iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

MELANOMA: A COMPUTER-AIDED DIAGNOSIS USING THE ABCD RULE

Dr Aziz Makandar
[1] Professor, Department of Computer Science,
Karnataka State Akkamahadevi Women's University, Vijayapura-586101
azizkswu@gmail.com

Mrs. Ayisha Soudagar
[2] Research Scholar, Department of Computer Science,
Karnataka State Akkamahadevi Women's University, Vijayapura-586101
ayishamudassar07@gmail.com

ABSTRACT
Malignant melanoma, an uncommon but far more dangerous form of skin cancer, is a highly serious dermatological condition. The most deadly type of skin cancer is melanoma, which can be cured if caught early. In addition to qualified experts who can accurately diagnose the illness, automated systems can also identify diseases, potentially saving lives and lowering expenses. In order to achieve this, a support vector machine (SVM) classifier model with a radial basis function kernel was fed image features and class labels in order to predict whether or not dermoscopy images included malignancy. In order to identify the best features that may improve the diagnosis process' accuracy, this classifier's ABCD features (Asymmetry, Border, Color, and Diameter) and texture features generated from the Haralick texture features estimated from the Gray Level Co-occurrence Matrix were examined. Twelve texture elements were ultimately found to have the maximum efficiency. Furthermore, the results indicate that the accuracy was improved by 9.6% using the newly added texture features, compared to 84% using the standard ABCD rule and 93.6% using it.

Keywords:
Melanoma Diagnosis, Texture Features, Support Vector Machine, ABCD Features.

I.INTRODUCTION

The most deadly kind of SKIN cancer, malignant melanoma (MM), accounts for almost four out of five fatalities. The only available treatments for melanoma in its advanced stages are palliative care, immunotherapy, surgery, and chemotherapy [1]. However, it is very possible to cure non-melanoma skin cancers, particularly squamous or basal cell cancers. If detected early and treated appropriately, melanoma is highly curable, even though it spreads quickly to other regions of the body [2]. According to the Skin Cancer Foundation, one in five Americans will probably develop skin cancer at some point in their lives. It is important to note that MM frequently causes injury to white people because of prolonged exposure to direct sunlight [3].

Melanoma and other pigmented skin lesions (PSLs) are identified using the dermoscopy pictures [4], [5]. Dermatologists use these photos to see the lesions clearly during clinical procedures. Even the majority of qualified doctors have accuracy levels below 85%, despite the fact that this equipment clearly shows the characteristics of skin lesions, making it extremely difficult to distinguish between different types of melanoma [4]. As a result, many occurrences of melanoma are not correctly diagnosed. Experienced dermatologists use test measurements, history, and pattern recognition to clinically identify lesions.

Various classification techniques, including Artificial Neural Network (ANN), Support Vector Machine (SVM), and feature selection, deep learning, or hybrid approaches, are used by computer-aided diagnostic (CAD) systems to automatically diagnose lesions [6, 7, 8, 9]. However, the primary steps in any melanoma recognition system are preprocessing, lesion segmentation, feature extraction, feature selection, and classification.

The primary goal of the study is to investigate textural characteristics that ought to be combined with ABCD characteristics in order to improve the entire process of diagnosing malignant melanoma.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

1.BACKGROUND

Research has recently focused on health data analysis because of the various challenges faced by analysts and big data miners [10], [11], and [12]. One of the focused areas [13] that aids doctors in their decision-making is classification issues. The classification of lesions is the main emphasis of this study. For any automated system to classify lesions, image processing and analysis are essential. The results of lesion classification can be ternary (common nevus/dysplastic nevus/melanoma), binary (benign/malignant or non-suspicious/suspicious for malignancy), or n-array, which denotes several skin conditions. Thus, these results identify the kinds of pigmented skin lesions (PSLs) that a system is trained to differentiate.

Using a collection of 3639 dermoscopic pictures, Maglogiannis and Doukas [14] presented a unified comparison of eleven classifiers on a group of feature descriptors (using multiple feature selection methodologies). They also introduced classification outputs of different CAD systems. The eleven chosen classifiers, which included decision trees, SVM, regression analysis, and neural networks, among others, were the most often used classifier sets in the PSL computerized study. Three sub-experiments that calculated the number of resultant classes were used to conduct the comparison. The third trial combined all three classes, whereas the first two were common nevus classes/common nevus and dysplastic/melanoma. SVM hence showed the best overall performance.

Several more research have compared two or three or more classifiers from the CAD systems and classification categories. Comparing SVM to artificial neural networks (ANN) has therefore been the subject of numerous papers [15], [16]. Overall, there was a modest variation in the classifiers' performance. In contrast, discriminant analysis (DA) performed even or marginally worse when compared to ANN in [17] and to both SVM and ANN in [18]. In [19], the Bayesian classifier was evaluated against SVM, and in [20], it was evaluated against ANN and k-nearest neighbor (KNN). It was demonstrated to perform better than the KNN approach but worse than the ANN. Kassem and associates. For the categorization of skin lesions, [21] contrasted deep learning techniques with conventional machine learning techniques. They examined the primary obstacles to assessing skin lesion segmentation and classification techniques, including racial bias, ad hoc image selection, and short datasets.

The realm of medical diagnosis and therapy has improved as a result of numerous additional studies from the CAD systems and classification categories. Specifically, in [22], the artificial neural networks' (ANN) performance has increased. Deep learning algorithms have been used at PSLs to extract texture and color information from photos of skin lesions, resulting in these improvements. The findings demonstrated that their system could assist clinical specialists and was more reliable than existing CBIR systems. Kharazmi et al. developed another decision assistance system in [23] that applied deep feature learning to dermoscopic pictures. They put out a novel paradigm for thorough cutaneous vascular detection. Because the classification was binary, the system could detect the presence or absence of vasculature. With a detection accuracy of 95.4%, the framework performed quite well.

A novel approach, based on the ABCD rule, was put out to categorize PSL as either benign or malignant [24]. In the tests, two classifiers were used: the KNN classifier by itself and the KNN classifier in conjunction with a Decision Tree classifier (KNN-DT) [25]. Two publicly accessible datasets of PSL photos were used to test the methodology.

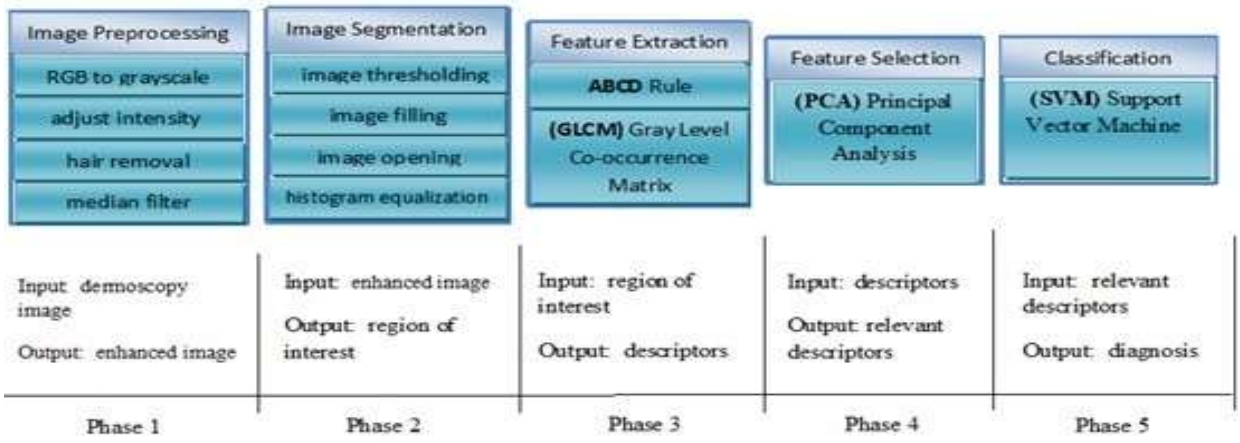


Figure 1. The research phases that are followed in this research

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

In conclusion, the authors said that the technique can achieve a classification accuracy of up to 96.71% and that the preliminary experimental findings were encouraging [26]. In [27], an additional enhancement to the ABCD rule was detailed. The new algorithm's primary feature is its deep application of the Persistent Homology method, which guarantees a thorough examination of the skin lesion. After testing the algorithm on a dataset of 107 melanocytic lesions, the authors came to the conclusion that further research and testing are necessary.
In conclusion, the need for more symptoms and characteristics to improve the diagnosis process is one of the shortcomings and gaps of the earlier studies. To close this gap, a study was published that looks at texture features that should be combined with ABCD features to improve the detection of malignant melanoma. The SVM classifier is used to predict the class labels of the images.

1. RESEARCH METHODOLOGY

Preprocessing, segmentation, feature extraction, feature selection, and classification are the five stages of operation that we have used to obtain an accurate retrieval of skin tumor photos. The key steps of the investigation are shown in Fig. 1.

A. DATA SAMPLE

The international skin imaging collaboration project (https://api.isic-archive.com/api/docs/swagger/) and the international atlas of dermoscopy and dermatoscopy website (https://www.dermoscopyatlas.com/), an educational initiative of the Skin Cancer College of Australia and New Zealand, were the two sources from which the digital dermoscopy images, case histories, symptoms, and treatment plans were collected. Finally, 220 photos were examined, scaled to 512×512 pixels, and evenly distributed between 110 photographs for common nevus and malignant melanoma. The gathered data was subjected to the aforementioned study phases, as detailed in the part that follows.

2. EXPERIMENTS

A. Preprocessing

Preprocessing is the initial stage that was applied to the raw imaging data with the goal of improving the image. It included the following steps:

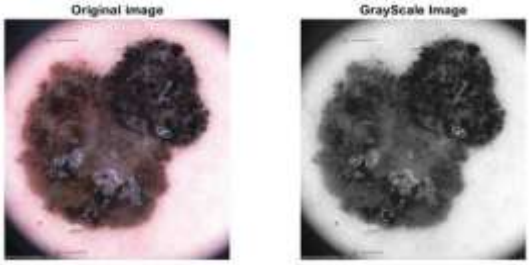


Figure2. Converting RGB image to grayscale image

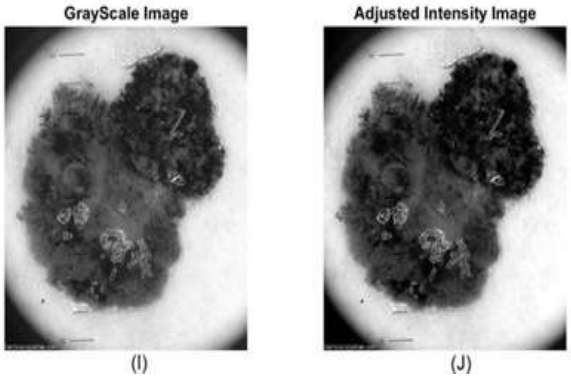


Figure3. Adjusting gray scale image intensity

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

1) **GRAY SCALE IMAGE**
The values in this array correspond to intensities within a certain range. However, identifying visual characteristics that are impossible to define in real colored images is the primary objective of transforming color map photos to grayscale. The outcome of converting an RGB image from the data sample to a grayscale image is displayed in Fig. 2
2) **INTENSITY ADJUSTMENT**
At this stage, as shown in Fig. 3, 1% of the top and bottom pixel values were saturated in order to change the intensity values in the grayscale image (Fig. 3.I) to new values (Fig. 3.J). Stated otherwise, the resulting image (Fig. 3.J) had a higher contrast.
3) **HAIR REMOVAL**
It is evident from Fig. 4 that, in addition to hair, superfluous little things were eliminated, including millimeter marks that were used to measure the diameter. The dull razor filter was applied to complete the hair removal process.

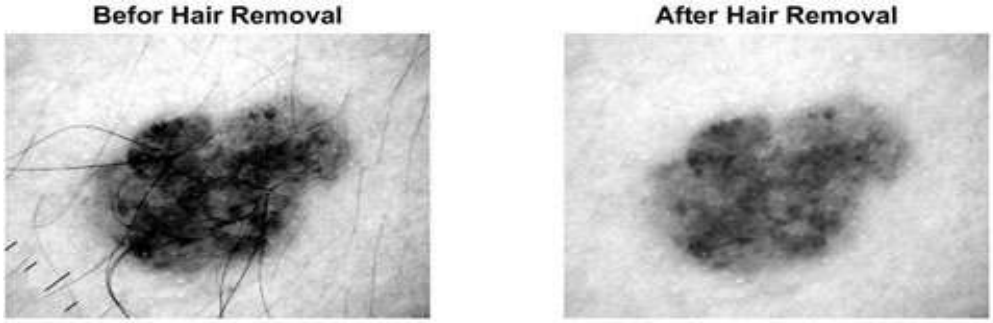


Figure4. Applying hair removal

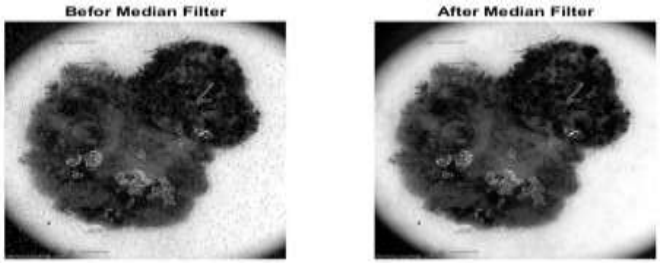


Figure5. Applying median filter to an image

Figure6. Image Segmentation example

4) **MEDIAN FILTER**
In image processing, median filtering is typically employed to remove noise, as shown in Fig. 5. It is a nonlinear approach with the primary goals being noise reduction and edge preservation. In a 3-by-3 region close to the linked pixel in the input image, each output pixel has the median value [28].

B. **SEGMENTATION**
Lesions segmentation, as shown in Fig. 6, joins pixels with similar characteristics to indicate the division of the image into separate regions (lesion/healthy skin). Humans can identify the boundaries of lesions with the aid of the segmentation process. Even though it is typically exceedingly difficult to partition a picture accurately, segmentation reliability is essential to the success of image analysis.

1. **THRESHOLDING**
Otsu's approach [29], which employs the threshold to reduce the intra-class variance (within a

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

certain segment) of the white and black pixels, was employed in this phase to calculate the level or global threshold that may be used to convert an intensity image into a binary image, as shown in Fig. 7.

2. **IMAGE FILLING**

Since white pixels will be utilized to fill in the holes and sections of the image, the binary image will first be inverted, as illustrated in Fig. 8. After that, a morphological reconstruction algorithm will be used to fill in the gaps and areas of the image. By default, the algorithm assigns a new value to a particular pixel in the 2-D inverted binary image using 4-connected neighborhoods.

3. **IMAGE OPENING**

In image processing and computer vision, the opening is regarded as a fundamental feature for eliminating morphological noise [30]. As a result, opening pushes little things to the background of an image from the foreground, which is typically thought of as the bright pixels. An illustration of image opening is shown in Fig. 9

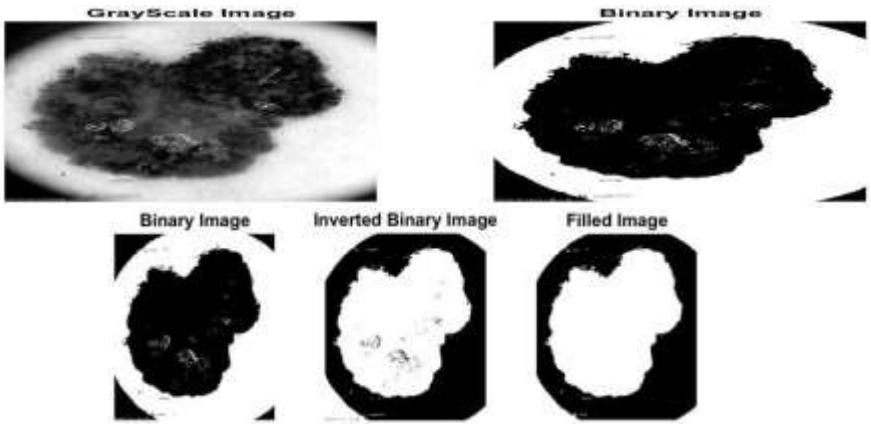


Figure 8. image Filling

1) **HISTOGRAM EQUALIZATION**

Prior to applying histogram equalization, we identified the image's largest item (the lesion) and its border. We then cropped the image to make it more suitable for the features extraction stage. The accuracy of the (GLCM) approach is then increased by employing histogram equalization [31] to enhance contrast.

C. **FEATURES EXTRACTION**

Choosing the appropriate features that best characterize each image class in a unique way is essential to developing an efficient retrieval system. For the description of various diseases, the traits must be adequate and discriminative. Using the ABCD rule for melanoma identification in conjunction with the common texture feature method (GLCM) and additional features like skewness, standard deviation, mean, entropy, and kurtosis, we try to represent the parameters of pigmented lesions in our study.

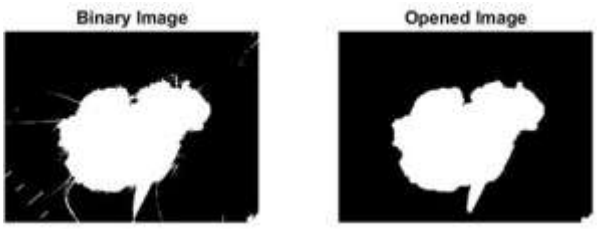


Figure 9. image Opening

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

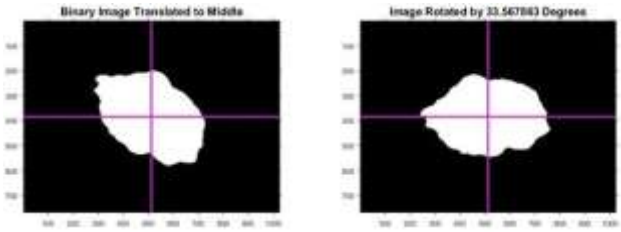


Figure 10. Lesion Rotation

1) **ABCD RULE OF DERMOSCOPY**

Four arguments were shown to be crucial in the diagnosis of melanoma. It is a scientific criterion that assigns points to the approach that includes the PSL's Asymmetry (A), Border (B), Colors (C), and Diameter (D). According to Equation 1, the Total Dermoscopic Score (TDS) is produced by multiplying each recognized criterion by a weight parameter.

$$TDS = 1.3 \times A + 0.1 \times B + 0.5 \times C + 0.5 \times D \quad (1)$$

A lesion is deemed likely to be melanoma if the TDS value is greater than 5.45. A TDS score of 4.75 to 5.45 is considered moderate, indicating the possibility of melanoma; a number below 4.75 implies benignity [25].

a) **ASYMMETRY**

We used a novel picture editing technique in our study to accurately evaluate the asymmetry. Based on examining the lesion's characteristics, we calculated the angle between the lesion's major axis and the x-axis using the lesion's orientation. In the end, we used this angle to determine the degree of rotation required to match the appropriate components that need to be compared in order to verify the asymmetry, which is visible in Fig. 10.

The lesion has since been cut along two axes that are positioned to provide the least amount of asymmetry in terms of colors, borders, and dermoscopic features. A point below one or more axes is used to test for asymmetry [14]. As shown in Fig. 11, in (A), there is no obvious asymmetry in either of the two axes, where (A,B,C,D) = (0,0,2,3), respectively. This indicates that the ABCD score is 2.5, and as a result, it is regarded as a typical nevus. The pigmented lesion in (B) has asymmetry in one axis, though; the ABCD score is 5.7 because (A,B,C,D) = (1,4,4,4), respectively. As a result, it is classified as malignant melanoma (MM). However, in (C), there is asymmetry of lesion in both axes at a 90-degree angle; the ABCD score is 7.9, the (A,B,C,D)=(2,8,5,4), and as a result, it is classified as malignant melanoma.

b) **BORDERS IRREGULARITY**

Eight pie-piece portions make up the lesion. The pigment pattern is next examined to determine if it is gradually or abruptly cut off at the lesion boundary. Then, border-based features that categorize the lesion's shape are computed. The border irregularity, area, largest diameter, thinness ratio, variation of the distance of the border lesion spots from the centroid position, circularity index (CIRC) [33], and symmetry distance (SD) [34] are the most prevalent border features.

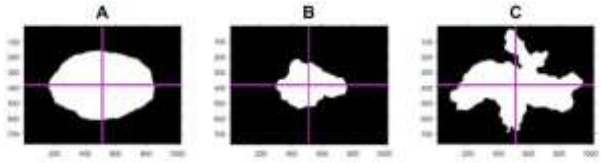


Figure 11. Evaluation of asymmetry in pigmented lesions

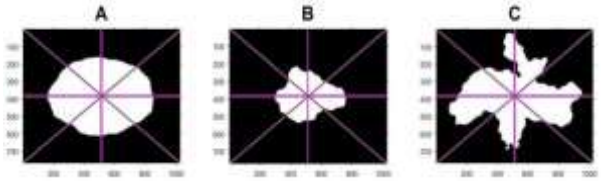


Figure 12. Evaluation of border in pigmented lesions

As illustrated in Fig. 12, the ABCD score is 2.5 in (A) due to the strong cutoff of the pigment pattern in one of eight sections/segments (border score = 0, typical nevus), where (A,B,C,D) = (0,0,2,3) respectively. But in (B),

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

the pigment pattern is sharply cut off in four of the eight portions (border score = 4, MM), so (A, B, C, D)=(1,4,4,4) respectively, and the ABCD score is 5.7. However, in (C), the ABCD score is 7.9, and the (A,B,C,D)=(2,8,5,4), with a severe cutoff of the pigment pattern in all eight areas (border score =8, MM).

c) **COLOR**

The amount of colors present is specified, and the color properties of the lesion are examined. Slate blue, dark brown, light brown, red, black, and white (if it is whiter than the surrounding skin) are the only six colors that will be examined. Additionally, the nature of pigmented skin lesions could be described using color and texture.
Prior studies have looked at the color of the lesion in a variety of ways, but the majority of them failed to remove the black color of the binary mask surrounding the lesion, which is irrelevant for the color of the lesion we are testing. As a result, black was counted in every image even though it wasn't visible in the lesion. By switching the background color from black to green, which wouldn't be taken into account when searching for the preceding six colors, this study did make a new contribution by totally eliminating the unnecessary black color (Fig. 13). As a result, color and texture feature extraction procedures are now more accurate.
As shown in Fig. 14, the ABCD score is 2.5 for the two colors (dark brown and light brown) in (A) where (A,B,C,D) = (0,0,2,3) correspondingly and the color score is 2. On the other hand, the ABCD score is 5.8 since (A,B,C,D)=(1,5,4,4) in (B) four colors (dark brown, light brown, black, and blue-gray) (color score = 4, MM). However, in (C), the ABCD score is 7.1 since the six hues (black, dark brown, light brown, red, blue-gray, and white) have color scores of 6 MM.

Figure 13. Changing background color to green

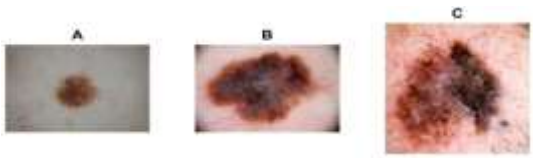


Figure 14. Evaluation of different colors in pigmented lesions

d) **DIAMETER**

Any mole growth should raise concerns because a width of more than 6 mm indicates a likely melanoma. We will use these scales to quantify the lesion's diameter because the dermatoscope allows for the tracking of marks on the picture through acquisition as demonstrated in [35]. However because they did not scale or convert from pixel to millimeter (mm) in a specific image—something that can only be done with the dermatoscope diameter marks—the majority of earlier work failed to estimate the actual word diameter size.
We took the following actions to achieve this: First, we determined the lesion diameter by taking the longest line that extends from the lesion center. The diameter length, expressed in pixels, was then measured. Lastly, to determine how many pixels are in (mm) for a given image, we scaled and translated an already known distance in (mm) to distance in pixels, as shown in Fig. 15.

2) **GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM)**

Texture analysis functions as a classification and image analysis technique. Classifying similar objects (regions) in various photographs is made much easier by this method of displaying the spatial distribution of intensities. Features of the Haralick texture The schieda2015diagnosis, which is derived from a GLCM, is a widely used method for characterizing image texture. 20 texture features, including cluster, were calculated in this study using Haralick texture features. Contrast, entropy, correlation, homogeneity, autocorrelation, cluster shade, sum variance, difference entropy, information measures of correlation 1 and 2, inverse difference and inverse difference moment, difference variance, Energy, maximum probability, sum average, sum of squares, sum entropy, and dissimilarity.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

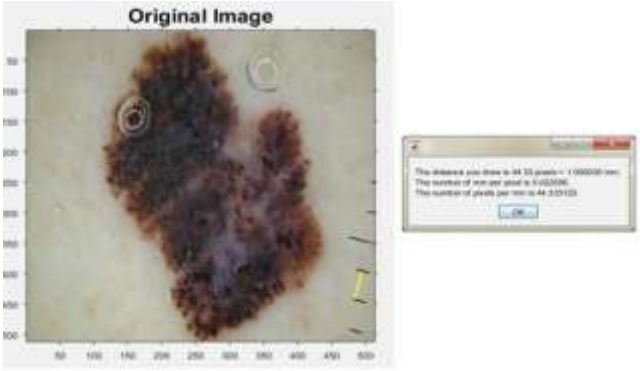


Figure 15. The scales used to calculate the actual diameter

D. **FEATURES SELECTION**

Selecting a subset of features from the prior features selection phase that may effectively describe the picture data while reducing negative effects from noise or irrelevant features and still producing respectable prediction results is the aim of feature selection. Principal component analysis (PCA) was employed in this investigation. One well-liked linear feature extractor that aids in data interpretation is PCA. It preserves trends and patterns while simplifying the intricacy of high-dimensional data. Better regression models may be produced by transforming the data with PCA, which will reduce dimensions. As an alternative to the covariance matrix, the correlation matrix was used with PCA.

A collection of key components was collected and arranged based on their capacity to distinguish between benign and malignant lesions following the use of this operation and the measurement of variances and eigenvalues [36].

In order to rank them according to their prediction outcomes, the features and their prediction efficiency were sorted. Finally, contrast, kurtosis, entropy, circulation, correlation, homogeneity, TDS, autocorrelation, cluster shade, sum variance, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference, and inverse difference moment were selected as the top 15 features with the highest efficiency.

E. **CLASSIFICATION**

The final step in the optimal workflow for computerized analysis is lesion categorization. Different skin illnesses can be defined by the binary, ternary, or n-array lesion categorization results. As a result, these results indicate the kinds (classes) of PSLs from which a system is trained. In this study, the SVM radial based function (RBF) kernel classifier will be used.

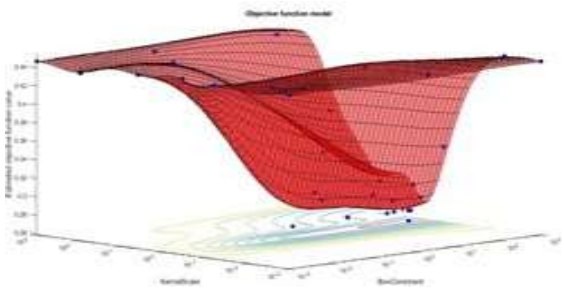


Figure 16. Hyperparameter optimization (kernelScale=gamma/ box- Constraint=cost)

1) SUPPORT VECTOR MACHINE (SVM)

We began by first creating the dataset, which includes 220 class labels and the 15 characteristics that correlate to them (chosen from the features selection phase). Second, 114 observations from the previous step were chosen for the training stage, and they were split into 85 training observations and 29 testing observations at random.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Then, we used k-fold Cross Validation (CV) with Bayesian optimization to optimize the fit of the SVM classifier. To prepare the validation set, we divided the training set into K equal partitions at random. Each partition was kept as the validation set for testing the model, and the remaining K-1 partitions were used as training. This process was repeated K times, or k folds, with each partition being used exactly once as the validation set. In order to avoid overfitting, we finally improved the fit to obtain a decent fit (one with a low cross-validation loss). Consequently, CV did assist in improving the hyperplane parameters (Gamma and Cost) [37].

While C, or the cost parameter, trades off training observation misclassification against decision boundary simplicity, the gamma parameter determines the extent to which a single training observation has an impact. The hyper parameter optimization is shown in Fig. 16.

Lastly, the classification model was constructed using kernel SVM with radial basis function because the data handled here is nonlinear. In order to handle situations when the relationship between attributes and class labels is nonlinear, this kernel nonlinearly translates data into a higher dimensional space.

F. **EXPERIMENT CASES**

Two tiers of experiments were carried out in order to assess the efficacy of the categorization model employed in this study. First, a comparison with one of the most advanced research is made [36]. Case 1: Eleven features (contrast, skewness, kurtosis, entropy, mean, standard deviation, circulation, energy, correlation, homogeneity, and TDS) and 76 data samples were given into the classification model. These eleven characteristics are identical to those utilized in [36]. Case 2: Despite using the same characteristics, there were 114 data samples instead of just 7.

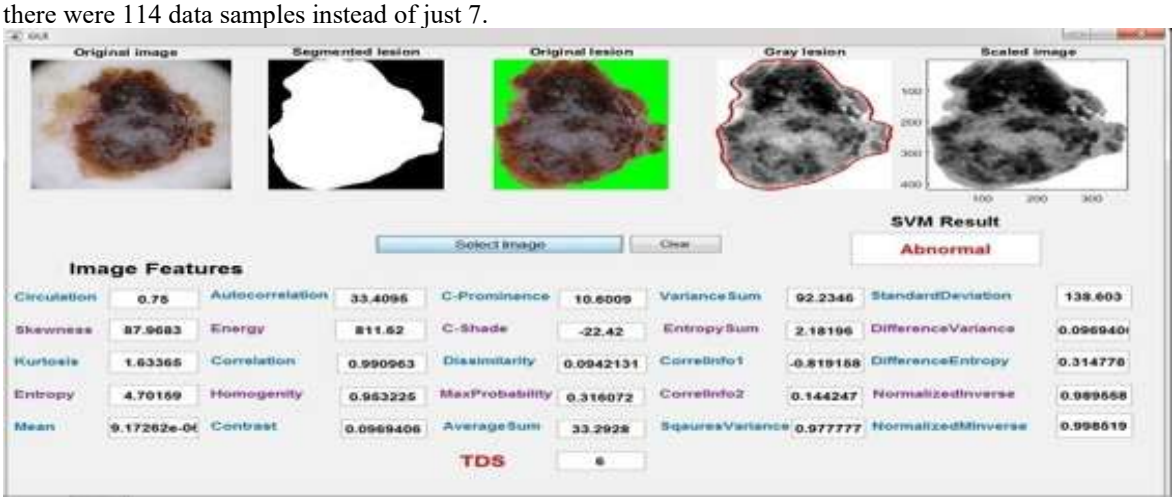


Figure 17. Graphical User Interface

Finding the features that will improve the CAD system's accuracy is the goal of the second stage of testing. In example 3, 114 data samples and the common TDS feature—which contains the ABCD rule features—were given into the classification model. Case 4 was created by comparing Case 3 with the same 114 data samples and the 15 characteristics selected using the PCA approach.

Each of the 220 images was chosen independently and diagnosed four times, each time using one of the previously listed cases, in order to assess the diagnostic role. The original picture, segmented lesion, original lesion, gray lesion, scaled image, 26 feature values, and SVM disease prediction result are all included in the program that was done with the interface that is displayed in Figure 17.

5) **RESULTS**

The confusion matrix for each of the four scenarios is displayed in Table I. For 79 observations, the model accurately predicted a normal lesion; however, it incorrectly labeled 31 normal observations as abnormal. However, the model misclassified 24 abnormal observations as normal while accurately predicting 86 abnormal lesion observations.

Experiment case

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Case 1	79	86	24	31
Case 2	101		94	16 9
Case 3	95	90	20	15
Case 4	106	100	10	4

TABLE I. The confusion matrix for the four cases of experiments

The model successfully predicted 101 normal observations in the case 2 experiment, which included 114 data samples and 11 features. The remaining 9 normal observations were incorrectly labeled as abnormal. However, the model misclassified 16 abnormal observations as normal while accurately predicting 94 abnormal lesion observations.

Regarding Case 3 (114 data samples and TDS feature) experiment. Results show that the model misclassified 15 normal data as abnormal while accurately predicting 95 normal lesion observations. However, the model misclassified 20 abnormal observations as normal while accurately predicting 90 abnormal lesion observations.

For the Case4 experiment, which includes 15 features and 114 data samples. Results show that the model misclassified four normal observations as abnormal while correctly predicting 106 normal lesion observations. However, the model misclassified the remaining 10 abnormal data as normal while accurately predicting 100 abnormal lesion observations.

Performance metrics for accuracy, recall, and precision were computed to assess the study's effectiveness. In general, the accuracy value indicates how frequently the classifier model is accurate. The recall value indicates how frequently the model predicts abnormal when the lesion's class label is in fact abnormal. The precision number indicates the frequency with which the model correctly predicts an aberrant image. These findings for the four tests are displayed in Table II.

Experiment Case	Accuracy	Recall	Precision
Case 1	75%	76.7%	71.8%
Case 2	88.6%	86.3%	91.8%
Case 3	84%	82.6%	86.3%
Case 4	93.6%	91.3%	96.3%

TABLE II. Accuracy, recall, and precision performance measures

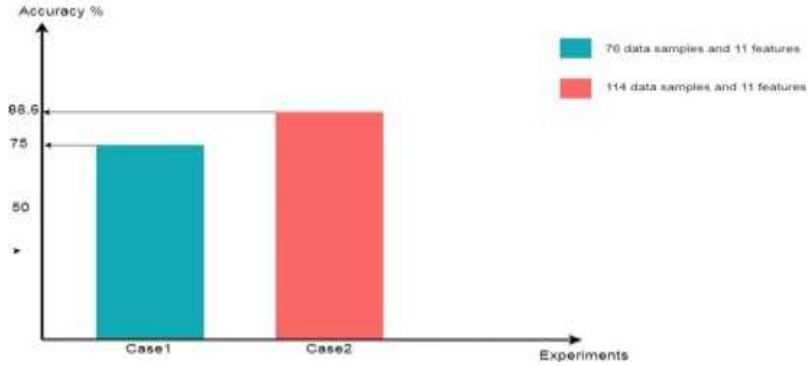
The weighted average of the precision and detection probability (recall) is the F1-score value. Additionally, it occasionally serves as a statistical measure of precision (Table III).

Experiment Case	F1 Score
Case 1	74.1%
Case 2	88.9%
Case 3	84.4%
Case 4	93.8%

TABLE III. F1-Score Results

At first, the study was successful in obtaining the ABCD characteristics. resulting from using a novel technique to extract the color feature that involves removing the black backdrop that isn't important. Additionally, by scaling it to a real-world distance, the study accurately determined the lesion width.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/



Figur 18. (Case1/ Case2) accuracy diagram

Second, PCA was used to prioritize the GLCM features and their prediction efficiency. Contrast, circulation, correlation, homogeneity, autocorrelation, cluster shade, sum variance, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference, and inverse difference moment were the features that were selected as having the maximum efficiency.

The accuracy diagrams for Case1 and Case2 tests are displayed in Fig. 18. Case1 had an accuracy of 75%, while Case2 had an accuracy of 88.6%. Since the number of data samples supplied into the classification model rose from 76 to 114 (to be handled as train set and test set), it is evident that the growing accuracy value for Case2 refers to the amount of data samples. This variation in values so demonstrated that the more data samples incorporated into the classifier, the more accurate the diagnosis will be.

Since the same number of characteristics and data samples were utilized as in [36], an additional finding from the Case1 experiment may be presented. In their study, the accuracy result was 92.1%; this fall in accuracy was brought about by the cross validation method used, namely k-fold Cross Validation (CV) with Bayesian optimization. Rather, they employed the widely used k-fold Cross Validation in their study.

In conclusion, because the Bayesian optimization algorithm performed exceptionally well in Cases 3 and 4, it demonstrated that it performs better with large sets of data samples than with small ones. The accuracy diagram for the Case3 and Case4 studies is displayed in Fig. 19, with Case3 having an accuracy of 84% and Case4 having 93.6%. Since the most pertinent 15 texture characteristics were retrieved and added to the classifier along with the ABCD features, it is clear that the growing accuracy value for Case4 reflects to the quantity of features that have been fed into the classification model.

However, because overfitting is more likely to result in incorrect classifications, this variance of values does not show that the accuracy of diagnosis should rise as the number of characteristics grows [38]. Ultimately, the correct selection of pertinent characteristics that enhance the diagnosis process and the rejection of features that result in overfitting and misclassification improved the accuracy.

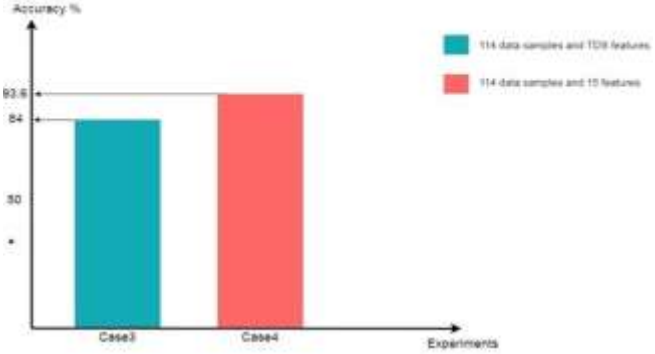


Figure19. (Case3/ Case4) accuracy diagram

causes over fitting, which results in incorrect classifications [38]. Ultimately, the correct selection of pertinent characteristics that enhance the diagnosis process and the rejection of features that result in over fitting and misclassification improved the accuracy.

Since this study deals with medical cases involving dangerous diseases, there is a high cost/risk associated with false negatives. The recall metric is the most effective model metric to choose the model specification. When it

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

is preferable to have fewer false negatives in exchange for more false positives, recall is more significant than precision. In other words, a false positive is less expensive than a false negative.

6. CONCLUSION

In order to identify the best features that could improve the diagnosis process' accuracy, this study examined ABCD features and other textural features using an SVM with RBF kernel classifier. In summary, the findings indicate the following: Compared to the standard ABCD rule, which increased accuracy from 84% to 93.6%, the newly added texture features did boost accuracy by 9.6%. Contrast, circulation, correlation, homogeneity, autocorrelation, cluster shade, sum variance, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference, and inverse difference moment were the texture features that performed the best. Furthermore, the suggested approach outperforms several algorithms found in the literature in terms of accuracy. [39].

The accuracy of the diagnosis will increase with the number of data samples added to the classifier. It has been demonstrated that the Bayesian optimization technique performs better with large data sets than with small data sample sets. Increasing the number of characteristics does not always improve the classification outcome. To enhance the diagnosis process and avoid over fitting, the classifier requires additional training data as the number of characteristics rises.

The study's findings indicate that more significant efforts are required in the field to develop an effective classifier with a respectable cross-validation technique. To lessen overfitting on models, it may also be advised to investigate methods like data augmentation, which meaningfully expands the size of melanoma training data. Additionally, it is necessary to investigate the development and/or emergence of new lesions, as this may aid in the early diagnosis of MM. In order to compare the accuracy results and choose the best artificial intelligence methodology for diagnosing with greater accuracy, future study will employ deep learning techniques to categorize such images in addition to another classification method, such as ANN or DT, in addition to SVM.

REFERENCES

[1] T. Tarver, “Cancer facts & figures 2012. american cancer society (acs),” Journal of Consumer Health on the Internet, vol. 16, no. 3, pp. 366–367, 2012. [Online]. Available: https: //doi.org/10.1080/15398285.2012.701177

[2] M. N. E. Arani and H. Ghasemian, “A hierarchical content-based image retrieval approach to assisting decision support in clinical der- matology,” Iranian journal of electrical and computer engineering, vol. 9, pp. 23–33, 2010.

[3] S. Pathan, K. G. Prabhu, and P. Siddalingaswamy, “Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review,” Biomedical Signal Processing and Control, vol. 39, pp. 237–262, 2018. [Online]. Available: https: //www.sciencedirect.com/science/article/pii/S1746809417301428

[4] G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, G. Ferrara et al., “Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet,” Journal of the American Academy of Dermatology, vol. 48, no. 5, pp. 679–693, 2003.

[5] F. Del Rosario, J. M. Farahi, J. Drendel, T. Buntinx-Krieg, J. Caravaglio, R. Domozych, S. Chapman, T. Braunberger, R. P. Dellavalle, D. A. Norris et al., “Performance of a computer-aided digital dermoscopic image analyzer for melanoma detection in 1,076 pigmented skin lesion biopsies,” Journal of the American Academy of Dermatology, vol. 78, no. 5, pp. 927–934, 2018.

[6] F. Warsi, R. Khanam, S. Kamya, and C. P. Sua´rez-Araujo, “An efficient 3d color-texture feature and neural network technique for melanoma detection,” Informatics in Medicine Unlocked, vol. 17, p. 100176, 2019.

[7] M. A. Kadampur and S. Al Riyaee, “Skin cancer detection: Apply- ing a deep learning based model driven architecture in the cloud for classifying dermal cell images,” Informatics in Medicine Unlocked, vol. 18, p. 100282, 2020.

[8] A. K. Verma, S. Pal, and S. Kumar, “Comparison of skin disease pre- diction by feature selection using ensemble data mining techniques,” Informatics in Medicine Unlocked, vol. 16, p. 100202, 2019.

[9] K. M. Nahar, R. M. Al-Khatib, M. A. Al-Shannaq, and M. M. Barhoush, “An efficient holy Quran

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

recitation recognizer based on SVM learning model," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 6, no. 04, 2020.

[10] J. Hacker, N. Wickramasinghe, and C. Durst, "Can health 2.0 address critical healthcare challenges? insights from the case of how online social networks can assist in combatting the obesity epidemic," Australasian journal of information systems, vol. 21, pp. 1–17, 2017.

[11] S. Pathan, K. G. Prabhu, and P. Siddalingaswamy, "Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review," Biomedical Signal Processing and Control, vol. 39, pp. 237–262, 2018.

[12] P. G. Cavalcanti and J. Scharcanski, "Automated prescreening of pigmented skin lesions using standard cameras," Computerized Medical Imaging and Graphics, vol. 35, no. 6, pp. 481–491, 2011.

[13] M. Ferri, I. Tomba, A. Visotti, and I. Stanganelli, "A feasibility study for a persistent homology-based k-nearest neighbor search algorithm in melanoma detection," Journal of Mathematical Imaging and Vision, vol. 57, no. 3, pp. 324–339, 2017.

[14] J. S. Lim, "Two-dimensional signal and image processing," Engle- wood Cliffs, 1990.

[15] N. Otsu, "A threshold selection method from gray-level histograms," IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1,pp. 62–66, 1979.

[16] K. M. Nahar, B. Abul-Huda, A. Abu Naser, and R. M. Al- Khatib, "Twins and Similar Faces Recognition Using Geometric and Photometric Features with Transfer Learning," International Journal of Computing and Digital System, vol. 11, no. 01, 2022.

[17] C. Y. Wong, S. Liu, S. C. Liu, M. A. Rahman, S. C.-F. Lin, G. Jiang, N. Kwok, and H. Shi, "Image contrast enhancement using histogram equalization with maximum intensity coverage," Journal of Modern Optics, vol. 63, no. 16, pp. 1618–1629, 2016.

[18] I. G. Maglogiannis and E. P. Zafiropoulos, "Characterization of digital medical images utilizing support vector machines," BMC Medical Informatics and Decision Making, vol. 4, no. 1, pp. 1–9, 2004.

[19] A. Bono, S. Tomatis, C. Bartoli, G. Tragni, G. Radaelli, A. Maurichi, and R. Marchesini, "The abcd system of melanoma detection: A spectrophotometric analysis of the asymmetry, border, color, and dimension," Cancer: Interdisciplinary International Journal of the American Cancer Society, vol. 85, no. 1, pp. 72–77, 1999.

[20] V. T. Ng, B. Y. Fung, and T. K. Lee, "Determining the asymmetry of skin lesion with fuzzy borders," Computers in biology and medicine, vol. 35, no. 2, pp. 103–120, 2005.

[21] T. Satheeshsa, D. Sathyanarayana, and M. Giriprasad, "Detection of in-situ melanoma using symmetry of data and color spread factor," International Journal of Engineering Research, vol. 3, no. SP 2, pp. 64–67, 2014.

[22] H. Alquran, I. A. Qasmieh, A. M. Alqudah, S. Alhammouri, E. Alawneh, A. Abughazaleh, and F. Hasayen, "The melanoma skin cancer detection and classification using support vector machine," in 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT). IEEE, 2017, pp. 1–5.

[23] H. Alquran, E. Shaheen, J. M. O'Connor, and M. Mahd, "En- hancement of 3d modeling and classification of microcalcifications in breast computed tomography (bct)," in Medical Imaging 2014: Image Processing, vol. 9034. SPIE, 2014, pp. 799–807.