**CHAPTER-1**

**INTRODUCTION**

Skin disease is one of the foremost common and difficult diseases for diagnosis because of its lack of awareness and ignorance. In many developing countries also people consult dermatologist for disease of the skin and prevention measures. The people are uncertain of the medicinal prescriptions provided by the dermatologist and there is no justification within this system. Importance of disease without ignoring at the primary stage is extremely important as skin plays a heavy role in protecting the shape against fungal and harmful bacterial infections. Many folks get disease of the skin through their inheritance, job, lack of nutrition, regular habitats, exposed to chemicals etc. Environmental factors also influence the existence of disease like climate, summer season, and winter season. Thus identifying disease and diagnosis at the primary stage is extremely crucial. Thus to provide feasible and efficient system and because of the emergence of smart phones, image processing based disease analysis is more remindful as this might provide promising finally ends up in less time. Utilization of camera technique, the people can provide the input and integration of image processing and machine learning techniques the respective disease of the skin is identified and diagnosis is sometimes recommended. The input analysis is performed using two staged approach to handle this problem. The first approach is that the image processing technique and second approach is that the machine learning technique to educate the model. This  
trained model is kept on training to predict different types of skin diseases. Because the characteristics and features of varied skin condition are different, the machine algorithm needs to be trained for efficient prediction.

Skin disease are mostly ignored and provided less importance at the primary stages. Some ignorance among people might cause carcinoma. Also the physician finds it difficult to identify the kind of disease of the skin and thus the stage of disease at the analysis stage. Thus making the medication prescription difficult. This concern is going to be addressed by usage of machine learning and deep learning techniques by analysing the microscope image. This proposed machine learning based approach is also an efficient tool to identify the clinical data and provide the lands up in a very brief period of some time. This approach can provide promising results by combining computer vision and machine learning techniques.

The identification of skin condition from the microscope images are provided to image processing model. Pre-processing, feature extraction are performed within the image processing stage.

**Different types of skin disorders**

1. **ACNE**

* Commonly located on the face, neck, shoulders, chest, and upper back.
* Breakouts on the skin composed of blackheads, whiteheads, pimples, or deep, painful cysts and nodules.
* May leave scars or darken the skin if untreated.

1. **COLD SORE**

* Red, painful, fluid-filled blister that appears near the mouth and lips.
* Affected area will often tingle or burn before the sore is visible.
* Outbreaks may also be accompanied by mild, flu-like symptoms such as low fever, body aches, and swollen lymph nodes.

1. **BLISTER**

* Characterized by watery, clear, fluid-filled area on the skin.
* May be smaller than 1 cm (vesicle) or larger than 1 cm (bulla) and occur alone or in groups.
* Can be found anywhere on the body.

1. **HIVES**

* Itchy, raised welts that occur after exposure to an allergen
* Red, warm, and mildly painful to the touch
* Can be small, round, and ring-shaped or large and randomly shaped.

1. **ACTENIC KERTOSIS**

* Typically less than 2 cm, or about the size of a pencil eraser
* Thick, scaly, or crusty skin patch.
* Appears on parts of the body that receive a lot of sun exposure (hands, arms, face, scalp, and neck).
* Usually pink in colour but can have a brown, tan, or gray base.

1. **ROSACEA**

* Chronic skin disease that goes through cycles of fading and relapse.
* Relapses may be triggered by spicy foods, alcoholic beverages, sunlight, stress, and the intestinal bacteria *Helicobacter pylori.*
* There are four subtypes of rosacea encompassing a wide variety of symptoms.
* Common symptoms include facial flushing, raised, red bumps, facial redness, skin dryness, and skin sensitivity.

1. **CARBUNCLE**

* Red, painful, and irritated lump under your skin.
* May be accompanied by fever, body aches, and fatigue.
* Can cause skin crustiness or oozing.

1. **LATEX ALLERGY**

* Rash may occur within minutes to hours after exposure to a latex product.
* Warm, itchy, red wheals at the site of contact that may take on a dry, crusted appearance with repeated exposure to latex.
* Airborne latex particles may cause cough, runny nose, sneezing, and itchy, watery eyes.
* A severe allergy to latex can cause swelling and difficulty breathing.

1. **ECZEMA**

* Yellow or white scaly patches that flake off.
* Affected areas may be red, itchy, greasy, or oily.
* Hair loss may occur in the area with the rash.

1. **PSORIASIS**

* Scaly, silvery, sharply defined skin patches.
* Commonly located on the scalp, elbows, knees, and lower back.
* May be itchy or asymptomatic.

1. **RINGWORM**

* Circular-shaped scaly rashes with raised border.
* Skin in the middle of the ring appears clear and healthy, and the edges of the ring may spread outward.
* Itchy
  1. **MOTIVATION:**

Skin diseases include all conditions that irritate, clog or damage your skin, as well as skin cancer. You may inherit a skin condition or develop a skin disease. Many skin disease cause itchiness, dry skin or rashes. Often, you can manage these symptoms with medication, proper skin care and lifestyle changes. Hence progress in the development of deep learning algorithms that would quickly and accurately identify skin damage on diagnostic tests can contribute to the prevention of skin damage.

* 1. **PROBLEM STATEMENT**

Automated detection of Skin disease using deep learning technique and provide Explain-ability techniques to help us better understands our model's predictions, and how we could further improve its performance.

1.3 **OBJECTIVES AND SCOPE OF THE PROJECT**

The main objective of the system is to detect any presence of Skin disease in the colour image. But to detect disease and giving an explanation effectively to the model is the key. Hence some of the objectives that help in effective detection are discussed in this section.

**1.3.1 OBJECTIVES**

* 1. To demonstrate the use of machine learning model on color images for the recognition of Skin disease.
  2. To assess the role of XAI based automated model for detection of Skin disease by color images.
     1. **SCOPE OF THE PROJECT**

Proposed system can be used as an application so that it can recognize the images and prevent Skin disease. Skin disease detection can help doctors to detect easily based on the models which give accurate results and also in understanding the cause of Skin disease detection. The system could bring improvements in Dermatology, such that the dermatologist can detect Skin disease before a Skin specialist can perform the skin examination for Skin disease diagnosis which can be expensive and time consuming. The system can also process different formats of colour images i.e., .tif, .jpeg or .png formats.

**CHAPTER 2**

**REQUIREMENT ANALYSIS**

**2.1 FUNCTIONAL REQUIREMENTS**

The system’s primary objective, as discussed, is to detect the presence of Skin disease in the colour image. The section describes what the system should do to detect skin disease.

System

• System shall be able to pre-process images as required by models.

• System shall be able to extract textural features from the colour images.

**2.2 NON FUNCTIONAL REQUIREMENTS**

The non-functional requirements describe mainly the performance of the system, quantifying them.

• The system should be able to grade a new input image.

• The input image should be a colour image belonging to one of the two classes.

**2.3 HARDWARE AND SOFTWARE REQUIREMENTS**

In addition to the functional and non-functional requirements as discussed in sections 2.1 and 2.2 respectively, below are a few hardware and software requirements of the project.

* A machine with significant RAM and GPU to process input and run the models.
* The system utilizes the feature sub-module in the ImageNet.
* An implementation of SHAP and GRAD-CAM, deep learning models.

**CHAPTER 3**

**SYSTEM DESIGN**

In this chapter, the suitable architectural framework for the skin disease detection and the further design of the system is discussed.

**3.1 ARCHITECTURAL FRAMEWORK**

The architecture framework of the Skin disease detection system can be that of a segmentation, in which the colour images (data) go through the Resnet-50 a deep learning module represented as filters for SHAP model and Resnet -18 for Grad-cam model.

On classification, the performance of the model is to be tested and validated. If suitable results are not obtained, feature extraction is revisited to identify the significant features to give the result.

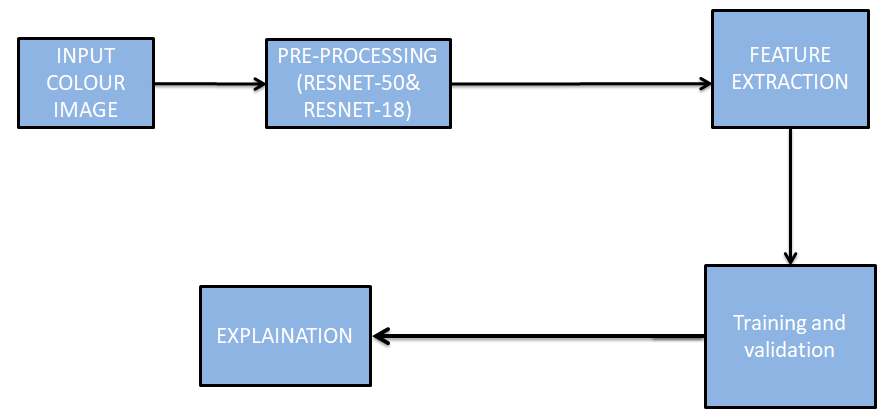


Fig 3.1: Structure of the model

The above fig 3.1 explains that the Resnet-50 and Resnet-18 models will be given the input colour images for pre-processing the data where the image will be segmented and feature extraction will be done later on, the classification will be made which will be trained and tested and gives the result.

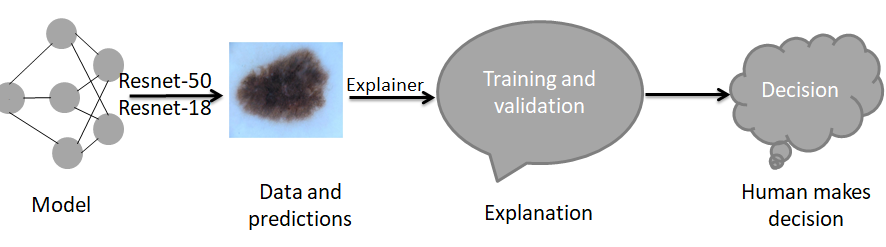


Fig 3.2: Architecture of the model

For image detecting, we are using a pre-trained model which is Resnet-50 and Resnet-18. SHAP and GRAD-CAM model will be used to explain the models prediction for better understanding of the Resnet-50 and Resnet-18 models we have used for training and testing the data set. SHAP and GRAD-CAM aims to attribute a model’s prediction to human-understandable features. In order to do this, we need to run the explanation model on a diverse but representative set of instances to return a non-redundant explanation set that is a global representation of the model.

**CHAPTER 4**

**IMPLEMENTATION**

In this chapter, the resnet50 and renet18 model will be discussed and SHAP and Grad-cam explain ability technique will be used for better predictions of the Image dataset.

**4.1 FEATURE EXTRACTION**

Texture is used in many computer vision systems as a key element. Texture is defined as a measure of coarseness, contrast, directionality, like-likeness, regularity, and roughness. The texture can also be seen as a similarity grouping in an image or as natural scenes containing semi-repetitive arrangements of pixels. In features include colour, texture and shapes in the image.

**4.2 RESNET-150 and RESNET-18**

As a part of our analysis, we examined ResNet50 and ResNet-18 model, possessed the highest prediction accuracy for the examined dataset. In this the deep learning model is used for Skin disease detection is ResNet-50 and ResNet-18. It is a convolutional neural network that is 50 layers and 18 layers deep respectively and the models used here has 3 convolution layers and 1 max-pooling layer in combination, a fully connected layer has 2 layers and the Relu activation has been used after convolution.

The models will be given the input colour images for pre-processing where the image will be resized where the image data will be trained with the pretrained ResNet-50 and ResNet-18 models on ImageNet. The training and validation of a dataset will be finely tuned on both the models. After training the dataset colour images will be tested.

|  |  |
| --- | --- |
| **HYPER PARAMETERS** | **VALUES** |
| Target size | 224,224 |
| Batch size | 32 |
| Epochs | 15 |

Table 4.1: Hyper parameters of ResNet-50 model

**4.3 EXPLAINABLE AI**

Explainable AI is an emerging field in machine learning that aims to address how black box decisions of AI systems are made. This area inspects and tries to understand the steps and models involved in making decisions. Little visibility and knowledge on how AI systems make the decisions they do. The lack of explainability and trust hampers our ability to fully trust AI systems. One way to gain explainability in AI system is to use machine learning algorithms that are inherently explainable. Simpler forms of machine learning algorithms will have certain amounts of traceability and transparency.

**Prediction**

**MODEL**

**DATA**

EXPLANATION

**SHAP Grad-CAM**

Fig 4.1: Architecture used to explain the models result.

**SHAP**

Shapely Additive Explanations (SHAP) is a collective strategy to understand or to clarify the outcome of any machine learning or deep learning standard. SHAP provides local model interpretability. SHAP modifies a single data sample by tweaking the feature values and observes the resulting impact on the output. Where the dataset will be given to the deep learning model (ResNet50) which will give the predicted output. The same dataset will be given to the SHAP model which will give the explanation for the predicted output given by the deep learning model with Shapely values.

We have used some model parameters, i.e. image size as 8,8, kernel size as 4 and batch-size as 10. In this model expected gradient combines ideas from Integrated Gradients, SHAP into a single expected value equation. This allows an entire dataset to be used as the background distribution and allows local smoothing. If we approximate the model with linear function between each background data sample and the current input to be explained, and we assume the input feature are independent then expected gradients will compute approximate SHAP values.

|  |  |
| --- | --- |
| **HYPER PARAMETERS** | **VALUES** |
| Image size | 8,8 |
| Kernel size | 4 |
| Batch-size | 10 |

Table 4.2: Hyper parameters of SHAP model

**GRAD-CAM**

Gradient-weighted Class Activation Mapping (Grad-CAM) uses the gradients of any target concept flowing into the final convolutional layer to produce a coarse localization map highlighting the important regions in the image for predicting the concept. Grad-CAM is a strict generalization of the Class Activation Mapping. Unlike CAM, Grad-CAM requires no re-training and is broadly applicable to any CNN-based architecture.

Grad-CAM is a popular technique for visualizing where a convolutional neural network model is looking. Grad-CAM is class-specific, meaning it can produce a separate visualization for every class present in the image. GRAD-CAM can be used for weakly-supervised localization, i.e. determining the location of particular objects using a model that was trained only on whole-image labels rather than explicit location annotations. GRAD-CAM can also be used for weakly –supervised segmentation, in which the model predicts all of the pixels that belong to particular objects, without requiring pixel-level labels for training. GRAD-CAM can be used to gain better understanding of a model, for example by providing insight into model failure modes.

The basic idea behind Grad-CAM is to exploit the spatial information that is preserved through convolutional layers, in order to understand which parts of an input were important for a classification decision. Grad-CAM uses the feature maps produced by the last convolutional layer of CNN and we can expect the last convolutional layers to have the best compromise between high-levels semantics and detailed spatial information.

We have used some model parameters, i.e. image size as 224,224, and torch size as 7, 7, and layer 2, layer 3, layer 4 are used to retrieve the CAM from several layers at the same time.

|  |  |
| --- | --- |
| **HYPER PARAMETERS** | **VALUES** |
| Image size | 224,224 |
| Torch size | 7,7 |
| Layers used | Layer 2, Layer 3, Layer 4 |

Table 4.3: Hyper parameters of Grad-CAM model

**CHAPTER 5**

**RESULTS AND DISCUSSIONS**

In this chapter, the results of the implementation methodology and the Explainable AI model will be discussed.

**5.1 DATASET DESCRIPTION**

The data utilized from an open source Kaggle dataset we use a dataset created by Skin Cancer MNIST: HAM10000 picture Archive communication system for a challenge based on a problem on Kaggle. The dataset is labeled dataset with acne and some skin allergies images rated with expert opinion on images.

Some sample images from the dataset can be seen in figure 5.1. The entire dataset itself is divided into training and validation datasets.

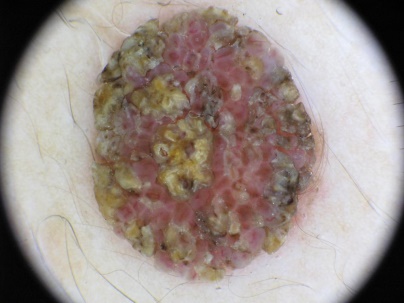
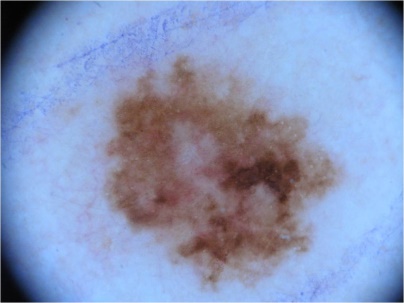
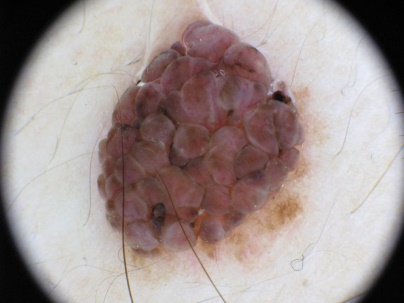


Fig 5.1: Sample images from the dataset

**5.2 CLASSIFICATION RESULTS**

In this section results of Explainability technique to help us better understand our model's predictions, and how we could further improve its performance will be discussed. The basic idea is to understand why a machine learning model predicts that an instance (image) belongs to a certain class.

**5.2.1 ResNet-50 and ResNet-18**

The ResNet-50 model consists of 5 stages each with a convolution and Identity block. Each convolution block has 3 convolution layers and each identity block also has 3 convolution layers. The ResNet-50 has over 23 million trainable parameters.

ResNet-18 is a convolutional neural network that is 18 layers deep. You can load a pretrained version of the network trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224.

The dataset is split into training and validation subsets with 70-30 split ratio. The model is then trained and validated for 15 epochs with batch size of 10. With ’Adam’ optimizer, the training accuracy was found to be 99.41% and validation accuracy was 95.86% and the training and validation loss is found to be 0.0443 and 0.1131 respectively.

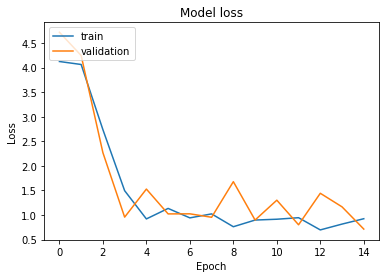
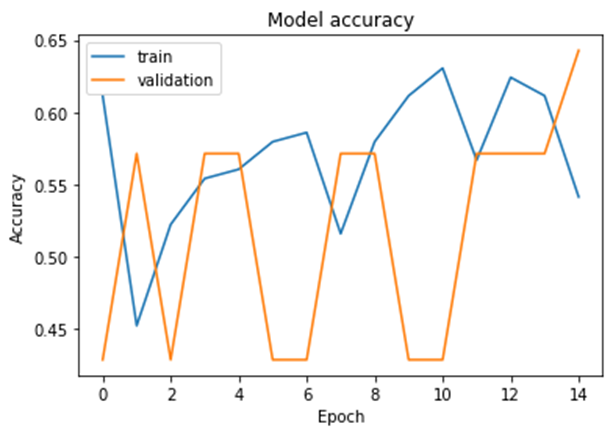


Fig 5.2: Training Vs. Validation Accuracy Fig 5.3: Training Vs. Validation Loss

|  |  |
| --- | --- |
| Sensitivity | 0.90983 |
| Specificity | 0.84922 |
| Training Accuracy = 0.9130 | Validation Accuracy = 0.8621 |
| Training Loss = 0.2250 | Validation Loss = 0.3263 |

Table 5.1: Results of ResNet-50 and ResNet-18 model

**5.2.2 EXPLAINABLE AI**

**SHAP**

XAI is a domain in which the techniques are developed and designed to explain the predictions by ML/DL systems. XAI task is to make AI, ML and DL more understanding to human users. An implementation of expected gradients to approximate SHAP values for deep learning models. It is based on connection between SHAP and the integrated Gradient algorithm. Deep SHAP is a high-speed approximation algorithm for SHAP values in deep learning models that builds on a connection with DeepLIFT described in the SHAP. The implementation here differs from the original DeepLIFT by using a distribution of background sample instead of a single reference value, and using Shapley equations to linearize components such as max, softmax, products, division, etc. note that some of these enhancements have also been since integrated into DeepLIFT. Tensor flow models and keras models using the tensor flow backend are supported.

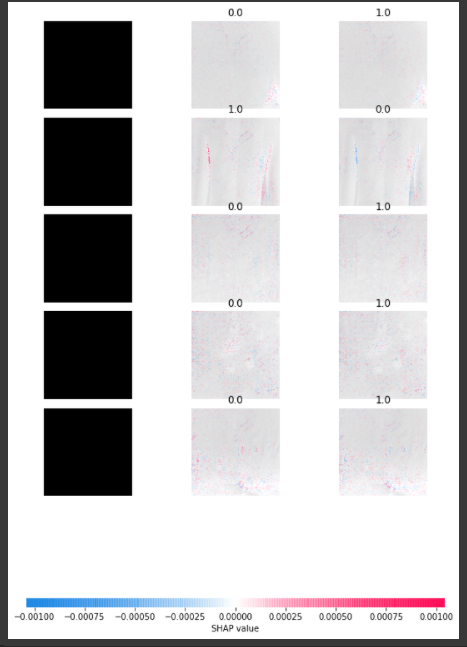


Fig 5.4 Feature attribution of predictions of the model

* The plot above explains outputs (levels of glaucoma) for three different images. Positive SHAP value increases the model's output while negative SHAP value decreases the output. The input images are shown on the left (they are black because most of the pixels are greater than 0), and as nearly transparent grayscale backings behind each of the explanations. The sum of the SHAP values equals the difference between the expected model output (averaged over the background dataset) and the current model output.
* Note that for the images that the label is greater than Zero that is positive SHAP value.
* Labels that have higher positive SHAP values as the correct one are labels that our model probably doesn't have a high confidence prediction.

**GRAD-CAM**

GRAD-CAM is a form of post-hoc attention, meaning that it is a method for producing heat maps that is applied to an already-trained neural network after training is complete and the parameters are fixed. This is distinct from trainable attention, which involves learning how to produce attention maps (heat maps) during training by learning particular parameters. For a more in-depth discussion of post-hoc vs. trainable attention.

GRAD-CAM does not require a particular CNN architecture. GRAD-CAM is a generalization of CAM (class activation mapping), a method that does require using a particular architecture.

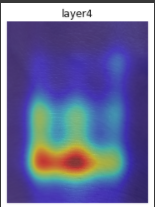
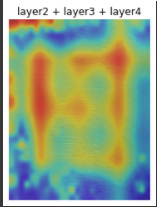
 

Fig 5.5(a): Healthy image on layer4 Fig 5.5(b): Healthy image on combination of layer 2, layer3, layer4

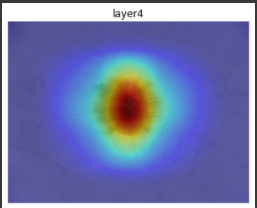
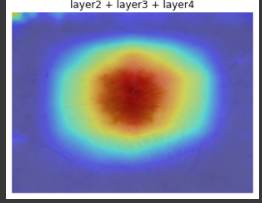
 

Fig 5.6(a): Diseased image on layer4 Fig 5.6(b): Diseased image on combination of layer2, layer3, layer4

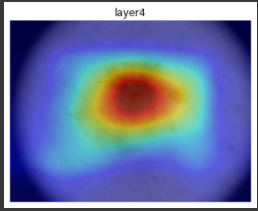
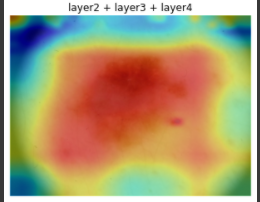
 

Fig 5.7(a): Diseased image on layer4 Fig 5.7(b): Diseased image on combination of layer2, layer3, layer4

Grad-CAM seeks to identify the parts of an input that have been important for a classification decision by utilizing the spatial information that is preserved through convolutional layers. We can expect that the final convolutional layers of CNN have the best balance between semantic relevance and spatial detail in Grad-CAM's feature maps.

By analysing the gradients of the classification score in relation to the final convolutional feature map, Grad-CAM identifies the input features most important to classification. It is precisely in these places where this gradient is largest that the final score is most dependent on the data. In gradCAM, the importance map is computed by taking the derivative of a reduction layer's output that corresponds to a convolutional feature map for a given class. GradCAM automatically selects suitable layers for computation of importance maps for classification tasks. With the name-value arguments 'Reduction Layer' and 'Feature Layer', you can also specify the layers.

**CHAPTER 6**

**CONCLUSION AND FUTURE SCOPE**

In this project, an approach for skin disease detection based on textural feature extraction was implemented. A review on Explainability techniques for skin disease detection using colour images was presented. We present potential explainable methods that, with future improvements in implementation, can be generalized to different medical data sets and can provide effective decision support for medical experts. From the viewpoint of users, our project offers deep insight into the details of explanation support and can be used as constructive feedback for the potential implementation of explainable machine learning methods in the future. Our findings suggest that there are notable differences in human decision-making between various explanation support settings showing the best decision support and performance. Additionally, in comparison to the no-explanation setting, explanation support proved to increase the number of correct decisions made by users in two out of three user studies. The presented work can thus give developers more confidence to further develop and utilize explainable methods, which, in turn, will instill users with more confidence and trust. The model was able to successfully extract feature from the dataset. The results obtained with deep learning based Resnet-50 model are of 91.62%. The output of SHAP and GRAD CAM is a list of explanations, reflecting the contribution of each feature to the prediction of a data sample. On Comparing the SHAP and GRAD CAM results the GradCAM gives the better explanation than SHAP model.

In future, we will use other different models to explain the things better by using a large number of databases. The application of explainable methods to other medical data sets, as well as further testing and improvements in the context of providing decision support to medical professionals and automating diagnostic procedures, may lead to solutions that are more broadly applicable.