An Interpretable AI Enabled Skin Cancer Detection Model using Convolutional Neural Networks
COMPUTER SCIENCE & ENGINEERING, SCHOOL OF TECHNOLOGY, PDEU.
TEAM MEMBERS : DHYEY SHAH, KRISHNA PATEL
## MENTOR: DR. RAJEEV GUPTA

SOT
SCHOOL OF TECHNOLOGY

INTRODUCTION

Skin cancer is a major public health concern, with over one million new cases reported in India alone each year. Skin cancer identification and categorization are critical for timely treatment and improved patient outcomes. Our project aims to perform the detection and classification of Skin Cancer using different Machine Learning and Deep Learning methods and comparing the results to deduce the best process possible.
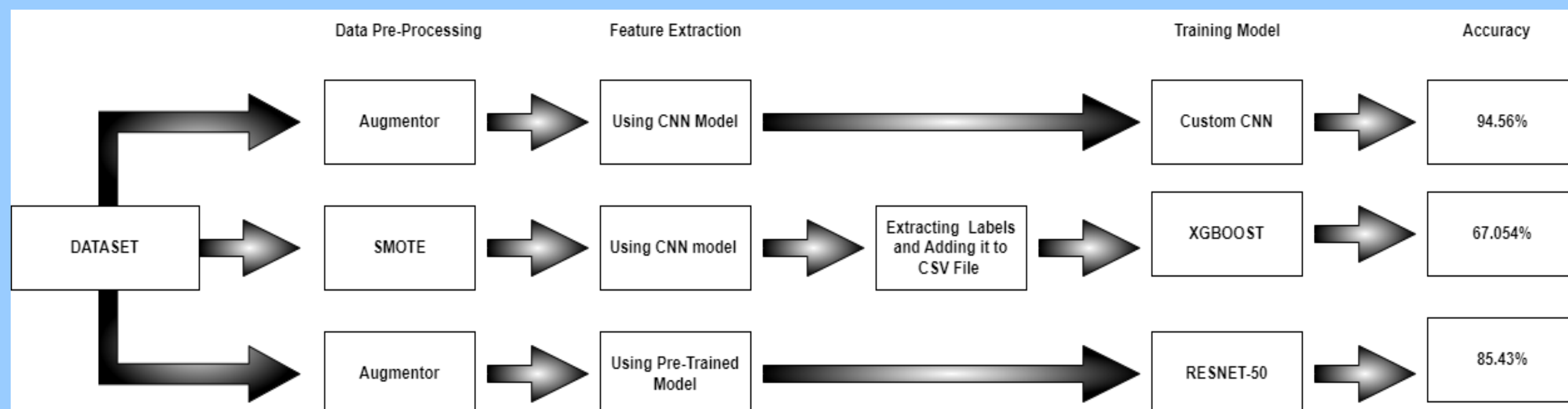
Flow Chart of the Proposed Methodology

DATASET OVERVIEW

Dataset : Skin cancer ISIC The International Skin Imaging Collaboration

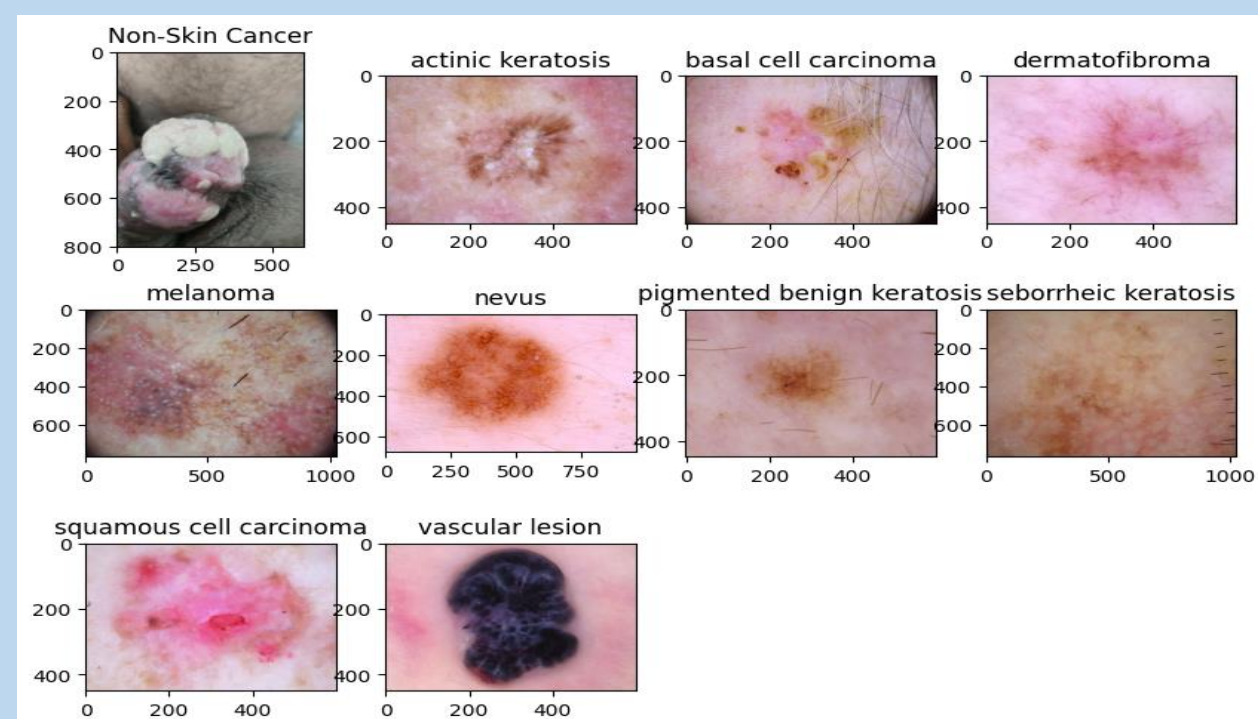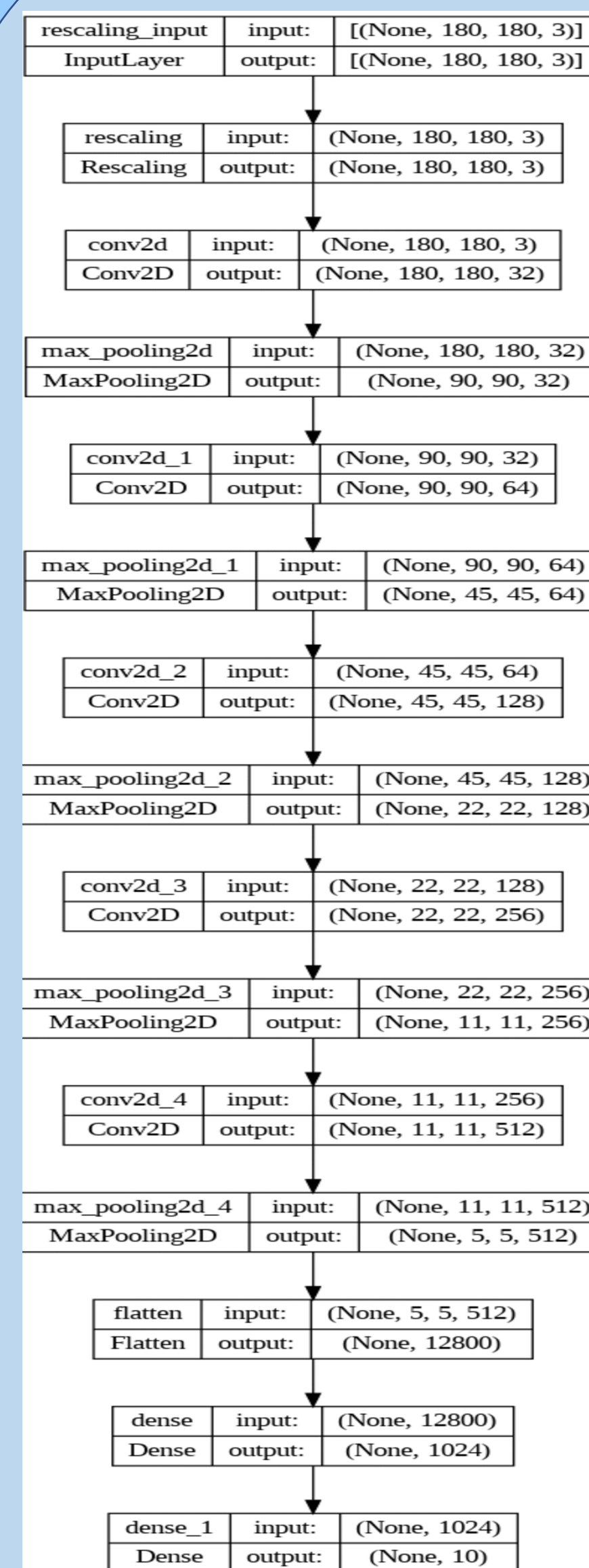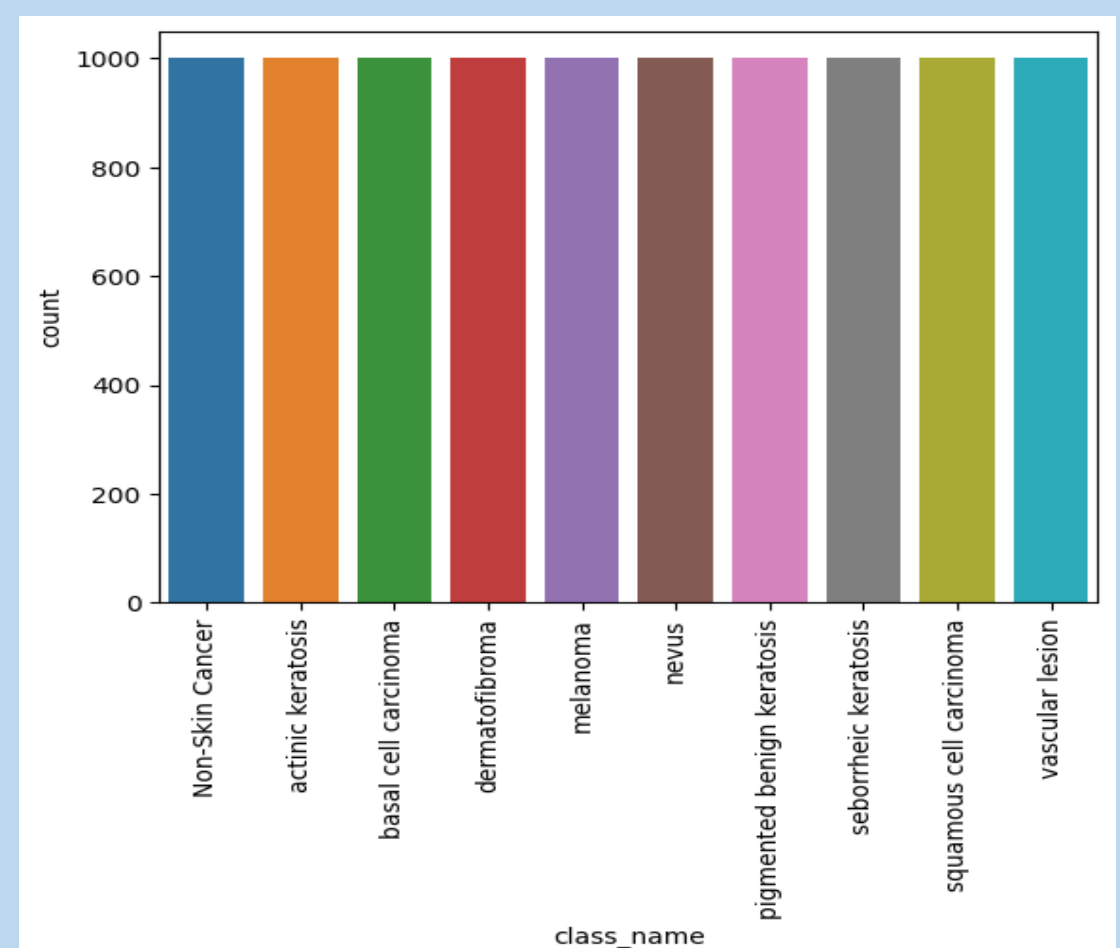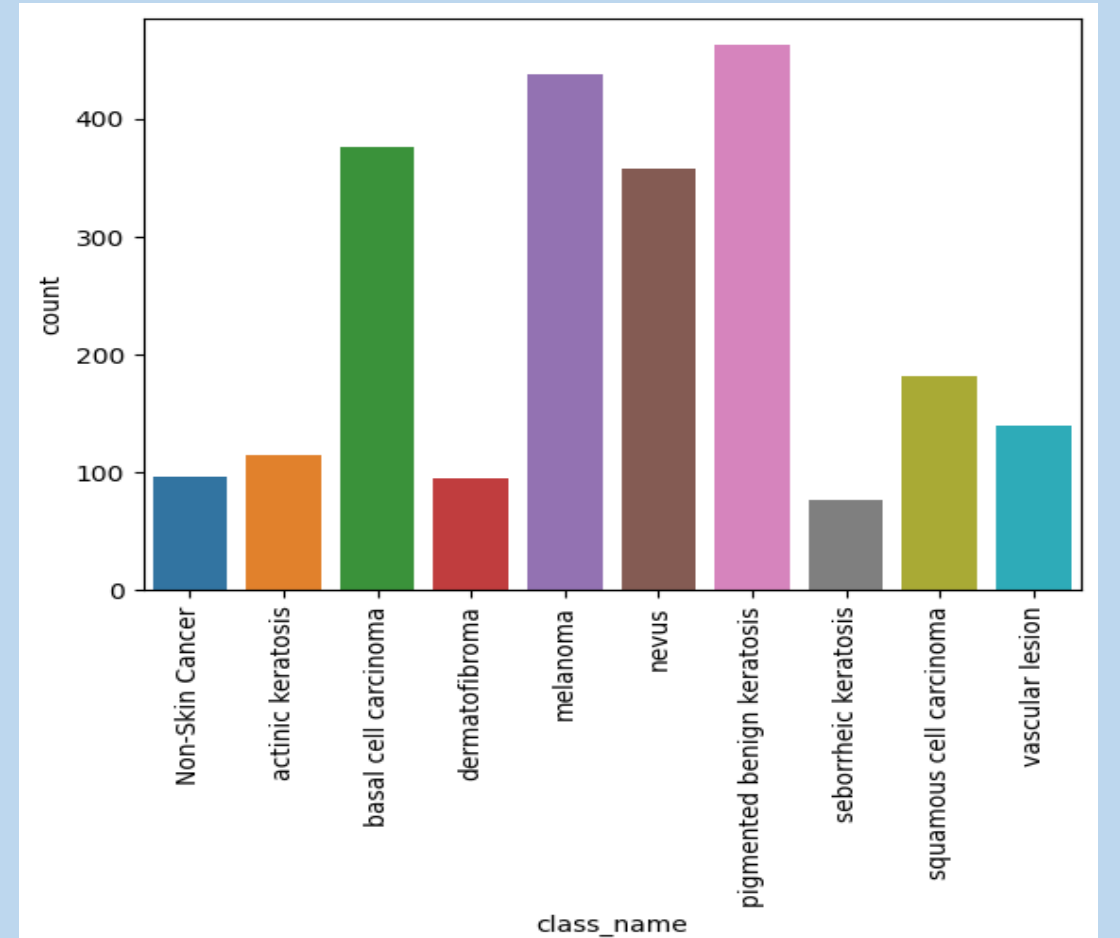Total 2415 images belonging to 10 different classes.

Figure 1: Dataset sample image for 10 different classes

Hybrid Model

In this approach a CNN model was used for the feaure extraction of the image dataset to a CSV file. A total of 1023 features were extracted from the dataset. Then a Image data generator was used to load the images and extract the labels. For handling the imbalance labels we had use SMOTE with the sampling strategy known as dictionary. Finally we applied the machine learning model XGBOOST Classifier on the dataset and acquired an accuracy of 64%.

RESNET-50 Model

ResNet50 is a Pre-Trained model which is a part of Transfer learning Keras Libraries, it allows to use of the weights from the image net directly which were calibrated for some similar problem. We have used this pre trained model for our dataset and used the Augmentor technique to balance the data.

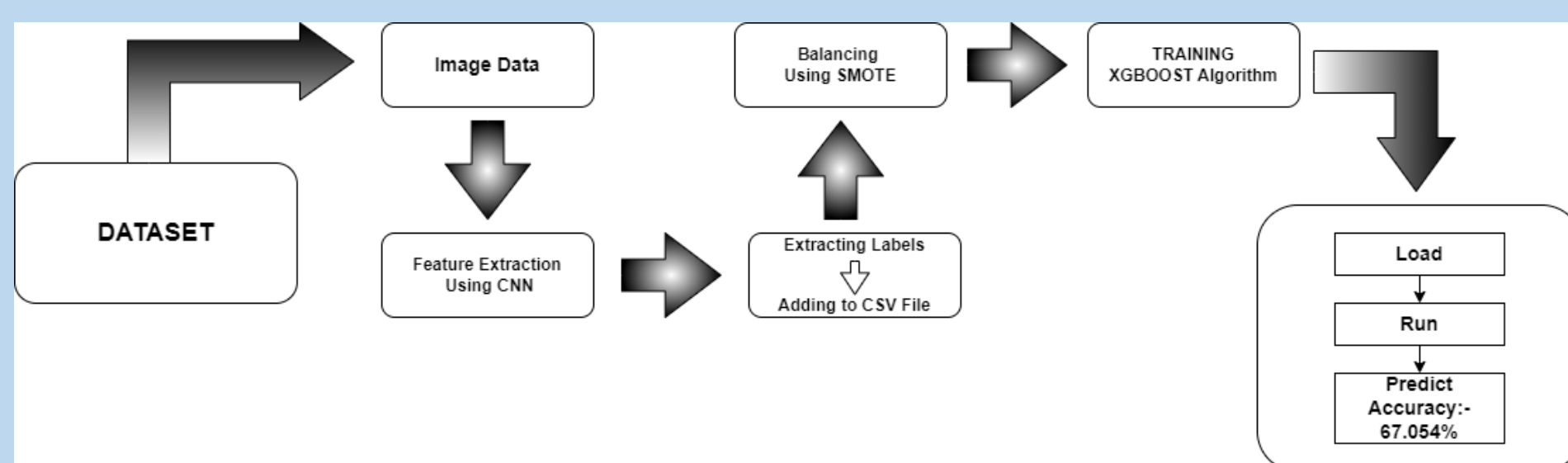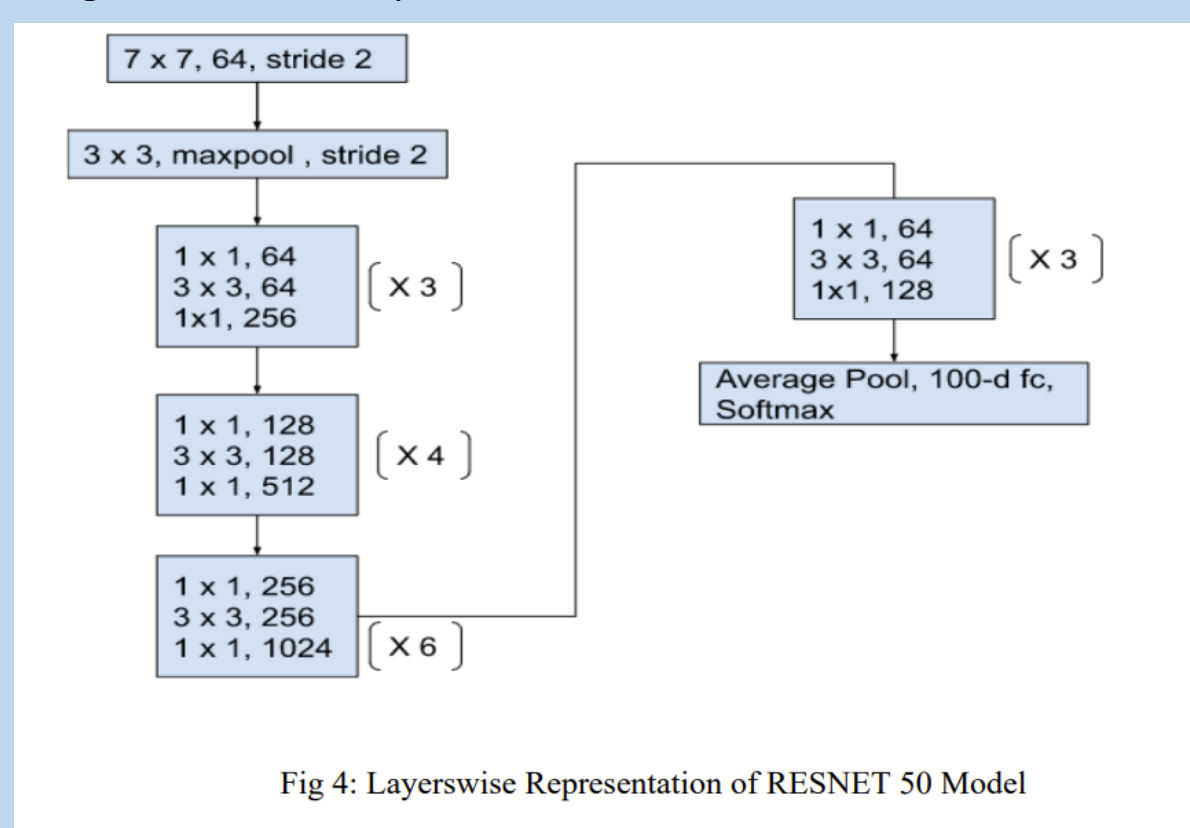Fig 4: Layerswise Representation of RESNET 50 Model

CNN Model on Balanced Dataset

We have used data augmentor to balance the dataset. The graphs of the number of images before and after the Augmentor can be seen below.

The Hyperparameters used for the CNN models are as follows:
- Activation: softmax
- Optimizer: adam
- Loss: sparse Categorical crossentropy
- Number of Epochs: 25

RESULTS & FINDING

No.	Model	Training	Validation	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1.	CNN Model with Balanced dataset	8000 images	2000 images	94.56	90.05	0.1424	0.4187
2.	CNN model without data balancing	1834 images	457 images	83.01	56.46	0.4381	2.0191
3.	Hybrid Model	2960 rows	740 rows	67.054			
4	RESNET-50 with Balanced Dataset	8000 images	2000 images	85.43	67.10	0.4167	1.2578