PAPER NAME

Group 4 PR Report V1.4.docx

WORD COUNT	CHARACTER COUNT
912 Words	**5298 Characters**
PAGE COUNT	FILE SIZE
10 Pages	**415.2KB**
SUBMISSION DATE	REPORT DATE
Apr 30, 2023 9:30 PM GMT+5:30	**Apr 30, 2023 9:31 PM GMT+5:30**

5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 0% Internet database
- 1% Publications database
- Crossref database
- Crossref Posted Content database
- 5% Submitted Works database

Contents

SKIN DISEASE DETECTION USING DEEP LEARNING

Abstract

Skin disease [1] is one of the most prevalent types of diseases worldwide. However, their diagnosis can be challenging and requires a high level of expertise. Dermatologists often conduct comprehensive tests[2] to determine the specific skin condition and the duration of the illness, which can vary depending on the practitioner's experience and the patient's individual case. To overcome these limitations, it is crucial to adopt a technique that is not restricted by these factors and can accurately diagnose skin diseases without any limitations. In this implementation a pre-trained ResNet50V2 has been used for classification of images into malignant and benign.

1. Introduction

Skin disease classification using deep learning attempts to provide precise and effective ways for diagnosing skin problems. Convolutional neural networks (CNNs)[3] and other deep learning algorithms have demonstrated promising results in reliably recognizing a variety of skin illnesses from image files.

The classification process includes training deep learning model on a large dataset of images that are labeled with their corresponding skin diseases. The CNN model extracts features from the images and uses them to identify patterns and features that are indicative of specific skin diseases. The model then assigns a probability score to each disease, indicating the likelihood of that disease being present in the image.

There are several challenges in skin disease classification using deep learning, including the variability in skin color, lighting, and imaging conditions. However, researchers have been able to address these challenges using data augmentation technique and transfer learning.

Overall, skin disease classification using deep learning has the potential to improve the accuracy and speed of diagnosis in skin disease, leading to efficient healthcare systems.

2. Literature

This implementation of Skin Disease Detection[1] is motivated from the eponymous paper published by Syed Inthiyaz et al in the Elsevier journal of Advances in Engineering Software in Nov 22.

3. Dataset

Dataset for this implementation has been obtained from Kaggle[7]. Dataset consists of images labeled as 'malignant' and 'benign'. Dataset was pre-divided into three folders as train, validation and test. Train data contained 9001 images, while the remaining two contained 1100 images in each folder. Originally the Train data was biased for benign class with around 56% of total distribution.

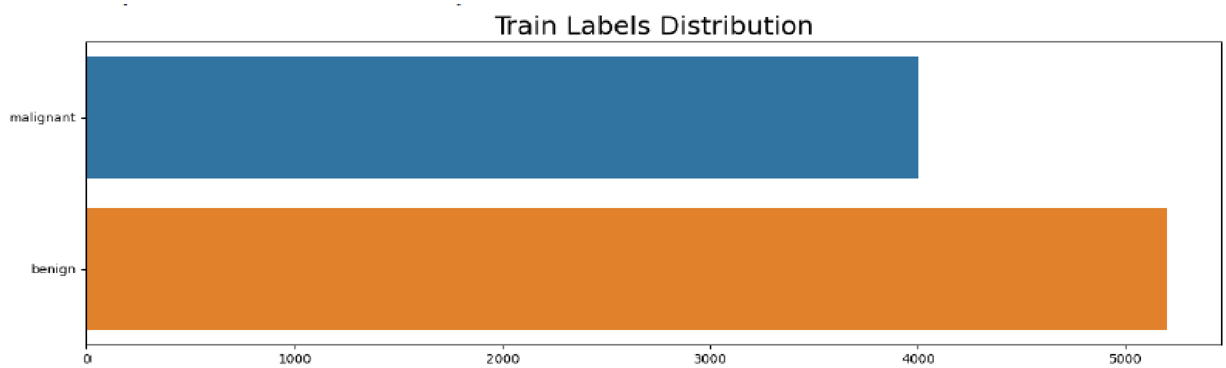


Fig. 1. Initial Class wise data distribution

Hence, to avoid overfitting in the model we have rebalanced the training dataset by making equal distribution of both benign and malignant class in the train set. The new train set consists of 4001 images per class.

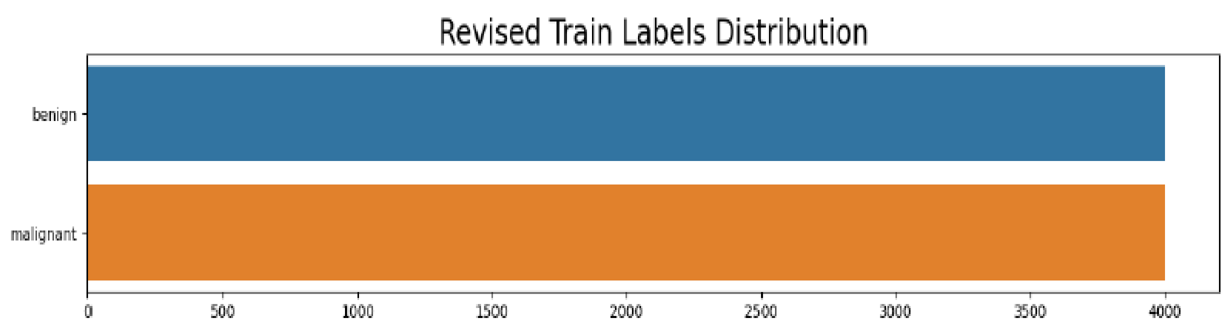


Fig. 2. Class wise data distribution after balancing the train set

4. Model

We have used a pre-trained *ResNet50V2* [6] model from Tensorhub. ResNet50V2 contains 50 layers. It is used as input layers, whose output is fed to a Convolutional layers followed by a max-pooling layers.

Total number of parameters were 24,894,018, out of which 24,847,938 were trainable.

```
Model: "ResNet50_Sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 resnet50v2 (Functional)     (None, 7, 7, 2048)        23564800

 conv2d (Conv2D)             (None, 7, 7, 64)          1179712

 max_pooling2d_3 (MaxPooling  (None, 3, 3, 64)         0
 2D)

 batch_normalization (BatchN  (None, 3, 3, 64)         256
 ormalization)

 dropout (Dropout)           (None, 3, 3, 64)          0

 flatten (Flatten)           (None, 576)               0

 dense (Dense)               (None, 256)               147712

 batch_normalization_1 (Batc  (None, 256)              1024
 hNormalization)

 leaky_re_lu (LeakyReLU)     (None, 256)               0

 dropout_1 (Dropout)         (None, 256)               0

 dense_1 (Dense)             (None, 2)                 514

=================================================================
Total params: 24,894,018
Trainable params: 24,847,938
Non-trainable params: 46,080
```

Fig 3. Parameters of model

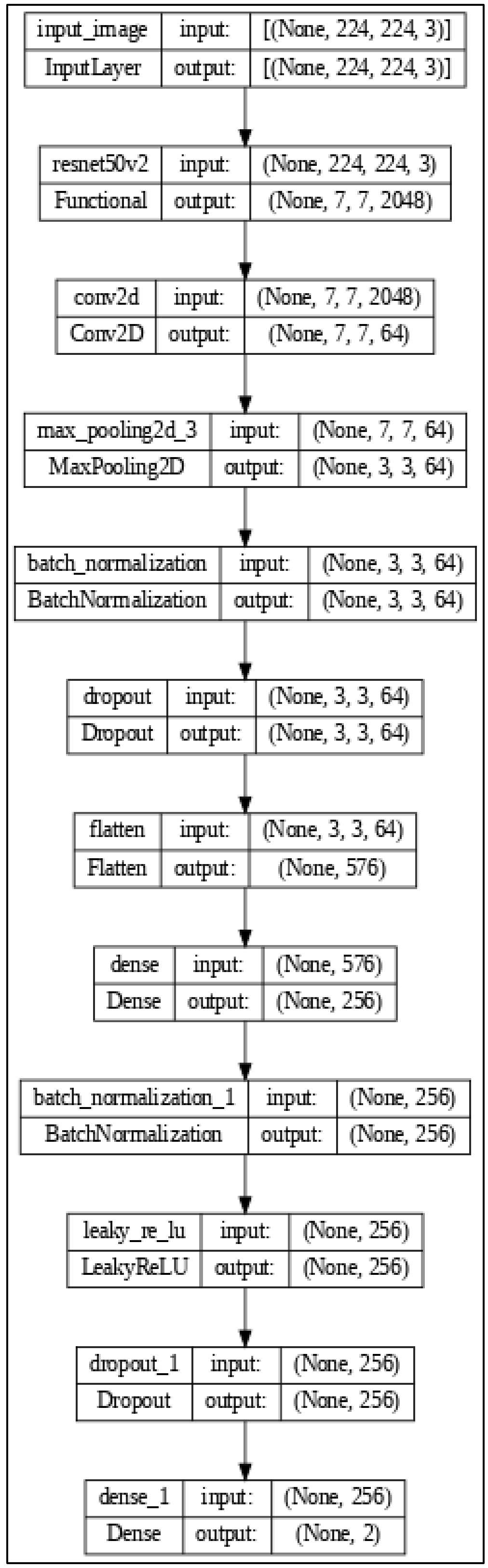


Fig 4: Block diagram of our model

5. Methodology

We have pre-processed the data which is made available to us and subsequently we carried out re-balancing of images in the two classes. We have used data augmentation[4][5] techniques like random zoom, and random rotate to increase the size of our training set images.

Subsequently we created a pipeline for chaining together multiple functions to include shuffling of images, pre-fetching, marking batch size, data augmenting and loading of images into datasets. It enabled us with modularity, scalability and simplicity in our implementations.

Image input is provided to ResNet50V2 whose output is given to a 256 unit convolution layer with ReLU activation function in our model. A binary cross entropy loss function is used to calculate total loss after each epoch.

5.1 Resnet50V2

For our implementation, we have been guided to use the ResNet50 model, which was pre-trained on 'Imagenet' datat. ResNet50V2 is a convolutional neural network (CNN) architecture introduced as an improvement to the original ResNet50 architecture.

ResNet50V2 architecture consists of 50 layers and uses skip connections, also known as residual connections, which allow the gradient to flow more easily through the network.The model uses a pre-activation approach, which applies batch normalization and activation before the convolutional layers instead of after. This helps to reduces number of parameters in model and has improved accuracy.

I.Optimizer – Adam

We have used a popular optimizer ADAM[5] (Adaptive Moment Estimation) known for its robustness and efficiency. ADAM is a stochastic gradient descent algorithm that computes individual adaptive learning rates for different parameters from estimates of first and second moments of the gradients.

It combines the advantages of AdaGrad[6] and RMSProp[5]. AdaGrad adapts the learning rate for each parameter based on the historical gradients, while RMSProp uses a moving average of the squared gradients to adjust the learning rate.

6. Result

We implemented the subjected paper "Skin disease detection using deep learning" wherein the Skin Cancer prediction was done using images using pre-trained ResNet50 model with activation functions such as ReLU. As an advancement to the proposed model in the paper, we have used pre-trained ResNet50V2 model in combination with Adam optimizer, Binary Cross entropy as loss function and activation functions such as Leaky ReLU and Sigmoid. The model has been trained for 25 epochs using batch normalization achieving train accuracy of 90% as well as test accuracy of 89%.

- **Accuracy as per original paper[1]:**

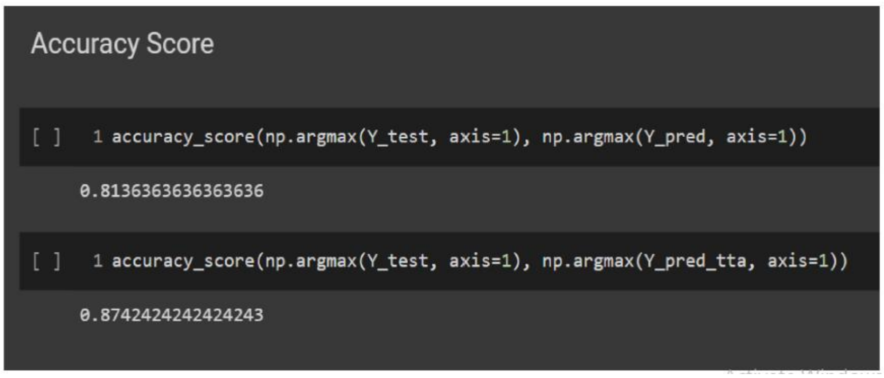
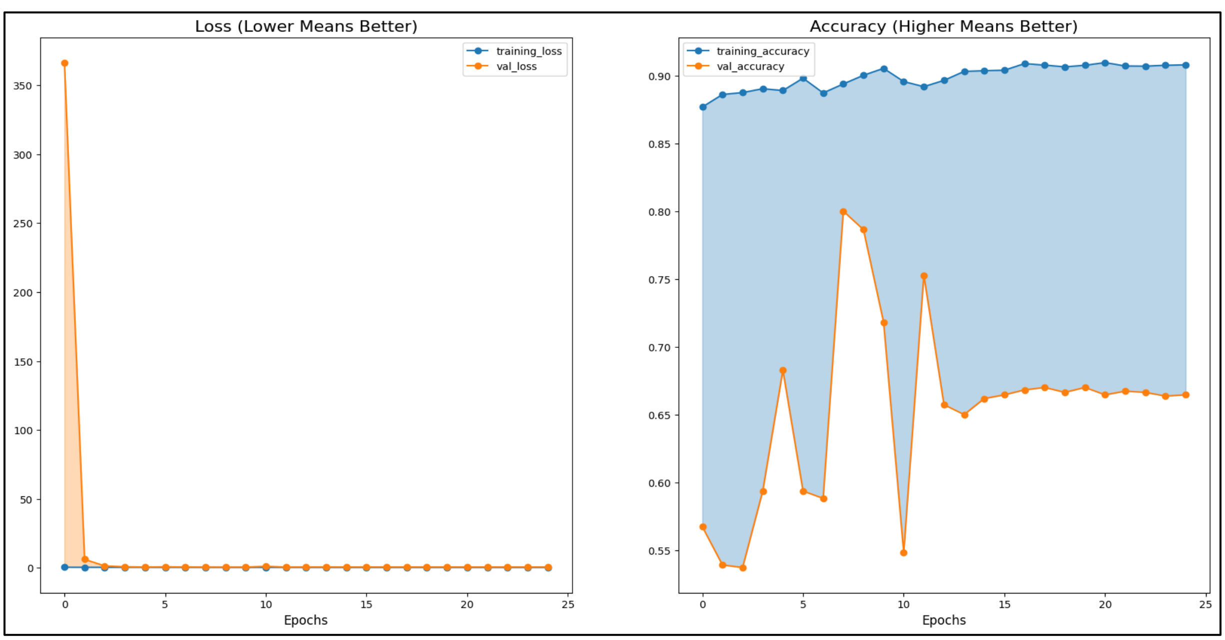


Fig 5: Accuracy of model in original paper

- **Our Model : Loss and Accuracy for Train and Validation Data**



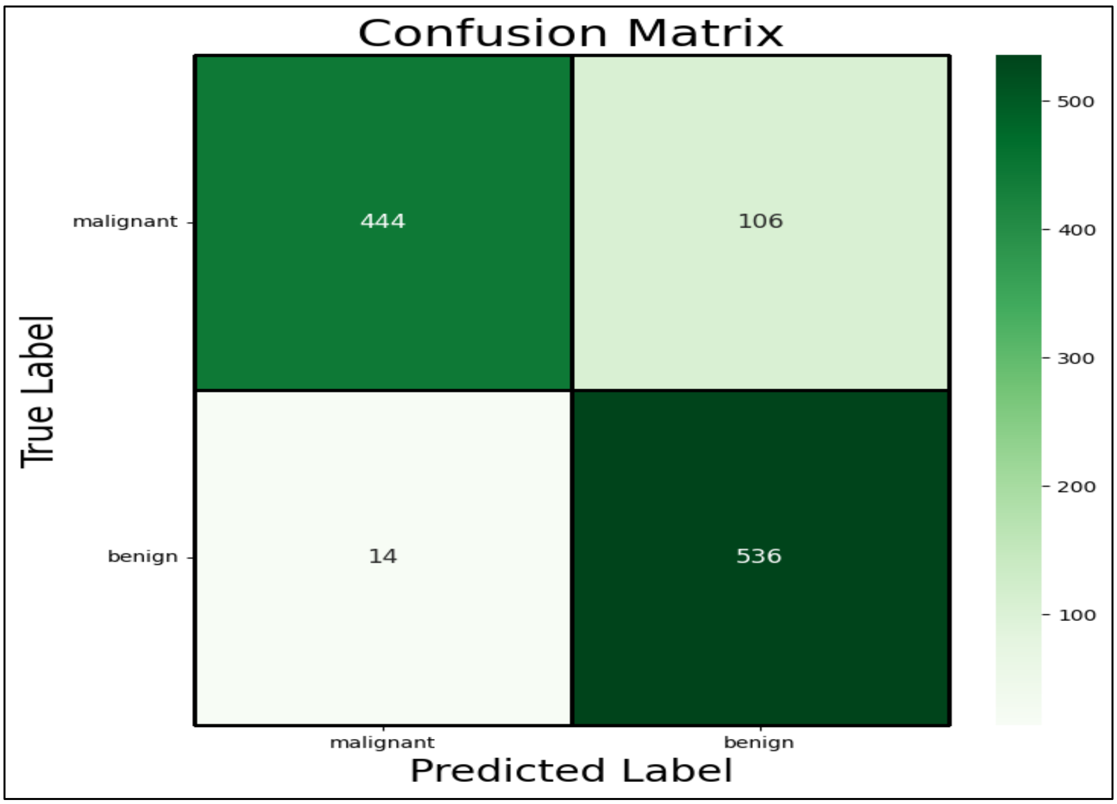
1 Fig 6: Loss and Accuracy of model for train and validation set in project

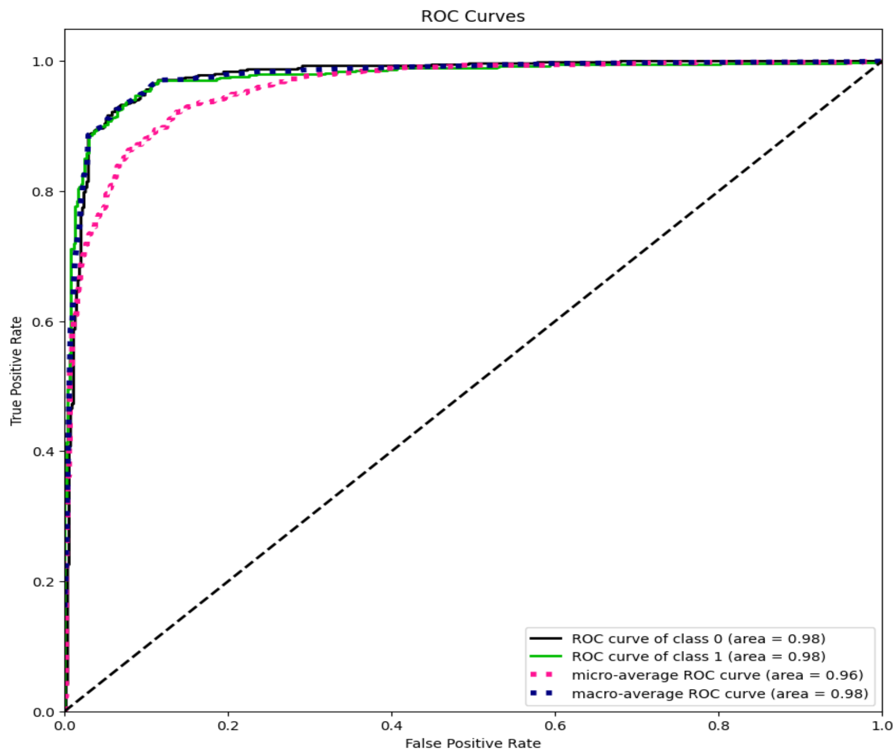
- **Accuracy for Test Dataset**

```
eval_result=model_resnet_local.evaluate(test_ds)

    35/35 [==============================] - 463s 13s/step - loss: 0.4120 - accuracy: 0.8955
```

- **Confusion Matrix for Test Dataset**





- **Performance metrics for Test Data**

```
=============================================

Performance Metrics:

=============================================
accuracy_score:          0.8909

_____________________________________________
precision_score:         0.9022

_____________________________________________
recall_score:            0.8909

_____________________________________________
f1_score:                0.8901

_____________________________________________
matthews_corrcoef:       0.7930

=============================================
```

Fig 9: Performance matrix for test set

Learning Outcome

- Familiarity with Google Colab
- Learnt use of Tensorflow for

- Handling image data
- Creation of pipeline
- Data augmentations
- Model building
- Model testing
- Analysis of results

7. References

1. Inthiyaz, Syed, et al. "Skin disease detection using deep learning." *Advances in Engineering Software* 175 (2023): 103361.
2. Rathod, Jainesh, et al. "Diagnosis of skin diseases using Convolutional Neural Networks." *2018 second international conference on electronics, communication and aerospace technology (ICECA)*. IEEE, 2018.
3. Wu, Z. H. E., et al. "Studies on different CNN algorithms for face skin disease classification based on clinical images." *IEEE Access* 7 (2019): 66505-66511.
4. Shanthi, T., R. S. Sabeenian, and R. Anand. "Automatic diagnosis of skin diseases using convolution neural network." *Microprocessors and Microsystems* 76 (2020): 103074.
5. Gu, Jiuxiang, et al. "Recent advances in convolutional neural networks." *Pattern recognition* 77 (2018): 354-377.
6. Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a convolutional neural network." *2017 international conference on engineering and technology (ICET)*. Ieee, 2017.
7. https://www.kaggle.com/datasets/sallyibrahim/skin-cancer-isic-2019-2020-malignant-orbenign

5% Overall Similarity

Top sources found in the following databases:

- 0% Internet database
- 1% Publications database
- Crossref database
- Crossref Posted Content database
- 5% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Bangladesh University of Professionals on 2021-10-27 Submitted works	2%
2	Tuan Nguyen, L. C. Ngoc, Tung Nguyen Son, Duc Ha Minh, T. Ha Phuon... Crossref	1%
3	Aberystwyth University on 2020-05-11 Submitted works	<1%
4	The Robert Gordon University on 2023-04-27 Submitted works	<1%
5	University of Lincoln on 2020-04-23 Submitted works	<1%