RNN Classification of Spectral EEG/EMG Data Associated with Facial Movements for Drone Control

Akintomide Adebile
Electrical and Computer Engineering
Georgia Southern University
aa22690@georgiasouthern.edu

Shaen Mehrzed
Mechanical Engineering
Georgia Southern University
sm30767@georgiasouthern.edu

Destinee Hicks
Mechanical Engineering
Kennesaw University
dhicks49@students.kennesaw.edu

Joshua Hale
Mechanical Engineering
Augusta University
johale@augusta.edu

Darryl Wiltz
Engineering
Savannah State University
dwiltz@student.savannahstate.edu

Rocio Alba-Flores
Electrical and Computer Engineering
Georgia Southern University
ralba@georgiasouthern.edu

***Abstract*—This research involves developing a drone control system that functions by relating EEG and EMG from the forehead to different facial movements using recurrent neural networks (RNN) such as long-short term memory (LSTM) and gated recurrent Unit (GRU). As current drone control methods are largely limited to handheld devices, regular operators are actively engaged while flying and cannot perform any passive control. Passive control of drones would prove advantageous in various applications as drone operators can focus on additional tasks. The advantages of the chosen methods and those of some alternative system designs are discussed. For this research, EEG signals were acquired at three frontal cortex locations (f_{p1}, f_{pz}, and f_{p2}) using electrode from an OpenBCI headband and observed for patterns of Fast Fourier Transform (FFT) frequency-amplitude distributions. Five different facial expressions were repeated while recording EEG signals of 0- 60Hz frequencies with two reference electrodes placed on both earlobes. EMG noise received during EEG measurements was not filtered away but was observed to be minimal. A dataset was first created for the actions done, and later categorized by a mean average error (MAE), a statistical error deviation analysis and then classified with both an LSTM and GRU neural network by relating FFT amplitudes to the actions. On average, the LSTM network had classification accuracy of 78.6%, and the GRU network had a classification accuracy of 81.8%.**

Keywords— RNN, LSTM, GRU, Drone Control, EEG, Machine Learning

I. INTRODUCTION

A. Motivation and Literature Review

Current abilities in automated control systems can be attributed to the availability of computing power, decades of research, and government interests. In principle, automated flight requires a control system which can sense, at minimum what a pilot can and replicate at minimum the logic and actuation which a pilot can perform. The difficulty does not lie in the formulation of such a system, moreover in the optimization thereof to face adverse conditions to adhere to safety standards. Today, machine-learning (ML) classification methods are frequently used in the medical field, advanced aircraft, and autonomous vehicles. The use of traditional computing logic to classify sensor inputs is robust and complicated while the use of novel machine learning algorithms is less robust but also less complicated. Their difference in setup complexity results from traditional computing logic requiring manual interpretations of optimal flying conditions and controls, while machine learning methods can automatically interpret these optimal conditions. Their difference in robustness results from manually set evidence or model-based thresholds being fully explainable compared to the complex automatic thresholds defined by a machine learning algorithm. Naturally, a manually configured classification system is fully understood as someone having to set up the thresholds. All applications must still be suitable for ML assistance, given reliability requirements. Considering traditional and machine learning classification advantages and disadvantages, they should be implemented in their ideal scenario.

Electroencephalography (EEG) signals are externally probed scalp voltage fluctuations (~100 μV amplitude) of general brain activity with limited applications due to high impedance and electromyography (EMG) noise [1]. However, the current understanding of EEG signals suggests that motor movements are captured in the rates of general and externally probed electric signal fluctuation coming from the human skull's skin, both EEG and EMG [1], [2], [3]. Given a machine learning algorithm's ability to find multi-dimensional, nonlinear correlations independently, it is a plausible method to classify features of an EEG signal. Other works have also effectively demonstrated multichannel EEG to emotion classification [3].

Funded by NSF-DoD Award #1950207

Brain EEG signals are a control method for aircraft that is not widely used to enhance safety systems. This work presents a working EEG signal classification method using RNN-LSTM neural networks and validates the application in drones able to receive flying commands. A computer-aided detection system developed for the detection of focal EEG signals was proven successful in the detection of concealed nonlinear features of the EEG signals [4]. Robust motor-imagery-related EEG signal feature extraction and classification were achieved using LSTM neural networks, and when the results were compared to other classification techniques, the LSTM method outperformed in accuracy and had low standard deviation [5]. An LSTM deep learning algorithm developed for the recognition of emotions performed with an average accuracy of 86.36% was developed in [6]. An effective EEG-based emotion recognition method uses a preliminary feature attention LSTM network to decipher emotion-related electrodes [7]. Similarly, LSTM RNNs were used to correctly classify five sleep stages and based on the time domain input of a single-channel EEG signal 86.7% of a test sequence was correctly classified [8].

Time is a very important factor to consider when processing EEG signals as most of the research points to the fact that time dependent neural networks can outperform non-time-dependent neural networks in the classification of EEG signals [9]. Furthermore, Large existing datasets of EEG signals can be implemented with transfer learning for better EEG classification accuracy of simple thoughts of left- and right-hand movements [10].

B. EEG Signals

Electroencephalography (EEG) is a non-intrusive method to measure electrical brain activity from electrodes resting on the scalp. Electroencephalograms have been used to diagnose and recognize brain-related medical conditions such as epilepsy and strokes [14]. During EEG, electrical activity from a person's brain is measured using electrodes that rest externally on the head. EEG signal measurement are possible by measuring a potential difference between a region on the skull and an earlobe. A depiction of the internationally accepted EEG electrode designations can be seen in Fig. 1.

EEGs are designed to detect general brain activity from desired brain sections, not specific neurons or their groups. Though more precise brain activity measurement methods exist to target specific neuron groups in the brain, they involve surgical implants. Nevertheless, a general understanding of EEG exists, categorizing signals in four main frequency ranges when measured in human brains. Delta is the range from 0.1 Hz to 4 Hz and is the highest amplitude wave. It is found in all sleep stages and represents the brain's gray matter. Theta range is from 3.5 to 8 Hz and relates to subconscious activity. It is abnormal for adults to have these signals, but they are expected for children under 13. Finally, the alpha range is between 8 and 13 Hz and represents white matter in the brain. It acts as a connection between the conscious and subconscious mind. The Beta range is from 13 to 30 Hz and is associated with actions and thinking. It is connected to the five senses and occurs in conscious states when problem-solving, talking, decision-making, or making judgements. Finally, the Gamma range is from 30 to 100Hz and is linked to perception and consciousness

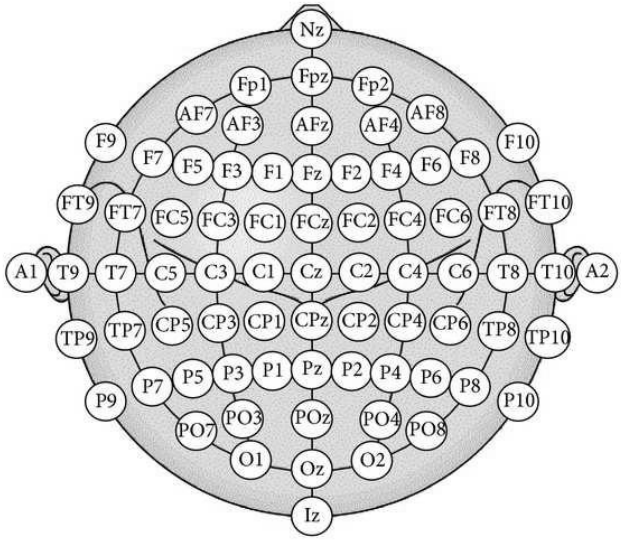

Fig. 1. 10-20 EEG Electrode Placement Designations [8]

(30-70 Hz) area. It connects senses and memory and occurs during hyper-alertness and when sensory inputs are integrated [13].

C. Existing Drone Control Methods

Manual human flight control has been the most prevalent method of aerial control throughout human flight history. The intricate details and problem-solving associated with maintaining a safe flight were trusted with highly trained pilots with minimal computational assistance. For commercial air travel, this is strictly still the standard even though companies are successfully developing and testing autonomous aircraft. As flight dynamics became better understood, technologies like autopilot became prevalent in commercial aircraft. These technologies added electronic actuators or an added layer to existing electronically actuated control systems to let a computer send commands based on sensor readings.

Current control developments and technology selling points are polarized in intelligent flight systems, fully allowing autonomous control, or disregarding it as unsafe for its application. Systems not adhering to either ultimatum could receive more approval than a fully autonomous control system while outperforming a human in speed and accuracy. This work's classification framework can be utilized in systems on the more manual side of the spectrum, as pilots must make an input to receive the desired output. While facial gestures are more suitable for this type of manual control, pattern recognition techniques presented are expected to have similar results in detecting other facial or mental states of the pilot.

D. Machine Learning Classification

Established EEG applications have mostly been limited to clinical settings where patients are screened for EEG signal patterns indicating a limited number of diseases. With today's improved computational power and neural network designs

relating EEG with emotions, movements, and visualizations is possible. EEG patterns also differ according to cognitive behavior and can act as an indicator of specific cognitive gestures.

Artificial neural networks (ANNs) are a subset of a machine learning system involving its underlying node-layered algorithms that are inspired by the human brain. These networks replicate the method in which biological neurons transmit signals to one another. First, nodes are connected and possess a specific threshold. It begins with inputs, follows with layers of nodes, and ends with an output. If any individual node surpasses its threshold, it is activated, and data is transferred to the next layer of nodes. Once all layers of nodes have been completed, outputs are determined, given the path of activated nodes.

ANNs are not limited to this simplified computing process; many modified versions of the same idea are designed for specific uses. For example, RNNs are modified ANNs that maintain history information from initial hidden layers and add them to current hidden layers during computation. This time-dependent looping constraint separates the normal ANNs from RNNs. Because of their design, RNNs are very good at feature recognition in instances where information is received over time and patterns exist relative to time. LSTM neural networks are a particular type of RNN known for their advantages in effective recurrent neural network implementation. Its design was motivated by the issues present in simple RNNs. For example, RNN issues involved failure to learn in the presence of time lags greater than 5-10 discrete steps [11]. LSTM cells mostly employ sigmoid and tangent functions.

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{1}$$

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

$$i_t = \sigma(w_{xi}^T x_{(t)} + w_{hi}^T h_{(t-1)} + b_i) \tag{3}$$

$$f_t = \sigma(w_{xf}^T x_{(t)} + w_{hf}^T h_{(t-1)} + b_f) \tag{4}$$

$$o_t = \sigma(w_{xo}^T x_{(t)} + w_{ho}^T h_{(t-1)} + b_o) \tag{5}$$

$$g_t = tanh(w_{xg}^T x_{(t)} + w_{hg}^T h_{(t-1)} + b_g) \tag{6}$$

$$c_t = f_t \otimes c_{(t-1)} + i_t \otimes g_{(t)} \tag{7}$$

$$h_t = o_t \otimes tanh\, c_{(t)} \tag{8}$$

Where:

$\sigma = sigmoid$

$tanh = tanget\ activation\ function$

$i = input\ gate$

$f = forget\ gate$

$o = output\ gate$

$c = intermidate\ gate$

$h = cell\ memory$

$t = time\ step$

$T = length\ of\ window$

$w = layer\ weight\ representing\ input\ x$

$b = threshold\ of\ the\ output\ gate$

LSTMs work well to allow the effective use of an RNN architecture. However, they involve considerable time-related complexity, which is only sometimes needed for a proper RNN performance. GRUs solve this problem by using a less complex function inspired by LSTM RNNs while ensuring comparable performance. The following are the ways in which GRU differs from LSTM:

$$r_t = \sigma(w_{xr}^T x_{(t)} + w_{hr}^T o_{(t-1)} + b_r) \tag{8}$$

$$z_t = \sigma(w_{xz}^T x_{(t)} + w_o^T zo + b_z) \tag{9}$$

$$o_t = z_t \otimes o_{(t-1)} + (1 - z_t) \otimes \tilde{o}_{(t)} \tag{10}$$

Where:

$r_t = reset\ gate$

$z_t = update\ gate$

$o_t = output\ gate$

$\otimes = element - wise\ multiplication$

$t = time\ step$

$T = length\ of\ window$

$w = layer\ weight\ representing\ input\ x$

$b = threshold\ of\ the\ output\ gate$

A simple diagram comparing the node architectures of an RNN, LSTM, and GRU can be seen in Fig. 2. The comparison in performance will be evaluated to better understand the EEG and EMG time-related complexity and give insight into the design of such classification systems.

II. METHODOLOGY

A. EEG Data Acquisition

EEG data for all experiments were taken using a dry electrode headband equipped with three electrodes and two reference electrode clips for both ears. The OpenBCI device measured the electrode's voltages and sampled data from this electrode at a sampling frequency of 250 Hz. The samples were then wirelessly transferred to the OpenBCI software, which applied a notch filter and a band pass filter to the received signal in real time. Three locations on the subject's forehead (Fp1, Fpz, and Fp2) were measured using electrodes in relation to the voltages of both earlobes. Depicted in Fig. 3 is the EEG headband described with each electrode’s placement. For each

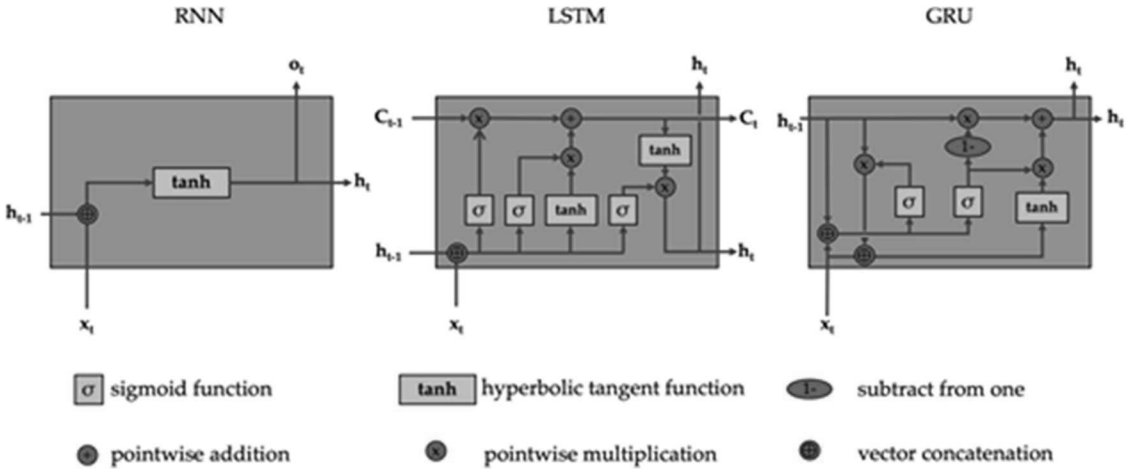

Fig. 2. Simple RNN, LSTM, and GRU Structure [13]

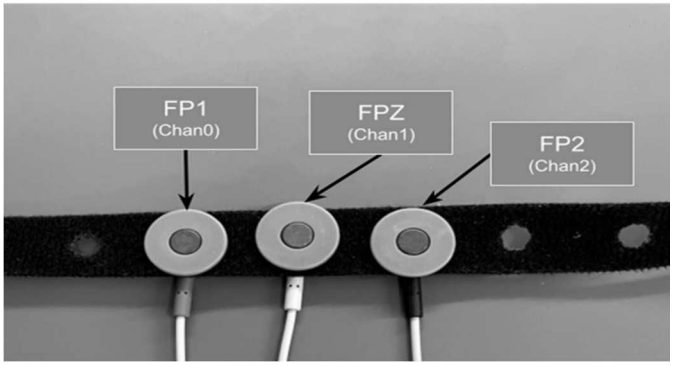

Fig. 3. EEG Headband and Electrode 10-20 Designations

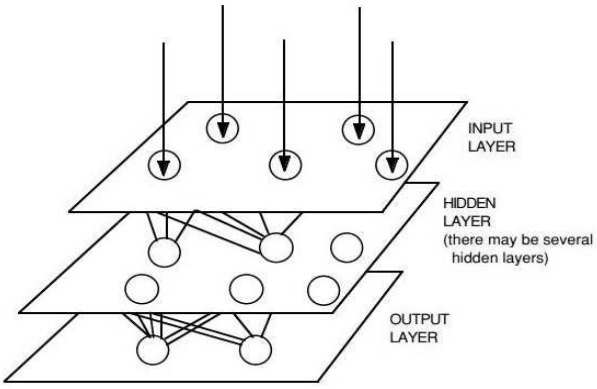

Fig. 4. Simple Neural Network Diagram [12]

gesture, the subject wearing the headset clenched their teeth firmly. The Blinking gesture was performed with two second gesture followed by a four second rest in between gestures. The subjects had to perform the gesture repetitively and rhythmically over an approximate interval of thirty seconds. The thirty-second recordings comprises of a two second gesture being performed followed by a four seconds rest, and that is repeated for the entire 30 seconds span . These gestures are chosen based on a visual distinction of EEG wave patterns when performing them. Other gestures are possible with the platform developed, however attention must be given to whether or not brain activity can be recognized at the electrode positions or with EEG. Each gesture's three voltage channels were recorded at 3Hz for their spectral transformations done with a Fast Fourier Transform on MATLAB.

B. Feature Extraction

A MATLAB script was developed to automate the dataset labeling process, and a baseline was created by finding the average resting signal over a period of 3 minutes, figure 9 shows this average data for two participants. This baseline FFT was compared to incoming FFT datasets each containing one action done repetitively with resting performed in between. The difference between gesture and baseline frequencies was established using three different error methods: root mean squared error (RSME), mean average error (MAE), and mean squared error (MSE). Incoming datasets were automatically labeled as the dataset's action or rest for network training by seeing if a data point's error relative to the baseline was large enough. The results for the feature extraction process can be seen in Fig. 5,6,7, and 8.

$$MSE = \sqrt{\frac{\sum_{i=1}^{N} ||y_i - \hat{y}_i||^2}{N}} \tag{11}$$

$$MAE = \frac{\sum_{i=1}^{N} |y_i - x_i|^2}{N} \tag{12}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 \tag{13}$$

C. LSTM and GRU Algorithm

In this stage, the GRU network's parameters were adjusted to employ a mini-batch size of 300, with each ideal parameter being discovered experimentally. To avoid overfitting, a drop frequency of 0.5 was employed. The schedule for the learning rate was set to piecewise, and the learning rate was set at 0.00099. Both the validation frequency and the validation patience were set to 30. 90 Maximum Epochs was chosen.

Additionally, a mini-batch size of 60 was used for the LSTM network, to avoid overfitting, a drop frequency of 0.5 was employed. The schedule for the learning rate was set to piecewise, and the learning rate was set at 0.000999. Both the validation frequency and the validation patience were set to 25, and 28 respectively. Maximum Epochs was set to 300.

III. Results

A. Datasets

Two volunteers provided the data set for this experiment, which was then randomly separated into training, testing, and validation sets. The total number of samples used and the data distribution for training, testing, and validation are shown in Table 1.

B. Discussion

As seen in Fig. 5, 6, 7, and 8, the feature extraction process's output completely distinguishes between an actual gesture being done and the rest of the data. Ignore-rest classification was used as an outlier since it was observed that certain spikes are present in Fig. 9 when compared to results from a single user in Fig. 10 when results from averaging rest data over ten samples for two groups of users. The two separate networks LSTM and GRU were trained using the categorized dataset. Although there are some variations, both networks exhibit adequate overall accuracy, with LSTM showing an accuracy of 78.6% and GRU to be 81.82% as seen in the confusion matrices in Fig. 11 and Fig. 12 for LSTM and GRU, respectively.

1. For predicted class, the outlier, ignore-rest for GRU will see around 82.6% of data classified as outliers, implying certain spikes different from rest and up to 73.5% for LSTM. This means that when user wear the EEG headband and are in rest position, and certain gestures are made unintentionally, these datasets will be ignored with LSTM performing better. For true classification for both networks, the percentage of true classification was somewhat similar, which was 78.1% correct classification with a 21.9% misclassification with a 6.8% misclassification for the LSTM network.
2. For predicted class, the outlier, ignore-rest for GRU will see around 82.6% of data classified as outliers, implying certain spikes different from rest and up to 73.5% for LSTM. This means that when user wear the EEG headband and are in rest position, and certain gestures are made unintentionally, these datasets will be ignored with LSTM performing better. For true classification for both networks, the percentage of true classification was somewhat similar, which was 78.1% correct classification with a 21.9% misclassification for the GRU network, and 93.2% correct classification with a 6.8% misclassification for the LSTM network.
3. Actual rest data classification is seen to perform better with GRU with up to 84.4% being classified correct as opposed to 94.3% accurate classification for the LSTM model when looking at predicted classification. For true classification for both networks, the percentage of true classification was somewhat similar, which was 79.1% correct classification with a 20.9% misclassification for the GRU network, and a lower 75.9% correct classification with a 24.1% misclassification for the LSTM network.
4. When comparing the predicted classification results for the bite gesture for the two networks, the GRU network had a classification accuracy of 80.5% and a misclassification rate of 19.5%, LSTM had a classification accuracy of 81.9% and a misclassification rate of 18.1%. In contrast, while examining real classification, it was found that biting gesture classification was correctly classified 88.9% of the time for GRU classification was incorrect 11.1% of the time and LSTM network classification was incorrect 9.6% of the time with correct classification of 90.4%.
5. For predicted classification for blink gesture for both networks, 87.5% of the dataset were correctly classified with a misclassification of 12.5% for the GRU network and 88.5.9% correct classification for LSTM and a misclassification of 11.5% observed. On the other hand, the results from true classification indicated a 88.8% correct classification with a 13.2% misclassification for GRU network, and 95.9% correct classification for LSTM network with a 4.1% misclassification for the same network.
6. Finally, the predicted classification for the gesture of raising the eyebrows was observed for the GRU network to be 87.2% correct with a misclassification of 17.4% and for the LSTM network to be 95.3% correct with a misclassification of 4.7%. While the GRU network had a correct classification rate of 99.2% and a misclassification rate of 0.8%, the LSTM network had a correct classification rate of

Table 1: Dataset Distribution

	Total data	Training data	Test data	Validation data	categories
Data1	1,250	70%	15%	15%	5

Table 2: Model for LSTM Network Algorithm

Input: the sequence of gesture FFTs
Output: categorization of gesture
Learning rate, batch size
Drop out Layer (0.5)
LSTM (hidden units, batch size)
Drop out Layer (0.5)
Softmax
Return Output

Table 3: Model for GRU Network Algorithm

Input: the sequence of gesture FFTs
Output: categorization of gesture
Learning rate, batch size
Drop out Layer (0.5)
GRU (hidden units, batch size)
Drop out Layer (0.5)
Softmax
Return Output

99.2% and a misclassification rate of 0.8% according to the results from the true classification.

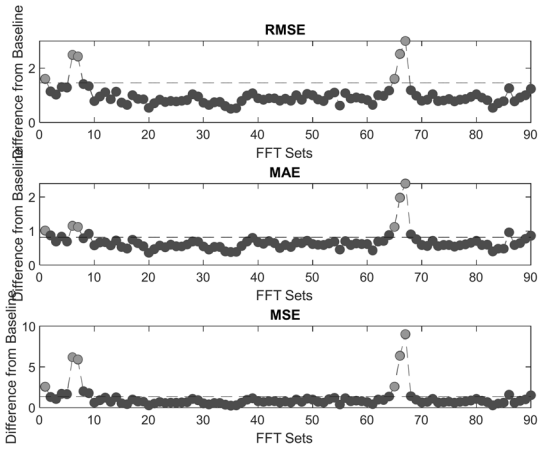

Fig. 5. Ignore Rest Pre-Training Recognition

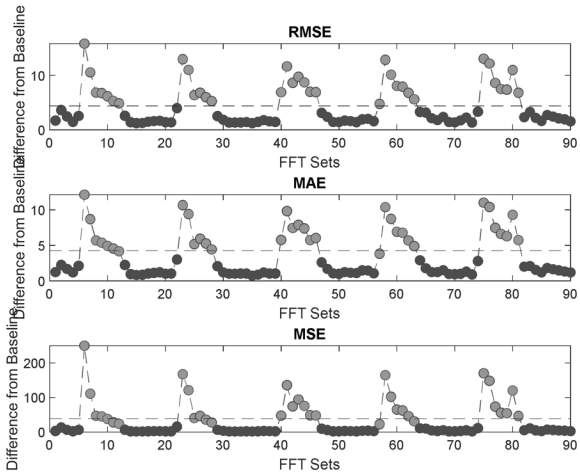

Fig. 6. Bite Pre-Training Recognition

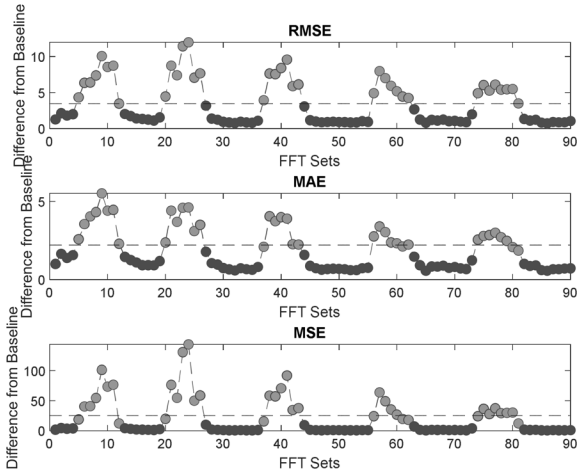

Fig. 7. Blink Pre-Training Recognition

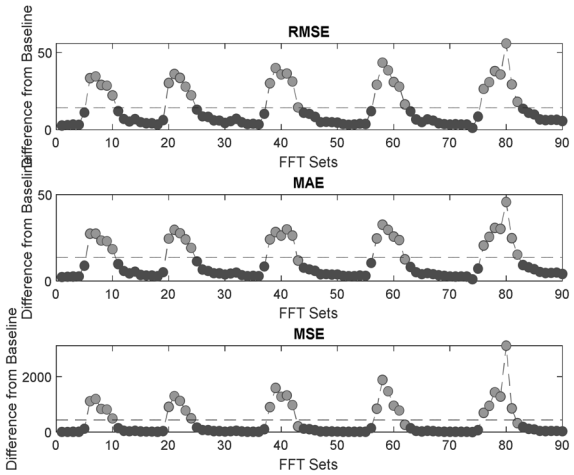

Fig. 8. Raise Eyebrows Pre-Training Recognition

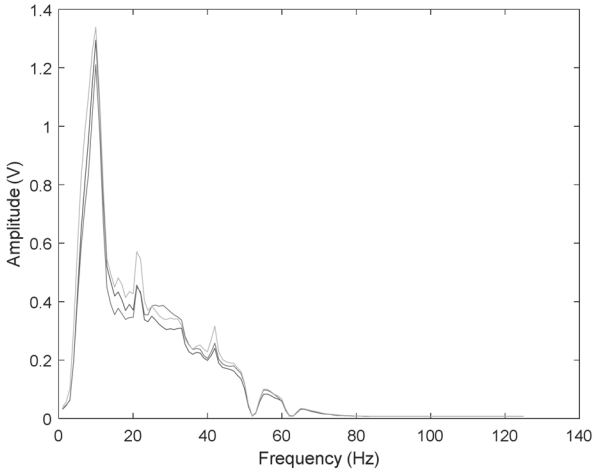

Fig. 9. Average of 10 rest data with little distortion from two participant

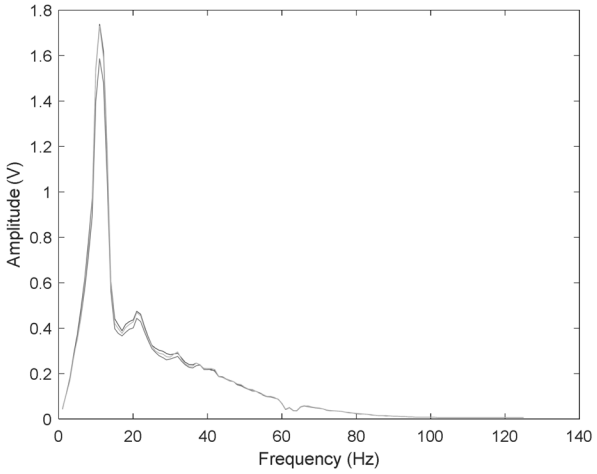

Fig. 10. Average of 20 rest data with lot of distortion for one participants

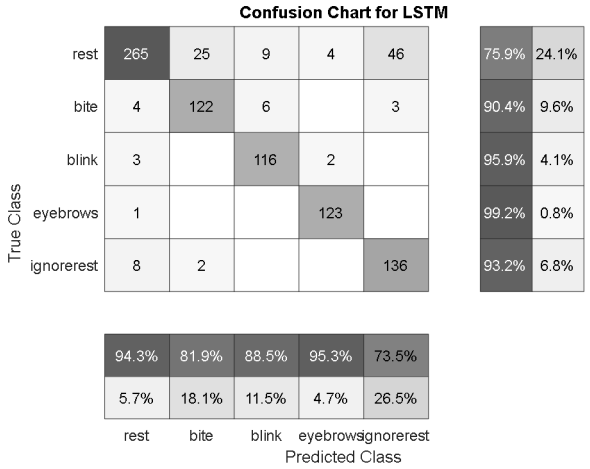

Fig. 11. LSTM Classification Confusion Matrix

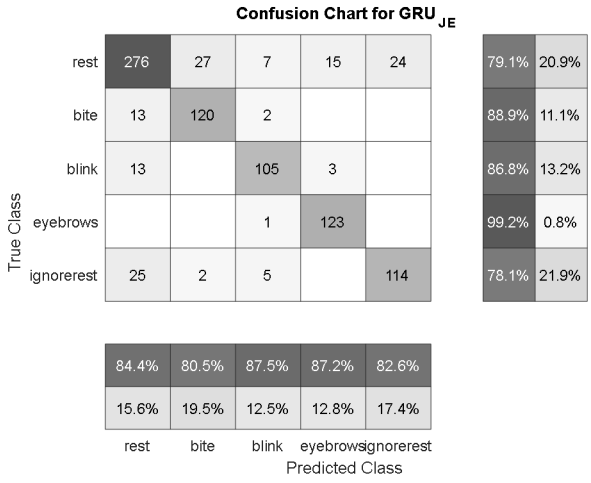

Fig. 12. GRU Classification Confusion Matrix

IV. Conclusion

A facial movement classification algorithm is possible when using recurrent neural networks to extract and classify based on spectral EEG and EMG data received from frontal lobe electrodes. The facial movements used in our experiments were raising eyebrows, biting, accidental gestures, and resting. Out of the utilized recurrent neural networks, LSTM and GRU, LSTM performed better. LSTM superiority shows that EEG/EMG signals received from these facial movements involve time-related complexity, making it likely worth the additional computational load of LSTM to better classify incoming spectral EEG data. Further work in this project includes adding other mental states and actions, RNN classification time response analysis, RNN classification process predictability, and further reliability testing.

ACKNNOWELEGEMENT

This work was possible through the technical and financial support of Georgia Southern University's Aerospace Engineering Research Experience for Undergraduates (REU), PI Dr. Valentin Soloiu. We appreciate the support and mentoring of the electrical and mechanical engineering departments at Georgia Southern.

REFERENCES

[1] Teplan M, "Fundamentals of EEG Signal Processing," *EEG Signal Processing and Machine Learning*, pp. 77–113, 2021.

[2] E. Habibzadeh Tonekabony Shad, M. Molinas and T. Ytterdal, "Impedance and Noise of Passive and Active Dry EEG Electrodes: A Review," in *IEEE Sensors Journal*, vol. 20, no. 24, pp. 14565-14577, 15 Dec.15, 2020, doi: 10.1109/JSEN.2020.3012394.

[3] R. Romo-Vazquez, R. Ranta, V. Louis-Dorr and D. Maquin, "EEG Ocular Artefacts and Noise Removal," *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 5445-5448, doi: 10.1109/IEMBS.2007.4353577.

[4] U. R. Acharya, Y. Hagiwara, S. N. Deshpande, S. Suren, J. E. Koh, S. L. Oh, N. Arunkumar, E. J. Ciaccio, and C. M. Lim, "Characterization of focal EEG signals: A Review," *Future Generation Computer Systems*, vol. 91, pp. 290–299, 2019.

[5] P. Wang, A. Jiang, X. Liu, J. Shang and L. Zhang, "LSTM-Based EEG Classification in Motor Imagery Tasks," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 11, pp. 2086-2095, Nov. 2018, doi: 10.1109/TNSRE.2018.2876129.

[6] S. Alhagry, A. Aly, and R. A., "Emotion recognition based on EEG using LSTM recurrent neural network," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, 2017.

[7] X. Du, C. Ma, G. Zhang, J. Li, Y.-K. Lai, G. Zhao, X. Deng, Y.-J. Liu, and H. Wang, "An efficient LSTM network for emotion recognition from multichannel EEG signals," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1528–1540, 2022.

[8] G. Korats, S. Le Cam, R. Ranta, and M. Hamid, "Applying ICA in EEG: Choice of the window length and of the decorrelation method," *Biomedical Engineering Systems and Technologies*, pp. 269–286, 2013.

[9] E. Haselsteiner and G. Pfurtscheller, "Using time-dependent neural networks for EEG classification," in *IEEE Transactions on Rehabilitation Engineering,* vol. 8, no. 4, pp. 457-463, Dec. 2000, doi: 10.1109/86.895948.

[10] W. Tu and S. Sun, "A subject transfer framework for EEG Classification," *Neurocomputing*, vol. 82, pp. 109–116, 2012.

[11] F. Altche and A. de La Fortelle, "An LSTM network for highway trajectory prediction," *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017.

[12] P. P., J. Gago, and M. Landi, "Artificial Neural Networks technology to model and predict plant biology process," *Artificial Neural Networks - Methodological Advances and Biomedical Applications*, 2011.

[13] M. Sit, B. Demiray, Z. Xiang, G. Ewing, Y. Sermet, and I. Demir, "A comprehensive review of deep learning applications in hydrology and Water Resources," https://doi.org/10.48550/arXiv.2007.12269, 2020.

[14] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, "Brain computer interfacing: Applications and challenges," *Egyptian Informatics Journal*, vol. 16, no. 2, pp. 213–230, 2015.