6th International Conference on Advances in Artificial Intelligence,
21-23 October 2022, Birmingham, UK
http://www.icaai.org/

Miscellaneous EEG Preprocessing and Machine Learning for Pilots' Mental States Classification: Implications

IBRAHIM M. ALRESHIDI

Centre for Computational Engineering Sciences, Cranfield University, UK

Machine Learning and Data Analytics Laboratory, Digital Aviation Research and Technology Centre (DARTeC)

College of Computer Science and Engineering, University of Ha'il, Ha'il, Saudi Arabia

IRENE MOULITSAS

Centre for Computational Engineering Sciences, Cranfield University, UK

Machine Learning and Data Analytics Laboratory, Digital Aviation Research and Technology Centre (DARTeC)

KARL W. JENKINS

Centre for Computational Engineering Sciences, Cranfield University, UK

Higher cognitive process efforts may result in mental exhaustion, poor performance, and long-term health issues. An EEG-based methods for detecting a pilot's mental state have recently been created utilizing machine learning algorithms. EEG signals include a significant noise component, and these approaches either ignore this or use a random mix of preprocessing techniques to reduce noise. In the absence of uniform preprocessing procedures for cleaning, it would be impossible to compare the efficacy of machine learning models across research, even if they employ data obtained from the same experiment. In this study, we intend to evaluate how preprocessing approaches affect the performance of machine learning models. To do this, we concentrated on fundamental preprocessing techniques, such as a band-pass filter and independent component analysis. Using a publicly accessible actual physiological dataset gathered from a pilot who was exposed to a variety of mental events, we explore the influence of these preprocessing strategies on two machine learning models, SVMs and ANNs. Our findings indicate that the performance of the models is unaffected by preprocessing techniques. Moreover, our findings indicate that the models were able to anticipate the mental states from merged data collected in two environments. These findings demonstrate the necessity for a standardized methodological framework for the application of machine learning models to EEG inputs.

CCS CONCEPTS • Computing methodologies~Machine learning~Machine learning approaches~Kernel methods~Support vector machines • Computing methodologies~Machine learning~Machine learning approaches~Neural networks • Hardware~Communication hardware, interfaces and storage~Signal processing systems~Noise reduction

Additional Keywords and Phrases: EEG preprocessing, Mental states classification, Machine Learning

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 INTRODUCTION

Detecting the pilot's mental state is critical to ensuring the safety of the plane's flight path [1]. Performance measures, questionnaires, and neurophysiological methods such as brain activity have all been shown by researchers to be effective in detecting people's mental states [2]. In particular, brain activity measures using electroencephalography (EEG) have been shown to be the most effective method of identifying the mental state of pilots. This is due to the fact that EEG signals can capture brain activity with great temporal resolution [3]. Thus, improved machine learning models have been built to reliably diagnose mental states by capturing variance properties in EEG data. Support Vector Machines (SVMs) and Artificial Neuronal Networks (ANNs) are examples of such efforts [4].

However, neuroscientists have demonstrated that EEG signals are susceptible to noise. Numerous efforts have been made in the past to improve data preprocessing procedures to reduce noise, and standardized protocols for cleaning EEG data have been constructed.

The machine learning algorithms that utilize EEG signals do not, however, adhere to a typical data decontamination approach. Due to these anomalies in their data preparation, it is impossible to quantify the true impact of machine learning models. Second, even when utilizing the same experiment's data, it is impossible to compare the outcomes of various experiments. In light of these facts, there is still work to be done on the standardized artefact removal procedure to be used in machine learning [5–7].

Consequently, the purpose of this work is to address the following research question: "What are the impacts of various preprocessing procedures on the performance of machine learning models that use EEG data to classify the pilot's mental states?" We examine the influence of applying various preprocessing approaches to EEG data using SVM and ANN algorithms. The SVM and ANN models were trained with data collected from a pilot in a non-flight environment, a flight environment, and merged flight and non-flight environment data using a 5-fold cross-validation method. The training data of the SVM and ANN models consists of unprocessed data, filtered data, and data that has been filtered and eye blinks were removed from it.

The paper is organized as follows: An overview of mental states and related research studies can be found in the second section. In Section 3, the methods will be described. Section 4 presents the findings and a discussion. Section 5 summarizes and concludes.

2 BACKGROUND

A substantial influence on a pilot performance within a complex working system has been recognized as mental workload. Overload or underload can both reduce performance and have a direct impact on the efficiency and quality of the system's operation. To ensure a functioning and successful human-system collaboration, monitoring and detecting mental states has become a critical job [8]. Self-report methods such as questionnaires or assessments of performance on a secondary task have long been used to assess mental states [9]. Individuals must, however, be trained to understand the instrument used to report their mental workload [10]. Therefore, such strategies may increase the subject's workload. Thus, recent research has studied the feasibility of classifying mental states using EEG recordings of brain activity. Neuroscientists have discovered, however, that EEG signals are sensitive to noise interference (artefacts), which could be internal or external sources. External artefacts may come from instruments or the subject's body, and internal artefacts include ocular, muscle, and cardiovascular activities [11]. There is a chance that noise in the EEG signal will degrade the quality of the data and thus the precision of the models used in the analysis. As a result, removing artefacts is a key requirement before using the acquired EEG data [12].

2.1 Mental states classification

EEG signals could efficiently indicate the pilot's mental state. Due to its ability to gather accurate data representations of features, machine learning has recently been successfully applied to EEG analysis [13]. For EEG preprocessing, some studies applied noise reduction methodologies; however, the efficacy of each strategy on machine learning models for mental state classification has not been evaluated. Such models cannot be compared due to the lack of a consistent preprocessing framework.

In a simulation environment, Chaudhuri et al. [14] applied a band-pass filter to the EEG data to remove extraneous signals and the SVM algorithm to identify normal and fatigue states. As a result, their classification accuracy has increased by an average of 86%. The band-pass filter was also utilized in the investigation by Han et al. [15]. Nonetheless, the frequency range has been adjusted to a different value. In this work, the sampling frequency was adjusted between 0.1 and 50 Hz, and the ICA components pertaining to eye blinks and movements were eliminated. Using SVMs, k-Nearest Neighbors (k-NN), Logistic Regression (LR), Random Forest (RF), shrinkage Linear Discriminant Analysis (sLDA), and deep Convolutional Neural Network (CNN) classifiers, the preprocessed data has been utilized to detect four distinct mental states induced by four benchmark activities. The obtained classification accuracy ranged between 64% and 83%. A notch filter is an additional type of filter that has been applied. In their investigation, Binias et al. [16] used the EPOC+ headset, which contains an integrated digital 5th-order Sinc filter, notch filters at 50Hz and 60Hz, and a band-pass filter between 0.16 and 43 Hz. LDA, k-NN, SVMs, RF, and ANNs have categorized two mental states that were induced to distinguish between states of brain activity associated with idle but concentrated anticipation of a visual cue and a reaction to it. The average accuracy of the proposed models ranges from 67 to 78%. The Butterworth band-pass filter with a high-pass cut-off frequency of 0.5 Hz and a low-pass cut-off frequency of 50 Hz was used in [17], but the authors left ocular artefacts in their data. The filtered data was then incorporated into a two-stream neural network (TSNN) model for a three-class mental workload classification task [18]. The model's average degree of accuracy is 91.9 percent. According to published studies, using preprocessing methods developed by neurobiologists, machine learning researchers have attempted to reduce noise from their EEG data. However, there is no standard preprocessing approach that everyone follows. In particular, the band-pass filter approach, which appears to be the most popular tool, has been characterized in a variety of ways.

In addition, event participants classify the same dataset using various preprocessing methods. Using their dataset, Harrivel et al. [19,20] performed attention-related human performance-limiting states (AHPLS) classification. The authors examined frequency domain components between 0 and 40 Hz using the A Lomb-Scargle frequency transform during the artefact removal phase. This method of spectral analysis takes sample rate abnormalities into consideration. The authors have achieved 82% AHPLS classification accuracy with a Deep Neural Network (DNN) model. Harrivel et al. [20] generated 40 power spectral density (PSD) features per channel to represent the EEG frequency bands between 1 and 40 Hz. Using gradient boosting, RF, and SVM classifiers, they achieved 50 to 78% classification accuracy. The AHPLS dataset was also utilized for AHPLS classification in the Terwilliger et al. [21]. However, no preprocessing approaches to eliminate signal artefacts were utilized. The proposed ResNet Autoencoder model has been fed with the original data for AHPLS detection. They have undertaken an analysis to determine whether or not there is an event. In their study, the proposed model showed a low rate of false positives and false negatives.

Although neuroscientists propose various EEG preprocessing guidelines [6,22,23], they are rather general and not universally accepted. Therefore, it is up to the researchers to determine which noise-removal method will be the most effective. In addition, the process for some existing pipelines includes a visual inspection and hand labeling [24]. While these techniques can be highly beneficial for reducing signal noise, they have three limitations. First, these techniques are time-intensive, especially when working with huge datasets. Then, this can introduce bias into the analysis [25]. Lastly, they restrict the use of such pipelines in automated procedures.

Machine learning articles for EEG analysis do not take into account the effect of preprocessing processes, hence findings cannot be compared across studies. In order to determine the effect of preprocessing strategies on machine learning models, we conducted nine experiments employing three experimental preprocessing scenarios on EGG data collected from a pilot in three distinct settings. Using a 5-fold cross-validation method, we created and trained SVM and ANN models with non-flight environment data, flight environment data, and merged flight and non-flight environment data. The training data of the SVM and ANN models comprises unfiltered data, filtered data, and data from which eye blinks have been removed. The following section presents the used cases as well as a framework for removing EEG artefacts for machine learning model evaluation.

3 METHODOLOGY

3.1 Data Acquisition

The dataset was acquired from the website of NASA's open data portal. It contains experimental EEG data gathered from a pilot who was required to complete tasks in two environments. Included in the tasks were resting tasks, benchmark tasks meant to elicit AHPLS, and experimental flight situations. LaRC's Research Flight Deck and Cockpit Motion Facility were utilized for the collecting of data. A psychophysiological sensor was utilized to measure electroencephalography (EEG) signals using an Advanced Brain Monitoring X24 EEG System. The EEG system consists of 20 electrodes in the standard 10-20 format + POz (Fz, Cz, Pz, F3, F4, C3, C4, P3, P4, O1, O2, T5, T3, F7, Fp1, Fp2, F8, T4, and T6 with Linked Mastoids) with sampling rate of 256 Hz. The signals were captured in two distinct contexts, a motion-based flight simulator and a non-flying environment. The pilot did a complete flying simulation using the motion-based flight simulator (take off, flight and landing) [26]. The pilot conducted three benchmark activities outside of the flight simulation environment. The situations that the pilot encountered during the experiments were designed to elicit one of the three cognitive states listed below:

1- Channelized Attention (CA): the state of concentrating on a single task. The benchmarking process is triggered by having the pilot play a puzzle-based video game.

2- Diverted Attention (DA): the state of having one's attention diverted by decisions-related behaviors or thinking processes. This is accomplished by having the pilot conduct a display monitoring task while math problems appear intermittently and must be solved before returning to the monitoring task.

3- Startle/Surprise (SS): it is induced by showing the pilot jump-scare movie clips.

The EEG data was downloaded in CSV format. To perform fundamental and sophisticated preprocessing approaches, we used an open-source library (MNE-Python) and generated an appropriate object for continuous EEG data's core data structures (i.e., raw object) [27]. The core data structures object is initialized with the necessary fields of information, including a list of channel names and types, the standard montage naming schemes, and the sampling frequency. The EEG data were then segmented into trials of one second with no overlap. We gathered 872 samples (i.e., CA: 352, DA: 46, SS: 15, No Event (NE): 459) from the non-flying environment data and 4053 samples (i.e., CA: 432, DA: 52, SS: 30, NE: 3539) from the flight simulator environment data.

3.2 EEG Preprocessing

EEG preprocessing involves multiple strategies. Some can reduce data noise automatically, while others must be performed manually. The purpose of this work is to examine the impact of preprocessing procedures that can only be performed automatically, i.e., without human interaction. The benefit of an automatic processing analysis is that it eliminates the problem of subjective marking of artefacts by visual inspection [25]. Consequently, we study the impact of the two most prevalent preprocessing approaches, a band-pass filter and ICA. As a result, we have three experimental cases for each type of environmental data: non-flight environment data, flight environment data, and flight and non-flight environment data combined. Following is a description of the three experimental cases:

Case 1 - Unprocessed Data: The data have not been preprocessed. Figure 1 illustrates a 10 second window of the unprocessed EEG data collected from the pilot in a non-flight environment.

Case 2 - Band-pass Filter: In this instance, a band-pass filter (finite impulse response (FIR), 1-50 Hz) was utilized to minimize artefacts and increase signal-to-noise ratio (SNR) appears in Figure 1. Figure 2 illustrates the filtered EEG data.

Case 3 - Band-pass Filter + ICA: To extract the artefact components from the EEG signals in this instance, ICA was applied to the filtered signals from the previous phase (Case 2) using the Fastica approach [12]. After partitioning the multichannel EEG into ICs, eye blinks, as it can be seen from Figure 2 at the fifth, seventh and eighth seconds in channel Fp1 and Fp2 which are the channels near the eyes, were automatically recognized and rectified. As shown in Figure 3, eye blinks are removed from the EEG signals. Table 1 is a summary of our three preprocessing cases.

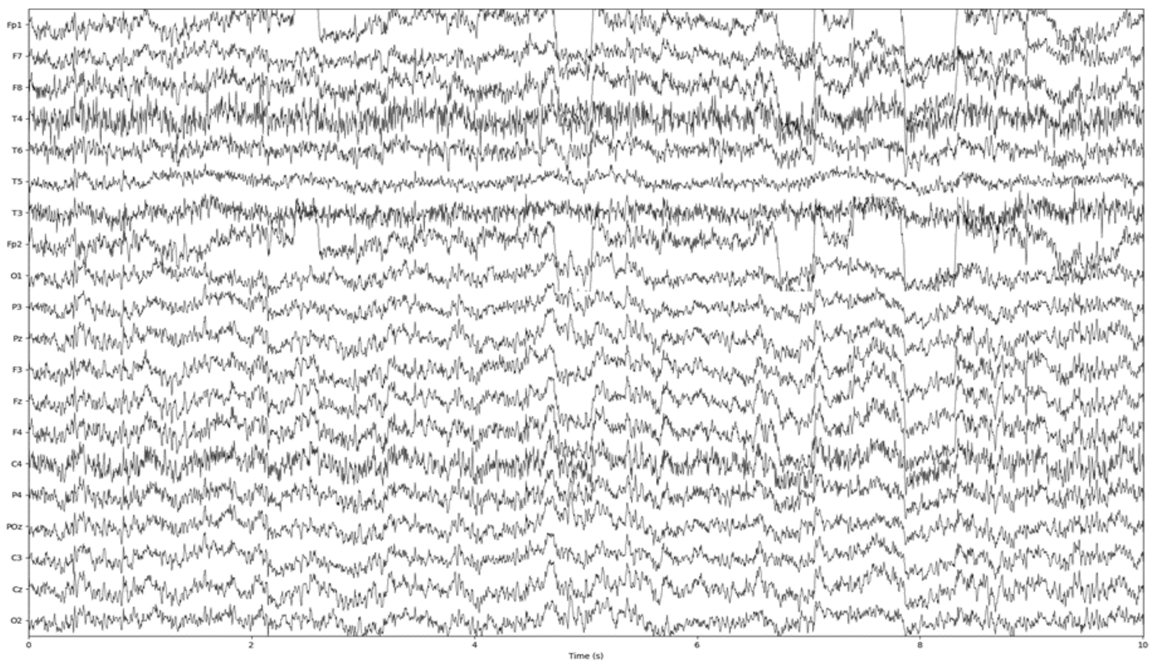

Figure 1: The unprocessed EEG signal

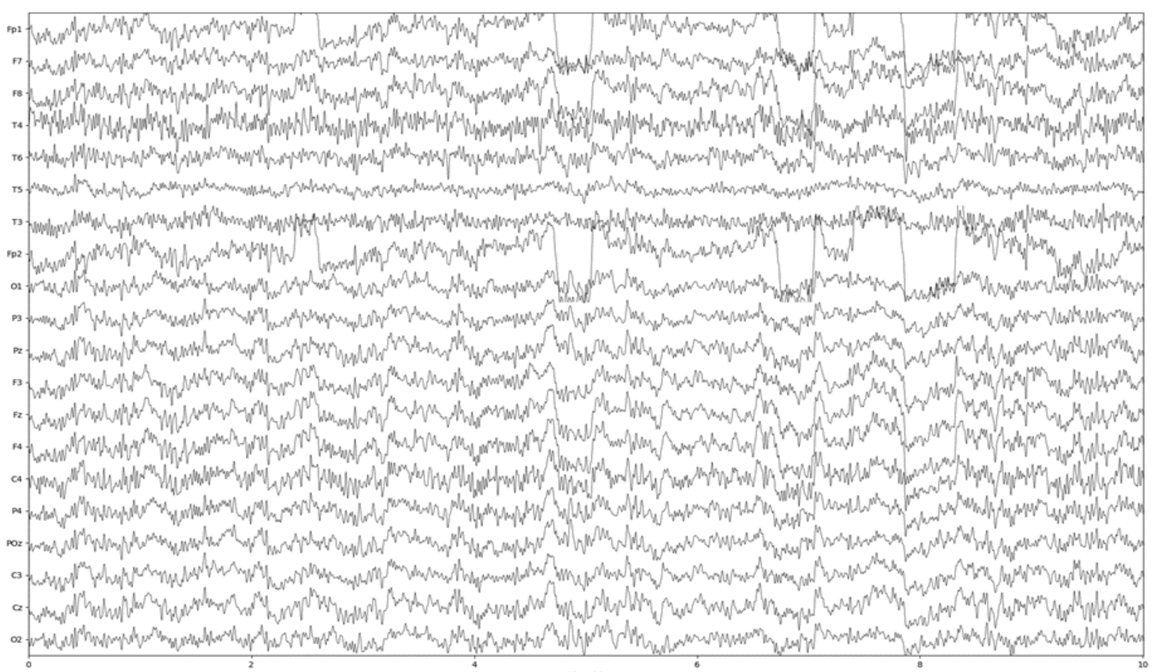

Figure 2: The filtered EEG signal

Table 1: Preprocessing cases

Case	Preprocessing procedure
1	None (EEG Raw data)
2	Band-pass filtering
3	Band-pass filtering and ICA

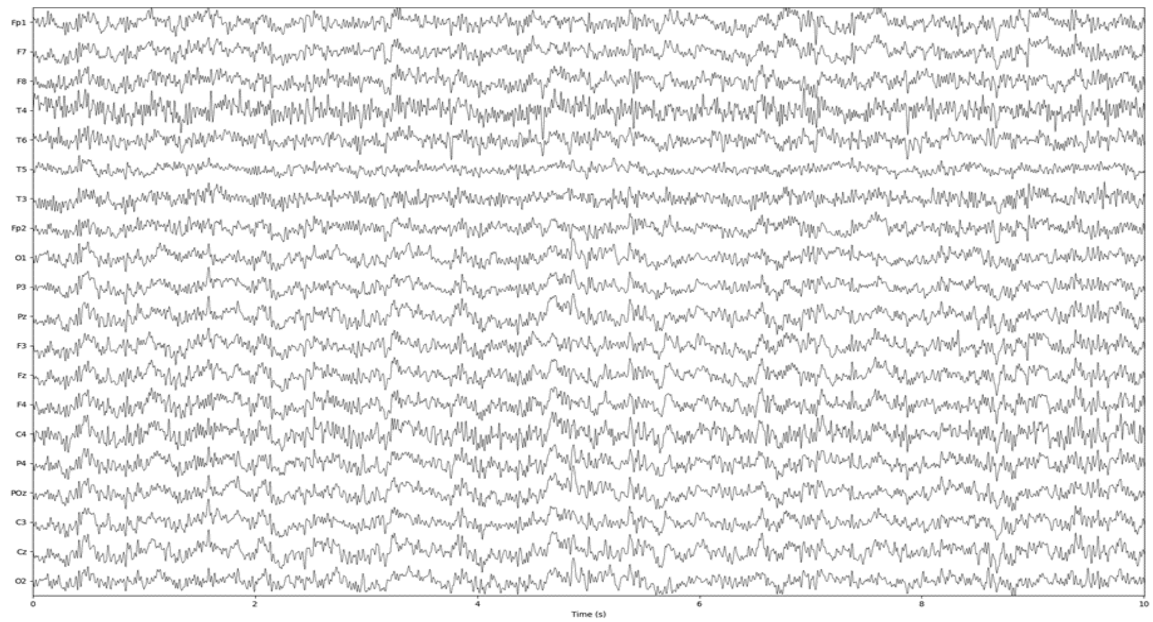

Figure 3: The EEG signal after filtering and removing eye-related artefacts

3.3 EEG Feature Extraction/Engineering

The objective of feature extraction is to identify EEG signal properties. Various variables have been utilized in the literature for EEG classification. Regarding our features, the data were spatially filtered using the Xdawn technique [28], and Tangent space projection was then used to transfer a set of covariance matrices belonging to a manifold into a Euclidean space where they can be considered as vectors [29]. This mapping process permits the application of cutting-edge classifiers within the Riemannian framework.

3.4 Classification Models

In general, machine learning is the generation of knowledge through experience by a computer system. This means that the system receives examples of actual situations and, rather than simply memorizing them, it identifies common principles among examples within the same class. The system can generalize and assign new events to predefined classes following a training phase. To investigate the effect of preprocessing techniques, we verified our cases in SVMs and ANNs models.

SVMs. As large margin classifiers, SVMs are widely employed. SVMs have the characteristic of attempting to classify objects into classes with the largest possible object-free area surrounding them. Within a vector space, every object is represented by a vector. A SVM searches for a hyperplane that separates classes in this space. The data separation procedure can be linear or non-linear, depending on the SVM kernel employed. In non-linear separation, the vector space and its objects are transformed into a higher-dimensional space, allowing for plane-based linear separation. After reverting to a space with fewer dimensions, the linear hyperplane becomes non-linear and may even be discontinuous [30,31]. In this study, we employed the Radial Basis Function (RBF), which is the preferred non-linear kernel.

ANNs. ANNs were presented for the first time in the early 1940s, and they are particularly useful in situations when little is known about the issue. The architecture of an ANN is determined by the number of layers, the number of specific neurons (nodes), and the manner in which they are linked (edges). There are input and output layers as well as one or more hidden layers in ANNs. The network's structure is dependent on the number of hidden layers. In this research, the ANN consisted of a one-layer network trained by 100 iterations of back-propagation [32] utilizing one layer.

4 RESULTS AND DISCUSSION

This section covers several experimental results alongside an analysis of the models' performances. Table 2 summarizes the classification results of the non-flight environment data, the flight environment data, and merged flight and non-flight environments data, respectively. The confusion matrix for 5-fold cross-validation results using non-flight environment data is depicted in Figure 4 (a,b,c), flight environment data is depicted in Figure 5 (a,b,c), and merged data (flight and non-flight environments) is depicted in Figure 6 (a,b,c); where a symbolizes the unprocessed data, b symbolizes the filtered data, and c symbolizes the data that has been filtered and eye blinks have been removed.

Table 2: The classification accuracies

Model\Cases	Unprocessed data	Filtered data	Filtered data and ICA
SVM	91.1%	90.9%	91.3%
ANN	90.6%	90.0%	88.6%

a) The effect of the preprocessing cases on the non-flight environment data

Model\Cases	Unprocessed data	Filtered data	Filtered data and ICA
SVM	91.7%	90.9%	90.8%
ANN	90.7%	90.0%	90.0%

b) The effect of the preprocessing cases on the flight environment data

Model\Cases	Unprocessed data	Filtered data	Filtered data and ICA
SVM	90.8%	90.2%	89.9%
ANN	88.9%	88.4%	88.2%

c) The effect of the preprocessing cases on the merged environments data

Case 1. Table 2 demonstrates that, for Case 1, in which we used only raw data without any preprocessing, all adopted machine learning models are able to record relevant information and classify with a high model performance score, particularly, the SVM model.

Non-flight Data. Table 2a demonstrates that the SVM model has the highest accuracy, at 91.1 %, closely followed by the ANN model, at 90.6 %. However, Figure 4a depicts the confusion matrix of the SVM model, from which a number of observations can be derived. First, the adopted model was incapable of accurately detecting the startle/surprise state. In fact, all fifteen samples of the startle/surprise state were anticipated to be normal (no event) states. In addition, 34 of the 46 samples for diverted attention states and 17 of the 352 samples for channelized attention states were incorrectly predicted as normal states. We hypothesize that the significantly class-imbalanced dataset is the cause of these inferior outcomes. As the majority of poorly predicted samples were predicted as normal states, the results imply that the signal-to-noise ratio of the dataset is low. The presence of noise in a data set can have a negative effect on the performance of learning algorithms, since it can make models more complex and lengthen learning times.

Flight Data. When examining model accuracy in Table 2b, we notice that SVM remains the best model. The model accuracy of the SVM and the ANN is 91.7% and 90.7%, respectively. Figure 4b depicts the confusion matrix of the SVM model from which a number of observations can be recovered. Even though roughly 87% of the data set consists of occurrences of the normal state, the model correctly predicted 201 out of 432 instances of the channelized attention state. Despite such outcomes, the SVM model proved incapable of predicting startle/surprise and diverted attention mental states. This dataset has a higher-class imbalance than the dataset from the non-flight environment. These unbalanced data can significantly diminish a model's accuracy.

Merged Data. It can be seen from the data in Table 2c that the highest accuracy score is still the SVM model which is 90.8% compared to the ANN model which is 88.9. Figure 4c depicts the confusion matrix of the SVM model, from which several observations can be reconstructed. We noticed that after merging the data collected from the pilot in different environments, the model performed slightly better compared to the result from Figure 4b. Specifically, 10 of the 85 samples of the diverted attention states were accurately classified. It may be assumed that the increased model precision was a result of the greater number of minority class samples.

Case 2. In Case 2, we used band-pass filtering to eliminate artefact components from our dataset. With a few exceptions, we notice the same pattern of Table 2 results as in the preceding case.

Non-flight Data. Table 2a shows that the SVM model has the highest accuracy. Nonetheless, the accuracy of the SVM and ANN models declined by 0.2% and 0.6%, respectively, when compared to the results from Case 1. The confusion matrix of the SVM model is illustrated in Figure 5a. Figure 5a's confusion matrix is relatively comparable to that of Figure 4a's.

Flight Data. Taking into account the accuracy score in Table 2b, the best model is the SVM. Nonetheless, it dropped by 0.8% when compared with its in Case 1. Figure 5b depicts the confusion matrix of the SVM model. We noticed a significant decrease in the precision of the correctly predicted samples of the channelized attention (29 fewer samples) compared with its in Case 1.

Merged Data. According to Table 2c, SVM is still the optimum model when compared to ANN. However, the accuracy of the SVM and ANN declined by 0.6% and 0.5%, respectively, when compared to their performance in Case 1. Figure 5c illustrates the SVM model's confusion matrix. The matrices of confusion in Figures 5c and 4c are comparable to some extent.

It appears from the data that band-pass filtering has no positive effect on the performance of the model. The model's performance did indeed slightly regressed. In addition, it could alter the EEG signal in some way, which would mean we would lose some of the data's content.

Case 3. We eliminated extra eye-related artefact components from our dataset in Case 3 using the ICA technique. With a few deviations, Table 2's results follow the same patterns as Case 2's, with few notable outliers.

Non-flight Data. According to the data in Table 2a, the SVM model has the highest accuracy score, followed by the ANN model. Although the accuracy of the SVM model increased by 0.2% compared to Case 1, the accuracy of the ANN model decreased by 2%. The SVM model's confusion matrix is depicted in Figure 6a. Figures 6a and 4a show two similar confusion matrices.

Flight Data. Table 2b shows that the SVM is the best model for flight environment data, followed by the ANN. In comparison to Case 1, the SVM and ANN models' accuracy dropped by 0.9% and 0.7%, respectively. The SVM model's confusion matrix is depicted in Figure 6b. In comparison to Case 2, a minor gain in precision was seen in Case 3 after eye-related artefacts had been removed, despite a large drop in precision in the correctly predicted samples of channelized attention when compared to Case 1.

Merged Data. With respect to Table 2c's accuracy score, the best model is the SVM, which decreased by 0.3 percent from its performance in Case 1. Figure 6c depicts the confusion matrix of the SVM model. Compared to Case 2, we noted that the precision of the channelized attention state samples accurately predicted in Case 3 rose slightly.

Generally, after applying the ICA technique to remove more eye-related artefacts from EEG signals, we find a further slight regression in model performance in this Case 3. The ICA does not appear to have had much of an impact on model performance. There appears to be a slight regression especially in the ANN model's performance in this case, despite the fact that ICA can remove unnecessary information from the data.

Lastly, with a few exceptions, these results indicate that the model's performance marginally deteriorates when more preprocessing approaches are introduced to the pipeline for two datasets gathered from different environments and a merged dataset. Case 3 was the exception, as a band-pass filter and ICA were applied to the non-flight environment dataset. In this instance, we saw a minuscule boost of 0.2% in the performance of the SVM model. As shown in Figure 6a's confusion matrix, the SVM model accurately predicted two more samples of channelized attention and one more sample of diverted attention compared to Case 1. EEG artefacts are typically removed using preprocessing techniques such as filters, which may also induce temporal distortions onto the signals. The visual interpretation of experts and the use of machine learning algorithmic processing and analysis are both hampered by artefacts in EEG data. In addition, the SVM and ANN models were incapable of predicting the startle/surprise and diverted attention mental states using any of the two preprocessing strategies reported in this study. A more advanced preprocessing technique on the EEG signal is therefore justified in identifying during an experiment brain electrical background activity that is unique to a cognitive task.

In addition, results show that standard machine learning models, such as SVMs and ANNs, were more effective at getting things started. However, neither model could detect the mental states of startle or surprise. This issue is mostly attributed to the dataset's high-class imbalance and signal contamination. The vast majority of machine learning techniques presume that data is uniformly distributed. Consequently, when used with a dataset that has an irregular distribution of classes, the classifier tends to be more biased towards the majority classes, resulting in an inaccurate classification of the minority classes. Therefore, the model's inability to distinguish between startle/surprise and the other mental states is a logical consequence.

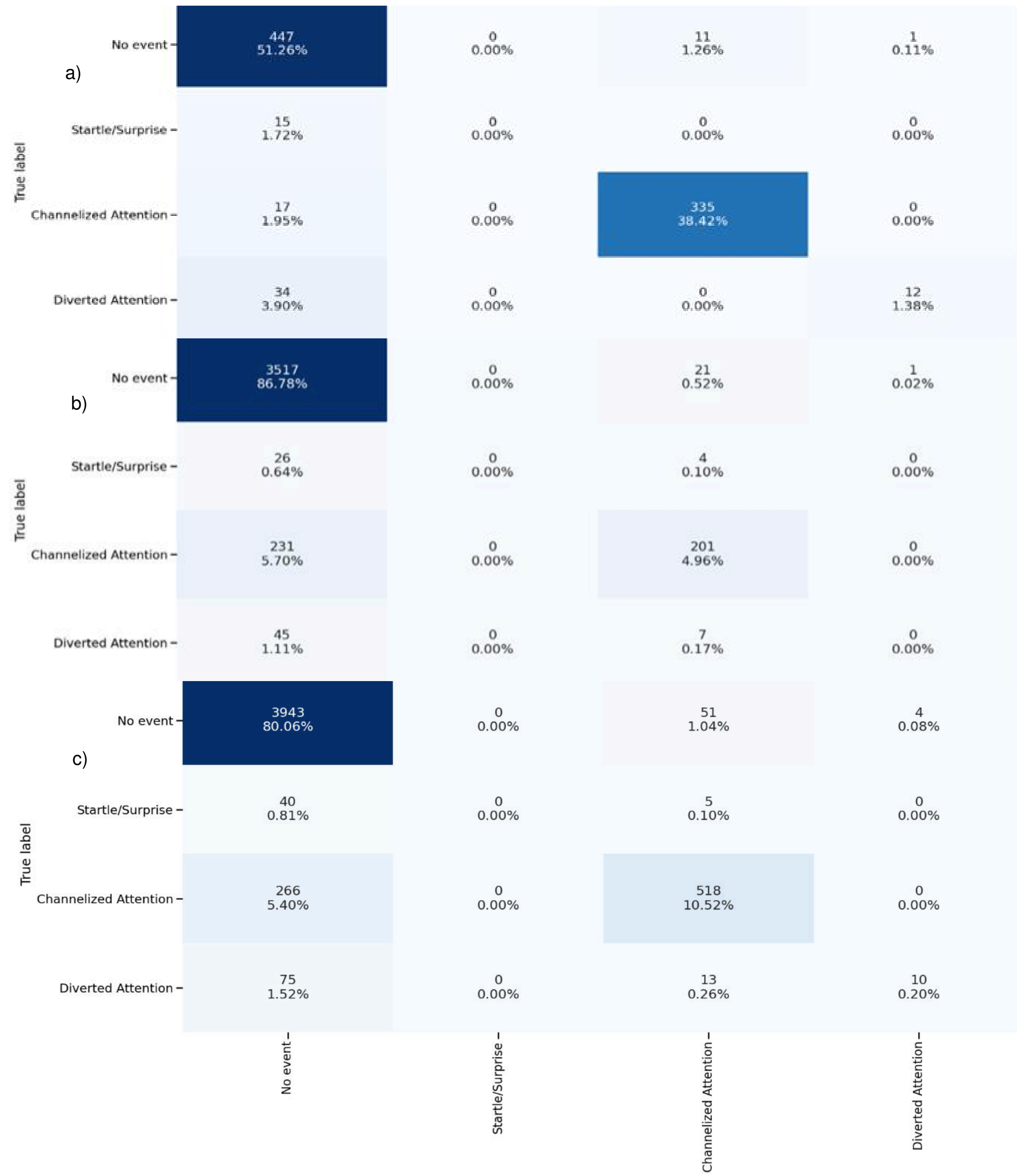

Figure 4: SVM model’s confusion matrix for unprocessed data collected from a pilot in a) non-flight environment, b) flight environment, and c) merged non-flight and flight environments

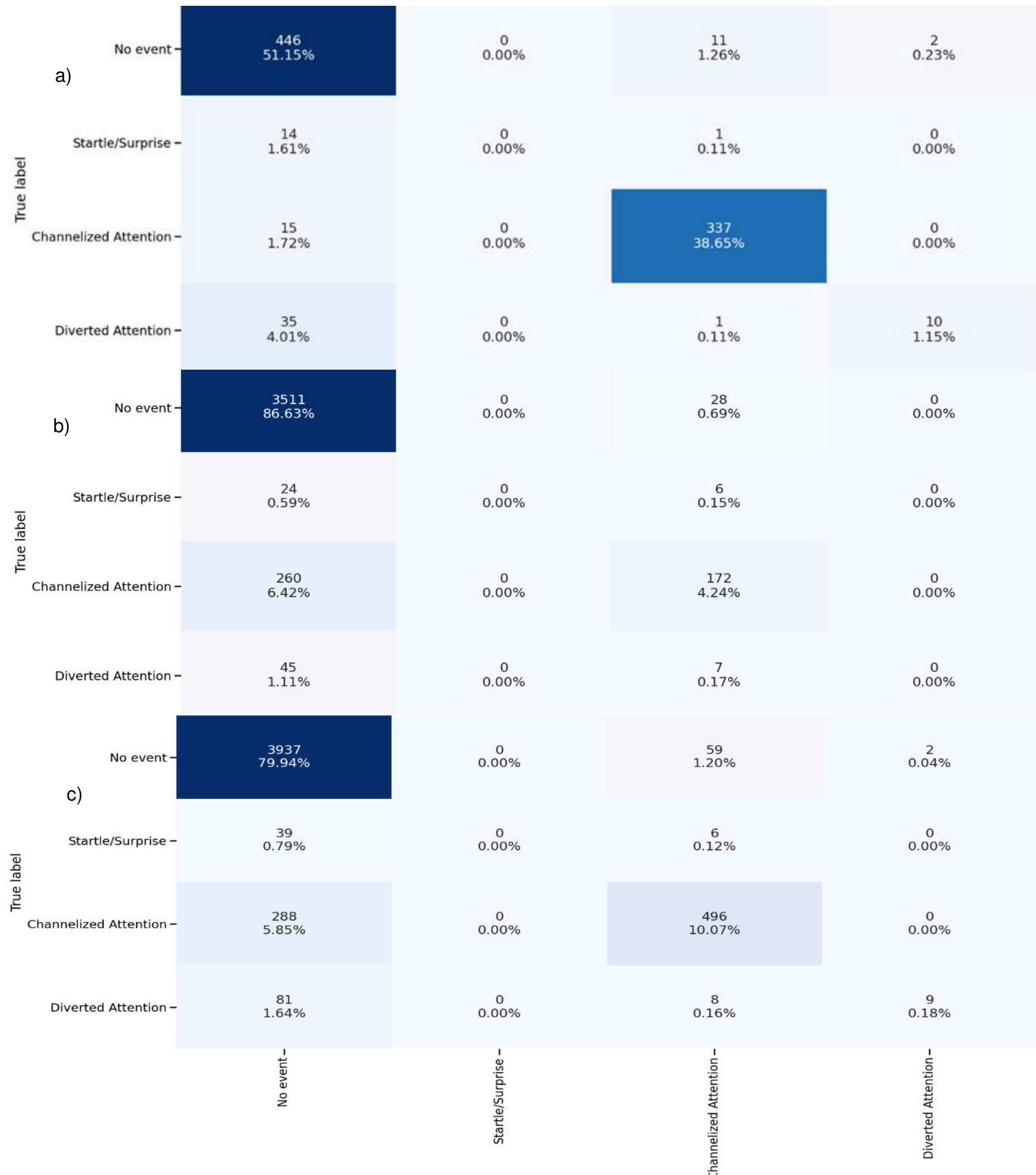

Figure 5: SVM model's confusion matrix for filtered data collected from a pilot in a) non-flight environment, b) flight environment, and c) merged non-flight and flight environments

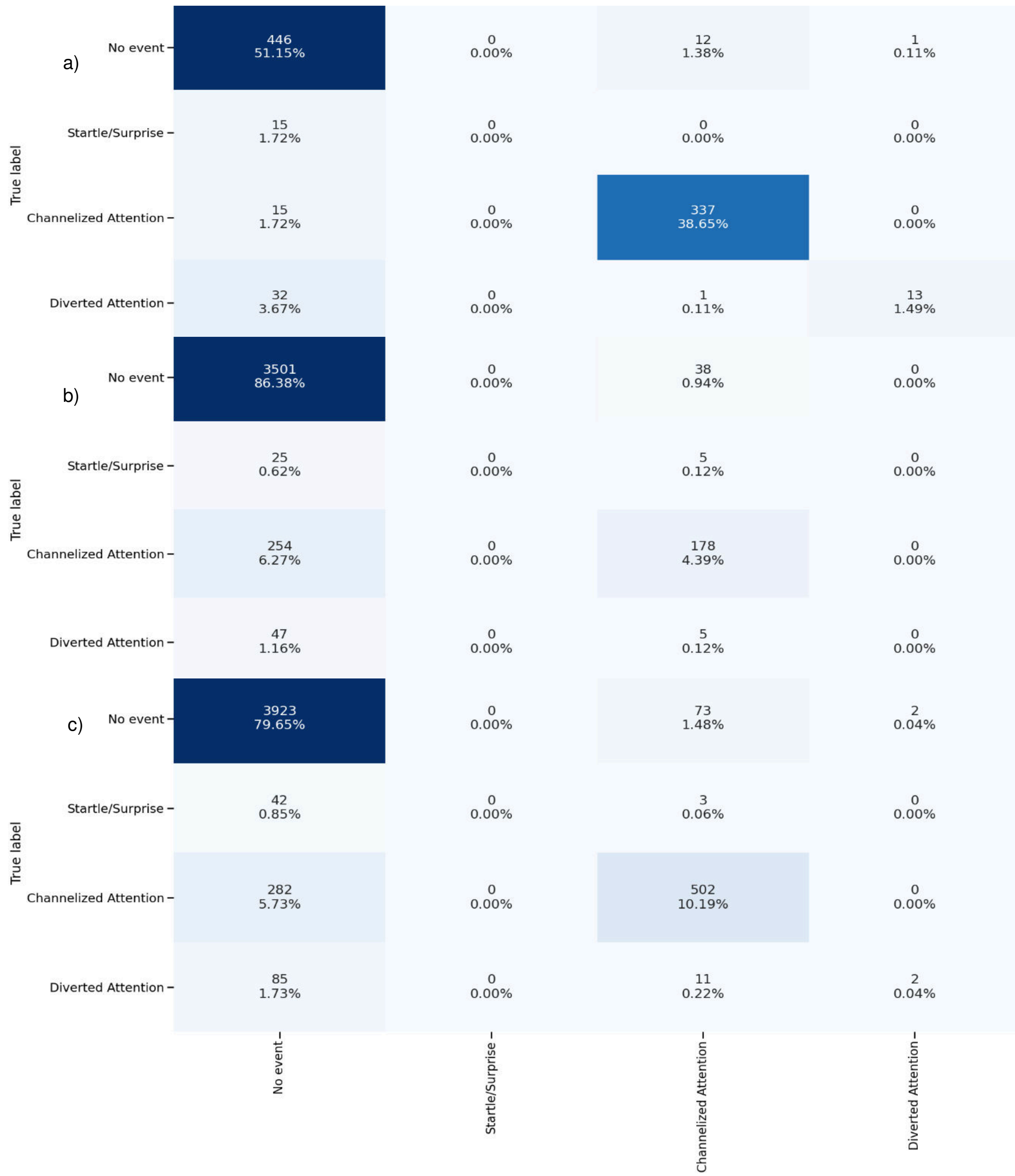

Figure 6: SVM model's confusion matrix for unprocessed data collected from a pilot in a) non-flight environment, b) flight environment, and c) merged non-flight and flight environments

5 CONCLUSION

In this study, we investigated the effect of neuroscientist-defined preprocessing approaches on the performance of machine learning models. To study the effect of preprocessing techniques, we restricted our attention to the band-pass filter and the ICA algorithm. Using a publicly available EEG signals dataset, we validated the impact of these preprocessing strategies on a dataset that is used to train two machine learning models called SVMs and ANNs. Our results indicate that simple preprocessing approaches, such as band-pass filtering and ICA for eye-related artefacts, have no positive effect on the performance of the machine learning models. Our results also indicate that both models were able to operate on a dataset containing information from two distinct environments. In addition, our results indicate that EEG signals preprocessed solely with a band-pass filter exhibited a small regression in classification performance across all models. Similarly, we discovered that our models were unable of differentiating the startle/surprise mental state class from the other classes due to the dataset's extreme class imbalance. While the overall performance of the models is satisfactory, additional work is required to effectively categorize the "startle/surprise," "channelized attention," and "distracted attention" classes. The extraction of brain activity that is unique to a particular cognitive task necessitates more complex strategies for locating the pertinent information in raw data. Future research should consider a sophisticated automated preprocessing pipeline to automatically detect signal artefacts and correct or delete them. In this study, the effects of two artefact removal strategies, a band-pass filter and the ICA algorithm, were examined. Our favorable results inspire additional investigation into the impact of different artefact removal strategies. In addition, we would like to assess the impact of these strategies on a wide variety of deep learning models, such as CNN. Finally, we will study how EEG preprocessing approaches can be included into the deep learning model.

REFERENCES

1. Oehling J., Barry DJ. Using machine learning methods in airline flight data monitoring to generate new operational safety knowledge from existing data. Safety Science. Elsevier; 2019; 114(May 2018): 89–104. Available at: DOI:10.1016/j.ssci.2018.12.018

2. Reid GB., Nygren TE. The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. Advances in Psychology. 1988. pp. 185–218. Available at: DOI:10.1016/S0166-4115(08)62387-0

3. Arico P., Borghini G., di Flumeri G., Colosimo A., Graziani I., Imbert JP., et al. Reliability over time of EEG-based mental workload evaluation during Air Traffic Management (ATM) tasks. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. 2015; 2015-Novem: 7242–7245. Available at: DOI:10.1109/EMBC.2015.7320063

4. Radüntz T., Scouten J., Hochmuth O., Meffert B. Automated EEG artifact elimination by applying machine learning algorithms to ICA-based features. Journal of Neural Engineering. IOP Publishing; 2017; 14(4). Available at: DOI:10.1088/1741-2552/aa69d1

5. Farzan F., Atluri S., Frehlich M., Dhami P., Kleffner K., Price R., et al. Standardization of electroencephalography for multi-site, multi-platform and multi-investigator studies: Insights from the canadian biomarker integration network in depression. Scientific Reports. Springer US; 2017; 7(1): 1–11. Available at: DOI:10.1038/s41598-017-07613-x

6. Bigdely-Shamlo N., Mullen T., Kothe C., Su KM., Robbins KA. The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. Frontiers in Neuroinformatics. 2015; 9(JUNE): 1–19. Available at: DOI:10.3389/fninf.2015.00016

7. Keil A., Debener S., Gratton G., Junghöfer M., Kappenman ES., Luck SJ., et al. Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. Psychophysiology. 2014; 51(1): 1–21. Available at: DOI:10.1111/psyp.12147

8. Longo L., Rusconi F., Noce L., Barrett S. THE IMPORTANCE OF HUMAN MENTAL WORKLOAD IN WEB DESIGN. Proceedings of the 8th International Conference on Web Information Systems and Technologies. SciTePress - Science and and Technology Publications; 2012. pp. 403–409. Available at: DOI:10.5220/0003960204030409

9. Young MS., Brookhuis KA., Wickens CD., Hancock PA. State of science: mental workload in ergonomics. Ergonomics. Taylor & Francis; 2015. pp. 1–17. Available at: DOI:10.1080/00140139.2014.956151

10. Wiebe EN., Roberts E., Behrend TS. An examination of two mental workload measurement approaches to understanding multimedia learning. Computers in Human Behavior. Elsevier Ltd; 2010; 26(3): 474–481. Available at: DOI:10.1016/j.chb.2009.12.006

11. Choi JW., Kim KH. Computational EEG Analysis. 2018; : 125–145. Available at: DOI:10.1007/978-981-13-0908-3

12. Urigüen JA., Garcia-Zapirain B. EEG artifact removal - State-of-the-art and guidelines. Journal of Neural Engineering. IOP Publishing; 2015; 12(3). Available at: DOI:10.1088/1741-2560/12/3/031001

13. Roy Y., Banville H., Albuquerque I., Gramfort A., Falk TH., Faubert J. Deep learning-based electroencephalography analysis: A systematic review. Journal of Neural Engineering. IOP Publishing; 2019; 16(5). Available at: DOI:10.1088/1741-2552/ab260c

14. Chaudhuri A., Routray A. Driver Fatigue Detection through Chaotic Entropy Analysis of Cortical Sources Obtained from Scalp EEG Signals. IEEE Transactions on Intelligent Transportation Systems. IEEE; 2020; 21(1): 185–198. Available at: DOI:10.1109/TITS.2018.2890332

15. Han SY., Kwak NS., Oh T., Lee SW. Classification of pilots' mental states using a multimodal deep learning network. Biocybernetics and Biomedical Engineering. Nalecz Institute of Biocybernetics and Biomedical Engineering of the Polish Academy of Sciences; 2020; 40(1): 324–336. Available at: DOI:10.1016/j.bbe.2019.12.002

16. Binias B., Myszor D., Cyran KA. A machine learning approach to the detection of pilot's reaction to unexpected events based on EEG signals. Computational Intelligence and Neuroscience. 2018; 2018: 1–9. Available at: DOI:10.1155/2018/2703513

17. Yang S., Yin Z., Wang Y., Zhang W., Wang Y., Zhang J. Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders. Computers in Biology and Medicine. Elsevier Ltd; 2019; 109(January): 159–170. Available at: DOI:10.1016/j.compbiomed.2019.04.034

18. Zhang P., Wang X., Chen J., You W., Zhang W. Spectral and Temporal Feature Learning with Two-Stream Neural Networks for Mental Workload Assessment. IEEE Transactions on Neural Systems and Rehabilitation Engineering. June 2019; 27(6): 1149–1159. Available at: DOI:10.1109/TNSRE.2019.2913400

19. Harrivel AR., Liles CA., Stephens CL., Ellis KK., Prinzel LJ., Pope AT. Psychophysiological sensing and state classification for attention management in commercial aviation. AIAA Infotech @ Aerospace Conference. 2016. Available at: DOI:10.2514/6.2016-1490

20. Harrivel AR., Stephens CL., Milletich RJ., Heinich CM., Last MC., Napoli NJ., et al. Prediction of cognitive states during flight simulation using multimodal psychophysiological sensing. AIAA Information Systems-AIAA Infotech at Aerospace, 2017. American Institute of Aeronautics and Astronautics Inc, AIAA; 2017. Available at: DOI:10.2514/6.2017-1135 (Accessed: 18 July 2020)

21. Terwilliger P., Sarle J., Walker S., Terwilliger P., Sarle J., Walker S. A ResNet Autoencoder Approach for Time Series Classification of Cognitive State A ResNet Autoencoder Approach for Time Series Classification of Cognitive State. 2020; (0053): 1–11.

22. Gabard-Durnam LJ., Leal ASM., Wilkinson CL., Levin AR. The harvard automated processing pipeline for electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. Frontiers in Neuroscience. 2018; 12(February): 1–24. Available at: DOI:10.3389/fnins.2018.00097

23. Groppe DM., Makeig S., Kutas M. Identifying reliable independent components via split-half comparisons. NeuroImage. May 2009; 45(4): 1199–1211. Available at: DOI:10.1016/j.neuroimage.2008.12.038

24. Mognon A., Jovicich J., Bruzzone L., Buiatti M. ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology. 2011; 48(2): 229–240. Available at: DOI:10.1111/j.1469-8986.2010.01061.x

25. Vaid S., Singh P., Kaur C. EEG signal analysis for BCI interface: A review. International Conference on Advanced Computing and Communication Technologies, ACCT. 2015; 2015-April: 143–147. Available at: DOI:10.1109/ACCT.2015.72

26. Wilkerson J. Training for Attention Management. 2010; : 1–6.

27. Gramfort A. MEG and EEG data analysis with MNE-Python. Frontiers in Neuroscience. 2013; 7. Available at: DOI:10.3389/fnins.2013.00267

28. Hajinoroozi M., Mao Z., Jung TP., Lin CT., Huang Y. EEG-based prediction of driver's cognitive performance by deep convolutional neural network. Signal Processing: Image Communication. September 2016; 47: 549–555. Available at: DOI:10.1016/j.image.2016.05.018

29. Barachant A., Bonnet S., Congedo M., Jutten C. Multiclass brain-computer interface classification by Riemannian geometry. IEEE Transactions on Biomedical Engineering. April 2012; 59(4): 920–928. Available at: DOI:10.1109/TBME.2011.2172210

30. Christopher J.C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery. 1998; 2: 121–167.

31. Bishop CM. Pattern Recognition and Machine Learning. Information Science and Statistics ed M Jordan et al (Berlin: Springer). 2006.

32. Rumelhart DE., Hintont GE., Williams RJ. Learning representations by back-propagating errors. Nature. 1986;