Computer Aided Diagnosis System for Skin Lesions Detection using Texture Analysis Methods

Habeba Mahmoud[1], Mohamed Abdel-Nasser[1], and Osama A.Omer[1,2]

[1]Electrical Engineering Department, Aswan University, 81542 Aswan, Egypt

[2] Department of Electronics and Communications, Arab Academy for Science, Technology and Maritime Transport, Aswan, Egypt

***Abstract*—Malignant melanoma is considered as one of the most dangerous type of skin cancers as it increases the mortality rate. Computer-aided diagnosis (CAD) systems can help to detect melanoma early. In this paper, we propose a new skin melanoma CAD system using texture analysis methods. The proposed CAD system consists of four steps: hair removal, filtering, feature extraction and classification. In the feature extraction step, we evaluate the performance of five widely used texture analysis methods: grey level co-occurrence matrix, Gabor filters, a histogram of oriented gradients, local binary pattern and local directional number. Our CAD system classifies between non-melanoma skin lesions (represented as common nevi or dysplastic nevi) and melanoma. The experimental results show that extracting HOG features after hair removal yields the best classification results. HOG gives an AUC of 0.9783 with melanoma/common nevi classification and an AUC of 0.9439 with melanoma/dysplastic nevi classification.**

***Index Terms*—Melanoma, skin Lesion, Common Nevi, Dysplastic Nevi, Feature Extraction, Classification**

I. Introduction

Malignant melanoma is considered to be the most dangerous form of skin cancer. This type of skin cancer occurs when the human skin is exposed to the ultraviolet radiation (UV) emitted from sunshine or tanning beds, which caused the damage to skin cells [1]. The rate of melanoma increased in the last 50 years, specially on fair-skinned populations, which increased the death rate. This type of skin cancer increases with age because the studies refer to decreasing disease at the people within 25-44 years [2]. Moreover, people how have a lot of moles on their skins are at risk of melanoma. In the USA, about 68,130 cases were diagnosed as melanoma and about 8,700 patients died because of the disease in 2010 [1], but in 2016 the infected cases have been increased to reach 76,380 and the estimated death grew up to 10,130 [3]. However, observing the disease by both doctors and patients is very difficult, so corrected diagnose is very important to help in treatment [4]. The color of patient skin helps doctors to determinate the type of skin lesion, if the skin lesion is diagnosed as melanoma, its color could be black, brown, pink, red, purple, blue or white [5].

As it becomes hard to treat melanoma, the researchers found another ways beside the clinical diagnosis to help in accurate detection. The dermoscopy technique is high spread skin imaging way that helps in skin lesion detection [6]. At this technique, a dermatoscope device take an image, known as dermoscopic image, with a low level noise to examine the skin lesion by magnifying and filtering the infected part of skin [4]. Another aiding way to detect the skin lesion at an early stage is the computer aided diagnosis (CAD) system. The dermoscopic images of skin lesions have been classified by Gonzalez-Castro et al.[7] into benign lesions and melanoma by using color texture descriptor. Also, Reda and Karim [8] used dermoscopic images as a way of classification of melanoma and benign skin lesions depends on ABCD rule, which refers to (asymmetry, border irregularity, colour and dermoscopic structure). Ferris et al. [9] used a decision forest classifier with 54 geometrical features, colour and texture analysis methods, to classify skin lesions. The authors of [10] made a description of an automatic system for inspection of pigmented skin lesions. Global and local feature extraction was proposed in [11] to extract a different features of an image such as color, texture, shape and domain specific features.

Before start explaining our work in this paper, we should clarify the difference between two types of nevi: common nevi and dysplastic nevi. Common nevi, also called common mole, refers to the usual birthmark that appears on any part of the human skin, which considered to be benign skin lesion. Dysplastic nevi, also called atypical mole, refers to the unusual birthmark on human skin as it could be larger than common nevi and grows with time, its border may have edges, and its color ranges from pink to dark brown. People who have this type of nevi are at risk of getting melanoma [12]. Both of these types are classified as non-melanoma skin lesion.

In this paper, we propose a new melanoma CAD system. Removing hair from skin lesion dermoscopic images will be the first step to apply, then using median filter with a suitable dimensions will be the next step to do, after that five different feature extraction methods, such as Gabor filters, histogram of oriented gradients (HOG), grey level co-occurrence matrix (GLCM), local binary pattern (LBP) and local directional number pattern (LDN) will be using to extract special features from the images. The final step will be classification using a multi-layer perceptron (MLP) classifier to determine the skin lesion. The final results evaluate the classifier performance by calculating some metrics such as AUC and sensitivity.

The rest of this paper is organized as follows. Section II illustrates the four main steps in the proposed CAD system (hair removal, image filtering, feature extraction and classification). Section III describes the dataset and the experimental

978-1-5386-0879-1/18/$31.00

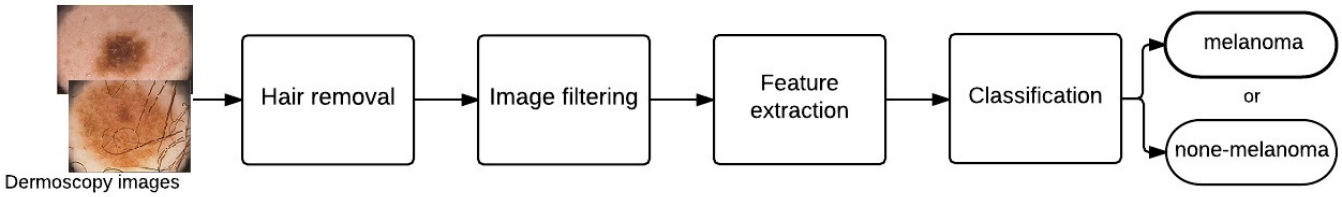

Fig. 1. Proposed CAD system

results. Finally, section IV provides the conclusion and the future work.

II. Proposed CAD System

Fig. 1 shows the steps of the proposed CAD system: hair removal, filtering, feature extraction, and classification. Below, we explain each step in detail.

A. Hair removal

When we take an image of the infected part of skin we observe that, there are some hair artifacts on that part of skin lesion. So, before working on the image we should remove hair from it to facilitate infected part detection. Therefore, we use morphological bottom-hat filter followed by closing operation as it remove the holes, fill them, from the image [13].

B. Image filtering

We used a 7×7 median filters, separately, for each channel of the colored image (RGB) to reduce the impact of the removed hair on lesion part [14]. After testing different filter sizes we found that 7×7 window size was the most appropriate. Fig. (2) shows the difference between the original image, the image after removing hair and the image after passing through the median filter.

C. Feature extraction

1) Gabor filters: This is the first chosen method for testing the metrics results, which able to extract features from our dermoscopic images. For each pixel of image [15], a bank of Gabor filters with 4 different scales and 6 orientations have been chosen to filter the dermoscopic image in spatial domain, thus we could extract the texture orientation from it. The following equation describes the kernal of the Gabor filter

$$g(x,y) = -\frac{1}{2\pi\sigma_x\sigma_y} exp[-\frac{1}{2}(\frac{x^2}{{\sigma_x}^2} + \frac{y^2}{{\sigma_y}^2})]cos(2\pi f x) \quad (1)$$

where, σ_x and σ_y represent the standard deviation values of x-axis and y-axis, respectively, and f represents the frequency information of the image. We calculate the mean (μ) and the standard deviation (σ) for each filtered image, and then concatenate them to form the feature vector of Gabor filters with the following length: 4 scales $\times$ 6 orientations $\times$ 2 (μ and σ) = 48

2) Grey level co-occurrence matrix (GLCM): This feature extraction method deals with the grey scale image, it is used to extract different features from the image by considering the relation between the reference pixel and its neighbors [11], depending on two parameters D and θ, where D refers to the distance (number of pixels) between the reference pixel and the neighboring pixel which equals one, and the parameter θ refers to the orientations of the neighboring pixel relative to the reference one, which are (0°, 45°, 90° and 135°). The size of GLCM matrix is considering as $N_x \times N_y$ [16], where N_x represents the number of rows and N_y represents the number of columns. To extract the GLCM features [17], the following steps have been applied in our CAD system:

- Dividing the dermoscopic image into four sub-regions.
- Reconstruct each sub-region with 8×8 matrix.
- Then, calculating 22 feature vector for each GLCM, these features illustrated in Fig. (3)
- After that, forming all features into one vector with a dimension of 352, calculated as: 4 sub regions $\times$ 4 orientations $\times$ one distance $\times$ 22 feature vector.

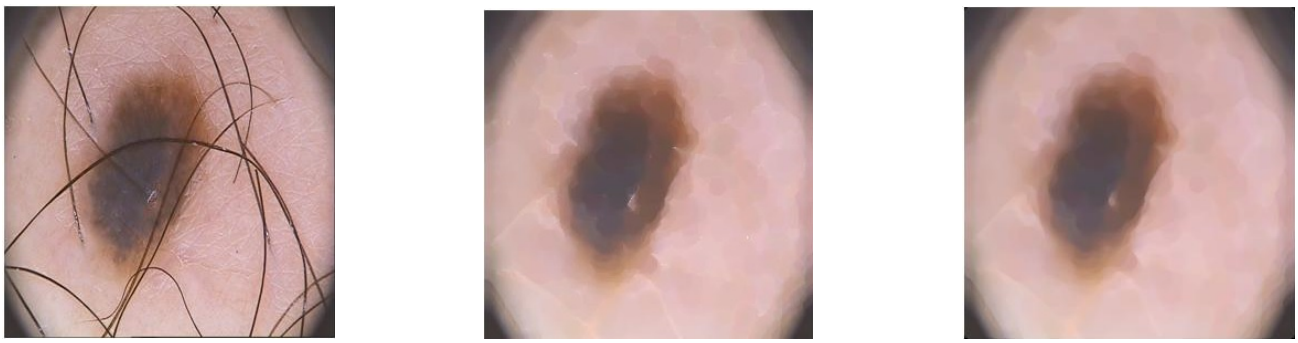

Fig. 2. Hair removal and filtered image results. (a) the original image, (b) image without hair, and (c) filtered image

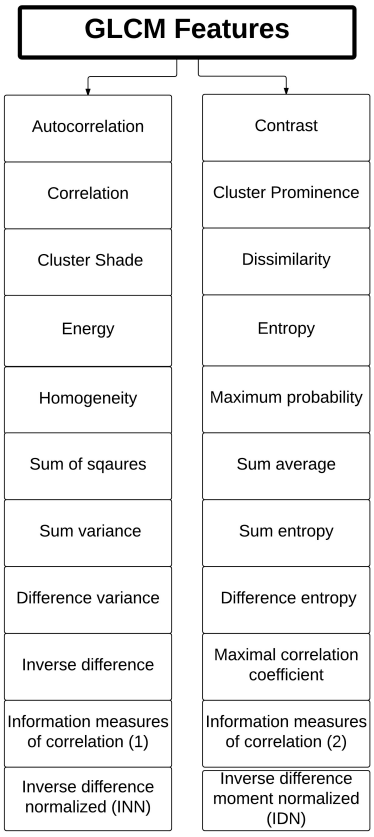

Fig. 3. GLCM Features

3) Histogram of oriented gradients (HOG): In order to their immovability to variance of illumination and their simplicity in performance, we used the HOG method [18]. This method divided the image into 4 blocks, with a dimension 4×4 cells for each block, each cell has a size of 8×8 pixels, shown in Fig. (4), then a 9-bin histogram (from 0 to 8) of gradient orientation is computed for each cell. The two neighboring blocks overlapping each other by 50%. For compensating the changing in illumination [17], the frequencies of all histograms are normalized, after that concatenated them into one vector with a dimension of 576, which equals to (8×8) pixels × (9-bit) histogram.

4) Local binary pattern (LBP): In LBP, the image is divided into blocks, each pixel in the block is surrounded by a window with a size 3×3. These eight pixels in the window are compared with the central pixel [19]. If the pixel value is greater than or equal the central pixel, it sets as 1, otherwise it sets as 0. Therefore, each pixel is represented by 8-bit, then the decimal value of the 8-bit represents local structural information about the central pixel [20], after that the histograms of LBP codes have been computed to represent the final feature vector of LBP. LBP pattern could be a uniform pattern or a non-uniform pattern which is mostly caused by the image noise [21]. If the transition in the pattern occurs at most two times from 0 to 1 or from 1 to 0, the pattern is considered to be uniform. For example, 00100000 is a uniform LBP, but the pattern 11011100 is a non-uniform LBP. LBP histogram have 58 bins for the uniform patterns and one bin for all the non-uniform patterns, so the whole dimension of histogram equal 59 bins.

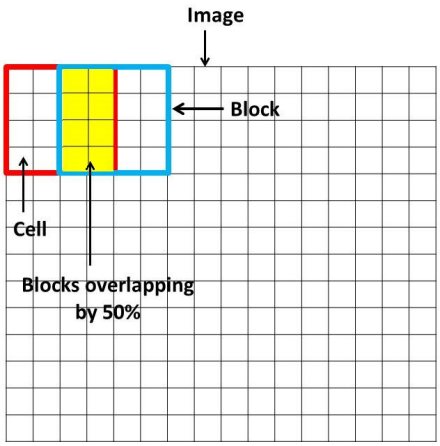

Fig. 4. Description of HOG method

5) Local directional number pattern (LDN): This method generate a 6-bit binary code for each pixel of the dermoscopic image [22]. Because of the sensitivity of edge to the lighting, Kirsch compass mask has been used to compute the edge response in eight different directions by taking the maximum positive number and the minimum negative number for the neighbor pixels which collected to generate the 6-bit code. The dimension of feature vector equals to 256 (8×8 mask × 4 blocks).

D. Classification

In the classification step, we usd a multilayer perceptron (MLP) classifier to distinguish data which collected from the previous texture extraction methods. The MLP classifier is formed by a neural network [23], which consists of three layers:

- The input layer, which contains the input data. The number of its neural nodes depends on the number of input feature vectors.
- The hidden layer, could be more than one, located between the input layer and the output layer, its function is to receive the weights from the input nodes to encode and deliver them to the output nodes.
- The output layer, has two neural nodes, related to the number of neural nodes that classify between positive (abnormal) and negative (normal) cases. Where positive cases refer to melanoma skin lesion and negative cases refer to common nevi or dysplastic nevi (non-melanoma).

III. Results and Discussion

A. Dataset description

In our paper, we use a PH^2 dermoscopic images with a resolution of 768x560 pixels, were obtained at the Dermatol-

ogy Service of Hospital Pedro Hispano, available in ADDI website [24]. This dataset consists of 200 images divided into 80 common nevi, 80 dysplastic nevi and 40 melanomas with colors of white, red, light-brown, dark-brown, blue-gray and black. This dataset is available for research and educational purposes.

B. Evaluation

Before describing our results we should illustrate that the MLP classifier has been evaluated by sensitivity and specificity [25]. Sensitivity indicates to the percentage of the correct detection of malignant melanoma (the percentage rate of true positive detection). Specificity indicates to the percentage of the correct detection of the non-melanoma skin lesion (the percentage rate of false positive detection). Both of sensitivity and specificity could be computed by using the following concepts [13]:

- True positive (TP): positive cases classified as positive.
- True negative (TN): negative cases classified as negative.
- False positive (FP): negative cases classified as positive.
- False negative (FN): positive cases classified as negative.

We also use the AUC, ACC, Precision, Recall and F-Score [26] to evaluate the performance of our CAD system. AUC refers to Area Under the Curve, where this curve describes the relation between sensitivity and specificity, which called the Receiver Operating Characteristic (ROC) summarized by the AUC number, the more ROC curve is closer to the top-left corner of figure, the better classification is performed. ACC measure the correct decision (accuracy) of the classifier. Precision could be defined as positive prediction value, Recall is the same as sensitivity, F-Score is the measure of classifier effectiveness which depends on both precision and recall as shown in the F-Score equation. The following equations describe the above-mentioned metrics:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Precision(P) = \frac{TP}{TP + FP} \tag{3}$$

$$Recall(R) = \frac{TP}{TP + FN} \tag{4}$$

$$F - Score = \frac{2 * P * R}{P + R} \tag{5}$$

C. Experimental results

The experimental show the analysis of the MLP classifier for each method of feature extraction which we used in our paper. The results divided into two parts, the first part is before removing hair from the dermoscopic images, the second part is after removing hair from the dermoscopic images. On the other hand, each part in the results consist of two tables, one classified between melanoma (the positive case) and common nevi (the negative case) and the other classified between melanoma (the positive case) and dysplastic nevi (the negative case).

1) Without hair removal method: Table I gives the results of melanoma/common nevi classification with a good rate values of sensitivity (Recall) and AUC when we used HOG, LBP and LDN with the MLP classifier. Table II shows the results of melanoma/dysplastic nevi classification which also indicates that HOG, LBP and LDN have good rate values with sensitivity and AUC. Moreover, we deduced that the CAD system gives the best results with HOG feature extraction method.

2) With hair removal method: Table III illustrates the results of melanoma/common nevi classification, HOG method produced very good AUC rate and sensitivity with optimal precision. As well, Table IV shows the results of melanoma/dysplastic nevi with a very good results of AUC and sensitivity when using HOG and LDN feature extraction with the MLP classifier.

TABLE I
RESULTS USING MLP CLASSIFIER BETWEEN MELANOMA(POSITIVE CASE) AND COMMON NEVI(NEGATIVE CASE) BEFOR HAIR REMOVAL

Methods	AUC	ACC	Recall	Precision	F-score
Gabor	0.5432	0.6778	0.6755	1	0.8035
GLCM	0.4982	0.6633	0.6633	1	0.7975
HOG	0.9356	0.8667	0.891	0.9186	0.9015
LBP	0.9169	0.8667	0.8938	0.9071	0.8989
LDN	0.8045	0.7	0.7008	0.965	0.8101

TABLE II
RESULTS USING MLP CLASSIFIER BETWEEN MELANOMA(POSITIVE CASE) AND DYSPLASTIC NEVI(NEGATIVE CASE) BEFOR HAIR REMOVAL

Methods	AUC	ACC	Recall	Precision	F-score
Gabor	0.5403	0.6667	0.6954	0.9375	0.7945
GLCM	0.5002	0.6629	0.6628	1	0.7972
HOG	0.928	0.85	0.8757	0.9188	0.8906
LBP	0.8601	0.8	0.8252	0.8962	0.8556
LDN	0.7634	0.65	0.6525	0.9573	0.7734

TABLE III
RESULTS USING MLP CLASSIFIER BETWEEN MELANOMA(POSITIVE CASE) AND COMMON NEVI(NEGATIVE CASE) AFTER HAIR REMOVAL

Methods	AUC	ACC	Recall	Precision	F-score
Gabor	0.9318	0.9111	0.8995	0.9818	0.9382
GLCM	0.9265	0.9	0.9046	0.9583	0.93
HOG	0.9783	0.9667	0.9557	1	0.9765
LBP	0.9476	0.8722	0.8829	0.9191	0.8986
LDN	0.9273	0.8833	0.9484	0.8686	0.9045

From the results we notice that the hair removal step is very important in this CAD system because it applied good performance of MLP classifier. The HOG feature extraction obtained the best metrics results because it deals with the gradient of the image which is less sensitive to noise, these results appeared as 97.8% AUC with common nevi and 95.4% AUC with dysplastic nevi. The accuracy of LBP and LDN were less than the other texture extraction methods given

TABLE IV
RESULTS USING MLP CLASSIFIER BETWEEN MELANOMA(POSITIVE CASE) AND DYSPLASTIC NEVI(NEGATIVE CASE) AFTER HAIR REMOVAL

Methods	AUC	ACC	Recall	Precision	F-score
Gabor	0.9017	0.8278	0.8098	0.9652	0.8785
GLCM	0.9136	0.8389	0.8779	0.8837	0.8783
HOG	0.9439	0.95	0.9504	0.9854	0.9665
LBP	0.8878	0.8722	0.8518	0.9544	0.8975
LDN	0.9056	0.8556	0.9217	0.8558	0.8821

values of 87.2% and 88.3%, respectively, with common nevi because they deal with pixels value which made them more sensitive to noise.

IV. CONCLUSION AND FUTURE WORKS

In this paper, we have proposed a new CAD system to differentiate between malignant melanoma and non-melanoma skin lesions. Our dataset consists of three groups of dermoscopic images, where 40 are melanoma, 80 are common nevi and 80 are dysplastic nevi. We have removed hair from these dermoscopic images to improve the classifier performance, and then a median filter with 7×7 window has been used to reduce artifacts after removing hair. Several textural analysis methods, such as Gabor filters, GLCM, HOG, LBP and LDN have been used in the texture analysis step. MLP classifier has been chosen to make a decision with the image case (positive case or negative case). The best performance of the proposed CAD system is achieved when we used HOG as a feature extraction. HOG obtained an AUC of 0.9783 with melanoma/common nevi classification task and an AUC of 0.9439 with melanoma/dysplastic nevi classification. Our future work will focus on using several combinations of texture analysis methods to further improve the results of the proposed CAD system.

REFERENCES

[1] Simes, M. C. F., J. J. S. Sousa, and A. A. C. C. Pais. "Skin cancer and new treatment perspectives: A review." Cancer letters 357.1 (2015): 8-42.

[2] Erdmann, Friederike, et al. "International trends in the incidence of malignant melanoma 19532008are recent generations at higher or lower risk?." International journal of cancer 132.2 (2013): 385-400.

[3] Siegel, Rebecca L., Kimberly D. Miller, and Ahmedin Jemal. "Cancer statistics, 2016." CA: a cancer journal for clinicians 66.1 (2016): 7-30.

[4] Sun, Xiaoxiao, et al. "A benchmark for automatic visual classification of clinical skin disease images." European Conference on Computer Vision. Springer International Publishing, 2016.

[5] Wolff, Klaus, and Richard Allen Johnson. Fitzpatrick's color atlas and synopsis of clinical dermatology. McGraw Hill, 2009.

[6] Faal, Maryam, et al. "Improving the diagnostic accuracy of dysplastic and melanoma lesions using the decision template combination method." Skin Research and Technology 19.1 (2013).

[7] Gonzalez-Castro, Victor, et al. "Automatic classification of skin lesions using color mathematical morphology-based texture descriptors." Twelfth International Conference on Quality Control by Artificial Vision. Vol. 9534. No. 953409. SPIE. Digital Library, 2015.

[8] Kasmi, Reda, and Karim Mokrani. "Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule." IET Image Processing 10.6 (2016): 448-455.

[9] Ferris, Laura K., et al. "Computer-aided classification of melanocytic lesions using dermoscopic images." Journal of the American Academy of Dermatology 73.5 (2015): 769-776.

[10] Alcn, Jos Fernndez, et al. "Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis." IEEE journal of selected topics in signal processing 3.1 (2009): 14-25.

[11] Kavitha, J. C., A. Suruliandi, and D. Nagarajan. "Melanoma Detection in Dermoscopic Images using Global and Local Feature Extraction." (2017).

[12] Eliades, Philip, and Hensin Tsao. "New Insights into the Molecular Distinction of Dysplastic Nevi and Common Melanocytic NeviHighlighting the Keratinocyte-Melanocyte Relationship." Journal of Investigative Dermatology 136.10 (2016): 1933-1935.

[13] Chatterjee, S., D. Dey, and S. Munshi. "Mathematical morphology aided shape, texture and colour feature extraction from skin lesion for identification of malignant melanoma." Condition Assessment Techniques in Electrical Systems (CATCON), 2015 International Conference on. IEEE, 2015.

[14] Kleefeld, Andreas, et al. "Adaptive Filters for Color Images: Median Filtering and Its Extensions." CCIW. 2015.

[15] Rangayyan, Rangaraj M., and Fbio J. Ayres. "Gabor filters and phase portraits for the detection of architectural distortion in mammograms." Medical and biological engineering and computing 44.10 (2006): 883-894.

[16] Shabat, Abuobayda M., and Jules-Raymond Tapamo. "A comparative study of the use of local directional pattern for texture-based informal settlement classification." Journal of Applied Research and Technology (2017).

[17] Abdel-Nasser, Mohamed, Antonio Moreno, and Domenec Puig. "Towards cost reduction of breast cancer diagnosis using mammography texture analysis." Journal of Experimental and Theoretical Artificial Intelligence 28.1-2 (2016): 385-402.

[18] Bakheet, Samy. "An SVM Framework for Malignant Melanoma Detection Based on Optimized HOG Features." Computation 5.1 (2017): 4. APA

[19] Stella, X. Arockia, TamilNadu Sivagangai, and India Dr N. Sujatha. "Performance Analysis of GFE, HOG and LBP Feature Extraction Techniques using kNN Classifier for Oral Cancer Detection." Journal of Network Communications and Emerging Technologies (JNCET) www. jncet. org 6.7 (2016). APA

[20] Kaya, Ylmaz, et al. "1D-local binary pattern based feature extraction for classification of epileptic EEG signals." Applied Mathematics and Computation 243 (2014): 209-219.

[21] Ren, Jianfeng, Xudong Jiang, and Junsong Yuan. "Noise-resistant local binary pattern with an embedded error-correction mechanism." IEEE Transactions on Image Processing 22.10 (2013): 4049-4060.

[22] Rivera, Adin Ramirez, Jorge Rojas Castillo, and Oksam Oksam Chae. "Local directional number pattern for face analysis: Face and expression recognition." IEEE transactions on image processing 22.5 (2013): 1740-1752.

[23] Islam, Md Shafiqul, et al. "Solid waste bin detection and classification using dynamic time warping and MLP classifier." Waste management 34.2 (2014): 281-290.

[24] Mendona, Teresa, et al. "PH 2-A dermoscopic image database for research and benchmarking." Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE. IEEE, 2013.

[25] Matthews, S., et al. "Predictive values, sensitivity and specificity of abdominal fluid variables in determining the need for surgery in horses with an acute abdominal crisis." Australian veterinary journal 80.3 (2002): 132-136.

[26] Mesaros, Annamaria, Toni Heittola, and Tuomas Virtanen. "Metrics for polyphonic sound event detection." Applied Sciences 6.6 (2016): 162.