International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Improving Skin Lesion Classification and Prediction through Data Augmentation for Enhanced Accuracy

V.Auxilia Osvin Nancy*[1], Meenakshi S Arya[2], Rajasekar V[3]

Submitted: 29/01/2024 **Revised**: 07/03/2024 **Accepted**: 15/03/2024

Abstract: Early identification of skin cancer is crucial for improved survival rates, emphasizing the need for accurate computer-aided systems in the diagnostic process. This paper presents a deep learning-based system designed for efficient classification and prediction of skin cancer through the analysis of skin lesions. The proposed system utilizes datasets; HAM10000 collected from the ISIC archive and addresses the class imbalance problem within the skin lesion classes through augmentation techniques. The balanced and processed dataset is used to train and test a fine-tuned 5-layered CNN model with optimized parameters. The evaluation of the model's performance is determined solely by its accuracy and other performance metrics specific to the skin lesion classes. Furthermore, the article includes a comparative analysis and visualization of the balanced and unbalanced datasets to provide insights into their characteristics. The proposed deep learning system offers promising potential for enhancing skin cancer diagnosis by accurately classifying skin lesions and predicting malignancy.

Keywords: *5 layered CNN, ISIC archive, Skin lesion, Malignancy, Augmentation*

1. Introduction

Melanoma is a prevalent and rapidly growing type of cancer worldwide. The identification of factors contributing to the development of the disease includes both abiotic and biotic elements [1]. Early detection is crucial for effective skin cancer treatment [2], as it significantly improves the 5-year survival rate. [3]. Delayed identification of melanoma leads to increased treatment costs [4]. Assessing dermoscopic photos is often considered time-consuming task for dermatologists and is susceptible to inaccuracies [5]. Visual assessment of skin lesions requires procedures such as dermoscopy and biopsy. However, the success rate of visual inspection, particularly in detecting melanoma tumors, can drop significantly, sometimes as low as 80 percent depending on the dermatologist's expertise [6].

Deep learning techniques in healthcare have shown promise in disease identification and prediction [7]. These techniques provide efficient results by minimizing the impact of humans. A key benefit of deep learning is its ability to diagnose without the need for invasive surgical procedures, thereby reducing computation time and diagnostic delays [8]. Surgical interventions are time-consuming and can cause discomfort for patients [9].

Melanoma refers to the development of abnormal moles from melanocytes [10]. The ISIC dataset is commonly used for skin lesion classification and detection, as it offers a wide range of cell types in skin lesions. Fig.1 [11] lists the categories of cutaneous lesions.

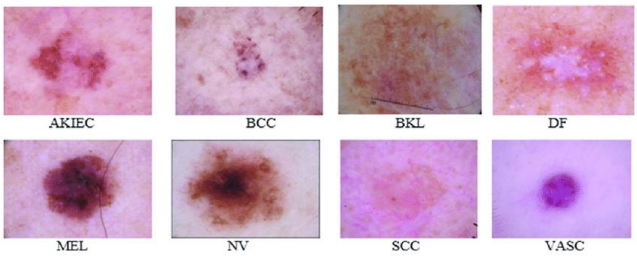


Fig.1 Skin lesion Classes [11]

The classification of skin lesions poses a significant challenge, and deep learning algorithms have shown promise in addressing this issue. However, several factors contribute to the complexity of the classification task. Firstly, the distribution of skin lesion classes in the dataset is heavily skewed, with certain classes such as NV (non-melanocytic nevi) being more prevalent compared to classes like SK (seborrheic keratosis) and MEL (melanoma). This class imbalance makes it difficult to accurately classify the different types of skin lesions. Additionally, the presence of intra-class and inter-class differences and similarities further complicates in identifying the types of skin lesion.

Pre-trained networks may not effectively address the specific concerns related to skin cancer classification. To overcome this challenge, the model is proposed to develop efficient network. The approach involves fine-tuning model with five layers of CNN. This model extracts essential features by passing the input through the connected blocks, utilizing various parameters. Both the original dataset and

[1] *Dept. of Computer science and Engineering, Rajalakshmi Engineering College, Chennai, Tamil Nadu, India*
ORCID ID : 0000-0002-4254-0537
[2] *Department of Transportation, IOWA State University, IOWA, United States*
ORCID ID : 0000-0001-6631-8810
[3] *Dept. of Computer science and Engineering, College of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani Campus,Chennai-26, Tamil Nadu,India*
** Corresponding Author Email: auxilianancy@gmail.com*

the augmented dataset are utilized in the experiments.

By addressing the class imbalance issue and leveraging the proposed CNN model with fine-tuning and data augmentation, the research is to increase the efficiency of skin lesion classification.

2. Relevant work

This section of the research examines the utilization of data augmentation techniques and CNN architectures for the purpose of skin lesion categorization. Valuable suggestions taken from research articles shed light on potential avenues for further analysis and improvement.

To enhance network performance, TJ Brinker et.al. [12] and Gessert et.al. [13] propose the incorporation of patient metadata during CNN training, which has shown promising results in improving classification accuracy for skin lesions.

Rezvantalab et al. [14] conducted skin lesion categorization using pre-trained CNN models on ImageNet. They employed four different CNN architectures and utilized two publicly available datasets, HAM10000 and PH2. The study revealed challenges in effectively differentiating between NV (Melanocytic Nevi) and MEL (Melanoma) classes, while also facing difficulties in predicting AKIEC and BKL.

Tschandl et al. [15] addressed class imbalance and improved model performance through data augmentation techniques implemented during CNN training. They observed significant enhancements in output metrics by combining various data augmentation methods with softmax adjustments across three different datasets.

In the work of Hosny et al. [16], AlexNet demonstrated remarkable accuracy over 95% in the classification of skin lesions. The study included three datasets, specifically ISIC 2017, MED-NODE, and DermIS, and showcased that implementing various data augmentation approaches and softmax changes resulted in significant enhancements in network performance.

Gessert et al. [13] Employed Efficient Net models trained on high-resolution pictures from the ISIC 2019 dataset to enhance network performance. By utilizing the scalability of EfficientNet, which adjusts the breadth and depth of the model according to the input size, and implementing loss balancing to tackle class imbalance, a notable enhancement in classification accuracy was achieved.

Hekler et al. [17] studied how label noise affects Convolutional Neural Networks (CNNs) in classifying skin cancer. They highlighted the importance of accurate ground truth data and examined the potential errors caused by non-biopsy-verified training images. The study compared the performance of trained models with that of dermatologists to assess their quality.

Zhang and colleagues [18] proposed a technique for emulating deep convolutional neural network similarity evaluation. This strategy entails generating a significant number of samples from a restricted number of individuals. The goal is to enhance the balance of the dataset and enhance the dependability of the outcomes. The authors Sucholutsky et al. [19] proposed a data augmentation method that enhanced the performance of the network by effectively extracting crucial characteristics from photos of skin lesions.

While these studies provide valuable insights and suggestions for further analysis, it is important to address the limitations identified, such as small training and testing image sets, reliance on a single architecture, and biases associated with biopsy-verified lesions. By overcoming these limitations and conducting further investigations, the accuracy and reliability of skin lesion classification using CNNs can be significantly improved.

3. Methodologies

Deep learning's capacity to efficiently evaluate vast volumes of medical data, has made it a prominent technique in medical diagnostics, particularly the categorization of skin cancer. These algorithms successfully identify and classify individuals with skin cancer, which can help clinicians provide better care. Improving diagnosis accuracy and categorization performance are the main objectives of applying deep learning models to skin cancer classification. Typically, a range of deep learning models are used to build an automated skin cancer classification system to accomplish these goals. The optimal model for the job is then identified by evaluating these models using criteria like accuracy, precision, recall, and F1 score.

This study's improved method for automatically classifying skin cancer entails dividing dermatoscopic picture data into four groups. The method of developing the model consists of four steps: The steps involved in creating a skin cancer dataset are as follows: (1) gathering a set of dermoscopic images. The process consists of three steps: first, preparing the dataset by rescaling, resizing, and normalising the data; second, extracting features by layer-wise relevance propagation; and finally, building a classifier by utilising optimisation techniques on the extracted feature vectors.

3.1. Datasets

The focus in this section is on the description of datasets suitable for classification and prediction, HAM10000 [20]. Each dataset is outlined below.

The ISIC challenge dataset is widely used in melanoma categorization research. The main goal of using this information is to aid in the automated detection of melanoma and other malignant diseases [20]. The ISIC databases contain images spanning the years 2018 to 2020, with a significant increase in the quantity of photographs

each year. For classification purposes, HAM10000 is considered.

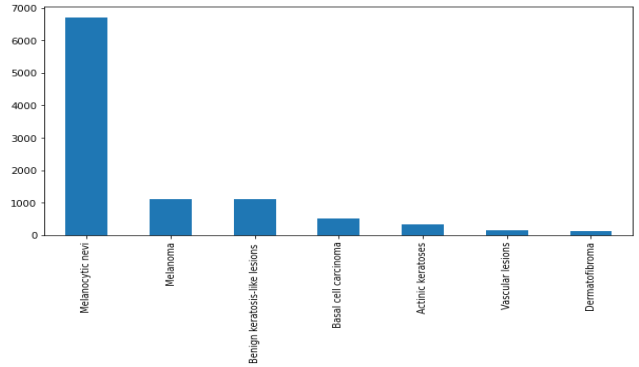


Fig.2 Count of Samples per class

The HAM10000 dataset has 10015 pictures that are labeled with seven distinct lesion classes [21]. Figure 3 shows some photos from the HAM10000 dataset.

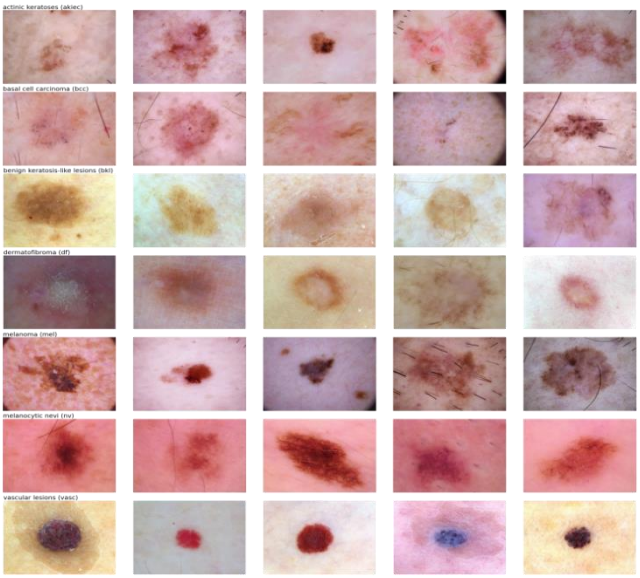


Fig. 3 Sample Images-HAM10000 [21]

3.2. Data Pre-processing

Contrast stretching, sometimes referred to as normalisation, is a method for enhancing an image's contrast that improves the image's visual quality. This method normalises the pixel intensity values in the image so that the highest and lowest intensity values are matched to the corresponding values in the output image.

3.3. Augmentation

Data augmentation is the process of adding variations or modifications to existing data to artificially increase the dataset size. By augmenting the data, additional examples are created, which helps in improving the model's generalization and robustness. Data augmentation techniques such as rotation, scaling, flipping, and adding noise can introduce diversity and capture different variations in the data. This leads to a more comprehensive and representative dataset, enabling the model to learn a broader range of patterns and improve its performance.

In the context of skin lesion classification and prediction, data augmentation techniques can be employed to generate variations in images, such as changing the position, size, or color, which can aid in improving the accuracy and robustness of the trained CNN model.

3.4. Data Augmentation techniques

3.4.1. Data Up-sampling

Data up-sampling is a technique used to increase the number of samples in the minority class. Oversampling or generating new instances of the minority class is used to balance the class distribution. Data up-sampling addresses the class imbalance issue by boosting the representation of the minority class, thereby offering the model additional training instances to learn from.

3.4.2. Cost-Sensitive Learning

Cost-sensitive learning is a technique that assigns suitable weights to different classes based on their sizes or importance. The concept involves assigning more weights to the minority class and lesser weights to the majority class, hence highlighting the significance of the minority class during the training of the model. The weight assigned to each class is determined using equation (1).

$$Weight_j = \frac{\sum_{i=1}^{7} n_i}{n_j} \quad \text{-------- (1)}$$

$$n_j = samples\ of j^{th}\ class\ of\ skin\ lesion$$

3.4.3. Random sampling

Random resampling strategies entail altering the dataset by oversampling the minority class or undersampling the majority class.

Random oversampling is a method that deals with class imbalance by increasing the number of instances in the minority class. This facilitates achieving a more equitable class distribution and offers more training data for the minority class.

Random under sampling involves randomly selecting instances from the majority class and then removing them from the training dataset. This method reduces the dominance of the majority class, addressing the issue of class imbalance by achieving a more balanced distribution of classes.

Data augmentation involves oversampling the minority class and under sampling the majority class. The strategy involves determining the value for each class or label to be determined from the dataset, based on the values of the minority class. The code below demonstrates the implementation of oversampling and under sampling techniques utilizing strategy values.

```
strategy_over = {0:2000, 1:2000, 2:2000, 3:1000, 5:1000, 6:2000}
oversample = RandomOverSampler(sampling_strategy=strategy_over)
undersample = RandomUnderSampler(sampling_strategy = {4: 3000})
```

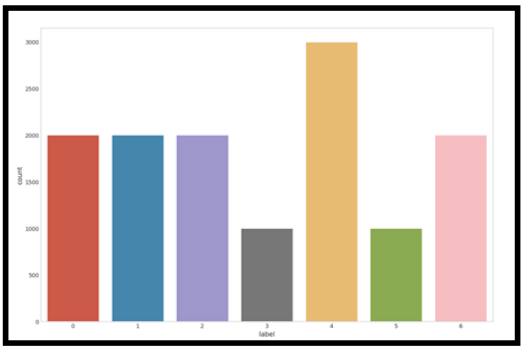


Fig.4 Random sampling

3.4.4. Random SMOTE

SMOTE is a commonly utilised technique for dealing with imbalanced datasets by creating artificial training instances for the minority class by linear interpolation. Synthetic training instances are created by randomly selecting one or more of the k-nearest neighbours for each example in the minority class. After oversampling, the data is recovered and can be classified using several methods. The main concept can be summarised as:

Algorithm SMOTE(N, T, k)
Input:
Number of minority class samples N;
Amount of SMOTE T%; Number of nearest neighbors k

Output: *(T/100) * N* synthetic minority class samples
1. (*Randomize the minority class samples if T is less than 100% so only a random percentage of them will be SMOTEd. *)
2. **if** T < 100
3. **then** Randomize the N minority class samples
4. *N= (T/100) *N*
5. *T = 100*
6. **endif**
7. *T= (int)(T/100)* (* The amount of SMOTE is assumed to be in integral multiples of 100. *)
8. k = Number of nearest neighbors
9. Num_Attr = Number of attributes
10. S[ ][ ]: array for original minority class samples
11. new_index: Number of synthetic samples generated, initialized to 0
12. Synthetic_sample[ ][ ]: array for synthetic samples (*Only for each sample of a minority class, compute the k nearest neighbours.*)
13. **for** i ? 1 to N
14. Compute k nearest neighbors for i, and save the indices in the narray
15. Populate(T, i, narray)
16. **endfor**
Populate(T, i, narray) (* Function to generate the synthetic samples. *)
17. **while** T ?0
18. Call it nn and choose a random number between 1 and k. During this stage, one of the k closest neighbours of i.
19. **for** attr ? 1 to Num_attr
20. Compute: dif = S[narray[nn]][attr] - S[i][attr]
21. Compute: gap = random number between 0 and 1
22. Synthetic_sample[new_index][attr] = S[i][attr] + gap * dif
23. **endfor**
24. new_index++
25. T = N- 1
26. **endwhile**
27. **return** (* End of Populate. *)
End of Pseudo-Code.

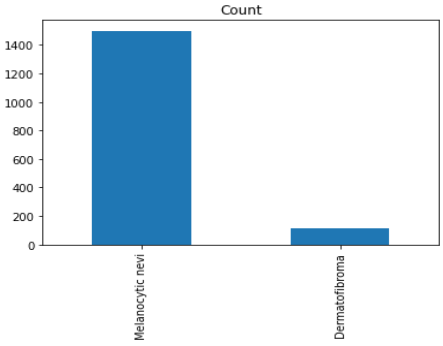


Fig.5 Combine majority class with downsampled minority class

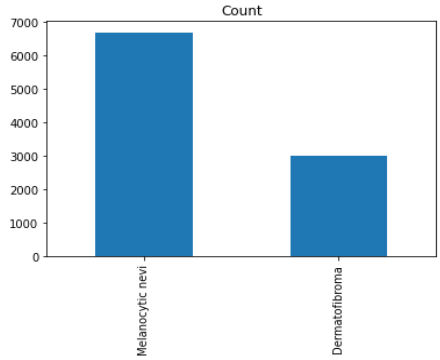


Fig.6 Combine majority class with upsampled minority class

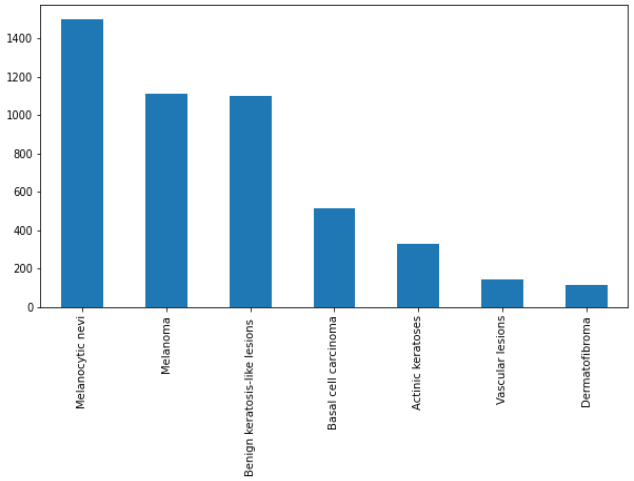


Fig.7 Resampled dataset

3.4.5. Augmentation with image data generator and normalize with class weight

Several augmentation techniques, including rotation, flipping, and noise addition, are applied in this work. As flipping and rotating are scale-invariant operations, the image's scale remains constant. The photographs are rotated 15 and 45 degrees clockwise and anticlockwise before being horizontally flipped. This leads to an increased number of samples in the dataset and introduces variations in the image orientation.

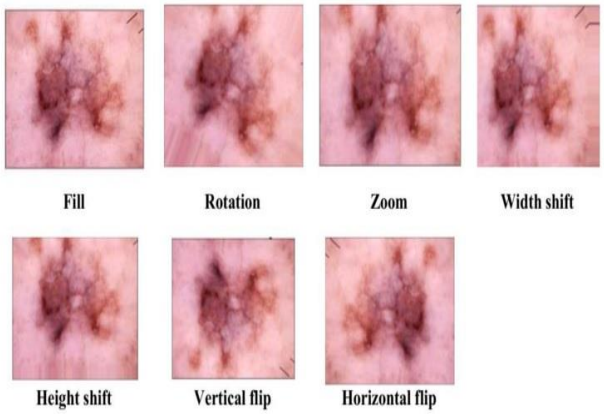


Fig.8 Augmentation with Image generator

Images have been generated by rotation and flip and the augmented dataset has 26565.The weights for the loss function have been computed as the dataset faces the data imbalance problem. Class weight has been calculated by

$$\boldsymbol{n_{samples}}/(\boldsymbol{n_{classes}} * \boldsymbol{np.bincount(y)}) \text{ --------(2)}$$

The class weight of the skin lesion class is tabulated in Table.1

Table.1 Normalize class using Class weight.

Lesion type	Lesion class	Class weight
Nv	melanocytic nevi	0.57
Mel	Melanoma	0.57
Bkl	benign keratosis-like lesions	0.58
Bcc	basal cell carcinoma	1.23
Akiec	Actinic keratosis	1.93
Vasc	Vascular lesions	4.45
Df	Dermatofibroma	5.5

3.4.6. Extreme learning machine

One approach to augmenting data using ELM is through random projections. This involves generating random vectors and projecting the original data onto these vectors. This can introduce variations in the data without requiring extensive computational resources. ELM can also be used to learn a transformation of the original features. By mapping the original features to a different space, new representations can be generated, effectively augmenting the dataset.

3.4.6.1. Algorithm

Input data matrix: $X \in R^{n \times m}$, where n is the number of samples and m is the number of features.

Output labels (for classification): $Y \in R^{n \times c}$, where c is the number of classes.

Weight matrix of input-hidden layer connections: $W \in R^{h \times m}$, where h is the number of hidden neurons.

Bias vector for the hidden layer: $b \in R^{h \times 1}$.

Activation functions for the hidden layer σ.

The output of the hidden layer, often denoted $H \in R^{n \times h}$, is calculated as follows:

$H=\sigma(XW^T+b)$

Here, W^T represents the transpose of the weight matrix W.

The output of the output layer, denoted as $O \in R^{n \times c}$, is calculated as:

$O=HV^T$

3.4.7. WELM-Weighted Extreme Learning Machine

Addressing imbalanced data by allocating greater weight to the minority class and lesser weight to the dominant class. The Enhance Remora Optimization algorithm has been employed to optimize the three parameters of WELM: input weight, bias, and the weight of imbalanced training data. The challenges posed by class imbalance can be addressed by utilizing a cost-sensitive variant of ELM. The training samples are assigned weights using generalized weighting approaches that rely on the class distribution. The weighted Extreme Learning Machine (ELM) can be extended to incorporate cost-sensitive learning. WELM optimization problem defined by

$$Mini_{mize} = \frac{1}{2} \| \alpha \|^2 + \frac{1}{2} DV \sum_{j=1}^{M} \| \tau_j \|^2) \quad \text{----------(3)}$$

α- output layer weight $g(y_j)$ - Hidden layer output y_j – Input feature vector τ_j - Training error vector P_j - output class label vector D – Regularization parameter

Equation exposed to

$$g(y_j)\alpha = P_j^p - \tau_j^\tau \quad \text{----------(4)}$$

j =1,2,…….M

3.4.7.1. Two weighting Techniques proposed.

MXM – Matrix

$V=diag(v_{jj})$ – The training samples represented by a matrix with diagonal members are given weights.

1. $V_{jj} = \frac{1}{fl}$ l =p_j [Total number of samples in the l^{th} class is fl]

2. $V_{jj} = \frac{0.618}{fl} if(fl \leq f_{avg})$

$\frac{1}{fl} if(fl \leq f_{avg})$ f_{avg} – Mean number of samples per class

j^{th} - The samples are assigned a weight of V_{jj}

In both weighting procedures, samples from the minority class will be assigned weights equal to 1/fj. The second

weighing approach assigns less weight to samples belonging to the majority class, in contrast to the first weighting scheme.

3.4.7.2. Algorithm for WELM

Step 1: Weighted Extreme Learning Machine (WELM), a diagonal matrix W is calculated, and weights are assigned to each instance of a class. Each element of the matrix represents the diagonal vector of a certain instance xi.

Step 2: The chosen activation function for WELM is the sigmoid function.

Step 3: Employ the suggested weighting technique to determine the magnitude of the lesion category.

Step 4: Assign the most effective weights and bias to the Weighted Extreme Learning Machine (WELM) to achieve optimal classification accuracy. The weight has been calculated for the lesion class using WELM.

Table.2 Class weight using WELM.

Lesion type	Lesion class	Class weight
Nv	melanocytic nevi	0.000092
Mel	Melanoma	0.000092
Bkl	benign keratosis-like lesions	0.000095
Bcc	basal cell carcinoma	0.00032
Akiec	Actinic keratosis	0.0005
Vasc	Vascular lesions	0.0011
Df	Dermatofibroma	0.0014

3.5. Classification

3.5.1. Proposed CNN

The suggested model is a sequential model consisting of five convolution layers, one flattened layer, and three dense layers. A max pooling layer immediately follows each convolution layer. The initial convolution layer defines the input shape, kernel size, and activation function (ReLU), and this procedure is iterated four times. The flattened layer is utilized to transform the two-dimensional array into a one-dimensional array. The flattened layer's output is subsequently sent to the dense layers. The final dense layer specifies the number of output classes and the activation function. Figure 8 presents a comprehensive outline of the architecture of the CNN model that has been suggested. These pictures provide a graphical depiction of the organizational structure of the proposed CNN model, illustrating the arrangement of its various layers and their interconnections.

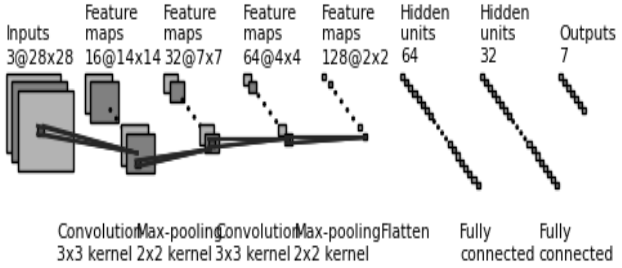


Fig. 9 Proposed CNN

3.5.2. System Flow without and with Augmentation

The datasets HAM10000 undergo preprocessing, including duplicate removal and ground truth table conversion. They are then split into train, test, and validation sets. Fig.10 visualize the system flow without and with augmentation.

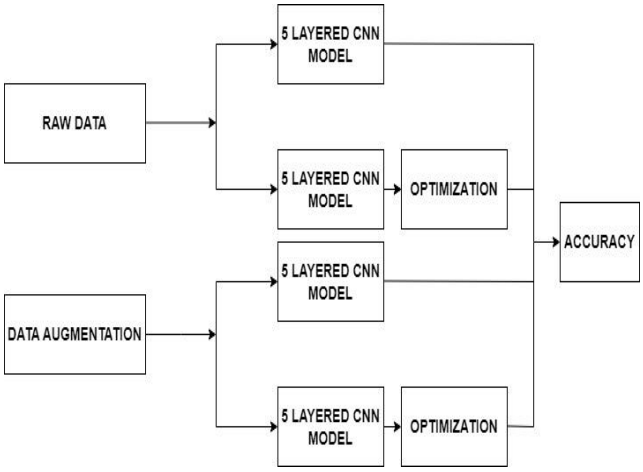


Fig.10 System flow without and with augmentation

3.5.3. Evaluation Metrics Used

Various metrics are accessible in deep learning algorithms for the classification process. The raw dataset is analyzed using metrics such as accuracy, precision, recall, and F1 score. To enhance the model's overall forecast, various metrics are assessed.

3.5.3.1. Accuracy

The ratio of correct forecasts to the total number of predictions.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

3.5.3.2. Precision

The accuracy of predicting circumstances is positively correlated with precision, resulting in a percentage of successfully anticipated situations. False positives are of more concern than false negatives.

$$Precision = \frac{TP}{TP + FP}$$

3.5.3.3. Recall

Recall quantifies the model's ability to properly predict the actual number of positive cases.

$$Recall = \frac{TP}{TP + FN}$$

3.5.3.4. F1-Score

The F1-Score is a metric that combines precision and recall measurements. When the values of recall and precision are the same, the F1 score reaches its highest possible value.

$$F1 = 2.\frac{Precision\ X\ Recall}{Precision + Recall}$$

4. Findings and Analysis

The experimental results demonstrate the successful implementation of the proposed CNN model for finding the suitable augmentation technique for skin lesion classification using HAM10000 dataset. A comparative analysis is conducted between raw and augmented datasets. The results are visualized to evaluate the performance of the classification. In the HAM10000 dataset, the label values represent various skin lesion classes. Fig.11 shows the pre-processed dataset.

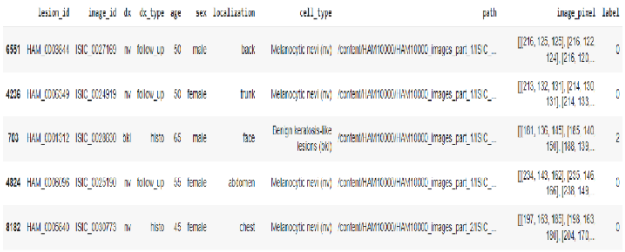


Fig. 11 Pre-processed Data-HAM10000 Dataset

The model's correctness is evaluated by analyzing its performance in training, testing, and validation. The accuracy score is derived from the count of accurate predictions made. Loss numbers represent the deviation from the intended target state(s). The precision of HAM10000 is 73.55%. The model's accuracy and losses prior to augmentation are displayed in Figure 12. The pre-augmentation model's efficiency, as measured by accuracy, is presented in Table 3.

Table.3 Before Augmentation – Accuracy

Accuracy	Value
Training	0.932
Validation	0.7355
Test	0.74039

Fig. 12 Accuracy and Loss before Augmentation

Table.5 Classification Report of skin lesion

SKIN LESION CLASS	HAM10000		
	PRECISION	RECALL	F1SCORE
NV	0.85	0.94	0.89
MEL	0.45	0.3	0.37
BKL	0.54	0.41	0.46
BCC	0.44	0.61	0.51
AKIEC	0.49	0.33	0.39
VASC	0.76	0.57	0.65
DF	0	0	0

4.1. Prediction Of Skin Lesion Class Without Augmentation

The model's accuracy on the HAM10000 dataset, without augmentation, is 73.55%. Despite the unbalanced distribution of skin lesion classes, the model's performance on the datasets is below 75%. The NV class in the ISIC archive dataset contains a greater quantity of photos compared to the other classes. As a result, the model produces inaccurate forecasts and operates with reduced efficiency. To enhance the model's performance, it is necessary to achieve a balanced dataset, which can be accomplished through augmentation techniques. The anticipated and projected classification of skin lesions in the HAM10000 dataset is depicted in Figure 13, revealing two instances of incorrect class predictions.

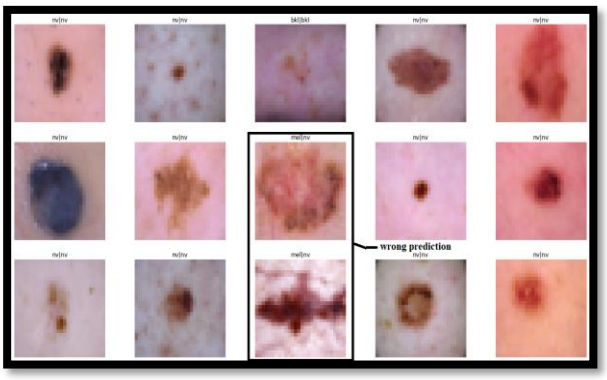


Fig.13 Expected and Predicted lesion without augmentation.

4.2. After Augmentation the Evaluation Metrics of the Model

To enhance the model's performance, the dataset is expanded according to the specific needs of the lesion classes. The inclusion of the specific lesion class alone contributed to achieving a balanced dataset. In Table.6 and Table.7, the results have been visualized with evaluation metrics by comparing different augmentation techniques.

4.3. Comparison On Augmentation Techniques for Class Imbalance Problem with Existing Study

The results were discussed with existing study and other augmentation techniques and identified that the proposed augmentation technique using WELM has been proved with highest accuracy of 98.75% and the result has been visualized in Table.8.

Table.8 Comparison with existing study

AUGMENTATION TECHNIQUE	ACCURACY(%)
Random Resampling	97.88
SMOTE	91.31
Image Generator with Class Weight	88.8
WELM(DSSA) Panneerselvam, R., & Balasubramaniam, S. (2023)	97.25
Weighted Extreme Learning Machine (WELM) With Weighting Proposed Technique	98.75

5. Conclusion and Future Work

Skin cancer remains a significant health concern globally, and early detection plays a crucial role in improving patient survival rates. However, there are challenges related to the availability of highly trained doctors and advanced medical technology, which can hinder effective skin cancer detection and treatment. To address this issue, the study focused on exploring state-of-the-art research projects that utilize Convolutional Neural Network (CNN) architectures to enhance skin cancer detection and classification. Various ISIC datasets were used to train the CNN model, and a comparative analysis of results was conducted to mitigate the limited availability of labeled data. To solve the data imbalance problem various augmentation techniques were analyzed to find the suitable technique for class imbalance. HAM10000 datasets were analyzed with augmentation techniques like, random resampling, SMOTE, normalize the class weight and by weighted extreme learning machine. The results show that WELM has the high accuracy when compared with other augmentation technique. The processed dataset trained with the proposed CNN model for classification. The study highlights the potential of augmentation technique in the future to further improve the effectiveness of the CNN model. This approach can help address the scarcity of labeled data and enhance the accuracy and reliability of skin cancer detection and classification methods. By leveraging advanced technology and research advancements, the aim is to facilitate early detection and improve outcomes for individuals affected by skin cancer.

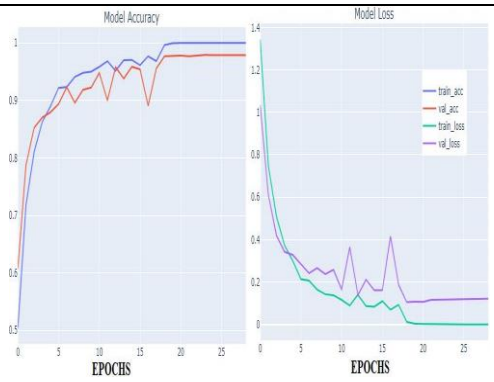
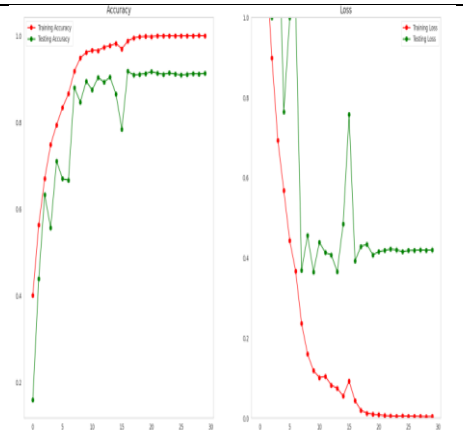
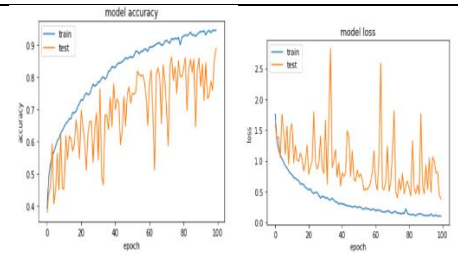
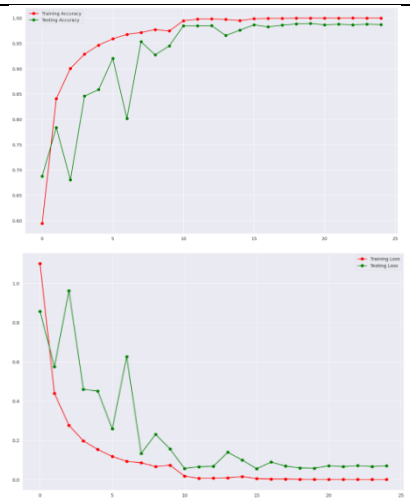
<table>
<tr><th>Augmentation Technique</th><th>Accuracy and Loss graph</th><th colspan="2">Validation</th><th>Classification Report</th></tr>
<tr><td rowspan="2">Random Resampling</td><td rowspan="2"> 
</td><td>Acc</td><td>Loss</td><td rowspan="2"><table>
<tr><th>Class</th><th>Precision</th><th>Recall</th><th>F1score</th></tr>
<tr><td>NV</td><td>0.99</td><td>0.88</td><td>0.93</td></tr>
<tr><td>MEL</td><td>0.95</td><td>0.99</td><td>0.97</td></tr>
<tr><td>BKL</td><td>0.95</td><td>0.99</td><td>0.97</td></tr>
<tr><td>BCC</td><td>0.99</td><td>1.00</td><td>0.99</td></tr>
<tr><td>AKIEC</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
<tr><td>VASC</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
<tr><td>DF</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
</table></td></tr>
<tr><td>97.88</td><td>0.12</td></tr>
<tr><td>SMOTE</td><td> 
</td><td>91.31</td><td>0.42</td><td><table>
<tr><th>Class</th><th>Precision</th><th>Recall</th><th>F1score</th></tr>
<tr><td>NV</td><td>0.90</td><td>0.83</td><td>0.86</td></tr>
<tr><td>MEL</td><td>0.86</td><td>0.88</td><td>0.87</td></tr>
<tr><td>BKL</td><td>0.86</td><td>0.85</td><td>0.86</td></tr>
<tr><td>BCC</td><td>0.97</td><td>0.99</td><td>0.98</td></tr>
<tr><td>AKIEC</td><td>0.97</td><td>1.00</td><td>0.98</td></tr>
<tr><td>VASC</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
<tr><td>DF</td><td>0.99</td><td>1.00</td><td>1.00</td></tr>
</table></td></tr>
<tr><td>Normalize With Class Weight</td><td> 
</td><td>88.80</td><td>0.37</td><td><table>
<tr><th>Class</th><th>Accuracy</th></tr>
<tr><td>NV</td><td>0.86</td></tr>
<tr><td>MEL</td><td>0.90</td></tr>
<tr><td>BKL</td><td>0.88</td></tr>
<tr><td>BCC</td><td>0.94</td></tr>
<tr><td>AKIEC</td><td>0.85</td></tr>
<tr><td>VASC</td><td>0.98</td></tr>
<tr><td>DF</td><td>0.87</td></tr>
</table></td></tr>
<tr><td>WELM</td><td>  </td><td>98.75</td><td>0.07</td><td><table>
<tr><th>Class</th><th>Precision</th><th>Recall</th><th>F1score</th></tr>
<tr><td>NV</td><td>0.99</td><td>0.92</td><td>0.96</td></tr>
<tr><td>MEL</td><td>0.96</td><td>0.99</td><td>0.98</td></tr>
<tr><td>BKL</td><td>0.97</td><td>0.99</td><td>0.98</td></tr>
<tr><td>BCC</td><td>0.99</td><td>1.00</td><td>1.00</td></tr>
<tr><td>AKIEC</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
<tr><td>VASC</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
<tr><td>DF</td><td>1.00</td><td>1.00</td><td>1.00</td></tr>
</table></td></tr>
</table>

Table.6 Evaluation Metrics of the Model

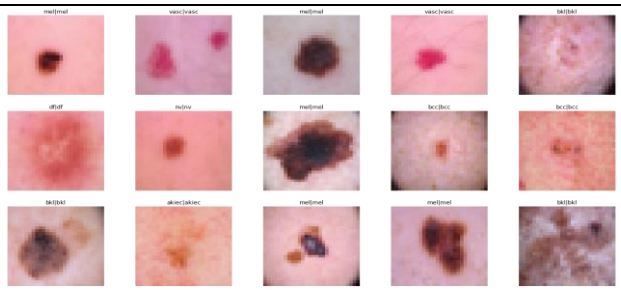
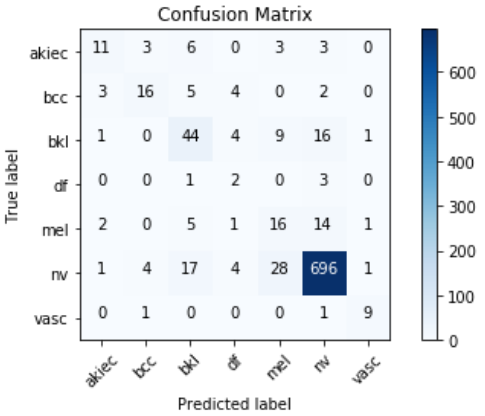
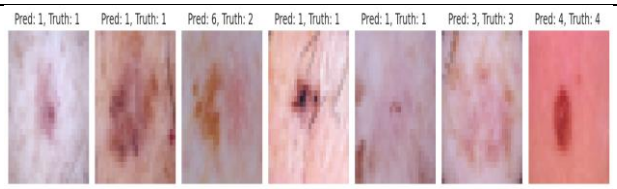
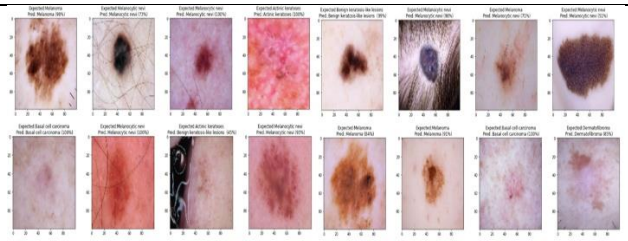
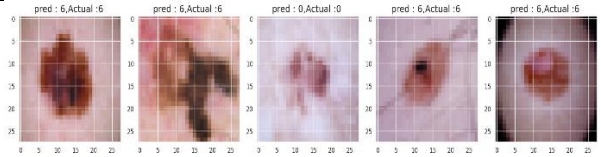
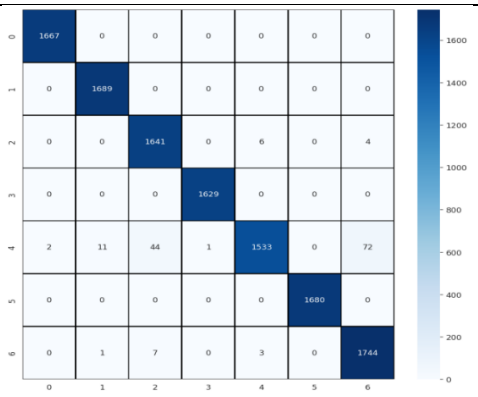
Augmentation Technique	Expected and Predicted Skin lesion	Confusion Matrix
Random Resampling		
SMOTE		
Normalize With Class Weight		Total-Test-Data:8767 Accurately-Predicted-Data: 7785 Wrongly-Predicted-Data: 982
WELM		

Table.7 Expected and Predicted Skin lesion-Augmentation Technique

Author contributions

V.Auxilia Osvin Nancy: Conceptualization, Methodology, Software, Field study **Meenakshi S Arya:** Data curation, Writing-Original draft preparation, Software, Validation., Field study **Rajasekar V:** Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Woo YR, Cho SH, Lee JD, Kim HS. The Human Microbiota and Skin Cancer. International Journal of Molecular Sciences. 2022; 23(3):1813. https://doi.org/10.3390/ijms23031813

[2] Akyel C, Arıcı N. LinkNet-B7: Noise Removal and Lesion Segmentation in Images of Skin Cancer. Mathematics. 2022; 10(5):736. https://doi.org/10.3390/math10050736

[3] Ünver, Halil Murat, and Enes Ayan. "Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm." Diagnostics 9, no. 3 (2019): 72.

[4] McNoe, Bronwen M., Kate C. Morgaine, and Anthony I. Reeder. "Effectiveness of Sun Protection Interventions Delivered to Adolescents in a Secondary School Setting: A Systematic Review." Journal of skin cancer 2021 (2021).

[5] Alom, Md Zahangir, Theus Aspiras, Tarek M. Taha, and Vijayan K. Asari. "Skin cancer segmentation and classification with NABLA-N and inception recurrent residual convolutional networks." arXiv preprint arXiv:1904.11126 (2019).

[6] Kadampur, Mohammad Ali, and Sulaiman Al Riyaee. "Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images." Informatics in Medicine Unlocked 18 (2020): 100282.

[7] Senan, Ebrahim Mohammed, and Mukti E. Jadhav. "Classification of dermoscopy images for early detection of skin cancer–a review." International Journal of Computer Applications 975 (2019): 8887.

[8] Gillmann, Christina, Dorothee Saur, and Gerik Scheuermann. "How to deal with Uncertainty in Machine Learning for Medical Imaging?." In 2021 IEEE Workshop on TRust and EXpertise in Visual Analytics (TREX), pp. 52-58. IEEE, 2021.

[9] Haggenmüller, Sarah, Roman C. Maron, Achim Hekler, Jochen S. Utikal, Catarina Barata, Raymond L. Barnhill, Helmut Beltraminelli et al. "Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts." European Journal of Cancer 156 (2021): 202-216.

[10] "Skin Cancer: Skin Cancer Facts: Common Skin Cancer Types." American Cancer Society. Accessed June 7, 2022. https://www.cancer.org/cancer/skin-cancer.html/.

[11] Kassem, Mohamed & Hosny, Khalid & Fouad, M.. (2020). Skin Lesions Classification Into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.3003890.

[12] Brinker TJ, Hekler A, Utikal JS, Grabe N, Schadendorf D, Klode J, Berking C, Steeb T, Enk AH, von Kalle C. Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review. J Med

[13] Internet Res. 2018 Oct 17;20(10):e11936. doi: 10.2196/11936. PMID: 30333097; PMCID: PMC6231861.

[14] Gessert, N., Nielsen, M., Shaikh, M., Werner, R., & Schlaefer, A. (2020). Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data. MethodsX, 7, 100864.

[15] Rezvantalab, A., Safigholi, H., & Karimijeshni, S. (2018). Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. arXiv preprint arXiv:1810.10348.

[16] Tschandl, P., Codella, N., Akay, B. N., Argenziano, G., Braun, R. P., Cabo, H., ... & Kittler, H. (2019). Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. The lancet oncology, 20(7), 938-947.

[17] Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2019). Classification of skin lesions using transfer learning and augmentation with Alex-net. PloS one, 14(5), e0217293.

[18] Hekler, A., Kather, J. N., Krieghoff-Henning, E., Utikal, J. S., Meier, F., Gellrich, F. F., ... & Brinker, T. J. (2020). Effects of label noise on deep learning-based skin cancer classification. Frontiers in Medicine, 7, 177.

[19] Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6848-6856).

[20] Sucholutsky, I., & Schonlau, M. (2020). Secdd:

Efficient and secure method for remotely training neural networks. arXiv preprint arXiv:2009.09155.

[21] Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... & Halpern, A. (2018, April). Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp. 168-172). IEEE.

[22] Tschandl, P. et al. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5:180161 doi: 10.1038/sdata.2018.161 (2018)

[23] What is a convolutional neural network? What is a Convolutional Neural Network? - MATLAB & Simulink. (n.d.). Retrieved June 14, 2022, from https://in.mathworks.com/discovery/convolutional-neural-network-atlab.html?s_tid=srchtitle_Convolutional+neural+network_1