Areej *et al.LGU (IJECI)* 2023

LGU (IJECI)
ISSN: 2522-3429 (Print)
ISSN: 2616-6003 (Online)

LGU International Journal for
Electronic Crime Investigation

Research Article

Vol. 7 issue 1 Year 2023

Skin Lesion Detection and Classification Using Deep Learning

Areej Fatima
Department of Computer Science, NCBA&E, Lahore, Pakistan.
Corresponding address: areejfatima@ncbae.edu.pk

Received: December 22, 2022; Accepted: January 30, 2023; Published: March 03, 2023

Abstract:

exposure to ionizing radiation (IR) can cause basal cell carcinoma (BCC) development. A skin lesion is a region that is differentiable from another skin surface which can occur because of skin damage, allergy, etc. Even though the majority of skin lesions are mild and are not that dangerous yet few of them are infectious and their severity can turn into skin cancer. In USA, 5.4 million people are analyzed with skin cancer. The diverse types of skin lesions result in an incorrect diagnosis because of their high similarity. Skin lesions can be treated by dermatologists. The current work proposes a model for the classification of skin lesions. The proposed methodology aims to detect and classify skin lesions using potential and different deep-learning algorithms. The research focuses to achieve state-of-the-art accuracy and compare the performance of algorithms.

Keywords: deep learning (DL), skin lesion, detection, classification, ionizing radiation

1. Introduction

NSkin cancer is common types of cancer which has a life-threatening effect. The atomic bombs exploded over Nagasaki and Hiroshima exposed the population to both neutrons and gamma rays. Epidemiological studies revealed that atomic bomb survivors living in Nagasaki and Hiroshima showed relationship between ionizing radiation and risk of skin cancer development (1). In united states, it has affected more than 9500 individuals on regular basis and around 3.6 million individuals suffered from basal cell skin cancer on annual basis. Most occurring form of skin cancer is Melanoma which grow in melanocytes cell and its severity can affect stomach, lungs and other body parts. According to report, the early detection of malignant melanoma can be treated 99% whereas; patients with progressive melanoma has 25% survival probability [2].

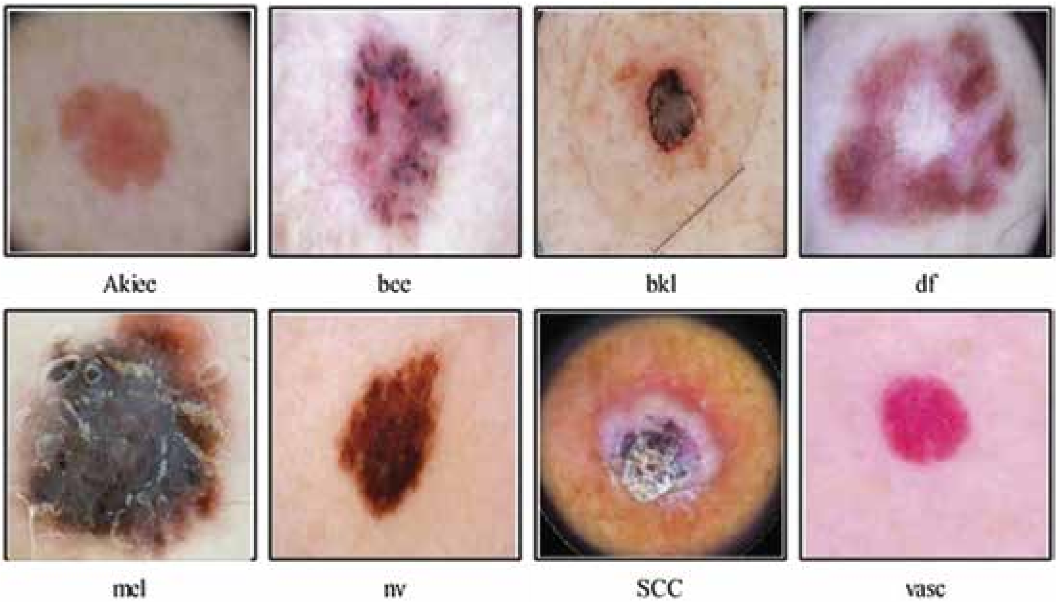


Figure 1: Sample images of Skin lesion from Ham1000 dataset

Early detection of skin lesion increases probability of patient's survival. So, the significance of detection and classification of different skin lesion has raised. The similar appearance of mild and severe skin lesion makes the detection and classification, a challenging task. The diagnostic task of skin lesion is based on ABCDE formula:

- A representing asymmetric skin surface
- B representing abnormal border
- C representing lesion color
- D representing lesion diameter
- E representing lesion enlargement

Different skin cancers seem alike with respect to above properties. There is a chance of error if detection is made via naked eye [3]. Dermoscopy is the significant method for detecting skin lesion as compared to clinical approach like biopsy etc. which is time taking procedure. But dermoscopy has some demerits as it is error prone. So, there is high need of effective technique for accurate detection with less error rate [4,5].

Deep learning (DL) is edge cutting technology which trains model for specified task more accurate as compared to machine learning. This research aims to utilize deep learning algorithm for skin lesion detection and classification through dermoscopic images as DL fast learning models are less error prone. Data augmentation is required as dataset is disproportional, it is done through affine transformation. Moreover, the model is cross-validated for effective performance [6].

2. Methodology

Figure 2 shows the proposed methodology of system.

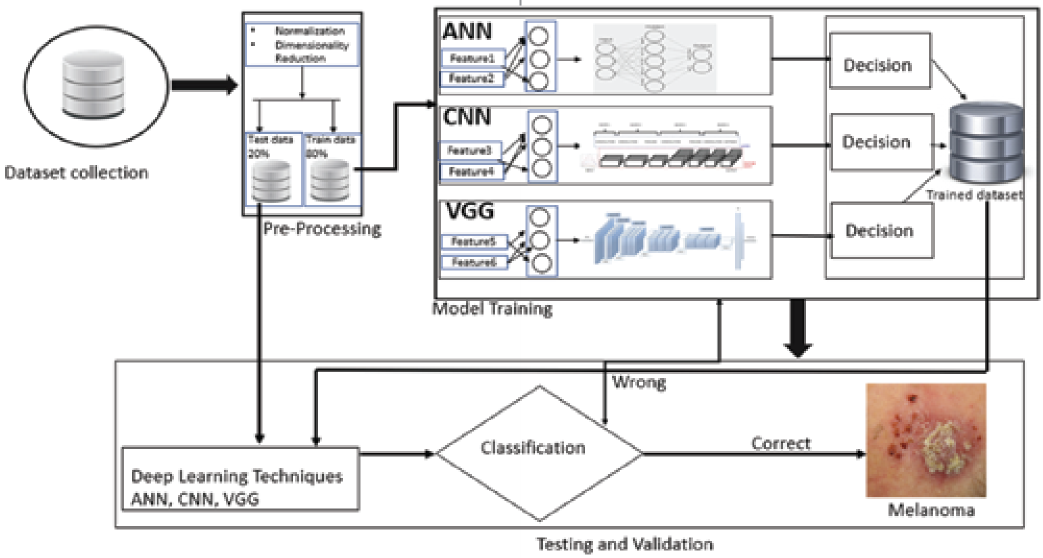


Figure 2: Proposed methodology

The methodology of the proposed research comprised of following phases:

2.1 Dataset collection

The targeted dataset is HAM1000 dataset. The HAM10000 ("Human Against Machine with 10,000 training images") dataset is a collection of multi-source dermatoscopic images of pigmented lesions,which is sourced from various populations and stored by using different techniques. It comprises of 10,015 images and 7 distinct categories of skin cancer. The seven categories of skin cancer are Melanocytic nevi, Melanoma, Dermatofibroma, Benign keratosis-like lesions, Actinic keratosis, Basal cell carcinoma, Vascular lesions.

2.2 Preprocessing

After collecting the target dataset, it is passed to a preprocessing phase where data cleansing is performed. The data underwent normalization through various methods, such as reducing missing values, resizing images, and properly labeling them. This normalization process is crucial for the success of the research as it minimizes the loss of information. This helps to concentrate on the region of interest. Figure 3 shows the dataset image size reduction to minimize the computational cost of the model.

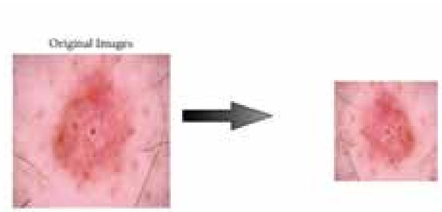


Figure 3: Image resizing

2.3 Training

After undergoing processing, the dataset was used to train three different neural network models. The training models include Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and Visual Geometry Group (VGG-16).

2.3.1. ANN

Artificial Neural Networks (ANNs) are mathematical models made up of interconnected processing units, also known as neurons. These neurons receive signals from other neurons, process the information by combining and transforming it, and produce a numerical output. The processing units in an ANN mimic the structure of biological neurons and are interconnected to form a network, creating the artificial neural network.

2.3.2. CNN

A Convolutional Neural Network (CNN) is a technique in computer vision that is designed to identify and distinguish the features of images. This architecture takes skin lesion images as input and passes them through a convolutional layer, where the weights are transformed into features. These features are further refined in the pooling layer before being transformed into a 1D representation in the fully connected layer. Finally, the features are classified using a Softmax layer.

2.3.3. VGG-16

The VGG16 architecture, a Convolutional Neural Network (CNN), won the 2014 ILSVRC (Imagenet) competition. It is considered to be one of the most advanced vision models developed to date. The VGG16 model consists of 16 layers and is known for its consistent placement of convolutional and max-pool layers throughout its architecture [7].

2.4. Model Validation

After training, the performance of the model is evaluated based on some parameters. It evaluates if the model is classifying the classes accurately or not. Figure 4 shows the block diagram of the proposed research.

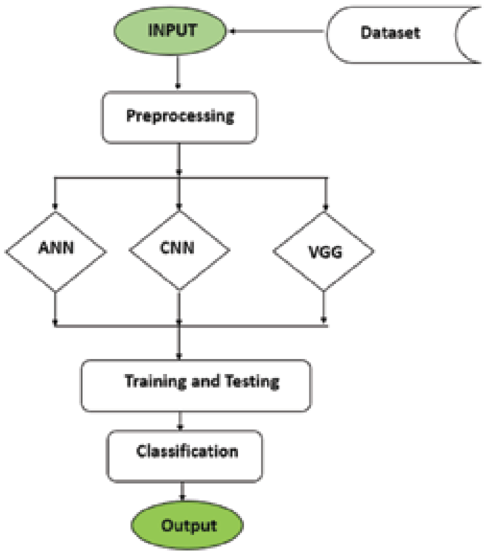


Figure 4: Block diagram of a system

3. RESULTS

In present section, the results of the proposed study are presented in numerical values and confusion matrices. Three classifiers were used for the experimental process, including Artificial Neural Network, Convolutional Neural Network, and Visual Geometry Group. First, the ANN model is considered which is comprised of three different layers (input, hidden, and output). Figure 5 shows the architecture of ANN.

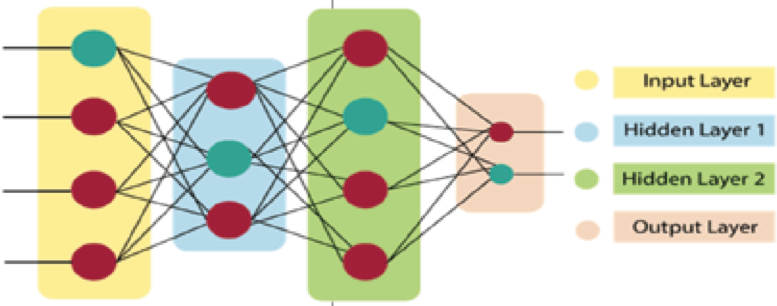
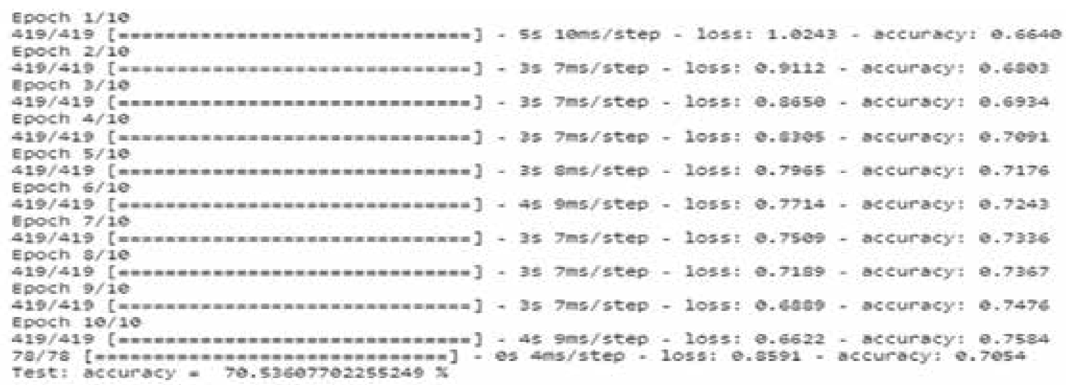


Figure 5: ANN architecture

ANN took input image from HAM1000 dataset and performance classification according to 7 different skin cancer classes. Figure 6 shows the accuracy of ANN model after 10 epochs. The average accuracy gained after ANN training is 70.5%



```
Epoch 1/10
419/419 [==============================] - 5s 10ms/step - loss: 1.0243 - accuracy: 0.6640
Epoch 2/10
419/419 [==============================] - 3s 7ms/step - loss: 0.9112 - accuracy: 0.6803
Epoch 3/10
419/419 [==============================] - 3s 7ms/step - loss: 0.8650 - accuracy: 0.6934
Epoch 4/10
419/419 [==============================] - 3s 7ms/step - loss: 0.8305 - accuracy: 0.7091
Epoch 5/10
419/419 [==============================] - 3s 8ms/step - loss: 0.7965 - accuracy: 0.7176
Epoch 6/10
419/419 [==============================] - 4s 9ms/step - loss: 0.7714 - accuracy: 0.7243
Epoch 7/10
419/419 [==============================] - 3s 7ms/step - loss: 0.7509 - accuracy: 0.7336
Epoch 8/10
419/419 [==============================] - 3s 7ms/step - loss: 0.7189 - accuracy: 0.7367
Epoch 9/10
419/419 [==============================] - 3s 7ms/step - loss: 0.6889 - accuracy: 0.7476
Epoch 10/10
419/419 [==============================] - 4s 9ms/step - loss: 0.6622 - accuracy: 0.7584
78/78 [==============================] - 0s 4ms/step - loss: 0.8591 - accuracy: 0.7054
Test: accuracy =  70.53607702255249 %
```

Figure 6: Accuracy of ANN

Figure 7 shows the confusion matrix of the ANN model.

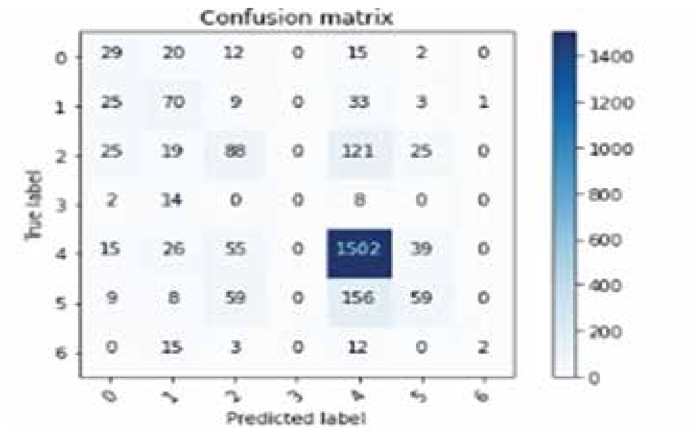


Figure 7: Confusion Matrix of ANN

Secondly, the CNN model is considered which is comprised of different layers. Figure 8 shows the general structure of CNN for skin lesion classification.

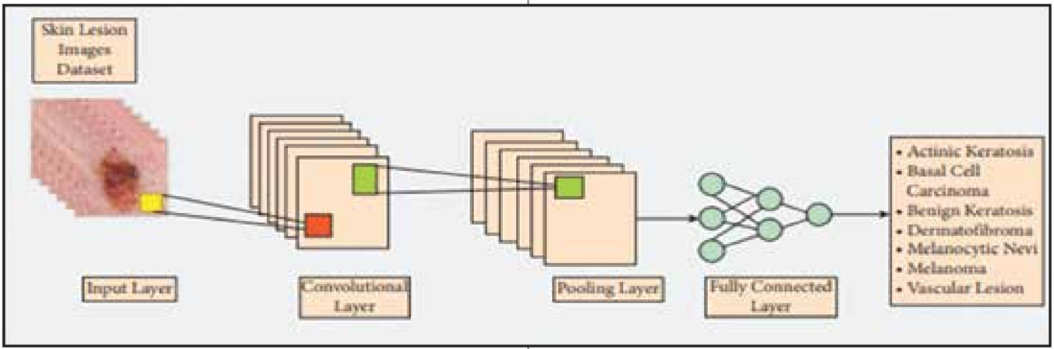
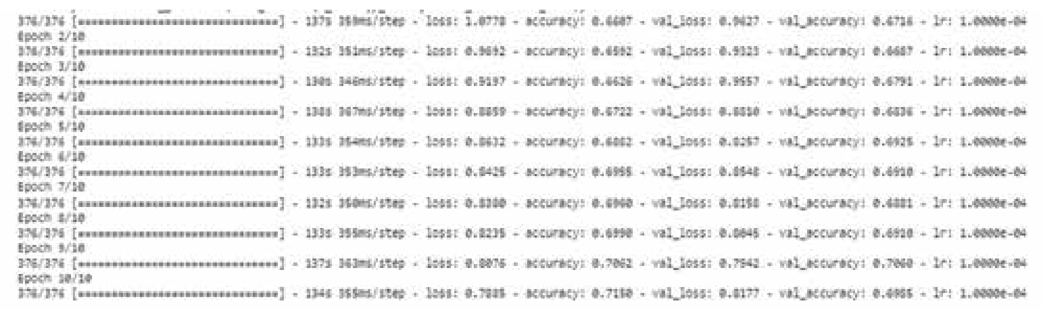


Figure 8: general structure of CNN model

CNN took input images from the HAM1000 dataset and pass them to convolutional, pooling, fully connected layers. After that, it performances classification according to 7 different skin cancer classes. Figure 9 shows the accuracy of CNN model after 10 epochs. The average accuracy gained after CNN training is 71.5%



```
376/376 [==============================] - 137s 358ms/step - loss: 1.0778 - accuracy: 0.6607 - val_loss: 0.9627 - val_accuracy: 0.6716 - lr: 1.0000e-04
Epoch 2/10
376/376 [==============================] - 132s 351ms/step - loss: 0.9692 - accuracy: 0.6592 - val_loss: 0.9323 - val_accuracy: 0.6687 - lr: 1.0000e-04
Epoch 3/10
376/376 [==============================] - 130s 346ms/step - loss: 0.9197 - accuracy: 0.6626 - val_loss: 0.9557 - val_accuracy: 0.6791 - lr: 1.0000e-04
Epoch 4/10
376/376 [==============================] - 138s 367ms/step - loss: 0.8859 - accuracy: 0.6722 - val_loss: 0.8850 - val_accuracy: 0.6836 - lr: 1.0000e-04
Epoch 5/10
376/376 [==============================] - 133s 354ms/step - loss: 0.8632 - accuracy: 0.6802 - val_loss: 0.8267 - val_accuracy: 0.6925 - lr: 1.0000e-04
Epoch 6/10
376/376 [==============================] - 133s 353ms/step - loss: 0.8425 - accuracy: 0.6955 - val_loss: 0.8548 - val_accuracy: 0.6910 - lr: 1.0000e-04
Epoch 7/10
376/376 [==============================] - 132s 350ms/step - loss: 0.8300 - accuracy: 0.6960 - val_loss: 0.8158 - val_accuracy: 0.6881 - lr: 1.0000e-04
Epoch 8/10
376/376 [==============================] - 133s 355ms/step - loss: 0.8235 - accuracy: 0.6990 - val_loss: 0.8045 - val_accuracy: 0.6910 - lr: 1.0000e-04
Epoch 9/10
376/376 [==============================] - 137s 363ms/step - loss: 0.8076 - accuracy: 0.7062 - val_loss: 0.7942 - val_accuracy: 0.7060 - lr: 1.0000e-04
Epoch 10/10
376/376 [==============================] - 134s 356ms/step - loss: 0.7885 - accuracy: 0.7150 - val_loss: 0.8177 - val_accuracy: 0.6965 - lr: 1.0000e-04
```

Figure 9: Accuracy of CNN

Figure 10 shows the confusion matrix of the CNN model.

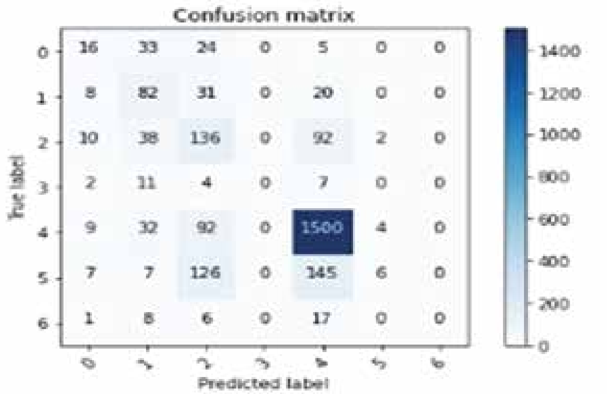


Figure 10: Confusion Matric of CNN

Thirdly, the VGG-16 model is considered which is comprised of 16 different layers. Figure 11 shows the basic structure of VGG-16.

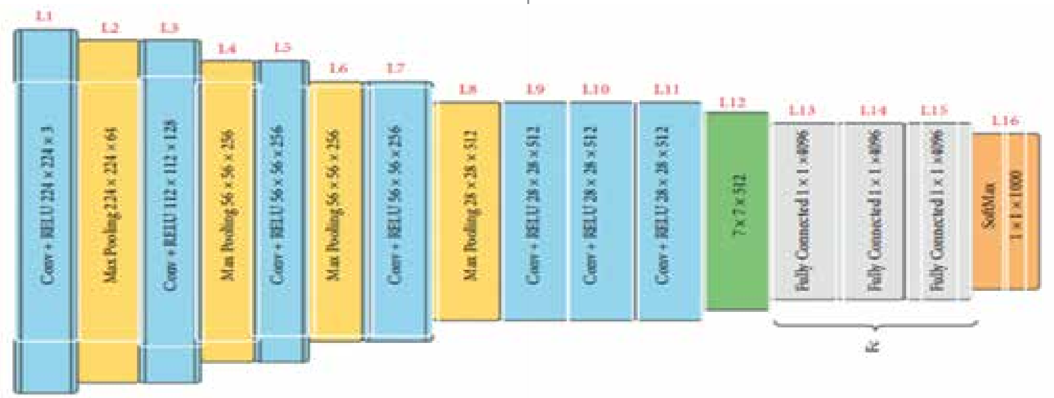
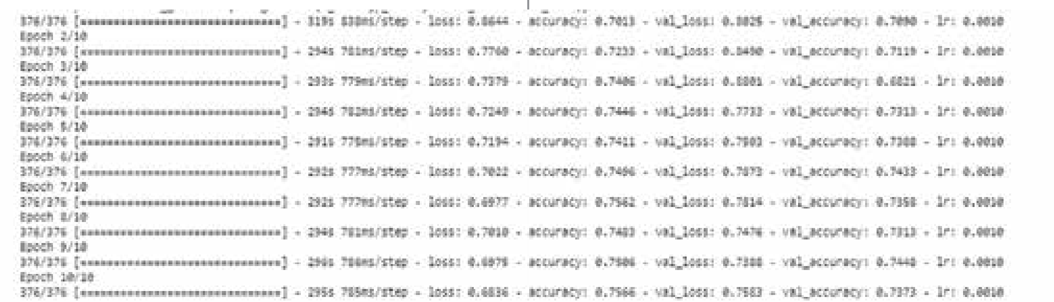


Figure 11: Basic structure of VGG-16

VGG-16 took input images from the HAM1000 dataset and pass them to its 16 different layers. After that, it performs classification according to 7 different skin cancer classes. Figure 12 shows the accuracy of VGG-16 model after 10 epochs. The average accuracy gained after VGG training is 75.6%



```
376/376 [==============================] - 319s 838ms/step - loss: 0.8644 - accuracy: 0.7013 - val_loss: 0.8026 - val_accuracy: 0.7090 - lr: 0.0010
Epoch 2/10
376/376 [==============================] - 294s 781ms/step - loss: 0.7760 - accuracy: 0.7233 - val_loss: 0.8400 - val_accuracy: 0.7119 - lr: 0.0010
Epoch 3/10
376/376 [==============================] - 293s 779ms/step - loss: 0.7379 - accuracy: 0.7406 - val_loss: 0.8801 - val_accuracy: 0.6821 - lr: 0.0010
Epoch 4/10
376/376 [==============================] - 294s 782ms/step - loss: 0.7240 - accuracy: 0.7446 - val_loss: 0.7732 - val_accuracy: 0.7313 - lr: 0.0010
Epoch 5/10
376/376 [==============================] - 291s 776ms/step - loss: 0.7194 - accuracy: 0.7411 - val_loss: 0.7503 - val_accuracy: 0.7388 - lr: 0.0010
Epoch 6/10
376/376 [==============================] - 292s 777ms/step - loss: 0.7022 - accuracy: 0.7496 - val_loss: 0.7073 - val_accuracy: 0.7433 - lr: 0.0010
Epoch 7/10
376/376 [==============================] - 292s 777ms/step - loss: 0.6977 - accuracy: 0.7562 - val_loss: 0.7814 - val_accuracy: 0.7358 - lr: 0.0010
Epoch 8/10
376/376 [==============================] - 294s 781ms/step - loss: 0.7010 - accuracy: 0.7403 - val_loss: 0.7476 - val_accuracy: 0.7313 - lr: 0.0010
Epoch 9/10
376/376 [==============================] - 296s 786ms/step - loss: 0.6876 - accuracy: 0.7506 - val_loss: 0.7388 - val_accuracy: 0.7448 - lr: 0.0010
Epoch 10/10
376/376 [==============================] - 295s 785ms/step - loss: 0.6836 - accuracy: 0.7566 - val_loss: 0.7583 - val_accuracy: 0.7373 - lr: 0.0010
```

Figure 12: accuracy of VGG-16 model

Figure 13 shows the confusion matrix of the VGG-16 model.

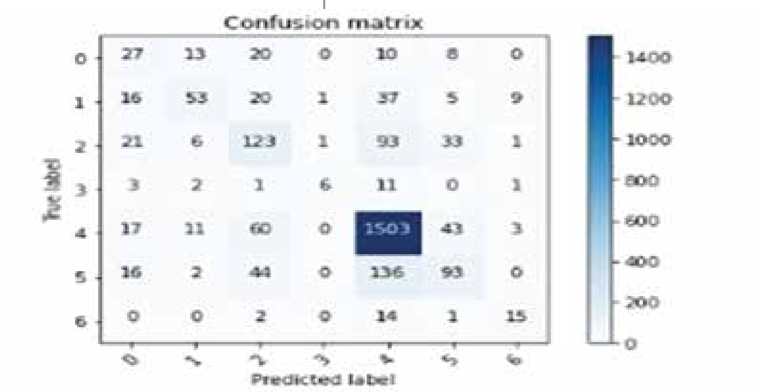


Figure 13: Confusion Matric of VGG-16

Figure 14 shows the comparison in performance of VGG, CNN and ANN.

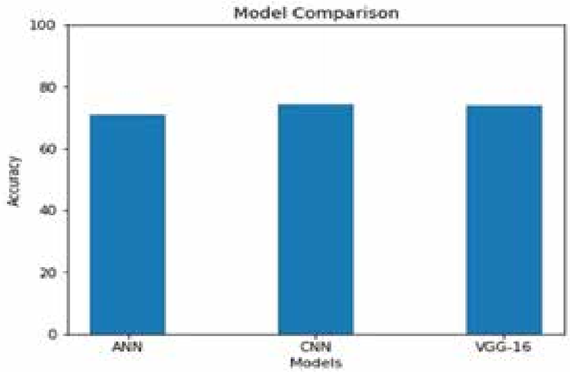


Figure 14: Performance comparison of 3 models

Figure 15 shows a comparison of the model performance of ANN with respect to Precision, Recall and F1-score.

```
Precision
['0.26%', '0.48%', '0.43%', '0.12%', '0.81%', '0.41%', '0.44%']
Recall
['0.29%', '0.51%', '0.22%', '0.08%', '0.92%', '0.25%', '0.34%']
F1 Score
['0.28%', '0.49%', '0.29%', '0.10%', '0.86%', '0.31%', '0.39%']
Specificty
0.46
```

Figure 15: Precision, Recall, F1-Score of 7 classes with respect to ANN

Figure 16 shows a comparison of the model performance of CNN with respect to Precision, Recall and F1-score.

```
Precision
['0.33%', '0.50%', '0.42%', '0.00%', '0.85%', '0.66%', '0.66%']
Recall
['0.38%', '0.65%', '0.51%', '0.00%', '0.91%', '0.21%', '0.59%']
F1 Score
['0.35%', '0.56%', '0.46%', '0.00%', '0.88%', '0.31%', '0.62%']
Specificty
0.5882352941176471
```

Figure 16: Precision, Recall, F1-Score of 7 classes with respect to CNN

Figure 17 shows a comparison of the model performance of VGG with respect to Precision, Recall and F1-score.

```
Precision
['0.32%', '0.53%', '0.46%', '0.62%', '0.83%', '0.54%', '0.76%']
Recall
['0.31%', '0.59%', '0.40%', '0.42%', '0.92%', '0.28%', '0.41%']
F1 Score
['0.31%', '0.56%', '0.43%', '0.50%', '0.87%', '0.37%', '0.53%']
Specificity
0.5714285714285714
```

Figure 17: Precision, Recall, F1-Score of 7 classes with respect to VGG

4. CONCLUSION

A method was developed in the proposed study for classifying seven types of lesions. The proposed method achieved high performance measures, including accuracy, sensitivity, specificity, and precision respectively. The performance of methods increased when the number of images in all classes decreased to address the imbalance issue. Fine-tuning all architecture layers resulted in higher performance measures compared to fine-tuning only the replaced layers. Additionally, the comparison has also performed to evaluate the performance of all three models (VGG, CNN, ANN). After experimental analysis, VGG outperform in classification as compared to other two models with 75.6% accuracy.

5. REFERENCES

[1]. H. Sugiyama, M. Misumi, M. Kishikawa, M. Iseki, S. Yonehara and T. Hayashi. “Skin cancer incidence among atomic bomb survivors from 1958 to 1996”. Radiat Res. vol. 181: pp. 531-539. 2014.

[2]. H. K. Koh, ‘‘Melanoma screening,’’ Arch. Dermatol., vol. 143, no. 1, pp. 101–103, Jan. 2007, doi: 10.1001/archderm.143.1.101.

[3]. Y. Zong, Y, Yang, and T. Hospedales. “MEDFAIR: Benchmarking Fairness for Medical Imaging”. arXiv preprint arXiv:2210.01725. 2022.

[4]. H. Basak, R. Kundu, and R. Sarkar. “MFSNet: A multi focus segmentation network for skin lesion segmentation”. Pattern Recognition, 128, 108673. 2022.

[5]. S. Singla. “Deep Learning for Medical Imaging From Diagnosis Prediction to its Counterfactual Explanation”. arXiv preprint arXiv:2209.02929. 2022.

[6]. A. Singh, S. Bera, P. Chaturvedi, P. Gadhave and C. S. Lifna,. “DermoCare. AI: A Skin Lesion Detection System Using Deep Learning Concepts”. In Data Intelligence and Cognitive Informatics”: Proceedings of ICDICI 2022

(pp. 39-51). Singapore: Springer Nature Singapore.

[7]. M. S. Khan, K. N. Alam, A. R. Dhruba, H. Zunair and N. Mohammed. "Knowledge distillation approach towards melanoma detection". Computers in Biology and Medicine, 105581. 2022.