Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Supplementary materials for

Li CHEN, Fan ZHANG, Guangwei XIE, Yanzhao GAO, Xiaofeng QI, Mingqian SUN, 2025. S3Det: a fast object detector for remote sensing images based on artificial to spiking neural network conversion. *Front Inform Technol Electron Eng*, 26(5):713-727. https://doi.org/10.1631/FITEE.2400594

Functional equivalence of ANN-to-SNN

In an ANN using ReLU, for a L^{th} layer network, the process of layer-l forward derivation can be described as

$$a_i^l = \max\left(0, \sum_{j=1}^{M^{l-1}} W_{ij}^l a_j^{l-1} + b_i^l\right), 1 \le l \le L\,, \tag{S1}$$

where a_i^l is the input vector to the l^{th} layer, $W^l, l \in \{1,2,\ldots,L\}$ represents the weight matrix from the $(l\text{-}1)^{\text{th}}$ to the l^{th} layer, and b_i^l denotes the bias term for the i^{th} neuron in the l^{th} layer.

In the SNN, we used the integrate-and-fire (IF) membrane equation to model neuron dynamics. The i^{th} neuron in the l^{th} layer has a membrane potential $V_i^l(t)$ that includes input current $z_i^l(t)$ and generated spike events $\Theta_i^l(t)$.

$$V_i^l(t) = V_i^l(t-1) + z_i^l(t) - V_{th}\Theta_i^l(t-1)\,, \tag{S2}$$

$$z_i^l(t) = V_{th}\left(\sum_{j=1}^{M^{l-1}} W_{ij}^l \Theta_j^{l-1}(t) + b_i^l\right), \tag{S3}$$

$$\Theta_i^l(t) = \begin{cases} 1 & \text{if } V_i^l(t) \ge V_{th} \\ 0 & \text{else.} \end{cases}, \tag{S4}$$

where V_{th} is the membrane voltage threshold after triggering neuron spikes $\Theta_i^l(t)$ and soft reset, and the input current $z_i^l(t)$ is a linear combination of the spikes $\Theta_j^{l-1}(t)$ from the previous layer.

For ANN-to-SNN conversion, the firing rate approximates the activation value of the ReLU. To achieve this approximation, the mapping function requires spikes to fall in the unit interval. Assuming a simulation duration of T, the firing rate can be expressed as

$$q_i^{(l)}(T) = \frac{\sum_{t=0}^{T} \mathbf{W}_i^{(l)} \Theta_i^{l-1}(t)}{T\theta_i^{(l)}} + \frac{V_i^{(l)}(0) - V_i^{(l)}(T)}{T}\,. \tag{S5}$$

If the threshold $\theta_i^{(l)}$ is set to 1 and the simulation duration is sufficiently long, the last term in Eq. (3) approximates to 0, while the first term approximates to a_i^l. This completes the equivalent mapping for ANN-to-SNN conversion. Notably, to ensure that spikes fall in the unit interval, the maximum weight is scaled to the p^{th} percentile.