VT-SNN: Variable Time-step Spiking Neural Network Based on Uncertainty Measure and Its Application in Brain Disease Diagnosis

Haonan Rao[1,*], Shaolong Wei[2,*], Shu Jiang[2], Mingliang Wang[3], Weiping Ding[2], and Jiashuang Huang[2,✉]

[1] College of Information Science and Technology, Nantong University, Nantong 226019, China
[2] School of Artificial Intelligence and Computer Science, Nantong University, Nantong 226019, China
[3] School of Computer and Software Nanjing University of Information Science and Technology, Nanjing 210044, China
hjshdym@163.com

Abstract. The integration of neural networks with Magnetic Resonance Imaging (MRI) data for brain disease diagnosis has become a significant research focus. However, the inherent complexity of 3D MRI data poses challenges for traditional models like CNNs and Transformers, leading to high computational costs and difficulties in clinical deployment. Spiking Neural Networks (SNNs), inspired by biological neurons, offer a promising alternative with enhanced efficiency and robustness. Yet, their application to MRI data is limited by fixed time-steps that fail to account for inter-sample variability. To address this, we propose a Variable Time-Step Spiking Neural Network (VT-SNN) that dynamically adjusts the time-step based on sample-specific uncertainty. Our method employs an SNN-based Transformer module to convert MRI data into spike form and extract features, followed by a variable time-step module that measures decision uncertainty using Fisher information and PAC-Bayes theory. Experiments on AHNU and AMRD datasets demonstrate superior classification performance and reduced computational costs. Our codes are available at https://github.com/UAIBC-Brain/MICCAI-2025-Paper-VT-SNN.

Keywords: Magnetic Resonance Imaging · Spiking Neural Networks · Variable time-step · Uncertainty.

1 Introduction

The application of neural networks to learn representations from Magnetic Resonance Imaging (MRI) scans and utilize these representations for diagnosing brain diseases has emerged as a prevalent approach in contemporary research [14, 15]. Unlike natural images, MRI data are typically three-dimensional and exhibit a

[*] Haonan Rao and Shaolong Wei — Equal first-author contribution.

higher degree of complexity [1]. When employing classical convolutional neural networks (CNNs) or Transformer models to analyze MRI data, the models become more intricate than those used for traditional natural images. This increased complexity leads to higher computational costs, causing difficulties deploying these models for clinical applications.

Spiking Neural Networks (SNNs) are a novel class of brain-inspired neural networks modelled after the spiking behaviour and information transmission of biological neurons [18]. By encoding information via spike sequences, SNNs achieve significant reductions in energy consumption and enhanced computational efficiency. Compared to traditional neural networks, SNNs offer advantages such as sparse computation, spatiotemporal integration, and noise robustness [4, 21], thereby alleviating the computational cost issues associated with conventional models. SNNs have shown broad application potential in various fields. In computer vision, they are used for tasks like image classification, object detection, and action recognition [20]. In healthcare, SNNs are applied to brain disease diagnosis and neural prediction.

Direct application of SNNs to MRI data analysis faces limitations in accounting for individual variability in MRI datasets [10]. Traditional SNNs employ a fixed time-step to enhance the capture of data details and dynamics. However, MRI data exhibit substantial inter-sample heterogeneity due to noise in data acquisition and pathological variations in brain diseases. A fixed time-step is not universally suitable for all samples [5,17]. For instance, smaller time-steps fail to provide adequate diagnostic information for complex samples, while larger time-steps increase computational costs and risk overfitting noise in simpler samples, thereby impairing diagnostic accuracy.

To this end, we propose the variable time-step spiking neural network (VT-SNN). It is a variable spiking neural network architecture in which each sample can decide at which time-step to terminate according to its own uncertainty information. Specifically, we first employ an SNN-based Transformer module as the main framework of the network to convert MRI data into spike form and extract features. Subsequently, we design a variable time-step module that calculates the confidence score. It measures the decision uncertainty for each sample after each time-step using Fisher information and PAC-Bayes theory and determines whether to terminate the computation based on these scores. Our proposed VT-SNN method achieves excellent classification performance and low computational cost on the AHNU and AMRD datasets.

2 Proposed Method

Our proposed VT-SNN consists of two parts: the SNN-based Transformer module and Vthe ariable Time-step module. The SNN-based Transformer module converts MRI inputs to spike-form data, extracts features from high to low dimensions, and reworks the self-attention mechanism with Leaky Integrate-and-Fire (LIF) neurons. The Variable Time-step module evaluates the output confi-

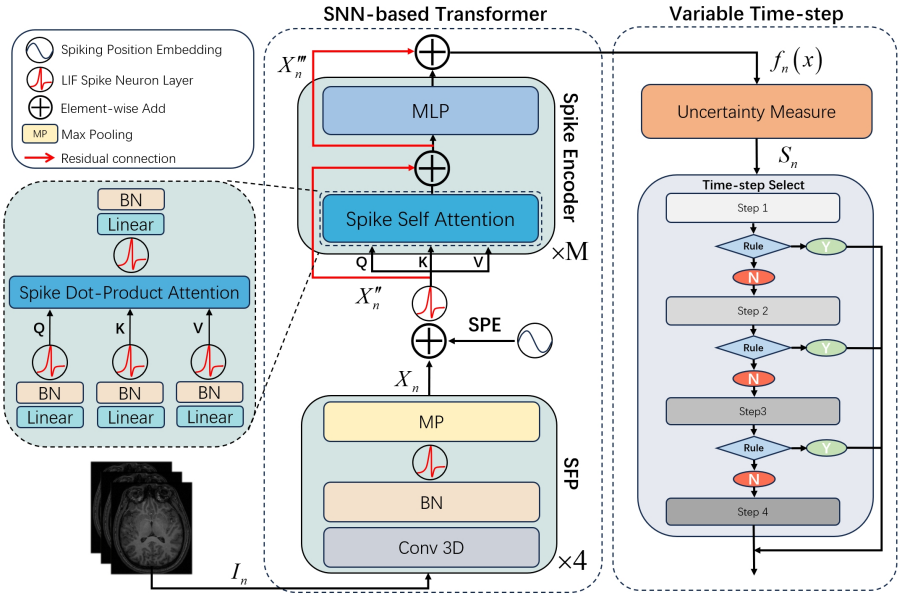

Fig. 1. Variable Time-step Spiking Neural Network Based on Uncertainty Measure (VT-SNN) Overview. The architecture mainly includes the SNN-based Transformer module and the Variable Time-step module. The SNN-based transformer module consists of an SFP (Spike Feature Patching) and M layers of Spike Encoder. The Variable Time-step module performs two operations: uncertainty measurement and time-step selection.

dence score at each step, allowing early inference termination and cost reduction through conditional early outputs. Details are as follows:

2.1 SNN-based Transformer Module

As shown in Fig. 1, the SNN-based transformer module consists of the Spike Feature Patching (SFP) and the Spike Encoder. Firstly given an MRI data $I_n \in \mathbb{R}^{T\times C\times D\times H\times W}$, where T represents the time-step (By encoding the MRI data into T pieces of spike data), C represents the channel dimension ($C = 1$ for MRI), and D, H, W correspond to the depth, height, and width of the MRI data respectively, n represents the n-th time-step T. After lightweight 4-layer SFP, we obtain $X_n \in \mathbb{R}^{T\times N\times K}$. The feature extraction process can be described as follows:

$$X_n = \mathrm{MP}(\mathcal{SNN}(\mathrm{BN}(\mathrm{Conv3D}(I_n)))), \quad I_n \in \mathbb{R}^{T\times C\times D\times H\times W} \tag{1}$$

After the input passes through the MP, the spike data will be converted into the float-form. Since SNNs cannot utilize float-form positional embeddings, a spike-encoded Positional Embedding (SPE), which comprises Conv3D, BN, and

SNN layers, generates X'_n. The positional embedding process can be described as follows:

$$X'_n = X_n + \text{SPE}, \quad X_n \in \mathbb{R}^{T\times N\times K}, \quad X'_n \in \mathbb{R}^{T\times N\times K} \tag{2}$$

where N represents the product of D, W, H, which is the size of the patch, and K denotes the dimension after spike feature extraction.

Then we pass the data through an M-layer spike encoder module, which includes a Spiking Self Attention (SSA) and an MLP block. The SSA computes attention using a spike dot-product method for the Query (Q), Key (K), and Value (V), for details, see the Spike Self Attention in the bold part. Moreover, we use MLP to generate the preliminary output $f_n(x)$ at time-step T=n. The process of obtaining $f_n(x)$ from X''_n through the spike encoder is as follows:

$$X''_n = \mathcal{SNN}(X'_n), \quad X''_n \in \mathbb{R}^{T\times N\times K}, \tag{3}$$

$$X'''_n = \text{SSA}(X''_n) + X''_n, \quad X'''_n \in \mathbb{R}^{T\times N\times K}, \tag{4}$$

$$f_n(x) = \text{MLP}(X'''_n) + X'''_n, \tag{5}$$

Spike Self Attention: It is distinct from the traditional method of using softmax in self-attention. For the input $X''_n \in \mathbb{R}^{T\times N\times K}$, we first compute its Q, K, V separately ($Q = \mathcal{SNN}_Q(\text{BN}(XW_Q))$, $K = \mathcal{SNN}_K(\text{BN}(XW_K))$, $V = \mathcal{SNN}_V(\text{BN}(XW_V))$, where $Q, K, V \in \mathbb{R}^{T\times N\times K}$), and then calculate the attention of X using a spike dot-product method, enabling X to encompass spatial structural information ranging from local to global scales. Subsequently, this is processed through an $\mathcal{SNN}$. The SSA operations are as follows:

$$\text{SSA}'(Q, K, V) = \mathcal{SNN}\left(Q\,K^T\,V\right), \tag{6}$$

$$\text{SSA}(Q, K, V) = \text{BN}(\text{Linear}(\text{SSA}'(Q, K, V))), \tag{7}$$

2.2 Variable Time-step Module

If we only use the SNN-based Transformer, the process will be repeated T times to obtain the final result, resulting in high computational costs. So, we designed a variable time-step module, which performs two operations: uncertainty measurement and time-step selection.

Uncertainty Measurement: The data uncertainty caused by the differences among different samples will lead to the uncertainty of the model's output results [11,16]. Therefore, we adopt reliable uncertainty decision-making to provide a more reliable confidence score.

As the amount of information carried by different MRI samples varies, leading to varying amounts of information carried by the classification probability p. So we propose using a Fisher Information-based evidential network as an Uncertainty Decision-Making to measure the information carried by the MRI classification probability p regarding the concentration parameter α of the Dirichlet distribution. We first define a Fisher Information Matrix (FIM) based on the

Dirichlet distribution by taking the expected value of the second-order derivatives of the log-likelihood function:

$$I(\alpha) = \mathbb{E}_{\mathrm{Dir}(p|\alpha)} \left[\frac{\partial \ell}{\partial \alpha} \frac{\partial \ell}{\partial \alpha^T} \right], \tag{8}$$

where p represents the classification probability, a represents the concentration parameter, and $\ell = \log \mathrm{Dir}(p \mid \alpha)$ represents the amount of information contained in the likelihood function. After introducing Fisher information and going through a series of derivations [3], it can be expressed as:

$$I(\alpha) = \mathrm{diag}\left([\psi^{(1)}(\alpha_1), \cdots, \psi^{(1)}(\alpha_K)] \right) - \psi^{(1)}(\alpha_0) 1 1^T, \tag{9}$$

where $\psi^{(1)}(x) = \frac{d\psi(x)}{dx} = \frac{d^2 \ln \Gamma(x)}{dx^2}$ represents the second derivative of the gamma function. Ultimately, we can obtain the target variable $Y_n \sim \mathcal{N}(p_n, \sigma^2 \mathcal{I}(\alpha_n)^{-1})$, where $\alpha_n = f_n(x) + 1,\ p_n \sim \mathrm{Dir}(\alpha_n)$, after incorporating the Fisher Information matrix.

In the case where the time-step $T = n$, then we can obtain the loss function, which consists of the MSE, the Kullback-Leibler (KL) divergence term, and the Fisher Information term. Ultimately, the Uncertainty Decision-Making is integrated through learning using the PAC-Bayesian bound [7, 9]. By introducing appropriate prior and posterior distributions, we can better utilize the complementary information in the data and enhance the model's generalization ability. We have integrated our loss function according to the PAC-Bayesian theory. Ultimately, we obtain the following formula for the Uncertainty Decision-Making:

$$\min_{\boldsymbol{\theta}} \frac{1}{P} \sum_{i=1}^{P} \mathcal{L}_i^{\mathcal{I}\cdot\mathrm{MSE}} - \lambda_1 \mathcal{L}_i^{|\mathcal{I}|} + \lambda_2 \mathcal{L}_i^{\mathrm{KL}}, \tag{10}$$

where $\mathcal{L}_i^{\mathcal{I}-\mathrm{MSE}} = \sum_{j=1}^{K} \left[\left(Y_{ij} - \frac{\alpha_{ij}}{\alpha_{i0}} \right)^2 + \frac{\alpha_{ij}(\alpha_{i0} - \alpha_{ij})}{\alpha_{i0}^2(\alpha_{i0}+1)} \right] \psi^{(1)}(\alpha_{ij})$, $\mathcal{L}_i^{|\mathcal{I}|} = \sum_{j=1}^{K} \log \psi^{(1)}(\alpha_{ij}) + \log\left(1 - \sum_{j=1}^{K} \frac{\psi^{(1)}(\alpha_{i0})}{\psi^{(1)}(\alpha_{ij})} \right)$, $\mathcal{L}_i^{\mathrm{KL}} = \log \Gamma\left(\sum_{i=1}^{K} \hat{\alpha}_{ij} \right) - \log \Gamma(K) - \sum_{j=1}^{K} \log \Gamma(\hat{\alpha}_{ij}) + \sum_{i=1}^{K} (\hat{\alpha}_{ij} - 1) \left[\psi(\hat{\alpha}_{ij}) - \psi\left(\sum_{k=1}^{K} \hat{\alpha}_{ik} \right) \right]$.

In order to ensure that the concentration parameters are not negative in the Dirichlet distribution, we use $\alpha_i = f_i(x) + 1$, we set $\hat{\alpha}_i = \alpha_i \cdot (1 - Y_i) + Y_i$ to remove the predicted concentration parameter of the true label corresponding to the sample x, α_{ij} represents the concentration parameters of the j-th category of the i-th sample calculated by the model. α_{i0} represents the sum of all categories of the i-th sample. We set $\lambda_1, \lambda_2 \geq 0$. The penalty term composed of ${L_i}^{|I|}$ and ${L_i}^{KL}$ can avoid overconfidence caused by overfitting. The weighted term of the Fisher Information Matrix can adaptively adjust the weights of other terms based on the uncertainty information contained in different classes of different samples.

Then we adopt the confidence score as the criterion for the uncertainty assessment of MRI samples. The formula for the confidence score is as follows:

$$S_n = \max(\mathrm{softmax}(f_n(x))) \tag{11}$$

where $f_n(x)$ denotes the preliminary output at time $T{=}n$ and S_n represents the confidence score at time $T{=}n$, which is the maximum value of the probability distribution of the preliminary output after softmax processing.
Time-step Selection: We select the time-step according to these confidence scores. Specifically, for these scores, it is necessary to set a threshold S_n to make a judgment on the uncertain results output by the model. The specific operation is as follows: Based on the confidence score S_n, we can decide when to output the final result by checking if S_n at time $T = n$ exceeds the threshold S_{th}. We set the maximum time-steps to 4. Starting from $T = 1$, if the condition isn't met, we increment T by 1 and recheck until the condition is satisfied or $T = 4$. In brief, when $S_n > S_{th}$, the model will output the final result.

3 Experiments

3.1 Subjects and Data Preprocessing

This study used two sMRI datasets: a dataset of 1410 cases from the Decoded Neurofeedback Project organized by the Japan Agency for Medical Research and Development (AMRD) [19], covering multiple mental disorders (MMD), and a Parkinson's disease (PD) dataset of 300 cases from the Affiliated Hospital of Nantong University (AHNU). Both datasets were linearly registered to standard templates via FSL, followed by AC-PC alignment, skull stripping, and removal of extraneous structures. Incomplete scan images were systematically excluded to ensure data quality.

Table 1. Five models are compared on AHNU and AMRD Datasets. Power is the average hardware consumption per unit time during the test. Time-step is the setting of the snn inference process, as well as Acc, Sen, Spe and F1 scores.

Dataset	Method	Power(mJ)	Time-step	Acc(%)	Sen(%)	Spe(%)	F1(%)
AHNU	ResNet3D	49.8	–	76.3	71.2	78.2	73.4
	Sfcn	29.8	–	78.5	74.1	83.1	77.6
	SEW ResNet	25.2	4	74.3	81.3	66.1	76.4
	Spikformer	27.4	4	78.3	75.2	78.7	76.6
	Spike-driven Transformer	26	4	80.2	82	62.2	80.4
	Ours	**12.7**	**1.7**	**84.3**	**84.2**	**84.2**	**83.4**
AMRD	ResNet3D	48.2	–	71.6	67.2	75.4	69.4
	Sfcn	26.9	–	73.3	67.4	79.6	71.2
	SEW ResNet	26.8	4	67.3	73.1	61.6	65.4
	Spikformer	27.6	4	71.6	69.8	72.4	70.8
	Spike-driven Transformer	26.4	4	74.6	**80.2**	68.4	72.4
	Ours	**13.6**	**1.9**	**77.3**	75.8	**80.6**	**77.2**

3.2 Experimental Setup

All experiments utilized five-fold cross-validation with 350 epochs, batch size 15, and initial learning rate 0.001 optimized via cosine annealing scheduler (decay

rate 0.01). We evaluated classification results of AHNU and AMRD datasets, assessing performance through time-step, power, accuracy (ACC), sensitivity (SEN), specificity (SPE), and F1 score.

For comparative experiments, we employed two ANNs (ResNet3D [8] and SFCN [13]) and three SNNs (SEW-ResNet [6], Spikformer [23], and Spike-Driven Transformer [22]) as baseline models.

3.3 Classification Performance

In the AHNU and AMRD datasets, the experimental performance of each method is shown in Table 1. Our proposed method achieved the best results across all evaluation metrics. Specifically, in the AHNU dataset, our VT-SNN method achieved an accuracy (ACC) of 84.3% for PD recognition, outperforming the Spike-Driven Transformer (80.2%) and other methods (all below 80%). In other metrics, VT-SNN improved sensitivity (SEN) by 2%-13%, with similar gains in specificity (SPE) and F1 score. In the AMRD dataset, VT-SNN maintained superior performance across all metrics. Overall, our method significantly outperformed others in brain disease identification.

3.4 Spike Activation Visualization Analysis

In Fig. 2, brain region structures in MRI are represented by spike form. AMRD shows the activated brain regions of MMD, with the highest spike intensity in the prefrontal region, followed by the parietal and occipital lobes. AHNU represents the spike activation regions of PD, located near the substantia nigra. Compared to previous studies, our results demonstrate equivalent effectiveness [2, 12].

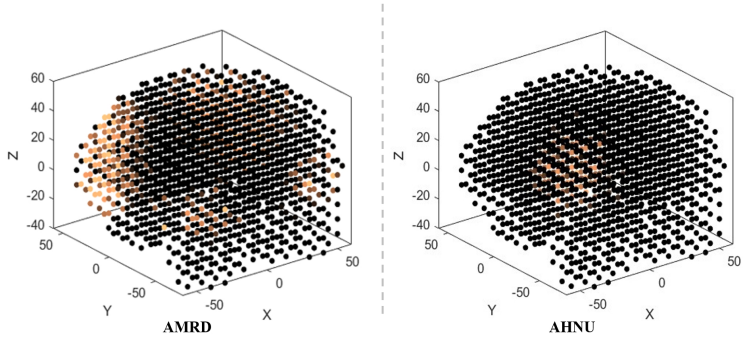

Fig. 2. The visualization of the results of VT-SNN in using spike representations to indicate the pathological locations for the classification of MMD and PD. Black dots denote unactivated regions, while bright dots signify spike-activated regions, which are employed to identify the pathological locations of diseases.

3.5 Latency and Hardware Power Consumption

We used CPUID HWMonitor Pro to measure the computational energy consumption of our method and baseline models. As shown in Table 1, our model achieves time-steps of no more than 2, reducing latency by over 50% compared to other SNN models, while consuming 52% to 73% less power. These results demonstrate that VT-SNN significantly lowers training costs through optimized computational resource utilization compared to conventional ANN and SNN models. Fig. 3 (b) shows the time-step size distribution for the AMRD and AHNU datasets. In AMRD, nearly 50% of samples have time-steps at T=1, with proportions decreasing as T increases: T=2 and T=3 each account for about 20%, and T=4 has the smallest share. AHNU follows a similar trend but with a higher proportion at T=1 (close to 60%) and less than 10% at T=4.

3.6 Ablation Study

Fig. 3 shows ablation studies on the AHNU dataset demonstrate that fixed time-step models achieve suboptimal accuracy (ACC < 80% at T=4), while the variable time-step mechanism in VT-SNN significantly improves performance (ACC > 84% with maximum T<2). Increasing the confidence score threshold S_{th} from 0.6 to 0.9 correlates with a rising time-step T and continuous ACC enhancement, validating the reliability of the confidence score generated by the variable time-step module. Similar trends are confirmed on the AMRD dataset.

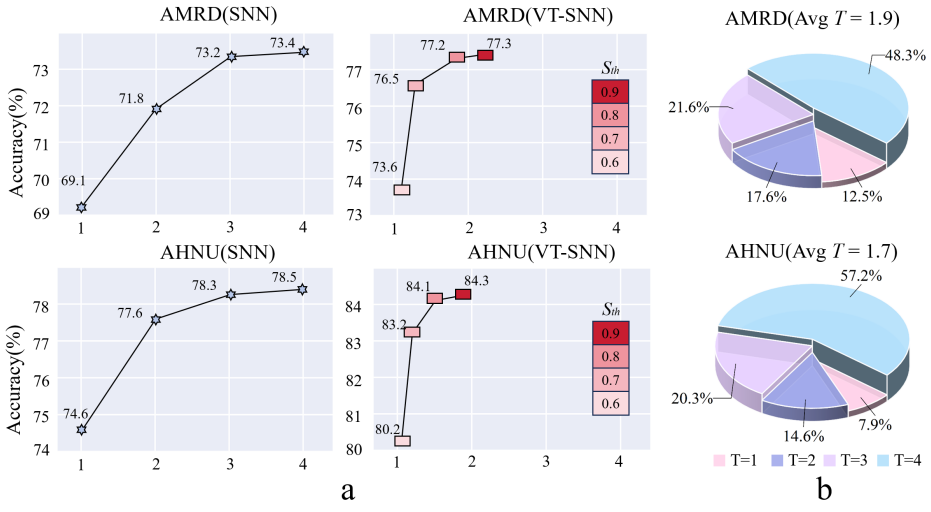

Fig. 3. (a) Accuracy vs. Time-step, where SNN baselines show degraded performance without variable time-steps, contrasted with VT-SNN's threshold-controlled progression (The red square: 0.6-0.9 thresholds). (b) Time-step distributions across datasets.

4 Conclusion

In this study, we propose VT-SNN, an adaptive early-stopping spiking neural network that dynamically selects optimal temporal steps based on model confidence score to achieve efficient inference. To enhance model reliability, we incorporate an uncertainty estimation module grounded in PAC-Bayesian theory, thereby improving generalization capabilities. When applied to auxiliary diagnosis of neurological disorders, our VT-SNN framework demonstrated a balance between high diagnostic accuracy and low computational latency.

Acknowledgements. This study was supported in part by the National Natural Science Foundation of China (No. 62471259, 62472228, 62406153), and the Natural Science Foundation of Jiangsu Higher Education Institutions of China (No. 24KJB520032, 23KJB520031).

Disclosure of Interests. The authors have no competing interests to declare relevant to this article's content.

References

1. Bullmore, E., Barnes, A., Bassett, D.S., Fornito, A., Kitzbichler, M., Meunier, D., Suckling, J.: Generic aspects of complexity in brain imaging data and other biological systems. Neuroimage **47**(3), 1125–1134 (2009)
2. Chakraborty, S., Aich, S., Kim, H.C.: Detection of parkinson's disease from 3t t1 weighted mri scans using 3d convolutional neural network. Diagnostics **10**(6), 402 (2020)
3. Deng, D., Chen, G., Yu, Y., Liu, F., Heng, P.A.: Uncertainty estimation by fisher information-based evidential deep learning. In: International conference on machine learning. pp. 7596–7616. PMLR (2023)
4. Deng, L., Wu, Y., Hu, X., Liang, L., Ding, Y., Li, G., Zhao, G., Li, P., Xie, Y.: Rethinking the performance comparison between snns and anns. Neural networks **121**, 294–307 (2020)
5. Deng, S., Gu, S.: Optimal conversion of conventional artificial neural networks to spiking neural networks. arXiv preprint arXiv:2103.00476 (2021)
6. Fang, W., Yu, Z., Chen, Y., Huang, T., Masquelier, T., Tian, Y.: Deep residual learning in spiking neural networks. Advances in Neural Information Processing Systems **34**, 21056–21069 (2021)
7. Germain, P., Bach, F., Lacoste, A., Lacoste-Julien, S.: Pac-bayesian theory meets bayesian inference. Advances in Neural Information Processing Systems **29** (2016)
8. Hara, K., Kataoka, H., Satoh, Y.: Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet? In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 6546–6555 (2018)
9. McAllester, D.A.: Some pac-bayesian theorems. In: Proceedings of the eleventh annual conference on Computational learning theory. pp. 230–234 (1998)
10. Meng, Q., Yan, S., Xiao, M., Wang, Y., Lin, Z., Luo, Z.Q.: Training much deeper spiking neural networks with a small number of time-steps. Neural Networks **153**, 254–268 (2022)

11. Nair, T., Precup, D., Arnold, D.L., Arbel, T.: Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation. Medical image analysis **59**, 101557 (2020)
12. Pagano, G., Niccolini, F., Politis, M.: Imaging in parkinson's disease. Clinical Medicine **16**(4), 371–375 (2016)
13. Peng, H., Gong, W., Beckmann, C.F., Vedaldi, A., Smith, S.M.: Accurate brain age prediction with lightweight deep neural networks. Medical image analysis **68**, 101871 (2021)
14. Rauschecker, A.M., Rudie, J.D., Xie, L., Wang, J., Duong, M.T., Botzolakis, E.J., Kovalovich, A.M., Egan, J., Cook, T.C., Bryan, R.N., et al.: Artificial intelligence system approaching neuroradiologist-level differential diagnosis accuracy at brain mri. Radiology **295**(3), 626–637 (2020)
15. Sadeghi, D., Shoeibi, A., Ghassemi, N., Moridian, P., Khadem, A., Alizadehsani, R., Teshnehlab, M., Gorriz, J.M., Khozeimeh, F., Zhang, Y.D., et al.: An overview of artificial intelligence techniques for diagnosis of schizophrenia based on magnetic resonance imaging modalities: Methods, challenges, and future works. Computers in Biology and Medicine **146**, 105554 (2022)
16. Sensoy, M., Kaplan, L., Kandemir, M.: Evidential deep learning to quantify classification uncertainty. Advances in neural information processing systems **31** (2018)
17. Shen, J., Xu, Q., Liu, J.K., Wang, Y., Pan, G., Tang, H.: Esl-snns: An evolutionary structure learning strategy for spiking neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 86–93 (2023)
18. Taherkhani, A., Belatreche, A., Li, Y., Cosma, G., Maguire, L.P., McGinnity, T.M.: A review of learning in biologically plausible spiking neural networks. Neural Networks **122**, 253–272 (2020)
19. Tanaka, S.C., Yamashita, A., Yahata, N., Itahashi, T., Lisi, G., Yamada, T., Ichikawa, N., Takamura, M., Yoshihara, Y., Kunimatsu, A., et al.: A multi-site, multi-disorder resting-state magnetic resonance image database. Scientific data **8**(1), 227 (2021)
20. Yamazaki, K., Vo-Ho, V.K., Bulsara, D., Le, N.: Spiking neural networks and their applications: A review. Brain sciences **12**(7), 863 (2022)
21. Yang, J., Zhang, P., Liu, Y.: Robustness of classification ability of spiking neural networks. Nonlinear Dynamics **82**, 723–730 (2015)
22. Yao, M., Hu, J., Zhou, Z., Yuan, L., Tian, Y., Xu, B., Li, G.: Spike-driven transformer. Advances in neural information processing systems **36**, 64043–64058 (2023)
23. Zhou, Z., Zhu, Y., He, C., Wang, Y., Yan, S., Tian, Y., Yuan, L.: Spikformer: When spiking neural network meets transformer. arXiv preprint arXiv:2209.15425 (2022)