STM JOURNALS

Journal of
Artificial Intelligence Research & Advances

ISSN: 2395-6720
Volume 8, Issue 3, 2021
DOI Journal: 10.37591/JoAIRA

http://computers.stmjournals.com/index.php?journal=JoAIRA&page=index

Research JoAIRA

DeMe-Net Model to Diagnose Skin Cancer Disease from Skin Lesion

Aakriti Yadav[1,*], Preeti Rai[2]

Abstract

Skin cancer is a quite common disease in the world and Melanoma is one of the serious types of skin cancer. It begins in melanocyte cells. Skin irritation or tanning due to UV radiation triggers changes in melanocytes, leading to uncontrolled growth of cells. It can potentially develop and spread if it is not treated early. The biopsy is a traditionally used method to diagnose skin cancer in which a sample of the skin lesion is sent to a laboratory for testing. This methodology of diagnosis is troublesome, tedious, and time-consuming. Deep Learning, a branch of artificial intelligence (AI) is one of the most promising researches and innovation areas in the field of medical imaging. In this research, we propose a DeMe-Net model to diagnose skin cancer with deep learning techniques. The overall accuracy of the model is 98.33% on the testing data and 99.85% on the training dataset.

Keywords: DeMe-Net, Deep Learning, skin cancer, Convolutional Neural Network, Image Processing

INTRODUCTION

Melanoma cancer is the fifth most common cancer in both men and women. The rate of development of melanoma cancer in people depends on age. But it is also seen that younger people (under the age of 30) also suffer from melanoma. In the survey for the year 2020, people aged 15 to 29 years were estimated to constitute about 2,400 cases of melanoma to be diagnosed [1]. With early detection, the most lethal form of melanoma has a 5-year survival rate of up to 99%, but a late diagnosis reduces the survival rate to 23% [2].

The traditional method to detect skin cancer is a biopsy. There are also visual inspection-based methods to diagnose skin cancer that can be detected by a dermatologist of skin lesions, like Dermoscopy and Computed Tomography Scan [3]. Dermoscopy is based on imaging technology for the examination of the skin using skin surface microscopy, it eliminates skin surface reflection. There will be dermatoscopy equipment available to operate on smartphones [4], and they provide opportunities for automated dermatology diagnostic algorithms that would create a positive impact on the medical field. Skin cancer can be cured if detected early. Early diagnosis is required through dermoscopy images of skin lesions with high precision.

Deep learning and convolutional neural networks (CNN) have become the standard approach for automated diagnosis based on medical images. For the problem of dermoscopy-based skin cancer diagnosis, a dataset has been published. The dataset consists of 10000 dermoscopic images of skin lesions with seven

***Author for Correspondence**
Aakriti Yadav
E-mail: aakriti.1410@gmail.com

[1]Student, Department of Computer Science & Engineering, Gyan Ganga Institute of Technology and Sciences, Jabalpur, Madhya Pradesh, India
[2]Associate Professor, Department of Computer Science & Engineering, Gyan Ganga Institute of Technology and Sciences, Jabalpur, Madhya Pradesh, India

Received Date: June 22, 2021
Accepted Date: October 07, 2021
Published Date: November 25, 2021

Citation: Aakriti Yadav, Preeti Rai. DeMe-Net Model to Diagnose Skin Cancer Disease from Skin Lesion. Journal of Artificial Intelligence Research & Advances. 2021; 8(3): 10–20p.

different diseases including Actinic keratosis/Bowen's disease, Basal cell carcinoma, Benign keratosis, Dermatofibroma, Melanocytic nevus, Vascular lesion, Melanoma. The dataset is highly biased based on the number of images of skin lesions in the different classes. We have used preprocessing techniques to make it usable for this research.

In this research, we provide a method to automatically diagnose melanoma cancer along with six other types of skin disease. We used deep learning algorithms and the concept of convolution neural networks and created a DeMe network with parallel convolutional layers. The network takes the images of skin lesions as input and provides predictions of seven diseases as output. DeMe Network provides state-of-the-art performance on the dermoscopic skin disease image dataset.

LITERATURE REVIEW

A. Polat and Koc proposed two different methods, CNN and one-versus-all (OVA) approach and on the HAM10000 dataset and achieved an accuracy of 92.90% [5].
B. Pham *et al.* proposed a hybrid method for class imbalance handling and the EfficNetB4 model classification [6]. They used 24,530 dermoscopic images for training and achieved an accuracy of 89.97%.
C. Mahajan *et al.* proposed the method Meta-Derm Diagnosis utilizing a meta-learning-based few-shot learning technique [7]. They evaluated the model on ISIC 2018, SD-198 datasets, and Derm7pt datasets. The model obtained an accuracy of 83.70%.
D. Majtner *et al*. proposed the methodology based on deep learning in their research [8]. They used VGG16 and GoogLeNet architectures and used the transfer learning technique for training the models. They evaluated their model on ISIC 2018 Skin Lesion dataset and obtain an accuracy of 80.1% for VGG16, 79.70% for GoogLeNe, and 81.50% for their ensemble.
E. Gessert *et al*. proposed the solution based on pre-trained ResNeXt, SENet and Densenet deep learning architectures [9]. The final ensemble contained 54 models and achieved 5-fold accuracy of 85.10%.

METHODOLOGY

Dataset

Deep Learning techniques required a large amount of data for training to obtain higher accuracy. In this research, we accumulated 10015 dermoscopic images of skin lesions of seven different. This data set is contributed by the ISIC, also known as International Skin Imaging Collaboration [2]. The ISIC dataset is the largest collection of publicly available dermoscopic images of skin lesions. Some images of skin lesions are shown in Figure 1. In this research, we have used skin lesions images of a size of 28×28 pixels with RGB format to train the model [10]. The number of images of different classes we accumulated is shown in Table 1.

Data Pre-processing

In order to train deep learning models well, we need to pre-process and clean the data. The flowchart of data pre-processing is shown in Figure 2.

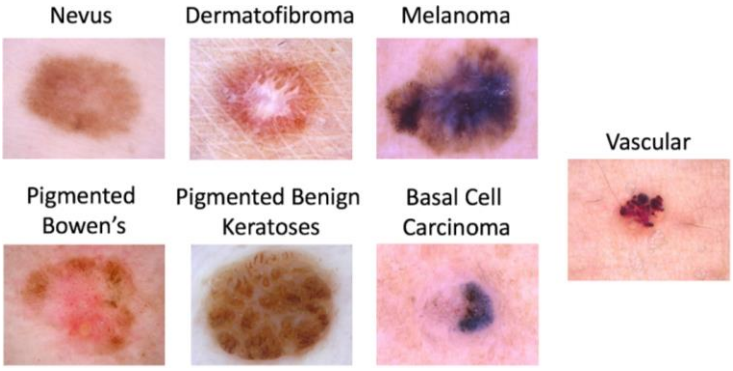


Figure 1. Skin leasion image dataset.

Table 1. Dataset.

S.N.	Classes	No. of Images
1	Actinic keratosis/Bowen's disease (akiec)	327
2	Basal cell carcinoma (bcc)	514
3	Benign keratosis (bkl)	1099
4	Dermatofibroma (df)	115
5	Melanocytic nevus (nv)	6705
6	Vascular lesion (vasc)	142
7	Melanoma (mel)	1113

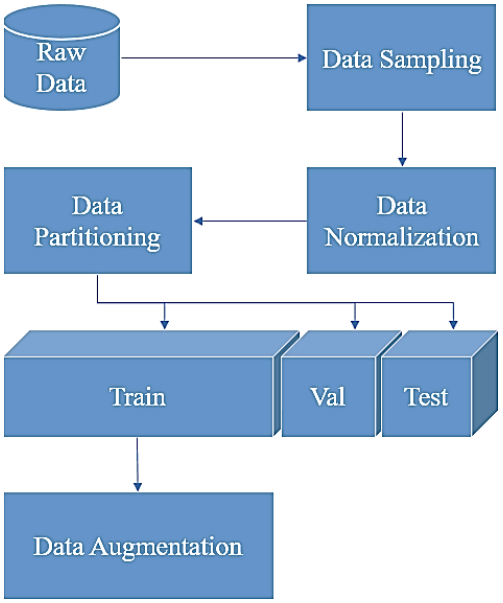


Figure 2. Flowchart of data preprocessing techniques.

Data Sampling

According to Table 1, we observed that the distribution of images of different classes is unequal, which causes a class imbalance problem. There are two techniques, undersampling and oversampling used to handle the class imbalance. In this study, we needed a large amount of data, for which it is appropriate to use oversampling. Figure 3 shows the dataset distribution with class imbalance problem and the dataset distribution after oversampling is shown in Figure 4.

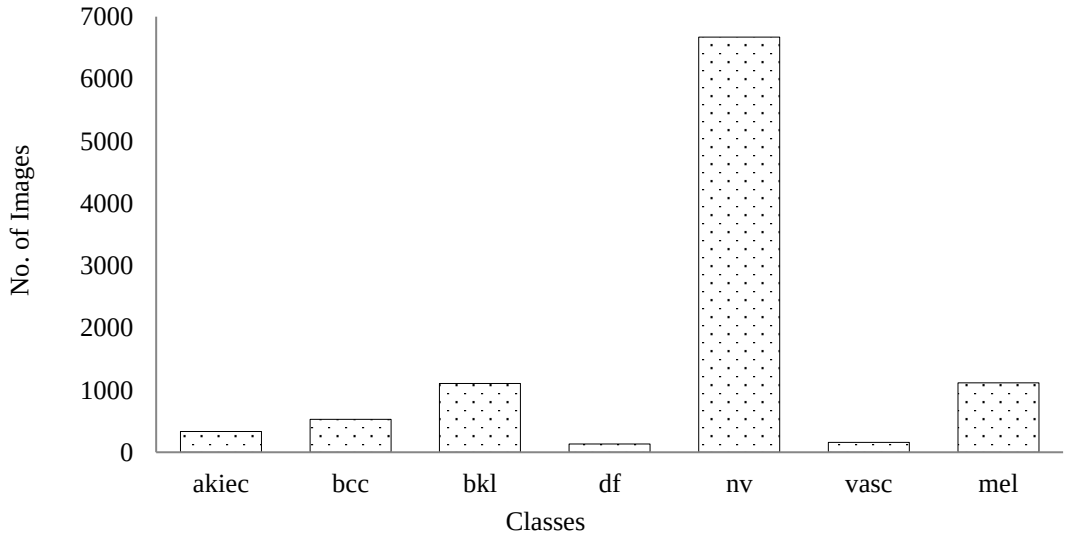


Figure 3. Dataset distribution with class imbalance problem.

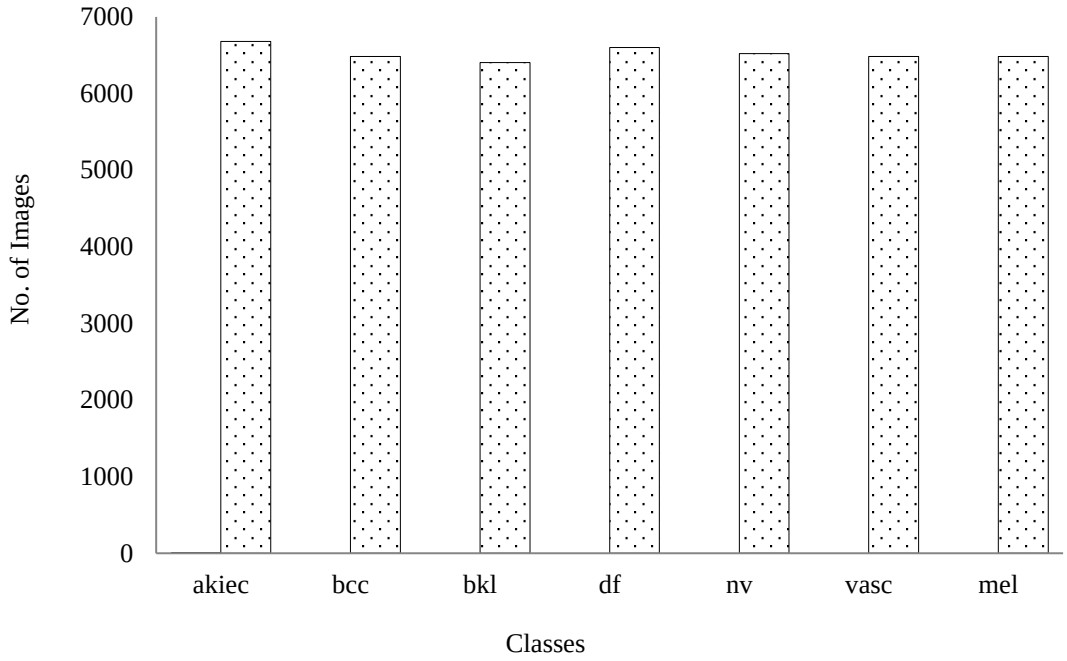


Figure 4. Dataset Distribution after Oversampling.

Data Normalization

Deep learning algorithms try to find patterns in the data by comparing the features of the data points. A problem occurred when features are at very different scales. In order to overcome this, we normalized the data with Z-Score Normalization. It is a strategy of normalizing data that avoids this outlier issue. The formula of Z-Score Normalization is shown in Eq. (1); where, μ and σ denote the mean and standard deviations respectively.

$$Z = \frac{x - \mu}{\sigma} \tag{1}$$

Data Partitioning

Data partitioning is a process of splitting data into train and test sets usually for cross-validation purposes. It helps to evaluate the performance of deep learning algorithms. In this research, we split the data into three parts, train, validate and test the set in the ratio 6:2:2. The train set is used for training the DeMe net model, the validation set is used for the evaluation of the model during training and hyperparameter tuning and the test set is used for prediction and analysis of the data not used to train the model.

Data Augmentation

In a real-world scenario, we only have a limited number of conditions in a dataset of images; but may exist in a variety of conditions in the medical landscape, such as different scales, rotation, orientation, location, color, and brightness, etc. We used image data augmentation to generate additional synthetically modified data to train the DeMe Net model. The different types of augmentation techniques that we implemented are shown in Table 2.

Table 2. Dataset for image augmentation.

S.N.	Augmentation Technique	No. of Images
1	Rescale	1/255
2	Rotation Range	10
3	Width Shift Range	0.2
4	Height Shift Range	0.2
5	Shear Range	0.2
6	Horizontal Flip	True
7	Vertical Flip	True

DeMe Net Architecture

The deep learning method is providing promising results in complex problems. State-of-the-art deep learning models have a large number of parameters, which take longer to train and consume a larger space and processing power [11]. In this research, we have proposed the DeMe Net architecture. DeMe Net is a Convolutional Neural Network (CNN) based architecture that contains only 788,279 trained parameters which are 175 times less than the VGG16 architecture [12].

The DeMe network takes the input of images with a size of 28×28 pixels and distributes it in four parallel blocks. Each block contains two convolutional layers that with 32 filters of size 3×3, 24 filters of size 5×5, 16 filters of size 7×7, and eight filters of size 11×11 respectively. All the convolutional layers use the same padding and stride of 1×1. After the convolution layers, one batch normalization layer and a max-pooling layer with a pool size of 2×2 are stacked. Then we concatenated all the outputs of the parallel blocks and again distributed them into two parallel blocks. These two blocks contain three convolutional layers with 64 filters of size 3×3 and 32 filters of size 5×5 respectively and the same padding and stride of 1×1. Also contains the batch normalization and the pooling layers of the size of 2×2. Then we concatenated the outputs of two parallel blocks and used a single convolution layer. All the convolution layers used the ReLu activation function. Then we used the feed-forward network with two hidden layers with 256 and 127 units. The last layer of the model contains seven units with a SoftMax activation function for the output. The architecture of DeMe Net is shown in Figure 5.

To train the model, we tuned the hyperparameters of the DeMe Network to achieve higher accuracy. We used categorical cross-entropy loss to minimize the cost of true value and predicted value and also used Adam optimizer to handle sparse gradients on noisy problems. To train the model we initially set a learning rate of 0.00075 which can automatically reduce during training based on the validation accuracy. We used 30 epochs to train the model (the entire dataset is passed through the DeMe Network 30 times). Table 3 shows the hyperparameters used while training.

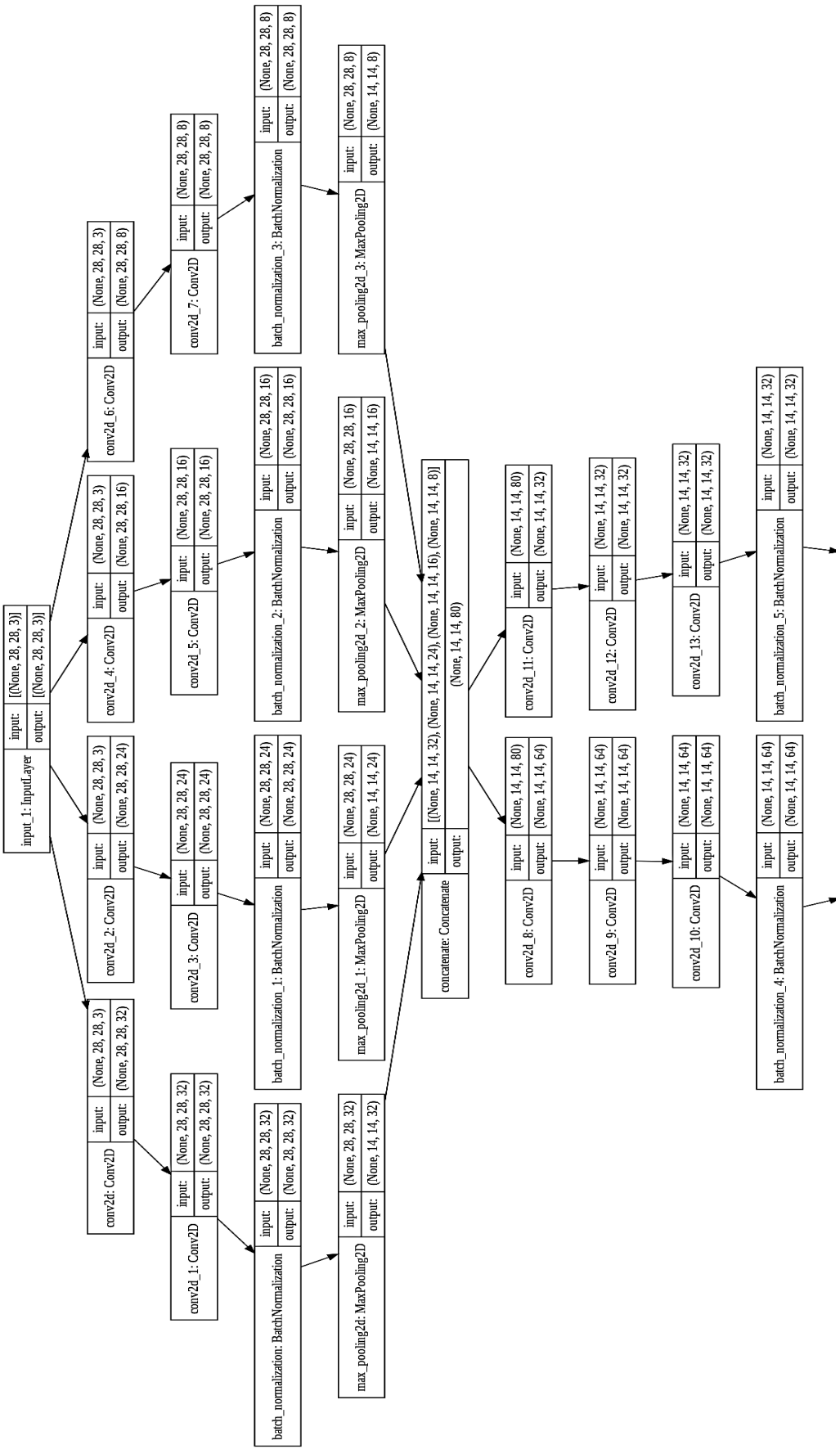
RESULT AND ANALYSIS

Experimental Analysis

In this research, we have created parallel convolutional neural networks to automatically diagnose dermatological melanoma disease. We have called this CNN network, DeMe Net. To train DeMe Network we used ISIC 2018 skin lesion image data set. After the pre-processing, we accumulated a total of 46985 skin lesion images. Then we split them into training and test set in the ratio of 8:2, we got 37588 and 9397 images. We trained the DeMe Net with a training set and converted this training set into new training and validation set in the ratio of 8:2 in real-time. While train the model, we used the learning rate reduction method and trained the network with 30 epochs to achieve the desired accuracy. The initial learning rate of 0.00075 provides stability in the accuracy and loss curve concerning epochs. Also, it prevents the model from underfitting and overfitting problems. The accuracy and loss graph for 30 epoches of DeMe Network training and validation is shown in Figure 6.

Table 3. Hyperparameters.

S.N.	Hyperparameter	Value
1	learning rate (α)	0.00075
2	β_1	0.9
3	β_2	0.999
4	Epochs	30
5	Batch Size	64
6	Optimizer	Adam
7	Loss	Categorical Cross Entropy



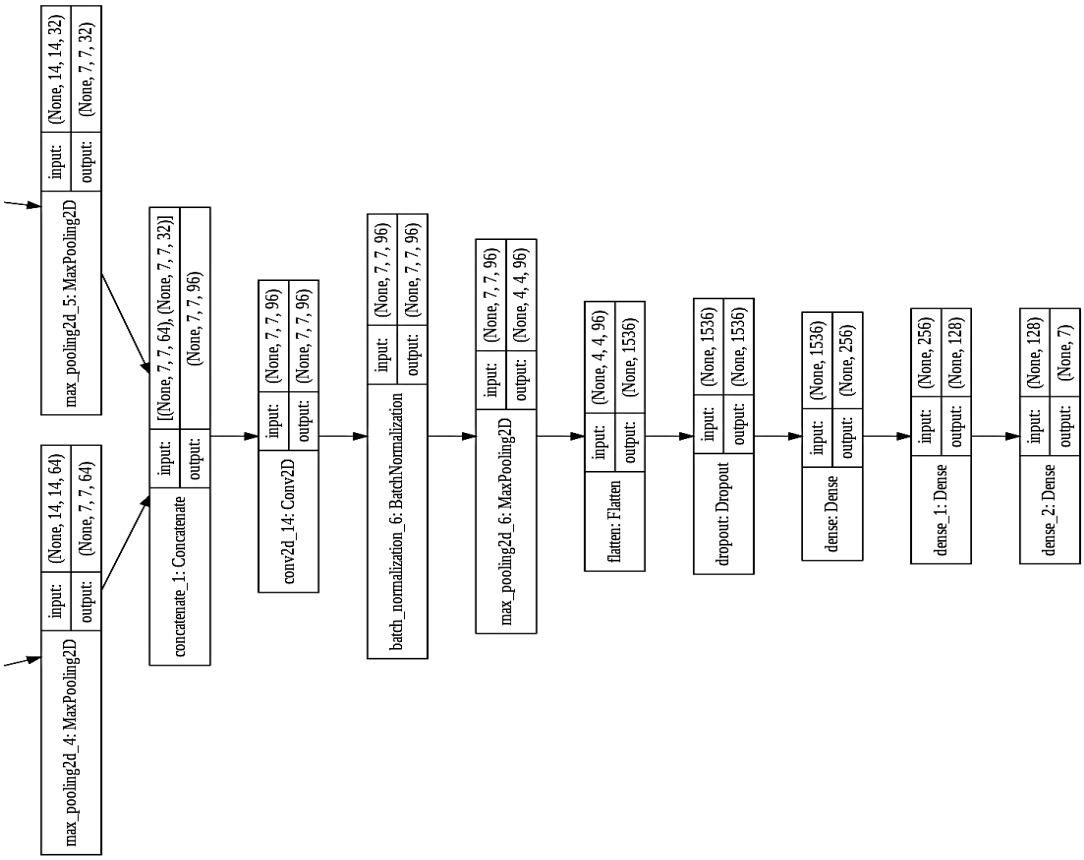


Figure 5. The Architecture of DeMe Net.

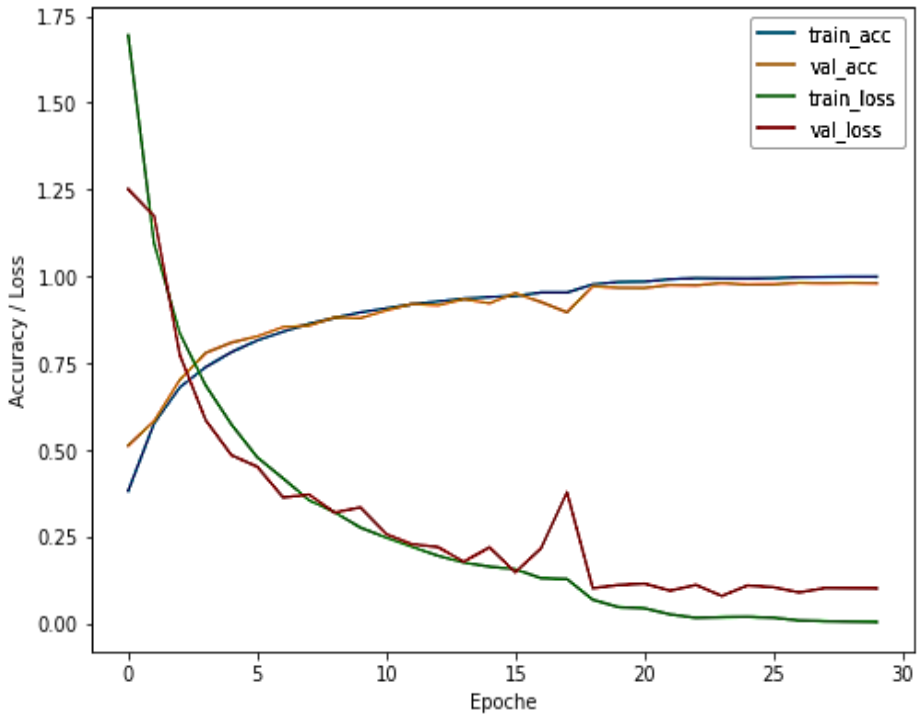


Figure 6. Accuracy and Loss Graph of the Training and Validation of DeMe Network.

Performance Analysis

Performance measurement is the way to evaluate the algorithm to the problem. This is the measurement of the predictions made by a trained network on the test dataset. There can be a variety of measurements used to evaluate the classification model's performance and accuracy. In this research, we used Confusion Matrix, Accuracy, Precision, Recall, and F1-Score as the measurement to evaluate the DeMe network.

- *True Positive (TP):* It defines the number of predictions where the model has been able to correctly predict the positive class as positive.
- *True Negative (TN):* It defines the number of predictions where the model has been able to correctly predict the negative class as negative.
- *False Positive (FP):* It defines the number of predictions where the model has been able to falsely predict the negative class as positive.
- *False Negatives (FN):* It defines the number of predictions where the model has been able to incorrectly predict the positive class as negative.
- *Accuracy:* It is a measure of the performance of the model. Accuracy is calculated as the total number of skins lesion images that were correctly predicted by the model.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP++FP+ TN+ FN}}$$

- *Precision:* It is the combination of precision and recall into a single measurement. Mathematically it is the harmonic mean of precision and recall.

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

- *Recall:* It is also known as True Positive Rate and Sensitivity. Recall is the proportion of true positive (TP) among all the images predicted correctly by the model.

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

- *F1-score:* It is the combination of precision and recall into a single measurement.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

We evaluate DeMe Network on the test set. According to the confusion matrix and the classification report, we achieved an accuracy of the 98.33% on the testing dataset and 99.85% on the training dataset. All the experiments were performed in Google Colab and the development codes uploaded on GitHub repository in https://github.com/Rajsoni03/DeMe-Net. The confusion matrix of the DeMe Network on the testing dataset shown in the Figure 7. The classification report of the DeMe Network on the testing dataset is shown on Figure 8.

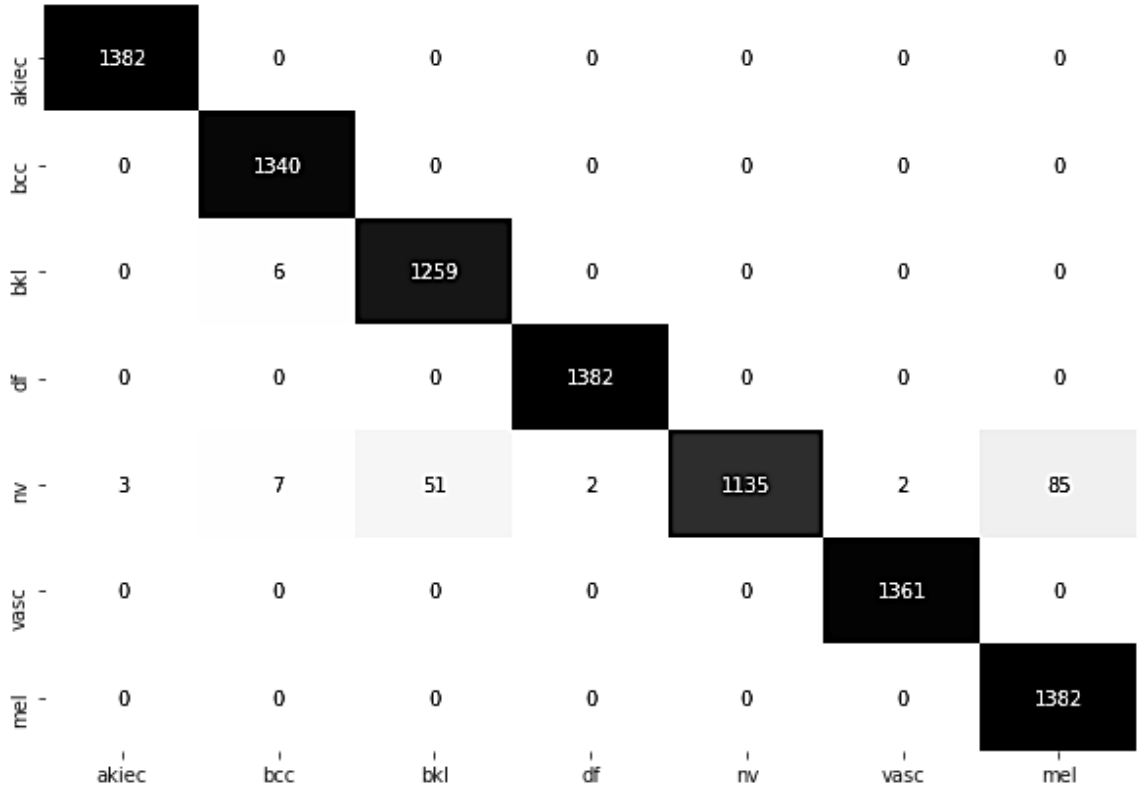


Figure 7. Confusion metrix.

	precision	recall	f1-score	support
akiec	1.00	1.00	1.00	1382
bcc	0.99	1.00	1.00	1340
bkl	0.96	1.00	0.98	1265
df	1.00	1.00	1.00	1382
nv	1.00	0.88	0.94	1285
vasc	1.00	1.00	1.00	1361
mel	0.94	1.00	0.97	1382
accuracy			0.98	9397
macro avg	0.98	0.98	0.98	9397
weighted avg	0.98	0.98	0.98	9397

Figure 8. Classification report on testing dataset.

Comparative Analysis

In the field of medical diagnosis, treatment of a disease requires an accurate diagnosis of that disease. We studied a variety of techniques for the automatic detection of skin cancer from images of

skin lesions. We compared these techniques based on their accuracy. We analysed that the performance of DeMe network is much higher than other techniques [10–12]. The comparative analysis of various deep learning techniques is shown in Table 4.

Table 4. Comparative analysis.

Authors	Methodology	Dataset	Accuracy
Polat and Koc [5]	CNN and OVA	HAM10000 [10]	92.90%
Pham *et al*. [6]	EfficientNetB4	HAM10000	89.97%
Mahajan *et al*. [7]	Meta-Derm Diagnosis network	ISIC 2018 [13], Derm7pt [14], SD-198 [15]	83.70%
Majtner *et al*. [8]	Ensemble of VGG16 and GoogLeNet	ISIC 2018	81.50%
Gessert *et al*. [9]	Ensembling	HAM10000, ISIC 2018	85.10%
Ours	DeMe Net	HAM10000	98.33%

CONCLUSION

In the medical field, early diagnosis is important in order to prevent the disease. The evaluation metric treats all classes equally but the ISIC dataset is highly imbalanced in terms of classes, which makes it very challenging. We used pre-processing (Data Sampling, Augmentation) techniques to make this data clean that helps to prevent overfitting and underfitting while training the model. The DeMe Net consists of a total of 788,823 parameters that make this model extremely compact in terms of size and it also reduces the computation cost that makes this model capable to work on low-end devices. Also, it reduces the prediction time with higher accuracy. The proposed DeMe Network evaluated using the test dataset consists of 9397 dermatology skin cancer images from the ISIC 2018 dataset and achieved an accuracy of 98.33%. This experiment shows that parallel processing can help to solve complex problems and reduce the size of the model and computation cost.

REFERENCES

1. Cancer.Net. Melanoma: Statistics [Online]. Availalbe from https://www.cancer.net/cancer-types/melanoma/statistics [Access Date: 01/Jan/2021]
2. Noel Codella, Veronica Rotemberg, Philipp Tschandl, Emre Celebi M, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, Harald Kittler, Allan Halpern. (2018). Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC). [Online]. https://arxiv.org/abs/1902.03368
3. https://www.cancercenter.com/cancer-types/skin-cancer/diagnosis-and-detection
4. Abbott LM, Smith SD. Smartphone apps for skin cancer diagnosis: Implications for patients and practitioners. Australas J Dermatol. 2018 Aug; 59(3): 168–170. https://doi.org/10.1111/ajd.12758, 2018
5. Polat K, Koc KO. Detection of Skin Diseases from Dermoscopy Image Using the combination of Convolutional Neural Network and One-versus All. Journal of Artificial Intelligence and Systems. 2020; 2: 80–97.
6. Pham TC, *et al*. Improving Skin-Disease Classification Based on Customized Loss Function. IEEE Access. 2020 Aug 26; 8: 150725–150737.
7. Kushagra Mahajan, Monika Sharma, Lovekesh Vig. Meta-DermDiagnosis: Few-Shot Skin Disease Identification using Meta-Learning. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2020; 730–731.
8. Majtner T, *et al.* (2018). Ensembling Convolutional Neural Networks for Skin Cancer Classification. [Online]. https://arxiv.org/abs/1808.05071.
9. Nils Gessert, *et al*. (2018). Skin Lesion Diagnosis using Ensembles, Unscaled Multi-Crop Evaluation and Loss Weighting. [Online]. arXiv:1808.01694 [cs.CV]
10. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data. 2018; 5: 180161 doi:10.1038/sdata.2018.161.

11. Heartbeat. Jameson Toole (5 Nov, 2019). Deep learning has a size problem: Shifting from state-of-the-art accuracy to state-of-the-art efficiency [Online]. Available from https://heartbeat.fritz.ai/deep-learning-has-a-size-problem-ea601304cd8
12. Keras. Keras Applications [Online]. Available from https://keras.io/api/applications/
13. Codella N, *et al*. (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC). [Online]. arXiv preprint arXiv: 1902.03368.
14. Kawahara J, *et al*. Seven-point checklist and skin lesion classification using multitask multimodal neural net. IEEE J Biomed Health Inform. 2018; 23(2): 538–546.
15. Xiaoxiao Sun, Jufeng Yang, Ming Sun, and Kai Wang. A benchmark for automatic visual classification of clinical skin disease images. In: European Conference on Computer Vision, Springer. 2016; 206–222.

