

# Skin Cancer Classification Using Dermoscopic Images based on Ranklet Transform, Co-occurrence Features and Random Forest Classifier

**Chithra Devi M**

*Assistant Professor, Annai Vailankanni Arts and Science College, Thanjavur, (Affiliated to Bharathidasan University)*

## **Abstract**

The skin cell which grows abnormally is known as skin cancer and it is caused due to sun rays in the uncovered skin. It spreads within a week, so the early diagnosis of skin cancer is required. Skin Cancer Classification (SCC) based on ranklet transform, co-occurrence features and random forest classifier is presented in this paper. Dermoscopic image in the PH2 database is used in this study for performance evaluation. Initially, the dermoscopic images of three categories normal benign and malignant images are preprocessed to smooth the images. Then the images are given to ranklet transform for decomposition. It produces subband coefficients. The Ranklet Features based Co-occurrence Matrix (RFCM) is used to extract the features and stored in database. The classification is made by Random Forest Classifier (RFC). The result shows better classification accuracy of 93.5% sensitivity is 92% and specificity is 95% is obtained by RFCM and RFC.

**Keywords:** *Dermoscopy images, Ranklet transform, Co-occurrence matrix, Random forest classifier*

## **Introduction**

Skin cancer caused due to ultraviolet rays from sun light. Mostly all color peoples are affected by skin cancer and people with light color have a higher risk because the sun burn easily in the light skin. Melanoma classification using contourlet transform And naïve bayes classifier is described in [1]. At first, the input dermoscopic images are extracted by non-subsampled contourlet transform. Finally, the naïve bayes classifier is used to classify the melanoma images.

Automatic skin lesion detection and classification using dermoscopy images is described in [2]. Initially, the dermoscopic images are preprocessed by median filter to remove the hairs. After that the segmentation is made by k-means clustering. Then the features are extracted by subregion, color and text features. Features are selected by wilkis lambda method. The skin image classification is made by Support Vector Machine (SVM) classifier.

Skin cancer detection and classification using SVM is discussed in [3]. The input skin images are preprocessed by histogram enhancement and noise filtering. Segmentation of skin images is made by thresholding. The features are extracted by statistical feature extraction, border, diameter, color, asymmetry and Gray-Level co-occurrence Matrix (GLM). Then the features are selected by principal component analysis. SVM technique is used to classify the skin images.

Classification of skin lesion by novel regularizer and Convolutional Neural Network (CNN) is discussed in [4]. The input skin images are extracted by different layers of CNN. Classification is made by novel regularizer. Skin cancer detection using snake model and SVM is discussed in [5]. Initially, the skin images are extracted by snake model algorithm. The skin cancer templates are classified by SVM. Automatic SCC using multiclass SVM and GLM is described in [6-8]. At first, the input skin images are converted into grayscale. The features are extracted by GLM and texture features. Multiclass SVM is used for classification.

Skin lesion analysis using image processing technique is discussed in [9]. The input skin images are extracted by geometric features like diameter, border and asymmetric parameters. Classification is made by k nearest neighbor classifier. SCC for a robust approach is discussed in [10]. Initially, the input skin images are segmented by k means clustering algorithm. The features are extracted by local binary patterns and color coherence vector. Classification is made by multiclass SVM.

Melanoma skin cancer detection using dermoscopy images is described in [11]. Preprocessing is made by Gaussian filter to remove hairs in the skin. The features like color, irregularity, area and texture features are extracted. SVM is used for classification. Skin cancer nuclear segmentation using histopathological images is described in [12]. Initially, preprocessing is made by contrast limited adaptive histogram enhancement technique. Fuzzy clustering means algorithm is used for segmentation. The features extraction is made by elliptical descriptor analysis.

Classification and segmentation of melanoma skin cancer images is discussed in [13]. Adaptive linear filter is used for pre-processing to remove hairs in the skin. K means clustering is used for segmentation. Features like border, asymmetry, diameter and color variation are extracted. SVM classifier is used for classification. Early detection of melanoma for skin lesion classification is discussed in [14]. At first, image acquisition is made to convert the image into grayscale. The GLM features are extracted. Multiclass SVM is used for classification. Segmentation is made by k-means clustering algorithm.

An efficient method for SCC using ranklet transform, RFCM and RFC is presented in this paper. The organization of paper is follows: Section 2 describes

the methods and materials used for SCC. In section 3 experimental results and discussion of SCC is presented. Section 4 concludes the SCC by ranklet transform based features and RFC.

## Methods and Materials

The overall workflow of SCC using ranklet features, RFCM and RFC is shown in figure 1. SCC system has three stages (i). Pre-processing, (ii). Feature extraction and (iii). Classification these stages are briefly explained below.

### Pre-processing using Wiener filter

The estimation of target random process for time invariant linear is made by wiener filter due to noisy process. It has an additive noise and noise spectra. Mean square error is minimized by the wiener filter between the random and desired process. Wiener filter is also used in processing of signals. Wiener filter used in other fields like image restoration [16] and video coding [15]. In this study, the input dermoscopic images are preprocessed by wiener filter to smooth the image and remove unwanted hairs in the skin.

### Feature extraction based on ranklet transform and RFCM

Ranklet transform is a non-parametric feature based on the non-parametric features and computation of Mann Whitney Wilcoxon in statistics domain. Ranklet has the similar form of Haar wavelets both has the same pattern of suitable notion of completeness, orientation selectivity and nature multiscale. The rank feature has high robustness in detecting the outliers and transformations like, gamma correction, brightness and contrast change.

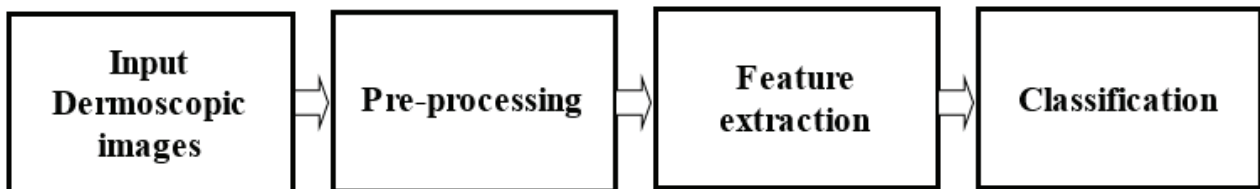


Figure 1 Block diagram of SCC system

The Wilcoxon ranklet transform is defined by,

$$R_s = \sum_{m=1}^K \pi_m L_m \tag{1}$$

where  $\pi_m$  is rank element of  $m$  and  $L_m$ . Then the Mann-Whitney statistics is defined by,

$$M = R_s - \frac{i(i+1)}{2} \tag{2}$$

where  $i$  is the no of treatment values. It is a statistical domain non-linear filter. The ranklets divides the image window  $Win$  into Two regions known as control and treatment which is shown in figure 2.

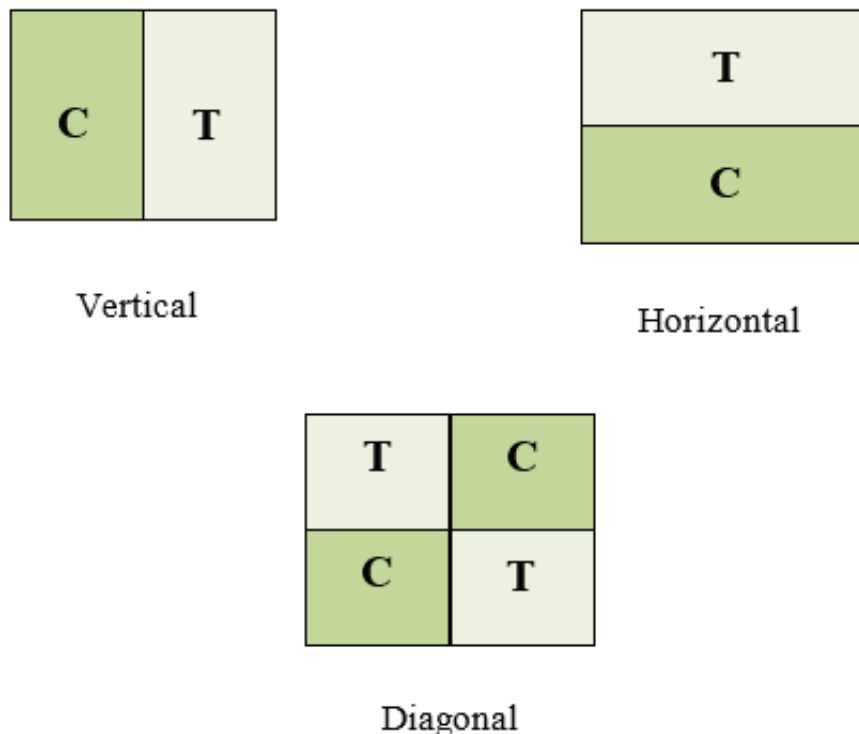


Figure 2 Image Windows ( $Win$ ) in ranklet transform

The rank test statistics are employed to identify the variations in the given region samples in  $Win$ . The variations in each region are replaced by ranking scores. There is a pairwise comparison between the control and treatment regions. The variations are shown by the horizontal, vertical and diagonal shapes. The ranklet transforms counts the no of control and treatment pairs. The treatment values are brighter than the control values. The co-occurrence matrix is a statistical method the neighborhood values are accessed by the central pixels. Ranklet transform is also used in digital watermarking [17], mammogram classification [18] and texture classification [19]. In this study, ranklet transform is used to decompose the preprocessed images and produces the

sub-band coefficients by different rank features. Then RFCM is used to extract the features and stored in the database for classification.

**RFC Classification**

RFC is an ensemble method for regression, classification and other tasks by constructing the multiple numbers of decision trees for the prediction of individual trees. Let us consider a decision tree combination of  $\{M(k, \phi_i)\}_{i=1}^R$ , where  $k$  is a input,  $\phi_i$  is denoted as the random split of independent vectors which has equal distribution to other trees in a forest  $\phi_1, \phi_2, \phi_3, \dots, \phi_{i-1}$ . Here,  $R$  is the bootstrap of training data. Every tree is built by different bootstrap samples. The RFC algorithm

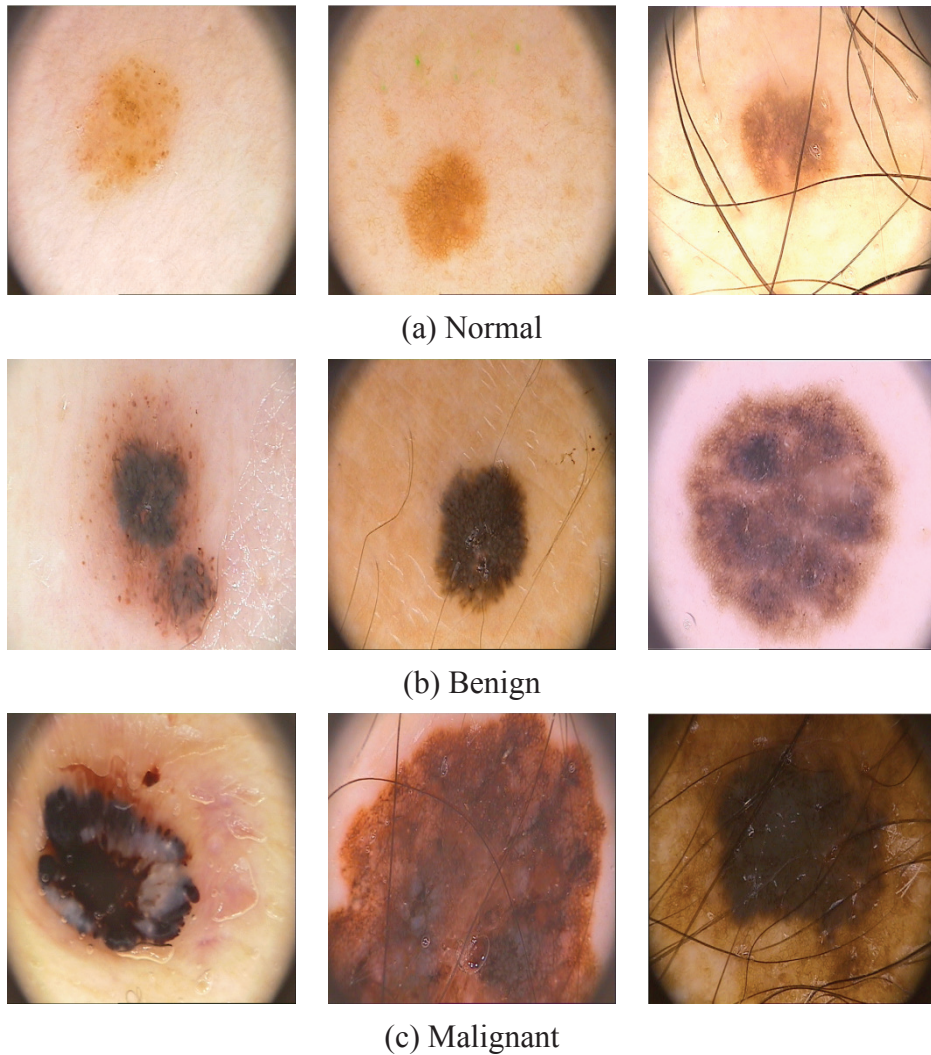
is defined by,

$$P(l)=1 - \sum_{i=1}^n s^2\left(\frac{i}{l}\right) \tag{3}$$

Here  $s^2\left(\frac{i}{l}\right)$  for  $i = 1, 2, 3, \dots, n$  is the probability estimation for splitting the node. RFC is also used in the glaucoma image classification [20] and brain tumor segmentation [21]. In this study RFC is used for the classification of skin images.

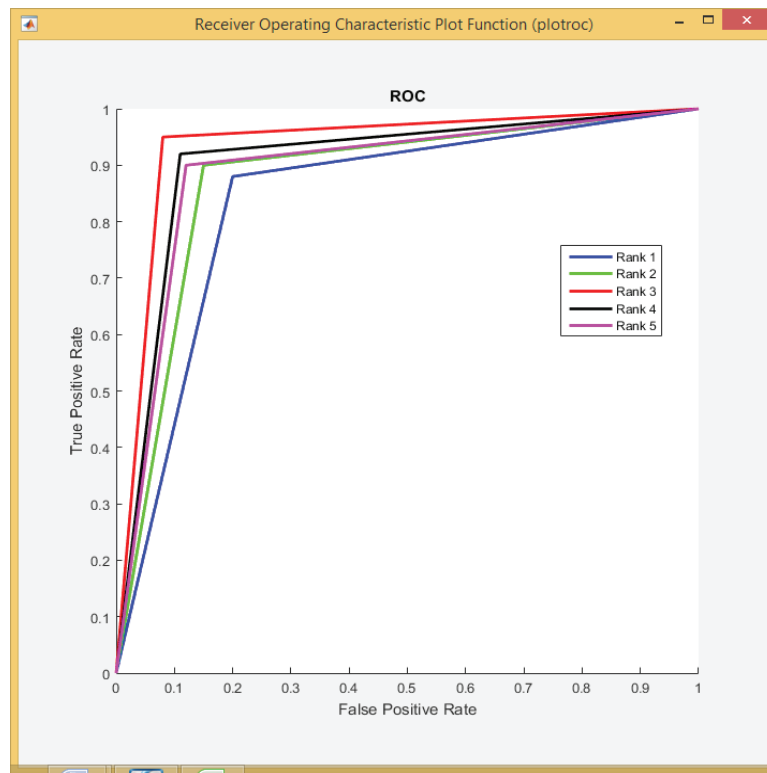
### Results and Discussion

The performance of SCC system is evaluated by PH2 database which is downloaded from PH2 database link [22]. Database contains 100 dermoscopic color image with lesions. The resolution of the image is 768x560 pixels. Figure 3 show the sample normal, benign and malignant images in the PH2 database.



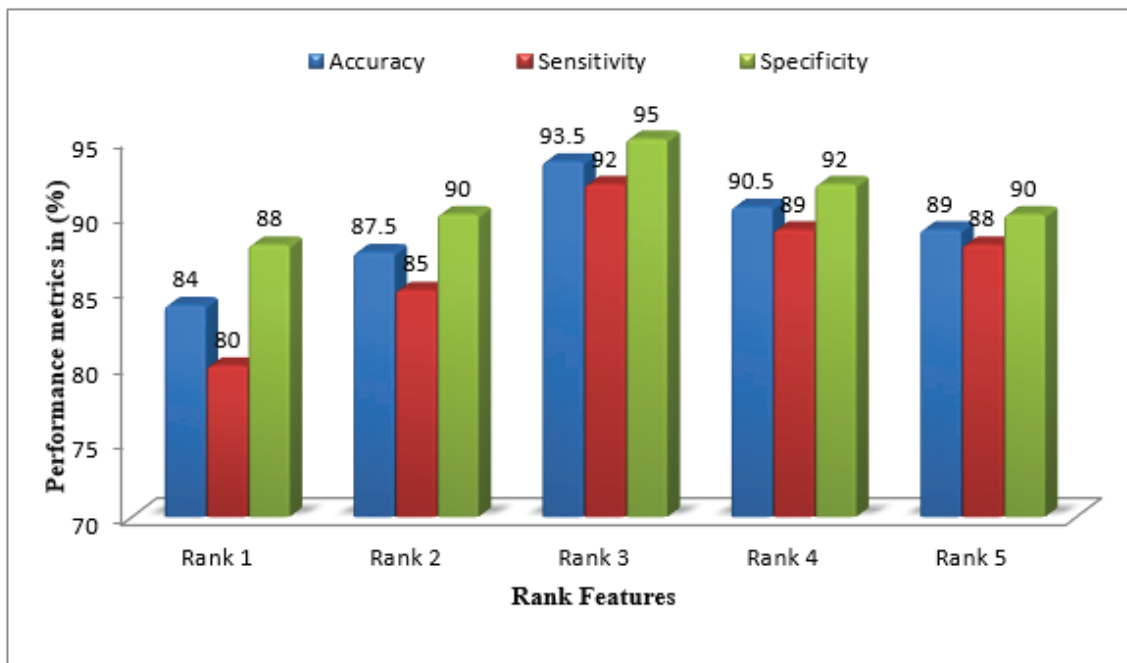
**Figure 3 Sample images in PH2 database**

The input skin images are preprocessed by wiener filter to remove noise and hairs in the skin. Then the ranklet transform is given to preprocessed images for decomposition. It produces sub-band coefficients with different rank features. These different rank features are extracted by RFCM and stored in the feature database. Finally, classification is made by RFC. Figure 4 shows the Receiver Operating Curve (ROC) for SCC system using RFCM and RFC.



**Figure 4 ROC for SCC using RFCM and RFC**

From the above figure it is observed that, maximum area of 0.93 is obtained by ROC curve for 3<sup>rd</sup> rank feature. The minimum area 0.84 is obtained by the 1<sup>st</sup> rank feature by using RFCM and RFC. Figure 5 shows the performance metrics of the SCC system.



**Figure 5 Performance metrics of SCC system**

It is observed from the above figure the maximum classification accuracy of SCC system is 93.5 % at the 3<sup>rd</sup> rank feature by using RFCM and RFC. Also, the sensitivity is 92 % and specificity is 95 %. The minimum classification accuracy is 84 % obtained at 1<sup>st</sup> rank feature by using RFCM and RFC and its sensitivity and specificity are 80% and 88%. The classification accuracy is increased at first 3 rank features and decreased at last 2 rank features.

### Conclusion

The SCC system using ranklet transform, RFCM and RFC is presented. The dermoscopic images in the PH2 database are used for performance evaluation. The input dermoscopic images are preprocessed by Wiener filter to remove noise and hairs in the skin. Ranklet transform is used to decompose the preprocessed images and produce subband coefficients with different rank features. Then RFCM method is used for extraction of subband coefficients with different rank features. These features are stored in the database and given as input for classification. The classification is made by RFC. Experimental results show the better classification accuracy of 93.5 % at the 3<sup>rd</sup> rank feature for SCC system.

**Ethical Clearance:** Taken from Annai Vailankanni Arts and Science College, Thanjavur, (Affiliated to Bharathidasan University)

**Source of Funding:** Self

**Conflict of Interest:** Nil

### References

1. Sonia, R. Melanoma image classification system by NSCT features and Bayes classification. *International Journal of Advances in Signal and Image Sciences*, 2016; 2(2), pp. 27-33.
2. Suganya R. An automated computer aided diagnosis of skin lesions detection and classification for dermoscopy images. In 2016 International Conference on Recent Trends in Information Technology (ICRTIT) 2016 Apr 8 (pp. 1-5). IEEE.
3. Alquran H, Qasmieh IA, Alqudah AM, Alhammouri S, Alawneh E, Abughazaleh A, Hasayen F. The melanoma skin cancer detection and classification using support vector machine. In 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT) 2017 Oct 11 (pp. 1-5). IEEE.
4. Albahar MA. Skin Lesion Classification Using Convolutional Neural Network With Novel Regularizer. *IEEE Access*. 2019 Mar 19;7:38306-13.
5. Bumrungkun P, Chamnongthai K, Patchoo W. Detection skin cancer using SVM and snake model. In 2018 International Workshop on Advanced Image Technology (IWAIT) 2018 Jan 7 (pp. 1-4). IEEE.
6. Maurya R, Singh SK, Maurya AK, Kumar A. GLCM and Multi Class Support vector machine based automated skin cancer classification. In 2014 International Conference on Computing for Sustainable Global Development (INDIACom) 2014 Mar 5 (pp. 444-447). IEEE.
7. Kumarapandian S. Melanoma Classification Using Multiwavelet Transform and Support Vector Machine. *International Journal of MC Square Scientific Research*. 2018 Sep 28;10(3):01-7.
8. Manikandan M. SKIN DETECTION UNDER VARYING ILLUMINATION. *International Journal of MC Square Scientific Research*. 2012 Dec 15;4(1):84-95.
9. Linsangan NB, Adtoon JJ, Torres JL. Geometric Analysis of Skin Lesion for Skin Cancer Using Image Processing. In 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (pp. 1-5). IEEE.
10. Singh SK, Jalal AS. A robust approach for automatic skin cancer disease classification. In 2016 1st India International Conference on Information Processing (IICIP) 2016 Aug 12 (pp. 1-4). IEEE.
11. Mane S, Shinde S. A Method for Melanoma Skin Cancer Detection Using Dermoscopy Images. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBE) 2018 Aug 16 (pp. 1-6). IEEE.
12. Ray PJ, Priya S, Kumar TA. Nuclear segmentation for skin cancer diagnosis from histopathological images. In 2015 Global Conference on Communication Technologies (GCCT) 2015 Apr 23 (pp. 397-401). IEEE.

13. Lynn NC, Kyu ZM. Segmentation and Classification of Skin Cancer Melanoma from Skin Lesion Images. In 2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT) 2017 Dec 18 (pp. 117-122). IEEE.
14. Sundar RS, Vadivel M. Performance analysis of melanoma early detection using skin lesion classification system. In 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT) 2016 Mar 18 (pp. 1-5). IEEE.
15. Trambadia S, Dholakia P. Design and analysis of an image restoration using Wiener filter with a quality based hybrid algorithms. In 2015 2nd International Conference on Electronics and Communication Systems (ICECS) 2015 Feb 26 (pp. 1318-1323). IEEE.
16. Siekmann M, Bosse S, Schwarz H, Wiegand T. Separable Wiener filter based adaptive in-loop filter for video coding. In 28th Picture Coding Symposium 2010 Dec 8 (pp. 70-73). IEEE.
17. Jan Z, Azeem N, Zahoor F. Digital watermarking using combination of Ranklets and wavelets. In 2015 First International Conference on Anti-Cybercrime (ICACC) 2015 Nov 10 (pp. 1-5). IEEE.
18. Masotti M. Exploring ranklets performances in mammographic mass classification using recursive feature elimination. In 2006 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing 2006 Sep 6 (pp. 265-270). IEEE.
19. Masotti M, Campanini R. Texture classification using invariant ranklet features. *Pattern Recognition Letters*. 2008 Oct 15;29(14):1980-6.
20. Ganeshbabu TR. Glaucoma Image Classification Using Discrete Orthogonal Stockwell Transform. *International journal of advances in signal and image sciences*. 2017 Jun 30;3(1):1-6.
21. Selvathi D, Selvaraj H. SEGMENTATION OF BRAIN TUMOR TISSUES IN MR IMAGES USING MULTIREOLUTION TRANSFORMS AND RANDOM FOREST CLASSIFIER WITH ADABOOST TECHNIQUE. In 2018 26th International Conference on Systems Engineering (ICSEng) 2018 Dec 18 (pp. 1-7). IEEE.
22. PH2 Database Link: <https://www.fc.up.pt/addi/ph2%20database.html>