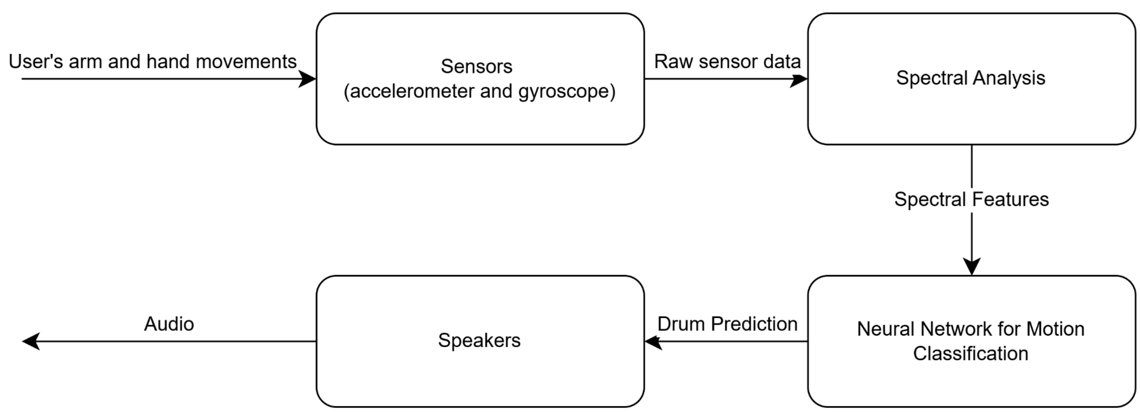
Max Tang
Jiarong Wang
18-848 Embedded Deep Learning
Project Report
Dec. 15, 2024

1. Abstract:
   AirDrums: a virtual drum set powered by deep learning. AirDrums packages an entire drumset into just a wristband with built-in sensors. By wearing the wristband and making the motions of hitting a drum set with a drumstick, we can use deep learning to identify the motion made and which drum the user pretended to hit to play the drum noise. Drum sets are large, unwieldy, and loud. AirDrums solves all of these problems by providing a new, portable, drum set that drummers can use to practice anywhere. What sets AirDrums apart from existing virtual drumsets is there there is no need for any extra equipment other than the wristband. Existing virtual drumsets use specialized drumsticks, a camera, and require a well-lit room to track the user's movements. AirDrums does not have any of those limitations, and is truly a portable drumset that allow users to use their favorite drumsticks to practice drumming on the go.

2. Block Diagram



3. Feature Extraction
   The spectral analysis block used in the AirDrums model takes in raw accelerometer and gyroscope data and processes it using a Fast Fourier Transform (FFT). It uses a FFT length of 32 and takes the logarithmic value of the spectrum for dynamic range compression. The output of this spectral analysis is processed features, which are used as inputs to the motion classification model. This allows the model to better identify patterns and classify motion based on its frequency characteristics.

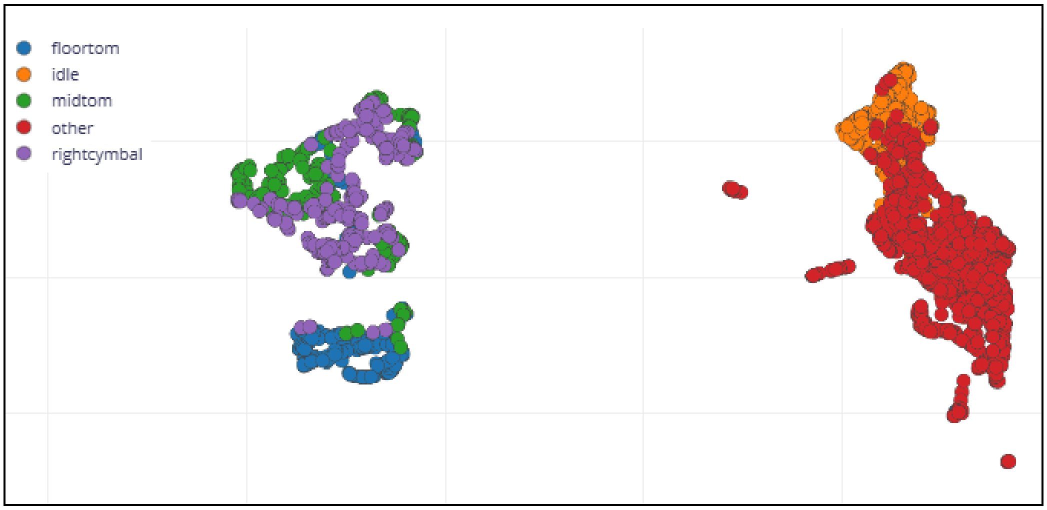


Figure 1. By applying the FFT to the raw data, clear decision boundaries between motion classes appear.

4. Classifier Architecture

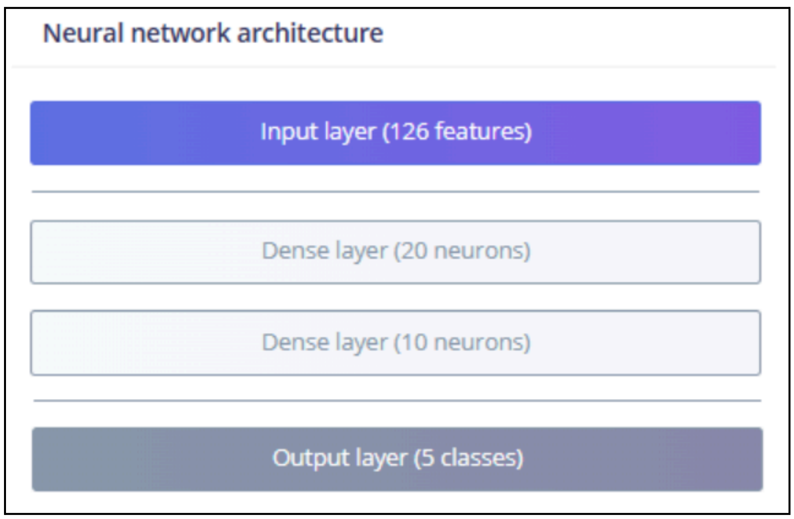


Figure 2. The neural network architecture for motion classification.

The motion classification model used by AirDrums is very simple and lightweight, consisting of just an input layer, two small dense layers, and one output layer. The reasoning behind this small neural network is because of the latency requirement. To achieve a realistic and optimal user experience, the AirDrums should have minimal latency and feel as responsive as hitting a real drum. Therefore, the neural network has to be as fast as possible. AirDrums can achieve high accuracy with such a small model because there are only three drum classes for the model to predict, as well as an idle and other class. This means that a smaller neural network can still achieve high accuracy.

5. Confusion Matrix and Overall Accuracy Metrics

Model Model version: Quantized (int8)

Last training performance (validation set)

ACCURACY
94.5%

LOSS
0.18

Confusion matrix (validation set)

	FLOORTOM	IDLE	MIDTOM	OTHER	RIGHTCYMBAL
FLOORTOM	97.7%	0%	0%	0%	2.3%
IDLE	0%	93.6%	0%	6.4%	0%
MIDTOM	0.3%	0%	99.2%	0%	0.5%
OTHER	0%	7.6%	0%	92.4%	0%
RIGHTCYMBAL	0.3%	0%	4.7%	0%	95%
F1 SCORE	0.99	0.89	0.97	0.95	0.96

Metrics (validation set)

METRIC	VALUE
Area under ROC Curve	1.00
Weighted average Precision	0.95
Weighted average Recall	0.94
Weighted average F1 score	0.95

Figure 3. Confusion matrix and accuracy metrics from the validation set. The AirDrums model was able to achieve very high accuracy across all classes.

6. Deployment Method
AirDrums uses an Arduino Nano 33 BLE Sense. The code is flashed to the Arduino board, and a USB connection to a laptop is required for power and playing sounds. During use, a separate Python script on the laptop runs to read the deployed model's predictions over a UART port and play audio files.

7. Significant Challenges
The main challenge faced during AirDrum's development was achieving high accuracy during live testing. The model often struggled to classify motions for drums that are closer to the user, such as the floor tom drum.

8. Images

Figure 4. You can use a pencil as a drumstick!

Demo Video: https://photos.app.goo.gl/w8MLmktgMDHbgLvu7