Tech Science Press

Skin Melanoma Classification System Using Deep Learning

R. Thamizhamuthu[*] and D. Manjula

Department of Computer Science and Engineering, CEG Campus, Anna University, Chennai, India
[*]Corresponding Author: R. Thamizhamuthu. Email: tamillec2011@gmail.com

Abstract: The deadliest type of skin cancer is malignant melanoma. The diagnosis requires at the earliest to reduce the mortality rate. In this study, an efficient Skin Melanoma Classification (SMC) system is presented using dermoscopic images as a non-invasive procedure. The SMC system consists of four modules; segmentation, feature extraction, feature reduction and finally classification. In the first module, k-means clustering is applied to cluster the colour information of dermoscopic images. The second module extracts meaningful and useful descriptors based on the statistics of local property, parameters of Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model of wavelet and spatial patterns by Dominant Rotated Local Binary Pattern (DRLBP). The third module reduces the features by the t-test, and the last module uses deep learning for the classification. The individual performance shows that GARCH parameters of 3rd DWT level sub-bands provide 92.50% accuracy than local properties (77.5%) and DRLBP (88%) based features for the 1st stage (normal/abnormal). For the 2nd stage (benign/malignant), it is 95.83% (GRACH), 90% (DRLBP) and 80.8% (Local Properties). The selected 2% of features from the combination gives 99.5% and 100% for 1st and 2nd stage of the SMC system. The greatest degree of success is achieved on PH^2 database images using two stages of deep learning. It can be used as a pre-screening tool as it provides 100% accuracy for melanoma cases.

Keywords: Dermoscopic images; skin cancer; melanoma; deep learning; autoregressive models

1 Introduction

The decrease in mortality rate due to skin cancer may be attributed to several treatment and detection factors. Due to the vast amount of research in both categories, significant advances have seen over the past 30 years. One of the prognostic factors for cancer cure is detection at the earliest. Currently, the advancement in imaging techniques and computerized system provides better results. The imaging techniques used for skin cancer diagnosis is dermoscopy, where a magnified visualization of the affected skin region is acquired. It shows the morphological structures that cannot be found by naked eyes. The accuracy of skin melanoma diagnosis has been improved with the use of many algorithms, such as ABCD rule [1], 7-point checklist [2] and pattern analysis [3]. However, the interpretation is time-consuming and also it is subjective based

on skills of dermatologists. To overcome these difficulties, computerized systems are developed to help the dermatologists.

Among the various components in the computerized system, feature extraction and classification are the main components. Many spatial and spectral features are utilized in the former steps, and many machine learning methods are developed in the later stage for the classification. Apart from ABCD rule [1], quantitative measures such as the shape of the affected region [4], colors and structural differences [5,6] are used for effective classification. Textural patterns play a vital role to extract dominant features in the medical domain. These patterns include Haralick features [7] from gray level co-occurrence matrix, Laws texture features [8], Local Binary Pattern (LBP) [9] and their extensions.

Due to the advancement in multiresolution analysis, frequency domain features are added to the feature vector to increase accuracy further. Some of them are Discrete Wavelet Transform (DWT) [10], Multi wavelet transform [11], Contourlet [12], Curvelet [13], and Shearlet [14]. They can be used for the diagnosis independently or combined with the statistical features for better performance. The novelty is the choice of feature vectors used for effective classification of dermoscopic images. DWT will be used in conjunction with GARCH [15] to generate one of the feature vectors used with other features. In the diagnosis stage, many machine learning approaches are used for classification such as SVM [11], Naive Bayes [12], k-nearest neighbour [16], and decision trees [17].

The evaluation of deep learning [18–20] helps the computerized system highly effective in melanoma diagnosis, and thus this study uses deep learning as a classifier. The objective is to design a computerized SMC system with a high level of sensitivity and specificity. The rest of the paper is as follows: Section 2 discusses SMC system's design and a summary of the results obtained in this study are discussed in Section 3. The conclusion is given in the last Section.

2 Related Works

This Section discusses the design of the SMC system. It consists of four sequential steps which are illustrated in Fig. 1. Section 2.1 explains the lesion segmentation by a clustering approach. Section 2.2 describes how the features are extracted from dermoscopic images. The feature reduction technique is described in Section 2.3, and Section 2.4 explains the deep learning approach for the classification.

2.1 Segmentation

The exact skin cancer region is segmented using k-means clustering approach on RGB colour image. Before segmentation, the noises and hair in the images are removed by averaging filter with a predefined window size of 21×21. The basic k-means clustering approach is applied to segment gray scale images [21]. As the dermoscopic images are colour images and the colour information is very useful for extracting skin lesions, the k-means clustering is modified to accept colour images. The visual differences can be easily quantifiable in $L * a * b*$ colour space than RGB mode. The conversion formulae can be found in [21] and it is known that all colour information can be visible in only two channels; a^* and b^*. k-means clustering is easily applied to cluster the colour information by Euclidean distance metric. In this study, k is set to 3, so that exact lesion area, background and unaffected skin areas are clustered. From the clusters, skin lesion area can be easily separated from the other two clusters. Fig. 2 shows the results of the segmentation approach.
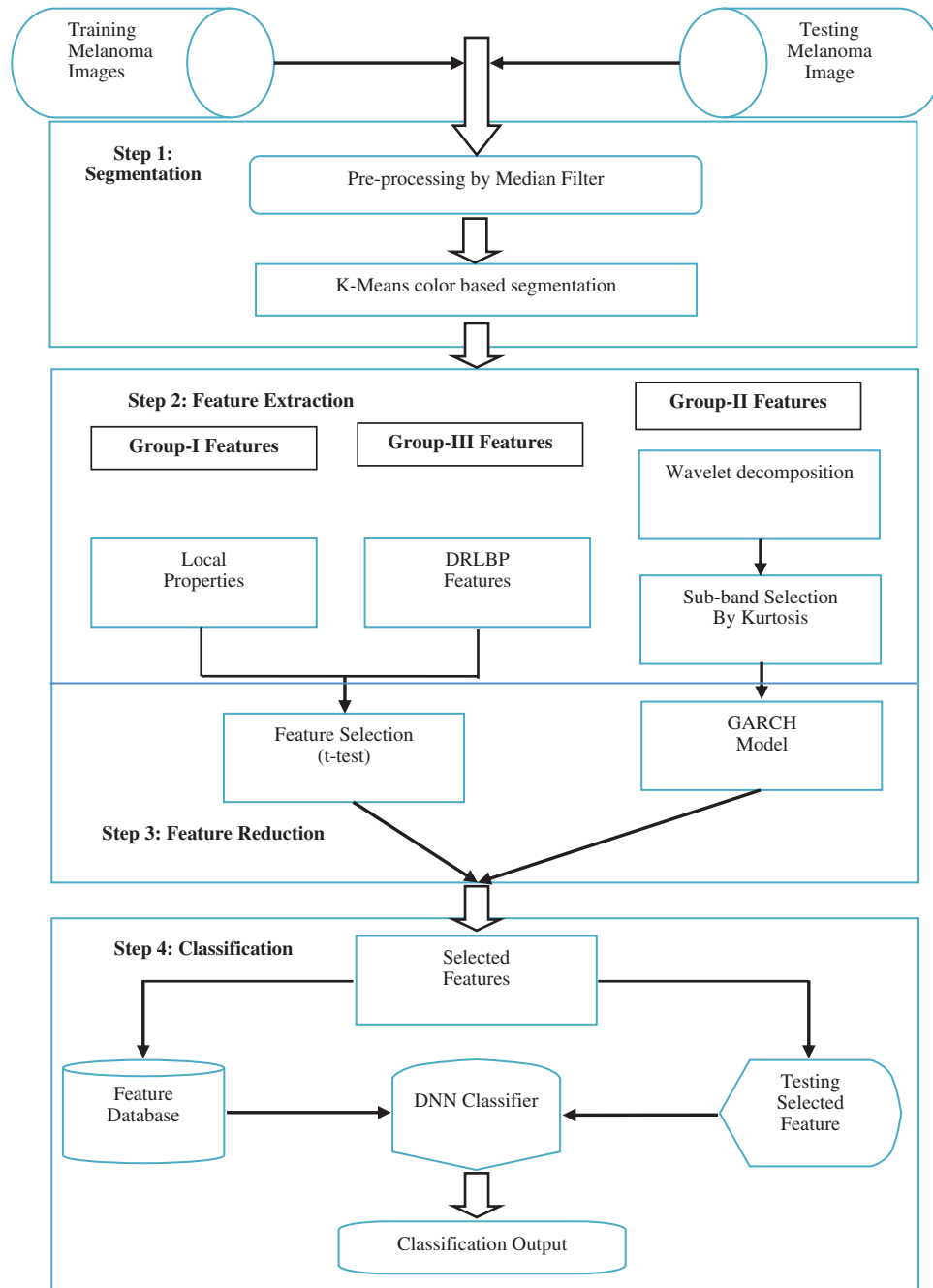
Figure 1: Block diagram of the SMC system

2.2 Feature Extraction

Feature extraction aims to preserve the class discriminating information so that best class separation is achieved for least computational complexity. A classifier then uses these to decide whether the region is normal or abnormal. The advantage of extracting descriptors is that they will be a more compact representation of the segmented region than the image pixels alone if carefully chosen. The features are usually chosen based on the domain under consideration and in

this study fall into three categories. The first group is based on the statistics of local property, and the second group consists of the parameters of the GARCH model of wavelet. Thirdly, spatial patterns of skin lesions are also recorded.
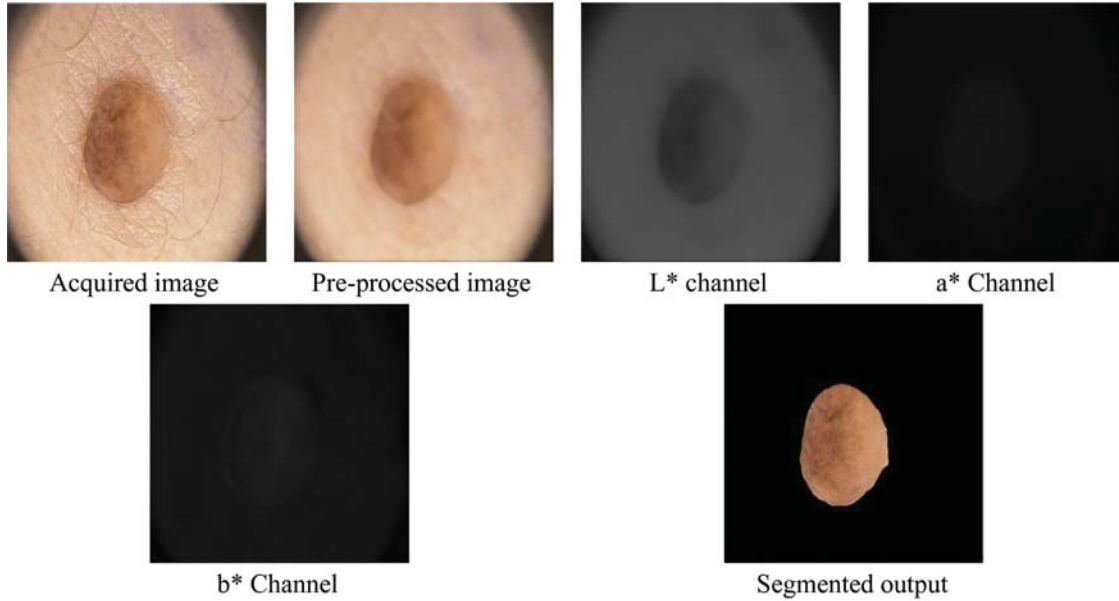

Figure 2: Outputs of segmentation of the SMC system

2.2.1 Local Properties

There are four local properties extracted in this study, mean (μ), standard deviation (σ), skewness (S) and kurtosis (K). They are computed from the central moments about the mean. The first moment, Mean (μ) is simply the total of pixel intensities (PI) divided by the number of pixels (n) in the dermoscopic images (Eq. (1)).

$$\mu = \frac{1}{n}\sum_{i=1}^{n} PI_i \tag{1}$$

The second moment is called variance and its positive square root is called σ (Eq. (2)). It measures how much a PI in the dermoscopic images can be expected to deviate from μ. A low μ indicates that the PI is clustered about σ while high μ means the opposite.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(PI_i - \mu)^2} \tag{2}$$

The third standardized central moment is skewness (Eq. (3)), and the fourth standardized central moment is called kurtosis (Eq. (4)).

$$S = \frac{\frac{1}{n} \sum_{i=1}^{n} (PI_i - \mu)^3}{\left(\frac{1}{n} \sum_{i=1}^{n} (PI_i - \mu)^2\right)^{\frac{3}{2}}} \tag{3}$$

$$K = \frac{\frac{1}{n} \sum_{i=1}^{n} (PI_i - \mu)^4}{\left(\frac{1}{n} \sum_{i=1}^{n} (PI_i - \mu)^2\right)^2} - 3 \tag{4}$$

Using Eqs. (1)–(4), the local proprieties are computed for each colour channel and stored in the feature database.

2.2.2 GARCH Features

To boost the performance of the SMC system, GARCH model is applied in the wavelet domain. The GARCH (p, q) for ε_t is as follows [15]:

$$\varepsilon_t = z_t \sigma_t \tag{5}$$

and

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{k-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{k-j}^2$$

$$\text{where } q > 0; \ p \geq 0; \ \alpha_0 > 0; \ \alpha_i \geq 0; \ \beta_j \geq 0; \ 1 \leq i \leq q; \ 1 \leq j \leq p; \ \text{and} \ \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1 \tag{6}$$

where α_i and β_j are GARCH parameters and these parameters are estimated using maximum likelihood estimation [15] for unit p and q. z_t and σ_t represents random variable and conditional standard deviation. These two variables are computed from a Gaussian distribution which has unit (1) variance and zero (0) mean.

At first, the dermoscopic images are transformed into DWT domain which is a powerful tool used in many pattern recognition techniques [22]. DWT is very useful to increase the accuracy of the SMC system as it provides localized frequency information. Fig. 3 shows a DWT decomposition of a dermoscopic image at two levels. While applying DWT on the dermoscopic image, it is observed that it provides four sub-bands by applying low pass and high pass filter in a predefined manner.

It is noted that the GARCH model is efficient only when the distribution of data has a heavy tail [23]. Thus, the coefficients' distribution in each sub-band in the DWT domain is tested for a heavy tail. It is achieved by calculating the K value using Eq. 4 which demonstrates whether they have a heavy tail or not. For a Gaussian distribution, K is three. If the K value of any data is greater than three, it indicates a heavier tail than the Gaussian distribution. Hence, the GARCH parameters are extracted from the sub-bands which have K of more than three.
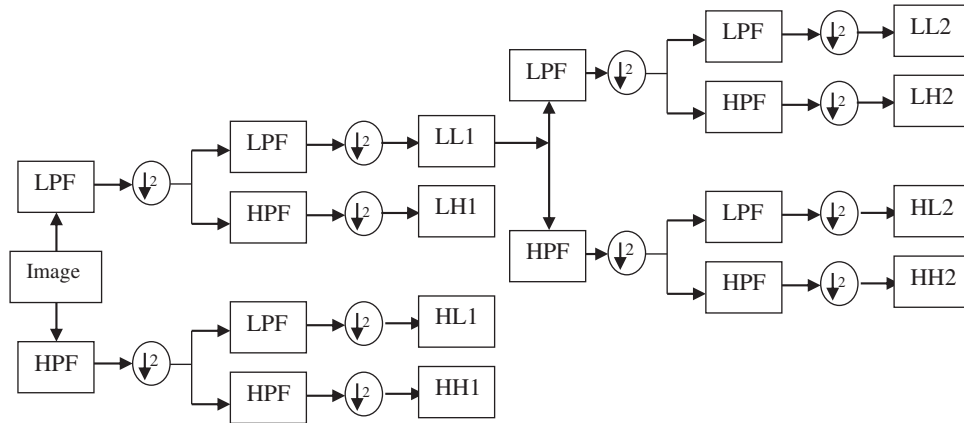
Figure 3: DWT decomposition of a dermoscopic image at two levels

2.2.3 DRLBP Features

DRLBP [24] is an extended version of LBP [9]. LBP features are rotational invariant due to the fixed order of the weights. To overcome this limitation, the weights are arranged based on a reference direction which is computed locally. It is defined as

$$D = \underset{n \in (0,1,\dots N-1)}{\arg\max} |i_n - i_c| \tag{7}$$

where i_c and i_n are intensity of central pixel and *nth* neighborhood pixels respectively. The DRLBP is defined based on D is given below.

$$DRLBP = \sum_{n=0}^{N-1} s(i_n - i_c) \cdot 2^{\mathrm{mod}(n-D,N)},$$

$$s(i_n - i_c) = \begin{cases} 1 & i_n \geq i_c \\ 0 & i_n < i_c \end{cases} \tag{8}$$

It is evident from the Eq. (8), the weights depend on D and thus DRLBP satisfies the rotation invariant property. DRLBP gives the spatial pattern in the dermoscopic images. In this study, they are computed for each colour channel and stored in the feature database. The feature extraction is also used in [25,26].

2.3 Feature Reduction

A large set of features makes the classification system extremely computationally intensive. The complexity of the SMC system increases more when the combination of features used in the classifiers. Thus a feature reduction step is necessary to eliminate the poor performing features that affect the classifier's performance. The significant features are identified using *t*-test [27].

Let us consider features from two classes; c_1 of n_1 samples and c_2 of n_2 samples. A statistical value called as t is computed for a particular feature (f) in the feature set is defined by

$$t(f) = \frac{\mu_{c1}(f) - \mu_{c2}(f)}{\sqrt{\left(\sigma_{c1}^2(f)/n_1 + \sigma_{c2}^2(f)/n_2\right)}} \tag{9}$$

where $\mu_x(f)$, and $\sigma_x(f)$ are the mean and standard deviation of f of xth class. The application of Eq. (9) for all features produces t value. This value indicates the significance of features. They are sorted and features having high t values are selected.

2.4 Classification by Deep Learning

The non-linear relationship between the features of different classes can be modeled by neural networks which consist of input layers (number of features), hidden layer (normally 1) and output layer (number of classes). The information in the bracket shows the number of layers in each layer. The relationship between the features can be effectively modeled if the number of the hidden layer is increased. This is called deep learning [28] and in this study, it is applied for the classification.

The error between the actual and desired output is computed at first. The weights are updated iteratively while computing the error signal in the training phase. The update is done using the mean-squared error function. In the output layer, the error is multiplied using the sigmoid activation function. This process is stopped when the error is minimized at a predefined level by using the backpropagation algorithm. It is a descent algorithm that propagates the error from the output layer to lower layers. The weights are adjusted for the dampening oscillations with the help of learning rate and momentum factor so that the error rate is reduced in a decent direction. As the SMC system outputs a binary decision, a linear function is used in the output layer.

3 Results and Discussion

The developed SMC system is analyzed using publically available dermoscopic image databases; PH^2 [29,30]. It is extremely useful for the development and testing of any computerized skin cancer classification system. Also, they included ground truth data that describes the types of abnormal severity present in the dermoscopic images. The system is applied to classify the dermoscopic images into normal and abnormal categories in the first stage and then classified into benign or malignant in the second stage. The various details about the databases are listed in Tab. 1.

Table 1: Basic information about the images used in this study

Description	Parameters
Name of the database	PH^2
#Normal images	80
#Benign images	80
#Malignant images	40
#Resolution	768×560
#Type of images	RGB

The accuracy (A_c) of the system can be broken into two important measures; sensitivity (S_n) and specificity (S_p). Before defining these two terms, four more variables are to be computed from the SMC system's outputs. When the system correctly identifies a positive result, it is referred to True Positive (TP) and if the system incorrectly identifies a positive result, it is referred as False Positive (FP). Similarly, two more terms can be defined and referred to True Negative (TP) and False Negative (FN) for identifying negative results. The definition of sensitivity and specificity are as follows:

$$S_n = \frac{TP}{TP + FN} \tag{10}$$

$$S_p = \frac{TN}{TN + FP} \tag{11}$$

$$A_c = \frac{TP + TN}{TP + FN + TN + FP} \tag{12}$$

There is no misclassification for a perfect system, which means that sensitivity and specificity will both be 100%. A high sensitivity measure can lead to a decrease the mortality rate. These measures are computed using k-fold (10-fold) cross-validation testing scheme where the classifier uses k-1 folds in training, and the remaining fold is tested.

As the GARCH features are extracted from DWT with many resolution levels of decomposition, the local properties and DRLBP are first analyzed independently. All normal images are considered a group of negative samples and abnormal images as positive samples in this stage. Then, k-fold cross-validation is employed for splitting images into these two groups for training and testing purposes. The confusion matrices obtained from these features are given in Fig. 4.

(a)

(b)

Figure 4: Confusion matrices of the 1st stage SMC system (a) local properties (b) DRLBP

It is evident from the Fig. 4 that the DRLBP features provide better performance than local properties of dermoscopic images. This is because DRLBP extracts spatial patterns that are available in the different classes of images effectively. It increases the performance by ~10% than the performance of local properties in terms of sensitivity, specificity and accuracy. The GARCH parameters are extracted from different DWT Levels (DWT-L), and their performances are analyzed. Tab. 2 shows the performance of the 1st stage SMC system using GARCH features with DWT levels.

Table 2: Performance of 1st stage SMC system using GARCH features with DWT

DWT-L	Performance measures						
	TP	FN	TN	FP	S_n	S_p	A_c
1	100	20	70	10	83.33	87.50	85.00
2	103	17	72	8	85.83	90.00	87.50
3	109	11	76	4	90.83	95.00	92.50
4	105	15	73	7	87.50	91.25	89.00

It is observed that over 90% of accuracy is obtained by GARCH parameters extracted from the sub-bands of 3rd level DWT. It is well known that more information can be obtained when increasing the resolution levels. However, the features obtained from higher resolution levels reduce the system's accuracy due to the redundant data that can be seen at 4th level DWT features. Also, it is evident from Tab. 2 and Fig. 4, the GARCH features have better performance than others.

Applying the 1st stage SMC system is reasonable while using the features independently but insufficient in the medical field that requires more accuracy to decrease the mortality. The redundant features in each group, which affects the performance, are eliminated by a feature reduction approach to obtain more accuracy. Tab. 3 shows the performance of the 1st stage SMC system after the feature reduction approach.

After feature reduction, the highest performance is 99.17% sensitivity and 100% specificity for 2% selected features. With more features, both performance measures are reduced and thus the system select only 2% features from the combination of features as the best set to classify abnormal images. Fig. 5 shows the SMC system's accuracy for all possible feature set used in the 1st stage. Fig. 6 Confusion matrices of 2nd stage SMC system obtained using local properties and DRLBP respectively.

Table 3: Performance of the 1st stage SMC system using after feature reduction

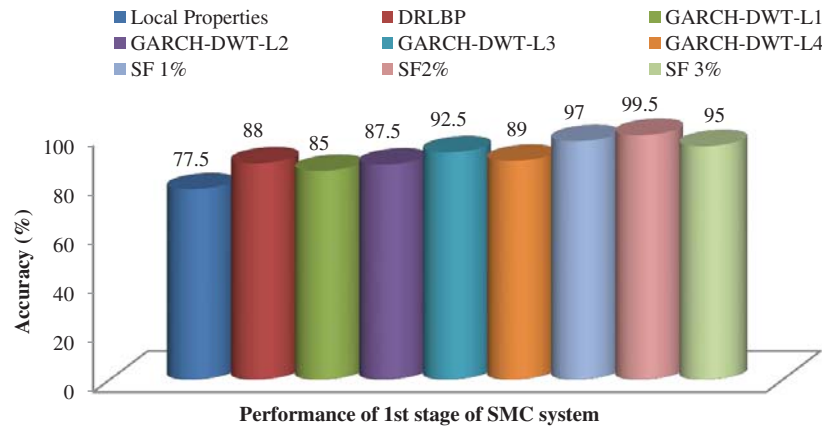| SF (%)	Performance measures						
	TP	FN	TN	FP	S_n	S_p	A_c
1	115	5	79	1	95.83	98.75	97.00
2	119	1	80	0	99.17	100	99.50
3	112	8	78	2	93.33	97.50	95.00

Figure 5: Accuracy of the SMC system for all possible feature set used in the 1st stage

Figure 6: Confusion matrices of the 2nd stage SMC system (a) local properties (b) DRLBP

It is evident from the 2nd stage SMC system; DRLBP has a maximum specificity of 86.8% and sensitivity of 92.5%. Tab. 4 shows the performance of the 2nd stage SMC system using GARCH features with DWT levels.

The best features which perform better than any other GARCH features are extracted from the 3rd level. The sensitivity of 3rd level GARCH features is increased ~5% than other features. Tab. 5 shows the 2nd stage SMC system's performance after feature reduction approach.

Table 4: Performance of the 2nd stage SMC system using GARCH features with DWT

DWT-L	Performance measures						
	TP	FN	TN	FP	S_n	S_p	A_c
1	35	5	72	8	87.50	90.00	89.17
2	36	4	75	5	90.00	93.75	92.50
3	38	2	77	3	95.00	96.25	95.83
4	36	4	74	6	90.00	92.50	91.67

Table 5: Performance of the 2nd stage SMC system using after feature reduction

SF-%	Performance measures						
	TP	FN	TN	FP	S_n	S_p	A_c
1	37	3	79	1	92.50	98.75	96.67
2	40	0	80	0	100.00	100.00	100.00
3	36	4	77	3	90.00	96.25	94.17

After feature reduction, the best performing features for 100% sensitivity and specificity are 2% features from the feature reduction approach. Fig. 7 shows the SMC system's accuracy for all possible feature set used in the 2nd stage.
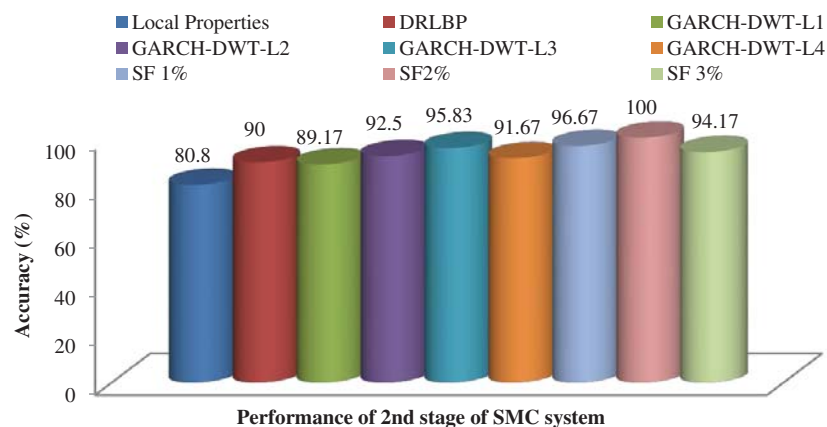

Figure 7: Accuracy of the SMC system for all possible feature set used in the 2nd stage

To visually analyze the SMC system, which uses three types of features and a combination of these features, ROC is used. Fig. 8 shows the ROCs of 1st and 2nd stage of SMC system. The classification is significantly better for the best-selected features (2% features) than others in both stages. The areas under the curve for the 1st stage and 2nd stage of SMC systems are 0.775 & 0.808 (Local properties), 0.88 & 0.9 (DRLBP), 0.925 & 0.9583 (GARCH of 3rd Level of DWT) and 0.99 & 1 (Best Selected Features) respectively.
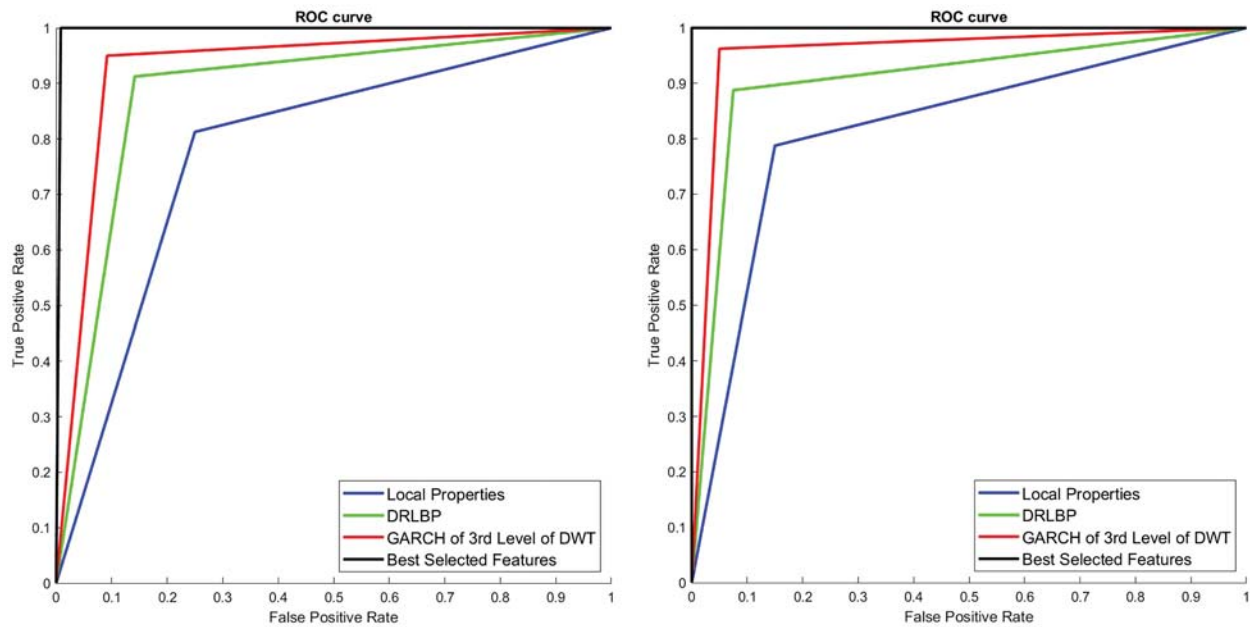
Figure 8: ROCs of 1st and 2nd stage SMC system

4 Conclusion

An efficient SMC system which combines segmentation, feature extraction, feature reduction and classification stages into one automated operation is developed and investigated for skin cancer diagnosis. The use of local properties, GARCH parameters from 3rd DWT level sub-bands and DRLBP to classify skin melanoma images is tested. Deep learning is tested using PH^2 database images and gives almost near-ideal system performance in terms of accuracy, sensitivity and specificity. Also, it is found that GARCH modelling can indeed be used for skin cancer diagnosis, and there are indeed performance differences in these features. The sensitivity of 1st stage and 2nd stage of the SMC system are 99.17% and 100% respectively, with all normal images are perfectly classified. The greatest degree of success is achieved on PH^2 database images using two stages of deep learning. It can be used as a pre-screening tool as it provides 100% accuracy for melanoma cases.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] W. Stolz, A. Riemann and A. Cognetta, "ABCD rule of dermatoscopy: A new practical method for early recognition of malignant melanoma," *European Journal of Dermatology*, vol. 4, no. 7, pp. 521–527, 1994.

[2] G. Argenziano, G. Fabbrocini and P. Carli, "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," *Architecture of Dermatology*, vol. 134, no. 12, pp. 1563–1570, 1998.

[3] H. Pehamberger, A. Steiner and K. Wolff, "In vivo epiluminescence microscopy of pigmented skin lesions—I: Pattern analysis of pigmented skin lesions," *Journal of American Academy of Dermatology*, vol. 17, no. 4, pp. 571–583, 1987.

[4] O. Abuzaghleh, B. D. Barkana and M. Faezipour, "Automated skin lesion analysis based on color and shape geometry feature set for melanoma early detection and prevention," in *Proc. of IEEE Long Island System and Application Technology Conf.*, Farmingdale, NY, USA, pp. 1–6, 2014.

[5] I. Maglogiannis and C. Doukas, "Overview of advanced computer vision systems for skin lesions characterization," *IEEE Transformation of Information Technology Biomedical*, vol. 13, no. 5, pp. 721–733, 2009.

[6] M. Celebi, H. Kingravi, B. Uddin, H. Iyatomi, Y. Aslandogan *et al.,* "A methodological approach to the classification of dermoscopy images," *Computing Medical Image Graphics*, vol. 31, no. 6, pp. 362–373, 2007.

[7] R. Haralick, "Statistical and structural approaches to texture," *Proc. of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[8] K. I. Laws, "Rapid texture identification," *Image Processing for Missile Guidance*, vol. 238, pp. 376–381, 1980.

[9] T. Ojala, M. Pietikäinen and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[10] R. Walvick, K. Patel, S. Patwardhan and A. Dhawan, "Classification of melanoma using wavelet-transform-based optimal feature set," *Proc. of SPIE Medical Image: Image Processing*, vol. 5370, pp. 944–951, 2004.

[11] S. Kumarapandian, "Melanoma classification using multiwavelet transform and support vector machine," *International Journal of MC Square Scientific Research*, vol. 10, no. 3, pp. 1–7, 2018.

[12] R. Sonia, "Melanoma image classification system by NSCT features and Bayes classification," *International Journal of Advances in Signal and Image Sciences*, vol. 2, no. 2, pp. 27–33, 2016.

[13] D. Donoho and E. Candes, "Continuous curvelet transform: II Discretization and frames," *Applied and Computational Harmonic Analysis*, vol. 19, no. 2, pp. 198–222, 2005.

[14] G. Easley, D. Labate and W. Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 25–46, 2008.

[15] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

[16] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder *et al.,* "Automated melanoma recognition," *IEEE Transaction of Medical Imaging*, vol. 20, no. 3, pp. 233–239, 2001.

[17] M. Wiltgen, A. Gergerb and J. Smolle, "Tissue counter analysis of benign common nevi and malignant melanoma," *International Journal of Medical Information*, vol. 69, no. 1, pp. 17–28, 2003.

[18] A. Mahbod, G. Schaefer, C. Wang, R. Ecker and I. Ellinge, "Skin lesion classification using hybrid deep neural networks," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Brighton, United Kingdom, pp. 1229–1233, 2019.

[19] M. A. Albahar, "Skin lesion classification using convolutional neural network with novel regularizer," *IEEE Access*, vol. 7, pp. 38306–38313, 2020.

[20] D. Moldovan, "Transfer learning based method for two-step skin cancer images classification," in *E-Health and Bioengineering Conf.*, Iasi, Romania, pp. 1–4, 2019.

[21] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proc. of the Eighteenth Annual ACM-SIAM Symp. on Discrete Algorithms*, New Orleans, Louisiana, pp. 1027–1035, 2007.

[22] K. Hackmack, F. Paul, M. Weygandt, C. Allefeld and J. D. Haynes, "Multi-scale classification of disease using structural MRI and wavelet transform," *Neuroimage*, vol. 62, no. 1, pp. 48–58, 2012.

[23] H. Kalbkhani, M. G. Shayesteh and B. Zali-Vargahan, "Robust algorithm for brain magnetic resonance image (MRI) classification based on GARCH variances series," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 909–919, 2013.

[24] R. Mehta and K. Egiazarian, "Dominant rotated local binary patterns (DRLBP) for texture classification," *Pattern Recognition Letters*, vol. 71, no. 2, pp. 16–22, 2016.

[25] C. Srinivasan, S. Dubey and T. R. Ganeshbabu, "Complex texture features for glaucoma diagnosis using support vector machine," *International Journal of MC Square Scientific Research*, vol. 7, no. 1, pp. 81–92, 2015.

[26] S. Murugan, A. Bhardwaj and T. R. Ganeshbabu, "Object recognition based on empirical wavelet transform," *International Journal of MC Square Scientific Research*, vol. 7, no. 1, pp. 74–80, 2015.

[27] W. Zhu, X. Wang, Y. Ma, M. Rao, J. Glimm *et al.,* "Detection of cancer-specific markers amid massive mass spectral data," *Proc. of the National Academy of Sciences of the United States of America*, vol. 100, no. 25, pp. 14666–14671, 2003.

[28] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.

[29] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal and J. Rozeira, "PH^2-A dermoscopic image database for research and benchmarking," in *35th Annual Int. Conf. on Engineering in Medicine and Biology Society*, Osaka, Japan, pp. 5437–5440, 2013.

[30] PH^2 Database Link: [Online]. Available: https://www.fc.up.pt/addi/ph2%20database.html.