International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V70I2P202

Volume 70 Issue 2, 11-18, February, 2022

Original Article

Machine Learning and Deep Learning Integration for Skin Diseases Prediction

Samir Kumar Bandyopadhyay[1], Payal Bose[2], Amiya Bhaumik[3], Sandeep Poddar[4]

[1]*Lincoln University College, Petaling Jaya, Selangor D.E, Malaysia*

[1]*1954samir@gmail.com*, [2]*bosepayal91@gmail.com*, [3]*amiya@lincoln.edu.my*

Abstract *- Living creature skin disease is a fairly prevalent ailment. In the medical world, monitoring dermatological disorders and classifying them is a complex process. Due to the sheer intricacy of individual skin tone and the visible proximity effect of infections, recognizing the precise type can be challenging at times. As a result, it is critical to diagnose and recognize skin disease as soon as possible. Artificial intelligence (AI) is quickly expanding in therapeutic areas in a modern context. For diagnostic purposes, much deep learning (DL) and machine learning (ML) methods are applied. These strategies drastically enhance the diagnosing process while also speeding it up. In this study, to improve disease detection, a model combining deep learning (DL) and machine learning (ML) has been developed. For classification, three sets of machine learning models were utilized, and for feature selection, four sets of pre-trained deep learning models were being used. For classification models, deep neural networks Alexnet, Googlenet, Resnet50, and VGG16 were used, while Support Vector Machine, Decision tree, and Ensemble boosting Adaboost classifier were applied for classification. To identify the best prediction model, a comparative study was carried out. The hybrid method Resnet50 with SVM produced the best results, with 99.11% accuracy.*

Keywords — *Deep Feature Extraction, Deep Learning (DL), Machine Learning (ML), Skin Diseases Detection, Support Vector Machine (SVM).*

I. INTRODUCTION

The epidermis is the most vital bodily component of the human body occupies an area of 21 to 23 square feet. It protects several critical organs of the human body from exterior harm as well as pathogens and also the weather, aids in thermoregulation, and enables sensation, warmth, and cold feelings. However, the skin can be altered by a range of extrinsic and genetic variables. In the human population, skin illnesses are the most frequent. Three forms of skin problems mostly impact human skin. 1) infectious skin disorder, 2) bacterial skin disorder, and 3) contact dermatitis disorder. The consistency or colour of the skin may alter as a result of a skin illness. Skin illnesses are persistent, contagious, and can lead to skin cancer in some cases. If fungal and allergy disorders are adequately diagnosed and recognized in their early stages, they can be easily treated. In the occurrence of viral diseases, however, immediate identification is essential.

AI models, ML [1,2], and DL techniques [3,4] have advanced quickly in the medical area during the last few years. Numerous ensemble-based [5] machine learning approaches and artificial neural network-based approaches [6] are also widely used to improve skin disease predictions. Computer vision, in addition to physical indications, is critical for detecting many skin diseases. The computer vision approach aids in the detection of skin diseases with greater precision.

To diagnose skin abnormalities, a fusion of ML and DL models is built in this study. For classification models, three sets of ML models were used: Decision Tree, SVM, and Adaboost Ensemble Classifier. For feature selection, four sets of pre-trained deep neural networks were used: Alexnet, Googlenet, Resnet50, and VGG16. Finally, a series of experiments were carried out to choose the best forecasting models.

II. LITERATURE REVIEW

Skin diseases are more common than other types of illnesses. A fungal infection, bacteria, viruses or allergies, along with other factors, can cause skin problems. Changes in the surface or tone of the skin can be caused by a skin disease. Skin infections are chronic, infectious, and can lead to skin cancer in some cases. As a result, skin diseases should always be discovered early to prevent their evolution and propagation. A skin illness takes longer to diagnose and cure, and it costs the patient both financially and physically. Numerous researchers examined skin conditions images in an attempt to establish a mechanism for diagnosing various skin disorders. In this section, various skin diseases detection techniques were investigated that have been documented in table 1 in a detailed manner.

TABLE 1. BRIEF SUMMARY OF EXISTING WORKS

Problem Statement	Method Used	Accuracy
Skin Diseases Detection [7]	Using image processing to detect the disease area and multi-class SVM for classification	Detect 3 types of skin diseases with 100% accuracy

Skin Diseases Detection [8]	Computer Vision Technique with a combination of different machine learning algorithms	Efficiency of 99%
Skin Diseases Detection [9]	Five different machine learning algorithms were used. Logistic Regression, Kernel SVM, Random Forest, Naive Bayes and Convolution Neural Network are examples of machine learning algorithms.	CNN gave the best result with 96% accuracy
Melanoma Detection [10]	Feature extraction using multi-direction three-dimensional colour-texture features selection and a multilayer back propagation neural network classifier	Accuracy 97.5%
Skin Diseases Detection [11]	For feature extraction, 2D Wavelet Transform algorithm and for classification Convolution network was used	good accuracy results
Skin Cancer Classification [12]	Using ResNet50	Accuracy almost 92%
Skin Diseases Classification [13]	Using Deep Neural Networks [LSTM and MobileNET V2]	Accuracy almost 85.34%
Melanoma Skin Cancer Detection [14]	Using SVM	Accuracy 96.9%
Skin cancer Detection [15]	Deep Learning Network	Accuracy 99.77%
Skin cancer Detection [16]	Using dermoscopic images for feature extraction and convolution neural network for classification	Accuracy 89.5%

III. BACKGROUND DETAILS

Artificial intelligence refers to the replication of human brain functions by computers, notably computer systems. To resolve real-world problems, practically DL [17,18] and ML models are used. In this work, a hybrid of the DL and ML models is preferred to handle the problem of skin disease diagnosis. Four different deep learning models were utilized to gather features from data samples for this investigation, and three different machine learning methods were employed to evaluate the input sample.

A. Deep Learning Model for Features Extraction

A feature in machine learning is a performance evaluation attribute of raw data. The selection of useful, discriminating, and independent features is a critical component of every classification method. Feature extraction from raw data is a dimension reduction approach. It is used to remove extraneous data from the original data and build a new, more compact data set for future usage.

Deep feature extraction [19] is the method of extracting features using deep learning. It gives the persistent response of a neural network layer inside a conceptual framework to an input yield a response related to the model's ultimate output. Depending on where the answer is invoked in the decision tree or other structure, one feature is regarded "deeper" over another.

A pre-defined neural network is utilized as a feature extractor in deep learning to accomplish feature extraction. This extractor allows the input picture to propagate to a pre-specified layer known as the max-pooling layer, and the outputs of that layer are used as the intended features. In this experiment, four pre-trained models were employed to do this task: Alexnet [20], Googlenet [21], Resnet50 [22], and VGG16 [23].

B. Machine Learning Models for Classification

Classification [24] is a type of computational modelling task in machine learning [25] that predicts a target class for a given sample of input data. The approach begins by attempting to predict the class labels of provided observations. The classifications are commonly referred to as a target. The objective of classification predictive modelling is to approximate the transformation matrix from categorical input variables to categorical output variables. The fundamental goal of classification is to establish which class the new information will go under. Three common machine learning classifiers were utilized in the experiment.

a) Support Vector Machine

It is a prominent algorithm [26] that falls within the supervised learning category. This approach is utilized in both classification and regression analysis. The entire aim of this method is to divide the target classes using a hyperplane with the greatest feasible divergence. This algorithm is employed as a multi-class classification model in this investigation.

b) Ada Boost Ensemble Classifier

This ensemble technique [27,28] is also known as Adaptive Boosting. The training weights were redistributed to each occurrence using this manner, with higher weights applied to improperly classified instances. Then, to remove bias, use the boosting strategy. It is based on the notion of progressive growth of learners. Thus, every subsequent learner, with the exception of the earliest, is generated from originally developed learners. In certain terms, weak learners are changed into powerful learners.

c) Decision Tree

This algorithm [29] belongs to the supervised learning algorithm. The purpose of this technique is to learn some basic decision rules from learning data to produce a learning model that predicts the correctly classified variable. In Decision Trees, predicting a class label for a record began at the tree's root. The elements of the root attribute are then compared to the values of the label's attribute. Each branch of the tree relates to that value and jump to the next node based on the correlation.

IV. PROPOSED METHODOLOGY

This research is being utilized to examine several skin problems. Four deep feature extraction approaches and three machine learning classification algorithms were used to carry out this investigation. Alexnet, Googlenet, Resnet50, and VGG16 were used to extract features from the diseased skin images, and Multi-Class Support Vector Machine, Decision Tree, and AdaBoost Ensemble were adopted as classification models. The algorithm for this investigation is shown in a part algorithm, and detailed performance analysis of each categorization model is described in section results and discussions.

A. Algorithm

Step-1: Collect the image dataset
Step-2: Pre-Process them.
Step-2.1: Remove Noises if present.
Step-2.2: Adjust Brightness and Contrast level.
Step-2.3: Adjust the Sharpness level to enhance the dark level edges.
Step-3: Divide the entire dataset in a ratio of 2:1. The primary partition is utilized as a training dataset, while the remainder should be used as a validation set.
Step-4: Now, deep learning models are used to feature extraction from the training sample, and ML models are used to categorize those.
Step-5: Now analyse the testing sample using this training subset.
Step-6: Calculate the accuracy of each disease class and visualize the result by plotting the receiver operating curve (ROC).
Step-7: Perform a quick examination of each classifier's performance using each feature selection technique.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. Experimental Dataset

About forty thousand photos were collected from the ISIC repository [30] for this project to accomplish skin disease detection.

B. Experimental Results

Tables 2, 3, and 4 describe the performance of three models using four distinct deep feature extraction strategies. There are nine disease classes in the experimental dataset, Actinic keratosis (class 1), Basal Cell Carcinoma (class 2), Dermatofibroma (class 3), Melanoma (class 4), Nevus (class 5), Pigmented Benign Keratosis (class 6), Seborrheic keratosis (class 7), Squamous Cell Carcinoma (class 8), and Vascular Lesion (class-9).

TABLE 2. PERFORMANCE ANALYSIS OF SUPPORT VECTOR MACHINE

Network Model	Classification Model: Support Vector Machine [Accuracy Observation %]									Mean Acc [%]
	class1	class2	class3	class4	class5	class6	class7	class8	class9	
AlexNet	85.54	93.65	95.97	90.12	87.58	83.95	85.62	85.47	99.08	89.66
GoogleNet	89.24	90.32	84.36	84.63	75.35	82.75	72.82	83.48	97.49	84.49
ResNet50	88.93	91.71	93.19	89.79	83.00	90.17	85.59	86.60	99.50	**89.83**
VGG16	81.30	85.76	90.42	84.39	78.28	81.70	67.85	78.62	99.34	83.07

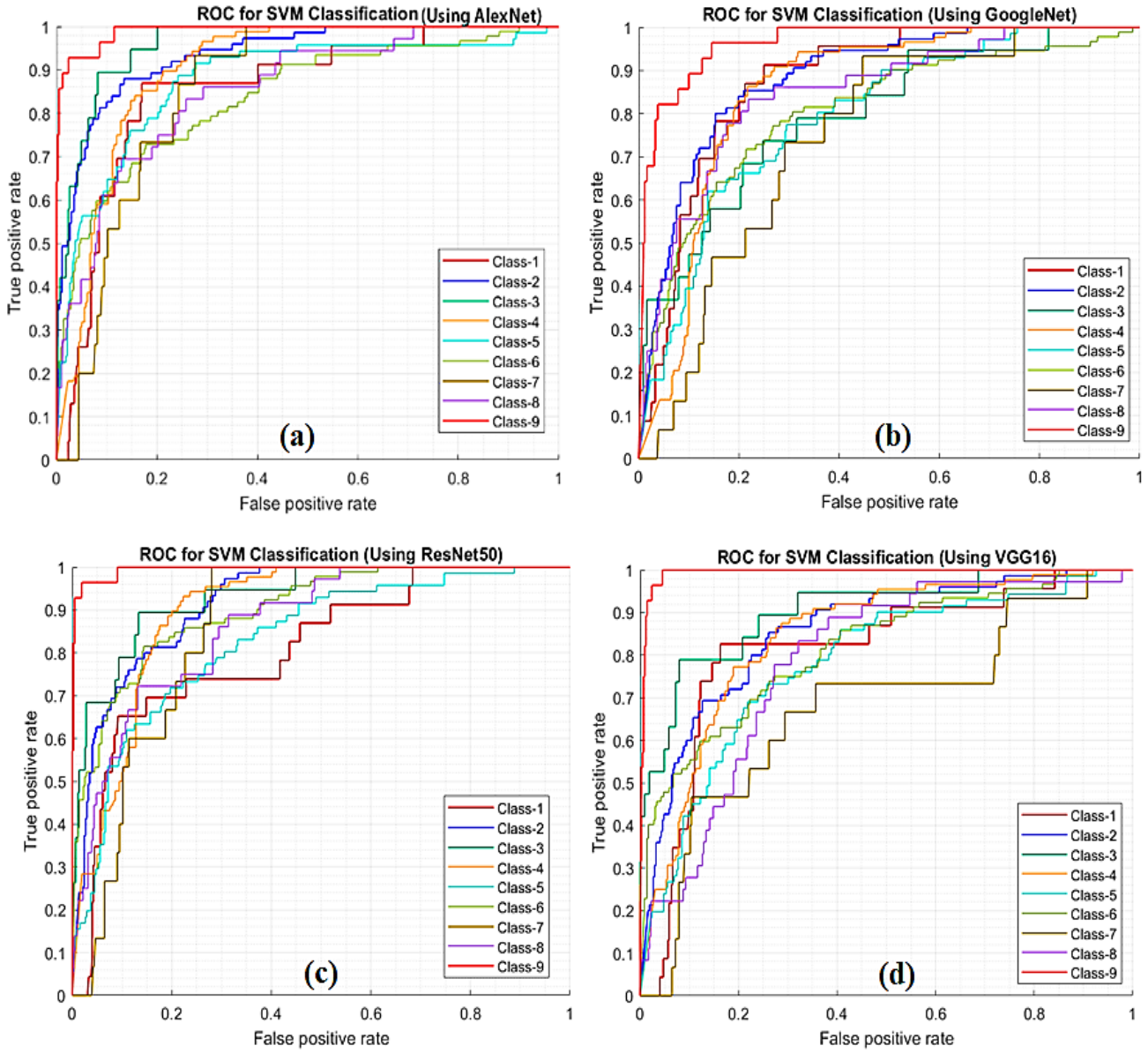

Fig 1. Characteristic Curves of SVM Classifier with Four Different Deep Feature Extraction Techniques for All Classes

TABLE 3. PERFORMANCE ANALYSIS OF ADABOOST ENSEMBLE CLASSIFIER

Network Model	Classification Model: AdaBoost Classifier [Accuracy Observation %]									Mean Acc [%]
	class1	class2	class3	class4	class5	class6	class7	class8	class9	
AlexNet	91.82	90.44	85.17	87.11	85.18	87.00	74.70	84.37	97.49	87.03
GoogleNet	85.57	86.10	80.72	85.73	81.29	76.30	84.24	78.39	96.12	83.83
ResNet50	91.06	84.96	86.65	81.63	84.13	87.74	83.45	85.62	98.87	**87.12**
VGG16	88.28	83.09	76.34	80.28	80.26	75.37	77.62	78.35	98.08	81.96

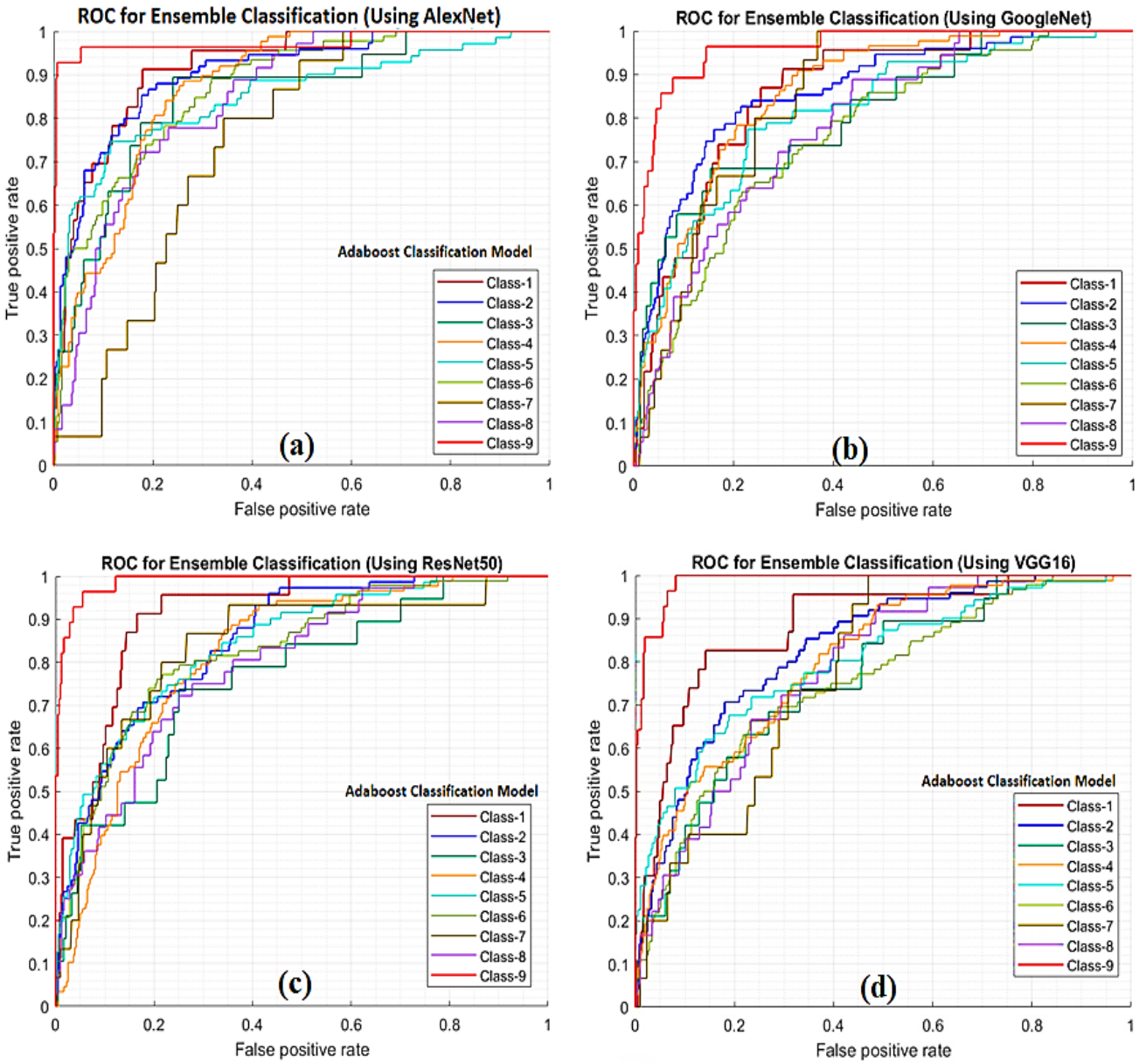

Fig 2. Characteristic Curves of AdaBoost Ensemble Classifier with Four Different Deep Feature Extraction Techniques for all classes

TABLE 4. PERFORMANCE ANALYSIS OF DECISION TREE CLASSIFIER

Network Model	Classification Model: Decision Tree [Accuracy Observation %]									Mean Acc [%]
	class1	class2	class3	class4	class5	class6	class7	class8	class9	
AlexNet	60.97	70.34	72.61	66.20	65.21	68.25	77.52	56.92	74.46	68.05
GoogleNet	70.61	68.25	61.19	61.64	62.95	64.41	82.13	59.74	84.04	68.33
ResNet50	61.16	69.22	75.47	72.93	64.61	65.90	78.24	53.18	78.60	**68.81**
VGG16	58.92	63.02	53.96	62.01	73.06	62.11	74.09	48.40	71.51	63.01

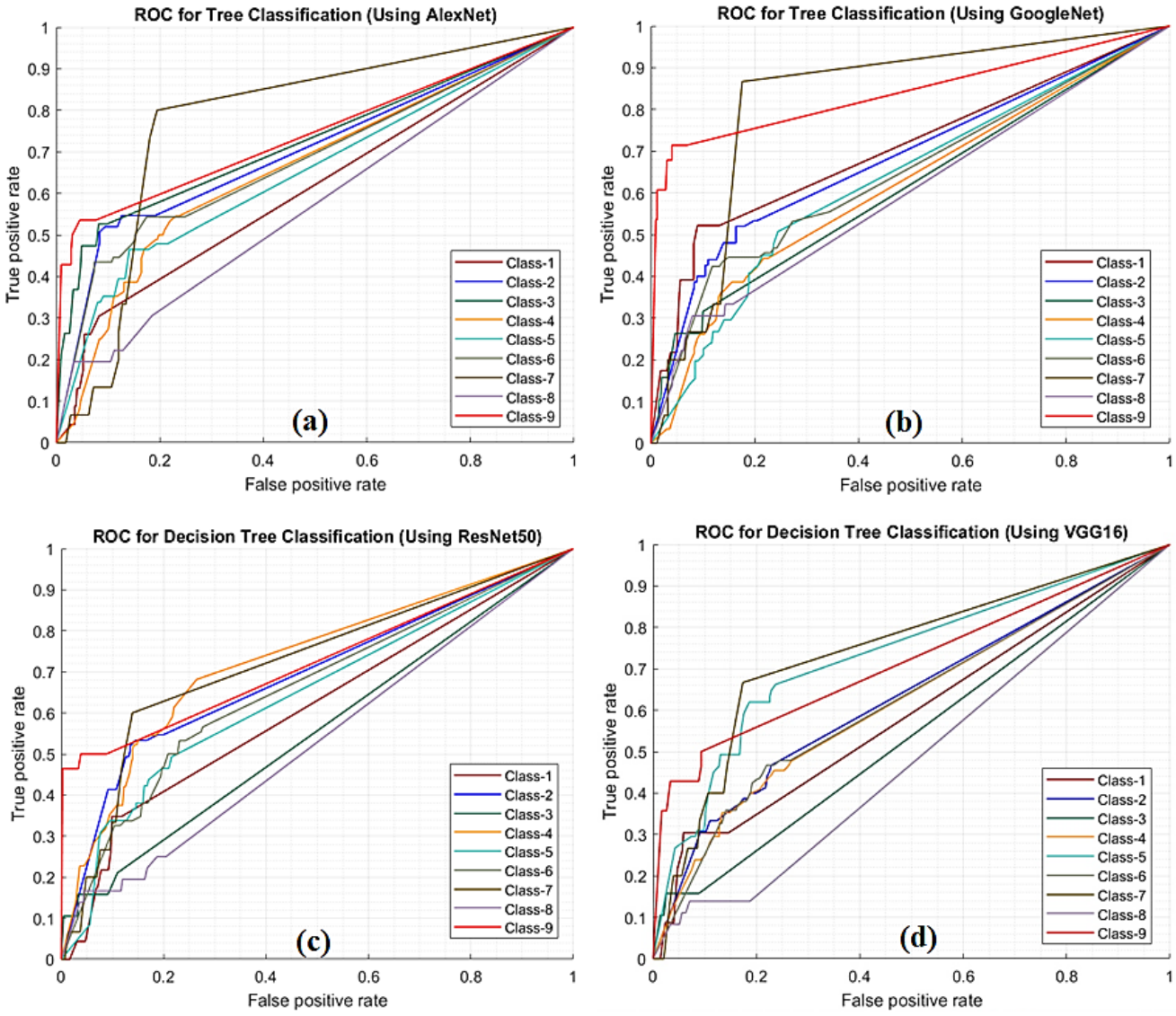

Fig 3. Characteristic Curves of Decision Tree Classifier with Four Different Deep Feature Extraction Techniques for All Classes

This study is motivated in applying an integrated approach for skin diseases detection. This integrated method incorporates Resnet50 for the feature selection process, which is further followed by employing the Support Vector Machine classification model. This hybrid model is quite efficient in identifying various skin diseases. The results are further described in table 2, 3, and 4. Table 5 further demonstrates the overall performance of the SVM as a classifier when integrated with other features selection models such as AlexNet, GoogleNet, Resnet50, and VGG16.

TABLE 5. OVERALL PERFORMANCE ANALYSIS OF SUPPORT VECTOR MACHINE CLASSIFIER

Deep Neural Network Models for Feature Extraction	Overall Calculated Accuracy [%]
AlexNet	91.62
GoogleNet	96.42
ResNet50	**99.11**
VGG16	97.84

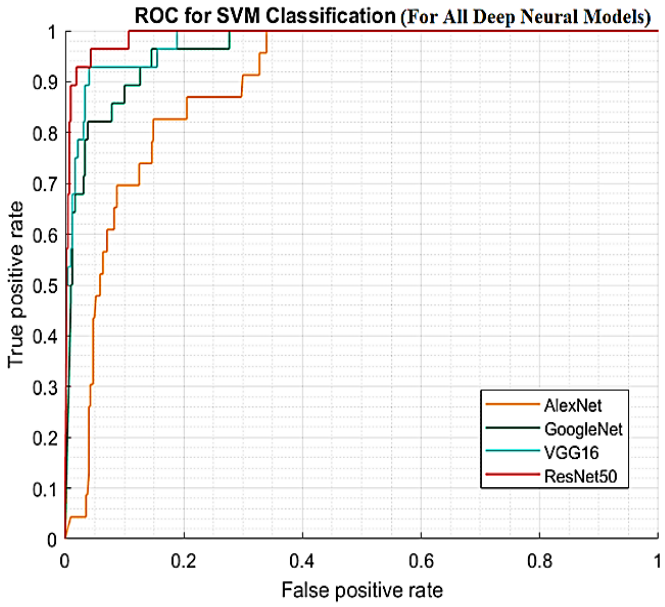

Fig 4. Characteristic Curves of SVM Classifier for Four Different Deep Feature Extraction Techniques

C. Discussions

An automated skin disease recognition system has been developed as part of this study to diagnose skin diseases. This analysis was conducted using the data from the ISIC repository. In this work, a combination of ML and DL models were employed to accomplish the diagnosing procedure. Deep learning models were employed to extract attributes from the source images, while ML models were used to determine disease categories. This study employed four distinct deep learning models and three distinct machine learning models. Instead of focusing on only one feature selection and classification technique, multiple feature selection models along with multiple classifier models were involved. As feature selection methods, four pre-trained deep neural networks, AlexNet, GoogleNet, ResNet50, and VGG16, were used, and for classification, SVM, Adaboost ensemble, and Decision tree algorithms were applied. All the different combinations are tried out, and the performance is measured in terms of accuracy. Finally, it has been observed SVM turns out to be the best classifier model. This model even outperforms well when the Resnet50 deep neural network is used as a feature selection model. The performance justification for all the combinations is provided in Tables 2, 3, and 4.

In terms of effectiveness, this combination (Resnet50 + SVM) exceeds all other classifiers. This aforementioned combination was found to have the higher precision, with 89.83 % for each disease class. Table 5 also revealed that the overall accuracy of the fusion technique, Resnet50 with SVM, delivers the maximum accuracy of 99.11%.

VI. CONCLUSION

The epidermis is the mortal's natural biggest entity. Skin illnesses can emerge as a result of a variety of environmental and internal factors. As a result, discovering skin illnesses is a critical component of medical science. Because it has the potential to diminish the number of people who die as a consequence of skin disease or exposure to infectious. The therapeutic approach takes a long time and does not always correctly diagnose the type of disease. In these circumstances, automatic skin disease diagnosis techniques are extremely beneficial. An electronic skin disease diagnosis approach is proposed in this investigation. A hybrid of ML and DL models were employed to complete this process. In this suggested strategy, four deep learning models for feature extraction from training data were combined with three prominent machine learning classifiers. Finally, a full comparative analysis was performed, and it was discovered that the deep network model Resnet50 with Support Vector Machine Classifier produces the best results for prediction with a 99.11% accuracy rate.

REFERENCES

[1] Mohammed, S. S., & Al-Tuwaijari, J. M. Skin Disease Classification System Based on Machine Learning Technique: A Survey. IOP Conference Series: Materials Science and Engineering, 1076 (1) (2021) 012045.

[2] Hashmani, M. A., Jameel, S. M., Rizvi, S. S. H., & Shukla, S. An adaptive federated machine learning-based intelligent system for skin disease detection: A step toward an intelligent dermoscopy device. Applied Sciences (Switzerland), 11(5) (2021) 1–19.

[3] Li, H., Pan, Y., Zhao, J., & Zhang, L. Skin disease diagnosis with deep learning: a review. Hongfeng Li. (2020).

[4] Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., Alsaiari, S. A., Saeed, A. H. M., Alraddadi, M. O., & Mahnashi, M. H. Skin cancer detection: A review using deep learning techniques. International Journal of Environmental Research and Public Health. 18(10) (2021).

[5] Verma, A. K., Pal, S., & Kumar, S. Comparison of skin disease prediction by feature selection using ensemble data mining techniques. Informatics in Medicine Unlocked. 16(April) (2019) 100202.

[6] Jaychandra Reddy, V., & Nagalakshmi, T. J. Skin disease detection using artificial neural network. Indian Journal of Public Health Research and Development, 10(11) (2019) 3829–3832.

[7] Alkolifi Alenezi, N. S. A Method of Skin Disease Detection Using Image Processing and Machine Learning. Procedia Computer Science. 163 (2019) 85–92.

[8] Santhiya, D. S., Pravallika, S. S. L., Sukrutha, M. A., Nishanth, I., Iswarya, N., & Aishwarya, D. Skin Disease Detection using V2 and V3 in Machine Learning. International Journal of Engineering Science and Computing. 9(4) (2019) 21343–21347.

[9] Bhadula, S., Sharma, S., Juyal, P., & Kulshrestha, C. Machine Learning Algorithms based Skin Disease Detection. International Journal of Innovative Technology and Exploring Engineering. 9(2) (2019) 4044–4049.

[10] Warsi, F., Khanam, R., Kamya, S., & Suárez-Araujo, C. P. An efficient 3D colour-texture feature and neural network technique for melanoma detection. Informatics in Medicine Unlocked, 17(November 2018) (2019) 100176.

[11] Leelavathy S, Jaichandran R, Shobana R, Vasudevan, S. S. P. and N. Skin Disease Detection Using Computer Vision and Machine Learning Technique. European Journal of Molecular & Clinical Medicine. 7(4) (2020) 2999–3003.

[12] Gouda, N., & Amudha, J. Skin Cancer Classification using ResNet. 2020 IEEE 5th International Conference on Computing Communication and Automation, ICCCA 2020. (2020) 536–541.
[13] Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., & Kang, J. J. Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. Sensors. 21(8) (2021) 2852.
[14] Banasode, P., Patil, M., & Ammanagi, N. A Melanoma Skin Cancer Detection Using Machine Learning Technique: Support Vector Machine. IOP Conference Series: Materials Science and Engineering. 1065(1) (2021) 0–5.
[15] Kadampur, M. A., & Al Riyaee, S. X. Skin cancer detection: Applying a deep learning-based model-driven architecture in the cloud for classifying dermal cell images. Informatics in Medicine Unlocked, 18(December 2019) (2021) 100282.
[16] Hasan, M., Barman, S. Das, Islam, S., & Reza, A. W. Skin cancer detection using convolutional neural network. ACM International Conference Proceeding Series: March 2020 (2019) 254–258.
[17] Gopalakrishnan, S., Abishek. B, Dr.E., Vijayalakshmi, Dr A., Rajendran, Dr V. Analysis And Diagnosis Using Deep-Learning Algorithm On Erythemato-Squamous Disease. International Journal of Engineering Trends and Technology 69(3) (2021) 52-57.
[18] Mohan, N. Recognition of Skin Diseases using Deep Neural Network Optimized by Group Teaching Algorithm. International Journal of Engineering Trends and Technology. 68(9) (2020) 109-120.
[19] Patel, K. Image Feature Extraction: Traditional and Deep Learning Techniques. Medium (2020).
[20] Deep Learning Based Image Segmentation with AlexNet Feature Extraction for Classification of Mammogram Images. International Journal of Pharmaceutical Research 13(01) (2021).
[21] Alake, R. Deep Learning: GoogLeNet Explained - Towards Data Science. Medium. (2020).
[22] Kaushik, A. Understanding ResNet50 architecture. OpenGenus IQ: Computing Expertise & Legacy (2020).
[23] Manasa, K., & Student, M. T Skin Cancer Detection Using VGG-16. European Journal of Molecular & Clinical Medicine 08(01) (2021) 1419–1426.
[24] Waseem, M. How To Implement Classification In Machine Learning? Edureka (2021).
[25] Team, E. What is the Definition of Machine Learning? Expert.Ai (2021)
[26] Ray, S. SVM | Support Vector Machine Algorithm in Machine Learning. Analytics Vidhya (2021).
[27] Great Learning Team. The Ultimate Guide to AdaBoost Algorithm | What is AdaBoost Algorithm? GreatLearning Blog: Free Resources What Matters to Shape Your Career! (2021).
[28] Narayanan A G, H., Singh, Dr J.A.P. Skin Disease Ensemble Classification Using Transfer Learning and Voting Classifier. International Journal of Engineering Trends and Technology. 69(12) (2021) 287-293.
[29] Saini, A. Decision Tree Algorithm - A Complete Guide. Analytics Vidhya (2021).
[30] ISIC Archive. (2016). ISIC Archive. Retrieved October 19, 2021, from https://www.isic-archive.com/