



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

journal homepage: [www.elsevier.com/locate/cose](http://www.elsevier.com/locate/cose)Computers  
&  
Security

# Neuronal Jamming cyberattack over invasive BCIs affecting the resolution of tasks requiring visual capabilities



Sergio López Bernal<sup>a,\*</sup>, Alberto Huertas Celdrán<sup>b</sup>, Gregorio Martínez Pérez<sup>a</sup>

<sup>a</sup>Department of Information and Communications Engineering, University of Murcia, Murcia 30100, Spain

<sup>b</sup>Communication Systems Group CSG, Department of Informatics Ifl, University of Zurich UZH, CH—8050 Zürich, Switzerland

## ARTICLE INFO

### Article history:

Received 19 January 2021

Revised 19 October 2021

Accepted 1 November 2021

Available online 8 November 2021

### Keywords:

Cybersecurity

Safety

Neuronal cyberattacks

Convolutional neural networks

Brain-Computer Interfaces

## ABSTRACT

Invasive Brain-Computer Interfaces (BCIs) are extensively used in medical application scenarios to record, stimulate, or inhibit neural activity with different purposes. An example is the stimulation of some brain areas to reduce the effects generated by Parkinson's disease. Despite the advances in recent years, cybersecurity on BCIs is an open challenge since attackers can exploit the vulnerabilities of invasive BCIs to induce malicious stimulation or treatment disruption, affecting neuronal activity. In this work, we design and implement a novel neuronal cyberattack called Neuronal Jamming (JAM), which prevents neurons from producing spikes. To implement and measure the JAM impact, and due to the lack of realistic neuronal topologies in mammals, we have defined a use case using a Convolutional Neural Network (CNN) trained to allow a simulated mouse to exit a particular maze. The resulting model has been translated to a biological neural topology, simulating a portion of a mouse's visual cortex. The impact of JAM on both biological and artificial networks is measured, analyzing how the attacks can both disrupt the spontaneous neural signaling and the mouse's capacity to exit the maze. Besides, another contribution of the work focuses on comparing the impacts of both JAM and FLO (an existing neural cyberattack), demonstrating that JAM generates a higher impact in terms of neuronal spike rate. As a final contribution, we discuss whether and how JAM and FLO attacks could induce the effects of neurodegenerative diseases if the implanted BCI had a comprehensive electrode coverage of the targeted brain regions.

© 2021 The Author(s). Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## 1. Introduction

Brain-Computer Interfaces (BCIs) are devices providing bidirectional communication channels between the brain and external devices. One of the primary uses of BCI technology

is in health scenarios, where clinicians acquire relevant information about the brain for diagnosis purposes (Lebedev and Nicolelis, 2017). Additionally, BCI systems enable artificial stimulation and inhibition of neuronal activity (Yao et al., 2019). In particular, neurostimulation has been used in a wide variety of medical scenarios, ranging from treating neurodegenerative diseases, such as Parkinson's or depression (Hartmann et al., 2019), to provide prosthetic users with feedback (O'Doherty et al., 2011). Within these systems,

\* Corresponding author.

E-mail address: [slopez@um.es](mailto:slopez@um.es) (S.López Bernal).

<https://doi.org/10.1016/j.cose.2021.102534>

0167-4048/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

there are two main categories based on their invasiveness. Non-invasive BCIs can externally stimulate the brain without surgery and, although some technologies can target small areas of the brain, non-invasive BCIs cover larger regions of the brain. In contrast, invasive systems can be applied to small areas, even with a single-neuron resolution, but introducing higher physiological risks (Ramadan and Vasilakos, 2017).

Based on the relevance and expansion of BCIs, new technologies and companies have emerged in recent years, focusing on developing new invasive systems to stimulate the brain with neuronal granularity. This is the case of Neuralink (Musk and Neuralink, 2019), a company that has designed disruptive BCI systems to record data at the neuronal level, and it is currently working on covering the stimulation functionality. Besides, Neural Dust (Seo et al., 2013) is an architecture of millions of nanoscale implantable devices located in the cortex that allow neural recording. Evolution of Neural Dust is the Wireless Optogenetic Nanonetworking device (WiOptND) (Wirdatmadja et al., 2017), which uses optogenetics to stimulate the neurons. Although these approaches are promising, the authors of Bernal et al. (2020) have shown that they have vulnerabilities that could allow attackers to control both systems and perform malicious stimulation actions, altering spontaneous neuronal signaling. Depending on the coverage of the attack, in terms of brain regions and number of neurons affected, cyberattackers could inflict permanent brain damage or even cause the death of the patients.

In the same direction, Bernal et al. (2021) identified that the field of cybersecurity in BCI is not mature enough, and non-sophisticated attacks can generate significant damage. In summary, the BCI vulnerabilities could be exploited by attackers to take advantage of these promising neurostimulation technologies. Taking the findings of these works as motivation, this manuscript focuses on the scarce research dealing with cyberattacks aiming to alter neuronal behavior. Additionally, new ways to measure and understand the impact of these attacks are also required. In particular, these issues gain special relevance due to the possibility of attacks being able to worsen or recreate the effects of common neurodegenerative diseases (Bernal et al., 2021).

Intending to improve the previous challenges, the main contribution of this work is the definition and implementation of a novel neuronal cyberattack, *Neuronal Jamming cyberattacks* (JAM), focused on the inhibition of neural activity. The present work aims to explore the impact that inhibitory neuronal cyberattacks can generate on the brain. Nevertheless, there is an absence in the literature of comprehensive neuronal topologies, and therefore, we simulate a portion of the visual cortex of mice, placed in the occipital region of the brain, defining a use case of a mouse trying to exit a given maze. The neuronal topology was built by using a Convolutional Neural Network (CNN) (Géron, 2019) trained to solve this particular use case. The second contribution of this work is the evaluation of the impact caused by JAM cyberattacks over both neuronal and artificial simulation in this specific scenario. To perform the analysis, we have used existing metrics but also defined a subset of new ones, concluding that JAM cyberattacks can alter spontaneous neuronal behavior and force the mouse to perform erratic decisions to escape the maze.

The third main contribution of this work is to compare the impact caused by JAM with an existing cyberattack named Neuronal Flooding (FLO) from the biological and artificial perspectives. We have observed that applying a FLO cyberattack over the last positions of the maze generates a reduction of its effectiveness from both biological and artificial approaches. Additionally, JAM cyberattacks are more damaging when increasing the number of consecutive positions under attack, translated into a reduction in the neural activity and an augmentation in the number of steps to find the exit. The fourth contribution is a comparison between biological and artificial scenarios based on linear correlation analysis between variables. In this sense, FLO presents a high Pearson correlation between experiments, of around 0.8, indicating a strong relationship. On its side, JAM presents worse results, which can be explained due to the particular restrictions during the implementation. Finally, we discuss the relationship that recent neuronal cyberattacks could have with neurodegenerative diseases.

These contributions aim to advance the current state of the art, which is limited to the references presented in this section. Compared to Bernal et al. (2020), which only characterized and measured the impact of two neural cyberattacks (Neural Flooding and Neural Scanning), this work further explores the impact of neural cyberattacks, presenting, for the first time, a comparison between the impact on neuronal and behavioral dimensions.

The remainder of the paper is structured as follows. Section 2 reviews the state of the art in cybersecurity oriented to BCI and neuronal cyberattacks. After that, Section 3 introduces the definition of the Neuronal Jamming cyberattack. Section 4 presents the experimental setup required to implement both JAM and FLO neuronal cyberattacks. Additionally, Section 5 and Section 6 describe, respectively, the results obtained after implementing JAM and FLO cyberattacks over multiple positions of the maze and the impact they cause. These two sections also include a comparison of the relationship between artificial and biological approaches. Subsequently, Section 7 discusses the impact that neuronal cyberattacks can have on neurodegenerative diseases. Finally, Section 8 presents conclusions and future work.

---

## 2. Related work

Cybersecurity applied to BCI is relatively recent, emerging in the last five years concepts such as brain-hacking or neurosecurity (Ienca, 2015; Ienca and Haselager, 2016). These publications identify that neurostimulation BCI devices present a high risk in patients' safety since an attacker could disrupt the treatment parameters. Additionally, they highlighted that attacks do not need to be complex to cause brain damage.

During these recent years, the academic literature has widely focused on the study of cybersecurity in health scenarios, aiming to preserve patients' privacy or improving the security of clinical devices (Huertas Celdrán et al., 2017; Huertas Celdrán et al., 2018). However, the literature has focused on particular cybersecurity aspects of BCI, mostly from theoretical and ethical perspectives. Although previous studies have highlighted the applicability of cryptographic and jam-

ming attacks (Ienca and Haselager, 2016), malware strategies (Bonaci et al., 2015), acquisition of sensitive data from neural signals (Quiles Pérez et al., 2021), disruption of neural signals (Martínez Beltrán et al., 2021), or potential attacks over BCI architectures (Ballarin Usieto and Minguez, 2018), these works are scarce and focus on particular privacy and security aspects, not addressing the physical safety dimension. Additionally, the authors of Takabi et al. (2016), Bonaci et al. (2015) identified that the platforms and frameworks used to develop BCI applications could be vulnerable to cyberattacks. Based on that, the authors of Bernal et al. (2021) performed a review of the state of the art in cybersecurity on BCI with a comprehensive analysis of physical safety issues, compiling already documented attacks over the BCI life-cycle, their impacts, and the countermeasures to detect and mitigate them. This work also studied the literature concerning attacks, impacts, and countermeasures from existing and prospecting architectural BCI deployments. Furthermore, they proposed applying well-known attacks, impacts, and countermeasures from the cybersecurity domain to BCI. In a nutshell, they identified an enormous absence of works addressing cybersecurity aspects in BCI technologies.

Regarding cyberattacks altering the behavior of neurons, the authors of Bernal et al. (2020) detected vulnerabilities in emerging neurostimulation technologies. They defined two neuronal cyberattacks, Neuronal Flooding (FLO) and Neuronal Scanning (SCA), aiming to disrupt the spontaneous behavior of the targeted zones of the brain. The FLO cyberattack consists in attacking, in a particular instant, a subset of neurons from the brain, while SCA targets one neuron per time instant, imitating the port scanning technique. They also defined several metrics to measure the impact of these attacks compared to spontaneous neuronal activity. In short, they identified that both neuronal cyberattacks induced a considerable alteration in the spontaneous neural signaling.

The neuronal cyberattacks presented in Bernal et al. (2020) demonstrate the feasibility of performing attacks over the brain aiming to disrupt its spontaneous neural activity. However, they do not explore the physiological or psychological consequences that an alteration in neural signaling can generate. In that direction, the authors of Bernal et al. (2021) theoretically proposed recreating the effect of neurodegenerative disorders such as Parkinson's and Alzheimer's diseases. For that, the neurostimulation system would be required to cover the brain regions naturally impacted by these diseases and present vulnerabilities that attackers can exploit. This work highlighted the high impact that recreating neurodegenerative disorders could have on users' physical safety.

To understand how cyberattacks could affect the brain and its relationship with degenerative diseases, it is essential to mention that, from a neurological point of view, most brain disorders are revealed as a dysfunction of communication between neurons or with other organs defining the term of *brain connectivity disorders*. Within this term, we can include neurodegenerative diseases. Alzheimer's Disease (AD) is a progressive neurodegenerative disorder that induces the degradation and death of brain cells. It seems that neurodegenerative diseases spread along structurally connected neural networks, known as *neuronal circuits*, presenting a functional relevance. There is a relationship between AD and changes in

neuronal activity in the Default Mode Network circuit (DMN), where parts of the DMN present increased connectivity at the beginning of the disease, indicating compensation for the failure of other regions of the circuit before they degenerate. During the progression of AD, the deactivation of the DMN is gradually more pronounced. Nevertheless, it is not clear if the circuit disruption is a cause or a consequence of the disease (Zott et al., 2018).

Amiotrophic lateral sclerosis (ALS) is a neurodegenerative disease affecting cortical and spinal neurons, which generates a loss of muscle control and paralysis. ALS is associated with a dysfunction of cortical circuits based on hyperexcitability of neuronal activity. Hyperexcitability can be understood as an exaggerated response to a stimulus, or the response to stimuli that generally do not induce a response. In this sense, ALS presents a perturbation in the excitatory/inhibitory balance, leading to pathological changes in cortical excitability (Brunet et al., 2020).

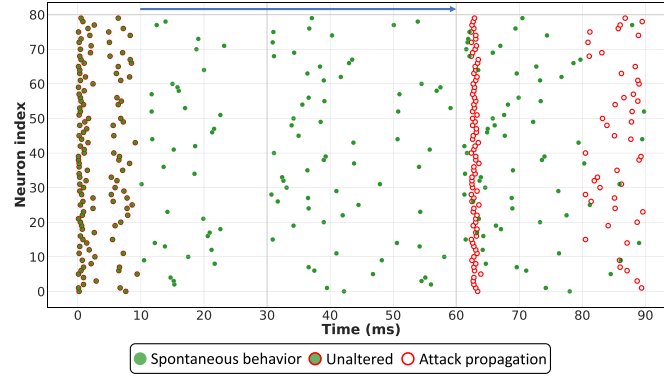
Despite the current knowledge about the behavior of neurodegenerative diseases, such as AD or ALS, there are no proposed cyberattacks in the literature trying to emulate the neuronal behavior of these conditions. Because of that, the current manuscript explores the possibility of inducing excitatory and inhibitory neuronal behavior to lay the foundation for future research aiming to recreate these conditions in the long term.

### 3. Neuronal Jamming cyberattack

This section presents the formal definition of the Neuronal Jamming cyberattack (JAM), including algorithmic and graphical representations to ease its understanding.

Jamming is a well-known cyberattack aiming to block the legitimate communication between elements of a system using malicious interference, resulting in the generation of a Denial of Service (DoS) over the communication. From a neurological perspective, we conceive a jamming cyberattack as an inhibition of the spontaneous activity of a set of neurons during a particular duration of time, preventing their interaction with other neurons. This attack does not need previous knowledge by the attacker about the status of the targeted neurons, presenting a low complexity compared to those that could require to study their previous and current status to determine the best instant to attack.

To formalize this attack, we denote  $\mathbb{NE} \subset \mathbb{N}$  as a subset of neurons from the brain, where  $n \in \mathbb{NE}$  expresses every single neuron.  $t^{\text{attk}}$  is the time instant when the cyberattack starts, and  $t^{\text{pulse}}$  is the duration of the attack. During that particular period, a subset of neurons  $\mathbb{AN} \subseteq \mathbb{NE}$  is attacked. The voltage of a single neuron in a specific instant of time is denoted as  $v_n \in \mathbb{R}$ , whereas  $v_{\min} \in \mathbb{R}$  indicates the minimum value of the voltage that the neuron can have, directly dependent on the neuronal model used in case of simulations. Moreover,  $t^{\text{win}}$  is the temporal window in which the cyberattack is evaluated, which corresponds to the duration of the simulation presented in subsequent sections.  $\Delta t$  is the amount of time between evaluations during the process, representing the duration of steps of the simulation in the implementation of the cyberattacks.



**Fig. 1** – Raster plot of a JAM cyberattack when the attack is performed between the instants 10ms and 60ms. This temporal window is represented by a blue arrow for clarity.

---

**Algorithm 1** JAM cyberattack execution.

---

```

t = 0
while t < twin do
  if t >= tattk AND t < (tattk + tpulse) then
    for all n ∈ ℕN do
      vn ← vmin
    end for
  end if
  t ← t + Δt
end while

```

---

As shown in [Algorithm 1](#), JAM cyberattacks are performed during a continuous duration of time, where the attacked neurons are forced to have their minimum voltage value. In other words, it avoids the targeted neurons to produce spikes, understood as the inhibition of the neurons.

To visually understand the behavior of a JAM cyberattack, [Fig. 1](#) presents the comparison between a JAM cyberattack and the spontaneous neuronal behavior for a simulation of 90ms. Until the instant 10 ms, green dots with a red outline can be appreciated, indicating that the attack has not altered those spikes. This attack, performed between the instants 10 ms and 60 ms, and indicated by a blue arrow, affects all 80 neurons represented in the figure. Because of that, during that temporal window, only green dots are presented, having an absence of neural activity during the application of the attack. After the instant 60ms, white dots with red outline appear, indicating the new spikes generated as a consequence of the attack. It is relevant to note that, from that moment until the instant 90ms, the neural signaling generated by the attack is completely different from the spontaneous behavior.

---

## 4. Experimental setup

Due to the lack of realistic and precise neuronal topologies in the literature, this section presents the methodology followed to create a neuronal topology used to evaluate the impact of JAM cyberattacks. For simplicity, we have summarized the explanations of this section, where a broader description is available at [Bernal et al. \(2020\)](#).

Nevertheless, it is relevant to indicate that the feasibility of neural cyberattacks was documented in [Bernal et al. \(2020\)](#), where we identified that novel neurostimulation technologies offering recording and neurostimulation capabilities with a single-neuron resolution, such as Neuralink, presented vulnerabilities that cyberattackers could exploit to gain access to the devices and, thus, disrupt the behavior of the brain. This work highlighted the sensitivity of using wireless communications, such as Bluetooth, between the implants and external devices controlling the implant. Thus, attackers could determine the instant (or instants) of attack, the list of targeted neurons, and the voltage used to affect the neurons.

It is essential to highlight that the knowledge of precise neocortical synaptic connections in mammalian is nowadays an open challenge [Gal et al. \(2017\)](#). Although artificial and biological networks cannot be comparable in complexity and functioning, there are works in the literature demonstrating that neurons in the visual cortex present certain similarities with a Convolutional Neural Network (CNN). In this sense, the visual recognition process operates incrementally in both networks, moving from simple to abstract ([Kuzovkin et al., 2018](#)). Based on that, we have trained a CNN using Keras on top of TensorFlow ([Chollet et al., 2015](#)) to solve a simplistic scenario based on a mouse trying to escape a maze from any position, inspired in the code from [Zafrany \(0000\)](#). The maze has a size of 7x7 positions with fixed obstacles that serve as walls, containing a single starting cell and an exit. [Fig. 2](#) presents the maze, indicating with numbers the optimal path to the exit, which has been determined during the training process of the CNN. It is essential to note that this process does not involve any real mouse since all this testing is based on simulations.

The CNN has been trained employing reinforcement learning ([Sutton and Barto, 2018](#)), using a topology consisting in three layers where the first two were convolutional layers, and the third one was dense. After the training process, a topology of interconnected nodes between layers was obtained, where each link had associated a filter weight. These weights represent the relevance that this connection has in the topology to solve the problem. [Table 1](#) summarizes the configuration used to define the CNN, composed of a total number of 276 nodes.

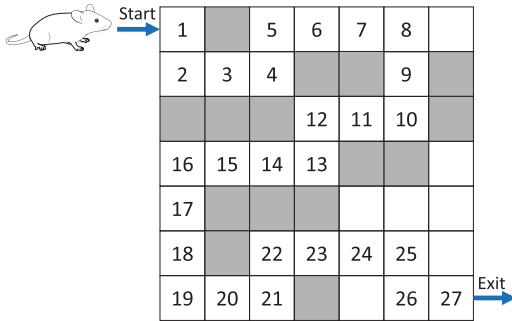
The resulting topology was translated to a biological neuronal network by keeping the exact number of layers and

**Table 1 – Summary of the layers of the CNN.**

Layer	Type	Filters	Input size	Output size	Kernel size	Stride	Activation function	Nodes
1	Conv2D	8	7×7×1	5×5×8	3×3	1	ReLU	200
2	Conv2D	8	5×5×8	3×3×8	3×3	1	ReLU	72
3	Dense	-	3×3×8	4	-	-	ReLU	4

**Table 2 – Parameters used in the Izhikevich model.**

Parameter	Description	Values
$v$	Membrane potential of a neuron	[-65, 30] mV
$u$	Membrane recovery variable providing negative feedback to $v$	(-16, 2) mV/ms
$a$	Time scale of $u$	0.02/ms
$b$	Sensitivity of $u$ to the sub-threshold fluctuations of $v$	0.2/ms
$c$	After-spike reset value of $v$	-65mV
$d$	After-spike reset value of $u$	8mV/ms
$I$	Injected synaptic currents	{10, 15} mV/ms

**Fig. 2 – Maze used to model the movement of the mouse, including the optimal path between the starting and final cells.**

nodes per layer and translating the filter weights to synaptic weights. These synaptic weights represent the influence that the firing of one neuron has on another neuron within a neuronal synapse. Particularly, this topology represents a small section of the visual cortex of a mouse, located in the occipital brain area. Once having the biological topology, we have used the Brian2 neural simulator (Stimberg et al., 2019) to represent the behavior of each individual neuron. In particular, we have implemented the Izhikevich neuronal model (Izhikevich, 2003), whose parameters are presented in Table 2, and Eqs. (1)–(3). It is relevant to highlight the functioning of the  $I$  parameter used in the experiments to model the visual stimuli received by the mouse in terms of free cells and walls in the biological simulation. To enclose the problem, we implemented and monitored a neuronal simulation with a total duration of 27 s, where the mouse stayed in one position of the optimal path for one second, and studied its spontaneous behavior and the behavior under attack. When the mouse is in a particular position, the *intervening neurons* associated with each *adjacent position* from the current cell were obtained. The concept of *intervening neurons* can be understood as the set of neurons influenced by the list of adjacent positions from the

**Table 3 – Parameters used in the analysis for JAM cyber-attacks.**

Parameter	Values
Number of consecutive attacked positions (Bio, CNN)	{1, 2, ..., 27}
Number of neurons/nodes (Bio, CNN)	{5, 35, 55, 75, 105}
Voltage under attack (Bio)	-65 mV
Output importance (CNN)	-1
Number of executions (Bio, CNN)	10

current cell. For those intervening neurons, the simulation assigns a value of 15mV/ms for the  $I$  parameter, keeping a value of 10mV/ms for the rest of the neurons. These particular implementation aspects are presented in-depth in Bernal et al. (2020).

$$v' = 0.04v^2 + 5v + 140 + u + I \quad (1)$$

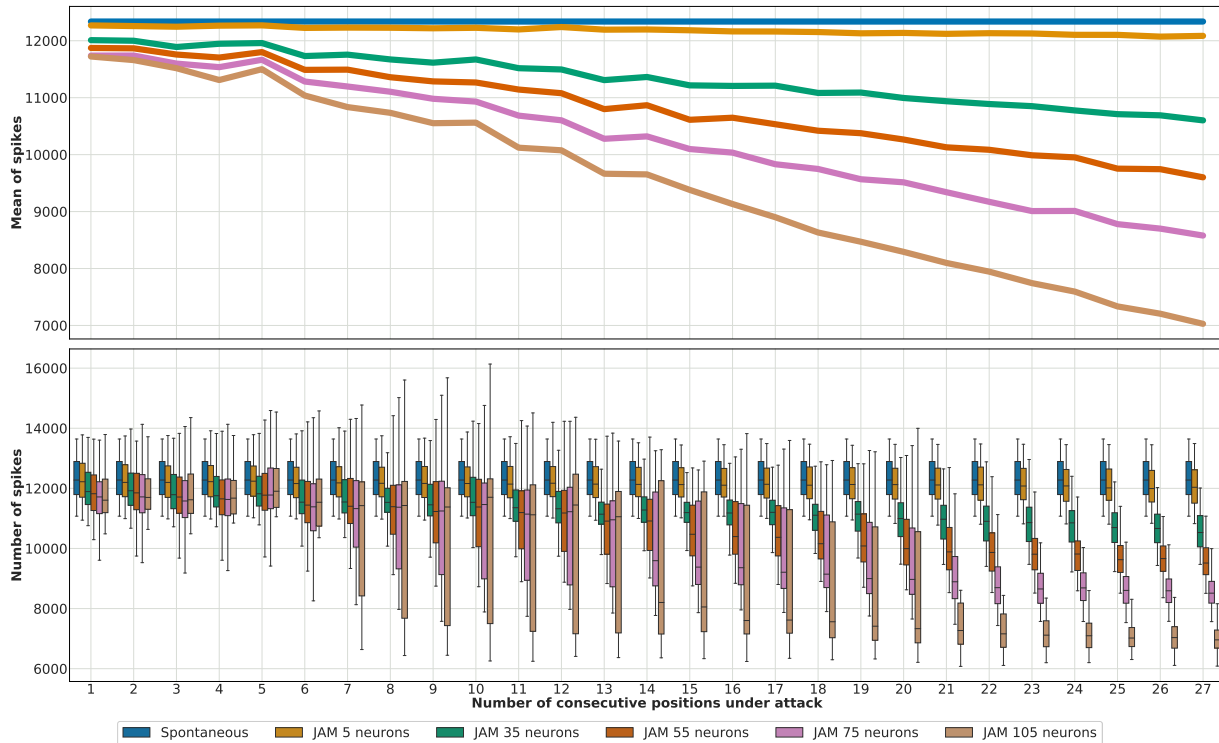
$$u' = a(bv - u) \quad (2)$$

$$\text{if } v \geq 30\text{mV, then } \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \quad (3)$$

## 5. Impact of JAM attacks over biological and artificial neural networks

Once explained the generation of the artificial and biological networks, this section measures and compares the impact generated by Neuronal Jamming cyberattacks (JAM) over biological and artificial networks. In particular, this analysis aims to study if an alteration in neuronal behavior can also impact the mouse's ability to solve the maze based on the evaluation of the CNN model.

Table 3 presents the parameters used to perform the experiments, indicating between parentheses if a parameter is



**Fig. 3 – Distribution of the number of spikes based on the consecutive number of positions attacked for JAM cyberattacks.**

common to both scenarios or specific to one of them. As can be seen, five number of simultaneously attacked neurons (named as nodes in the CNN) have been tested, probing several consecutively attacked positions ranging from one to all the positions of the optimal path of the maze. Additionally, each combination of parameters is executed ten times, where each execution targets a different set of randomly selected neurons. The meaning of these parameters will be presented throughout this section.

### 5.1. JAM cyberattacks over the biological network

Focusing on the biological perspective, attacked neurons are forced to the minimum voltage value of the model, which corresponds to  $-65$  mV, as indicated in Table 3. Fig. 3 presents the experiment consisting in augmenting the number of consecutive positions of the optimal path under attack, always initiating the attack in the first position, and evaluating different numbers of simultaneously attacked neurons. The variability shown corresponds to the ten executions performed per combination of parameters. In particular, this figure highlights how augmenting the number of consecutive positions of the labyrinth under attack impacts in terms of the number of spikes metric. The upper sub-figure depicts that increasing the number of simultaneously attacked neurons considerably reduces the mean of spikes, reaching a difference of 5000 spikes in the most damaging situation compared to spontaneous behavior. The bottom sub-figure shows that the distribution of the number of spikes presents small variability during the first six positions. More consecutive positions under attack generate a progressive reduction in the dispersion,

particularly for higher numbers of attacked neurons, indicating that JAM cyberattacks cause an enormous impact on the spike metric. Nevertheless, increasing the number of consecutive positions over more than 20 generates a progressive reduction in the distributions when attacking more than 75 neurons. This situation is explained by many neurons without activity during most of the simulation, decreasing their variability in the number of spikes.

Moving to the temporal dispersion of spikes, Fig. 4 depicts that attacking a higher number of neurons reduces the temporal dispersion. It is relevant to highlight that targeting a reduced number of neurons (up to 35) produces a slightly higher dispersion than the spontaneous behavior, where these peaks can be produced by the slight variations generated by the attack. Nevertheless, increasing the number of selected neurons gets a substantial reduction. In particular, attacking 105 neurons achieves the most damaging configuration, causing a reduction from 36% of instants with spikes to an approximate 28%. It is also important to note that, in the bottom sub-figure, the distribution of targeting 105 neurons significantly decreases compared to other numbers of attacked neurons, indicating the importance of this parameter of the attack.

### 5.2. JAM cyberattacks over the artificial network

In the artificial scenario, the attack consists in modifying the targeted nodes of the trained model, affecting their normal functioning. For that, the concept of *output importance* refers to the value used to alter the output of the nodes targeted by the attack, thus affecting their relevance in the network. In JAM, the value used to attack the nodes is  $-1$ , which indicates

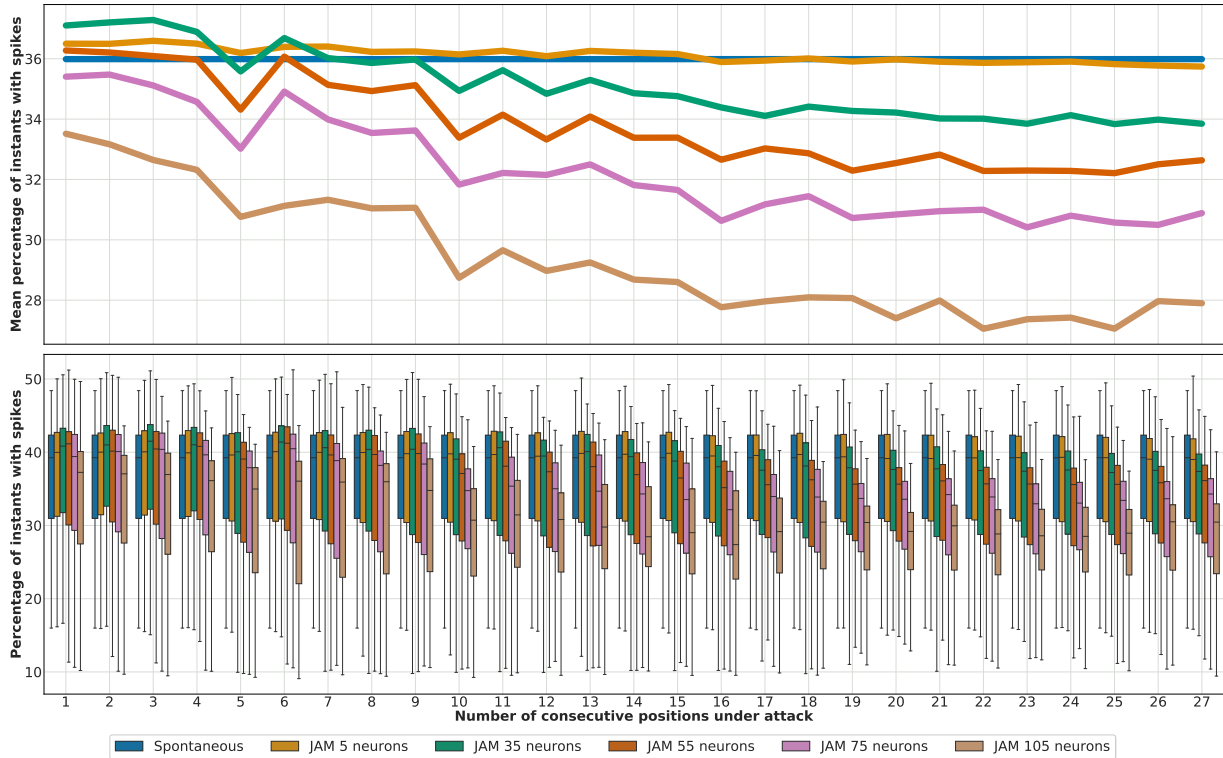


Fig. 4 – Distribution of the temporal dispersion based on the consecutive number of positions attacked for JAM cyberattacks.

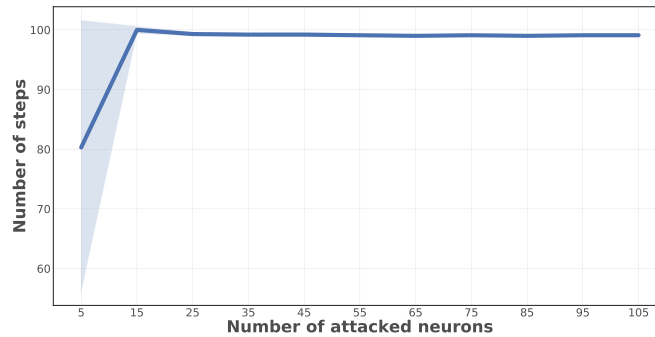


Fig. 5 – Number of steps for different number of neurons between five and 105, with ten executions, for JAM cyberattacks.

that those nodes do not have any relevance in the network, representing their inhibition. This forces the network to find alternative paths to solve the problem, deactivating the paths from the affected nodes to later layers.

The first approach followed was to apply the attacked model for the targeted consecutive positions, restoring it to the non-altered model after the duration of the attack. Although the mouse performed erratic decisions across the maze during the attack, once the model without attacks was restored, the mouse could always find the exit position ultimately. To better measure the impact of this attack in terms of percentage of success and number of steps, we decided to continuously perform the attack for all 27 positions of the maze. These experimentation results are represented in Fig. 5, which indicates that simultaneously attacking more than 15 nodes does not generate any difference since the number of steps

gets stabilized in around 100 steps. It is worthy to note that the success percentage is not studied as both variables are highly correlated, with a -0.99 Pearson correlation.

Based on the decision to attack during the whole simulation (27 positions), and compare these results with the biological simulation, we decided to focus the analysis of both scenarios on a number of attacked neurons between one and 20. From the CNN point of view, this decision is motivated by Fig. 6, which indicates that this particular range reflects variations in the number of steps and that further increments in this variable do not offer new variability.

After defining the range, the biological experiments were adapted to be comparable with those from the CNN scenario. For that, a number of attacked neurons between one and 20 were selected, setting the attack to cover all 27 consecutive positions of the optimal path of the maze, starting in the in-

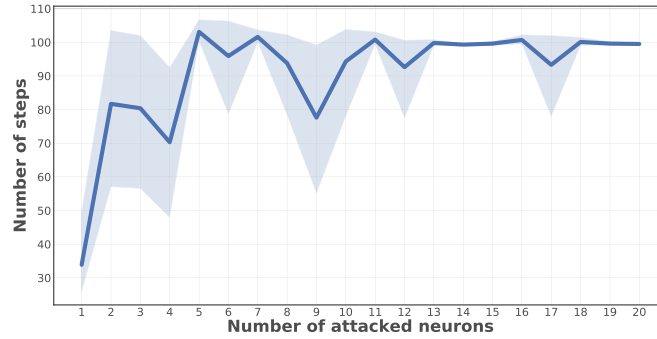


Fig. 6 – Number of steps for a range between one and 20 attacked neurons, with ten executions, for JAM cyberattacks.

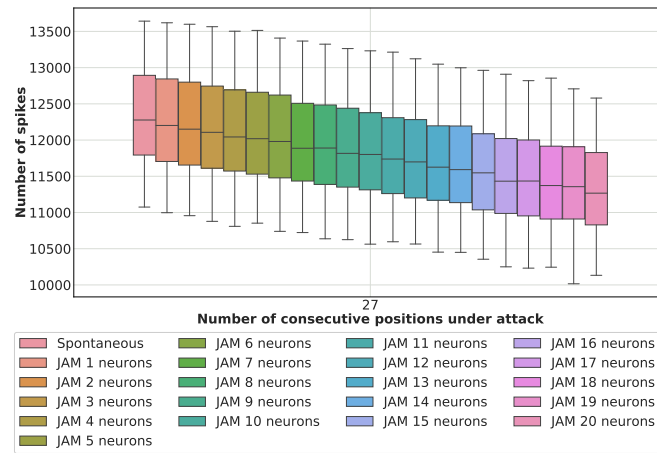


Fig. 7 – Number of spikes for a range of attacked neurons between one and 20 for JAM cyberattacks.

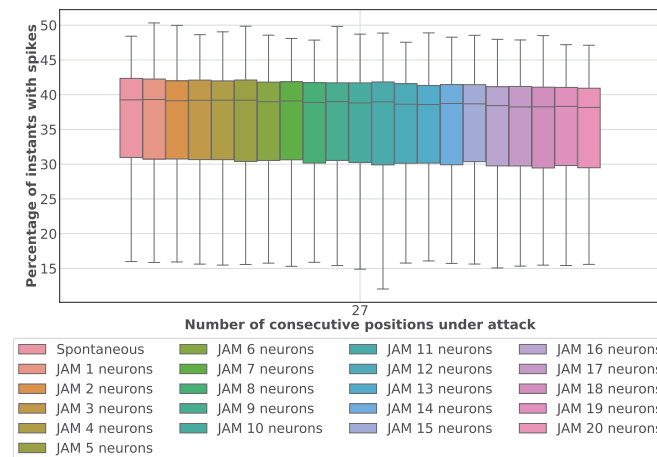


Fig. 8 – Temporal dispersion for a range of attacked neurons between one and 20 for JAM cyberattacks.

stant 50 ms. Figs. 7 and 8 present, respectively, the results for the number of spikes and the temporal dispersion. It is important to highlight that these plots present the same trend as described in Fig. 3 and Fig. 4, respectively, for the analysis between five and 105 attacked neurons.

Finally, Table 4 compares the Pearson correlation between both scenarios, which determines a correlation between the

number of steps and the number of spikes of -0.66, which indicates that these variables have a 66% linear correlation in an inversely proportional way. A similar situation happens between the number of steps and the percentage of dispersion, with a -0.59 Pearson correlation. This indicates, in general, a low correlation between scenarios. However, this can be explained due to the reduction in the number of attacked neu-

**Table 4 – Correlation of relevant features between CNN and biological experiments for JAM cyber attacks.**

	# spikes	% dispersion	# steps	# neurons
# spikes	1.00	0.98	-0.66	-0.99
% dispersion	0.98	1.00	-0.59	-0.98
# steps	-0.66	-0.59	1.00	0.66
# neurons	-0.99	-0.98	0.66	1.00

rons considered. As indicated before, the number of neurons has been limited to a range between one and 20. Although these values offer variability in the CNN, there is not much difference in the distribution between these close sizes in the biological simulation.

Nevertheless, the individual analysis performed in this section for both biological and artificial scenarios presents the high impacts that JAM cyberattacks generate over these scenarios. The Spearman correlation values have also been calculated, studying the non-linearity of the data. Since the values obtained were similar to those presented for the Pearson correlation, we opted to include the latter for concision.

Finally, it is interesting to present the performance of the attacked model in terms of ROC curves. First, it is essential to highlight that the model has four different outputs (up, down, left, right), corresponding to the direction to perform the next step within the maze. Based on that, the ROC curves present the relationship between erroneous and correct predictions when the model is not under attack and when different configurations of the attacks are applied.

Focusing on JAM cyberattacks, and since they affect multiple positions, it is not possible to know the number of steps correctly performed to obtain the True Positive Rate (TPR) and False Positive Rate (FPR). Based on that limitation, we could assume a TPR equal to zero and FPR of 1, according to the configuration of the attack.

## 6. Comparison of JAM and FLO cyberattacks

This section compares the impact caused by JAM cyberattacks with FLO, a neuronal cyberattack existing in the literature. For that, we first introduce FLO cyberattacks, moving to the analysis of their impacts, and later we compare it with JAM. This section also provides an in-depth study of the results of individually performing FLO cyberattacks in different positions of the optimal path, comparing the results of biological and artificial networks.

### 6.1. Definition of Neuronal Flooding cyberattacks

Neuronal Flooding cyberattacks (FLO) were defined in our previous work (Bernal et al., 2020) as a way to overstimulate targeted neurons. In that work, we just explored the cyberattacks for the first position of the maze, whose behavior is formally represented by Algorithm 2. In particular, it indicates that the attack over the targeted neurons is performed in a particular instant of time  $t^{attk}$ , in contrast to JAM, which is executed within a determined temporal period.

**Algorithm 2** FLO cyberattack execution.

```

t = 0
while t < twin do
  if t == tattk then
    for all n ∈ AN do
      vn ← vn + vin
    end for
  end if
  t ← t + Δt
end while

```

**Table 5 – Parameters used in the analysis for FLO cyberattacks.**

Parameter	Values
Positions attacked (Bio, CNN)	{1, 2, ..., 27}
Number of neurons/nodes (Bio, CNN)	{5, 35, 55, 75, 105}
Voltage increment (Bio)	40 mV
Output importance (CNN)	60 %
Number of executions (Bio, CNN)	10

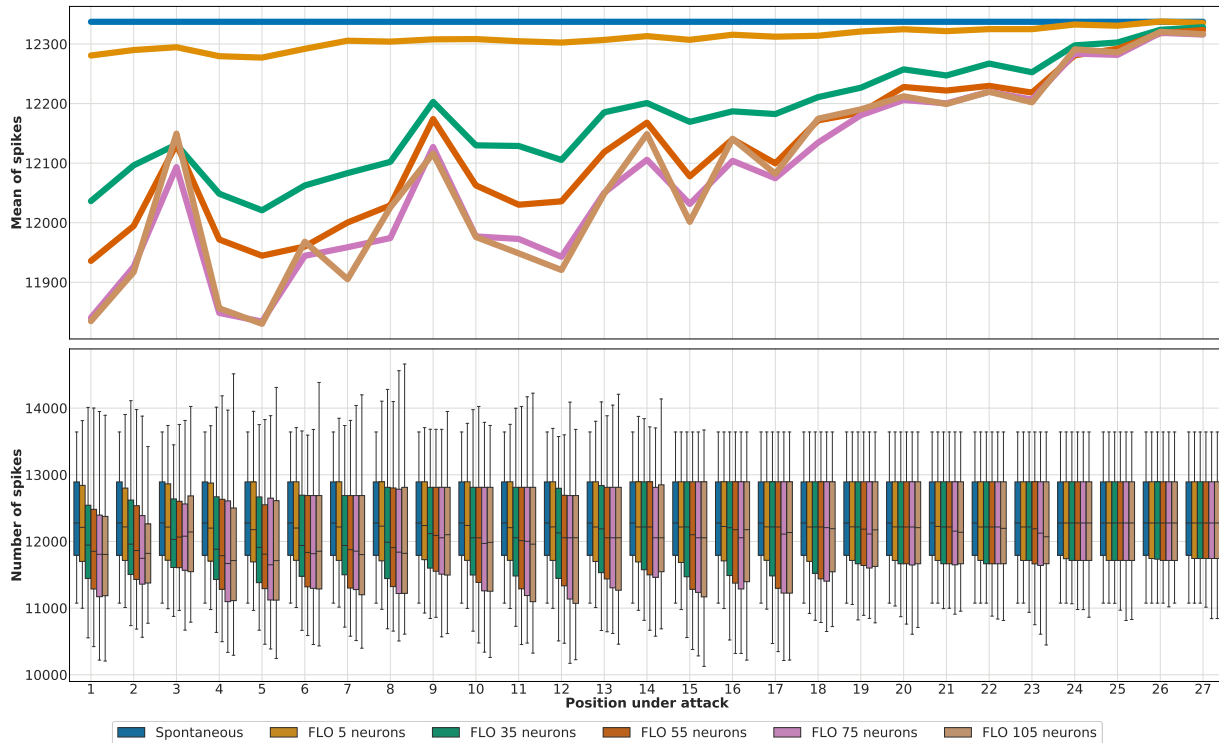
In contrast, the current work performs FLO cyberattacks over each individual position of the optimal maze path, evaluating a different number of simultaneously attacked neurons and multiple increment voltages per position. The parameters used for this experiment are indicated in Table 5, having five different values of simultaneously attacked neurons (or nodes) and a single value of voltage increment. The use of just one voltage value is based on the experiments performed in Bernal et al. (2020), which concluded that, for FLO cyberattacks, the usage of different voltages did not have a substantial impact. Besides, each combination of parameters is executed ten times.

### 6.2. FLO cyberattacks over the biological network

In the biological scenario, we perform a FLO cyberattack individually over each position of the optimal path of the maze, at the instant 50ms after reaching a targeted position, evaluating the impact of the attack during the complete simulation (27 s, until the mouse reaches the exit) based on the number of spikes and temporal dispersion metrics.

Fig. 9 presents the evolution of the number of spikes according to the individual position of the optimal path under attack. As previously indicated, the voltage used to increment the targeted neurons is 40 mV.

It is worthy to note that, for each attacked position, the represented values correspond to the number of spikes over the complete simulation. The upper sub-figure presents the mean of spikes for each position under attack, where each line represents a different number of attacked neurons. The effect of FLO cyberattacks to reduce the temporal dispersion was already documented in Bernal et al. (2020). In Fig. 9, we can observe that performing the attack in later positions of the optimal path generates a lower impact since in the positions before the cyberattack the spikes are not altered and, thus, the spiking behavior is the same as the spontaneous behavior. Particularly, it can be observed that attacking 105 neurons in



**Fig. 9 – Distribution of the number of spikes according to the position under attack for FLO cyberattacks.**

the first position generates an approximate reduction of 500 spikes. These results also indicate that this attack causes a desynchronization of neuronal activity over time, presenting a higher variability when the attack is performed in the first positions. This variability is also benefited by the particular model used and the propagation of the spikes.

Additionally, attacking a broader number of neurons produces, in general, a higher reduction in the mean of spikes. Nevertheless, we can observe no significant differences between attacking 75 and 105 simultaneous neurons in terms of the mean of spikes. Regardless of these similarities, there are variations in their maximum and minimum values, indicating variations in their distributions. These data correspond to the mean of the distribution represented in the bottom sub-figure, where we can see a higher variability in the number of spikes when the attack is applied in the first positions. This figure also highlights that the maximum and minimum values of the distribution have a significant variability compared to the spontaneous behavior, stabilized when we attack in later positions.

After analyzing the behavior of the FLO cyberattack in terms of the number of spikes, Fig. 10 presents its impact focusing on the temporal dispersion metric. As can be seen, the dispersion is higher when attacking the first positions due to the same reasons addressed for the number of spikes metric. Additionally, attacking a broader number of neurons derives in a higher percentage of instants with spikes. Specifically, simultaneously attacking 75 neurons reaches the highest impact, augmenting the initial 36% of instants with spikes to an approximate 40%. Finally, it is worthy to note that these two

metrics are highly related, with a Pearson correlation value of -0.97.

### 6.3. FLO cyberattacks over the artificial network

In terms of attacks over the CNN, it is essential to note that the voltage increment used to attack the biological network has been proportionally adapted to the CNN scenario, corresponding to the output importance indicated in Table 5. Based on that, the value of 40mV used in the biological scenario represents a 60% from the voltage range defined by the Izhikevich model used. This 60% is the equivalent value used to increment the importance of the targeted nodes during the attack to the CNN.

Fig. 11 presents the evolution of the mean number of steps among the ten executions per number of consecutively attacked nodes. This figure indicates the impact caused by attacking the mouse when it is placed in each individual position of the optimal path of the maze. When the simulated mouse is placed in a particular position, we obtain the number of steps required to reach the exit from the position attacked. To this resulting number of steps, we add the number of steps correctly performed until the attacked position, which corresponds to correctly performed decisions before the attack. It is essential to note that, once the model is attacked, it is used until the end of that particular execution.

In this figure, each color indicates a different number of simultaneous neurons attacked. It can be appreciated that the number of steps remains constant in the spontaneous behavior of the CNN, requiring 26 steps to find the exit. These 26

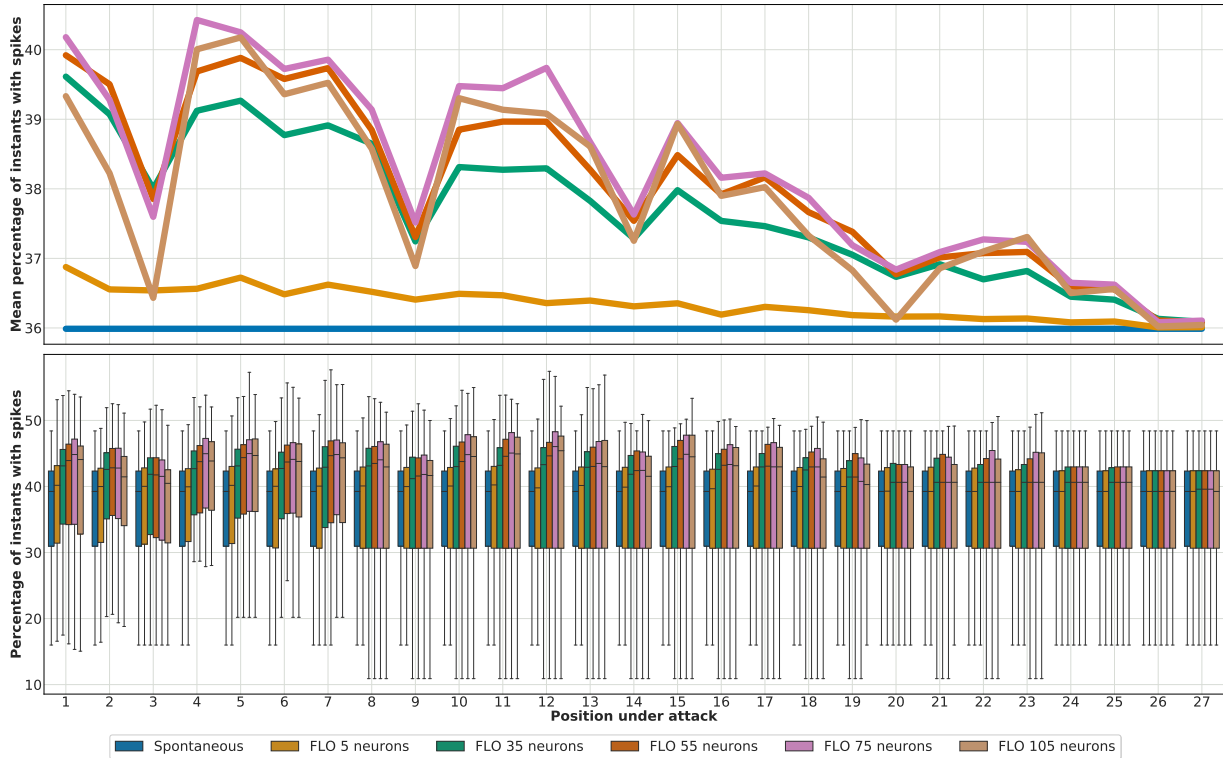


Fig. 10 – Percentage of instants with spikes according to the position under attack for FLO cyberattacks.

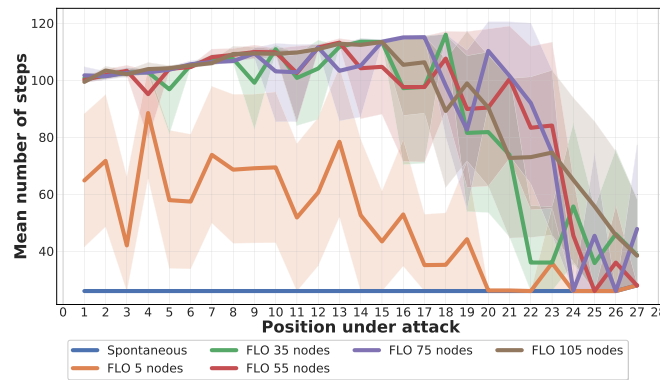


Fig. 11 – Mean of steps when we perform a FLO cyberattack in each position of the optimal path of the maze, considering five different number of simultaneously attacked neurons.

steps are determined by the model resulting from training the CNN, which concluded an optimal path of 27 positions to exit the exit and, thus, 26 steps between them. There is an exception in position 27, where the mouse needs to move to an adjacent cell in the maze to finally reach the exit since the mouse initially started in the exit position. This figure highlights that augmenting the number of attacked neurons increases the number of steps until position 21. From that position, the trend decreases since the closer the mouse is to the exit, the easier it is to solve the maze by probability, even if the mouse suffers an alteration in its decision ability.

Another relevant metric to study this situation is the percentage of times in which the mouse finds the exit. The Pearson correlation has been calculated between the number of

steps and the success rate, obtaining a value of -0.99, meaning that they present a trend almost identical in an inversely proportional way. That is to say, we have observed that the number of steps increases when the percentage of success decreases. Based on that, the number of steps will be the sole metric used to evaluate the CNN in this analysis.

It is interesting to consider the relationship between the results obtained from attacking the biological and artificial scenarios to help understand the behavior in the biological network. To perform this comparison, Table 6 presents the Pearson correlation between the relevant features considered in these domains. In particular, we are interested in the relationship between the number of steps and the number of spikes, and between the number of steps and the percentage of dis-

**Table 6 – Correlation of relevant features between CNN and biological experiments for FLO cyberattacks.**

	position of attack	# spikes	% dispersion	# steps	# neurons
position attack	1.00	0.53	-0.53	-0.42	-0.0
# spikes	0.53	1.00	-0.97	-0.82	-0.66
% dispersion	-0.53	-0.97	1.00	0.81	0.56
# steps	-0.42	-0.82	0.81	1.00	0.65
# neurons	-0.0	-0.66	0.56	0.65	1.00

persion. Based on that, it can be determined that the CNN and biological approaches have a high correlation, with an approximate 80% correlation in both of them.

Based on the above, we can conclude a significant relationship between the results obtained in both experimental dimensions. These results suggest that performing attacks over the brain of the mouse could not only alter its spontaneous neuronal behavior but also affect its decisions to solve the maze, increasing the number of steps to find the exit and decreasing its chances to exit the maze. Nevertheless, these results are limited to our use case, the neuronal topology, and the use of a CNN to model a portion of the mouse's visual cortex.

Once presented the relationship between the biological and artificial scenarios, this section compares the results of both attacks. Since the approaches followed between these attacks are not directly comparable, where FLO focuses on individually attacking different positions and JAM affects multiple consecutive positions, this study focuses on analyzing the correlations obtained for each attack. In FLO, the Pearson correlation obtained was -0.82 for the relationship between the number of steps and number of spikes and 0.81 between steps and temporal dispersion. On the contrary, a value of -0.66 was obtained between the steps and the spikes and -0.59 for the relationship between steps and dispersion for JAM. These values indicate that the relationship between the biological and artificial networks is closer in the FLO situation, despite the analysis for the JAM cyberattack presented some limitations as stated in Section 5.

Finally, and as previously presented for JAM cyberattacks, we offer the performance of the attacked model based on ROC curves. In particular, for FLO cyberattacks, we have obtained two ROC curves. The first curve presents the TPR and FPR for aggregation of positions 24 to 27. We have included this range since in these positions, the mouse is able, on average, to always exit the maze (see Fig. 11). This ROC curve, subsequently presented in Fig. 12, indicates that since the mouse can always find the exit of the maze, the TPR will always be 1. Moreover, the FPR ranges from close to zero (perfect value) when attacking five simultaneous nodes to more than 0.8 when attacking 105. The FPR is determined based on the number of decisions incorrectly taken compared to the decisions performed by the spontaneous behavior.

The second ROC curve obtained for FLO presents an aggregation between positions one to 23 since we can observe in Fig. 11 that performing attacks in those positions is more damaging, and thus, the mouse is not always able, on average, to exit the maze. Because of that, the TPR decreases, where attacking five neurons presents the best TPR. From its part, the

FPR is considerably high for a number of simultaneously attacked neurons higher than five, as presented in Fig. 13.

## 7. Neural cyberattacks and neurodegenerative diseases

This section discusses the results obtained in this work, aiming to understand the impact of these attacks better, their possible consequences in the real world, and defend against them. Additionally, if we could reproduce the effect of neurodegenerative diseases with these attacks, we could generate databases containing multiple attack configurations, study their impact, and propose mechanisms to reduce these impacts.

Previous sections have highlighted the enormous impact that neuronal cyberattacks can cause over spontaneous neuronal activity, affecting the amount, periodicity, and even the presence of spikes. Additionally, we have observed that these cyberattacks could also alter the simulated mouse's decision ability, forcing it to make mistakes in the resolution of the labyrinth. Furthermore, these cyberattacks possess differences based on their action mechanisms. JAM cyberattacks focus on continuously inhibiting the neuronal activity of the targeted neurons, suppressing this signaling along with the duration of the attack. On the contrary, FLO cyberattacks aim to overstimulate a set of neurons in a particular instant, extending its impact after its application.

Based on these action mechanisms, we identify that the behavior of the previous attacks has similarities with the effects and consequences that certain neurodegenerative diseases generate. As indicated in Section 2, neurodegenerative diseases can be included within the concept of brain connectivity disorders. In particular, for Alzheimer's Disease (AD), the deactivation of the Default Mode Network (DMN) could be reproduced by an attacker able to target individual neurons, reproducing or accelerating the effects of the disease. We identify that JAM, focused on neuronal activity inhibition, could be used for these purposes. On the contrary, Amyotrophic lateral sclerosis (ALS) is based on neuronal activity hyperexcitability, where FLO could be applied to periodically stimulate the targeted neurons and thus produce a perturbation in the excitatory/inhibitory balance of cortical neurons.

Although neuronal cyberattacks are promising mechanisms aiming to extend our knowledge about cybersecurity on BCI, further research is required to study the impact these cyberattacks can cause over neural circuits and cognitive and behavioral functions. The study of neuronal cyberattacks could help identify particular characteristics helping to detect

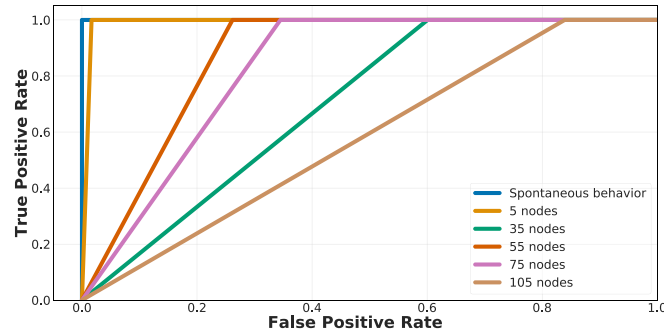


Fig. 12 – ROC curve for an aggregation of positions 24 to 27 of the optimal path of the maze.

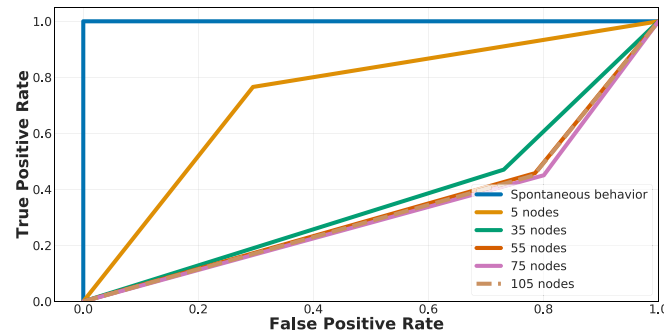


Fig. 13 – ROC curve for an aggregation of positions one to 23 of the optimal path of the maze.

prospect threats on BCI systems. Additionally, the application of neuronal cyberattacks could be beneficial in neurological research, using these cyberattacks to control the spread of the disease in neural models or even in vivo trials.

## 8. Conclusion

This work introduces the Neuronal Jamming cyberattack (JAM), consisting in the inhibition of neuronal activity. To implement this attack, and due to a lack of realistic neuronal topologies, a Convolutional Neural Network (CNN) has been trained to generate a neuronal topology based on a use case of a mouse trying to exit a maze. Once having both topologies, we analyze the impact that JAM cyberattacks present over biological and artificial scenarios. Additionally, this manuscript offers a comparison between JAM and FLO cyberattacks. For that, we have implemented several configurations of FLO, a cyberattack already existing in the literature aiming to overstimulate neural activity. To measure their impact, we have studied multiple metrics in the biological scenario (number of spikes and temporal dispersion) and in the CNN (number of steps and success rate in solving the problem).

The obtained results highlight that, in JAM cyberattacks, increasing the number of consecutive positions under attack reduces the spikes and temporal dispersion. In the artificial network, attacking up to 20 nodes is enough to prevent the mouse from completing the labyrinth. Moreover, a contribution of this work is the comparison between scenarios based

on the study of linear correlation between variables. This analysis indicates that this attack could affect the mouse's ability to escape the maze. We have obtained a Pearson's correlation of 0.6, a low value explained due to the restriction of the number of neurons used to compute the correlations.

Additionally, we have observed for FLO experiments that delaying the instant of attack to later positions reduces the impact from both biological metrics. Moreover, delaying the attack until position 21 generates an increase in the number of steps. From this position, delaying the instant of attack decreases the number of steps since it is more probable to find the exit by probability. Pearson's correlation between variables for this cyberattack was approximately 0.8, highlighting a closer relationship between scenarios. Finally, we have discussed the similarities between neurodegenerative diseases and the neuronal cyberattacks studied.

In future work, we plan to investigate new neuronal cyberattacks with different action mechanisms and impacts. Additionally, since the main limitation of this work is the use of a neuronal topology extracted from a CNN, we aim to explore the possibility of having realistic topologies, which are currently very limited, to simulate existing and prospecting cyberattacks. Finally, as the present work only focuses on the characterization of these cyberattacks, we want to focus our efforts on designing and implementing detection mechanisms to identify the initiation of a neuronal cyberattack and propose mitigation techniques to reduce their impact or even neutralize it.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRedit authorship contribution statement

**Sergio López Bernal:** Methodology, Writing – original draft, Data curation, Software. **Alberto Huertas Celdrán:** Methodology, Conceptualization, Writing – review & editing. **Gregorio Martínez Pérez:** Supervision, Project administration, Funding acquisition.

## Acknowledgment

This work has been partially supported by (a) Bit & Brain Technologies S.L. under the project CyberBrain, associated with the University of Murcia (Spain), by (b) the Swiss Federal Office for Defense Procurement (armasuisse) with the CyberSpec (CYD-C-2020003) project, and by (c) the University of Zürich UZH.

## REFERENCES

- Ballarin Usieto, P., Minguez, J., 2018. Avoiding brain hacking - challenges of cybersecurity and privacy in Brain Computer Interfaces.
- Bernal SLópez, Huertas Celdrán A, Fernández Maimó L, Barros MT, Balasubramaniam S, Martínez Pérez G. Cyberattacks on miniature brain implants to disrupt spontaneous neural signaling. *IEEE Access* 2020;8:152204–22. doi:[10.1109/ACCESS.2020.3017394](https://doi.org/10.1109/ACCESS.2020.3017394).
- Bernal SLópez, Huertas Celdrán A, Martínez Pérez G, Barros MT, Balasubramaniam S. Security in brain-computer interfaces: state-of-the-art, opportunities, and future challenges. *ACM Comput. Surv.* 2021;54(1). doi:[10.1145/3427376](https://doi.org/10.1145/3427376).
- Bonaci T, Calo R, Chizeck HJ. App stores for the brain : privacy and security in Brain-Computer Interfaces. *IEEE Technol. Soc. Mag.* 2015;34(2):32–9. doi:[10.1109/ETHICS.2014.6893415](https://doi.org/10.1109/ETHICS.2014.6893415).
- Brunet A, Stuart-Lopez G, Burg T, Scekcic-Zahirovic J, Rouaux C. Cortical circuit dysfunction as a potential driver of amyotrophic lateral sclerosis. *Front. Neurosci.* 2020;14:363. doi:[10.3389/fnins.2020.00363](https://doi.org/10.3389/fnins.2020.00363).
- Chollet, F., et al., 2015. Keras. <https://keras.io>.
- Gal E, London M, Globerson A, Ramaswamy S, Reimann MW, Muller E, Markram H, Segev I. Rich cell-type-specific network topology in neocortical microcircuitry. *Nat. Neurosci.* 2017;20(7):1004–13. doi:[10.1038/nn.4576](https://doi.org/10.1038/nn.4576).
- Géron A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media; 2019.
- Hartmann CJ, Fliegen S, Groiss SJ, Wojtecki L, Schnitzler A. An update on best practice of deep brain stimulation in Parkinson's disease. *Ther. Adv. Neurol. Disord.* 2019;12:1756286419838096.
- Huertas Celdrán A, Gil Pérez M, García Clemente FJ, Martínez Pérez G. Preserving patients' privacy in health scenarios through a multicontext-aware system. *Ann. Telecommun.* 2017;72(9):577–87. doi:[10.1007/s12243-017-0582-7](https://doi.org/10.1007/s12243-017-0582-7).
- Huertas Celdrán A, Gil Pérez M, García Clemente FJ, Martínez Pérez G. Sustainable securing of medical cyber-physical systems for the healthcare of the future. *Sustain. Comput. Inform. Syst.* 2018;19:138–46. doi:[10.1016/j.suscom.2018.02.010](https://doi.org/10.1016/j.suscom.2018.02.010).
- Ienca M. Neuroprivacy, neurosecurity and brain-hacking: emerging issues in neural engineering. *Bioeth. Forum* 2015;8(2):51–3.
- Ienca M, Haselager P. Hacking the brain: brain-computer interfacing technology and the ethics of neurosecurity. *Eth. Inf. Technol.* 2016;18(2):117–29. doi:[10.1007/s10676-016-9398-9](https://doi.org/10.1007/s10676-016-9398-9).
- Izhikevich EM. Simple model of spiking neurons. *IEEE Trans. Neural Netw.* 2003;14(6):1569–72. doi:[10.1109/TNN.2003.820440](https://doi.org/10.1109/TNN.2003.820440).
- Kuzovkin I, Vicente R, Petton M, Lachaux J-P, Baciau M, Kahane P, Rheims S, Vidal JR, Aru J. Activations of deep convolutional neural networks are aligned with gamma band activity of human visual cortex. *Commun. Biol.* 2018;1(1):107. doi:[10.1038/s42003-018-0110-y](https://doi.org/10.1038/s42003-018-0110-y).
- Lebedev MA, Nicolelis MAL. Brain-machine interfaces: from basic science to neuroprostheses and neurorehabilitation. *Physiol. Rev.* 2017;97(2):767–837. doi:[10.1152/physrev.00027.2016](https://doi.org/10.1152/physrev.00027.2016).
- Martínez Beltrán ET, Quiles Pérez M, López Bernal S, Huertas Celdrán A, Martínez Pérez G. Noise-based cyberattacks generating fake p300 waves in brain-computer interfaces. *Clust. Comput.* 2021. doi:[10.1007/s10586-021-03326-z](https://doi.org/10.1007/s10586-021-03326-z).
- Musk E, Neuralink. An integrated brain-machine interface platform with thousands of channels. *bioRxiv* 2019. doi:[10.1101/703801](https://doi.org/10.1101/703801). <https://www.biorxiv.org/content/early/2019/08/02/703801.full.pdf>
- O'Doherty JE, Lebedev MA, Ifft PJ, Zhuang KZ, Shokur S, Bleuler H, Nicolelis MAL. Active tactile exploration enabled by a brain-machine-brain interface. *Nature* 2011;479:228–31. doi:[10.1038/nature10489](https://doi.org/10.1038/nature10489).
- Quiles Pérez M, Martínez Beltrán ET, López Bernal S, Huertas Celdrán A, Martínez Pérez G. Breaching subjects' thoughts privacy: a study with visual stimuli and brain-computer interfaces. *J. Healthc. Eng.* 2021;2021:5517637. doi:[10.1155/2021/5517637](https://doi.org/10.1155/2021/5517637).
- Ramadan RA, Vasilakos AV. Brain computer interface: control signals review. *Neurocomputing* 2017;223:26–44. doi:[10.1016/j.neucom.2016.10.024](https://doi.org/10.1016/j.neucom.2016.10.024).
- Seo, D., Carmena, J. M., Rabaey, J. M., Alon, E., Maharbiz, M. M., 2013. Neural dust: An ultrasonic, low power solution for chronic brain-machine interfaces. *arXiv:1307.2196*.
- Stimberg M, Brette R, Goodman DF. Brian 2, an intuitive and efficient neural simulator. *eLife* 2019;8:e47314. doi:[10.7554/eLife.47314](https://doi.org/10.7554/eLife.47314).
- Sutton RS, Barto AG. *Reinforcement Learning: An Introduction. second. The MIT Press; 2018.*
- Takabi H, Bhalotiya A, Alohaly M. Brain computer interface (BCI) applications: privacy threats and countermeasures. In: *Proceedings of the 2016 IEEE 2nd International Conference on Collaboration and Internet Computing, IEEE CIC 2016; 2016.* p. 102–11. doi:[10.1109/CIC.2016.24](https://doi.org/10.1109/CIC.2016.24).
- Wirdatmadja SA, Barros MT, Koucheryavy Y, Jornt JM, Balasubramaniam S. Wireless optogenetic nanonetworks for brain stimulation: device model and charging protocols. *IEEE Trans. NanoBiosci.* 2017;16(8):859–72. doi:[10.1109/TNB.2017.2781150](https://doi.org/10.1109/TNB.2017.2781150).
- Yao L, Sheng X, Mrachacz-Kersting N, Zhu X, Farina D, Jiang N. Sensory stimulation training for BCI system based on somatosensory attentional orientation. *IEEE Trans. Biomed. Eng.* 2019;66(3):640–6. doi:[10.1109/TBME.2018.2852755](https://doi.org/10.1109/TBME.2018.2852755).
- Zafrazy, S., Deep reinforcement learning for maze solving.
- Zott B, Busche MA, Sperling RA, Konnerth A. What happens with the circuit in Alzheimer's disease in mice and humans? *Annu. Rev. Neurosci.* 2018;41(1):277–97. doi:[10.1146/annurev-neuro-080317-061725](https://doi.org/10.1146/annurev-neuro-080317-061725). PMID: 29986165



**Sergio López Bernal** received the B.Sc. and M.Sc. degrees in computer science from the University of Murcia, and the M.Sc. degree in architecture and engineering for the IoT from IMT Atlantique, France. He is currently pursuing the Ph.D. degree with the University of Murcia. His research interests include ICT security on braincomputer interfaces and network and information security.



**Alberto Huertas Celdrán** received the M.Sc. and Ph.D. degrees in computer science from the University of Murcia, Spain. He is currently a postdoctoral fellow associated with the Communication Systems Group (CSG) at the University of Zurich UZH. His scientific interests include medical cyber-physical systems (MCPS), braincomputer interfaces (BCI), cybersecurity, data privacy, continuous authentication, semantic technology, context-aware systems, and computer networks.



**Gregorio Martínez Pérez** is Full Professor in the Department of Information and Communications Engineering of the University of Murcia, Spain. His scientific activity is mainly devoted to cybersecurity and networking, also working on the design and autonomous monitoring of real-time and critical applications and systems. He is working on different national (14 in the last decade) and European IST research projects (11 in the last decade) related to these topics, being Principal Investigator in most of them. He has published 160+ papers in national and international conference proceedings, magazines and journals.