

# Smart Home-Healthcare for Skin Lesions Classification with IoT Based Data Collection Device.

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

26-10-2021 / 08-11-2021

CITATION

Islam, Md Khairul; Kaushal, Chetna; Amin, Md Al (2021): Smart Home-Healthcare for Skin Lesions Classification with IoT Based Data Collection Device.. TechRxiv. Preprint.  
<https://doi.org/10.36227/techrxiv.16870729.v1>

DOI

[10.36227/techrxiv.16870729.v1](https://doi.org/10.36227/techrxiv.16870729.v1)

# Smart Home-Healthcare for Skin Lesions Classification with IoT Based Data Collection Device.

MD KHAIRUL ISLAM<sup>1</sup>, CHETNA KAUSHAL<sup>2</sup>, MD AL AMIN<sup>3</sup>

<sup>1</sup>Dept. of Information Communication Technology, Islamic University, Kushtia, Bangladesh)

<sup>2</sup>Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

<sup>3</sup>Prime University, Dhaka-1216

**ABSTRACT** Skin lesions or malignancies have been a source of worry for many individuals in recent years. The underline reason behind it mainly the diet and environmental pollution. Yet many individuals are unaware of the issue and, more importantly, many people do not want to visit a hospital for diagnostic or therapeutic purposes. So, we have come up with a pipeline to diagnose skin lesions at home. Firstly, we proposed a IoT base data collection device which is accessible by patient to capture skin lesions image. This IoT device will encrypt and send the collected image towards a cloud storage; then it will decrypt the image send to the computer assisted diagnosis system. In CAD, we have implemented ensemble classifier. Ensemble classifier created depending on four deep learning classifiers namely VGG16, DenseNet201, Inception V3 and Efficient B7; whereas encryption and decryption performed in order to secure a patient data from unauthorized access. For skin lesions classification, we have used "HAM10000" dataset where 7 kind of skin lesions data included; Although DenseNet201 performed well, the ensemble model provides the highest accuracy with 87.22% as well as its test loss/error is lower than others with 0.4131.

**INDEX TERMS** Skin lesions, IoT, classification, cloud, ensemble

## I. INTRODUCTION

**S**kin lesion is a condition which is different from normal skin. In another way, it can be described as an abnormal change of the skin compared to the surrounding tissue. The American Society defines skin lesion as an abnormal lump, bump, ulcer, sore, or colored area of the skin. Skin lesions act as a risk factor for skin cancers, as shown by the fact that they were closely linked to an increased risk of squamous-cell carcinoma [1].

Literature review indicates that the overall number of instances of non-melanoma skin cancer is 1,042,056, and melanoma skin cancer is 287,723 [2]. These statistics surely show the risk of being diagnosed with melanoma. However, Skin cancer prevention would be possible if early detection of various skin lesions were possible in any case. In this study seven skin lesions described and detected including: Melanocytic nevus, melanomas, benign keratotic lesions, basal cell carcinoma, actinic ketarosis, and dermator fibroma. Melanocytic nevus is a type of melanocytic tumor that develops in the cells of the nevus where it acts as a precursor to cutaneous melanoma and indicates an increased chance of acquiring the cancer [3]. Melanocytic nevus that occurs in newborns and is present in the head and neck is known as

congenital melanocytic nevus. Congenital melanocytic nevi affect 0.2 to 6 percent of newborns around the world [4]. It usually begins on the skin, but it may also affect the mouth, intestines, and eyes. Seborrheic keratosis is a benign skin tumor that develops from cells and found across the external surface of the skin. Basal cell carcinoma is a skin cancer that grows on the skin's surface and it is the most prevalent form of skin cancer, and perhaps the most common type of cancer in general [5]. Actinic keratoses, alternatively referred to as solar keratoses, sun spots, or precancerous spots, are premalignant squamous lesions. [6]. Vascular lesions are a type of skin and basement membrane anomaly that can occur during or shortly after birth. Therefore, the treatment of acquired vascular birthmarks is one of the most frequently requested and performed cutaneous laser procedures. [7]. Dermatofibromas are benign fibrous nodules that are formed when extra cells accumulate in the deeper layers of the skin and predominantly found in the tissues including its lower legs [8].

The biopsy technique is widely used in almost every medical lab to diagnose skin cancer, but an automated skin cancer detection technique is used in this work to detect skin cancer.

A biopsy is a treatment modality in which a small amount of tissue is removed and observed under the microscope. If an initial examination reveals that an area of the body's tissue is irregular, a doctor should suggest a biopsy. Punch biopsy is a form of biopsy used to diagnose skin cancer. It is the most common biopsy procedure, and it involves cutting a cylindrical specimen of skin tissue with a circular blade [9]. Several complications may occur during biopsy to diagnose skin cancer, such as pain, bleeding, damage to other tissues, incorrect biopsy site, and difficulty covering the biopsy area, as well as other factors such as taking more time, going to the operating room, and costing for biopsy [10]. Even an attack can occur during biopsy which is known as a vasovagal attack. The most obvious type of syncope is vasovagal syncope, which occurs most often in anxious patients as well as those with an abnormal autonomic nervous system.

However, in this study, automated skin lesion detection is used. Automatic skin lesion detection is performed using a methodology in which firstly a computer with an algorithm takes a snapshot of the target region of skin and then calculates the true result [11]. After passing through many essential steps including image preprocessing, image segmentation, and image extraction, a Convolutional Neural Network (CNN) is used as a method to detect the skin lesion.

## II. LITERATURE REVIEW

Deep/Machine learning is an emerging technology to automate our day-to-day life task and reduce time and cost. They have been widely used in various fields like home/office, agriculture, industry, health etc. But in the medical sector, its implantation is increasing nowadays which is also known as Medical imaging. Medical imaging is implemented in different image processing or diagnosing techniques like MRI, X-Ray, Ultrasound, Endoscopy, Electrocardiography, etc. In medical imaging, computer vision technology techniques are highly used to classify among many types/targets.

In [12] presents a study of medical image processing such as classification, pattern recognition, segmentation, etc. using deep neural networks in the contemporary medical arena. They specify several state-of-the-art methods for carrying out the job. In addition, certain disadvantages are mentioned, such as data limitation, enhanced feature extraction and representation techniques for better precision, or black box-like character of deep learning methods, etc. In [13], the application of deep learning approaches in the various health sectors (medical imaging) such as histological/microscopic elements identification, gastrointestinal early diagnosis, cardiac/tumor detection, Alzheimer/Parkinson detection, etc are explored. They looked at a number of studies that used either pre-trained (GoogLeNet, Alex Net, and LeNet) or personalized deep learning models to detect. They also defined various datasets that are often used to train a reliable and accurate model for a particular health problem. Using machine learning techniques, khan2020healthcare updated an IoMT (Internet of Medical Things) for heart disease diagnosis. To

increase the accuracy of the prediction, they used salp swarm optimization and an adaptive neuro-fuzzy inference method. They assessed the risk of heart disease using data on chest pain, cholesterol levels, blood pressure, sex, age, and blood sugar levels. In [14], "Wisconsin Diagnostic Breast Cancer" dataset and machine learning algorithm to detect breast cancer has been employed. Both of them did not, however, have any information about how the data was gathered using IoT devices or how IoT device helped/linked to their proposed machine learning approaches.

In case of skin lesions classification, 20 dermatologists and a faster region-based CNN (FRCNN) model are compared (malignant and benign tumors). Where FRCNN has outperformed accredited dermatologists. However, they did not use any other pre-trained models or fine tuning to increase accuracy. Another study li2018skin, classified skin lesions melanoma using deep learning methods. They used a fully connected residual network for segmentation and classification. Further lesion index calculation unit has been used to validate the result of the classification. In their approaches, no pretrained weights value was used to increase the accuracy. Besides they used only shifting procedure for data augmentation. In [15], the authors used image enhancement techniques, segmentation techniques and then extracted 15 features from each image which they fed to the deep neural network and hybrid AdaBoost-SVM model. They didn't employ any data augmentation techniques. VGG-16 is used to train the skin melanoma detector classifier in [16]. They used random weight initialization for training and fine-tuning on the VGG-16 network to improve accuracy. But no experiments about fine-tuning the different layers has been carried out and their accuracy metric is very low compared to the recent study.

IoT is an integration of sensors and software to automate our daily life. In [17], an E-health monitoring system is created that tracks a person's life cycle and links them to a health management service provider. Their IoT device is used for tracking diabetes, pulse rate, and kidney function, etc. However, they failed to mention the IoT device that would be used to test their hypothesis. Another study by pinto2017we employed a health care system for aging people to assist them with regular monitoring. They developed an android app called 'We-Care' with a secure gateway which get information from 'We-Watch' wristband. But they didn't mention a proper description of what types of data will be collected and also no specific task of the app has been mentioned.

masood2018computer discussed about IoT based computer assisted pulmonary cancer detection system which gather data using wearable sensors and classification of computed tomography images found from online sources to evaluate cancer stage. The study used fully connected CNN along with a medical advisor to prescribe the further procedures. They didn't propose any automated device to capture CT images from patient.

An Iot-based device created by kodali2015implementation

was used to collect medical information and archive it in the cloud for further assessment. A Cloud-Fog computing was employed to address the data availability to the IoT based healthcare by mahmud2018cloud. The fog cloud aids in cost reduction and closing of data sources to the monitoring body.

### III. METHODOLOGY

In this section, the method used to detect skin cancer using deep neural network based on IoT has been described. For capturing the image of the infected skin Raspberry pi sensor board is used.

After capturing infected skin image from patient, two major steps are undertaken to send the data into a central repository. These are: OS Installation and Internet Setting.

The captured image is then submitted to the CAD (Computer Assisted Diagnosed) system. CAD system preprocess the image for further processing. Then an image segmentation technique is used to separate the target object from the image. Feature extraction technique is then extended to the segmented area of interest. Further the data is separated into train, validation and test. At last, the Classifier model is used to identify each patient's current condition. The output of the classifier is submitted to the central repository and patients will be recommended according to the CAD outcome by the recommendation body and the data will also be saved for future model training. This helps to make our infrastructure more stable.

#### A. PATIENT CONNECTION WITH THE SMART HEALTHCARE

A patient is the beneficiary of all medical services rendered by health providers. In this study the patients are monitored to identify the existence of Skin Cancer using modern technology (IoT and deep learning) without the involvement of pathologists. Therefore, the patient can easily utilize the IoT based data device, which is connected with CAD system, to evaluate the skin lesions condition and prescribe accordingly at any time without going to the hospitals.

#### B. IOT BASED DATA COLLECTION SYSTEM

##### C. IOT

The Internet of Things corresponds to a collection of interconnected devices that can capture and transmit data over a mobile or wifi network without human interference. IoT allows objects to be controlled accurately from a remote area with the help of network infrastructure and computer-assisted systems. In the future, it will cross the border, even the earth, and be used for a new and important purpose. And often incorporate into a broad range of fields and devices to work in a wide range of areas [18]. Raspberry Camera: The Pi camera chipset is a compact, portable camera that uses the Mobile Industry Processor Interface (MIPI) serial protocol to support the Raspberry Pi. And the decreased cost of the Raspberry Pi Foundation delivers high-performance single-board computers to teach and address real-world issues. It is now feasible to use multiple lenses, automatically change the

focus and exposure, and mount stable video. Used in deep learning, machine learning-based projects such as recognition, classification, and other functions. [19].

#### D. DATA COLLECTION

Data collection is the most critical step in this study. Without a large volume of data and careful identification of each data point, it is very difficult to collect information and learn more. We have therefore suggested a IoT base stable, sensible, and simple pi camera for our data collection method in our study. Although we initially train our classifiers with a publicly accessible data set. There are many publicly available datasets. Like "BCN20000" which consists of almost 194k dermoscopic images collected from Barcelona's hospital from 2010 to 2016 [20]. "Human Against Machine with 10000 training images (HAM10000)" consisted of almost 10015 images of dermatoscopic skin lesions tagged with the correct labeling by a pathologist or multi-step follow-up, expert opinion, or in-vivo confocal microscopy validation [21]. "PAD-UFES-20" is comparatively new dataset composed in 2020 of almost 2298 images from 1373 using smartphones [22].

#### E. IMAGE PREPROCESSING

The captured image has to be preprocessed before feeding the image into the neural network, so that the neural network can filter further information into each of the layer. Traditional methods for image pre-processing refer to the transformation of images into raw data that can be fed by the neural network. The aim of pre-processing is to optimize image data that eliminates unwilling information or improves those image features that are necessary for further processing. In one respect, choosing the very first set of pixels in a image in the RGB color space is a procedure performed by different algorithm operating on a computer device. [23]. Techniques such as resize, mean normalization, standardization, smoothening, blurring, etc. are used in image preprocessing. The HAM10000 images are resized to 64\*64 and 96\*96 in the RGB color space. We also conducted Data-Augmentation technique to supply the neural network with a wide variety of images as a batch. Data augmentation includes rotation, scaling, shift and fill mode etc.

#### F. FEATURE EXTRACTION

The patterns of the object in an image are referred to as features. A triangle, for example, has three corners and three sides, which are the characteristics that our eyes use to recognize the triangle. Similarly, the convolutional network can remove key features from each image to aid in the identification of a certain class or target value. Convolutional kernels remove valuable features from the original image, reducing the dimension and resulting in more efficient features with fewer redundant data. This method is known as a feature mapping process. This reduced representation of the original image would provide a detailed understanding and more accuracy. [24].

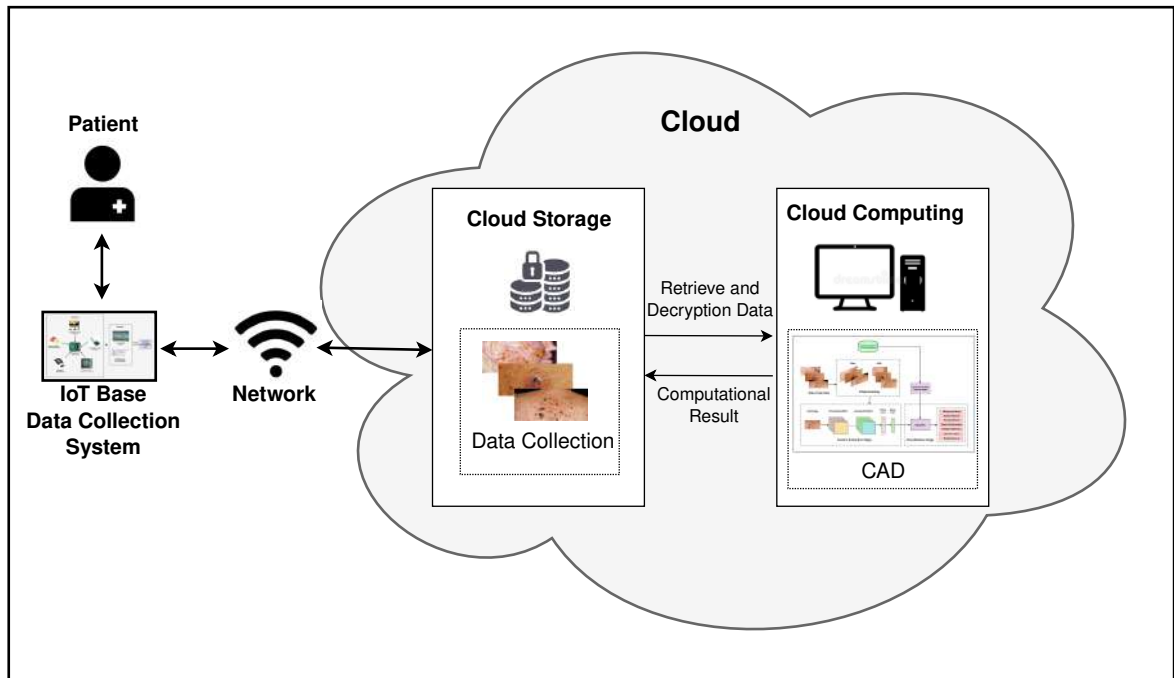


FIGURE 1. Proposed Methodology for IoT-Deep Learning based Skin Lesion Classification and Recommendation.

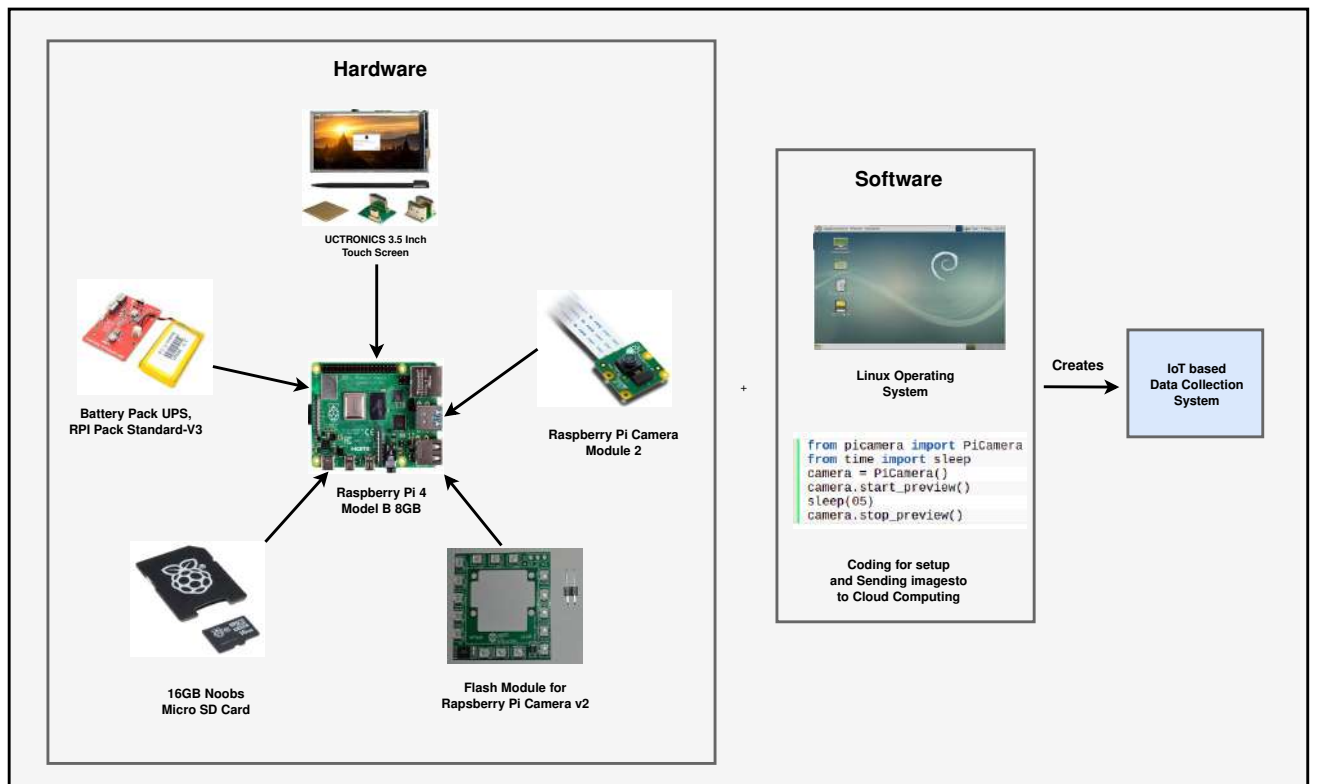


FIGURE 2. IoT based Data Collection System.

### G. CLASSIFIER

A classifier is an algorithm that converts data into one of several categories or classes. And convolutional neural



network or ConvNet is one type of neural network mostly used in image classification. There are several layers in the convolutional neural network, such as: input layer, hidden layer, output layer, etc. Each layer has several neurons or nodes that take information from the previously hidden layer or input layer that used the mapping function to process the data and transfer it to the next layer [25]. One of the categories is generated by the output layer. The input layer takes raw data, which in our case would be the tensor of the original image. And this overall mechanism is supervised learning because computer trains using the target value of the training data set [26].

### 1) Convolutional Layer

Convolution layer used to extract information/features using kernel/filters. Kernel size is smaller than the input image which scans the images spatial position step by step. Then bias and other required elements add, also weighted sum are calculated. At last, the output of the layer is passing through the nonlinear activation function to get new features for next convolutional layer [25]. Generally used activation function are ReLU, sigmoid, tanh etc. The function is processed in convolutional layer denoted by:

$$x_j^l = f\left(\sum_i^m (x_i^{l-1} * w_{ij}^l + b_j)\right) \quad (1)$$

$x_j^l$  denotes the  $j^{th}$  features of  $l_{th}$  layer where  $w_{ij}$  is the weight between the  $j^{th}$  feature of  $l_{th}$  and  $i_{th}$  feature of  $l-1_{th}$  feature. And  $b_j$  is the bias of  $j^{th}$ ,  $m$  is the number of features created in  $l_{th}$  layer.

Traditionally, convolutional layer kernel size is 3\*3. Because it derives the most important features and less information loss occurred. Beside it reduce the number of parameters and hence reduce the time of calculation.

### 2) Pooling layer

Pooling layer generally used between two convolutional layer. It tries to compress the features which found from the previous convolutional layer. Compression is done by taking the max or average value from a particular region. Max pooling is mostly used as it produce the best result [25].

### 3) Fully Connected Layer

The convolutional and pooling layers provide room for the features of the images. Fully connected layer are layer where all the previous layer neurons are connected to the next layer. It can be considered an affordable way to learn a linear function from the feature region [27].

### 4) Output Layer

Softmax classifier is used as output layer usually for classification of multiple class. There are  $J$  images  $x_i^J$ , where each image labelled with value  $y_i \in (1, 2, 3, \dots, k)$ .  $k$  is the total number of classes in the dataset and  $y_i$  is the targeted label.

For image  $x_i$  there will be probability for each classes of  $k$  is denoted by equation:

$$h_\theta(x_i) = \begin{bmatrix} p(y_i = 1|x_i; \theta) \\ p(y_i = 2|x_i; \theta) \\ p(y_i = 3|x_i; \theta) \\ \dots\dots\dots \\ \dots\dots\dots \\ \dots\dots\dots \\ p(y_i = k|x_i; \theta) \end{bmatrix} \quad (2)$$

In (2), the sum of all the classes probability is equal to 1 and  $\theta$  represent the parameter of the classifier (softmax classifier).

There are several ConvNet classifiers in Deep Learning, e.g. Xception, VGG16, ResNet50, ResNet101, DensNet121, MobileNet and so on.

#### a: VGG16

The VGG16 architecture is a convolutional neural network, and it was used in 2014 in an annual competition called the "ImageNet Large Scale Visual Recognition Challenge." The ImageNet dataset includes RGB-channel images with a fixed size of 224\*224; in our case, it is 96\*96. In this architecture, 16 layers are used, including convolution layers of 3x3 filter with stride 1 and always the same padding and max pool layer of 2x2 filter with stride 2.

#### b: DenseNet201

One of the neural networks for visual object recognition is DenseNet-201, which has 201 layers. The network's image input size is 224\*224 pixels. The name DenseNet arises from the fact that each layer in a DenseNet architecture is linked to every other layer. Instead of using summation, the DenseNet paper suggests concatenating outputs from previous layers. Transition layers are used by DenseNet. Convolution with a kernel size of 1 is accompanied by 2x2 average pooling with a stride of 2.

#### c: Inception V3

Inception-v3 is a convolutional neural network architecture from Google's Inception family V3. The third version in a series of Deep Learning Convolutional Architectures makes many improvements, such as Label Smoothing and Factorized 7 x 7 convolutions. It was trained using a dataset of 1,000 classes from the original ImageNet dataset, which was trained with over 1 million trajectories.

#### d: Efficient B7

EfficientNet-B7 achieves new state-of-the-art 84.4 percent top-1 / 97.1 percent top-5 accuracy despite being 8.4% smaller than the best current CNN by scaling up the baseline network. The total number of layers in EfficientNet-B7 is 813, and all of these layers can be created using only five

modules. Module 1 serves as a starting point for the sub-blocks, module 2 serves as a starting point for the first sub-block of each of the seven main blocks except the first, module 3 serves as a skip link for all of the sub-blocks, and module 4 helps to combine the skip connections in the first sub-blocks. Finally, module 5 connects each sub-block to the one before it in a skip link, and this module is used to merge them.

#### e: Ensemble Model

Ensemble Model is a single model where multiple model's (in our case IceptionV3, VGG16, DensNet121, Efficient B7 and a custom model) prediction combined together and predict one single outcome. Its like decision making board which decide based on the predictions of different models [28]. Generally a base model have unique error on a dataset samples, but ensemble model use all the models prediction which reduce the error than a base model. Different machine and deep learning based study used ensemble model for better accuracy and to get robust model for classification task [29] [30] [28]. In our case we have used averaging ensemble model as shown in figure 3. In averaging ensemble model we take average of all models probability for classes and make final decision based on the average outcome of each class. Suppose, we have  $L$  models where each model  $l_i$  predict  $k$  probability as we have  $k$  classes. Then the final class of ensemble model for a particular image  $x_i$  will be:

$$E(x_i) = \frac{\sum_i^L h_{\theta}^{l_i}(x_i)}{L} \quad (3)$$

In (3),  $E(x_i)$  denotes Ensemble outcome for  $x_i$  image where  $h_{\theta}^{l_i}(x_i)$  defines  $i_{th} \in L$  model's classifier function outcome.

#### 5) Central Repository

Repository is the disk where all data will be stored. Central Repository means the Criminal Justice Information System Central Repository of the Department of Public Safety and Correctional Services. In the central repository, we maintain our data in an uncompressed form for easy handing, accessing and viewing of the data. The central repository is usable by a computing device via a communication network. The approach involves the steps of connecting via the communication network to the central repository depending on the user request [31].

#### 6) Recommendation System

### IV. EXPERIMENTS AND RESULTS

This section describes the result of the proposed methods in our methodology and their experimental comparison.

#### A. DATASET

HAM10000 is the selected dataset for our proposed methodology. We have split our dataset into train (7210), test (1002) and validation (1803).

HAM10000 dataset has almost 10015 dermoscopic images of seven different skin lesions disease. In Figure 4 shows

the number and percentage of dermoscopic image regarding each skin lesion disease. And it is clearly noticeable that the dataset is imbalanced.

The dermoscopic images of seven diseases are shown in the figure 5.

#### B. EXPERIMENTAL SETUP

In the experimental setup, we have discussed computational and structural configuration for our proposed methodology in detail. We have used the support of Numpy, Scikit learn, Pandas, Keras, Tensorflow, etc. Keras is extensively used for neural network training with TensorFlow as a backend.

The dataset contains images of different sizes. So, we resized the images to 96\*96 pixel size so that it can be easily feed to neural network. Then we converted the ground truth to categorical values. The train and validation set is used to learn our models. The test set is remained to validate our models performance on unseen data. Different pretrained models are loaded to get the best classification result out of the dataset. But the top classification layer (including flatten, fully connected layer, softmax) is customized to get the best results. Imagenet' weights are initialized for training. Then data augmentation techniques used in each epoch before feeding to the neural network. In each epoch there are 144 steps for training dataset and 36 steps for validation dataset with batch size of 50. Different optimizers has been used alternatively like Adam, RMSprop etc. 'categorical crossentropy' is used as a loss function. We also used learning rate reduction function to reduce the learning rate as the epoch increases.

#### C. EVALUATION METRIC

In this section we have described the evaluation metrics used to validate the performance of the proposed methodology. In general, accuracy is used to define the performance of a particular model. But in medical imaging accuracy of model is not enough to have a exact idea of the a model. So, there are many more metric to evaluate a deep learning model like precision, recall, ROC curve and F1-score. We have used all of these metrics to evaluate and understand a model performance.

Confusion matrix is considered the most comprehensive way to describe all the metrics and is determined by:

- **True Positive** - predicted target label true and ground label also true
- **False Positive** - predicted target label true but ground label actually negative
- **True Negative** - predicted target label negative and ground label also negative
- **False Negative** - predicted target label negative but ground label actually true

So, a target label correctly classified means that should be True Positive (TP) or True Negative (TN). Similarly, when target label wrongly classified it will be False Positive (FP) or False Negative (FN).

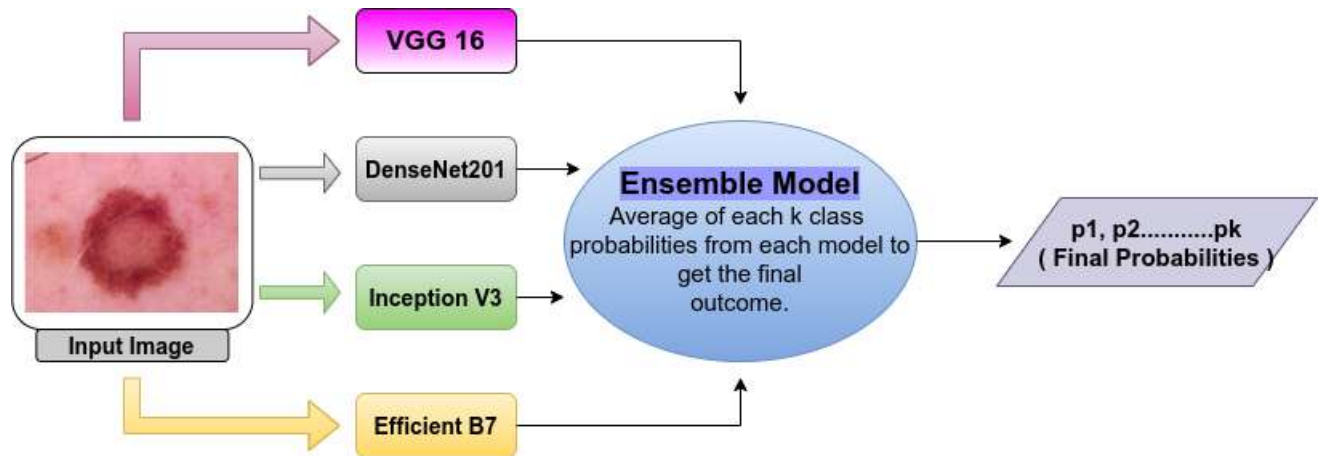


FIGURE 3. Flow diagram of Ensemble model.

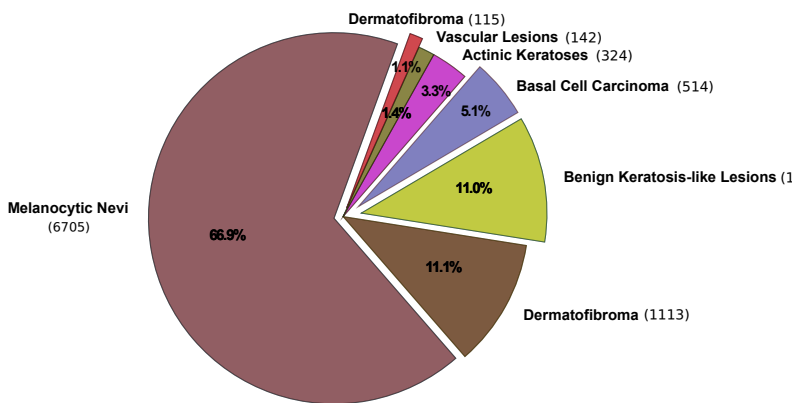


FIGURE 4. A probabilistic representation of the seven classes in the HAM10000 dataset.

Accuracy is the sum of all correctly classified data points divided by all data points :

$$Accuracy = \frac{\sum (\text{True Positive} + \text{True Negative})}{\sum (\text{True Positive} + \text{False Positive} + \text{True Negative} + \text{False Negative})} \quad (4)$$

Precision defines the proportion of patients correctly predicted as having skin cancer and predicted all cancers patients. That means sum of all true positives labels by sum of all true predicted true labels. Where recall/sensitivity specify the proportion of patients correctly predicted as having skin cancer and actually all cancers patients. So, sum of all true positives labels by sum of all actual true labels. They are determined by :

$$Precision = \frac{\sum \text{True Positive}}{\sum \text{True Positive} + \text{False Positive}} \quad (5)$$

$$Recall = \frac{\sum \text{True Positive}}{\sum \text{True Positive} + \text{False Negative}} \quad (6)$$

Precision tries to minimize the false positives rates. So, in our case with the increase of precision, it tries reduce wrongly predicted cancer patient. But recall is most important for our

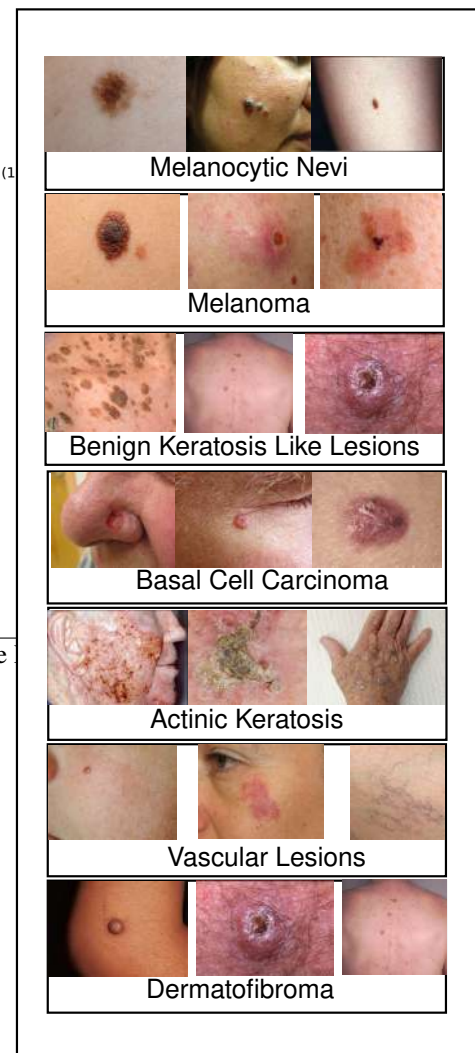


FIGURE 5. Different Skin lesions.

case. As we don't want to miss any cancer patients predicted as normal. Recall tries to minimize the false negative rates.



F1 score helps to understand both precision and recall from one score. We can take the average of both precision and recall, but it reduce the score drastically if any smaller (either precision or recall) value appear. So, they come up with a balance equation which tries to more closer to the smaller value and give more accurate score that is :

$$f1 - score = \frac{\sum(2 \times precision \times recall)}{\sum(precision + recall)} \quad (7)$$

A function which we try to minimize or maximize is called objective function. However, when we sought to mitigate it precisely, we called it a cost function or a loss function. The cost function or loss function is used to describe all the good or bad points using a single scalar value. It helps to understand the error of a particular model and help to rank all the models. In our case we have used categorical cross-entropy. Categorical cross-entropy explicitly used for multi-class classification and described as:

$$Categorical-cross-entropy = - \sum_{i=1}^{output\_size} y_i \log \hat{y}_i \quad (8)$$

Roc auc score helps to understand the separable capability of a particular classifier.

In table 1, dataset's train and validation accuracy and train and validation loss for VGG-16, Inception-V3, Efficient-B7 and DenseNet-201 are shown.

Table 2 shows the testing data's accuracy, loss, precision, recall, f1 score and roc auc score for all the deep learning models.

## V. DISCUSSION

In this research, we tried to develop a Computer Assisted Diagnosis (CAD) system which can detect skin diseases automatically using the image of skin lesions. First of all we have collected dermoscopic images from patient using IoT device. In CAD we have then implemented a deep learning based image classifier using transfer learning and ensembling of different models (as shown in fig 3) for skin lesions classification. For transfer learning approach we have used popular pre-trained models like VGG-16, Inception-V3, Efficient-B7 and DenseNet-201 etc. We have performed 50 epoch for each model on HAM10000 dataset. The final outcome of transfer learning is quite satisfactory.

The figure 6 depicts the relative positions of training and validation accuracy in each epoch for all four pretrained models, where figure 7 shows the relative positions losses. In table 1 show training and validation accuracy, loss etc. In table 2 show the test's metrics, the visual representation of test accuracy and loss is shown in figure ??.

The difference of VGG-16 model's train, validation and test accuracy is low, so the model neither overfit nor underfit. InceptionV3 test accuracy is comparatively low than other models. Both VGG-16 and InceptionV3 took less time to train in each epoch than the other two models. The test loss for InceptionV3 is higher than VGG-16.

Efficient-B7 took more times than other models because of its huge number of parameters to train. The model also shows

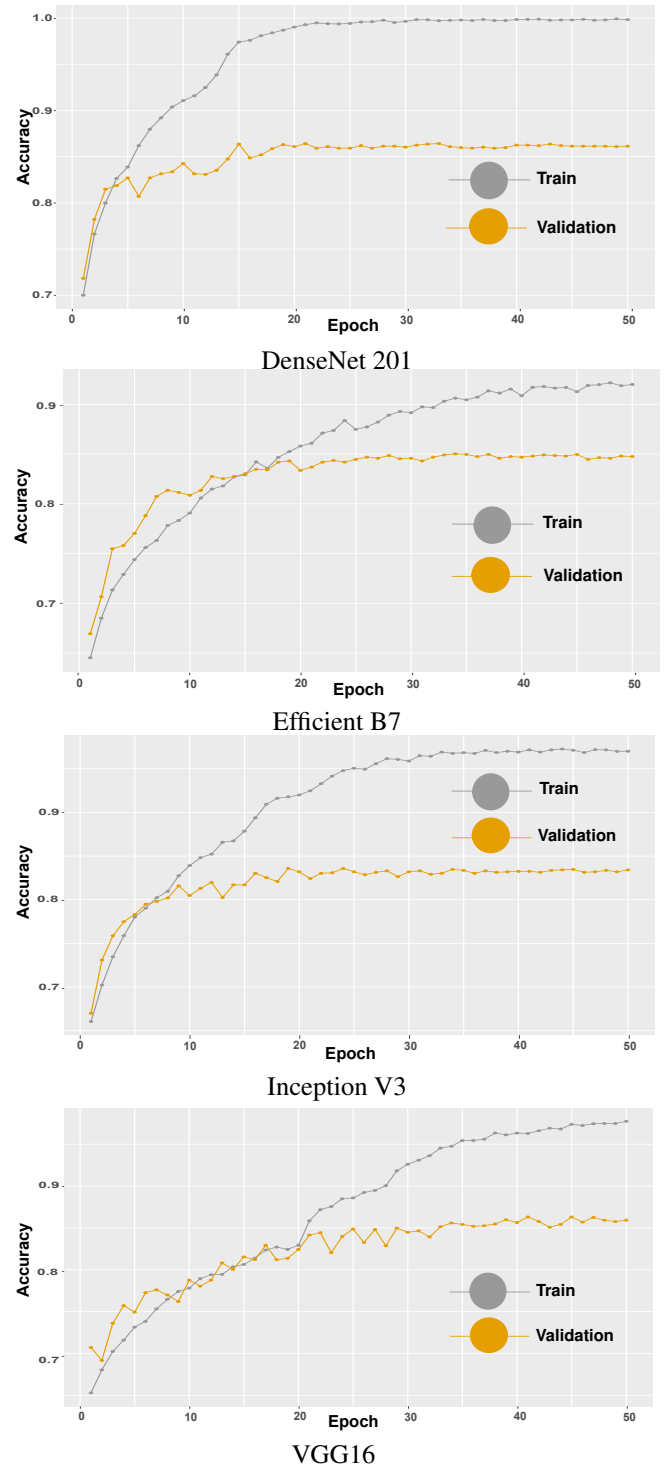


FIGURE 6. Expression of accuracy in each epoch both for training and validation dataset.

a good accuracy (train = 91.82, validation = 85.14 and test = 82.93) with low loss (train = 0.2173, validation = 0.4829 and test = 0.6954). Efficient-B7's roc auc score also good that means its separable/classification quality considerable. But DenseNet-201 has the best separable quality than any others model we have applied. DenseNet-201 have the best train

**TABLE 1.** Train and Validation metrics of Deep learning models.

Model	Train Accuracy	Train Loss	Validation Accuracy	Validation Loss
VGG-16	92.74	0.1976	83.97	0.5837
Inception-V3	94.75	0.1444	83.58	0.6582
Efficient-B7	91.82	0.2173	85.14	0.4829
DenseNet-201	99.71	0.0086	86.36	0.8200
Ensemble Model	-	-	89.36	0.3369

**TABLE 2.** Test metrics of Deep learning models.

Model	Test Accuracy	Test Loss	Precision	Recall	F1 Score	Roc Auc Score
VGG-16	83.33	0.7207	83.33	83.33	83.33	0.9074
Inception-V3	80.94	0.7858	80.93	80.93	80.93	0.9321
Efficient-B7	82.93	0.6954	82.93	82.93	82.93	0.9546
DenseNet-201	85.33	0.9501	85.32	85.32	85.32	0.9599
Ensemble Model	87.22	0.4131	86.72	87.22	86.60	0.9745

(99.71%), validation (86.36%) and test (85.33%) accuracy. But its validation and test loss is higher among all the models.

In case of precision and recall, DenseNet-201 win the race with 85.32% and 85.32%, respectively. Precision defines the correctly predicted true positives ( present of skin malignant) against total predicted true positives. Where recall/sensitivity defines the correctly predicted skin malignant against the total actual skin malignant cases.

Transfer learning took comparatively less time to train than baseline models. Because we have used pre-initialized weights from "ImageNet". ImageNet was trained on almost 14 million data for 20,000 classes [32]. So, it has a well established weights to train.

We also performed ensemble of our trained models. Ensemble model is like a decision making board which reduce the generalisation error. In our case we have ensemble VGG16, DenseNet201, InceptionV3 and EfficientB7 together using average ensembles methods. And our ensemble model is outperform than all other models on test dataset as shown in figure ?? . Not only loss is reduced to 0.41 but also accuracy is increased to 87.22% for test. Both loss and accuracy is improved than any other base pretrained models. The accuracy, recall, and f1-score of the ensemble model outperform than that of other models. Ensemble model's roc curve shown in figure 10. In case of skin lesions classification, roc-auc score (with 0.9745 ) of ensemble model defines the best separable ability than any other models as far as we know.

Our CAD system use the classifier to classify the skin lesions type and then pass the patients result to the central repository. A recommendar team have the access to the central repository. And they will prescribed or recommend the patients about next step. So, our IoT and deep learning base computer assisted diagnosis (CAD) will help the patients to diagnosis skin diseases at home. It will reduce both the cost and time for diagnosis.

## VI. CONCLUSION

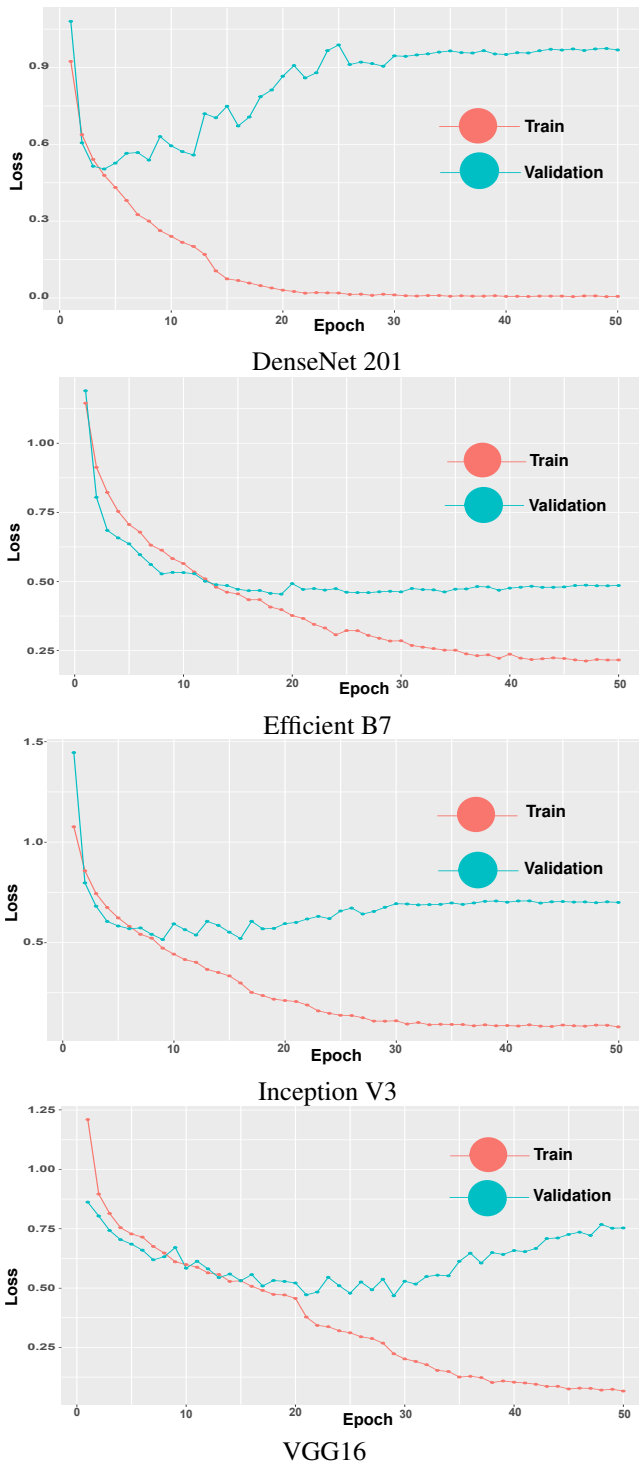
In this research we proposed an IoT based Computer Assisted Diagnosis system. A deep learning model implemented in the CAD system which can classify skin lesions from skin images.

We have performed transfer learning models like VGG-16, Inception-V3, Efficient-B7 and DenseNet-201 etc. This models show exceptional performance for classifying skin lesions. Parameters tuning help us to increase our accuracy and produce more robust model on the dataset. We then ensemble all the models which outperform than any other models. Our deep learning method and pre-trained models are available for more use of the classifier, as well as for researchers and students to help with more precise improvement.

Our IoT based CAD system also connected to a central repository and a recommend team. A central repository is used to store data for further processing and use, and the recommend team assists a specific patient in learning about his or her current conditions and referring him or her to a specialist doctor. Overall, our IoT-based CAD system connects the patient, diagnostic center, and doctor in a chain.

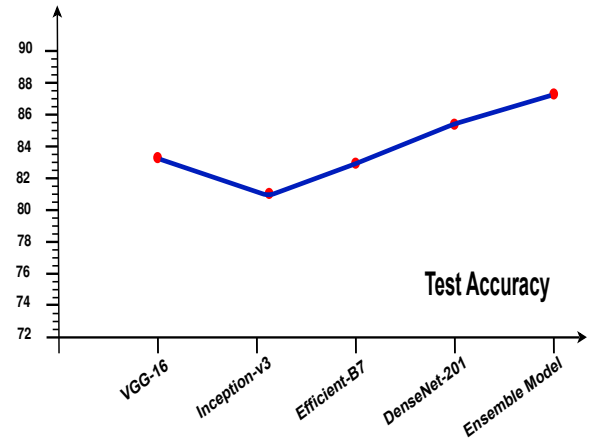
## REFERENCES

- [1] Jan N Bouwes Bavinck, Sylvie Euvrard, Luigi Naldi, Ingo Nindl, Charlotte M Proby, Rachel Neale, Damiano Abeni, Gian P Tessari, Mariet CW Feltkamp, Alain Claudy, et al. Keratotic skin lesions and other risk factors are associated with skin cancer in organ-transplant recipients: a case-control study in the netherlands, united kingdom, germany, france, and italy. *Journal of Investigative Dermatology*, 127(7):1647–1656, 2007.
- [2] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L Siegel, Lindsey A Torre, and Ahmedin Jemal. *Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA: a cancer journal for clinicians, 68(6):394–424, 2018.
- [3] Hensin Tsao, Caroline Bevona, William Goggins, and Timothy Quinn. The transformation rate of moles (melanocytic nevi) into cutaneous melanoma: a population-based estimate. *Archives of dermatology*, 139(3):282–288, 2003.
- [4] Vito Ingordo, Ciro Gentile, Silvia S Iannazzone, Francesco Cusano, and Luigi Naldi. Congenital melanocytic nevus: an epidemiologic study in italy. *Dermatology*, 214(3):227–230, 2007.
- [5] CSM Wong, RC Strange, and JT Lear. Basal cell carcinoma. *Bmj*, 327(7418):794–798, 2003.
- [6] Jeffrey P Callen, David R Bickers, and Ronald L Moy. Actinic keratoses. *Journal of the American Academy of Dermatology*, 36(4):650–653, 1997.

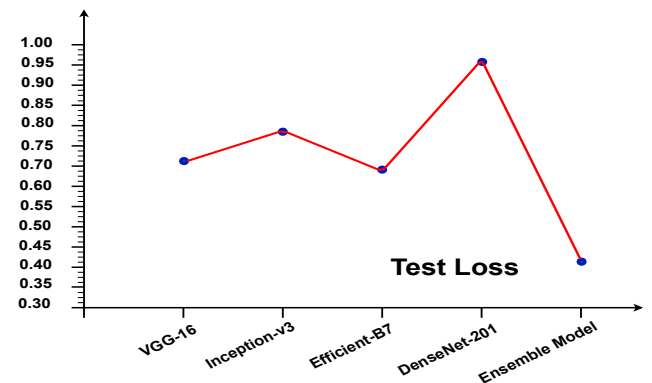


**FIGURE 7.** Expression of loss in each epoch both for training and validation dataset.

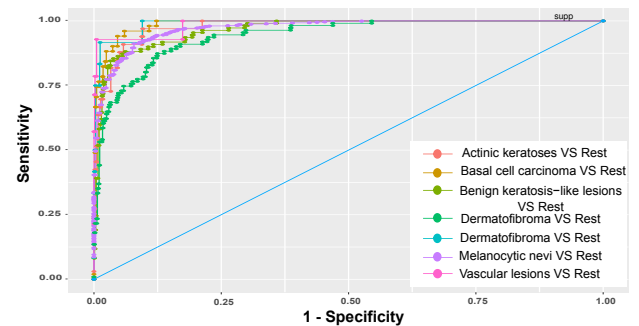
- [7] Susanne Astner and R Rox Anderson. Treating vascular lesions. *Dermatologic therapy*, 18(3):267–281, 2005.
- [8] David J Myers and Eric P Fillman. Dermatofibroma. 2017.
- [9] Thomas J Zuber. Punch biopsy of the skin. *American family physician*, 65(6):1155, 2002.
- [10] Kumar Abhishek and Niti Khunger. Complications of skin biopsy. *Journal of cutaneous and aesthetic surgery*, 8(4):239, 2015.
- [11] Yading Yuan. Automatic skin lesion segmentation with



**FIGURE 8.** Accuracy on the test dataset of our trained model



**FIGURE 9.** Loss on the test dataset of our trained model



**FIGURE 10.** Roc Curve of the ensemble model on test data

fully convolutional-deconvolutional networks. arXiv preprint arXiv:1703.05165, 2017.

- [12] Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19:221–248, 2017.
- [13] Muhammad Imran Razzak, Saeeda Naz, and Ahmad Zaib. Deep learning for medical image processing: Overview, challenges and the future. *Classification in BioApps*, pages 323–350, 2018.
- [14] Muhammad Hammad Memon, Jian Ping Li, Amin Ul Haq, Muhammad Hunain Memon, and Wang Zhou. Breast cancer detection in the iot health environment using modified recursive feature selection. *wireless communications and mobile computing*, 2019, 2019.
- [15] J Premaladha and KS Ravichandran. Novel approaches for diagnosing

- melanoma skin lesions through supervised and deep learning algorithms. *Journal of medical systems*, 40(4):1–12, 2016.
- [16] Adria Romero Lopez, Xavier Giro-i Nieto, Jack Burdick, and Oge Marques. Skin lesion classification from dermoscopic images using deep learning techniques. In *2017 13th IASTED international conference on biomedical engineering (BioMed)*, pages 49–54. IEEE, 2017.
- [17] Dharmendra Singh Rajput and Rakesh Gour. An iot framework for healthcare monitoring systems. *International Journal of Computer Science and Information Security*, 14(5):451, 2016.
- [18] Pradyumna Gokhale, Omkar Bhat, and Sagar Bhat. Introduction to iot. *International Advanced Research Journal in Science, Engineering and Technology*, 5(1):41–44, 2018.
- [19] Mary A Pagnutti, Robert E Ryan, Maxwell J Gold, Ryan Harlan, Edward Leggett, James F Pagnutti, et al. Laying the foundation to use raspberry pi 3 v2 camera module imagery for scientific and engineering purposes. *Journal of Electronic Imaging*, 26(1):013014, 2017.
- [20] Marc Combalia, Noel CF Codella, Veronica Rotemberg, Brian Helba, Veronica Vilaplana, Ofer Reiter, Cristina Carrera, Alicia Barreiro, Allan C Halpern, Susana Puig, et al. Bcn20000: Dermoscopic lesions in the wild. *arXiv preprint arXiv:1908.02288*, 2019.
- [21] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5(1):1–9, 2018.
- [22] Andre GC Pacheco, Gustavo R Lima, Amanda S Salomão, Breno Krohling, Igor P Biral, Gabriel G de Angelo, Fábio CR Alves Jr, José GM Esgario, Alana C Simora, Pedro BC Castro, et al. Pad-ufes-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones. *Data in brief*, 32:106221, 2020.
- [23] Robert Mikio Free. Image preprocessing, January 22 2013. US Patent 8,358,813.
- [24] Cory McKay, Ichiro Fujinaga, and Philippe Depalle. jaudio: A feature extraction library. In *Proceedings of the International Conference on Music Information Retrieval*, pages 600–3, 2005.
- [25] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6. Ieee, 2017.
- [26] Ryo Asaoka, Hiroshi Murata, Aiko Iwase, and Makoto Araie. Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. *Ophthalmology*, 123(9):1974–1980, 2016.
- [27] Wei Ma and Jun Lu. An equivalence of fully connected layer and convolutional layer. *arXiv preprint arXiv:1712.01252*, 2017.
- [28] Ashnil Kumar, Jinman Kim, David Lyndon, Michael Fulham, and Dagan Feng. An ensemble of fine-tuned convolutional neural networks for medical image classification. *IEEE journal of biomedical and health informatics*, 21(1):31–40, 2016.
- [29] Peijun Du, Junshi Xia, Wei Zhang, Kun Tan, Yi Liu, and Sicong Liu. Multiple classifier system for remote sensing image classification: A review. *Sensors*, 12(4):4764–4792, 2012.
- [30] William H Beluch, Tim Genewein, Andreas Nürnberger, and Jan M Köhler. The power of ensembles for active learning in image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9368–9377, 2018.
- [31] John Hind and Marcia Stockton. System and method for managing and securing meta data using central repository, July 17 2003. US Patent App. 10/051,951.
- [32] Simon Kornblith, Jonathon Shlens, and Quoc V Le. Do better imagenet models transfer better? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2661–2671, 2019.



MD. KHAIRUL ISLAM Department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh. Md. Khairul Islam graduated from the Department of Information and Communication Technology, Islamic University, Bangladesh. His research priorities cross many fields, including Bioinformatics, Machine learning, Deep learning, and Medical Imaging.