

AUTO-PCOS CLASSIFICATION CHALLENGE

OVERVIEW:

Polycystic Ovarian Syndrome (PCOS) is one of the most prevalent amongst the young generation, comprising of a combination of symptoms of androgen excess and ovarian dysfunction. The aetiology for this condition remains largely unknown and also absence of specific diagnostic criteria.

The aim is to develop a model that would classify healthy and unhealthy frames of ultrasound images extracted from the ultrasound videos.

DATA:

The data consists of 3200 frames of ultrasound images extracted from ultrasound videos. Each of the images were labelled into either of the two classes “Healthy” or “Unhealthy”, listed in a .xls file. This data was exclusively used for training and validation of the model.

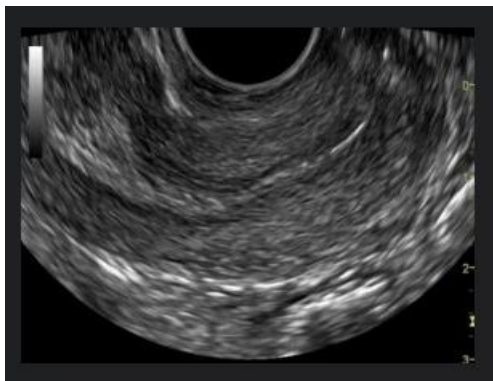


Figure 2- "Healthy"- 0



Figure 1- "Unhealthy"-1

METHOD:

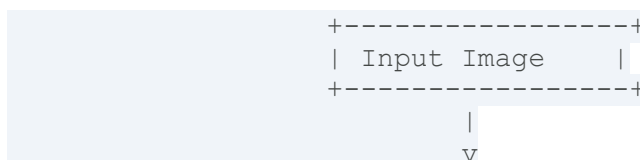
The given data was pre-processed (augmented, resized and converted to grey-scale images) and further segregated into the two classes as per the given label.

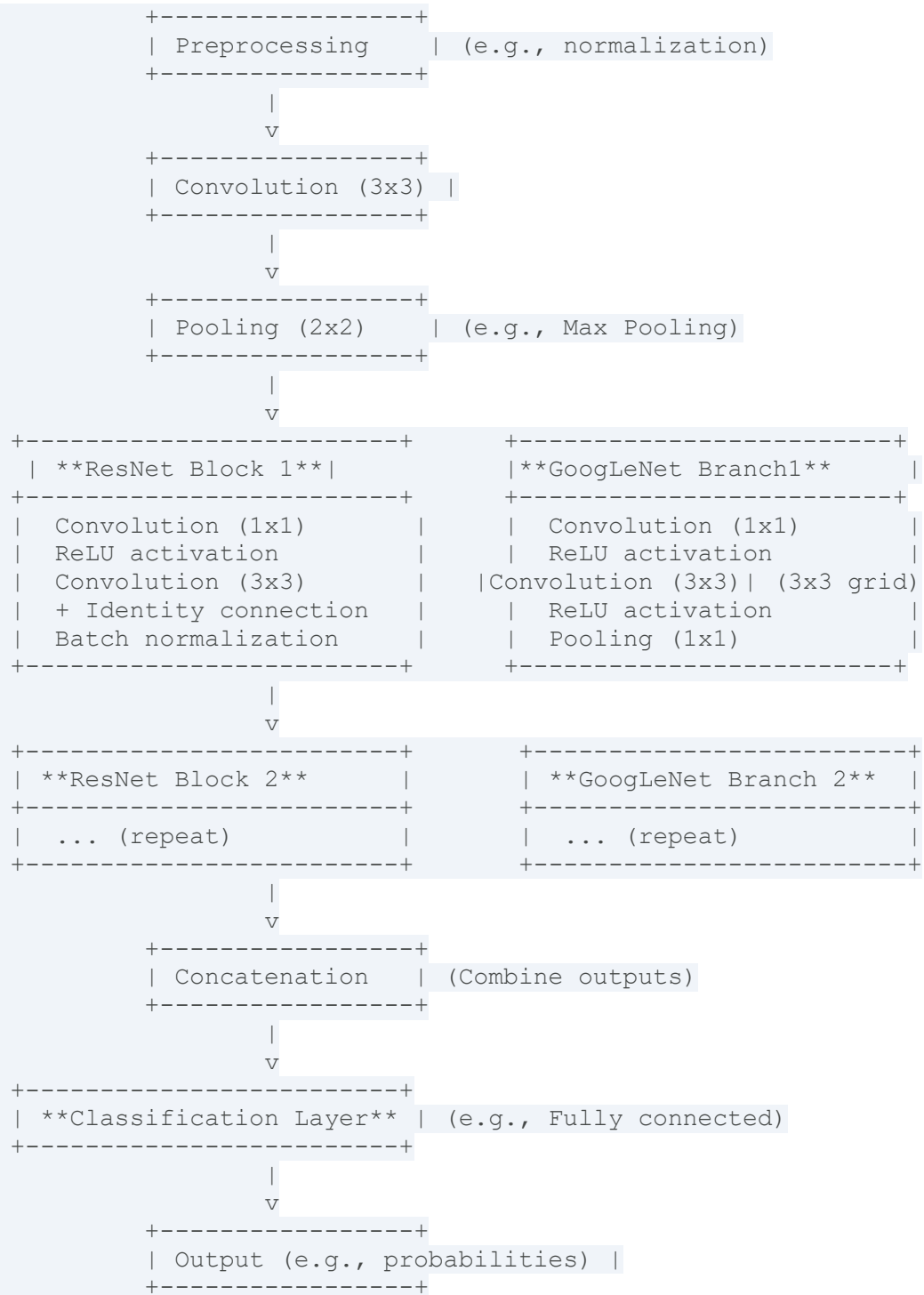
This segregated data is now used for training and validation of our model. Eighty percent of the data is used for training and rest for validation from each class as stated.

The model used in our project is a deep learning based on GoogLeNet Model and ResNet Model, which after training will take an ultrasound image frame as input and tell whether it is “Healthy” or “Un healthy”.

Regarding this required problem, we have developed a deep learning model by combining identity block of ResNet model with GoogLeNet model, aiming to leverage ResNet's skip connection for improved gradient flow, mitigating vanishing gradient issues. The addition of ResNet's identity block to GoogLeNet introduces feature diversity by preserving input features, complementing GoogLeNet's multi-scale representation. This fusion is justified by the potential synergy between ResNet's gradient flow and GoogLeNet's feature diversity, aiming for enhanced training stability and representational power.

Outline of our model implanted as follows:





TRAINING:

We are using Accuracy, Precision, Recall, F1-Score, AUC-ROC as the evaluation metrics.

Confusion Matrix:

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

A tabular representation that summarizes the performance of a classification model by showing the count of true positive, true negative, false positive, and false negative predictions. It is a valuable tool for assessing the effectiveness of a model in terms of precision, recall, and overall accuracy.

Accuracy:

Accuracy provides a general measure of a model's correctness across all classes.

Accuracy=Number of Correct Predictions/Total Number of Predictions

Accuracy = (TP+TN) / (TP+TN+FP+FN)

Precision:

Precision quantifies the accuracy of positive predictions made by a model. It is calculated using the following formula:

Precision=TP/(TP+FP)

Recall:

Recall quantifies the ability of a model to capture all the relevant positive instances. It is calculated using the following formula:

Recall= TP/(TP+FN)

F1-Score:

F1 score is a metric in classification that combines both precision and recall into a single value. It is particularly useful when there is an uneven class distribution. The F1 score is calculated using the following formula:

F1-Score= (2*Precision*Recall)/(Precision+Recall)

F1-Score=2*TP/(2*TP+FP+FN)

AUC-ROC:

Area under Receiver Operating Characteristic Curve measures the ability of the model to distinguish between positive and negative instances across various threshold values.

AUC-ROC= $\int_0^1 \text{Sensitivity (True Positive Rate)} d(1 - \text{Specificity (False Positive Rate)})$

RESULTS:

While Training the model , the following results were obtained:

Model	Accuracy (from training data)	Accuracy (from validation data)
Our developed Model	71.55	71.00