

## Research Article

# A Computer-Aided Diagnosis System Using Deep Learning for Multiclass Skin Lesion Classification

**Mehak Arshad,<sup>1</sup> Muhammad Attique Khan <sup>1</sup> Usman Tariq <sup>2</sup> Ammar Armghan,<sup>3</sup> Fayadh Alenezi <sup>3</sup> Muhammad Younus Javed,<sup>1</sup> Shabnam Mohamed Aslam <sup>4</sup> and Seifedine Kadry<sup>5</sup>**

<sup>1</sup>Department of Computer Science, HITEC University Taxila, Taxila, Pakistan

<sup>2</sup>College of Computer Engineering and Science, Prince Sattam Bin Abdulaziz University, Al-Kharaj, Saudi Arabia

<sup>3</sup>Department of Electrical Engineering, Jouf University, Sakaka 75471, Saudi Arabia

<sup>4</sup>Department of Information Technology, College of Computer and Information Sciences, Majmaah University, Al-Majmaah 11952, Saudi Arabia

<sup>5</sup>Faculty of Applied Computing and Technology, Noroff University College, Kristiansand, Norway

Correspondence should be addressed to Muhammad Attique Khan; [attique@ciitwah.edu.pk](mailto:attique@ciitwah.edu.pk) and Shabnam Mohamed Aslam; [s.aslam@mu.edu.sa](mailto:s.aslam@mu.edu.sa)

Received 7 August 2021; Revised 28 October 2021; Accepted 10 November 2021; Published 6 December 2021

Academic Editor: Ahmed A. Abd El-Latif

Copyright © 2021 Mehak Arshad et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the USA, each year, almost 5.4 million people are diagnosed with skin cancer. Melanoma is one of the most dangerous types of skin cancer, and its survival rate is 5%. The development of skin cancer has risen over the last couple of years. Early identification of skin cancer can help reduce the human mortality rate. Dermoscopy is a technology used for the acquisition of skin images. However, the manual inspection process consumes more time and required much cost. The recent development in the area of deep learning showed significant performance for classification tasks. In this research work, a new automated framework is proposed for multiclass skin lesion classification. The proposed framework consists of a series of steps. In the first step, augmentation is performed. For the augmentation process, three operations are performed: rotate 90, right-left flip, and up and down flip. In the second step, deep models are fine-tuned. Two models are opted, such as ResNet-50 and ResNet-101, and updated their layers. In the third step, transfer learning is applied to train both fine-tuned deep models on augmented datasets. In the succeeding stage, features are extracted and performed fusion using a modified serial-based approach. Finally, the fused vector is further enhanced by selecting the best features using the skewness-controlled SVR approach. The final selected features are classified using several machine learning algorithms and selected based on the accuracy value. In the experimental process, the augmented HAM10000 dataset is used and achieved an accuracy of 91.7%. Moreover, the performance of the augmented dataset is better as compared to the original imbalanced dataset. In addition, the proposed method is compared with some recent studies and shows improved performance.

## 1. Introduction

The development of skin cancer has risen throughout the previous decade [1]. Ultraviolet rays in the sun damage the skin over time and cause cancer cells to develop [2]. Usually, such conditions have hidden risks that lead to a lack of confidence and psychological distress in humans and to skin cancer risks. Several types of skin cancer exist,

including basal cells, melanoma, actinic keratosis, and squamous cell carcinoma [3]. The squamous cell carcinoma is contrasted against actinic keratosis (solar keratosis) [4]. Each year, the incidence rate of both melanoma and nonmelanoma continues to grow [2]. The deadliest form of skin cancer is melanoma and quickly spread to other body parts due to the malignancy of neural crest neoplasia of melanocytes [5].

In the United States, almost 5.4 million new cases of skin cancer are detected each year. Due to melanoma, more than 10,000 deaths are registered every year in the USA [6]. In the USA, 104,350 new cases of skin cancers were diagnosed during the year 2019, where the numbers of deaths were 7230. In the year 2020, 196,060 Americans are diagnosed with melanoma. According to these facts, melanoma cases are increasing approximately 2% [7]. Recently, in the year 2021, 207.39 K peoples are diagnosed with skin cancer whereas the numbers of deaths are 70.18 K. According to the facts, when the lesion is detected earlier, the survival rate increases approximately 98% [7]. The summary of diagnoses and deaths due to skin cancer is illustrated in Figure 1.

Dermatologists diagnose malignant lesions via a dermoscopic visual examination technique [8]. Diagnosis of skin cancer using dermoscopy is challenging due to various textures and wounds [9]. However, the manual inspection of dermoscopic images makes it difficult to diagnose skin cancer with better accuracy. The accuracy of the lesion diagnosis depends on the dermatologist's experience [9]. Few other techniques are available for diagnosing skin cancer, such as biopsy [7] and macroscopic [10]. Due to the complex nature of skin lesions, the clinical methods need more attention and time [11, 12].

The computer-based detection (CAD) techniques are introduced by several researchers in medical imaging [7, 13]. They introduced CAD techniques for several cancers such as skin cancer [14], brain tumor [15, 16], lung cancer [17, 18], COVID-19 [19, 20], and more [21–23]. A simple CAD technique consists of four key steps such as preprocessing of input images, detection of infected parts, features extraction, and classification. A computerized method can be helpful as a second opinion for dermatologists to verify the manual diagnosis results [8]. The advancement in machine learning, like deep learning, has shown much achievement in medical imaging in the last couple of years. Convolutional Neural Network (CNN) is a form of deep learning used for automated features extraction [6]. A convolutional neural network is a computer vision technique that automatically distinguishes and recognizes images' features [24]. Due to its high accuracy, it has attracted interest in medical image processing, agriculture, biometric, and surveillance, to name a few. A simple CNN typically entails a series of layers such as a convolutional layer, ReLU layer [25], normalization layer, pooling layer [26], fully connected layer, and Softmax layer [27]. In many techniques, researchers used some pretrained deep learning models for the classification tasks. A few publically available pretrained deep learning models are AlexNet, VGG, GoogleNet, InceptionV3, and ResNet to name a few [28]. They used these models through transfer learning [7]. Few researchers used feature selection and fusion techniques to improve recognition accuracy [29, 30].

The computer-aided diagnostic systems can allow dermatologists and physicians to make decisions, decrease diagnostic costs, and increase diagnostics reliability [31]. An automated skin lesion identification mechanism is challenging due to several challenges such as changing appearance and imbalanced datasets to name a few [32]. Chaturvedi et al. [6] presented an automated framework for

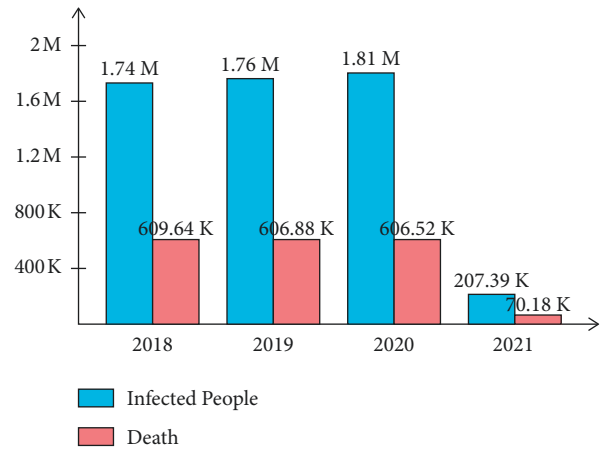


FIGURE 1: Graph of infected and death cases of skin lesion.

multiclass skin cancer classification. Five steps were involved in the presented method: dataset preprocessing, classification models (pretrained deep learning), fine-tuning, feature extraction, and performance evaluation. During the evaluation process, it is noted that the maximum accuracy of 93.20% was achieved for an individual model (ResNet-101), whereas a complete precision of 92.83% was performed on the ensemble model (InceptionResNetV2 + ResNet-101). In the end, they concluded that the training of deep learning models with the best setup of hyperparameters could be performed better than even ensemble models. Hsin et al. [33] presented the automatic lightweight diagnostic algorithm for skin lesion diagnosis. The presented algorithm was more reliable, feasible, and easy to use. For the experimental process, the HAM10000 dataset was used and achieved an accuracy of 85.8%. Besides, this method was tested on a five-class KCGMH dataset and achieved an accuracy of 89.5%. Kumar et al. [9] presented an automated electronic device. They considered numerous challenges such as skin cancer injuries, skin colors, asymmetric skin, and the shape of the area affected. They used fuzzy C-means to divide homogeneous image regions. Then, some texture features are extracted and trained with the Differential Evolution (DE) algorithm. The experimental process was conducted on HAM10000 and achieved an accuracy of 97.4%.

Afshar et al. [8] presented a computerized method for lesion localization and identification. For the lesion localization, they used RCNN architecture and extract deep features. Later, the best features are selected using Newton-Raphson (IcNR) and artificial bee colony (ABC) optimization. Daghrir et al. [5] developed a hybrid approach for diagnosing suspect lesions that may be checked for melanoma skin cancer. They used a coevolutionary neural network and two classical classifiers in three different methods. Shayini [2] presented a classification framework using geometric and textural information. They used ANN for the final features classification. Results showed improved accuracy as compared to the existing techniques. Akram et al. [7] presented deep learning-based lesion segmentation and classification process. They used Mask RCNN architecture

for lesion segmentation. Later, a 24-layered CNN architecture was designed for the multiclass skin lesion classification.

Moreover, many other techniques are introduced such as deep learning and improved moth-flame optimization [34], teledermatology-based architecture [35], hierarchical three-step deep framework [35], and more [36, 37].

*1.1. Challenges.* Several challenges affect the multiclass lesion classification accuracy. As compared to binary class classification, the multiclass problem is a complex and challenging recognition process. The following challenges are considered in this research work:

- (i) Classifying multiple skin lesions into a correct class is challenging due to the high similarity among different lesions.
- (ii) The imbalanced dataset classes increase the probability of a higher sample class.
- (iii) Multiclass skin lesion types have similar shapes, colors, and textures, which also extract similar features. In the later stage, those features are classified into an incorrect skin class.
- (iv) In the fusion step, multiproperties features are fused in one matrix for better accuracy, but it is a high chance that several redundant features are also added. This kind of problem later increases the computational time.
- (v) In the feature extraction step, several essential features are also removed, which may cause a problem of misclassification. Therefore, a good feature optimization technique is required [38].

*1.2. Major Contributions.* In this work, an automated technique has been proposed for multiclass skin lesion classification. The significant contributions in this work are as follows:

- (i) Intra-class pixel change operations are implemented for data augmentation based on the left to right flip, up-to-down flip, and rotation at 90 degrees. This step shifts entire image pixels for differentiating the images from each other for a fair training of a deep model.
- (ii) A modified serial-based approach is proposed for the fusion of extracted deep features.
- (iii) A novel skewness-controlled SVR approach is proposed for the best feature selection. The best-selected features are finally classified using supervised learning algorithms.

The rest of the manuscript is organized in the following order. Section 2 presented the proposed methodology including deep feature, selection of best features, and fusion process. Results and comparisons with existing techniques are presented in Section 3. Finally, the manuscript is concluded in Section 4.

## 2. Proposed Methodology

For the multiclass skin lesion classification, a new framework was proposed using deep learning and features selection. The proposed framework consists of a series of steps such as data augmentation, model fine-tuning, transfer learning, feature extraction, the fusion of extracted features, and selection of best features. In the augmentation phase, three operations are performed: rotate 90, right-left flip, and up and down flip. In the fine-tuning model step, two models are opted, such as ResNet-50 and ResNet-101, and updated their layers. Later, transfer learning is applied to train both fine-tuned deep models on augmented datasets. In the subsequent step, features are extracted and performed fusion using a modified serial-based approach. Finally, the fused vector is further enhanced by selecting the best features using the skewness-controlled SVR approach. The main architecture diagram of the proposed framework is illustrated in Figure 2.

*2.1. Data Augmentation.* Data augmentation is a vital information extension approach in machine learning (ML). Data augmentation showed much importance in deep learning due to a massive amount of data for training a model. In this article, the HAM10000 dataset is selected for the experimental process. This dataset consists of seven highly imbalanced classes. Initially, the HAM10000 dataset includes more than 10,000 images of seven skin classes such as 6705 images of melanocytic nevi, 1113 images in melanomas, 1099 images in benign keratoses, 514 images in basal cell carcinomas, 327 images of actinic keratoses, 142 images in vascular lesions, and 115 images in dermatofibromas [39]. From this information, it is noted that few classes are highly imbalanced; therefore, it is essential to balance this dataset. On imbalanced datasets, the deep learning models are not trained for better performance. A few sample images are shown in Figure 3.

Three operations are performed in the data augmentation phase: rotate 90, right-left flip (LR), and up and down flip (UD). These operations are applied multiple times until the number of images in each class reached 6000. In the end, the numbers of images in the newly updated dataset are 42,000, which are previously 10,000. Mathematically, these operations are performed as follows.

Consider an image dataset  $\rho = \{a_1, \dots, a_k\}$  [40], where  $a_k \in U$  is an example image from the dataset. Let  $a_k$  have fully  $N$  pixels; then, the homogeneous pixel matrix coordinates  $C_k$  or  $a_k$  is defined as follows:

$$C_k = \begin{bmatrix} Y_1 & Z_1 & 1 \\ Y_2 & Z_2 & 1 \\ \vdots & \vdots & 1 \\ Y_n & Z_n & 1 \end{bmatrix}, \quad (1)$$

where each row of single-pixel indicates the exact coordinates. Consider that the size of an input image is  $256 \times 256 \times 3$ , represented by  $U_{i,j,k}$  having  $i$ th rows,  $j$ th columns, and  $k$ th channels, where  $U_{i,j} \in R^{i \times j}$ . The flip-up (UD) operation is formulated as follows [41]:

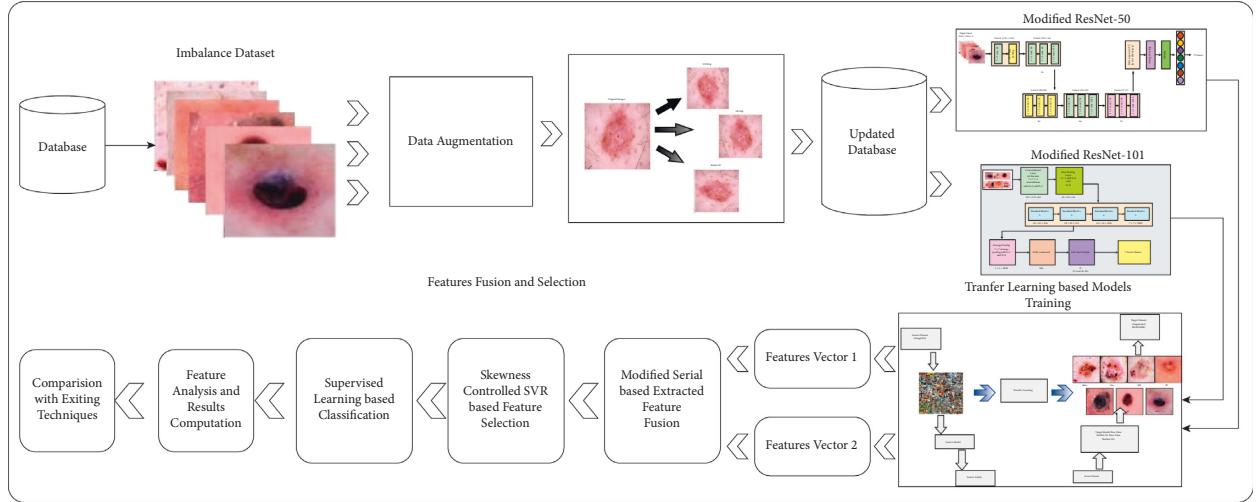


FIGURE 2: Architecture of proposed methodology.

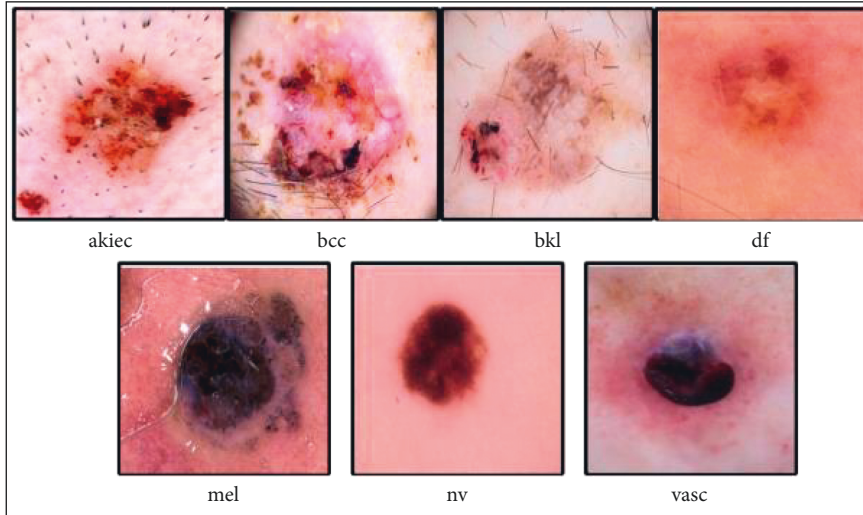


FIGURE 3: Sample skin lesion images of the HAM10000 dataset [7].

$$U^t = U_{j,i}, \quad (2)$$

where  $U^t$  denotes the transposition of the original image. This image is further updated as follows:

$$U^V = U_{(m+1-i)j}, \quad (3)$$

where  $U^V$  denotes the vertical flip image. The horizontal flip (LR) operation is performed as follows:

$$U^H = U_{i(n+1-j)}, \quad (4)$$

where  $U^H$  denotes the horizontal flip image. The third operation, named rotate 90, is formulated as follows:

$$\text{Rot} = \begin{bmatrix} \cos \beta & -\sin \beta & 0 \\ \sin \beta & \cos \beta & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad (5)$$

where Rot denotes the rotation matrix of the image. Visually, these operations are illustrated in Figure 4. This figure shows that three operations are performed on each original image: vertical flip (UD), horizontal flip (LR), and rotate 90.

**2.2. Convolutional Neural Networks.** A convolutional neural network (CNN) is a computer vision technique that automatically distinguishes and recognizes images' features [24]. A simple CNN architecture for image classification is illustrated in Figure 5. In this figure, skin lesion images are considered as input, passed to the convolutional layer. In this layer, weights are transformed into features that are further refined into the pooling layer. Later, the features are transformed into 1D in a fully connected layer. The features of this layer are finally classified through the Softmax layer.

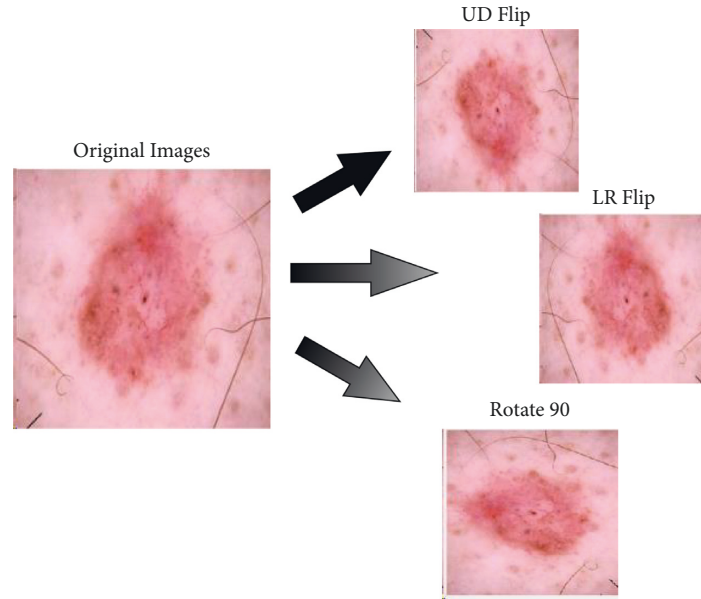


FIGURE 4: Flip operations in data augmentation.

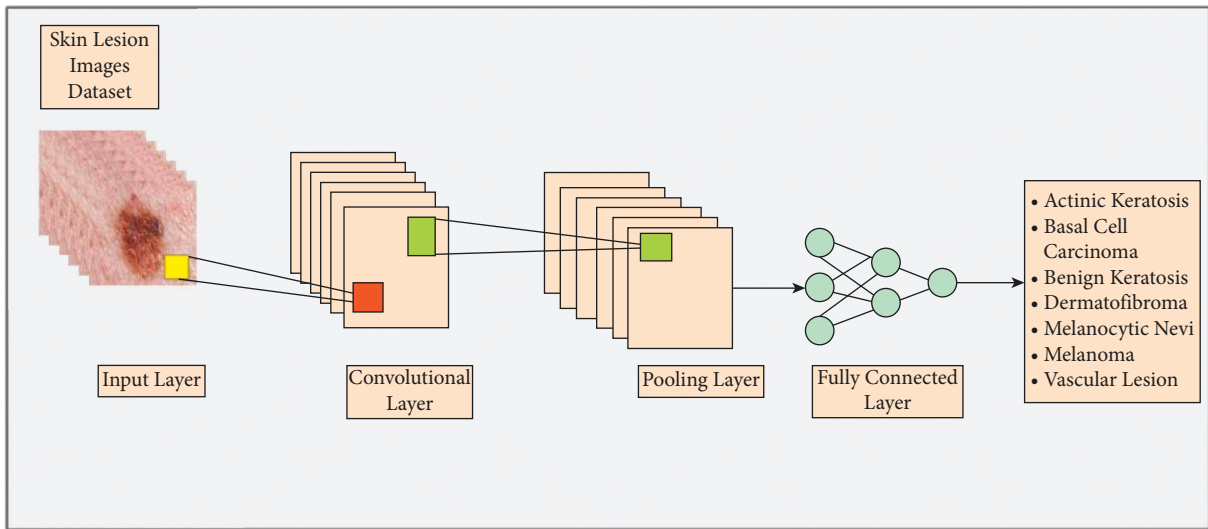


FIGURE 5: A general structure of a CNN model for skin lesion classification.

2.3. *Transfer Learning.* Transfer learning is a technique to define applied knowledge based on one or more source activities. Consider a domain  $M$  consisting of two parts:

$$M = \{z, Q(Z)\}, \quad (6)$$

where  $y$  is a feature space, and the distribution is marginal:

$$q(Z), \quad Z = \{z_1, \dots, z_n\}, z_i \in Z. \quad (7)$$

Given a two-component task  $U$  and  $X$ ,

$$\begin{aligned} U &= \{\varphi, Q(X|Z)\} = \{\varphi, \delta\}; \\ X &= \{x_1, \dots, x_n\}, \quad x_i \in \varphi, \end{aligned} \quad (8)$$

where  $\varphi$  is label space containing a prediction function; then,  $\delta$  is trained as

$$(x_i, z_i), \quad z_i \in Z, x_i \in \varphi. \quad (9)$$

Each vector of features in the  $M$  domain and  $\delta$  represents an appropriate label.

$$\delta(x_i) = (z_i). \quad (10)$$

Suppose the source domain  $M_S$  and an objective domain  $M_T$ , where  $M = \{z, Q(Z)\}$  and the task is  $U_S$  and  $U_T$ , where  $U = \{\varphi, Q(\varphi|Z)\}$ . Hence, TL is defined as follows:

- (i)  $y_S \neq y_T$ : different feature space
- (ii)  $Q(Z_S) \neq Q(Z_T)$ : different marginal possibilities
- (iii)  $\varphi_S \neq \varphi_T$ : different label spaces
- (iv)  $Q(Z_S|Y_S) \neq Q(Z_T|Y_T)$ : different conditional probabilities

Visually, this process is illustrated in Figure 6. This figure describes that the ImageNet dataset used as source data has 1000 object classes. After transferring knowledge of the source model to the target model, the weights and labels are updated according to the target dataset. The HAM10000 skin cancer dataset is utilized as a target dataset with seven skin classes in this work.

**2.4. Fine-Tuned ResNet-50 Deep Features.** Residual Network (ResNet) is a traditional neural network model for many computer vision tasks utilized as an integrated network element. The network has a depth of 50 layers and a size of  $224 \times 224$  pixels in the input [42]. When it comes to residual learning functions, ResNet may reformulate network layers given an input mapping reference. The layers are stacked directly within ResNet. The basic idea of ResNet-50 is to use identity mapping to anticipate what is required to obtain the final prediction of previous layer output [43]. ResNet-50 reduces the disappearing gradient effect by applying an alternative bypass shortcut. It may help the model overcome the overfitting training problem. Visually, it is shown in Figure 7.

Moreover, a complete architecture is also given in Figure 8. This figure describes that five residual blocks are used in this network, and in each residual block, multiple layers are added to convolve hidden layer features. Overall, this network includes 50 deep layers with a  $7 \times 7$  input layer receptive field, followed by a max-pooling layer of  $3 \times 3$  kernel size.

The last fully connected (FC) layer is removed, and a new FC layer is added in the fine-tuning process. Then, the new FC layer is connected with the Softmax layer and final classification output layer. The fine-tuned architecture is shown in Figure 9. This figure describes that the augmented skin lesion dataset is considered an input to this network, and in the output, seven classes of different skin cancer types are gotten. After this, the TL technique is employed to train this network, and a new modified network is obtained. In the training process, the following parameters are initialized; for example, the learning rate is 0.0001, the epochs are 100, the minibatch size is 64, and the learning method is Stochastic Gradient Descent (SGD). Features are extracted from the global average pooling layer, which is later utilized for the classification process. The dimension of an extracted feature on this layer is  $N \times 2048$ , where  $N$  denotes the dermoscopy images.

**2.5. Fine-Tuned ResNet-101 Deep Features.** ResNet-101 consists of 104 layers composed of 33 squares, of which the previous blocks use 29 squares directly [44]. Figure 10 shows a brief description of the ResNet-101 CNN model. In this figure, it is described that the output of the first residual block is  $112 \times 112$ . After the first convolutional layer, a max-pooling layer is added of filter size  $3 \times 3$  and stride 2. Using the same sequence, four more residual blocks are added, and each block consists of several layers, as given in Figure 11. This model was initially trained on the ImageNet dataset; therefore, the output was 1000D.

In this work, this model is fine-tuned according to the target dataset named HAM10000 having seven skin classes. The FC layer is removed in the fine-tuning process and a new FC layer is added with seven outputs. Later, the FC layer is connected with the Softmax layer and output layer and trained using TL. The following parameters are initialized in the training process: the learning rate is 0.0001, epochs are 100, the minibatch size is 64, and the learning method is Stochastic Gradient Descent (SGD). Features are extracted from the average pooling layer, which is later utilized for the classification process. On this layer, the dimension of extracted features is  $N \times 2048$ .

**2.6. Feature Fusion.** Feature fusion is an essential topic in pattern recognition, where multisource features are fused in one vector. The main purpose of feature fusion is to increase the object information for accurate classification. In this work, we consider the idea of a serial-based approach named modified serial-based feature fusion. The proposed fusion approach works in two sequential steps. In the first step, all features of vectors are fused in one matrix, and later on, a standard error mean- (SEM-) based threshold function is proposed.

Assume that  $P$  and  $Q$  are two function rooms on the sample size pattern  $\Delta$ . The corresponding two characteristic vectors  $\delta \in P$  and  $\gamma \in Q$  for an arbitrary sample are  $f \in \Delta$ . The serial-based feature combination of  $f$  is defined as  $\omega = \begin{pmatrix} \delta \\ \gamma \end{pmatrix}$ . Of course, if the vector feature  $\delta$  is  $n$ -dimensional and  $\gamma$  is  $m$ -dimensional, then the combined serial feature  $\omega$  is  $(n + m)$ -dimensions [45]. A serial combined feature space is created by combining all serially merged feature vectors of pattern samples of  $(n + m)$ -dimensions. The resultant  $\omega$  vector has dimension  $N \times 4096$ . After this step, SEM is computed of  $\omega$  using the following formulation:

$$\text{SEM} = \frac{s}{\sqrt{n}},$$

$$s = \sqrt{\frac{\sum_{i=1}^n (\omega_i - \text{Mean})^2}{n - 1}}, \quad (11)$$

$$\text{Thr} = \begin{cases} \text{Fus}(i), & \text{for } \omega_i \geq \text{SEM}, \\ \text{Nfus}(j), & \text{elsewhere,} \end{cases}$$

where Thr denotes the threshold function,  $\text{Fus}(i)$  is fused feature vector of dimension  $N \times 2506$ ,  $\text{Nfus}(j)$  is a feature that is not considered in the fused vector, and  $s$  is a standard deviation value. The output of this step is further refined in the feature selection step, as given below.

**2.7. Feature Selection.** The goal of feature selection is to reduce input variables when a predictive model is developed. This process minimizes the computational time of a proposed system and improves classification accuracy. In this work, a new heuristic search-based feature selection method is proposed named skewness-controlled SVR. In the first step, a skewness feature vector is extracted from the fused

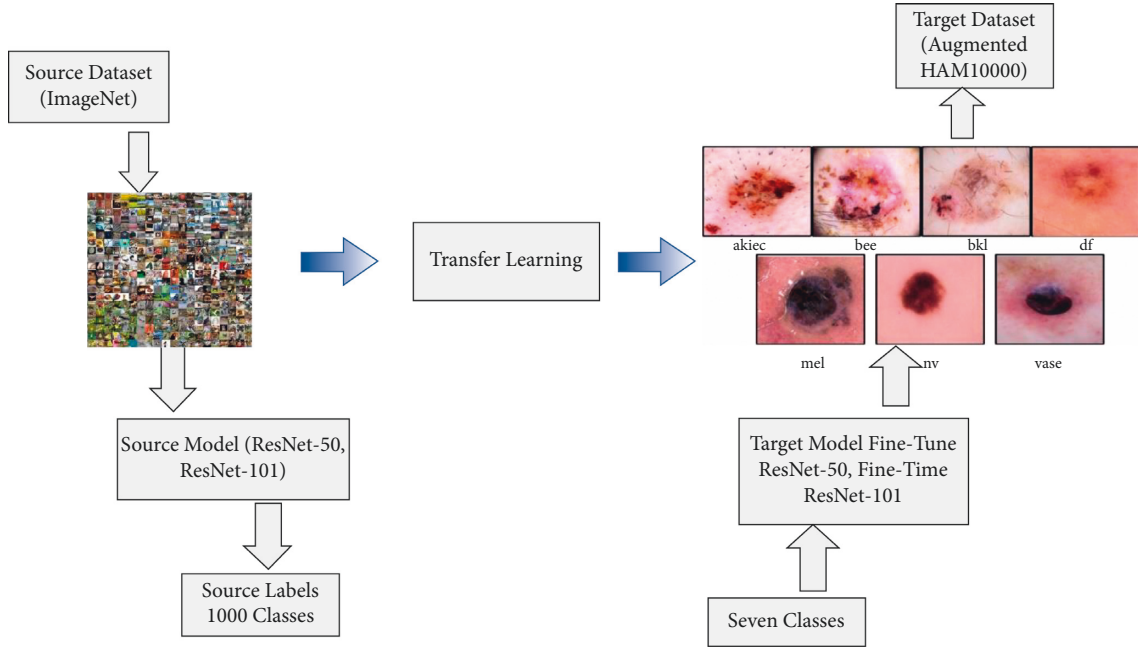


FIGURE 6: Transfer learning-based training a new target model for the classification of skin lesions.

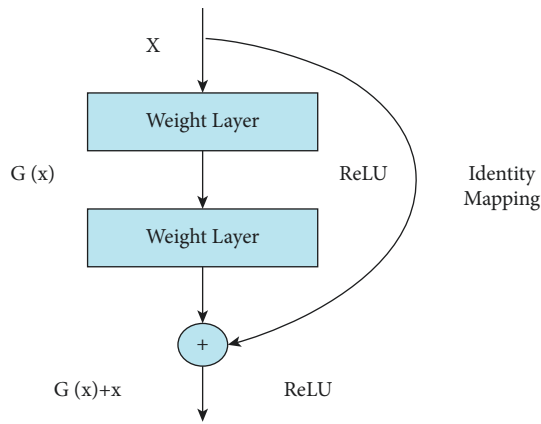


FIGURE 7: Residual identity mapping.

vector  $Fus(i)$ . This step aims to find the likelihood of the features falling in the specific probability distribution. Mathematically, skewness is computed as follows:

$$Skew = \frac{3\overline{Fus(i)} - Median}{s}, \quad (12)$$

where Skew is the skewness feature vector,  $\overline{Fus(i)}$  is the mean value of the fused feature vector, and  $s$  is the standard deviation. Using this skewness value, a threshold function is defined to select features at the first stage.

$$Thr 1 = \begin{cases} Sel(i), & \text{for } Fus(i) \geq Skew, \\ Ignore, & \text{elsewhere.} \end{cases} \quad (13)$$

Using this threshold function, features are selected at the initial phase. The selected features of this phase are later validated using a fitness function Support Vector Regression (SVR). The SVR is formulated as follows.

Layer Name	Output Size	50-Layer
Conv1	112×112	7×7,64, stride2
Conv2	56×56	3×3 max pool, stride2
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
		$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 512 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv4	14×14	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2,048 \end{bmatrix} \times 3$
		$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2,048 \end{bmatrix} \times 3$
Conv5	7×7	Average pool, 7D fc, Softmax
		1×1
FLOPs		3.8×10 <sup>9</sup>

FIGURE 8: Architecture of ResNet-50.

Assume that the dataset for training  $Q$  comprises the instances  $q$ , each having an attribute  $u_i$ , an associated class, and  $v_i$ .  $u_i \in Sel(i)$  is a selected feature and  $v_i$  represents labels; i.e.,  $\{(u_1, v_1), (u_2, v_2), \dots, (u_q, v_q)\}$ . On the dataset  $D$ ,  $b$  is a bias, and the linear function  $f(x)$  may be defined as follows:

$$f(u) = \delta_1 u_1 + \delta_2 u_2 + \dots + \delta_d u_d + b, \quad (14)$$

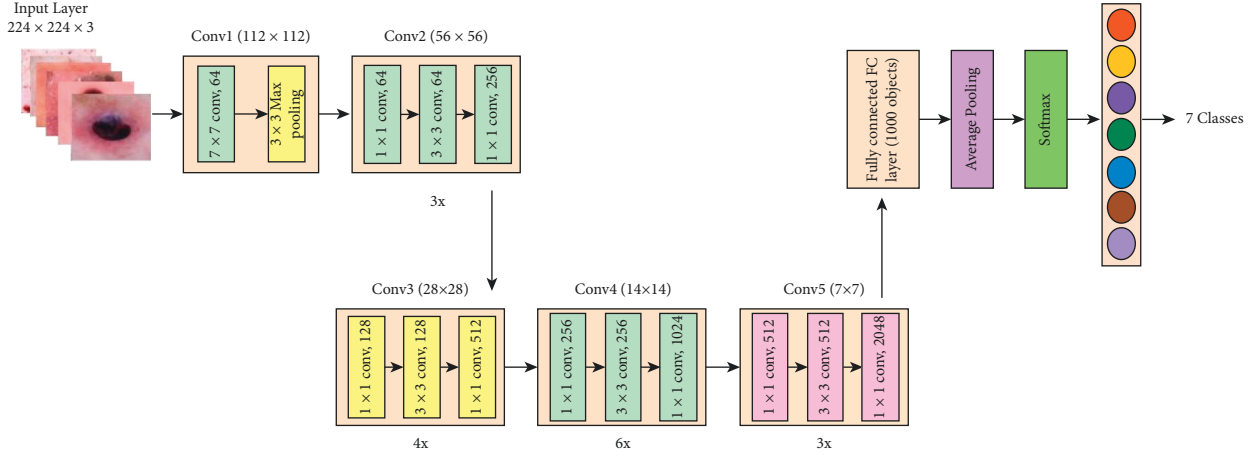


FIGURE 9: Block diagram of fine-tuned ResNet-50.

Layer Name	Output Size	100-Layer
Conv1	112x112	7x7,64,stride2
Conv2	56x56	3x3max pool,stride2 $\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3	28x28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Conv4	14x14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 512 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$
Conv5	7x7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1x1	Average pool,7D fc,Softmax
FLOPs		$7.6 \times 10^9$

FIGURE 10: Fine-tuned architecture of ResNet-101.

where the weight  $\delta_i$  is defined as input space  $S^d$ ; i.e.,  $\delta_i \in S^d$ . The maximum margin size is determined by the Euclidean weight ( $\|Y\|$ ). The flatness, therefore, requires a minimum weight standard in the case of the following equation. Here, the definition of ( $\|Y\|$ ) is

$$\|Y\|^2 = Y_1^2 + Y_2^2 + \dots + Y_d^2. \quad (15)$$

Each training data error may be represented as  $\langle u_i, v_i \rangle$ .

$$\text{Err}_i(u_i) = v_i - (\delta_i u_i + b). \quad (16)$$

If there is error  $\text{Err}_i(u_i)$ , the deviation is permitted to be within it, and the previous equation may be expressed as  $\eta$ .

$$\begin{aligned} v_i - (\delta_i u_i + b) &\leq \eta, \\ (\delta_i u_i + b) - v_i &\leq \eta. \end{aligned} \quad (17)$$

Using these two equations, the minimization issue for  $\delta$  can be formulated as follows:

$$\text{minimized: } \frac{1}{2} \|Y\|^2, \quad (18)$$

subject to

$$\begin{aligned} v_i - (\delta_i u_i + b) &\leq \eta, \\ (\delta_i u_i + b) - v_i &\leq \eta. \end{aligned} \quad (19)$$

The restrictions of the above equation imply that the function  $f$  corresponds to all pairings  $(u_i, v_i)$  with a deviation of  $\eta$ . However, the assumption is not accepted in all instances when the slack variables  $\delta_i, \delta_i^*$  are neither required nor necessary in case of violation of the assumption. The optimization problem may be reformulated using slack variables as follows:

$$\text{minimize: } \frac{1}{2} \|Y\|^2 + C \sum_{i=0}^d (\delta_i + \delta_i^*), \quad (20)$$

subject to

$$\begin{aligned} \forall_i: v_i - (\delta_i u_i + b) &\leq \eta + \delta_i, \\ \forall_i: (\delta_i u_i + b) - v_i &\leq \eta + \delta_i^*, \\ \forall_i: \delta_i &\geq 0, \\ \forall_i: \delta_i^* &\leq 0, \end{aligned} \quad (21)$$

where  $C$  is the penalty constant, which does not meet the constraints. It also helps in reducing overfitting. The Kernel is defined by the input data  $K(u_i, u_j)$  and can substitute the occurrence of the dot product between the tuples to avoid the dot product on a data tuple changed. All computations are therefore done in the original input areas. In this work, a radial basis Kernel/Gaussian function is utilized:



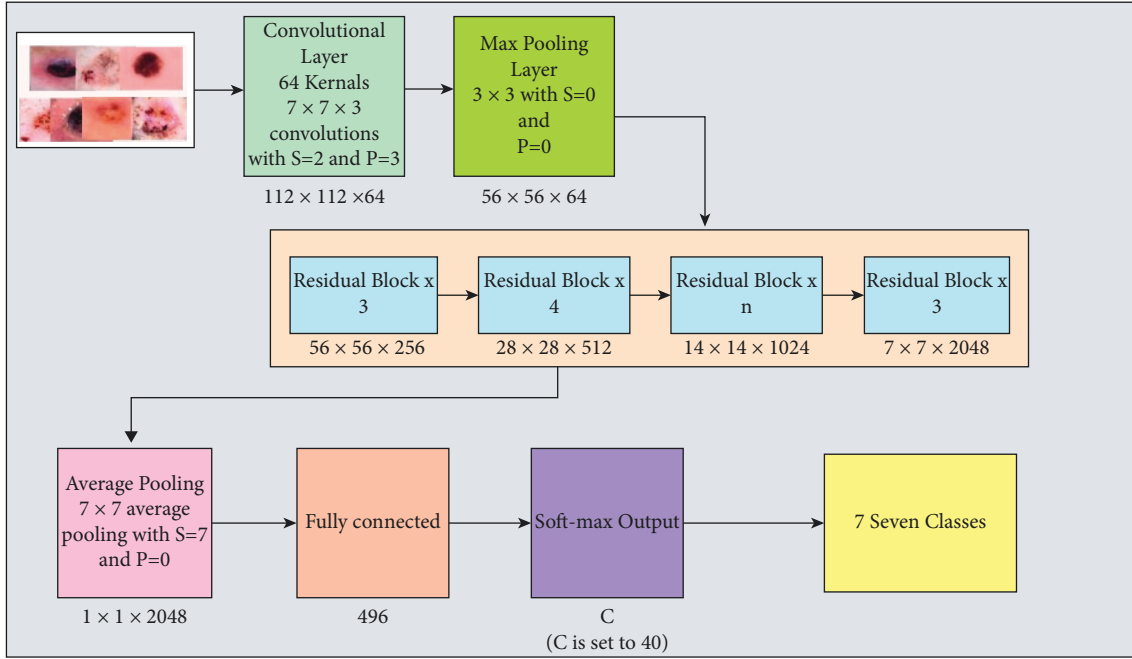


FIGURE 11: Transfer learning architecture with modified ResNet-101 model ( $n=23$  for ResNet-101; stride  $\ast S$  and padding  $\ast P$ ).

$$K(u_i, u_j) = \exp - \frac{\left( \|u_i, u_j + 1\|^2 \right)}{2\rho^2}. \quad (22)$$

The accuracy is computed using SVR, and if accuracy is less than the target accuracy value, then  $Sel(i)$  is again updated. This process is continued until the maximum number of iterations is performed. In this work, the target accuracy is 90%, and the numbers of iterations are 5. Following this process, a feature vector is obtained called the best-selected feature vector of dimension  $N \times 1456$  and further fed to supervised learning algorithms for final classification.

### 3. Experimental Results and Discussion

The proposed method is evaluated on the augmented HAM10000 dataset. Dataset is divided into 70 : 30, where the 70% data is used for the training of a model, and the rest of the 30% is utilized for the testing process. The other training hyperparameters; for example, epochs are 100, the mini-batch size is 64, and the learning rate is 0.0001. The 10-fold method was carried out for cross-validation [46]. Seven performance measures are used for the experimental process: recall rate, precision rate, false-negative rate (FNR), Area under Curve (AUC), accuracy, time, and F1-score. The proposed method is implemented in MATLAB 2020b, Corei7, with a RAM 16GB and 8GB graphics card.

**3.1. Results.** In this section, the proposed method results are described in numerical values (Tables) and confusion matrixes. Total ten classifiers are utilized for the experimental process, such as Linear Support Vector Machine (LSVM), Quadratic SVM (QSVM), Cubic SVM (CSVM), Medium

Gaussian SVM (MGSVM), Cosine K-Nearest Neighbor (CKNN), Weighted KNN (WKNN), Coarse KNN (CKNN), Ensemble Subspace Discriminative (ESD), Ensemble Boosted Tree (EBT), and Ensemble Subspace KNN (ESKNN). Five experiments are performed for the validation of the proposed framework such as (i) Experiment # 1: classification using fine-tuned ResNet-50 CNN model, (ii) Experiment # 2: classification using Fine-Tuned ResNet-101CNN model, (iii) Experiment # 3: perform features fusion of Fine-Tuned ResNet-50 and ResNet-101 CNN models, and (iv) Best Features (BF) selection.

**3.1.1. Experiment # 1.** In the first experiment, features are extracted using fine-tuned ResNet-50 CNN model, and results are computed. The augmented dataset was used for the experimental process. The results of this experiment are given in Table 1. CSVM has the highest accuracy of 92.7% in this table, with computational time 1190.3 (sec). Figure 12 shows the confusion matrix of CSVM for this experiment. In this figure, the diagonal values represent the correct predicted values such as AKIEC (96%), BCC (93%), BKL (87%), DF (97%), MEL (86%), NV (94%), and VASC (99%), respectively. Moreover, the recall rate is 93.14, the precision rate is 93.14, and F1-score is 93.14%, respectively. Compared with the rest of the classifiers, it is noticed that the CSVM showed better classification accuracy. Moreover, the computational time of each classifier is also noted and plotted in Figure 13. This figure shows that the CKNN has the lowest computational time of 274.55 (sec).

**3.1.2. Experiment # 2.** Table 2 presents the results of fine-tuned ResNet-101 CNN features using the augmented HAM10000 dataset. This table shows that the best accuracy

TABLE 1: Classification accuracy of fine-tuned ResNet-50 deep features using augmented HAM10000 dataset.

Classifier	Recall rate (%)	Precision rate (%)	FNR (%)	AUC	Accuracy (%)	Time (sec)	F1-score (%)
LSVM	86.42	86.85	13.57	0.988	86.5	742.5	86.63
QSVM	92.00	92.14	8.00	0.992	91.7	1046.1	92.07
<b>CSVM</b>	<b>93.14</b>	<b>93.14</b>	<b>6.858</b>	<b>0.994</b>	<b>92.7</b>	1190.3	<b>93.14</b>
MGSVM	89.57	90.00	10.42	0.988	89.3	1906.8	89.78
CKNN	53.25	65.28	46.75	0.898	60.8	<b>274.5</b>	58.65
CKNN	80.42	79.00	19.57	0.967	78.7	287.6	79.70
WKNN	85.14	84.42	14.85	0.98	83.6	262.3	84.78
ESKNN	93.14	92.57	6.858	0.99	92.3	4514.7	92.85
EBT	55.28	86.71	44.71	0.974	57.1	1546.0	86.49
ESD	86.28	57.00	13.71	0.85	86.1	839.8	56.12

The bold value represents best ones.

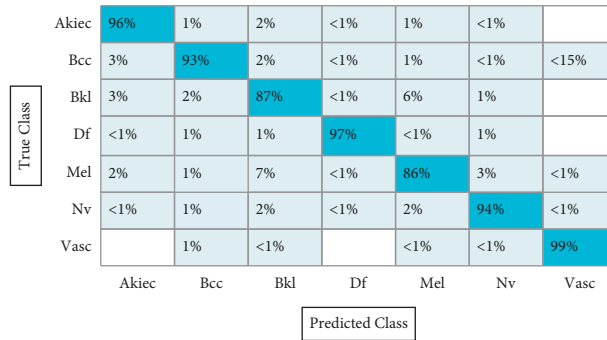


FIGURE 12: Confusion matrix of CSVM using ResNet-50 model for HAM10000 dataset.

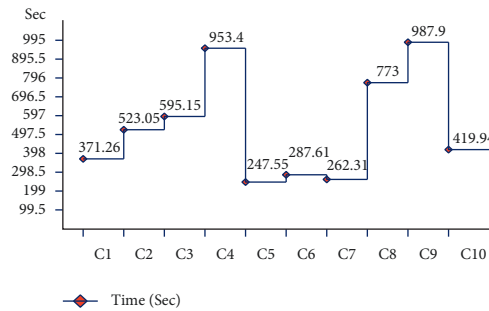


FIGURE 13: Time plot for fine-tuned ResNet-50 CNN model using augmented HAM10000 dataset.

TABLE 2: Classification accuracy of fine-tuned ResNet-101 deep features using augmented HAM10000 dataset.

Classifier	Recall rate (%)	Precision rate (%)	FNR (%)	AUC	Accuracy (%)	Time (sec)	F1-score (%)
LSVM	86.00	86.28	14.00	0.98	85.5	746.2	86.14
QSVM	91.57	91.71	8.428	0.992	91.1	1010.2	91.64
<b>CSVM</b>	<b>92.71</b>	<b>92.42</b>	<b>7.285</b>	<b>0.992</b>	<b>92.1</b>	11321.1	<b>92.56</b>
MGSVM	89.42	89.42	10.57	0.988	88.9	1919.4	89.42
CKNN	78.87	77.85	21.14	0.961	77.2	268.2	78.36
CKNN	58.42	63.00	41.27	0.887	58.9	263.6	60.62
WKNN	84.85	84.00	15.14	0.977	83.3	<b>260.5</b>	84.42
EBT	56.57	56.57	43.42	0.855	57.2	1544.5	56.57
ESKNN	80.28	92.28	19.71	0.99	92.1	4590.9	85.86
ESD	85.85	86.28	14.14	0.98	85.6	821.79	86.06

The bold value represents best ones.

achieved by CSVM is 92.1%, with a computational time of 11321.1 (sec), recall rate is 92.7, the precision rate is 92.42, and F1-score is 92.56%, respectively. Figure 14 shows the

confusion matrix of CSVM. In this figure, the diagonal values represent the correct predicted values such as AKIEC (96%), BCC (92%), BKL (85%), DF (98%), MEL (86%), NV

True Class	Akiec	96%	1%	1%		1%	<1%	
	Bcc	3%	92%	2%	1%	2%	1%	<1%
	Bkl	2%	3%	85%	<1%	7%	2%	<1%
	Df	1%	1%	<1%	98%		<1%	
	Mel	2%	1%	7%	<1%	86%	3%	<1%
	Nv	<1%	1%	3%	<1%	2%	93%	<1%
	Vasc		1%	<1%			<1%	99%
			Akiec	Bcc	Bkl	Df	Mel	Nv
		Predicted Class						

FIGURE 14: Confusion matrix of CSVM using ResNet-101 model for HAM10000 dataset.

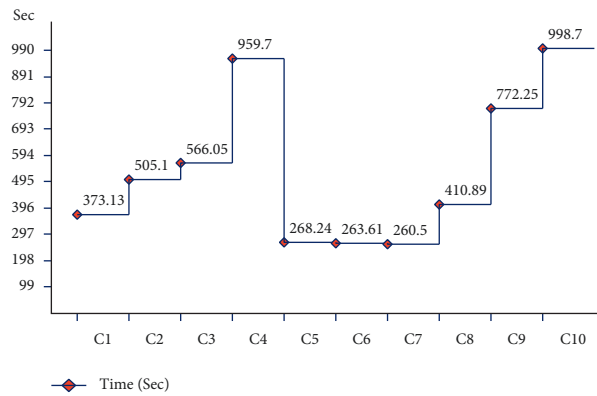


FIGURE 15: Time plot for fine-tuned ResNet-101 CNN model using augmented HAM10000 dataset.

(93%), and VASC (99%), respectively. As given in this table, a few other classifiers are also implemented and show that the CSVM gives better accuracy. Moreover, the computational time is computed for each classifier, and the minimum noted time is 260.5 (sec) for the W-KNN classifier. The noted time is also plotted in Figure 15.

**3.1.3. Experiment # 3.** In the next experiment, features are fused using the serial-based extended (SbE) approach. Results are given in Table 3. This table represents the best accuracy achieved by the ESD classifier of 95%, further demonstrating in a confusion matrix, given in Figure 16. This figure represents the correct predicted values such as AKIEC (97%), BCC (94%), BKL (89%), DF (98%), MEL (89%), NV (99%), and VASC (99%), respectively. The other computed measures are recall rate, precision rate, FNR, AUC, and F1-score of 95.0, 95.0, 5.00, 0.99, and 95.0%, respectively. The CSVM achieved the second-best accuracy of 94.9%, whereas the recall rate and precision rates are 95.0%. Comparison with the rest of the classifiers shows the superiority of the ESD classifier. Moreover, the computational time is also noted, as illustrated in Figure 17.

Compared with the results of this experiment with Tables 1 and 2, it is noticed that the fusion using the SbE approach significantly improves the classification accuracy. The limitation of this step increases computational time, which needs to be minimized.

**3.1.4. Experiment # 4.** Finally, the proposed feature selection algorithm is applied on the fused feature vector and achieved an accuracy of 91.7% on the ESD classifier, where the computational time is 1367 (sec), given in Table 4. The recent time was 4118 (sec), which is significantly minimized after the selection algorithm. This table also showed that the proposed accuracy decreases, but on the other side, it helps to minimize the computational time. The accuracy of the ESD classifier is further verified using a confusion matrix given in Figure 18. In this figure, the diagonal values represent the correct predicted values such as AKIEC (94%), BCC (91%), BKL (85%), DF (93%), MEL (83%), NV (97%), and VASC (99%), respectively.

The F1-score-based analysis is also conducted and plotted in Figure 19. In this figure, it is illustrated that the value of the F1-score is improved after the feature fusion process except the CKNN and EBT classifier. Moreover, the feature selection approach reduced the computational time but accuracy is degraded. Overall, the proposed framework performed well on the selected dataset. In the last, the proposed method accuracy is compared with some recent techniques, as given in Table 5. In this table, Khan et al. [7] presented a deep learning method for skin lesion classification. They used the HAM10000 dataset and achieved an accuracy of 88.5%. The recent best-reported accuracy was 91.5%, achieved by Sevli [47]. The proposed accuracy is 91.7% and 95% for the best feature selection approach and fusion approach. Based on this accuracy, it is noted that the proposed method showed improved accuracy.

TABLE 3: Classification results using SbE approach-based deep features fusion on augmented HAM10000 dataset.

Classifier	Recall rate (%)	Precision rate (%)	FNR (%)	AUC	Accuracy (%)	Time (sec)	F1-score (%)
LSVM	92.71	92.85	7.285	0.992	92.50	1303.8	92.78
QSVM	94.85	94.85	5.142	0.997	94.80	2400.4	94.75
CSVM	95.00	95.00	5.00	0.854	94.90	2868.5	95.00
MG SVM	92.85	93.14	7.142	0.995	92.60	4501.6	92.99
CKNN	61.71	73.14	38.28	0.910	62.20	562.0	66.94
CKNN	84.14	83.57	15.85	0.975	82.60	550.5	83.85
WKNN	83.42	84.57	16.27	0.971	82.10	<b>530.7</b>	83.99
<b>ESD</b>	<b>95.00</b>	<b>95.00</b>	<b>5.00</b>	<b>0.997</b>	<b>95.00</b>	<b>4118.2</b>	<b>95.00</b>
FKNN	88.14	87.57	11.85	0.931	87.00	543.2	87.85
EBT	62.00	62.22	38.00	0.855	62.80	69886.0	62.11

The bold value represents best ones.

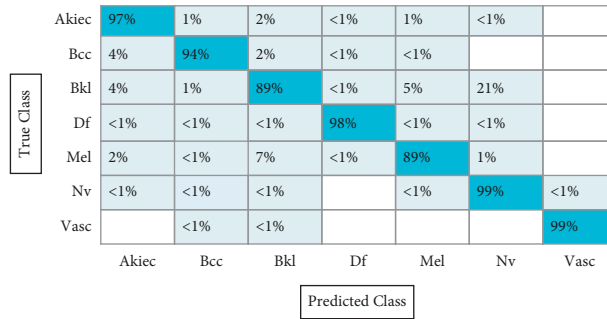


FIGURE 16: Confusion matrix of ESD using ResNet-50 and ResNet-101 model for HAM10000 dataset.

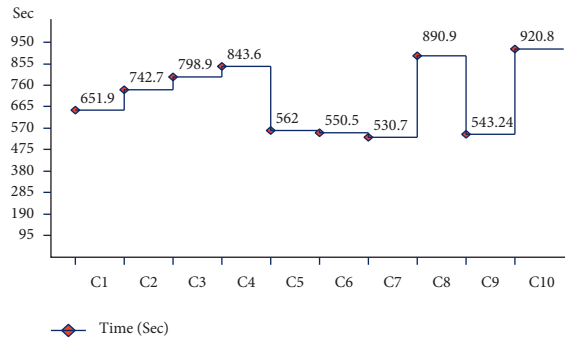


FIGURE 17: Time plot for the fusion of ResNet-50 and ResNet-101 using augmented dataset (HAM10000).

TABLE 4: Classification results using proposed feature selection algorithm on augmented HAM10000 dataset.

Classifier (ESD)	Recall rate (%)	Precision rate (%)	FNR (%)	AUC	Accuracy (%)	Time (sec)	F1-score (%)
500	78.42	79.28	21.57	0.954	78.3	102.4	78.85
1000	64.85	66.71	35.14	0.902	65.1	<b>132.5</b>	65.77
1500	87.14	87.57	12.85	0.978	86.8	406.0	87.35
2000	83.57	84.42	16.42	0.97	83.3	748.4	83.99
2500	89.71	90.14	10.28	0.984	89.5	1293.5	89.92
<b>3000</b>	<b>91.85</b>	<b>92.00</b>	<b>8.142</b>	<b>0.99</b>	<b>91.7</b>	<b>1367.6</b>	<b>91.92</b>

The bold value represents best ones.

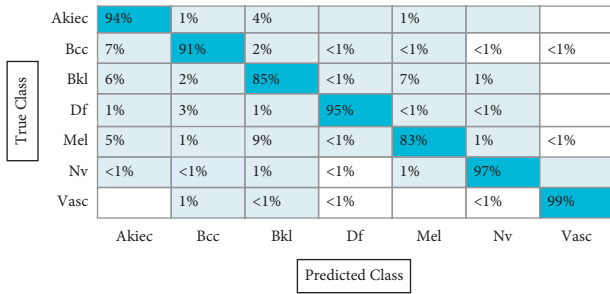


FIGURE 18: Confusion matrix of ESD classifier using a proposed feature selection algorithm.

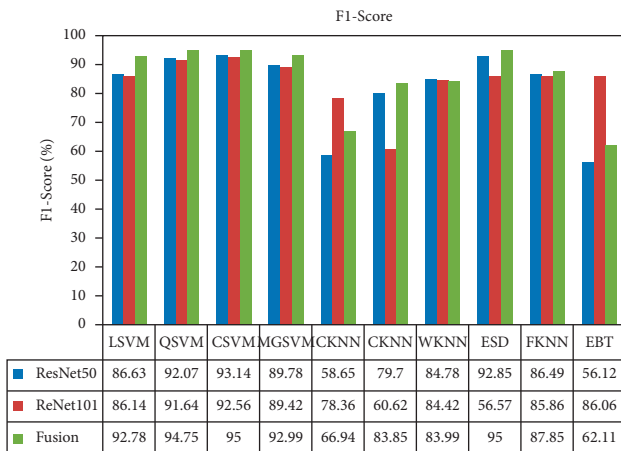


FIGURE 19: F1-score-based analysis of middle steps like ResNet-50, ResNet-101, and fusion.

TABLE 5: Comparison with existing techniques.

Reference	Year	Dataset	Accuracy (%)
[7]	2020	HAM10000	88.5
[48]	2020	HAM10000	86.1
[49]	2020	HAM10000	83.1
[30]	2021	HAM10000	85.50
[47]	2020	HAM10000	91.5
<b>Proposed (fusion)</b>	<b>2021</b>	<b>HAM10000</b>	<b>95.0</b>
<b>Proposed (feature selection)</b>	<b>2021</b>	<b>HAM10000</b>	<b>91.7</b>

The bold value represents best ones.

#### 4. Conclusion

In this work, a new framework is presented for multiclass skin lesion classification using deep learning. The proposed method consisted of a series of steplike data augmentation, feature extraction using deep learning models, the fusion of features, selection of parts, and classification. The experiment was performed on an augmented HAM10000 dataset. The number of experiments was performed, such as non-augmented and augmented datasets, and achieved accuracy with a nonaugmented dataset of 64.36% using ResNet-50 and 49.98% using ResNet-101. The augmented dataset achieved an accuracy of 95.0% for feature fusion and 91.7% for feature selection. The results show that the augmentation process helps improve the classification accuracy for a complex dataset.

Moreover, the fusion process increases the performance but also increases the computational time. This process can be further refined through a feature selection process. However, according to the results, the feature selection process decreases the computational time and reduces accuracy. But from the overall comparison with recent techniques, feature fusion and feature selection technique both perform better than previous techniques. The new datasets ISBI 2020 and ISIC 2020 can be used for the experimental process in future work. Latest deep learning models can be used as feature extraction. Fusion can be performed using parallel approaches. The selection process can be refined, which not only reduces the time but also increases accuracy.

#### Data Availability

The HAM10000 dataset is utilized in this work for the experimental process. The dataset is publically available at <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T>.

#### Conflicts of Interest

All authors declare that they have no conflicts of interest in this study.

#### Acknowledgments

The work of Shabnam M. Aslam was supported by the Majmaah University’s Deanship of Scientific Research under Project 155/46683.

#### References

- [1] M. H. Trager, L. J. Geskin, F. H. Samie, and L. Liu, “Bio-markers in melanoma and non-melanoma skin cancer prevention and risk stratification,” *Experimental Dermatology*, 2020.
- [2] R. Shayini, “Classification of skin lesions in digital images for the diagnosis of skin cancer,” in *Proceedings of the 2020 International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 162–166, Trichy, India, 2020.
- [3] B. Ahmad, M. Usama, C.-M. Huang, K. Hwang, M. S. Hossain, and G. Muhammad, “Discriminative feature learning for skin disease classification using deep convolutional neural network,” *IEEE Access*, vol. 8, pp. 39025–39033, 2020.
- [4] M. Fernandez-Figueras, C. Carrato, X. Sáenz, L. Puig, E. Musulen, and C. Ferrándiz, “Actinic keratosis with atypical basal cells (AK I) is the most common lesion associated with invasive squamous cell carcinoma of the skin,” *Journal of the European Academy of Dermatology and Venereology*, vol. 29, pp. 991–997, 2015.
- [5] J. Daghrir, L. Tlig, M. Bouchouicha, and M. Sayadi, “Melanoma skin cancer detection using deep learning and classical machine learning techniques: a hybrid approach,” in *Proceedings of the 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 1–5, Sousse, Tunisia, 2020.
- [6] S. S. Chaturvedi, J. V. Tembhurne, and T. Diwan, “A multi-class skin cancer classification using deep convolutional

- neural networks,” *Multimedia Tools and Applications*, vol. 79, no. 39-40, pp. 28477–28498, 2020.
- [7] T. Akram, Y.-D. Zhang, and M. Sharif, “Attributes based skin lesion detection and recognition: a mask RCNN and transfer learning-based deep learning framework,” *Pattern Recognition Letters*, vol. 143, pp. 58–66, 2021.
  - [8] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, “Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images,” *Pattern Recognition Letters*, vol. 138, pp. 638–643, 2020.
  - [9] M. Kumar, M. Alshehri, R. AlGhamdi, P. Sharma, and V. Deep, “A de-ann inspired skin cancer detection approach using fuzzy c-means clustering,” *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1319–1329, 2020.
  - [10] A. R. Lopez, X. Giro-i-Nieto, J. Burdick, and O. Marques, “Skin lesion classification from dermoscopic images using deep learning techniques,” in *Proceedings of the 2017 13th IASTED international conference on biomedical engineering (BioMed)*, pp. 49–54, Innsbruck, Austria, 2017.
  - [11] S. Roy, T. D. Whitehead, S. Li, F. O. Ademuyiwa, R. L. Wahl, and F. Dehdashti, “Co-clinical FDG-PET radiomic signature in predicting response to neoadjuvant chemotherapy in triple negative breast cancer,” *European Journal of Nuclear Medicine and Molecular Imaging*, 2021.
  - [12] K. Dutta, S. Roy, T. D. Whitehead et al., “Deep learning segmentation of triple-negative breast cancer (TNBC) patient derived tumor xenograft (PDX) and sensitivity of radiomic pipeline to tumor probability boundary,” *Cancers*, vol. 13, no. 15, p. 3795, 2021.
  - [13] A. Sedik, A. Ilyasu, M. E El-Rahiem et al., “Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections,” *Viruses*, vol. 12, no. 7, p. 769, 2020.
  - [14] T. Saba, M. A. Khan, A. Rehman, and S. L Marie-Sainte, “Region extraction and classification of skin cancer: a heterogeneous framework of deep CNN features fusion and reduction,” *Journal of Medical Systems*, vol. 43, pp. 289–319, 2019.
  - [15] I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, and A. Rehman, “Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists,” *Diagnostics*, vol. 10, p. 565, 2020.
  - [16] U. Nazar, M. A. Khan, I. U. Lali et al., “Review of automated computerized methods for brain tumor segmentation and classification,” *Current Medical Imaging Formerly Current Medical Imaging Reviews*, vol. 16, no. 7, pp. 823–834, 2020.
  - [17] S. A. Khan, M. Nazir, M. A. Khan et al., “Lungs nodule detection framework from computed tomography images using support vector machine,” *Microscopy Research and Technique*, vol. 82, no. 8, pp. 1256–1266, 2019.
  - [18] T. Meraj, H. T. Rauf, S. Zahoor, A. Hassan, M. I. Lali, and L. Ali, “Lung nodules detection using semantic segmentation and classification with optimal features,” *Neural Computing & Applications*, vol. 33, pp. 10737–10750, 2021.
  - [19] A. Sedik, M. Hammad, F. E. Abd El-Samie, B. B. Gupta, and A. A. Abd El-Latif, “Efficient deep learning approach for augmented detection of Coronavirus disease,” *Neural Computing and Applications*, Springer, Berlin, Germany, pp. 1–18, 2021.
  - [20] H. T. Rauf, M. I. U. Lali, S. Kadry, H. Alolaiyan, and A. Razaq, “Time series forecasting of COVID-19 transmission in Asia Pacific countries using deep neural networks,” *Personal and Ubiquitous Computing*, pp. 1–18, 2021.
  - [21] M. Hammad, M. H. Alkinani, B. Gupta, and A. A. Abd El-Latif, “Myocardial infarction detection based on deep neural network on imbalanced data,” *Multimedia Systems*, pp. 1–13, 2021.
  - [22] M. Hammad, A. M. Ilyasu, A. Subasi, E. S. Ho, and A. A. Abd El-Latif, “A multitier deep learning model for arrhythmia detection,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2020.
  - [23] A. Alghamdi, M. Hammad, H. Ugail, A. Abdel-Raheem, K. Muhammad, and H. S. Khalifa, “Detection of myocardial infarction based on novel deep transfer learning methods for urban healthcare in smart cities,” 2019, <https://arxiv.org/abs/1906.09358>.
  - [24] X. Yu, X. Wu, C. Luo, and P. Ren, “Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework,” *GIScience and Remote Sensing*, vol. 54, no. 5, pp. 741–758, 2017.
  - [25] A. F. Agarap, “Deep learning using rectified linear units (relu),” 2018, <https://arxiv.org/abs/1803.08375>.
  - [26] L. Liu, C. Shen, and A. Van Den Hengel, “The treasure beneath convolutional layers: cross-convolutional-layer pooling for image classification,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4749–4757, Boston, MA, USA, 2015.
  - [27] S. Sharma and S. Sharma, “Activation functions in neural networks,” *Data Science*, vol. 6, pp. 310–316, 2017.
  - [28] P. Marcelino, “Transfer learning from pre-trained models,” *Data Science*, 2018.
  - [29] A. Majid, N. Hussain, M. Alhaisoni, Y.-D. Zhang, and S. Kadry, “Multiclass stomach diseases classification using deep learning features optimization,” *Computers, Materials and Continua*, vol. 67, no. 3, 2021.
  - [30] F. Afza, M. Sharif, M. Mittal, and D. J. Hemanth, “A hierarchical three-step superpixels and deep learning framework for skin lesion classification,” *Methods*, 2021.
  - [31] A. K. Verma, S. Pal, and S. Kumar, “Comparison of skin disease prediction by feature selection using ensemble data mining techniques,” *Informatics in Medicine Unlocked*, vol. 16, Article ID 100202, 2019.
  - [32] T. Akram, H. M. J. Lodhi, S. R. Naqvi, S. Naeem, M. Alhaisoni, and M. Ali, “A multilevel features selection framework for skin lesion classification,” *Human-Centric Computing and Information Sciences*, vol. 10, pp. 1–26, 2020.
  - [33] H. W. Huang, B. W. Y. Hsu, C. H. Lee, and V. S. Tseng, “Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers,” *The Journal of Dermatology*, vol. 48, no. 3, 2020.
  - [34] M. Sharif, T. Akram, R. Damaševičius, and R. Maskeliūnas, “Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization,” *Diagnostics*, vol. 11, p. 811, 2021.
  - [35] K. Muhammad, M. Sharif, T. Akram, and V. H. C. de Albuquerque, “Multi-class skin lesion detection and classification via teledermatology,” *IEEE Journal of Biomedical and Health Informatics*, 2021.
  - [36] K. Melbin and Y. J. V. Raj, “Integration of modified ABCD features and support vector machine for skin lesion types classification,” *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 8909–8929, 2021.
  - [37] H. Zanddzari, N. Nguyen, B. Zeinali, and J. M. Chang, “A new preprocessing approach to improve the performance of CNN-based skin lesion classification,” *Medical, & Biological Engineering & Computing*, vol. 59, no. 5, pp. 1123–1131, 2021.
  - [38] H. T. Rauf, U. Shoaib, M. I. Lali, M. Alhaisoni, M. N. Irfan, and M. A. Khan, “Particle swarm optimization with

- probability sequence for global optimization,” *IEEE Access*, vol. 8, pp. 110535–110549, 2020.
- [39] H. Younis, M. H. Bhatti, and M. Azeem, “Classification of skin cancer dermoscopy images using transfer learning,” in *Proceedings of the 2019 15th International Conference on Emerging Technologies (ICET)*, pp. 1–4, Peshawar, Pakistan, 2019.
- [40] N. Hussain, M. Sharif, S. A. Khan, A. A. Albeshier, and T. Saba, “A deep neural network and classical features based scheme for objects recognition: an application for machine inspection,” *Multimedia Tools and Applications*, Springer, Berlin, Germany, pp. 1–23, 2020.
- [41] O. A. Shawky, A. Hagag, E.-S. A. El-Dahshan, and M. A. Ismail, “Remote sensing image scene classification using CNN-MLP with data augmentation,” *Optik*, vol. 221, Article ID 165356, 2020.
- [42] H. Touvron, A. Vedaldi, M. Douze, and H. Jégou, “Fixing the train-test resolution discrepancy,” 2019, <https://arxiv.org/abs/1906.06423>.
- [43] J.-A. Almaraz-Damian, V. Ponomaryov, S. Sadovnychiy, and H. Castillejos-Fernandez, “Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures,” *Entropy*, vol. 22, no. 4, p. 484, 2020.
- [44] A. Demir, F. Yilmaz, and O. Kose, “Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3,” in *Proceedings of the 2019 Medical Technologies Congress (TIPTEKNO)*, pp. 1–4, Izmir, Turkey, 2019.
- [45] S. Mustafa, “Feature selection using sequential backward method in melanoma recognition,” in *Proceedings of the 2017 13th International Conference on Electronics, Computer and Computation (ICECCO)*, pp. 1–4, Abuja, Nigeria, 2017.
- [46] D. Zhao, G. Yu, P. Xu, and M. Luo, “Equivalence between dropout and data augmentation: a mathematical check,” *Neural Networks*, vol. 115, pp. 82–89, 2019.
- [47] O. Sevli, “A deep convolutional neural network-based pigmented skin lesion classification application and experts evaluation,” *Neural Computing and Applications*, Springer, Berlin, Germany, pp. 1–12, 2021.
- [48] M. A. Khan, Y.-D. Zhang, M. Sharif, and T. Akram, “Pixels to classes: intelligent learning framework for multiclass skin lesion localization and classification,” *Computers & Electrical Engineering*, vol. 90, Article ID 106956, 2021.
- [49] S. S. Chaturvedi, K. Gupta, and P. S. Prasad, “Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using MobileNet,” in *Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications*, pp. 165–176, Cairo, Egypt, 2020.