frontiers | Frontiers in Neuroscience

CORRECTION
published: 19 May 2022
doi: 10.3389/fnins.2022.920292

Corrigendum: A Brain-Inspired Theory of Mind Spiking Neural Network for Reducing Safety Risks of Other Agents

Zhuoya Zhao[1,2†], *Enmeng Lu*[1†], *Feifei Zhao*[1†], *Yi Zeng*[1,2,3,4,5*] *and Yuxuan Zhao*[1]

[1] *Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences, Beijing, China,* [2] *School of Future Technology, University of Chinese Academy of Sciences, Beijing, China,* [3] *School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China,* [4] *National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China,* [5] *Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China*

Keywords: brain-inspired model, safety risks, SNNs, R-STDP, theory of mind

A Corrigendum on

A Brain-Inspired Theory of Mind Spiking Neural Network for Reducing Safety Risks of Other Agents

by Zhao, Z., Lu, E., Zhao, F., Zeng, Y., and Zhao, Y. (2022). Front. Neurosci. 16:753900. doi: 10.3389/fnins.2022.753900

OPEN ACCESS

Edited and reviewed by:
Georgios Ch. Sirakoulis, Democritus University of Thrace, Greece

***Correspondence:**
Yi Zeng
yi.zeng@ia.ac.cn

These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Neuromorphic Engineering, a section of the journal Frontiers in Neuroscience

Received: *14 April 2022*
Accepted: *03 May 2022*
Published: *19 May 2022*

Citation:
Zhao Z, Lu E, Zhao F, Zeng Y and Zhao Y (2022) Corrigendum: A Brain-Inspired Theory of Mind Spiking Neural Network for Reducing Safety Risks of Other Agents. Front. Neurosci. 16:920292. doi: 10.3389/fnins.2022.920292

In the original article **Wu et al. (2002)** was not cited in the article. The citation has now been inserted in ***Methods***, ***Encoding and Decoding Schemes***, ***Paragraph 1*** and should read:

Spiking neural networks need effective encoding methods to process the input stimulus and decoding methods to represent the output stimulus to handle various stimulus patterns. Population coding is "a method to represent stimuli by using the joint activities of a number of neurons. Experimental studies have revealed that this coding paradigm is widely used in the sensor and motor areas of the brain" (Wu et al., 2002). Besides, population coding tries to avoid the ambiguity of the messages carried within a single trial by each neuron (Panzeri et al., 2010).

Additionally, in the original article, **Rabinowitz et al. (2018)** should be referenced more than once. The citation has now been inserted in ***Methods***, ***The Architecture of the ToM-SNN***, ***Paragraph 2*** and should read:

Our model is a multiple brain areas coordination model composed of multiple modules. It is not an end-to-end multilayer neural network. The advantages of a multiple brain areas coordinationmodel are reflected in two aspects. First, inspired by brain structure and function, modules in the ToM-SNN corresponding to specific brain areas have specific functions. The end-to-end neural networks are "regularly described as opaque, uninterpretable black-boxes" (Rabinowitz et al., 2018). Our model is more biologically plausible and more interpretable. Second, a multiple brain areas coordination model can reduce the burden of training. When a new feature appears in the task, only the module for this feature needs to be retrained. So this structure can reduce the amount of calculation and improve efficiency. The policy inference module, the action prediction module, and the state evaluation module are fully connected SNNs with two layers. Details of the two-layers SNNs are as follows. The input current of the input layer and the output layer is denoted by I^{in} and I^{out}, respectively. The output spikes of the input layer and the output layer are denoted by S^{in} and S^{out}, respectively. Section 3.1 describes

the neural spiking process. At each time step *t*, the input current to neuron *j* at the output layer is integrated as Equation (5).

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

REFERENCES

Panzeri, S., Montani, F., Notaro, G., Magri, C., and Peterson, R. S. (2010). "Population coding," in *Analysis of Parallel Spike Trains* (Boston, MA: Springer US), 303-319.

Rabinowitz, N., Perbet, F., Song, F., Zhang, C., Eslami, S. A., and Botvinick, M. (2018). "Machine theory of mind," in *International Conference on Machine Learning* (Stockholm: PMLR), 4218–4227.

Wu, S., Amari, S.-I., and Nakahara, H. (2002). Population coding and decoding in a neural field: a computational study. *Neural Comput.* 14, 999–1026. doi: 10.1162/08997660275363 3367