CLASSIFICATION BY A STACKING MODEL USING CNN FEATURES FOR MEDICAL IMAGE DIAGNOSIS

Baidaa MUTASHER RASHED[1]*, Nirvana POPESCU[2]

Medical imaging coupled with Artificial Intelligence (AI) applications, in particular Deep learning (DL) and Machine Learning (ML), can speed up the disease diagnostic process. The purpose of this work is to present a novel disease detection system by suggesting a new Convolutional Neural Network (CNN) model and combining the CNN features with three of ML classifiers and suggesting a new classifier using the stacking model. The proposed system was used in binary and multiclassification and applied to two different medical datasets. The proposed model was evaluated using accuracy, sensitivity, specificity, precision, recall, F1 score, and AUC, achieving robust results.

Keywords: Deep learning; CNN; Machine learning, Stacking ensemble model

1. Introduction

In recent years, CNN has achieved remarkable outcomes in image analysis [1]. In comparison to the handcrafted feature extraction-based methods, the CNN method can automatically learn to extract the features from large-scale datasets [2]. CNN extracts features through its structure of deep and different layers. CNN architectures, according to the planned architecture, extract features from images and then categorize them using fully connected layers, the work of which corresponds to the ML method. As a result, instead of these layers, the ML algorithm can be used to generate effective classifications [3]. The proposed system was applied to two medical datasets to diagnose two types of dangerous diseases spread in the world, especially in recent times due to negative changes in the climate, environment, and human lifestyle, which are lung diseases and melanoma skin cancer that are treated if detected early [4, 5].

The purpose of this work is to build a novel disease diagnosis system that depends on deep learning (DL) represented by convolutional neural network (CNN) and machine learning (ML) methods represented by Support Vector Machine (SVM), and Naive Bayes (NB), and Decision Tree (DT). In this system, we first designed 37 layers CNN structure, then extracted features from the CNN model. After that, the three ML algorithms are applied to classify these features. Finally,

[1] Computer Science Dept., National University of Science and Technology POLITEHNICA of Bucharest, Romania, contact author: rashed.baidaa@stud.acs.upb.ro, baidaaalsafy@utq.edu.iq

[2] Prof., Computer Science Dept., University POLITEHNICA of Bucharest, Romania, nirvana.popescu@upb.ro

we proposed a new classifier using the stacking model, which combines three hybrid classifiers (CNN-SVM, CNN-NB, CNN-DT). The proposed system was used in binary and multiclassification and applied to two different sets of medical data (chest X-ray and dermoscopy melanoma skin cancer) to increase generalizability. Experimental results revealed that the suggested system attained good results.

Recent research has concentrated on creating an ensemble of multiple models to obtain great accuracy with medical imaging. The Ensemble approach has been shown to be effective in boosting the overall accuracy of several applications. In [6], the authors proposed a unique stacked ensemble-based architecture by combining fine-tuned pre-trained CNN models like Xception, InceptionResNet-V2, Inceptionv3, DenseNet201, and DenseNet121 for acral lentiginous melanoma classification. The suggested method's performance was evaluated using a Figshare benchmark dataset. The results show that the suggested method achieved 97.93% accuracy. In [7] the authors suggested a stacking-ensemble model that combines six pre-trained CNN models, comprising EfficientNetV2-B0, Efficient- NetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S, and EfficientNetV2-M on the categorization of chest X-ray and CT images, on the chest X-ray dataset; the suggested ECA-EfficientNetV2 model achieved the greatest accuracy (99.21%), on the chest CT dataset; the suggested ECA-EfficientNetV2 model achieved the greatest accuracy (99.81%). The study in [8] aimed to classify chest X-ray images into COVID-19, normal and viral pneumonia using a stacking model constructed by integrating three single ML (SVM, ANN, LR) classifiers. Features were extracted from the CNN-based VGG19 structure, the SVM, ANN, LR, and stacking models attained classification accuracy of 90.2%, 96.2%, 96.7%, and 96.9%, respectively.

2. Materials and Methods

The suggested system of this work consists of six main stages: the first stage is responsible for loading medical datasets and dividing them into 70% for the training and 20% for the validation, and 10% for the testing, the second stage is responsible for preprocessing; the third stage is responsible for suggesting a new CNN structure for feature extraction. The fourth stage is responsible for building hybrid classifiers (CNN-ML); The fifth stage is responsible for applying a stacking ensemble model for medical datasets classification. Finally, evaluation of the model by using evaluation metrics. Fig. 1 shows the workflow of the proposed system.

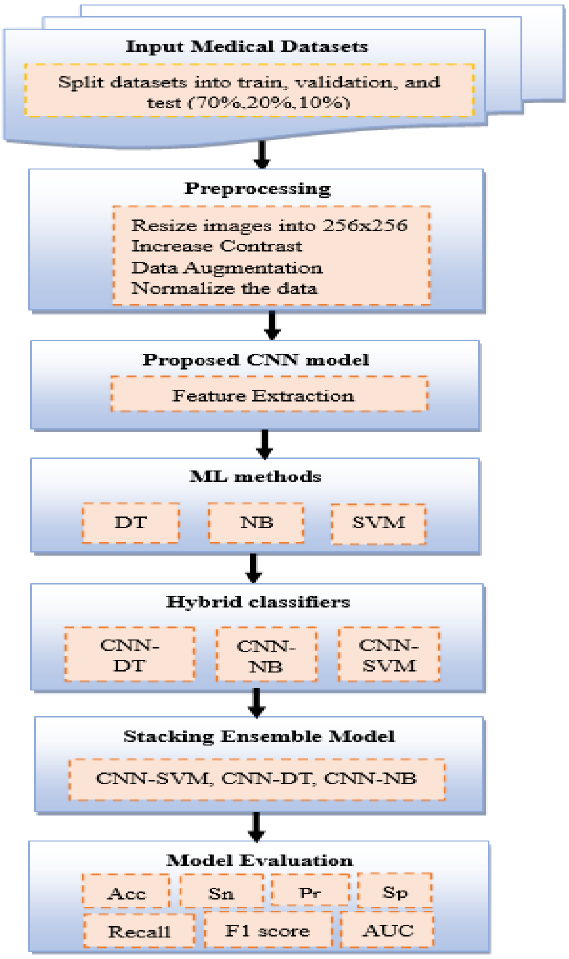


Fig. 1. Workflow of the suggested system

2.1 Datasets

In this work, two groups of medical databases were used for binary and multiclassification; the first dataset (DB1) is comprised of images for chest X-rays obtained from Kaggle [9, 10]. The second dataset (DB2) is comprised of dermoscopy images for melanoma skin cancer obtained from [11, 12]. For binary classification, 1050 X-ray images (500 normal, 550 abnormal) and 550 dermoscopy images (250 benign, 300 malignant) were used. For multiclassification,1060 X-ray images (338 bacterial, 368 covid-19, 354 viral Pneumonia) and 585 dermoscopy images (145 acral, 150 lentigo, 150 nodular, superficial) were used. Images were captured in JPG format, with different resolution sizes.

2.2 Datasets Preprocessing

Image preprocessing for the datasets comprises four main processes: First, resizing all images to 256 x 256. Second, increasing the contrast by using gamma-correcting the intensity of the images with a gamma value of 0.7 to make the images clearer [13]. Third, enhancing the number of training samples for each category by data augmentation to minimize overfitting and boost overall model performance [14]. To expand the training dataset, two image augmentation techniques (rotation and translation) were utilized. Finally, normalizes the entered data to reduce unwanted characteristics and data redundancies by dividing them by 255 [15].

2.3 The proposed CNN Architecture

The proposed CNN consists of 37-layer comprising the input layer, 8 convolutional layers, 8 batch normalization layers, 8 activation functions (ReLU), 8 max-pooling layers, and a fully connected layer with one dropout layer, a SoftMax layer, and a classification layer. Details of the information of the layers in the suggested CNN structure are given in Table 1. The learning rate (LR) was set to 0.001, the model was trained for 50 epochs, the loss function was binary and multi-cross-entropy, and Adam was used as an optimizer which has two advantages: it uses less memory and fewer computing resources [16, 17].

Table 1

Information detail of the layers in the suggested CNN structure

No. layer	Name Layer	Info	Value	No. layer	Name Layer	Info	Value
1	Input layer	Size	256x256	15	Batch_Norm_5	Channels	64
2	Conv_1	Filters Kernel Size Activation	8 3x3 Relu	16	Maxpool_5	Kernel Size Stride	2x2 2x2
3	Batch_Norm_1	Channels	8	17	Conv_6	Filters Kernel Size Activation	64 3x3 Relu
4	Maxpool_1	Kernel Size Stride	2x2 2x2	18	Batch_Norm_6	Channels	64
5	Conv_2	Filters Kernel Size Activation	16 3x3 Relu	19	Maxpool_6	Kernel Size Stride	2x2 2x2
6	Batch_Norm_2	Channels	16	20	Conv_7	Filters Kernel Size Activation	128 3x3 Relu
7	Maxpool_2	Kernel Size Stride	2x2 2x2	21	Batch_Norm_7	Channels	128
8	Conv_3	Filters Kernel Size Activation	32 3x3 Relu	22	Maxpool_7	Kernel Size Stride	2x2 2x2

9	Batch_Norm_3	Channels	32	23	Conv_8	Filters Kernel Size Activation	128 3x3 Relu
10	Maxpool_3	Kernel Size Stride	2x2 2x2	24	Batch_Norm_8	Channels	128
11	Conv_4	Filters Kernel Size Activation	32 3x3 Relu	25	Maxpool_8	Kernel Size Stride	2x2 2x2
12	Batch_Norm_4	Channels	32	26	Dropout	‘dropout’	0.2
13	Maxpool_4	Kernel Size Stride	2x2 2x2	27	Fc	Activation	Softmax
14	Conv_5	Filters Kernel Size Activation	64 3x3 Relu	28	Output	Classification	Cross entropy

2.4 The CNN-ML Classifiers

In this work, we used three ML classifiers (SVM, DT, NB) to classify the medical datasets using the features extracted from the proposed CNN model.

SVM is a supervised ML method that is utilized for classification [18]. The SVM algorithm's fundamental idea is to locate the plane with the highest margin, or the greatest distance between data points from classes [19]. DT is a type of supervised machine learning that is utilized to address classification problems [20]. It is referred to as a decision tree because it begins with the root node and then branches out to generate multiple branches and a tree-like structure [21]. NB is a statistical classifier based on the Bayes theorem [22]. NB is divided into two steps: learning and testing. In learning, an estimation is generated according to the applied characteristics; in testing, predictions are generated according to the learning phase [23]. The CNN-ML model work is explained as follows: First, the proposed CNN is trained with whole layers. Then, the classification layers are eliminated from the proposed CNN architecture, and 128 features are acquired from the dataset from the last max pool layer (Maxpool_8), where this layer possesses important characteristics that help in the classification. After, these training features are trained with SVM, DT, and NB classifiers. Finally, the classification of the testing features is acquired with these classifiers. Fig. 2 demonstrates the scheme of the proposed CNN architecture to build the structure of the CNN-ML model.

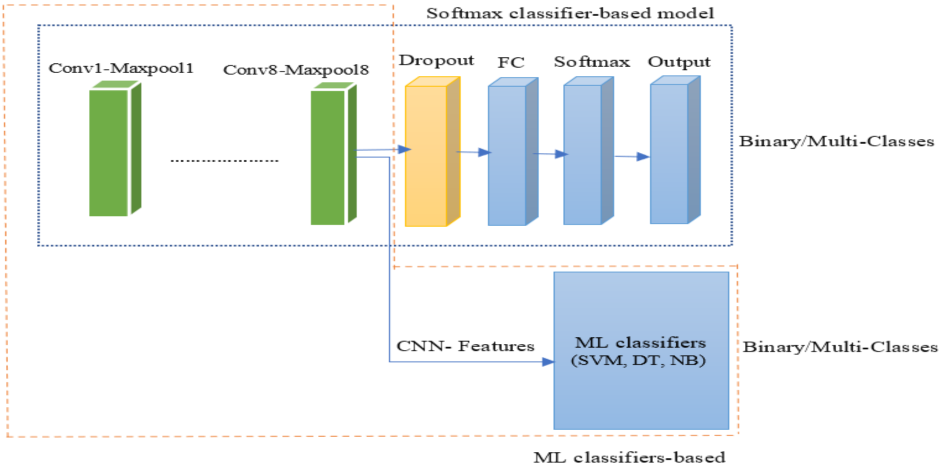


Fig. 2. Scheme of suggested CNN with ML classifiers architecture

2.5 The Stacking Ensemble Model

Stacking's major role is to integrate different methods to create predictions. Stacking is based on the idea of mixing classifiers to create a new classification model [24]. The stacking-ensemble model works as follows: First, the predictions obtained from the DT, SVM, and NB models are given as input to the stacking model. Second, the basis models (NB, SVM, and DT) are trained using training data, and the training set is trained using the five k-fold cross-validations. Following training, the model's performance is evaluated using test data, with each model providing an individual forecast. The forecasts of these models serve as an extra input to our ensemble learning, which functions as a combined model trained to generate the final prediction. Fig. 3 shows a flow chart of the stacking model employed in the work.

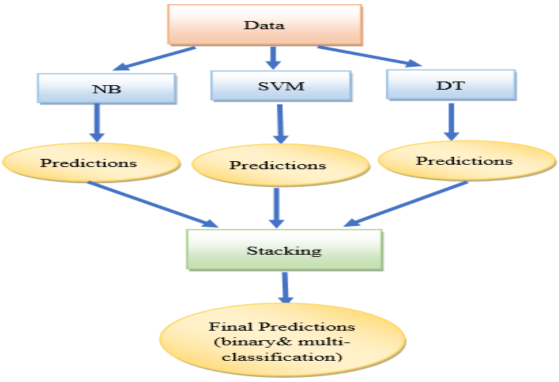


Fig. 3. Flow diagram of the proposed stacking model

2.6 Model Evaluation

Common performance measures such as accuracy (Acc), sensitivity (Sn), specificity (Sp), precision (Pr), recall, F1-score, and AUC values were utilized for evaluation [25]. Also, ROC (Receiver Operator Characteristics Curve) curves were utilized to measure the performance of classification models [8]. Performance measures are computed with the formulas in Table 2:

Table2

Formulas for performance metrics

Performance Measures	Formula
Accuracy	$(TP+TN)/(TP+TN+FN+FP)$
Sensitivity	$TP/\ (TP+FN)$
Specificity	$TN/(TN+FP)$
Precision	$TP/(TP+FP)$
Recall	$TP/(TP+FN)$
F1-score	$(2TP)/(2TP+FN+FP)$

TP and TN denote the number of correctly predicted positive and negative samples; FP and FN denote the number of incorrectly predicted positive and negative samples.

3. Experimental Results

Accuracy, Specificity, Sensitivity, Recall, precision, F1-score, and AUC values were computed for each hybrid classifier and stacking model for the two datasets as shown in Table 3.

Table3

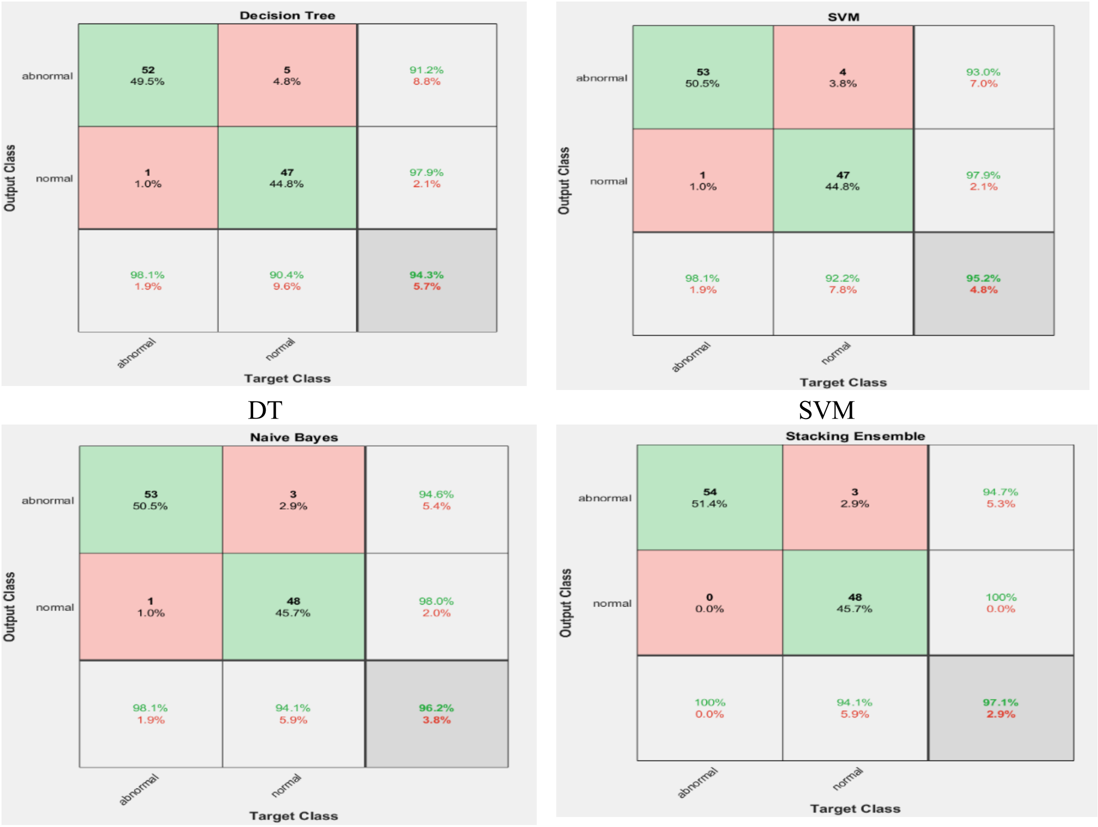
Results of performance metrics for CNN-ML models and stacking model

Data set	Class	Classifier	Acc %	Sn%	Sp%	Pr	Recall	F1-score	AUC
DB1	Binary	CNN-DT	94.3	98.2	90.3	0.94	0.94	0.94	0.94
		CNN-SVM	95.2	98.1	92.8	0.95	0.95	0.95	0.97
		CNN-NB	96.2	98.7	94.9	0.96	0.96	0.96	0.97
		Stacking	97.1	100	94.1	0.97	0.97	0.97	0.98
	Multi	CNN-DT	92.5	94.5	82.3	0.92	0.92	0.92	0.94
		CNN-SVM	90.6	89.4	82.3	0.90	0.90	0.90	0.94
		CNN-NB	91.5	92	82.3	0.91	0.91	0.91	0.95
		Stacking	94.3	96	90.3	0.94	0.94	0.94	0.97
DB2	Binary	CNN-DT	96.4	100	93.3	0.96	0.96	0.96	0.97
		CNN-SVM	96.4	96	96	0.96	0.96	0.96	0.96

		CNN-NB	94.5	95.8	93.5	0.94	0.94	0.94	0.95
		Stacking	98.2	100	96.6	0.98	0.98	0.98	0.98
	Multi	CNN-DT	93.1	93.1	81.2	0.93	0.93	0.93	0.94
		CNN-SVM	94.8	95.4	86.6	0.95	0.94	0.98	0.95
		CNN-NB	96.6	95.4	87.5	0.97	0.96	0.96	0.97
		Stacking	98.3	100	92.8	0.98	0.98	0.98	0.99

Each model showed success in binary and multiclassification and the stacking model has proven to be the most successful in classification.

We used a confusion matrix (CM) and ROC curves with AUC values to evaluate the performance of each classifier in the testing stage. In Figs. 4–7, the CM of each classifier; ROC curves, and AUC values are shown in Figs. 8–11.



Decision Tree

Output Class	abnormal	normal	
abnormal	52 49.5%	5 4.8%	91.2% 8.8%
normal	1 1.0%	47 44.8%	97.9% 2.1%
	98.1% 1.9%	90.4% 9.6%	94.3% 5.7%

Target Class

DT

SVM

Output Class	abnormal	normal	
abnormal	53 50.5%	4 3.8%	93.0% 7.0%
normal	1 1.0%	47 44.8%	97.9% 2.1%
	98.1% 1.9%	92.2% 7.8%	95.2% 4.8%

Target Class

SVM

Naive Bayes

Output Class	abnormal	normal	
abnormal	53 50.5%	3 2.9%	94.6% 5.4%
normal	1 1.0%	48 45.7%	98.0% 2.0%
	98.1% 1.9%	94.1% 5.9%	96.2% 3.8%

Target Class

NB

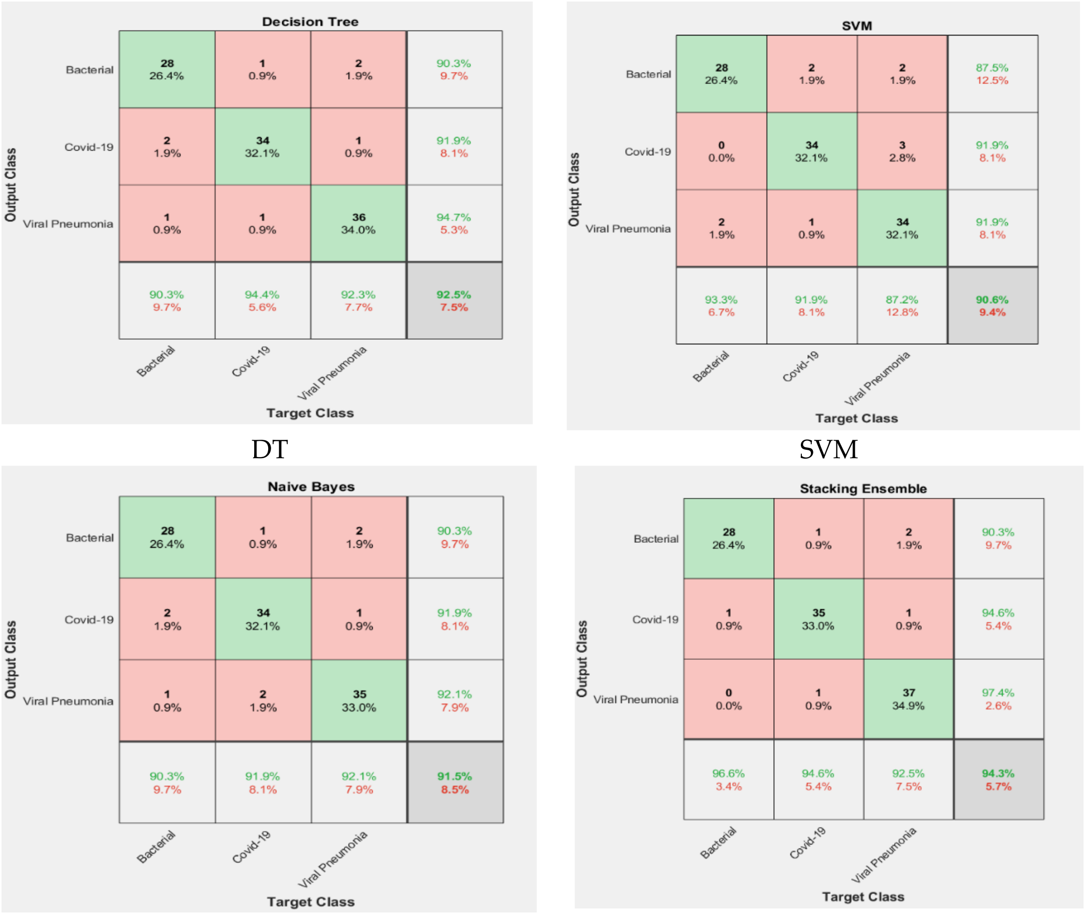
Stacking Ensemble

Output Class	abnormal	normal	
abnormal	54 51.4%	3 2.9%	94.7% 5.3%
normal	0 0.0%	48 45.7%	100% 0.0%
	100% 0.0%	94.1% 5.9%	97.1% 2.9%

Target Class

Stacking model

Fig. 4. CM for DB1 in binary classification



	Bacterial	Covid-19	Viral Pneumonia	
Bacterial	28 26.4%	1 0.9%	2 1.9%	90.3% 9.7%
Covid-19	2 1.9%	34 32.1%	1 0.9%	91.9% 8.1%
Viral Pneumonia	1 0.9%	1 0.9%	36 34.0%	94.7% 5.3%
	90.3% 9.7%	94.4% 5.6%	92.3% 7.7%	92.5% 7.5%

	Bacterial	Covid-19	Viral Pneumonia	
Bacterial	28 26.4%	2 1.9%	2 1.9%	87.5% 12.5%
Covid-19	0 0.0%	34 32.1%	3 2.8%	91.9% 8.1%
Viral Pneumonia	2 1.9%	1 0.9%	34 32.1%	91.9% 8.1%
	93.3% 6.7%	91.9% 8.1%	87.2% 12.8%	90.6% 9.4%

DT

SVM

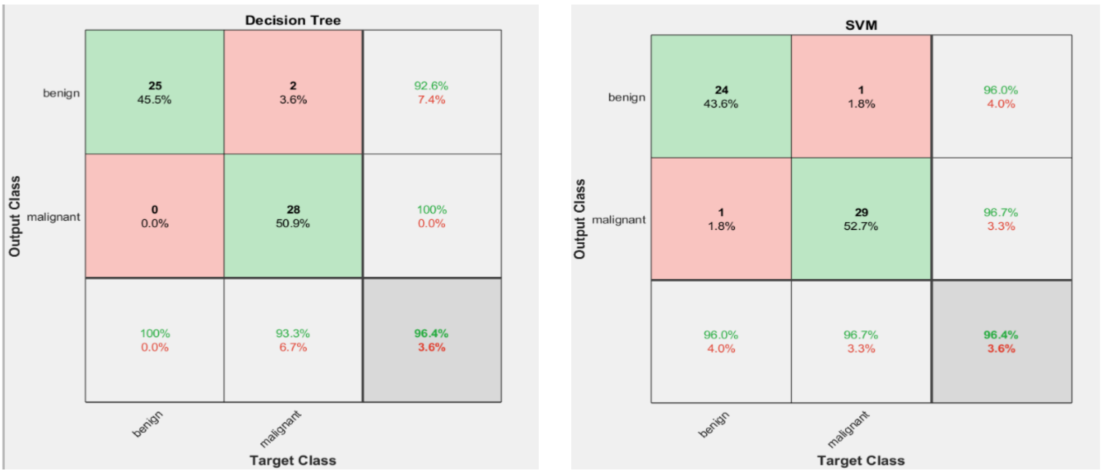
	Bacterial	Covid-19	Viral Pneumonia	
Bacterial	28 26.4%	1 0.9%	2 1.9%	90.3% 9.7%
Covid-19	2 1.9%	34 32.1%	1 0.9%	91.9% 8.1%
Viral Pneumonia	1 0.9%	2 1.9%	35 33.0%	92.1% 7.9%
	90.3% 9.7%	91.9% 8.1%	92.1% 7.9%	91.5% 8.5%

	Bacterial	Covid-19	Viral Pneumonia	
Bacterial	28 26.4%	1 0.9%	2 1.9%	90.3% 9.7%
Covid-19	1 0.9%	35 33.0%	1 0.9%	94.6% 5.4%
Viral Pneumonia	0 0.0%	1 0.9%	37 34.9%	97.4% 2.6%
	96.6% 3.4%	94.6% 5.4%	92.5% 7.5%	94.3% 5.7%

NB

Stacking model

Fig. 5. CM for DB1 in multi-classification

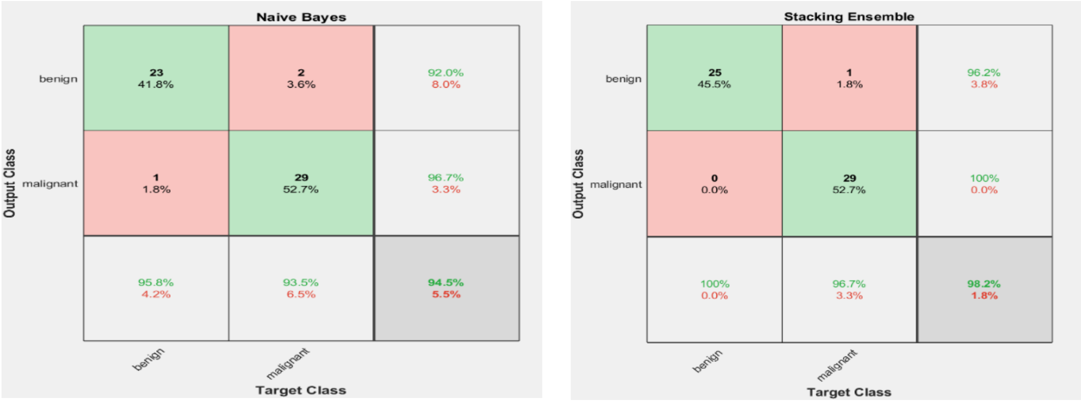


	benign	malignant	
benign	25 45.5%	2 3.6%	92.6% 7.4%
malignant	0 0.0%	28 50.9%	100% 0.0%
	100% 0.0%	93.3% 6.7%	96.4% 3.6%

	benign	malignant	
benign	24 43.6%	1 1.8%	96.0% 4.0%
malignant	1 1.8%	29 52.7%	96.7% 3.3%
	96.0% 4.0%	96.7% 3.3%	96.4% 3.6%

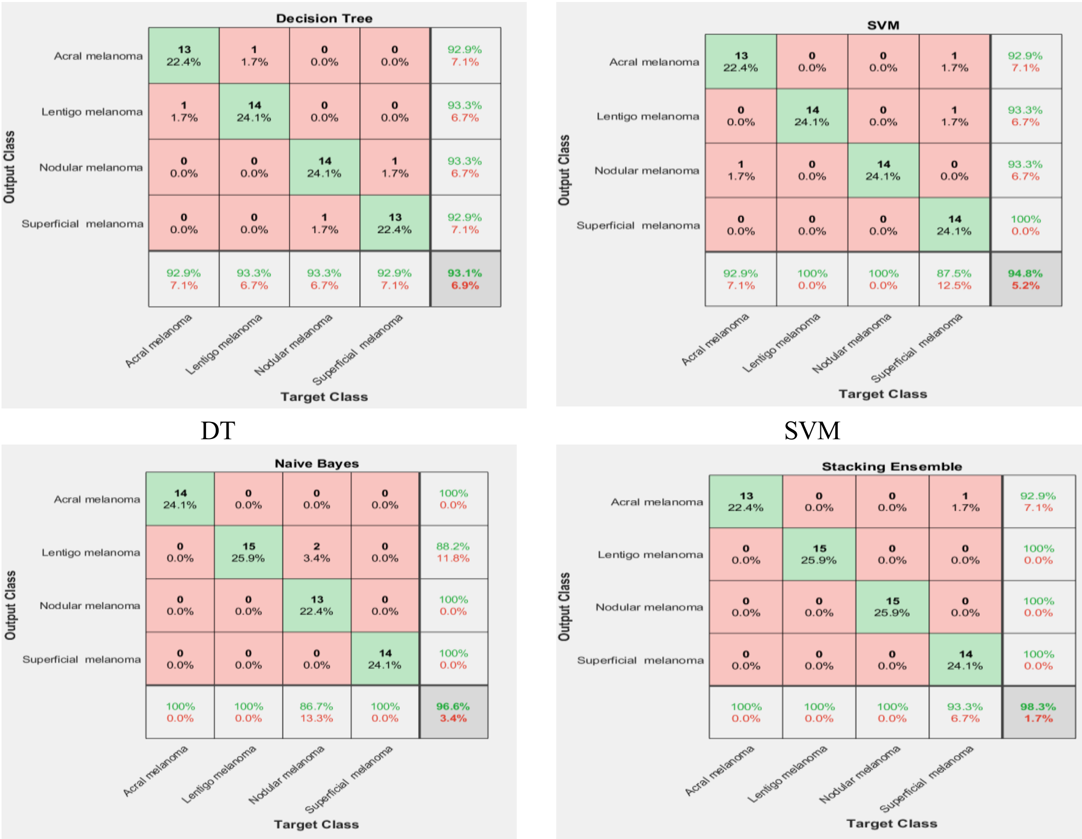
DT

SVM



NB Stacking model

Fig. 6. CM for DB2 in binary classification



DT SVM

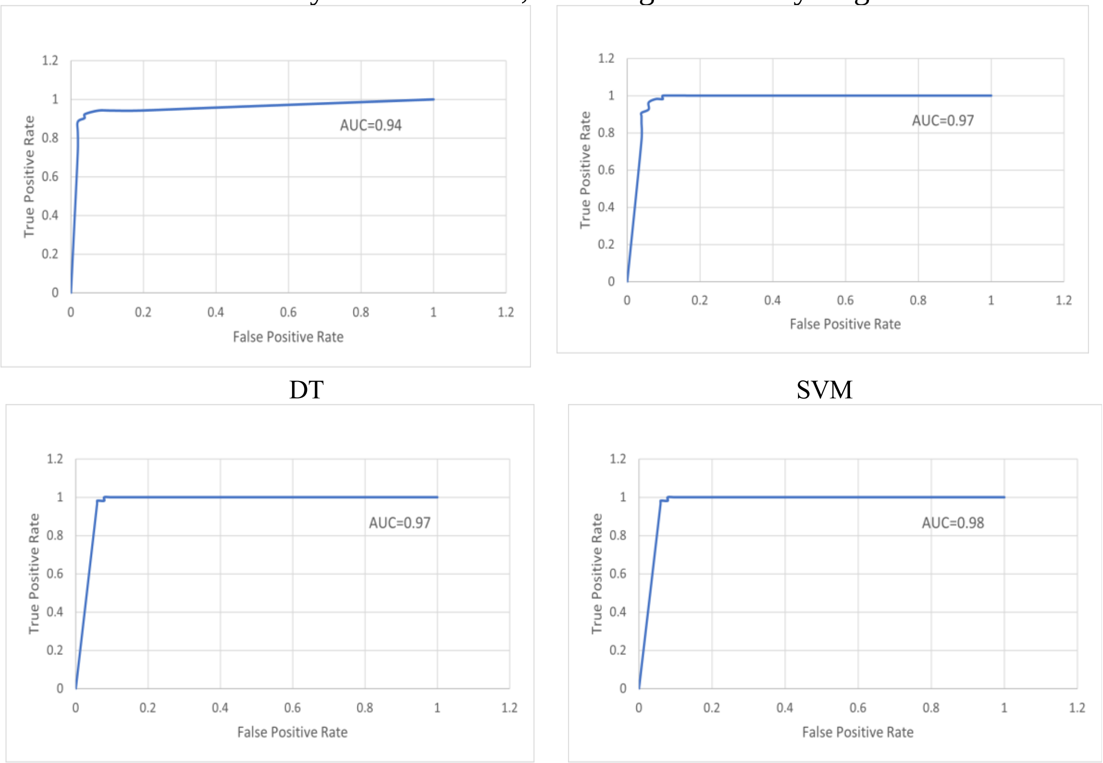
NB Stacking model

Fig. 7. CM for DB2 in multi-classification

The CM for DB1 and DB2 datasets can be observed in Figs. 4–7.

For the lung dataset, in binary classification, we utilized 1050 samples for training and the number of samples for testing was 105. In multi-classification, we used 1060 samples for training and the number of samples for testing was 106. For the skin cancer dataset, in binary classification, we utilized 550 samples for training and, the number of samples for testing was 55; in multi-classification, we used 585 samples for training and the number of samples for testing was 58.

From Fig. 4, we notice the classification accuracy reached 94.3%, 95.2%, and 96.2% for CNN-DT, CNN-SVM, and CNN-NB respectively for binary classification for DB1. For multiclassification, we notice in Fig. 5 that classification accuracy reached 92.5%, 90.6%, and 91.5% for CNN-DT, CNN-SVM, and CNN-NB respectively. In addition, the proposed stacking model achieved an accuracy of 97.1% for DB1 in binary classification while in multi-classification the accuracy reached 94.3%, much higher than any single classifier. From Fig. 6, we notice the classification accuracy reached 96.4%,96.4%, and 94.5% for CNN-DT, CNN-SVM, and CNN-NB respectively for binary classification to DB2. For multiclassification, we notice in Fig. 7 that classification accuracy reached 93.1%,94.8%, and 96.6% for CNN-DT, CNN-SVM, and CNN-NB respectively. In addition, the proposed stacking model achieved an accuracy of 98.2% for DB2 in binary classification while in multi-classification the accuracy reached 98.3%, much higher than any single classifier.



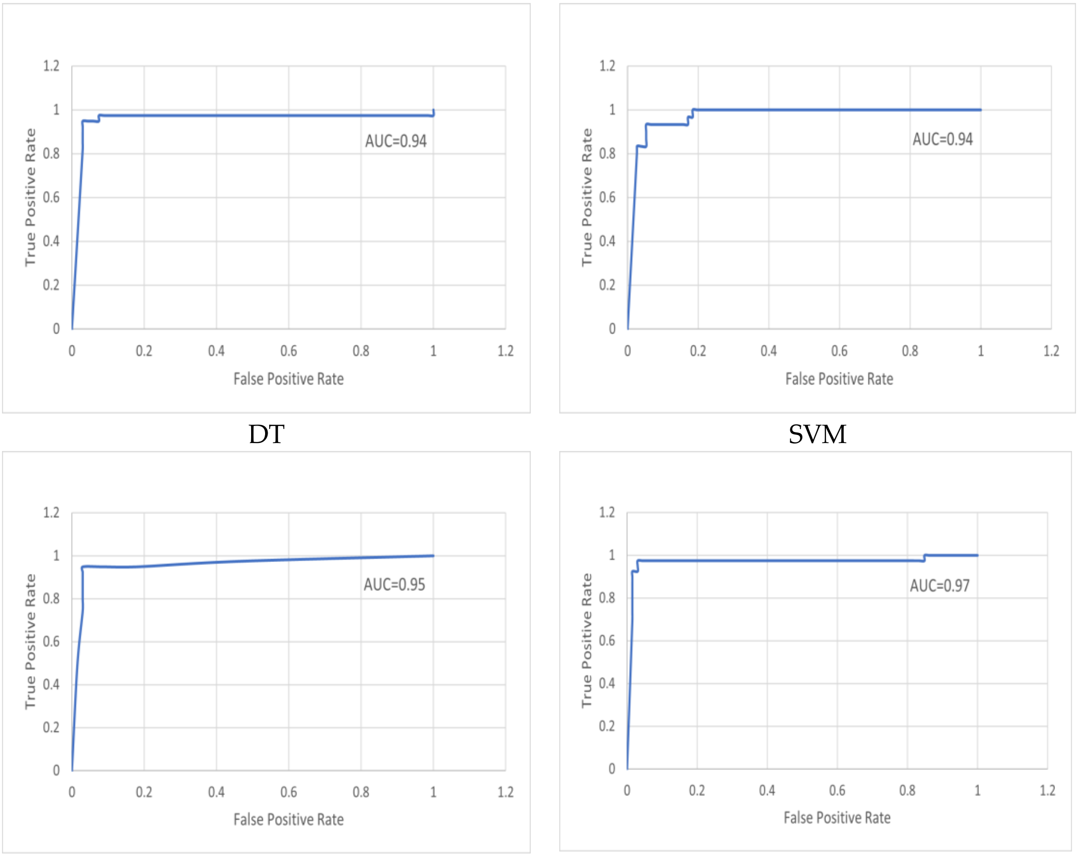
DT

SVM

NB

Stacking model

Fig. 8. ROC curves for DB1 in binary classification



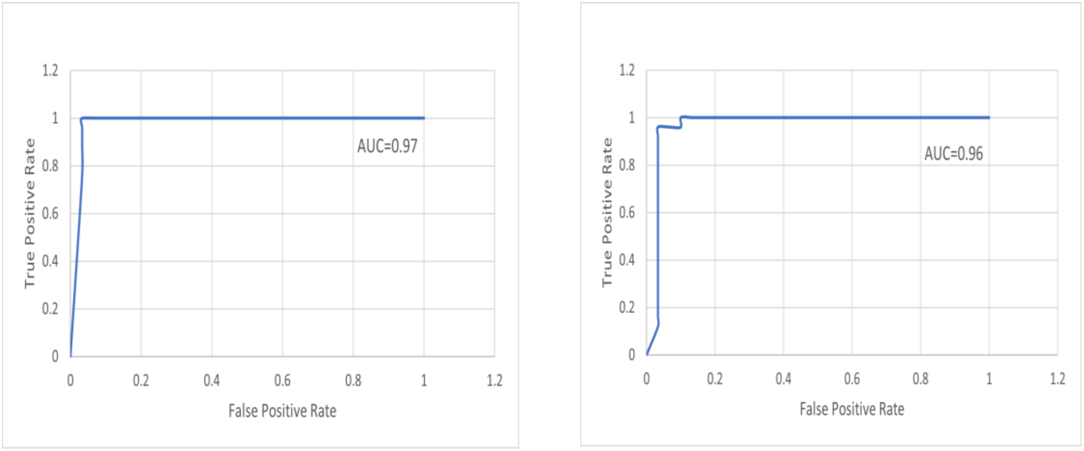
DT

SVM

NB

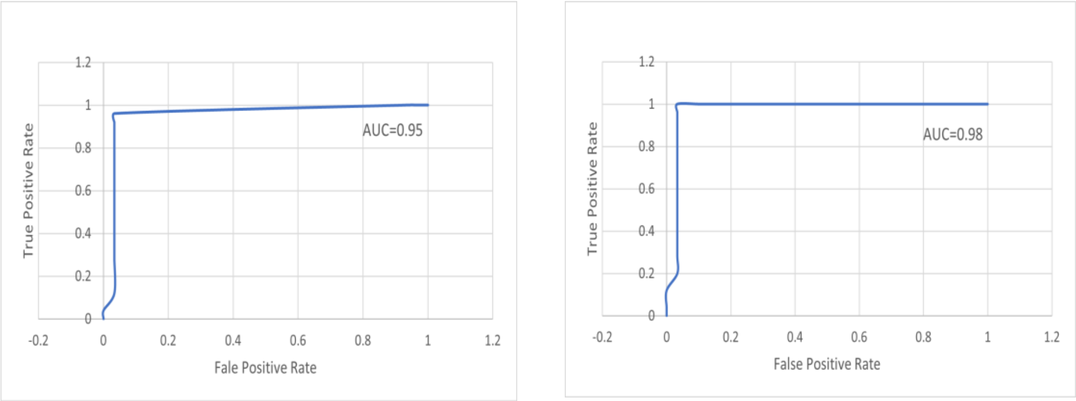
Stacking model

Fig. 9. ROC curves for DB1in multi-classification



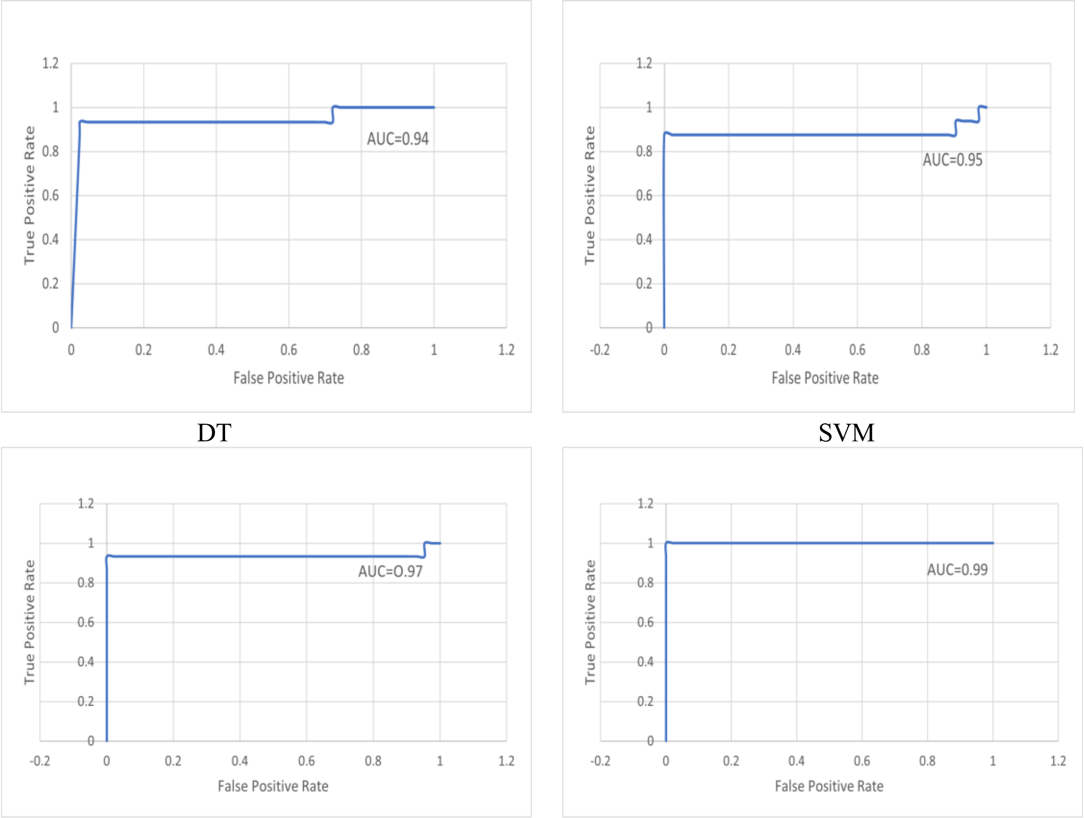
DT

SVM



NB Stacking model

Fig. 10. ROC curves for DB2 in binary classification



DT SVM

NB Stacking model

Fig. 11. ROC curves for DB2 in multi-classification

Figs. 8–11 illustrate the ROC curve for the proposed model performance evaluation for the two medical datasets in binary and multiclassification. The ROC curve denotes the tradeoff between true positive rate (TPR) and false positive rate

(FPR). The area under the curve is referred to as AUC and its value ranges between 0 and 1. The predictive value increases as this value approaches one and drops as it approaches zero. The curves in Figs. 8–11 are very close to the upper left corner, indicating high performance in classification for the two datasets (DB1, DB2).

Finally, we compare the classification performance of the suggested proposed stacking ensemble model with previous related work. The comparison of the performance is shown in Table 4.

Table 4

Comparison stacking ensemble model with previous related studies

Studies [Ref]	Method	Database	Results
[6]	Combining CNN models (Exception, InceptionResNet-V2, Inceptionv3, DenseNet201, and DenseNet121)	Acral melanoma (Binary classification)	Acc (97.9), Sn (97.8), Sp (97.5)
[7]	Combining CNN models, including (EfficientNetV2-B0, Efficient- NetV2-B1, EfficientNetV2-B2, Efficient NetV2-B3, EfficientNetV2-S and EfficientNetV2-M)	Chest X-ray CT images (multi-classification)	On the chest X-ray dataset, Acc (99.21%) On the chest CT dataset, Acc (99.81%)
[8]	Used VGG19 model to features extraction, Combining three single ML (SVM, ANN, LR) classifiers	Chest X-ray (multi-classification)	Acc (96.9%)
The Proposed Model	Used a new CNN structure to features extraction, Combining three single ML (SVM, DT, NB) classifiers	Chest X-ray Dermoscopy melanoma skin cancer (binary and multi-classification)	For DB1 in binary classification, Acc (97.1%), Sp (94.1%), Sn (100%), Precision (0.97), Recall (0.97), F1-score (0.97), AUC (0.98) For DB1 in multi-classification, Acc (94.3%), Sp (90.3%), Sn (96%), Pr (0.94), Recall (0.94), F1-score (0.94), AUC (0.97) values For DB2 in binary classification, Acc (98.2%), Sp (96.6%), Sn (100%), Pr (0.98), Recall (0.98), F1-score (0.98), AUC (0.98) For DB2 in multi-classification, Acc (98.3%), Sp (92.8%), Sn (100%), Pr (0.98), Recall (0.98), F1-score (0.98), AUC (0.99)

4. Discussion and Future directions

The results indicate that in this work, we have been able to build a new CNN model to extract deep features and build hybrid classifiers that combine the features extracted from the proposed CNN with ML classifiers. The feature extraction from the maxpool8 layer for the proposed CNN with ML classifiers achieved high performance. A stacking ensemble model was applied to obtain a new classifier to classify the medical datasets with high accuracy. The proposed model was evaluated using accuracy, sensitivity, specificity, precision, recall, F1 score, and AUC. We were able to acquire satisfactory outcomes for the proposed system in binary and multi-classification. The proposed stacking model achieved an accuracy of 97.1% for DB1, and 98.2% for DB2 in binary classification while in multi-classification the accuracy reached 94.3% for DB1 and 98.3% for DB2, much higher than any single classifier. In the future, we will try applying the proposed model to more medical datasets to diagnose more diseases and increase the generalizability of the model.

5. Conclusions

This work provides important insights into recent ML/DL approaches that are now used in illness research. The work outcomes indicate that the proposed stacking model is a quick and low-cost way with the potential to help in better and more efficient diagnosing. Most of the hybrid classifiers that were applied to the two chosen databases gave a good accuracy, reaching the highest accuracy (96.6%) to CNN-NB classifier in multi-classification to classifying the malignant melanoma skin cancer, and the lowest accuracy (90.6%) to CNN-SVM classifier in multi-classification to classifying the lung diseases. A stacked model has shown even better results, with accuracy for DB1 and DB2 binary classification reaching 97.1% and 98.2%, respectively, and for multi-classification reaching 94.3% and 98.3.%, respectively.

REFERENCES

1. *Khozeimeh, F., et al.*, RF-CNN-F: random forest with convolutional neural network features for coronary artery disease diagnosis based on cardiac magnetic resonance. Scientific Reports, 2022. **12**(1): p. 11178.
2. *Ismail, A., et al.*, Improving convolutional neural network (CNN) architecture (miniVGGNet) with batch normalization and learning rate decay factor for image classification. International Journal of Integrated Engineering, 2019. **11**(4).
3. *Alzubaidi, L., et al.*, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 2021. **8**(1): p. 53.
4. *Alsharif, R., et al.*, PneumoniaNet: Automated detection and classification of pediatric pneumonia using chest X-ray images and CNN approach. Electronics, 2021. **10**(23): p. 2949.
5. *Alif, M.A.R., S. Ahmed, and M.A. Hasan*. Isolated Bangla handwritten character recognition with convolutional neural network. in 2017 20th International conference of computer and information technology (ICCIT). 2017. IEEE.
6. *Raza, R., et al.*, Melanoma classification from dermoscopy images using ensemble of convolutional neural networks. Mathematics, 2022. **10**(1): p. 26.
7. *Huang, M.-L. and Y.-C. Liao*, Stacking Ensemble and ECA-EfficientNetV2 Convolutional Neural Networks on Classification of Multiple Chest Diseases Including COVID-19. Academic Radiology, 2022.
8. *Taspinar, Y.S., I. Cinar, and M. Koklu*, Classification by a stacking model using CNN features for COVID-19 infection diagnosis. Journal of X-ray science and technology, 2022. **30**(1): p. 73-88.
9. Data Availability: Data available for free at the Kaggle repository. https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database (accessed on 2 April 2023).

10. Data Availability: Data available for free at the Kaggle repository. https://www.kaggle.com/datasets/unaissait/curated-chest-xray-image-dataset-for-covid19 (accessed on 11 April 2023).
11. Dermatology Online Atlas: http://homepages.inf.ed.ac.uk/rbf/DERMOFIT/ (accessed on 12 April 2023).
12. Data Availability on : https://dermnetnz.org/images (accessed on 12 April 2023).
13. *Shabana, D.F., et al.*, an Image Enhancement Algorithm Using Gamma Correction By Swarm Optimization. Int Res J Eng Technol, 2020. **7**(9).
14. *Maharana, K., S. Mondal, and B. Nemade*, A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings, 2022.
15. *Islam, M.K., et al*. Melanoma Skin Lesions Classification using Deep Convolutional Neural Network with Transfer Learning. in 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA). 2021. IEEE.
16. *Zhao, H.-h. and H. Liu*, Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition. Granular Computing, 2020. **5**(3): p. 411-418.
17. *Yaqub, M., et al.*, State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images. Brain Sciences, 2020. **10**(7): p. 427.
18. *Koklu, M., et al.*, A CNN-SVM study based on selected deep features for grapevine leaves classification. Measurement, 2022. **188**: p. 110425.
19. *Ali, O.M.A., S.W. Kareem, and A.S. Mohammed*. Evaluation of Electrocardiogram Signals Classification Using CNN, SVM, and LSTM Algorithm: A review. in 2022 8th International Engineering Conference on Sustainable Technology and Development (IEC). 2022. IEEE.
20. *Kumari, L. and Y.P. Sai*, Classification of ECG beats using optimized decision tree and adaptive boosted optimized decision tree. Signal, Image and Video Processing, 2022. **16**(3): p. 695-703.
21. *Bansal, M., A. Goyal, and A. Choudhary*, A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. Decision Analytics Journal, 2022. **3**: p. 100071.
22. *Dhobale, N., S.S. Mulik, and S.P. Deshmukh*, Naïve Bayes and Bayes net classifier for fault diagnosis of end mill tool using wavelet analysis: A comparative study. Journal of Vibration Engineering & Technologies, 2022. **10**(5): p. 1721-1735.
23. *El Boujnouni, M*. A study and identification of COVID-19 viruses using N-grams with Naïve Bayes, K-nearest neighbors, artificial neural networks, decision tree and support vector machine. in 2022 International Conference on Intelligent Systems and Computer Vision (ISCV). 2022. IEEE.
24. *Kumar, M., et al.*, Optimized Stacking Ensemble Learning Model for Breast Cancer Detection and Classification Using Machine Learning. Sustainability, 2022. **14**(21): p. 13998.
25. *Ramadhan, A.A. and M. Baykara*, A Novel Approach to Detect COVID-19: Enhanced Deep Learning Models with Convolutional Neural Networks. Applied Sciences, 2022. **12**(18): p. 9325.