Digital Twin for Drone Control through a Brain-Machine Interface

Diana Ramos
Faculdade de Engenharia da Universidade do Porto
Capgemini Engineering
Porto, Portugal
email: dianacristina.teixeiraramos@altran.com

Gil Gonçalves
Faculdade de Engenharia da Universidade do Porto
Porto, Portugal
email: gil@fe.up.pt

Ricardo Faria
R&D Tech Leader
Capgemini Engineering
Lisboa, Portugal
email: ricardoandre.pintofaria@altran.com

***Abstract*—Drones enable humans to perform certain high-risk and attention operations and safety-critical tasks remotely, which are boosted by the use of Brain-Computer Interfaces. However, these technologies are correlated with the cognitive state of the operator, who is prone to stress and diversions, which brings instability to drone control. In this paper, we propose a decision making system aiming to decide, upon the operator's emotional state, whether the command should or should not be sent to the drone. By building a predictive operator's digital twin for cognitive emotional detection and by benefiting from a visual facial expression classifier, this system computes the coordinates and sends them to the drone through a Robot Operating System 2 client. Results show that both the digital twin and the facial expression classifier are capable of detecting emotions in a real-time setting and the system provides a reliable and secure way of commanding drones through the mind.**

***Keywords*—drone; Brain-Computer Interface; digital twin; Robot Operating System 2.**

I. Introduction

The drone sector has been growing with higher demand through the years. The common belief is that drones are singularly used for military affairs; however, they are functional and versatile systems. One major use is providing monitoring services, i.e., target searching, surveillance for security purposes and others. Even thought they have attracted companies due to their visionary application, the most significant change is how civilians have been incorporating them for entertainment purposes or as assistive devices to help with their routines [1]. Still, the most impactful usage of drones is their applicability to complete high-risk and safety-critical operations with success, often in locations unreachable and/or dangerous to humans.

A. Problem Overview

Controlling one drone is already a complex task. Operators are responsible for, not just to perform standard operations (takeoff and landing) with success, but also to safely execute them. When adding unsafe and critical operations to the task log, the control complexity increases significantly. The operator needs abilities at their peak, full attention/focus when performing these operations, to provide a reliable and stable control.

Hand control allows operators to remotely send commands to drones; however, as these are critical systems, operators need to be cautious with the commands they deliver. The Brain-Computer Interface (or BCI), or Brain-Machine Interface, is an alternative control mechanism. As humans are prone to fatigue, increasing mental workload and emotions, the control will become uncertain and insecure, that is, the operator can potentially mislead the drone with the wrong commands. In addition, BCIs require operators to have previous experiences to learn to formulate commands, which implies multiple training sessions. Even so, these command classifications are error-prone and contribute to unreliable control.

B. Proposed Solution

The hypothesis of this work is that by adopting a digital twin [4] to virtually represent the operator and by using machine learning techniques, it is possible to process, filter and predict whether the human operator has high mental workload and/or impactful emotions and decide whether the commands produced by the operator should or should not be sent to the drones. With the goal of validating the formulated commands, the digital twin is complemented with a visual emotion recognizer that will classify the operator's visual facial expression into a set of emotional states. Additionally, a Robot Operating System 2 (or ROS2) client node can be used in order to send the commands to the drone.

II. State-of-Art

A BCI is defined as "a device that connects the brain to a computer and decodes in real time a specific, predefined brain activity" [2]. This technology can use direct or indirect methods to do so, namely by evaluating the nerve cells activity or by assessing the levels of blood oxygen for these cells [2]. This technology has proved its relevance in many areas, for instance, there was a study aiming to deliver accurate real-time and precise command classification for drone reliable control. An Electroencephalography (or EEG) headset was used to record the brain activity, followed by a motor imagery acquisition. This mechanism involved four tasks, based on the subject visualizing physical movements instead of performing

 ISBN: 978-1-61208-882-2

them. Then, a classification methodology was developed by combining the Common Spatial Paradigm (or CSP) and the Linear Discriminant Analysis algorithms (LDA) [3]. Using this method, the authors were able to improve classification precision in real time. The solution was validated using a fixed-wing drone use case [3].

Another crucial component of this work is the digital twin. Nowadays, a virtual twin is described as a virtual representation that carries information to realistically behave and change as a physical hardware [4]. This technology is constantly evolving to serve each project needs. One variant that derives from it is the digital twin environment [4] with predicting capabilities. The main goal is to train the digital twin to gain predictive capabilities in order to anticipate the hardware's response or behaviour in situational events during run time. One example is a research work that proposes a framework to improve the estimates of certain measurements of physical systems, more specifically a drone, by implementing a virtual layer, i.e., a digital twin, that would represent the real device and predict its performance [5]. This approach implies that each piece of the drone has its own prediction models that should learn and be updated through time to, ultimately, accurately anticipate some metrics that are valuable to the end-user.

III. Implementation

As a starting point, we chose the *Emotiv Epoc+* headset for data acquisition due to its portability and reliability as a commercial BCI. In addition, this headset is connected with an application, *EmotivBCI*, that aids users as it provides a platform for direct command and facial expression's training and monitoring of multiple data streams. For accurate detection of command patterns, the application is supported by a machine learning prediction model to build a profile and refine it each time the user has a training session. For the purpose of this work, it was necessary that the operator was subjected to multiple sessions of command training to ensure its accuracy. After this stage, the operator was able to formulate *right* and *left* commands, and establish a *neutral* one (stationary state).

As illustrated by Figure 1, this system is composed by 4 components: (1) the digital twin, (2) the visual classifier/component, (3) the decision making component and (4) the ROS2 component.

The digital twin is a virtual representation of the operator and its goal is to classify, in real-time, the operator's emotional state. It relies on data collected by the BCI to build a cognitive profile, adapted to the operator. It is the core component of the proposed solution and provides decisive information to ascertain the destination of the command. The remaining components that follow are designated to support the digital twin and add complementary information for the decision. Three of the data streams, recorded by the *Emotiv Epoc+* headset, are collected: the band power, i.e., power of the EEG data according to the sensor and frequency band; the motion, based on the built-in gyroscope of the headset and the facial expressions, recorded from facial muscle motions. The system communicates with the cortex API to send requests for these data subscriptions and receive JSON responses with the resulting data streams as well as the classified commands, for the time period of the subscription. Since each request and response are unique to each stream, the newly collected data is integrated according to the nearest point in time of each observation, resulting into a single dataset. Columns with unique values are eliminated from this dataset, as well as features that do not add any value for the resolution of the problem. Motion related-features (categorical data) are transformed into binary columns, representing each type, through an one hot encoder. In addition, 2 features were added to the dataset: *arousal* and *valence* values, that are computed according to certain values of band power (according to [6]).

For the classification of the operator's emotional states, a set of classes were selected to represent positive and negative states. The positive classes are *calm* and *focused*, representing a stable cognitive state to send commands to the drone, as opposed to the negative classes (i.e., *stressed* and *distracted*) that detail an unstable cognitive state and, therefore, unacceptable state to send commands. In this work, the same operator simulated all the four emotions, at multiple days, in sessions of 8 seconds, reproducing a balanced dataset of about 19500 observations per emotion. In this work, we split 70% of the data for training the algorithms and 30% for testing and evaluated 8 classifiers in 4 performance metrics. As presented in Table I, Random Forest outperforms the remaining algorithms and is chosen for the training and modeling of the digital twin.

TABLE I
Evaluation of Algorithms

Algorithms	**Performance Metrics**			
	Accuracy	***Precision***	***Recall***	***f1-Score***
Decision Tree	0.995	0.995	0.995	0.995
k-NN	0.997	0.997	0.997	0.997
LDA	0.911	0.916	0.911	0.912
Naive Bayes	0.614	0.645	0.614	0.617
Random Forest	0.999	0.999	0.999	0.999
Support Vector Machine (or SVM) (linear kernel)	0.994	0.994	0.994	0.994
SVM (rbf kernel)	0.888	0.923	0.888	0.894
Neural Networks	0.948	0.949	0.948	0.948

Regarding the visual component of the system, a camera captures the real-time image of the operator and uses a Convolutional Neural Network based-prediction model [7] from an open-source project [7], modeled and trained with the FER-2013 emotion dataset, to classify the visual expressions of the operator as a set of emotions. This component can output: *positive* emotions as *happy* and *neutral* and *negative* emotions as *angry, disgust, fear, sad and surprise*.

While the *EmotivBCI* application classifies the reproduced commands from the operator, the digital twin receives information from the headset and classifies the cognitive state of the operator. The visual component receives the image from the camera and classifies the facial expressions. This process results in three input variables for the decision component. This decision module will decide whether the operator is stable

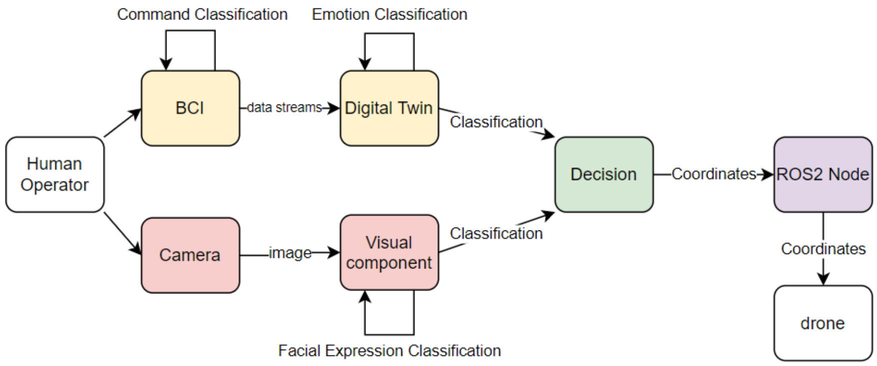

Figure 1. System architecture.

by mentally and visually evaluating his state. Only *positive* emotions detected on both components will allow the operator to send the command. Considering the confidence percentage of the command and the digital twin upon the classification, the decision module, in case of an overall positive emotion detection, will compute the drone coordinates accordingly and send them through a ROS2 node.

For the connection between the system and the drone, a ROS2 client-server architecture is created between what is called the *base station*, meaning the server machine that manages the drone, and the client node that sends requests through a service scheme, specific to a certain operation (takeoff, relative motions and landing). The client node is implemented as a gateway of the decision module, sending a request with the coordinates; the server receives the coordinates and forwards them to the drone in real-time.

IV. EXPERIMENTS

It is expected that, after the operator training session and digital twin training, the system is capable of detecting multiple emotional states of the operator in real-time and handle the drone accordingly. So, to validate this approach, it was assembled a physical environment for secure and controlled drone flight demonstrations, composed by a four square meter indoor zone, called the *arena*, and a pair of *anchors* on each vertex, forming a positioning system that locates the drone by referencing its absolute coordinates. This *arena* was used for operating a *crazyflie* quadcopter in real-time.

To evaluate the different impacts of the solution, functionalities were split in a multi-level manner that go from the lowest experiment to the highest level (solution as a whole) to emphasize its value on securing a stable control environment for the drone. These experiments are: (1) the *baseline test*, defining the current state of drone control without the support of emotion recognition, (2) the *level 1 test*, representing the implementation of the core of the solution which is the cognitive emotion recognition system, (3) the *level 2 test*, the same as the previous test but with the addition of the computation of distances according to the confidence of both the mental command and classified cognitive state of the subject and (4) the *full test*, having all the above functionalities and the support of the visual emotion recognition.

With the exception of the *baseline test*, which gives no importance to the mental state of the subject, each test covers the four mental states (*focused, calm, distracted and stressed*) individually, each one with sessions of 8 seconds. The subject had to be put under the same conditions in which he used to simulate the four emotions on the training phase.

V. RESULTS AND DISCUSSION

Given the environment set-up described in Section IV, the number of observations per emotion and per experiment, for the same subject that trained the commands in the *EmotivBCI* application, are described in Table II:

TABLE II
NUMBER OF OBSERVATIONS PER EMOTION

Emotions	Group of Test		
	Level 1 Test	***Level 2 Test***	***Full Test***
Calm	142	120	85
Focused	94	101	90
Distracted	124	135	104
Stressed	134	120	112

From the number of observations, it was computed the success rate, or accuracy, for each emotion and per experiment (Figure 2). This metric is calculated by dividing the number of correctly classified observations by the total amount of observations. For the *calm* state, the highest accuracy of the digital twin was 87,5%, for the *focused* state a 98,8%, for the *distracted* state a 93,5% and for the *stressed* state a 100%, as described by the round values on Figure 2.

Even with a high average of success rate for detecting the subject's mental states, the most accurately classified emotion was the *stressed* state. The difference between them can be due to the distinct way the model is trained in this segment, which involves more physical movement to denote agitation, rather than a low on motion condition on the remaining ones.

Lower success rate depicted on *level 1* for the *focused* state can be explained by the different background noise and movement between the training and test phase. This caused the subject to deviate his attention, explaining the occurrences of *distracted* classifications during this period. In the next test levels, this value is no lower that 80%, which is explained by the calmer environment. As opposed to this situation, the lower

 ISBN: 978-1-61208-882-2

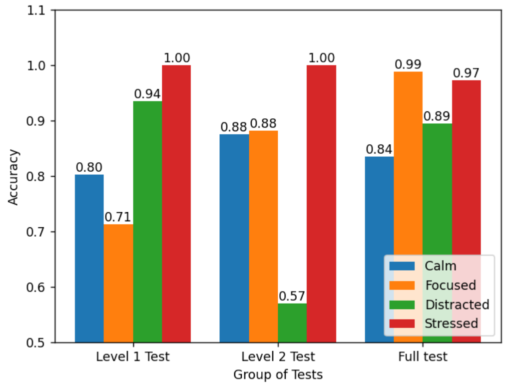

Figure 2. Success rate bar chart.

success rate on *level 2* for the *distracted* state classification can be explained by the lower amount interference or other diversions derived from background movement which led to short occurrences of focus by the subject.

Regarding the classification of the negative emotional spectrum (*distracted* and *stressed* states), Tables III and IV give some insight about the number of sent commands under an incorrect classification.

TABLE III
DISTRACTED EMOTION RECOGNITION

Positive Detections	**Group of Test**		
	Level 1 Test	***Level 2 Test***	***Full Test***
Total number	6	11	10
Nº of neutral commands	4	7	6
Nº of sent commands	2	4	1
BCI positive, visual negative	N/A	N/A	3

As registered in Table III, at *level 1* were detected 6 positive states, 2 of them sent; at *level 2* were detected 11 positive emotions, 4 were sent and at the *full test*, 10 positive emotions were detected, 1 command was sent to the drone and 3 were prevented due to the detection of a negative emotion by the visual component.

TABLE IV
STRESSED EMOTION RECOGNITION

Positive Detections	**Group of Test**		
	Level 1 Test	***Level 2 Test***	***Full Test***
Total number	0	0	2
Nº of neutral commands	0	0	1
Nº of sent commands	0	0	0
BCI positive, visual negative	N/A	N/A	1

As registered in Table IV, at the *full test* were detected 2 positive emotions and none were sent to the drone. One of them was a neutral command and the other was associated with a negative visual emotion, detected by the visual emotion component.

Since the training of mental commands is a task that requires time to practice and refine, it is challenging to reproduce a command at a live setting and in an equivalent environment the subject trained. Even with a digital twin inaccurate classification, most commands detected by the BCI are neutral ones, which have no impact on the trajectory of the drone. However, the command classifier can incorrectly output a *right* or *left* commands and these can potentially be sent to the drones. With the extra layer of the visual component, these unique situations are assessed by it and some of those errors are prevented. At a mission environment, where the operator needs to follow a sequence of commands, if there is a cancellation of a certain command, the operator will observe it and has enough time to reproduce the needed operation.

Considering that this is a 4-class classification problem, there is a probability of 25% that a baseline classifier correctly categorizes the subject emotion state and, in the *baseline test* characterized by the lack of machine learning, all commands are sent to the drones, regardless of the operator's emotional state, which could only be beneficial if the subject has perfect cognitive condition at all times.

VI. CONCLUSION

In this work, we analysed EEG data captured by the BCI *Emotive Epoc+* of a drone operator and, using machine learning techniques, we were able to build a digital twin of the operator capable of predicting its emotional state and decide whether the commands should be sent to the *crazyflie* quadcopter. The classification of the emotional state not only is supported by EEG data but also by a visual component that analyses the facial expressions. In addition, the communication between the system and the drone is done through a ROS2 client node. Multiple machine learning algorithms were validated and Random Forest was the best fitted and therefore used for training the digital twin. Results showed that the digital twin can accurately discriminate the operator's emotional states at a live setting and the combination of classification models improved the security and reliability of the system to decide upon the broadcasting of the reproduced commands. In the future, the goal is to adapt the current system to a swarm of drones, improving the training of the mental commands, the digital twin's accuracy and the efficiency of the digital twin's training and validate this approach with a larger number of subjects with different demographics.

REFERENCES

[1] A. Kreilinger et al., "Switching between manual control and brain-computer interface using long term and short term quality measures", *Frontiers in Neuroscience*, vol. 5, p. 147, 01 2011.

[2] A. Kubler, "The history of bci: From a vision for the future to real support for person hood in people with locked-in syndrome", *Neuroethics*, vol. 13, no. 2, p. 163–180, 07 2020.

[3] R. M. Vishwanath, S. K. Saksena, and S. Omkar, "A real-time control approach for unmanned aerial vehicles using brain-computer interface", *CoRR*, vol. abs/1809.00346, 2018.

[4] M. Grieves, "Origins of the digital twin concept", Available online: https://www.researchgate.net/publication/307509727_Origins_of_the_Digital_Twin_Concept, 08 2016.

[5] H. Y. Jeon, C. Justin, and D. Mavris, "Improving prediction capability of quadcopter through digital twin", in *AIAA Scitech 2019 Forum*, p. 1365, 01 2019.

[6] R. Ramirez and Z. Vamvakousis, "Detecting emotion from eeg signals using the emotive epoc device", in *Brain Informatics*, vol. 7670, pp. 175–184, 12 2012.

[7] U. Gogate, A. Parate, S. Sah and S. Narayanan, ”Real Time Emotion Recognition and Gender Classification”, *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC)*, pp. 138-143, doi: 10.1109/ICSIDEMPC49020.2020.9299633, 2020.