



Deep learning architecture using transfer learning for classification of skin lesions

S. P. Godlin Jasil¹ · V. Ulagamuthalvi¹

Received: 27 August 2020 / Accepted: 2 March 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Skin cancer is one of the most dangerous health problems in many countries the development in automatic medical image analysis technique leads to accurate classification of diseases. Deep learning is a recent technology which solves the complexity in diagnosing the skin cancer. Due to high dissimilarity of skin lesion, it is challenging to classify the Image automatically. By using Convolutional Neural Network the classification of skin lesions can be done with high accuracy. The proposed models was developed by transfer learning concept with three popular architectures: inception V3, VGG16 and VGG19. Our models were trained by ISIC dataset that contain 2487 train images and 604 test images of seven skin lesion classes. Our models give the best performance of test data with 74%, 77% and 76% accuracy for Inception V3, VGG16 and VGG19.

Keywords Convolutional neural network · Deep learning · Skin lesion · Transfer learning · Diseases

1 Introduction

Cancer is a significant health problem in the world nowadays. Every sixth death in the world is due to cancer, making it the second leading cause of death after cardiovascular diseases. Cancer (malignant or tumor) is an abnormal growth of cells, which cells reproduce in an uncontrolled way that spread to other tissues (Roser and Ritchie 2019). Skin cancer is an unusual growth in skin cells. It effects the skin that exposed to sunlight. Skin cancer can affect all people and have the higher risk to those who have light skin. Skin cancer can occur any time of life, and it can grow to other inner cells of the skin and can damage its DNA. The major reason is because of exposing to Ultraviolet (UV) radiation than other reasons such as ozone coating exhaustion, chemical exposure and genetic issues (Yogita et al. 2018; Maiti and Chatterjee 2019).

Deep learning neural network based medical image analysis techniques are helpful for dermatologist in skin lesion

segmentation, classification and detection (Indra et al. 2020). Machine learning techniques dynamically classify, predict, and detect different skin lesion using annotated skin lesions images produce by dermoscopy (Kittler et al. 2002; Salerni et al. 2013; Pan and Yang 2009; Tschandl et al. 2018). On the other hand, automatic recognition of dermoscopy images is a difficult task because of some factor such as skin lesion segmentation as it differs in its texture, color, size, location and shape. Second, the surroundings such as hair, ruler marks and veins (Hosny et al. 2018).

In recent years Deep Learning neural network has been used for Medical Image Segmentation. Convolutional Neural Networks (CNN) which works like feedforward neural network with additional of techniques such as convolutions, max-pooling, batch-normalization, dropout, etc. Generally CNN takes 2D input image and apply 2D filters. Transfer Learning is a technique that can be used for challenging problem which has the large medical dataset. The models are trained by freezing few layers and the last layer is replaced by a new softmax layer that classifies the lesion images into seven different classes.

Due to the development in image processing and machine learning techniques, classification tasks can be done better than humans by artificial convolutional neural network (Gavrilov 2017; Nagarajan et al. 2020). In New York Oncology Center, the decision-making in diagnostics and treatment of the patients with ontological diseases are done

✉ S. P. Godlin Jasil
spgodlinjasil@gmail.com

V. Ulagamuthalvi
muthalvi73@gmail.com

¹ Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology,
Chennai 600119, India

by deep learning methods of artificial intelligence techniques in medicine is the IBM Watson system which is the best example (Tsvetkova 2017) for the decision-making systems.

We proposed a deep learning architecture with transfer learning technique that classifies seven different classes of skin lesion. We select three pre-trained deep learning architecture namely VGG19, VGG16 and Inception V3 which are pre-trained architecture on ImageNet that achieves higher accuracy. In this study first the preprocessing such as normalization, resizing the image and data augmentation is done. Second the deep neural networks was developed by using pre-trained networks. Finally, the performance of three pre-trained architectures is evaluated.

The remaining part of the paper is organized as follows: a brief discussion on Background work in Sect. 2. Dataset are explained in Sects. 3 and 4 on details of Methodology. The experimental results of the models are presented in Sect. 5. Finally, conclusion on Sect. 6.

2 Related work

Hosny et al. (2018) proposed a AlexNet pretrained deep learning architecture by replacing the last layer with the softmax layer that has three classes namely common nevus, atypical nevus and melanoma.

Dorj et al. (2018) deals with rapid classification system of skin cancer by using the efficient deep convolutional neural network which mainly focus on classifying the skin cancer using ECOC SVM, and deep convolutional neural network. The collected images have noises that are cropped to remove the noise which gives better results. A pre-trained AlexNet convolutional neural network model is used in Transfer learning. The results obtained for a total of 3753 images, which include four kinds of skin cancers images. Gavrilov et al. (2019) proposed an Inception V3 architecture by removing the last layer and adding one neuron to classify the skin lesion. This model archives the accuracy of 91%, that is compared with the diagnostic done by extremely qualified dermatologists.

Guo et al. (2018) propose a novel multiple convolution neural network model (MCNN) which classify dermoscopic images with different disease. Using an additive sample learning strategy several models are trained. Alheejawi et al. (2019) propose a deep learning algorithm which is used for automatic measurement of PI values in Ki-67 stained biopsy image. By using trained convolutional neural network (CNN) model it segments the nuclei and measures the PI. Manually segmented nuclei of Ki-67 images are used to train the CNN model. Ech-Cherif et al. (2019a, b) propose a binary classification of skin lesions into benign or malignant by manually trained CNN MobileNetV2 model using transfer learning. A mobile application for iOS devices was

developed using the Core ML library. This model gives a high accuracy.

Najat et al. (2019) proposed an Improved U-Net architecture consists of 46 layers separated into two parts Contraction path and Expansion path that contains convolutions, activation functions, down-sampling, and up-sampling operator which is used to predict the value of a particular pixel belongs to the skin lesion. Vesal et al. (2018) proposed a modified version of U-Net called skinNet was developed with dilated and densely block convolutions which helps to include multi-scale and global context information when training the model.

According to Zhang et al. (2019) a fully automatic framework was developed by coupling a deep FCN with a shallow network contains textons derived from domain specific filter kernels. The shallow network helps to add clinical prior knowledge (textures) with deep FCN. A fusing strategy is used to combine domain-specific hand-crafted texton features into a deep network. Abbes and Sellami (2019) developed a deep neural network by fuzzy modelling of Bag-of-Words (BoW) and the features are extracted by ABCD rules. For each BoW of lesion the membership degrees are determined by Fuzzy C-Means (FCM). This model produces the accuracy of 87.5%. Kavimathi (2016) developed a model in which the evaluation has been done with different classifiers like ensemble classifier, Support Vector Machine, probabilistic neural network, adaptive neuro-fuzzy inference system and come out with the best performance classifier.

Akram et al. (2020) developed a pre-trained model that is used to generate fused feature vectors then a framework was developed for feature selection and dimensionality reduction that is used for selecting the principle components and extricating the redundant and irrelevant data. This frameworks shows the accuracy of 98.8%, 99.2%, 97.1% and 95.9%. Thao and Quang (2017) developed a architecture to segment skin tumor from the surrounding skin and a pretrained network VGG-16 for classification of skin lesion. Ahmed et al. (2019a) propose a pre-trained architecture MobileNetV2 that is used to classify the skin lesion into benign or malignant classes. A mobile application for iOS devices was developed by using the trained model. This model produce the accuracy of 91.33%. According to Mahbod et al. (2019) three pre-trained networks such as VGG16, AlexNet and ResNet-18 are used to extracted features that is used by support vector machine classifiers for training then the classifier outputs are combined for classifying the skin lesion.

3 Dataset

Our pre-trained architectures are trained by the dataset taken from International Skin Imaging Collaboration as a part of ISIC 2018 (ISIC Archive) that consist of seven different

types of skin lesion. The models classify the input image into a category as follows (Fig. 1): Melanoma (MEL) 271 images; Melanocytic Nevus (NV) 2061 images; Basal Cell Carcinoma (BCC) 151 images; Benign Keratosis (BKL) 345 images; Dermatofibroma (DF) 36 images; Actinic Keratosis (AKIEC) 113 images; Vascular Lesion (VASC) 45 images.

3.1 Pre-processing

The images in our dataset are resized to 224×224 pixel. We applied the standard pre-processing steps such as normalization, resizing and data argumentation. We have normalized the images because the pre trained network was already trained with those images. The modified pre-trained architecture was trained with 2418 training images with seven different types of skin lesion that is used to distinguish different types of skin cancer. To increase the size of the dataset, data argumentation techniques such as shift, zoom, compression, flip and altering lighting are used. Data argumentation helps the network to identify new training images that give more accuracies of the model. After the data argumentation we get 2487 images.

4 Methodology

Convolution neural network is a most effective method for image classification and feature extraction in deep learning architecture (LeCun et al. 2015). The dermoscopic images might have noise (Nirmalraj and Nagarajan 2020, 2019), aberrations, and artefacts in addition with the variation of

same feature and visual similarity of various lesion types. To overcome this problem, a huge dataset should be used for training (Yang et al. 2018; Greenspan et al. 2016).

The performance of the CNN can be improved by using pre-trained architecture like VGG16, VGG19, Resnet50, MobileNet and Inception V3. In this proposed CNN network VGG19, VGG16 and Inception V3 has been used which gives a better performance.

Due to insufficient dataset to train the model from the scratch, a pre-trained architecture is used for skin lesion classification. Transfer learning is a technique that allows to retrain the last layer of existing architecture which is used to decrease the training time with limited dataset. Adam optimizer is used for optimization with the initial learning rate of $4e-5$. The most popular models in pre-trained architectures are Inception V3, VGG16 and VGG19 which gives an excellent performance in skin lesion classification.

4.1 Inception V3 CNN model

Inception.v3 network (Fig. 2) demonstrates image classification model of various images for ILSVRC (ImageNet Large Scale Visual Recognition Challenge) competition. Inception.v3 network consist of 42 layers. The network includes symmetric and asymmetric building blocks that consist of convolution layer, average and max pooling, dropouts, and fully connected layers. In modified pre-trained Inception.v3 neural network the last layer of the network was removed and replaced with a softmax layer that classifies seven classes of skinlesion.

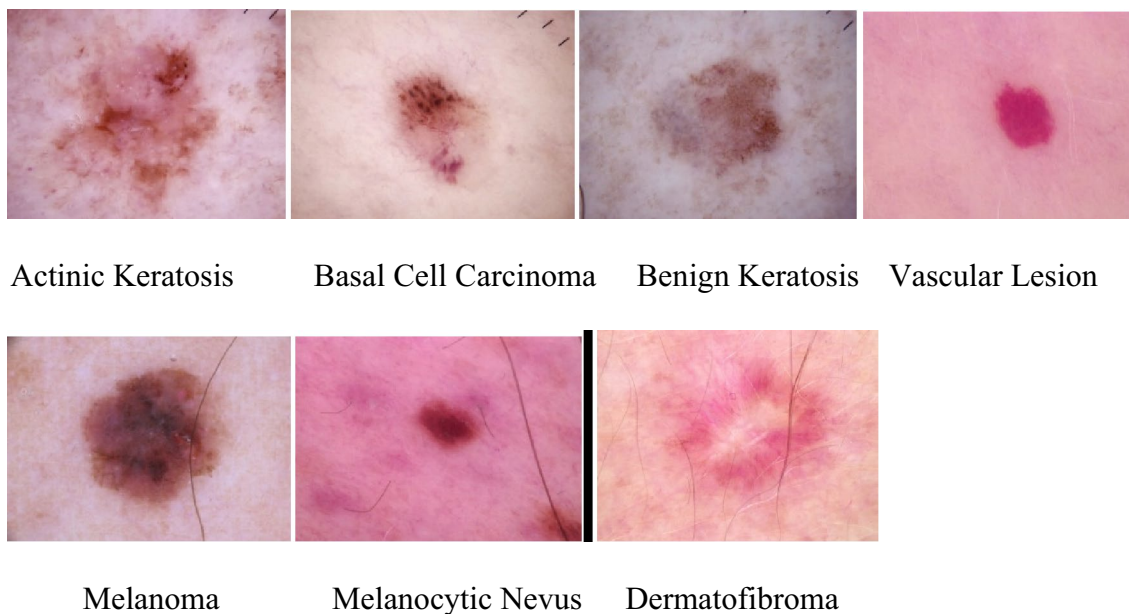


Fig. 1 Sample images of ISIC dataset

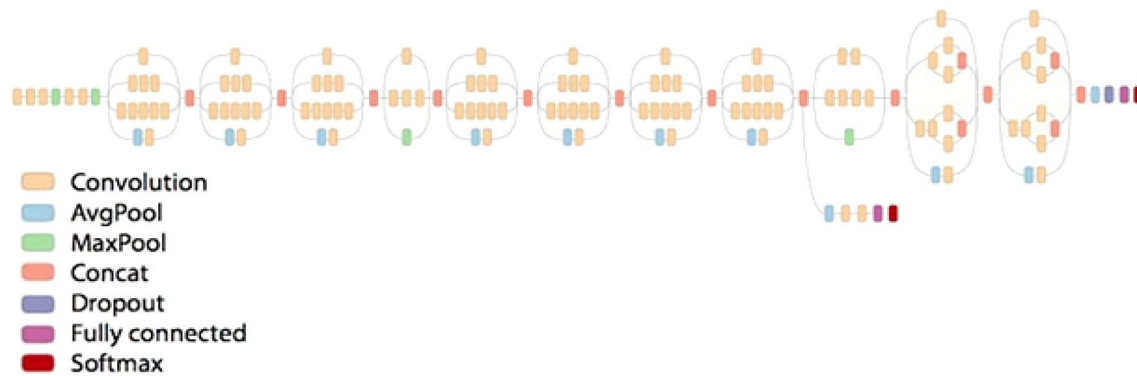


Fig. 2 Architecture of inception V3 (Szegedy et al. 2015)



Fig. 3 VGG19 framework

4.2 VGG19 CNN model

VGG19 network (Fig. 3) consist of 19 layers including sixteen convolution layers, three Fully connected layers, five MaxPool layers and a SoftMax layer. It takes the input image of 224×224 size into the network. The network uses the kernel size of 3×3 with stride 1 and spatial padding. Max pooling was done with stride 2. There are two fully connected layers and the last layer was removed and replaced with a softmax layer which classifies 7 classes of skin lesion.

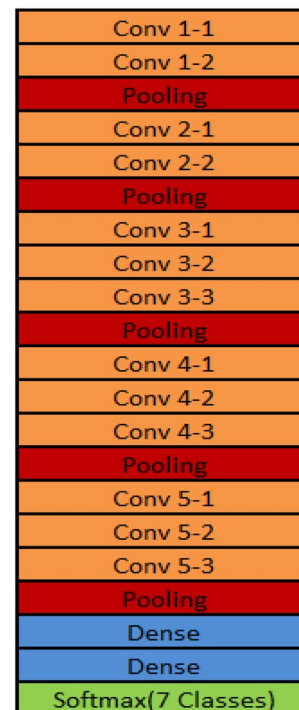


Fig. 4 VGG16 framework

4.3 VGG16 CNN model

VGG-16 (Fig. 4) is a simpler architecture mode with convolution layer that consist of 3×3 filter and stride 1. It consist of pooling layers 2×2 with SAME padding of stride 2. There are two fully connected layers and the last layer was removed and replaced with a softmax layer which classifies 7 classes of skin lesion.

5 Results

The proposed modified architecture used in our work are Inception V3, VGG16 and VGG19 network to extract features that can classify skin lesion in to seven classes. To obtain a better result, deep learning model need large dataset. The dataset used in our architecture consist of

2487 training images and 604 testing images. The dataset consist of various dimension photographic images. Our dataset is divided into 80% of training images and 20% of testing images. The performance of three different pre-trained architecture is evaluated by using skin lesion dataset of ISIC archives (Figs. 5, 6).

Fig. 5 Classification rates for the model

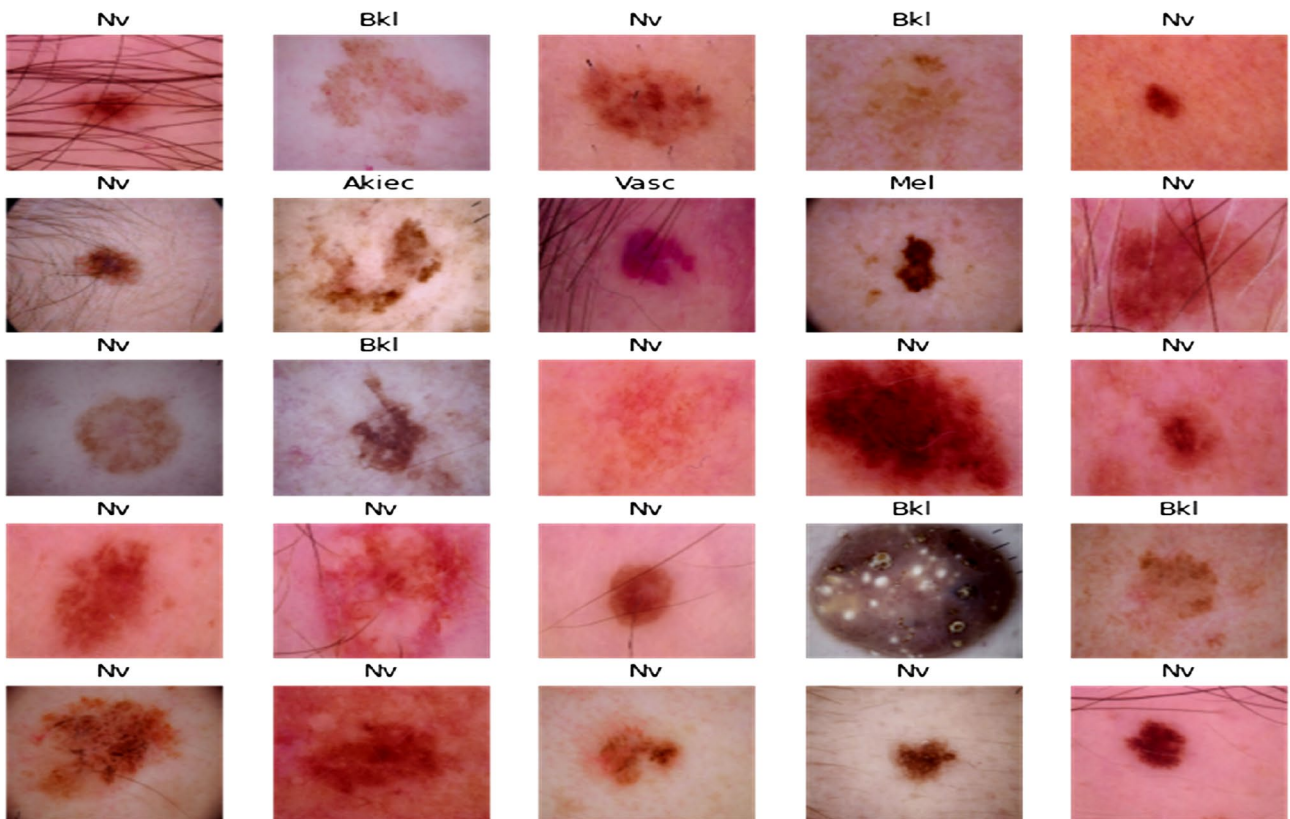
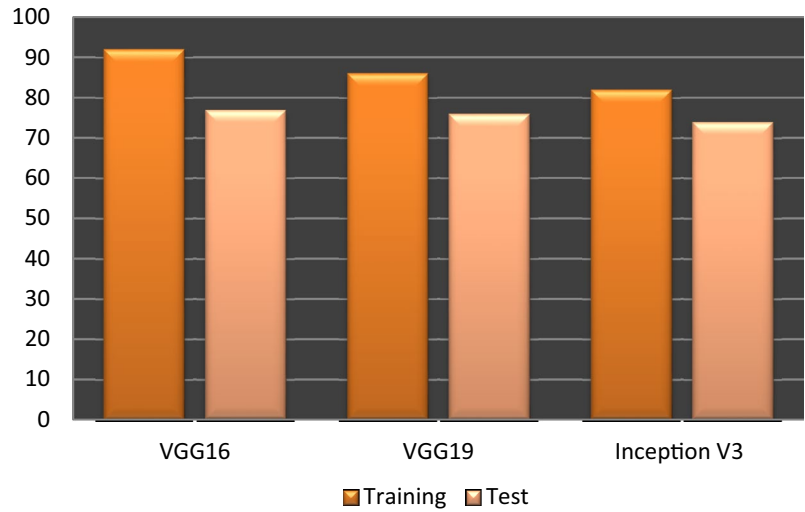


Fig. 6 Sample visualized output of the model

5.1 Classification results

5.1.1 Inception V3

The training set consist of 2487 images that are trained for 60 epochs with the batch size of 32. Figure 7 shows the accuracy and loss values produced by Inception V3 pre-trained architecture. From the graph, the train accuracy and test accuracy gradually increasing and loss decreases for both train and test data. This model gives the best accuracy of 82% of training accuracy and 74% of test accuracy.

5.1.2 VGG19

The training set consist of 2487 images that are trained for 60 epochs with the batch size of 32. Figure 8 shows the accuracy and loss values of VGG19 pre-trained architecture. From the graph, the train accuracy and test accuracy gradually increasing and loss decreases for both train and

test data. This model gives the best accuracy of 86% and test accuracy of 76%

5.1.3 VGG16

The training set consist of 2487 images that are trained for 50 epochs with the batch size of 32. Figure 9 shows the accuracy and loss values of VGG16 pre-trained architecture. From the graph, the train accuracy and test accuracy gradually increasing and loss decreases for both train and test data. This model gives the best accuracy of 92% for training and 77% of testing accuracy.

5.2 Recall, precision and F-1 score

The results of the models can be analyzed by calculating its values such as recall, precision and F-1 score.

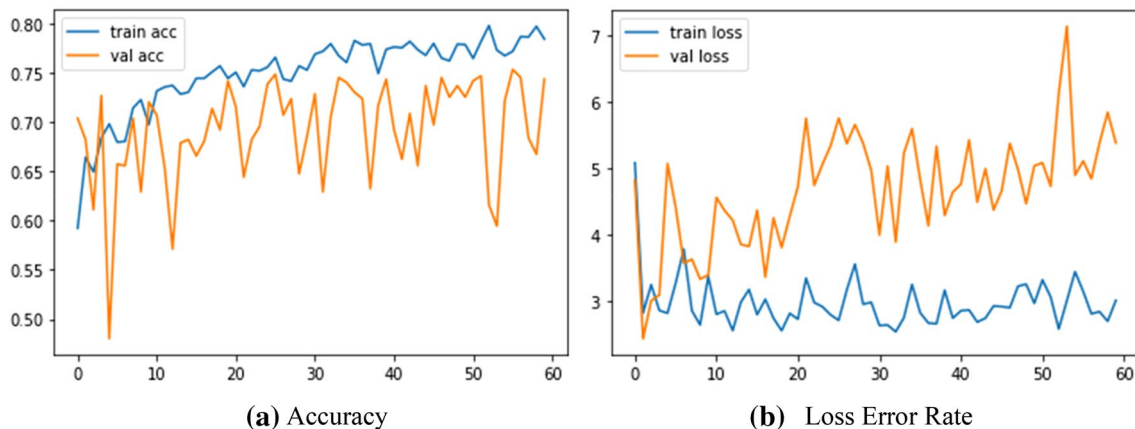


Fig. 7 Accuracy and loss graph on modified inception V3 network

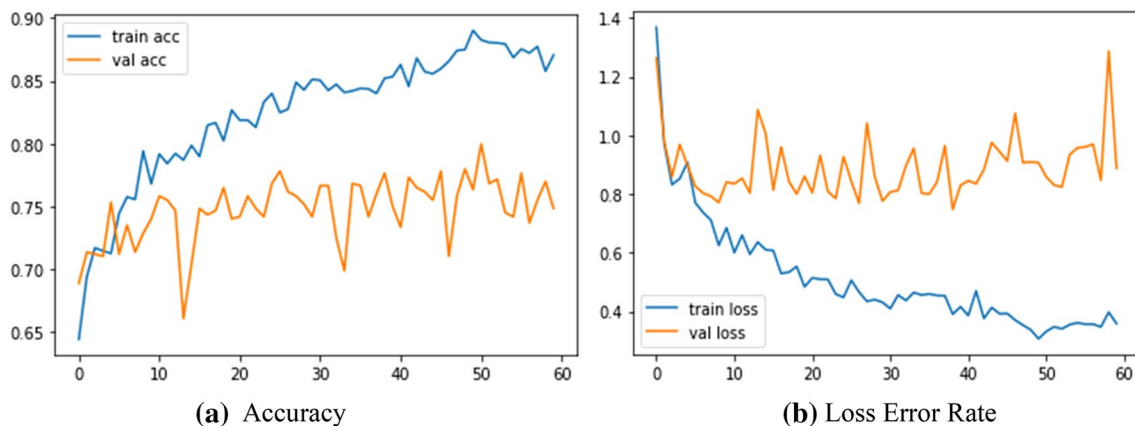


Fig. 8 Accuracy and loss graph on modified Vgg19 network

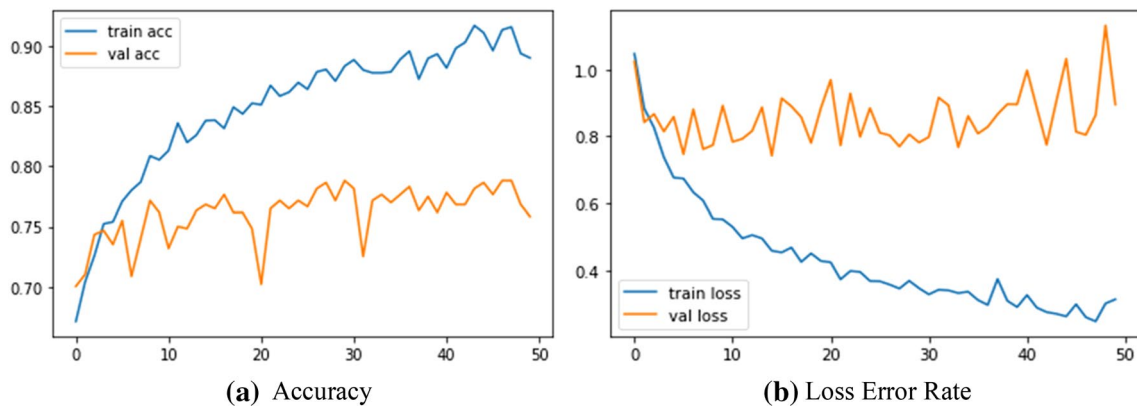


Fig. 9 Accuracy and loss graph on modified Vgg16 network

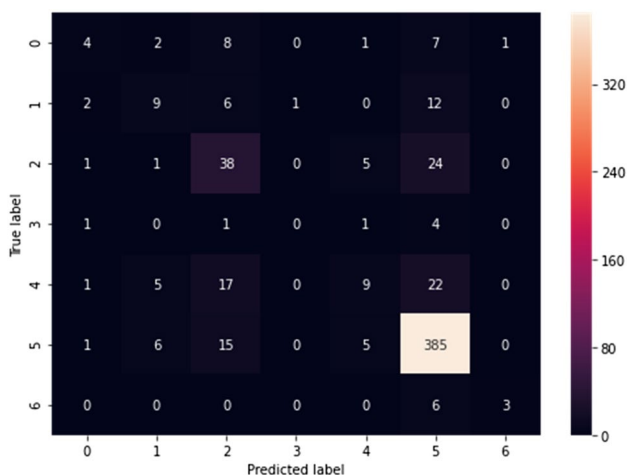


Fig. 10 Confusion matrix for Inception v3 architecture

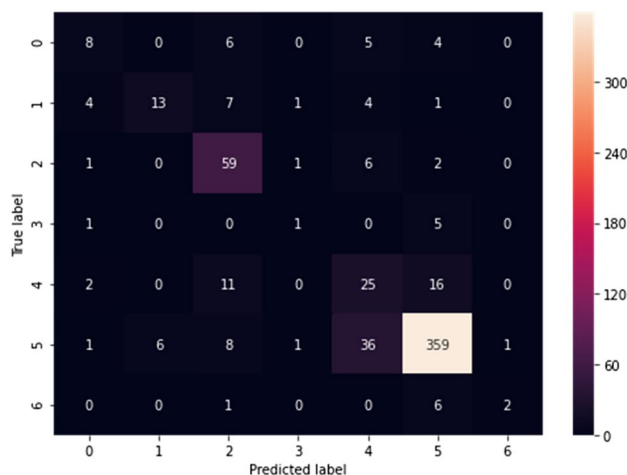


Fig. 11 Confusion matrix for VGG16 architecture

$$Precision = \frac{true\ positive}{true\ positive + false\ positive} \tag{1}$$

$$Recall = \frac{true\ positive}{true\ positive + false\ negative} \tag{2}$$

$$F - measure = 2 \times \frac{precision \times recall}{precision + recall} \tag{3}$$

5.3 Confusion matrix

The performance of the models can be evaluated by confusion matrix. In Figs. 10, 11, 12 we show the Confusion Matrix of three models. In this figures Actinic Keratosis=0; Basal Cell Carcinoma=1; Benign Keratosis=2; Dermatofibroma=3; Melanoma=4; Melanocytic Nevus=5; Vascular Lesion=6.

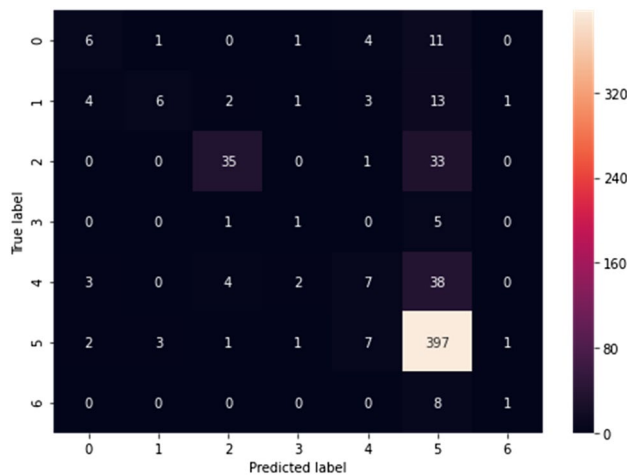


Fig. 12 Confusion matrix for VGG19 Architecture

6 Conclusion

In the proposed deep learning architecture we used three pre-trained model to classify seven classes of skin lesion. We obtained a good accuracy rates with our dataset. Our results of test data show that, when we compare Vgg19, vgg16 and Inception pre-trained architecture, the performance of VGG16 network gives the better accuracy of 77% and Vgg19 give the accuracy of 76%, Inception V3 give the accuracy of 74%. Since the pre-trained deep learning architecture obtains a good accuracy in medical image processing, the proposed work can be used in classifying different types of skin lesions.

References

- Abbes W, Sellami D (2019) Deep neural network for fuzzy automatic melanoma diagnosis. In: 14th international conference on computer vision theory and applications, VISAPP
- Ahmed E-C, Mohammed M, Mohammed E-C (2019) Deep neural network based mobile dermoscopy application for triaging skin cancer detection. IEEE
- Akram T, Lodhi HJM, Naqvi SR, Naeem S, Alhaisoni M, Ali M, Haider SA, Qadri NN (2020) A multilevel features selection framework for skin lesion classification. *Hum Cent Comput Inf Sci* 2020:5
- Alheejawi S, Mandal M, Berendt R, Jha N (2019) Automated melanoma staging in lymph node biopsy image using deep learning. *IEEE Can Conf Electr Comput Eng* 2019:5
- ISIC Archive [electronic resource] (2019) <https://www.kaggle.com/nodoubtome/skin-cancer9-classesisic>
- Dorj U-O, Lee K-K, Choi J-Y, Lee M (2018) The skin cancer classification using deep convolutional neural network. *Multimed Tools Appl* 2018:5
- Ech-Cherif A, Misbhauddin M, Ech-Cherif M (2019) Deep neural network based mobile dermoscopy application for triaging skin cancer detection. *IEEE* 2019:5
- Gavrilov DA (2017) Artificial Intelligence based skin lesions photo recognition. In: AMEC LIVE and VISAGE Joint Meeting, Monaco
- Gavrilov DA, Shchelkunov NN, Melerzanov AV (2019) Deep learning based skin lesions diagnosis. *Rem Sens Spatial Inf Sci* 2019:12
- Greenspan H, Ginneken B, Summers R (2016) Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans Med Imag* 35(5):2115–2115
- Guo Y, Ashour AS, Si L, Mandalaywala DP (2018) Multiple convolutional neural network for skin dermoscopic image classification. *IEEE* 2018:365–369
- Hosny KM, Kassem MA, Foad MM (2018) Skin cancer classification using deep learning and transfer learning. In: The 9th Cairo international biomedical engineering conference (CIBEC 2018)
- Hosny KM, Kassem MA, Foad MM (2019) Classification of skin lesions using transfer learning and augmentation with Alex-net. *PLoS ONE* 2019:0217293
- Indra MR, Govindan N, Satya RKDN, Thanasingh SJS (2020) Fuzzy rule based ontology reasoning. *J Ambient Intell Hum Comput* 2020:1–7
- Kavimathi P (2016) Comparative analyses of classifiers for diagnosis of skin cancer using dermoscopic images. *Indian J Sci Technol* 9(43):2016
- Kittler H, Pehamberger H, Wolff K, Binder M (2002) Diagnostic accuracy of dermoscopy. *Lancet Oncol* 3(3):159–165
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444. <https://doi.org/10.1038/nature14539>
- Mahbod A, Schaefer G, Wang C, Ecker R, Ellinger I (2019) Skin lesion classification using hybrid deep neural networks. *IEEE* 2019:5
- Maiti A, Chatterjee B (2019) Improving detection of Melanoma and Naevus with deep neural networks. *Multimedia Tools Appl* 2019:11042
- Nagarajan G, Minu RI, Devi AJ (2020) Optimal nonparametric bayesian model-based multimodal BoVW creation using multilayer pLSA. *Circ Syst Signal Process* 39(2):1123–1132
- Najat S, Hasan Murat GEZER, Raghada AA, Sevinc A (2019) Skin lesion segmentation by using deep learning techniques. In: *IEEE*
- Nirmalraj S, Nagarajan G (2019) An adaptive fusion of infrared and visible image based on learning of sparse fuzzy cognitive maps on compressive sensing. *J Ambient Intell Hum Comput* 2019:1–11
- Nirmalraj S, Nagarajan G (2020) Biomedical image compression using fuzzy transform and deterministic binary compressive sensing matrix. *J Ambient Intell Hum Comput* 2020:5
- Pan SJ, Yang Q (2009) A survey on transfer learning. *IEEE Trans Knowl Data Eng* 22(10):1345–1359
- Roser M, Ritchie H (2019) Cancer. <https://ourworldindata.org/cancer>
- Salerni G, Terán T, Puig S, Malvey J, Zalaudek I, Argenziano G, Kittler H (2013) Meta-analysis of digital dermoscopy follow-up of melanocytic skin lesions: a study on behalf of the International Dermoscopy Society. *J Eur Acad Dermatol Venereol* 27(7):805–814
- Szegedy C et al (2015) Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1–9
- Thao LT, Quang NH (2017) Automatic skin lesion analysis towards melanoma detection. In: 21st Asia Pacific symposium on intelligent and evolutionary systems (IES)
- Tschandl P, Rosendahl C, Kittler H (2018) The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* 5:180161
- Tsvetkova LA (2017) Artificial intelligence technology as a factor in digitization of the economy in Russia. *Ekonomika Nauki* 2:126144
- Vesal S, Ravikumar N, Maier A (2018) SkinNet: a deep learning framework for skin lesion segmentation. *IEEE* 2018:5
- Yang X, Zeng Z, Yeo S, Tan C, Tey H, Su Y (2018) A novel multitask deep learning model for skin lesion segmentation and classification. *CORR Preprint*. arXiv:1703.01025
- Yogita R, Samina S, Pooja S, Jyotsana G (2018) Review on skin cancer. *Asian J Res Pharmaceut Sci* 8(2):100–106
- Zhang L, Yang G, Ye X (2019) Automatic skin lesion segmentation by coupling deep fully convolutional networks and shallow network with textons. *J Med Imaging* 6(2):024001