International Journal on Advances in Intelligent Systems, vol 14 no 1 & 2, year 2021, http://www.iariajournals.org/intelligent_systems/

Building a Decision-Making System for Handling a Drone Operator's Emotional States Using a Brain-Computer 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. Drone swarms could be integrated as this solution eases the addition of more ROS2 client nodes.**

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

I. Introduction

This paper showcases the implementation of a drone operator digital twin that relies on a brain-computer interface available data streams in order to evaluate whether the operator is in a suitable condition to send commands to the drone. This work is an extended version of [1] (that detail preliminary results regarding human emotion recognition); thus, details regarding the modelling of the cognitive digital are further analyzed and explained in this document as well as different validation tasks that comply with other requirements adjacent to the main purpose of the work (i.e., experiments with more than one drone as proof-of-concept of a ROS2 client-subscriber communication system for multi-drone control).

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 adopting this technology in their lives. Photography and cinematography are key activities that lead to potential customer interest because of the accessibility drones provide to reach high places, enough to capture a panoramic shot. Either way, drones are essentially useful for people with less motor skills that costly go through their everyday routines. In this case, a drone could serve as an assistive device [2].

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. Although there is a risk of compromising the mission due to faulty hardware or control management, the drone operator is isolated from the target site, which ensures the safety of all the stakeholders involved.

A. Problem Overview

Drones are considered critical systems. By definition, a critical system is a technology that brings its inherent risks while executing, whose failures could be significantly damaging [3]. For instance, the failure of these systems could lead to financial loss where the hardware could be damaged by a collision with other objects or could compromise the mission itself by a poor performance from the operator. Nonetheless, controlling one drone is already a complex task. Operators are responsible for, not just to perform standard operations (i.e., 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 and 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) is an alternative mechanism that aims to optimize this control. As humans are prone to fatigue, increasing mental workload and emotions, the control will become uncertain and insecure, that is, the operator cannot be fully consistent with his performance for longer periods of time due to the organic degrading factors of its human condition. These situations can lead to events of disruption and/or disasters, such as the collision with objects that the system cannot autonomously detect.

All things considered, the problem of controlling a drone with a BCI rises on the addition of the human factor and his involuntary natural incapability of managing a critical system consistently. To address this issue, this work focuses, primarily, on how to reduce or avoid the operational impact on the drone when the operator sends a command under the influence of a negative emotion.

B. Proposed Solution

The hypothesis of this work is that, by adopting a digital twin [6] 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 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.

C. Structure of the Document

This document is structured in seven sections: (I) the current section detailing the context and proposed solution to accommodate the urged necessities mentioned; (II) literature review of the state-of-the-art key technologies selected in the proposed solution; (III) that details the methodology adopted for the development of the digital twin; (IV) implementation insights for the development of the solution; (V) validation of the solution including an incremental test environment for measuring the accuracy and other suitable test scenarios for showcasing the value of the developed system; (VI) conclusion regarding the system and its performance and (VII) listing goals to achieve in future work.

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" [4]. 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 [4]. 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 them. Then, a classification methodology was developed by combining the Common Spatial Paradigm (or CSP) and the Linear Discriminant Analysis algorithms (LDA) [5]. 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 [5].

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 [6]. This technology is constantly evolving to serve each project needs. One variant that derives from it is the digital twin environment [6] 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 [7]. 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.

The goal of this work is to predict the emotional state of a drone operator in real-time. In this specific area, there are some articles that detail visionary approaches on how to analyse the mental workload of a subject or, more importantly, the emotional spectrum. One of these works [8] makes use of a dataset composed by sounds with the goal of triggering certain emotions on the subject. The *arousal* (or excitement) and *valence* (defines the whether the emotion is positive or negative) (Figure 1) are measurements that are computed according to a set of frequency channels of the brain-wave activity, for instance, the *alpha* and *beta* signals. By using machine learning techniques and by training with the Linear Discriminant Analysis (or LDA) and Support Vector Machine (or SVM) algorithms, the authors were able to classify the emotional states of the subjects in the following categories: *happiness, anger, sadness*, and *calm*.

Even though there is a wide range of projects that target these key technologies, this particular research area that crosses the digital twin framework with a BCI for drone control is undeveloped. While the BCI mechanism is being targeted for scientific purposes, it is possible to bring it as a control mechanism of drone systems (as validated in [9]). However, these studies suggest the remote control of drones whenever as far as the human operator can continuously formulate commands. In contrast, this work allows the evaluation of the human emotional state to permit or break this cycle of continuous commands, in the context of the execution of high-risk tasks (where the error margin is very limited). Ultimately, the purpose is to establish a mental control mechanism and assessment of the human condition, while incorporating an efficient communication channel with ROS2.

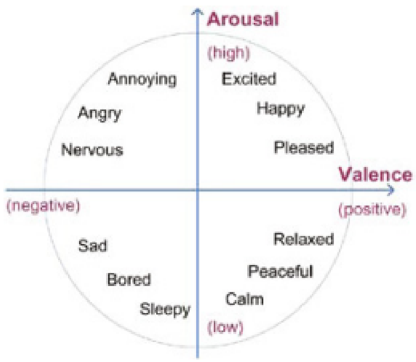

Figure 1. Emotional map based on *arousal* and *valence* (from [8]).

III. Research Methodology

A central piece of this work is the applicability of machine learning techniques to build a cognitive profile of a drone operator (the digital twin). In this work, the Cross Industry Standard Process for Data Mining (or CRISP-DM) methodology [10] was used. This dictates a systematic approach for building an intelligent software component by means of data analysis and application of machine learning algorithms for knowledge gathering. The main steps adopted are the following:

1) **Data understanding:** data acquisition from known and reliable sources that delivers important information for the problem, in addition to analysis of this data to acquire knowledge on its quality and required processing (i.e., finding missing values).
2) **Data Preparation:** prepared the acquire data for the modeling phase based on prior knowledge gathered from the data analysis. This task consumes the most resources, because it has a direct impact on the quality and performance of the final digital twin.
3) **Modeling:** manipulate the previous dataset to build a prediction model, i.e., the digital twin. In this phase, training multiple machine learning algorithms and validating in a set of performance metrics is crucial to assess which algorithm best fits the proposed problem.
4) **Evaluation:** the best fitted prediction model and its outcomes provides a way of gathering more knowledge and validate if all goals established are being met.

A. Brain-Computer Interface Headset

As a starting point, the *Emotiv Epoc+* [11] headset was chosen (Figure 2), developed by the *Emotiv* company, for data acquisition due to its portability and reliability as a commercial BCI. *Emotiv Epoc+* provides a source of data of interest since EEG signals to the motion of the head according to a set of 3-axis. It is composed by *arms* that fit on specific locations on the scalp and are arranged accordingly to ease the subject while placing it on the head. In addition, each sensor is embedded with a felt tip. Dumping this felt tip is crucial for the signal quality since it will connect the whole assembly to the scalp and allow brain activity to be detected (for more information check [11]).

Figure 2. *Emotiv Epoc+* hardware.

Between the multiple software that allow the exploration of data streams, it was used the *EmotivBCI*, developed by the *Emotiv* company. This provides a platform for direct command and facial expression's training and monitoring of multiple data streams. Fundamentally, in order to control a drone with a BCI, a first approach is to create a set of commands to be operational during experiments. This task implies the creation of strategies to be mentally reproducible whenever the operator desires. In this case, the operator could reproduce a certain command by visualizing the movement to the matching direction, followed by the natural eye movement. This procedure works around a method of training the command and testing in a live environment, both provided by the platform. 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. This allows the definition of a command through the identification of patterns the operator organically produces while training. 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). For more information check [12].

B. Experimental Setup

This work falls within the context of a real-world use case and validated as such. In this environment, the drone used was the *crazyflie 2.1* quadcopter, developed by the *Bitcraze* company [13]. Measuring 92 mm of width, 92 mm of height and 29 mm of depth, this drone model is as lightweight as 27 g and holds on air for about 7 minutes. *Crazyflie 2.1.* is a suitable model for validating this work, because it is easier to control upon unexpected behaviors derived from hardware or communication failures, without having significant impact on their surroundings.

The second part of this experimental setup is the zone for flight demonstrations. As explained above, a drone is a critical system and can result in unpredictable behaviors. To validate the proposed solution, a dedicated physical area was assembled, called the *arena* (Figure 3). The *arena* is a four-meter indoor area, gathered by net for safety precautions and a position system, similar to a GPS. The *loco positioning system* [14] was developed by the *Bitcraze* company and aims for locating the drone in the 3-dimensional space. For this purpose, *anchors* are placed on each vertex of the *arena*, serving as reference guides.

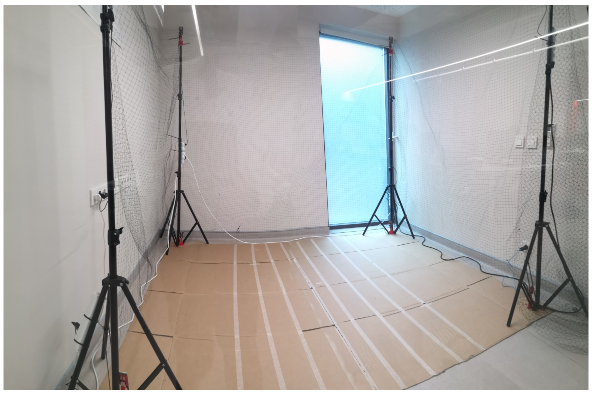

Figure 3. Drone Arena.

IV. Implementation

As mentioned in Section I-B, the core goal of this work is the development of a software platform capable of: (1) acquiring data from a BCI, (2) pre-process it, (3) identify the current emotional state of a drone operator and (4) decide whether he is in a suitable emotional condition to send commands. Even though brain-activity is the primary source of data for identifying the cognitive emotional level, a real-time visual input is also taken into account as a way of validating or invalidating this result. Nonetheless, as a decision is computed, it is equally crucial to establish a communication channel with the drone in order to forward all necessary information. This section will detail the implementation of such system and all necessary components that function together for these purposes.

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

A. Digital Twin

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.

According to Section III, this work comprehends that the digital twin, by means of machine learning techniques, will be built in an iterative manner following the tasks of: data acquisition, data preparation, modeling and evaluation. After data acquisition, procedures that follow are executed offline. The goal is to produce a digital twin, or a prediction model, that is going to be built under a set of training and testing tasks. Training requires data to be pre-processed and *cleaned* so that a machine learning algorithm can learn from it and find patterns on behaviors that correspond to certain emotions. This will result in an intelligent twin, which will be submitted to a testing session to validate whether the predictions match the correct emotion.

1) Data Acquisition: In a first approach, data acquisition was performed based on the analysis of available data streams from the BCI and fetched by a subscription procedure, which allows recording in real-time. For this purpose, three data streams, recorded by the *Emotiv Epoc+* headset, are collected: (1) the band power, i.e., power of the EEG data according to the sensor and frequency band; (2) the motion, based on the built-in gyroscope of the headset and (3) the facial expressions, recorded from facial muscle motions. The system communicates with the cortex API [15] 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.

2) Data Preparation: Since each request and response are unique to each stream, records are written to different files as separate datasets matching their type of data stream. In order to manipulate data on further tasks and, as needed by algorithms, these datasets need to be merged together. The newly collected data is integrated according to the nearest point in time of each observation (according with the *timestamp* feature), resulting into a single dataset. One particularity of this joint transaction is that it fills missing values of whatever feature with the closest value. This is useful due to different frequencies of the subscriptions and the asynchronous timestamps, which may not match exactly.

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 (i.e., the *timestamp* feature).

Another strategy to improve the quality of the dataset is to perform feature engineering. This method transforms the current features and/or includes domain knowledge to increase their value and impact when solving the problem at hand. For this purpose, *one-hot-encoding* was performed on motion related-features (categorical data). This process consists on selecting each categorical column and transform it into a binary value. Motion related-features are transformed into binary columns, representing each type, through an one hot encoder. In addition, two features were added to the dataset: *arousal* and *valence* values, that are computed according to certain values of band power (according to [16] and equations (1) and (2)).

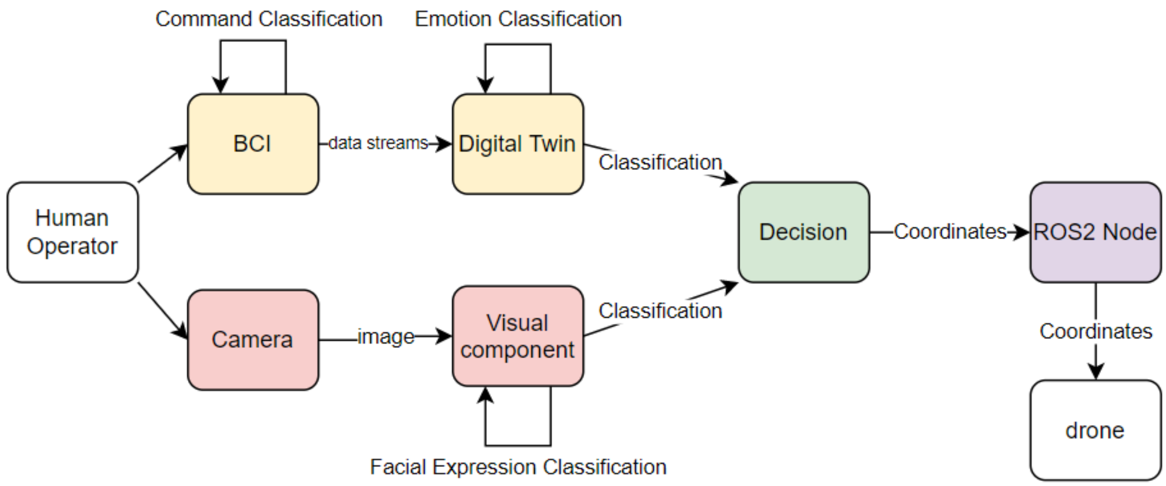

Figure 4. System architecture.

$$arousal = \frac{(F3/betaL + F4/betaL)}{(F3/alpha + F4/alpha)} \quad (1)$$

$$valence = \frac{F4/alpha}{F4/betaL} - \frac{F3/alpha}{F3/betaL} \quad (2)$$

3) Modeling and Evaluation: 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 19 600 observations per emotion (78 400 total). In this work, 70% of the data was split for training the algorithms and 30% for testing and evaluated 8 classifiers in four performance metrics (Table I).

TABLE I
EVALUATION OF ALGORITHMS

Algorithms	Performance Metrics			
	Accuracy	***Precision***	***Recall***	***f1-Score***
Decision Tree	0.995	0.995	0.995	0.995
k-Nearest Neighbors	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
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

As presented in Table I, Random Forest outperforms the remaining algorithms and is chosen for the training and modeling of the digital twin. In addition, the resulting confusion matrices are further analysed. These matrices display the true and false positives and negatives as way of assessing the level of confusion between classes from the machine learning algorithms. This work focuses primarily on a secure platform to send commands only when the operator is in a suitable condition to do so. Nevertheless, it also aims for minimizing the impact on the drone upon inaccurate classifications from the digital twin. Taking this into consideration, a confusion matrix is useful to assess the proportion of observations that were incorrectly classified as *calm* or when the operator was indeed *distracted* or *stressed*.

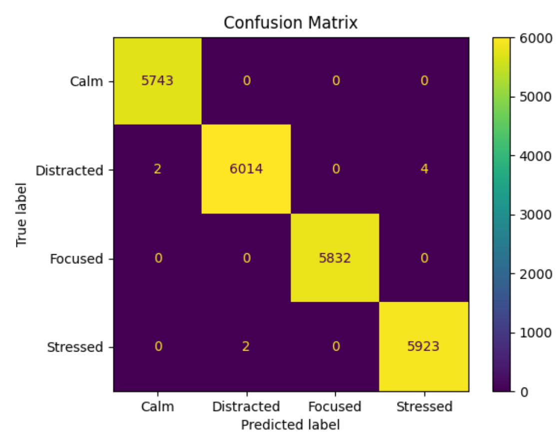

Figure 5. Random Forest confusion matrix (from [12]).

As illustrated by Figure 5, almost all observations were correctly identified by the Random Forest-based digital twin. Still, by analysing the true condition of the operator, in the negative spectrum, against the predicted labels, two observations that belonged to the *distracted* class were incorrectly classified as the *calm* class. From the testing set (30% of the whole dataset, about 23 520 observations), this error represents 0.009% of the sample. Although this demonstrates a flaw from the digital twin, the probability of occurrence is minimum. For further details, check [12].

B. Visual Component

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 [17] from an open-source project [17], 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*.

C. Decision Component

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 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. These coordinates are based on the distance to be travelled by the drone.

$$d = (c * i) * e \tag{3}$$

Equation (3) showcases the computation of the distance, where c is the confidence of the *EmotivBCI* classifier upon the identified command, e is the confidence of the digital twin upon the classified cognitive emotion and i is a safety increment with value 0.25 (meters). This means that classification errors from the prediction models might have an impact on the drone but is not decisive that it will be catastrophic due to filtering on the computation of coordinates.

D. ROS 2 Communication 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 (Figure 6). Nodes are software units that are responsible for the execution of some task. Here, there are two nodes, the client and the server. For these nodes to communicate, service schemes are created specifically to portray each task. The client instantiates the required scheme with the desired values and sends the request message to the other node. In this case, the operations available for the drone are: takeoff, relative motions and landing. These actions are translated into service schemes that are used by the solution, through the client node, with the desired values. In this context, 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. Code listing 1 is an example of service that depicts the relative motion operation that matches the mental command formulated by the operator.

```
float32 x
float32 y
float32 z
float32 yaw
float32 duration
---
int8 ret
```

Listing 1. *GoTo* service scheme

V. Results and Discussion

In this section, are presented the results and their discussion. This validation is divided into three parts: (V-A) where each emotion is experienced isolated, (V-B) a free environment with unexpected occurrence of emotions and (V-C) ROS2 communication validation with two drones to ensure a message can reach both. 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.

A. Isolated Emotion Validation

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:

- The ***baseline test***, defining the current state of drone control without the support of emotion recognition;
- The ***level 1 test***, representing the implementation of the core functionality which is the cognitive digital twin and the ROS2 client node;
- The ***level 2 test***, the cognitive digital twin with the addition of the computation of coordinates according to the prediction's confidence and the ROS2 client node;
- 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.

One particularity is that levels 1 and 2 differentiate only in the computation of coordinates, which is a more visible advancement while operating the drone rather than an improvement of accuracy of the system. In this context, as a first validation session, the goal is to compute the accuracy

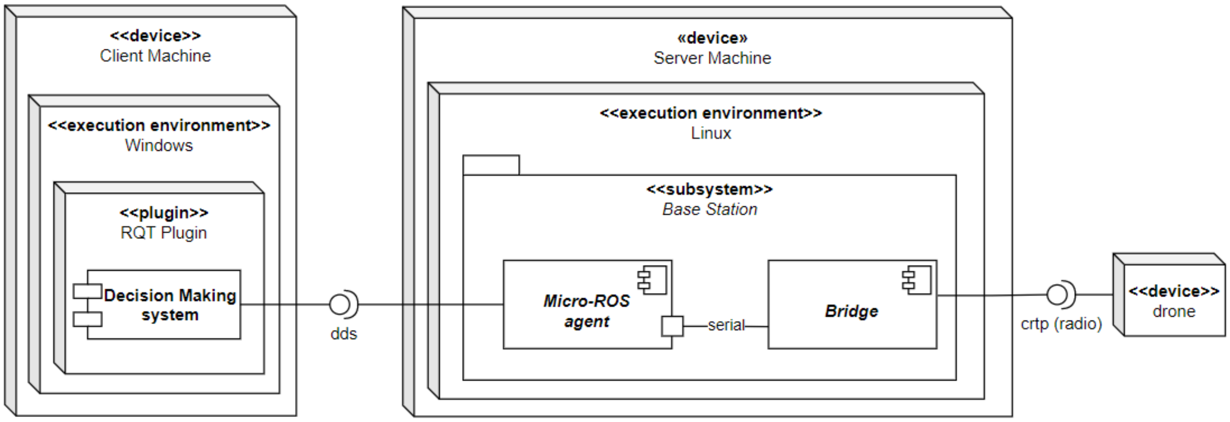

Figure 6. ROS 2 System architecture (from [12]).

of the digital twin when classifying the operator's mental emotions (experimental levels are seen as different temporal spaces). The second approach showcased in this section, is the analysis of observations of the negative spectrum. Here, the experimental levels are more significant to discuss the value of the functionalities of the system.

Given the environment set-up described in Section V, 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 7). 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 7.

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

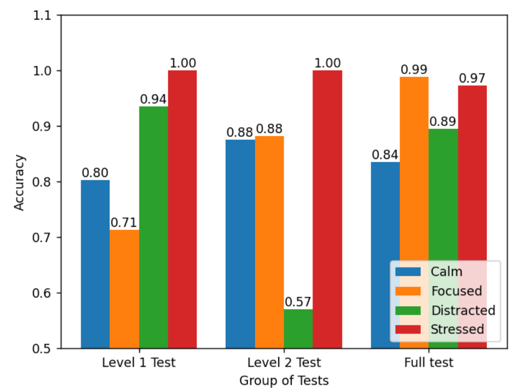

Figure 7. Accuracy bar chart (from [12]).

levels, this value is no lower that 80%, which is explained by the calmer environment. As opposed to this situation, the lower 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

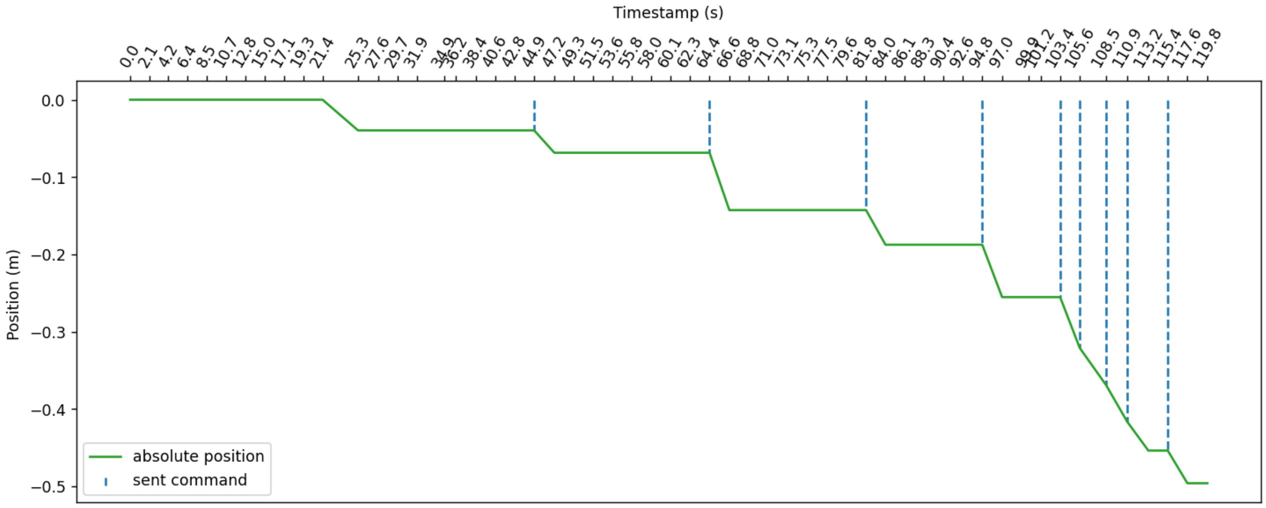

Figure 8. Real-time mission with distracting external events with a single drone (from [12]).

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.

B. Free Mission Flight

Even though previous section already validates the goal of this work, an isolated environment, where the operator simulates the four emotions, is a controlled scenario. In a free environment, realistically, is it common to occur unexpected events and different reactions from the operator. In this section, it was conducted a 2-minute validation session were the operator was able to send whatever command. For the purpose of validating the detection of mood swings, for instance between the *focused* and *stressed* states, two alarms were set to trigger at specific timestamps (27.6 and one minute and twenty eight seconds) after the experiment initiated.

Figure 8 illustrates the distance travelled by the drone, referencing the absolute position on the y axis (only *right* commands were sent). During this experiment, the operator was able to focus on the drone and send multiple commands, each one with a different distances. Even though the operator was consistently focused, at the trigger of the first alarm, the system detected a *distracted* mental state and a *surprised* visual state. The operator continued to be either *stressed* or *distracted* afterwards but was able to refocus on the drone and send new commands. At the sound of the second alarm, the system did not detect immediately a mental reaction but identified the physical reaction as *fear*. This validates the value and robustness of the solution when handling incorrect classifications by including a second classifier, the visual component, to break the command cycle. For more information regarding this experiment, check [12].

C. ROS2 Swarm Management Validation

Although a drone is a fundamental piece for the execution of complex tasks, a swarm of drones aims to optimize resource allocation to more efficiently perform these high-risk tasks. This section details a third experiment composed to include two drones to receive and execute the same operations. The following code listings demonstrate the communication of two client nodes with the server node.

```
[INFO] [1623235179.701766700] [minimal_service]:
    Take off incoming request
height: 0.500000
duration: 2.000000
response: 1

[INFO] [1623235179.704947100] [minimal_service]:
    Take off incoming request
height: 0.500000
duration: 2.000000
response: 1
```

Listing 2. *Server node take-off* message.

```
[INFO] [1623235180.412227300] [drone1]: Sending
    information to server: height: 0.500000
duration: 2.000000
response: 1

[INFO] [1623235180.413246332] [drone2]: Sending
    information to server: height: 0.500000
duration: 2.000000
response: 1
```

Listing 3. *Client nodes take-off* messages.

```
[INFO] [1623235924.239647700] [minimal_service]: Go
    to incoming request
x: 0.000000 y: 0.041117 z: 0.000000 yaw: 0.000000
    duration: 1.000000

[INFO] [1623235924.244461100] [minimal_service]: Go
    to incoming request
x: 0.000000 y: 0.041117 z: 0.000000 yaw: 0.000000
    duration: 1.000000
```

Listing 4. *Server node go to* message.

```
[INFO] [1623235925.104762600] [drone1]: Sending
    information to server: x coordinate: 0.000000
y coordinate: 0.041117
z coordinate: 0.000000
rotation: 0.000000
duration: 1.000000
response: 1

[INFO] [1623235925.106783200] [drone2]: Sending
    information to server: x coordinate: 0.000000
y coordinate: 0.041117
z coordinate: 0.000000
rotation: 0.000000
duration: 1.000000
response: 1
```

Listing 5. *Client nodes go to* messages.

```
[INFO] [1623231818.296960000] [minimal_service]:
    Land incoming request
height: 0.000000
duration: 2.000000
response: 1

[INFO] [1623231818.407826700] [minimal_service]:
    Land incoming request
height: 0.000000
duration: 2.000000
response: 1
```

Listing 6. *Server node land* message.

```
[INFO] [1623235180.412227300] [drone1]: Sending
    information to server: height: 0.000000
duration: 2.000000
response: 1

[INFO] [1623235180.413246332] [drone2]: Sending
    information to server: height: 0.000000
duration: 2.000000
response: 1
```

Listing 7. *Client nodes land* messages.

As expected, the inclusion of an additional client node (one more drone) does not disrupt the functionality of the overall system and it is equally possible to send requests and receive messages from the same server client in parallel with the execution of other client nodes. This experiment demonstrates that is possible to work with a swarm of drones with ROS2 without significant structural efforts.

VI. CONCLUSION

In this work, it was analysed EEG data captured by the BCI *Emotiv Epoc+* of a drone operator and, using machine learning techniques, we were able to build a digital twin of the operator capable of predicting his 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. Considering the research question pointed in Section I-A, 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 reliability of the system to decide upon the broadcasting of the reproduced commands.

While uniquely relying on the cognitive classifier, the digital twin, even with an overall satisfactory performance, allows inaccuracies to happen unexpectedly while the operator is not at his best state of mind. Including the visual component minimizes the impact of these situations. As part of our human condition, emotional reactions are often accomplished by mental and physical responses. While the digital twin could produce an incorrect classification, the visual component will likely detect a negative emotion and the command will not be forwarded to the drone. A potential obstacle to the control of the drone could be the fact that the accumulative complexity of be mentally prepared to send commands, formulate a command with success and be physically stable could disrupt the execution of the command or sequence of command unnecessarily. From what could be a quick set of commands to perform a simple operation, it could demonstrate to be the opposite and even leading to the frustration of the operator. Regarding this issue, the visual component best detects the *happy* and *neutral* states. Without a proper physical reaction, this component will not detect other states and, therefore, these disruptions will most likely not happen.

Additionally, the implementation of a ROS2 framework in this work has proven to be crucial for the satisfactory functionality of the system as it allows the management and communication of information for one drone without explicit hardware/firmware constraints. This work also validates the ease of adding more drones (client nodes) for the execution of tasks of higher demand.

VII. Future Work

As future work, we aim for collecting more data to train the digital twin regarding the four emotional states and ensure it keeps improving its real-time detection.

Due to *Covid-19*, it was not possible to validate this solution with a wider range of subjects. Since a digital twin is adapted to each subject, we plan on creating more profiles of people with different demographics.

Emotiv Epoc+ was the target BCI for this work; thus, we would like to experiment this platform with other commercialized and, perhaps, open-source devices such as the *OpenBCI* headset.

References

[1] D. Ramos, G. Gonçalves, and R. Faria, ”Digital Twin for Drone Control through a Brain-Machine Interface”, *INTELLI 2021, The Tenth International Conference on Intelligent Systems and Applications*, 2021.

[2] M. A. Soto and M. Funk, ”Look, a guidance drone! Assessing the Social Acceptability of Companion Drones for Blind Travelers in Public Spaces”, *20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18)*, Association for Computing Machinery, New York, NY, USA, pp. 417–419, 2018.

[3] M. Hinchey and L. Coyle, ”Evolving Critical Systems: a Research Agenda for Computer-Based Systems”, *2010 17th IEEE International Conference and Workshops on Engineering of Computer Based Systems*, pp. 430-435, 2010.

[4] 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, pp. 163–180, 2020.

[5] 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.

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

[7] H. Y. Jeon, C. Justin, and D. Mavris, “Improving prediction capability of quadcopter through digital twin”, in *AIAA Scitech 2019 Forum*, pp. 1365, 2019.

[8] R. Ramirez and Z. Vamvakousis, ”Detecting emotion from eeg signals using theemotive epoc device”, *Brain Informatics*, pp. 175–184, Berlin, Heidelberg, 2012.

[9] R. Vishwanath, S. Saksena, and S. Omkar. ”A real-time control approach for unmanned aerial vehicles using brain-computer interface”, CoRR, abs/1809.00346, 2018.

[10] C. Manna, ”An intro to the crisp-dm methodology”. Available at https://medium.com/@chrismanna/an-intro-to-the-crisp-dm-methodology-c58cbe0371a3, Jun. 2019. Accessed on 21.03.2021.

[11] EMOTIV, ”Emotiv epoc 14-channel wireless eeg headset”. Available athttps://www.emotiv.com/epoc/, Sep. 2020. Accessed on 13.07.2021.

[12] D. Ramos, ”Digital Twin for Drone Control using a Brain-Computer Interface” (Unpublished master's thesis). Faculty of Engineering of the University of Porto, 2021.

[13] Crazyflie 2.1, ”Crazyflie 2.1”. Available at https://www.bitcraze.io/products/crazyflie-2-1/, Sep. 2020. Accessed on 7.09.2021.

[14] Loco Positioning System, ”Loco Positioning System”. Available at https://www.bitcraze.io/documentation/system/positioning/loco-positioning-system/, Sep. 2020. Accessed on 7.09.2021.

[15] Cortex API, ”Getting Started”. Available at https://emotiv.gitbook.io/cortex-api/, Jun. 2021. Accessed on 7.09.2021.

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

[17] 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.