DETECTING IN-SITU MELANOMA USING MULTI PARAMETER EXTRACTION AND NEURAL CLASSIFICATION MECHANISMS

Ruksar Fatima¹, Dr. Mohammed Zafar Ali Khan², Dr. A. Govardhan³ and Kashyap Dhruve⁴

¹(Head Dept. of Bio-Medical Engg, KBN College of Engg, Gulbarga, India, Ruksar_blue@yahoo.co.in)
²(Associate Professor, IIIT Hyderabad, Hyderabad, India, zafar@iith.ac.in)
³(Director of Evaluation of JNTU, Hyderabad, India, dejntuh@jntuh.ac.in)
⁴(Technical Director, Planet-i Technologies, Bangalore, India, kashyapdhruve@hotmail.com)

ABSTRACT

Computer aided diagnostics systems are widely used for diagnosis of skin lesions. This paper discusses a Multi Parameter Extraction and Classification System (MPECS) to aimed to enable early detection of skin cancer melanoma. The MPECS adopts a supervised machine learning algorithm for classification. The dermoscopic images are represented by extensive parameter sets extracted using a six phase approach. The paper discusses the operation of the MPECS in the training and testing phase. The adoption of the Multilayer Feed Forward Neural Network (MFNN) classifier is justified, its proficiency in accurately diagnosing and classifying early signs of skin cancer melanoma in dermoscopic images of skin lesions is proved through the experimental results discussed in the manuscript.

Keywords: Skin Cancer, Melanoma, Early Detection, In Situ Melanoma, MPECS, Multi Parameter Extraction, Classification, Computer Aided Diagnostics, Image Processing, Skin Lesions, Feature Extraction, Multilayer Feed Forward, Neural Network, Back Propagation, SVM Classifier, ROC, Receiver Operating Characteristics

1. INTRODUCTION

Computer vision based diagnosis systems which are non-invasive have experienced a wide acceptability ratio in the recent years. Deaths arising from skin cancers have increased tremendously in the past decade [1] [2][3]. Skin Cancer Melanoma is dangerous and accounts for maximum number of deaths arising due to skin cancers [3][4]. A recent survey [5] conducted in 2012 for the United States of America alone, states that more than 75% of the deaths arising from skin cancers are due to melanotic skin lesions. Skin cancer melanomas
are generally diagnosed by adopting biopsy procedures recommended by dermatologists. Studies have proved that the mortality rate could be reduced if melanotic lesions are detected at a very early stage [6].

The research work presented here puts forth the MPECS to aid early detection of in situ melanoma and classification of skin lesions. The dermoscopic images of skin lesions are represented as a set of 21 features or parameter extracted phase wise. The skin lesions represented as a set of features need to be classified. Statistical analysis carried out for on the parameter set obtained post extraction were inconclusive in classifying the skin lesions [7] hence there was an requirement for advanced classification mechanisms to be incorporated into the MPECS. Machine learning algorithms are adopted for accurate classification. A MFNN classifier trained using the back propagation algorithm is adopted. The training phase of the MPECS enables the system to understand the features that exhibit properties of specific classes defined during training. Subsequently the MPECS can be utilized for classification of the testing data set. The proposed methodology is compared with the popular SVM classifier.

This research manuscript is organized as follows. Section 2 discusses the literature survey about skin cancer melanoma diagnosis systems and classification methodologies adopted by researchers. The next section discusses the proposed MPECS at length along with the MFNN classifier. The experimental evaluation and comparisons are discussed in the penultimate section of this paper. The conclusion of the research work presented in this paper is discussed in the fifth section.

2. LITERATURE SURVEY

The Asymmetry Border Color and Differential Structure (ABCD) methodology is the most commonly used in the diagnosis of skin cancer melanoma [8]. To enhance the efficiency additional optimizations such as fuzzy border extraction [9], “JSEG” based optimization [10]; hybrid techniques [11] have been incorporated. A classification accuracy of about 60% was achieved by incorporating wavelet decomposition techniques along with the ABCD principle. Skin lesion classification can also be achieved using the Menzies Method [12]. The ABCD based methods and the Menzies Method exhibit lesser accuracy in detection of early stages of skin cancer melanoma. G. Betta at el [13] introduced a seven point checklist method to diagnose skin lesions which proved to exhibit higher classification accuracy than the other methodologies described. The computational complexity of the seven point checklist method is considered as a major drawback. The, methods discussed above exhibited less of low classification accuracy. The need of machine learning algorithms to accurately classify and diagnose early symptoms of skin cancer melanoma arose. The use of advance clustering techniques [14], k Nearest Neighborhood [6], classification/regression trees [15] has been closely studied. The use of the support vector machines (SVM) for classification [16] [17] has been proposed by most of the researchers even for early detection of skin cancer melanoma [18]. The MPECS adopts the MFNN based classifier [23]. The above mentioned methodologies are inaccurate to detect and classify early stages of skin cancer melanoma. A few systems proposed to detect early symptoms exhibit low classification accuracy as the early stages of skin cancer melanoma described through features belong to the non-melanoma and melanoma classes cumulatively making classification a very difficult or cumbersome task.
3. MULTI PARAMETER EXTRACTION AND CLASSIFICATION SYSTEM (MPCS) MODELING

Appropriate skin lesions diagnosis at early stages enable possible cure to melanoma skin cancers. Skin lesions of the melanotic type in the in situ stages are easily mistaken with other skin related ailments by even experienced physicians. This paper introduces a computer aided skin lesion diagnosis system named Multi Parameter Extraction and Classification System (MPCS) aiding early detection of skin cancer melanoma. Skin cancer melanoma accounts for the highest number of deaths amongst skin cancers. The MPCS relies on the machine learning techniques incorporated to accurately classify the skin lesions defined by a set of parameters extracted.

3.1. MPCS – PRELIMINARY NOTATION

Let \( I_{SK} \) represent a set of dermoscopic skin lesion images for analysis defined as

\[
I_{SK} = \{i_1, i_2, i_3, i_4, \ldots, i_n\}
\]

Where

\( i \) represents an image

\( n \) represents the total number of images to be analyzed.

The MPECS system proposed in this paper aids the early diagnosis of skin cancer melanoma skin lesions. Let the classification set of the skin lesions be defined as

\[
C = \{c_1, c_2, c_3, \ldots, c_c\}
\]

Where \( c \) represents the class and \( c \) represents the total number of classes considered.

Based on the above mentioned definitions it could be stated that \( \forall \ i \in I : \exists ! \ c \in C \mid i \in c \).

The MPCS aids the classification of skin lesions from dermoscopic images. The skin lesions are defined as a set of parameter \( P \) that it exhibits. Let the Parameter set be defined by

\[
P = \{p_1, p_2, p_3, \ldots, p_p\}
\]

Where \( p \) represents the total number of \( p \) parameters

In the MPECS an image \( i \in I \) is represented by a set of parameters \( P \) is used for classification. The image \( I_n \) could now defined as

\[
I_n = \{p_{n_1}, p_{n_2}, p_{n_3}, \ldots, p_{n_p}\}
\]

An image \( I_n \) could be represented as a matrix and is defined as
Where \( H \) represents the height and \( W \) represents the width of the image \( I \), and \( x \) represents the pixel value at the given location.

The MPECS adopts a 6 phase approach to extract the parameters set \( P \forall i \in \mathbb{R}^k \). The system considers 21 parameters for classification i.e. \( a = 21 \). Each phase provides towards the construction of the parameter set \( P \). The MPECS relies on the supervised machine learning algorithm for classification. The use of multi-layer feed forward neural networks (MFNN) with back propagation [24] training is adopted for classification. The trainings set \( T_R \) is defined as

\[
T_R = \{R_{TR1}, R_{TR2}, R_{TR3}, \ldots, R_{TRn}\}
\]

Where \( R_{TR} \) represents the \( r^{th} \) training record.

The training record is constructed from the parameter set and the corresponding class \( c \) from the training image. I.e. \( R_{TR} = \{I_r \cup c_r\} \) where \( I_r \) represents the parameter set that defines the \( r^{th} \) training image and \( c_r \) represents the corresponding classification class. Similarly the test set \( T_t \) is defined as

\[
T_t = \{R_{TT1}, R_{TT2}, R_{TT3}, \ldots, R_{TTt}\}
\]

The objective of the MPECS is to identify \( c_t \mid R_{TTt} \in c_t \ and \ c_t \in C \)

### 3.2 OPERATION OF THE MPECS

The MPECS relies on supervised learning techniques for classification. To impart intelligence the MPECS is trained using the pre-classified dermoscopic images. The training operation of the MPECS is described by the algorithm given below where \( MFNN_{TR}(X) \) represents the back propagation training function based on the inputs set \( X \). The parameters extracted from the phase \( p \) is represented by \( p_{TRi} \).

**Algorithm Name:** Training Phase of the MPECS

**Input:**
1. \( r \) training dermoscopic images
2. classification set \( C = \{c_1, c_2, c_3 \ldots c_n\} \)

**Output:**
1. Trained MFNN classifier.
Algorithm
Step 1: **Training Phase Begin**
Step 2: \(T_R = \emptyset\)
Step 3: For all training images \(i : 1 \text{ to } r\)
Step 4: \(R_{TRi} = \emptyset\)
Step 5: **MPECS Parameter Extraction Begin**
Step 6: Parameter Set \(P_{TRi} = \emptyset\)
Step 7: For phases \(p : 1:6\)
Step 8: \(P_{TRi} = P_{TRi} \cup p_{TRi}\)
Step 9: End phase
Step 10: Parameter Extraction End
Step 11: \(R_{TRi} = P_{TRi} \cup c_i\)
Step 12: \(T_R = T_R \cup R_{TRi}\)
Step 13: End For
Step 14: Perform Training \(MFNN_{TR}(T_R)\)
Step 15: **End Training Phase**

The **MFNN** post training could be used for classification of the skin lesions represented in the testing dermoscopic image dataset. The testing phase of the **MPECS** is described in the algorithm given below where \(p_{TTi}\) represents the parameters extracted in the \(p^{th}\) extraction phase and \(MFNN_{TT}(Y)\) is the classification function of \(Y\) defined as

\[
MFNN_{TT}(Y) = c_t
\]

Where \(c_t \in C\)

**Algorithm Name:** Testing Phase of the **MPECS**

**Input:**
1. \(t\) testing dermoscopic images
2. Trained \(MFNN_{TR}(T_R)\)

**Output:**
1. Classification Output \(c_t \in C\)

Algorithm
Step 1: **Testing Phase Begin**
Step 2: \(T_T = \emptyset\)
Step 3: For all testing images \(i : 1 \text{ to } t\)
Step 4: \(R_{TTi} = \emptyset\)
Step 5: **MPECS Parameter Extraction Begin**
Step 6: Parameter Set \(P_{TTi} = \emptyset\)
Step 7: For phases \(p : 1:6\)
Step 8: \(P_{TTi} = P_{TTi} \cup p_{TTi}\)
Step 9: End phase
Step 10: Parameter Extraction End
Step 11: \(R_{TTi} = P_{TTi}\)
Step 12: \(T_T = T_T \cup R_{TTi}\)
Step 13: End For
Step 14: Perform Classification Query \(MFNN_{TT}(T_T) = c_t\)
Step 15: **End Testing Phase**
The training and testing phases of the MPECS have been discussed above which explains the operation. Parameter extraction is an integral part of the MPECS which defines the skin lesions and aids in training and classification. The 21 parameter extraction of the MPECS is achieved using a 6 phase approach described in the subsequent section of this paper.

3.3. PARAMETER EXTRACTION IN THE MPECS

The dermoscopic images of skin lesions is represented as a set of 21 parameters extracted in 6 phases described below [7]

3.3.1. MPECS PARAMETER EXTRACTION – PHASE 1

This preliminary phase is primarily designed towards identifying the skin lesions present in an image and extracting the Lesion Borders, the symmetry of the lesions and the color spreading factor of the skin lesions.

To extract the parameters discussed the images are resized using a bicubic interpolation technique. The bicubic interpolation technique is a weighted average of the pixels in the nearest proximity and is defined as

\[ p_{\text{Res}}^{\text{in}}(x, y) = \sum_{i=1}^{3} \sum_{j=1}^{3} (w_{ij} \times x'^{i} y'^{j}) \]

Where \( x, y \) represent the pixel position and \( w_{ij} \) is the weight at position \( i,j \) of the image \( i_n \in I_{5K} \).

Grey scaling is performed on the resized pixels, and the resultant grey scale pixel is computed based on the following definition.

\[
\begin{align*}
P_{\text{GreyScale}}^{\text{in}}(x, y) &= [0.2989 \times R(p_{\text{Res}}^{\text{in}}(x, y))] + [0.5870 \times G(p_{\text{Res}}^{\text{in}}(x, y))] \\
&+ [0.11400 \times B(p_{\text{Res}}^{\text{in}}(x, y))]
\end{align*}
\]

Where \( R(p_{\text{Res}}^{\text{in}}(x,y)) \) represents the red channel value, \( G(p_{\text{Res}}^{\text{in}}(x,y)) \) represents the green channel value and \( B(p_{\text{Res}}^{\text{in}}(x,y)) \) represents the blue channel value of the resized pixel \( p_{\text{Res}}^{\text{in}}(x,y) \).

Binirazation is performed on the grayscale image utilizing the adaptive thresholding algorithm. The image noise elimination property of the adaptive thresholding based binirazation algorithm is an additional parameter for adopting this algorithm. The adaptive thresholding based binirazation is an iterative process wherein the thresholds are adapted based on the regions it is being performed on. The iterative process is terminated when convergence is achieved. Let \( P_{\text{Bin}}^{\text{in}}(x, y) \) represent the pixels of the binirized image \( i_n \). The regions of interest (roi) is extracted using the connected component labeling algorithm. The roi’s obtained hold the information required and are identified by labels \( L_{ROI-l}^{\text{in}} \). The image
i_n ∈ I_{SK} based on the ROI’s is defined as

\[ i_{n-\text{roi}} = \{L_{\text{roi}-1}^{i_n} \cup L_{\text{roi}-2}^{i_n} \cup L_{\text{roi}-3}^{i_n} \cup \cdots \cup L_{\text{roi}^{-1}}^{i_n}\} \]

Where \( l \) represents the total number of roi’s extracted.

The centroids of the roi’s are computed based on the area enclosed by the roi’s. The centroids are computed to cluster the similar roi’s together in the same vicinity as it is observed that the roi’s are similar in nature and can be clustered not effecting the classification accuracy. This operation adopted enables roi’s exhibiting similar properties to be clustered into a single cluster together represented as ROI. The labels \( L_{\text{roi}^{-1}}^{i_n} \) used to address the roi’s are clustered based on the energy computed and is defined as

\[ E_{\text{ROI}^{-1}}(L_{\text{ROI}^{-1}}^{i_n}) = \sum_{\text{pos}} \sum_{l \in S_{roi}} \Delta(L_{\text{roi}^{-1}}^{i_n}, L_{\text{roi}^{-1}}^{i_n}) \]

Where \( \Delta(L_{\text{roi}^{-1}}^{i_n}, L_{\text{roi}^{-1}}^{i_n}) \) represents a function \( \Delta(L_{\text{roi}^{-1}}^{i_n}, L_{\text{roi}^{-1}}^{i_n}) = \{0,1\} \) i.e. if the roi’s \( L_{\text{roi}^{-1}}^{i_n} \) and \( L_{\text{roi}^{-1}}^{i_n} \) are similar, 0 is returned and 1 in the case of dissimilarity. \( S_{roi} = \{1,2,3,\ldots,i\} \) represents the set of roi’s obtained and \( a \in S_{roi} \) and \( b \in S_{roi} \). pos represents the position of the \( i \)th clustered ROI represented as \( L_{\text{ROI}^{-1}}^{i_n} \).

The clustered ROI’s obtained define the image \( i_n \) as

\[ i_{n-\text{ROI}} = \{L_{\text{ROI}^{-1}}^{i_n} \cup L_{\text{ROI}^{-2}}^{i_n} \cup L_{\text{ROI}^{-3}}^{i_n} \cup \cdots \cup L_{\text{ROI}^{-1}}^{i_n}\} \]

Sobel edge detection is performed on the image \( i_{n\text{ROI}} \). Post the edge detection the distortion on the basis of the coordinate displacement and the frequency of distortion observed is computed using the Fast Fourier Transforms. The axis based asymmetry parameter of the lesion is obtained based on the principal component decomposition. The skin lesion extracted is rotated such that the principal component coincides with the horizontal axis. The skin lesion axis based asymmetry parameter is defined as

\[ P_{\text{Axis Symmetry}} = \left(\frac{\sum |i_n(x) - \bar{i}_n(x)|}{\sum i_n(x)}\right) \]

Where \( i_n(x) \) the original skin lesion and its reflected version is represented as \( \bar{i}_n(x) \).

The symmetry parameter is computed for all the pixels within the lesion along the horizontal and vertical axis respectively. The color spreading factor of the lesions is computed based on the similarity of a pixel in a \( n \times n \) neighbourhood. The standard deviation of the pixels in the neighborhood based on the Red channel, Green channel and Blue channel is computed using

\[ \sigma_{\text{Chnl}} = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n^2} (\text{Chnl}_i - \overline{\text{Chnl}})^2} \]

Where \( \text{Chnl} = \{R_{\text{Chnl}}, G_{\text{Chnl}}, B_{\text{Chnl}}\} \) and \( \overline{\text{Chnl}} \) represents the mean value defined as
The cumulative standard deviation for all the channels is defined as
\[
\sigma_{\text{Cumulative}} = \sigma_{R_{\text{Chnl}}} + \sigma_{G_{\text{Chnl}}} + \sigma_{B_{\text{Chnl}}}
\]

\[
\sigma_{\text{Cumulative}} = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n^2} (R_{\text{Chnl}} - \overline{R_{\text{Chnl}}})^2 + \frac{1}{n^2} \sum_{i=1}^{n^2} (G_{\text{Chnl}} - \overline{G_{\text{Chnl}}})^2 + \frac{1}{n^2} \sum_{i=1}^{n^2} (B_{\text{Chnl}} - \overline{B_{\text{Chnl}}})^2}
\]

The color spreading factor of the lesion is computed using
\[
P_{\text{Clr-Spread}} = 1 - \sigma_{\text{Norm}}
\]

Where \(\sigma_{\text{Norm}}\) is the normalized value of \(\sigma_{\text{Cumulative}}\) between the range 0 and 1 and is obtained as
\[
\sigma_{\text{Norm}} = \frac{\sigma_{\text{Cumulative}}}{\sigma_{\text{Cumulative-Max}}}
\]

In order to extract the lesion border parameters the image is filtered using the Gaussian and laplacian filter. The parameter describing the lesion borders is obtained by computing the mean of the resultant pixels and is defined as
\[
P_{\text{Lesion-Boundary}} = \left[ \sum_{x=1}^{Wd} \sum_{y=1}^{Ht} P^{in}(x,y) \right] / (Wd \times Ht)
\]

In this phase the parameters of the lesions such as the asymmetry along the horizontal and vertical axis, the color spreading factor of the lesion and the boundary parameters is discussed.

3.3.2. MPECS PARAMETER EXTRACTION – PHASE 2

The second phase of the MPECS enables the extraction of the area, perimeter and the eccentricity. This phase considers the binirized and clustered ROI image as an input and computes the parameters based on the binirized image obtained from phase 1 defined as
\[
i_{n-ROI} = \{ L_{ROI-1}^{i_n} \cup L_{ROI-2}^{i_n} \cup L_{ROI-3}^{i_n} \cup \cdots \cup L_{ROI-\ell}^{i_n} \}
\]

Where \(\ell\) represents the total number of ROI’s identified for the image \(i_n\).

The ROI’s of image \(i_{n-ROI}\) contain binirized pixel points i.e. black or white pixels. The area parameter defines the white pixels that are encapsulated by the skin lesion and the lesion area parameter is computed using
Where $Pnts$ is the number of points enclosed by the skin lesion defined by the $l^{th}$ ROI. Also $i, j$ represent the $Pnts$ location and $(x_i, y_j) \in \{(x_i, y_j) | f(x_i, y_j) = 1\}$ as $x_i, y_j$ represents a white pixel.

The perimeter defines the boundary $\Delta B$ length of the skin lesion. The lesion boundary is computed using

$$\Delta B(l) = (B(2; l, :) - B(1; l - 1, :))^2$$

Where $l$ represents the number of ROI's and its corresponding boundary is represented as $B$. The parameter defining boundary of the skin lesion is defined as

$$P_{\text{Lesion-Bound}} = \sum \sqrt{\sum \Delta B(l)}$$

The aspect ratio of the skin lesion is defined as the eccentricity of a skin lesion. The eccentricity is obtained by obtaining the moments represented as $M$. The major axis is defined as

$$A_{\text{Major}} = 2\sqrt{2(M_{xx} + M_{yy} + M_{comm})}$$

The minor axis is defined as

$$A_{\text{Minor}} = 2\sqrt{2(M_{xx} + M_{yy} - M_{comm})}$$

Where the moments $M_{xx}, M_{yy}$, $M_{xy}$ and $M_{comm}$ are defined as

$$M_{xx} = \left[\left(\sum x^2\right)/N + 1/12\right]$$

$$M_{yy} = \left[\left(\sum y^2\right)/N + 1/12\right]$$

$$M_{xy} = \left[\left(\sum x \times y\right)/N\right]$$

And

$$M_{comm} = \sqrt{(M_{xx} - M_{yy})^2 + 4M_{xy}^2}$$

The eccentricity parameter is computed using

$$P_{\text{Lesion-Eccent}} = \left[\sqrt{(A_{\text{Major}}^2 - A_{\text{Minor}}^2)}/A_{\text{Major}}\right]$$

### 3.3.3. MPECS PARAMETER EXTRACTION – PHASE 3

Researches have well understood the importance to 3D depth parameters to enhance the classification accuracy [6] [19]. To obtain the 3D depth of the skin lesions the MPECS introduced in this paper considers creating a $3 \times 3$ masked windows represented as

$$i_{n_{3x3-Mask}} = \begin{bmatrix} W_1 & W_2 & W_3 \\ W_4 & W_5 & W_6 \\ W_7 & W_8 & W_9 \end{bmatrix}$$

Where $W_i$ represent the weights of the pixel

The projection filter is defined as
Where $i_i$ represents the intensity of the $i^{th}$ pixel.

The 3D depth projection parameter $p_{3D\text{-}Depth}$ is obtained by considering the geometric mean defined as

$$p_{3D\text{-}Depth} = \frac{1}{9} \sum_{i=1}^{9} W_i i_i$$

Where $W_i$ represents the horizontal and vertical pixel position.

### 3.3.4. MPECS PARAMETER EXTRACTION – PHASE 4

The fourth phase of the MPECS is targeted towards obtaining the color components of the skin lesions. The color components are extracted for the red channel blue channel and the green channel. The mean and the variance parameters of each channel are considered as the parameters to be utilized for classification. For the red channel the mean parameter is defined as follows

$$p_{RChnl\text{-}Mean} = \frac{\sum_{x=1}^{Wd} \sum_{y=1}^{Ht} R_{Chnl}(x, y)}{Wd \times Ht}$$

Where $x, y$ represent the pixel positions.

The variance parameter is computed by obtaining the mean of all the $B_{Chnl}$ columns, the difference amongst them and then summed square difference divided by the height $Ht$

$$Clm\_mean_{RChnl}(i) = \frac{\sum_{i=1}^{Wd} R(i)}{Wd}$$

$$Diff_{RChnl}(i) = RChnl(:, i) - Clm\_mean_{RChnl}(i)$$

$$p_{RChnl\text{-}var} = \frac{\sum_{i} Diff_{RChnl}(i)}{(Ht - 1)}$$

Similarly the green channel and blue channel parameters are defined as

$$p_{GChnl\text{-}Mean} = \frac{\sum_{x=1}^{Wd} \sum_{y=1}^{Ht} G_{Chnl}(x, y)}{Wd \times Ht}$$

$$p_{GChnl\text{-}var} = \frac{\sum_{i} Diff_{GChnl}(i)}{(Ht - 1)}$$

$$p_{BChnl\text{-}Mean} = \frac{\sum_{x=1}^{Wd} \sum_{y=1}^{Ht} B_{Chnl}(x, y)}{Wd \times Ht}$$

$$p_{BChnl\text{-}var} = \frac{\sum_{i} Diff_{BChnl}(i)}{(Ht - 1)}$$

$$i_{3D\text{-}Depth} = \sum_{i=1}^{9} W_i i_i$$

Where $i_i$ represents the intensity of the $i^{th}$ pixel.

The 3D depth projection parameter $p_{3D\text{-}Depth}$ is obtained by considering the geometric mean defined as

$$p_{3D\text{-}Depth} = \frac{1}{9} \sum_{i=1}^{9} W_i i_i$$

Where $x, y$ represent the horizontal and vertical pixel position.
3.3.5. MPECS PARAMETER EXTRACTION – PHASE 5

This phase of the MPECS discusses the smoothening process of the RChnl and GChnl using the 3 × 3 masking procedure discussed in the third phase. The smoothening is not adopted for the BChnl as the blue veils of the skin lesions is an important parameter for early detection of skin cancer melanoma [20]. The smoothening filter for the RChnl and GChnl is defined as follows:

\[ i_{n_{RChnl-Smooth}} = \sum_{i=1}^{9} W_i RChnl_i \]
\[ i_{n_{GChnl-Smooth}} = \sum_{i=1}^{9} W_i GChnl_i \]

Where \( W_i \) represent the weights of the pixel and \( RChnl_i, GChnl_i \) is red channel intensity and green channel intensity of the \( i^{th} \) pixel.

The red channel smoothened parameter \( p_{RChnl-Smooth} \) is defined as:

\[ p_{RChnl-Smooth} = \left[ \frac{1}{9} \right] \\
\times \left[ RChnl_{x-1,y-1} + RChnl_{x-1,y+1} + RChnl_{x-1,y+1} + RChnl_{x,y-1} + RChnl_{x,y} + RChnl_{x,y+1} + RChnl_{x+1,y-1} + RChnl_{x+1,y+1} \right] \]

The green channel smoothened parameter \( p_{GChnl-Smooth} \) is defined as:

\[ p_{GChnl-Smooth} = \left[ \frac{1}{9} \right] \\
\times \left[ GChnl_{x-1,y-1} + GChnl_{x-1,y+1} + GChnl_{x-1,y+1} + GChnl_{x,y-1} + GChnl_{x,y} + GChnl_{x,y+1} + GChnl_{x+1,y-1} + GChnl_{x+1,y+1} \right] \]

3.3.6. MPECS PARAMETER EXTRACTION – PHASE 6

For early detection of skin cancer melanoma it is essential to extract all the color components of the skin lesion to obtain accurate classification results [21][22]. The MPECS discusses the cylindrical coordinate representation of pixels in the fourth and fifth phase. In order to extract accurate and elaborate color parameters this phase of the MPECS considers the Cartesian representation of pixels. Hue, Saturation and Value representation of the pixels are considered as the Cartesian representations. The mean and variance parameters of the hue channel, variance channel and the value channel are extracted from the \( RChnl, GChnl, BChnl \) pixel values of the skin lesions. The hue value of a pixel is computed using the following definition:

\[ \text{Hue}_{i,j} = \begin{cases} 
0 + \frac{43 \times |GChnl - BChnl|}{\text{MaxVal} - \text{MinVal}} & , \text{MaxVal} = \text{RChnl} \\
85 + \frac{43 \times |BChnl - RChnl|}{\text{MaxVal} - \text{MinVal}} & , \text{MaxVal} = \text{GChnl} \\
171 + \frac{43 \times |RChnl - GChnl|}{\text{MaxVal} - \text{MinVal}} & , \text{MaxVal} = \text{BChnl}
\end{cases} \]

Where \( \text{MaxVal} = \text{Max}(RChnl, GChnl, BChnl) \) is the maximum value of the red and blue channel of the pixel.
MinVal = Min(RChnl, GChnl, BChnl) is the minimum value of the red green and blue channel of the pixel.
The Saturation and the value parameter of the pixel is computed using

$$Saturation_{i,j} = 255 \times \frac{MaxVal - MinVal}{MaxVal}$$

$$Value_{i,j} = MaxVal$$

The mean of the hue channel and the is defined as

$$p_{Hue_{Chnl}-Mean} = \frac{\sum_{x=1}^{Wd} \sum_{y=1}^{Ht} Hue_{Chnl}(x, y)}{Wd \times Ht}$$

The variance is computed in a manner similar to the procedure described in phase 4 and is defined as

$$p_{Hue_{Chnl}-Var} = \frac{\sum_{i} Diff_{Hue_{Chnl}}(i)}{(Ht - 1)}$$

The mean and the variance of the saturation channel is defined as

$$p_{Saturation_{Chnl}-Mean} = \frac{\sum_{x=1}^{Wd} \sum_{y=1}^{Ht} Saturation_{Chnl}(x, y)}{Wd \times Ht}$$

$$p_{Saturation_{Chnl}-Var} = \frac{\sum_{i} Diff_{Saturation_{Chnl}}(i)}{(Ht - 1)}$$

Accordingly the Value channel parameters are obtained using the following equations

$$p_{Value_{Chnl}-Mean} = \frac{\sum_{x=1}^{Wd} \sum_{y=1}^{Ht} Value_{Chnl}(x, y)}{Wd \times Ht}$$

$$p_{Value_{Chnl}-Var} = \frac{\sum_{i} Diff_{Value_{Chnl}}(i)}{(Ht - 1)}$$

The MPECS discussed in this paper discusses the extraction 21 vital parameters which would enable for classification of skin lesions especially targeted towards early detection of skin cancer melanoma. The parameter extraction and the procedure involved in extraction are achieved adopting a six phase approach discussed above.

3.4. MFNN CLASSIFIER

![Figure 1: A MFNN Classifier Structure](image)
The classification of the testing data set of dermoscopic images requires a supervised training procedure to be incorporated. A multilayer feed forward neural network (MFNN) with the back propagation algorithm for training is adopted for the purpose of classification. The parameters extracted from the testing image set are directly provided to the input neurons of the neural network. A sample MFNN is shown in Fig 1. The MFNN consists of a set of input neurons, M number of hidden layers and an output layer. The MFNN classifier is a machine learning algorithm. The learning and classification is achieved based on the weights which are changed in the neighboring layers. The input to the neuron in every layer represented as $\gamma$ where $\gamma = 1,2,3, ..., c_r$ and $c_r$ is the trained $r$ classes.

The output of the neuron is defined as

$$a_j^{my} = f(\phi_j^{my}) = f \left( \sum_{i=0}^{n_{m-1}} W_{ji} a_i^{(m-1)y} \right)$$

Where

$m$ represents the index of the hidden layer i.e. $1 \leq m \leq M$
$n_m$ represents the number of neurons in $m^{th}$ layer
$j$ represents the index $j$ | $1 \leq j \leq n_m$
$\phi_j^\gamma$ represents the weighted sum of input values for $j^{th}$ neuron in layer $a$
$W_{ji}^m$ represents the weight between $j^{th}$ neuron in $m^{th}$ layer and $i^{th}$ neuron in $(m-1)^{th}$ layer

The classification error $\delta_j^{my}$ observed at the output layer is obtained from the following definition

$$a_j^{my} = f'(\phi_j^{my}) \delta_j^\gamma = f'(\phi_j^{my})(y_j^{2\gamma} - y_j^\gamma)$$

The sum of square errors observed through the MFNN is defined as

$$\xi_y = \frac{1}{2} \sum_{j=1}^{a} (\delta_j^\gamma)^2$$

The error existent at the output layers is propagated back and is defined as

$$\delta_j^{my} = f'(\phi_j^{my}) \sum_{\ell=1}^{d_{m+1}} \delta_{\ell}^{(m+1)\gamma} W_{\ell j}^{(m+1)}$$

The errors propagated is minimized by altering the weights of the neighboring neurons and the weight update function is defined as

$$\Delta y W_{ji}^m = \eta \delta_j^{my} u_i^{(m-1)\gamma}$$

The classification error that exist through

$$\xi_{class} = \sum_{\mu=1}^{c_r} \xi_{\mu} = \frac{1}{2} \sum_{\mu=1}^{c_r} \sum_{j=1}^{a} (y_j^{2\gamma} - \phi_j^\gamma)^2$$

The MFNN Classifier is designed to minimize the classification error by effective weight adjustments achieved due to the back propagation algorithm.
The MPECS adopts machine learning techniques for classification. The MFNN based classifier is adopted for classification which classifies based on the supervised learning imparted using the back propagation algorithm. Parameter extraction is achieved using a six phase approach to define the skin lesion. The training dataset constitutes of a set of dermoscopic images which are pre-classified with the aid of dermatologists. The images defined in terms of the parameters and the classified type constitute towards the training dataset. The MFNN classifier is trained using the training data set. The testing set consists of dermoscopic images that need to be diagnosed. The diagnostic results are represented as classes. To enable diagnosis the test images are represented as a set of parameters extracted using the MPECS parameter extraction phases. The resulting dataset is termed as the test dataset. The test dataset is provided to the trained MFNN for classification. The subsequent section of this paper discusses the experimental verification and comparisons with the SVM classifiers.

4. EXPERIMENTAL STUDY AND COMPARISONS

The MPECS discussed in this paper was realized on the MATLAB platform. In order to evaluate the MPECS training and testing dataset are to be created. To construct the training and testing datasets dermoscopic images are obtained from the “The Atlas of Dermoscopy” [25]. The atlas considered is a collection of over 2000 skin lesion images obtained from various patients. The atlas also provides the classification and diagnosis data which is used as a reference to evaluate the classification accuracy of the MPECS system proposed. The use of the atlas is charted as others researchers too have used the same atlas to evaluate their proposed diagnosis mechanisms especially skin cancer melanoma [20] [26]. The MPECS proposed is compared with the popular SVM classifiers as proposed by Xiaojing Yuan et al.[18].

![Classification Error](image-url)
The training dataset consists of pre-classified dermoscopic images. We have considered a set of 45 dermoscopic images from the “Atlas of Dermoscopy” as the training data set. The training dataset consist of skin lesion images of primarily 3 classes namely “Advanced Melanoma”, “Early Melanoma” and “Non-Melanoma”. The 21 parameters extracted per image using the 6 phase approach contribute to each record of the training data set along with the classified name. A similar approach is followed to construct the testing dataset. The MPECS and the SVM Classifier was trained using similar training datasets. Similar test data sets were evaluated used the MPECS and SVM classifier. The classification results obtained are provided to the Receiver Operating Characteristic (ROC) analysis tool. The error in classification obtained is graphically shown in Fig 2. From the figure it is evident that the classification error of the SVM classifier is greater than the error exhibited by the proposed MPECS system. The MPECS and the SVM classifier classifiers are compared on the basis of the receiver operating characteristic (ROC) analysis. The ROC analysis is carried out using a tool developed using C#.Net programming language on the Visual Studio 2010 platform. The ROC analysis plot obtained is shown in Fig 3.

![ROC Curve Analysis](image)

Figure 3 : ROC Curve Analysis

The classification accuracy and the efficiency of the MPECS Classifier along with the SVM Classifier are diagrammatically shown in Fig 4 and Fig 5. The MPECS system exhibits an 81% accuracy in classification as compared to the 75% accuracy exhibited by the SVM Classifier. The MPECS system proposed in this paper improves the efficiency in classification by about 7.4% when compared to the SVM classifier. The experimental study conducted and from the results discussed in this section of the paper, it can be concluded that the MPECS system proposed in this paper is capable of accurately classifying skin lesions exhibiting early melanotic symptoms when compared to the popularly used SVM classifier systems.
5. CONCLUSION

This paper highlights the benefits and the requirement for early detection of skin cancer melanoma. The Multi Parameter Extraction and Classification System (MPECS) discussed in this paper relies on a supervised MFNN based classifier to accurately classify skin lesions and enable early detection of skin cancer melanoma. The MPECS comprehensively defines skin lesions in terms of a parameter sets extracted using the six phase approach. The training and testing phase algorithms of the MPECS is discussed in this paper. The MPECS system is trained using the Back Propagation Algorithm. The trained classifier of the MPECS is used for analysis and diagnosis of skin lesions on unknown types. The classification preeminence of the MPECS is proved over the popular SVM classifier systems. The results obtained and discussed prove that the MPECS system proposed in this paper aids the detection of skin cancer melanoma in the naïve stages thereby improving computer aided diagnosis mechanisms and early treatment initiatives to be adopted by dermatologists for the benefit of patients suffering from the life threatening melanotic skin cancer diseases.
REFERENCES


[28] S. Shenbaga Ezhil and Dr. C. Vijayalakshmi, “Prediction of Colon-Rectum Cancer Survivability Using Artificial Neural Network” International journal of Computer Engineering & Technology (IJCET), Volume 3, Issue 1, 2012, pp. 163 - 168, Published by IAEME