J.Natn.Sci.Foundation Sri Lanka 2022 50 (3): 705 - 716
DOI: http://dx.doi.org/10.4038/jnsfsr.v50i3.10403

RESEARCH ARTICLE

Biomedical Engineering

Non invasive automated approach for eczema lesions segmentation using colour space normalization

H Nisar*, Y-K Ch'ng and KH Yeap
Department of Electronic Engineering, Faculty of Engineering and Green Technology, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia.

Submitted: 9 February 2021; Revised: 17 September 2021; Accepted: 21 October 2021

Abstract: Eczema is a common type of atopic dermatitis. Eczema skin lesions can be identified visually by observing the difference between the colour and texture of the lesions and the normal skin. Dermatologists assess eczema by direct visual assessment and record their observations in specialized forms. These methods are tedious as well as time consuming and introduce inter-rater and intra-rater variability in the results. To make the assessments objective and easier for the dermatologists, we have proposed a framework for the segmentation of eczema skin lesions. Red-Green-Blue (RGB), CIELab and their normalized colour spaces were considered. For segmentation, a two step K-means algorithm was proposed. In the 1st step, a conventional K-means algorithm segments the image into three regions, *i.e.*, skin, lesion, and mixed region. This is followed by a 2nd K-means segmentation step. The performance of this method was better than the conventional methods. The algorithm was evaluated using 85 eczema images of different severity and grades. To assess the performance of the algorithm, the gold standard segmentation for eczema lesions was manually drawn and verified by a dermatologist. The Green-channel of normalized (CSN-I) RGB colour space provided the best result for a semi-supervised approach giving the segmentation accuracy of 88.28% whereas for fully automated approach a segmentation accuracy of 84.43% was achieved using support vector machine (SVM).

Keywords: Atopic dermatitis, K-means segmentation, selection modes, skin lesions.

INTRODUCTION

Atopic eczema is a chronic, relapsing and inflammatory skin condition that is widespread in all age groups and may cause social and psychological problems (Carroll *et al.*, 2005). Measuring the severity of a skin disease is the main objective in dermatology. It is essential to assess the course of the disease before, during and after the treatment. In skin diseases like atopic eczema, there are no serologic markers to identify the severity and extent of the disease. Hence evaluation of the disease is based on the visual signs and symptoms (Finlay, 1996). The general symptoms of atopic eczema are dry reddish skin with signs of inflammation. Due to scratching, the skin patches may be thickened, blistered, and sometimes infected. The different symptoms on eczema plaques are: "erythema, edema/swelling, oozing/crusting, excoriation, and lichenification" (HOME, 2014). Subjective symptoms are itching and loss of sleep which also includes difficulty in falling asleep and/or waking up at night. Eczema is commonly found at the front side of elbows and wrists, around the neck, and back of the knees.

The prevalence of eczema appears to vary across the world as observed in key international studies (Lee &

* Corresponding author (humaira@utar.edu.my; https://orcid.org/0000-0003-2026-5666)

Detzel, 2015). The latest data based on the phase three of the International Study of Asthma and Allergies in Childhood (ISAAC) (Mallol *et al.*, 2013) showed that atopic dermatitis (AD) has reached a plateau in countries like the UK and New Zealand, but it continues to increase in prevalence, specifically in younger children (age 6–7 years) in low-income countries, in regions such as South East Asia and Latin America. This skin condition affects about 31.6 million people in the United States, and one in ten individuals will develop eczema during their lifetime (Silverberg & Hanifin, 2013). In the Asia Pacific, the 12-month prevalence of AD in children (age 13–14 years) was reported to be as high as 9% in Malaysia and Singapore, and as low as 0.9% in China, which has the lowest prevalence in the world. The reasons for the differences in prevalence are not fully understood, but industrialization and socio-economic factors have been implicated (Lee & Detzel, 2015).

The main objective of measuring the severity level of eczema is for assessing the progression of the disease in patients. The quantitative scales used for measuring eczema severity are difficult to interpret in clinical practice. Doctors use systems like Eczema Area and Severity Index (EASI) (Barbier *et al.*, 2004), and Patient Oriented Eczema Measure (POEM) (Charman *et al.*, 2002) to assess the severity of eczema. The process of manual form filling is time consuming, and involves subjective assessment that results in inter-rater and intra-rater variability. Therefore, the objective of this research is to develop an automated Eczema lesion segmentation system based on the EASI index, for assessment and grading of atopic eczema. The symptoms and severity levels for EASI scoring scheme are shown in Figure 1.

A comprehensive literature review of eczema has shown that very little research has been done for the development of a non-invasive, automated, objective method for the assessment and grading of the severity of eczema. A study conducted by Tremp *et al*. (2011) proposed the use of digital imaging techniques to assess the severity of atopic eczema using the EASI scoring system. They called their method EASIdig. In this method digital camera is used to take images of the patients; these images were evaluated by general physicians trained in EASI and SCORAD scoring systems. It took about thirty minutes to calculate the score using EASIdig. Another study reported by Alam *et al.* (2016) used 85 images to develop an automated approach for eczema detection and severity measurement. K-means clustering, and morphological image processing were used to segment eczema lesions using CIELab colour space. For classification ten colour and four texture features were used, reporting an accuracy of 90% using SVM classifier.

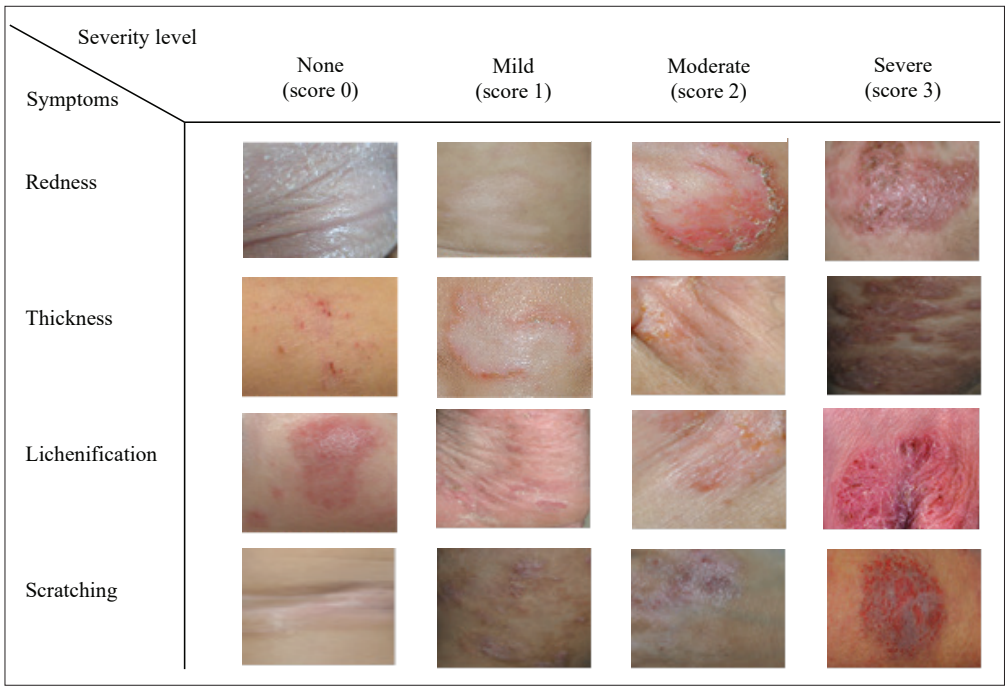


Figure 1: Symptoms and severity levels used for assessment in EASI scoring scheme (HOME, 2014)

Based on a literature review it is observed that other skin conditions such as psoriasis, acne vulgaris, and melanoma have been studied extensively. The research is normally conducted for the classification of skin lesions originating from different diseases. For example, Hegde *et al.* (2018) used RGB colour and gray level co-occurrence matrix (GLCM) texture features to classify different types of skin lesions, *i.e.*, chronic eczema, lichen planus, and plaque psoriasis. The machine learning methods SVM, naïve bayesian classifier (NBC), artificial neural networks (ANN), and linear discriminant analysis (LDA) were used. The classification accuracy using SVM classifier was 81.61%. Hameed *et al.* (2020) have proposed a framework for classification of healthy, benign, and malignant melanoma, and eczema images. A classification accuracy of 96.47% was achieved using colour and texture features.

Recently neural networks were investigated for melanoma lesion segmentation. Zafar *et al.* (2020) segmented dermoscopic images of melanoma using convolutional neural networks (CNN). They have combined UNet and ResNet architectures to form a novel architecture known as Res-Unet. A Jaccard Index of 77.2% was obtained. Abraham and Khan (2018) proposed U-Net and attention U-Net for segmentation of melanoma lesions. They achieved a dice score of 0.838 by U-Net and 0.856 by attention U-Net. The above literature review shows that tremendous research has been carried out for the classification of diseases where lesions are significantly different from each other in terms of surface properties, size, and shape.

In this paper our aim is to segment eczema skin lesions having the characteristic feature of no specific size and shape with a fuzzy boundary. The shape, size and colour of eczema lesions are not well defined; it can take any shape, spread to any size, and its colour varies a lot with the increase in the severity level. Some lesions are concentrated in a smaller area whereas others are scattered in a larger area. This makes the segmentation and classification of eczema lesions extremely difficult.

In the following sub-sections, we will give a brief overview of the pre-processing methods, colour spaces, segmentation, and classification techniques that we found useful for Eczema lesion segmentation.

Image pre-processing

Image pre-processing removes noise and the effect of irregular illumination from an image. Ambient light condition is a very important criterion for image acquisition. The image segmentation may be affected by over-exposure or under-exposure of light. Two illumination correction techniques are found suitable for eczema lesion image illumination correction; these are adaptive light compensation (ALC) (Ch'ng *et al.*, 2014a) and Frankle-McCann retinex (retinex) (Funt *et al.*, 2004). In ALC the value of the luminance denoted as Y is extracted from the input image and average luminance (Y_{avg}) is computed. The maximum value of 200 and minimum value of 80 for good luminance exposure is obtained empirically from the images. A luminance factor is computed for RGB image correction as given in equation (1).

$$Factor = \begin{cases} 200 / Y_{avg}; & \text{if } Y_{avg} > 200 \\ 1 \;; & \text{if } 80 < Y_{avg} < 200 \\ 80 / Y_{avg}; & \text{if } Y_{avg} < 80 \end{cases} \quad ...(1)$$

The retinex method is designed to correct the illumination of under-exposed (dark) areas of an image while maintaining the visual discrimination of lighter areas (Funt *et al.*, 2004).

Colour spaces

Colour is an important property in visual perception. Different colour spaces are useful for different applications. A preliminary study with four colour spaces (HSI, YCbCr, CMY, CIELab) for eczema lesion segmentation using k-means showed that hue (H) channel of HSI has the highest segmentation accuracy of 76.63% (Nisar *et al.*, 2013).

Yang *et al.* (2010) proposed a colour space normalization (CSN) technique to improve the face recognition rate. His analyses of different colour spaces showed that those colour spaces perform better when the sum of elements in the 2nd and 3rd row of transformation matrix are zero. The RGB colour space does not possess this property; however, the YUV colour space has this property as shown in equations (2) and (3).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad ...(2)$$

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1471 & -0.2888 & 0.4359 \\ 0.6148 & -0.5148 & -0.1000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad ...(3)$$

The CSN technique can normalize any colour space and

the normalized colour space has a transformation matrix in which the sum of elements in the 2nd and 3rd rows are zero. Colour space normalization-I (CSN-I) is also known as within-colour-component normalization. It is obtained by removing the means of 2nd and 3rd row vectors respectively. As a result, the sum of the row vectors is equal to zero in the transformation matrix. Equations (4) and (5) (Yang *et al.*, 2010) show the transformation matrix for the CSN-I RGB, $\widetilde{A}_{RGB-I}$ and CSN-I CIE-XYZ, $\widetilde{A}_{XYZ-I}$ colour spaces, respectively. Colour space normalization-II (CSN-II) is known as across-colour-component normalization. In this technique the zero-mean vector of the 2nd and 3rd rows is obtained via linear combination. Equations (6) and (7) (Yang *et al.*, 2010) are the transformation matrices obtained for CSN-II RGB, $\widetilde{A}_{RGB-I}$ and CSN-II CIE-XYZ, $\widetilde{A}_{XYZ-I}$ colour spaces, respectively. Figure 2 gives an overview of different colour components.

$$\widetilde{A}_{RGB-I} = \begin{bmatrix} 1 & 0 & 0 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix} \quad ...(4)$$

$$\widetilde{A}_{XYZ-I} = \begin{bmatrix} 0.6070 & 0.1740 & 0.2 \\ -0.0343 & 0.2537 & -0.2193 \\ -0.3940 & -0.3280 & 0.7220 \end{bmatrix} \quad ...(5)$$

$$\widetilde{A}_{RGB-I} = \begin{bmatrix} 1 & 0 & 0 \\ -0.5774 & 0.7887 & -0.2113 \\ -0.5774 & -0.2113 & 2/3 \end{bmatrix} \quad ...(6)$$

$$\widetilde{A}_{XYZ-I} = \begin{bmatrix} 0.6070 & 0.1740 & 0.2 \\ -0.0901 & 0.3631 & -0.2730 \\ -0.4600 & -0.1986 & 0.6586 \end{bmatrix} \quad ...(7)$$

Segmentation and classification methods

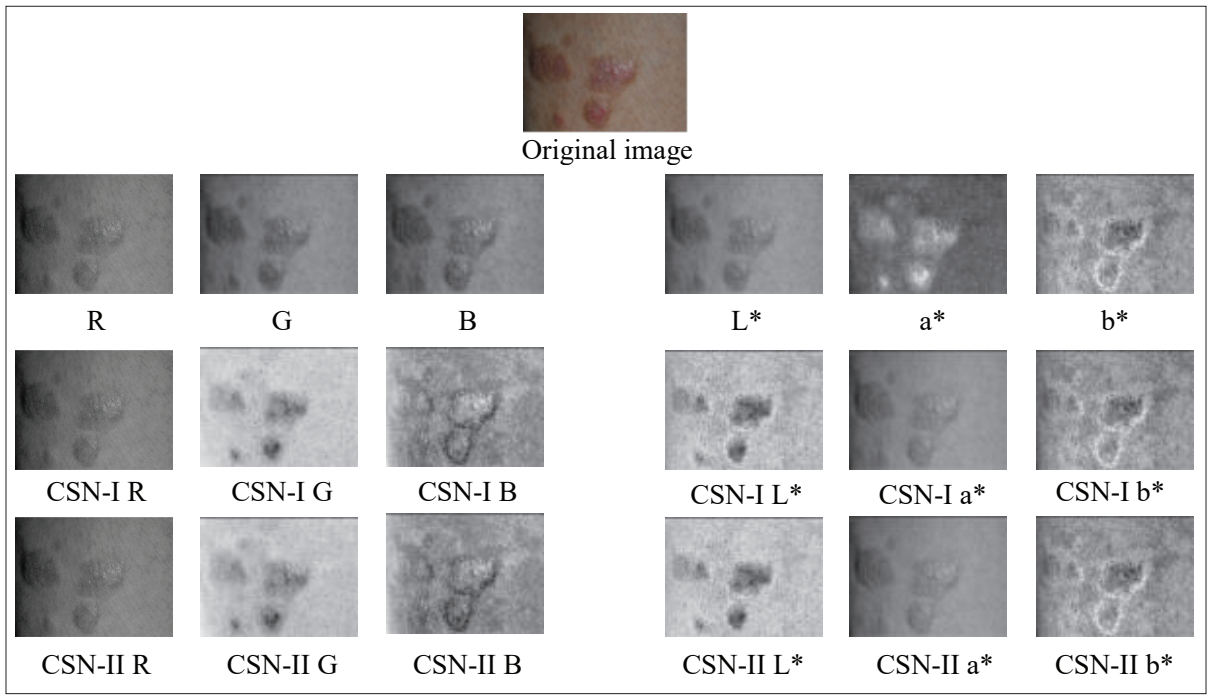


Figure 2: Illustration of different colour components for RGB, CIE Lab and their normalized CSN-I and CSN-II colour spaces

In image processing, clustering methods are commonly used for segmentation of the desired regions of interest (ROI). The classification methods are divided into two categories: unsupervised learning and supervised learning classification methods. Unsupervised learning classifiers do not require any reference data for carrying out the classification. The classifier groups the data to a class itself. Clustering requires the user to define the hypothesized number of clusters. Commonly used methods of clustering are K-means clustering (Hartigan & Wong, 1979) and Fuzzy C-mean (FCM) (Bezdek

et al., 1984). However, supervised learning classifiers require a set of training data to work as a reference for classification. Supervised learning classification groups similar data based on the training data groups. Commonly used approaches are region-based, energy-based, and region and contour-based approaches. There are also some other classifiers such as SVM (Hegde *et al.*, 2018).

MATERIALS AND METHODS

The first step in any image processing-based application is image acquisition. It is followed by image pre-processing for image enhancement and noise removal. This is followed by image segmentation to obtain the ROI. In this paper we have proposed an algorithm to segment eczema skin lesions. We will discuss all the steps in the algorithm in the following sub-sections.

Data acquisition

The images were acquired using a digital single lens reflex (DSLR) camera. A dataset of 85 images in JPEG format is available for this study, the images were voluntarily provided. The gold standard for the skin lesions was manually drawn using the GNU Image Manipulation Program (GIMP) (The GIMP Team, 2014); and verified by a dermatologist. For performance assessment of segmentation, the gold standard images were used.

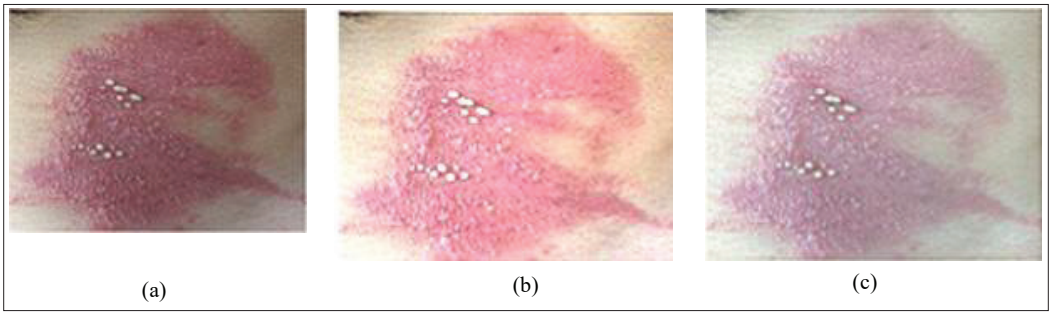


Figure 3: Comparison of pre-processing algorithms (a) input image; (b) adaptive light compensation (ALC) and (c) Frankle-McCann retinex (retinex)

Image pre-processing

Colour invariance due to uneven illumination makes the image segmentation more challenging. Hence, an ALC method (Ch'ng *et al.*, 2014a) was used to pre-process the images. We have also used another pre-processing method, Frankle-McCann retinex (Funt *et al.*, 2004) to compare the performance of pre-processing as illustrated in Figure 3.

Proposed segmentation approach

The K-means segmentation method is commonly used in unsupervised segmentation. We have proposed a bi-level K-means segmentation that divides the segmentation process into two steps (Ch'ng *et al.*, 2014b). The proposed technique is shown in in Figure 4. In the first step, it divides the image into three regions: the skin region, the mixed region, and the lesion region. Then the intensity feature [mean (μ) ± standard deviation (SD)], of skin or lesion region can be used as a reference value for the calculation of Mahalanobis distance to segment the mixed region into three sub-regions. We have used skin region as the reference as it gives better results.

For user supervised selection, the segmented cluster can be obtained by using one of the four different combinations as shown in the flowchart in Figure 4. The user will choose the most appropriate segmentation combination based on visual observation as shown in Figure 5. However, for the unsupervised method the SVM classifier with ranked features as training data will be used.

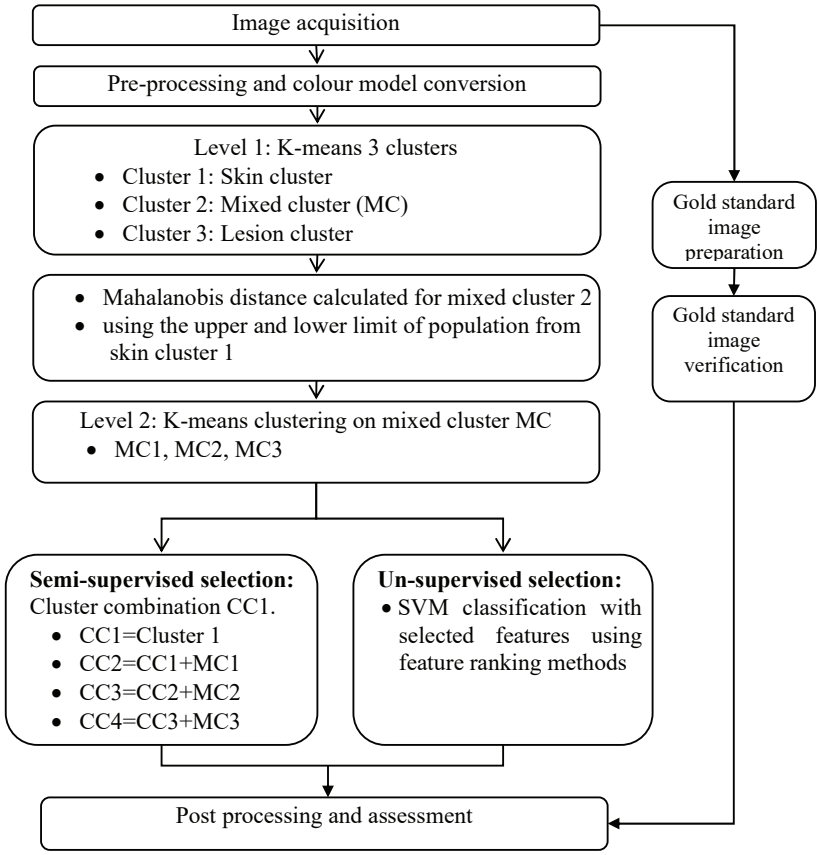


Figure 4: The proposed bi-level segmentation algorithm

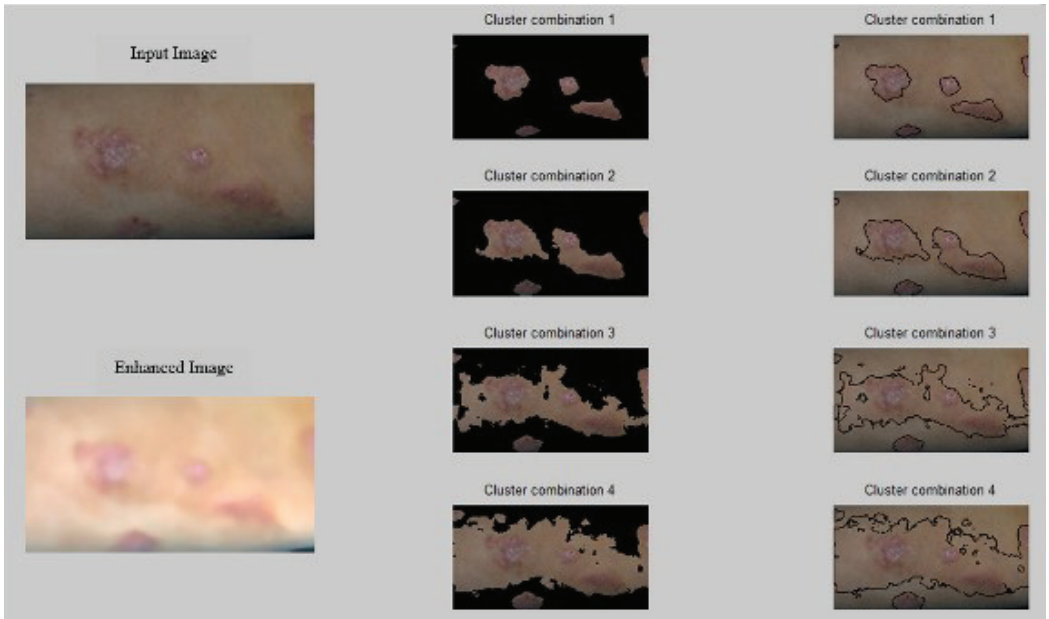


Figure 5: Window for the operator to select a suitable cluster combination for manual selection in level 2

Level 1 clustering

After the pre-processing, the image was segmented into three regions using K-means clustering. The regions were labelled as skin region (cluster 1), mixed region (cluster 2), and lesion region (cluster 3). K-means clustering approach proposed by Yao *et al.* (2013) is used. Equation (8) shows the Euclidean distance (Brindha *et al.*, 2013) used in the 1st level of K-means clustering.

$$Euclidean\ \ Dist, d(c,x) = \sqrt{\sum_{u=1}^{v} (x_u - c_u)^2} \qquad ...(8)$$

where *d* represents the distance, *x* and *c* represent the data points and centroid respectively, and *u* and *v* are the data dimensions used (here only 1 dimension of data is used).

Level 2 clustering

In Level 2 the mixed region is divided into three sub-regions (skin, mixed and lesion) using K-means clustering. The lesion is now composed of 5 clusters. Here the Mahalanobis distance (McLachlan, 1999) was used for the mixed region segmentation as it gives better results than the Euclidean distance (empirically observed). The Mahalanobis distance is given in equation (9).

$$Mahalanobis\ \ Dist, d(x,y) = \sqrt{(y - \mu_x)/\mathrm{cov}(x) \times (y - \mu_x)^T} \qquad ...(9)$$

where *y* represents the data point and *x* represents the vector (mean ± standard deviation) that is used as the reference in equation (9).

The segmentation of the final (mixed) cluster was the most difficult. It was done, using semi-supervised (operator input) or un-supervised (SVM) method.

Cluster combination for manual selection

For semi-supervised selection, operator input is required. The operator subjectively selects the most suitable combination of clusters that represents the segmented lesion. From Level 2 segmentation one cluster was combined with the lesion region at a time, and the option for four types of segmentation selection is present, as shown in the selection window in Figure 5. In the selection window the input image and pre-processed image are shown in the left column. The middle and right columns show the cluster combinations that were selected by the operator. In the right column, the boundary of the lesion was drawn using a black line. For the middle column the region that does not belong to lesion region in the segmented image was excluded. The operator selects one of the four cluster combinations that best represents the segmented lesion, based on his judgement.

SVM classifier for automated selection

SVM was trained using selected features that are ranked by feature ranking methods. In this study, the T-statistic scoring method was used for feature ranking. The T-statistic scoring method calculates a score for the capability, C_T, of each feature to separate them between classes. Based on the value of C_T, the scoring of each feature was ranked in descending order. The two classes are denoted as 'a' and 'b'. For each feature, the mean, μ_a and μ_b; standard deviation, σ_a and σ_b; and the number of images used, n_a and n_b were calculated (Liu *et al*., 2002). The mathematical expression for calculating C_T is given in equation (10).

$$C_T = \frac{\mu_a - \mu_b}{\sqrt{(\sigma_a^2/n_a) + (\sigma_b^2/n_b)}} \qquad ...(10)$$

The score value obtained in equation (10) was used to sort the features in descending order. The combination of features was done by increasing one feature at a time. Finally, validation was done using 53 gold standard images; the validation error rate was obtained using equation (11).

$$\text{Error rate} = \frac{\text{Falsely classified data}}{\text{Overall data}} \qquad ...(11)$$

Feature extraction and selection

For SVM training features were extracted from the skin and lesion regions of 20 gold standard images. These images were classified as mild images in context of the redness symptom. Five features extracted from the colour layers of four-colour spaces (RGB, CIELab, CSN-I RGB, CSN-I CIELab) are mean, standard deviation (SD), mode, skewness, and kurtosis, whereas the four features extracted from GLCM are contrast, correlation, energy, and homogeneity (Nisar *et al.*, 2020). Colour layers of colour spaces were used for creating the GLCM; the intensity of the colour layer was normalized to '0' to '1'. A symmetrical GLCM was used with the horizontal direction only. The number of levels used for GLCM is 256. All extracted features were ranked using the T-statistical score ranking method and the combination of ranked features with lowest error was selected for SVM training.

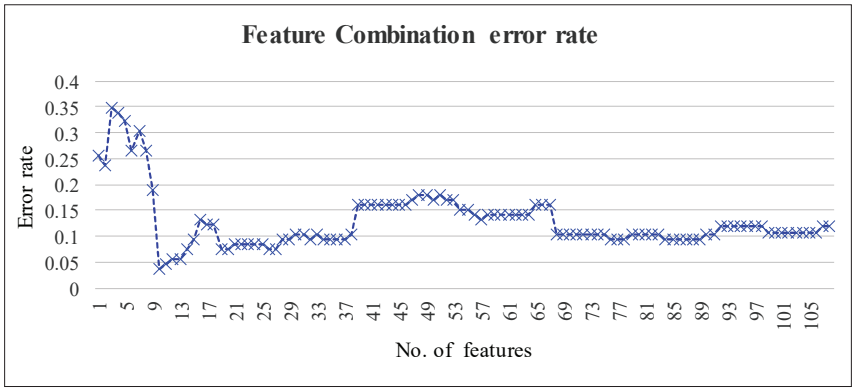


Figure 6: Feature combination error rate graph

Figure 6 shows the feature combination error rate obtained during the training and classification of gold standard data. In this study, 108 features are used, of which 60 (5 features × 4 colour spaces × 3 layers) are colour features and 48 (4 features × 4 colour spaces × 3 layers) are texture features. It is observed from Figure 6 that the lowest error rate (less than 5%) is achieved for 10 features. Hence, we have used these 10 features for segmentation and classification, which are SD and kurtosis of the R channel, the energy of the G channel, correlation and homogeneity from CSN-I R, homogeneity from CSN-I G, Mean, skewness and homogeneity from CSN-I a*, and skewness from CSN-I b*.

RESULTS AND DISCUSSION

In this paper we have proposed an algorithm for segmentation of eczema skin lesions. For pre-processing, the ALC method is used, and the results are compared with the retinex algorithm. The performance of the proposed segmentation algorithm is compared with K-means and FCM algorithms with three and five clusters. Two core colour spaces, RGB and CIELab, and two normalization techniques for colour spaces, CSN-I and CSN-II, are studied. The green channel of RGB colour space is selected for segmentation as it gives better segmentation performance when compared with red and blue channels; we denote it as RGB(G) (Ch'ng *et al.*, 2014a). Similarly, the green channel is also chosen for CSN-I RGB and CSN-II RGB, and we denote it as CSN-I RGB(G) and CSN-II RGB(G), respectively. For CIELab colour space, chrominance channel a* is selected, and we denote it as CIELab(a*); for CSN-I CIELab and CSN-II CIELab illumination channel L* is selected and we denote it as CSN-I CIELab(L*) and CSN-II CIELab(L*), respectively.

For performance assessment, different metrics are used. True positive, true negative, false positive and false negative are denoted as TP, TN, FP, and FN respectively. These are used to calculate the precision, recall/ sensitivity (Sens), specificity (Spec), and accuracy (Acc), as given in equations (12)–(15). All the three parameters should have a high value for good classification results. The DICE coefficient is calculated using equation (16) to compare the similarity between the gold standard and segmented images.

$$Precision = TP/(TP + FP) \quad ...(12)$$

$$Recall/Sensitivity\ (Sens) = TP/(TP + FN) \quad ...(13)$$

$$Specificity\ (Spec) = TN/(TN + FP) \quad ...(14)$$

$$Accuracy\ (Acc) = (TP + TN)/(TP + TN + FP + FN) \quad ...(15)$$

$$DICE\ Coefficient = (2 \times TP)/(2 \times TP + FP + FN) \quad ...(16)$$

Table 1 shows the performance comparison of segmentation algorithms using the semi-supervised method: ALC as pre-processing technique and CSN-I RGB(G) colour channel. It is observed that all methods can achieve above 98% accuracy for the best-case segmentation, image 1; however, for image 2, the proposed algorithm has the highest accuracy whereas K-means and FCM with 5 clusters have better accuracy than with 3 clusters.

Tables 2 to 4 show the performance assessment of different segmentation algorithms. Out of the six

colour spaces evaluated, results for only top three best colour spaces are shown. Table 2 shows the overall segmentation performance for the proposed technique using semi-supervised and un-supervised methods. Among the three-colour spaces, CSN-I RGB(G) has the highest accuracy. Between pre-processing methods, ALC has a higher accuracy as compared to retinex.

Tables 3 shows the segmentation assessment for K-means using five and three clusters respectively. It is observed that the ALC pre-processing technique gives better segmentation accuracy. In K-means, for both pre-processing methods CSN-I RGB(G), CSN-II RGB(G) and CIELab(a*) are the three top ranked colour spaces that give the highest segmentation accuracy. Finally, K-means with k = 5 gives higher segmentation accuracy as compared to the k = 3 case.

Table 4 shows the segmentation performance for FCM using five and three clusters respectively. For both pre-processing techniques, ALC has a better segmentation accuracy for FCM compared to retinex. For FCM the top three ranked colour spaces are the same as with K-means, as in Table 3. For both FCM (k = 5) and FCM (k = 3), CSN-I RGB(G) has highest segmentation accuracy. However, for the semi-supervised method using the retinex pre-processing technique there is an exception, as the CIELab(a*) colour space has a slightly higher segmentation accuracy compared to other colour spaces. Table 5 shows the comparison in terms of DICE coefficient. When comparing between pre-processing methods, ALC performs better than retinex. CSN-I RGB(G) colour space outperforms all the other colour spaces with highest performance accuracy. For segmentation techniques, the proposed bi-level K-means segmentation gives a better

Table 1: Performance comparison of segmentation algorithms (semi-supervised method) using CSN-I RGB (G channel).

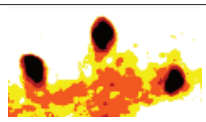
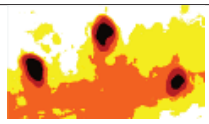
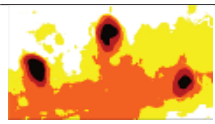
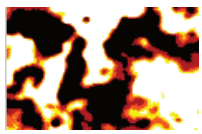
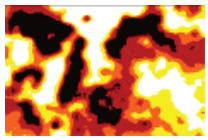
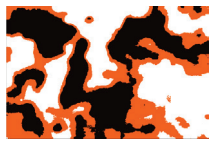
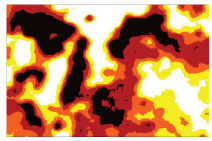
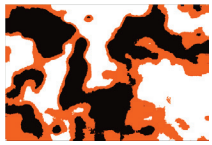
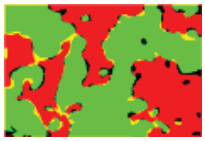
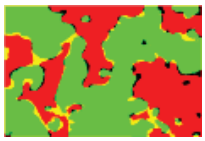
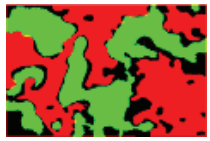
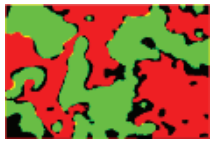
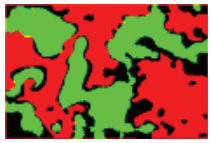
	Proposed	K-means with 5 clusters	K-means with 3 clusters	FCM with 5 clusters	FCM with 3 clusters
Clustered image 1					
Performance of segment. (Best case)					
Accuracy	98.91%	98.20%	98.91%	98.32%	98.90%
Clustered image 2					
Performance of segment. (random sample)					
Accuracy	91.22%	90.79%	79.15%	84.31%	78.42%

Table 2: Segmentation performance of the proposed two-level K-means algorithm (%)

	ALC				Retinex			
	Precision	Recall/Sens	Spec	Acc	Precision	Recall/Sens	Spec	Acc
Semi-supervised selection								
CSN-I RGB(G)	86.78	81.76	90.37	88.28	87.78	76.86	87.93	86.40
CSN-II RGB(G)	82.52	80.34	87.77	85.98	80.07	82.08	82.79	84.63
CIELab(a*)	79.93	82.45	84.91	85.27	76.31	85.09	78.12	83.08
Un-supervised selection								
CSN-I RGB(G)	80.05	84.98	84.28	84.43	77.91	81.80	81.96	81.99
CSN-II RGB(G)	73.52	86.80	77.87	81.04	73.73	84.51	77.91	80.28
CIELab(a*)	71.95	87.17	75.33	80.20	74.22	84.25	79.10	80.81

Table 3: Segmentation performance of the K-means with 5 and 3 clusters (%)

	ALC				Retinex			
	Precision	Recall/Sens	Spec	Acc	Precision	Recall/Sens	Spec	Acc
Semi-supervised selection (5clusters)								
CSN-I RGB(G)	87.51	79.23	90.70	87.51	83.38	78.44	87.01	84.87
CSN-II RGB(G)	84.24	77.52	88.41	85.77	80.15	80.94	84.24	84.25
CIELab(a*)	81.04	79.94	85.33	84.67	80.06	82.73	83.65	84.79
Un-supervised selection (5 clusters)								
CSN-I RGB(G)	76.71	84.77	76.04	79.78	73.89	83.24	73.47	78.23
CSN-II RGB(G)	68.64	88.15	68.19	76.25	70.28	85.52	68.68	76.39
CIELab(a*)	68.12	88.33	67.28	75.61	71.02	85.46	71.03	77.08
Semi-supervised selection (3 clusters)								
CSN-I RGB(G)	86.88	77.84	87.37	86.23	79.91	81.30	80.41	82.49
CSN-II RGB(G)	81.33	80.08	84.07	84.32	80.70	77.95	84.13	82.84
CIELab(a*)	75.98	84.91	78.34	82.91	80.21	81.29	83.25	83.67
Un-supervised selection (3 clusters)								
CSN-I RGB(G)	80.01	82.55	81.68	82.46	78.88	79.26	79.31	80.31
CSN-II RGB(G)	72.91	86.25	74.78	79.50	76.78	80.87	78.60	79.77
CIELab(a*)	71.91	87.72	73.61	79.66	73.51	82.74	76.70	78.33

Table 4: Segmentation performance of Fuzzy C mean with 5 and 3 clusters (%)

	ALC				Retinex			
	Precision	Recall/ Sens	Spec	Acc	Precision	Recall/ Sens	Spec	Acc
Semi-supervised selection (5 clusters)								
CSN-I RGB(G)	87.52	78.60	91.37	87.36	84.25	78.59	88.65	85.65
CSN-II RGB(G)	85.02	77.15	89.18	86.08	81.10	81.16	84.27	84.44
CIELab(a*)	80.13	81.30	85.11	84.63	78.89	83.57	83.24	84.31
Un-supervised selection (5 clusters)								
CSN-I RGB(G)	77.79	84.16	76.71	79.90	74.48	82.57	73.76	78.10
CSN-II RGB(G)	68.17	87.80	67.96	75.94	72.21	84.55	71.62	77.02
CIELab(a*)	67.83	88.62	66.90	75.61	70.36	86.09	70.20	76.75
Semi-supervised selection (3 clusters)								
CSN-I RGB(G)	87.78	76.86	87.93	86.40	83.34	78.02	83.05	83.43
CSN-II RGB(G)	80.07	82.08	82.79	84.63	80.59	79.08	83.66	83.11
CIELab (a*)	76.31	85.09	78.12	83.08	80.49	80.92	81.73	83.60
Un-supervised selection (3 clusters)								
CSN-I RGB(G)	79.80	81.74	81.43	81.70	80.00	79.74	79.87	80.90
CSN-II RGB(G)	72.57	86.38	73.97	78.90	76.39	81.66	77.62	79.38
CIELab(a*)	71.11	87.94	72.63	79.06	72.68	83.42	74.84	77.59

Table 5: Comparison of the segmentation algorithms in terms of DICE coefficient (%)

		Semi-supervised selection					Un-supervised selection				
		Proposed	K-means (k=5)	FCM (k=5)	K-means (k=3)	FCM (k=3)	Proposed	K-means (k=5)	FCM (k=5)	K-means (k=3)	FCM (k=3)
ALC	CSN-I RGB(G)	83.47	81.72	81.74	79.96	79.73	80.17	76.74	77.01	77.41	76.44
	CSN-II RGB(G)	80.56	79.70	79.69	78.51	79.06	77.28	73.94	73.47	75.87	75.46
	CIELab(a*)	80.62	79.21	79.52	78.48	78.72	76.47	73.30	73.32	76.27	75.67
Retinex	CSN-I RGB(G)	80.41	79.34	80.29	77.75	77.88	76.99	74.09	73.65	74.97	76.02
	CSN-II RGB(G)	79.94	79.32	79.96	76.71	77.61	76.13	73.79	74.14%	75.65	75.60
	CIELab(a*)	80.61	80.09	79.77	78.46	78.13	76.31	74.00	73.76	73.63	73.34

accuracy for most of the cases. There are only six cases out of 24 where bi-level K-means does not give the best accuracy; these are RGB(G) in ALC for both the semi-supervised and un-supervised selection mode, CSN-II RGB(G) in ALC, and CIELab(a*) colour space with both normalized techniques in retinex for semi-supervised selection mode. Conventional K-means seems to have a slightly better performance than the proposed method for those cases. Finally, we can say that the proposed method gives an overall best accuracy for the semi-supervised method, which is 88.28%, whereas the un-supervised selection mode has an accuracy of 84.43%.

For the semi-supervised selection mode, segmentation with a higher cluster number (5) gives a better accuracy compared to fewer clusters (3) for both K-means and FCM, whereas the accuracy of the unsupervised selection mode tends to be lower for the higher number of clusters as compared to the lower number of clusters used. For the proposed segmentation algorithm, the input image will first be segmented into three regions in the first step and then the mixed region will further be segmented into another three sub-regions. For the unsupervised selection mode, the poor performance with a higher number of clusters for K-means and FCM may be due to the features selected for training by SVM, despite using the features with lowest validation error rate.

Hence from the results it is observed that the CSN-I RGB(G) colour space has better segmentation accuracy resulting in the highest similarity between the gold standard and segmented images. The colour space normalization technique proposed by Yang *et al.* (2010) shown in equation (4) amplifies the difference between normal skin and an eczema lesion in the green channel of CSN-I RGB colour space, that leads to a better segmentation for the proposed bi-level K-means segmentation algorithm.

CONCLUSION

In this paper we have investigated the performance of colour spaces, pre-processing methods, segmentation techniques and selection modes. Investigations show that G channel of CSN-I RGB outperforms all other colour spaces for the Eczema skin lesion segmentation. The proposed segmentation method and pre-processing technique ALC performed better in most cases. The highest segmentation accuracy achieved is 88.28% for proposed bi-level segmentation method using CSN-I RGB (G channel) colour space for semi-supervised selection. While highest accuracy achieved for un-supervised selection is 84.43% with the proposed segmentation method using the CSN-I RGB (G channel) colour space. In future we plan to increase the image database for an effective evaluation. Secondly more features should be considered for feature selection; used for training the supervised learning classifier. Deep learning methods will also be considered.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgement

This work is supported by Universiti Tunku Abdul Rahman Research Fund (UTARRF) Grant Number: IPSR/RMC/UTARRF/2020-C1/Y03. The authors would also like to thank Dr Jyh Jong Tang, Department of Dermatology, Hospital Raja Permaisuri Bainun, Ipoh, Malaysia, for his help in verifiying the gold standard for eczema skin lesions

REFERENCES

Abraham N. & Khan N.M. (2018). A Novel Focal Tversky Loss Function with Improved Attention U-Net for Lesion Segmentation. *arXiv preprint [cs.CV]*. arXiv:1810.07842

Alam M.N., Munia T.T.K., Tavakolian K., Vasefi F., MacKinnon N. & Fazel-Rezai R. (2016). Automatic detection and severity measurement of eczema using image processing. In: *2016 IEEE 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, USA, 16–20 August. IEEE, pp. 1365–1368.
DOI: https://doi.org/10.1109/EMBC.2016.7590961

Barbier N., Paul C., Luger T., Allen R., De Prost Y., Papp K., Eichenfield L.F., Cherill R. & Hanifin J. (2004). Validation of the eczema area and severity index for atopic dermatitis in a cohort of 1550 patients from the pimecrolimus cream 1% randomized controlled clinical trials programme. *British Journal of Dermatology* **150**(1):96–102.
DOI: https://doi.org/10.1111/j.1365-2133.2004.05696.x

Bezdek J.C., Ehrlich R. & Full W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences* **10**(2–3):191–203.
DOI: https://doi.org/10.1016/0098-3004(84)90020-7

Brindha M., Tamilselvan G.M., Valarmathy S., Kumar M.A. & Suryalakshmipraba M. (2013). A comparative study of face

authentication using Euclidean and Mahalanobis distance classification method. *International Journal of Emerging Technology and Advanced Engineering* **3**(1):263–268.

Carroll C.L., Balkrishnan R., Feldman S.R., Fleischer A.B. & Manuel J.C. (2005). The burden of atopic dermatitis: impact on the patient, family, and society. *Pediatric Dermatology* **22**(3):192–199.
DOI: https://doi.org/10.1111/j.1525-1470.2005.22303.x

Ch'ng Y.K., Nisar H., Yap V.V., Yeap K.H. & Tang J.J. (2014). Segmentation and grading of eczema skin lesions. In: *Signal Processing and Communication Systems (ICSPCS), 8th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, Australia, 15-17 December. IEEE, pp. 1–5.
DOI: https://doi.org/10.1109/ICSPCS.2014.7021131

Ch'ng Y.K., Nisar H., Yap V.V. & Tang J.J. (2014). A two level k-means segmentation technique for eczema skin lesion segmentation using class specific criteria. In: *Biomedical Engineering and Sciences (IECBES), IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, Kuala Lumpur, Malaysia, 8-10 December. IEEE, pp. 985–990.
DOI: https://doi.org/10.1109/IECBES.2014.7047659

Charman C.R., Venn A.J. & Williams H.C. (2002). Reliability testing of the six area, six sign atopic dermatitis severity score. *British Journal of Dermatology* **146**(6):1057–1060.
DOI: https://doi.org/10.1046/j.1365-2133.2002.04644.x

Finlay A.Y. (1996). Measurement of disease activity and outcome in atopic dermatitis. *British Journal of Dermatology* **135**(4): 509–515.
DOI: https://doi.org/10.1046/j.1365-2133.1996.d01-1031.x

Funt B., Ciurea F. & McCann J. (2004). Retinex in MATLAB™. *Journal of Electronic Imaging* **13**(1):48–57.
DOI: https://doi.org/10.1117/1.1636761

Hameed N., Shabut A.M., Ghosh M.K. & Hossain M.A. (2020). Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. *Expert Systems with Applications* **141**(1 a): 112961.
DOI: https://doi.org/10.1016/j.eswa.2019.112961

Hartigan J.A. & Wong M.A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **28**(1): 100–108.
DOI: https://doi.org/10.2307/2346830

Hegde P.R., Shenoy M.M. & Shekar B.H. (2018). Comparison of machine learning algorithms for skin disease classification using color and texture features. In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, India, 19–22 September. IEEE, pp. 1825–1828.
DOI: https://doi.org/10.1109/ICACCI.2018.8554512

HOME (Harmonising Outcome Measures for Eczema). (2014). EASI For Clinical signs. Available at *http://www.homeforeczema.org/research/easi-for-clinical-signs.aspx*

Lee B.W. & Detzel P.R. (2015). Treatment of childhood atopic dermatitis and economic burden of illness in Asia Pacific countries. *Annals of Nutrition and Metabolism* **66**(Suppl 1): 18–24.
DOI: https://doi.org/10.1159/000370221

Liu H., Li J. & Wong L. (2002). A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns. *Genome Informatics* **13**: 51–60.
DOI: https://doi.org/10.11234/gi1990.13.51

Mallol J., Crane J., von Mutius E., Odhiambo J., Keil U., Stewart A. & ISAAC Phase Three Study Group (2013). The international study of asthma and allergies in childhood (ISAAC) phase three: a global synthesis. *Allergologia et Immunopathologia* **41**(2): 73–85.
DOI: https://doi.org/10.1016/j.aller.2012.03.001

McLachlan G.J. (1999). Mahalanobis distance. *Resonance* **4**(6): 20–26.
DOI: https://doi.org/10.1007/BF02834632

Nisar H., Ch'ng Y.K., Chew T.Y., Tang J.J., Yap V.V & Yeap K.H. (2013). A color space study for skin lesion segmentation. In: *Circuits and Systems (ICCAS), IEEE International Conference on Circuits and Systems (ICCAS)*, Kuala Lumpur, Malaysia, 18–19 September. IEEE, pp. 172–176.
DOI: https://doi.org/10.1109/CircuitsAndSystems.2013.6671629

Nisar H., Ch'ng Y.K. & Ho Y.K. (2020). Automatic segmentation and classification of eczema skin lesions using supervised learning. In: *IEEE Conference on Open Systems (ICOS)*, Kota Kinabalu, Malaysia, 17–19 November. IEEE, pp. 25–30.
DOI: https://doi.org/10.1109/ICOS50156.2020.9293657

Silverberg J.I. & Hanifin J.M. (2013). Adult eczema prevalence and associations with asthma and other health and demographic factors: A US population–based study. *Journal of Allergy and Clinical Immunology* **132**(5): 1132–1138.
DOI: https://doi.org/10.1016/j.jaci.2013.08.031

The GIMP Team. (2014). *GNU Image Manipulation Program ver 2.8*. Available at *https://download.gimp.org/mirror/pub/gimp/v2.8/* , Accessed 24 January 2021.

Tremp M., Knafla I., Burg G., Wüthrich B. & Schmid-Grendelmeier P. (2011). EASIdig A digital tool to document disease activity in atopic dermatitis. *Dermatology* **223**(1): 68–73.
DOI: https://doi.org/10.1159/000330333

Yang J., Liu C. & Zhang L. (2010). Color space normalization: Enhancing the discriminating power of color spaces for face recognition. *Pattern Recognition* **43**(4): 1454–1466.
DOI: https://doi.org/10.1016/j.patcog.2009.11.014

Yao H., Duan Q., Li D. & Wang J. (2013). An improved K-means clustering algorithm for fish image segmentation. *Mathematical and Computer Modelling* **58**(3-4): 790–798.
DOI: https://doi.org/10.1016/j.mcm.2012.12.025

Zafar K., Gilani S.O., Waris A., Ahmed A., Jamil M., Khan M.N. & Kashif A.S. (2020). Skin lesion segmentation from dermoscopic images using convolutional neural network. *Sensors* **20**(6): 1601.
DOI: https://doi.org/10.3390/s20061601