Diyala Journal for Pure Science

A Hybrid System for Classification of Skin Cancer Images Using Artificial Neural Network and Support Vector Machine

Raghad Majeed Azawi

College of Medicine – University of Diyala

mraghad699@gmail.com

Received: 20 August 2020 **Accepted: 8 October 2022**

DOI: https://dx.doi.org/10.24237/djps.1804.667A

Abstract

Skin cancer is a very serious disease and identifying it early is therefore very important to take the appropriate treatment before the situation worsens. Melanoma cancer is one of the most prevalent malignant types in recent years because it spreads to the rest of the body quickly, it must be detected early and treated. Computer Vision in medical imaging analysis is important in the diagnosis of various diseases at many stages. These methods give a high-resolution assessment of the disease by analyzing the images digitally. The steps followed in this paper involve collecting an image dataset from a global website for skin cancer about 1500 photos of different sizes, preprocessing, segmentation using threshold, feature extraction for Asymmetry, Border, Color, Diameter, (ABCD), etc., calculating the total of dermatoscopy score and then classification using Artificial Neural Network (ANN) and Support Vector Machine (SVM) in the same of the stage. The results of the system display that the attained classification accuracy is 98.38%.

Keywords: Hybrid System, Artificial Neural Network, Support Vector Machine, Skin Cancer, Computer Vision, threshold, Texture Analysis.

Diyala Journal for Pure Science

A Hybrid System for Classification of Skin Cancer Images Using Artificial Neural Network and Support Vector Machine

Raghad Majeed Azawi

نظام هجين لتصنيف صور سرطان الجلد باستخدام

رغد مجيد عزاوي

كلية الطب – جامعة ديالى

الخلاصة

يعتبر سرطان الجلد مرض خطير للغاية، ولذلك فإن تحديده مبكراً أمر مهم للغاية لاتخاذ العلاج المناسب قبل أن يزداد الوضع سوءاً. سرطان الجلد هو واحد من أكثر أنواع الأورام الخبيثة انتشارا في السنوات الأخيرة لأنه ينتشر بسرعة إلى بقية اجزاء الجسم، ويجب اكتشافه مبكراً وعلاجه. تقوم رؤية الحاسوب Computer Vision في تحليل الصور الطبية بشكل دقيق وكذلك تشخيص الأمراض المختلفة في مراحل عديدة منها. تعطي هذه التقنية تقييماً عالي الدقة لهذا المرض من خلال تحليل الصور رقمياً. تتضمن الخطوات المتبعة في هذه الورقة جمع مجموعة من بيانات الصورلسرطان الجلد من موقع عالمي خاص بسرطان الجلد حوالي 1500 صورة مختلفة الاحجام، والمعالجة المسبقة، والتجزئة باستخدام threshold واستخراج الميزات لعدم التماثل، والحدود، واللون، والقطر، (ABCD)وما إلى ذلك، وحساب مجموع نقاط تنظير الجلد ثم التصنيف باستخدام الشبكة العصبية الاصطناعية (ANN) و (SVM)في نفس المرحلة اظهرت نتائج النظام أن دقة التصنيف التي تم تحقيقها هي ٪98.38.

الكلمات المفتاحية:

Introduction

In the past years, the incidence of skin cancers has increased in the world and a large percentage of these people have a chance of recovering from the disease [1]. The most dangerous type is melanoma because it can easily become infected the skin and its incidence has continued to rise in the past 3 decades in several countries of the world [2]. To identify melanoma skin cancer, various papers and research studies that are used in the detection of this kind of disease have been proposed with medical image processing techniques to become effective and approved diagnostic tools for the early detection of patients and to give the appropriate treatment to the condition [3]. After obtaining the infected images, these methods give an effective diagnostic tool to compare the patient's condition over time and provide clear images for explanation and teaching which may be economically useful in hospitals [4].

There are 2 kinds of skin malignant tumors, “Basel cell cancer and squamous cell carcinoma”, these two don't grow rapidly in the skin, and the 3’rd kind is melanoma cancer growth quickly, in addition to the presence of 4 kinds of benign skin tumors [5]. In the classification schemes and analysis of the features of medical images found the best ANN and SVM algorithms [6].

Literature Review

The use of artificial intelligence and image processing techniques are precise systems to detect the location of the tumor in the skin. The researchers developed a set of automatic detection systems for skin cancer for easier clinical treatment later.

In 2021, Balaji et al. [7] utilized the NNA and analyzed it into several tools to improve the skin cancer classification system across an optimal strategy. The scheme includes segments by “fuzzy c-means”, then feature selection by firefly-optimized. The aim of typical was calculated in terms of accuracy, sensitivity, specificity, and satisfactory results were obtained at 96.1%. In 2021, Thomas et al. [8] utilized simulation technology in deep learning DL at the same time to segment the skin image and classify malignant melanoma.

They have made it into 12 types according to the images given. These classifier modules by networks achieved (> 93%) accuracy while the entire tissue was classified. In this scheme, the way utilized was able to achieve routine pathologist tasks, such as surgical margin clearance assessments, etc.

In 2021, Tumpa and Kabir [9] In their work, they proposed a technique for developing a neural network to intelligently classify melanin cancers. The steps begin with image processing and optimization across gradient intensity. Then you segmented the image by the threshold way to filter the important area within it. Then calculate the features of LBP, and GLCM to train the NNA in the next step. An effective 97% accuracy of the neural network was achieved for the global ISIC dataset. This project has proven to be highly accurate as it involves analyzing the information more clearly. In 2020, Shunichi et al. [10].

They proposed an intelligent scheme for classification through CNN with DL and compared several results from this scheme with the techniques utilized at the time and the intelligent method got the highest accuracy 94% and percentage of results so it was actually used in hospitals.

In 2019, Kumar et al. [11] Suggest a method that detects the presence or absence of cancer in the skin. In this study the following steps were adopted: preprocessing for images was utilized to assemble categorized data before they were flattened and subjected to pixels extraction. The extracted pixel intensities could then be collected into an array and kept in a definite dataset.

An SVM with categorized data utilized a convenient kernel to effectively classify the tasters by trusting previously training data. The suggested scheme presented (90%) accuracy for classification.

The Proposed System

The goal of developing the proposed scheme is in to help in the right diagnosis of illustrated skin diseases and then verify the extracted results. The image is processed from noise and then segment the image before extracting features from it.

A hybrid classifier (ANN-SVM) was selected for the final classification stage. The hybrid classifier scheme of SCs identified as an illustration in figure (1). In this work, all of the research implements and techniques backgrounds are deliberated in the next sections.

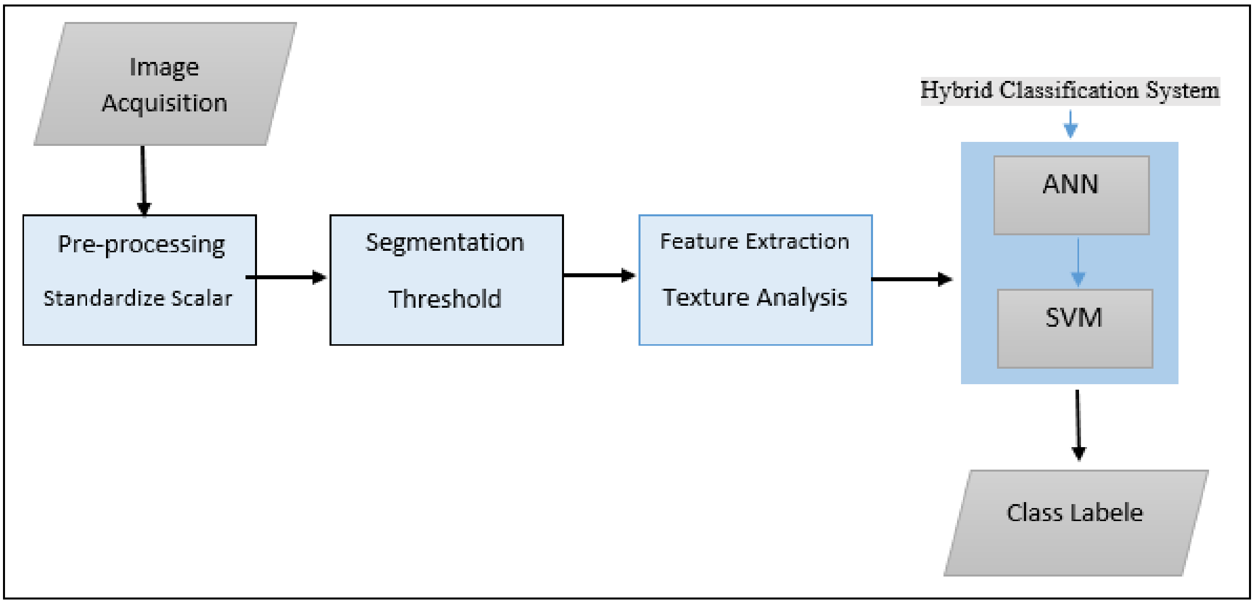


Figure1: Hybrid System of Skin Cancer Classification.

Image Acquisition Stage

A collection of 1500 skin images was obtained from the International Skin Imaging Collaboration (ISIC) Archive [12] and American Cancer Society ACS [13]. In this system skin images with .jpeg extension was used.

The images used are colored and in different sizes. In this study, there are three types of abnormal skin cancer for each type (300 images) and four types of normal skin images for each type (150 images) the figure (2) shows that.

(A)

(B)

Figure 2: (a) Malignant Skin Tumor and (b) Benign Skin Tumor.

Pre-processing

In this manner modifying the image from noise, and distortion, adjusting of brightness of the image and focusing on the important data it contains. First, they are resized to (256*256) for support and prominence of the desired features. Then it converts from RGB to gray level for easy representation in the next step. Approaches in re-processing phases as per remove-elements or characteristics missing value, Standardizations Scalar (SS), in Max Scaling have been useful to the dataset. The SS defenses that individual feature has in the collection between (0,1) variance, attractive exact features for the identical coefficient. Also, in Min-Max Scalar transfers the data analysis such that very features are between “0 & 1”. It lost values feature row is fair canceled as of the dataset [14]. In this work, techniques were used for processing as displayed in figure (3).

(a): RGB image **(b): Gray Image** **(c): After SS Applied**

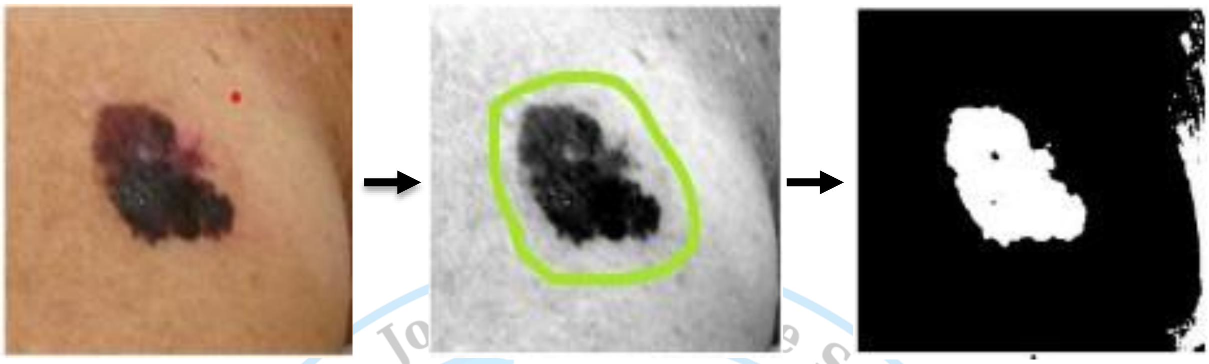
Figure 3: Pre-processing Stage

Segmentation Image

The purpose of image segmenting is to modify and simplify its recommendation and accurately analyze the task data. In computer vision techniques, it is the segmentation of an image into many images to determine the object, curves, and borders in the image.

More precisely, segmentation is the way to assign an attribute to each pixel and give similar pixels a uniform value. This method is called the threshold where the values are applied to other pixels and then the dynamic threshold value is adjusted with the adaptability of the image as displayed in figure (4).

At this specific stage the adaptive threshold is used to remove the skin from the main image [15].



(a): Original Image **(b): Gray Image** **(c): Segment Image**

Figure 4: After Segmentation Step.

Feature Extraction

Finding the known qualities for texture in images is the 1'st phase towards construction calculated modeling of texture. As for the character's possessions in image concentration variation, the texture possessions are typically observable of interior physical characters. Texture involves spatial dissemination for gray levels. Thus, 2-dimension histograms (co-occurrence matrices) are suitable texture analysis tools. As shown in the equations below [16].

The homog. (IDE) $= \sum_i \sum_j \frac{1}{1+(i-j)^2}\, p\,(i,j)$ .... **(1)**

Energy (ASM) $= \sum_i \sum_j \; p(i,j)$ .... **(2)**

Correlation $= \frac{\sum_i \sum_{j\,(i,j)} p(i,j)\, \mu x\, \mu y}{\sigma x\, \sigma y}$ .... **(3)**

Where σ_x, σ_y, μ_x and μ_y are the standard deviations and means of p_x and p_y.

$\mu x = \sum_i \sum_j \; p(i,j)$ .... **(4)**

$\sigma_x = \sum_a (a- \mu x)^2 \sum_b p(a - b)$ .... **(5)**

Entropy $= -\sum_i \sum_j p(i,j) \log_2 p(i,j)$ .... **(6)**

Contrast $= \sum_i \sum_j p(i,j) * (i-j)^2$ .... **(7)**

Where i, j, a, b, x, and y are counters, and p (i,j) is a matrix of the position of pixels in an image.

A Hybrid Classification System

In order to obtain classifications of all kinds of skin diseases quickly and with high quality, whether natural or abnormal diseases, the hybrid system is briefly discussed in this part of this study.

ANN-SVM Algorithms

In this proposed scheme, the mix of (ANN-SVM) algorithms is utilized in recognized and categorize the SCs analyses in several ways. In the last step of the proposal ANN is utilized to analyze the features from the previous step and then convert data to SVM to classify it into kinds of dermatology (SCs) thru a linear SVM kernel. In her SVM kernel will stay hyper flat mid collections into data views thru the approach of the conclusion border. In this process, there are two categories for skin cancer, a malignant category and a benign category that were determined by mixing two algorithms [17].

ANN is a mathematical representation and a set of layered neurons that allow messages to pass through under the supervision ML. It consists of three components: inputs, outputs, and vectors across a given function [15]. In the network, the input substances give special values and weights, and it's transformed over the training scheme. The outputs are data analysis, determining the item and calculating weights again using the error edge or boundary amid the output into predicted or real class. ANN calculates and outputs results over the multi-layer neuron cell [17].

SVM is an ML classifier that has been usually applied to solve classification difficulties. It exploited a max-edge style that transmutes for determining some complex quadratic program problems. For the high action of SVM in a classification manner, various requirements are normally useful. After analyzing and simplifying the data from ANN_algorithm it is converted to SVM the final classification and verified it's labeled [16].

System Performance Evaluation

Several measures should be used to verify the efficiency of the scheme. In this study the use of confusion 2×2-matrix, in a test case, the result of the classification is placed in a box according to a mathematical equation. Furthermore, it provides two kinds of right prediction of the classifier and two varieties of classifier of an incorrect prediction. The accurate illustrations of the overall presentation of the scheme are as follows:

$$System\ accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \qquad (8)$$

Sensitivity (Sn): it is the ratio of the lately classified SCs images to the total number of SCs patients. The Sn of the classifier for detecting positive instances is known as the "true positive rate." It can be written as follows [18]:

$$Sn\ or\ positive\ rate = \frac{TP}{TP + FN} \times 100\% \qquad (9)$$

Specificity (Sp): an analytic test is negative and the person is healthy and is arithmetically written as follows [19]:

$$Sp = \frac{TN}{TN + FP} \times 100\% \qquad (10)$$

Where TN is True Negative, TP is True Positive, FN is False Negative, and FP is False Positive.

Results and Discussion

This step includes discussions of detection schemes and products of dissimilar perspectives. All image samples were taken from two international websites ISIC and ACS Archive [12, 13]. To increase the efficiency of the system's performance, 30% of the samples were used to train the hybrid system, and the remaining 70% to test algorithms. By utilizing sample images, the scheme extracts the features of an image and classifies the cancerous (thee types) and non-cancerous skin (four types) and segments.

The System Performance

System performance checked for hybrid algorithms such as ANN and SVM, on the skin cancer dataset and utilized features extraction tools on the segment image data store to check the accuracy performance of algorithms performance assessment metrics were useful. All features were regularized and consistent beforehand relating to the algorithm as shown in tables (1) and (2) and figures (5). All computations were done in MATLAB on Intel(R) Core i5 -1500CPU @2.10 GHz.

Table 1: Test with Cancerous Skin Images

IMAGE TYPE	NO. OF IMAGES	SUCCESSFULLY	ACCURACY (%)
Melanoma	210	207	98.57%
Basal Cell Carcinoma	200	200	100%
Squamous Cell Carcinoma Skin cancer	220	216	98.18%
Total	630	623	98.91%

Table 2: Test with Normal Skin Images

IMAGE TYPE	NO. OF IMAGES	SUCCESSFULLY	ACCURACY (%)
Normal Skin type-1	105	101	96.19%
Normal Skin type-2	105	100	95.23%
Normal Skin type-3	105	105	100%
Normal Skin type-4	105	105	100%
Total	420	411	97.85%

Diyala Journal for Pure Science

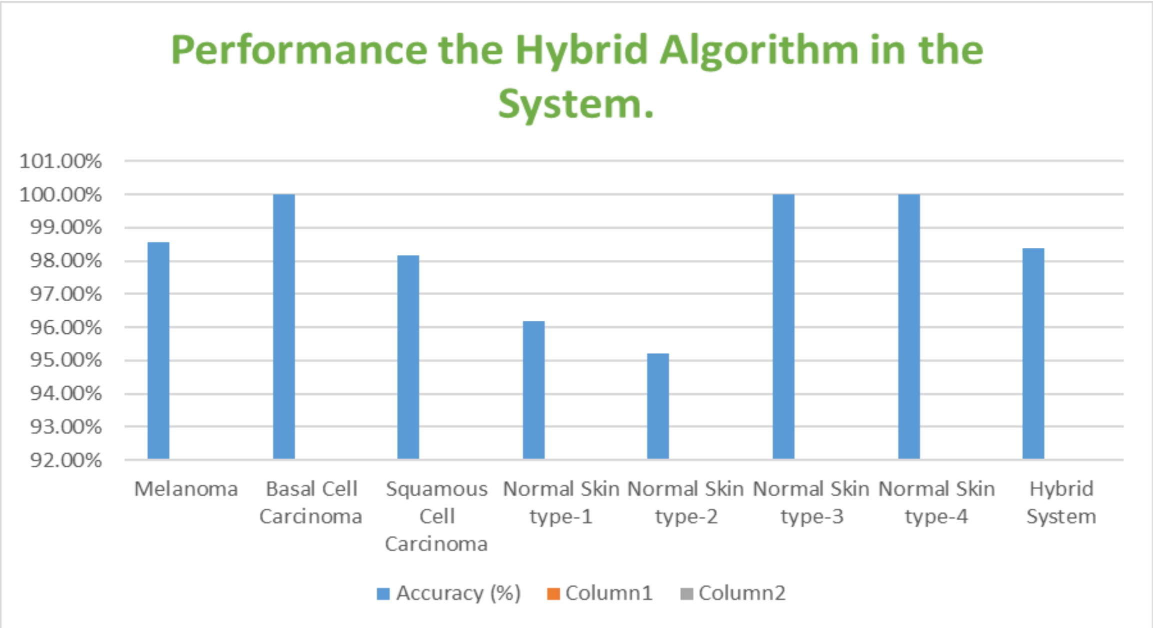


Figure 5: Performance of the Hybrid Algorithm in the Scheme.

As shown in figure (6) the accuracy of the hybrid system is 98.38 % considered the highest result obtained from the performance of every single algorithm and by analyzing information more accurately. Table (3) shows a comparison with the algorithms utilized to classify skin cancer.

Table 3: Comparative Analysis of the System.

ALGORITHMS	ACCURACY	SENSITIVITY	SPECIFICITY
ANN	91.66%	91.66%	82.72%
SVM	94.50%	86.75 %	88.82%
ANN-SVM	98.38%	96.80%	94.88%

Diyala Journal for Pure Science

A Hybrid System for Classification of Skin Cancer Images Using Artificial Neural Network and Support Vector Machine

Raghad Majeed Azawi

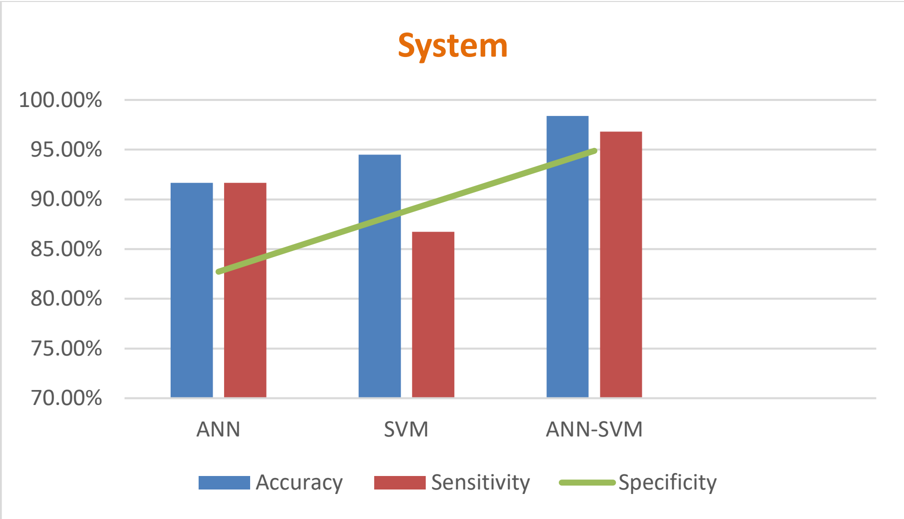


Figure 6: Accuracy of the Hybrid System.

Compared to Other Works

In this section, compared the results of the proposed scheme with Hiam et al, [16] and Mohammad Z. et al. [17] as displayed in the following table (4).

Table 4: Compared Our Proposed Hybrid Scheme Results with [16, 17].

System	Algorithm	Accuracy	Error
Hiam A. and et al, [16]	SVM	92.1%	7.9%
Zakareya and et al. [17]	ANN	98%	2%
The proposed scheme	ANN-SVM	98.38%	1.62%

Conclusion

This paper proves that the hybrid system (ANN-SVM) is stronger than the classification of medical images of skin cancer. The disease is considered a danger in some types of cancers and is easy to spread to the rest of the body so the accuracy of its detection is very important. The proposed Hybrid Classifier method provides upper accuracy ratio of about 98.38 % than earlier proposed techniques. The result of the proposed hybrid system is more practical and reliable in the possibility of adopting it with doctors in the diagnosis of the disease easily this is the main important goal of the study in this paper.

References

1. Seeja R. and Suresh A., "Deep Learning Based Skin Lesion Segmentation and Classification of Melanoma Using Support Vector Machine (SVM)", Asian Pacific Journal of Cancer Prevention APJCP, Vol. 20, 2019.
2. Xie F., Fan H., Li Y. and Jiang Z., "Melanoma classification on dermatoscopy images using a neural network ensemble model", IEEE Trans Med Imaging, Vol. 36, 847-58, 2017.
3. Al-Amin M., Alam M. and Mia R., "Detection of Cancerous and Non-Cancerous Skin by using GLCM Matrix and Neural Network Classifier", International Journal of Computer Applications, Vol. 132(8), pp. 44, 2015.
4. Goel R., and Singh S., "Skin Cancer Detection using GLCM Matrix Analysis and Back Propagation Neural Network Classifier", International Journal of Computer Applications, Vol. 112, pp. 9, 2015.
5. Raja, C., and Jeyaprakash M., "Skin Disease Diagnosis Using Texture Analysis", Seth Institute of Technology and Chettinad College of Engg. and Tech, India, International Journal of Advanced Research in Computer Science and Software Engineering India, Vol. 4(1), 2014.
6. Santosh A. and G. Sadashivappa S., "Skin Cancer Detection and Diagnosis using Image Processing and Implementation using Neural Network and ABCD parameters", International Journal of Electronics, Communication and Instrumentation Engineering Research and Development (IJECIERD), Vol. 4, Issue 3, Jun, 2014.

7. Balaji M., Saravanan S., Chandrasekar M., Rajkumar G. and Kamalraj S., "Analysis of basic neural network types for automated skin cancer classification using Firefly optimization method", Journal Ambient Intel Human Computer, vol. 12(7), 2021.
8. Thomas S., Lefevre J., Baxter G. and Hamilton N., "Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer", Med Image Anal, Vol. 6, 2021.
9. Tumpa P. and Kabir M., "An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features", Sensors International,Vol. 2, pp. 33, 2021.
10. Shunichi J., Naoya Y., Yuichiro H., Yohei S., Yuichiro O. and Ryuji H., "The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning", Biomolecules Journals, Vol. 10, 2020.
11. Kumar N., Kumar P., Pramodh K. and Karuna Y., "Classification of skin diseases using image processing and SVM", In: Abstracts of the international conference on vision towards emerging trends in communication and networking, IEEE, Vol. 3, March, 2019.
12. https://coh.centre.uq.edu.au/project/international-skin-imaging-collaboration.
13. https://www.cancer.org.
14. Mustafa Q. , "Skin lesion classification system using a K-nearest neighbor algorithm", Hatem Visual Computing for Industry, Biomedicine, and Art, Vol. 5,pp 26, 2022.
15. Patnaik S., Sidhu M., Gehlot Y., Sharma B. and Muthu P., "Automated skin disease identification using deep learning algorithm", Biomed Pharmacol Journal, Vol. 11(3),1429–1436, 2018.
16. Hiam A., Ali M., Sajidah A., Esraa A., Ammar A., and Firas H., "The Melanoma Skin Cancer Detection and Classification using Support Vector Machine", IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Vol. 4, October, 2017.

A Hybrid System for Classification of Skin Cancer Images Using Artificial Neural Network and Support Vector Machine

Raghad Majeed Azawi

17. Mohammad Z., Mohammad B. and Arafat U., “Classification of Cancerous Skin using Artificial Neural Network Classifier”, International Journal of Computer Applications, Vol. 181, No. 22, October 2018.
18. Raghad M., Dhahir A., Jamal M. and Ibrahim T., "A Hybrid Approach for Classification of MRI Brain Tumors Using Genetic Algorithm, K-Nearest Neighbor and Probabilistic Neural Network," International Journal of Computer Science and Information Security, Vol. (3), 2018.
19. Raghad M., Dhahir A., Jamal M. and Ibrahim T., "Brain Tumors Classification by Using Gray Level Co-occurrence Matrix, Genetic Algorithm and Probabilistic Neural Network," Diyala Journal of Medicine, Vol. 14, no. 2, June, 2018.