Improved Gesture Recognition

Team B:
Thasina Tabashum
Arif Ahmed
Dan Waters
Ryan Moye

Abstract

According to NIDCD, approximately 7 million people in the U.S. struggle and experience difficulty in using their voices[1]. Vocalization difficulty can lead to several social and daily life complications. With this in mind, our motivation is to use proper validation techniques as well as hypertune the parameters.

We did GridSearch to better tune our model and improve the overall accuracy. We also validated our model by doing Group KFold Subject-Wise Cross Validation to ensure it does not overfit, as well as to prove our model is not biased to the training data.

Motivation

- According to NIDCD, more than seven million people have problems with speech.
- Difficulty with speech can lead to social and job related complications.
- By developing an assistive technology that allows the user to communicate more effectively, we can improve the daily life of the user.
- This project will help the user reduce the conversational response time and increase their ability to communicate more naturally.

Previous Work by Team A

- In the previous iteration of this project by Team A, they created an LSTM model for gesture recognition based on the labeled dataset of user activities.
- The dataset consists of the timestamp, activity, and x, y and z of each activity.
- Team A created a confusion matrix of their model.

Design and Milestones

- The previous group built an LSTM model which we needed to further validate.
- We did hyper-parameter tuning; previously the model was giving around 92% accuracy when we did subject-wise cross validation. But after hyperparameter tuning we achieved 97% on that. Further extensive tuning can be done but it might over-fit the model. If we have more data from the new subjects further hyper-parameter tuning can be route to take.
- We have also did groupkfold validation.
- We made a code repository for the future groups to carry on the project.

Data Specifications

Our data set comes from the last team's github, it is a CSV file including 133,136 samples with 6 subjects defined as user number 25-30 with 11 gestures.

The data type is integer values and each sample represents 5 main labels; timestamp, gesture, X, Y, and Z values regarding each gesture. The dataset has been labeled and as for now, we are approaching a supervised learning method. The main goal is to use the accelerometer information to classify a gesture into one of the eleven categories of the gestures that have been provided in the data set.

LSTM Model (in summary)

The model layers in summary

- Bidirectional LSTM layer: Input
- Dropout Layer
- Dense Layer
- Denser Layer : Output

Hyper parameter Tuning

We have performed GridSearch technique to search best hyper parameters, by splitting the dataset on train set, test set and validation set.

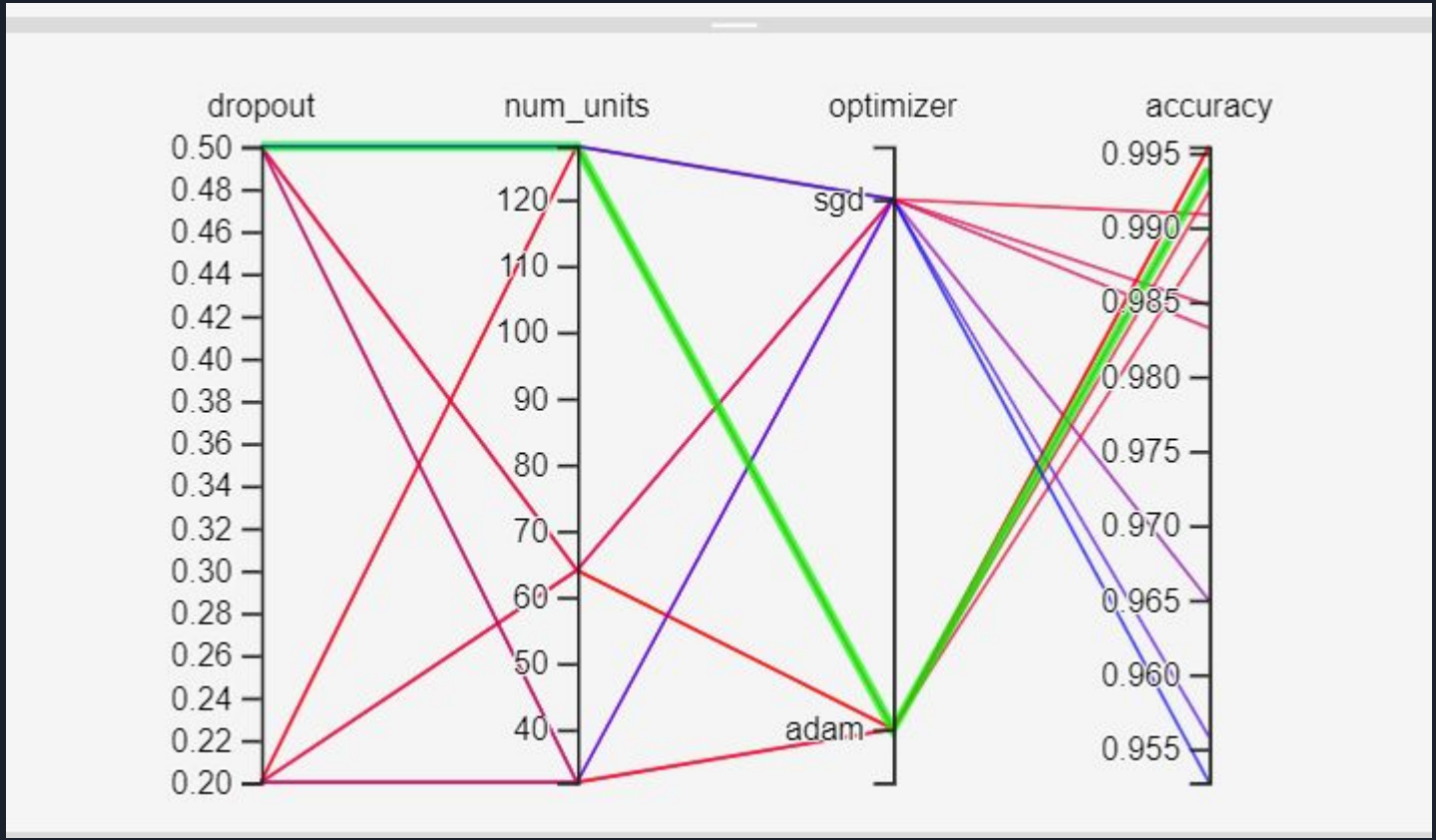
- We have tried 32, 64 and 128 units for LSTM and Dense layers
- Drop-out : 0.2 , 0.5
- Optimizer: adam , sgd

Picking Hyper-Parameters from GridSearch

The best pick from the GridSearch is 64 units and 0.2 drop-out. But after we did subject-wise validation it did not work well.

Then we picked the 128 units and 0.2 dropout. Which gave around 97% accuracy on subject-wise cross-validation.

Tensorboard: Parallel Coordinates View



Tensorboard: Hyperparameter Tuning Table View

dropout	num_units	optimizer	accuracy
0.20000	32.000	sgd	0.96494
0.20000	32.000	adam	0.98933
0.20000	128.00	sgd	0.99085
0.20000	64.000	adam	0.99543
0.50000	64.000	adam	0.99543
0.50000	64.000	sgd	0.98476
0.20000	128.00	adam	0.99390
0.20000	64.000	sgd	0.98323
0.50000	32.000	sgd	0.95579
0.50000	32.000	adam	0.99238
0.50000	128.00	adam	0.99390
0.50000	128.00	sgd	0.95274

Why We Need Further Validation

- Validation is useful in determining the effectiveness of a model.
- It is also used to ensure there is no overfitting of the model.
- Further validation is important in our model to ensure that the model correctly identifies the gesture with high accuracy.

Group K-Fold Cross Validation

- GroupKFold is a variation of k-fold which ensures that the same group is not represented in both testing and training sets.
- For example our dataset is obtained from 6 different subjects with several samples per-subject and if the model is flexible enough to learn from highly person specific features it could fail to generalize to new subjects. GroupKFold makes it possible to detect this kind of overfitting situations.

Accuracy from Subject-Wise Cross Validation

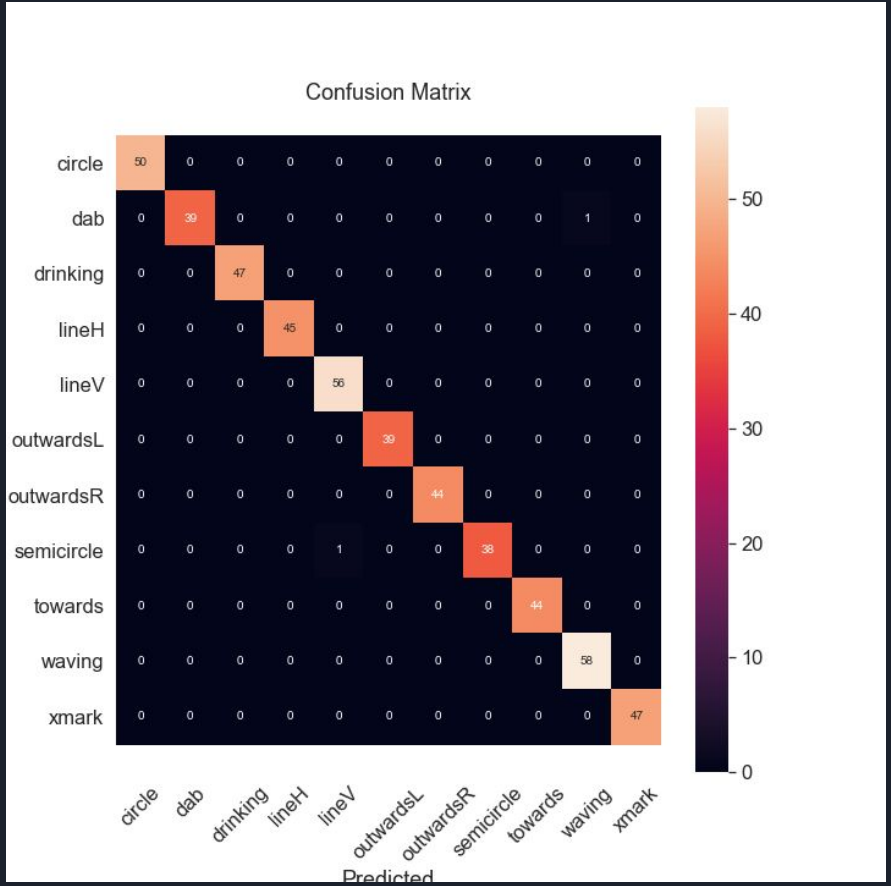
- After performing hypertuning and validation techniques, we were able to get an accuracy of approximately 97%.
- Some subjects had a higher accuracy than others, which is expected as different people will gesture in different manners. The standard deviation of accuray is 3.5%.
- Note, only one subject showed around 90% accuracy, but for others it was around 97-99%
- We used confusion matrices to have better understanding how our model is performing

Confusion Matrix:

We have achieved the best accuracy on the 4th and 2nd subject and shows a relatively high accuracy on predicting the gesture that the subject displayed. (accuracy 99.7%)

We demonstrated confusion Matrix for 4th subject, here.

Two of the mis-classifications are
--Semicircle predicted lineV
-- Dab as waving

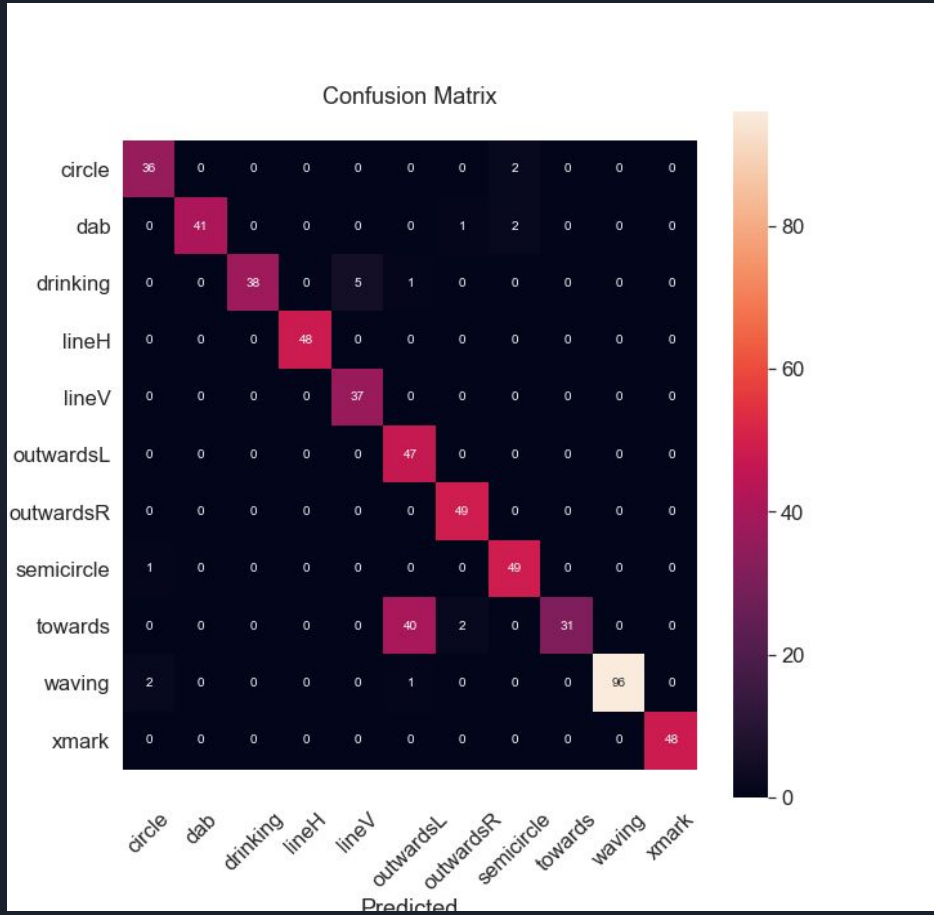


F1-Score and Recall

	precision	recall	f1-score	support
circle	1.00	1.00	1.00	50
dab	1.00	0.97	0.99	40
drinking	1.00	1.00	1.00	47
lineH	1.00	1.00	1.00	45
lineV	0.98	1.00	0.99	56
outwardsL	1.00	1.00	1.00	39
outwardsR	1.00	1.00	1.00	44
semicircle	1.00	0.97	0.99	39
towards	1.00	1.00	1.00	44
waving	0.98	1.00	0.99	58
xmark	1.00	1.00	1.00	47
accuracy			1.00	509
macro avg	1.00	1.00	1.00	509
weighted avg	1.00	1.00	1.00	509

Confusion Matrix:

The 3rd subject had the worst accuracy. As we can see from the confusion matrix, there is a lot of confusion with the ‘towards’ and ‘outwardsL’ gestures.



Future Direction

- Further hyper parameter tuning can be done
- Clustering similar kind of activity in one to see how much accuracy they can achieve.

Repository & Archive

Our Github Repository can be found at:

https://github.com/danwaters/improved-gesture-recognition

Resources

Google Drive link of the project:

https://drive.google.com/file/d/130i_S9aILLGgfRVPtPoREW5MKo9nMreP/view?usp=sharing

Previous group repository:

https://github.com/sridharnandigam/GestureRecognition