Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 7, No.1, January 2025, pp. 0250116-01 ~ 0250116-10
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v7i1.1154

Advancing Dermatological Image Classification: GLCM-Based Machine Learning Insights

Rania R. kadhim[1]*, Mohammed Y. Kamil[1]

[1]College of Science, Mustansiriyah University, 9C93+623، مجسر المستنصرية, Baghdad, Baghdad Governorate, Iraq

*rania.r90@uomustansiriyah.edu.iq

Abstract. The prospects to improve skin illness via the utilization of artificial intelligence algorithms is what renders this study economically important. Machine learning may assist physicians detect people quicker and more accurately. The effective identification of skin disorders using machine learning could result in the development of large and readily available digital tests. A model was used in the present study to analyze the HAM 10000 data. Two hundred images in total were chosen at random; one hundred showed dermatofibroma diseases, whereas the other hundred displayed benign keratosis. Subsequently, these images were resized to prepare for additional examination. The statistical features of the gray level co-occurrence matrix were calculated from the image dataset by changing the distances 0, 5, 10, 15 and angles 0°, 45°, 90°, 135°. Five different machine learning models were subsequently trained and assessed based on these features. The study shows that the logistic regression model accurately detects and classifies various skin diseases. The logistic regression model showed exceptional performance, exceeding the expected results in terms of accuracy 91.50%, sensitivity 93.00%, and F1-score 91.36. The results of the study were most favorable when using an angle measurement of 135°.

Keywords: Skin cancer, AI in healthcare, Classification, HAM dataset, sustainable technology.

(Received 2024-10-25, Accepted 2024-12-31, Available Online by 2025-01-10)

1. Introduction

Sustainable technological advancements in healthcare aim to improve reduce costs and efficiency via innovations, like Artificial Intelligence (AI). These technologies provide scalable solutions, more accurate making healthcare and accessible in underserved areas. By automating diagnostics, reduce pressure improve patient outcomes on healthcare systems, promoting long-term sustainability. The sustainability of healthcare infrastructures is supported by this move toward growing robotic, AI-driven medical systems[1]. Skin diseases include a diverse range of health issues that impact the internal organs, the human body possesses an organ of considerable size that serves the crucial functions of protecting internal organs and creating a protective barrier between the body and the outside world [2]. These diseases include a diverse range of problems, from small and transient afflictions to severe and persistent diseases of the skin, one such class of illnesses includes benign keratoses and dermatofibroma [3, 4]. Benign keratoses, commonly referred to as seborrheic keratoses, are prevalent skin proliferations typically exhibiting shades of brown, light or black tan. Frequently confused as moles, warts, or skin cancer, these growths are indeed benign. Benign keratoses generally exhibit features such as a waxy texture, scaly surface, or a little elevation [5, 6]. The clinical images exhibit certain limitations, including low resolution and the presence of various defects such as skin lines, darkness duplications, and hair, because of these challenges, the study of skin diseases is extremely challenging. AI plays an essential role in the medical field, machine learning (ML) models were utilized to assist with detection [7, 8]. AI could offer rapid and precise disease diagnoses based on complete clinical analysis of data this helps prevent delays in diagnosis and ensures rapid therapy initiation. Radiology is an essential part of medical diagnosis, AI has the potential to enhance the capabilities of radiologists by allowing faster and more accurate interpretation and analysis of medical images, hence leading to enhanced accuracy in diagnosis. AI monitor the health of patients, to assure treatment response, it can evaluate chronic disease data and monitor medical image changes [9, 10]. Additionally, it enhances access, which ultimately results in an improvement in the general level of medical care. The main goal of this research effort is to improve the diagnostic accuracy of skin diseases, with a particular focus on correctly distinguishing between benign keratoses and dermatofibroma. In this study, GLCM method has been applied to the HAM 10000 dataset for the purpose to perform image classification. Five various models for classification are employed in order get the most effective results. Despite advances in AI for healthcare diagnosis, difficulties like low-resolution images and trouble identifying identical disorders render it challenging to correctly classify diseases of the skin such benign keratoses and dermatofibroma. The following parts of this study are structured in the following style: Section 2, discusses the suggested method that was adopted in this study. Section 3 displays the used datasets. Finally, Section 4 discusses the conclusions derived from the research.

2. Methods

The present section provides a description and analysis of the applied methodologies for the classification of skin diseases .The data processing involves three main steps: image resizing to standardize input, feature extraction using GLCM to extract texture features like homogeneity and contrast, classification by ML models. GLCM was chosen because it effectively highlights important texture patterns in skin images, which are crucial for accurate disease classification. This approach improves diagnostic accuracy sustainably by automating the feature extraction and classification processes, enabling scalability, reducing human error, and ensuring continuous improvements as more data is processed [11]. The entire procedure includes the following parts: The initial phase involves image resizing, then feature extraction to extract statistical features, and at last classification, Figure 1 illustrates the flowchart of work steps.

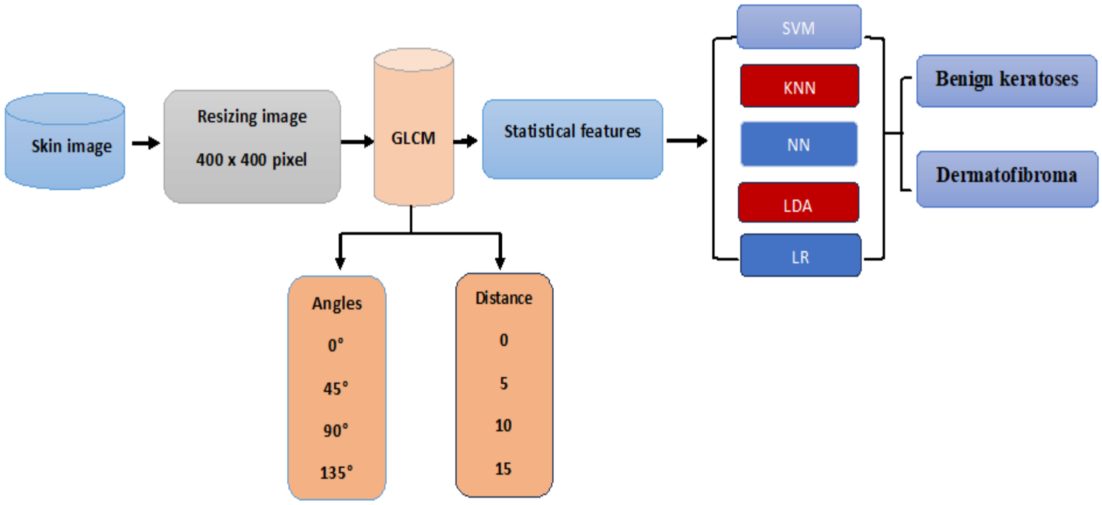


Figure 1: Flowchart of the proposed method

2.1 Image resizing

The number of features obtained from images of varying sizes is not constant. Hence, the images supplied change in size, either increasing or decreasing, to address this problem. Additionally, it reduces the time required for processing and enhances the overall efficiency of the system. In this study, the input images were uniformly resized to a scale of 400 × 400 pixels [12].

2.2 Feature Extraction

The process of feature extraction provides necessary work in the analysis and exploration of the fundamental connections among different entities. The models for image classification, prediction, and recommendation fail to directly understand the images. Therefore, feature extraction is required to transform into the forms that can be applied. The thermoscopic image could be described using a variety of features. Fortunately, not every attribute can be used to classify skin diseases. As a result, the classifier gets more complicated and requires more processing work for a number of unimportant characteristics, limiting the classification accuracy [13]. In images of skin cancer, the most effective features have to represent the region's features. Because of this, many features need to be extracted to tell images to separate as correctly as possible. In this study, we implemented the GLCM to identify the type of skin disease by extracting several statistical features [14]. In the field of research, feature extraction is of the greatest importance; The GLCM processing in an image depends on the distance and angle. The measurement of distance is employed to define the degree of distance between the pixels under consideration, whereas the angle parameter specifies the relative orientation of such pixels [15]. To be more precise, the spacing between pixels affects the degree to which surrounding pixels influence the calculations. Thus, we applied a set of four various distance values (1, 5, 10, and 15) together with four different angle values (0^o, 45^o, 90^o, and 135^o). GLCM can be used to get various types of data regarding the inside of an image by determining different distances and angles, enabling a more precise and comprehensive examination of textural features and co-occurrences across various grey levels. The most effective features of images of skin cancer ought to accurately represent the distinctive features of the areas affected. As a result, for this to, it is necessary to extract sufficient features. The GLCM includes a wide variety of textural elements which are provided from GLCM [16].

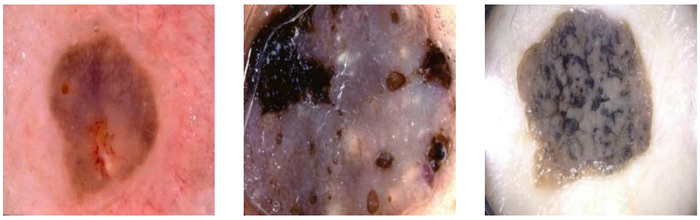
2.3. Classification

The last step of the process is classification. It involves classifying a set of data into two distinct classes. benign keratoses and dermatofibroma. This study detected a type of skin disease by employing extracted image features. Different classification models are utilized according to the application and

data set type. In the current study, we classified skin diseases using five different models, which are outlined below:
Support vector machine (SVM), linear discriminant analysis (LDA), K nearest neighbour (KNN), logistic regression (LR) and neural network (NN) [17, 18].

2.4 Dataset

The dataset utilized in this study has been obtained and made available by the HAM10000 project, which focuses on the comparison between the performance of humans and machines using a collection of 10,000 training images. The HAM 10000 dataset includes different features, showcasing diversity in this field, and helping researchers to examine and identify various medical diseases, such as warts, melanoma, common nevi, and other medical conditions [19]. The HAM 10000 dataset has a significant quantity of dermatological images with an average total of tens of thousands of images, this dataset proves to be extremely effective for helping with the training of AI models [20]. The use of this dataset helps the training of ML models to successfully recognize and classify different diseases, this instrument can assist physicians in enhancing accurate diagnosis and improving care for patients [21]. There were 200 images selected randomly, including 100 images from the category of benign keratoses and 100 images from the category of dermatofibroma. Figure 2 shows a dataset-selected skin disease sample. The dataset can be downloaded by using the given link :https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000



Benign keratoses

Dermatofibroma

Figure 2: Two classes of skin disease in the HAM10000 dataset (a) Benign keratoses and (b)Dermatofibroma [22]

2.5 Performance analysis

The proposed methodology involves the calculation of six metrics for efficiency, including Sensitivity, Specificity, Accuracy, Precision, and F1-score, these evaluation metrics are calculated using the following formulas [10, 23]

$$Accuracy=(TN+TP)/(TN+TP+FP+FN) \quad (1)$$

$$Sensitivity= TP/ (TP+FN) \quad (2)$$

$$Specificity=TP/(TN+TP) \quad (3)$$

$$Precision= TP/(TP+FP) \quad (4)$$

$$F1\text{-}Score=2TP/(2TP+FP+FN) \quad (5)$$

3. Results and Discussion

All tasks have been successfully implemented using the MATLAB platform. The model was applied to the training dataset using the HAM 10000 dataset. 200 images were randomly chosen, composed of 100 images of benign keratoses disease and 100 images of dermatofibroma disease. After resizing the images, we selected four distinct angle values (0º,45º,90º,135º) and four distance values (0,5,10,15) for the GLCM statistical features were obtained by employing the GLCM method, and then employed the statistical features for the five ML models with the greatest performance. This study employed a collection of skin disease images, which concentrated on two types of skin diseases: benign keratoses and dermatofibroma. The GLCM method was utilized to extract statistical features from these images. Various numbers were selected for the angles and distance. The angle values of 0°, 45°, 90°, and 135°, and also the distance values of 1, 5, 10, and 15, were selected to examine their impact on the performance of the five ML models. Which, when it came to classifying diseases of the skin, produced the most effective results. The performance of the models was examined using six criteria, namely accuracy, sensitivity, specificity, precision, F-1 score, and AUC. Figure 3 illustrates the accuracy score for the five models at angle values of 0°, 45°, 90°, and 135°, at distance values 1, 5, 10 and 15, it was observed that the LR model achieved its highest accuracy of 91.50% when the distance value was specified as 1 and the angle value was equal to 135°.

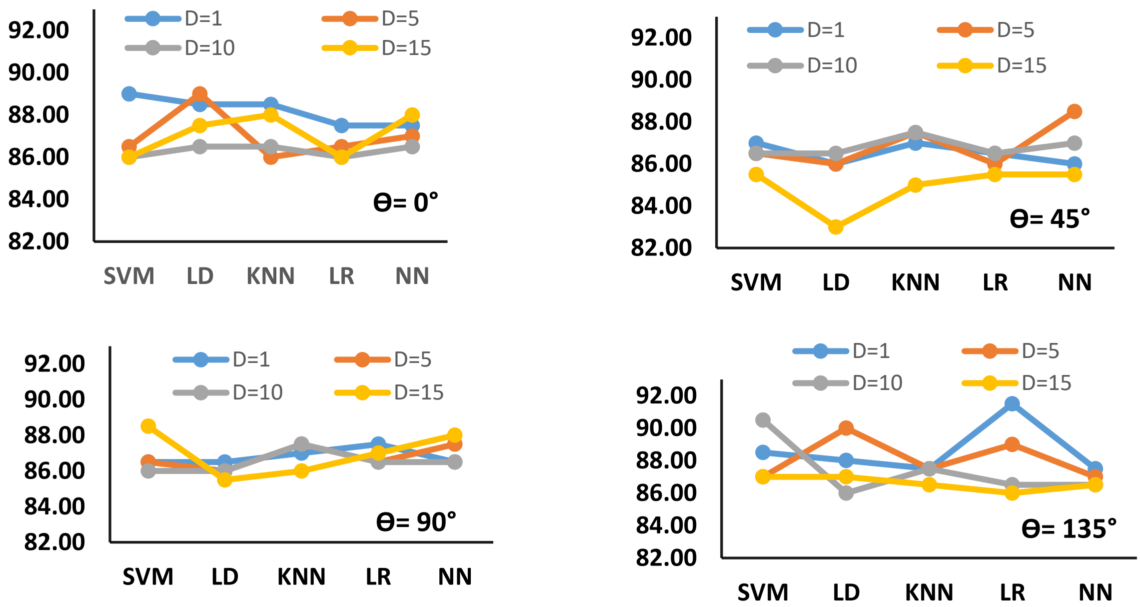


Figure 3. Accuracy of various models

Figure 4 displays the five models' sensitivity scores for angles of 0°, 45°, 90°, and 135° with distances of 1, 5, 10, and 15. When the distance value was equal to 1 and the angle value was equal to 135°, the LR classifier's sensitivity got its highest value of 93.00%.

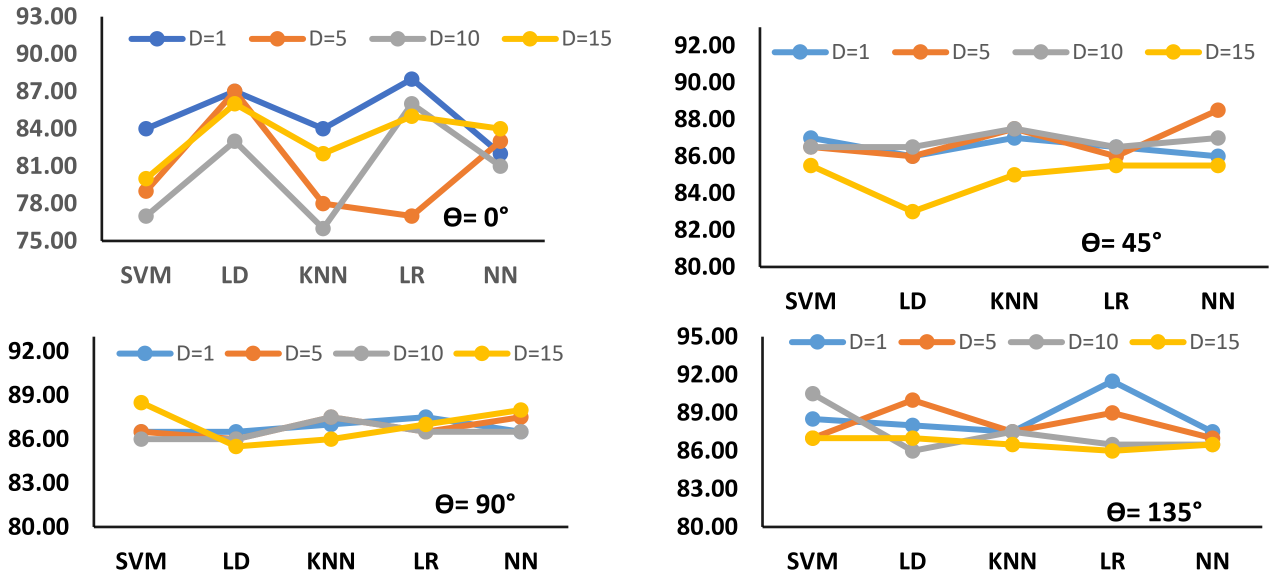


Figure 4. Sensitivity of various models

The KNN classifier achieved its highest degree of specificity equal to 97.00% when the distance value was set to 10 and the angle was set to 0° as shown in Figure 5.

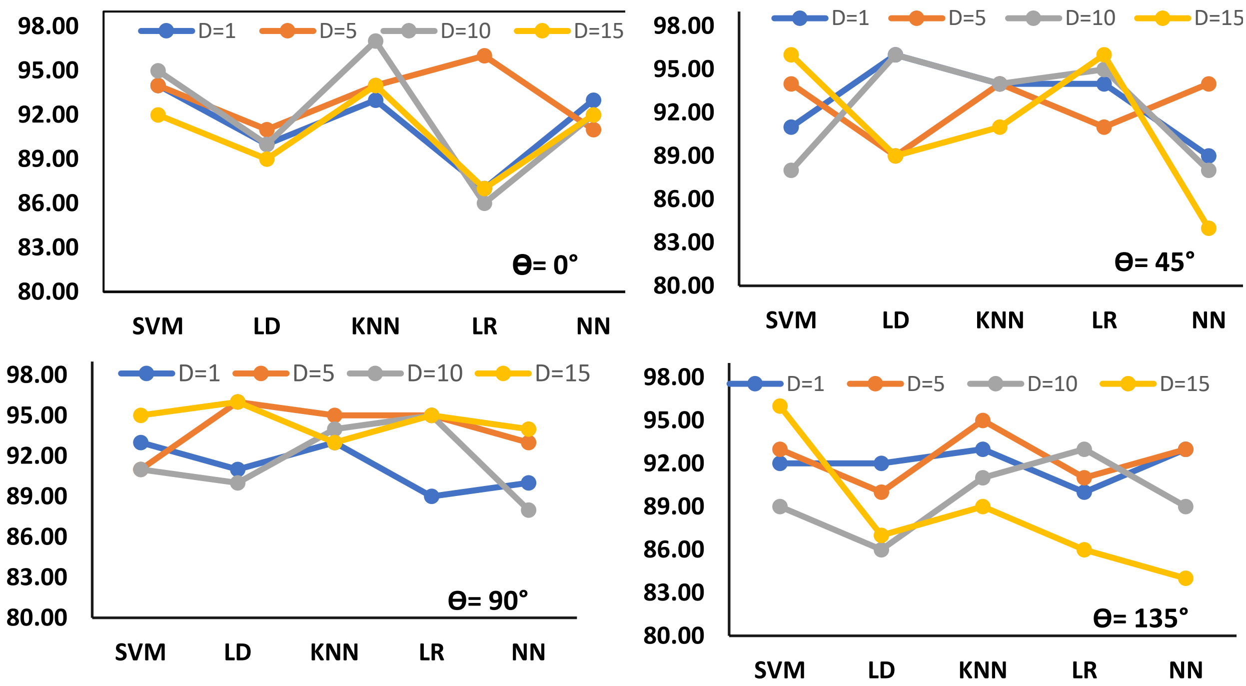


Figure 5. Specificity of Various Models

With respect to the precision criteria, the KNN method had the highest possible effectiveness, with a score of 96.20%. This result occurred when the distance value was set to 10 and the angle value was set to 0° as illustrated in Figure 6.

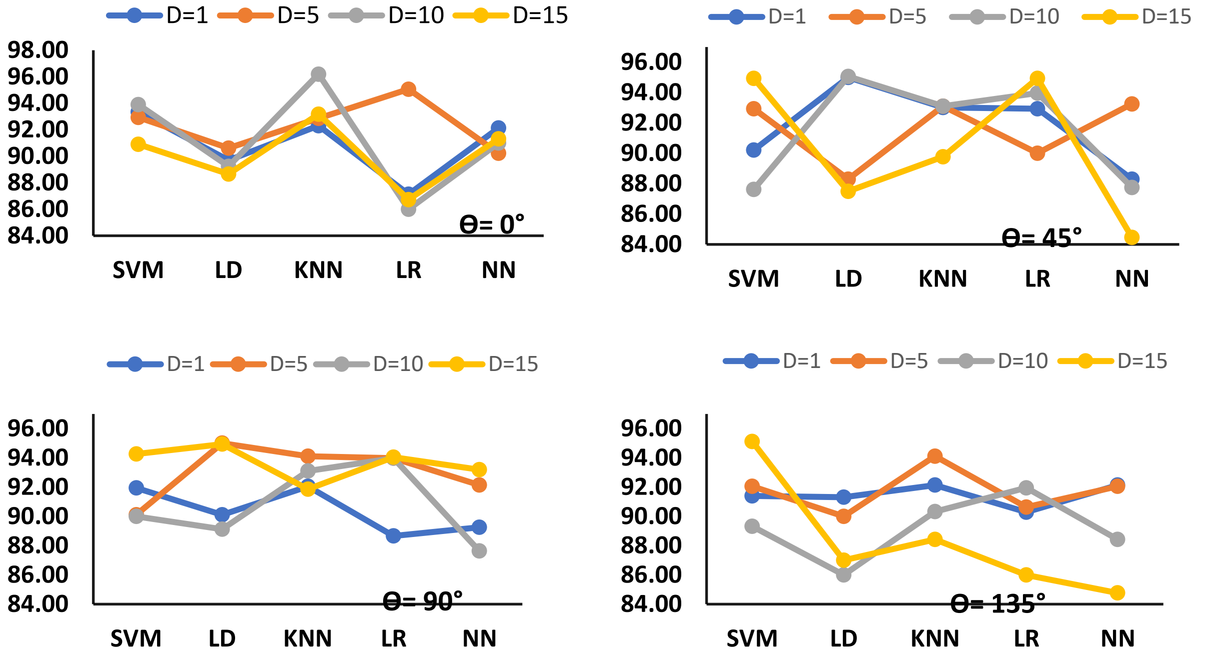


Figure 6. Precision Of Various Models

The F1-Score measurement produced the greatest results for the LR model, achieving a value of 91.36%. This result was achieved when the distance value was set to 1 and the angle element was equal to 135° according to Figure 7.

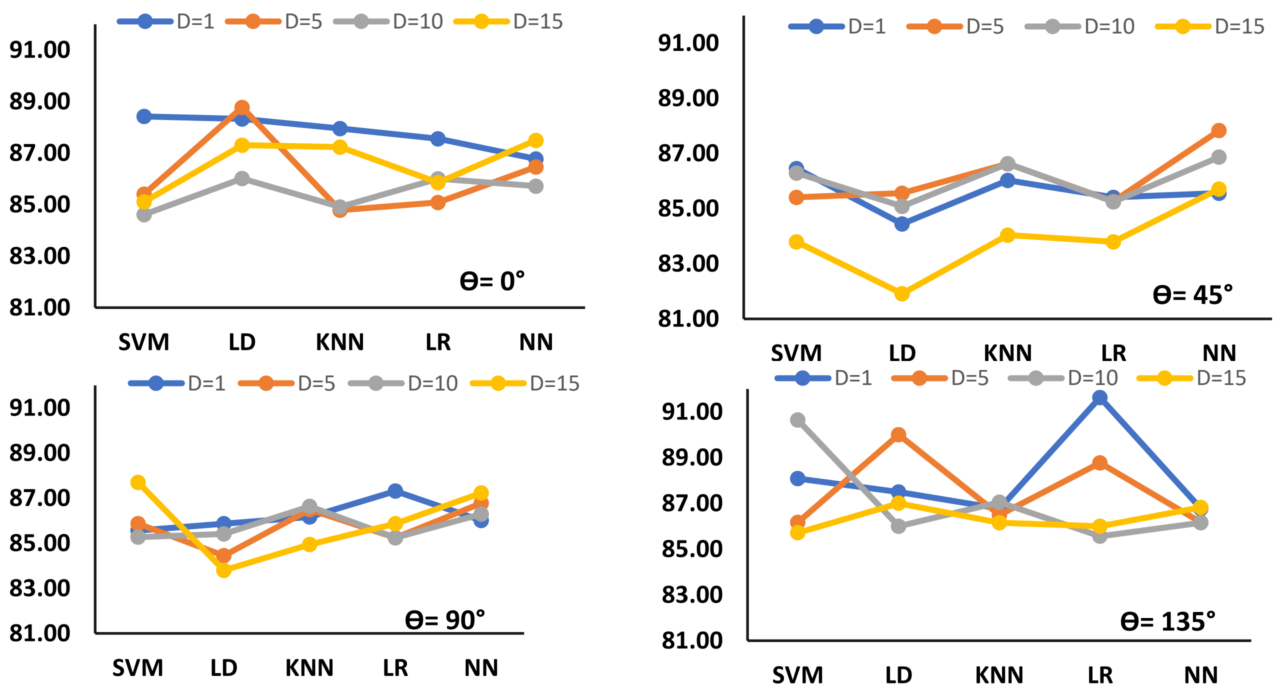


Figure 7. F1-Score of Various Models

It is seen from Figure 8 that when the distance was set to 1 and the angle was set to 135°, the AUC standard for the LD classifier obtained the highest value which was equal to 0.95.

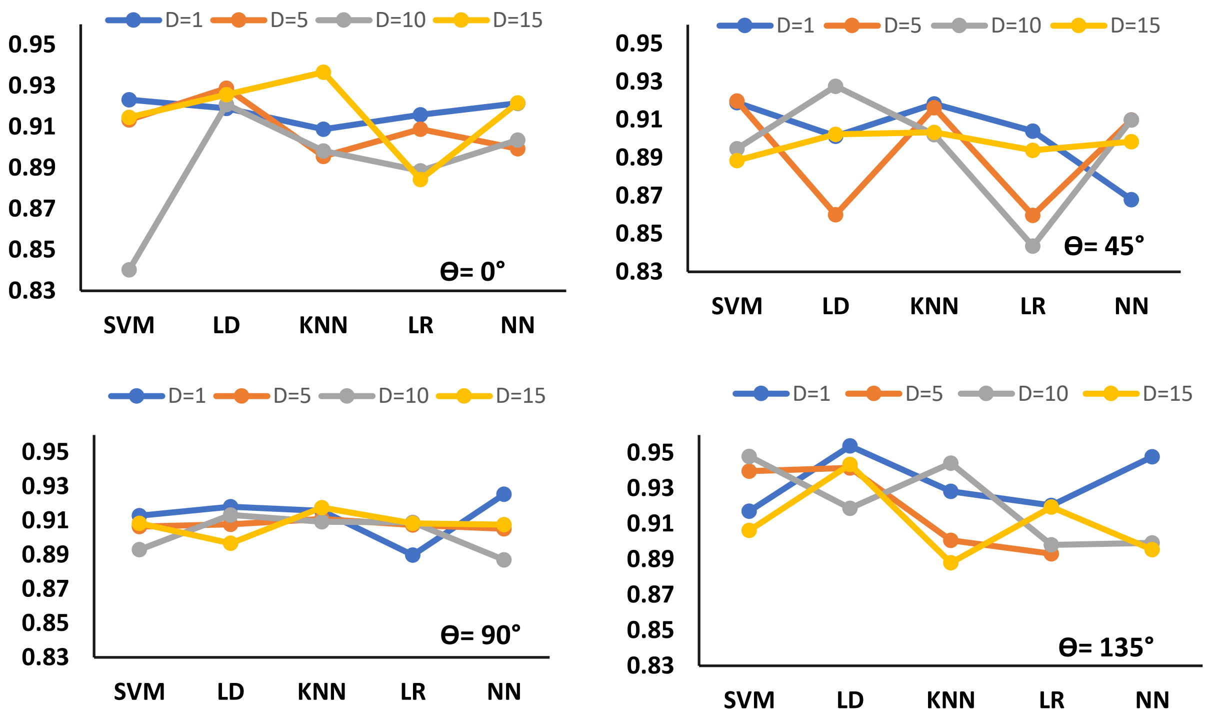


Figure 8. AUC of various models

The research results of this study enhance the argument by showing the relevance of ML models in sustainable medical devices in addition to establishing their accuracy in classifying skin disorders. AI lessens the strain on health care systems by automating the diagnostic process, allowing for quicker and more affordable assessment, especially in areas with limited resources. This is consistent with the expanding demand for effective, promising longevity, scalable healthcare solutions and better patient results worldwide.

4. Conclusions

Based on the results and evaluation obtained from the present study, utilizing a database including dermatological images that specifically target two distinct skin diseases, namely benign keratosis and dermatofibroma, the results show that using GLCM analysis for the extraction of statistical information from skin images can be an effective and efficient strategy. The inclusion of a wider range of values for distances and angles serves to improve the system's ability to accurately represent the structural characteristics of diseases. A range of values were chosen for angles (0°, 45°, 90°, 135°) and distances (1, 5, 10, 15) to evaluate the impact of this variety on the effectiveness of ML models. The results obtained show that such variety is of major significance for improving the efficacy of ML strategies. The evaluation of model performance was done using six primary criteria, including accuracy, sensitivity, specificity, precision, AUC, and F1- score. These criteria offer a thorough perspective on the effectiveness of the models. LR model is famous for its exceptional performance, exhibiting outstanding outcomes in terms of accuracy, sensitivity, and F1- score. This finding illustrates the effectiveness of the model in precisely identifying and classifying skin diseases. The results suggest that distance 1 demonstrated greater effectiveness in the majority of instances, emphasizing the significance of meticulously identifying the optimal distance for improving model performance. The angle value of 135°.The angle measure of 135° yielded the most optimal results.

The results of this study offer chances for further examinations, including the improvement of analytical techniques and expanding understanding of the impact of different factors on the effectiveness of models. Integrating this AI-driven technology into real-world diagnostic tools can enhance healthcare delivery by providing faster, more accurate diagnoses, especially in underserved areas. It improves accessibility, and streamlines the diagnostic process, reduces reliance on specialists, this contributes to the sustainability of healthcare systems by optimizing resources and lowering cost while continuously improving accuracy.

Acknowledgments

With great appreciation, I would like to thank Mustasiriyah University for all of its assistance and assets during my research journey.

References

[1] Salsabila, A.S., C.A. Sari, and E.H. Rachmawanto, *Classification of Movie Recommendation on Netflix Using Random Forest Algorithm.* Advance Sustainable Science, Engineering and Technology, 2024. **6**(3).

[2] Edwar, D.A., I.N. Naji, and H.M. Aboul-Ela, *Investigation the Role of Various Antiseptics on the Prevalence of Skin Microbiota and Post Cesarean Surgery Infections.* Al-Mustansiriyah Journal of Science, 2023. **34**(3): p. 1-9.

[3] Hosny, K.M., M.A. Kassem, and M.M. Fouad, *Classification of Skin Lesions into Seven Classes Using Transfer Learning with AlexNet.* Journal of Digital Imaging, 2020. **33**(5): p. 1325-1334.

[4] Keerthana, D., et al., *Hybrid convolutional neural networks with SVM classifier for classification of skin cancer.* Biomedical Engineering Advances, 2023. **5**: p. 100069.

[5] Mobiny, A., A. Singh, and H. Van Nguyen, *Risk-aware machine learning classifier for skin lesion diagnosis.* Journal of Clinical Medicine, 2019. **8**(8).

[6] Hasan, M.K., et al., *A survey, review, and future trends of skin lesion segmentation and classification.* Computers in Biology and Medicine, 2023: p. 106624.

[7] Tschandl, P., et al., *Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study.* The Lancet Oncology, 2019. **20**(7): p. 938-947.

[8] Shames, M.A. and M.Y. Kamil, *Lung Infection Detection via CT Images and Transfer Learning Techniques in Deep Learning.* Journal of Advanced Research in Applied Sciences and Engineering Technology, 2025. **47**(1): p. 206-218.

[9] Ahammed, M., M.A. Mamun, and M.S. Uddin, *A machine learning approach for skin disease detection and classification using image segmentation.* Healthcare Analytics, 2022. **2**.

[10] Kadhim, R.R. and M.Y. Kamil, *Breast invasive ductal carcinoma diagnosis using machine learning models and Gabor filter method of histology images.* International Journal of Reconfigurable and Embedded Systems, 2023. **12**(1): p. 9-18.

[11] Alam, L. and A.Z. Fanani, *Implementation of the Adaboost Method to Increase the Accuracy of Early Diabetes Predictions to Prevent Death Decision Tree-Based.* Advance Sustainable Science, Engineering and Technology, 2024. **6**(2).

[12] Qiu, X., et al., *Raman spectroscopy combined with deep learning for rapid detection of melanoma at the single cell level.* Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2023. **286**: p. 122029.

[13] Omeroglu, A.N., et al., *A novel soft attention-based multi-modal deep learning framework for multi-label skin lesion classification.* Engineering Applications of Artificial Intelligence, 2023. **120**: p. 105897.

[14] Al-Khuzaay, H.M., et al., *Evaluation of Effect of β-Glucan on Cancer Cell Lines In vitro.* Al-Mustansiriyah Journal of Science, 2024. **35**(1): p. 17-20.

[15] Choudhary, P., J. Singhai, and J.S. Yadav, *Skin lesion detection based on deep neural networks.* Chemometrics and Intelligent Laboratory Systems, 2022. **230**: p. 104659.

[16] Kamil, M.Y. and A.L.A. Jassam. *Analysis of Tissue Abnormality in Mammography Images Using Gray Level Co-occurrence Matrix Method*. in *Journal of Physics: Conference Series*. 2020.

[17] Hashem, S.A. and M.Y. Kamil, *Segmentation of Chest X-Ray Images Using U-Net Model.* Mendel, 2022. **28**(2): p. 49-53.

[18] Obaid, A.S., M.Y. Kamil, and B.H. Hamza, *People Recognition via Tongue Print Using Deep and Machine Learning.* Journal of Artificial Intelligence and Technology, 2023. **3**(3): p. 119-125.

[19] Said, R.A., et al. *Skin Cancer Detection and Classification Based on Deep Learning*. in *International Conference on Cyber Resilience, ICCR 2022*. 2022.

[20] Mohammed, N.B., et al., *Quantitative Analysis of Blurry Color Image Fusion Techniques using Color Transform.* Al-Mustansiriyah Journal of Science, 2023. **34**(3): p. 132-140.

[21] Thabit, Z.H., S.A. Mehdi, and B.M. Nema, *Enhancing Color Image Security: Encryption with Dynamic Chaotic Three-Dimensional System and Robust Security Analysis.* Al-Mustansiriyah Journal of Science, 2023. **34**(4): p. 87-95.

[22] Nakai, K., Y.-W. Chen, and X.-H. Han, *Enhanced deep bottleneck transformer model for skin lesion classification.* Biomedical Signal Processing and Control, 2022. **78**: p. 103997.

[23] Shames, M.A. and M.Y. Kamil, *Early Diagnosis of Lung Infection via Deep Learning Approach.* International Research Journal of Multidisciplinary Technovation, 2024. **6**(3): p. 216-224.