



Prediction of Skin Lesions Using Integrated Multi-Layered Network Model with Baseline Learning Approaches

Arpita Roy¹, Shaik Razia^{2,*}

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India

Emails: arpitaroy@kluniversity.in; skrazia@kluniversity.in

Abstract

Skin cancer has become more common in recent decades, raising severe concerns about world health. Creating an automated system to distinguish between benign and malignant images is challenging because of the subtle variations in how skin lesions appear. This study introduces Computer-Aided Diagnosis (CAD) system that offers high classification accuracy while maintaining low computing complexity for categorizing skin lesions. The system incorporates a pre-processing stage that uses morphological filtering to remove hair and artefacts. With the least minimum of human interaction, deep learning techniques are employed to separate skin lesions automatically. Image processing methods are currently being utilized to investigate the automated implementation of the prediction criteria for distinguishing between benign and malignant melanoma lesions. Various pre-trained convolutional neural networks (CNNs) with multi-layered (ML-CNN) are under examination for the classification of skin lesions as either benign or malignant. The best performance is achieved when RF, k-NN and XGBoost are combined, according to average 5-fold cross-validation findings. The outcomes also demonstrate that data augmentation works better than acquiring novel images for training and testing purposes. The experiment results show that the suggested diagnostic framework performs better than existing methods when used on actual clinical skin lesions, with accuracy at 97.5%, F1-score at 91.3%, precision at 96.5%, sensitivity at 89.2% and specificity at 96.7%. It also takes 2.6 seconds to complete with the MNIST dataset and accuracy at 98.2%, F1-score at 92.5%, precision at 98.4%, sensitivity at 92.3% and specificity of 97.2% with the ISIC dataset. This indicates that medical professionals can benefit from using the suggested framework to classify various skin lesions.

Keywords: skin lesions; prediction; accuracy; deep learning; malignant image

1. Introduction

For many years, both men and women have been developing skin cancer globally. There were around 8,790 new fatalities from melanoma and about 76,250 new cases of the disease. In India, predictions of new instances of non-melanoma skin cancer were recorded [1]. Many factors, including extended life expectancy, sun exposure, and early skin cancer detection, contribute to the growth of the disease. Dermoscopy is an efficient technique for detecting skin cancer in its prior stages, a non-invasive skin imaging technique [2]. The condition of the skin can significantly influence dermoscopic images of skin lesions. Moreover, several artefact sources, including skin tone, air bubbles, and hair, can make it difficult to distinguish between the rest of the healthy skin and skin lesions [3]. Dermatologists with expertise may find it difficult to distinguish benign skin lesions from malignant melanoma when analyzing many dermoscopic images, even though dermoscopic diagnostics effectively detect skin cancer [4]. Consequently, developing a non-invasive CAD system for categorizing skin lesions is imperative. The critical stages of a CAD system include image segmentation, image pre-processing, image classification and feature extraction. The classification performance of the entire CAD system

is significantly impacted by each stage, as stated in [5] – [6]. Therefore, practical algorithms should be used for high diagnosis performance at every stage.

Numerous research studies [7] have examined various machine-learning techniques for identifying different kinds of cancer. In most research, machine learning models were utilized to classify data based on manually extracted image features. For a successful diagnosis, most machine learning algorithms have a high computational time requirement, and their effectiveness depends on the characteristics chosen to describe the malignant region [8] best. CNNs and deep learning approaches have gained importance in the automated diagnosis of many cancer forms [9]. Significant progress has been made using deep learning in image classification challenges, particularly image categorization [10]. Data augmentation is commonly employed in image classification to address challenges such as limited data availability and memory requirements [11] – [12]. Transfer learning is helpful for classification problems when few datasets are available in imaging applications. Utilizing a pre-trained CNN architecture and fine-tuning its performance is more computationally efficient than training a CNN from the ground up, necessitating substantial data and computational resources. Various pre-trained CNN models, including Inception, AlexNet, DenseNet, and ResNet, have been previously trained on the provided dataset. Specific pre-trained CNNs were employed in various experiments, demonstrating promising capabilities in diagnosing skin cancer [13].

The "data augmentation" method can enhance the incoming data by producing data from the initially collected data. The lack of datasets on skin cancer can be addressed by employing image enhancement techniques. Many ways to enhance data include scaling, rotation, random cropping, and colour corrections. CNN architectures that have already been trained are widely used for data augmentation. To assess the effectiveness of skin lesion classification, a study trained three CNNs using thirteen data augmentation methods - DenseNet, ResNet, and Inception-v4, as outlined in [14]. As the results showed, training and testing stages can also benefit from using data augmentation. The author proposed an automated approach for classifying skin lesions using the Alex-Net CNN architecture. With the architecture weights adjusted and the datasets supplemented with fixed rotation angles, the average accuracy was 95%.

DL-based on NN ensemble was used in a study to identify skin lesions from dataset images. The melanoma categorization is described to distinguish between malignant and benign melanoma skin lesions. This was achieved by analyzing characteristics such as asymmetry, irregular borders, colour variations, size, and texture of the skin lesion and utilizing the SVM classifier. A deep CNN prediction model was used with a unique regularizer technique to classify skin lesions, achieving an accuracy of 97.49%. The provided model for skin lesion diagnosis based on Gabor wavelets was presented in a different paper [15]. This model used deep CNN models to assess the skin images after applying Gabor filters to separate certain features of the lesions. The classification of skin lesions was done through decision fusion using the sum rule, resulting in an accuracy rate of 83%. When separating benign lesions from malignant melanoma, most techniques now used to diagnose skin lesions in use provide feasible classification findings. However, developing a reliable automated system that can function in real-time with low technological needs and produce highly accurate diagnosis results remains a significant issue. This research aims to provide a proper diagnostic technique for categorizing skin lesions observed in dermoscopic images.

The proposed methodology utilizes advanced image processing techniques to enhance data augmentation, image segmentation, and pre-processing procedures. It also looks into various CNN architectures already trained for classifying skin lesions. The following briefly describes the primary contributions of this work: It is recommended to use a new automated real-time computer-aided design (CAD) system for the effective and high-performing classification of skin lesions with little complexity. This proposed method utilizes deep learning for segmenting lesions, novel pre-processing methods for identifying and removing hair, and automatic ABCD feature detection to distinguish between benign and malignant melanoma lesions. Pre-trained CNNs with multi-layered convolutional networks are being studied for their ability to classify skin lesions. An examination was undertaken to assess the effects of employing data augmentation on pre-trained CNN models with baseline classifiers like RF, k-NN, and XGBoost, which shows substantial performance metrics like sensitivity, accuracy, specificity, computational time, precision, and F1-score via 5-fold cross-validation. The outcomes demonstrated the benefits of using data augmentation to raise the accuracy of diagnosis. The effectiveness of the proposed framework was evaluated by a series of experiments demonstrated the recommended method gives better in sensitivity, classification accuracy, computing efficiency,

and specificity than other well-known advanced diagnostic systems. This suggests that the proposed framework is a helpful resource for physicians to employ when they need to classify skin lesions quickly.

The work is drafted as follows: section 2 depicts the detailed analysis of diverse approaches. The methodology is demonstrated in section 3, and graphical outcomes are shown in section 4. The conclusion is summarized in section 5.

2. Related works

As mentioned in [16], the COVID-19 pandemic and intelligent cities have made IoT and AI in healthcare a critical necessity in recent years. The term "cancer" describes a collection of illnesses brought on by uncontrollably high cell proliferation. Due to abnormal cell growth, cancer can infiltrate and metastasize to different parts of the body. There has been a notable rise in cancer-related deaths, with skin cancer recently becoming leading mortality, specifically with high levels of sun exposure. Identifying melanoma amid benign skin lesions is still a significant problem [17]. Different types of skin lesions exist, such as SCC, BCC, BKL, mild (nevus), and malignant (melanoma). The visual investigation is necessary because skin lesions are similar [18]. The naked eye is difficult to use and needs highly skilled dermatologist. Dermoscopy is a noninvasive imaging technology that improves the diagnosis of melanoma instead of taking a direct visual inspection of the human skin with the naked eye. However, the dermatologist's accuracy in identifying melanoma from dermoscopic images in routine clinical settings was less than 80% [19]. As a result, researchers have directed their efforts towards detecting melanoma to aid healthcare providers in distinguishing between melanoma and non-cancerous growths in the initial stages, thereby safeguarding the patient's well-being.

Two main methods exist for classifying images. The first method relies on manually created features that are extracted from images. The second method uses hierarchical learning of features using Deep Convolutional Neural Networks (DCNN). Ravi et al. list the advantages of DCNN compared to the initial approach for classifying medical images in [20]. The development of computer-aided diagnosis systems faces two significant obstacles. The need for more data and image processing techniques are these issues [21]. An accurate diagnosis is crucial for any early skin cancer detection and treatment. CNN-based methods significantly improve prediction accuracy, which is why several academics have recently focused. In particular, Researchers are conducting thorough investigations into DL algorithms for tasks such as detecting skin cancer and cardiovascular events due to their ability to create features and learn independently and autonomously. High performance using deep neural networks can be achieved at the cost of expanding, deepening, and raising the CNN's resolution, necessitating more parameters in the design and high processing power for training and testing [22] – [25].

Because of distortions, differences in image clarity, and weaker characteristics distinguishing between different kinds of cancer, classifying skin cancer is typically challenging. Patients report that clinical procedures for skin lesions are difficult and unpleasant, and they struggle to distinguish between different types of lesions accurately [26]. Computer vision methods show potential in addressing various difficulties associated with classifying skin lesions [27]. Creating features for dividing lesions into pathological and normal categories is the first step in CAD systems. CAD devices help detect skin cancer early, leading to a decrease in mortality rates. Identifying skin cancer remains challenging due to issues such as artefacts, variations in image quality, and the similarity between various lesions [28]. These challenges motivated the authors to develop a novel RDCNN. Using this architecture, we set a high classification rate and a dependable diagnostic method for early skin cancer detection [29]. However, the suggested approach is a superior resource for identifying and evaluating lesions to expedite treatment and increase the probability of survival [30].

3. Methodology

The proposed method adjusts the relationships within different granularity levels to achieve the skin lesion classification goal. Fig 1 illustrates the use of pre-trained multi-layered CNNs for feature extraction. The following section details each pre-trained multi-layered CNN baseline used for this purpose. These designs showed excellent performance in classifying images on the ConvNet-based skin lesions dataset, leading to their feature extraction application. Next, we contrast the four linear base classifiers to classify skin lesion images like RF, k-NN, and Xgboost. The general feature extraction framework is shown in Fig 1, along with a comparison of its phases using different classifiers. In the next phase, we will discuss methods of transfer learning for acquiring features like

edges, shapes, and colours. Using the multi-layered level architecture of deep learning with precisely calibrated multi-granularity levels can improve the accuracy, built upon pre-trained layers and supplemented with extra fully connected layers. We created a single network by combining several activation functions (AF) from various architectures into other architectures. Finally, we merge these trained networks with diverse optimization functions and architectures. The model got better outcomes than pre-trained, ultimately linked, additional connected layers and architectural aspects. Fig. 4 demonstrates the seamless integration of pre-trained multi-layered CNN with enhanced two-tier architecture by adding a fully connected layer. The feature maps from previous convolutional blocks serve as inputs for the dense blocks, which are then received by the layer. All successive layers of LI have been receiving its feature maps. After reading the state from the beginning layers, each layer executes to the next. It transfers information that needs to be preserved but replaces the form.

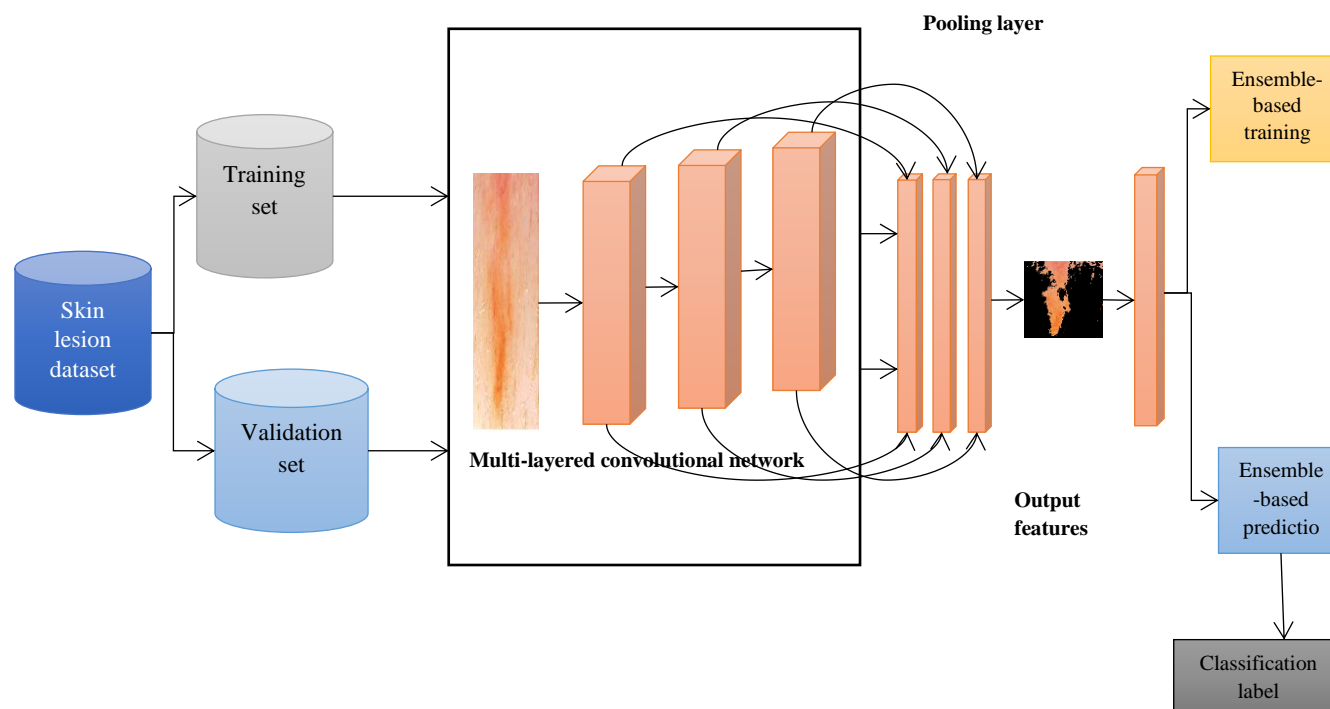


Figure 1: Structure of the proposed network model

Unlike ResNet, which sums features, the dense Net structure concatenates features to clearly distinguish between information held and information added to the network. We fixed the training settings for our model and the baseline models during the experiments. Additionally, MATLAB 2020a has been used to implement all of the models. Fig 1 shows an illustration of the ML-CNN third stage. The properties of a fully linked layer with x neurons are transferred to level ' i '. After being combined, the probability class values from the level $i - 1$ output layer are sent to the group. Fig 1 summarizes ML-CNN, which increases deep learning accuracy, depending on various CNN groupings. Some novel research is done on this technique. It is based on improving performance accuracy depending on computational power, batch size, training size, geometric filter and number of pre-trained networks. Numerous prior studies have been conducted to enhance accuracy performance. The deep learning architecture that is suggested in Fig 1 makes use of the features that are extracted by various networks. The idea is to use deep learning (ML_{ij}) to break out the network's parameter optimization into multi-granularity levels where performance is increased by individual training and optimizing each block. In summary, additional AFs are available to networks with varying design levels for subsequent feature extraction. Finding more useful AFs outside ReLU has become more and more motivating. One of the main factors leading to the primary activation problem in many studies is the vanishing gradient, which involves the AF in neural network computing. The hyper-parameters influencing the activation issue and the network's overall performance include the learning rate, filter size, quantization level position and regularization strength. The optimization of an AF with a narrow range is the motivating factor behind ML-CNN. Before predicting each label's class score, a linear model

performs an input function in the hidden layers. The following transformation is applied to the input vector x :

$$f(x) = w^T x + b \quad (1)$$

The input is denoted by x in this formula, the weights by w , and the biases by b . In this case, neural networks produce linear results. The AF is crucial in this process as it transforms linear results into non-linear outputs, enabling the network to recognize patterns in data for more advanced computations. The result is provided by:

$$y = (w_1x_1 + w_2x_2 + \dots + w_nx_n + b) \quad (2)$$

For multilayers, to get the final outcomes which is linear by default, outputs are fed into the layer afterwards. After applying the Activation Function (AF), which is given by, the linear model outputs are changed into non-linear outputs. The AFs are transfer functions.

$$y = \alpha(w_1x_1 + w_2x_2 + \dots + w_nx_n + b) \quad (3)$$

Here, the AF is α . Each input element undergoes a threshold operation where values below zero are set to zero from a more rapidly learning activation function (AF), which is the basis for ReLU.

$$f(x) = \max(0, x) = \begin{cases} 0 & x \leq b \\ n \left(\frac{x-b}{d-b} \right)^{n-1} & a < x < d \\ 1 & x \geq d \end{cases} \quad (4)$$

By driving the inputs that are less than zero to zero through rectifying them, this function solves the vanishing gradient issue that was present in earlier AF types. Different architectures and AFs can aid in extracting distinct features at varying granularities. In a two-level design, the ReLU is where the second level's input is provided by

$$f(x) = \begin{cases} 0 & x \leq b \\ n \left(\frac{x-b}{d-b} \right)^{n-1} & a < x < d \\ 1 & x \geq d \end{cases} \quad (5)$$

Where the parameters to be set are b, d , and n . The derivative function of this function can be expressed as follows:

$$f(x) = \begin{cases} 0 & x \leq b \\ n \left(\frac{x-b}{d-b} \right)^{n-1} & a < x < d \\ 0 & x \geq d \end{cases} \quad (6)$$

By adjusting the parameters, pre-trained ML-CNN can extract a wide range of information. ReLU AF layer modified by the proposed function is applied. Eq. (7) modifies the suggested function for results.

$$f(x) = \begin{cases} 0 & cx \leq a = b \\ d \left(\frac{cx-b}{d-b} \right)^n & b < cx < d \\ cx & cx \geq d \end{cases} \quad (7)$$

Eq. (7) derivative is represented in Eq. (8):

$$f(x) = \begin{cases} 0 & cx \leq a = b \\ dn \left(\frac{cx - b}{d - b} \right)^{n-1} & b < cx < d \\ c & cx \geq d \end{cases} \quad (8)$$

Here, 11 training-on-ConvNet inception blocks from baseline classifiers were used in this research. Starting with the first two blocks, the first experiment focuses on optimizing Incthe model. It conducts two different tests. The second model improves the pre-trained one using training data from these baseline models developed using ReLU as an AF. Moreover, transfer learning was applied from the pre-trained network after 30 epochs to enhance the model's convergence, maintaining the same weights for the first three layers. Mini-batch statistics were utilized consistently throughout the training process, regardless of whether the BN layer was frozen. The network used the frozen BN layers' statistics to make inferences. Additionally, modifications can be applied to the weight of the top layer, which is currently adjusted based on the variance and mean of the new dataset. Further, they got data scaled differently as original ImageNet dataset mean and variance will be used during inference. Due to IncceptionV3's poor validation accuracy, we cannot fine-tune VGG16 using the same procedure.

3.1. Ensemble model

In this segment, we aim to compare kNN, random forest, and XGBoost model-building procedures. We initiate the process by developing tuning wrappers that encapsulate each learner and its corresponding hyperparameter tuning approach. Subsequently, these wrapper learners are compiled and inputted into the benchmark() function. Recognizing the time-consuming nature of this procedure, we opt to define and implement a holdout cross-validation method for assessing the performance of each wrapper. However, the ideal choice would involve using k-fold or repeated k-fold techniques. Both regression and classification tasks can leverage k-NN and tree-based algorithms. In predicting a continuous outcome variable, kNN produces predictions by calculating the mean outcome values of the k-NN. On the other hand, tree-based algorithms determine predictions for a continuous outcome variable by considering the mean of the cases within the leaves of the trees. For assessing the adequacy of random forest and XGBoost ensembles in regression problems, the out-of-bag and Root Mean Square Error (RMSE) can still serve as valuable metrics. These metrics help determine whether there are enough trees in the ensembles to optimize predictive performance.

Algorithm 1: Prediction algorithm

Input: Trained ML-CNN network model

Output: allocate the proposed model for classification and prediction

BEGIN

Read the input dataset

T = no. of dataset images

For $C \leq$ no. of classes

C = Evaluate the class weights by partitioning the no. of images based on input classes by the total images;

END

//Load the multi-layered CNN model

Read the intermediate features from the layered network

Read the weights of every middle layer

For $I \leq$ analyze the last layers

If image ($C, 1$)

Analyze the AF of input images based on Eq. (7) and Eq. (8);

Perform ensemble model; //RF, k-NN and Xgboost to eliminate the weak learners

Repeat the process

End if

End for

4. Numerical results and discussion

Some 580 pictures (270 malignant and 310 benign) are available in the ISIC 2017 database. MNIST: We analyzed 350 photos (135 malignant and 215 benign) from the HAM10000 database to compute the efficacy of the anticipated skin lesion categorization. The information being analyzed includes various instances of patient gender, age, and abnormalities, as detailed in Tables 1 to 4. By creating four different images of each original image with varying rotation angles, 2320 images were generated from the initial 1400 images from the ISIC 2017 database within the MNIST: HAM10000 database. As a result, there are more images overall and more data accessible for testing and training. The key metrics used to evaluate picture categorization like false negatives (FN), false positives (FP), true positives (TP) and true negatives (TN). These metrics are essential for assessing classification performance, along with other measures such as processing time, recall (Se), precision (Pr), accuracy (ACC) and F1-score. Recall (Se) represents the proportion of actual malignant cases correctly identified, precision (Pr) indicates the percentage of determined malignant cases that are accurate, and ACC measures the accuracy of predictions. F1-score is a combined measure of precision and recall, while AUC evaluates performance across various categorization criteria. Here is how each of these measures is defined:

$$Acc = (TP + TN)/(TP + FP + TN + FN) \quad (9)$$

$$Pr = TP/(TP + FP) \quad (10)$$

$$Se = TP/(TP + FN) \quad (11)$$

$$F1 = 2 * (Pr * Se)/(Pr + se) \quad (12)$$

Each of the two primary partitions contains the 2320 enhanced images from ISIC 2017 and 1400 augmented images from MNIST: are being examined: 280 for testing and 1120 for validation and training employing the MNIST: 1856 for validation and training, ISIC 2017 database and HAM10000 database for testing purposes. To ensure unbiased results in diagnosing performance, the training and validation data were divided into five equal groups, with one group used for validation and the other four for training. This method is known as a 5-fold CV. Classification metrics for each fold were calculated for both datasets to evaluate the test dataset. All the pre-trained ML-CNN architectures are analyzed for accuracy in training validating and loss curves. The results indicate that the ML-CNN design works best. For instance, Fig 7 displays the training progress for the ML-CNN architecture about the number of epochs. Compared to previous CNN networks, the training of the pre-trained network has advanced and improved more quickly.

Furthermore, as Fig 4 and 5 illustrate, it outperforms other CNN architectures' validation accuracy. Comparable performance is achieved using the proposed architecture. Thus, a critical comparison of the pre-trained architectures is made to classify skin lesions under different parameters. Some existing VGG-16 and ResNet50 models were used to assess the proposed system's classification performance for ISIC 2017 database. The 5-fold benign-malignant classification outcomes and associated standard deviations are shown in Tables 1 to 4. By averaging 5-fold findings, the average classification outcomes for ACC, AUC, Se, F1-score, and Pr were determined. The average metrics are calculated for every pre-trained CNN architecture that was considered. It showed that the ResNet50 architecture outperformed the other models regarding AUC, ACC, Se, F1-score and Pr. For all CNN designs, equivalent results are obtained using the MNIST: HAM10000 database; Table 3 shows the results,

highlighting the architecture's better performance. To improve the anticipated system, we integrated kernel functionality of the classifier with various CNN models and evaluated their classification outcomes. Table 4 displays the SDs of the average 5-fold classification outcomes for ResNet50, VGG-16, InceptionV3, MobileNet, and ResNetX designs, both with and without pre-trained ML-CNN, on the MNIST and ISIC 2017: HAM10000 datasets, with and without data augmentation. Interestingly, employing a baseline classifier combined with the pre-trained ML-CNN architecture produced more significant results than using these designs alone.

Table 1: Accuracy prediction with two datasets

Methods	MNIST	ISIC
ResNet50	87	90
VGG-16	89	91
InceptionV3	82	94
MobileNet	85	96
ResNetX	86	96.2
VGG-16+SVM	88	91
Proposed	97.5	98.2

Table 2: Loss measurements comparison

Methods	Validation	Test	Testing loss
ResNet50	86.8	89.9	0.005
VGG-16	86.6	88.8	0.003
InceptionV3	86.9	89.7	0.009
MobileNet	96.8	87.8	0.006
ResNetX	86.4	87.3	0.0081
VGG-16+SVM	89.6	89.1	0.0060
Proposed	98.5	99.5	0.002

Table 3: Overall performance comparison with MNIST

Methods	Accur acy	Specific ity	Sensitiv ity	Precisi on	F1- sco re
ResNet5 0	87	88.2	88.4	74.2	81
VGG-16	89	90.1	70.4	75.6	75. 8
Inceptio nV3	82	94.2	70.78	70.5	75. 6

MobileNet	85	87.6	72.5	69.7	72.5
ResNetX	86	88.2	74.6	65.8	67.8
VGG-16+SV M	88	89.1	77.5	62.7	66.33
Proposed	97.5	96.7	89.2	96.5	91.3

Table 4: Overall performance comparison with ISIC

Methods	Accuracy	Specificity	Sensitivity	Precision	F1-score
ResNet50	90	90.1	89	76.2	83
VGG-16	91	92	72.4	76.6	77
InceptionV3	94	94	73.78	72.5	79
MobileNet	96	89	74.5	70.7	73
ResNetX	96.2	90	75.6	61.8	69
VGG-16+SV M	91	90.1	79.5	65.7	68
Proposed	98.2	97.2	92.3	98.4	92.5

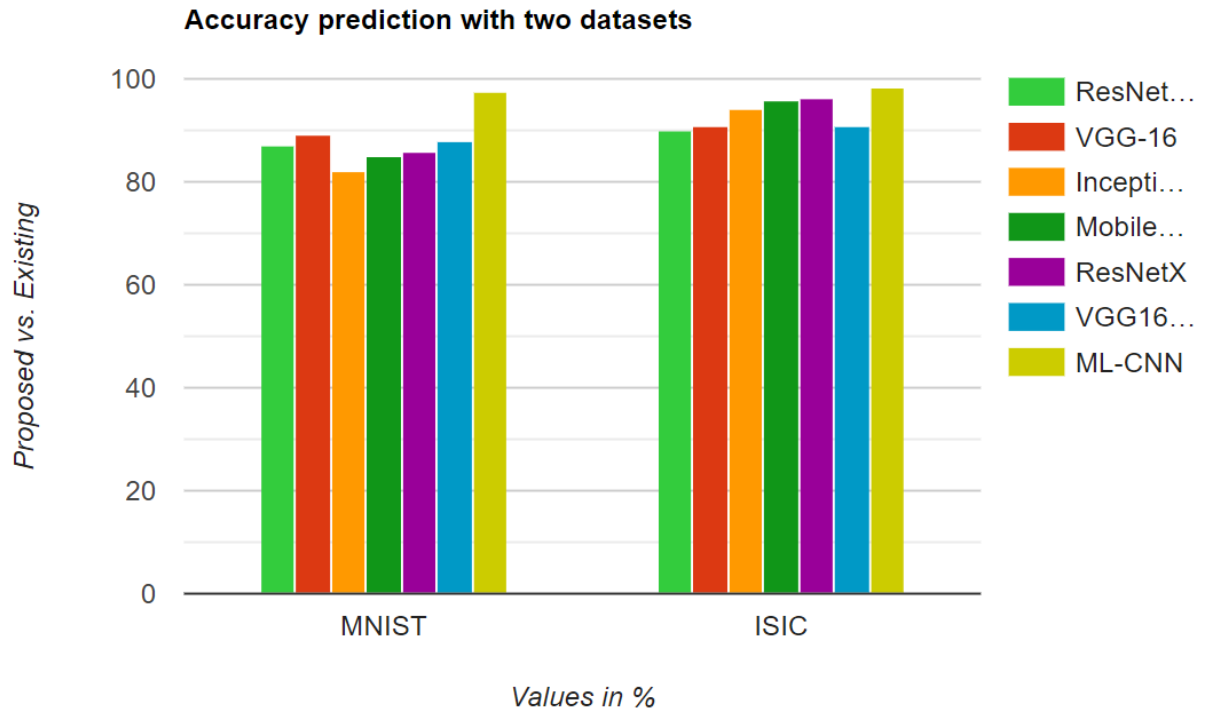


Figure 2: Accuracy prediction with two datasets

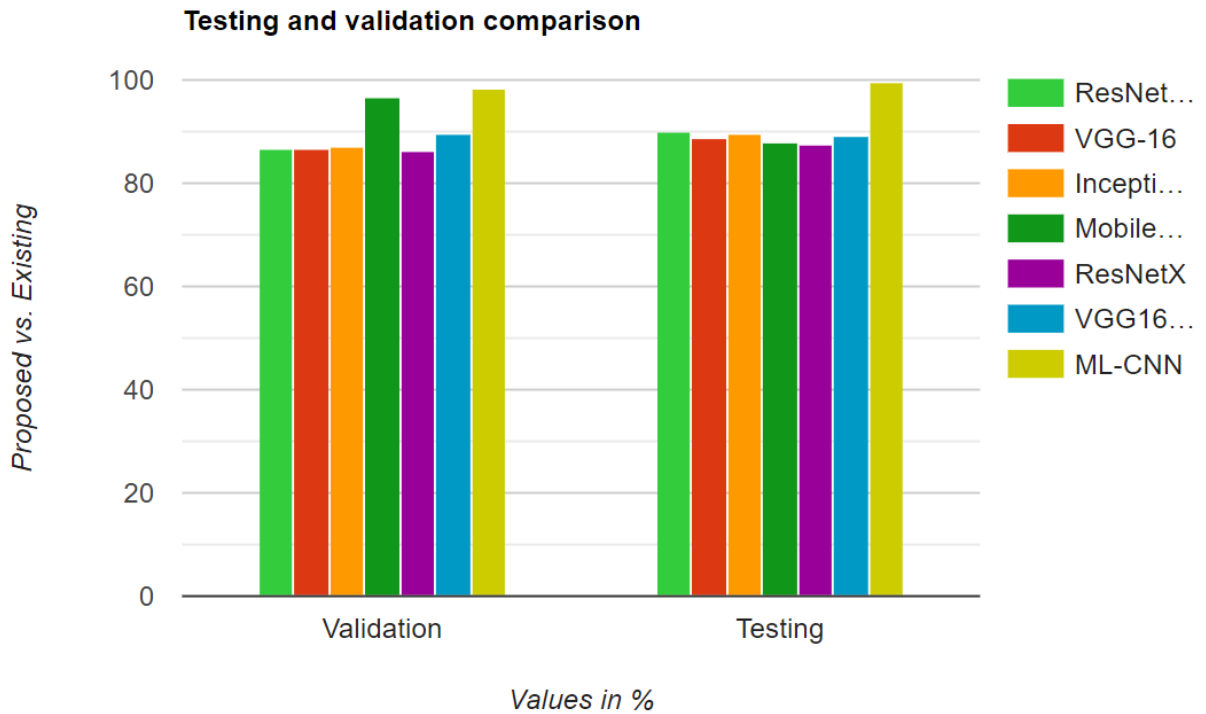


Figure 3: Testing and validation comparison

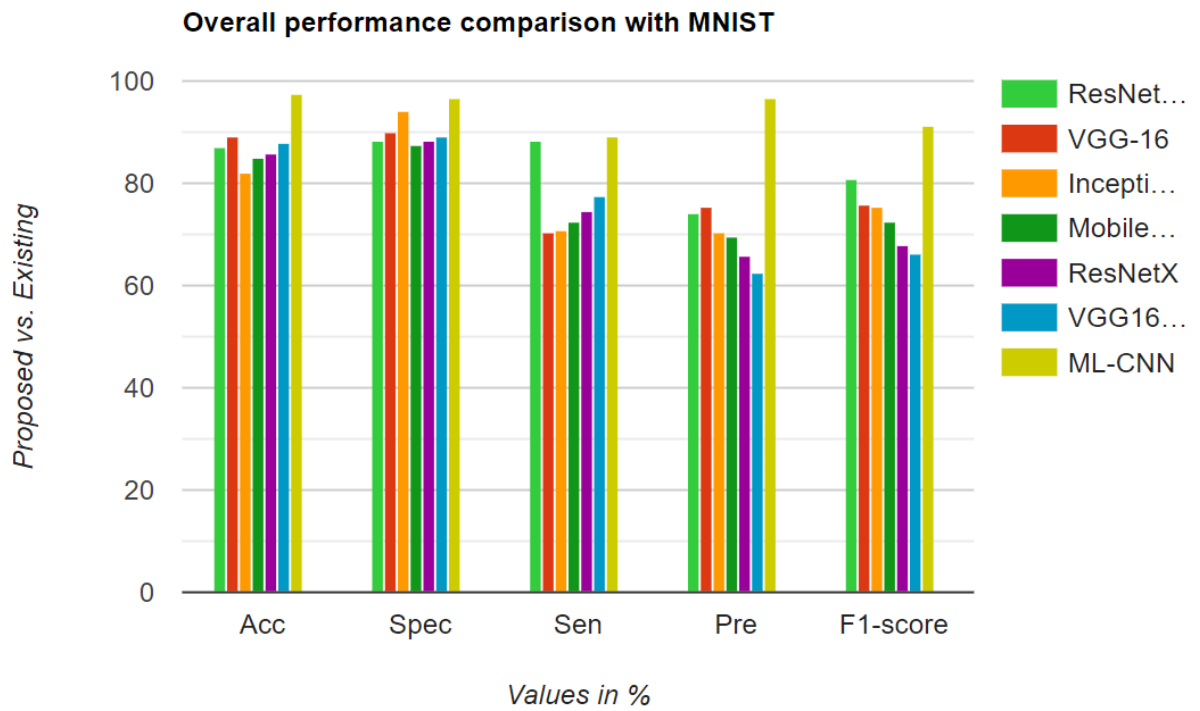


Figure 4: Overall performance comparison with MNIST

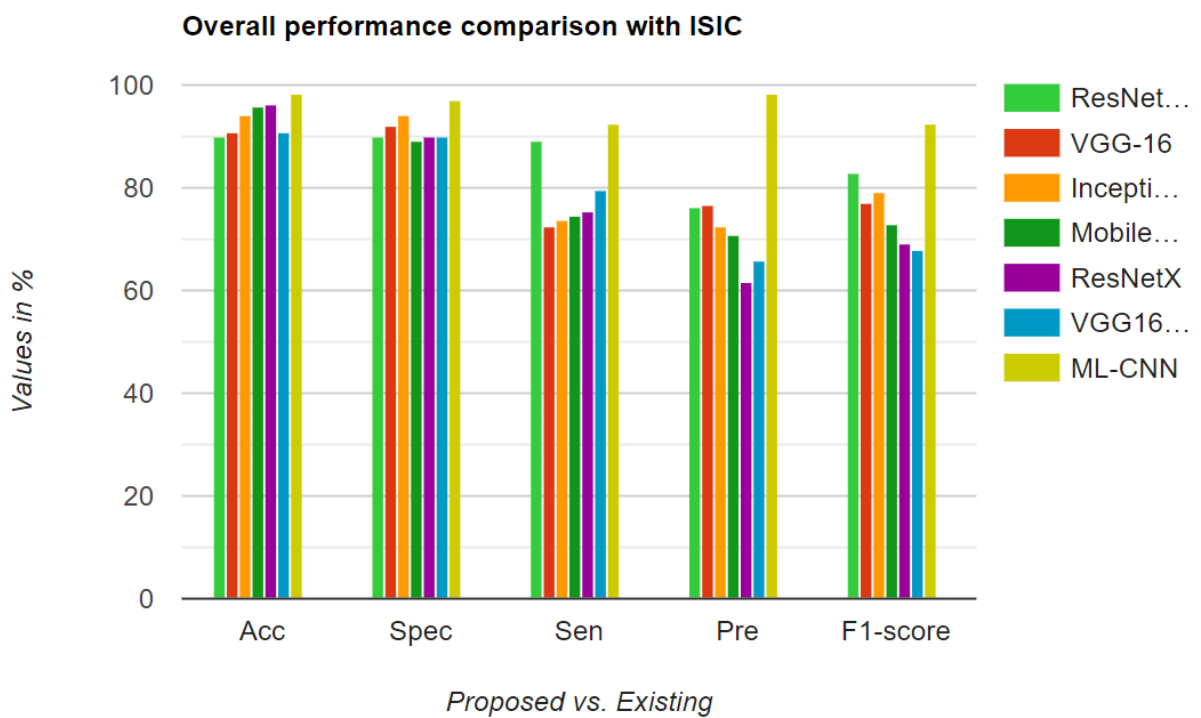


Figure 5: Overall performance comparison with ISIC

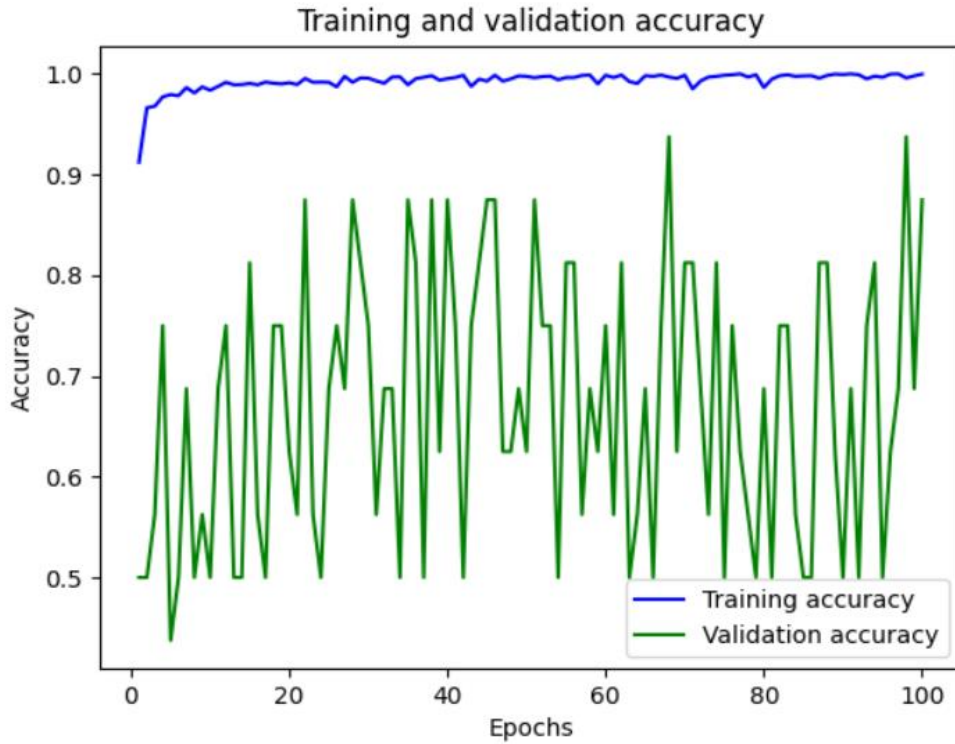


Figure 6: Accuracy comparison

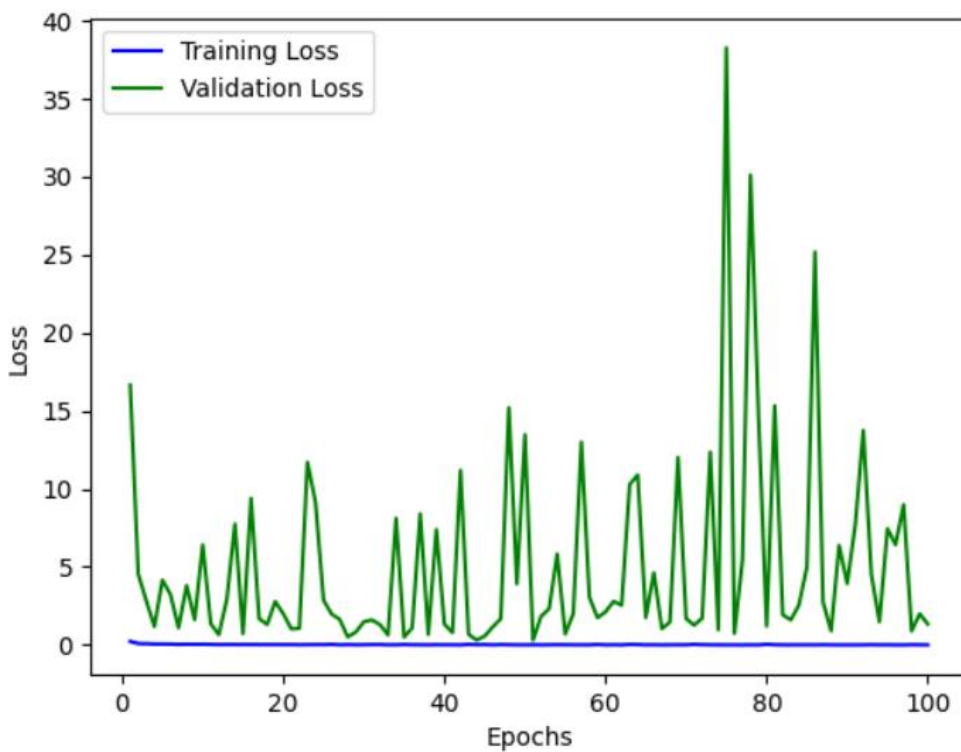


Figure 7: Loss prediction

Additionally, utilizing the ISIC 2017 database, the ML-CNN architecture with baseline classifiers performs better than the VGG-16 design with SVM regarding average AUC, ACC, Pr, F1-score, and

Se. The existing ResNet50 architecture with SVM achieves the same excellent performance for the MNIST: HAM10000 database. Compared to ML-CNN with or without a baseline classifier, reductions in standard deviation values indicate that the proposed architecture with a baseline classifier consistently and accurately classifies skin lesions. This is an important finding. When the classification results for both databases are compared with and without data augmentation, it becomes evident that more data can be used for training and testing thanks to data augmentation, which improves the classification results overall [31] – [32].

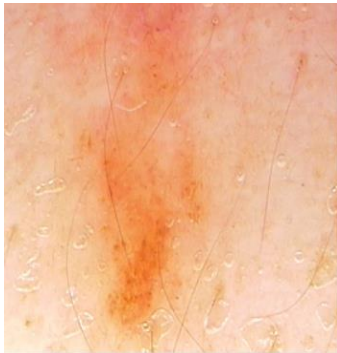


Figure 8a: Input image

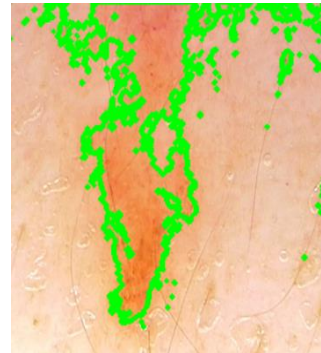


Figure 8b: Output image

Table 3 shows how the suggested pre-trained ML-CNN architecture for skin lesion prediction compares to other existing systems. In contrast to other techniques, the proposed system's efficacy is confirmed by displaying the 5-fold CV average classification results. Table 4 shows that compared to most other procedures, we examined a more significant number of lesion images. The suggested method performs superior to alternative diagnosis approaches concerning ACC (98.97% for HAM10000 and 99.87% for ISIC 2017) and AUC (97.91% for HAM10000 and 99.52% for ISIC 2017), respectively. Compared to alternative methods, the suggested approach achieves the highest sensitivity, demonstrating its efficacy in classifying skin cancer patients. For dermatologists and other healthcare professionals, it should be noted that evaluation parameters such as sensitivity are more important than others because low sensitivity or a high false negative rate (malignant results presented as benign) could prove fatal. Because so few incorrectly identified malignant lesions are in this study, the suggested technique yields higher sensitivity results. Regarding diagnosis findings, Table 3 demonstrates how the proposed CAD system performs better than the current methods.

The recommended ML-CNN architecture with baseline classifier for skin lesion classification uses a Kaggle Notebook GPU cloud with thirteen gigabytes of RAM and two CPU cores. The total average compute time is approximately 2.5 seconds. Compared to other studies in the same field, most did not include any analysis of computing time. The model was trained in roughly 30 hours using the NVIDIA GTX Titan XP GPU, with each patch requiring an average of 0.02 seconds to test. The computation time was disclosed. Although existing approaches have a shorter calculation time than the suggested method, they need better classification performance and high hardware requirements. With minimal hardware specifications, the recommended method uses poor computing efficiency to achieve good diagnosis performance. During the testing phase, the proposed framework demonstrated better computational efficiency than the deep learning model in a shorter time (2.5 seconds versus 6 seconds). Despite the DL model achieving a high accuracy of 95%, the outcomes highlight the potential of the anticipated architecture to enhance the real-time system performance in classifying skin lesions. The suggested skin lesion classification method performs well, although a few issues may be investigated further. First, other pre-trained ML-CNN architectures may be analyzed to include more modern and complex pre-trained models. Second, there may be certain practical drawbacks to extracting comprehensive features, such as misclassifying melanoma with uniform colour and regular shape. New deep-learning approaches may be used in future research on this topic to enhance further the removal of artefacts and the detection of structures. Future research may also aim to build and construct the hardware for the suggested CAD system employing high-dimensional chaotic circuits. These issues are obstacles that the upcoming segment will investigate and report on.

5. Conclusion

This study suggests a revolutionary computer-aided design (CAD) system with low software and hardware requirements to classify skin lesions automatically. The proposed methodology examines novel pre-processing methods for identifying benign lesions from malignant melanoma, enhancing image data and pre-trained ML-CNN architectures trained for categorization. The pre-trained ML-CNN is evaluated for their effectiveness in categorizing skin lesions with baseline classifiers like RF, k-NN, and Xgboost. The suggested approach's efficacy is assessed by examining authentic skin images encompassing various cases from diverse datasets. For all pre-trained architectures under examination, the effectiveness of benign-malignant categorization is examined through data augmentation to resolve the data restriction issue. Multiple measures are used to determine the classification performance: The experiment results show that using ML-CNN architecture along with data augmentation and baseline classifiers, the model leads to the exact classification of skin lesions. This approach also requires minimal computational time. The model gives an accuracy of 97.5%, F1-score of 91.3%, precision of 96.5%, sensitivity of 89.2% and specificity of 96.7%. It also takes 2.6 seconds to complete with the MNIST dataset and accuracy at 98.2%, F1-score at 92.5%, precision at 98.4%, sensitivity at 92.3% and specificity of 97.2% with the ISIC dataset. The results showed how augmenting data might help diagnose patients more accurately. The suggested method for classifying skin lesions is compared to current systems, with the comparison indicating that the recommended method outperforms existing ones in terms of classification performance. This illustrates how different skin lesions can be categorized using the suggested framework by dermatologists and professionals with less training.

References

- [1] Yang Y, Wang W, Yin Z, Xu R, Zhou X, Kumar N, Alazab M, Gadekallu TR. Mixed game-based AI optimization for combating COVID-19 with AI bots. *IEEE J Sel Areas Commun.* 2022;40(11):3122–38.
- [2] Hwang SM, Pan HC, Hwang MK, Kim MW, Lee JS. Malignant skin tumours are misdiagnosed as benign skin lesions. *Arch Craniofac Surg.* 2016;17(2):86–9
- [3] Eltoukhy MM, Hosny KM, Kassem MA. Classification of multiclass histopathological breast images using residual deep learning. *Comput Intell Neurosci.* 2022;10(2022):9086060
- [4] Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, Yang GZ. Deep learning for health informatics. *IEEE J Biomed Health Inform.* 2017;21(1):4–21.
- [5] Catarina B, Emre Celebi M, Marques JS. Development of a clinically oriented system for melanoma diagnosis. *Pattern Recogn.* 2017;69:270–85
- [6] Zortea M, Flores K, Scharcanski J. A simple weighted thresholding method for segmenting pigmented skin lesions in macroscopic images. *Pattern Recogn.* 2017;64(8):92–104.
- [7] Pandya S, Gadekallu TR, Reddy PK, Wang W, Alazab M. InfusedHeart: a novel knowledge-infused learning framework for diagnosis of cardiovascular events. *IEEE Trans Comput Soc Syst.* 2022.
- [8] Hu Z, Tang J, Wang Z, Zhang K, Sun Q. Deep learning for image-based cancer detection and diagnosis—a survey. *Pattern Recogn.* 2018;83:134–49.
- [9] Bi L, Kim J, Ahn E, Kumar A, Dagan F, Fulham M. Step-wise integration of deep class-specific learning for dermoscopic image segmentation. *Pattern Recogn.* 2019;85:78–89
- [10] Hosny KM, Kassem MA. Refined residual deep convolutional network for skin lesion classification. *J Digit Imaging.* 2022;35(2):258–80.
- [11] Ozkan IA, Koklu M. Skin lesion classification using machine learning algorithms. *Int J Intell Syst Appl Eng.* 2017;5(4):285–9.
- [12] Singh R, Bharti V, Purohit V, Kumar A, Singh AK, Singh SK. MetaMed: few-shot medical image classification using gradient-based meta-learning. *Pattern Recognit.* 2021;120:108111.
- [13] Fu Z, An J, Yang Q, Yuan H, Sun Y, Ebrahimian H. Skin cancer detection using Kernel Fuzzy C-means and developed red fox optimization algorithm. *Biomed Signal Proc Cont.* 2022;71:103160
- [14] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115–8
- [15] Yu L, Chen H, Dou Q, Qin J, Heng PA. Automated melanoma recognition in dermoscopy images via intense residual networks. *IEEE Trans Med Imaging.* 2017;36(4):994–1004.
- [16] Amin J, Sharif A, Gul N, Anjum MA, Nisar MW, Azam F, Bukhari SA. C, Integrated design of deep features fusion for skin cancer localization and classification. *Pattern Recogn Lett.* 2020;131:63–70.
- [17] Mahbod A, Schaefer G, Ellinger I, Ecker R, Pitiot A, Wang C. Fusing fine-tuned deep features for skin lesion classification. *Comput Med Imaging Graph.* 2019;71:19–29

- [18] Hosny KM, Kassem MA, Fouad MM. Classification of skin lesions using transfer learning and augmentation with Alexie. *PLoS ONE*. 2019;14(5):e0217293
- [19] Hosny KM, Kassem MA, Fouad MM. Classification of skin lesions into seven classes using transfer learning with Alexie. *J Digit Imaging*. 2020;33(5):1325–34
- [20] Hosny KM, Kassem MA, Fouad MM. Skin melanoma classification using ROI and data augmentation with deep convolutional neural networks. *Multimedia Tools Appl*. 2020;79(33):24029–55.
- [21] Yu Z, Jiang X, Zhou F, Qin J, Ni D, Chen S, Lei B, Wang T. Melanoma recognition in dermoscopy images via aggregated deep convolutional features. *IEEE Trans Biomed Eng*. 2019;66(4):1006–16
- [22] Majtner T, Yildirim-Yayilgan S, Hardeberg JY. Optimized deep learning features for improved melanoma detection. *Multimedia Tools Appl*. 2019;78:11883–903.
- [23] Albert BA. Deep learning from limited training data: novel segmentation and ensemble algorithms applied to automatic melanoma diagnosis. *IEEE Access*. 2020;8:31254–69
- [24] Harangi B, Baran A, Hajdu A. Assisted deep learning framework for multiclass skin lesion classification considering a binary classification support. *Biomed Signal Proc Cont*. 2020;62:102041.
- [25] Benyahia S, Meftah B, Lézoray O. Multi-features extraction based on deep learning for skin lesion classification. *Tissue Cell*. 2022;74:101701.
- [26] Sathya Preiya V, Kumar VDA. Deep Learning-Based Classification and Feature Extraction for Predicting Pathogenesis of Foot Ulcers in Patients with Diabetes. *Diagnostics*. 2023; 13(12):1983. <https://doi.org/10.3390/diagnostics13121983>.
- [27] Balakrishnan C, Ambeth Kumar VD. IoT-Enabled Classification of Echocardiogram Images for Cardiovascular Disease Risk Prediction with Pre-Trained Recurrent Convolutional Neural Networks. *Diagnostics (Basel)*. 2023 Feb 18;13(4):775. doi: 10.3390/diagnostics13040775. PMID: 36832263; PMCID: PMC9955174.
- [28] Hemamalini, Selvamani, and Visvam Devadoss Ambeth Kumar. 2022. "Outlier Based Skimpy Regularization Fuzzy Clustering Algorithm for Diabetic Retinopathy Image Segmentation" *Symmetry* 14, no. 12: 2512. <https://doi.org/10.3390/sym14122512>.
- [29] Kumar, V.D.A., Sharmila, S., Kumar, A. et al. A novel solution for finding postpartum haemorrhage using fuzzy neural techniques. *Neural Comput & Applic* 35, 23683–23696 (2023). <https://doi.org/10.1007/s00521-020-05683-z>
- [30] V. D. A. Kumar, M. Raghuraman, A. Kumar, M. Rashid, S. Hakak and M. P. K. Reddy, "Green-Tech CAV: Next Generation Computing for Traffic Sign and Obstacle Detection in Connected and Autonomous Vehicles," in *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 3, pp. 1307-1315, Sept. 2022, doi: 10.1109/TGCN.2022.3162698.
- [31] Ahmed Abdelaziz, Alia N. Mahmoud, Skin Cancer Detection Using Deep Learning and Artificial Intelligence: Incorporated model of deep features fusion, *Journal of Fusion: Practice and Applications*, Vol. 8, No. 2, (2022) : 08-15 (Doi : <https://doi.org/10.54216/FPA.080201>)
- [32] S. Zhang, Y. Li, and X. Wang, "Deep Learning for Skin Lesion Analysis: A Comprehensive Review," *Journal of Medical Systems*, vol. 46, no. 3, pp. 1-12, Mar. 2022. DOI: 10.1007/s10916-022-01847-3
- [33] Dai D, Dong C, Xu S, Yan Q, Li Z, Zhang C, Luo N. Ms RED: a novel multi-scale residual encoding and decoding network for skin lesion segmentation. *Med Image Anal*. 2022;75:102293
- [34] Tang P, Yan X, Liang Q, Zhang D. AFLN-DGCL: adaptive feature learning network with difficulty-guided curriculum learning for skin lesion segmentation. *Appl Soft Comput*. 2021;110:107656.
- [35] Mahbod A, Tschandl P, Langs G, Ecker R, Ellinger I. The effects of skin lesion segmentation on the performance of dermatoscopic image classification. *Comput Met Progr Biomed*. 2020;197:105725
- [36] Shetty B, Fernandes R, Rodrigues AP, Chengoden R, Bhattacharya S, Lakshmana K. Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Sci Rep*. 2022;12(1):18134
- [37] Ahmed Abdelhafeez, Hoda K. Mohamed, Skin Cancer Detection using Neutrosophic c-means and Fuzzy c-means Clustering Algorithms, *Journal of Intelligent Systems and Internet of Things*, Vol. 8, No. 1, (2023) : 33-42 (Doi : <https://doi.org/10.54216/JISIoT.080103>)

- [38] P. Sherubha, P Amudhavalli, SP Sasirekha, “Clone attack detection using random forest and multi-objective cuckoo search classification”, International Conference on Communication and Signal Processing (ICCSP), pp. 0450-0454, 2019.
- [39] S. Dinesh, K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji, “Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms”, Hindawi Journal of Healthcare Engineering, Volume 2022.