BIOMEDICAL SYSTEM DESIGN CAPSULE

PROJECT REPORT
Furkan Samed Şansaçar - Fatih Çetin

Furkan Samed Şansaçar
Department of Electrical and Electronics Engineering
Abdullah Gul University
Kayseri/ Turkey
furkansamed.sansacar@agu.edu.tr

Fatih Çetin
Department of Electrical and Electronics Engineering
Abdullah Gul University
Kayseri/ Turkey
fatih.cetin@agu.edu.tr

Abstract

Skin cancer is a significant health concern worldwide, and early diagnosis is crucial. Traditional diagnostic methods have limitations in distinguishing between basal cell carcinoma and squamous cell carcinoma. This article aims to utilize image processing and artificial intelligence techniques to determine whether skin cancer is benign or malignant. The dataset used in this study was obtained from Kaggle, and MATLAB and Python were employed for image processing, including denoising. The artificial intelligence techniques utilized the auto Weka library in Weka, employing more than 250 combinations of classifiers and attribute selectors to automatically detect the relationship between extracted features. Cross-validation was employed to evaluate performance, resulting in accuracy rates of 79.20% (RFE and XGB Classifier) and 85.93% (Multilayer Perceptron and Greedy Stepwise) in two distinct models. The findings of this study demonstrate that these artificial intelligence techniques can effectively distinguish between benign and malignant skin cancer with an accuracy rate of approximately 85%, which can be further improved by utilizing more data, advanced algorithms, and enhanced image processing techniques. Although a model with 100% accuracy is not currently available, this approach offers a rapid and cost-effective diagnostic method, reducing the need for invasive procedures such as biopsy and minimizing patient discomfort. While there is currently no public application, the proposed techniques can be easily utilized through a website or mobile application. However, it should be acknowledged that artificial intelligence does not always provide complete and accurate results, as errors may occur due to the training dataset or algorithmic biases, potentially leading to misdiagnosis and patient harm. The study was identified certain features to determine skin cancer's class, standard deviation, skewness, perimeter, gray level variance, and homogeneity (GLCM), that are associated with skin cancer. The literature encompasses numerous studies focusing on skin cancer diagnosis, noise removal from cancer images (e.g., hair, skin artifacts), segmentation, and machine learning. Further improvements can be made in various aspects. The primary beneficiaries of this study are hospitals, doctors, and institutions that lack access to pathology laboratories or have limited resources. Furthermore, the study was designed to alleviate the workload of doctors. In conclusion, this study presents a practical approach that enhances the skin cancer diagnostic process by leveraging artificial intelligence techniques, providing rapid results within seconds.

Index Terms
Skin Cancer, Benign and Malignant Skin Cancer

1. INTRODUCTION

Skin cancer is a significant public health issue, and early and accurate diagnosis is crucial for effective treatment. Traditional diagnostic methods for distinguishing between different types of skin cancer, such as basal cell carcinoma and squamous cell carcinoma, have limitations. This calls for a faster, cost-effective, and reliable diagnostic approach. In this study, we propose a novel approach that combines image processing and artificial intelligence techniques to determine whether a skin lesion is benign or malignant. By leveraging a dataset obtained from Kaggle, we perform image preprocessing, including denoising, followed by feature extraction. Subsequently, we employ artificial intelligence algorithms for feature selection and classification, enabling the identification of malignant skin cancer. The primary beneficiaries of this research are hospitals, clinics, and healthcare providers with limited access to pathology laboratories or financial resources. Additionally, this approach aims to alleviate the workload of healthcare professionals. Traditional pathology results can take anywhere from 2 to 7 days, whereas our proposed method provides near-instantaneous results, obtained within seconds. By offering a fast and cost-effective diagnostic solution, this study aims to contribute to improved skin cancer management and patient outcomes.

Please refer to Figure 1 for the workflow illustrating the different stages of the proposed methodology. The workflow encompasses the following steps: skin cancer image acquisition, preprocessing, segmentation, feature extraction, feature selection algorithms, algorithm selection, training and evaluation, and classification into benign

or malignant. The iterative process of feature selection is employed until the best accuracy is achieved. This workflow visually demonstrates the sequential and iterative nature of the proposed approach, highlighting the essential components and steps involved in accurately categorizing skin cancer lesions.

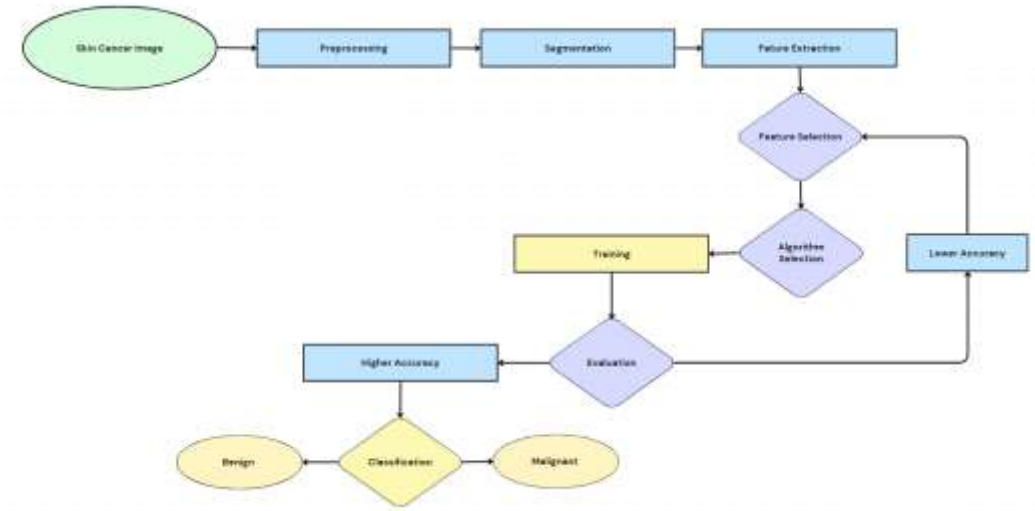

Figure 1: Workflow illustrating the stages of the proposed methodology for skin cancer classification.

2. DATA ACQUISITION

The "Skin Cancer: Malignant vs. Benign" [6] data set used in this study was obtained from Kaggle. The data set consists of two distinct folders, namely "train" and "test." The "train" folder contains a total of 1197 images depicting aggressive skin cancers and 1440 images representing benign skin cancers. On the other hand, the "test" folder comprises 299 photographs of malignant skin cancers and 360 images of benign skin cancers.

These images were collected from various clinical sources, ensuring a diverse representation of skin cancer cases. The inclusion of both benign and malignant skin cancer images allows for a comprehensive analysis and evaluation of the proposed methodology.

3. PRE-PROCESSING

Preprocessing plays a crucial role in the accurate analysis of skin cancer images using image processing and artificial intelligence techniques. In the presented study, a series of preprocessing steps were performed to enhance the quality and relevance of the images, the images were converted from RGB to grayscale to reduce the color information and focus solely on the structural details related to the cancerous regions. By carefully evaluating the impact of each preprocessing technique on the overall accuracy, adjustments were made to refine the preprocessing pipeline and optimize the subsequent analysis. The meticulous preprocessing steps employed in this study were instrumental in preparing the skin cancer images for subsequent feature extraction and classification using artificial intelligence algorithms, ultimately leading to more accurate and reliable predictions. Additionally, the images were resized to a standardized dimension of 244 x 244 pixels to ensure consistency throughout the dataset. Furthermore, a series of morphological operations, such as opening and closing, were initially applied, but due to the observed loss of information in cancerous areas, these operations were omitted from the preprocessing pipeline. Subsequently, denoising operation was performed to original images (Figure 2) the initial-colored photographs of skin cancer used in the study [6].

Figure 2 : The initial-colored photographs of skin cancer used in the study (benign) [6].

First step converting the images to grayscale (Figure 3) and applying a black hat operation (Figure 4), creating mask (Figure 5) and applying denoising to the original images.

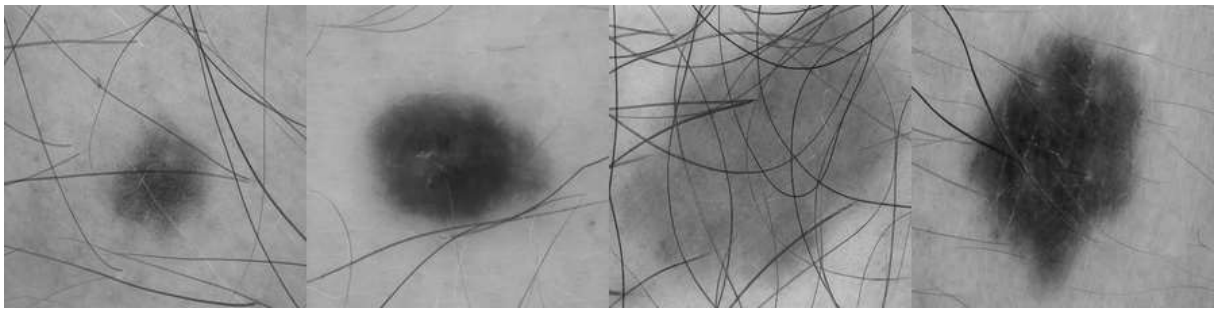

Figure 3: The images obtained by converting the original images to grayscale during the preprocessing stage.

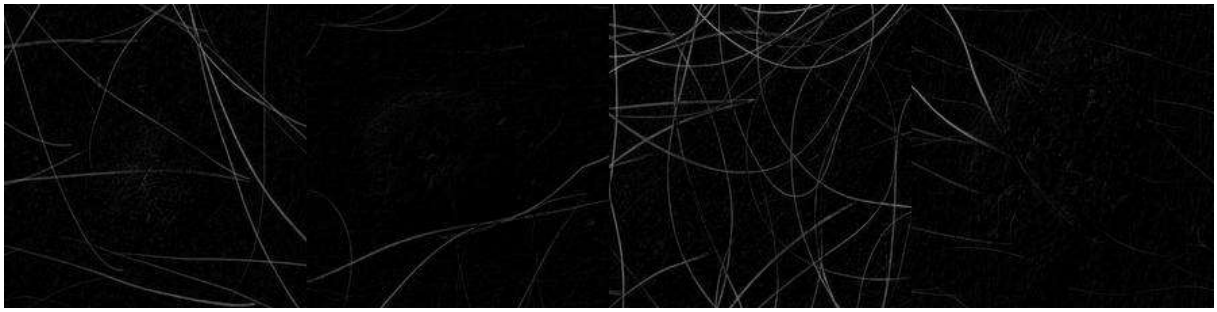

Figure 4: The images obtained after applying the black-hat transformation during the preprocessing stage

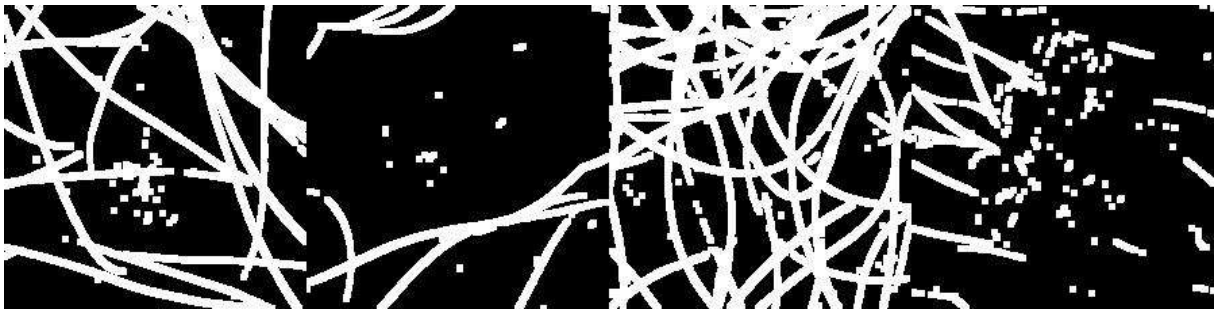

Figure 5: The images obtained by creating masks during the preprocessing stage, where only the cancerous areas are preserved

This step effectively eliminated unwanted artifacts such as hairs, skin blemishes, or marker rulers, while preserving the cancerous regions of interest. Denoised images are obtained after denoising through the preprocessing process at Figure 6 with the comparison of original image.

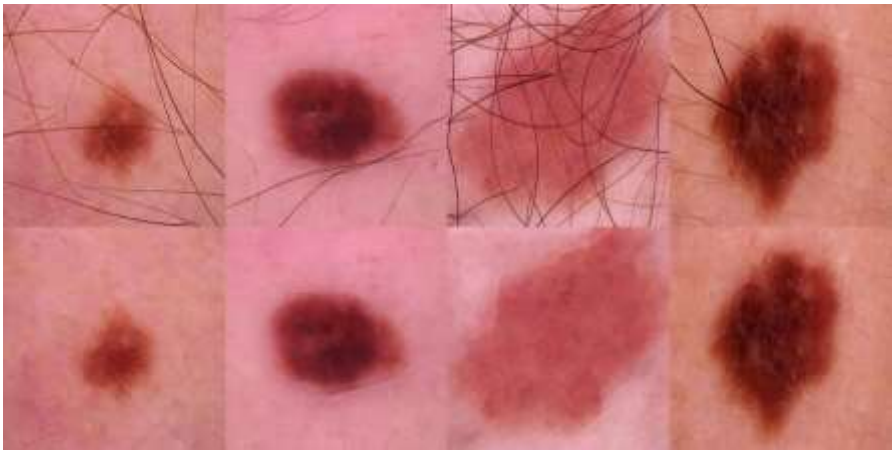

Figure 6: The comparison of images obtained after denoising through the preprocessing process with original images

4. SEGMENTATION

Segmentation was performed using the active contour model with the Chan-Vese method. Active contour model, also known as snakes, is a popular technique used in image processing and computer vision for object segmentation. It is a deformable model that iteratively adjusts its shape to fit the contours of objects in an image. The active contour model is driven by energy minimization principles, aiming to find the optimal boundary that separates the object of interest from the background. The Chan-Vese algorithm is a specific formulation of the active contour model that is widely used for image segmentation tasks. It was proposed by Tony Chan and Luminita Vese in their paper titled "An Active Contour Model Without Edges" published in 2001 [4]. The Chan-Vese method is particularly suitable for segmenting objects with smooth boundaries or objects with weak edges. It formulates the energy functional to be minimized as a combination of region-based and edge-based terms. By iteratively minimizing this energy functional, the algorithm evolves the contour to achieve a segmentation that optimally separates the object from the background.

The images were first converted to grayscale, and then a mask was created using the active contour model. The original image was then multiplied by the mask to apply the segmentation. The Chan-Vese method with 100 iterations was used, taking into account the computational power and time constraints figure 7 shows the original images, masks and segmentation results.

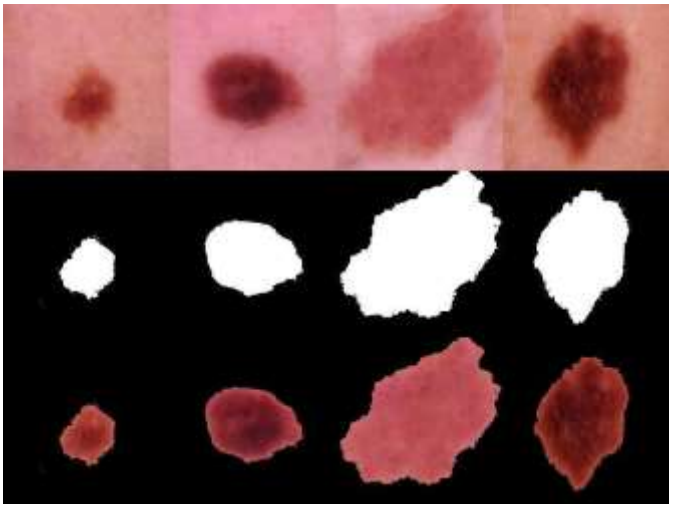

Figure 7

The segmentation process involved different numbers of iterations, including 10, 50, 100, and 200 iterations. The images were first converted to grayscale, and then a mask was created using the active contour model for each iteration. The original image was then multiplied by the respective mask to apply the segmentation. The purpose of using different iteration values was to evaluate the impact on the segmentation results.

Figure 8 shows the segmentation results obtained after 10 iterations. Despite the limited number of iterations, the method was able to capture the main cancer regions effectively. However, there were some inaccuracies and incomplete segmentation in smaller or more complex areas.

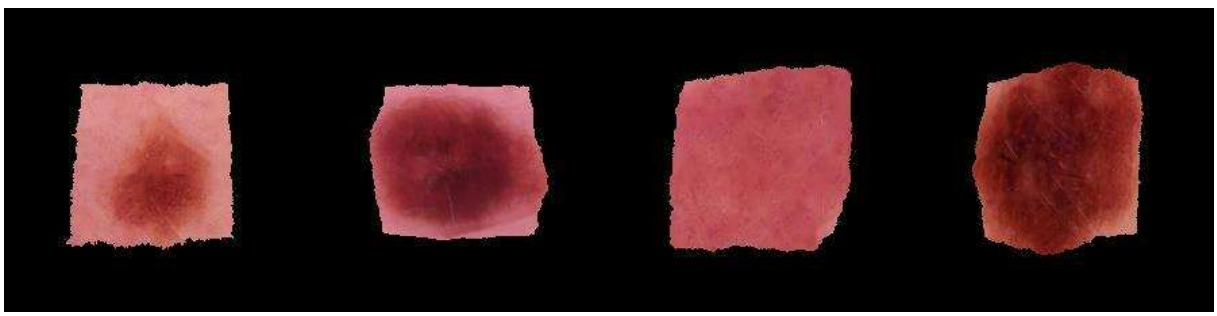

Figure 8: The segmentation results obtained after 10 iterations

Figure 9 illustrates the segmentation results achieved after 50 iterations. With increased iterations, the method exhibited improved accuracy and better delineation of cancerous regions. Smaller cancer areas showed better segmentation, although there were still some minor imperfections.

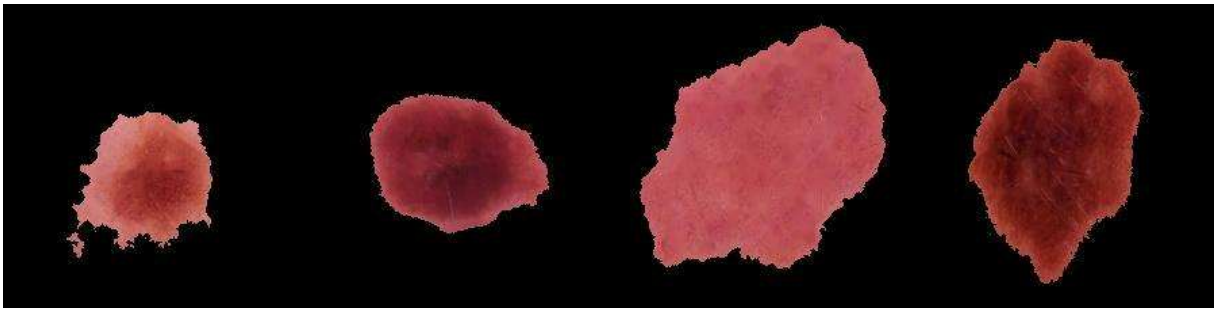

Figure 9: The segmentation results obtained after 50 iterations

Figure 10 displays the segmentation results obtained after 100 iterations. At this stage, the method achieved even higher accuracy and more precise delineation of cancerous areas. The majority of the cancer regions were successfully segmented, with minimal errors or false positives.

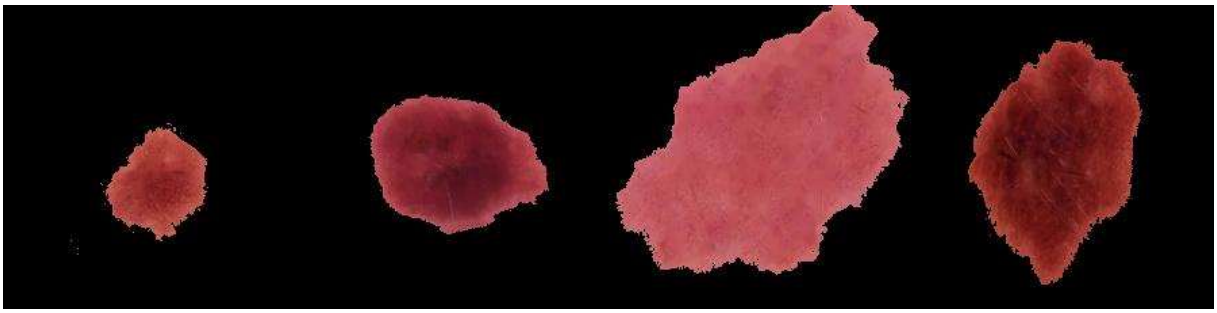

Figure 10: The segmentation results obtained after 100 iterations

Figure 11 demonstrates the segmentation results obtained after 200 iterations. With a higher number of iterations, the method reached its optimal performance, producing highly accurate segmentation results. The method effectively captured intricate details and boundaries of cancerous regions, resulting in reliable segmentation outcomes.

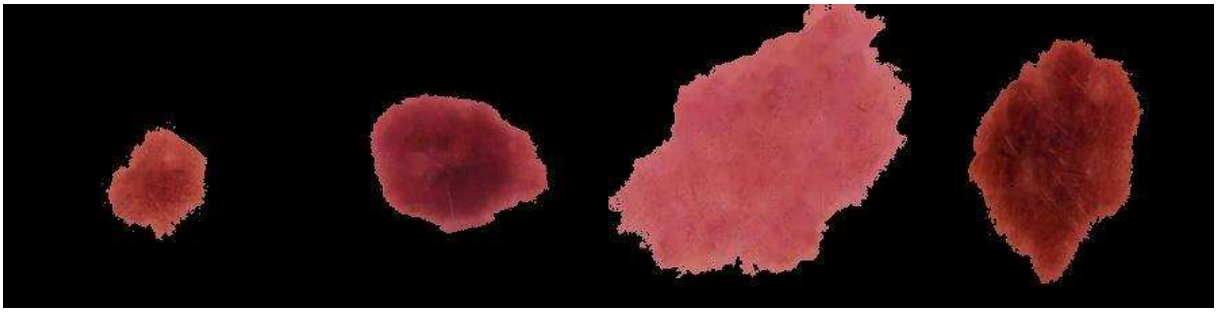

Figure 11: The segmentation results obtained after 200 iterations

The comparison of these figures indicates that increasing the number of iterations leads to improved segmentation quality. While the segmentation results after 200 iterations might showcase slightly improved delineation of intricate details and boundaries, the difference may not be substantial enough to justify the additional computational resources and time required for the extra iterations. It is worth noting that the choice of the optimal number of iterations depends on the specific requirements of the application and the desired trade-off between accuracy and efficiency.

However, it was observed that the method provided high accuracy for medium and large-sized cancer images, but performed poorly for small cancer regions some examples shown at figure 12. There were no specific constraints or limitations applied during the segmentation process. The segmented images were used for further feature extraction. It should be noted that for small-sized cancer regions, a different algorithm may be required for better segmentation results. Additionally, there were instances where non-cancerous boundaries remained in the segmented images, which could be further improved in future work. The performance and reliability of the segmentation results were evaluated, but there were no specific studies on the impact of different image conditions such as lighting, resolution, or perspective. Other alternative segmentation method thresholding is also used but not used due to low segmentation accuracy.

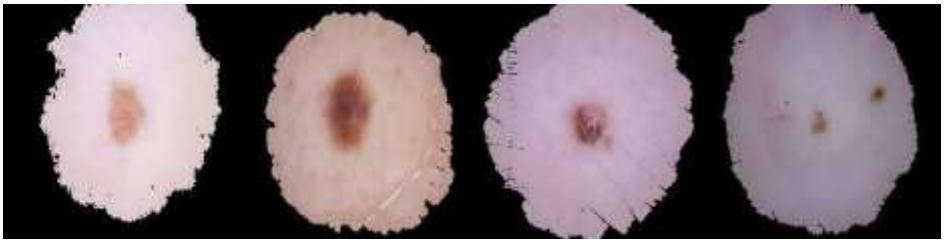

Figure 12: Malignant Examples

In summary, although there may be some nuanced differences between the segmentation outcomes obtained after 200 and 100 iterations, the overall quality and effectiveness of the segmentation process remain relatively consistent.

5. FEATURE EXTRACTION

In medical image processing studies, feature extraction methods play a crucial role in the diagnosis of skin cancer. In this study, feature extraction was performed to differentiate between malignant and benign skin lesions. Firstly, the preprocessing stage of the images was conducted, and the region of interest was determined through segmentation. Subsequently, feature extraction was performed using various feature extraction methods calculated using MATLAB for each image.

In medical image processing studies, different features are used to distinguish between malignant and benign skin lesions. These features listed below:

- Entropy (GLCM): Calculates the entropy of features obtained from GLCM.
- Flatness: Computes the flatness property of an object.
- Entropy (GLSZM): Calculates the entropy of features obtained from the gray-level size zone matrix (GLSZM).
- Energy (GLSZM): Computes the energy of features obtained from GLSZM.
- Entropy (GLDM): Calculates the entropy of features obtained from the gray-level distance matrix (GLDM).
- Energy (GLDM): Computes the energy of features obtained from GLDM.
- Intensity Range: Computes the range between the pixel values of an image.
- Mean Intensity: Calculates the mean value of pixel intensities in an image.
- Median Intensity: Calculates the median value of pixel intensities in an image.
- Ferrets Diameter: Calculates the diameter of an object based on the longest straight line.
- Perimeter: Computes the perimeter of an object.
- Solidity: Calculates the solidity (compactness) property of an object.
- Eccentricity: Calculates the eccentricity property of an object.
- Area: Calculates the area of an object.
- Elongation: Computes the elongation property of an object.

- Fractal Dimension: Calculates the fractal dimension of an object.
- Roundness: Computes the roundness property of an object.
- Minimum Intensity: Calculates the minimum pixel value of an image.
- Maximum Intensity: Calculates the maximum pixel value of an image.
- Convexity: Computes the convexity property of an object.
- Dissimilarity: Computes the dissimilarity property of an image.
- Maximum Probability: Calculates the maximum probability among the pixel values of an image.
- Gray Level Variance: Computes the variance of gray-level values in an image.
- Short Run Low Gray Level Emphasis: Calculates the emphasis of low gray-level values in short runs of an image.
- Short Run High Gray Level Emphasis: Calculates the emphasis of high gray-level values in short runs of an image.
- Long Run Low Gray Level Emphasis: Calculates the emphasis of low gray-level values in long runs of an image.
- Long Run High Gray Level Emphasis: Calculates the emphasis of high gray-level values in long runs of an image.
- Contrast GLCM: Calculates the contrast of features obtained from GLCM.
- Homogeneity GLCM: Calculates the homogeneity of features obtained from GLCM.
- Correlation GLCM: Calculates the correlation of features obtained from GLCM.
- Short Run High Gray Level Emphasis (GLRLM): Calculates the emphasis of high gray-level values in short runs of an image using GLRLM.
- Long Run High Gray Level Emphasis (GLRLM): Calculates the emphasis of high gray-level values in long runs of an image using GLRLM.
- Total Energy: Calculates the total energy of an image.
- Interquartile Range: Computes the range between the first and third quartiles of an image.
- Range: Computes the value range of an image.
- Mean Absolute Deviation (MAD): Calculates the mean absolute deviation of an image.
- rMAD: Calculates the normalized mean absolute deviation of an image.
- RMS: Calculates the root mean square of an image.
- Uniformity: Computes the similarity property of pixel values in an image.

These features are used to capture the textural, color, and edge information of skin lesions. It is mentioned that the obtained features were saved to Excel files and in some features, zero or NaN values were replaced with the mean value for further data usability later on. These extracted features can be used to achieve better results in the diagnosis of skin cancer. Selected features and how are they selected is discussed at feature extraction part. The extracted features of 4 images which used at this study can be shown at table 1.

Table 1

Images/Fe	Mean	Median	Standard_	Skewness	Kurtosis	Energy	Volume	Surface_A	Compactnes	Sphericity	Contrast	Homogen	Energy_(GLCM)	Correlatio
154.jpg	19.40842235	0	38.71725	2.265689316	7.254137109	11647480	11456.9059	52257	0.00085538	4.377335	780.5	48698.35	1303134912	6.390813
18.jpg	8.078058567	0	30.36727	4.338277567	22.05196565	3371337	4768.52549	16830	0.00309288	2.266132	780.5	49379.18	2098324098	6.502698
54.jpg	15.3504265	0	36.9429	2.714112458	9.799620492	7750010	9061.44706	34779	0.00129644	0.694451	780.5	49292.72	1638936934	6.779953
61.jpg	65.03750797	74	70.27817	0.663543807	2.139960559	20903306	38392.0235	89495	0.0004739	0.479917	780.5	48499.22	1019231996	4.288808

Images/Fe	Elongation	Fractal_Di	Roundnes	Minimum_Inte	Maximum_Inter	Perimeter	Convexity	Dissimilar	Maximum_P	Gray_Leve	Short_Run	Short_Run	Long_RunLow_G	Long_Run_
154.jpg	4.377334775	-2.02536	3.002495	0	0.517647059	981.5	14691	0.531871	0.76405349	0.013858	0.933345	0.076153	0.000120115	0.010939
18.jpg	2.266131749	-2.08545	3.258321	0	0.592156863	323.5	4063	0.189822	0.91373719	0.010685	0.976194	0.027206	4.00384E-05	0.003917
54.jpg	0.694451392	-5.49539	3.517477	0	0.564705882	428.625	9693	0.261811	0.84491111	0.013895	0.967213	0.037436	8.00769E-05	0.005363
61.jpg	0.479916727	-6.12168	2.457872	0	0.650980392	915.5	26517	0.839206	0.61446989	0.052133	0.895004	0.119975	8.00769E-05	0.017174

Images/Fe	Entropy_(GLC	Flatness	Entropy_(	Energy_(GLSZM	Entropy_(GLDM)	Energy_(GLDM	Intensity_Ra	Mean_Inte	Median_Inte	Ferets_Dia	Perimeter	Solidity	Eccentricity	Area
154.jpg	-488637.209	0.549837	0.928362	1303134912	0.942163169	1527404390	0.51764706	0.073334	0	224	78.03663	0.989216	0.398509882	1836.375
18.jpg	-518671.274	0.404756	0.912999	2098324098	0.811278124	2092882728	0.59215686	0.030867	0	224	19.90294	0.982383	0.242539062	253.9375
54.jpg	-500722.388	0.23508	0.942163	1638936934	0.69621226	1817491418	0.56470588	0.055944	0	224	139.5877	0.982252	0.225408389	3231
61.jpg	-485343.708	0.195424	0.965202	1019231996	0.757878463	1135867978	0.65098039	0.243677	0.40784314	224	833.267	0.90983	0.777594199	26517

Features	ContrastGLCM	Homogen	Correlatio	Short_Run_Hig	Long_Run_High_	Total_Energy	Interquartile	Range	Mean_Absol	rMAD	RMS	Uniformit	Spherical_Dispr	GLN
154.jpg	0.105861627	0.974903	0.907634	2.13058088	186914261.3	255	22	202	18106.0934	0.050007	14.07693	4342.11	0.003742611	4.609224
18.jpg	0.069346573	0.988533	0.929751	1.054227957	62421034.67	255	0	218	9807.55794	0.000159	14.03218	2285.803	0.006478079	4.59892
54.jpg	0.063380846	0.986802	0.946972	1.573631591	114791040	255	0	217	16355.7643	0.032327	14.06094	3704.816	0.004809378	4.515764
61.jpg	0.212003523	0.970916	0.957739	2.45385643	277068288	255	96	247	40554.0095	0.202553	14.25771	21225.24	0.004893687	3.217874

6. FEATURE SELECTION

The extracted features from the skin cancer images contain a large amount of information, but not all of them may be relevant for distinguishing between benign and malignant lesions. Feature selection is performed to identify the most informative and discriminative features that contribute significantly to the classification task. In this study, various feature selection algorithms were employed to select the optimal subset of features. Also In this study, the auto Weka library in Weka, a popular machine learning toolkit, was utilized for automated feature selection and classification. The auto Weka library offers a wide range of classifiers and attribute selectors that can be combined in various ways to automatically detect the relationship between the extracted features and the target class (benign or malignant). More than 250 combinations of classifiers and attribute selectors were employed to find the optimal configuration that maximizes the classification accuracy. Based on the applied methods, the top 5 features that effectively and efficiently train the algorithm can be ranked as follows: Standard Deviation, Skewness, Perimeter, Gray Level Variance, and Homogeneity GLCM.

Addition to Weka, various models were utilized in the project, and the following is a list of the models used:

- RandomForestClassifier
- SVC
- XGBClassifier
- AdaBoostClassifier
- KNeighborsClassifier
- GradientBoostingClassifier
- DecisionTreeClassifier
- GaussianNB
- LogisticRegression
- MLPClassifier

By employing these models, the important features determined through feature extraction methods were utilized, and the performance of different algorithms was evaluated.

For feature selection variety of attribute selection methods used:

a) Chi-Square: Chi-Square test is utilized to assess the relationship between categorical independent variables and the target variable. By employing this method, the relationship between categorical variables in the dataset and the target variable was analyzed to identify significant features results of using this method can be seen at table 2.

Feature Se	Classifier	Accuracy	Precision	Recall (Se	Specificity	Confusion Matrix
Chi-Squar	RandomFo	0.694	0.663	0.666	0.719	[[259, 101], [100, 199
Chi-Squar	SVC	0.686	0.661	0.632	0.731	[[263, 97], [110, 189]]
Chi-Squar	XGBClassi	0.666	0.629	0.645	0.683	[[246, 114], [106, 193
Chi-Squar	AdaBoost	0.689	0.651	0.679	0.697	[[251, 109], [96, 203]]
Chi-Squar	KNeighbo	0.701	0.666	0.686	0.714	[[257, 103], [94, 205]]
Chi-Squar	GradientB	0.695	0.661	0.672	0.714	[[257, 103], [98, 201]]
Chi-Squar	DecisionT	0.619	0.585	0.552	0.675	[[243, 117], [134, 165
Chi-Squar	GaussianN	0.628	0.584	0.629	0.628	[[226, 134], [111, 188
Chi-Squar	LogisticRe	0.663	0.681	0.485	0.811	[[292, 68], [154, 145]]
Chi-Squar	MLPClassi	0.69	0.661	0.652	0.722	[[260, 100], [104, 195

Table 2

b) Mutual Information: Mutual Information is an information measure used to evaluate the dependency between variables. Using this method, the relationship between variables and the target variable was assessed, and informative features were selected results of using this method can be seen at table 3.

Feature Se	Classifier	Accuracy	Precision	Recall (Se	Specificity	Confusion Matrix
Mutual Int	RandomFo	0.697	0.684	0.615	0.764	[[275, 85], [115, 184]]
Mutual Int	SVC	0.725	0.714	0.659	0.781	[[281, 79], [102, 197]]
Mutual Int	XGBClassi	0.707	0.691	0.642	0.761	[[274, 86], [107, 192]]
Mutual Int	AdaBoost(	0.715	0.696	0.659	0.761	[[274, 86], [102, 197]]
Mutual Int	KNeighbo	0.718	0.7	0.662	0.729	[[268, 92], [102, 197]]
Mutual Int	GradientB	0.711	0.69	0.648	0.749	[[270, 90], [106, 193]]
Mutual Int	DecisionT	0.619	0.585	0.552	0.675	[[243, 117], [134, 165
Mutual Int	GaussianN	0.635	0.607	0.624	0.646	[[230, 130], [111, 188
Mutual Int	LogisticRe	0.672	0.667	0.547	0.79	[[281, 79], [140, 159]]
Mutual Int	MLPClassi	0.695	0.661	0.643	0.722	[[260, 100], [104, 195

Table 3

c) RFE (Recursive Feature Elimination): RFE is a backward selection method used to determine the importance ranking of features. This approach enabled the identification of the most important features by ranking the features based on their importance results of using this method can be seen at table 4.

Feature Se	Classifier	Accuracy	Precision	Recall (Se	Specificity	Confusion Matrix
RFE	RandomFo	0.702	0.686	0.622	0.761	[[276, 84], [114, 185]]
RFE	SVC	0.729	0.721	0.668	0.786	[[283, 77], [100, 199]]
RFE	XGBClassi	0.705	0.69	0.637	0.759	[[275, 85], [105, 194]]
RFE	AdaBoost(	0.717	0.7	0.664	0.759	[[274, 86], [101, 198]]
RFE	KNeighbo	0.718	0.7	0.662	0.73	[[268, 92], [102, 197]]
RFE	GradientB	0.713	0.694	0.654	0.753	[[271, 89], [105, 194]]
RFE	DecisionT	0.617	0.588	0.546	0.678	[[242, 117], [135, 164
RFE	GaussianN	0.638	0.61	0.624	0.652	[[231, 128], [111, 188
RFE	LogisticRe	0.674	0.672	0.556	0.792	[[282, 78], [138, 161]]
RFE	MLPClassi	0.694	0.658	0.643	0.724	[[261, 99], [104, 195]]

Table 4

d) L1-based methods: L1 regularization is a technique employed to achieve low-dimensional and sparse solutions. By employing L1-based methods, important features in the dataset were identified, and dimensionality reduction was performed while eliminating irrelevant features results of using this method can be seen at table 5.

Feature Se	Classifier	Accuracy	Precision	Recall (Se	Specificity	Confusion Matrix
L1-based	RandomFo	0.721	0.71	0.654	0.776	[[282, 78], [109, 190]]
L1-based	SVC	0.737	0.728	0.677	0.798	[[287, 73], [101, 198]]
L1-based	XGBClassi	0.713	0.695	0.643	0.758	[[275, 85], [105, 194]]
L1-based	AdaBoost(	0.724	0.71	0.678	0.756	[[277, 83], [101, 198]]
L1-based	KNeighbo	0.715	0.697	0.66	0.746	[[270, 90], [102, 197]]
L1-based	GradientB	0.719	0.703	0.66	0.757	[[273, 87], [102, 197]]
L1-based	DecisionT	0.615	0.584	0.542	0.675	[[241, 119], [135, 164
L1-based	GaussianN	0.642	0.612	0.629	0.654	[[233, 127], [109, 190
L1-based	LogisticRe	0.676	0.672	0.556	0.792	[[282, 78], [138, 161]]
L1-based	MLPClassi	0.696	0.659	0.645	0.727	[[261, 99], [103, 196]]

Table 5

e) Greedy Stepwise: The Greedy Stepwise method aims to sequentially select the best feature among candidate features during feature selection. By using this method, features in the dataset were selected step by step to determine the best combination.

Feature Se	Classifier	Accuracy	Precision	Recall (Se	Specificity	Confusion Matrix
Greedy St	RandomFo	0.724	0.711	0.665	0.769	[[283, 77], [107, 192]]
Greedy St	SVC	0.736	0.727	0.674	0.796	[[286, 74], [102, 197]]
Greedy St	XGBClassi	0.715	0.699	0.655	0.752	[[276, 84], [104, 195]]
Greedy St	AdaBoost(	0.72	0.709	0.665	0.761	[[281, 79], [107, 192]]
Greedy St	KNeighbo	0.711	0.695	0.655	0.743	[[271, 89], [104, 195]]
Greedy St	GradientB	0.72	0.705	0.662	0.76	[[280, 80], [106, 193]]
Greedy St	DecisionT	0.617	0.584	0.546	0.677	[[241, 119], [133, 166
Greedy St	GaussianN	0.638	0.608	0.627	0.649	[[232, 128], [110, 189
Greedy St	LogisticRe	0.674	0.669	0.56	0.788	[[281, 79], [136, 163]]
Greedy St	MLPClassi	0.692	0.656	0.643	0.721	[[259, 101], [103, 196

Table 6

7. EVALUATION

Using these techniques, the relationships between features in the dataset were analyzed, important features were identified, and the most suitable features for the model were selected. To evaluate the performance of the selected models, cross-validation was used. The dataset was randomly divided into several subsets (folds), and the models were trained and tested on different combinations of these subsets. This process was repeated multiple times to ensure reliable and unbiased performance evaluation. Two distinct models were developed in this study. The first model utilized the Recursive Feature Elimination (RFE) attribute selector combined with the eXtreme Gradient Boosting (XGB) classifier. The second model employed the Multilayer Perceptron classifier and the Greedy Stepwise attribute selector. The performance of the models was measured using accuracy, which represents the percentage of correctly classified instances. The first model achieved an accuracy rate of 79.20%, while the second model achieved an accuracy rate of 85.93%.

As can be seen from Figure 13, when the Chi-Square selection method is used, the accuracy value generally increases as the number of features increases. In general, the highest accuracy is observed with the Random Forest classifier and MLP classifier.

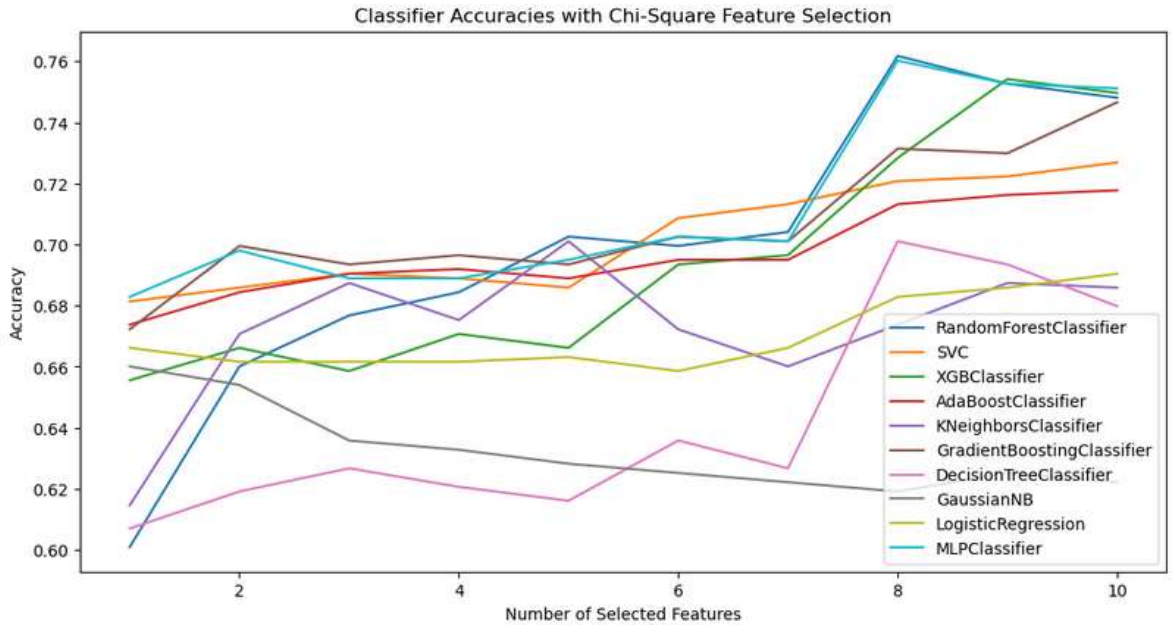

Figure 13 : Classifier Accuracies with Chi-Square Feature Selection

In Figure 14, when looking at the precision values, it can be observed that the precision value generally increases as the number of features increases. Many classifiers have shown peaks at different values. Precision achieved its highest value with a feature count of 9 in the MLP classifier.

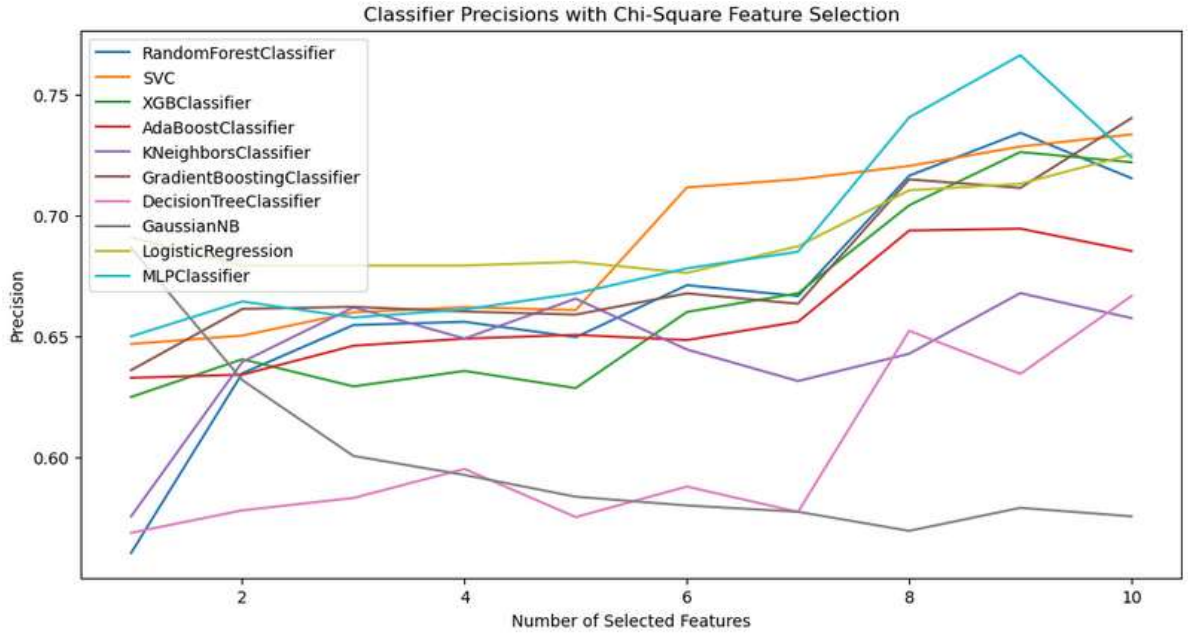

Figure 14: Classifier Precisions with Chi-Square Feature Selection

In Figure 15, the accuracy values of different classifiers using the mutual information method have been observed. When looking at the graph, the difference between classifiers is less compared to previous methods. The highest accuracy value was achieved in the Random Forest classifier.

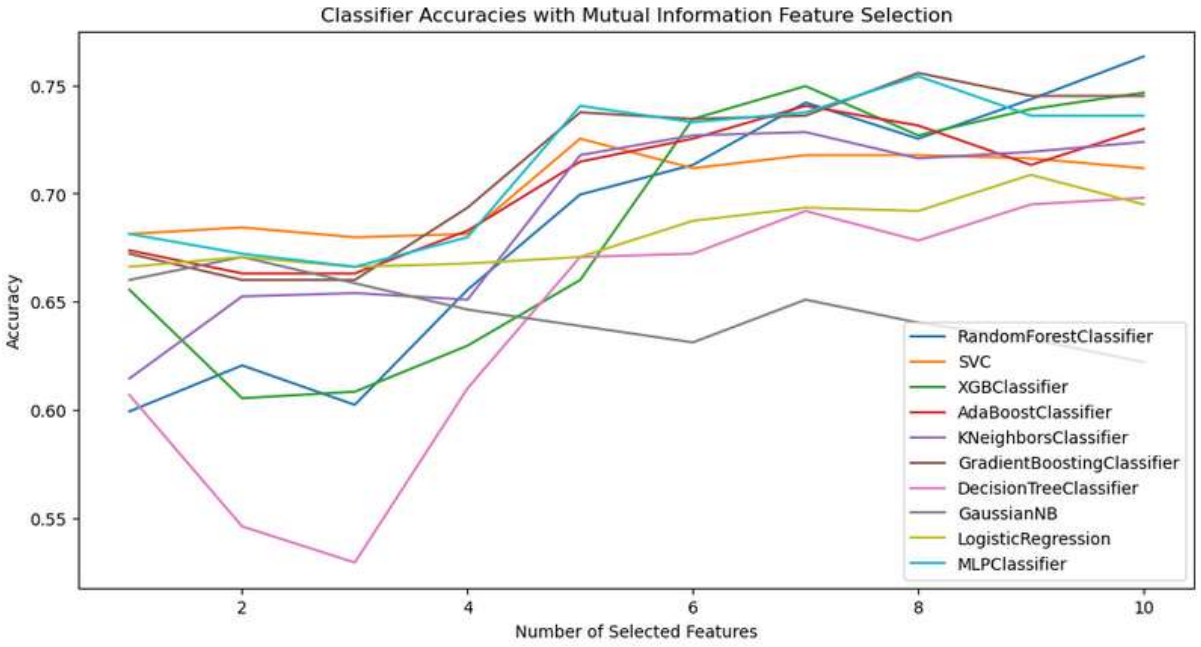

Figure 15:Classifier Accuracies with Mutual Information Feature Selection

The graph of Precision values using the mutual information method is shown in Figure 16. The highest value, exceeding 0.75, was achieved in the Random Forest classifier.

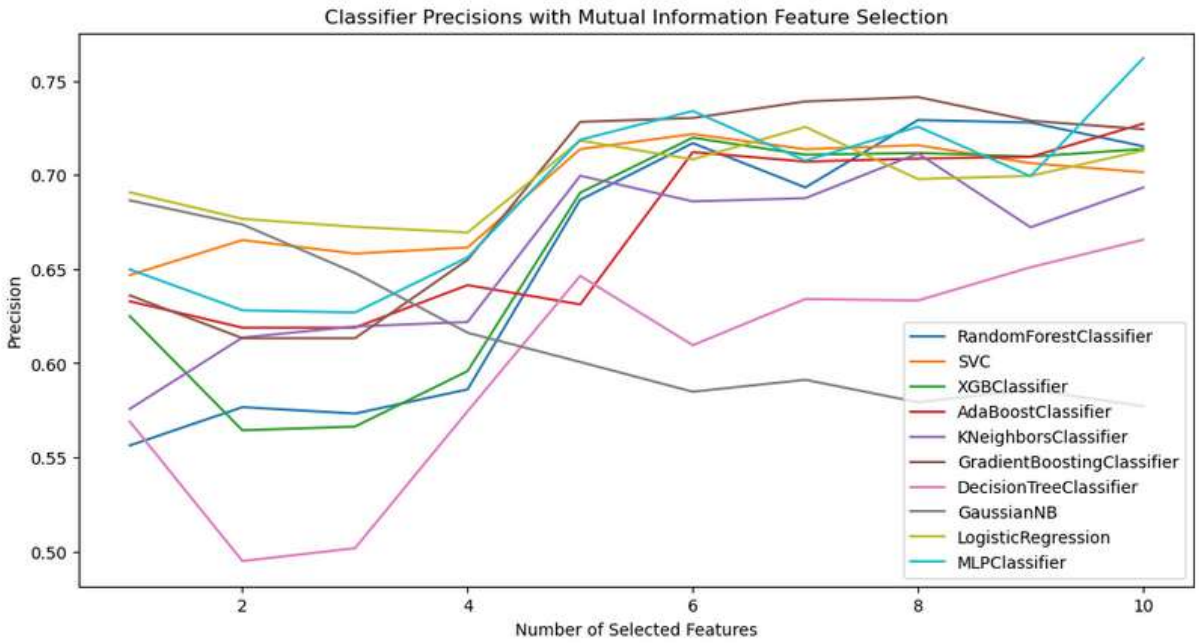

Figure 16:Classifier Precisions with Mutual Information Feature Selection

The accuracy values calculated using the L1-based feature selection method can be observed in Figure 17. After the number of features exceeds 4, the linear increase diminishes and reaches a plateau. The highest accuracy value was achieved when using the Random Forest classifier.

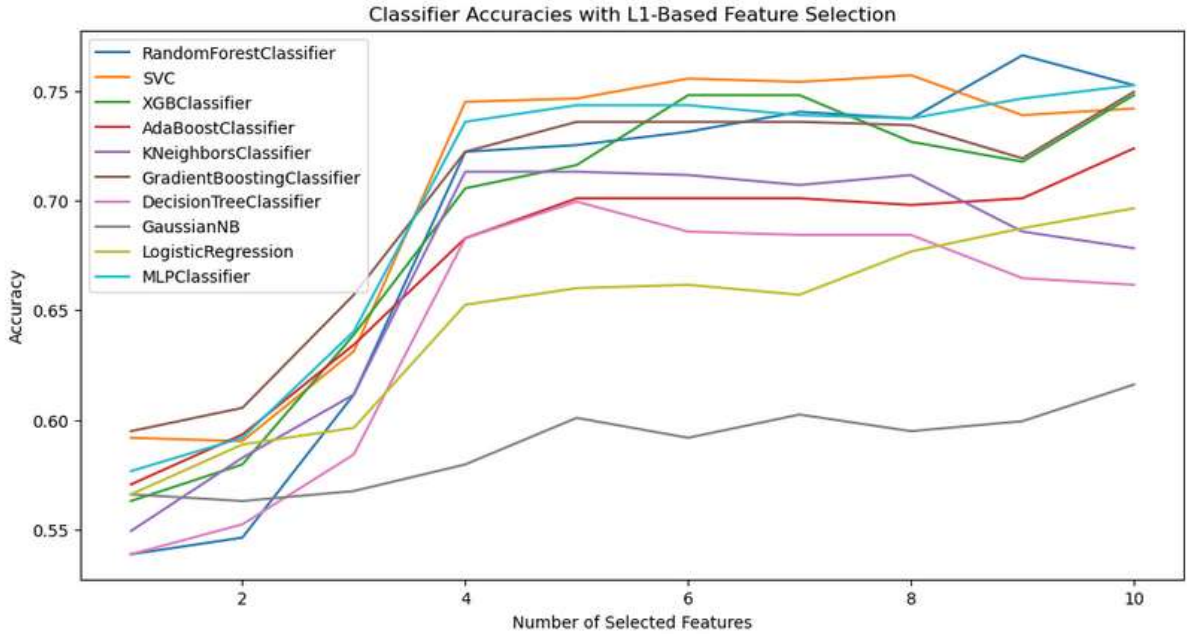

Figure 17: Classifier Accuracies with L1-Based Feature Selection

Similar observations were made in Figure 18, where Precision values were calculated using the L1-based feature selection method. A similar pattern was observed, where Precision reaches its highest value when using the Random Forest classifier.

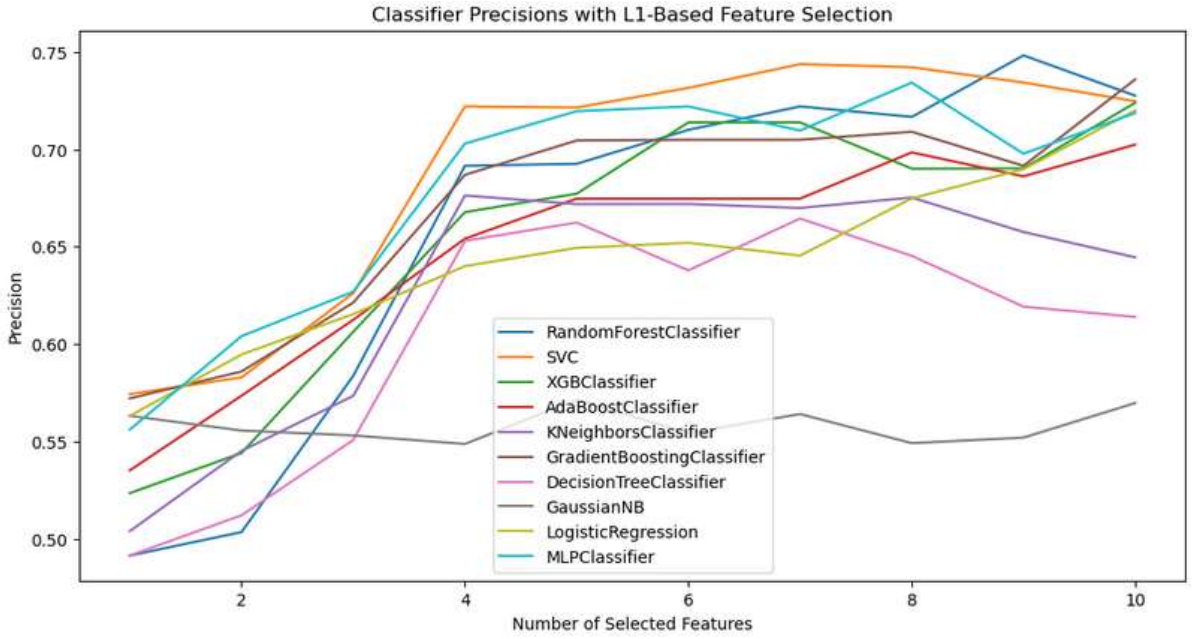

Figure 18:Classifier Precisions with L1-Based Feature Selection

Upon examining the graphs, it is quite natural to observe that each classifier exhibits peaks at different feature counts. The chosen selection methods also have an impact on the performance of the classifiers used. When analysing the graphs, it can be observed that both precision and accuracy values generally increase as the feature count increases up to 4. However, upon surpassing 4 features, the linear increase diminishes, and the values tend to stabilize. Beyond 10 features, there is a general trend of decreasing values. It is crucial to train the machine learning algorithm specifically on selected features that are distinct and informative in order to achieve accurate results.

8. RESULTS AND DISCUSSION

The findings of this study demonstrate that the proposed approach, combining image processing techniques, feature extraction, and artificial intelligence algorithms, can effectively distinguish between benign and malignant skin cancer with an accuracy rate of approximately 85%. This represents a significant improvement compared to traditional diagnostic methods and offers a rapid and cost-effective alternative to invasive procedures like biopsy. It should be noted that while the achieved accuracy rates are promising, there is still room for improvement. The performance can be further enhanced by incorporating more data, advanced algorithms, and enhanced image processing techniques. Additionally, fine-tuning the parameters of the classifiers and attribute selectors may lead to better results.

Furthermore, it is important to acknowledge the limitations and potential risks associated with artificial intelligence in medical diagnosis. AI models are only as reliable as the data they are trained on, and errors or biases in the training dataset can lead to inaccurate results. Algorithmic biases can also occur, potentially resulting in misdiagnosis and patient harm. Therefore, it is essential to validate and verify the performance of the AI models using diverse and representative datasets and to exercise caution in relying solely on automated diagnostic methods.

While there is currently no public application available for the proposed techniques, they can be easily implemented in a website or mobile application to provide accessible and user-friendly diagnostic tools for healthcare professionals. This can greatly benefit hospitals, doctors, and institutions that lack access to pathology laboratories or have limited resources.

9. CONCLUSION

This study presents a practical approach that enhances the diagnostic process of skin cancer by leveraging image processing and artificial intelligence techniques. By utilizing a dataset obtained from Kaggle, image preprocessing, segmentation, feature extraction, feature selection, classification and validation were performed to distinguish between benign and malignant skin lesions. The results demonstrate that the proposed approach can achieve an accuracy rate of approximately 85% in distinguishing between different types of skin cancer. This offers a rapid and cost-effective diagnostic solution, reducing the need for invasive procedures and minimizing patient discomfort.

While the achieved accuracy rates are promising, further improvements can be made by incorporating more data, advanced algorithms, and enhanced image processing techniques. Additionally, the limitations and potential risks associated with artificial intelligence in medical diagnosis should be considered and addressed.

The primary beneficiaries of this research are hospitals, clinics, and healthcare providers with limited access to pathology laboratories or financial resources. The proposed approach aims to alleviate the workload of healthcare professionals and contribute.

References:

[1] Ahmed, H. M., Al-azawi, R. J., & Abdulhameed, A. A. (2018). Evaluation methodology between globalization and localization features approaches for skin cancer lesions classification. Journal of Physics: Conference Series, 1003, 012029. https://doi.org/10.1088/1742-6596/1003/1/012029

[2] Akyel, C., & Arıcı, N. (2022). LinkNet-B7: Noise Removal and lesion segmentation in images of skin cancer. Mathematics, 10(5), 736. https://doi.org/10.3390/math10050736

[3] Barata, C., Celebi, M. E., & Marques, J. S. (2019). A survey of feature extraction in dermoscopy image analysis of skin cancer. IEEE Journal of Biomedical and Health Informatics, 23(3), 1096–1109. https://doi.org/10.1109/jbhi.2018.2845939

[4] Chan, T.F., & Vese, L.A. (2001). An Active Contour Model Without Edges. In IEEE Transactions on Image Processing, 10(2), 266-277. DOI: 10.1109/83.902291

[5] Javed, R., Rahim, M. S., Saba, T., & Rehman, A. (2019). A comparative study of features selection for skin lesion detection from dermoscopic images. Network Modeling Analysis in Health Informatics and Bioinformatics, 9(1). https://doi.org/10.1007/s13721-019-0209-1

[6] Kaggle. (Year). Skin Cancer: Malignant vs. Benign [Data set]. Retrieved from [Skin Cancer: Malignant vs. Benign | Kaggle]

[7] Mete, M., & Sirakov, N. M. (2014). Optimal set of features for accurate skin cancer diagnosis. 2014 IEEE International Conference on Image Processing (ICIP). https://doi.org/10.1109/icip.2014.7025457

[8] Thanh, D. N., Prasath, V. B., Hieu, L. M., & Hien, N. N. (2019). Melanoma skin cancer detection method based on adaptive principal curvature, colour normalisation and feature extraction with the ABCD Rule. Journal of Digital Imaging, 33(3), 574–585. https://doi.org/10.1007/s10278-019-00316-x