9991668, <https://ieeexplore.ieee.org/document/9991668>

Papers published in volumes of international conferences

1. **Felicia Anișoara Damian**, Simona Moldovanu, Luminița Moraru, *Skin lesions analysis based on significant features selection algorithm*, 8th International Conference on Experiments/ Process/System/ Modeling/ Simulation/ Optimization (IC-EPSMSO), 3-6 July 2019, Athens, Greece, Pages 380-387, ISSN: 2241-9209, https://www.researchgate.net/publication/349806616_IC-EPSMSO_8_thInternational_Conferenceon_Experiments_Process_System_Modeling_Simulation_Optimization_Proceedings_VOLUME_II_Pages_200-475
2. **Damian, F. A.**, Moldovanu, S., & Moraru, L., *Colour histogram analysis for skin lesions discrimination*, SGEM Conference Proceedings: 19th International Multidisciplinary Scientific GeoConference SGEM 2019, Section Advances in Biotechnology, 9 - 11 December 2019, Vienna, Austria, Volume 19, Issue 6.3, Pages 75-82, doi:10.5593/sgem2019V/ 6.3/S08.010, <https://search.proquest.com/docview/2348209566/C48A0D4612324515PQ/2>

3. S. Moldovanu, **F. A. Damian**, L. Moraru, *A kNN approach for melanoma diagnosis based on color cluster features*, 3rd International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAl' 2021), 17-19 November 2021, Porto, Portugal, Pages 6-8, eISSN: 978-84-09-35865-6,
https://www.sensorsportal.com/ASPAl_2021/ASPAl_2021_Conference_Proceedings.pdf

Papers published in BDI-listed journals

1. Moraru, L., **Damian, F. A.**, Moldovanu, S., *Analysis of Common Melanomas and Nevi Using Geometrical Features*, Annals of "Dunarea De Jos" University Of Galati Mathematics, Physics, Theoretical Mechanics, Fascicle II, YEAR XI(XLII) 2019, Volume 42, No. 1, Pages 27-33,
<https://doi.org/10.35219/ann-ugal-math-phys-mec.2019.1.04>
2. **Damian, F. A.**, Moldovanu, S. and Moraru, L., *Comparison of distance metric measures in differentiation of non - dermoscopic skin lesions*, Annals Of "Dunarea De Jos" University of Galati Mathematics, Physics, Theoretical Mechanics, Fascicle II, YEAR XII(XLIII) 2020, Volume 43, No. 1, Pages 20- 27, <https://doi.org/10.35219/ann-ugal-math-phys-mec.2020.1.03>
3. **Damian, F.**, Moldovanu, S., Moraru, L., *Color space influence on ANN skin lesion classification using statistics texture feature*, Annals of the "Dunarea de Jos" University of Galati. Fascicle II, Mathematics, Physics, Theoretical Mechanics, YEAR XIII(XLIV) 2021 Volume 44, No.1, Pages 53-62. <https://doi.org/10.35219/ann-ugal-math-phys-mec.2021.1.08>

Papers presented at national and international conferences

1. **Damian Felicia Anișoara**, Moldovanu Simona, Moraru Luminița, *Analizarea nevilor pe baza caracteristicilor semnificative*, Conferința științifică: "Fizica medicală: simbioză între Fizică, Medicină și Mediu", 28 Februarie 2019, Galați, Secțiunea 1: Fizică medicală; aplicații ale Fizicii în Medicină,
https://www.ugal.ro/files/stiri%20si%20evenimente/program_conferinta_Fizica_medicala_2019.pdf
2. Luminița Moraru, **Felicia Anișoara Damian**, Simona Moldovanu, *Analysis of Common Melanomas and Nevi Using Geometrical Features*, Scientific Conference of Doctoral Schools from "Dunărea de Jos" University of Galati (SCDS-UDJG 2019), 7th Edition, 13th-14th of June 2019, Galați, Section 2. Advanced investigation methods in environment and biohealth,
http://www.cssd-udjg.ugal.ro/files/2019/Program_detaliat_al_conferintei_nou.pdf
3. **Damian Felicia Anișoara**, Moldovanu Simona, Moraru Luminița, *Skin lesions analysis based on significant features selection algorithm*, 8th International Conference on Experiments/ Process/System/ Modeling/ Simulation/ Optimization, 3-6 July, Athens, Greece,
<https://lfme.gr/8th-ic-epsmsso-program/#1645452905480-7ab7aae2-2cc5>
4. **Felicia Anișoara Damian**, Simona Moldovanu, Luminita Moraru, *Colour histogram analysis for skin lesions discrimination*, 19th International Multidisciplinary Scientific GeoConference SGEM, Section Advances in Biotechnology, 9 - 11 December 2019, Vienna, Austria,
https://www.sgemviennagreen.org/images/deadlines/conference_programme/Day2_Programme_POSTER.pdf
5. **Felicia Anișoara Damian**, Simona Moldoveanu, Luminița Moraru, *Comparison of distance metric measures in differentiation of non-dermoscopic skin lesions*, Scientific Conference of Doctoral Schools from "Dunărea de Jos" University of Galati SCDS-UDJG, 8th

Edition, 18th-19th of June 2020, Galați, Section 2. Advanced investigation methods in environment and biohealth,

http://www.cssd-udjg.ugal.ro/files/2020/program/CSSD_Section2_Program_2020_draft.pdf

6. Moldovanu, S., **Damian, F. A.**, Pană, L., Moraru, L., *Machine Learning and Artificial Intelligence for health monitoring*, International Conference Environmental Challenges in the Black Sea Basin: Impact on Human Health, 23 – 26 September 2020, Galati, Romania, Pages 47-47a, https://ibn.idsi.md/vizualizare_articol/115939

7. **Damian, F.**, Moldovanu, S., Moraru, L., *Color space influence on ANN skin lesion classification using statistics texture feature*, Scientific Conference of Doctoral Schools from “Dunărea de Jos” University of Galati SCDS-UDJG, 9th Edition, 10th- 11th of June 2021, Galați, Section 2. Advanced investigation methods in environment and biohealth, <http://www.cssd-udjg.ugal.ro/index.php/2020-2/programme-22>

8. S. Moldovanu, **F. A. Damian**, L. Moraru, *A kNN approach for melanoma diagnosis based on color cluster features*, 3rd International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAI' 2021), 17-19 November 2021, Porto, Portugal,

https://www.sensorsportal.com/ASPAI_2021/ASPAI_2021_Conference_Programme.pdf

9. **F. A. Damian**, S. Moldovanu, L. Moraru, *Skin Lesions Asymmetry Estimation Using Artificial Neural Networks*, 25th International Conference on System Theory, Control and Computing (ICSTCC), Iași, România, 2021,

https://icstcc2021.ac.tuiasi.ro/wp-content/uploads/2021/10/Book-of-Abstracts-ICSTCC-2021_V5-online.pdf

10. **Felicia Anișoara Damian**, Simona Moldovanu, Luminița Moraru, *Melanoma Detection using a Random Forest Algorithm*, The 10th IEEE International Conference on E-Health and Bioengineering - EHB 2022, Grigore T. Popa University of Medicine and Pharmacy, 17-19 November 2022, Iași, Romania,

http://www.ehbconference.ro/Portals/18/EHB2022_Detailed_Program.pdf

Citations:

Damian, F. A., Moldovanu, S., Dey, N., Ashour, A. S., Moraru, L., *Feature Selection of Non-Dermoscopic Skin Lesion Images for Nevus and Melanoma Classification*, MDPI, Journals Computation, Volume 8, 2020, Issue 2: 41, eISSN: 2079-3197, doi:10.3390/computation8020041, <https://www.mdpi.com/2079-3197/8/2/41> (IF=0, emerging).

Cited by:

1. Ashour, A. S., Nagieb, R. M., El-Khobby, H. A., Abd Elnaby, M. M., & Dey, N, *Genetic algorithm-based initial contour optimization for skin lesion border detection*, Multimedia Tools and Applications, Volume 80, 2020, Issue 2, Pages 2583-2597, eISSN: 1573-7721, doi: 10.1007/s11042-020-09792-8, <https://dl.acm.org/doi/abs/10.1007/s11042-020-09792-8>

2. Srivastava, R., Ong, E. P., Lee, B. H., Tan, L. S., Tey, H. L., *Quantitative Comparison of Color Asymmetry Features for Automatic Melanoma Detection*, 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, Pages 3753-3756, eISSN:1558-4615, doi: 10.1109/EMBC46164.2021.9631103,

<https://ieeexplore.ieee.org/document/9631103>

3. Hu, L., Chen, Q., Qiao, L., Du, L., Ye, R., *Automatic Detection of Melanins and Sebums from Skin Images Using a Generative Adversarial Network*, Cognitive Computation, Volume 14, 2022, Issue 5, Pages 1599-1608, eISSN: 1866-9964, doi: 10.1007/s12559-021-09870-5,

<https://link.springer.com/article/10.1007/s12559-021-09870-5>

4. Moldovanu, S., Obreja, C. D., Biswas, K. C., Moraru, L., *Towards Accurate Diagnosis of Skin Lesions Using Feedforward Back Propagation Neural Networks*, *Diagnostics*, Volume 11, 2021, Issue 6: 936, Pages 1-16, eISSN: 1573-7721, doi: 10.3390/diagnostics11060936, <https://www.mdpi.com/2075-4418/11/6/936>
5. Barhoumi, W., Khelifa, A., *Skin lesion image retrieval using transfer learning-based approach for query-driven distance recommendation*, *Computers in Biology and Medicine*, Volume 137, 2021, 104825, eISSN: 1879-0534, doi: 10.1016/j.combiomed.2021.104825, <https://www.sciencedirect.com/science/article/abs/pii/S0010482521006193?via%3Dihub>
6. Mello-Román, J.C., Vázquez Noguera, J.L., Legal-Ayala, H., García-Torres, M., Facon, J., Pinto-Roa, D.P., Grillo, S.A., Salgueiro Romero, L., Salgueiro Toledo, L.A., Bareiro Paniagua, L.R. and Leguizamon Correa, D.N., *Dermoscopy Images Enhancement via Multi-Scale Morphological Operations*, *Applied Sciences*, Volume 11, 2021, Issue 19: 9302, Pages 1-16, eISSN: 2076-3417, doi: 10.3390/app11199302, <https://www.mdpi.com/2076-3417/11/19/9302>
7. Wang, J., Yan, Y., Kim, J., *Classification of melanoma images using 2D multifractal detrended cross-correlation analysis*, *Modern Physics Letters B*, Volume 36, 2022, Issue 09: 2150619, eISSN: 1793-6640, doi:10.1142/S0217984921506193, <https://www.worldscientific.com/doi/abs/10.1142/S0217984921506193>
8. Anghelache Nastase, I. N., Moldovanu, S., Moraru, L., *Image Moment-Based Features for Mass Detection in Breast US Images via Machine Learning and Neural Network Classification Models*, *Inventions*, Volume 7, 2022, Issue 2: 42, eISSN: 2411-5134, doi: 10.3390/inventions7020042, <https://www.mdpi.com/2411-5134/7/2/42>
9. Gulzar, Y., & Khan, S. A., *Skin Lesion Segmentation Based on Vision Transformers and Convolutional Neural Networks-A Comparative Study*, *Applied Sciences*, Volume 12, 2022, Issue 12: 5990, eISSN: 2076-3417, doi: 10.3390/app12125990, <https://www.mdpi.com/2076-3417/12/12/5990>
10. Radhika, V., & Chandana, B. S., *Skin Melanoma Classification from Dermoscopy Images using ANU-Net Technique*, *International Journal of Advanced Computer Science and Applications*, Volume 13, 2022, Issue 10, ISSN: 2158-107X, doi: 10.14569/IJACSA.2022.01310109, https://thesai.org/Downloads/Volume13No10/Paper_109-Skin_Melanoma_Classification_from_Dermoscopy_Images.pdf
11. Hasan, M. K., Ahamad, M. A., Yap, C. H., & Yang, G., *A survey, review, and future trends of skin lesion segmentation and classification*, *Computers in biology and medicine*, Volume 155, 2023, ISSN: 0010-4825, doi:10.1016/j.combiomed.2023.106624, <https://arxiv.org/abs/2208.12232>

Citations:

Moldovanu, S., **Damian Michis, F. A.**, Biswas, K. C., Culea-Florescu, A., Moraru, L., *Skin lesion classification based on surface fractal dimensions and statistical color cluster features using an ensemble of machine learning techniques*, *Cancers*, Volume 13, 2021, Issue 21: 5256, eISSN: 2072-6694, doi: 10.3390/cancers13215256, <https://www.mdpi.com/2072-6694/13/21/5256> (IF=6,639 Clarivate).

Cited by:

1. Kara, A. C., Hardalaç, F., *Detection and Classification of Knee Injuries from MR Images Using the MRNet Dataset with Progressively Operating Deep Learning Methods*, *Machine Learning and Knowledge Extraction*, Volume 3, 2021, Issue 4, Pages 1009-1029, eISSN: 2504-

4990, doi: 10.3390/make3040050, <https://www.mdpi.com/2504-4990/3/4/50>

2. Anghelache Nastase, I. N., Moldovanu, S., Moraru, L., *Image Moment-Based Features for Mass Detection in Breast US Images via Machine Learning and Neural Network Classification Models*, Inventions, Volume 7, 2022, Issue 2: 42, eISSN: 2411-5134, doi: 10.3390/inventions7020042, <https://www.mdpi.com/2411-5134/7/2/42>

3. Li, S., Wang, H., Xiao, Y., Zhang, M., Yu, N., Zeng, A., & Wang, X., *A Workflow for Computer-Aided Evaluation of Keloid Based on Laser Speckle Contrast Imaging and Deep Learning*, Journal of Personalized Medicine, Volume 12, 2022, Issue 6: 981, eISSN: 2075-4426, doi: 10.3390/jpm12060981, <https://www.mdpi.com/2075-4426/12/6/981>

4. Popecki, P., Kozakiewicz, M., Ziętek, M., & Jurczyszyn, K., *Fractal Dimension Analysis of Melanocytic Nevi and Melanomas in Normal and Polarized Light-A Preliminary Report*, Life-Basel, Volume 12, 2022, Issue 7: 1008, eISSN: 2075-1729, doi: 10.3390/life12071008, <https://www.mdpi.com/2075-1729/12/7/1008>

5. Hussain, A., Alam, S., Ghauri, S.A., Ali, M., Sherazi, H.R., Akhunzada, A., Bibi, I., Gani, A., *Automatic Modulation Recognition Based on the Optimized Linear Combination of Higher-Order Cumulants*, Sensors, Volume 22, 2022, Issue 19: 7488, eISSN: 1424-8220, doi: 10.3390/s22197488, <https://www.mdpi.com/1424-8220/22/19/7488>

6. Panthakkan, A., Anzar, S.M., Jamal, S., Mansoor, W., *Concatenated Xception-ResNet50-A novel hybrid approach for accurate skin cancer prediction*, Computers in Biology and Medicine, Volume 150, 2022, 106170, eISSN: 1879-0534, doi: 10.1016/j.compbiomed.2022.106170,

<https://www.sciencedirect.com/science/article/abs/pii/S0010482522008782?via%3Dihub>

7. Maiti, A., Sultana, M. and Bhattacharya, S., *Detection of skin cancer through hybrid color features and soft voting ensemble classifier*, Innovations in Systems and Software Engineering, 2022, Pages 1-14, eISSN: 1614-5054, <https://link.springer.com/article/10.1007/s11334-022-00498-8>

8. Hasan, M. K., Ahamad, M. A., Yap, C. H., & Yang, G., *A survey, review, and future trends of skin lesion segmentation and classification*, Computers in biology and medicine, Volume 155, 2023, ISSN: 0010-4825, doi:10.1016/j.compbiomed.2023.106624,

<https://arxiv.org/abs/2208.12232>

Selected bibliography

- [36] M. E. Celebi, N. Codella, A. Halpern, *Dermoscopy Image Analysis: Overview and Future Directions*, IEEE Journal of Biomedical and Health Informatics, Volume 23, 2019, Issue 2, Pages 474–478, eISSN 2168-2208, doi: 10.1109/JBHI.2019.2895803, <https://pubmed.ncbi.nlm.nih.gov/30703051/>;
- [37] A. W. Choi, R. S. Xu, S. Jacob, B. O. Dulmage, M. L. Colavincenzo, J. K. Robinson, S. Xu, *Visual perception training: a prospective cohort trial of a novel, technology-based method to teach melanoma recognition*, Postgraduate Medical Journal, Volume 95, 2019, Issue 1124, Pages 350-352, eISSN 1469-0756, doi: 10.1136/postgradmedj-2018-136379 <https://pubmed.ncbi.nlm.nih.gov/31266882/>;
- [38] D. Coppola, H. K. Lee, C. Guan, *Interpreting mechanisms of prediction for skin cancer diagnosis using multi-task learning*, Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, Pages 3162-3171, doi: 10.1109/CVPRW50498.2020.00375, <https://ieeexplore.ieee.org/document/9150967>;
- [39] S. W. Menzies, C. Ingvar, K. A. Crotty, W. H. McCarthy, *Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features*, Archives of dermatology, Volume 132, 1996, Issue 10, Pages 1178–1182, ISSN 0003-987X, <https://pubmed.ncbi.nlm.nih.gov/8859028/>;
- [40] J. S. Henning, S. W. Dusza, S. Q. Wang, A. A. Marghoob, H. S. Rabinovitz, D. Polsky, A. W. Kopf, *The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy*, Journal of the American Academy of Dermatology, Volume 56, 2007, Issue 1, Pages 45-52, eISSN 1097-6787, doi: 10.1016/j.jaad.2006.09.003, <https://pubmed.ncbi.nlm.nih.gov/17190620/>;
- [41] R. Ramji, G. Valdes-Gonzalez, A. Oakley, M. Rademaker, *Dermoscopic ‘Chaos and Clues’ in the diagnosis of melanoma in situ*, Australas Journal Dermatology, Volume 59, 2018, Issue 3, Pages 201–205, eISSN 1440-0960, doi: 10.1111/ajd.12740, <https://pubmed.ncbi.nlm.nih.gov/29094749/>;
- [42] P. Bourne, C. Rosendahl, J. Keir, A. Cameron, *BLINCK-A diagnostic algorithm for skin cancer diagnosis combining clinical features with dermatoscopy findings*, Dermatology Practical & Conceptual, Volume 2, 2012, Issue 2, Pages 202a12, eISSN 2160-9381, doi: 10.5826/dpc.0202a12, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3663344/>;
- [85] N. Otsu, *A Threshold Selection Method from Gray-Level Histograms*, Transactions on Systems, Man, and Cybernetics, Volume 9, 1979, Issue 1, Pages 62-66, eISSN: 2168-2909, doi: 10.1109/TSMC.1979.4310076, <https://ieeexplore.ieee.org/document/4310076>;
- [116] ***, https://download.cnet.com/DermLite/3000-2129_4-76936578.html;
- [117] ***, <https://hub.fotofinder.de>;
- [118] ***, <https://apps.apple.com/us/app/skin-cancer-app-myskinpal-map-your-skin-moles/id955071565>;
- [119] ***, <https://apps.apple.com/ca/app/molescope/id1003576096>
- [120] ***, <https://appadvice.com/app/apreskin-you-track-your-moles-skin-cancers-melanoma/1131517287>;
- [121] ***, <https://apps.apple.com/us/app/derma-analytics/id1174696625>;
- [122] ***, <https://miiskin.com>;
- [123] ***, <https://apkpure.com/apd-skin-monitoring-app/apd.bii.wound>;
- [124] ***, <https://www.skinvision.com>;

- [125] A. Udrea, M. Tanase, D. Popescu, *Nonlinear deterministic methods for computer aided diagnosis in case of kidney diseases*, 9th International Conference on Informatics in Control, Automation and Robotics, Rome, Italy, Volume 1, 2012, Pages 511–516, <https://pdfs.semanticscholar.org/29ac/2b0f2fb8922584a0cf282db9b3eec02f431e.pdf>;
- [126] T. Maier, D. Kulichova, K. Schotten, R. Astrid, T. Ruzicka, C. Berking, A. Udrea, *Accuracy of a smartphone application using fractal image analysis of pigmented moles compared to clinical diagnosis and histological result*, Journal of the European Academy of Dermatology and Venereology, Volume 29, 2015, Issue 4, Pages 663-667, ISSN 1468-3083, doi: 10.1111/jdv. 12648, <https://pubmed.ncbi.nlm.nih.gov/25087492/>;
- [127]***, <https://skinmdnow.com/freeapp.html>;
- [162] U.K. Acharya, S. Kumar, *Directed searching optimized mean-exposure based sub-image histogram equalization for grayscale image enhancement*, Multimedia Tools and Applications, Volume 80, 2021, Issue 16, Pages 24005–24025, ISSN 1380-7501, doi: 10.1007/s11042-021-10855-7, <https://link.springer.com/article/10.1007/s11042-021-10855-7>;
- [163] H. Chugh, S. Gupta, M. Garg, D. Gupta, S. Juneja, H. Turabieh, Y. Na, Z. Kiros Bitsue, *Image Retrieval Using Different Distance Methods and Color Difference Histogram Descriptor for Human Healthcare*, Journal of Healthcare Engineering, Volume 2022, 2022, Pages 1-10, ISSN 2040-2295, doi:10.1155/2022/9523009, <https://www.hindawi.com/journals/jhe/2022/9523009/>;
- [164] V. Rajinikantha, M. S. Couceirob, *RGB Histogram based Color Image Segmentation Using Firefly Algorithm*, Procedia Computer Science, Volume 46, 2015, Pages 1449-1457, ISSN 1877-0509, doi:10.1016/j.procs.2015.02.064, <https://www.sciencedirect.com/science/article/pii/S1877050915001283>;
- [167] G.Tzortzis, A. Likas, *The MinMax k-Means clustering algorithm*, Pattern recognition, Volume 47, 2014, Issue 7, Pages 2505-2516, ISSN 0031-3203, doi: 10.1016/j.patcog.2014.01.015, <https://www.sciencedirect.com/science/article/pii/S0031320314000338>;
- [168] X. Liu, L. Song, S. Liu, Y. Zhang, *A Review of Deep-Learning-Based Medical Image Segmentation Methods*, Sustainability, Volume 13, 2021, Issue 3: 1224, Pages 1-29, ISSN 2071-1050, doi: 10.3390/su13031224, <https://www.mdpi.com/2071-1050/13/3/1224>;
- [172] D. J. Bora, A. K. Gupta, *A new approach towards clustering based color image segmentation*, International Journal of Computer Applications, Volume 107, 2014, Issue 12, Pages 23-30, ISSN 0975 –8887, doi: 10.5120/18803-0329, <https://research.ijcaonline.org/volume107/number12/pxc3900329.pdf>;
- [173] M. R. Ferreira, F. D. A. De Carvalho, *Kernel fuzzy c-means with automatic variable weighting*, Fuzzy Sets and Systems, Volume 237, 2014, Pages 1–46, ISSN 0165-0114, doi: 10.1016/j.fss.2013.05.004, <https://www.sciencedirect.com/science/article/pii/S0165011413002054>;
- [174] K. Rajkumar, A. Yesubabu, K. Subrahmanyam, *Fuzzy clustering and Fuzzy C-Means partition cluster analysis and validation studies on a subset of CiteScore dataset*, International Journal of Electrical and Computer Engineering, Volume 9, 2019, Issue 4, Pages 2760-270, ISSN: 2088-8708, doi: 10.11591/ijece.v9i4.pp2760-2770, <https://www.mendeley.com/catalogue/4e436394-dd3c-3e2a-9a5f-1900ab10368a/>;
- [175] Z. H. Hoo, J. Candlish, D. Teare, *What is an ROC curve?*, Emergency Medicine Journal, Volume 34, 2017, Issue 6, Pages 357-359, ISSN 1472-0205, doi: 10.1136/emj-2017-206735, <https://emj.bmj.com/content/emj/34/6/357.full.pdf>;

- [176] N. A. Obuchowski, J. A. Bullen, *Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine*. Physics in medicine and biology, 2018 Mar Volume 63, Issue 7, Pages 07TR01, ISSN 1361-6560, doi: 10.1088/1361-6560/aab4b1, <https://dx.doi.org/10.1088/1361-6560/aab4b1>;
- [177] D. İlhan Topcu, H. Can Çubukçu, *Optimization of patient-based real-time quality control based on the Youden index*, Clinica Chimica Acta, Volume 534, 2022, Pages 50-56, ISSN 0009-8981, doi: 10.1016/j.cca.2022.06.028, <https://www.sciencedirect.com/science/article/pii/S0009898122012189>;
- [178] M. W. Browne, *Cross-validation methods*, Journal of mathematical psychology, Volume 44, 2000, Issue 1, Pages 108-132, ISSN 0022-2496, doi: 10.1006/jmps.1999.1279, <https://www.sciencedirect.com/science/article/pii/S0022249699912798>;
- [179] C. Bergmeir, R. J. Hyndman, B. Koo, *A note on the validity of cross-validation for evaluating autoregressive time series prediction*, Computational Statistics & Data Analysis, Volume 120, 2018, Pages 70–83, ISSN 0167-9473, doi: 10.1016/j.csda.2017.11.003, <https://www.sciencedirect.com/science/article/pii/S0167947317302384>;
- [180] F. M. Bayer, F. Cribari-Neto, *Model selection criteria in beta regression with varying dispersion*, Communications in Statistics-Simulation and Computation, Volume 46, 2017, Issue 1, Pages 729-746, ISSN 1532-4141, doi: 10.1080/03610918.2014.977918, <https://doi.org/10.1080/03610918.2014.977918>;
- [181] N. Khasawneh, M. Fraiwan, L. Fraiwan, B. Khasawneh, A. Ibrani, *Detection of COVID-19 from Chest X-ray Images Using Deep Convolutional Neural Networks*, Sensors, Volume 21, 2021, Issue 17, Pages 5940, doi: 10.3390/s21175940, ISSN 1424-8220, <https://europepmc.org/articles/PMC8434649>;
- [182] H. Ullah, M. Bhuiyan, *Performance Evaluation of Feed Forward Neural Network for Image Classification*, Journal of Science and Technology, Volume 10, 2018, Issue 1, Pages 19-27, doi:10.30880/jst.2018.10.01.004 https://www.researchgate.net/publication/325258560_Performance_Evaluation_of_Feed_Forward_Neural_Network_for_Image_Classification;
- [183] L. Abdel-Ilah, H. Šahinbegović, *Using machine learning tool in classification of breast cancer*, International Conference on Medical and Biological Engineering (CMBEBIH), Sarajevo, Volume 62, 2017, Pages 3-8, ISBN 978-981-10-4166-2, doi: 10.1007/978-981-10-4166-2_1, https://link.springer.com/chapter/10.1007/978-981-10-4166-2_1;
- [184] R. Kaur, H. GholamHosseini, R. Sinha, M. Lindén, *Melanoma Classification Using a Novel Deep Convolutional Neural Network with Dermoscopic Images*, Sensors, Volume 22, 2022, Issue 3: 1134, doi:10.3390/s22031134, ISSN 1424-8220, <http://europepmc.org/article/MED/35161878>;
- [185] D. S. K. Karunasingha, *Root mean square error or mean absolute error? Use their ratio as well*, Information Sciences, Volume 585, 2022, Pages 609-629, ISSN 0020-0255, doi: 10.1016/j.ins.2021.11.036, <https://www.sciencedirect.com/science/article/pii/S0020025521011567>;
- [186] A. Alexandridis, E. Chondrodima, *A medical diagnostic tool based on radial basis function classifiers and evolutionary simulated annealing*, Journal of biomedical informatics, Volume 49, 2014, Pages 61-72, ISSN 1532-0464, doi:10.1016/j.jbi.2014.03.008, <https://www.sciencedirect.com/science/article/pii/S1532046414000653>;

- [187] A. H. Fath, F. Madanifar, M. Abbasi, *Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems*, Petroleum, Volume 6, 2020, Issue 1, Pages 80-91, ISSN 2405-6561, doi:10.1016/j.petlm.2018.12.002, <https://www.sciencedirect.com/science/article/pii/S2405656118301020>;
- [188] M. R. Bonyadi, Q. M. Tieng, D. C. Reutens, *Optimization of Distributions Differences for Classification*, IEEE Transactions on Neural Networks and Learning Systems, Volume 30, 2018, Issue 2, Pages 511-523, eISSN 2162-2388, doi: 10.1109/TNNLS.2018.2844723, <https://ieeexplore.ieee.org/document/8401711>;
- [189] Z. Tembusai, H. Mawengkang, M. Zarlis, *K-Nearest Neighbor with K-Fold Cross Validation and Analytic Hierarchy Process on Data Classification*, International Journal of Advances in Data and Information Systems, Volume 2, 2021, Issue 1, Pages 1-8, ISSN 2721-3056, doi: 10.25008/ijadis.v2i1.1204, <https://media.neliti.com/media/publications/396954-k-nearest-neighbor-with-k-fold-cross-val-1a5c3b3f.pdf>;
- [212] A. Tyagi, K. Miller, M. Cockburn, *e-Health Tools for targeting and improving melanoma screening: a review*, Journal of skin cancer, Volume 2012, 2012, Pages 1-8, ISSN 2090-2913, doi: 10.1155/2012/437502 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3530856>;
- [213] A. N. MacLellan, E. L. Price, P. Publicover-Brouwer, K. Matheson, T. Y. Ly, S. Pasternak, N. M. Walsh, C. J. Gallant, A. Oakley, P. R. Hull, R. G. Langley, *The use of noninvasive imaging techniques in the diagnosis of melanoma: a prospective diagnostic accuracy study*, Journal of the American Academy of Dermatology, Volume 85, 2021, Issue 2, Pages 353-359, ISSN 0190-9622, doi: 10.1016/j.jaad.2020.04.019, <https://www.sciencedirect.com/science/article/pii/S0190962220305594>;

Priority axis 6- Education and skills

Project title: "Programme for increasing performance and innovation in doctoral and postdoctoral research of excellence - PROINVENT"

Contract No: 62487/03.06.2022 POCU/993/6/13 - SMIS Code: 153299

The views expressed in the paper, are those of the author and do not necessarily reflect the views of the European Commission and the University "Dunărea de Jos" of Galati, the beneficiary of the project.