Evaluating Deep Learning EEG-Based Anxiety Classification in Adolescents with Autism for Breathing Entrainment BCI


¹ The Nueva School, San Mateo, CA 94033, USA
{bripenc, avisund, seacho}@nuevaschool.org
² Muvik Labs, LLC, Locust Valley, NY 11560, USA
³ Center for Computer Research in Music and Acoustics, Stanford University, Stanford, CA 94305, USA
vic@muviklabs.io
⁴ Department of Anatomy and Neurobiology, Laboratory of Cerebral Dynamics, Boston University School of Medicine, Boston, MA 02118, USA
⁵ Cognitive Neuroscience and Information Technology Research Program, Open University of Catalonia (UOC), Barcelona, Spain
⁶ Causal Brain Dynamics, Plasticity and Rehabilitation Team, Frontlab, Brain and Spine Institute, ICM, CNRS UMR 7225, Paris, France
{antoni.valerocabre, adrien.martel}@icm-institute.org

Abstract. Anxiety is one of the most common comorbidities in youth with autism spectrum disorder (ASD), severely limiting academic opportunities and overall quality of life. In the present study, we compared several machine learning classifiers, namely support vector machine (SVM) and deep learning methods, in order to evaluate the feasibility of an EEG-based BCI for the real-time assessment and mitigation of anxiety through a closed-loop adaptation of respiration entrainment. We trained a total of eleven subject-dependent models—four with conventional BCI methods and seven with deep learning approaches—on the EEG of neurotypical (n=5) and ASD (n=8) participants performing alternating blocks of mental arithmetic stress induction, guided breathing, and unguided breathing. We propose a multiclass two-layer LSTM RNN deep learning classifier capable of identifying anxious states from ongoing EEG with an overall accuracy of 93.27%. Our study is the first to successfully apply an LSTM RNN classifier to identify anxious states from EEG in both neurotypical adolescents and adolescents with ASD.

Keywords: Anxiety · Autism · EEG · Deep learning · Breathing entrainment.
In addition to the core symptomatology of autism spectrum disorder (ASD), characterized by severe deficits in social communication skills and restricted or stereotyped behaviors and interests [1], individuals with ASD frequently suffer from co-occurring anxiety symptoms [2]. The comorbidity of anxiety and ASD in children and adolescents has been studied extensively with 40% to 85% of individuals with ASD aged 6 to 18 having at least one form of anxiety [3-4]. Importantly, anxiety has been identified as a significant hurdle in education for students [5], in particular for students with autism [6]. Given the frequency in which anxiety co-occurs in ASD, in conjunction with the long-term functional impairments, reduction in quality of life, and increased caregiver burden [7-10], a more comprehensive understanding of comorbidities in ASD as well as personalized intervention methods to relieve clinical symptoms of the disorder and improve emotional and physical well-being for individuals with ASD is needed.

Concurrently, a growing number of studies have demonstrated the efficacy of stand-alone meditation, relaxation and breathing practices for improving well-being, mental health and anxiety [12-14]. Although the underlying mechanisms are not yet fully understood, breathing practices such as ‘anatomically optimised respiration’, i.e. controlled, slow diaphragmatic breathing through the nose in the range of 6-10 breaths per minute, have been found to procure significant physiological benefits [15], reduce physiological and psychological stress [16] and even improve sustained attention performance [17]. Prior studies have shown that breath-control can address physiological correlates of anxiety, including heart rate variability (HRV) [18], a well-validated quantitative stress indicator [19]. Notably, breath-control has been found to significantly decrease test anxiety in students in an educational setting [20]. Moreover, a recent review by Zaccaro et al. [21] found that controlled slow breathing (<10 breaths per min) had a significant impact on autonomic nervous system activity, especially in the theta (3-7 Hz), alpha (8-14 Hz) and beta (15-30Hz) bands of the electroencephalogram (EEG), linked to improved cognitive performance during attentional and executive functions [22]. Although these findings, taken together, speak to the promise of using controlled slow breathing as a simple, low-cost and non-pharmacologic intervention [16] to mitigate anxiety, optimized efficacy hinges on assessing an individual’s current level of anxiety and ideal respiration parameters in real-time.

Although cognitive or affective states such as anxiety are not directly observable externally nor reliably measurable through behavioral measures or subjective reports, developments in EEG-based brain-computer interfaces (BCIs) have increasingly permitted the continuous and real-time monitoring of mental states. Neuroadaptive technologies and passive brain-computer interfaces (pBCIs) aim at intelligent forms of adaptation in response to cognitive state assessments [23-24]. The field of EEG-based BCIs has blossomed in recent years, largely on account of EEG’s high temporal resolution, non-invasiveness, relatively low
cost, and novel advances in the effectiveness and usability of acquisition systems [25-26]. While BCIs have historically been employed in the context of assistive technologies for severely impaired individuals [27-28], pBCIs have mainly been aimed at developing adaptive automation for real-world applications [29-31].

The central challenge of EEG-based pBCIs is to account for the high inter- and intra-individual variability of neurophysiological signals exhibited under particular cognitive states [26]. However, by averaging over a large enough number of samples, it is possible to distill sufficiently specific brain activity patterns and train a machine learning classifier to learn to discern these patterns in real-time [32-33]. This approach has already been successfully applied to monitor several cognitive states such as workload [34-36], vigilance [37-39], and fatigue [40-41].

Neurofeedback involves monitoring a user’s mental state with EEG and providing feedback through a variety of modalities (visual, audio, tactile, etc.) in order to modulate particular biomarkers [42]. In conjunction with breath control, neurofeedback training has been shown to be a promising mitigatory tool for anxiety. For example, White et al. [43] demonstrated that breathing-based visual neurofeedback reduces symptoms in patients with anxiety and depression, while acoustically-mediated deep breathing neurofeedback was shown by Crivelli et al. [44] to diminish long-term stress and anxiety levels in young adults. The first step towards an EEG-based BCI able to monitor anxiety levels, to identify an individual’s optimal breathing patterns, and to adapt breathing entrainment parameters in real-time, is to determine whether anxiety can be classified on the basis of ongoing EEG data.

Classification algorithms are key elements of any EEG-based BCI’s ability to recognize users’ EEG patterns and associated cognitive states. Among the large diversity of existing architectures and types of classifiers (for a review see Lotte et al. 2018 [45] and Roy et al. 2019 [46]), deep learning methods have recently emerged as methods of analysis that can consider neurophysiological data in its entirety, including the time domain [46]. Convolutional neural networks are the most widely used deep learning algorithms in EEG analysis [47], and have been shown to be effective in emotion detection [48-49] and anxiety classification [50] in particular. Further, deep learning with convolutional neural networks (CNNs) have recently been shown to outperform the widely-used filter bank common spatial pattern (FBCSP) algorithm [51] by extracting increasingly more complex features of the data [52]. Accordingly, we aimed at comparing several classifiers previously used in EEG-based BCIs for the classification of different states of anxiety in ASD and neurotypical adolescents. We employed classical machine learning methods, specifically support vector machines (SVMs) combined with FBCSP, which have been successfully applied to detect a wide range of covert cognitive and emotional states [45], including mental stress detection [53-55]. Although classical classifiers present several drawbacks compared to deep learning (e.g. elaborate feature extraction and extensive prior knowledge about the dataset [56, 47]), SVMs remain a useful benchmark against which deep learning methods can be evaluated.
For deep learning methods we selected EEGNet, a recently developed compact CNN for EEG-based BCIs [57], as well as the Deep ConvNet and the Shallow ConvNet developed by Schirrmeister et al. [51]. Moreover, we also applied long short-term memory recurrent neural networks (LSTM RNNs), which are a type of neural net with the ability to “remember” long-term dependencies far better than traditional RNNs, without the loss of short-term memory [58], and enable robust analysis of temporal trends in EEG data [59]. LSTM RNNs have also shown high accuracy in emotion detection [60], with LSTM RNN architectures performing better than CNN architectures on DEAP, a major EEG emotion analysis dataset [47]. Hybrid deep neural networks combining both LSTM RNN and CNN architectures have also shown promising results on DEAP [61-62]. Building upon these recent advancements, we implemented an LSTM RNN and a hybrid long short term memory fully convolutional network (LSTM-FCN) [63] to classify anxious states from EEG.

The primary purpose of the present study is to assess the feasibility of real-time anxiety detection based on EEG signals and the identification of a robust classifier for prospective use in a pBCI able to identify the optimal breathing patterns and alleviate anxiety in students with and without ASD. To our knowledge, this is the first study to examine the efficacy of deep learning-based EEG anxiety classifiers in comparison to classical methods. In addition, ours is the first attempt of EEG-based anxiety classification for both adolescents with autism and neurotypical adolescents.

2 Methods

2.1 Participants and Data Acquisition

Eight students (1 female M: 15.13 SD: 1.45) diagnosed with autism, designated as participants L1-L8, from Learning Farm Educational Resources based in Menlo Park, (California), and five students (1 female M: 16.6 SD: 0.55) with no known mental or neurological disorders, designated as participants T1-T5, from The Nueva School in San Mateo, (California), voluntarily enrolled in the study. Participants and their parents or legal guardians were informed extensively about the experiment and all gave written consent. The study was approved by an Institutional Review Board composed of an educator from Learning Farm Educational Resources, an administrator from The Nueva School, and a licensed mental health professional at The Nueva School.

Participants were seated in an isolated and dimly lit room at a viewing distance of approximately 70 cm of a 16” LCD monitor with a refresh rate of 60Hz. 16-channel EEG data was acquired at 125Hz using an OpenBCI system (Ag/AgCl coated electrodes + Cyton Board; www.openbci.com) placed according to the international 10-20 system (channels: ‘Fp1’, ‘Fp2’, ‘C3’, ‘C4’, ‘P7’, ‘P8’, ‘O1’, ‘O2’, ‘F7’, ‘F8’, ‘F3’, ‘F4’, ‘T7’, ‘T8’, ‘P3’, ‘P4’). Participants were fitted with passive noise-canceling headphones to isolate them from ambient noise and to interact with the stress and breath modulating interface. The audio-visual stimuli was designed in close collaboration with Muvik Labs.
EEG-Based Anxiety Classification for Breathing Entrainment BCI

The stimuli featured sequential trials of stressor, guided breathing, and unguided breathing sections (Figure 1). The stimuli were procedurally generated by Muvik Labs’ Augmented Sound™ engine to ensure timing precision and effectiveness through evidence-backed breathing interventions driven by principles of psychoacoustics and behavioral psychology [64].

Prior to the main procedure, participants were asked to complete the trait anxiety component of Spielberger’s State-Trait Anxiety Inventory for Children (STAI-C) [65], a well-validated state and trait anxiety screen used for typically developing youth that can also be accurately used to assess trait anxiety in children and adolescents with autism [66].

2.2 Stress Induction and Alleviation

Following an initial EEG baseline recording for 120 sec (‘Baseline’), participants performed a 25-min session featuring stress induction and breath modulation tasks consisting of four main blocks. Each block began with a stressor (‘Stressor’) featuring an augmented arithmetic number task, intensified by bright contrasting colors displaying numbers appearing sequentially, coupled with audible sonified timers mapped to rising pitches similar to Shepard tones (powered by Muvik Labs Augmented Sound™) [64], with a 90 second time constraint. Timed mental arithmetic has been widely used to induce stress [67-68]. The stress induction was followed by a period of breathing for 200 seconds. The first and third breathing periods had participants breathe at their own pace (‘Unguided Breathing’) while the second and fourth breathing periods presented participants with a custom-generated breathing entrainment system, guiding breath airflow in and out of lungs at a relaxing pace of around 6 breaths per minute [69-70] (‘Guided Breathing’) with both visual (i.e. growing/shrinking circle outlining the air flow volume of target respiration speed) and auditory guides (musical patterns featuring nature sounds that mimic the sound of inhalation and exhalation; Muvik Labs Augmented Sound™).

2.3 EEG Signal Processing and Training Data Selection

MNE [71], an open-source Python tool for EEG analysis, was employed to filter EEG data from all 16 channels. In preparation for classification analysis, EEG time-courses were high-pass filtered at 1 Hz to remove slow trends and subsequently low-pass filtered at 50Hz to remove line noise. The routine clinical bandwidth for EEG is from 0.5Hz to 50Hz [72]. However, significant sinusoidal drift was observed on the 0.5Hz - 1Hz interval and therefore the interval was excluded in the selected bandpass filter range.

The data of two participants were rejected from all analyses due to unusually high impedances at the time of recording, which was confirmed offline by visual inspection: participant L1 from the ASD group, and participant T3 from the neurotypical group. Preprocessing of the EEG data was kept to a minimum in order to mimic online conditions found in a real-time BCI scenario.
Fig. 1. Experimental design of the procedure. Participants performed four blocks, each consisting of a mental arithmetic task followed by an anxiety self-report, a period of rest, either guided breathing entrainment or unguided breathing, a second anxiety self-report and lastly another rest period.

For training sample preparation, a cropped training strategy was employed. Samples with a length of 1s and 5s were extracted per participant from the EEG recorded during the ‘Stressor’, ‘Guided Breathing’, ‘Unguided Breathing’, and ‘Baseline’ periods of the procedure and were assigned corresponding labels.

2.4 Neural Signal Classification

We performed classification analysis on the selected EEG training samples using an SVM model with FBCSP, three CNN models, three LSTM RNN models, and a hybrid LSTM-FCN model. While all deep learning models were multiclass (‘Stressor’, ‘Baseline’, ‘Guided Breathing’, and ‘Unguided Breathing’), the SVM classifiers were binary (‘Guided Breathing’ vs ‘Stressor’, ‘Unguided Breathing’ vs ‘Stressor’, ‘Unguided Breathing’ vs ‘Guided Breathing’, & ‘Baseline’ vs ‘Stressor’), as is conventional for the classification of multiple classes with SVMs [73]. We opted to avoid using calculated features as inputs in favor of an end-to-end learning method with filtered EEG signal value inputs from all 16 channels. In addition, as different EEG channels represent neural signals from different areas of the brain, we elected not to combine channel data in order to preserve spatial information.

For the FBCSP-SVM, the EEG recording was divided in the time domain into samples of 1s and partitioned in the frequency domain using 9 filter bank band-pass filters from 4Hz to 40Hz prior to feature extraction, which was achieved with the common spatial pattern (CSP) algorithm, i.e. a linear map maximizing the variance difference between two classes [74]. The binary SVM classifiers used a radial basis function kernel with a gamma value of 1/360 and a regularization parameter (C) of 1.6. We implemented a number of FBCSP-SVM variants, such as multiclass SVM (baseline, stressor, guided and unguided breathing) with polynomial or sigmoid kernels and a 5 sec EEG sample length; these models were not.
Fig. 2. Diagram of the model architectures for the (A) Deep ConvNet, (B) Shallow ConvNet, (C) LSTM RNN, (D) EEGNet and (E) LSTM-FCN. Note: the first grayed layer of the LSTM RNN was only implemented for the two- and three-layer LSTM while the second grayed layer is only applicable to the three-layer LSTM.
included in the comparison due to lower classification performance. The FBCSP-SVMs were implemented with the sklearn library in Python and for validation samples were apportioned at a ratio of 80:20 for train and test dataset.

The Deep ConvNet CNN architecture [52] is composed of 4 convolution-max-pooling blocks. The first block, with 25 2D temporal convolutional filters of size (1, 5), 25 2D spatial convolutional filters of size (1, 64), and a max pooling layer, was especially designed to process the EEG input. The subsequent convolution-max-pooling blocks each have a 2D convolutional layer and a max pooling layer, with 50, 100 and 200 convolutional filters per block, respectively. Finally, a 4 neuron dense layer with softmax activation produces the output (see Figure 2A). The Shallow ConvNet CNN architecture is a modification of the Deep ConvNet to mimic the transformations of FBCSP. The Shallow ConvNet retains the first convolution-max-pooling block of the Deep ConvNet, albeit with a larger kernel size of 13 for the temporal convolution layer. This block performs similar transformations as the bandpass filter and CSP spatial filtering algorithm of the FBCSP workflow. Following the convolution-max-pooling block, the architecture contains a squaring nonlinearity function, an average pooling layer, and a logarithm activation function [52]. A 4 neuron dense layer with softmax activation produces the output (see Figure 2B).

We trained three LSTM RNN models, with one, two and three LSTM layers, respectively. The three-layer LSTM RNN model consists of three LSTM layers, two dense hidden layers, and a dense output layer. The first LSTM layer, containing 50 neurons, receives the input. The second and third LSTM layer contain 40 neurons. The number of neurons in the LSTM layers was informed by the amount used by Alhagry et al. [60] and adjusted to prevent underfitting and overfitting to our EEG data. Following the third LSTM layer, we include a dropout layer [75] with a dropout rate of 0.5 to prevent overfitting. The first dense layer contains 20 neurons and uses a sigmoid activation function. Following the first dense layer, we include a dropout layer with a dropout rate of 0.5. The second dense layer consisted of 10 neurons and used a rectified linear unit (ReLU) as an activation function. The dense output layer of 4 neurons used softmax activation. The two-layer LSTM architecture is obtained by omitting one 40 neuron LSTM layer, and the one-layer LSTM architecture is obtained by omitting both 40 neuron LSTM layers (see Figure 2C).

The EEGNet CNN architecture [57] used is composed of 8 2D convolutional filters of size (1, 64), a Depthwise Convolution layer of size (16, 1) to learn multiple spatial filters for each temporal filter, a Separable Convolution layer of size (1, 16), and a 4 neuron dense layer with softmax activation (see Figure 2D). In the LSTM-FCN [63] architecture, EEG time series input is simultaneously fed into an LSTM block, composed of an 8 neuron LSTM layer and a dropout layer with rate of 0.8, and an FCN block composed of 128 1D temporal convolutional layers of size 8, 256 1D temporal convolutional layers of size 5, and 128 1D temporal convolutional layers of size 3. The outputs of the LSTM and FCN blocks are then concatenated and passed into a 4 neuron dense output layer with softmax activation (see Figure 2E).
All deep learning architectures were implemented with the Keras machine learning library in Python, and were trained over 1000 EEG epochs with a batch size of 200. While training, we implemented the Adam optimization algorithm [76] with a learning rate of 0.001 in place of the standard stochastic gradient descent (SGD) algorithm. During validation, EEG samples for deep learning were apportioned at a ratio of 70:30 to the train dataset and test dataset, respectively.

3 Results

The performance for the FBCSP-SVM classifiers of each participant are shown in Table 1. The average classification accuracy was highest for the binary classification of ‘Baseline’ vs ‘Stressor’ (87.88%) and lowest for ‘Unguided Breathing’ vs ‘Stressor’ (78.28%). That the baseline and unguided breathing conditions had the lowest level of demands imposed on participants, and yet classification against the stessor condition yielded both the highest and lowest accuracy, suggests that muscle activity did not bias the classifiers significantly.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Guided vs Stressor (%)</th>
<th>Baseline vs Stressor (%)</th>
<th>Unguided vs Guided (%)</th>
<th>Unguided vs Stressor (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>98.75</td>
<td>97.92</td>
<td>100.00</td>
<td>96.30</td>
</tr>
<tr>
<td>L3</td>
<td>80.00</td>
<td>95.83</td>
<td>75.00</td>
<td>80.56</td>
</tr>
<tr>
<td>L4</td>
<td>68.75</td>
<td>91.67</td>
<td>70.00</td>
<td>75.93</td>
</tr>
<tr>
<td>L5</td>
<td>81.25</td>
<td>91.67</td>
<td>87.50</td>
<td>77.78</td>
</tr>
<tr>
<td>L6</td>
<td>67.50</td>
<td>81.25</td>
<td>57.50</td>
<td>71.30</td>
</tr>
<tr>
<td>L7</td>
<td>87.50</td>
<td>81.25</td>
<td>97.50</td>
<td>90.74</td>
</tr>
<tr>
<td>L8</td>
<td>90.00</td>
<td>93.75</td>
<td>86.25</td>
<td>70.37</td>
</tr>
<tr>
<td>T1</td>
<td>90.00</td>
<td>93.75</td>
<td>76.25</td>
<td>71.30</td>
</tr>
<tr>
<td>T2</td>
<td>83.75</td>
<td>89.58</td>
<td>91.25</td>
<td>83.33</td>
</tr>
<tr>
<td>T4</td>
<td>72.50</td>
<td>79.17</td>
<td>73.75</td>
<td>69.44</td>
</tr>
<tr>
<td>T5</td>
<td>65.00</td>
<td>70.83</td>
<td>71.25</td>
<td>74.04</td>
</tr>
<tr>
<td>Average</td>
<td>80.45</td>
<td>87.88</td>
<td>80.57</td>
<td>78.28</td>
</tr>
</tbody>
</table>

The three CNN models, Deep ConvNet, Shallow ConvNet, and EEGNet, yielded average classification accuracies of 58.80%, 62.84%, and 61.18%, respectively (see Table 2). The LSTM-FCN yielded an average classification accuracy of 62.97% across all four classes. The two-layer LSTM RNN classifier yielded an average accuracy of 93.27% on the test data across all four classes, outperforming the 73.53% average accuracy of the one-layer LSTM RNN and the 76.06% average accuracy of the three-layer classifier (see Table 3).
It is important to note that due to the longer length of the unguided and guided breathing periods compared to the stressor and the baseline periods, more samples were extracted from the unguided and guided breathing periods, creating an unbalanced dataset. Although this can lead to issues since an unbalanced dataset can artificially inflate the accuracy metric, the two-layer LSTM RNN model used here demonstrated high class-wise sensitivity and specificity during validation (see Figure 3), leading us to the conclusion that the unbalanced dataset was not a cause for concern.

Table 2. Class-wise and overall accuracies for the Deep ConvNet, Shallow ConvNet, and EEGNet CNN classifiers.

<table>
<thead>
<tr>
<th>Average Accuracy (%)</th>
<th>Deep ConvNet (%)</th>
<th>Shallow ConvNet (%)</th>
<th>EEGNet (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stressor (%) 60.73</td>
<td>49.14</td>
<td>59.31</td>
<td></td>
</tr>
<tr>
<td>Unguided Breathing (%) 59.38</td>
<td>56.25</td>
<td>60.38</td>
<td></td>
</tr>
<tr>
<td>Guided Breathing (%) 53.76</td>
<td>81.07</td>
<td>59.99</td>
<td></td>
</tr>
<tr>
<td>Baseline (%) 61.34</td>
<td>64.90</td>
<td>61.18</td>
<td></td>
</tr>
<tr>
<td>Average (%) 58.80</td>
<td>62.84</td>
<td>60.21</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Class-wise and overall accuracies for the 1-Layer, 2-Layer, and 3-Layer LSTM RNN classifiers and the hybrid LSTM-FCN classifier.

<table>
<thead>
<tr>
<th>Average Accuracy (%)</th>
<th>1-Layer LSTM (%)</th>
<th>2-Layer LSTM (%)</th>
<th>3-Layer LSTM (%)</th>
<th>LSTM-FCN (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stressor (%) 63.95</td>
<td>90.82</td>
<td>74.75</td>
<td>57.19</td>
<td></td>
</tr>
<tr>
<td>Unguided Breathing (%) 62.89</td>
<td>91.19</td>
<td>70.43</td>
<td>57.52</td>
<td></td>
</tr>
<tr>
<td>Guided Breathing (%) 80.09</td>
<td>94.57</td>
<td>67.76</td>
<td>72.32</td>
<td></td>
</tr>
<tr>
<td>Baseline (%) 73.53</td>
<td>96.50</td>
<td>76.07</td>
<td>64.84</td>
<td></td>
</tr>
<tr>
<td>Average (%) 70.12</td>
<td>93.27</td>
<td>72.26</td>
<td>62.97</td>
<td></td>
</tr>
</tbody>
</table>

3.1 LSTM RNN Performance with Individual Variation

We were interested in investigating the relationship between the classification accuracy of the two-layer LSTM model (the best performing model) and pre-existing mental conditions. First, we wished to see if there was a significant difference between model accuracy for participants with autism and neurotypical participants. On average, the two-layer LSTM model accuracy for a participant with autism was 93.33%, while the model accuracy for a neurotypical participant was 93.15%. A Mann-Whitney U test was conducted to compare model accuracy
Fig. 3. A) 2-Layer LSTM RNN model confusion matrix. B) 2-Layer LSTM RNN model precision-recall curve.

Table 4. 2-Layer LSTM RNN classification accuracy and trait anxiety per participant.

<table>
<thead>
<tr>
<th>Participant</th>
<th>2-Layer LSTM RNN Classification Accuracy (%)</th>
<th>STAI-C Trait Anxiety Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>92.83</td>
<td>37</td>
</tr>
<tr>
<td>L3</td>
<td>92.83</td>
<td>32</td>
</tr>
<tr>
<td>L4</td>
<td>93.72</td>
<td>24</td>
</tr>
<tr>
<td>L5</td>
<td>93.69</td>
<td>30</td>
</tr>
<tr>
<td>L6</td>
<td>92.57</td>
<td>33</td>
</tr>
<tr>
<td>L7</td>
<td>93.72</td>
<td>32</td>
</tr>
<tr>
<td>L8</td>
<td>93.97</td>
<td>46</td>
</tr>
<tr>
<td>T1</td>
<td>92.83</td>
<td>25</td>
</tr>
<tr>
<td>T2</td>
<td>93.08</td>
<td>37</td>
</tr>
<tr>
<td>T4</td>
<td>93.47</td>
<td>42</td>
</tr>
<tr>
<td>T5</td>
<td>93.24</td>
<td>28</td>
</tr>
</tbody>
</table>
for the participants with autism and neurotypical participants and found no significant difference between model accuracy for the two groups (p=0.566), indicating that the two-layer LSTM model performed similarly regardless of whether the participant had autism. We also wished to understand whether an individual’s persistent (trait) anxiety can influence the performance of the two-layer LSTM RNN. We employed Spearman correlation to compare model accuracy and individual STAI-C trait anxiety scores (see Table 4); higher STAI-C scores indicate higher trait anxiety. The analysis yielded a Spearman’s rho of 0.0393, indicating virtually no correlation between trait anxiety and the two-layer LSTM RNN performance.

4 Discussion

To the best of our knowledge, in this study we propose for the first time a deep learning-based classifier for decoding stress-induced anxiety, a complex and covert mental state, from scalp EEG signals. Our results show that anxious states can be accurately assessed in adolescents with and without ASD as well as in adolescents with varying levels of baseline anxiety. We compared classification accuracy of 4 binary FBCSP-SVM models and 7 multiclass deep learning models. These classifiers were employed to classify the EEG recorded from ASD and neurotypical adolescents performing a task with periods of stress induction (‘Stressor’), resting state (‘Baseline’), guided breathing (‘Guided Breathing’) and unguided breathing (‘Unguided Breathing’), which were converted into four distinct classes, one “anxious” class and three “relaxed” classes. The best classification accuracy was achieved with the multiclass two-layer LSTM at 93.27%.

The 4 binary FBCSP-SVM classifiers performed as follows: 80.45% for ‘Guided Breathing’ vs ‘Stressor’, 87.88% for ‘Baseline’ vs ‘Stressor’, 80.57% for ‘Unguided Breathing’ vs ‘Guided Breathing’, and 78.28% for ‘Unguided Breathing’ vs ‘Stressor’. The FBCSP-SVM performed best when classifying between the pre-task onset resting state epoch (‘Baseline’) and the stress induction (‘Stressor’) conditions, which could be due to the rest period imposing the least, and the stressor condition the most, demands on the participants. Interestingly, the classifier performed relatively well for ‘Unguided Breathing’ vs ‘Guided Breathing’ classes, although these two conditions were similar in terms of stimuli and demands imposed on the participants. Despite our binary FBCSP-SVM classifiers reaching a satisfactory overall classification accuracy of around 82% across all 4 condition pairs, there are several trade-offs pertaining to the use of SVM when compared to deep learning. Although SVMs require less optimizing parameters, do not suffer from the problem of local minima, and are less computationally demanding than neural networks, they are constrained to a small number of features [77], even when these features are extracted by algorithms [78]. In addition, SVMs cannot consider a robust set of EEG timepoints, rendering them unable to examine the EEG time domain, which is a critical dimension for analyses [78]. Contrastingly, LSTMs are well able to handle temporal information, given their ability to choose to remember or discard information depending on
contextual information. Nonetheless, due to their low computational complexity, SVMs remain one of the most popular types of classifiers for EEG-based BCI, in particular for online scenarios. Notwithstanding that adaptive implementations of SVM have been found to be superior to their static counterparts, they often require fully retraining the classifier with new incoming data, resulting in a much higher computational complexity and thus a lack of online applicability [45]. Conversely, with deep learning methods, adaptation can be achieved by retraining the input layer with new incoming data. LSTMs in particular are inherently adaptive and thus well suited for real-time scenarios, as their predictions are conditioned by past input. In addition, unlike SVMs, deep learning networks can automatically adjust and optimize their parameters, essentially alleviating the need for feature extraction and requiring far less processing and prior knowledge regarding the original EEG dataset [56, 47]. Lastly, while some multiclass SVMs have been found to outperform neural networks [79], our attempts with multiclass FBCSP-SVMs produced inconsistent results with accuracies ranging between chance-level and 90%.

With regard to the multiclass deep learning models, the Deep ConvNet CNN performed with an overall accuracy of 58.80%, the Shallow ConvNet CNN with 62.84%, the EEGNet CNN with 61.18%, the LSTM-FCN with 62.97%, the one-layer LSTM with 73.53%, the two-layer LSTM with 93.27% and the three-layer LSTM with 72.26%. The high classification accuracies achieved with the LSTM architecture presumably is a result of its ability to learn time dependencies within the data. Indeed, the retention property is useful in mental state monitoring, as considering the past activations of the EEG can drastically improve the prediction of target variables and the brain activity patterns leading up to, and associated with, specific cognitive states. However, stacking LSTM layers renders the model susceptible to the vanishing gradient problem, in which network weights fail to update significantly over time and model training becomes stagnant [80], as well as prone to overfitting [81]. These phenomena that could explain the lower classification accuracy of the three-layer LSTM compared to the two-layer architecture.

There are some caveats to consider in the interpretation of our results. First, given that anxiety varies significantly with context and individual, and cannot therefore be induced reliably and equivalently across participants, we utilized mental stress induction via mental arithmetic as a proxy for anxiety. Second, our experiment was designed to induce anxiety as efficiently as possible and thus minimize the time under stress to avoid any undue strain on the participants. Conversely, more time was required for relaxation to set in and the breathing rate to normalize. As a result, the time for the mental arithmetic task and the guided or unguided breathing differed. Thus, learning models were trained on an unbalanced dataset, with more ‘Unguided Breathing’ and ‘Guided Breathing’ EEG samples than ‘Stressor’ and ‘Baseline’ samples, with the potential of artificially inflating model accuracy. However, this is unlikely to be a concern for the two-layer LSTM RNN model, which exhibited high sensitivity and specificity metrics across all classes. Lastly, it should be noted that a potential drawback of LSTM
RNNs, and of deep learning algorithms in general, is over-reliance upon large datasets. To this regard, the same classifications performed with smaller datasets including only 2 or 3 conditions led to poorer performance (data not shown). Indeed, one major drawback of deep learning is the need for large amounts of data, an issue we aim to remedy in a future study involving a much larger set of participants, both neurotypical and ASD, as well as a more diverse set of stress induction tasks. Given that we have identified a viable classifier for the monitoring of cognitive states related to anxiety, the goal of forthcoming studies will be to refine and validate the two-layer LSTM RNN deep learning model for prospective implementation in a personalized pBCI. Such a system will be capable of monitoring for periods of anxiety and hone in on an individual’s optimal respiration patterns by adapting the breathing entrainment parameters in a closed-loop manner.

To recapitulate, the goal of this study was to compare several learning models or classifiers on their ability to assess anxiety levels from EEG recordings performed on adolescent students in order to determine the feasibility of an EEG-based BCI capable of real-time identification and the mitigation of anxiety through optimised respiration entrainment. Of the different classifiers we compared, the two-layer LSTM yielded the highest classification accuracy (93.27%), opening new avenues of decoding covert mental states for BCI-based neuroadaptive applications to benefit youth with autism.

5 Declarations

5.1 Availability of data and materials

The informed consent forms signed by the study participants and their legal guardians assured anonymity and utmost confidentiality of participant data, so it is not possible to make the data publicly available at this time.

5.2 Competing interests

The authors declare that they have no competing interests.

5.3 Funding

The contributions of Dr. Adrien Martel and Prof. Valero-Cabrè have been supported by research grants ANR (French National Research Agency) *projet Générique ‘OSCILOSCOPUS’* and Flag-era-JTC-2017 CAUSALTOMICS to AV-C. Dr. Martel is currently supported by a Marie Curie Fellowship (EU project 898813-CLONESA-DLV-898813).
5.4 Authors’ contributions

BP, AS, and SC contributed equally to this work. AM, BP, AV-C, and VG conceived of the study idea. BP coordinated and managed the collaborations. AM supervised the project with the help of AV-C. AM and VG designed and implemented the experimental paradigm. BP, AS, SC, and AM carried out the experiment, developed the theory and performed the computations and analyses. SC processed the experimental data and implemented the machine learning algorithms with the close help of AS and the guidance of AM. AS took the lead in writing the manuscript under the supervision of AM and support from BP, SC, and AV-C. All authors provided critical feedback and helped shape the research, analysis and manuscript.

5.5 Acknowledgements

We would like to thank The Nueva School, Learning Farm Educational Resources, OpenBCI, and Muvik Labs for their kind and invaluable support. The authors would also like to thank the Naturalia & Biologia Foundation for financial assistance for traveling and attending meetings.

References


