MATLAB SPEECH RECOGNITION USING DEEP LEARNING TO INSTRUCT ROBOTS IN THE INDUSTRIES

A.VISVA SAI SRIDHAR **B.SREE HARSHA** **K.TARUN SAI CHOWDARY**

GUIDE:Dr.Malay kumar hota.

ABSTRACT

Robots are perfect replacement for humans in the industry.But there are some areas where humans need to instruct the robots to place objects.One such application is in the diary industry and the automobile industry.This projects aims at instructing the robots to place the objects according to the wish of the controller present in the control room.We use matlab for simulation and Artificial neural network to train the robots.

Introduction

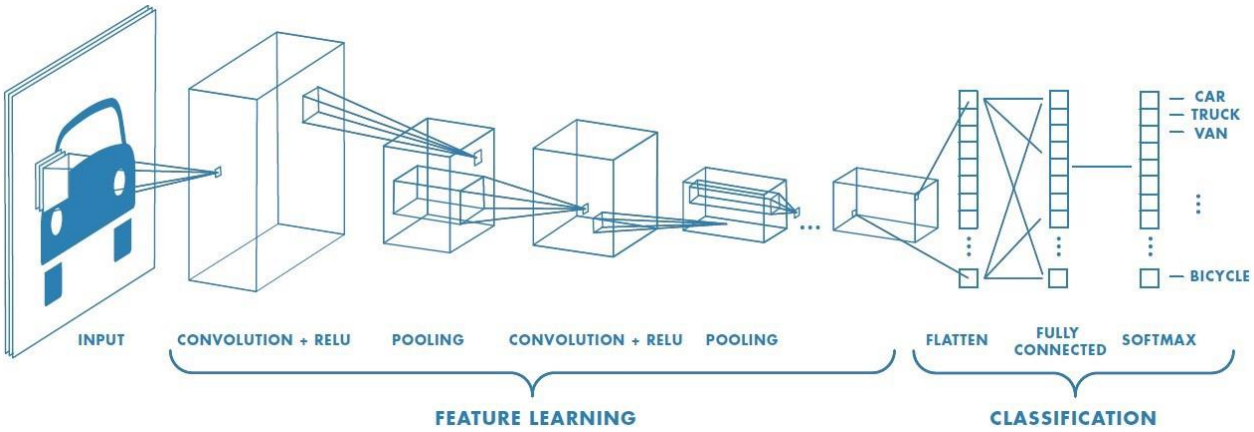
Speech recognition is one of the next generation technologies for human-computer interaction. Speech recognition has been researched since the late 1950s but due to its computational complexity and limited computing capabilities of the last few decades, its progress has been impeded. In laboratory settings automatic speech recognition systems (ASR) have achieved high levels of recognition accuracies, which tend to degrade in real world environment

The latest trend in automation comes from the fact that deep learning is the new technology adopted by many industries.This can also be used for speech recognition.This ensures that computers/robots learn the signals before hand and thus identify them when the instructor uses the same language that was pre heard by the robot.

Deep learning consists of a multiple of machine learningalgorithms fed with inputs in the form of multiple layeredmodels. These models are usually neural networks consistingof different levels of non-linear operations. The machinelearning algorithms attempt to learn from these deep neuralnetworks by extracting specific features and information .Prior to 2006, searching deep architecture inputs was not apredictable straight forward task; however, the developmentof deep learning algorithms helped resolve this issue andsimplified the process of searching the parameter space ofdeep architectures . Deep learning models can also operateas a greedy layerwise unsupervised pre-training. This meansthat it will learn hierarchy from extracted features from eachlayer at a time.

METHODOLOGY USED

TRAINING A CONVOLUTION NEURAL NETWORK USING MATLAB



Imagecredits: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

A **Convolutional Neural Network (ConvNet/CNN)** is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

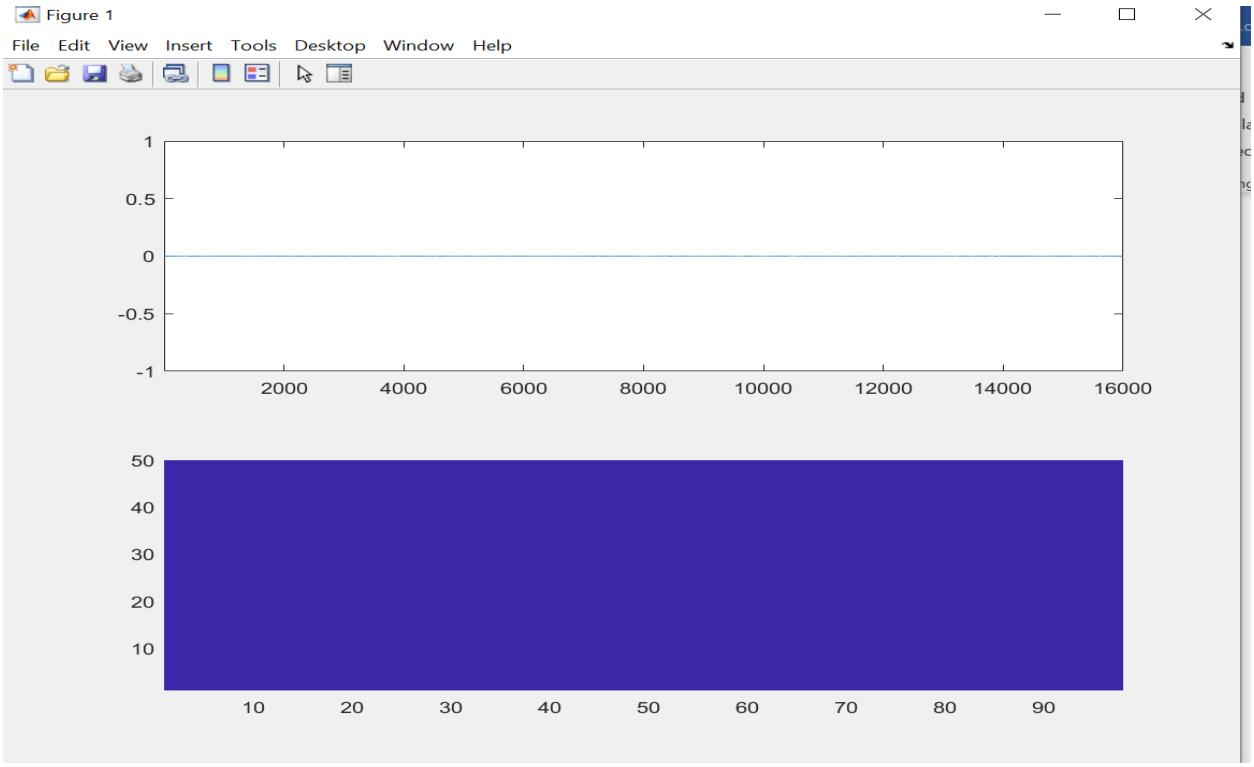
The commands that were to be identified by the robot are given to the convolution neural network and we Use a weighted cross entropy classification loss. weightedClassificationLayer(classWeights) creates a custom classification layer that calculates the cross entropy loss with observations weighted by classWeights. Specify the class weights in the same order as the classes appear in categories(YTrain). To give each class equal total weight in the loss, use class weights that are inversely proportional to the number of training examples in each class. When using the Adam optimizer to train the network, the training algorithm is independent of the overall normalization of the class weights.

Dataset used for training:Google voice data set 2017.

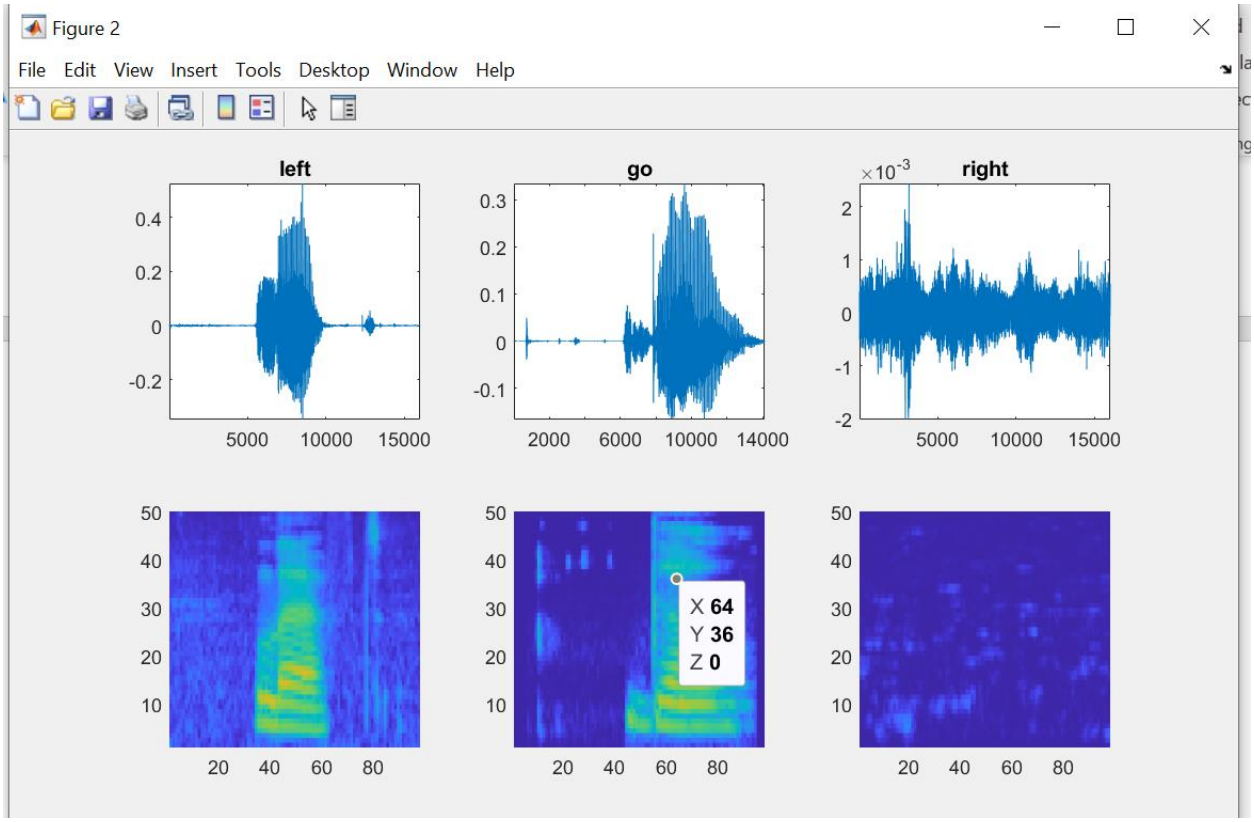
Matlab code for the project: https://drive.google.com/file/d/1RD6-ZcGl6tkXI4kwKKS3T4yBLUhiBzQv/view?usp=sharing

RESULTS

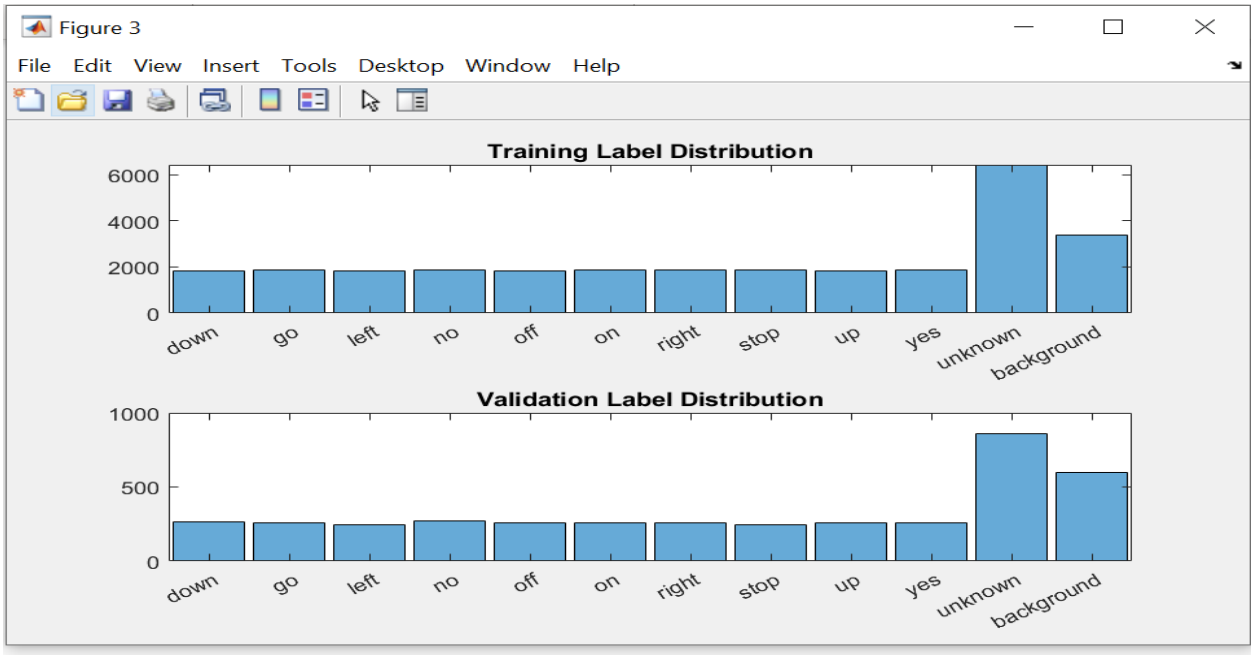
RECOGNITION OF THE SOUND FROM MICROPHONE



RECOGNIZING THE COMMANDS THAT WERE GIVEN



THE DISTRIBUTION OF COMMANDS FOR THE INPUT VOICE

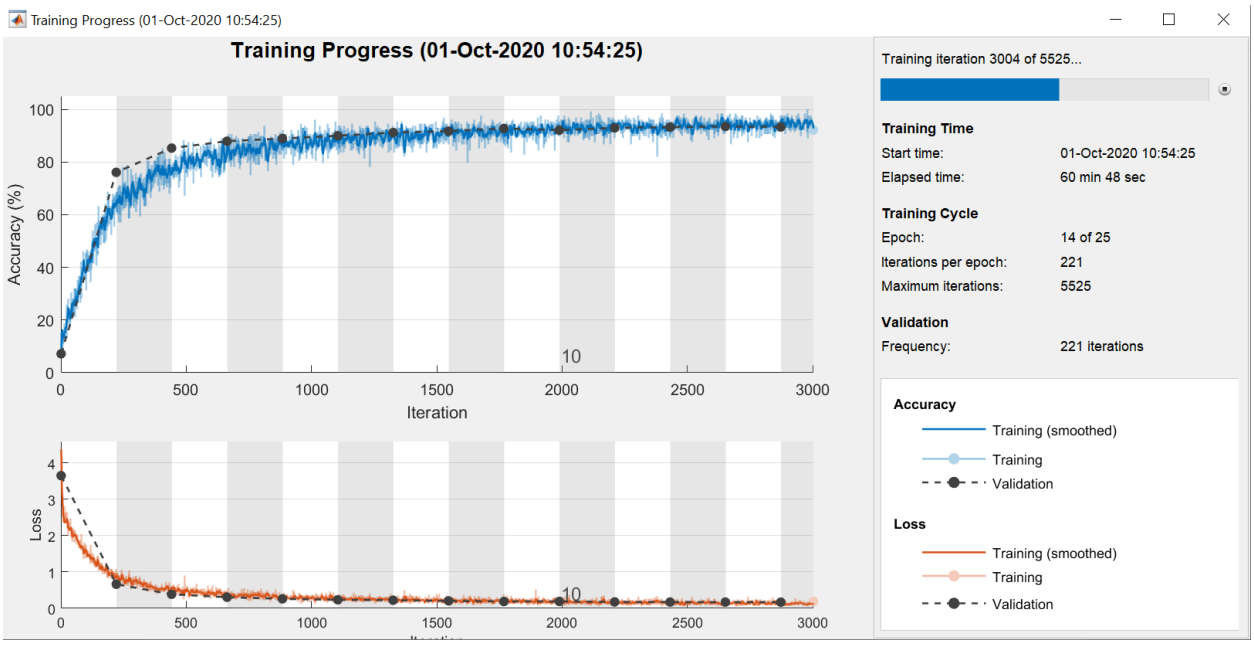


TRAINING PROGRESS

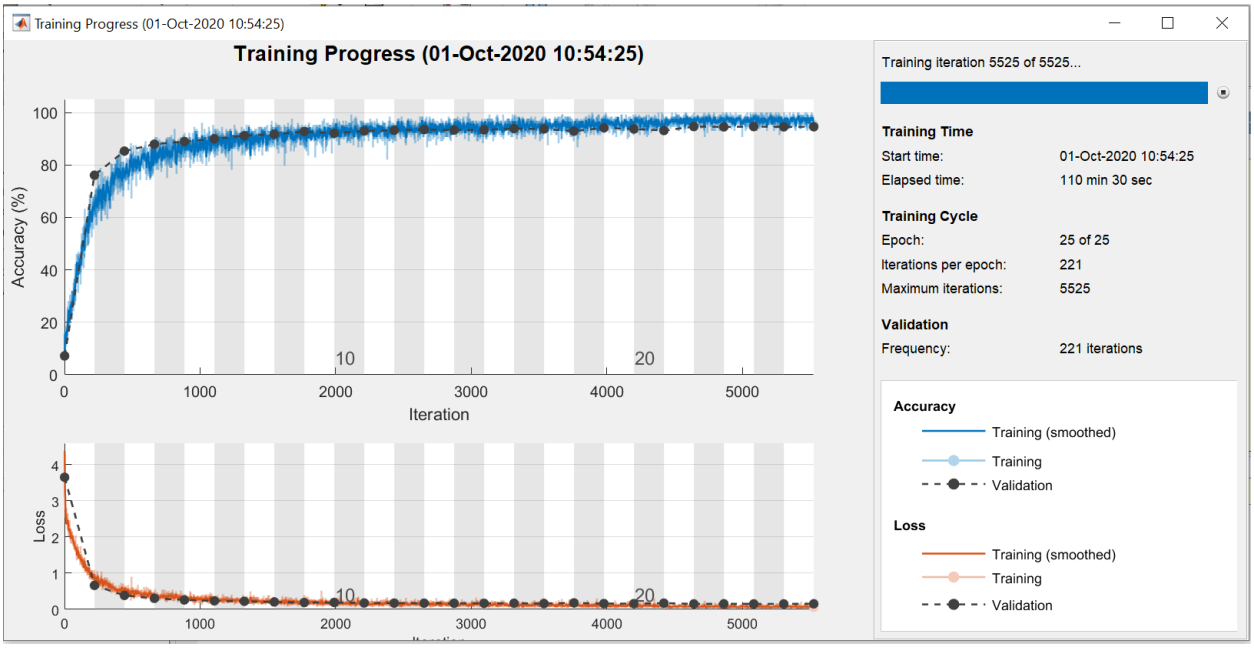
NO OF EPOCHES:25

NO OF ITERATION:5522

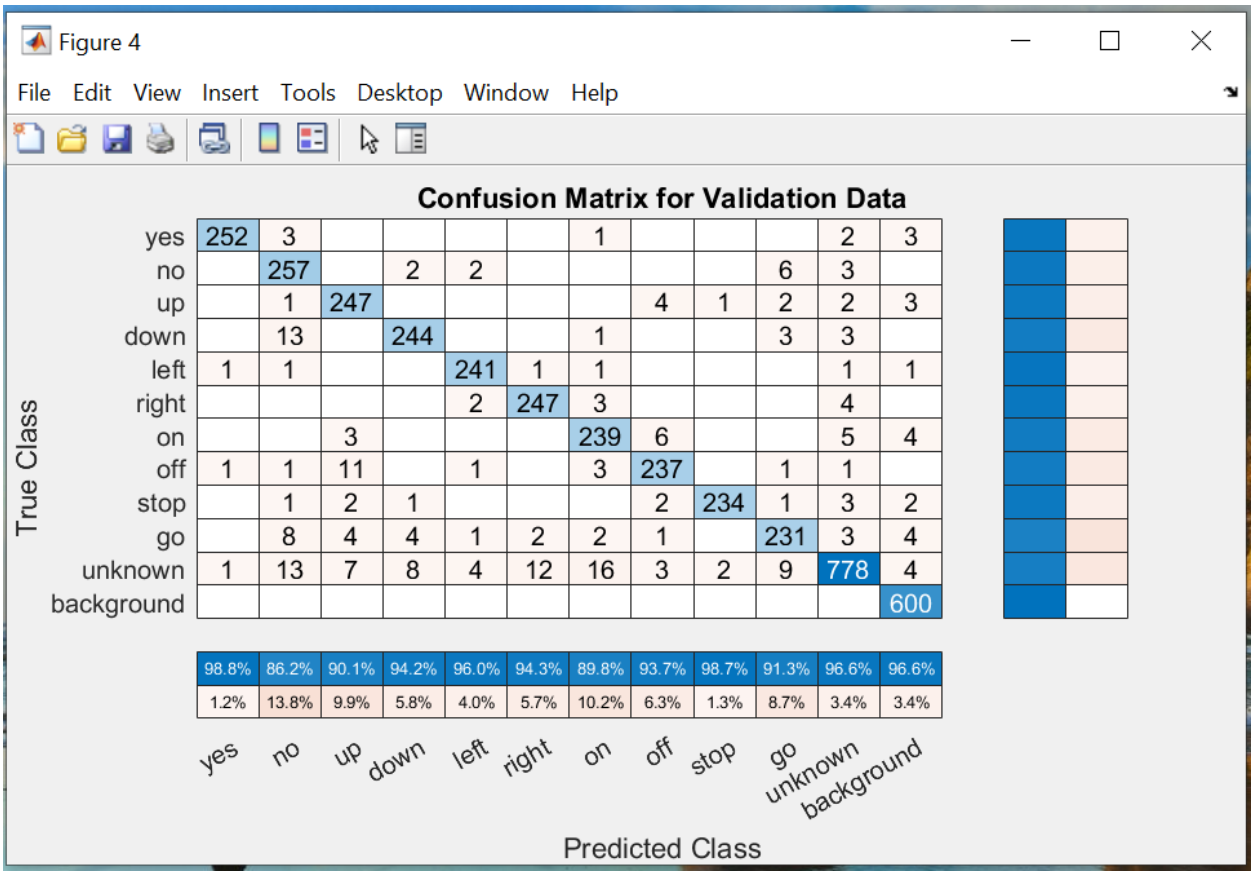
THE BELOW IS THE PROGRESS AT THE 14TH EPOCH AND 3004TH ITERATION



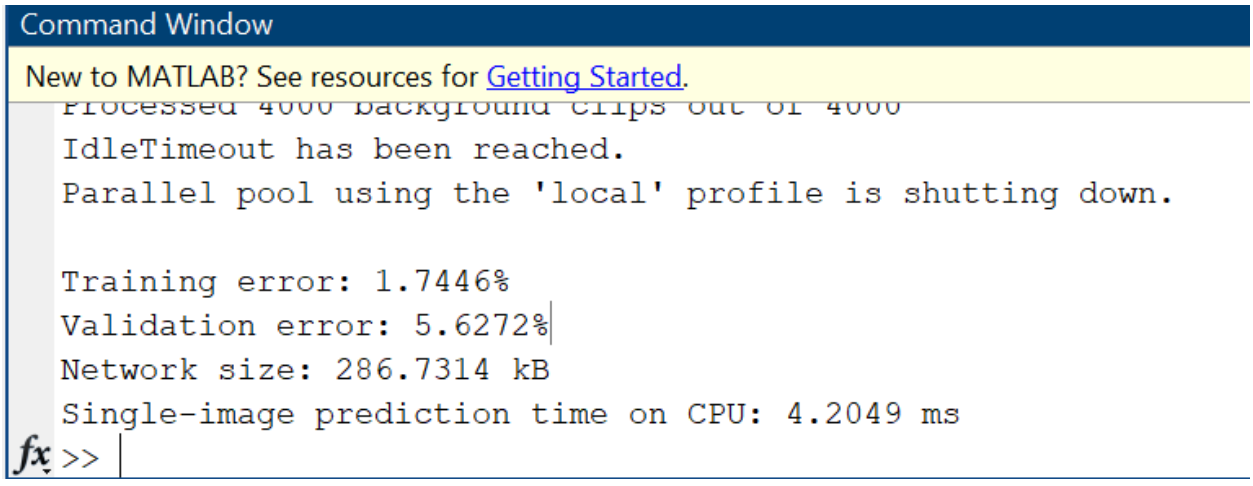
END RESULT AFTER TRAINING



CONFUSION MATRIX TO TEST THE ACCURACY



OUTPUT IN THE COMMAND WINDOW



For the entire coomad window code refer the link: https://drive.google.com/file/d/10OZwf-E3_RsDL_jKTPeWn0sqEzIb2zAQ/view?usp=sharing

CONCLUSION

Since the neural network is now trained.This can be deployed into a robot working in the industry to efficiently follow the instructor at the workstation.This not only ensures the accuracy of the robot but also helps the industry to use humans for more productive work.

The accuracy in the project is satisfactory and there were enough no of epochs performed to train the network ensuring that there wont be any fluctuation in the future.

After 5525 iteration and 25 epochs the accuracy is maximum and the loss is minimum hence the model is satisfactory

References

1. [1] Warden P. "Speech Commands: A public dataset for single-word speech recognition", 2017. Available from https://storage.googleapis.com/download.tensorflow.org/data/speech_commands_v0.01.tar.gz. Copyright Google 2017. The Speech Commands Dataset is licensed under the Creative Commons Attribution 4.0 license, available here: https://creativecommons.org/licenses/by/4.0/legalcode.
2. https://www.researchgate.net/publication/330815113_Speech_Recognition_Using_Deep_Neural_Networks_A_Systematic_Review
3. https://www.researchgate.net/publication/335106017_Speech_Recognition_Using_MATLAB_and_Cross-Correlation_Technique