ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

CLASSIFICATION OF DERMOSCOPIC IMAGES USING NEURAL NETWORKS

M.Sc. THESIS

Enes ALBAY

Department of Computer Engineering

Computer Engineering Programme

MAY 2016

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

CLASSIFICATION OF DERMOSCOPIC IMAGES USING NEURAL NETWORKS

M.Sc. THESIS

Enes ALBAY
(504141533)

Department of Computer Engineering

Computer Engineering Programme

Thesis Advisor: Assoc. Prof. Dr. Mustafa KAMAŞAK

MAY 2016

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

DERMOSKOPİK GÖRÜNTÜLERİN SİNİR AĞLARI KULLANILARAK SINIFLANDIRILMASI

YÜKSEK LİSANS TEZİ

Enes ALBAY
(504141533)

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Assoc. Prof. Dr. Mustafa KAMAŞAK

MAYIS 2016

Enes ALBAY, a M.Sc. student of ITU Graduate School of ScienceEngineering and Technology 504141533 successfully defended the thesis entitled “CLASSIFICATION OF DERMOSCOPIC IMAGES USING NEURAL NETWORKS”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assoc. Prof. Dr. Mustafa KAMAŞAK** Istanbul Technical University	.....
Jury Members :	**Assoc. Prof. Dr. Gözde Ünal** Istanbul Technical University	.....
	Assist. Prof. Dr. Gökhan Bilgin Yıldız Technical University	.....

Date of Submission : **05 April 2016**
Date of Defense : **04 May 2016**

To my spouse,

FOREWORD

This thesis has been written as a completion to the master in Computer Engineering, at Istanbul Technical University. The subject of the thesis is on classifications of dermoscopic images that contains skin lesions. Automatically detection of melonama is very popular subject in recent years. Also it is important for human health.

I would like to thank my supervisor Assoc. Prof. Dr. Mustafa E. Kamaşak of being great help during the development of this thesis. In this period, I can not forget support of my dear friend Bilal Aydos to complete my thesis. Also, I would like to give my special thanks to my wife for her patient, neverending support.

May 2016 Enes ALBAY

TABLE OF CONTENTS

Page

FOREWORD ..... **ix**
TABLE OF CONTENTS ..... **xi**
ABBREVIATIONS ..... **xiii**
LIST OF TABLES ..... **xv**
LIST OF FIGURES ..... **xix**
SUMMARY ..... **xxi**
ÖZET ..... **xxiii**
1. INTRODUCTION **1**
1.1 Purpose of Thesis 2
1.2 Literature Review 3
1.3 Hypothesis 4
2. FEATURE EXTRACTION **7**
2.1 ABCD Rule Motivated Features 8
2.1.1 Fourier descriptors extraction 8
2.2 Statistical Features 10
2.3 Wavelet Features 14
2.4 Artificial Neural Networks 16
3. CLASSIFICATION OF LESIONS **21**
3.1 ABCD Rules-Motivated Classification 23
3.2 Classification of Using Statistical Features 26
3.2.1 Classification of using first-order statistical features 26
3.2.2 Classification of using second-order statistical features 28
3.2.3 Wavelet classification 30
3.2.4 Classification after PCA reduction 32
3.2.5 Classification using boxplot filtered features 33
4. CONCLUSIONS AND RECOMMENDATIONS **35**
4.1 ABCD Rules-Motivated Classification 35
4.2 First-Order Statistics Based Classification 37
4.3 Second-Order Statistics Based Classification 40
4.4 Wavelet Transformation 43
4.5 Classification After PCA Reduction 45
4.6 Classification Using Boxplot Filtered Features 47
4.7 Discussion 47
4.8 Conclusion 48
REFERENCES **49**
CURRICULUM VITAE **53**

ABBREVIATIONS

ABCD Rule	**:** Asymmetry, Border, Color, Diameter Rule
ANN	**:** Artificial Neural Network
BP	**:** Back-propagation
DA	**:** Diagnosis Accuracy
DI	**:** Doubt Index
ITU	**:** Istanbul Technical University
kNN	**:** k-Nearest Neighbor
MLP	**:** Multilayer Perceptron
NPV	**:** Negative Predictive Value
NPV	**:** Positive predictive value
PCA	**:** Principal Component Analysis
PPV	**:** Negative Predictive Value
SE	**:** Sensitivity
SP	**:** Specificity
SVM	**:** Support Vector Machine

LIST OF TABLES

Page

Table 3.1 : Contingency table. 22
Table 3.2 : Using Fourier descriptors magnitude and phase of border, histogram of lesion radius and size of lesion. 24
Table 3.3 : Using Fourier descriptors magnitudes. 24
Table 3.4 : Using Fourier descriptor angles. 25
Table 3.5 : Using Fourier descriptor magnitudes and angles. 25
Table 3.6 : Using Fourier descriptor magnitudes and the histogram of the radius. 25
Table 3.7 : Using Fourier descriptor magnitudes and the size of the lesion. 25
Table 3.8 : Comparison of all classification results. 26
Table 3.9 : Classification of lesions using first-order statistical features: Mean, Variance, Skewness, Kurtosis, Energy and Entropy. 26
Table 3.10 : Classification of lesions using first-order statistical features: Mean... 27
Table 3.11 : Classification of lesions using first-order statistical features: Variance. 27
Table 3.12 : Classification of lesions using first-order statistical features: Skewness. 27
Table 3.13 : Classification of lesions using first-order statistical features: Kurtosis. 27
Table 3.14 : Classification of lesions using first-order statistical features: Energy. 28
Table 3.15 : Classification of lesions using first-order statistical features: Entropy. 28
Table 3.16 : Comparison of all classification results. 28
Table 3.17 : Classification of lesions using second-order statistical features: Correlation, Contrast, Inverse difference, Entropy, Maximum probability. 28
Table 3.18 : Classification of lesions using second-order statistical features: Correlation. 29
Table 3.19 : Classification of lesions using second-order statistical features: Contrast. 29
Table 3.20 : Classification of lesions using second-order statistical features: Inverse difference. 29
Table 3.21 : Classification of lesions using second-order statistical features: Entropy. 29
Table 3.22 : Classification of lesions using second-order statistical features: Maximum probability. 30
Table 3.23 : Comparison of all second-order statistical the classification results... 30
Table 3.24 : Classification of lesions using haar wavelet transformation features.. 31
Table 3.25 : Classification of lesions using Daubechies 2 wavelet transformation features. 31

Table 3.26 : Classification of lesions using Daubechies 3 wavelet transformation features. 31
Table 3.27 : Classification of lesions using Daubechies 4 wavelet transformation features. 31
Table 3.28 : Classification of lesions using Daubechies 5 wavelet transformation features. 32
Table 3.29 : Comparison of all wavelet transformation classification results. 32
Table 3.30 : Classification of lesions after PCA dimension reduction. 32
Table 4.1 : Validation of ANN that is trained using Fourier descriptors magnitude and phase, the histogram of the radius of the lesion, and the size of the lesion. 35
Table 4.2 : Validation of ANN that is trained using Fourier descriptors magnitude. 35
Table 4.3 : Validation of ANN that is trained using Fourier descriptors phase. 35
Table 4.4 : Validation of ANN that is trained using Fourier descriptors magnitude and phase. 36
Table 4.5 : Validation of ANN that is trained using Fourier descriptors magnitude and histogram of lesion radius. 36
Table 4.6 : Validation of ANN that is trained using Fourier descriptors magnitude and lesion size. 36
Table 4.8 : Validation of ANN that is trained using first-order statistics: Mean, Variance, Skewness, Kurtosis, Energy, and Entropy. 37
Table 4.9 : Validation of ANN that is trained using first-order statistics: Mean... 37
Table 4.10 : Validation of ANN that is trained using first-order statistics: Variance. 37
Table 4.7 : Comparison of validation set results of ABCD rule motivated the classifiers (T: True, F: False, B: Benign, M: Malignant). 38
Table 4.11 : Validation of ANN that is trained using first-order statistics: Skewness. 39
Table 4.12 : Validation of ANN that is trained using first-order statistics: Kurtosis. 39
Table 4.13 : Validation of ANN that is trained using first-order statistics: Energy. 39
Table 4.14 : Validation of ANN that is trained using first-order statistics: Entropy. 39
Table 4.16 : Validation of ANN that is trained using second-order statistics: Correlation, Contrast, Inverse Difference, Entropy, and Maximum Probability. 40
Table 4.17 : Validation of ANN that is trained using second-order statistics: Correlation. 40
Table 4.18 : Validation of ANN that is trained using second-order statistics: Contrast. 40
Table 4.15 : Comparison of validation set results first-order features based classifiers (T: True, F: False, B: Benign, M: Malignant). 41
Table 4.19 : Validation of ANN that is trained using second-order statistics: Inverse Difference. 42
Table 4.20 : Validation of ANN that is trained using second-order statistics: Entropy. 42
Table 4.21 : Validation of ANN that is trained using second-order statistics: Maximum Probability. 42

Table 4.23 : Validation of ANN that is trained using haar transformation..... 43
Table 4.24 : Validation of ANN that is trained using Daubechies 2 transformation. 43
Table 4.25 : Validation of ANN that is trained using Daubechies 3 transformation. 43
Table 4.22 : Comparison of validation set results of second-order based classifiers (T: True, F: False, B: Benign, M: Malignant). 44
Table 4.26 : Validation of ANN that is trained using Daubechies 4 transformation. 45
Table 4.27 : Validation of ANN that is trained using Daubechies 5 transformation. 45
Table 4.28 : Comparison of validation set results of wavelet transformation based classifiers (T: True, F: False, B: Benign, M: Malignant). 46
Table 4.29 : Validation of ANN that is trained using PCA dimension reducted features. 47
Table 4.30 : Validation of ANN that is trained using boxplot filtered features..... 47

LIST OF FIGURES

Page

Figure 2.1 : Dermoscopic image and manually segmented image..... 8
Figure 2.2 : Examples of reconstructions from Fourier descriptors. M is the number of descriptors taken to reconstruct [**?**]. 10
Figure 2.3 : Example image 12
Figure 2.4 : Structure of co-occurrence matrix 12
Figure 2.5 : Co-occurrence matrix 13
Figure 2.6 : Hierarchical decomposition of an image [**?**] 14
Figure 2.7 : Artificial neurons [**?**]. 17
Figure 2.8 : General structure of the ANN [**?**]. 18
Figure 2.9 : Single connection in BPN [**?**]. 19
Figure 3.1 : An illustrative collection of images from PH^2 database, including common nevi (1st row), atypical nevi (2nd row) and melanomas (3rd row) 21
Figure 3.2 : WEKA classification tool main window 23
Figure 3.3 : WEKA classification tool main explorer window 23
Figure 3.4 : Manual segmentation of three melanocytic lesions..... 24
Figure 3.5 : Example of feature that have similar distribution for benign and malignant melanomas 33
Figure 3.6 : Example of feature that have different distribution for benign and malignant melanomas 33

CLASSIFICATION OF DERMOSCOPIC IMAGES USING NEURAL NETWORKS

SUMMARY

Melanoma is amongst one of the deadliest form of cancer. American Cancer Society reports that in 2016, about 76,380 new melanomas will be diagnosed and about 10,130 people are expected to die of melanoma in U.S. Even though melanoma constitutes less than 2% of skin cancer, this type of skin cancer represents the majority of deaths caused by skin cancer.

The most basic way to reduce mortality due to melanoma is early diagnosis. Dermatologists, use dermoscopy for scanning the skin and as a result of this scan melanomas can be detected. A gel is applied on the lesion, and depending on the lesion enlargement devices, the lesion is magnified 6-100 times. This magnified view is used to evaluate based on various medical diagnostic procedures, such as ABCD rule, seven points check-list, and Menzies method. ABCD rule allows assigning a score for each lesion. This score is calculated using four different characteristics (Asymmetry, Border, Colors, Differential Structures) of the lesion. Menzies method includes two different features: negative (symmetrical patterns, color) and positive (blue-white veil, atypical structures, etc.). The presence of positive features indicate melanoma. Finally, the lesions are scored by seven point check-list. However, in this method, only, lesions are checked for whether there is anyone of different structures (atypical pigment network, irregular streaks, etc.) and assigned scores 2 or 1 for each structure. If the total score is higher than 3 this may indicate that the lesion is melanoma.

These methods are still applied subjectively. Dermoscopy maximized rate of identification of the disease, however usually diagnosis of the disease is based on physician's experience. Since the determination is based on the human vision and experience, there is huge effort to determine of the disease using digital techniques. Dermascopic image analysis with computers can be used to prevent subjective decisions and to maximize the accuracy of melanoma detection.

In this thesis, new methods that are based on ABCD rules-motivated features, statistical features and textural features and used to classify dermoscopic lesion images are proposed and compared. Dermoscopic lesion images are classified using different features.

ABCD Rules-motivated features contain radius histogram of lesions, relational size of lesions and the magnitude and the phase of Fourier descriptors of the border of segmented lesions. All of these features are used to classify the lesion images together and separately.

Statistical features of texture contain first-order statistical features and second-order statistical features of the lesion images. The lesion images are divided 128x128 patches

and first-order and second-order features are extracted from these patches. First-order statistical features contain mean, variance, skewness, kurtosis, energy, and entropy. Second-order features contain correlation, contrast, inverse difference, entropy, and maximum probability. These statistical features are used to classify lesion images.

To extract textural features, we use wavelet transformation. In wavelet transformation features contains the percentages of energy corresponding to the horizontal, vertical, and diagonal details. These energy percentages are used as dermoscopic image features.

PH^2 dermoscopic images are used as database. This database contains 200 images in total. 40 of them are melanoma, 160 of them are not melanoma. 180 of them are used to train ANN with 10-fold cross validation. 20 of them are used as a test set. According to our result the best method to classify our dermoscopic image database is ABCD rules-motivated methods. Using ABCD Rules-motivated methods, sensitivity=72.22% and specificity=93.75% are accomplished.

DERMOSKOPİK GÖRÜNTÜLERİN SİNİR AĞLARI KULLANILARAK SINIFLANDIRILMASI

ÖZET

Kanser türleri arasında melanom en ölümcül olanlarından bir tanesidir. Amerikan Kanser Topluluğunun 2016 raporuna göre Birleşik Devletler'de yaklaşık 76380 kişide yeni vaka teşhis edileceği ve bunlardan yaklaşık 10130 kişinin bu sebepten dolayı hayatını kaybedeceği tahmin edilmektedir. Melanom cilt kanserlerinin %2'sinden daha azını oluştursa da, cilt kanserinden ölenlerin çoğuna melanom sebep olmaktadır.

Melanom kaynaklı ölümleri azaltmanın en temel yolu melanomu erken teşhis etmektir. Dermatologlar melanomların teşhisi için yaygın olarak dermoskopi isimli yöntemi kullanmaktadır. Bu yöntemde hastaya cerrahi bir müdahalede bulunulmaz. Lezyon üzerine bir jel sürülür ve büyütme cihazıyla, kullanılan cihaza bağlı olarak, lezyon 6-100 kat büyütülür. Bu büyütülmüş görüntüler üzerinde ABCD kuralı, Menzies yöntemi ve yedi nokta kontrol listesi gibi tanı yöntemleri kullanılarak cilt lezyonunu değerlendirmeye tabi tutulur. ABCD kuralı her lezyona bir skor atanmasını sağlar. Bu skor dört farklı özelliğin ayrı ayrı puanlandırılmasının (asimetri (Asymmetry), kenar (Border), renk (Colors), farklı yapılar (Differential Structures)) ardından skorların birleştirilmesinden oluşur. Menzies yöntemi negatif (simetrik desenler, tek renk) ve pozitif (mavi-beyaz peçe, atipikal yapılar vb.) olmak üzere iki farklı özelliğe dayanmaktadır. Pozitif özelliklerin varlığı melanoma işaret eder. Son olarak yedi nokta kontrol listesi de lezyonları puanlandırır. Fakat bu yöntemde sadece farklı yapıların olup olmadığına bakılır ve her yapıya 2 ya da 1 olarak puan atanır. Eğer toplam skor 3'ün üzerinde ise bu durum lezyonun melanom olduğuna dair önemli bir işarettir.

Bahsedilen yöntemler kullanılsa bile melanomların tespiti yine de sübjektif olarak yapılmaktadır. Çünkü tespit edilmesi doktorun görmesine ve tecrübesine dayanmaktadır. Bilgisayarlarla yapılacak dermoskopik görüntü analizi bu tür sübjektif durumların önüne geçmek için ve melanom tespitinin duyarlılığını artırmak için kullanılabilir.

Son yıllarda bu hastalığın bilgisayarlı teknolojilerle tespit edilmesi için büyük bir çaba gösterilmektedir. Bu tezde, dermoskopik görüntülerin farklı özelliklerini kullanan yeni yöntemler önerilmekte ve bu yöntemler karşılaştırılmaktadır. Dermoskopik lezyon görüntülerinin sınıflandırılmasında kullanılan yöntemler arasında ABCD kuralı tabanlı yöntem, istatiksel özelliklere dayanan yöntem ve doku özelliklerini kullanan yöntem bulunmaktadır. Dermoskopik görüntülerin özellikleri çıkarılarak bu yöntemlerle sınıflandırılmıştır. Bunların dışında dermoskopik görüntünün renk özellikleri ve desen özellikleri çıkartılarak kutu grafik yöntemiyle el yordamıyla bütün öznitelikler kontrol edilmiş ve bu özniteliklerden en ayırt edici olanlar seçilerek ayrıca lezyon görüntüleri sınıflandırılmıştır. Bu yöntemin de kısmen öznel bir tarafı bulunmasına rağmen yine de özniteliklerin ayırt ediciliğinden faydalandığı için nesnel tarafı da bulunmaktadır.

Bu öznitelikler ayrıca otomatik olarak da ayırt ediciliklerine karar verilebilir. Bu kullanılan kutu grafik yönteminin de sonuçların doğruluğunu artırdığı gösterilmiştir.

ABCD kuralı tabanlı özellikler, lezyon kenarlarının Fourier tanımlayıcılarının büyüklük ve fazlarını, lezyonun yarıçap histogramını ve lezyonun orantısal büyüklüğünü içermektedir. Bu özelliklerin hepsi bütün olarak ve ayrı ayrı test edilmiştir.

Dokunun istatiksel öznitelikleri lezyon görüntüsünün birinci-derece istatiksel özniteliklerini ve ikinci derece istatiksel özniteliklerini içermektedir. Lezyon görüntüleri 128x128 yamalar olarak bölünmüş ve birinci-derece ve ikinci-derece öznitelikleri çıkartılmıştır. Birinci-derece öznitelikleri, ortalama, varyans, kayıklık, kürtosis, enerji ve entropiyi içermektedir. İkinci-derece öznitelikler korelasyon, kontrast, ters fark, entropi ve maksimum olasılığı içermektedir.

Doku özniteliklerinin çıkarılması için dalgacık dönüşümünden yararlanılmıştır. Dalgacık dönüşümü öznitelikleri, 2-boyutlu dalgacık dönüşümü 4. seviyeye kadar uygulandıktan sonra, yatay, dikey ve köşegen detayların enerji yüzdelerinin çıkartılmasından oluşur. Bunlar dermoskopik görüntü özellikleri olarak kullanılmıştır.

Kutu grafik yönteminde ise lezyon görüntülerinin renk öznitelikleri ve desen özniteliklerinden yararlanılmıştır. Bu öznitelikler daha sonra kutu grafik yöntemiyle değerlendirmeye tabi tutulmuş ve sadece ayırt edici öznitelikler kullanılarak sınıflandırma yapılmıştır.

Sınıflandırma için çok katmanlı yapay sinir ağlarından(ÇKYSA) yararlanılmıştır. ÇKYSA üzerinde geri yayılımlı öğrenme algoritması kullanılmıştır. Gizli katmanın boyutu en iyi sonucu verecek şekilde belirlenmiştir.

Bu tezde, PH^2 dermoskopik görüntü veri tabanı kullanılmıştır. Bu veri tabanında toplamda 200 görüntü bulunmaktadır. Bunlardan 40 tanesi melanom ve 160 tanesi melanom değildir. Bu görüntülerin 180 tanesi 10-katlı çapraz doğrulama ile YSA'yı eğitmek için kullanılmıştır. 20 tanesi ise test kümesi olarak kullanıldı. Yapılan testler sonucunda en başarılı sonuçlar ABCD kuralı tabanlı sistemden alınmıştır. ABCD kuralı tabanlı sistemde başarı duyarlılık = %72.22'ye ve özgüllük = %93.75'e ulaşmıştır.

Kutu grafik metodu kullanılarak yapılan sınıflandırmada ise duyarlılık oranı %92.5 ve özgüllük oranı %87.5 olarak tespit edilmiştir. Bu yöntemde genel olarak doğruluk yükselmiş olsa da özgüllük oranın düştüğü görülmüştür.

Duyarlılık oranın düşük olmasının temel sebepleri arasında veri kümesinin boyutunun küçük olması yatmaktadır. Bu tür veri kümelerinde genellikle pozitif örneklerin frekansı düşük olmaktadır. Bu nedenle veri kümesinde negatif örneklere doğru kaçınılmaz bir meyil oluşmaktadır. Bu da özgüllük oranının duyarlılık oranına göre nispeten fazla olmasına neden olmaktadır.

ABCD kuralı tabanlı yöntemde kullanılan lezyon kenarlarının Fourier tanımlayıcılarının alınan lezyon görüntülerin bazı lezyonlarda lezyonun tamamını içermediği için tam olarak istenen sonucu vermediği söylenebilir. Bu nedenle lezyon görüntüleri alınırken lezyonun bütünlüğünün bozulmaması doğruluğu artırıcı önemli bir faktör olarak görülmektedir.

Bunun yanında bazı melanom cilt lezyonları iyi huylu lezyonlara fiziksel olarak ciddi oranda benzemektedir. Aynı şekilde bazı iyi huylu lezyonlar da melanom lezyonlara

benzemektedir. Bu nedenle bunları dermatolojinin şimdiye kadar ortaya koyduğu teorik alt yapı ile belirlenebilmesi mümkün olmayabilir. Dermatolojide geliştirilecek olan yeni yöntemlerin bu tür bilgisayar destekli sistemlerin doğruluğunu artıracağı düşünülebilir.

1. INTRODUCTION

Melanoma is the one of the most deadliest form of cancer. American Cancer Society reports that in 2016, about 76,380 new melanomas will be diagnosed and about 10,130 people are expected to die of melanoma in U.S [**?**]. Even though melanoma constitutes less than 2% of skin cancer, this type of skin cancer represents the majority of deaths caused by skin cancer.

The most basic way to prevent melanoma is early diagnose. Dermatologists, use dermoscopy to scan the skin for the diagnosis of melanomas. In this method, a gel is applied on the lesion, and depending on the lesion enlargement devices it is magnified by 6-100 times [**?**]. This magnified view allows identification of skin lesions using various medical diagnostic procedures, such as ABCD rules [**?**], Menzies method [**?**], and seven point checklist [**?**]. ABCD rule allows assigning a score for each lesion. This score is calculated using four different characteristics (Asymmetry, Border, Colours, Differential Structures) of the lesion. Menzies method includes two different features: negative (symmetrical patterns, color) and positive (blue-white veil, atypical structures, etc.). The presence of positive features may indicate melanoma. Finally, the lesions are scored by seven point check-list (atypical pigment network, irregular streaks, Irregular dots, blue - whitish veil, regression structures, irregular pigmentation, atypical vascular pattern). However, in this method, it is checked for whether there is any construction of different structures and assigned scores of 2 or 1. If the total score is higher than 3 it indicates that the lesion is melanoma.

Dermoscopy techniques has raised the diagnosis rate of melanoma between 10% and 27% [**?**]. But dermoscopic diagnosis rate increases only if dermatologists have received formal education [**?**]. Even though mentioned methods are used in the detection of melanoma, it is still a subjective process since the detection of melanoma is based on the human vision and experience. Dermoscopic image analysis with computer can be used to prevent such subjective decisions and to increase the sensitivity of the melanoma detection. In the near future, computerized dermoscopy

will become very important assistant to clinicians as a diagnostic tool and to follow up on dangerous skin lesions.

1.1 Purpose of Thesis

In this study, a novel computerized automated method is developed for detection of melanoma. There are many different features that are based on methods used by clinicians, such as ABCD rules. Also there are statistical and textural features that can be used to classify the lesion images.

Besides, there are many techniques to process these features to classify lesions, such as k-NN classifier, support vector machines (SVM), artificial neural network(ANN) etc. These classifiers have certain advantages and disadvantages with respect to each other. A multilayer perceptron (MLP) is a ANN classifier that can be used as a nonlinear classifier. It creates nonparametric models of data and can be used in solving many classification problems that involves nonlinear data.

Different features such as Fourier descriptors, symmetry, size of the lesion, statistical and textural features of the images are used to classify dermoscopic images.

ABCD rules-motivated features contain radius histogram of lesions, relational size of lesions, and magnitude and phase of Fourier descriptors of border of segmented lesions. All of these features are used together and separately to classify lesion images.

Statistical features of texture contain first-order statistical features and second-order statistical features of lesion images. Lesion images are divided 128x128 patches and first-order and second-order features are extracted from the patches. First-order statistical features contain mean, variance, skewness, kurtosis, energy, and entropy and second-order features contain correlation, contrast, inverse difference, entropy, and maximum probability.

Textural features are also important to describe and classify images. However there is no universal definition of the texture of an image and different definitions are given depending on the application [**?**]. But in general, texture is defined as the spatial variation of pixel intensities. Texture analysis is used for image classification frequently [**?**]. Texture analysis methods are convenient for classification because they provide exclusive information on where they are applied. In this thesis, in extracting

textural features wavelet transformation is used. Wavelet transformation features contains, after a two-dimensional wavelet decomposition of an image at level 4, the percentages of energy corresponding to the horizontal, vertical, and diagonal details. These percentages are used as dermoscopic image features.

Multilayer perceptron (MLP) is used to classify lesions and the results of the classification with different features are compared. Back-propagation algorithm is also used to train MLP.

1.2 Literature Review

In order to assist the clinical decision of dermatologists, computer aided diagnosis systems of dermoscopic images have been developed [**?**]. Since 1987, such systems can be used as assistant devices by inexperienced doctors to make basic evaluation [**?**]. These systems use different features of the lesions but in general they consist of two main groups that concerns the extraction of image features to discriminate them. One of groups tries to follow medical diagnostic procedures and the other one is based on statistical pattern recognition and machine learning methods. The systems, in the latter groups, try to learn discrimination factors that discriminate melanoma and benign lesions from statistical properties of lesions.

The systems that try to mimic the performance of clinicians by detecting and extracting different dermoscopic structures, such as pigment network [**?**, **?**, **?**], irregular streaks [**?**], granularities [**?**], regression structures [**?**], blue-white veil [**?**] and blotches [**?**]. These structures can be used to score lesions in a analogous method that is followed by clinicians.

In the literature, the large part of melanoma detection systems are based on a pattern recognition approach [**?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**]. These systems are usually motivated by ABCD rules [**?**]. The most common features used in these studies include shape features such as compactness, aspect ratio, and maximum diameter, which represent both asymmetry and border; color features in several color spaces such as mean and standard deviation and texture features such as gray-level cooccurrence matrix.

Promising results are obtained by many researchers with different extracted features and different classifiers. k-Nearest Neighbor (kNN) classifier is used to distinguish

between melanoma and benign nevi. With a large data set with more than 5300 dermoscopic images, Sensitivity (SE) of 87% and a Specificity (SP) of 92% are achieved [**?**]. With shape and color features and added texture features, using an artificial neural network classifier and choosing different probability thresholds, a SE = 96% and a SP = 93% are achieved in a data set with 217 melanomas and 588 images [**?**]. Support vector machine (SVM) based classifier is proposed to identify melanomas [**?**]. A huge feature vector, that contains shape, color, and texture features, are used and SE = 93% and SP = 92% are achieved. An internet-based melanoma screening system that uses a feature vector of over 400 features, divided in shape, texture, and color features SE = 86% and SP = 86% are achieved. Bag-of-features (BoF) approach that uses Gabor-like and wavelets achieve 82% on a data set that contains 100 images, 30 of them were melanomas [**?**].

1.3 Hypothesis

There are many different features that describe the lesion images. ABCD rules-motivated features such as symmetricity, border, size of the lesions are very important features to classify skin lesions as melanoma or benign. Besides that features there are many other local or global features exist. The most important other properties are texture features. Global texture features or local texture features of the lesion images can be extracted by using different techniques. Statistical approaches and transform methods are two of these approaches for texture analysis.

In this study, it is assumed that dermoscopic lesion images can be classified correctly by using different features of dermoscopic images. ABCD Rules-motivated features and textural features are used. As textural features, statistical information of textures such as first-order statistics, second order statistics, wavelet transformation of textures and color features are utilised.

Edge information can be used in matching or distinguishing the images. Many objects can be separated using edge information in images. For discrimination of melanoma and non-melanoma lesions, the edge information is very important factor which also can be used for the classification of melanomas. After segmentation, the edge information of the lesions are extracted. The image edges can be expressed by

using the pixel coordinates. However, it is generally preferred to use other methods such as Fourier descriptors except the pixel coordinates [**?**, **?**].

ABCD Rules motivated features contain radius histogram of lesions, relational size of lesions, and the magnitude and the phase of Fourier descriptors of the border of segmented lesions. All of these features are used together and separately to classify the lesion images.

Textural analysis techniques are important in classifying lesions. Statistical approaches use different statistics as features to represent lesion images in textural analysis. The first-order and second order statistics are mainly used as texture features.

In first-order statistical texture analysis, information on texture is extracted from the histogram of image intensity. The histogram contains frequency of a particular gray-level information and does not include correlations or co-occurrences between pixels. Histogram is also referred as pixel occurrence probability. The lesion images are divided into 128x128 patches and first-order and second-order features are extracted from these patches. First-order statistical features contain mean, variance, skewness, kurtosis, energy, and entropy.

Texture pairs that have matching second-order statistics cannot be seperated by human vision [**?**]. Second-order statistical texture analysis is based on the probability of finding a pair of grey levels at random distances over an entire image. In this technique, pixel cooccurrence matrices, generally named as grey-tone spatial dependencies matrices (GTSDM), are computed. Second-order features contain correlation, contrast, inverse difference, entropy, and maximum probability.

In texture analysis, another method is the transform methods. There are different transform methods shuch as Fourier transforms, Gabor transforms and wavelet transforms. Fourier transform, in practice, does not give good results because of the loss of spatial localisation. Wavelet transform gives better result compared with Gabor transform because of the varying the spatial resolution and wide range of wavelet functions [**?**].

In wavelet transformation features contains, after a two-dimensional wavelet decomposition of image at level 4, the percentages of energy corresponding to the

horizontal, vertical, and diagonal details. These percentages are used as dermoscopic image feature.

In addition to these methods, color features and statistical features of the images are combined and by using boxplot some of the features are eliminated. After that ANN is trained filtered out features.

In this thesis, PH^2 dermoscopic images database is used. This database contains 200 images in total. 40 of them are melanoma, 160 of them are not melanoma. We use 180 of them to train ANN with 10-fold cross validation. 20 of them are used as a test set.

2. FEATURE EXTRACTION

As mentioned before, there are many different feature extraction methods. Mainly, ABCD rules-motivated features and texture analysis methods are commonly used to extract features of the lesion images. Approaches to texture analysis are usually categorised into four methods: structural, statistical, model-based, and transform.

In structural approaches, unambiguous basic elements define texture and a order of spatial compositions of those basic elements. The elements and the placement orders are necessary to define textures. This approach is more suitable for synthesis than analysis. For natural textures the abstract definitions can be ambiguous because there is no fine distinction between micro and macro-structure [**?**].

Statistical methods represent the texture by the non-deterministic properties that manage the distributions and relevance between grey levels of an image. Second-order statistics obtain more advanced differentiation rates than the power spectrum [**?**]. Differentiation of two image textures highly rely on the distinction of the second-order statistics of textures [**?**]. In texture analysis, the most prominant second-order statistical features are extracted from the cooccurrence matrix [**?**].

Model based texture analysis attempt to interpret an image texture by use of generative image model and stochastic model, respectively. It is demonstrated that the fractal model is useful for modelling some natural textures. It can also be exploited for textural analysis and differentiation [**?**, **?**].

Fourier, Gabor and wavelet transforms are transform methods of texture analysis. These transforms describe an image in a space that has a coordinate system that its interpretation is strictly linked to the characteristics of a texture. Wavelet transforms have several advantages. Over the other two varying the spatial resolution enables it to define textures at the most appropriate scale and there is a huge range of choices for the wavelet function and best suited wavelets can be chosen for texture analysis in a specific application [**?**].

Figure 2.1 : Dermoscopic image and manually segmented image.

The orthonormal Daubechies basis performs best at invariant texture classification. The Haar orthonormal basis which is also from the same family of basis functions perform close to the Daubechies basis [**?**].

In this thesis, as a textural methods, statistical methods and transform based methods are focused. First-order and second order statistical properties and different wavelet transforms are given as input to the ANN classifier to train and test multilayer perceptron ANN.

2.1 ABCD Rule Motivated Features

Fourier descriptors, radius histogram distribution and the size of the lesion images are used as features. These features are combined to generate feature vectors that are shown below. For N segmented lesion image, Fourier descriptors, radius histogram and the size of the lesion are denoted by v_i, h_i, and r_i, respectively

$$n_i = (\mathbf{v_i}, \mathbf{h_i}, \mathbf{r_i}), \qquad \mathbf{i} = \mathbf{1}, ..., \mathbf{N} \tag{2.1}$$

Multilayer Perceptron Artificial Neural Networks (MLP-ANN) is used in the classification of the data set as melanoma and non-melanoma lesions. Obtained feature vectors are given to train the multilayer neural network. Then, trained neural network is tested using the test lesion images features.

2.1.1 Fourier descriptors extraction

An example from our dermoscopic image database and its segmented image are shown in Figure **??**. The edge of the manually segmented image is selected by expert dermatologists.

In Fourier theory, any function can be expressed by allowing the different frequencies of the sines and cosines. In computer vision, the images related to the spatial domain can be converted into the frequency domain by Fourier transformation [**?**].

Pixel locations of the lesion boundaries can be extracted using segmented image of the lesion. After a complete set of coordinates describing the boundary is extracted, each coordinate (x_k, y_k) can be taken as an imaginary number as follow (where x represents horizontal location, y represents vertical location):

$$s(k) = x_k + iy_k, \qquad k = 0, 1, 2, ..., N-1 \tag{2.2}$$

Even though the interpretation of the sequence is remoulded, the nature of the boundary itself is not changed. The advantage of this declaration is that it reduces a 2-D into a 1-D problem, i.e. you now have N complex numbers instead of $2 * N$ real numbers. The Discrete Fourier Transform of $s(k)$ is

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) e^{i2\pi uk/N} \tag{2.3}$$

for $u = 0, 1, 2, ..., N-1$. The complex coefficient $a(u)$ are called the Fourier descriptors of the boundary. Applying inverse Fourier Transform to $a(u)$ restores $s(k)$:

$$s(k) = \sum_{u=0}^{N-1} a(u) e^{i2\pi uk/N} \tag{2.4}$$

for $k = 0, 1, 2, ..., N-1$. Beginning value and the changed values are exactly the same. Taking all N values to change the original image is not required. There is very little contribution to the image. Hence the number of Fourier descriptors can be reduced. Expressing this as an equation,

$$\hat{s}(k) = \sum_{u=0}^{M-1} a(u) e^{i2\pi uk/N} \tag{2.5}$$

where $k = 0, 1, 2, ..., M-1$. This is equivalent to setting $a(u) = 0$ for all terms where $k > M-1$. If you use much more descriptors to rebuild the original image, i.e. the bigger M in (**??**), the result will be closer to the original image (See Figure **??**.).

Geometrical centroid of the shape is given by the descriptor $a(0)$. Examining only having the value of $a(0)$ as the Fourier descriptor, the original shape can still be rebuilt

Figure 2.2 : Examples of reconstructions from Fourier descriptors. M is the number of descriptors taken to reconstruct [**?**].

from it. The result is located in the center of its original shape. This shape is a circle and it is referred as the geometrical centroid. Actually, removing all but the first two Fourier descriptors always generate a circle. $F(u)$ represents $f(k)$ in the frequency space. $F(0)$ is the minimum frequency component which adds to the image the average brightness. Among Fourier descriptors, just $a(0)$ is the only element that depends on the location of the shape. The same image can be translated to another place on the plane without needing to recalculate Fourier descriptors. The only thing needed is to set $a(0)$. Furthermore, the Fourier descriptors are indifference to geometrical changes like rotation, scale and the choice of the beginning point. It was shown that the Fourier descriptors are not tighly indifference to these geometrical changes, but the changes are relevant to simple modifications on the descriptors. [**?**]

2.2 Statistical Features

The basic statistical feature of a texture is first-order statistics. The main advantage of first-order statistics is its ease to use it. In first-order statistics, only the histogram or pixel occurrence probability is measured.

Assume the image as a function $f(i,j)$. It has two space variables i and j. i can take values of $i = 0,1,....N-1$ and j can take values of $j = 0,1,....M-1$. The function $f(i,j)$ can take discrete values $0,1,...G-1$, where G is the total number of intensity

values. The histogram of an image is a function showing the number of pixels in the whole image for each intensity level.

$$h(i) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \delta(f(i,j), i) \tag{2.6}$$

where $\delta(i,j)$ is

$$\delta(i,j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{2.7}$$

By dividing the values $h(i)$ by the total number of pixels in the image one obtains the approximate probability density of occurrence of the intensity levels

$$p(i) = \frac{h(i)}{NM} \qquad i = 0, 1, 2, ..., G-1 \tag{2.8}$$

But this approach does not take spatial relationship and correlation between pixels into account. In general, the main features of histogram is used to characterize an image. These features are mean, variance, skewness, kurtosis, energy, and entropy. The equations are given below:

$$\text{Mean:} \qquad \mu = \sum_{i=0}^{G-1} ip(i) \tag{2.9}$$

$$\text{Variance:} \qquad \sigma^2 = \sum_{i=0}^{G-1} (i-\mu)^2 p(i) \tag{2.10}$$

$$\text{Skewness:} \qquad \mu_3 = \sigma^{-3} \sum_{i=0}^{G-1} (i-\mu)^3 p(i) \tag{2.11}$$

$$\text{Kurtosis:} \qquad \mu_4 = \sigma^{-4} \sum_{i=0}^{G-1} (i-\mu)^4 p(i) - 3 \tag{2.12}$$

$$\text{Energy:} \qquad E = \sum_{i=0}^{G-1} [p(i)]^2 \tag{2.13}$$

$$\text{Entropy:} \qquad H = -\sum_{i=0}^{G-1} p(i) log_2[p(i)] \tag{2.14}$$

Although, textures cannot be fully characterized by them. Features, the most important advantage is certainly their ease. Texture pairs that have matching second-order statistics cannot be separated by human vision [**?**]. Second-order statistical texture

0	2	2	2
0	3	2	2
1	2	2	2
3	0	3	0

Figure 2.3 : Example image

i\j	0	1	2	3
0	# of (0,0)	# of (0,1)	# of (0,2)	# of (0,3)
1	# of (1,0)	# of (1,1)	# of (1,2)	# of (1,3)
2	# of (2,0)	# of (2,1)	# of (2,2)	# of (2,3)
3	# of (3,0)	# of (3,1)	# of (3,2)	# of (3,3)

Figure 2.4 : Structure of co-occurrence matrix

analysis is based on the probability of finding a pair of grey levels at random distances over an entire image. In this technique, pixel cooccurrence matrices generally named as grey-tone spatial dependencies matrices (GTSDM) are computed $h_{d\theta}(i,j)$ [**?**]. If this matrix is divided by the total number of neighbouring pixels $R(d,\theta)$ in the image, it gives us the estimate of the joint probability, $p_{d\theta}(i,j)$, of two pixels, a distance d apart along a given direction θ having special pair values i and j.

There are two types of GTSDM, one of them is symmetric on which couple of two pixels are separated by d and $-d$ for the direction θ are counted, and the other only pairs that are separated by distance d are counted.

With a set of G discrete intensity values, when the image $f(x,y)$ is given, the GTSDM $h_{d\theta}(i,j)$ is defined as its (i,j)th entry is equal to the number of times that

$$f(x_1,y_1) = i \quad \text{and} \quad f(x_2,y_2) = j \tag{2.15}$$

where

$$(x_2,y_2) = (x_1,y_1) + (d\cos(\theta), d\sin(\theta)) \tag{2.16}$$

For each distance d and orientation θ, this results a matrix that has a dimension which is equal to the number of intensity levels in the image. Because of the computational reasons, usually only $d = 1$ and $d = 2$ pixels with angles $\theta = 0°, 45°, 90°$ and $135°$ are computed [**?**].

i\j	0	1	2	3
0	0	0	1	4
1	0	0	1	0
2	1	1	10	1
3	4	0	1	0

Figure 2.5 : Co-occurrence matrix

A number of features can be extracted by using cooccurrence matrix for the purpose of texture classification [**?**]. Haralick et. al. introduced 14 local features especially created for the same reason. Following equations are a few of them (where μ_x, μ_y and σ_x and σ_y stand for the mean and standard deviation of the row and column sums of the matrix, respectively.):

Angular second moment (energy),

$$f_1 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [p(i,j)]^2 \tag{2.17}$$

Correlation,

$$f_2 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{ijp(i,j) - \mu_x y_y}{\sigma_x \sigma_y} \tag{2.18}$$

Contrast,

$$f_3 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i-j)^2 p(i,j) \tag{2.19}$$

Absolute value,

$$f_4 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} |i-j| p(i,j) \tag{2.20}$$

Inverse difference,

$$f_5 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{p(i,j)}{1+(i-j)^2} \tag{2.21}$$

Entropy,

$$f_6 = -\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} p(i,j) \log_2 [p(i,j)] \tag{2.22}$$

Figure 2.6 : Hierarchical decomposition of an image [**?**]

Maximum probability,

$$f_7 = \max_{i,j} p(i,j) \tag{2.23}$$

2.3 Wavelet Features

All the main weaknesses in the methods are that the image is analyzed in a single scale. This limitation can be overcome by using multiscale representations [**?**]. In the last decades, wavelet theory has appeared and for multiscale image analysis, it has become formal model.

Basically, the wavelet transform removes the low resolution image and a collection of detailed images from an image. The reduced resolution image is produced by iteratively blurring the image. The hierarchical wavelet transform uses a family of wavelet functions and its associated scaling functions to separate the original signal into different subbands. To produce the next hierarchy level, decomposition operation is executed iteratively for the sub-band. Information lost during this operation is contained in the detailed images. The most often used features for texture categorization and segmentation processes are the energy and mean deviation [**?**]. The 2-D discrete wavelet transform is calculated by applying a distinguishable filter bank to the image [**?**]:

$$L_n(b_i,b_j) = [H_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i,b_j) \tag{2.24}$$

$$D_n1(b_i,b_j) = [H_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i,b_j) \tag{2.25}$$

$$D_n2(b_i,b_j) = [G_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i,b_j) \tag{2.26}$$

$$D_n3(b_i,b_j) = [G_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i,b_j) \tag{2.27}$$

where $*$ indicate the convolution operator, $_{\downarrow 2,1}(_{\downarrow 1,2})$ subsampling along the rows (columns) and L_0 is the original image. H and G are a low-pass and high-pass filters. L_n is obtained by low pass filtering and hence referred to as the low resolution image at scale n. The detail images D_{ni} are extracted by bandpass filtering in a particular direction and therefore at that scale it involves positional detail information n. Hence, the original image I is showed by a set of subimages at a number of different scales $L_d, D_{ni_{n=1,\ldots,d,i=1,2,3}}$ which is a multiscale illustration of depth d of the image I (Figure **??**). Initially the texture is separated with the discrete wavelet transform to extract wavelet texture features. Once the image is decomposed, each subband is characterized by features. The energy of a subimage D_{ni} containing N coefficients is defined as

$$E_{ni} = \frac{1}{N}\sum_{j,k}(D_{ni}(b_j,b_k))^2 \tag{2.28}$$

The variance of energy is expressed by the wavelet energy features $E_{ni_{n=1,\ldots,d,i=1,2,3}}$ through the frequency axis over scale and orientation and these features have been verified to be very powerful for texture identification. Because most related texture information has been eliminated by recursive low pass filtering, the energy of the low resolution image L_d is usually not taken into account as a texture feature.

The mean deviation is a different measure which is sometimes regarded as a texture feature:

$$MD_{ni} = \frac{1}{N}\sum_{j,k}|D_{ni}(b_j,b_k)| \tag{2.29}$$

Both **??** and **??** are measure of the dispersion of the wavelet coefficients, they are strongly correlated.

Because G is a high-pass filter, the mean of the wavelet detail coefficients equals zero. Therefore the energy is exactly their variance. Therefore, utilizing energy as a texture feature is correspond to defining the detail histogram by a Gaussian. Utilizing the histogram as a support for feature taking out has a meaning because this outputs rendering invariant features. From the previous discussion, it is straightforward that texture identification can be clarified by developing the model for the histogram, cause a more appropriate definition of the wavelet coefficients' first order statistics. This may e.g. be done by computing higher order moments from it. Nevertheless, if an empirical parametric model of the histogram was available, then all first order statistical information incorporate in the detail histogram could be captured in the parameters of this model.

2.4 Artificial Neural Networks

Inspired by the general neuron model calculations artificial neurons are created. Typical neurons receive signals through synapses located on the dendrites or membrane of the neuron. When the entering signals to a neuron are forceful enough (exceed a certain threshold), the neuron is triggered and give out a signal through the axon. This signal may be received by another neuron, and may trigger other neurons.

When modelling artificial neurons the complexity of natural neurons is highly abstracted. These primarily involves inputs, which are weighted by the weights and then computed by a mathematical function which decide the activation of the neuron. Another function computes the output of the artificial neuron (sometimes depended on a defined threshold) (Figure **??**).

If the weight of an artificial neuron is higher, the input which is weighted by it will be more powerful. Weights can also be negative, so it can be said that the signal is restrained by the negative weight. The calculation of the neuron will be different based on the weights. By tuning the weights of an artificial neuron outputs can be reached for particular inputs. When there is an ANN of numerous of neurons, it would be very difficult to find all the necessary weights. However, algorithms can be developed which

Figure 2.7 : Artificial neurons [**?**].

can tune the weights of the ANN in order to get the needed output from the network. This process of tuning the weights is referred as learning or training. Numerously different models considered as ANNs have been discovered from the first neural model. They vary in the functions, the accepted values, the topology, the learning algorithms, etc. Also many hybrid models where each neuron has more features. MLP-ANN which learns using backpropagation algorithm for learning is used in this thesis. [**?**].

The general structure of ANN is demonstrated in Figure **??**. ANN is controlled by adjusting and tuning the weights between nodes. Initially the weights are commonly adjust at some random numbers and then they are tuned during ANN training. All weights are changed coherently but some ANN are coping with thousands, moreover millions of nodes. Therefore adjusting one or two at time would not help in tuning ANN to get target results in time while the ANN training weights are updated after iterations. Then if results of ANN after the weight updates are better than previous set of weights, the new values of weights are kept and iteration continue. Minimizing the error should be basic objective when adjusting and tuning the weights.

Working the network consist of a forward pass and a backward pass. Outputs are calculated in the forward pass and compared with desired outputs. Error from target and actual output are calculated. In the backward pass this error is used to change the weights in the network in order to reduce the size of the error. Backward and forward pass are repeated as far as the error is low enough. When training ANN, network is fed with set of examples that have inputs and desired outputs. If there is a set of 500 samples, 50 of them could be used to train the network and 450 of them could be used test the model. Choosing the right learning rate and right momentum will support the weight tuning. If the learning rate is too small, adjusting right learning rate could be difficult and algorithm might take long time to converge. Otherwise, selecting

Figure 2.8 : General structure of the ANN [**?**].

huge learning rate could have counter effect, algorithm could diverge. In ANN every weight sometimes has its own learning rate. Momentum term describes inertia. Large values of momentum term will affect the adjustment in the current weight to move into same direction as previous adjustment. One of the most popular ANN algorithms is back propagation algorithm. The back propagation algorithm can be separated in the following four steps:

1. Feed-forward computation
2. Back propagation to the output layer
3. Back propagation to the hidden layer
4. Weight updates

A backpropagation network learns by example. Network algorithm examples are given to do something that is targeted and it adjust the network's weights. When training is terminated, it will give the required output for a particular input. Once the network is trained, it will give the target output for any of the input [**?**].

The algorithm works like this for single neuron in Figure **??**:

1. Apply the inputs to the network and find out the output.

Figure 2.9 : Single connection in BPN [**?**].

2. Next find out the error for neuron B. Error for neuron

$$E_B = Output_B(1 - Output_B)(Target_B - Output_B) \tag{2.30}$$

3. Change the weight. Let W_{AB}^{+} be the new (trained) weight and W_{AB} be the initial weight.

$$w_{AB}^{+} = w_{AB} + (E_B x Output_A) \tag{2.31}$$

4. Error of the hidden layer neurons are figured out. These cannot be directly computed, it is different than the output layer, therefore we backpropagate them from the output layer. This is performed by receiving the errors from the output neurons and sending them back through the network to compute the hidden layer errors.

$$E_A = Output_A(1 - Output_A)(E_B w_{AB} + E_C w_{AC}) \tag{2.32}$$

5. Having computed the error for the hidden layer neurons now continue like in step 3 to adjust the hidden layer weights. By iterating this method, a network of any number of layers can be trained.

The network continues training all the patterns over and over up to the total error decrease to some particular low targeted value and then it terminates. Once the network has been trained, it should be able to recognise not just the ideal patterns, but also noisy forms [**?**].

3. CLASSIFICATION OF LESIONS

The PH^2 database was built up through a joint research collaboration between the Universidade do Porto, Tecnico Lisboa, and the Dermatology service of Hospital Pedro Hispano in Matosinhos, Portugal. The dermoscopic images were obtained under the same conditions through Tuebinger Mole Analyzer system using a magnification of 20x. They are 8-bit RGB color images with a resolution of 768x560 pixels. Several experiments are conducted using dermoscopic image database (PH^2). The PH^2 database contains a total number of 200 melanocytic lesions, including 40 melanomas [**?**]

Figure 3.1 : An illustrative collection of images from PH^2 database, including common nevi (1st row), atypical nevi (2nd row) and melanomas (3rd row)

As illustrated in Figure **??**, the images of the database were selected taking their quality, resolution and dermoscopic features into account. Every image was evaluated by an expert dermatologist with regard to the following parameters:

- Manual segmentation of the skin lesion;

- Clinical and histological (when available) diagnosis;
- Dermoscopic criteria (Asymmetry; Colors; Pigment network; Dots/Globules; Streaks; Regression areas; Bluewhitish veil).

In this database, images have already been segmented by specialists in as reference. 40 melanoma and 160 non-melanoma dermoscopic image exist in the database. A number of different measurements are used to calculate the accuracy of the classification. The measurements that are used is shown below:

Table 3.1 : Contingency table.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	TP	FN	$TP+FN$
	Negative	FP	TN	$FP+TN$
	Total	$TP+FP$	$FN+TN$	

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{3.1}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{3.2}$$

$$\text{Positive predictive value} = \frac{TP}{TP+FP} \tag{3.3}$$

$$\text{Negative predictive value} = \frac{TN}{TN+FN} \tag{3.4}$$

$$\text{Doubt index} = \frac{TP+FP}{TP+FN} \tag{3.5}$$

$$\text{Diagnosis accuracy} = \frac{TP}{TP+FP+FN} \tag{3.6}$$

TP, FN, TN and FP represent the number of true positive, the number of false negatives, shows the number of true negatives and the number of false positives, respectively. Positive classifies a lesion as melanoma. Sensitivity and specificity, respectively, measures the percentage of correct classification for melanoma, and non-melanoma cases. Positive predictive value (PPV) is actually classified as a proportion of true classification melanoma lesions. Negative predictive value (NPV) is the opposite of the PPV. Doubts index indicates the awareness of the possibility that a lesion of melanoma. If the ratio is greater than 100% over-diagnosis, while less than 100% indicates that the least-diagnostics. The diagnostic accuracy indicates the accuracy loss due to incorrect classification.

Figure 3.2 : WEKA classification tool main window

Figure 3.3 : WEKA classification tool main explorer window

Back-propagation algorithm is used during the classification. The classification experiments is performed using WEKA classification tool (See Figure **??** and **??**).

3.1 ABCD Rules-Motivated Classification

Border of the lesions are extracted from segmented lesion images that are manually segmented by dermatologists. Segmented images are shown in Figure **??**. Fourier descriptors of the borders of the lesion, radius histogram distribution and size of the lesions of the data set that we use is proportionally combined to generate feature vectors that are shown in Eq. **??**. For N segmented lesion image, Fourier descriptors, radius histogram and size of the lesion are denoted by v_i, h_i, and r_i, respectively.

$$n_i = (\mathbf{v_i}, \mathbf{h_i}, \mathbf{r_i}), \qquad \mathbf{i = 1, ..., N} \tag{3.7}$$

Artificial neural networks(ANN) is used in the classification of the data set as melanoma and non-melanoma lesions. For this obtained feature vectors are given to the multi-layer neural network.

Figure 3.4 : Manual segmentation of three melanocytic lesions

Using Fourier descriptors magnitude and angle of border, histogram of the lesion radius and the size of the lesion the lesion classification results are shown in Table **??**

Table 3.2 : Using Fourier descriptors magnitude and phase of border, histogram of lesion radius and size of lesion.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	26	10	36
	Negative	9	135	144
	Total	35	145	180

Using the magnitudes of Fourier descriptors of border, the lesion classification results are shown in Table **??**

Table 3.3 : Using Fourier descriptors magnitudes.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	24	12	36
	Negative	6	138	144
	Total	35	165	180

Using the angles of Fourier descriptor of the border, the lesion classification results are shown in Table **??**

Table 3.4 : Using Fourier descriptor angles.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	14	22	36
	Negative	28	116	144
	Total	42	138	180

Using the magnitudes and the phases of Fourier descriptors of the border, the lesion classification results are shown in Table **??**

Table 3.5 : Using Fourier descriptor magnitudes and angles.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	24	12	36
	Negative	12	132	144
	Total	36	144	180

Using the magnitudes of Fourier descriptors of the border and the histogram of the radius the lesion classification results are shown in Table **??**

Table 3.6 : Using Fourier descriptor magnitudes and the histogram of the radius.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	23	13	36
	Negative	6	138	160
	Total	29	151	180

Using the magnitudes of Fourier descriptors of the border and the size of the lesion classification results are shown in Table **??**

Table 3.7 : Using Fourier descriptor magnitudes and the size of the lesion.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	24	12	36
	Negative	8	136	160
	Total	32	148	180

All classification results are shown in Table **??** (M: Magnitude of Fourier Descriptors, P: Phase of Fourier Descriptors, H: Histogram of Radius, S: Size of Lesion):

Table 3.8 : Comparison of all classification results.

Comparison of the results						
	SE	SP	PPV	NPV	DI	DA
M+P+H+S	**72.22**	93.75	74.29	93.10	0.97	0.58
M	66.67	**95.83**	80	92	0.83	0.57
P	38.89	80.56	33.33	84.06	1.17	0.22
M+P	66.67	91.67	66.67	91.67	1.00	0.50
M+H	63.89	**95.83**	79.31	91.39	0.81	0.55
M+S	66.67	94.44	75.00	91.89	0.89	0.55

3.2 Classification of Using Statistical Features

Statistical features of texture contain first-order statistical features and second-order statistical features of the lesion images. Lesion images are divided into 128x128 patches and first-order and second-order features are extracted from these patches. First-order statistical features contain mean, variance, skewnees, kurtosis, energy, and entropy. Second-order features contain correlation, contrast, inverse difference, entropy, and maximum probability.

In first-order statistical texture analysis, information on texture is extracted from the histogram of image intensity. The histogram contains frequency of a particular gray-level information and does not include correlations or co-occurrences between pixels. It also referred as pixel occurrence probability.

Texture pairs that have matching second-order statistics cannot be seperated by human vision. Second-order statistical texture analysis is based on the probability of finding a pair of grey levels at random distances over an entire image. In this technique, pixel cooccurrence matrices generally named as grey-tone spatial dependencies matrices (GTSDM) are computed.

A number of features can be extracted using co-occurrence matrix for the purpose of texture classification [**?**]. We used correlation, contrast, inverse difference, entropy and maximum probability features of cooccurrence matrix.

3.2.1 Classification of using first-order statistical features

Using first-order statistical features, mean, variance, skewness, kurtosis, energy and entropy, the classification results are shown in Table **??**.

Table 3.9 : Classification of lesions using first-order statistical features: Mean, Variance, Skewness, Kurtosis, Energy and Entropy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	24	12	36
	Negative	13	131	144
	Total	37	143	180

Using first-order statistical feature, mean, classification results are shown in Table **??**.

Table 3.10 : Classification of lesions using first-order statistical features: Mean.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	23	13	36
	Negative	14	130	144
	Total	37	143	180

Using first-order statistical feature, variance, the classification results are shown in Table **??**.

Table 3.11 : Classification of lesions using first-order statistical features: Variance.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	18	18	36
	Negative	15	129	144
	Total	33	147	180

Using first-order statistical feature, skewness, the classification results are shown in Table **??**.

Table 3.12 : Classification of lesions using first-order statistical features: Skewness.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	22	14	36
	Negative	12	132	144
	Total	34	146	180

Using first-order statistical feature, kurtosis, the classification results are shown in Table **??**.

Table 3.13 : Classification of lesions using first-order statistical features: Kurtosis.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	17	19	36
	Negative	15	129	144
	Total	33	147	180

Using first-order statistical feature, energy, the classification results are shown in Table **??**.

Table 3.14 : Classification of lesions using first-order statistical features: Energy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	17	19	36
	Negative	16	128	144
	Total	33	147	180

Using first-order statistical feature, entropy, the classification results are shown in Table **??**.

Table 3.15 : Classification of lesions using first-order statistical features: Entropy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	21	15	36
	Negative	14	130	144
	Total	35	145	180

All classification, based on first order features, results are shown in Table **??** (M: Mean, V: Variance, S: Skewness, E: Energy, Ent: Entropy).

Table 3.16 : Comparison of all classification results.

Comperison						
	SE	SP	PPV	NPV	DI	DA
M+V+S+K+E+Ent	**66.67**	90.97	64.86	91.61	1.03	0.49
M	63.89	90.28	62.16	90.91	1.03	0.46
V	50.00	89.58	54.55	87.76	0.92	0.35
S	61.11	**91.67**	64.71	90.41	0.94	0.46
K	47.22	89.58	53.13	87.16	0.89	0.33
E	47.22	88.89	51.52	87.07	0.92	0.33
Ent	58.33	90.28	60.00	89.66	0.97	0.42

3.2.2 Classification of using second-order statistical features

Using second-order statistical features, correlation, contrast, inverse difference, entropy, maximum probability, the classification results shown in Table **??**.

Table 3.17 : Classification of lesions using second-order statistical features: Correlation, Contrast, Inverse difference, Entropy, Maximum probability.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	16	20	36
	Negative	16	128	144
	Total	32	148	180

Using second-order statistical features, correlation, the classification results shown in Table **??**.

Table 3.18 : Classification of lesions using second-order statistical features: Correlation.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	17	19	36
	Negative	20	134	144
	Total	37	153	180

Using second-order statistical features, contrast, the classification results shown in Table **??**.

Table 3.19 : Classification of lesions using second-order statistical features: Contrast.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	10	26	36
	Negative	29	115	144
	Total	39	141	180

Using second-order statistical features, inverse difference, the classification results shown in Table **??**.

Table 3.20 : Classification of lesions using second-order statistical features: Inverse difference.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	9	27	36
	Negative	26	118	144
	Total	35	145	180

Using second-order statistical features, entropy, the classification results shown in Table **??**.

Table 3.21 : Classification of lesions using second-order statistical features: Entropy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	16	20	36
	Negative	16	128	144
	Total	32	148	180

Using second-order statistical features, maximum probability, the classification results shown in Table **??**.

Table 3.22 : Classification of lesions using second-order statistical features: Maximum probability.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	13	23	36
	Negative	20	124	144
	Total	33	147	180

All classification, based on second order features, results are shown in Table **??**: (Corr: Correlation, Cont: Contrast, Inv Diff: Inverse Difference, Ent: Entropy, Max Pro: Maximum Probability)

Table 3.23 : Comparison of all second-order statistical the classification results.

Comperison						
	SE	SP	PPV	NPV	DI	DA
Corr+Cont+Inv Diff+Ent+Max Pro	44.44	**88.89**	50.00	86.49	0.89	0.31
Corr	**47.22**	86.11	45.95	86.71	1.03	0.30
Cont	27.78	79.86	25.64	81.56	1.08	0.15
Inv Diff	25.00	81.94	25.71	81.38	0.97	0.15
Ent	44.44	**88.89**	50.00	86.49	0.89	0.31
Max Pro	36.11	86.11	39.39	84.35	0.92	0.23

3.2.3 Wavelet classification

The wavelet transform removes the low resolution image and provides a collection of detailed images from an image. The reduced resolution image is produced by blurring the image iteratively. The hierarchical wavelet transform uses a family of wavelet functions and its associated scaling functions to separate the original signal into different subbands. To produce the next hierarchy level, decomposition process is applied recursively for the sub-band. Information lost during this operation is contained in the detailed images. The most often used features for texture categorization and segmentation processes are the energy or mean deviation [**?**].

In this thesis, we use haar and daubechies 2 - 5 wavelets.

Using haar wavelet transformation features the classification results shown in Table **??**.

Table 3.24 : Classification of lesions using haar wavelet transformation features.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	23	13	36
	Negative	5	139	144
	Total	28	152	180

Using daubechies 2 wavelet transformation features the classification results shown in Table **??**.

Table 3.25 : Classification of lesions using Daubechies 2 wavelet transformation features.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	21	15	36
	Negative	9	135	144
	Total	30	150	180

Using daubechies 3 wavelet transformation features the classification results shown in Table **??**.

Table 3.26 : Classification of lesions using Daubechies 3 wavelet transformation features.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	23	13	36
	Negative	9	135	144
	Total	32	148	180

Using daubechies 4 wavelet transformation features the classification results shown in Table **??**.

Table 3.27 : Classification of lesions using Daubechies 4 wavelet transformation features.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	17	19	36
	Negative	9	135	144
	Total	26	154	180

Using daubechies 5 wavelet transformation features the classification results shown in Table **??**.

Table 3.28 : Classification of lesions using Daubechies 5 wavelet transformation features.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	22	14	36
	Negative	9	135	144
	Total	31	149	180

Comparison of all wavelet transformation the classification results in Table **??**.

Table 3.29 : Comparison of all wavelet transformation classification results.

Comparison						
	SE	SP	PPV	NPV	DI	DA
Haar	**63.89**	**96.53**	82.14	91.45	0.77	0.56
Db2	58.33	93.75	70	90	0.83	0.47
Db3	**63.89**	93.75	71.88	91.21	0.89	0.51
Db4	47.22	93.75	65.38	87.66	0.72	0.38
Db5	61.11	93.75	70.97	90.60	0.86	0.49

3.2.4 Classification after PCA reduction

After separate classification, ABCD rules-motivated features, first-order features, second order features and wavelet transform features are combined. But combination of these features gives us very large feature set. So, in order to reduce the dimension of features, we use Principal Component Analysis (PCA) method. PCA is mathematically defined as a linear transformation that transforms the data into a new coordinate system such that the greatest variance by some projection of the data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so on.

ABCD rules-motivated features that are fourier transform of borders combines, the histogram of the radius and the size of the lesions. First-order features are mean, variance, skewness, kurtosis, energy and entropy. Second-order features are correlation, contrast, inverse difference, entropy and maximum probability. After this combination we take 5 principal components of the data and classify it using ANN. Classification of lesions after PCA dimension reduction results are shown in Table **??**

Table 3.30 : Classification of lesions after PCA dimension reduction.

SE	SP	PPV	NPV	DI	DA
69.44	98.61	92.6	92.81	0.75	0.66

3.2.5 Classification using boxplot filtered features

RGB histogram, normalized RG histogram, HS histogram, HSV values and also, using k-means algorithm, color quantums of image are extracted. These values are used as features. After the extraction process, boxplot is used to get the best features and to reduce dimension of feature size.

Figure 3.5 : Example of feature that have similar distribution for benign and malignant melanomas

An example of feature that have similar distribution for benign and malignant melanomas is shown Figure **??** and example of feature that have different distribution for benign and malignant melanomas is shown Figure **??**. After the selection of features with different distribution, they are used to train neural net.

Figure 3.6 : Example of feature that have different distribution for benign and malignant melanomas

4. CONCLUSIONS AND RECOMMENDATIONS

The results of the validations of ANNs are shown next sections separately for different features.

4.1 ABCD Rules-Motivated Classification

Validation of ANN that is trained by using Fourier descriptors magnitude and phase, the histogram of the radius of the lesion, and the size of the lesion are shown in Table **??**.

Table 4.1 : Validation of ANN that is trained using Fourier descriptors magnitude and phase, the histogram of the radius of the lesion, and the size of the lesion.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	1	15	16
	Total	4	16	20

Validation of ANN that is trained using Fourier descriptors magnitude is shown in Table **??**.

Table 4.2 : Validation of ANN that is trained using Fourier descriptors magnitude.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	0	16	16
	Total	3	17	20

Validation of ANN that is trained using Fourier descriptors phase is shown in Table **??**.

Table 4.3 : Validation of ANN that is trained using Fourier descriptors phase.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	2	2	4
	Negative	1	15	16
	Total	3	17	20

Validation of ANN that is trained using Fourier descriptors magnitude and phase is shown in Table **??**.

Table 4.4 : Validation of ANN that is trained using Fourier descriptors magnitude and phase.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	1	15	16
	Total	4	16	20

Validation of ANN that is trained using Fourier descriptors magnitude and histogram of lesion radius is shown in Table **??**.

Table 4.5 : Validation of ANN that is trained using Fourier descriptors magnitude and histogram of lesion radius.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	2	2	4
	Negative	0	16	16
	Total	2	18	20

Validation of ANN that is trained using Fourier descriptors magnitude and lesion size is shown in Table **??**.

Table 4.6 : Validation of ANN that is trained using Fourier descriptors magnitude and lesion size.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	0	16	16
	Total	3	17	20

We use ABCD rules-motivated features together and separately to classify lesions. Validation of the classifiers shows that the best classification results are given by fourier descriptor magniteudes of borders of images. The classifier achieves the process results correctly for 19 of 20 lesions. Table **??** shows results for 8 different validation images. First 4 of them benign and last 4 of them malignant lesions.(1st: The classifier that uses the magnitude and the phase of Fourier descriptors, the histogram of the radius and the size of the lesions, 2nd: The classifier that uses the magnitude of Fourier descriptors and the histogram of the radius, 3rd: The classifier that uses the phase of Fourier

descriptors and the histogram of radius, 4th: The classifier that uses the magnitude and the phase of Fourier descriptors, 5th: The classifier that uses the magnitude Fourier descriptors and the histogram of the radius and 6th: The classifier that uses the magnitude of Fourier descriptors and the size of the lesions)

4.2 First-Order Statistics Based Classification

Validation of ANN that is trained using first-order statistics, mean, variance, skewness, kurtosis, energy, and entropy is shown in Table **??**.

Table 4.8 : Validation of ANN that is trained using first-order statistics: Mean, Variance, Skewness, Kurtosis, Energy, and Entropy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	2	2	4
	Negative	2	14	16
	Total	4	16	20

Validation of ANN that is trained using first-order statistics, mean, is shown in Table **??**.

Table 4.9 : Validation of ANN that is trained using first-order statistics: Mean.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	2	14	16
	Total	5	15	20

Validation of ANN that is trained using first-order statistics, variance, is shown in Table **??**.

Table 4.10 : Validation of ANN that is trained using first-order statistics: Variance.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	0	16	16
	Total	3	17	20

Validation of ANN that is trained using first-order statistics, skewness, is shown in Table **??**.

Image	True	1st	2nd	3rd	4th	5th	6th
	B	T	T	**F**	T	T	T
	B	T	T	T	T	T	T
	B	T	T	T	T	T	T
	B	**F**	T	T	T	T	T
	M	**F**	T	T	T	**F**	T
	M	T	T	T	T	T	T
	M	T	**F**	T	**F**	**F**	**F**
	M	T	T	**F**	T	T	T

Table 4.7 : Comparison of validation set results of ABCD rule motivated the classifiers (T: True, F: False, B: Benign, M: Malignant).

Table 4.11 : Validation of ANN that is trained using first-order statistics: Skewness.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	1	15	16
	Total	4	16	20

Validation of ANN that is trained using first-order statistics, kurtosis, is shown in Table **??**.

Table 4.12 : Validation of ANN that is trained using first-order statistics: Kurtosis.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	1	3	4
	Negative	0	16	16
	Total	1	19	20

Validation of ANN that is trained using first-order statistics, energy, is shown in Table **??**.

Table 4.13 : Validation of ANN that is trained using first-order statistics: Energy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	1	3	4
	Negative	1	15	16
	Total	2	18	20

Validation of ANN that is trained using first-order statistics, entropy, is shown in Table **??**.

Table 4.14 : Validation of ANN that is trained using first-order statistics: Entropy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	4	0	4
	Negative	1	15	16
	Total	2	18	20

First-order statistical features are used to classify lesions together and separately. Validation of the classifiers shows that the best classification results are given by entropy of images. The classifier achieves the process results correctly for 19 of 20 lesions.Table **??** shows results for 8 different validation images. First 4 of them are benign and last 4 of them are malignant lesions.(1st: The classifier that uses all of the

first-order features, mean, variance, skewness, kurtosis, energy, and entropy, 2nd: The classifier that uses all of the first-order feature mean, 3rd: The classifier that uses all of the first-order feature variance, 4th: The classifier that uses all of the first-order feature skewness, 5th: The classifier that uses all of the first-order feature kurtosis, 6th: The classifier that uses all of the first-order feature energy, 7th: The classifier that uses all of the first-order feature entropy)

4.3 Second-Order Statistics Based Classification

Validation of ANN that is trained using second-order statistics correlation, contrast, inverse difference, entropy, and maximum probability is shown in Table **??**.

Table 4.16 : Validation of ANN that is trained using second-order statistics: Correlation, Contrast, Inverse Difference, Entropy, and Maximum Probability.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	1	3	4
	Negative	0	16	16
	Total	1	19	20

Validation of ANN that is trained using second-order statistics, correlation, is shown in Table **??**.

Table 4.17 : Validation of ANN that is trained using second-order statistics: Correlation.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	1	3	4
	Negative	0	16	16
	Total	1	19	20

Validation of ANN that is trained using second-order statistics, contrast, is shown in Table **??**.

Table 4.18 : Validation of ANN that is trained using second-order statistics: Contrast.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	1	3	4
	Negative	5	11	16
	Total	6	14	20

Image	True	1st	2nd	3rd	4th	5th	6th	7th
	B	T	T	T	T	T	T	T
	B	T	T	T	T	T	**F**	T
	B	T	T	T	T	T	T	**F**
	B	T	T	T	T	T	T	T
	M	T	**F**	T	**F**	**F**	**F**	T
	M	**F**	**F**	**F**	T	**F**	**F**	T
	M	**F**	T	T	T	**F**	T	T
	M	T	T	T	T	T	**F**	T

Table 4.15 : Comparison of validation set results first-order features based classifiers (T: True, F: False, B: Benign, M: Malignant).

Validation of ANN that is trained using second-order statistics, inverse difference, is shown in Table **??**.

Table 4.19 : Validation of ANN that is trained using second-order statistics: Inverse Difference.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	3	13	16
	Total	6	14	20

Validation of ANN that is trained using second-order statistics, entropy, is shown in Table **??**.

Table 4.20 : Validation of ANN that is trained using second-order statistics: Entropy.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	2	2	4
	Negative	1	15	16
	Total	3	17	20

Validation of ANN that is trained using second-order statistics, maximum probability, is shown in Table **??**.

Table 4.21 : Validation of ANN that is trained using second-order statistics: Maximum Probability.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	0	4	4
	Negative	3	13	16
	Total	3	17	20

Second-order statistical features are used to classify lesions together and separately. Validation of the classifiers shows that the best classification results are given by combination of all of the features and only correlation feature of images. The classifiers achieve the process reuslts correctly for 19 of 20 lesions. Table **??** shows results for 8 different validation images. First 4 of them benign are and last 4 of them are malignant lesions.(1st: The classifier that uses second-order features, correlation, contrast, inverse difference, entropy, and maximum probability, of the lesion images, 2nd: The classifier that uses second-order features, correlation, of the lesion images, 3rd: The classifier that uses second-order features, contrast, of images, 4th: The

classifier that uses second-order features, inverse difference, lesin of images, 5th: The classifier that uses second-order features, entropy, of the lesion images and 6th: The classifier that uses second-order features, maximum probability, of the lesion images)

4.4 Wavelet Transformation

In wavelet transformation features contain, after a two-dimensional wavelet decomposition of an image at level 4, the percentages of energy corresponding to the horizontal, vertical, and diagonal details. These percentages are used as dermoscopic image feature.

Validation of ANN that is trained using haar transformation is shown Table **??**.

Table 4.23 : Validation of ANN that is trained using haar transformation.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	2	2	4
	Negative	0	16	16
	Total	2	18	20

Validation of ANN that is trained using Daubechies 2 transformation is shown Table **??**.

Table 4.24 : Validation of ANN that is trained using Daubechies 2 transformation.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	4	0	4
	Negative	0	16	16
	Total	4	16	20

Validation of ANN that is trained using Daubechies 3 transformation is shown Table **??**.

Table 4.25 : Validation of ANN that is trained using Daubechies 3 transformation.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	4	0	4
	Negative	3	13	16
	Total	7	13	20

Validation of ANN that is trained using Daubechies 4 transformation is shown Table **??**.

Image	True	1st	2nd	3rd	4th	5th	6th
	B	T	T	**F**	T	T	T
	B	T	T	T	T	T	**F**
	B	T	T	T	T	T	**F**
	B	T	T	T	T	T	T
	M	**F**	**F**	**F**	**F**	**F**	**F**
	M	T	**F**	**F**	**F**	T	**F**
	M	**F**	**F**	T	T	T	**F**
	M	**F**	T	**F**	**F**	**F**	**F**

Table 4.22 : Comparison of validation set results of second-order based classifiers (T: True, F: False, B: Benign, M: Malignant).

Table 4.26 : Validation of ANN that is trained using Daubechies 4 transformation.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	1	15	16
	Total	4	16	20

Validation of ANN that is trained using Daubechies 5 transformation is shown Table **??**.

Table 4.27 : Validation of ANN that is trained using Daubechies 5 transformation.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	3	1	4
	Negative	2	14	16
	Total	5	15	20

Five different wavelet transformation, Haar, Daubechies 1, Daubechies 2, Daubechies 3 and Daubechies 4, are used to classify lesions. As mentioned before, 180 lesion is used as training set. 20 of them used as validation set. 10-fold classification is used to train ANN. In training phase, maximum sensitivity and specificity, 63.89 and 96.53, respectively. Validation results also shown in Tables **??**, **??**, **??**, **??**, and **??**.

The best validation results are obtained using Daubechies 2 transformation.

Table **??** shows results for 8 different validation images. First 4 of them are benign and last 4 of them are malignant lesions.(1st: The classifier that uses wavelet transformation features, Haar, Daubechies 2, Daubechies 3, Daubechies 4, and Daubechies 5, of the lesion images, 2nd: The classifier that uses wavelet transformation features, Haar, of the lesion images, 3rd: The classifier that uses wavelet transformation features, Daubechies 3, of images, 4th: The classifier that uses wavelet transformation features, Daubechies 4, lesin of images, and 5th: The classifier that uses wavelet transformation features, Daubechies 5, of the lesion images)

4.5 Classification After PCA Reduction

After separate classification, ABCD rules-motivated features, first-order features, second order features and wavelet transform features combined. But combination of these features gives us very big feature set. So, in order to reduce dimension of

Image	True	1st	2nd	3rd	4th	5th
	B	T	T	T	T	T
	B	T	T	T	T	T
	B	T	T	T	T	T
	B	T	T	T	T	T
	M	T	T	T	T	T
	M	**F**	T	T	T	**F**
	M	**F**	T	**F**	T	T
	M	T	T	T	T	T

Table 4.28 : Comparison of validation set results of wavelet transformation based classifiers (T: True, F: False, B: Benign, M: Malignant).

features, we use Principal Component Analysis (PCA) method. PCA is mathematically defined as a linear transformation that transforms the data into a new coordinate system such that the greatest variance by some projection of the data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so on.

ABCD rules-motivated features which are fourier transform of the borders, the histogram of the radius and the size of the lesions, first-order features that are mean, variance, skewness, kurtosis, energy and entropy are combined with second-order features which are correlation, contrast, inverse difference, entropy and maximum probability. After that 5 principal components of data are taken and classified using MLP-ANN. Classification of lesions after PCA dimension reduction results are shown in Table **??**.

Table 4.29 : Validation of ANN that is trained using PCA dimension reducted features.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	25	11	36
	Negative	2	142	144
	Total	27	153	180

4.6 Classification Using Boxplot Filtered Features

Color features and statistical features are described in section 3 are filtered out using boxplots. Then, they are used to train neural net. That is shown, classification results are improved using boxplot filtering. Classification of lesions after boxplot filtering results are shown in Table **??**

Table 4.30 : Validation of ANN that is trained using boxplot filtered features.

		Prediction outcome		
		Positive	Negative	Total
Actual value	Positive	35	12	47
	Negative	5	148	153
	Total	40	160	200

4.7 Discussion

In this thesis, problem of classification of skin lesions is studied using different features of the lesion images. Prominent results are displayed by some of the methods. The results presented above indicate that ABCD rule-motivated classification gives

the best specificity results but it gives lower sensitivity results than boxplot filtered classification.

It is shown that classification using only small feature set is not enough. Different features of lesion images such as color features, statistical features, wavelet features must be used together to increase the rate of accuracy. Because descriptor of border of lesion is used in ABCD rule-motivated classification, lesion images must contain lesion entirely. Partiality in lesion images reduce rate of accuracy in ABCD rule-motivated method. Another problem in classification of lesion images is that there is not enough sample of lesions. To get better results, larger data sets are needed. Moreover, in general, frequency of malignant lesion is lower than benign lesion. Therefore, there is bias in training sets.

Some of the malignant lesions shows very similar properties with benign lesions. In dermatology, feature works may put theoretical differences these malignant melanomas from benign lesions so, these differences make them more distinguishable from benign lesions.

4.8 Conclusion

According to results, different features gives different results. Boxplot filtered features increase sensitivity and specificity of classification. Without boxplot filtering the best classification method is ABCD motivated method that has sensitivity= 72.22% and specificity = 95.83%. But in validation set the neural network that is trained using Daubcehies 2 is shown 100% success. Using boxplot the results are increased to sensitivity= 92.5% and specificity = 87.5%.

REFERENCES

[1] **Yoo, Y.**, (2001), Tutorial on Fourier Theory.

[2] **Manian, V. and Vasquez, R.** (1998). Scaled and rotated texture classification using a class of basis functions, *Pattern Recognition*, *31*(2), 1937–1948.

[3] **Ripley, B.D.** (2008). *Pattern Recognition and Neural Networks*, Cambridge Universtiy Press, Cambridge.

[4] (2016). Cancer facts and figures 2016, **Technical Report**, American Cancer Society, American Cancer Society, Atlanta, GA, USA.

[5] (2015), Dermoscopy Tutorial, `http://www.dermoscopy.org/atlas/base.htm`.

[6] **Stolz, W.**, **Riemann, A. and Cognetta, A.B.** (1994). ABCD rule of dermatoscopy: A new practical method for early recognition of malignant melanoma, *Eur. J. Dermatol*, *4*(7), 521–527.

[7] **Menzies, S.**, **Ingvar, C.**, **Crotty, K. and McCarthy, W.** (1996). Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features, *Arch. Dermatol*, *132*(10), 1178–1182.

[8] **Argenziano, G.**, **Fabbrocini, G.**, **Carli, P.**, **De Giorgi, V.**, **Sammarco, E. and Delfino, M.** (1998). Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis, *Arch. Dermatol*, *134*(12), 163–1570.

[9] **Mayer, J.** (1997). Systematic review of the diagnostic accuracy of dermatoscopy in detectingmalignant melanoma, *Med. J. Aust*, *167*(4), 206–210.

[10] **Binder, M.**, **Puespoeck-Schwartz, M.**, **Steiner, A.**, **Kittler, H.**, **Wolff, K.**, **Muellner, M. and Pehamberger, H.** (1997). Epiluminescence microscopy of small pigmented skin lesions: Short-term formal training improves the diagnosis performance of dermatologists, *J. Amer. Acad. Dermatol*, *36*(2), 197–202.

[11] **Tuceryan, M. and Jain, A.K.**, (1998). Texture analysis, World Scientific Publishing, 2 edition.

[12] (2000). *Texture analysis in machine vision*, World Scientific Publishing, Singapore.

[13] **Cascinelli, N.**, **Ferrario, M.**, **Tonelli, T. and Leo, E.** (1987). A possible new tool for clinical diagnosis of melanoma: the computer, *Journal of the American Academy of Dermatology*, *16*(2), 361–367.

[14] **Korotkov, K. and Garcia, R.** (2012). Computerized analysis of pigmented skin lessions: A review, *Artificial Intelligence in Medicine*, *56*, 69–90.

[15] **Di Leo, G.**, **Paolillo, A.**, **Sommella, P. and Liguori, C.** (2008). An improved procedure for the automatic detection of dermoscopic structures in digital elm images of skin lessions, *Proc. 2008 IEEE Comput. Soc. VECIMS*, pp.190–194.

[16] **Sadeghi, M.**, **Razmara, M.**, **Lee, T.K. and Atkins, M.S.** (2011). A novel method for detection of pigment network in dermoscopic images using graphs, *Comput. Med. Imaging Graph*, *35*(2), 137–143.

[17] **Barata, C.**, **Marques, J. and Rozeira, J.** (2011). Detecting the pigment network in dermoscopy images: A directional approach, *Proc. 33rd IEEE EMBS Annu. Int. Conf.*, pp.5120–5123.

[18] **Fabbrocini, G.**, **Betta, G.**, **Di Leo, G.**, **Liguori, C.**, **Paolillo, A.**, **Pietrosanto, A.**, **Sommella, P.**, **Rescigno, O.**, **Cacciapuoti, S.**, **Pastore, F.**, **Vita, V.**, **Mordente, I. and Ayala, F.** (2010). Epilluminescence image processing for melanocytic skin lesion diagnosis based on 7-point check list: A preliminary discussion on three parameters, *Open Dermatol. J.*, *4*, 110–115.

[19] **Stoecker, W.**, **Wronkiewicz, M.**, **Chowdhury, R.**, **Stanley, R.J.**, **Xu, J.**, **Bangert, A.**, **Shrestha, B.**, **Calcara, D.A.**, **Rabinovitz, H.S.**, **Oliviero, M.**, **Ahmed, F.**, **Perry, L.A. and Drugge, R.** (2011). Detection of granularity in dermoscopy images of malignant melanoma using color and texture features, *Comput. Med. Imaging Graph*, *35*(2), 144–147.

[20] **Celebi, M.E.**, **Iyatomi, H.**, **Stoecker, W.**, **Moss, R.**, **Rabinovitz, H.**, **Argenziano, G. and Soyer, H.** (2008). Automatic detection of blue-white veil and related structures in dermoscopy image, *Comput. Med. Imaging Graph*, *32*(8), 670–677.

[21] **Stoecker, W.**, **Gupta, K.**, **Stanley, R.**, **Moss, R. and Shrestha, B.** (2005). Detection of asymmetric blotches in dermoscopy images of malignant melanomas using relative color, *Skin Res. Technol.*, *11*(3), 179–184.

[22] **Ganster, H.**, **Pinz, A.**, **Wildling, E.**, **Binder, M. and Kittler, H.** (2001). Automated melanoma recognition, *IEEE Trans. Med. Imaging*, *20*(3), 233–239.

[23] **Elbaum, M.** (2002). Computer-aided melanoma diagnosis, *Dermatol. Clin.*, *20*(4), 735–747.

[24] **Rubegni, P.**, **Cevenini, G.**, **Burroni, M.**, **Perotti, R.**, **Dell'Eva, G.**, **Sbano, P. and Miracco, C.** (2002). Automated diagnosis of pigment skin lesions, *Int. J. Cancer*, *101*(6), 576–580.

[25] **Blum, A.**, **Luedtke, H.**, **Ellwanger, U.**, **Schwabe, R.**, **Rassner, G. and Garbe, C.** (2004). Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions, *Brit. J. Dermatol*, *151*, 1029–1038.

[26] **Burroni, M.**, **Sbano, P.**, **Cevenini, G.**, **Risulo, M.**, **Dell'eva, G.**, **Barbini, P.**, **Miracco, C.**, **Fimiani, L.**, **Andreassi, M. and Rubegni, P.** (2005). Dysplastic naevus versus in situ melanoma: Digital dermoscopy analysis, *Brit. J. Dermatol.*, *152*(4), 679–684.

[27] **Seidenari, S.**, **Pellacani, G. and Grana, C.** (2005). Pigment distribution in melanocytic lesion images: A digital parameter to be employed for computer-aided diagnosis, *Skin Res. Technol.*, *11*(4), 236–241.

[28] **Celebi, M.E.**, **Kingravi, H.A.**, **Uddin, B.**, **Iyatomi, H.**, **Aslandogan, Y.**, **Stoecker, W. and Moss, R.** (2007). A methodological approach to the classification of dermoscopy images, *Computerized Medical Imaging and Graphics*, *31*(6), 362–373.

[29] **Stanley, R.**, **Stoecker, W. and Moss, R.** (2007). A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images, *Skin Res. Technol.*, *13*(1), 67–72.

[30] **Iyatomi, H.**, **Oka, H.**, **Celebi, M.E.**, **Hashimoto, M.**, **Hagiwara, M.**, **Tanaka, M. and Ogawa, K.** (2008). An improved Internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm, *Comput. Med. Imaging Graph.*, *32*(7), 566–579.

[31] **Situ, N.**, **Yuan, X.**, **Chen, G. and Zouridakis, J.** (2008). Malignant melanoma detection by bag-of-features classification, *Proc. 30th IEEE EMBS Annu. Int. Conf*, pp.3110–3113.

[32] (1996). *Fundamentals of Electronic Image Processing*, IEEE Press.

[33] **Zahn, C.T. and Roskies, R.Z.** (1972). Fourier descriptors for plane closed curves, *IEEE Trans. Computers*, *21*, 269–281.

[34] **Julesz, B.** (1975). Experiments in the Visual Perception of Texture, *Scientific American*.

[35] **Materka, A. and Strzelecki, M.** (1998). Texture Analysis Methods – A Review, **Technical Report**, Technical University of Lodz, Institute of Electronics, Brussels.

[36] **Weskza, J.**, **Deya, C. and Rosenfald, A.** (1976). A Comparative Study of Texture Measures for Terrain Classification, *IEEE Trans. System, Man and Cybernetics*.

[37] **Haralick, R.** (1979). Statistical and Structural Approaches to Texture, *Proc. IEEE*, *67*(5), 786–804.

[38] **Pentland, A.** (1984). Fractal-Based Description of Natural Scenes, *IEEE Trans. Pattern Analysis and Machine Intelligence*, *6*(6), 661–674.

[39] **Chaudhuri, B. and Sarkar, N.** (1995). Texture Segmentation Using Fractal Dimension, *IEEE Trans. Pattern Analysis and Machine Intelligence*, *17*(1), 72–77.

[40] **de Wouwer, G.V.**, **Scheunders, P. and Dyck, D.V.** (1999). Statistical texture characterization from discrete wavelet representations, *IEEE Trans. Image Proc.*

[41] **Gershenson, C.**, (2001). Artificial Neural Networks for Beginners.

[42] **Cilimkovic, M.** Neural Networks and Back Propagation Algorithm.

[43] **MacLeod, C.** An Introduction to Practical Neural Networks and Genetic Algorithms For Engineers and Scientists.

[44] **Mendonça, T.**, **Ferreira, P.M.**, **Marques, J.**, **Marcal, A.R.S. and Rozeira, J.** (2013). PH^2 - A dermoscopic image database for research and benchmarking, *35th International Conference of the IEEE Engineering in Medicine and Biology Society*, pp.5437–5440.

CURRICULUM VITAE

Name Surname: Enes Albay

Place and Date of Birth: Bursa, 1988

Adress: ITU, Faculty of Computer and Informatics Engineering, Ayazaga, Istanbul

E-Mail: albay@itu.edu.tr

B.Sc.: Istanbul Technical University, Department of Computer Engineering

M.Sc.: Istanbul Technical University, Department of Computer Engineering

PUBLICATIONS/PRESENTATIONS ON THE THESIS

- **Albay E.**, Kamasak M., 2015: Skin Lesion Classification Using Fourier Descriptors of Lesion Borders. *TıpTekno'15*, October 15-18, 2015 Bodrum, Turkey.
- **Albay E.**, Kamasak M., 2016: Improved Classification of Skin Lesions Using Shape and Color Features. *SIU 2016*, May 16-19, 2016 Zonguldak, Turkey.