P-38 **Audio Based Psychological Analysis Using Face Recognition**

Project by: *Vallabh Bhat, Sakshi Devale, Mayur Kulam*

Project Coordinator: *Dr. Rajashri Khanai, HOD ECE*

KLE Technological University's
Dr. M. S. Sheshgiri College of Engineering and Technology, Belagavi

---

Abstract: Emotion recognition using multiple information channels has recently received significant attention. A widely accepted prediction is that computing will move to the background, weaving itself into the fabric of our everyday living and projecting the human user into the foreground. To realize this goal, next-generation computing will need to develop human-centred user interfaces that respond readily to naturally occurring, multimodal, human communication. Some examples of multimodal emotion recognition application include online learning, Entertainment, biometric systems, human computer interaction, interpersonal relation and behavior prediction. Two natural and effective ways to express emotion in human-human interaction are speech and facial expression. In this work we will try to extract the facial features of a person using a hybrid computer vision and image processing techniques. The trained result thus obtained will be used to provide assistance using a audio feedback.

Project Objectives: To realize the next-generation computing we aim to perceive, understand intentions and emotions as communicated by individual and following would be our objectives,

- To develop a system for efficient image processing using raw real time images.
- To develop a system which can detect faces from any image format.
- To develop a neural network model which can get trained at acceptable high rate of accuracy.
- To develop the trained model on hardware device.
- To have an efficient system which can recognize the trained faces.
- To develop a highly efficient system for audio responses using natural language processing.

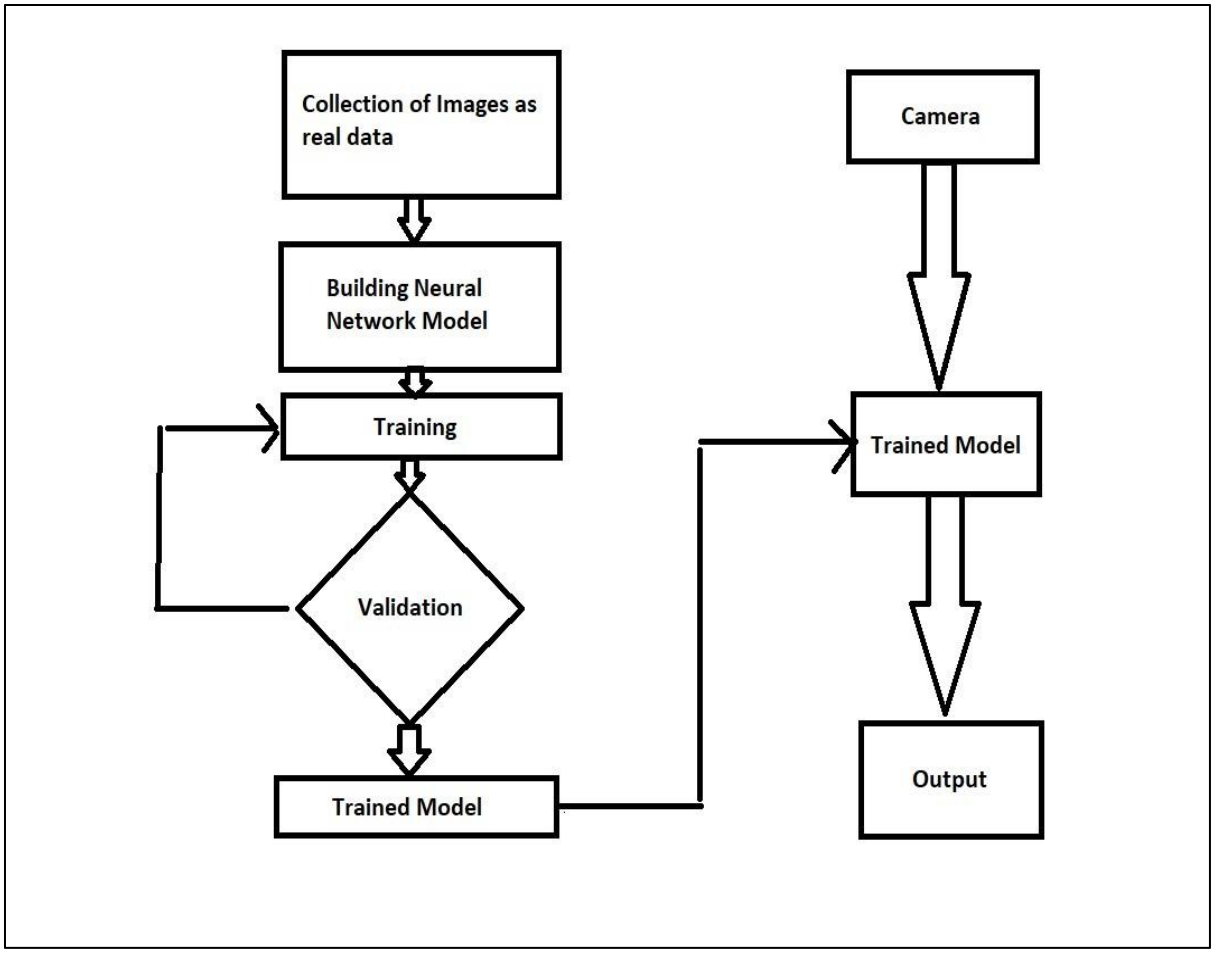
Methodology:

- ***Image Input:*** This involves capturing of real-time raw images via feeds such as webcam, CCTV cameras etc.,
- ***Image Processing:*** This image will be processed using various image processing techniques such as restoration, edge detection, image enhancement, image segmentation, filtering & noise removal.
- ***Design and Train:*** The network will be designed by creating input, hidden, neural structures. This is done for feature extraction data separation, activation and training along with validation for face recognition.
- ***Deployment on Raspberry Pi:*** The trained model will be deployed on hardware such as raspberry pi 4, model B.
- ***Audio Output:*** Using natural language processing (NLP), we will be automating a system to covert all kinds of texts, images, videos or program feedbacks in the form of speech.

Deliverables:

- Increased computational speed with large percentage accuracy in face detection and recognition.
- Multi language response with clear voice for kids and adults.
- Rechargeable battery with good battery backup.
- User Face data encryption within the limits of the respective system.
- Instant training with new faces.
- Has a feedback system for error correction.

Block Diagram: The detailed block diagram specification of the project is as shown in the figure bellow



Each of the above blocks are explained as follows,

1. ***Collection of images as real data:*** Initially a set of resources that are collection of raw statistics or direct information generated from various case study are collected in the form of dataset.
2. ***Building Neural Network Model:*** The neural network model is the built using various trial and error methods and the neural network makes some random predictions, these predictions are matched with the correct output and the error or the difference between the predicted values and the actual values is calculated. The function that finds the difference between the actual value and the propagated values.
3. ***Training:*** Training a model is done to learn and determine good values for all the weights and bias for all the labeled examples. It consists of the sample output data and the corresponding sets of input data that have an influence on the output. The training model is used to run the input data through the algorithm to correlate the processed output against the sample output. The result from this correlation is used to modify the model.
4. ***Validation and Trained Model Generation:*** Validation is the process of evaluating the trained model on the test data.This provides the generalization ability of trained model.
5. ***Parallel Processing Using Camera:*** The camera is then used to collect the input data in real-time.
6. ***Generation of real model and comparison:*** Here the real model is trained keeping the camera captured data as input and this is compared with the pre-trained model. This result is used to run the input data through the algorithm to correlate the processed output against the sample output. The result from this correlation is used to obtain the desired results.
7. ***Output:*** The output so obtained is required result.

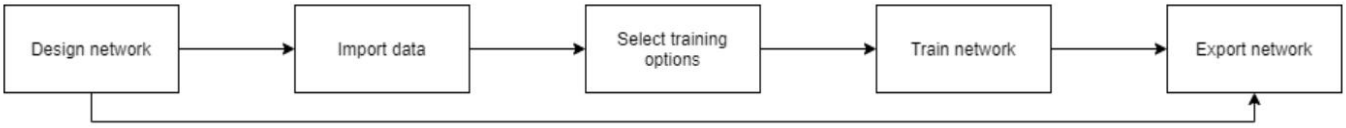
The ground-work that is going on / done so far:

The entire work progress has been divided into different categories as per the block diagram flow,

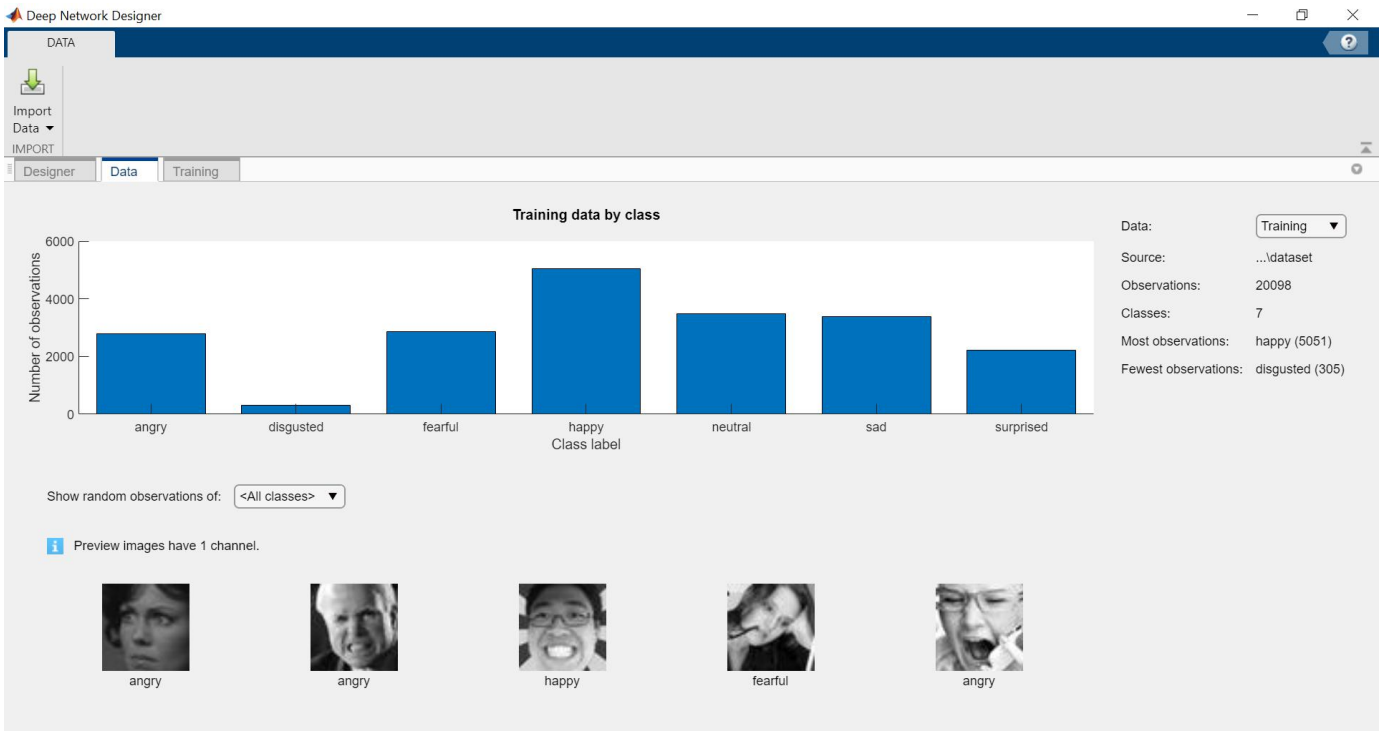
Collection of dataset: A dataset consisting of 35,906 test images was taken from kaggle after many trail and errors. This further had 7 categories of human facial emotions like 5000 images expressing human anger, 500 different disgusted human faces,4000 images displaying human fearfulness, 8000 happy expressions, 5000 neutral emotions, 5000 sad facial expressions and finally 3000 surprised human faces. All these were grouped to form a single data cluster and were used for further process. The link for availing the dataset used is given bellow:

https://www.kaggle.com/datasets/ananthu017/emotion-detection-fer/metadata

Building a neural network model: MATLAB provides Deep Learning Toolbox which is a framework for designing and implementing deep neural networks with algorithms, pretrained models, and apps. You can use convolutional neural networks (ConvNets, CNNs) and long short-term memory (LSTM) networks to perform classification and regression on image, time-series, and text data. One such tool is Deep Network Designer. This is used to Build, visualize, and train deep learning networks interactively. This also helps to Analyze your network to check that you have defined the architecture correctly and detect problems before training. Import and visualize training data, specify training options, and track network training using animated plots of training progress. Generate code to recreate network construction and training, and export trained networks.



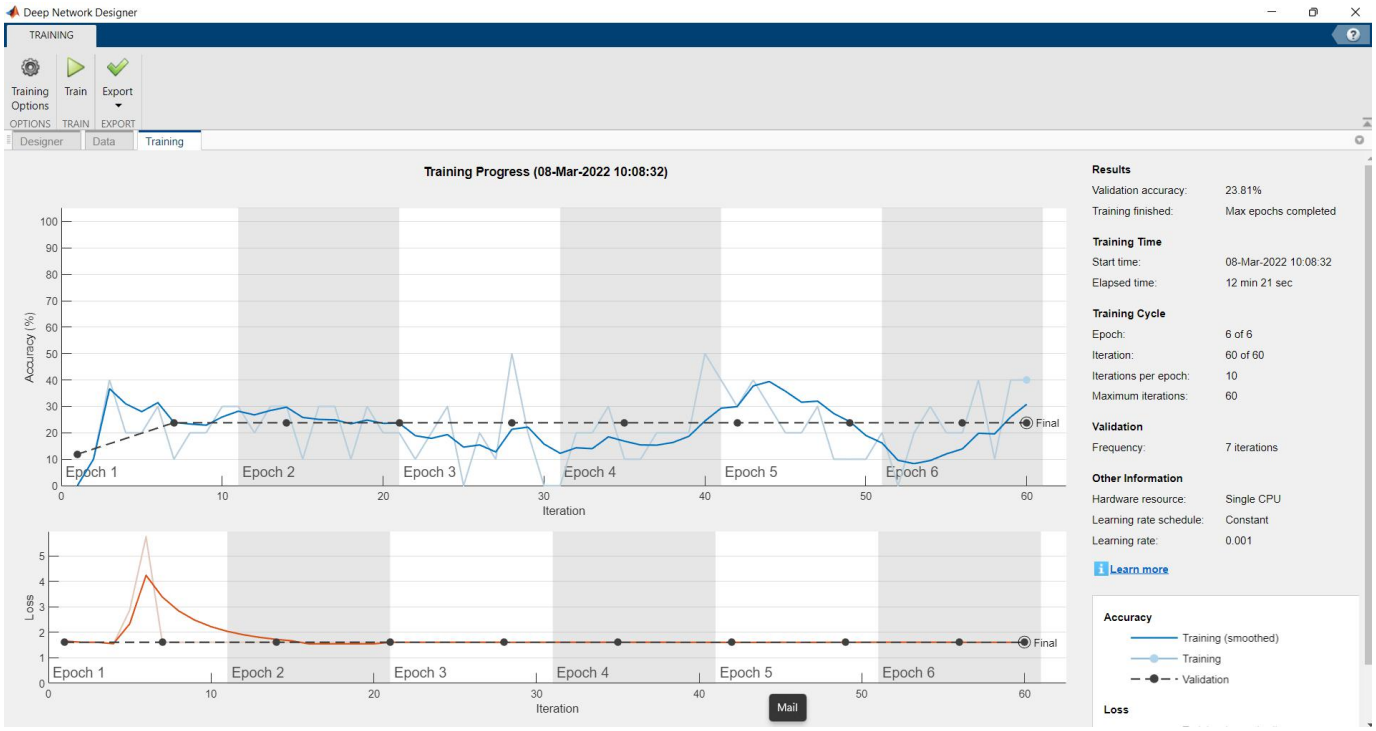
The data import to the deep network designer is displayed as follows,

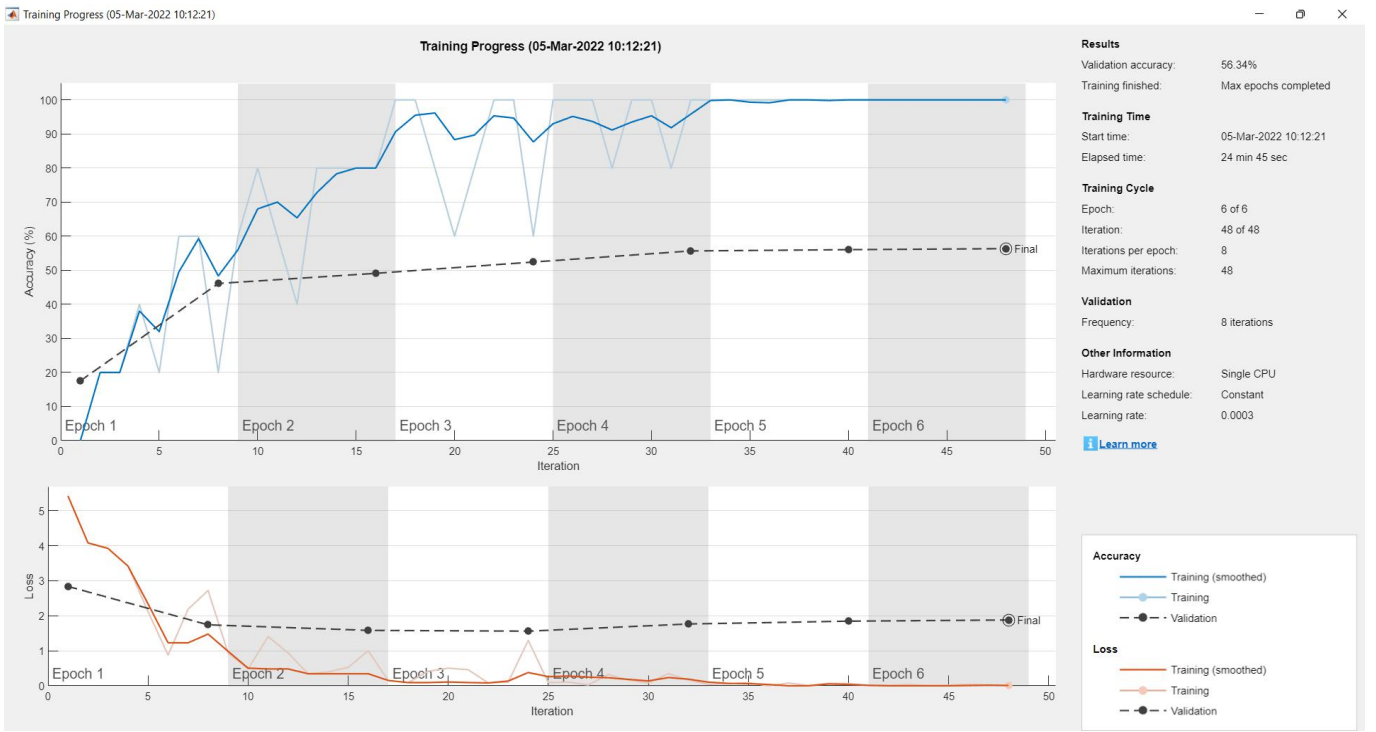


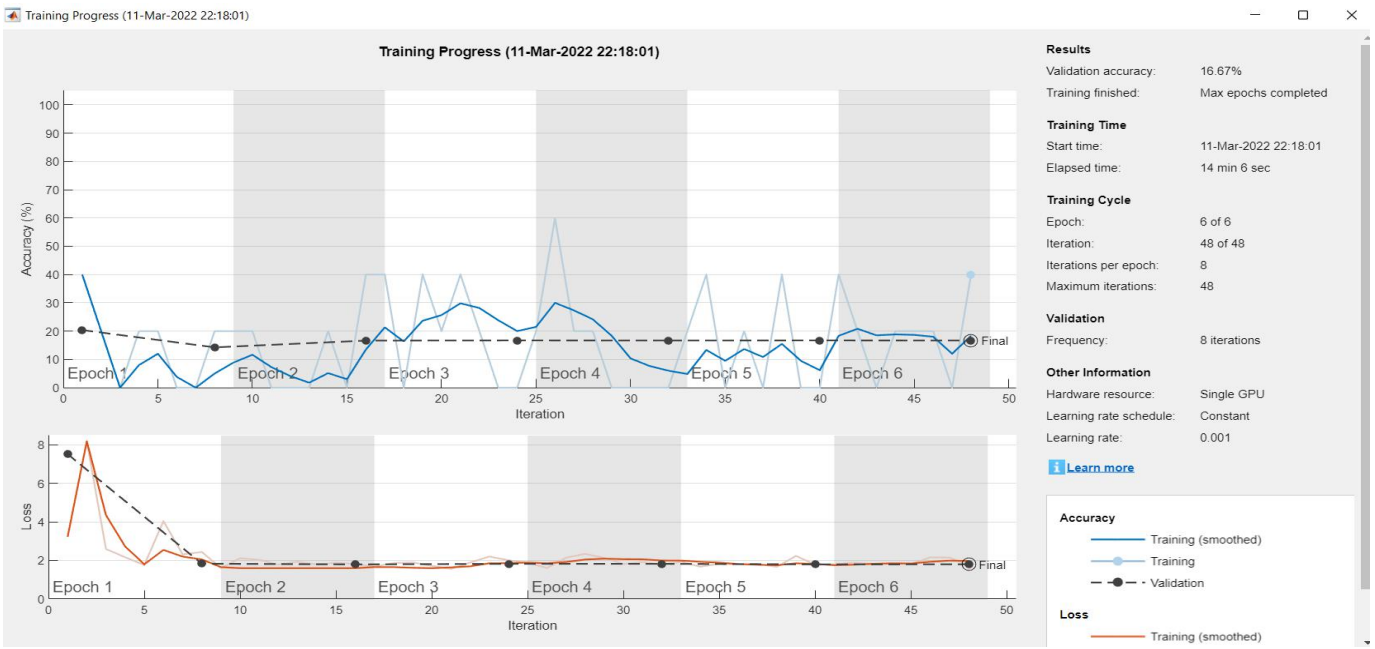
Input Dateset Configuration

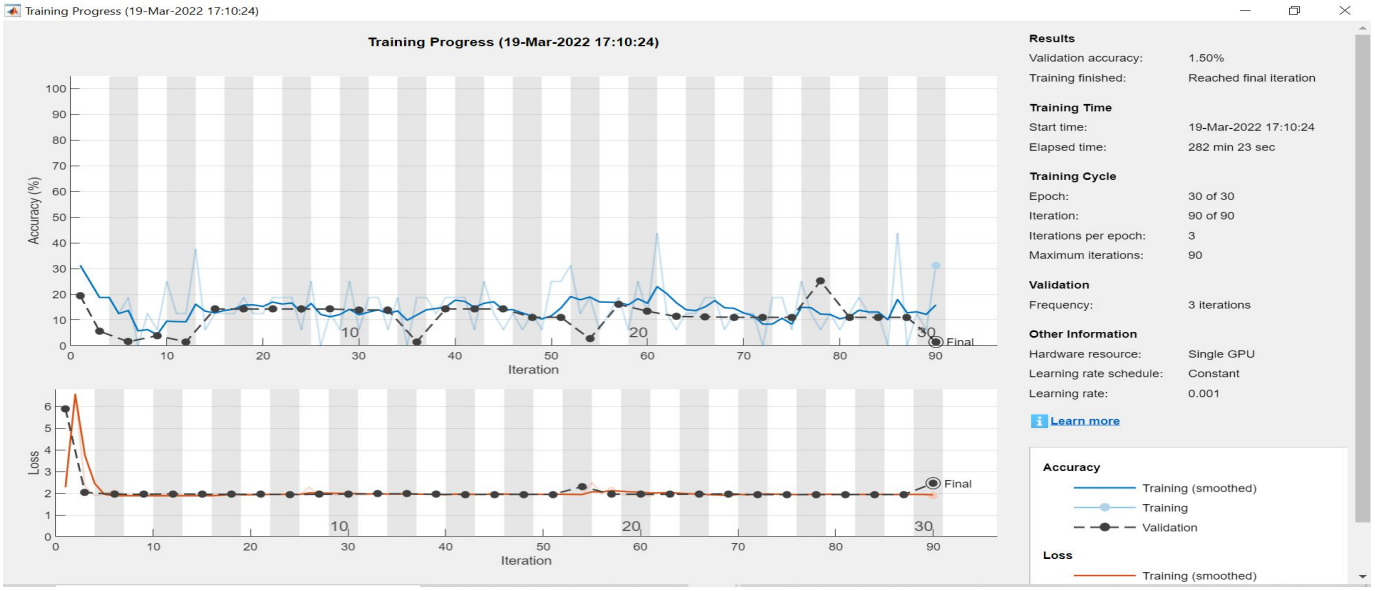
We tried using the same method for our dataset and we could observe the following results,

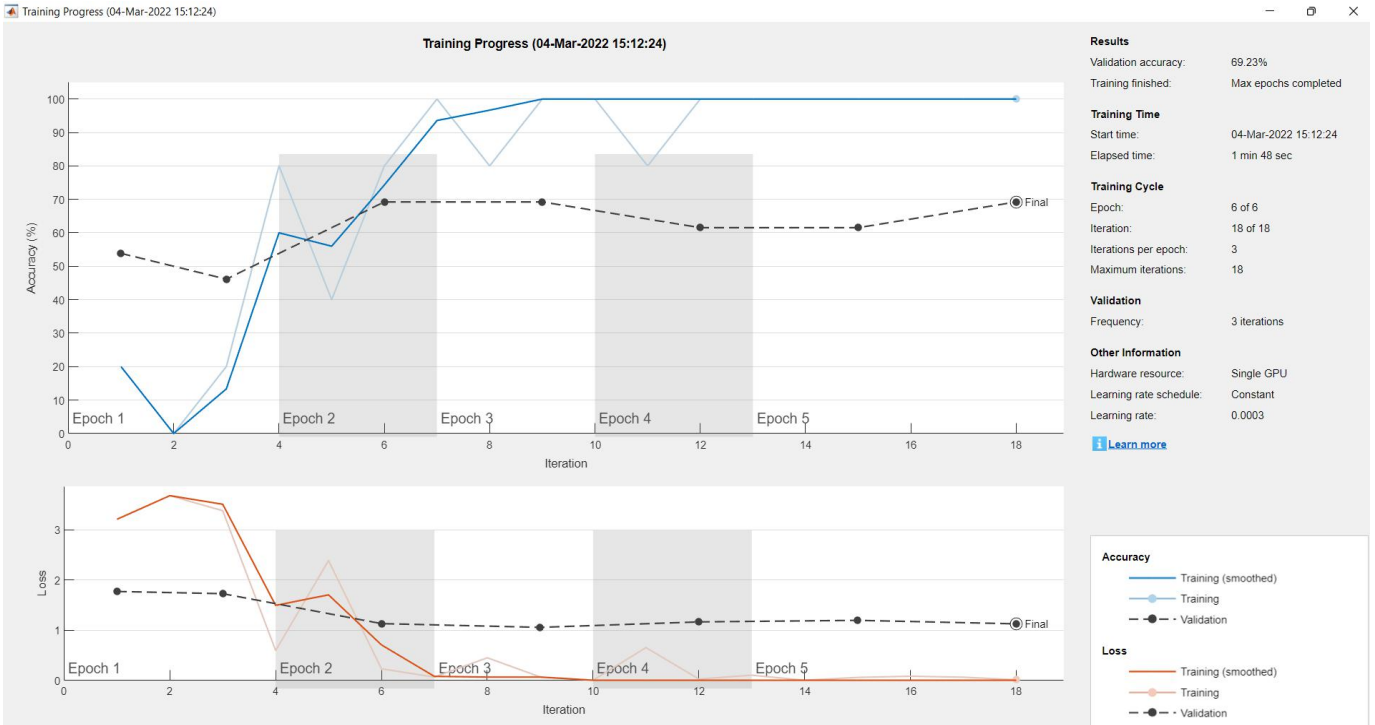
S No.	Trained Model	Epoch	minbatchsize	Accuracy Score in %
1.	ResNet-18	30	16	23.81
2.	ResNet-50	30	16	56.54
3.	GoogLeNet	30	126	16.67
4.	AlexNet	30	256	1.56
5.	Self Trained Model	30	16	69.23



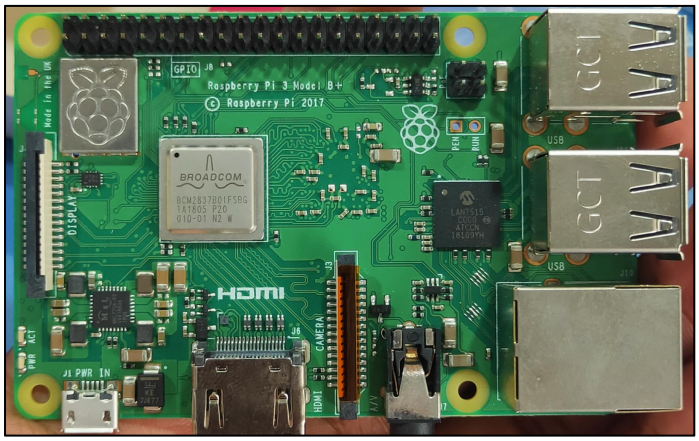


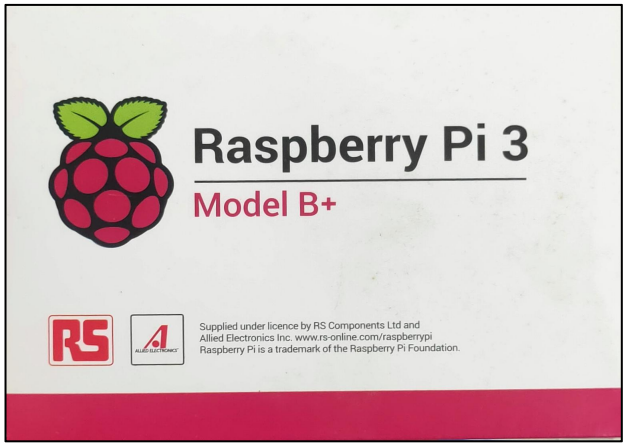

Training Progress (11-Mar-2022 22:18:01)
Training Progress (11-Mar-2022 22:18:01)
Accuracy (%)
Iteration
Epoch 1
Epoch 2
Epoch 3
Epoch 4
Epoch 5
Epoch 6
Final
Loss
Results
Validation accuracy: 16.67%
Training finished: Max epochs completed
Training Time
Start time: 11-Mar-2022 22:18:01
Elapsed time: 14 min 6 sec
Training Cycle
Epoch: 6 of 6
Iteration: 48 of 48
Iterations per epoch: 8
Maximum iterations: 48
Validation
Frequency: 8 iterations
Other Information
Hardware resource: Single GPU
Learning rate schedule: Constant
Learning rate: 0.001
Learn more
Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)


Training Progress (19-Mar-2022 17:10:24)
Training Progress (19-Mar-2022 17:10:24)
Accuracy (%)
Iteration
10
20
30
Final
Loss
Results
Validation accuracy: 1.50%
Training finished: Reached final iteration
Training Time
Start time: 19-Mar-2022 17:10:24
Elapsed time: 282 min 23 sec
Training Cycle
Epoch: 30 of 30
Iteration: 90 of 90
Iterations per epoch: 3
Maximum iterations: 90
Validation
Frequency: 3 iterations
Other Information
Hardware resource: Single GPU
Learning rate schedule: Constant
Learning rate: 0.001
Learn more
Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)


Training Progress (04-Mar-2022 15:12:24)
Training Progress (04-Mar-2022 15:12:24)
Accuracy (%)
Iteration
Epoch 1
Epoch 2
Epoch 3
Epoch 4
Epoch 5
Final
Loss
Results
Validation accuracy: 69.23%
Training finished: Max epochs completed
Training Time
Start time: 04-Mar-2022 15:12:24
Elapsed time: 1 min 48 sec
Training Cycle
Epoch: 6 of 6
Iteration: 18 of 18
Iterations per epoch: 3
Maximum iterations: 18
Validation
Frequency: 3 iterations
Other Information
Hardware resource: Single GPU
Learning rate schedule: Constant
Learning rate: 0.0003
Learn more
Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)
Training
Validation

With these results in our hand we have been making multiple attempts to improve the accuracy score and parallelly we have also procured some of the hardware components such as Raspberry Pi 3 Model B+, Adaptor 5V/2.5V, HDMI to VGA Connector and 64GB microSD card.





The role and responsibilities of each member in the team

Under the guidance of our mentor Dr. Rajashri Khanai madam each member of the team is made actively work throughout the phase of the project. They could be directly involved in campus or sometimes in online modes in their absence. Each member of the team has the following roles and responsibilities,

- Contributing to overall project objectives.
- Completing individual deliverables.
- Deliver project responsibility within deadline.
- Providing expertise.
- Working with mentors to establish and meet project needs.
- Documenting the progress, setback and new processes.

The action plane of the recent future

As displayed in the block diagram the collection of dataset, building the neural network model and training is been achieved, but the results aren't satisfactory. Significant improvements must be made to improve the accuracy scores with the different set of models to get satisfactory results. As soon as the acceptable score is achieved the next step would be to deploy the trained model through matlab on to raspberry Pi 3 B+ model. Attempt are to be made to achieve the audio feedback.

The expectation from mentor for the team

We believe mentor will facilitate our thinking.Good engagement with our mentor on an ongoing basis is crucial for the success of the scheme for all parties. Basically a mentor should be able to meet and participate in research program in activities. To help the students to learn and grow to his or her potential. To support in the development of the students profession and career. To provide honest and constructive feedback. To set examples of values, ethics and professional practices. Moreover **Mr MGPL Narayana sir** is an amazing mentor. He helped us tremendously to gather and analyze our data, as well as interpreting the results and communicating these results. He is instrumental to our research and we are lucky to have him as our advisor.