Skin Cancer Classification Using Image Processing and Machine Learning

Arslan Javaid,
Faculty of Engineering and Applied Sciences
Riphah International University
Islamabad, Pakistan

Muhammad Sadiq
Faculty of Engineering and Applied Sciences
Riphah International University
Islamabad, Pakistan

Faraz Akram
Faculty of Engineering and Applied Sciences
Riphah International University
Islamabad, Pakistan

2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST) | 978-1-6654-0516-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/IBCAST51254.2021.9393198

***Abstract*—one of the most rapidly spreading cancers among various other types of cancers known to humans is skin cancer. Melanoma is the worst and the most dangerous type of skin cancer that appears usually on the skin surface and then extends deeper into the layers of skin. However, if diagnosed at an early stage; the survival rate of Melanoma patients is 96% with simple and economical treatments. The conventional method of diagnosing Melanoma involves expert dermatologists, equipment, and Biopsies. To avoid the expensive diagnosis, and to assist dermatologists, the field of machine learning has proven to provide state of the art solutions for skin cancer detection at an earlier stage with high accuracy. In this paper, a method for skin lesion classification and segmentation as benign or malignant is proposed using image processing and machine learning. A novel method of contrast stretching of dermoscopic images based on the methods of mean values and standard deviation of pixels is proposed. Then the OTSU thresholding algorithm is applied for image segmentation. After the segmentation, features including Gray level Co-occurrence Matrix (GLCM) features for texture identification, the histogram of oriented gradients (HOG) object, and color identification features are extracted from the segmented images. Principal component analysis (PCA) reduction of HOG features is performed for dimensionality reduction. Synthetic minority oversampling technique (SMOTE) sampling is performed to deal with the class imbalance problem. The feature vector is then standardized and scaled. A novel approach of feature selection based on the wrapper method is proposed before classification. Classifiers including Quadratic Discriminant, SVM (Medium Gaussian), and Random Forest are used for classification. The proposed approach is verified on the publicly accessible dataset of ISIC-ISBI 2016. Maximum accuracy is achieved using the Random Forest classifier. The classification accuracy of the proposed system with the Random Forest classifier on ISIC-ISBI 2016 is 93.89%.The proposed approach of contrast stretching before the segmentation gives satisfactory results of segmentation. Further, the proposed wrapper-based approach of feature selection in combination with the Random Forest classifier gives promising results as compared to other commonly used classifiers.**

Keywords—Skin lesion segmentation, contrast stretching, features extraction, features reduction, features normalization, features scaling, wrapper method, SMOTE sampling, skin cancer classification, random forest classifier.

I. INTRODUCTION

With rapidly increasing global air pollution and damage to the ozone layer an alarming number of the human population is diagnosed to develop skin cancer as compared to any other type of cancer combined. Melanoma has a very high death ratio as compared to other types of skin cancer. The study of the science behind skin cancer reveals that melanin is present in human skin andmelanocytes are cells in the skin layer that produce melanin. The amount and kinds of melanin produced by melanocytes of different human bodies vary from person to person.In addition to coloring our skin, it also protects from ultraviolet rays of the sun.

The factors contributing to skin cancer include prolonged exposure to direct sunlightultraviolet (UV) rays, the presence of many or unusual moles, skin types, and also if there is a history of melanoma that runs in family. The mortality rate due to melanoma is usually very high butif diagnosed at an early stage has a 99% probability of survival[1], [2]. In many cases, it is a difficult task even for expert dermatologists to make decisions whether a lesion is benign or malignant because of the high resemblance of malignant with benign. Dermatologists use a few techniques such as ABCD rule (Atypical, Border, Color, and Diameter) to get better classification accuracy, but still, human expertise is required[3].

Frequent use of biopsies is also not encouraged by dermatologists. According to International Skin Imaging Collaboration, the number of unnecessary culture tests which are being performed vastly varies depending upon various parameters which include clinical setup,expertise of dermatologist, and the technology applied.For illustration, consider the cases of youngsters in which melanoma rates are significantly low, 500000 culture tests a year have been performed to analyze roughly 400 melanomas[4].

Computer procedures and advancements in machine learning not only aid the dermatologists in early detection of melanoma but also avoid heavy expenses of melanoma detection and unnecessary biopsies. Novel automatic melanoma detection systems save a lot of time, money, and effort. Machine learning has proven to provide melanoma classification with improved and higher accuracies.

II. RELATED WORKS

Although the advancement in the dermatological equipment has increased the classification accuracy of melanoma, the technological developments and improvements in the area of machine learning and image processing have resulted in a medical breakthrough in diagnosis, detection, and classification of melanoma with much more accuracy and reliability.

The literature review reveals that different practices have been used to develop computer-aided automatic diagnostics systems for the classification of skin cancer which take

978-0-7381-0535-2/21/$31.00 ©2021 IEEE

dermoscopic images as input and give classification results as benign or malignant at the output.

HiamAlquran et al. [5] proposed a skin cancer detection system based on OTSU Thresholding. After Image pre-processing and segmentation, they extracted features including ABCD (Atypical, Border, Color, and Diameter), Total dermoscopic score (TDS), Circulation, and texture features using Gray Level Co-occurrence Matrix (GLCM). They used the SVM classifier with the RBF Kernel algorithm for classification and achieved an accuracy of 92.1%.

UzmaBano et al. [6] developed a system for detecting skin cancer by image processing. After pre-processing, the resultant was subjected to maximum entropy thresholding for lesion segmentation. They extracted a few texture features using GLCM. They used the SVM classifier and achieved a classification accuracy of 95%.

M. Attique Khan et al. [7]developed an automatic skin cancer detection and classification system based on the normal distribution for image segmentation after some pre-processing. They extracted shape features using a Histogram of Gradients (HOG), texture features using Harlick Features, and color features. For dimensionality reduction, they performed an entropy controlled feature selection. They utilized several classifiers to check classification accuracy and achieved maximum accuracy using a multi-class SVM classifier. They achieved a classification accuracy of 97.5%, 97.75%, and 93.2% on PH2, ISIC (UDA, MSK-2), and Combined (ISBI 2016 -17) datasets respectively.

Vijayalakshmi M M [8] developed an automatic skin cancer detection system using image processing and artificial intelligence. She utilized 1000-1500 images from publicly available ISIC dataset. After initial Pre-processing, lesion segmentation was performed using the OTSU segmentation method, Modified Otsu segmentation method, and watershed segmentation method. She used multiple classifiers to evaluate the performance including the Back Propagation Algorithm (Neural Networks), SVM, and CNN. She achieved maximum accuracy of 85% with the SVM classifier.

Dalila et al. [9] introduced a system for the Segmentation and classification of melanoma and benign skin lesions based on Ant-Colony based segmentation algorithm for the lesion. They extracted features of shape, texture, and color. For classification, they used KNN and ANN classifiers and achieved accuracies of classification of 85.22% and 93.60% respectively.

Sumithra R et al. [10]developed a skin cancer diagnosis system based on the region-growing segmentation techniques. After Pre-Processing and segmentation of images, they extracted color, texture, and RGB Histogram Features. In the final phase of classification, they used three classifiers i.e. KNN, SVM, and combined SVM+KNN, and achieved classification accuracies of 86%, 87.5%, and 94% respectively.

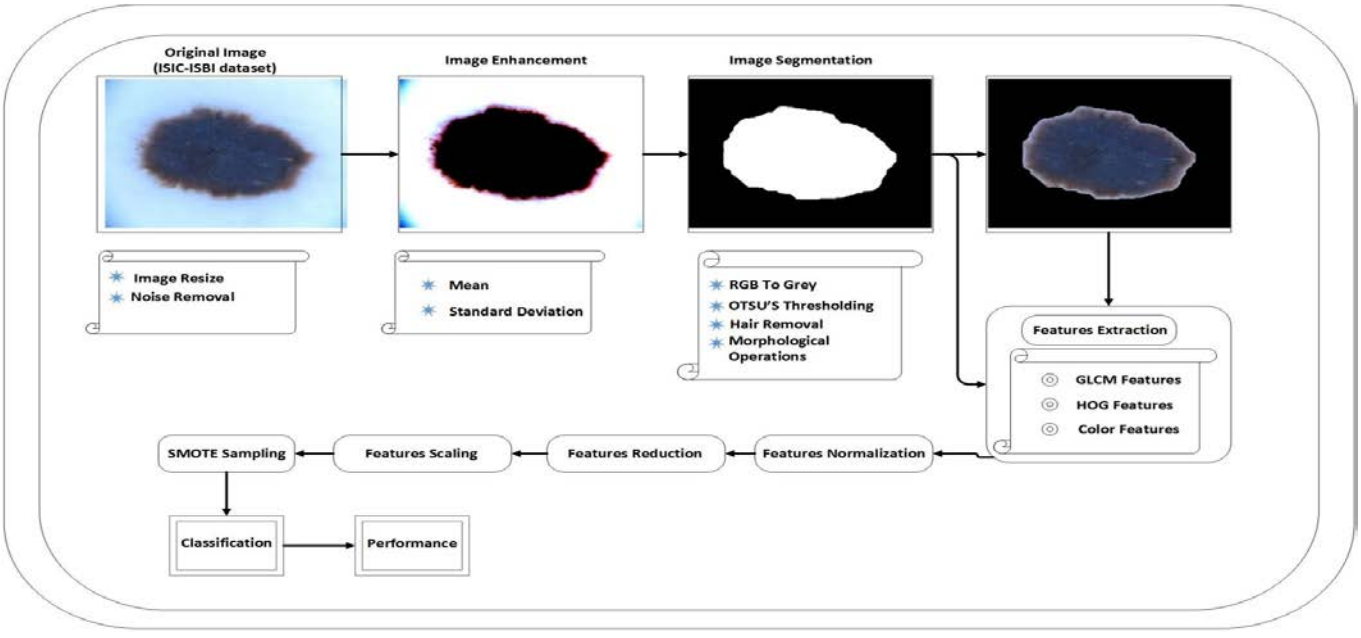
Most of the researchers have used their own data collected from hospitals. But their dataset is too limited for proper training during the machine learning phase. Few researchers have used the ISIC dataset which is a handsome amount of data and also it is more practical.

The purpose of this research is to find improved and more effective ways to detect skin cancer using digital image processing and machine learning techniques. The final objective is to assist the doctors in the detection of skin cancer at an early stage by providing improved and reliable results.

The suggested system gives high accuracy of classification of the lesion as benign or malignant which will be very helpful for the identification of patients efficiently and our results are better than earlier work on the ISIC dataset.

III. METHODOLOGY

In this paper, we propose a system that classifies dermoscopic images as benign or malignant by using image processing and machine learning.Fig.1 shows the block diagram of the proposed system.



A. Image Database

The database is downloaded from publicly available images on the ISIC website. The database included the ISBI-2016 challenge which has RGB dermoscopic images along with their labels and segmentation ground truths.

B. Pre-Processing

As ISIC-ISBI dataset images have different artifacts, so pre-processing is done on them to make them more meaningful. Pre-processing included image resizing, noise removal, contrast stretching, RGB to Gray conversion, and hair removal. We resized all images to 767x1022. For noise removal median filter of 3-by-3-by-3 is used. For contrast enhancement, a new method is proposed. Contrast enhancement significantly improves the results of segmentation in the next phase. Proposed contrast stretching is based on the mean and standard deviation of pixels intensities of images. Minimum and maximum intensity values i.e. "Low in" and "High in" of input images values of input images are calculated using formulas defined as in (1) and (2).

$$Low\ in = Avg - \sigma * N \quad (1)$$

$$High\ in = Avg + \sigma * N \quad (2)$$

Where,

Avg = Average

σ = Standard Deviation

N = 0.4

Intensity values are then mapped in the output image from 0 to 255. Contrast stretching is performed on R, G, and B channels separately which are later concatenated to contrast stretched RGB images. RGB to gray conversion is performed next. For removing hairs bottom-hat filtering is performed, then those pixels which have been filtered are replaced by neighboring pixels.

C. Segmentation

After the Pre-Processing, image segmentation is performed to extract the region of interest i.e. skin lesion. We used the OTSU thresholding algorithm for the segmentation of grayscale Pre-Processed images. OTSU thresholding aims to find the threshold and classify pixels into two classes so that within-class variance is minimum (between-class variance is maximum).

Background triangles that exist in a few images on four sides are then removed by creating a mask and adding it to the binary image. Later the image is inverted to make the skin lesion white and the background to be black. Flood fill operation on 4-connected pixels in the background is performed to remove holes. A morphological opening is then done to remove small objects having pixels fewer than 2000. The lesion border is smoothened using opening operation first with a disk-shaped structuring element of radius 20 followed by a closing operation with the same structuring element.

D. Features extraction and reduction

After the segmentation, features of texture, shape, and color are extracted from segmented skin lesions. To extract the texture features, segmented lesions mapped on original images are converted from RGB to grayscale first. Then two-dimensional wavelet decomposition of level 3 with mother wavelet as Daubechies (db4) is performed on grayscale images. Low pass approximation coefficients matrix (CA) of level 3 wavelet decomposition result in new images of reduced size. Afterward, texture features are obtained via gray level co-occurrence matrix of approximation coefficients matrix. A total of 13 texture features are extracted as shown in Table 1.

A total of 36 color features are extracted using three different color spaces i.e. RGB, HSV, and LAB color space as shown in Table 1. For extracting shape features, Histograms of Gradient features (HOG) are extracted. The size of all segmented lesion images is reduced first to 96 X 128 to reduce the dimensionality of images. Then Histogram of Gradients "HOG" feature vector is determined. The size of the "HOG" feature vector for a single image is 1x3456 because each image is of dimension 96 X 128 and the block size is chosen to be [8 8]. For reducing the dimensionality of a large data array of HOG vector, PCA is used. The size of the HOG feature vector using PCA is reduced from 1x3456 to 1x100 for a single image. PCA increases understandability and reduces information loss.

TABLE 1. EXTRACTED TEXTURE AND COLOR FEATURES

Texture features	Skewness	Mean	Contrast	Energy	Homogeneity	Standard
	Root Mean Square	Variance	Smoothness	Kurtosis	Correlation	Entropy
	Inverse Difference Movement					
Color Features	Mean R	Mean G	Mean Blue	Variance R	Variance G	Variance Blue
	Kurtosis R	Kurtosis G	Kurtosis	Skewness R	Skewness G	Skewness Blue
	Mean hue	Mean Saturation	Mean	Variance Hue	Variance	Variance Value
	Kurtosis H	Kurtosis S	Kurtosis V	Skewness H	Skewness S	Skewness V
	Mean L	Mean A	Mean B	Variance L	Variance A	Variance B

	Kurtosis L	Kurtosis A	Kurtosis B	Skewness L	Skewness A	Skewness B

13 GLCM texture features, 36 color features, and 100 shape features are all then concatenated to make a single feature vector of dimension 1x149 for a single image.

E. SMOTE Sampling

Data of ISIC- ISBI has a class imbalance problem. Almost 80% of data is "benign" and 20% of data is "malignant". So any classifier trained on this data will be more biased or skewed to have more training on benign data. As a result, it would be poor training, poor training accuracy, and poor classification accuracy on test data. Hence data needs to be balanced first. For this purpose Synthetic Minority Over Sampling (SMOTE) technique is implemented on the final feature vector [11]. SMOTE technique resulted in data being 50% benign and 50% malignant.

All of the training and test data of the ISIC-ISBI data set is used to extract the feature vector. Next, the SMOTE technique is used for data balancing. After that feature vector or predictors which are a total of 149 predictors of all images along with their given labels are randomized. Next data is broken down in such a way that 80% is utilized for purpose of training and 20% is utilized for testing. Now the ratio of benign and malignant images in training data is 50:50. Similarly, the ratio of benign and malignant images in test data is 50:50.

F. Features Standardization

Features vector is standardized using the method of "Linear scaling to unit variance" [12]. Average and standard deviation of each column of the feature vector is computed. Then the average of each column is subtracted from the original feature vector's respective column to get zero mean feature vector. Afterward zero mean feature vector's each column is divided by the standard deviation of that column to get the feature vector of unit variance.

G. Features Scaling and selection

Features are scaled in such a way that in the feature vector each feature is normalized in a range from 0 to 7 by using the Min-Max Mapping Algorithm [13].

In this paper, a novel approach of using wrapper methods as feature selection methods for skin cancer classification is proposed. The wrapper method aims at finding the subset of features that are most strongly relevant and also weakly relevant features that improve performance, using the training algorithm itself as a part of the evaluation function. For the problem of finding a subset of features that are most relevant as well as to improve accuracy, the wrapper method was proposed by George H. John et al [14]. Feature selection is used to improve the performance of the machine learning models, increase the accuracy of the classification, and to aid the analysis of the results. Normalized features are used to determine the most prominent features out of 149 features in hand." FEATURESELECT" is free open source software under MIT license [15]. A total of 95 most optimum features are selected by using optimization algorithms as a feature selection method and decision tree as a base classifier. We considered classification accuracy an objective function. "LA" [15]algorithm determined 95 features and gave the most promising results in terms of sensitivity, specificity, precision, FPR, and accuracy as compared to other algorithms.

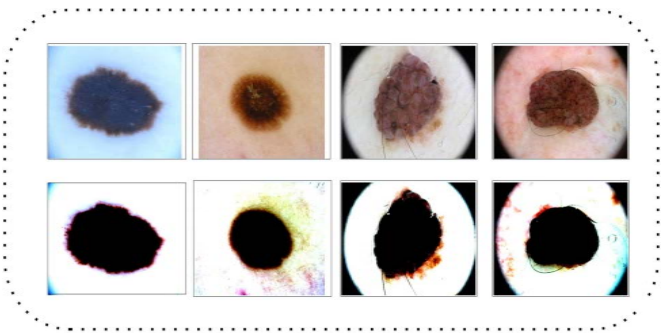
H. Classification

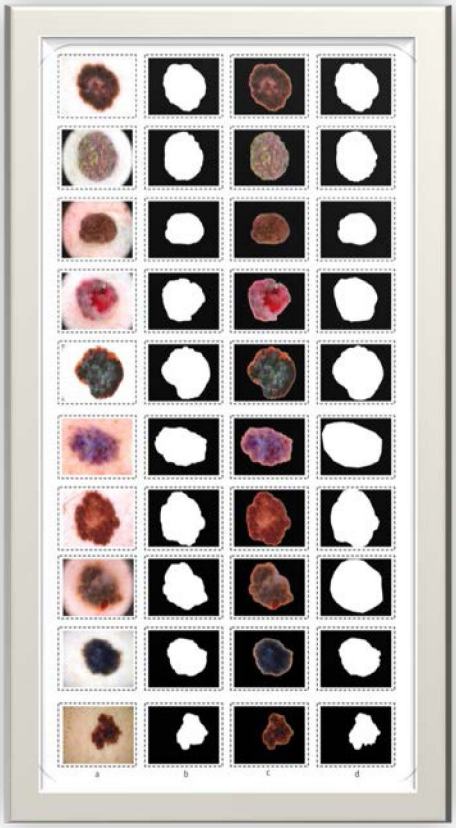
In the final step, we classified data as benign or malignant. Numbers of classifiers including Quadratic Discriminant, SVM (Medium Gaussian), and Random Forest are then used for training and their performance is evaluated on test data. Classifier models are later evaluated on test data. SVM divides the data into distinguishable classes to determine hyperplane with maximum margin. Proposed features selection based on wrapper methods gives remarkable and promising results when Random Forest classification is performed. 500 decision trees are used in the Random Forest classifier and classification is done based on the majority votes of decisions given by different trees. The Random Forest technique generates decision trees on samples of data and predictions from each tree undergo the process of voting for classification. More votes from different trees mean more likely it belongs to that particular class.

IV. RESULTS

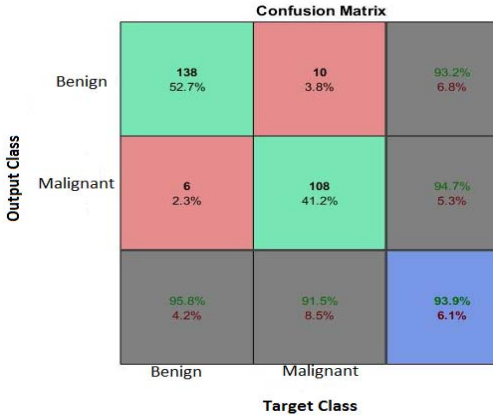
Fig.2 shows the results of contrast stretching. Results of segmentation compared with ground truths are shown in Fig.3.

Classification accuracies achieved on the ISIC-ISBI2016 dataset using SVM (Medium Gaussian), Quadratic Discriminant, and Random Forest classifiers are 85.50%, 88.17%, 90.84%, and 93.89% respectively. The confusion matrix using Random Forest classifier for ISIC-ISBI 2016 is shown in Fig.4. Among 114 malignant guesses, 94.7% are accurate and 5.3% are erroneous. Among 148 benign guesses, 93.2% are accurate and 6.8% are erroneous. Among 144 benign subjects, 95.8% are properly foreseen as benign and 4.2% are foreseen as malignant. From 118 malignant subjects, 91.5% are properly categorized as malignant and 8.5% are categorized as benign. 93.9% of the complete guesses are correct and 6.1% are erroneous.





Classification accuracies using the SVM, Quadratic Discriminant, and Random Forest are compared in Table 2.



The proposed system gives remarkable results using the Random Forest classifier in terms of accuracy, sensitivity, precision, and AUC-ROC.

V. CONCLUSION

In this work, a novel method of skin cancer classification using machine learning and image processing is implemented.

In the first step, a novel method of contrast stretching based on the mean and standard deviation of pixels for dermoscopic images enhancement is proposed. Then OTSU thresholding is performed for segmentation.

Shape, color, and texture features are extracted in the second step, and shape features are then reduced using the PCA. The class imbalance problem of the ISIC dataset is overcome using the SMOTE sampling technique.

In the third step, features are standardized and scaled and then a novel approach of features selection based on wrapper methods for selecting the most optimum features is proposed.

The recommended system is tested on an openly available dataset i.e. ISIC-ISBI 2016 and it is concluded that the proposed wrapper method for feature selection in combination with the Random Forest classifier gives promising results as compared with other classifiers.

TABLE 2 CLASSIFICATION RESULTS USING DIFFERENT CLASSIFIERS

Classification Algorithm	**Accuracy (%)**
SVM	88.17
Quadratic Discriminant	90.84
Random Forest	93.89

REFERENCES

[1] D. Schadendorf *et al.*, "Melanoma," *Lancet*, vol. 392, no. 10151, pp. 971–984, 2018.

[2] S. Gupta and H. Tsao, "Epidemiology of melanoma," *Pathol. Epidemiol. Cancer*, pp. 591–611, 2016.

[3] F. Nachbar *et al.*, "The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions," *J. Am. Acad. Dermatol.*, vol. 30, no. 4, pp. 551–559, 1994.

[4] "The International Skin Imaging Collaboration (ISIC)." https://www.isic-archive.com/#!/topWithHeader/tightContentTop/about/isicArchive (accessed Oct. 04, 2020).

[5] H. Alquran *et al.*, "The melanoma skin cancer detection and classification using support vector machine," *2017 IEEE Jordan Conf. Appl. Electr. Eng. Comput. Technol. AEECT 2017*, vol. 2018-Janua, pp. 1–5, 2017.

[6] U. B. Ansari and M. E. Student, "Skin Cancer Detection Using Image Processing Tanuja Sarode 2," *Int. Res. J. Eng. Technol.*, vol. 4, no. 4, pp. 2395–56, 2017, [Online]. Available: https://www.irjet.net/archives/V4/i4/IRJET-V4I4702.pdf.

[7] M. A. Khan *et al.*, "An implementation of normal distribution based segmentation and entropy controlled features selection for skin lesion detection and classification," *BMC Cancer*, vol. 18, no. 1, pp. 1–20, 2018.

[8] V. M. M, "Melanoma Skin Cancer Detection using Image Processing and Machine Learning," *Int. J. Trend Sci. Res. Dev.*, vol.

Volume-3, no. Issue-4, pp. 780–784, 2019.
[9] F. Dalila, A. Zohra, K. Reda, and C. Hocine, "Segmentation and classification of melanoma and benign skin lesions," *Optik (Stuttg).*, vol. 140, pp. 749–761, 2017.
[10] R. Sumithra, M. Suhil, and D. S. Guru, "Segmentation and classification of skin lesions for disease diagnosis," *Procedia Comput. Sci.*, vol. 45, no. C, pp. 76–85, 2015.
[11] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.
[12] S. Aksoy and R. M. Haralick, "Feature normalization and likelihood-based similarity measures for image retrieval," *Pattern Recognit. Lett.*, vol. 22, no. 5, pp. 563–582, 2001.
[13] X. H. Cao and Z. Obradovic, "A robust data scaling algorithm for gene expression classification," *2015 IEEE 15th Int. Conf. Bioinforma. Bioeng. BIBE 2015*, no. 1, 2015.
[14] G. John, R. Kohavi, and K. Pfleger, "IrreleJohn, G., Kohavi, R., & Pfleger, K. (1994). Irrelevant Features and the Subset Selection Problem. Icml, 121–129. Retrieved from http://machine-learning.martinsewell.com/feature-selection/JohnKohaviPfleger1994.pdfvant Features and the Subset Selectio," *Icml*, pp. 121–129, 1994, [Online]. Available: http://machine-learning.martinsewell.com/feature-selection/JohnKohaviPfleger1994.pdf.
[15] Y. Masoudi-Sobhanzadeh, H. Motieghader, and A. Masoudi-Nejad, "FeatureSelect: A software for feature selection based on machine learning approaches," *BMC Bioinformatics*, vol. 20, no. 1, pp. 1–17, 2019.