



OPEN ACCESS

EDITED BY

Minoru Asada,
Osaka University, Japan

REVIEWED BY

Shuangming Yang,
Tianjin University, China
Beck Strohmmer,
Technical University of Denmark,
Denmark

*CORRESPONDENCE

Yi Zeng
yi.zeng@ia.ac.cn

RECEIVED 22 August 2022

ACCEPTED 30 November 2022

PUBLISHED 20 December 2022

CITATION

Feng H and Zeng Y (2022) A
brain-inspired robot pain model based
on a spiking neural network.
Front. Neurobot. 16:1025338.
doi: 10.3389/fnbot.2022.1025338

COPYRIGHT

© 2022 Feng and Zeng. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

A brain-inspired robot pain model based on a spiking neural network

Hui Feng^{1,2} and Yi Zeng^{1,2,3,4*}

¹Brain-inspired Cognitive Intelligence Lab, Institute of Automation, Chinese Academy of Sciences, Beijing, China, ²School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China, ³Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China, ⁴National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China

Introduction: Pain is a crucial function for organisms. Building a “Robot Pain” model inspired by organisms’ pain could help the robot learn self-preservation and extend longevity. Most previous studies about robots and pain focus on robots interacting with people by recognizing their pain expressions or scenes, or avoiding obstacles by recognizing dangerous objects. Robots do not have human-like pain capacity and cannot adaptively respond to danger. Inspired by the evolutionary mechanisms of pain emergence and the Free Energy Principle (FEP) in the brain, we summarize the neural mechanisms of pain and construct a Brain-inspired Robot Pain Spiking Neural Network (BRP-SNN) with spike-time-dependent-plasticity (STDP) learning rule and population coding method.

Methods: The proposed model can quantify machine injury by detecting the coupling relationship between multi-modality sensory information and generating “robot pain” as an internal state.

Results: We provide a comparative analysis with the results of neuroscience experiments, showing that our model has biological interpretability. We also successfully tested our model on two tasks with real robots—the alerting actual injury task and the preventing potential injury task.

Discussion: Our work has two major contributions: (1) It has positive implications for the integration of pain concepts into robotics in the intelligent robotics field. (2) Our summary of pain’s neural mechanisms and the implemented computational simulations provide a new perspective to explore the nature of pain, which has significant value for future pain research in the cognitive neuroscience field.

KEYWORDS

brain-inspired intelligent robot, robot pain, spiking neural network, free energy principle, spike-time-dependent-plasticity

1. Introduction

Pain is vital to individual organisms’ survival and the social life between organisms (Walters and Williams, 2019). The International Association for the Study of Pain (IASP) defines pain as “an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage” (Raja et al., 2020).

Pain can alert to actual injuries and prevent potential injuries, help organisms protect themselves (Hardcastle, 1997; Loeser and Melzack, 1999), and synergistically trigger various cognitive functions, such as pain memory and pain empathy (Jackson et al., 2006; Wiech and Tracey, 2013; Asada, 2019). Designing Robot Pain inspired by the organisms' pain has essential implications for the survival and longevity of robots.

Previous works related to robots and pain have focused on how robots recognize pain signals, such as humans' painful expressions, and realize pain empathy for humans (Cao et al., 2021; Werner et al., 2022). Kuehn and Haddadin (2016) designed the robot's neural reflex behavior for pain, which is only a pain-induced avoidance response and not a human-like pain capacity. Sur and Amor (2017) attempted to model the robot's pain capacity, but the network structure they used is different from the biological mechanism. We explored the neural mechanisms of pain and established a Brain-inspired Robot Pain Spiking Neural Network (BRP-SNN) to simulate the brain regions involving pain, using the spike-timing-dependent-plasticity (STDP) learning rule to train the connection weights. Our model makes the robots have human-like pain capacity and has greater biological plausibility.

We believe that the definition of Robot Pain should be inspired by the nature of organisms' pain. First, the pain mechanism in organisms has evolved over thousands of years, and it is closely related to physical injury (Bagnato et al., 2015). Broom proposed that pain is the neural activity at the brain level that accompanies physical injury (Broom, 2001). This neural activity occurs primarily in the anterior cingulate cortex (ACC; Frankland and Teixeira, 2005). In addition, this neural activity can be associated with injury-related cues, such as scenes of physical injury or a dangerous object. When a similar cue occurs, the brain will generate the same neural activity and avoid potential injury in time (Karsdorp and Vlaeyen, 2009). This phenomenon is called pain memory (Eich et al., 1985). Therefore, pain is a passive response to actual physical injury and active response to potential physical injury (Sur and Amor, 2017). We argue that the Robot Pain should be designed in line with the neural mechanism of pain mentioned above—it should first respond to actual machine injury and then respond to potential machine injury.

For the cognitive process of actual physical injury, we take inspiration from the brain's Free Energy Principle (FEP; Friston, 2010). FEP proposes that the brain constantly makes predictions about the outside world, and all senses will receive real sensory information from the external world. If the predictions are consistent with the real sensory information, the brain is in a low-entropy state; if the predictions are inconsistent with the real sensory information, the brain is in a high-entropy state (Friston et al., 2006; Karl, 2012). Physical injury and other unexpected phenomena all belong to the high-entropy state (Peters et al., 2017), which indicates that the brain's predictions and sensations are inconsistent at a certain level. As the main

brain region associated with pain, the ACC has also been confirmed to be associated with prediction error of different levels by multiple neuroscience studies (Oliveira et al., 2008; Castellar et al., 2010; Jessup et al., 2010). Inspired by the FEP and ACC function, we summarize a functional connection map of brain regions for cognizing actual physical injury and generating pain experience. In addition, we explored the cognitive processes for potential physical injuries. The experience of pain can be associated with injury-related cues, such as scenes or sounds of injury. The visual or auditory cortex of the brain fires accordingly and establishes synaptic connections with the ACC through associative learning. When a similar cue reappears, the brain uses these synaptic connections to generate a rapid pain experience and avoid potential injury (Wiech and Tracey, 2013). We simulated the pain-related mechanisms mentioned above with a Spiking Neural Network (SNN), and experimentally validated our model on a real robot.

This paper proposed a Brain-inspired Robot Pain Spiking Neural Network (BRP-SNN) model inspired by the neural mechanisms of organisms' pain. This model has three major contributions: (1) We summarized a neural mechanism of pain emergence from the perspective of pain evolution and the brain's Free Energy Principle (FEP). (2) We use an SNN to build a Brain-inspired Robot Pain model that can respond to both actual and potential injury, and obtain a result curve similar to the neuroscience experiment. (3) We apply the proposed model to the robot and complete the alerting actual injury task and the preventing potential injury task.

2. Materials and methods

2.1. The neural mechanism of pain

Most of the existing neuroscience literature describes the neural mechanism of pain: nociceptive stimuli specifically activate nociceptors in the skin, and nociceptive information is transmitted through the spinal cord, thalamus, and brainstem to the Anterior Cingulate Cortex (ACC) and other areas of the cerebral cortex, forming a painful experience. It is a specific pathway that has evolved over thousands of years of evolution (Sneddon, 2019). However, this pathway is not sufficient to support the modeling of Robot Pain.

Broom (2001) proposed that the emergence of pain was first related to body injury during evolution. Pain is a neural activity accompanying the body injury, and it is preserved and internalized in the brain because of its survival benefits. Nociceptors have also evolved (Perl and Kruger, 1996). In further learning, corresponding brain regions (e.g., visual and auditory cortex) will capture cues related to the injury when the body injury actually occurs. These brain regions will be connected to pain regions through associative learning (Schlund

et al., 2011; Wiech and Tracey, 2013). This connection ensures a rapid avoidance response to a potential body injury.

The first step of pain emergence is the cognition of the actual body injury. The brain's Free Energy Principle (FEP) proposes that organisms are constantly maintaining the low-entropy state, and abnormal phenomena are the high-entropy state (Friston et al., 2006; Karl, 2012; Peters et al., 2017), such as body injury. Entropy cannot be quantified, but Free Energy (FE) can be approximately equal to entropy. Therefore, the FEP can be seen as the theoretical basis for cognitive actual body injury. Bogacz describes the details of the FEP and gives a formula derivation (Bogacz, 2017): the brain can never directly obtain the real state ϕ of the external world, but can only constantly estimate the external world state $\hat{\phi}$, and the body's senses can receive real sensory information s about the real external world state and use it to verify brain's estimate. FE in Bogacz (2017) is simplified as a negative logarithm of the joint probability distribution of external state estimation and real sensory information $-\ln p(s, \hat{\phi})$, which means that the greater the joint probability distribution, the more accurate the estimate, the smaller the FE, and the smaller the entropy of the organism. The FE expression is expanded with the Bayesian formula $p(s, \hat{\phi}) = p(s|\hat{\phi})p(\hat{\phi})$:

$$\begin{aligned} FE &= -\ln p(s, \hat{\phi}) \\ &= -\ln p(s|\hat{\phi}) - \ln p(\hat{\phi}) \end{aligned} \quad (1)$$

The probability distribution is replaced with a Gaussian distribution $f[s; g(\cdot), \sigma]$ with mean $g(\cdot)$ and variance σ , supposing $\sigma = 1$:

$$\begin{aligned} FE &= -\ln f_1(s; g_s(\hat{\phi}), \sigma_s) - \ln f_2(\hat{\phi}; \mu_\phi, \sigma_\phi) \\ &= \frac{[s - g_s(\hat{\phi})]^2}{2\sigma_s} + \frac{[\hat{\phi} - \mu_\phi]^2}{2\sigma_\phi} - \frac{1}{2} \ln \sigma_s - \frac{1}{2} \ln \sigma_\phi \quad (2) \\ &= \frac{1}{2} [s - g_s(\hat{\phi})]^2 + \frac{1}{2} [\hat{\phi} - \mu_\phi]^2 \end{aligned}$$

For the final derived expression, the first item represents the prediction error, where $g_s(\cdot)$ is a generative function representing the mapping between world states and sensory information that needs to be learned empirically in advance. The brain actively estimates the state of the world $\hat{\phi}$ and predicts the sensory information through the generation function $g_s(\hat{\phi})$, and the senses receive the real sensory information s , subtracting the two parts can produce prediction error. The second item represents the prior error, which indicates that the brain's estimates must be compared with the prior knowledge μ_ϕ stored in the brain. Prediction error and prior error together determine the value of the FE. The FE has three main functions: 1. It can be used to guide the brain to change its estimates. 2. It can be used to guide the brain to perform actions. 3. As an internal variable, it can reflect the internal state of the brain and can be used to study a variety of cognitive functions.

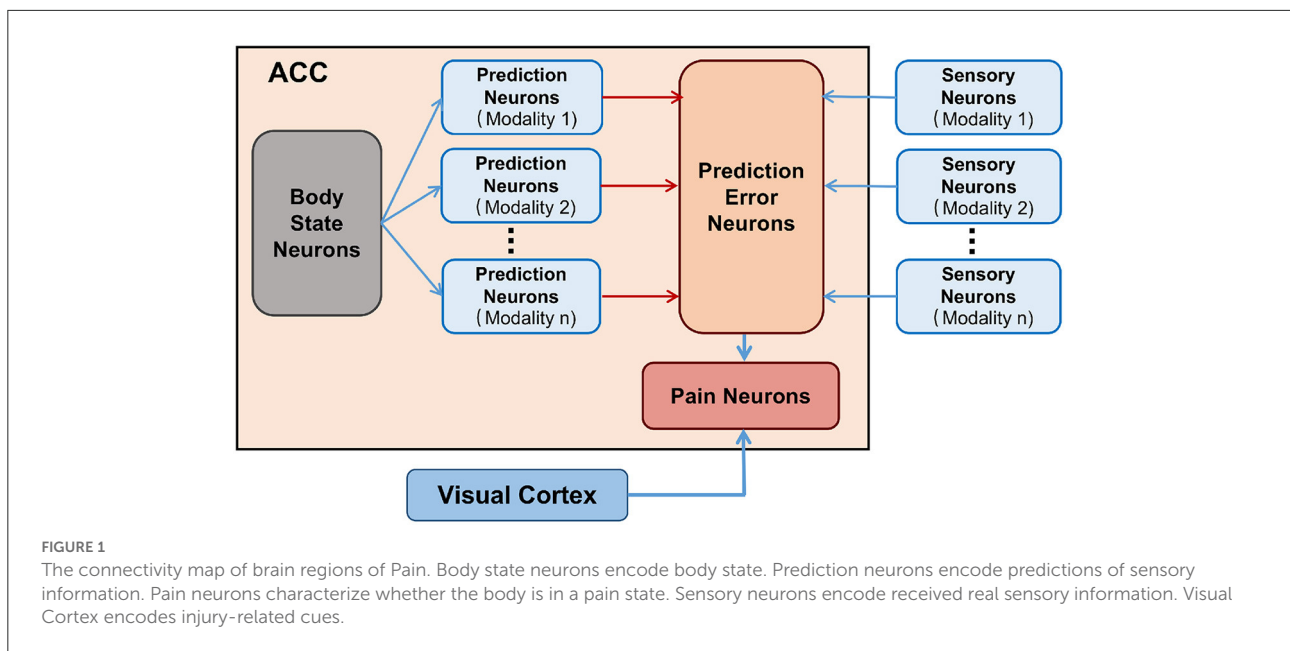
The cognitive process of body injury and pain emergence can be explained by FEP: The brain can predict the body sensory information of all modalities $[s_1, s_2, \dots, s_n]$ based on the current body state $\hat{\phi}$. This body state can be known based on prior knowledge stored by the brain (e.g., the previous moment's body state ϕ' and the performed action a'). Equation (3) shows the calculated rule of the FE in this scenario. If FE is 0, it means that the body is in a normal state. If FE is >0 , it means that the prediction error arises and the body is in an injured state, which leads to the pain experience.

$$\begin{aligned} FE &= [s_1 - g_{s_1}(\hat{\phi})]^2 + [s_2 - g_{s_2}(\hat{\phi})]^2 + \dots \\ &\quad + [s_n - g_{s_n}(\hat{\phi})]^2 + [\hat{\phi} - g_\phi(\phi', a')]^2 \end{aligned} \quad (3)$$

It is worth noting that the brain has many prediction processes at different levels in addition to the above prediction from body state to body sensation. In the case of an injection, the prediction of whether or not be pain is essentially an event-level prediction, which is a prediction of the pain event in the context where the pain has evolved. This is a different level from the prediction of cognitive body injury we described above.

Numerous neuroscientific studies have shown that pain is associated with the Anterior Cingulate Cortex (ACC) of the brain (Davis et al., 1997; Frankland and Teixeira, 2005; Du et al., 2020). The ACC has also been shown to be responsible for prediction error computation and conflict detection (Silvetti et al., 2013; Wiech and Tracey, 2013). Silvetti et al. proposed that the ACC contains neurons representing event prediction and neurons representing prediction errors. They argued that the ACC can receive feedback from the external environment, which together with the prediction neurons leads to the firing of the prediction error neurons (Silvetti et al., 2011). Inspired by the FEP and ACC function, we summarize the connectivity map of brain regions that cognize actual physical injuries and generate pain experiences.

As shown in Figure 1. The body state neurons of ACC constantly infer self body state and predict multi-modality sensory information, and the corresponding prediction neurons of ACC are activated. The senses of the body can receive the real sensory information and the corresponding sensory neurons are activated. When the body state is normal, the information from the sensory neurons and the prediction neurons are consistent, which does not cause the prediction error neurons of ACC to fire, and the body in a low-entropy state. When the body is injured, it definitely causes the sensory information of some modalities to be inconsistent with the predicted information, causing prediction error neurons to fire and the brain cognizes that the body is injured. Then pain neurons are activated to fire and produce pain experiences. The body in a high-entropy state. We also investigated the neural mechanisms associated with pain memory. Here we use vision as an example, as shown in the gray module in Figure 1. Visual Cortex has a direct connection to the pain-related region in the ACC (Wiech and Tracey, 2013).



The Visual Cortex will capture the injury-related cues when the body injury occurs. Due to the temporal correlation, the weight connections between the corresponding visual cortex neurons and the pain-related neurons of the ACC will be enhanced by association learning. The brain will use these connections to identify potential injury and avoid it.

2.2. Brain-inspired robot pain spiking neural network

This subsection introduces the implementation details of the BRP-SNN inspired by the neural mechanisms of pain. We first introduce the basic tools and methods, then introduce the overall architecture of the BRP-SNN.

2.2.1. Neuron model and synapse learning method

The emergence of a spiking neural network (SNN) has facilitated the development of neuromorphic computing (Yang et al., 2021). In recent years, SNNs have been successfully applied to many aspects, such as meta-learning and few-shot learning (Yang et al., 2022b,c), working memory and decision making (Zhao et al., 2018; Yang et al., 2022a). The Leaky Integrate-and-Fire (LIF) neuron is the most common neuron model for SNN (Tal and Schwartz, 1997; Gerstner and Kistler, 2002), and we use it as the basic unit of our BRP-SNN. The LIF neuron dynamics are described by the following Equations (3) and (4):

$$\tau_m \frac{du}{dt} = -[u_t - u_{rest}] + RI(t) \tag{4}$$

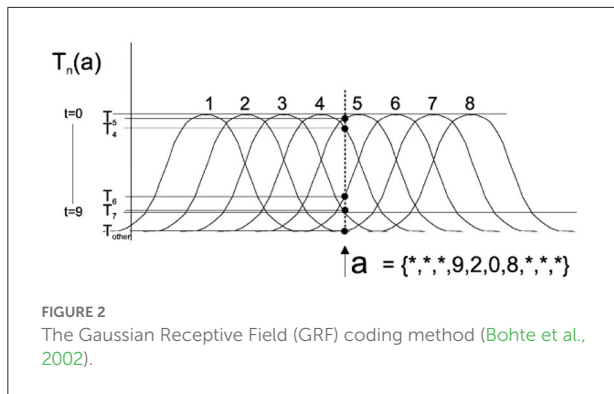
$$\lim_{\delta \rightarrow 0} u(t_f^+ + \delta) = u_{reset} \tag{5}$$

u_t is the membrane potential of the neuron at time t , u_{rest} is the membrane potential at steady-state, R is the resistance, $I(t)$ is input current, and τ_m is the time constant. When the membrane potential u_t exceeds a certain threshold u_{th} , the neuron fires, and t_f is the firing time. Once the neuron fired, the membrane potential returns to its reset state u_{reset} . In this paper, the parameters of the LIF neuron model are: $u_{rest} = u_{reset} = -65mV$, $\tau_m = 10ms$, $u_{th} = -50mV$.

We use spike-timing-dependent-plasticity (STDP) as a synapse learning rule to update synaptic weights. STDP is the most basic learning method in the brain, which relies on the time difference between the firing of pre-synaptic neurons and post-synaptic neurons to train the synaptic weight (Dan and Poo, 2004; Sjöström and Gerstner, 2010; Markram et al., 2012; Chevchenko and Ludermir, 2020). The weights update rule is written as Equation (5):

$$\Delta\omega = \begin{cases} A^+ \exp(\frac{\Delta t}{\tau_+}), & \Delta t < 0 \\ -A^- \exp(\frac{-\Delta t}{\tau_-}), & \Delta t > 0 \end{cases} \tag{6}$$

With $\Delta t = t_i - t_j$, where t_i and t_j are the firing time of the pre-synaptic neuron and the post-synaptic neuron respectively. A^+ and A^- are the learning rates. τ_+ and τ_- are the time constants. According to this rule, the connections will be strengthened when pre-synaptic neurons fire before post-synaptic neurons and will be weakened when pre-synaptic neurons fire after post-synaptic neurons. In this paper, $\tau_+ = \tau_- = 10ms$, $A^+ = 0.5$, $A^- = 0.1$. The choice of these parameters is based on the settings in reference (Sjöström and



Gerstner, 2010). The values of τ_+ and τ_- are consistent with the values in that literature. The A^+ and A^- reflect the strength of LTP and LTD, respectively, and the effect of LTP is considered more in this paper.

2.2.2. Population coding method

Population coding is a neuron coding method of SNN. The representation information of one neuron is limited, and the population coding is more effective and biologically plausible (Fang and Zeng, 2021; Fang et al., 2021).

Population coding uses a neuron population to encode input data, which can represent continuous values. In this paper, we adopted the Gaussian receptive field (GRF) population coding method mentioned in Bohte et al. (2002). Each dimensional receptive field is a group of neurons, where each neuron corresponds to a Gaussian activation function. Each input will cause all neurons in the receptive field to fire at different times through a different Gaussian activation function, Figure 2 shows eight neurons per dimensional receptive field. For each input pattern, the response time t is calculated for each neuron, and neurons with response times >9 are coded as not triggered because they are considered insufficiently excitable. The single input $a = \{*, *, *, 9, 2, 0, 8, *, *, *\}$ was encoded by eight Gaussian activation functions, and the black dot marking T_i indicates the firing time of the i -th neuron involved in its encoding.

2.2.3. Model implementation

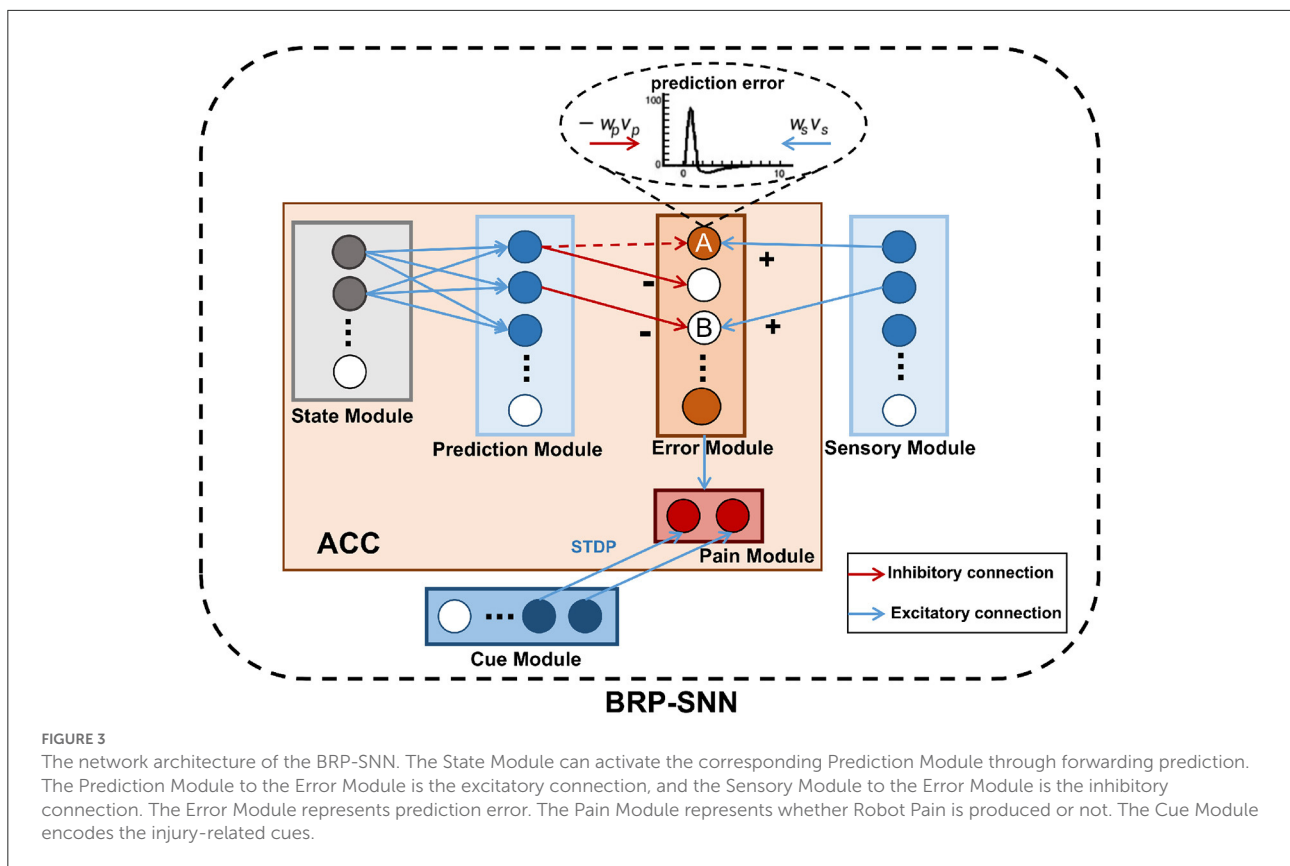
When body states are known, the prediction error generated by the prediction process from the body state to body sensations of all modalities can lead to pain. According to this mechanism, the specific cognitive process of Robot Pain is as follows: the robot continuously infers the current body state through the previous moment state and the executed action, and predicts the sensory information that will be received by all robot sensors. When the prediction error occurs, it indicates body injury, which in turn leads to Robot Pain.

As shown in Figure 3, The BRP-SNN simulates the function and connection of the brain regions mentioned in Figure 1, which contains the State Module, Prediction Module, Sensory Module, Error Module, Pain Module, and Cue Module. Each module is a neuron population. The State Module encodes different body states. The Prediction Module represents the predictive sensory information of different modalities. The Sensory Module encodes the real sensory information of different modalities. The Error Module represents prediction error. The Pain Module represents whether Robot Pain is produced or not. The Cue Module encodes the injury-related cues. The blue arrows in Figure 3 indicate excitatory connections, and the red arrows in Figure 3 indicate inhibitory connections.

The mapping relationship between the State Module and the Prediction Module needs to be learned empirically by collecting robot body data. The connection weights are established through STDP in the process of learning. After learning, the State Module can activate the corresponding Prediction Module through forwarding prediction. The Error Module represents the prediction error, once any neuron in this module fires, indicating that the prediction error arises. According to the excitatory and inhibitory synaptic connections in the ACC mentioned in Silveti et al. (2011), we designed that the Prediction Module to the Error Module is the inhibitory connection, and the Sensory Module to the Error Module is the excitatory connection. The dynamics of neurons in the Error Module are described by Equation (6):

$$\Delta u = -\omega_p v_p + \omega_s v_s (\omega_p, \omega_s > 0) \tag{7}$$

Δu represents an update of the membrane potential of the neurons in the Pain Module. $-\omega_p$ represents inhibitory weights. ω_s represents excitatory weights. v_p is the input from the Prediction Module. v_s is the input from the Sensory Module. When the firing pattern of the Prediction Module is consistent with the Sensory Module, the effects of excitatory and inhibitory on the neuron of the Error Module are counteracted, and the Error Module does not fire, such as the B neuron in Figure 3, indicating a normal body state. When the firing pattern of the Prediction Module is inconsistent with the Sensory Module, the effects of excitatory and inhibitory are not counteracted and specific neurons in the Error Module will fire. For example, the excitatory effect from the Sensory Module acts on A neuron, but the inhibitory effect from the Prediction Module acts on the other neuron. The red dashed arrow in Figure 3 indicates an inhibitory effect that has not yet been established. A neuron receives only excitatory effects, so it fires. The Error Module to the Pain Module is excitatory and fully connected. When the robot body state is known, any neuron fire of the Error Module causes the Pain Module to fire, indicating the Robot Pain state.



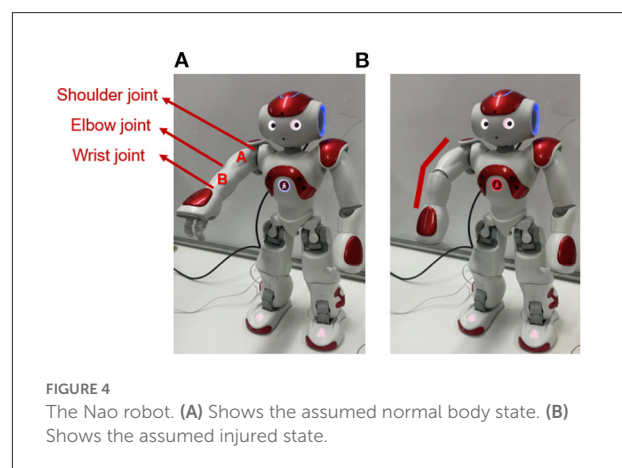
The Cue Module encodes the injury-related cues, such as a dangerous object captured by the camera. The connections between the Cue Module and Pain Module can be established through STDP by association learning. This BRP-SNN can not only recognize actual machine injury and produce Robot Pain, but also recognize injury-related cues and avoid potential machine injury.

3. Experiment

3.1. Experiment settings

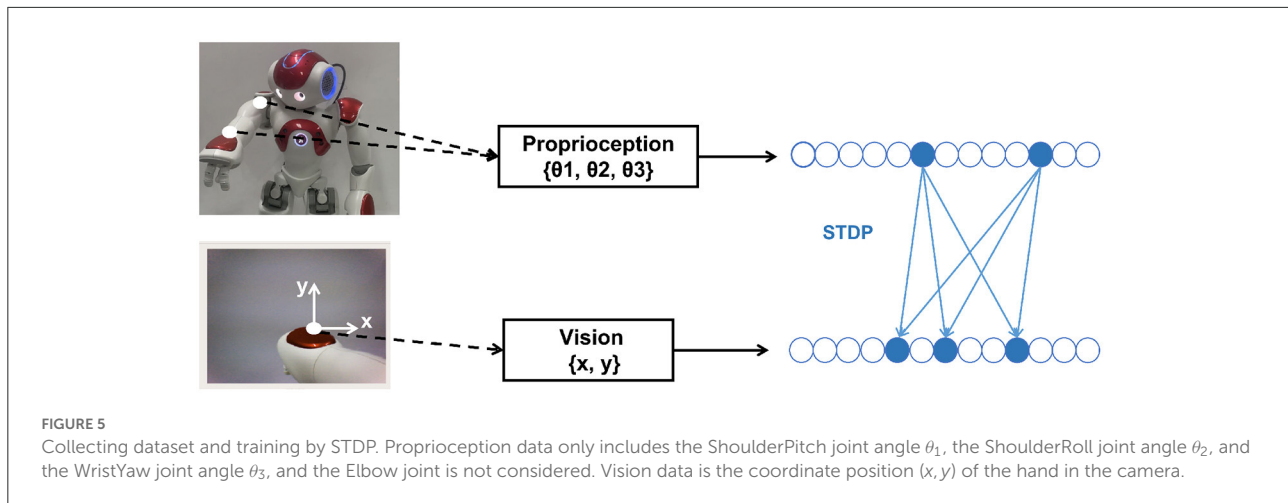
We use the Nao robot as an experimental platform and apply the BRP-SNN to implement two tasks: the alerting actual injury task and the preventing potential injury task.

As shown in Figure 4A, the left arm of the Nao robot contains three joints: shoulder joint, elbow joint, and wrist joint, and contains two links: link A and link B. In order to mimic the robot injury process without actually damaging the robot. Our experiments assume that the left arm is in a normal body state when it is straight and in an injured state when it is bent. This means that it is assumed that link A and link B are an integral link. The elbow joint sensor (between linkA and linkB) will not receive any action command from the robot, and the robot



will not process any input information from this elbow joint sensor. Under this assumption, the bending elbow as shown in Figure 4B is an undesirable deformation of the robot that is similar to a human fracture.

In this experiment, we only consider two sensory modalities of the robot: proprioception s_p and vision s_v , which were the angle data of each joint of the arm and the position coordinate data of the hand in the camera. As in Equation (8), a special case



of Equation (3).

$$FE = [s_v - g_{s_v}(\hat{\phi})]^2 + [s_p - g_{s_p}(\hat{\phi})]^2 + [\hat{\phi} - g_{\phi}(\phi', a')]^2 \quad (8)$$

The State Module encodes the body state. Lanillos and Cheng (2018) propose that the robot should determine body state through multi-modal sensations using the FEP, corresponding to the FE's function 1 described in Section 2.1. They set the body state as a variable that requires continuous learning. The robot will actively estimate the current body state and verify this estimate by sensory information received from sensors, continuously updating this estimate until the prediction error (FE) is zero, to determine the current body state. Inspired by this, we divide our experiment into two steps: Step 1 is to determine the robot's initial body state using the method proposed by Lanillos and Cheng (2018). This corresponds to the FE's function 1. Step 2 is to infer the body state at each moment by the initial state and the executed actions during the task, predict each sensation, and calculate the prediction error at each moment, to determine whether the robot is in the injury state. This corresponds to the FE's function 3. Step 1 uses the State Module, Prediction Module, Sensory Module, and Error Module of BRP-SNN. Step 2 uses the State Module, Prediction Module, Sensory Module, Error Module, and Pain Module of BPR-SNN.

The body state is a high-dimensional variable in the brain. For computational purposes, Lanillos et al. defined body state in the proprioception data space. This does not mean that body state is equivalent to proprioception, but only consistent with its data dimension and the range of values. We follow this assumption that body states are defined in the proprioception data space and encoded by the State Module of BRP-SNN. The Sensory Module encodes real proprioception and vision data. We set the initial value of the body state and continuously make predictions to both sensor data. The prediction error (Error Module fires) will guide the body state estimation update

using the exhaustive search algorithm until the Error Module no longer fires. After the initial robot body state is determined, the subsequent body states can be inferred by prior knowledge (the body state and the executed actions at the previous moment). In a normal situation, there is no prediction error when the current body state makes predictions about the current sensation, the Error Module does not fire. When the body is injured, the Error Module fires, which in turn causes the Pain Module fires, and the robot is in a pain state.

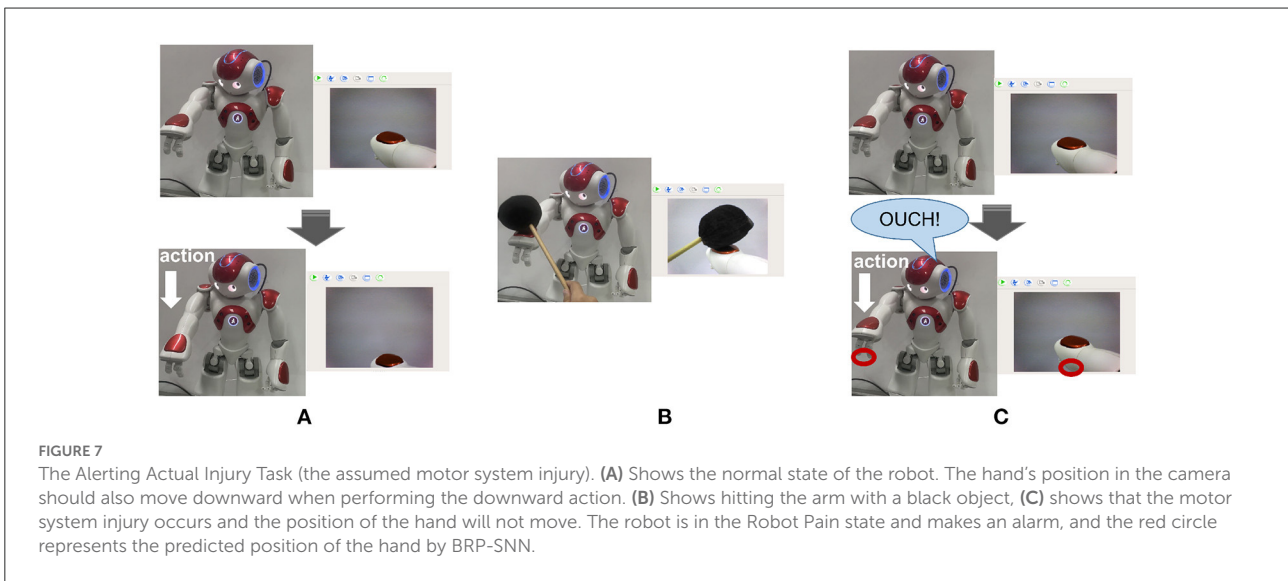
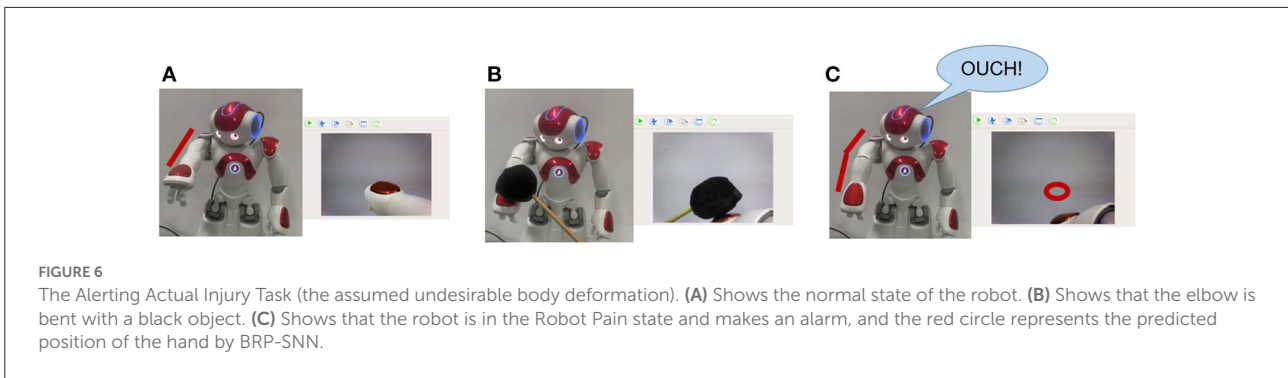
The mapping relationship between body state and proprioception $g_{s_p}()$ and vision $g_{s_v}()$ should be learned in advance for the forwarding prediction of the State Module to the Prediction Module. The robot collects training data through several random movements (motor babbling) while the left arm is straight. As shown in Figure 5. Proprioception data only includes the ShoulderPitch joint angle θ_1 , the ShoulderRoll joint angle θ_2 , and the WristYaw joint angle θ_3 , and the Elbow joint is not considered. Vision data is the coordinate position (x, y) of the hand in the camera, obtained by grayscale processing of the image captured by the camera and calculating the median value of the corresponding pixel coordinates. The body state ϕ is set in the proprioception data space, so the mapping relationship $g_{s_p}(\phi) = \phi$. The mapping relationship $g_{s_v}()$ needs to be learned using STDP.

3.2. The alerting actual injury task

3.2.1. Experiment results

We designed two actual injury paradigms in the alerting actual injury task: undesirable body deformation and motor system injury.

In the case of undesirable body deformation, when the elbow bending caused by a black object occurs, the robot's prediction will be inconsistent with the real sensory data. Then the robot



turns into the Robot Pain state and says “OUCH!” for alarm. Figure 6A shows the normal state of the robot. Figure 6B shows that the elbow is bent with a black object. Figure 6C shows that the robot is in the Robot Pain state and makes an alarm, and the red circle represents the predicted position of the hand by BRP-SNN.

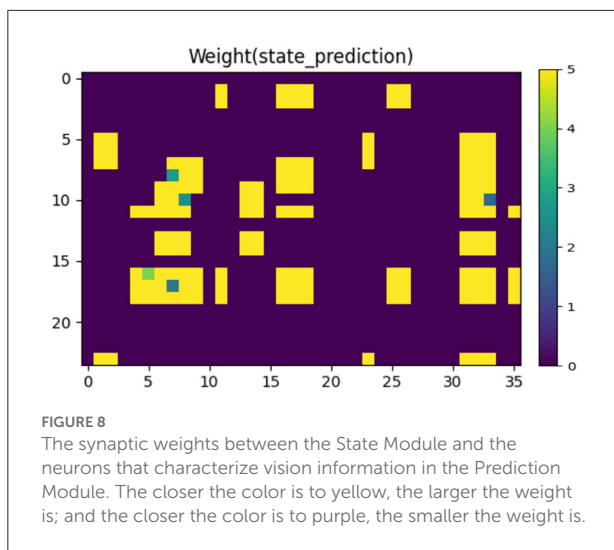
In the case of motor system injury, we made the following assumption: when a black object hits the robot’s arm and touches the tactile sensor, the robot does not execute any action at the program level, simulating motor system injury. In this experiment, the robot constantly does random movements and predicts the sensory information at the next moment by its current body state and the action command (prior knowledge). When the injury occurs, the predicted position is inconsistent with the real sensory. Then the robot is in the Robot Pain state and says “OUCH!” for alarm. Figure 7A shows the normal state of the robot. The hand’s position in the camera should also move downward when performing the downward action. Figure 7B shows hitting the arm with a black object, Figure 7C shows that the motor system injury occurs and the position of the hand will not move. The robot is in a Robot Pain state and makes an alarm,

and the red circle represents the predicted position of the hand by BRP-SNN.

3.2.2. Model analysis

Figure 8 shows the synaptic weights between the State Module and the neurons that characterize vision information in the Prediction Module, indicating the mapping relationship between the body state and the prediction of vision information. These weights are learned by STDP training rules from the data set collected in advance (as shown in Figure 5). The closer the color is to yellow, the larger the weight is; and the closer the color is to purple, the smaller the weight is.

Figure 9 represents the spike diagram of modules of the BRP-SNN in the actual injury condition. The X-axis represents the time, and the Y-axis represents the neuron index. Figure 9A represents the firing pattern of the State Module. Figure 9B represents the firing pattern of the Prediction Module when predicting sensory information. Figure 9C represents the firing pattern of the Sensory Module when receiving the real sensory information. It is shown that the firing neuron index of the



Prediction Module is inconsistent with the Sensory Module, so the Error Module fires. Figure 9D shows the firing pattern of the Error Module. Due to the excitatory connection between the Error Module and the Pain Module, the Pain Module then fires, indicating the Robot Pain state. In BRP-SNN, we design the time step as 100 ms, and the robot collects sensor information per 100 ms. Figure 9 represents the firing of each module during 1,000 ms when the robot experiences actual injury. The sensor acquires abnormal sensory information every 100 ms, which is encoded by the Sensory Module, which causes the Error Module to fire. The Error Module is fired 10 times in 1,000 ms.

Figure 10 shows the comparative analysis of our proposed model with previous neuroscience studies about ACC. Silvetti et al. (2011, 2013) proposed a neural model of the ACC—the Reward Value and Prediction Model (RVPM) and validated it in model-based human fMRI experiments. Figure 10A represents the average activation intensity of the ACC module of the RVPM over time. In this experiment, the cue signal is first fed into the model, which is a square wave with a 2,000 ms duration and unit amplitude. Then, a prediction process with a 2,000 ms duration starts. The reward feedback signal with a 400 ms duration is added at 1,600 ms. Each experiment is 2,500 ms and is performed 20 times. We simulated the same experiment steps in the BRP-SNN. As shown in Figure 10B, body state information (robot proprioceptive information) was first fed to the State Module to start a prediction process with a 2,000 ms duration. Then, a feedback signal (real visual information) of 400 ms duration is added to the Sensory Module at 1,600 ms. The total duration of the experiment is also 2,500 ms. We performed it 20 times and calculated the average activation intensity of the whole ACC. It is worth noting that each unit of RVPM represents a continuous value, and the calculation rule for the average activation intensity is the sum of the amplitudes of multiple

signals. However, each module of our BRP-SNN is a neuron group with discrete spike signals, and the calculation rule for the average activation intensity is the sum of the firing rates of all modules in ACC every 100 ms. The different data morphology and calculation rules result in inconsistent limits on the Y-axis in the two figures. But, both figures reflect the same change trend of average activation intensity: At the beginning of the prediction, both curves increase significantly, indicating that the prediction module is firing. At the reception of an environmental feedback signal that is inconsistent with the prediction, both curves increase again, indicating that the prediction error module is firing. At the end of the prediction, both curves decrease to 0, indicating that the whole ACC is at rest state. This result indicates that our BRP-SNN has the same change trend as RVPM under the same experimental conditions.

3.3. The preventing potential injury task

3.3.1. Experiment results

In both actual injury paradigms, the black object is captured by the camera and encoded by the Cue Module, and establishes the connection between the Cue Module and the Pain Module by association learning. When the robot sees the black object again, even if the actual injury has not occurred, the Pain Module also fires and then executes the avoidance behavior. Figure 11A shows that the dangerous black object approaches the robot, and the robot recognizes it. Figure 11B shows that the robot executes avoidance behavior to prevent potential injury.

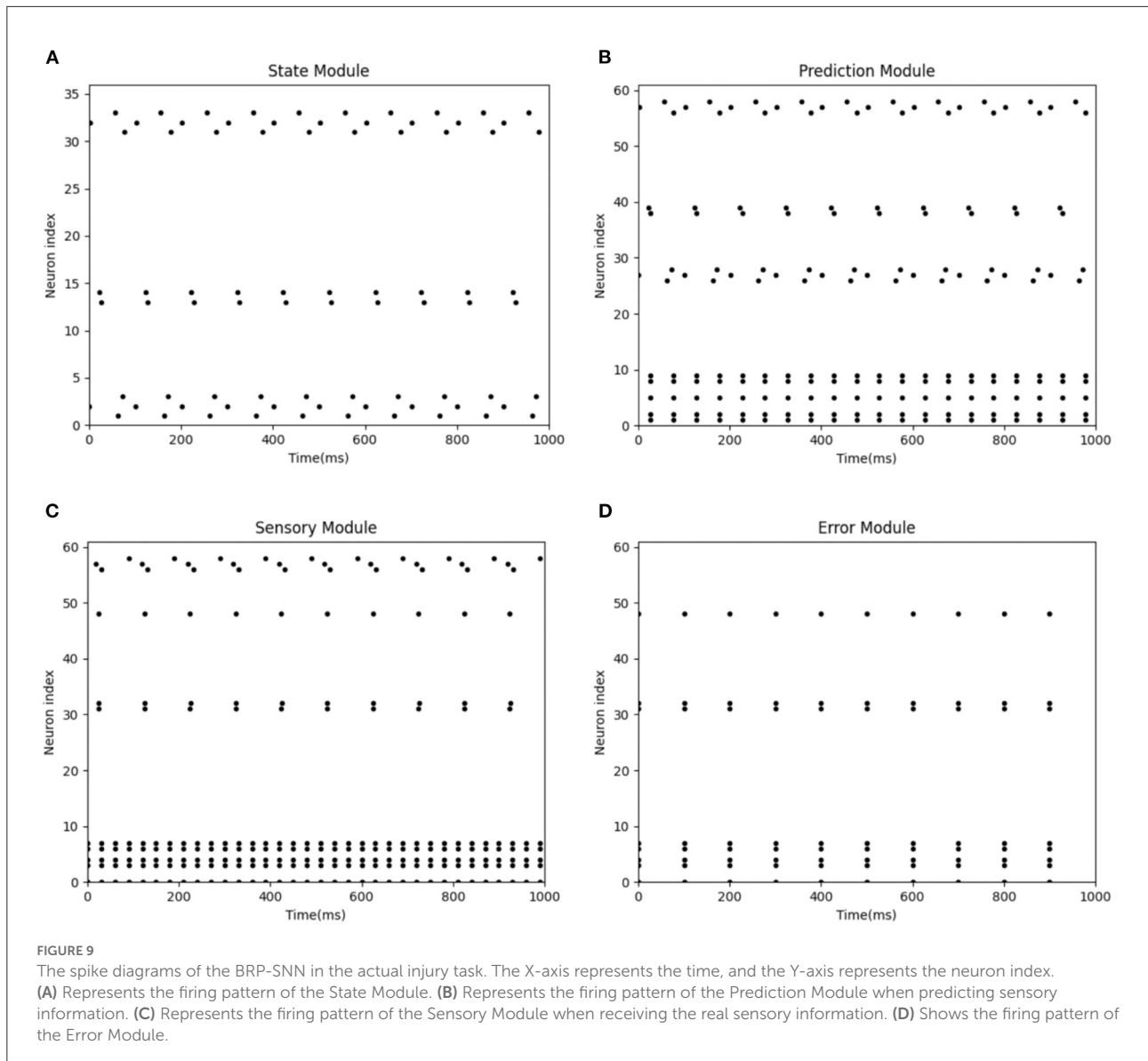
3.3.2. Model analysis

When the camera of the robot detects the black object, this black object is encoded by the Cue Module. When the actual injury occurs, the connection between the Cue Module and the Pain Module can be established. When the robot sees the black object again, the Cue Module fire, as shown in Figure 12A. Then the Pain Module fire, as shown in Figure 12B, and avoid potential injury.

4. Discussion

This paper proposes a Brain-Inspired Robot Pain model that simulates the neural mechanism and behavioral results of organisms' pain. This model can alert to actual injury and avoid potential injury, and obtain the result curves similar to previous neuroscience experiments.

In this paper, we argue that pain is a kind of neural activity occurring at the brain level accompanied by physical injury, and it is a subjective reflection of the objective physical injury event. It means that the organism has a self-body model, which reflects the coupling relationship between the multi-modality



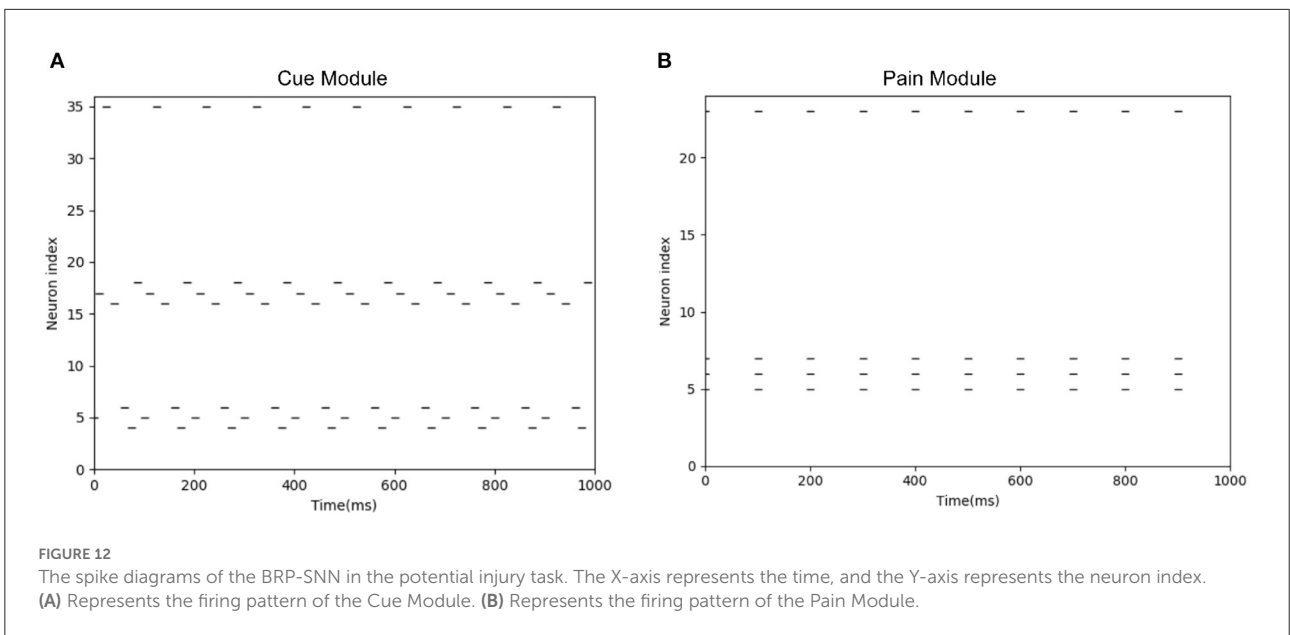
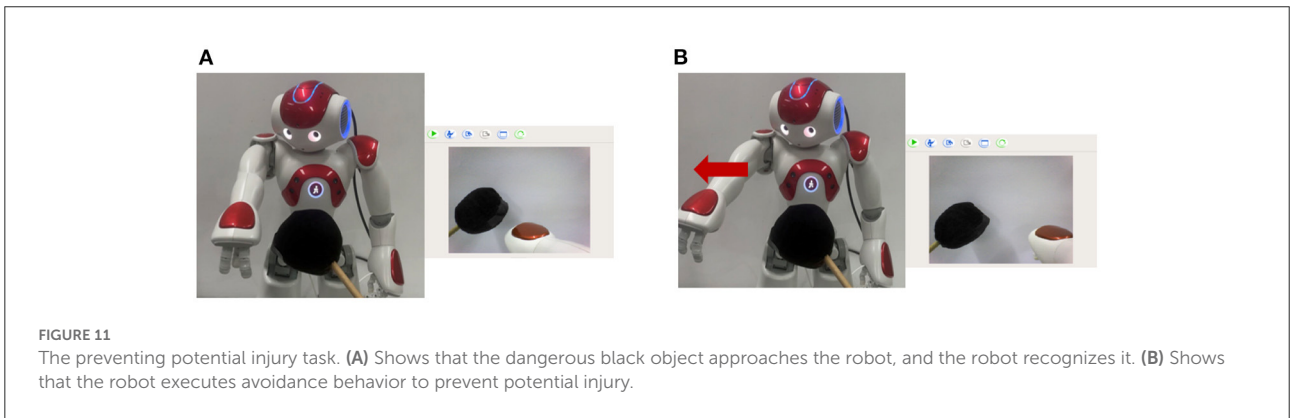
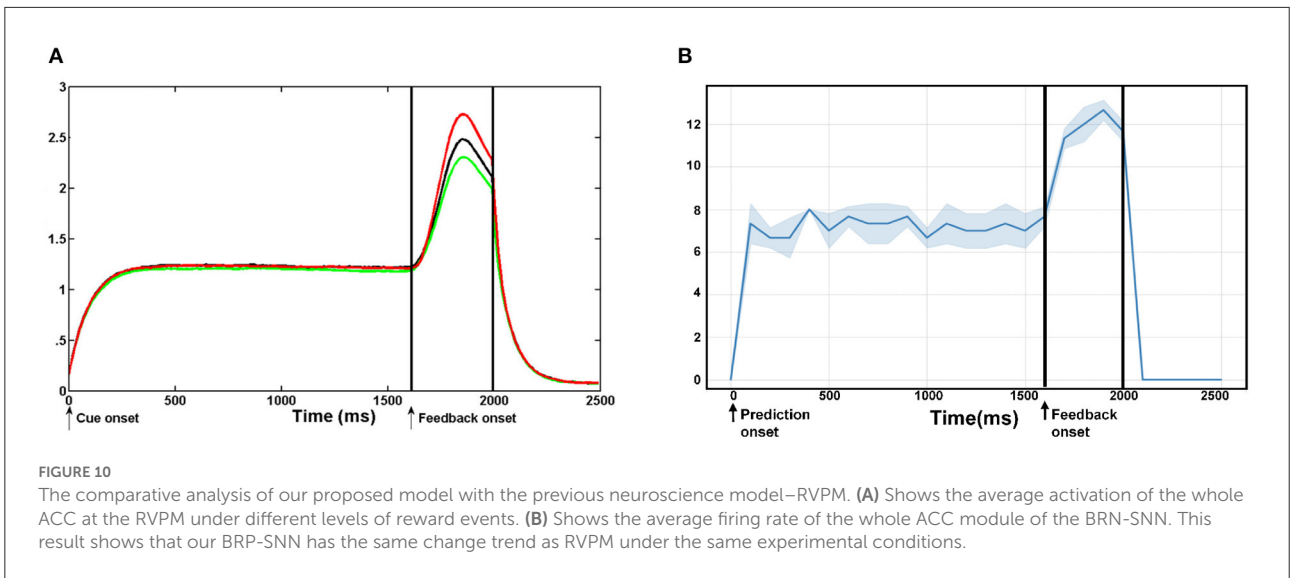
information of the body, such as the multisensory modalities and motor modality. Once a physical injury event occurs, the coupling relationship will be broken, and the organism will be in a state of increased entropy. Then the brain will perceive this abnormal state and react to it, gradually evolving into pain. In our work, only two sensory modalities and one motion modality of the robot are considered. If robots have more sensors in the future, our model can also be generalized to more modalities.

We have made certain assumptions in robot experiments, such as taking normal elbow bending as an abnormal deformation of robot and designing motor system injury at the program level. That is because we cannot actually harm the robots in the lab. The two tasks of the robot are only a validation of our BRP-SNN model. We also hope to use the principle of this

model to make the robots produce 'Robot Pain' in a real work environment, to realize the alarm and prevention of injury in the future.

In the training phase of the connection weights between the State Module and the Prediction Module, we set the initial weights to 0 and use the STDP to adjust the weights according to the correlation of the firing times of the pre-synaptic and post-synaptic spikes. In this paper, we mainly emphasize the effect of LTP, and the weight values that meet the LTP condition will be increased. We manually restricted the size of the weights to the range [0, 5] to facilitate observation and debugging.

There are still some limitations to our work. First, our model implements the task on a robot under the specific conditions we designed, so its robustness has not been verified yet. In future



work, we will validate the utility of our model under other experimental conditions (e.g., robot data from other modalities). Second, in the preventing potential injury task, the process from the firing of the Cue Module (potential danger detection) to triggering the firing of the Pain Module (pain state activation) is trained by STDP through association learning. However, the transition from the pain state activation to the execution of the avoidance action is artificially set by us. In fact, for an organism, the avoidance action is learned by trial and error. In the next step, we will explore the learning process of pain-triggered self-defensive actions and the relationship between pain models and intrinsic rewards in the reinforcement learning method. In addition, we plan to further investigate more neural mechanisms and computational models of cognitive functions related to pain, such as pain empathy. Pain empathy is an important factor to promote harmonious coexistence among social groups, and robots with empathic abilities will be more moral (Asada et al., 2009; Asada, 2015). In the future, we will explore the computational models of pain empathy and the altruistic behavior of robots.

This paper proposes a Brain-Inspired Robot Pain Spiking Neural Network inspired by the neural mechanism of organisms' pain, enabling the robot to have a human-like pain capacity. We explored the neural mechanism of pain emergence from the perspective of pain evolution and the brain's Free Energy Principle, and we used SNN to simulate relevant brain regions' functions and connections to build a Robot Pain model with the STDP method and the population coding method. Our model is inspired by the pain's neural mechanisms and achieves not only the alarming of actual machine injuries but also the prevention of potential danger, which has positive implications for the integration of pain concepts into the robotics field. Our work is a meaningful step toward creating more brain-like intelligent robots in the future.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

HF and YZ designed the study, developed the algorithm, performed the result analysis and experiments, and wrote and

revised the paper. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the new generation of artificial intelligence major project of the Ministry of Science and Technology of the People's Republic of China (Grant No. 2020AAA0104305) and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB32070100).

Acknowledgments

We appreciate Enmeng Lu, Feifei Zhao, Dongcheng Zhao, Yuxuan Zhao, and Hongjian Fang for valuable discussions. We would like to thank all the reviewers on their help for shaping and refining the paper.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnbot.2022.1025338/full#supplementary-material>

References

- Asada, M. (2015). Towards artificial empathy. *Int. J. Soc. Robot.* 7, 19–33. doi: 10.1007/s12369-014-0253-z
- Asada, M. (2019). Artificial pain may induce empathy, morality, and ethics in the conscious mind of robots. *Philosophies.* 4, 38. doi: 10.3390/philosophies4030038

- Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Yoshikawa, Y., Ogino, M., et al. (2009). Cognitive developmental robotics: a survey. *IEEE Trans. Auton. Mental Dev.* 1, 12–34. doi: 10.1109/TAMD.2009.2021702

- Bagnato, C., Takagi, A., and Burdet, E. (2015). "Artificial nociception and motor responses to pain, for humans and robots," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 7402–7405. doi: 10.1109/EMBC.2015.7320102
- Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. *J. Math. Psychol.* 76, 198–211. doi: 10.1016/j.jmp.2015.11.003
- Bohte, S. M., Kok, J. N., and La Poutre, H. (2002). Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing* 48, 17–37. doi: 10.1016/S0925-2312(01)00658-0
- Broom, D. M. (2001). Evolution of pain. *Vlaams Diergeneesk. Tijdschr.* 70, 17–21.
- Cao, S., Fu, D., Yang, X., Barros, P., Wermter, S., Liu, X., and Wu, H. (2021). How can AI recognize pain and express empathy.
- Castellar, E. N., Kühn, S., Fias, W., and Notebaert, W. (2010). Outcome expectancy and not accuracy determines posterror slowing: ERP support. *Cogn. Affect. Behav. Neurosci.* 10, 270–278. doi: 10.3758/CABN.10.2.270
- Chevchenko, S. F., and Ludermit, T. B. (2020). "Learning from sparse and delayed rewards with a multilayer spiking neural network," in *2020 International Joint Conference on Neural Networks (IJCNN)* (Glasgow), 1–8. doi: 10.1109/IJCNN48605.2020.9206846
- Dan, Y., and Poo, M.-m. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron* 44, 23–30. doi: 10.1016/j.neuron.2004.09.007
- Davis, K. D., Taylor, S. J., Crawley, A. P., Wood, M. L., and Mikulis, D. J. (1997). Functional MRI of pain- and attention-related activations in the human cingulate cortex. *J. Neurophysiol.* 77, 3370–3380. doi: 10.1152/jn.1997.77.6.3370
- Du, J., Fang, J., Xu, Z., Xiang, X., Wang, S., Sun, H., et al. (2020). Electroacupuncture suppresses the pain and pain-related anxiety of chronic inflammation in rats by increasing the expression of the NPS/NPSR system in the ACC. *Brain Res.* 1733, 146719. doi: 10.1016/j.brainres.2020.146719
- Eich, E., Reeves, J. L., Jaeger, B., and Graff-Radford, S. B. (1985). Memory for pain: relation between past and present pain intensity. *Pain* 23, 375–380. doi: 10.1016/0304-3959(85)90007-7
- Fang, H., and Zeng, Y. (2021). "A brain-inspired causal reasoning model based on spiking neural networks," in *2021 International Joint Conference on Neural Networks (IJCNN)* (Shenzhen), 1–5. doi: 10.1109/IJCNN52387.2021.9534102
- Fang, H., Zeng, Y., and Zhao, F. (2021). Brain inspired sequences production by spiking neural networks with reward-modulated STDP. *Front. Comput. Neurosci.* 15, 8. doi: 10.3389/fncom.2021.612041
- Frankland, P. W., and Teixeira, C. M. (2005). A pain in the acc. *Mol. Pain* 1, 1744–8069. doi: 10.1186/1744-8069-1-14
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J. Physiol. Paris* 100, 70–87. doi: 10.1016/j.jphysparis.2006.10.001
- Gerstner, W., and Kistler, W. M. (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press. doi: 10.1017/CBO9780511815706
- Hardcastle, V. G. (1997). When a pain is not. *J. Philos.* 94, 381–409. doi: 10.2307/2564606
- Jackson, P. L., Rainville, P., and Decety, J. (2006). To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain* 125, 5–9. doi: 10.1016/j.pain.2006.09.013
- Jessup, R. K., Busemeyer, J. R., and Brown, J. W. (2010). Error effects in anterior cingulate cortex reverse when error likelihood is high. *J. Neurosci.* 30, 3467–3472. doi: 10.1523/JNEUROSCI.4130-09.2010
- Karl, F. (2012). A free energy principle for biological systems. *Entropy* 14, 2100–2121. doi: 10.3390/e14112100
- Karsdorp, P. A., and Vlaeyen, J. W. (2009). Active avoidance but not activity pacing is associated with disability in fibromyalgia. *Pain* 147, 29–35. doi: 10.1016/j.pain.2009.07.019
- Kuehn, J., and Haddadin, S. (2016). An artificial robot nervous system to teach robots how to feel pain and reflexively react to potentially damaging contacts. *IEEE Robot. Autom. Lett.* 2, 72–79. doi: 10.1109/LRA.2016.2536360
- Lanillos, P., and Cheng, G. (2018). "Adaptive robot body learning and estimation through predictive coding," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid), 4083–4090. doi: 10.1109/IROS.2018.8593684
- Loeser, J. D., and Melzack, R. (1999). Pain: an overview. *Lancet* 353, 1607–1609. doi: 10.1016/S0140-6736(99)01311-2
- Markram, H., Gerstner, W., and Sjöström, P. J. (2012). Spike-timing-dependent plasticity: a comprehensive overview. *Front. Synapt. Neurosci.* 4, 2. doi: 10.3389/fnbot.2022.1025338
- Oliveira, F., McDonald, J., and Goodman, D. (2008). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J. Cogn. Neurosci.* 19, 1994–2004. doi: 10.1162/jocn.2007.19.12.1994
- Perl, E. R., and Kruger, L. (1996). "Nociception and pain: evolution of concepts and observations," in *Pain and Touch* ed L. Kruger (Elsevier: Academic Press), 179–211. doi: 10.1016/B978-012426910-1/50006-6
- Peters, A., McEwen, B. S., and Friston, K. (2017). Uncertainty and stress: why it causes diseases and how it is mastered by the brain. *Prog. Neurobiol.* 156, 164–188. doi: 10.1016/j.pneurobio.2017.05.004
- Raja, S. N., Carr, D. B., Cohen, M., Finnerup, N. B., Flor, H., Gibson, S., et al. (2020). The revised IASP definition of pain: concepts, challenges, and compromises. *Pain* 161, 1976. doi: 10.1097/j.pain.0000000000001939
- Schlund, M. W., Magee, S., and Hudgins, C. D. (2011). Human avoidance and approach learning: evidence for overlapping neural systems and experiential avoidance modulation of avoidance neurocircuitry. *Behav. Brain Res.* 225, 437–448. doi: 10.1016/j.bbr.2011.07.054
- Silvetti, M., Seurinck, R., and Verguts, T. (2011). Value and prediction error in medial frontal cortex: Integrating the single-unit and systems levels of analysis. *Front. Hum. Neurosci.* 5, 75. doi: 10.3389/fnhum.2011.00075
- Silvetti, M., Seurinck, R., and Verguts, T. (2013). Value and prediction error estimation account for volatility effects in ACC: a model-based fMRI study. *Cortex* 49, 1627–1635. doi: 10.1016/j.cortex.2012.05.008
- Sjöström, J., and Gerstner, W. (2010). Spike-timing dependent plasticity. *Scholarpedia* 5, 1362. doi: 10.4249/scholarpedia.1362
- Sneddon, L. U. (2019). Evolution of nociception and pain: evidence from fish models. *Philos. Trans. R. Soc. B* 374, 20190290. doi: 10.1098/rstb.2019.0290
- Sur, I., and Amor, H. B. (2017). "Robots that anticipate pain: anticipating physical perturbations from visual cues through deep predictive models," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Vancouver, BC), 5541–5548. doi: 10.1109/IROS.2017.8206442
- Tal, D., and Schwartz, E. L. (1997). Computing with the leaky integrate-and-fire neuron: logarithmic computation and multiplication. *Neural Comput.* 9, 305–318. doi: 10.1162/neco.1997.9.2.305
- Walters, E. T., and Williams, A. C. d. C. (2019). Evolution of mechanisms and behaviour important for pain. *Philos. Trans. R. Soc. B* 374, 20190275. doi: 10.1098/rstb.2019.0275
- Werner, P., Lopez-Martinez, D., Walter, S., Al-Hamadi, A., Gruss, S., and Picard, R. W. (2022). Automatic recognition methods supporting pain assessment: a survey. *IEEE Trans. Affect. Comput.* 13, 530–552. doi: 10.1109/TAFFC.2019.2946774
- Wiech, K., and Tracey, I. (2013). Pain, decisions, and actions: a motivational perspective. *Front. Neurosci.* 7, 46. doi: 10.3389/fnins.2013.00046
- Yang, S., Gao, T., Wang, J., Deng, B., Azghadi, M. R., Lei, T., et al. (2022a). SAM: a unified self-adaptive multicompartmental spiking neuron model for learning with working memory. *Front. Neurosci.* 16, 850945. doi: 10.3389/fnins.2022.850945
- Yang, S., Gao, T., Wang, J., Deng, B., Lansdell, B., and Linares-Barranco, B. (2021). Efficient spike-driven learning with dendritic event-based processing. *Front. Neurosci.* 15, 601109. doi: 10.3389/fnins.2021.601109
- Yang, S., Linares-Barranco, B., and Chen, B. (2022b). Heterogeneous ensemble-based spike-driven few-shot online learning. *Front. Neurosci.* 16, 850932. doi: 10.3389/fnins.2022.850932
- Yang, S., Tan, J., and Chen, B. (2022c). Robust spike-based continual meta-learning improved by restricted minimum error entropy criterion. *Entropy* 24, 455. doi: 10.3390/e24040455
- Zhao, F., Zeng, Y., and Xu, B. (2018). A brain-inspired decision-making spiking neural network and its application in unmanned aerial vehicle. *Front. Neurobot.* 12, 56. doi: 10.3389/fnbot.2018.00056