Egyptian Computer Science Journal (ISSN: 1110 – 2586)
Volume 40 – Issue 03, September 2016

Malignant Melanoma Detection Based on Machine Learning Techniques: A Survey

[1]Munya A. Arasi, [2]El-Sayed A. El-Dahshan, [3]El-Sayed M. El-Horbaty, [4]Abdel-Badeeh M. Salem
[1,3,4]Dept. of Computer Science, Faculty of Computer and Information Sciences
Ain Shams University, Abbassia, Cairo, Egypt
[2]Egyptian E-Learning University (EELU), , Eldoki, Giza, Egypt
munya_arasi@yahoo.com,seldahshan@eelu.edu.eg, shorbaty@cis.asu.edu.eg,absalem@cis.asu.edu.eg

Abstract

Skin cancer is one of the most growing types and dangerous cancer in the world; the important of these cancers are malignant melanoma. The early diagnosis of malignant melanoma is a critical issue for dermatologists. In this paper, we present an overview of recent the state of the art in Computer-aided detection/diagnosis (CAD) systems in identifying and diagnosing malignant melanoma of dermoscopy images and describe its steps starting with image acquisition, preprocessing; and finishing with malignant melanoma classification of dermoscopic images. The comparative study shows that the most common methods for features extraction are the Discreet Wavelet Transform (DWT) and the method which combines both texture and color features resulting in output of very high accuracy. The methods for the classification:*K*-Nearest Neighbor, Artificial Neural Networks, and Support Vector Machines are very well in the range **[%90 –% 97, 5].**

Keywords: *Skin cancer, Malignant melanoma, Machine learning, Medical Knowledge-Based systems, Medical imaging, Medical informatics.*

1. Introduction

Dermatology is the branch of medicine that concerned with the diseases' diagnosis of skin, hair and nails, the skin is the most important part in the human body which protects the internal parts from the outside world. The skin cancer is the most important of these diseases; it can be growing at any part in the body and occurring from non-pigmented cells [1].Skin cancer is one of the most growing types and the most dangerous in the world of cancer; the important of these tumors is malignant melanoma, the rates of melanoma have been rising for at least 30 years.

The main risk of melanoma is could be spread entire the body by lymphatic vessels and blood vessels [2], thus the main strategy is the early diagnosis of melanoma and removal of thin melanoma; it is the most common cancer if diagnosed at an early stage can be cured without complications. Therefore, the early diagnosis of cancer malignant melanoma is a critical issue and the main challenges for dermatologists to reduce mortality and morbidity [3]. It is very difficult the diagnosis of melanoma by naked eye using the features that recognize a benign from malignant melanoma when using the clinical practice.

Nowadays the computer technology in medical decision support is used widespread and pervasive across a wide range of medical area, such as cancer research, dermatology[4], thus

the computerized methods analysis of dermoscopy images allows overcome various issues which help the dermatologist to take less time and high efficiency in diagnosis of skin cancer. It is important to develop of various computerized methods for clinical diagnosis using computer aided diagnosis (CAD) systems, which these systems give to dermatologist different analysis steps such as border detection, features extraction and diagnostic by using classification methods [5]. Studies have shown that CAD systems for melanoma diagnosis is still complex problem, the problem of lighting effects the resulting from the reflection surface of the skin, this affecting on all steps of diagnostic [6,7]in addition there is difficulties for information contents of dermoscopy images which they are not clear due to the noises and presence the hair. Therefore, it has become the extraction of the features a complex issues, hence there are many researchers are interested to the development of systems for automated diagnosis of malignancy in skin lesions.

The rest of the paper is organized as follows: the generic methodology of Malignant Melanoma diagnosis "CAD" scheme is presented in section 2. In section 3the review of CAD systems for Malignant Melanoma Detection is presented. Section 4 provides the discussion of image acquisition and preprocessing; and finishing with skin lesions classification. Finally, conclusions and future work are given in section 5.

2. Generic methodology of Malignant Melanoma diagnosis (CAD) scheme

The goal of CAD is to automatically decide Type of lesion, i.e. melanoma or benign by examining various features of lesion and examining these features by using machine learning techniques that help the dermatologist to take less time and high efficiency in diagnosis of skin cancer and enhancement the diagnostic accuracy of physicians and reduce the overall rate of misdiagnosis [8]. The system [9] uses image processing techniques for improving and segmentation the images for detection the melanoma in early stages without the need for biopsy.

The main advantages of CAD systems that the features extracted from lesion based on computer provide high accuracy than the features extraction based on dermoscopy algorithms. The general scheme of a CAD system for the skin lesions is shown in figure 1. It consists of four main components steps: Image acquisition and preprocessing, feature extraction, classification and evaluation.

The Basic techniques in CAD systems for Malignant Melanoma Detection in previous studies that published during 2001–2015 are highlighted in figure 2, which presents clearly all the techniques that used in this field. The inputs to the computer aided system are dermoscopy images; in the first phase preprocessing of image is done that allows reducing the noise effects and various artifacts like hair that may be present in the dermoscopic images and improve the image quality such as image cropping[10], gradient operation, morphological operation, scaling color space transformation [11, 12], color quantization, contrast enhancement [6,7];[13,14], Filters [3],[15, 16, 17, 18,19,20,21], it is the most important step for a successful feature extraction and diagnosis.

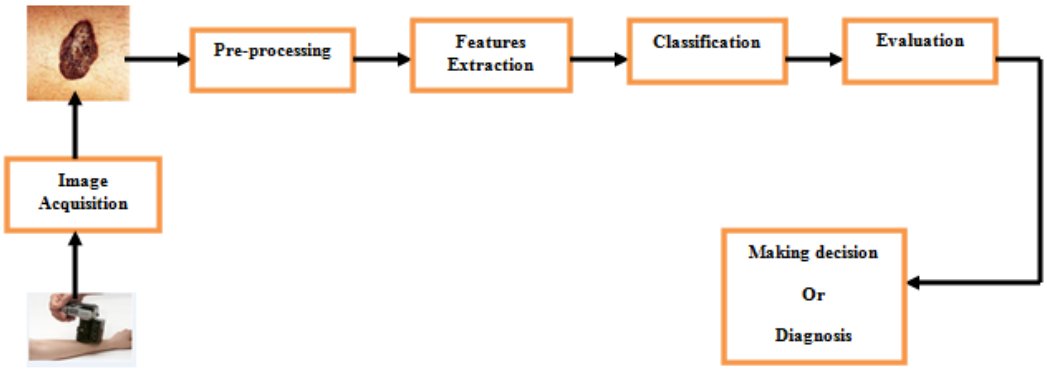

Figure 1.General scheme of a CAD system for malignant melanoma diagnosis

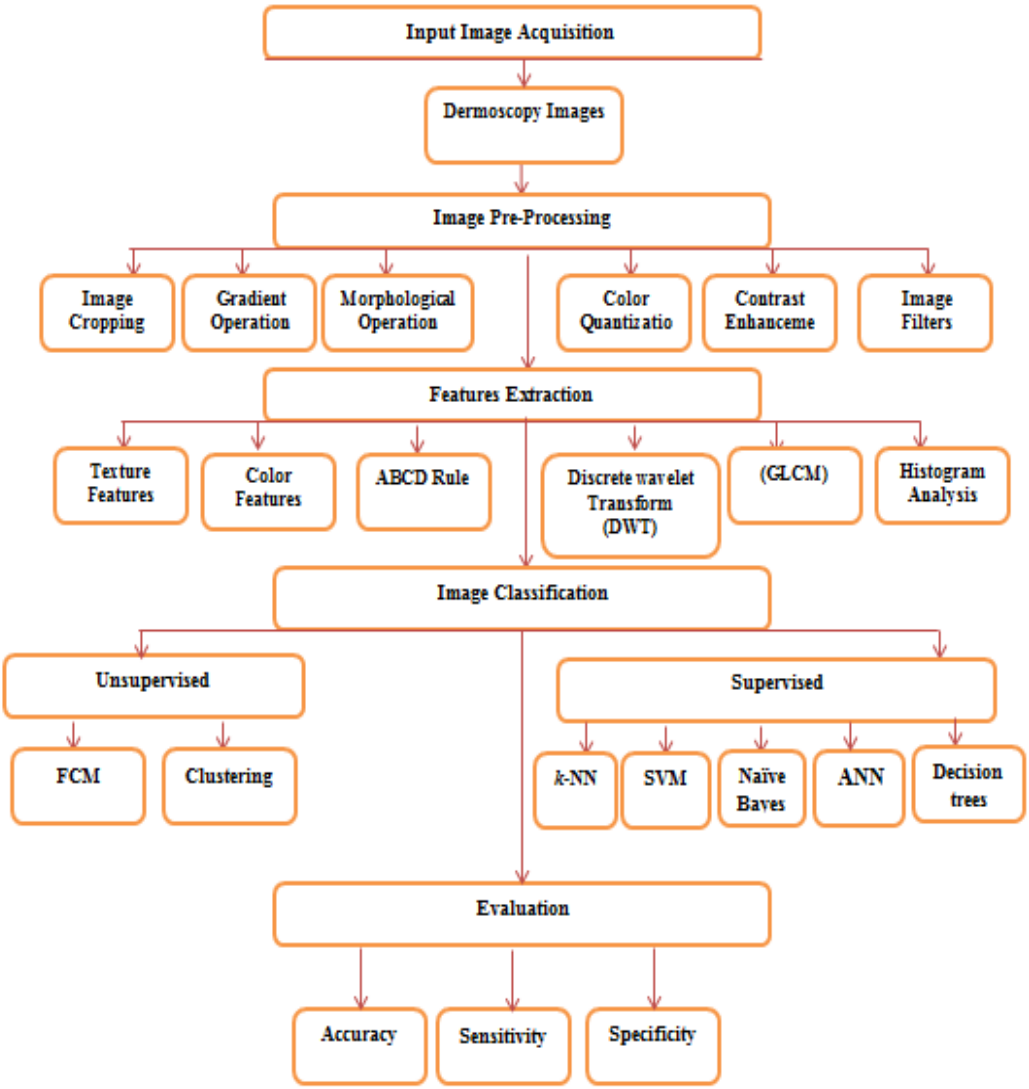

Figure2.Various Techniques in Malignant Melanoma diagnosis

After preprocessing the image, the features are extracted for classification of lesion. Feature extraction focuses on analysis the properties of skin lesions for a single skin lesion, there are many techniques, e.g. the textural features [14],[17],[19],[22, 23, 24], shape features [16], [18], color features [16, 17], [20],[22], [25, 26], [18, 19], [24]Asymmetry (A), Boundary (B), Color (C), and Diameter of a lesion (D) (ABCD rule)[6,7], [14], DWT features [15] [27, 28, 29], Gray Level Co-occurrence Matrix(GLCM)[7], [12], [21], histogram analysis [30], and High-level intuitive features (HLIF) [31].After the feature extraction step for extraction the attributes of skin lesions, it is important to classify the attributes of these regions for proper recognition of melanoma. The aim of the classification step is to distinguish benignity and malignancy by making use of extracted features. This step uses machine learning algorithms to reach a decision[32], the most recent published classification techniques, e.g.supervised techniques such as k-NN [10] ,[12] support vector machine[7],[13, 14];[17]; [31],[35], Naïve Bayes (NB) [3], [20], Artificial Neural Networks [3],[12],[15, 16, 17],[[27],[33],Multilayer Perceptron (MLP) [12],[16], Logistic Model Tree (LMT) [28], Hidden Naive Bayes (HNB) [28] and Decision Tree [11], [24 25],and unsupervised classification techniques such as Clustering [27] and fuzzy C-means. Also, hybrid intelligent systems using soft computing techniques are used for diagnosis of melanoma in dermoscopy images like Neuro-fuzzy [20]. It gives high accuracy **%91.26**. The evaluation step measures the performance of classification techniques for the Malignant Melanoma Detection based on the Parameters like Accuracy (AC), Sensitivity (SE), and Specificity (SP).

3. Review of CAD systems for Malignant Melanoma Detection

Most studies revolve around analysis of skin lesion images taken using dermatoscope, table1 gives the summary of many different techniques are used for developing Computer-aided diagnosis for Malignant Melanoma, generally CAD systems are included these stages: image acquisition, image pre-processing, feature extraction, classification and Diagnostic Evaluation.

Many existing techniques have been employed in recent years for the prediction of skin cancer at initial stage (see table 1). It gives an overview the classification and image processing techniques that published during 2001–2015, which used in the field for diagnosis Malignant Melanoma. The results obtained from these methods are used to recognize the patterns which are aiming to help the doctors for classifying the malignant and benign cases. Hence, the Malignant Melanoma diagnostic problems are basically in the scope of the classification problems.

Table1. Classification performance for different studies

Author,year	Dataset	Preprocessing	Features Extraction	Classifier	Diagnostic Evaluation
(Elbaum et al., **2001**)**[13]**	from 4 Clinical Centers,	NA	Statistical Features	Linear or nonlinear Classifiers	**SE** =95% **SP** = 70%
(Dreiseitl et al., **2001**)**[11]**	Dermatology, University of Vienna Medical School.	Gray level tresholding to	Global and Local Features	*k*-NN Log Regression ANN Decision Trees SVM	**AC**=%0.933 **AC**=%0.967 **AC**=%0.968 **AC**=%0.885 **AC**=%0.970
(Rubegni et al., **2002**)**[17]**	Dermatology of Siena University	Filtering, Detection of Borders	Geometry, Color, Texture	ANN	**AC**=%93
(Stanley et al., **2003**)**[30]**	A data set of 258 clinical images.	NA	Color Histogram Analysis	A fuzzy logic	Higher Accuracy.
(Barzegari et al., **2005**)**[33]**	Images of pigmented skin lesions	NA	Software of Visio med AG	ANN	**SE**=%83 **SP**=%96
(Yuan et al., **2006**)**[23]**	from University of Texas	Resizing, Cropping, Hair Removal	Texture Feature	SVM	**AC**=%.70
(M. E. Celebi, **2006**)**[25]**	Interactive Atlas of Dermoscopy	Second order B-spline Function	Color Feature	Decision Tree	Less time
(Grammatiko poulos, **2006**)**[14]**	NA	Histogram Equalization , Prewitt Method.	ABCD rule	TDS Calculation	NA
(Stanley, Stoecker, & Moss, **2007**)**[26]**	EDRA interactive Atlas	Histogram Quantization	Color Feature	Thresholds Discriminate	**AC**=%.87.7
(M. Emre Celebi et al., **2007**)**[18]**	Atlas of Dermoscopy	Median Filter	Shape, Color and Texture Features	SVM	**SP**=%.92.34 **SE**=%.93.33
(M. Emre Celebi et al., **2008**)**[24]**	Atlas of Dermoscopy	Manual Border Determination	Color Features and Texture Features	Decision Tree	**SE**= %69.35 **SP** =%89.97
(M Emre Celebi et al., **2009**)**[19]**	public dermoscopy image set	Gaussian Filter(GF), Median Filter(MF)	Color and Texture Features	Euclidean Distance Transform	Higher Diagnostic Accuracy.
(Garnavi et al**., 2010)[28]**	Interactive Atlas of Dermoscopy	the lesion borders were obtained manually	Wavelet-based Texture Analysis	SVM (RBF kernel), LMT HNB	**AC**=%88.24 by LMT

Follow Table1. Classification performance for different studies

Author,year	Dataset	Preprocessing	Features Extraction	Classifier	Diagnostic Evaluation
(Salah et al., **2011**)**[20]**	NA	Blure filter	Area and Color Features	Neuro-fuzzy	**AC=**%91.26
(Sheha et al., **2012**)**[12]**	102 dermoscopy Atlases	Resizing and Color space Transformation	GLCM	MLP	**AC** =%92
(KumarJain & Jain, **2012**)**[27]**	From different sources	Image contour Tracing Algorithm	DWT	Clustering, PNN	**AC**=%0.92 **AC** =%0.95
(Ramteke & Jain, **2013**)**[29]**	from available digital camera	Watershed method	DWT	Fuzzy System	**AC** =%90.
(Elgamal, **2013**)**[15]**	from a digital camera with dermoscope	Gaussian, Median Filter	PCA,DWT	FP-ANN, k-NN	**AC** =%95 **AC** =% 97.5
(Scharcanski et al., **2014**)**[31]**	Derm**IS**and DermQuest	Illumination Correction Algorithm	HLIFs	Linear soft-Margin SVM	High Accuracy
(Science & Engineering, **2014**)**[34]**	Health care Centers in Kottayam	NA	21 Medical Attributes	NB	High Accuracy
(Li et al., **2014**)**[3]**	spectroscopic system	Median filter and	Statistical Variables	ANN and NB	**AC** =%88.4 **AC** =%89.2
(Gajbar & Deshpande, **2015**)**[21]**	National Cancer Institute.	Median Filter, FCM	GLCM.	SVM	High Accuracy
(Jaiswar, Kadri, & Gatty, **2015**)**[6]**	Using the technique Dermoscope	Histogram Equalization Histogram Thresholding	ABCD rule	TDS Calculation	High Reliable and Robust.
(Mengistu, **2015**)**[35]**	Dermquest, Dermnet	Median Filtering	GLCM and Color features	SOM and RBF	**AC** =%96.15
(Amarathunga , Ellawala, Abeysekara, (**2015**)**[16]**	Using the technique Dermoscope	Median Filtering	Color and shape features	MLP J48	**AC** =%85 **AC** =%85
(Kaur, **2015**)**[7]**	NA	Histogram Equalization	ABCD rule and GLCM	Boosting Algorithm	Good Accuracy
(Immagulate & Vijaya, **2015**)**[22]**	Dermnet Dermofit.	Image resizes	Color and Texture Features	SVM, ASVM and PSVM	**AC** =%86 **AC** =%92 **AC** =%93

Not Available (NA)

4. Discussion

As it is mentioned in Table 1, the step of image acquisition, dermascopy images are collected from different sources for decoding the characteristics of human skin. The preprocessing and feature extraction techniques of the automatic system are proposed in previous works must be improved in order to become an effective tool and robust in the diagnosis of skin lesions.

The classification methods that have been used in computer-aided diagnosis of melanoma in this study involve *K* -Nearest Neighbour, Decision Tree, Support Vector Machine, Artificial Neural Network (ANN), Neuro-Fuzzy, Fuzzy C-Mean (FCM), Naïve Bayes and Clustering. For comparison purpose, the best three classifiers with excellent results like (k-NN, artificial neural networks, and support vector machines) When implementing support vector machine and its variants such as proximal support vector machine (PSVM), it based classification model yields a better performance and effectual when compared to other models, the predictive accuracy of PSVM is **93% [22]**.

Several studies have shown that the diagnostic accuracy of artificial neural networks is improved. From the results the hybrid techniques are robust and effective, it gives a higher accuracy is **91.26%** which achieved by using of Area Features and Color Features and Neuro-fuzzy classifier. That means the results of hybrid classifiers are obtained very well, due to the integration of their performance and combining their advantages. It was also showed that the GLCM and Color feature were used together the classification accuracy was increased. The best classification accuracy **96.15%** for Melanoma was obtained using combining SOM and RBF **[35]**. The single layer perceptron (SLP) type of artificial neural network was designed to estimate the probability of melanoma risk, which obtained a maximum accuracy in distinguishing melanoma from benign lesions is **93% [17].** The Multilayer Perceptron (MLP) is a feed forward network, capable of generating nonlinear boundaries. The accuracy **100%** and **92%** for training and testing respectively, in this study shows that combination between co-occurrence matrix and ANN is a promising technique for discrimination between malignant melanoma and melanocytic nevi dermoscopy images **[12]**.

Probabilistic Neural Network (PNN) is a feed forward neural network, which was derived from Bayesian network and the training consists essentially of incorporating the training cases into the pattern layer. The classification using Probabilistic Neural Network is better than other types of artificial neural networks (ANNs) proposed in the same domain and have shown excellent classification performance **95%** as compared to clustering Classifier **[27]**. It has got the same accuracy **95%** by combination between discrete wavelet transforms DWT and back-propagation (BP) algorithm whereas the combination between discrete wavelet transforms DWT and k-nearest neighbor algorithm has **97.5%** of accuracy for the same domain **[15]**.

5. Conclusions and Future Work

The most common methods for the features extractions are Discrete Wavelet Transforms (DWT), the combination between texture feature and color feature providing very high accuracy. The results show that the methods for the classification k-NN, Artificial Neural Networks, and Support Vector Machines are very well in the range **[%90 – % 97, 5]**.

Using Naïve Bayesian Classification is effective model for diagnosis Malignant Melanoma, but the decision tree algorithm is not well suited for this domain. The classification using Probabilistic Neural Network is better than the Clustering classifier. This is a significant improvement as compared to the earlier techniques proposed in the same domain. The PNN, back-propagation (BP), and the combining SOM and RBF classifiers show excellent classification performance of Artificial Neural Networks (ANNs).

In future we are looking for constructing a better framework for development a CAD system automatically accurate diagnosis of melanoma in dermoscopy images with extract the statistical and texture features from 2Dwavelet transform. We hope that will get the valuable information and overcome the drawbacks and enabling well delivery of medical fields.

References

[1]. K.Korotkov andR. Garcia, "Computerized analysis of pigmented skin lesions: A review", Artificial Intelligence in Medicine, 56 (2),2012,pp. 69–90.

[2]. J.Scharcanski,M. E. Celebi and S. Service, "Online. Computer Vision Techniques for the Diagnosis of Skin Cancer",2014,pp. 193–219.

[3]. L. Li, Q. Zhang, Y. Ding, H. Jiang, B. H. Thiers andJ. Z. Wang, "Automatic diagnosis of melanoma using machine learning methods on a spectroscopic system", BMC Medical Imaging, 14 (1),2014, 36

[4]. E. Barati, M. Saraee, A. Mohammadi, N. Adibi andM.R. Ahamadzadeh,"A Survey on Utilization of Data Mining Approaches for Dermatological (Skin) Diseases Prediction", Journal of Selected Areas in Health Informatics, 2 (3),2011,pp. 1–11.

[5]. W.-Y. Chang, A. Huang, C.-Y. Yang, C.-H. Lee, Y.-C. Chen,T.-Y. Wu, andG.-S. Chen, "Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study", PloS One, 8 (11), 2013, e76212.

[6]. S. Jaiswar, M. Kadri andV. Gatty, "Skin Cancer Detection Using Digital Image Processing", 3 (6), ,2015.

[7]. K. Kaur, "A Collaborative Biomedical Image-Mining Framework along with Image Annotation", 116 (13), 2015, pp. 25–28.

[8]. L.S. Goggin, R. H. Eikelboom and M. D. Atlas"Clinical decision support systems and computer-aided diagnosis in otology", Otolaryngology--Head and Neck Surgery, 136(4 suppl), 2007, pp.s21-s26.

[9]. D. F. Nikhil Cheerla, "Automatic Melanoma Detection Using MultiStage Neural Networks", International Journal of Innovative Research in Science, Engineering and Technology, 3 (2), 2014, pp. 9164–9183.

[10]. S. KhakAbi,P. Wighton, T. K. Lee andM. S. Atkins, "Multilevel feature extraction for skin lesion segmentation in dermoscopic images" , SPIE Medical Imaging, 2012, pp. 83150E–83150E.

[11]. S. Dreiseitl, L. Ohno-Machado, H. Kittler, S. Vinterbo, H. Billhardt andM. Binder, "A comparison of machine learning methods for the diagnosis of pigmented skin lesions", Journal of Biomedical Informatics, 34 (1), 2001, pp.28–36.

[12]. M. Sheha, M. Mabrouk andA. Sharawy,"Automatic detection of melanoma skin cancer

using texture analysis", International Journal of Computer Applications, 42 (20), 2012, pp. 22-26.

[13]. M. Elbaum, A. W. Kopf, H. S.Rabinovitz, R. G. Langley, H. Kamino, M. C. Mihm, S. Wang," Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: A feasibility study" Journal of the American Academy of Dermatology, 44 (2), 2001, pp. 207–218.

[14]. G. Grammatikopoulos, "Automated malignant melanoma detection using MATLAB", Proc. 5th Int. Conf. on Data Networks, Communications and Computers,2006, pp.91–94.

[15]. M. Elgamal, "Automatic Skin Cancer Images Classification", International Journal of Advanced Computer Science and applications, 4(3), 2013.

[16]. A. A. L. C. Amarathunga, E. P. W. C. Ellawala, G. N. Abeysekara andC. R. J. Amalraj, "Expert System For Diagnosis Of Skin Diseases", International Journal of Scientific & Technology Research, 4 (01), 2015, pp. 174–178.

[17]. P. Rubegni, M. Burroni, G. Cevenini, R. Perotti, G. Dell'Eva, P. Barbini, M. Fimiani, and L. Andreassi,"Digital dermoscopy analysis and artificial neural network for the differentiation of clinically atypical pigmented skin lesions: A retrospective study", Journal of Investigative Dermatology, 119 (2), 2002, pp.471–474.

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

[19]. M. E. Celebi, H. Iyatomi, G. Schaefer, and W. V. Stoecker, "Lesion border detection in dermoscopy images", Computerized Medical Imaging and Graphics : The Official Journal of the Computerized Medical Imaging Society, 33 (2), 2009, pp. 148–153.

[20]. B. Salah, M. Alshraideh, R. Beidasand F. Hayajneh, "Skin cancer recognition by using a neuro-fuzzy system", Cancer Informatics, 10, 2011, pp. 1–11.

[21]. A. M. Gajbar and P. A. S. Deshpande, "Detection and Analysis of Skin Cancer in Skin Lesions by using Segmentation", 5 (4), 2015, pp. 1173–1178.

[22]. I. Immagulate and M. S. Vijaya,"Categorization of Non-Melanoma Skin Lesion Diseases Using Support Vector Machine and Its Variants", 3 (2), 2015, pp. 34–40.

[23]. X. Yuan, Z. Yang, G. Zouridakis and N. Mullani, "SVM-based texture classification and application to early melanoma detection", Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings, 2006, pp.4775–4778.

[24]. M. E. Celebi, H. Iyatomi, W. V. Stoecker, R. H. Moss, H. S. Rabinovitz, G. Argenziano, and H. P. Soyer, "Automatic detection of blue-white veil and related structures in dermoscopy images", Computerized Medical Imaging and Graphics, 32 (8), 2008, pp. 670–677.

[25]. M. E. Celebi, "Detection of blue-white veil areas in dermoscopy images using machine learning techniques", Proceedings of SPIE, 2006, pp.61445T–61445T.

[26]. R. J. Stanley, W. V. Stoecker and R. H. Moss, "A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images", Skin Research and Technology, 13 (1), 2007, pp.62–72.

[27]. Y. KumarJain and M. Jain, "Comparison between Different Classification Methods with

Application to Skin Cancer", International Journal of Computer Applications, 53 (11), 2012, pp.18–24.

[28]. R. Garnavi, M. Aldeen and J. Bailey, "Classification of melanoma lesions using wavelet-based texture analysis", Proceedings - 2010 Digital Image Computing: Techniques and Applications, DICTA 2010, pp.75–81.

[29]. N. S. Ramteke and S. V. Jain, "ABCD rule based automatic computer-aided skin cancer detection using MATLAB", ®, 4(August), 2013, pp.691–697.

[30]. R. J. Stanley, R. H. Moss, W. Van Stoecker and C. Aggawal, "A fuzzy-based histogram analysis technique for skin lesion discrimination in dermatology clinical images", Computerized Medical Imaging and Graphics, 27 (5),2003, pp. 387–396.

[31]. J. Scharcanski, M. E. Celebi and S. Service, "Computer Vision Techniques for the Diagnosis of Skin Cancer", 2014, pp.193–219.

[32]. S. N. Deepa and B. Aruna Devi, "A survey on artificial intelligence approaches for medical image classification", Indian Journal of Science and Technology, 4 (11), 2011, pp.1583–1595.

[33]. M. Barzegari, H. Ghaninezhad, P. Mansoori, A. Taheri, Z. S. Naraghi and M. Asgari, "Computer-aided dermoscopy for diagnosis of melanoma", BMC Dermatology,5(1), 2005.

[34]. C. Cience and S. Engineering, "Prediction of Different Dermatological Conditions Using Naïve Bayesian Classification", International Journal of Advanced Research in Computer Science and Software Engineering, 4 (1), 2014, pp.864–868.

[35]. A. D. Mengistu, "Computer Vision for Skin Cancer Diagnosis and Recognition using RBF and SOM", (9), 2015, pp. 311–319.