Biomedical Signal Processing and Control 28 (2016) 27–40

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

A novel approach to multiclass psoriasis disease risk stratification: Machine learning paradigm

Vimal K. Shrivastava [a], Narendra D. Londhe [a,c], Rajendra S. Sonawane [b], Jasjit S. Suri [c,d,*]

[a] Electrical Engineering Department, National Institute of Technology, Raipur, India
[b] Psoriasis Clinic and Research Centre, Psoriatreat, Pune, Maharashtra, India
[c] Skin Point-of-Care Division, Global Biomedical Technologies, Inc., Roseville, CA, USA
[d] Electrical Engineering Department, Idaho State University (Aff.), Pocatello, ID, USA

ARTICLE INFO

Article history:
Received 27 October 2015
Received in revised form 29 March 2016
Accepted 11 April 2016
Available online 30 April 2016

Keywords:
Dermatology
Psoriasis skin disease
Color features
Texture features
Machine learning
Multiclass

ABSTRACT

The stage and grade of psoriasis severity is clinically relevant and important for dermatologists as it aids them lead to a reliable and an accurate decision making process for better therapy. This paper proposes a novel psoriasis risk assessment system (pRAS) for stratification of psoriasis severity from colored psoriasis skin images having Asian Indian ethnicity. Machine learning paradigm is adapted for risk stratification of psoriasis disease grades utilizing offline training and online testing images. We design four kinds of pRAS systems. It uses two kinds of classifiers (support vector machines (SVM) and decision tree (DT)) during training and testing phases and two kinds of feature selection criteria (Principal Component Analysis (PCA) and Fisher Discriminant Ratio (FDR)), thus, leading to an exhaustive comparison between these four systems.

Our database consisted of 848 psoriasis images with five severity grades: healthy, mild, moderate, severe and very severe, consisting of 383, 47, 245, 145, and 28 images respectively. The pRAS system computes 859 colored and grayscale image features. Using cross-validation protocol with *K*-fold procedure, the pRAS system utilizing the SVM with FDR combination with combined color and grayscale feature set gives an accuracy of 99.92%. Several performance evaluation parameters such as: feature retaining power, aggregated feature effect and system reliability is computed meeting our assumptions and hypothesis. Our results demonstrate promising results and pRAS system is able to stratify the psoriasis disease.

1. Introduction

Psoriasis is a chronic skin disease and characterized as a reddish and scaly lesion on the skin surface [1]. Statistics show that it affects 125 million people worldwide [2], but its prevalence differs depending on the geographical regions. The prevalence of psoriasis in Europe, USA, Malaysia and India is about 0.6–6.5% [2], 3.15% [2], 3% [3] and 1.02% [4], respectively. It may also affect the quality of life due to its discomforting social appearance [5]. Its consequences is increased risk of suicidal attempts (about 30%) making this disease an equally dangerous, at par with depression, heart disease, and diabetes [6]. The cause of psoriasis is still unknown but most researchers agree that the prime reason for this disease is genetics [7]. It is incurable disease but can be controlled by long and careful treatment. Among the various types of psoriasis namely plaque, guttate, inverse, pustular, and erythrodermic, plaque psoriasis is most commonly found in 80% of all psoriasis cases [8]. Thus, this study is focused on plaque psoriasis. An example of plaque psoriasis lesions on several regions of the human body is shown in Fig. 1.

Dermatologists are mainly interested in monitoring a stage and grade of psoriasis disease for betterment of patient's therapy. The current most widely used standard for measurement of psoriasis severity is "psoriasis area and severity index" (PASI) score [9]. This computation is based on the area, erythema, scaling and thickness of the lesion. This score is derived subjectively from the visual and haptic inspection of the dermatologist. This is critically based on the acquired expertise by the dermatologist. Both inspections suffer from inter- and intra-observer variability in severity scores. Hence, this process is very subjective, laborious, time consuming and brings unreliability to the process of decision making, leading to inaccurate. So, we here propose an objective computer-aided assessment system.

* Corresponding author at: Skin Point-of-Care Division, Global Biomedical Technologies, Inc., Roseville, CA, USA.
E-mail addresses: lky.vml@gmail.com (V.K. Shrivastava), nlondhe.ele@nitrr.ac.in (N.D. Londhe), drrajss@gmail.com (R.S. Sonawane), jsuri@comcast.net (J.S. Suri).

http://dx.doi.org/10.1016/j.bspc.2016.04.001
1746-8094/

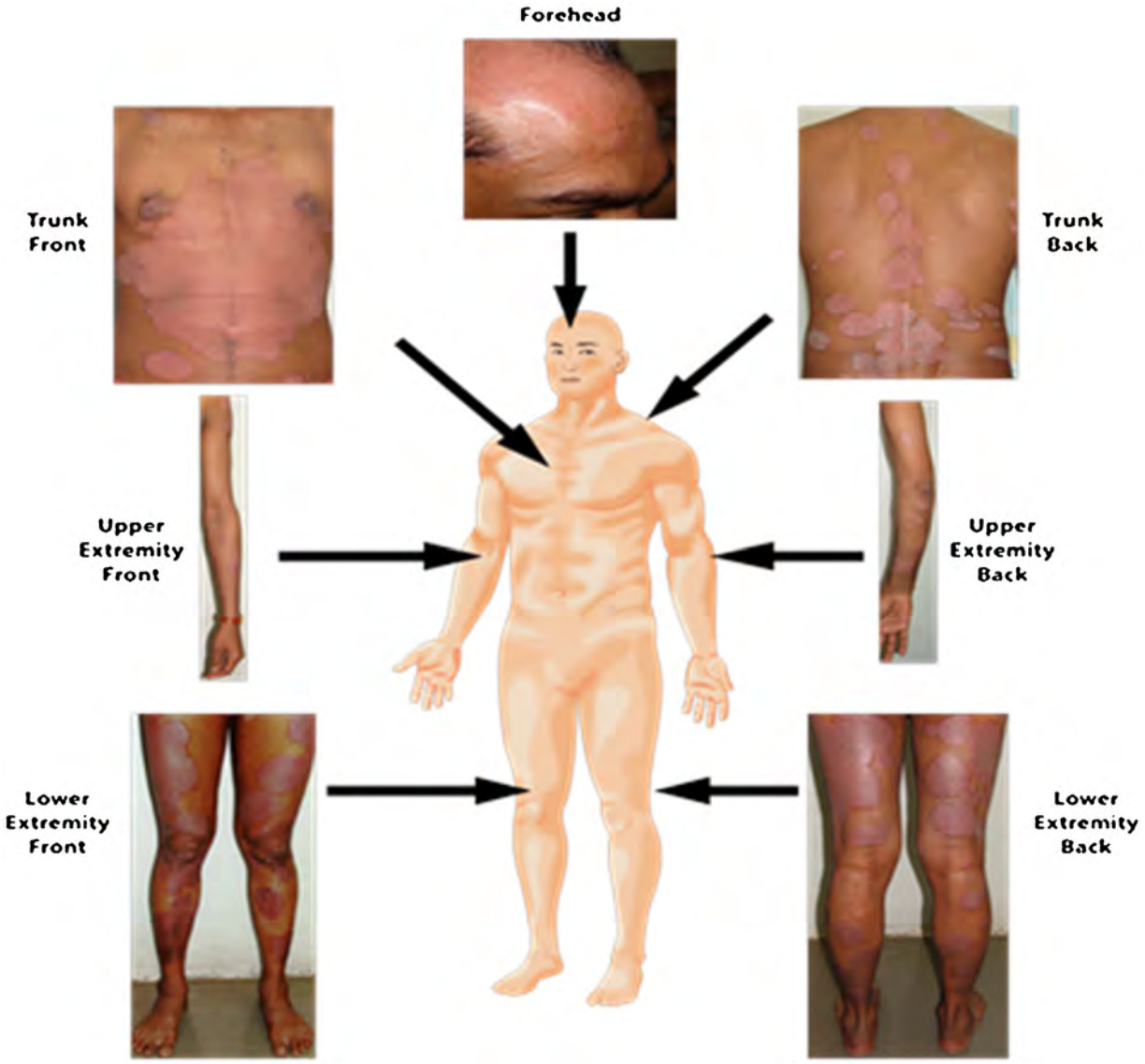

Fig. 1. Examples of plaque psoriasis disease on several body regions (psoriasis images—Courtesy of Psoriasis Clinic and Research Centre, Psoriatreat, Pune, Maharashtra, India; center body image copyright but courtsey from Global Biomedical Technologies, Inc., Roseville, CA, USA).

Biomedical computer-based methods have shown a promising sign in diagnosis and monitoring of diseases and in general healthcare. The current challenge lies under the same paradigm which can help dermatologists as a second guess for the disease severity evaluation and improving the work flow leading to better therapy. We thus propose a machine learning-based solution for severity risk assessment and stratification. To understand the severity of the psoriasis disease, dermatologists preferably analysis the lesion by its color, scales on the lesion surface and the lesion size or regional area spread by the lesion. The color of lesion varies from light red to dark red (purplish). The light red indicates improved condition of the disease while dark red indicates lesion severity [10]. So, hypothetically, color strength and its variations in a lesion are mainly adapted as a criterion for classification of psoriasis disease. Our first hypothesis is that there is strong color information in psoriasis images and machine learning model would be able to stratify the psoriasis disease by learning different color features.

Further, from the genetic nature of the disease, the severity starts to develop gradually by changing the lesion characteristics from grayscale to light red color to dark brown color (see Fig. 2). Such a skin characteristics can be reverse engineering modeled by converting the color lesions back to grayscale space and understanding their characteristics. Thus we hypothesis, that the psoriasis grayscale images have valuable lesion severity information helpful in risk stratification. Based on the hypothesis one and two, we further assume that the features of the psoriasis lesions by combining the grayscale features and color features can further lead to better decision making and risk assessment process. We therefore pose the third hypothesis that combined color and grayscale information can further improve the risk stratification. If our methods prove to be stratifying different grades of the lesions correctly, then the pRAS system can be proven to be reliable. We thus hypothesize if the reliability of the pRAS system is above a certain threshold, then; the pRAS system is able to correctly stratify the lesion grades.

The application of machine learning in skin cancer images has recently dominated. This comes from the spirit and initiatives of Suri and his team where machine learning was adapted dating back to early 90's [11–14]. The fundamental challenge in current psoriasis disease risk assessment systems is the lack of availability of risk stratification with varying degree of severity. Second serious issue is the lack of adaptation of risk assessment systems on multi-center

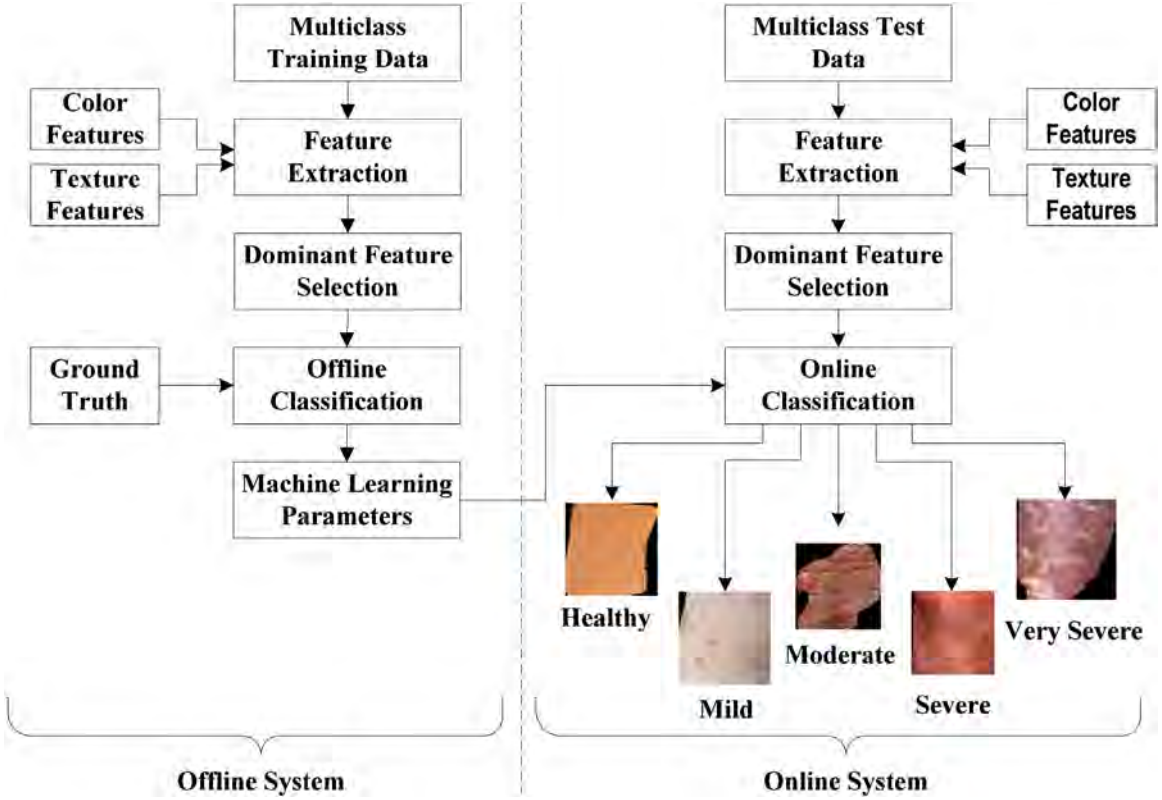

Fig. 2. Architecture of the proposed "pRAS" system.

clinical trials basis which have ability to handle large databases in an automated fashion. Third major issue is how to evaluate the risk assessment systems under the machine learning paradigms using patient images that are an amalgamation of color and grayscale at global level. Thus, the current published literature does not address any of these issues and the one which does, are very shallow, limited in technical merits and further, not applicable to psoriasis skin disease. We want to emphasize that psoriasis severity assessment technique of PASI is very time consuming and unreliable leading to subjective results and impossible to extend to large databases at a stage where it stands. Our recent review highlights the general framework of pRAS system [15]. Extending the above work of Suri's team, a two class psoriasis disease classification was presented in Refs. [16,17]. Detail analysis of the above methods will be presented in the benchmarking methods in the last section.

In this paper, we propose a psoriasis risk assessment system (pRAS) for stratification of psoriasis severity. Our database consists of 848 samples characterized into five grades or classes (hence the name multiclass) namely healthy, mild, moderate, severe and very severe on the basis of lesion severity. An expert dermatologist (RSS) with 27 years of experience used his visual judgment to stratify the lesions. Our research study is focused on lesion databases which are multiclass in nature. Thus, we are not comparing our paradigm and system design to past literature and studies which are developed for stratification into two classes (say normal and diseased).

The power of the color features is realized by taking into consideration four statistics each taken from 14 color spaces and each color space has three channels except one color space (CMYK) which has four channels, thus 172 color features has been extracted. Further, we have extracted 687 grayscale texture features utilizing nine grayscale texture feature extraction techniques. For simplicity here on, we call grayscale texture as texture features. Thus the combined feature set has: 859 features, most powerful set of system in psoriasis features. Thus the paper consists of three features set combinations (FC) and they are namely: FC1: color-172, FC2: texture-687, and FC3: color-texture-859. The numeric values at the end of FCs represent the number of features consists in that particular FC.

We summarize the novelties of the paper as under:

i. Generalized multiclass pRAS system on a multiclass colored psoriasis database for psoriasis severity risk stratification.
ii. Adaption of two kinds of machine learning classifiers in machine learning framework for psoriasis risk stratification: support-vector-machine (SVM) and decision tree (DT).
iii. Adaption of two kinds of feature selection techniques for choosing dominant features: Principal Component Analysis (PCA) and Fisher Discriminant Ratio (FDR) used in conjunction with two machine learning classifiers.
iv. Design and development of four kinds of machine learning systems which criss-crosses the above machine learning classifiers and feature selection methods. This constitutes 4 sets of pRAS systems for risk stratification: pRAS1: SVM-PCA, pRAS2: SVM-FDR, pRAS3: DT-PCA and pRAS4: DT-FDR.
v. Design of three set of comprehensive features, namely color (FC1), grayscale (FC2), and combined color-grayscale (FC3) as follows: FC1-172, FC2-687, and FC3-859.
vi. Unique Higher-Order-Spectrum (HOS) features extraction from individual color channel (R,G,B) of psoriasis images (never implemented before).
vii. Design of three kinds of performance evaluation techniques: reliability, feature retaining power and aggregate feature effect.
viii. Spirited towards completely automated system design for avoiding subjectivity.

2. Methods

The proposed pRAS system as shown in Fig. 2 works on the state of art machine learning paradigm in multiclass framework. The multiclass framework has the same spirit as two class paradigm [15–19], *i.e.*, the system consists of two phases segregated by the dotted line called as offline and online phases that perform training and testing paradigms, respectively. The key difference between the left and right sides is the training system based on the apriori ground truth information for the left and risk prediction on the right. Based on the rationale of hypothesis one and two discussed before, we build the feature extraction subsystem which involves the color feature and grayscale feature extraction process. The second fundamental block in the pipeline is the dimensionality reduction block, whose role is to select the dominant features and optimized features in multiclass framework. This is the first time a PCA-based and FDR-based framework is adapted and compared for psoriasis risk assessment in multiclass paradigm.

2.1. Data acquisition and preparation

The psoriasis image database required for this study was collected from Psoriasis Clinic and Research Centre, Psoriatreat, Pune, Maharashtra, India. Images were acquired by digitally photographing the patients of Indian ethnic origin using Sony NEX-5 camera with 22 mm lens and 350 dpi under the supervision of our dermatologist (RSS). The captured images were in Joint Photographic Expert Group (JPEG) format with color depth of 24 bits per pixel. The ethics committee approval was granted, informed consent was taken, and patient data was anonymized.

The preparation of database of healthy and diseased skin samples is done manually using a freehand image cropping tool in MATLAB. The steps followed were: (i) enlargement of an image to an optimum extends, (ii) cropping of lesion from its inner boundary to avoid false inclusion of background, (iii) cropped images were subdivided into five classes (grades) by our dermatologist on the basis of severity of the lesions. These were used as a ground truth for our pRAS system. Overall, 848 image samples were acquired from the images of 65 patients which are categorized into five grades as healthy (383), mild (47), moderate (245), severe (145) and very severe (28). Fig. 3 shows examples of sample images taken from five grades of lesion classes.

2.2. Color features

The color of the lesion is one of the criteria for dermatologists to decide the stage and grade of the psoriasis disease [9]. The stratification of lesion is judged on the color strength of that lesion. Since we hypothesize that color strength and its variations in a lesion are mainly adapted as a criterion for classification of psoriasis disease, we then assume that strong color information in psoriasis images and machine learning model would be able to stratify the psoriasis disease by learning different color features. Here, four statistics, *i.e.*, mean, standard deviation, skewness and kurtosis were calculated for each color channel in 14 color spaces [17]. We have used the following color spaces which are well available in industry: (i) RGB, (ii) normalized-RGB, (iii) YCbCr, (iv) HSV, (v) HSI, (vi) CIE XYZ, (vii) CIE Lab, (viii) CIE Lch, (ix) CIE Luv, (x) Hunter-Lab, (xi) SCT, (xii) opponent, (xiii) CMY and (xiv) CMYK. For the convenience of the readers, we have provided the mathematical formulations for conversion from RGB to remaining 13 color spaces in Supplementary material (Appendix A). As each color space contains three color channels except CMYK which has four color channels, thus, totaling to 172 color features ({13 color spaces × 3 color channels × 4 statistics} + {1 color spaces (CMYK) × 4 color channels × 4 statistics}) were extracted for each sample image. We will use these color features in machine learning pRAS system architecture of Fig. 2 corresponding in both unique training and testing data sets. The main novelty lies in the prediction of the lesion severity having multiple psoriasis grades using color feature spaces derived from these 14 color spaces.

2.3. Grayscale texture (GS) features

2.3.1. Statistical-based GS Features

Texture has proven to be one of the key components in medical imaging [12,13,16–19]. They have a strong ability to distinguish benign vs. malignant lesions [20]. Using the same spirit, our extension of texture is applied to the psoriasis framework. They key novelty here lies in the extraction of linear and non-linear texture features which actually evolves of the disease process. This novelty has extended our search to over eight Statistical-based grayscale texture spaces leading to 60 grayscale feature [16,21–24]. The most traditional and prominent texture analysis techniques are Gray-Level Co-occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM). We have extracted 18 features using GLCM and 11 features using GLRLM [21]. Other than these two well-known texture feature extraction techniques, Intensity Histogram (IH) [21], Invariant Moment (IM) [21], Gray Level Difference Statistics (GLDS) [22], Neighborhood Gray Tone Difference Matrix (NGTDM) [23], Statistical Feature Matrix (SFM) [24] and Fractal Dimension [16,25] have been used. These texture features have been described in detail in Supplementary material (Appendix B).

2.3.2. HOS-based GS texture features

Another technique based on higher order spectra (HOS) was utilized to analyze the random distribution of pixels of psoriatic lesions [26]. HOS are spectral representations of higher order cumulants among which bispectrum is one of the most commonly used for HOS features. Bispectrum is defined as the Fourier transform of the third-order cumulant sequence given by Eq. (1) [27]. But, before bispectrum computation, psoriasis color images were converted to grayscale images and then radon transformation [13] was performed to transform the 2D image into one dimensional signal at various angles. We have extracted the HOS features for every 10° of the Radon transform between 0 and 180° *i.e.*, for 19 angles.

$$B(\omega_1, \omega_2) = X(\omega_1)X(\omega_2)X^*(\omega_1 + \omega_2) \quad (1)$$

Eq. (1) represents the bispectrum of signal $x(k)$ for $|\omega_1| \leq \pi$, $|\omega_2| \leq \pi$ and $|\omega_1 + \omega_2| \leq \pi$. We have extracted 11HOS features in this paper namely; mean of spectral magnitude for power spectrum, mean of spectral magnitude for HOS, linear entropy, quadratic entropy, cubic entropy, bispectral phase entropy, sum of the logarithmic amplitudes of the bispectrum, sum of the logarithmic amplitudes of the diagonal elements in the bispectrum and first, second and third-order spectral moments of the amplitudes of the diagonal elements of the bispectrum [13,26]. These 11 features have been extracted from each color channel of RGB color space and for 19 different angles. In this way, a total of 627 (11 features × 3 color channels × 19 angles) HOS features have been extracted from each image. The derivation for bispectrum and HOS features has been provided in Supplementary material (Appendix C). Accumulating the total GS texture features derived from statistical-based features (60) and spectrum-based features (627), we have a total of 687 GS texture features. This is the largest feature set ever derived in grayscale framework for psoriasis risk stratification.

2.4. Feature selection

Due to extensive feature space of 859 features (172 color and 687 texture), there is a need for the data mining framework to be able to select the dominant features and avoids the redundant features.

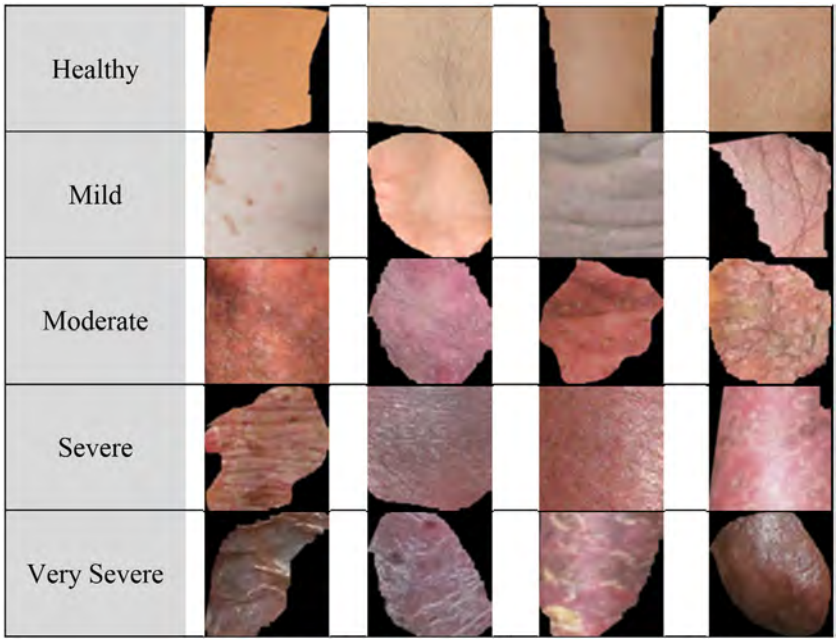

Fig. 3. Examples of sample images taken from five grades of lesion classes: normal, mild, moderate, severe, very severe.

It reduces the dimensionality of the feature space which results in high classification performance with the optimal number of features. In this paper, two feature selection techniques namely; PCA and FDR have been applied and the performance of pRAS system has been compared with both techniques. In PCA, the transformation of set of features to reduced number of uncorrelated features (principal components), is in such a way that the first principal component has the largest possible variance. But, since PCA is mainly used for dimensionality reduction and not able to perform feature selection, PCA with polling contribution is used in this research study [28]. Here, the cutoff factor (R) plays an important role in the selection of number of dominant features which extract the indices of the dominant features and the original feature values are never lost in the process of dominant feature selection.

In FDR, the objective is that the distances between data points in inter-classes should be as large as possible, while the distances between data points in the intra-class should be as small as possible [29]. The features were sorted in descending FDR value and a set of first m features were selected which have p-value of less than 0.0001 obtained using t-test.

2.5. Risk stratification

Classification is the final stage of pRAS system. The risk prediction involves learning based on different classification paradigms and testing based on similar classification methods. In a multiclass framework, our model of pRAS system adapted two different set of classifiers to understand how the feature selection criteria bridge with the training–testing paradigms. Our multiclass approach used two standardized approaches: one based on support vector machines (SVM) and second using a decision tree (DT), along the lines of previously adapted classifiers in binary framework [12,18,30–32]. The SVM is a state of the art classification technique based on maximum margin classifier [33]. It builds the separating hyper-plane on the basis of support vectors. Kernel functions can be used to map the input data to a high-dimensional space to classify the nonlinear data. A DT classifier uses trees built using the input features. A series of rules for the different classes derived from the built tree are used to predict the class label of the test data [30]. We have adapted "one-against-all" approach to implement multiclass problem for both the classifiers *i.e.*, SVM and DT. Further, a 10-fold cross validation protocol has been adapted in order to establish robust and generalized classification and we repeated this protocol to $T = 20$ number of trials to reduce the variability due to random partition in cross-validation protocol.

3. Results

We have designed four kinds of pRAS systems by using a combination of two classifiers (SVM and DT) with two feature selection techniques (PCA and FDR). They are namely: (1) pRAS1: SVM-PCA, (2) pRAS2: SVM-FDR, (3) pRAS3: DT-PCA and (4) pRAS4: DT-FDR. Further, three sets of feature combinations *i.e.*, (i) FC1: color-172; (ii) FC2: texture-687; (iii) FC3: color-texture-859; were utilized as an input feature sets sequentially for each of the pRAS systems. We have performed two experiments: (i) to analyze the system performance of pRAS systems for validation of hypothesis-1, 2 and 3 and (ii) to provide evidence of system consistency for validation of hypothesis-4. The list of symbols has been provided in Appendix D for more clarity to readers.

3.1. Performance analysis of pRAS system's for validation of hypothesis-1, 2 and 3

In this experiment, the data size of $N_p = 848$, has been kept fixed for all pRAS systems. The very first objective is the selection of the best SVM kernel function for each feature combination. The best SVM kernel function was chosen on the basis of classification accuracy among the five different conventional SVM kernels *i.e.*, linear, radial basis function (RBF) and polynomial kernel of order one, two and three. We observed that the polynomial kernel of order two showed the best classification accuracy for pRAS1: SVM-PCA system while polynomial kernel of order one produced best classification accuracy for pRAS2: SVM-FDR system. As a result, we fixed our analysis on subsequent experiments using this polynomial kernel of order two for pRAS1: SVM-PCA and polynomial kernel of order one for pRAS2: SVM-FDR.

Our next objective was to compare the performance of four pRAS systems for all feature combinations. Fig. 4 shows the classification accuracy for all four pRAS systems using all three feature

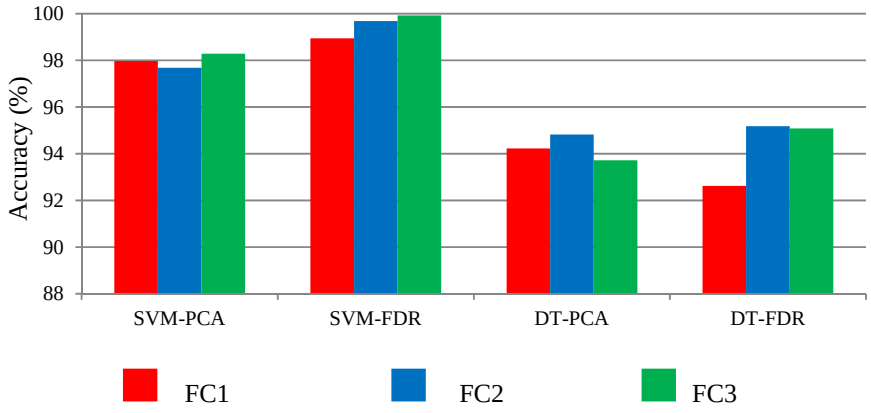

Fig. 4. Comparison between 4 different pRAS systems: pRAS1: SVM-PCA, pRAS2: SVM-FDR, pRAS3: DT-PCA and pRAS4: DT-FDR and 3 different feature combinations: (i) FC1: color-172; (ii) FC2: texture-687; (iii) FC3: color-texture-859, while keeping the data size fixed to 848.

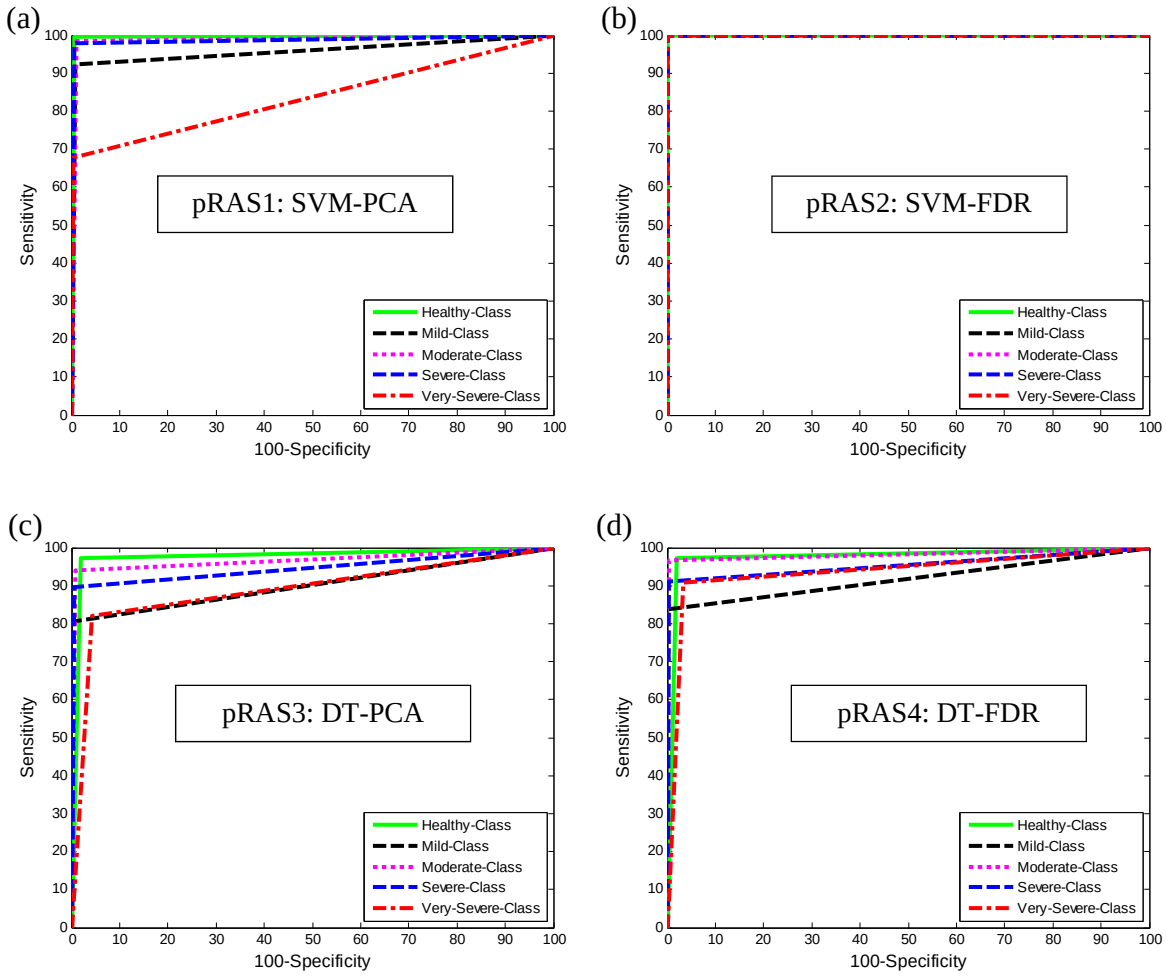

Fig. 5. ROC curves for all five classes using FC3 for different pRAS: (a) pRAS1: SVM-PCA (b) pRAS2: SVM-FDR (c) pRAS3: DT-PCA (d) pRAS4: DT-FDR.

combinations. It can be easily observed from Fig. 4, that the highest classification accuracy has been obtained using pRAS2: SVM-FDR system for all three feature combinations. Moreover, the feature combination FC3 has produced highest accuracy for two pRAS systems that has SVM-based classifier and FC2 has produced highest accuracy for other two pRAS systems that has DT-based classifier. Since the highest accuracy has been obtained with pRAS2: SVM-FDR system, we have shown other performance parameters *i.e.*, sensitivity (SE), specificity (SP), positive predictive value (PPV), individual class accuracy (Class-ACC) and area under the receiving operating characteristic curve (AUC) for this pRAS2: SVM-FDR system and using FC3 in Table 1. Moreover, the receiving operating characteristic (ROC) curves has been shown in Fig. 5 for all four pRAS systems using FC3 which again demonstrate that pRAS2: SVM-FDR system outperforms the other systems.

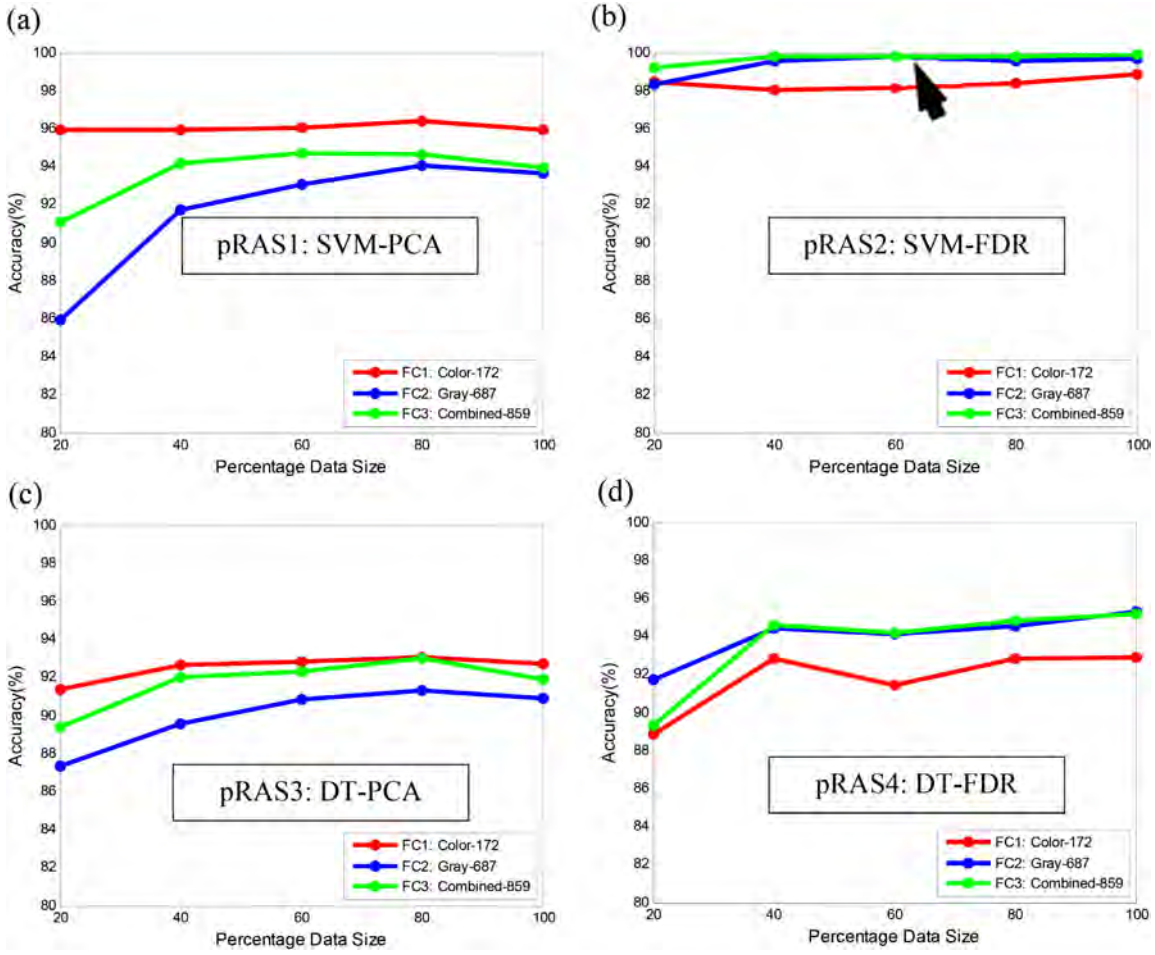

Fig. 6. Accuracy vs. percentage data size for different FCs and for different pRAS: (a) pRAS1: SVM-PCA (b) pRAS2: SVM-FDR (c) pRAS3: DT-PCA (d) pRAS4: DT-FDR.

Table 1
Best classification performance parameters for pRAS2: SVM-FDR system using FC3.

	SE (%)	SP (%)	PPV (%)	Class-ACC (%)	AUC
Healthy	99.80	100.00	100.00	99.91	0.999
Mild	100.00	100.00	100.00	100.00	1.000
Moderate	99.94	99.96	99.90	99.95	0.999
Severe	99.93	100.00	100.00	99.99	1.000
Very Severe	100.00	99.91	98.05	99.91	0.999

Table 2
Inter-system accuracy using all three FCs.

	pRAS1: SVM-PCA	pRAS2: SVM-FDR	pRAS3: DT-PCA	pRAS4: DT-FDR
FC1	96.06	98.40	92.54	91.77
FC2	91.70	99.38	90.00	94.02
FC3	93.72	**99.69**	91.72	93.62

The Bold value represents the highest system accuracy.

3.2. Evidence of system consistency for validation of hypothesis-4

The objective of this experiment was to analyze how the behavior of the pRAS machine learning systems is influenced by number of subjects (multi-class data size). We anticipate having a protocol which can determine the classification accuracy of multiclass lesions with increasing data size. We had derived five data sets and each data set took a pool of the following percentage from each class: $\mathbb{N} = 20\%$, 40%, 60%, 80% and 100%. This means the five data sets consisted of 170, 339, 509, 678 and 848 psoriasis samples. Note that each data size represented all five grades of psoriasis lesions. Fig. 6 illustrates the effect of increasing data size on the system performance (classification accuracy). Four pRAS systems were executed with all three features set (FC) combination. Following are our observations: (a) All 4 pRAS systems gave similar and consistent behavior for this machine learning protocol *i.e.*, with increasing sample size, the pRAS systems increase accuracy reach generalization stage from memorization; (b) All pRAS systems show FC3 as a best feature set except pRAS1: SVM-PCA and pRAS3: DT-PCA where FC1 has produced better accuracy than FC3 with the mean difference of nearly 3% and 1% respectively; (c) pRAS2: SVM-FDR system using FC3 (combined color and grayscale) shows the best performance among all the 4 pRAS systems.

For deeper understanding of feature sets for these 4 pRAS systems, we decided to demonstrate a complementary way of depiction in Fig. 7. This illustrates the system-wise comparison showing the classification accuracy with increasing data sizes for all 4 pRAS systems using all three FCs. Following were the observations: (a) the behavior of all four pRAS systems was consistent; (b) pRAS2: SVM-FDR has a larger improvement in combined color-grayscale feature set (FC3); (c) pRAS2: SVM-FDR produced the highest classification accuracy for all considered data sizes and for all three FCs. Moreover, system accuracy was computed by taking the average of accuracies corresponding to all data sizes as shown in Table 2.

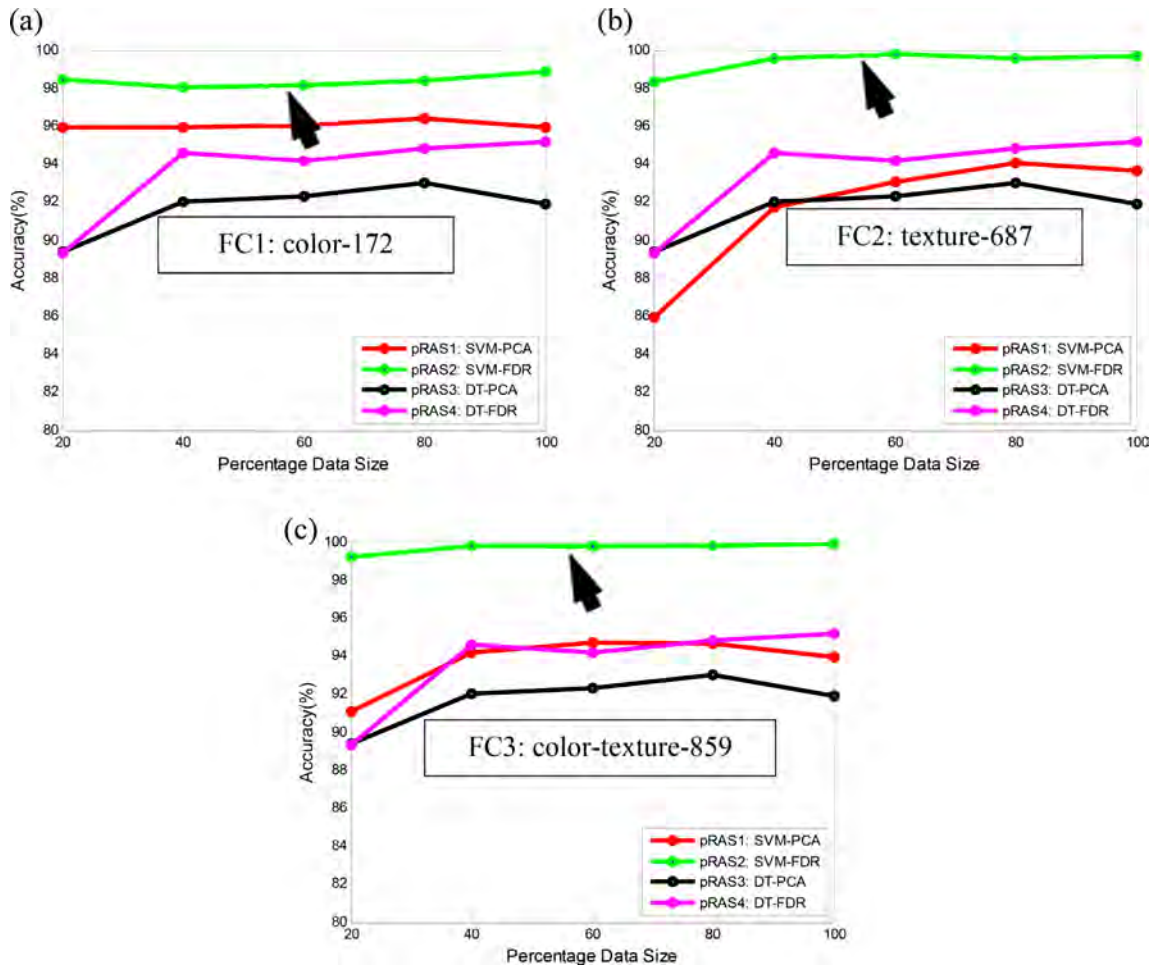

Fig. 7. Accuracy vs. percentage data size for different pRAS systems using different FCs: (a) FC1: color-172 (b) FC2: texture-687 (c) FC3: color-texture-859.

3.3. *Visual representation of the multi-grade lesions feature values*

We have shown the top 10 dominant features for five grades of psoriasis skin disease in Table 3. These feature values were calculated using pRAS2: SVM-FDR machine learning system utilizing the FC3 combination (consisting of grayscale and color gamut features). Note that ANOVA test was adapted for the analysis demonstrating that they are clinically significant with low p-values (<0.0001). As can be seen in the table, the top three features were: (i) skewness statistics from cyan color channel of CMYK color space; (ii) kurtosis statistics from cyan color channel of CMYK color space and (iii) mean value of hue color channel of CIE Lch color space. These features represent color information and this validates our assumptions (Hypothesis 1) that color features have ability to stratify the different grades of severity of the psoriasis skin disease. The top two features are actually representing the grayscale feature (as assumed using Hypothesis 2), which shows that the power to stratify also lies in the grayscale images when transformed from color framework. The GLCM feature information measure of correlation-1 which is a weighted version of the entropy measure shows here that randomness present is opposite with respect to the reference. This is demonstrating negative values which are not unusual for such a set up. Among the list of top 10 features, HOS is also listed which shows the spectrum behavior contributes in stratification, which is consistent to our previous publications [13,34,35]. As anticipated that the features do not have a very large absolute difference, and such fine line difference are anticipated in such classification techniques [31,35], however necessary to note that it requires the combination of features for stratification, and hence the usage of the 3rd assumption (Hypothesis 3) which allows us to merge the grayscale and color features leading to a better stratification.

3.4. *Performance evaluation*

An important characterization of the machine learning system for multi-grade lesion classification is to understand and evaluate its performance. We have introduced three novel methods strategies for evaluating its performance. They are: (i) reliability index, (ii) feature retaining power (FRP) and (iii) aggregate feature effect (AFE).

3.4.1. *Reliability index*

The concept of the reliability is based on the foundation taken from US food and drug regulatory foundation (www.fda.gov), which states that how the behavior of the system changes by addition of more population during the clinical trials. Thus, using this concept, we derive the reliability index by understanding the deviation of the classification accuracy with respect to its mean as more samples are added in the multi-grade pRAS system. The reliability index ζ_{N_i} is mathematically given as:

$$\zeta_{N_i}(\%) = \left(1 - \frac{\sigma_{N_i}}{\mu_{N_i}}\right) \times 100 \tag{2}$$

where σ_{N_i} and μ_{N_i} represent standard deviation and mean of all accuracies for data size N_i. A comprehensive analysis of the relia-

Table 3
Mean ± standard deviation values of the top 10 dominant features selected by pRAS2: SVM-FDR system using FC3 for all five classes.

Dominant features	Feature category	Healthy	Mild	Moderate	Severe	Very severe	*p*-value (ANOVA test)
Information measure of correlation-1	Texture (GLCM)	-0.35 ± 5.94	-0.82 ± 0.44	-0.86 ± 1.62	-0.88 ± 0.68	-0.87 ± 0.06	< 0.0001
Skew invariant moment (I7)	Texture (IM)	$0.08 \pm 1.26 \times 10^{-13}$	$2.3 \pm 12.1 \times 10^{-13}$	$0.08 \pm 4.91 \times 10^{-13}$	$-5.85 \pm 3.54 \times 10^{-13}$	$-1.11 \pm 4.45 \times 10^{-13}$	< 0.0001
Skewness-cyan	Color (CMYK)	2.25 ± 0.56	2.03 ± 0.57	1.97 ± 1.20	2.15 ± 0.51	2.27 ± 0.32	< 0.0001
Kurtosis-cyan	Color (CMYK)	6.51 ± 3.57	5.34 ± 2.56	6.21 ± 14.03	5.86 ± 2.27	6.35 ± 1.69	< 0.0001
Mean-Hue	Color (CIE Lch)	67.66 ± 8.04	51.49 ± 10.73	69.10 ± 55.42	50.77 ± 42.14	67.03 ± 62.03	< 0.0001
H4-G-160[a]	Texture (HOS)	$1.86 \pm 1.77 \times 10^{9}$	$1.49 \pm 1.06 \times 10^{9}$	$1.30 \pm 0.73 \times 10^{9}$	$1.27 \pm 0.79 \times 10^{9}$	$1.31 \pm 0.70 \times 10^{9}$	< 0.0001
H3-B-150	Texture (HOS)	$1.60 \pm 0.77 \times 10^{9}$	$1.48 \pm 0.93 \times 10^{9}$	$1.21 \pm 0.77 \times 10^{9}$	$1.07 \pm 1.23 \times 10^{9}$	$1.10 \pm 0.83 \times 10^{9}$	< 0.0001
H5-G-150	Texture (HOS)	$1.87 \pm 1.0 \times 10^{9}$	$1.47 \pm 0.94 \times 10^{9}$	$1.27 \pm 0.800 \times 10^{9}$	$1.19 \pm 0.91 \times 10^{9}$	$1.26 \pm 0.70 \times 10^{9}$	< 0.0001
$M_{avg\text{-}HOS}$-R-160	Texture (HOS)	$2.25 \pm 1.16 \times 10^{9}$	$1.94 \pm 0.86 \times 10^{9}$	$1.57 \pm 0.96 \times 10^{9}$	$1.69 \pm 0.91 \times 10^{9}$	$1.69 \pm 0.60 \times 10^{9}$	< 0.0001
E2-G-160	Texture (HOS)	$1.93 \pm 0.74 \times 10^{9}$	$1.53 \pm 0.89 \times 10^{9}$	$1.30 \pm 0.81 \times 10^{9}$	$1.26 \pm 1.03 \times 10^{9}$	$1.36 \pm 0.62 \times 10^{9}$	< 0.0001

[a] H4-G-160: here, H4 represents one of the HOS features (please see Supplementary material, Appendix C), G represents feature extracted from green color channel of RGB color space and 160 indicates the angle of radon transform.

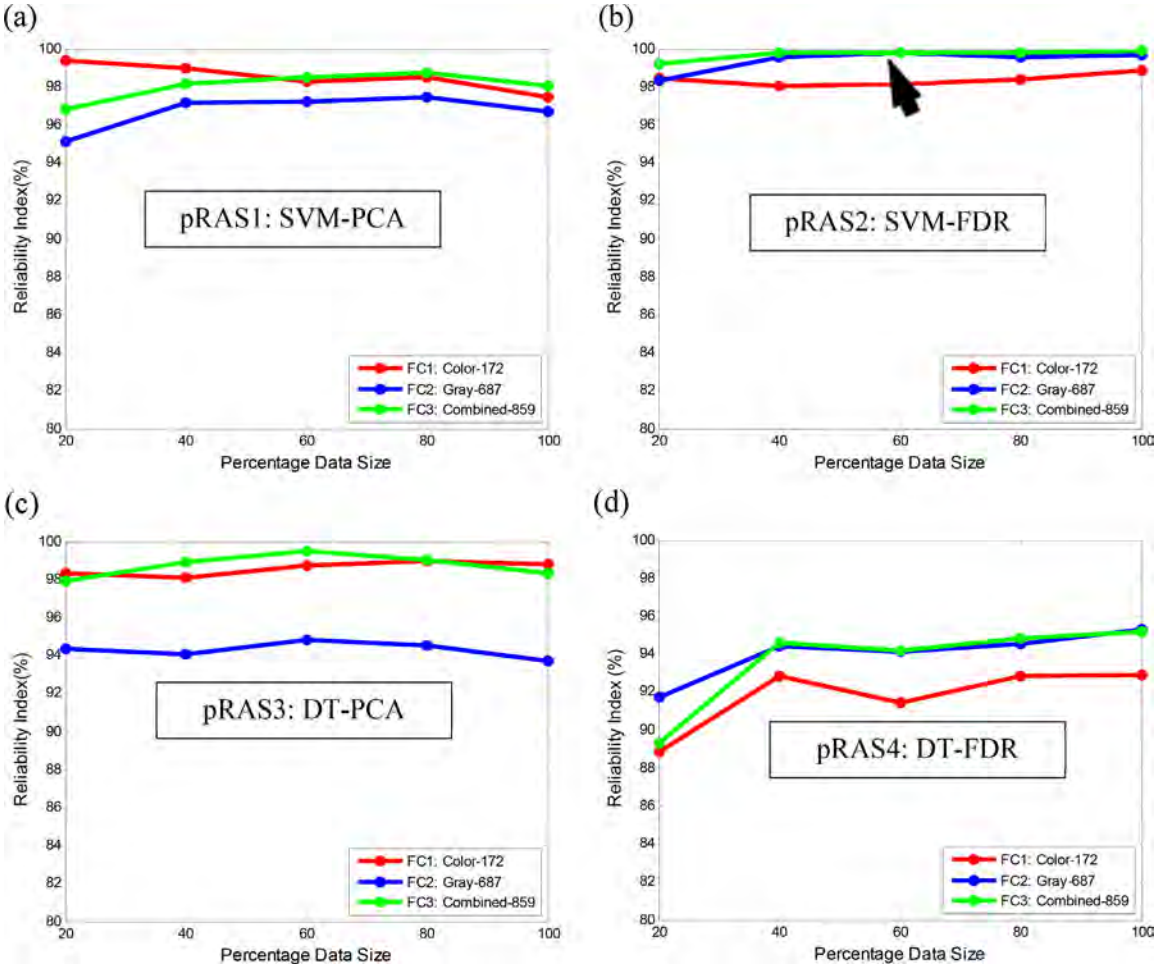

Fig. 8. Reliability Index vs. percentage data size for different FCs and for different pRAS: (a) pRAS1: SVM-PCA (b) pRAS2: SVM-FDR (c) pRAS3: DT-PCA (d) pRAS4: DT-FDR.

bility system was attempted. Fig. 8 shows the reliability index for all 4 pRAS systems using all three FCs as the data size increases and demonstrates that the pRAS2: SVM-FDR using FC3 has produced highest reliability throughout the considered data sizes. Further, we have calculated the system reliability by taking the average of reliability indexes for each pRAS system as shown in Table 4.

3.4.2. Feature retaining power

The second index of performance evaluation is to understand the effect of PCA cutoff on the machine learning systems. PCA-based dominant feature selection allows selecting the best feature and retaining them as well. Thus, the ability of the pRAS system to retain the features with increase in PCA cutoff values can be used as a criterion for performance evaluation. This ability is mathematically expressed as feature retaining power (FRP) and is given as:

$$\mathrm{FRP}(\%) = \left(\frac{\mathrm{SF}_{i-j}}{\#\ \text{Features in } R_i} \right) \times 100 \tag{3}$$

Table 4
Inter-system reliability using all three FCs.

	pRAS1: SVM-PCA	pRAS2: SVM-FDR	pRAS3: DT-PCA	pRAS4: DT-FDR
FC1	98.51	99.68	98.59	98.11
FC2	96.73	99.39	94.30	98.57
FC3	98.05	**99.73**	98.74	97.40

The Bold value represents the highest system accuracy.

Table 5
Feature retaining power (FRP) calculation for different cutoffs (R) using FC3.

Inputs for FRP test $F_{R_{i-j}}$	# Features in R_i	# Features in R_j	Similar Features	SF_{i-j}	FRP (%)
$F_{R_{1-2}}$	33	36	30		90.91
$F_{R_{2-3}}$	36	39	31		86.11
$F_{R_{3-4}}$	39	44	30		76.92
$F_{R_{4-5}}$	44	48	38		86.36
$F_{R_{5-6}}$	48	55	37		77.08
$F_{R_{6-7}}$	55	63	44		80.00
$F_{R_{7-8}}$	63	74	53		84.13
$F_{R_{8-9}}$	74	92	70		94.59
$F_{R_{9-10}}$	92	126	86		93.48

where, SF represents the number of similar features for consecutive cutoffs, R represents the PCA cutoff. Among the two consecutive cutoffs, the first and second cutoff is indicated by i and j, respectively. Thus the goal is to estimate how many features are retained at each consecutive PCA cutoff. Since the feature selection technique has been applied prior to the classifier, its output does not depend on classifier. Further, since PCA-based feature selection is applied to only 2 systems (pRAS1: SVM-PCA and pRAS3: DT-PCA), so we show the FRP for these systems only. A representative example is shown in Table 5. The inputs for FRP test are $F_{R_{i-j}}$ which represents features at consecutive cutoffs $\left(R_{i-j}\right)$. For example, $F_{R_{1-2}}$ denotes the features at cutoff R_1 (0.90) and R_2 (0.91). R represents cutoff varying from 0.90 to 0.99 in the step size of 0.01. Finally, the average of FRP of all PCA cutoffs has been calculated and it is found to be 85.51% using FC3.

The concept of FDR is based on p-value selection for the dominant features and two of our pRAS systems (pRAS2: SVM-FDR and pRAS4: DT-FDR) are based on FDR, unlike PCA. For these two systems the variability and stability criteria is actually jugged around the central selection point which is extraction based on the p-value. We thus choose the closest 10 values around the central point, typically adapted. These 10 points are ±five values (±15%, ±35%, ±55%, ±75%. and ±95%) around 0.0001 as considered by FDR. This complements the cutoffs of PCA in pRAS1: SVM-PCA and pRAS3: DT-PCA. Our observations demonstrate that the FRP for all consecutive p-values were nearly 100%. This is because the dominant features which are selected by FDR are always retained, even though new features were added. This ensures the stability of the system. This observation was for both the pRAS systems (pRAS2: SVM-FDR and pRAS4: DT-FDR).

3.4.3. Aggregate feature effect

The third and the final innovation of our pRAS evaluation is via understanding the net effect of change in machine learning classification accuracy as number of dominant features (D) increases with change in the consecutive PCA-based cutoff values in two of the pRAS systems or in FDR-based variation near the p-value selection. This is mathematically expressed as:

$$\text{AFE}\,(\%) = \left(100 - \frac{\text{ACC}_j - \text{ACC}_i}{\text{ACC}_i}\right) \times 100 \tag{4}$$

where ACC_i and ACC_j represent the pRAS system classification accuracies for two consecutive cutoffs (PCA) or p-values (FDR). Fig. 9 demonstrates the AFE with the increase in the number of dominant features due to change in PCA cutoffs or FDR p-values variation. As the number of D increases, the AFE for all 4 pRAS systems increases showing the consistency and stability. We demonstrated for FC3 combination as earlier we had shown that FC3 is the best feature combination. Further, note that the AFE reaches the point of diminishing returns under the current pool of multi-grade stratification using the current setup.

4. Discussion

4.1. Our system

The paper presented a risk stratification system (pRAS) on the basis of five grades of risk severity such as: healthy, mild, moderate, severe and very severe. The system design was based on three assumptions utilizing tissue characteristics. It used three set of feature combinations: color features, grayscale features and combined color and grayscale features namely: FC1, FC2 and FC3, consisting of 172 color features, 687 grayscale features and 859 combined color and grayscale features. Using the above gamut of features, we developed a foundational tissue characterization system that would learn using the training system for learning and predict the risk on the test images. Four kinds of pRAS systems were developed using two sets of classifiers (SVM and DT) and two sets of feature selection techniques (PCA and FDA) namely: pRAS1: SVM-PCA, pRAS2: SVM-FDR, pRAS3: DT-PCA and pRAS4: DT-FDR.

In case of SVM, best SVM kernel has been selected among five SVM kernel functions on the basis of classification accuracy. To select the number of dominant features in case of PCA, 10 cutoffs ranging from 0.90 to 0.99 with an interval of 0.01 has been considered while in case of FDR, t-test has been applied and p-value of less than 0.0001 has been kept as a threshold. We have adapted "one-against-all" approach to implement multiclass problem for both the classifiers *i.e.*, SVM and DT. Further, in order to establish robust and generalized classification, we adapted a 10-fold cross validation protocol for training and testing phases. Since the partitions are random, we therefore repeated this protocol to $T = 20$ number of trials. The higher success rate of such a set-up is attributed due to the unique feature set and the combination of best selected features for our pRAS systems. We have performed two experiments. First experiment demonstrates the performance analysis of pRAS system's for validation of hypothesis-1, 2 and 3 while showing the performance parameters such as accuracy, sensitivity, specificity and AUC. Second experiment provides evidence of system consistency for validation of hypothesis-4 and showed the protocol of classification accuracy while increasing the population of the data size to establish the high reliability index and stability of the system. We also introduced the FRP and AEF-new concept for evaluating the pRAS systems. The paper showed four machine learning systems and solid comparisons were established.

4.2. Brief survey

It is observed that the psoriasis public data bases are not prevalent and there is not much literature available with machine learning paradigm specifically for psoriasis disease. Hence, we have presented a comparative performance of proposed model against existing works on skin cancer classification in multiclass scenario as shown in Table 6. Further, studies which deal with two class stratification of dermatology images [16,17,36–41] are not included in our current survey. Shimizu et al. [42] classified dermoscopy images into melanoma, nevus, basal cell carcinoma and seborrheic kerato-

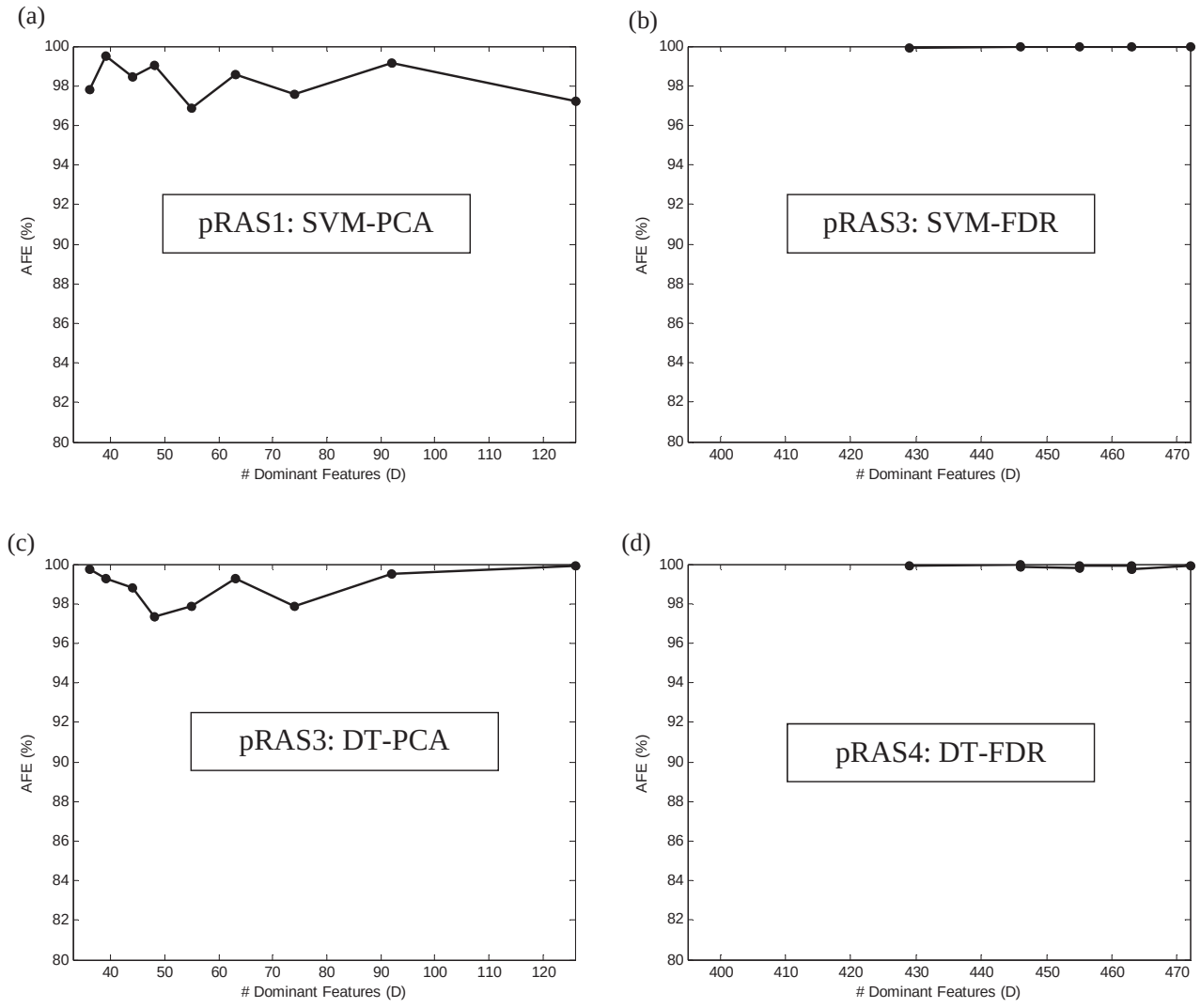

Fig. 9. Aggregate feature effect (AFE) vs. # dominant features (*D*) using FC3 for different pRAS: (a) pRAS1: SVM-PCA (b) pRAS2: SVM-FDR (c) pRAS3: DT-PCA (d) pRAS4: DT-FDR.

sis. Overall 828 features were extracted in Ref. [42]. The authors reported melanomas classification accuracy of 90% while for other type of skin lesions was 80%. The objective of our system was to stratify the risk severity based on 859 features and using pRAS system consisting of SVM and FDA. The system reported an accuracy of 99.92%.

4.3. System reliability and stability

As per our first hypothesis, there is intense color information in psoriasis images which may provide significant information for stratification of psoriasis severity. From Fig. 4, it can be clearly observed that very high accuracies have been obtained using FC1 which consist color features alone for all 4 kinds of pRAS systems. Moreover, the FC2 which consist texture features alone were also provide reasonably high accuracy for all system combinations. This validated our second hypothesis that was about texture features and its power.

Our third hypothesis was about combined color and texture features and it can be clearly seen from Fig. 4 that FC3 which contains all color and texture features has produced highest accuracies for all system combinations except DT-PCA. Moreover, we have shown mean and standard deviation values of the top 10 dominant features selected by pRAS2: SVM-FDR system using FC3 for all five classes in Table 3. It demonstrates the ability of our pRAS system of stratification with subtle difference in feature values of distinct classes. Lastly, we believed that our pRAS is a reliable system as demonstrated in Fig. 8. The system showed a reliability of 99.73% for pRAS2: SVM-FDR using FC3 (Table 4).

We have shown the feature-wise comparison in Fig. 6 and system-wise comparison in Fig. 7. From these two figures, it can be clearly observed that the combination of all features has produced best performance because of combined effect of segregation power of individual feature sets (Fig. 6). Among the classifiers, SVM classifier showed better performance compared to decision tree based classification because of its ability to operate with large set of features (Fig. 7). In case of feature selection technique, FDR showed better performance compared to PCA as demonstrated using FRP (Section-3.4.2). Therefore, the best performance has been achieved for the system combination pRAS2: SVM-FDR with FC3 and it can be validated from Figs. 5 to 9. Due to large number of features, we adapted the two feature selection method PCA and FDR. Under the changing conditions of the training data such as data size, lesion grades, image resolutions, these two techniques can handle such fluctuations elegantly. No system is fool proof, but under the normal assumptions of our study, these techniques have shown a promising accurate and reliable results and stable machine learning solution to risk stratification. Even though, accuracies are very

Table 6
Comparative performance of proposed model with existing works on dermatology diagnosis in multiclass scenario.

Authors (year)	DB	DS	C	F	F_{cat}	FST	Cl	ICC	IFC	ACDS	PE
Ubeyli [43]	Erythemato-squamous diseases	365	6	34	Clinical, histopathological	NA	MLPNN, RNN, SVM	Yes	No	No	ACC: 85.47% (MLPNN) 96.65% (RNN) 98.32% (SVM)
Polat et al. [44]	Erythemato-squamous diseases	365	6	34	Clinical, histopathological	NA	C4.5 decision tree with one-against-all method	No	No	No	ACC: 96.71%
Abbas et al. [45]	Dermoscopy images	180	6	8	Color, texture	NA	SVM	No	No	No	SE: 91.64%, SP: 94.14%, AUC:0.948
Sharma et al. [46]	Erythemato-squamous diseases	365	6	34	Clinical, histopathological	NA	SVM, ANN, Ensemble of ANN and SVM	Yes	No	No	ACC: 99.25% (training) ACC: 98.99% (testing) (For ensemble model)
Huang et al. [47]	Erythemato-squamous diseases	365	6	34	Clinical, histopathological	SVM-REF	SVM	No	No	No	ACC: 95.38%
Ding et al. [48]	Melanoma Images	46	4	10	Asymmetry, border irregularity, color and diameter	Forward Selection	SVM	No	Yes	No	ACC: 87.8%
Menai [49]	Erythemato-squamous diseases	366	6	34	Clinical, histopathological	NA	RF, DT	Yes	Yes	No	ACC: 98.00%
Shimizu et al. [42]	Dermoscopy images	964	4	828	Color, subregion, texture	Wilks' Lambda stepwise	Linear	No	No	No	90.48% for melanomas and over 80% for each of the three other types of skin lesions
Proposed method	Psoriasis images	848	5	859	Color, texture, HOS	PCA, FDR	SVM, DT	Yes	Yes	Yes	ACC: 99.92% SRI: 99.73% FRP(PCA):85.51% FRP(FDR):100%

Abbreviations: DB, database; DS, data size; C, number of classes; NA, not applied; F, number of features; F_{cat}, feature category; FST, feature selection technique; Cl, classifier; ICC, inter classifier comparison; IFC, inter feature comparison; ACDS, analysis with changing data size; PE, performance evaluation; SE, sensitivity; SP, specificity; AUC, area under the receiver operating characteristics curve; SVM, support vector machine; DT, decision tree; RF, random forests; ANN, artificial neural network; MLPNN, multilayer perceptron neural network; RNN, recurrent neural network; REF, recursive feature elimination; SRI, system reliability index; FRP: feature retaining power.

promising, larger clinical setups do have provisions for validating psoriasis grades and stages using biopsy based methods. Further, there are factors which can affect the machine learning paradigm such as (i) devices used for data acquisition, (ii) external factors like lightening conditions under which the image data was acquired, and (iii) the gold standard adapted during the training phase of the machine learning system [50]. Though these three factors are interesting to study, but currently it is out of the scope for this publication.

4.4. Strengths, weakness and extensions

The pRAS system proposed here to stratify psoriasis severity has been found successful. Though dataset belongs to class-5 (very severe class) has very less number of samples (28), it has achieved an accuracy of 99.91% (Table 1). Further, our system provides an automated approach of segregating the disease severity in large population. The clinical application of our pRAS system is two folds. The first clinical application is adaption of the system in a routine mode *i.e.*, our pRAS system can be adapted as a second opinion for dermatologist to assess severity of psoriasis disease for single patient at a time. It would help dermatologists to provide better medication and save the time of dermatologists as well as patients. The second important application of our pRAS system is in pharmaceutical mode, where psoriasis multi-center clinical trials can be conducted for drug evaluation and therapy for better management of the psoriasis disease and its risk. Furthermore, the scoring of psoriasis severity in this study is analogues to current gold standard, PASI. In PASI, the severity of psoriasis disease has been stratified using four different score levels, *i.e.*, mild, moderate, severe and very severe [51,52]. We have also categorized the severity of psoriasis disease into these same four categories. The only difference is the addition of a healthy class representing the control group for the pRAS system. This avoids any misclassification in case if the patient image is healthy. Thus, we have five classes, namely: healthy, mild, moderate, severe and very severe. Therefore, we believe that our pRAS system is acceptable for a pilot study, a prototype and assistive clinical tool for dermatologist community. However, we believe that there is still a potential to improve the current design. Since the psoriasis research in machine learning paradigm has not been very popular due to unavailability of psoriasis database compared to other skin diseases like melanoma, the benchmarking or standardization has not been achieved. Further, more generalized diverse psoriasis databases are needed which may have hair lines or restricted lesion sizes which thus requires our system to be modified to accommodate better preprocessing approaches such as illumination models and automatic segmentation. Currently this is out of scope of the current study and we focused on multi-stage disease characterization. Secondly, to avoid inter-observer variability analysis, one of the possible extensions of this study involves grading of psoriasis image data by another dermatologist or a pathologist. Along with these, better image acquisition devices, more collaboration are required to access diversified data sets. Lastly, there are other combinations of methods for feature extraction, feature selection, training-testing classifier and one can

replace the current setup of blocks to compare and contrast with new blocks.

5. Conclusion

The paper presented a generalized multi-grade psoriasis disease risk stratification system using machine learning paradigm. The fundamental contribution of this paper is in ability to extract three set of comprehensive feature sets: (1) color; (2) grayscale; (3) combined color and grayscale and then adapt to four set of machine learning systems with two types of training-testing classifiers and two types of feature selection methods. Thus the 4 pRAS systems are: (1) support vector machine with Principal Component Analysis; (2) support vector machine with Fisher Discriminant Analysis; (3) decision tree with Principal Component Analysis and (4) decision tree with Fisher Discriminant Analysis.

Our pRAS2: SVM-FDR system concluded the best performance using combined feature set of 99.92% classification accuracy using tenfold cross-validation protocol keeping our three hypotheses into consideration. We evaluated the pRAS machine learning systems using three novel performance indices: (1) reliability index; (2) feature retaining power and (3) aggregate feature effect. The overall system showed encourage stratification accuracies and demonstrated the reliability and stability of the pRAS systems. This is the first time such a prototype has been developed and can be adapted in clinical settings while applying to other skin cancer applications.

Appendix A, B, C and D Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bspc.2016.04.001.

References

[1] C. Leonardi, R. Matheson, C. Zachariae, G. Cameron, L. Li, E. Edson-Heredia, B. Daniel, B. Subhashis, Anti-interleukin-17 monoclonal antibody ixekizumab in chronic plaque psoriasis, N. Engl. J. Med. 366 (13) (2012) 1190–1199.
[2] National Psoriasis Foundation, Statistics. Available at: hyperlink https://www.psoriasis.org/cure_known_statistics, 2015 (accessed 14.07.15).
[3] V. Chandran, S.P. Raychaudhuri, Geoepidemiology and environmental factors of psoriasis and psoriatic arthritis, J. Autoimmun. 34 (3) (2010) J314–J321.
[4] Persatuan Dermatologi Malaysia, Overview of Psoriasis in Malaysia, 2015, Available at: hyperlink http://www.dermatology.org.my (accessed 14.07.15).
[5] G. Krueger, J. Koo, M. Lebwohl, A. Menter, R.S. Stern, T. Rolstad, The impact of psoriasis on quality of life: results of a 1998 National Psoriasis Foundation patient-membership, Arch. Dermatol. 137 (3) (2001) 280–284.
[6] C. Olivier, P.D. Robert, D. Daihung, G. Urba, M.P. Catalin, W. Hywel, S.K. Kurd, A.B. Troxel, P. Crits-Christoph, J.M. Gelfand, The risk of depression, anxiety, and suicidality in patients with psoriasis: a population-based cohort study, Arch. Dermatol. 146 (8) (2010) 891–895.
[7] T. Henseler, Genetics of psoriasis, Arch. Dermatol. Res. 290 (9) (1998) 463–476.
[8] T. Morrow, Evaluating new therapies for psoriasis, Managed Care 13 (2004) 34–40.
[9] C. Camisa, Handbook of Psoriasis, second ed., Blackwell Publishing Ltd., UK, 2004.
[10] J. Lu, E. Kazmiercazk, J.H. Manton, R. Sinclair, Automatic scoring of erythema and scaling severity in psoriasis diagnosis, in: AI 2012: Advances in Artificial Intelligence, Springer, Berlin Heidelberg, 2012, pp. 73–84.
[11] J.S. Suri, Computer vision, pattern recognition and image processing in left ventricle segmentation: the last 50 years, Pattern Anal. Appl. 3 (3) (2000) 209–242.
[12] U.R. Acharya, M.R.K. Mookiah, S.V. Sree, D. Afonso, J. Sanches, S. Shafique, A. Nicolaides, L.M. Pedro, J. Fernandes e Fernandes, Jasjit S. Suri, Atherosclerotic plaque tissue characterization in 2D ultrasound longitudinal carotid scans for automated classification: a paradigm for stroke risk assessment, Med. Biol. Eng. Comput. 51 (5) (2013) 513–523.
[13] U.R. Acharya, O. Faust, A.P.C. Alvin, G. Krishnamurthi, J.C. Seabra, J. Sanches, J.S. Suri, Understanding symptomatology of atherosclerotic plaque by image-based tissue characterization, Comput. Methods Prog. Biomed. 110 (1) (2013) 66–75.
[14] L.M. Pedro, J.M. Sanches, J. Seabra, J.S. Suri, J. Fernandes e Fernandes, Asymptomatic carotid disease—a new tool for assessing neurological risk, Echocardiography 31 (3) (2014) 353–361.
[15] V.K. Shrivastava, N.D. Londhe, R.S. Sonawane, J.S. Suri, First review on psoriasis severity risk stratification: an engineering perspective, Comput. Biol. Med. 63 (2015) 52–63.
[16] V.K. Shrivastava, N.D. Londhe, R.S. Sonawane, J.S. Suri, Reliable and accurate psoriasis disease classification in dermatology images using comprehensive feature space in machine learning paradigm, Expert Syst. Appl. 42 (15) (2015) 6184–6195.
[17] V.K. Shrivastava, N.D. Londhe, R.S. Sonawane, J.S. Suri, Exploring the color feature power for psoriasis risk stratification and classification: a data mining paradigm, Comput. Biol. Med. 65 (2015) 54–68.
[18] U.R. Acharya, G. Swapna, S.V. Sree, F. Molinari, S. Gupta, R.H. Bardales, A. Witkowska, J.S. Suri, A review on ultrasound-based thyroid cancer tissue characterization and automated classification, Technol. Cancer Res. Treat. 13 (4) (2014) 289–301.
[19] M. Mirmehdi, X. Xie, J.S. Suri, Handbook of Texture Analysis, Imperial College Press, 2009.
[20] S.F. Huang, R.F. Chang, W.K. Moon, Y.H. Lee, D.R. Chen, J.S. Suri, Analysis of tumor vascularity using three-dimensional power Doppler ultrasound images, IEEE Trans. Med. Imaging 27 (3) (2008) 320–330.
[21] K. Kalyan, B. Jakhia, R.D. Lele, M. Joshi, A. Chowdhary, Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images, Adv. Bioinform. 2014 (2014).
[22] J.S. Weszka, C.R. Dyer, A. Rosenfeld, A comparative study of texture measures for terrain classification, IEEE Trans. Syst. Man Cybern. SMC-6 (4) (1976) 269–285.
[23] M. Amadasun, R. King, Textural features corresponding to textural properties, IEEE Trans. Syst. Man Cybern. 19 (5) (1989) 1264–1274.
[24] C.M. Wu, Y.C. Chen, Statistical feature matrix for texture analysis, CVGIP: Graphical Models Image Process. 54 (5) (1992) 407–419.
[25] A. Piantanelli, P. Maponi, L. Scalise, S. Serresi, A. Cialabrini, A. Basso, Fractal characterisation of boundary irregularity in skin pigmented lesions, Med. Biol. Eng. Comput. 43 (4) (2005) 436–442.
[26] K.C. Chua, V. Chandran, U.R. Acharya, C.M. Lim, Application of higher order statistics/spectra in biomedical signals—a review, Med. Eng. Phys. 32 (7) (2010) 679–689.
[27] C.L. Nikias, M.R. Raghuveer, Bispectrum estimation: a digital signal processing framework, Proc. IEEE 75 (7) (1987) 869–891.
[28] F. Song, Z. Guo, D. Mei, Feature selection using principal component analysis, System Science, Engineering Design and Manufacturing Informatization (ICSEM), International Conference on IEEE 1 (2010) 27–30.
[29] J.I. Godino-Llorente, P. Gomez-Vilda, M. Blanco-Velasco, Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-term cepstral parameters, IEEE Trans. Biomed. Eng. 53 (October (10)) (1953) 1943.
[30] U.R. Acharya, F. Molinari, S.V. Sree, S. Chattopadhyay, K.H. Ng, J.S. Suri, Automated diagnosis of epileptic EEG using entropies, Biomed. Signal Process. Control 7 (4) (2012) 401–408.
[31] U.R. Acharya, O. Faust, A.P.C. Alvin, S.V. Sree, F. Molinari, L. Saba, A. Nicolaides, J.S. Suri, Symptomatic vs. asymptomatic plaque classification in carotid ultrasound, J. Med. Syst. 36 (3) (2012) 1861–1871.
[32] L. Saba, U.R. Acharya, S. Guerriero, J.S. Suri, Ovarian Neoplasm Imaging, Springer Science and Business Media, 2014.
[33] V. Vapnik, Statistical Learning Theory, Wiley, New York, NY, 1998.
[34] M.R.K. Mookiah, U.R. Acharya, C.M. Lim, A. Petznick, J.S. Suri, Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features, Knowl. Based Syst. 33 (2012) 73–82.
[35] U.R. Acharya, S.V. Sree, R. Ribeiro, G. Krishnamurthi, R.T. Marinho, J. Sanches, J.S. Suri, Data mining framework for fatty liver disease classification in ultrasound: a hybrid feature extraction paradigm, Med. Phys. 39 (7) (2012) 4255–4264.
[36] A. Sboner, C. Eccher, E. Blanzieri, P. Bauer, M. Cristofolini, G. Zumiani, S. Forti, A multiple classifier system for early melanoma diagnosis, Artif. Intell. Med. 27 (1) (2003) 29–44.
[37] M. Zortea, T.R. Schopf, K. Thon, M. Geilhufe, K. Hindberg, H. Kirchesch, K. Møllersen, J. Schulz, S.O. Skrøvseth, F. Godtliebsen, Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists, Artif. Intell. Med. 60 (1) (2014) 13–26.
[38] C. Barata, M. Ruela, M. Francisco, T. Mendonça, J.S. Marques, Two systems for the detection of melanomas in dermoscopy images using texture and color features, Syst. J. IEEE 8 (3) (2014) 965–979.
[39] M.E. Celebi, A. Zornberg, Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesion classification, Syst. J. IEEE 8 (3) (2014) 980–984.
[40] W.Y. Chang, A. Huang, C.Y. Yang, C.H. Lee, Y.C. Chen, T.Y. Wu, G.S. Chen, Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study, PLoS One 8 (11) (2013) e76212.
[41] Z. Liu, J. Sun, L. Smith, M. Smith, R. Warr, Distribution quantification on dermoscopy images for computer-assisted diagnosis of cutaneous melanomas, Med. Biol. Eng. Comput. 50 (5) (2012) 503–513.
[42] K. Shimizu, H. Iyatomi, M.E. Celebi, K.A. Norton, M. Tanaka, Four-class classification of skin lesions with task decomposition strategy, IEEE Trans. Biomed. Eng. 62 (1) (2015) 274–283.
[43] E.D. Ubeyli, Multiclass support vector machines for diagnosis of erythemato–squamous diseases, Expert Syst. Appl. 35 (4) (2008) 1733–1740.

[44] K. Polat, S. Güneş, A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems, Expert Syst. Appl. 36 (2) (2009) 1587–1592.
[45] Q. Abbas, M.E. Celebi, I. Fondón, Computer-aided pattern classification system for dermoscopy images, Skin Res. Technol. 18 (3) (2012) 278–289.
[46] D.K. Sharma, H.S. Hota, Data mining techniques for prediction of different categories of dermatology diseases, Acad. Inf. Manag. Sci. J. 16 (2) (2013) 103.
[47] M.L. Huang, Y.H. Hung, W.M. Lee, R.K. Li, B.R. Jiang, SVM-RFE based feature selection and Taguchi parameters optimization for multiclass SVM classifier, Sci. World J. 2014 (2014) 1–11.
[48] Y. Ding, N.W. John, L. Smith, J. Sun, M. Smith, Combination of 3D skin surface texture features and 2D ABCD features for improved melanoma diagnosis, Med. Biol. Eng. Comput. (2015) 1–14.
[49] M.E.B. Menai, Random forests for automatic differential diagnosis of erythemato–squamous diseases, Int. J. Med. Eng. Inform. 7 (2) (2015) 124–141.
[50] C.M. Wu, Y.C. Chen, Statistical feature matrix for texture analysis, CVGIP: Graphical Models Image Process. 54 (5) (1992) 407–419.
[51] R.J. Chalmers, Assessing psoriasis severity and outcomes for clinical trials and routine clinical practice, Dermatol. Clin. 33 (1) (2015) 57–71.
[52] S.R. Feldman, G.G. Krueger, Psoriasis assessment tools in clinical trials, Ann. Rheum. Dis. 64 (Suppl. 2) (2005) ii65–ii68.

Vimal K. Shrivastava has received the BE degree in Electronics and Telecommunication engineering from the Chhattisgarh Swami Vivekanand Technical University, Bhilai, Chhattisgarh in 2009 and MTech degree in Electronics Instrumentation engineering from National Institute of Technology Warangal, Andhra Pradesh in 2011. He has been working toward the Ph.D. degree since 2012 from Department of Electrical Engineering of National Institute of Technology, Raipur, India.

Dr. Narendra D Londhe received his BE degree from Amravati University in 2000. Later he received his MTech and Ph.D. degrees in the year 2004 and 2011, respectively from Indian Institute of Technology Roorkee. He is presently working as Assistant Professor in Department of Electrical Engineering of National Institute of Technology, Raipur, India.

Dr. Rajendra S. Sonawane is graduated from National Institute of Homeopathy, Kolkata, India. Later he received his M.D. and D.I. in homeopathy from London. He has treated over 15,000 psoriasis patients personally from all over India and many countries of the world since 27 years with homeopathic medicines only. He is the ex-professor of Homeopathic Medical Colleges, Malkapur, Amravati, Dhule and Shirpur.

Dr. Jasjit S. Suri, PhD, MBA, Fellow AIMBE is an innovator, visionary, scientist, and an internationally known world leader. Dr. Suri received the Director General's Gold medal in 1980 and the Fellow of American Institute of Medical and Biological Engineering, awarded by National Academy of Sciences, Washington DC in 2004. He is currently Chairman of Global Biomedical Technologies, Inc., Roseville, CA, USA. He has published over 500 peer reviewed articles and book chapters and over 100 innovations/trademarks. He is currently Chairman of Global Biomedical Technologies, Inc., Roseville, CA, USA.