IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 23, 2020, accepted December 4, 2020, date of publication December 14, 2020,
date of current version December 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3044646

An Imbalanced R-STDP Learning Rule in Spiking Neural Networks for Medical Image Classification

QIAN ZHOU, CONG REN, AND SAIBING QI
State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300130, China
Tianjin Key Laboratory of Bioelectromagnetic Technology and Intelligent Health, Hebei University of Technology, Tianjin 300130, China

Corresponding author: Qian Zhou (qzhou@hebut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61305077 and Grant 51977060, and in part by the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology under Grant EERI_PI2020006.

ABSTRACT Spiking neural networks (SNNs) have the advantages of inherent power-efficiency, biological plausibility and good image recognition performance. They are good candidates for medical image classification especially when the labeled training data are limited. In medical image classification, one of the major challenges is the highly class imbalanced problem which causes deep learning networks to bias towards the majority class and poorly recognizes the minority class. Despite that there are some methods for addressing this problem, very few algorithm-level methods exist for SNNs. In this work, we propose an imbalanced reward-modulation spike-timing-dependent plasticity (R-STDP) learning rule for SNNs to solve the medical image class imbalanced problem. We introduce an imbalanced reward coefficient for the R-STDP learning rule to set the reward from the minority class to be higher than that of the majority class, and this reward coefficient can help to set the class-dependent rewards according to the data statistic of the training dataset. Experiment results on three benchmark datasets with imbalanced splits show that our method significantly improves the performance than that of the baseline SNNs and outperforms the compared state-of-the-art methods addressing medical image class imbalanced problem including data-level and algorithm-level methods. Moreover, our method achieves excellent classification performance on the imbalanced medical dataset ISIC-2018. The results show that the proposed method can well help SNNs in classifying imbalanced medical image datasets. Besides, our proposed method can obtain high sensitivity to disease class by adjusting the reward coefficient, which is very useful for identifying disease samples in medical diagnostic tasks.

INDEX TERMS Class imbalance, spiking neural networks (SNNs), reward-modulated spike-timing-dependent plasticity (R-STDP), medical image.

I. INTRODUCTION

The future of medicine places much emphasis on early diagnosis and personalized medicine [1]. Medical image has always been personalized and plays an important role in disease diagnosis. It is also very useful for disease progress assessment, treatment planning, and effectiveness evaluation. With the increasing availability and development of medical devices, the volume and complexity of medical image are increasing at a speed faster than the availability of human expertise to interpret it [2]. In some specialties, there is even a relative shortage of professionals to give a timely diagnosis for patients [3]. Meanwhile, deep learning (DL) with its impressively good performance in image recognition has rapidly risen to be a prominent method to analyze medical images [4].

Deep learning can automatically extract complex features at high levels of abstraction from raw data, so it can learn and make predictions based on input data in an automated fashion [5]. In fact, DL algorithms have already surpassed human performance in some tasks such as image classification on the large scale ImageNet dataset [6] and Atari 2600 computer games [7]. With the current availability of big healthcare data, enhanced computing power, and new training algorithms, these DL methods are currently gaining increasing attention in the medical image applications, and they

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar.

mainly focus on classification [8], [9], localization [10], [11], detection [12], [13], segmentation [14], [15], and registration [16], [17]. In some specialties, DL methods have shown the potential to become substitutions of trained physicians to make diagnosis. For example, Nam *et al*. [13] proposed a DL-based automatic detection algorithm, which demonstrated higher performance than thoracic radiologist group in radiograph classification and nodule detection for malignant pulmonary nodules. Tschandl *et al*. [9] compared the diagnostic accuracies of 511 human physicians versus 139 DL models in the seven classification tasks of dermatoscopic images and found that the DL models consistently outperformed physicians.

Despite that DL has achieved impressive performance in some medical applications, there are still some challenges [18]. One of the major challenges is that medical image datasets are usually highly imbalanced because pathologic findings are generally rare in the diagnosis [18], [19]. The datasets are often predominantly composed of ''normal'' samples with only a very small portion of ''abnormal'' ones (target classes), leading to the so-called class imbalance problem [20]. For example, in a dataset of breast cancer images called BreaKHis [21] which contains microscopic biopsy images of benign and malignant breast tumors, the number of images of malignant tumors is twice that of benign ones. In the ISIC dataset [22] used for skin lesion analysis towards melanoma detection, melanoma lesion images account for only 11% of the total. It has been established that class imbalance has a detrimental effect on the training of DL networks [23]–[25]. Since most classifiers are modeled by exploring data statistics, a DL model trained with class imbalanced data may bias towards learning the majority class, and hence show very poor classification accuracy on the minority class [24]–[27]. However, in the case of medical diagnosis, misclassification costs of the two classes are often unequal, and classifying the minority samples (images containing cancerous tumor) as the majority (images for healthy patient or benign tumor) implies serious consequences. At present, there are generally two kinds of methods for handling class imbalanced problem in DL methods: data-level methods and algorithm-level methods [23]–[25]. Data-level methods aim to alter the training data distribution and to reduce the level of imbalance by re-sampling the train data, usually by under-sampling and over-sampling. These methods have shown helpful for imbalanced classification, but they also have some obvious drawbacks. Under-sampling may remove some important samples and thus lead to reduced classification performance, and over-sampling increases computation because it introduces additional training instances [24], [28]. Different from data-level methods, algorithm-level methods use all the samples from the training data, and they try to alleviate the bias towards the majority class by modifying the learning process of networks [29]. They have little effect on training times because they do not introduce additional training instances [24].

Another challenge for DL methods in the medical image applications is that at present there is still a lack of high-quality medical image datasets for training. This is due to the lack of labeled data or lack of images for the pathologic classes [4], [18]. DL-based image recognition and analysis require enough high-quality, annotated datasets, and a DL model with simple architecture is very likely to overfit on a small training dataset, and it will result in poor classification performance [30]. While using very deep network architecture (for example, Resnet152, which has 152 layers) can improve the classification performance, it greatly increases computing costs, which is another barrier for the clinical applications of DL methods [12]. Deep transfer learning has recently been used in the field of medical image analysis which avoids the problem of overfitting and improves the classification performance of networks [8]. Nevertheless, the network architectures are still very complex and consume a lot of computing costs. Therefore, networks that can achieve higher performance on small datasets and do not require a large amount of computational cost are in great need for analyzing medical images.

Spiking neural networks (SNNs) use spike trains to represent and process information in an asynchronous manner. They have been emerged as an ideal bio-inspired neuromorphic computing paradigm which well mimic the inherent spike-based and event-driven computations in the brain [31]. SNNs have obvious advantages of low power consumption and hardware friendliness, which makes them very promising in realizing energy-efficient on-chip intelligence hardware [32]. There are some SNNs whose classification performances are comparable to traditional convolutional neural networks (CNNs) for small-scale image recognition tasks [33]–[35]. Another advantage of SNNs is that they can learn and achieve better results than deep learning models when learning with fewer data [32]. This is especially fit for the application of the SNNs in the medical image domain. Recently SNNs has been applied to early stroke prediction [36], epilepsy detection [37], skin lesions classification [38], and CT and MR medical image fusion [39]. The learning rule in SNNs usually is bio-inspired spike-timing-dependent-plasticity (STDP) learning rule which modifies synaptic weights according to the temporal order and time difference between presynaptic and postsynaptic spikes [40]. Studies have shown that SNNs with STDP can learn and detect repeating spatio-temporal patterns [41], [42]. And an external classifier such as support vector machines (SVMs) and radial basis functions (RBFs) is generally required for SNNs with unsupervised STDP learning rule. Recently, some studies have shown that the modulation of STDP by a global reward signal enables reward-based reinforcement learning in SNNs, and this learning rule is called reward-modulated STDP [43]–[45]. R-STDP learning rule combines Hebbian and anti-Hebbian modulations of STDP using a reinforcement learning framework by a global reward signal [45], [46]. And it can help SNNs to detect rare but diagnostic features for decision-making and to make classification in an

end-to-end learning manner with no need for an extra classifier. Many hierarchical SNNs with R-STDP learning rule are used to solve the object recognition task and achieved excellent performance [33], [47]–[49]. Vasquez Tieck *et al*. [48] used SNNs with R-STDP to learn the movement target of robot arm, this networks can learn without a large amount of data compared with DL methods. Mozafari *et al*. [33] trained the deep SNNs using a combination of STDP for the lower layers and R-STDP for the higher layers, and this SNNs led to an accuracy of 97.2% on MNIST, without requiring an external classifier.

SNNs can learn from small training dataset quickly and effectively and do not suffer from the overfitting challenge caused by small training set size. Moreover, with spike-based and event-driven communication framework, they have obvious advantages in the computational cost and hardware friendliness. Therefore, SNNs can well solve the problems of limited labeled training data and high computing costs of DL methods for their clinical applications in the field of medical image analysis. However, current studies based on SNNs for medical classification mostly use balanced data [36], [37] or synthetic balanced training data through data-level methods such as undersampling [38], which will reduce the total amount of information learned by the networks and affect their classification performance or increase training time. On the other hand, similar to DL models, imbalanced training data will make the SNNs bias toward the majority classes, and features of the minority classes will not be adequately learned. Therefore, solving the class imbalanced problem from the algorithm-level for SNNs is of great significance to the future application of SNNs, especially when it comes to the medical image classification.

In this paper, to solve the imbalanced problem in medical images, we propose an imbalanced R-STDP learning rule to overcome the effect of class imbalance on SNNs. To our knowledge, this is the first work that modifies the learning algorithm of SNNs to solve the problem of class imbalance. We introduce an imbalanced reward coefficient to set the reward (or punishment) from the minority class to be higher than that of the majority class. In this way, the networks will reduce the bias to the majority class, thereby increase sensitivity to the minority class. The proposed method has been extensively tested on three benchmark datasets (Fashion-MNIST, CIFAR-10, CIFAR-100) and one medical image dataset (ISIC-2018), and is compared with the state-of-the-art methods used to solve class imbalanced problem including data-level and algorithm-level methods. Experiment results show that our method significantly improves the performance of the baseline SNNs on all datasets, and outperforms the compared state-of-the-art methods addressing class imbalanced problem. Moreover, our SNNs with imbalanced R-STDP achieved excellent classification performance on the imbalanced medical dataset ISIC-2018.

Our contribution consists of the following: 1) we study the impact of class imbalance on the classification performance of SNNs. 2) We propose an imbalanced R-STDP learning rule by introducing the reward coefficient, and prove its effect on the learning of SNNs through derivation. 3) The proposed method has been extensively tested on three benchmark datasets and one medical dataset.

The remainder of this paper is organized as follows. We review general methods of DL for medical image classification, the methods of DL for addressing medical image imbalanced problem, and the applications of SNNs on the medical data classification in Section II. In Sections III, we introduce our proposed imbalanced R-STDP learning rule and the architecture of convolutional SNNs. Section IV shows the experiment results and evaluates the performance of our method compared with the state-of-the-art methods used to solve class imbalanced problem including data-level and algorithm-level methods. And this paper concludes in section V.

II. RELATED WORKS

A. THE DEEP LEARNING METHODS FOR MEDICAL IMAGE CLASSIFICATION

It has been proven that DL models are very good at discovering intricate structures in high-dimensional data and is therefore applicable to medical image classification [50]. There are mainly two ways of applying DL to medical image classification. One method is training the DL model from scratch. Nasr-Esfahani *et al*. [51] trained the CNNs model from scratch to distinguish between melanoma and benign nevi. Their method achieved a sensitivity of 81%, a specificity of 80%, and an accuracy of 81% on the public image archive of the Department of Dermatology of the University Medical Center Groningen. Roth *et al*. [11] trained a deep CNNs model using 4298 images to learn 5 anatomical classes and achieved an anatomy-specific classification error of 5.9% in testing. However, due to the high-quality published medical image datasets are limited, training from scratch will make the network models prone to overfitting, using pre-trained DL models is also a popular method in medical image analyze domain. Menegola *et al*. [52] used the VGG16 network pretrained on diabetic retinopathy dataset and ImageNet respectively, and their results showed that the network pretrained on ImageNet had better performance, and the AUC on the 2016 ISIC dataset reached 80.9%. Mahbod *et al*. [53] combined three pretrained CNNs (Alexnet, VGG16 and Resnet18) to discriminate between malignant melanoma, seborrheic keratosis and benign nevi. Their method obtained an AUC of 0.838 for melanoma classification and it was more effective than a single network.

B. DEEP LEARNING METHODS FOR ADDRESSING MEDICAL IMAGE IMBALANCE PROBLEM

One of the most important challenge when it comes to medical image datasets is the class imbalanced problem. Previous researches on how to resolve the class imbalanced problem can be divided into two categories according to their operating level: the data-level methods and the algorithm-level methods [23]–[25].

Data-level methods manipulate the class representations in the original dataset by re-sampling to make the data distribution balanced, including over-sampling, under-sampling, and a combination of the above two methods [24]. Random over-sampling and random under-sampling are the simplest and most popular re-sampling techniques. The former randomly duplicates a certain number of samples from the minority class, while the latter randomly removes a certain number of samples from the majority class. They have been applied to solve the class imbalance problem since 1990s [54], [55]. Japkowicz *et al*. [56] discussed the effects of different re-sampling methods and noted that both the random over-sampling and random under-sampling were effective on their created domains. However, both these methods have known drawbacks. Under-sampling may throw away potentially useful data, and over-sampling can increase the likelihood of overfitting since most oversampling methods make exact copies of the minority class instances [26], [27]. In order to address the overfitting problem of random over-sampling, Chawla *et al*. [20] proposed an over-sampling method called SMOTE which generates new samples by linear interpolation between closely lying minority class samples. Wang *et al*. [57] combined the SMOTE algorithm and the locally linear embedding algorithm to solve the medical image imbalanced problem, and their method achieved a good performance on chest X-ray image databases. In addition, some studies made over-sampling by using data augmentation which generates the minority samples by flipping, rotation or generative adversarial networks [58], [59]. Huang *et al*. [58] applied data augmentation to eliminate the effect of the class imbalance, and using Inception-v4 networks they achieved better performance over the state-of-the-art methods on the public available ISIC skin lesion challenge datasets in 2018. Iqbal *et al*. [60] generated the minority samples by generative adversarial networks and their model achieved a dice coefficient of 0.837 on STARE dataset and 0.832 on DRIVE dataset which is state-of-the-art performance on both datasets. However, a big drawback of over-sampling methods is that by creating additional training samples it leads to a significantly increased computational cost [61].

Unlike data-level methods, algorithm-level methods do not alter the distribution of training data, but directly modify the learning process to improve the sensitivity of the networks toward minority class. Compare with the over-sampling methods, the algorithm-level methods have the minimal impact on training times of the networks [24]. These methods can be further divided into reshaping loss functions and cost-sensitive learning [25]. Reshaping loss function methods are proposed to allow the minority samples to contribute more to the loss. Wang *et al*. [26] proposed the mean false error loss function and its improved version, the mean squared false error loss function, for training neural networks on imbalanced data. These two loss functions can equally capture the errors from the majority and minority classes when the datasets are imbalanced. Lin *et al*. [62] proposed one DNNs model based on focal loss function which addressed the extreme class imbalance by reducing the impacts of easily classified samples on the loss. Then, Tran *et al*. [63] used focal loss and data augmentation to counter class imbalance problem in the pulmonary nodules detecting task, and achieved an AUC of 98.2% on the LIDC/IDRI dataset extracted by the LUNA16 challenge. Cost-sensitive learning methods assign different costs to misclassified samples of different classes by cost matrix [64]. Li *et al*. [65] proposed the discriminative cost-sensitive learning by introducing an expert-provided cost matrix, and improved the screening accuracy of COVID-19 using chest X-ray images. On the large-scale multi-class imbalanced dataset they collected, their method achieved an accuracy of 97.01%, a precision of 97%, a sensitivity of 97.09%, and a F_1-score of 96.98%. However, the appropriate cost matrix of cost-sensitive learning is often unavailable because the misclassification cost is unknown from the data and difficult to obtain [36], [37].

C. SNNs FOR MEDICAL DATA CLASSIFICATION

Compared to traditional DL models which usually have high computing costs, SNNs have obvious advantages of low power consumption and hardware friendliness. And when learning with fewer data, SNNs can learn and achieve better results than traditional DL models. Therefore, with current limited labeled training datasets, SNNs are good candidates for the clinical application of medical image classification. Recently SNNs have been applied to classify hand gestures using EMG signals [66], skin lesions classification [38], early stroke prediction [36], and epilepsy detection [37] in the medical data classification tasks. Elisa *et al*. [66] built a neuromorphic hardware base on SNNs to extract the spatio-temporal information of local EMG signals and classify hand gestures with very low power consumption. Although the classification accuracy of SNNs is slightly lower than that of traditional algorithms, it saves a lot of power consumption. Zhou *et al*. [38] used convolutional SNNs with STDP learning rule to make malignant melanoma and benign melanocytic nevi skin lesions classification on ISIC-2018 dataset. Comparing to DNNs, their methods achieved much better classification accuracies and have much better runtime efficiency. Kasabov *et al*. [36] proposed an evolving SNNs reservoir system to apply for early prediction of individual events such as stroke, compared with machine learning methods it has a significant improvement in accuracy and time of event prediction. Ghosh-Dastidar *et al*. [37] developed an efficient SNNs model for epilepsy and epileptic seizure detection using electroencephalograms and it yields a high classification accuracy of 92.5%.

However, imbalanced training dataset can potentially have a severely negative effect on the overall performance in SNNs. At present, these medical data classification tasks using SNNs are all performed on balanced training data or synthetic balanced training data by data-level methods such under-sampling. Data-level methods have some known

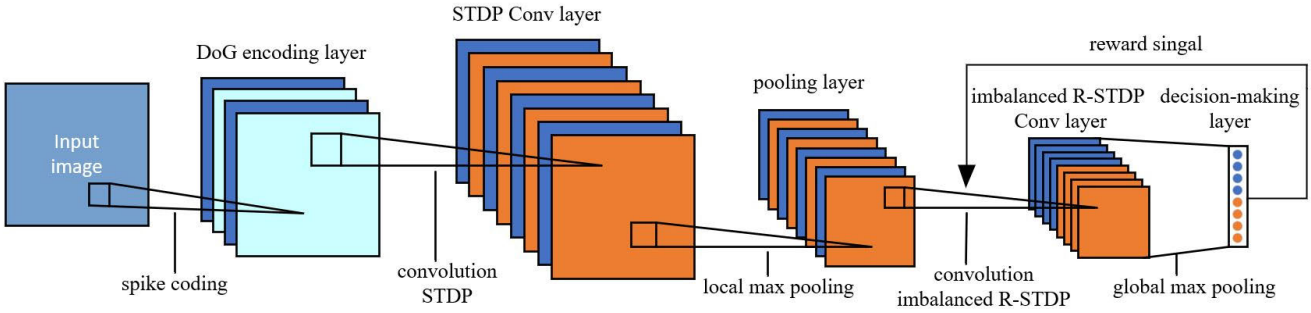

FIGURE 1. Overall structure of our proposed convolutional SNNs with imbalanced R-STDP learning rule.

drawbacks such as loss important training samples or increasing training time. And algorithm-level methods for SNNs to solve class imbalanced problem have not been studied so far.

Therefore, in this paper we try to modify the learning algorithm of SNNs to reduce the effect of class imbalance on their classification performance. We propose an imbalanced R-STDP learning rule for SNNs by introducing an imbalanced reward coefficient. The proposed algorithm has been extensively tested on three benchmark datasets and one medical image dataset, and is compared with the state-of-the-art methods used to solve class imbalanced problem including data-level and algorithm-level methods.

III. OUR PROPOSED CONVOLUTIONAL SNNs WITH IMBALANCED R-STDP LEARNING RULE

A. OVERALL STRUCTURE

The architecture of our proposed convolutional SNNs with imbalanced R-STDP learning rule is shown in Fig. 1. It consists of a Difference of Gaussians (DoG) encoding layer, two convolutional layers (STDP convolutional layer and imbalanced R-STDP convolutional layer), a pooling layer, and a decision-making layer. Except for the DoG encoding layer, the other layers of the networks are composed of integrate and fire (IF) neurons, which exchange information via spikes, and are primarily used in spiking neural networks [32]. The main advantages of IF neuron are that it is computationally inexpensive, and boost the importance of the first pre-synaptic spikes [67]. The latter is particularly important because each neuron in our networks emits at most one spike per stimulus. The DoG encoding layer uses the DoG filters and intensity-to-latency coding scheme to convert the input image into spike trains. Each convolutional layer learns features from its input by STDP or imbalanced R-STDP learning rule, in combination with winner-takes-all (WTA) weight updating strategy and lateral inhibition mechanism. The pooling layer performs spike-based local max pooling to provide translation invariance and compress data. The ultimate layer of the networks is the decision-making layer in which each neuron is preassigned to a certain input category, thereby the neuron that fires first indicates the networks' decision about the category of input image. In addition, the reward signal is generated according to the decisions made by the networks and is fed back to the neurons of the imbalanced R-STDP convolutional layer to adjust its learning process.

B. DoG ENCODING LAYER

In the input spike encoding of SNNs, DoG filters are often used to convert an input image into a spike train [68]. The filters are believed to well mimic the center-surround properties of retinal ganglion cells and can detect contrasts in the input [69]. In our networks, an input image is converted with four DoG filters to extract edges and then encoded into spike trains by intensity-to-latency encoding scheme. This scheme is shown to be effective to detect V1 like oriented edge features and complex visual features in higher cortical areas [34]. A DoG cell emits a spike whenever it detects contrasts in its convolutional window, and the higher the contrast, the earlier it fires. Therefore, the earliest spikes correspond to the most salient regions of an image, and are thus usually the most informative. Note that each DoG cell emits at most one spike. The output spikes are grouped into 15 equal-size packets in the order of time. The conversion of the input image to the spike trains is shown in Fig. 2. At each time step, spikes of one packet are propagated simultaneously and the next spike packet will be processed in the next time step. All of these spikes will to be processed in the following convolutional layer.

C. CONVOLUTIONAL LAYER

A convolutional layer detects more complex features from the output of its previous layer. It has several feature maps to learn different image features which are represented by different input synaptic weights. A convolutional IF neuron receives input spikes from the neurons located in the same convolutional window of all the feature maps of its previous layer. Hence, features extracted by a convolutional layer are a combination of simpler features extracted by its previous layer. At each time step, the membrane potential of a convolutional IF neuron is updated as follows:

$$V(t) = V(t-1) + \sum_j W_j \times S_j(t-1) \tag{1}$$

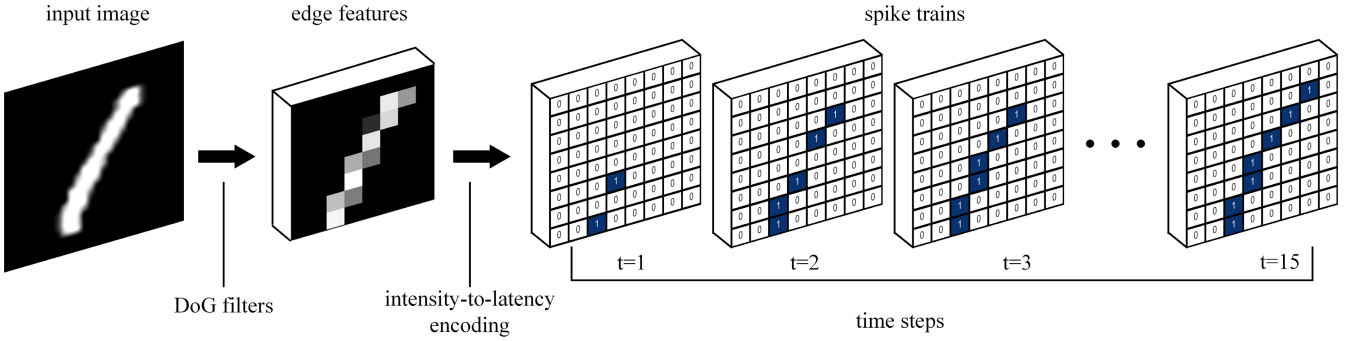

FIGURE 2. An example of generating spike trains from an input image. Firstly, four DoG filters are used to extract the edge features from the input image, then each edge feature is encoded into a spike train through intensity-to-latency encoding scheme. A grid in the figure represents a DoG cell. If it emits a spike at time step t = i (i = 1, 2, 3, ... 15), the corresponding position will be set to 1 from time step t = i to the final time step (t = 15).

where $V(t)$ is the membrane potential of this neuron at time step t, W_j is the synaptic weight between the *jth* presynaptic neuron and this neuron, and S_j is the spike train of the *jth* presynaptic neuron. $S_j(t-1) = 1$, if this presynaptic neuron has fired at time step $(t-1)$, otherwise $S_j(t-1) = 0$. After the updation, if the membrane potential of this convolutional IF neuron exceeds its set threshold, it fires a spike, and then the values of its corresponding S and V are reset as $S(t) = 1, V(t) = 0$.

Here lateral inhibition mechanism [70] is applied to the neurons of each convolutional layer in order to promote neurons in the same location of different feature maps to learn different features. That is, a neuron that fires will inhibit the neurons in the same location but on other feature maps to fire. In addition, each neuron is allowed to fire only once. Neurons belonging to the same feature map share the same input synaptic weights, and compete with each other to update their shared input synaptic weights according to a WTA mechanism [71]. The neuron that fires the earliest is the winner which is qualified to modify their shared weights according to STDP learning rule or imbalanced R-STDP learning rule.

There are two kinds of convolutional layers in our network, STDP convolutional layer and imbalanced R-STDP convolutional layer. The STDP convolutional layer uses the STDP learning rule to update its input synaptic weights between neurons in an unsupervised manner, and the imbalanced R-STDP convolutional layer uses our proposed imbalanced R-STDP learning rule to update its input synaptic weights in a form of reinforcement learning. The number of STDP convolutional layer can be increased or decreased depending on the complexity of the datasets.

D. POOLING LAYER

A pooling layer helps the networks to gain invariance and eliminates the redundancies by doing a local maximum pooling operation over each feature map of its previous convolutional layer. Maximum pooling has demonstrated faster convergence and better performance in comparison to average pooling and its other variants [72], and there is evidence that this maximum operation exists in the complex cells of the visual cortex [73]. Each neuron in a feature map of the pooling layer receives input spikes from the neurons located in the pooling window of the corresponding feature map of its previous convolutional layer. In addition, no learning occurs in the pooling layer because of its input synaptic weights and thresholds are all set to one. Hence, the pooling neurons are activated when they receive first input spike and then give an output spike.

E. DECISION-MAKING LAYER

The last layer of our networks is a decision-making layer. This layer performs global maximum pooling over the feature maps of the imbalanced R-STDP convolutional layer. Each decision-making layer neuron only propagates the first spike that is received from its corresponding feature map of the R-STDP convolutional layer. Before the learning of the networks, these output neurons are divided into several groups and each group is labeled to a particular input category. In this way, the decision of our networks on the category of one input image is assumed to be the label of the group which propagates the earliest spike among the output neurons. If the networks' decision matches the correct category of the input image, it will receive a reward, otherwise, it will receive a negative reward (penalty). The reward or penalty is then used to modulate the modification of input synaptic weights of the imbalanced R-STDP convolutional layer.

F. STDP LEARNING RULE FOR STDP CONVOLUTIONAL LAYER

The presynaptic weights of the STDP convolutional layer neurons are updated according to a simplified STDP learning rule in an unsupervised manner. The change of each synaptic weight depends on the temporal order of the firing times of presynaptic and postsynaptic neurons. The modification of each synaptic weight is calculated as follows:

$$\Delta W_{ij} = \begin{cases} a^+ \times W_{ij} \times (1 - W_{ij}), & if\ t_j - t_i \leq 0 \\ a^- \times W_{ij} \times (1 - W_{ij}), & if\ t_j - t_i > 0 \end{cases} \quad (2)$$

where W_{ij} is the synaptic weight from the *jth* neuron in the previous layer (presynaptic neuron) to the *ith* neuron (postsynaptic neuron) in the STDP convolutional layer, t_j, t_i are the corresponding spike times of presynaptic neuron and postsynaptic neuron, and a^+, a^- are the two learning rate parameters of STDP rule with $a^+ > 0$, $a^- < 0$. Similar to the Hebbian STDP mechanism, when the presynaptic neuron fire before postsynaptic neuron, the synaptic weight is strengthened, that is long-term potentiation (LTP), otherwise, the synaptic weight is weakened, that is long-term depression (LTD). The multiplicative term $W_{ij}(1 - W_{ij})$ ensures the weights remain in the range of [0, 1].

G. OUR PROPOSED IMBALANCED R-STDP LEARNING RULE FOR IMBALANCED R-STDP CONVOLUTIONAL LAYER

When training data is class imbalanced, similar to DL networks, SNNs with commonly used R-STDP learning rule [33], [44], [45], [47], [48], will exhibit a bias towards the majority class, causing the minority class to be more misclassified. This is because the SNNs with commonly used R-STDP try to maximize the cumulative rewards during the training process. They implicitly assign an identical reward to all classes of samples assuming their equivalent importance. As a result, the classification model tends to correctly classify and favor the more frequent classes. In order to solve this problem, we propose an imbalanced R-STDP learning rule. The principle is to give more reward or punishment to the SNNs when it classifies a minority class sample than when it classifies a majority class sample.

1) IMBALANCED REWARD COEFFICIENT

For commonly used R-STDP learning rule, the synaptic weight changes depend not only on the pre- and post-synaptic spiking times, but also on a reward signal feedbacked by the decision-making layer, and the feedback signal does not differentiate the minority class samples and the majority class samples. In order to improve the networks' sensitivity to the minority class, we introduce an imbalanced reward coefficient to give more reward or punishment to the networks when they deal with the minority class samples. The imbalanced reward coefficient is used to multiply with the Hebbian/anti-Hebbian synaptic weight modification δ_{ij} which depends on the temporal order of the firing times of presynaptic and postsynaptic neurons. The reward coefficient is defined as follows:

$$r = \begin{cases} 1, & if\ x \in D_{min} \\ \lambda, & if\ x \in D_{maj} \end{cases} \tag{3}$$

where λ is a parameter to adjust the importance of majority class, and its value is between 0 and 1. x represents the current input sample of the networks. D is a training dataset consisting of tuples comprising of data and label, $D = \{x_n, d_n\}_{n=1}^{N}$, so D_{min} is the minority class sample set and D_{maj} is the majority class sample set, they also consist of tuples comprising of data and label. The value of the reward coefficient is 1 when the networks classify a minority class sample, and is λ when the networks classify a majority class sample.

2) IMBALANCED REWARD-MODULATED STDP

By combining the imbalanced reward coefficient with the Hebbian/anti-Hebbian STDP rule, the modification of each input synaptic weight of imbalanced R-STDP convolutional layer is calculated as follows:

$$\Delta W_{ij} = r \times \delta_{ij} \tag{4}$$

$$\delta_{ij} = \begin{cases} (\alpha \times a_r^+ + \beta \times a_p^-) W_{ij}(1 - W_{ij}), & if\ t_j - t_i \le 0 \\ (\alpha \times a_r^- + \beta \times a_p^+) W_{ij}(1 - W_{ij}), & if\ t_j - t_i > 0 \end{cases} \tag{5}$$

where W_{ij} is the synaptic weight from the *jth* neuron in the previous layer (presynaptic neuron) to the *ith* neuron (postsynaptic neuron) in the imbalanced R-STDP convolutional layer. t_j, t_i are the corresponding firing times of presynaptic neuron and postsynaptic neuron. α, β are two controlling factors which have different values for different feedback signals from the decision-making layer. If the networks recognize a sample correctly, the feedback signal is a reward, then $\alpha = 1$ and $\beta = 0$, and the synaptic weight modification δ_{ij} is determined according to Hebbian rule. Whereas if the networks recognize a sample incorrectly, the feedback signal is a punishment, then $\alpha = 0$ and $\beta = 1$, and δ_{ij} is determined according to anti-Hebbian rule. a_r^+, a_r^-, a_p^+, a_p^- are the learning rates of the imbalanced R-STDP learning rule which can scale the magnitude of weight change.

3) CUMULATIVE EFFECT OF IMBALANCED R-STDP ON THE LEARNING OF SNNs

In this section, we discuss the cumulative effect of our proposed imbalanced R-STDP learning rule on the learning of the SNNs. In our networks, for the synapse weight from presynaptic neuron j and postsynaptic neuron i, its modification caused by all the training samples is calculated as follows:

$$\begin{aligned} \Delta W_{ij}^N &= \sum_{n=1}^{N} \Delta W_{ij}(x_n) \times S_j(x_n) \\ &= \sum_{n=1}^{N} r \times \delta_{ij}(x_n) \times S_j(x_n) \\ &= \sum_{n_1=1}^{N_{min}} r \times \delta_{ij}(x_{n_1}) \times S_j(x_{n_1}) + \sum_{n_2=1}^{N_{maj}} r \times \delta_{ij}(x_{n_2}) \times S_j(x_{n_2}) \\ &= \sum_{n_1=1}^{N_{min}} 1 \times \delta_{ij}(x_{n_1}) \times S_j(x_{n_1}) + \sum_{n_2=1}^{N_{maj}} \lambda \times \delta_{ij}(x_{n_2}) \times S_j(x_{n_2}) \end{aligned} \tag{6}$$

where N is the number of training samples, N_{min}, N_{maj} is the number of minority class and majority class training samples, respectively. x_n represents the *nth* input sample of the

networks, and $S_j(x_n)$ is the spike train of the *jth* presynaptic neuron generated by the *nth* input sample.

For class imbalanced training dataset, N_{min} is much smaller than N_{maj}. If the reward coefficient is balanced, that is $\lambda = 1$, the effect of the majority class on the weight change of this synapse (the second term of (6)) will be much larger than that of the minority class (the first term of (6)). This will make the learning of the networks favors the majority class. By introducing our imbalanced reward coefficient r and make λ take value between 0 to 1, the effect of the majority class on the weight change of this synapse and that of the minority class can be greatly balanced. Thus, SNNs with our proposed imbalance R-STDP learning rule can work well on imbalanced data by mitigating the importance of majority class.

4) ADAPTIVE LEARNING RATES

Similar to DL networks, a learning rate that is too large will result in training error increase and even oscillation over training epochs, and too small will make the convergence get very slow or no convergence [74]. Thus, it is necessary to adjust the learning rate in the process of training. Besides, adaptive learning rate can also tackle the problem of overfitting. For our imbalanced R-STDP learning rule, the four learning rates a_r^+, a_r^-, a_p^+, a_p^- are dynamically adjusted after each epoch. These learning rates of the *Mth* epoch are determined by multiplying an adjustment factor with the corresponding learning rates of the $(M-1)th$ epoch as follows:

$$a_r^{\pm}(M) = (ER_{min} \times \rho_{maj} + ER_{maj} \times \rho_{min}) \times a_r^{\pm}(M-1) \quad (7)$$

$$a_p^{\pm}(M) = (CR_{min} \times \rho_{maj} + CR_{maj} \times \rho_{min}) \times a_p^{\pm}(M-1) \quad (8)$$

where ER_{min} and CR_{min} are the classification error rate and classification correct rate of the minority class respectively in the $(M-1)th$ epoch. ER_{maj} and CR_{maj} are the classification error rate and classification correct rate of the majority class respectively in the $(M-1)th$ epoch. ρ_{min}, ρ_{maj} are the ratio of size of the minority class and the majority class to the overall training sample dataset, respectively.

At the beginning of the training, neurons in the networks are not yet selective to any particular pattern and they may respond to different patterns. Therefore, the misclassification rate is very high, and the networks receive many punishment signals which rapidly weakens synaptic weights and generates highly selective or dead neurons. It will cause the networks to overfit and lead to performance degradation. This problem can now be tackled by our adaptive learning rates. Although the classification error rates (ER_{min}, ER_{maj}) is higher than classification correct rates (CR_{min}, CR_{maj}) when the networks begin training, according to (7) and (8) the learning rates a_r^+, a_r^- will larger than the learning rates a_p^+, a_p^-, and consequently according to (4) and (5) the magnitudes of weight change by a reward signal is larger than those by a punishment signal. In this way, the networks can avoid performance decline at the beginning of the learning process and can keep on classify more samples correctly in the following epochs. As training trials go on, the classification correct rate (CR_{min}, CR_{maj}) increases, the networks will receive more reward signals. According to (7) and (8) the learning rates a_p^+, a_p^- are larger than the learning rates a_r^+, a_r^- and consequently according to (4) and (5) the magnitudes of weight change by a punishment signal will be larger than those by a reward signal. So that the networks can correctly recognize the misclassified samples in the following epochs. By using the weighted classification error rate $ER_{min} \times \rho_{maj} + ER_{maj} \times \rho_{min}$ and weighted classification correct rate $CR_{min} \times \rho_{maj} + CR_{maj} \times \rho_{min}$ of both the minority class and the majority class, the impact of correctly and incorrectly classified training samples to the learning of the networks is balanced over the trials.

The pseudo-code of our proposed imbalanced R-STDP learning rule is introduced in detailed step-by-step in Algorithm 1. This implementation can be very easily implemented using popular deep learning frameworks such as PyTorch.

Algorithm 1 Weight Update of Imbalabced R-STDP Convolutional Layer

Input: Training dataset, $D = \{x_n, d_n\}_{n=1}^{N}$; Number of epochs, M_{ep}; Intial learning rates, $\{a_r^{\pm}, a_p^{\pm}\}$; Intial weight matrix, W_{int}; Reward coefficient, $r = \{1, \lambda\}$;

Output: Weight matrix of imbalanced R-STDP convolutional layer, W;

for *epoch* $= 1$ *to* M_{ep} **do**
  shuffle the training data D
  for $n = 1$ *to* N **do**
    $Pd_n \leftarrow$ convolution (x_n, W_{int}) ▷ Pd_n is predict label of the SNNs
    if $Pd_n = d_n$ **then**
      if $x_n \in D_{min}$ **then**
        $\Delta W \leftarrow 1 \times$STDP $(x_n, a_r^{\pm})$
      end if
      if $x_n \in D_{maj}$ **then**
        $\Delta W \leftarrow \lambda \times$STDP $(x_n, a_r^{\pm})$
      end if
    else
      if $x_n \in D_{min}$ **then**
        $\Delta W \leftarrow 1 \times$anti-STDP $(x_n, a_p^{\pm})$
      end if
      if $x_n \in D_{maj}$ **then**
        $\Delta W \leftarrow \lambda \times$anti-STDP $(x_n, a_p^{\pm})$
      end if
    end if
    $W(n) \leftarrow W(n-1) + \Delta W$
  end for
  calculation WER, WCR ▷ WER, WCR is weighted classification error rate and weighted classification correct rate
  $a_r^{\pm} \leftarrow$ adaptive learning rates$(a_r^{\pm}, WER)$ ▷ (7)
  $a_p^{\pm} \leftarrow$ adaptive learning rates$(a_p^{\pm}, WCR)$ ▷ (8)
end for

IV. EXPERIMENTAL RESULTS

In this section, we perform extensive experiments on three benchmark datasets, Fashion-MNIST (F-MNIST), CIFAR-10, CIFAR-100, and one medical image dataset ISIC-2018 to evaluate the effectiveness of our proposed SNNs with imbalanced R-STDP learning rule. We also compare our classification results with those of state-of-the-art methods for class-imbalance learning. And we analyze the effectiveness of our proposed imbalanced R-STDP learning rule on improving the classification performance of SNNs when training data is imbalanced.

A. DATASETS AND EXPERIMENTAL SETTINGS

In our experiments, we study the performance of the SNNs with imbalanced R-STDP and compare it with the state-of-the-art methods. There are two kinds of class imbalanced datasets used in our experiments. One is balanced datasets with an equal representation of each class in the training and testing splits (F-MNIST, CIFAR-10, and CIFAR-100). We alter the training splits of these balanced datasets to obtain imbalanced datasets with different imbalanced levels. Another is imbalanced datasets with a significant class imbalance in the training split (ISIC-2018 dataset). The details of these binary-class imbalanced datasets in our experiments are shown in Table 1. In order to clearly show the impact of the class imbalance problem on the performance of different methods of addressing the imbalance problem, for each of the three benchmark datasets we use three different imbalance ratios (IR), which is defined as $IR = N_{min}/N_{maj}$.

F-MNIST: It contains 70000 grayscale images of various products of clothing and accessories such as t-shirts, shoes, and other fashion articles. There are 10 classes of fashion product images and each class has 7000 images (6000 training and 1000 test examples) related with a label from 0 to 9. Each grayscale image has a resolution of a resolution of 28×28. It is a more challenging classification task than the original MNIST dataset. In our experiments, we use a binary-class imbalanced dataset constructed from F-MNIST dataset. We use the images labeled by 0, 2 (*T − Shirt*, *Pullover*) as the minority class and the images labeled by 1, 3 (*Trouser*, *Dress*) as the majority class. To implement the training distribution to different imbalance ratios, we reduce the training samples of the minority class to 4%, 2% and 1% of the majority class, respectively.

CIFAR-10: It is a more complex image classification dataset than F-MNIST. It contains 32×32 RGB color images of 10 classes of natural objects labeled from 0 to 9. The standard train/test split for each class is 5000/1000. In our experiments, we use the images labeled by 0 (*airplane*) as the minority class and the images labeled by 1 (*automobile*) as the majority class. To implement the training distribution to different imbalance ratios, we reduce the training sample of the minority class to 10%, 5% and 2% of the majority class, respectively.

CIFAR-100: It contains 60000 color images of natural objects belonging to 100 classes (600 images/class). There are 500 training and 100 testing examples in each class. The 100 classes can be further divided into 20 super-classes such as *household electrical devices*, *household furniture* and *trees*. In our experiments, we construct three binary-class imbalanced datasets from this dataset. The first dataset is denoted as CIFAR-100-household in the experiments, which contains two super-classes, *household electrical devices* and *household furniture*. We use *household electrical devices* as the minority class and *household furniture* as the majority class. The second dataset is denoted as CIFAR-100-tree1 which is a mixture of *maple tree* class and *oak tree* class, and we set *maple tree* as the minority class and the *oak tree* as the majority class. The third dataset is a mixture of *maple tree* class and *palm tree* class, and this dataset is denoted as CIFAR-100-tree2. The *maple tree* class is the minority class and the *palm tree* class is the majority class. For all these three constructed datasets, we reduce the training sample of the minority class to 20%, 10% and 5% of the majority class, respectively.

ISIC-2018 dataset [22]: It is currently the largest public dermoscopic image collection of skin lesions in the world, which includes 7 common skin disease images. In our experiments, we use the images of *melanoma* (*MEL*) lesion and those of *melanocytic nevi* (*NV*) lesion which are visually similar and very difficult for clinical specialists to distinguish. There are 1113 images of *MEL* and 6705 images of *NV*. We randomly sample 305 images in both categories for testing, and the remaining images for training. Therefore, the training samples of the two class has an imbalanced ratio of 12.6%. Since dermoscopy images often contain artifacts such as air bubbles and hairs which will results in detection accuracy loss [75], all the training and testing samples are pre-processed by using the DullRazor hair removing algorithm [76] which is widely used to preprocess skin images [38], [77].

TABLE 1. Details of the Binary-class Imbalanced Datasets in Experiments.

Dataset	Imbalance ratio	Number of training data		Number of testing data	
		Minority	Majority	Minority	Majority
F-MNIST	4%	480	12000	2000	2000
	2%	240			
	1%	120			
CIFAR-10	10%	500	5000	200	200
	5%	250			
	2%	100			
CIFAR-100-household	20%	500	2500	500	500
	10%	250			
	5%	125			
CIFAR-100-tree1	20%	100	500	100	100
	10%	50			
	5%	25			
CIFAR-100-tree2	20%	100	500	100	100
	10%	50			
	5%	25			
ISIC-2018	12.6%	808	6400	305	305

B. COMPARISON METHODS AND NETWORK ARCHITECTURES

Using the datasets listed in Table 1, we compare the classification performance of our proposed algorithm with those commonly used methods of addressing the imbalance problem for deep learning including both algorithm-level methods and data-level methods. For algorithm-level methods, we use three algorithms including deep Q-learning network based imbalanced classification (DQNimb), focal loss (FL), and mean squared false error loss (MSFE) for comparison, and all of these three methods are used on all the datasets in our experiments.

DQNimb [78]: This algorithm improves the classification performance of DQN on imbalanced datasets by using reward function to reduce the reward obtained by the majority class samples.

FL [62]: This algorithm improves the classification performance on imbalanced datasets by reshaping the standard cross entropy loss function, and reduces the loss assigned to well-classified examples.

MSFE [26]: This algorithm improves the classification performance on imbalanced datasets by using MSFE loss function to effectively capture classification errors from both majority class and minority class equally.

The network structure and parameters of these CNNs for comparison are shown in Table 2. The baseline CNNs are all trained by the cross-entropy loss function. For fairness and conclusive comparisons, the network structures of all the CNNs we used for comparison are kept as close as possible to our SNNs.

TABLE 2. Network Architecture Parameters of the Comparison Algorithms for Different Datasets.

Dataset	Number of convolutional layers	Number of neurons on convolutional layers (from bottom to up)
F-MNIST	3	25088, 6272, 1568
CIFAR-10	3	22500, 84500, 1800
CIFAR-100-household	3	1000,300,100
CIFAR-100-tree1	3	1000,100,10
CIFAR-100-tree2	3	1000,100,10
ISIC-2018	3	77760,14720,100

In addition to the algorithm-level methods, we also compared our algorithm with two data-level methods, random over-sampling (ROS) and random under-sampling (RUS). The network structures of these two methods are the same as the baseline SNNs and both of these two methods are used on all the datasets in our experiments.

ROS [56]: A re-sampling method to make the number of instances of two classes approximately equal by randomly replicating the minority samples.

RUS [56]: A re-sampling method to make the number of instances of two classes approximately equal by randomly removing the majority samples.

In our experiment, our SNNs with imbalanced R-STDP learning rule are evaluated on all the datasets in Table 1. For datasets of different complexity, our SNNs have different numbers of STDP convolutional layers. The baseline SNNs have the same network structure except that the last convolutional layer is a commonly used R-STDP convolutional layer which does not differentiate the majority and minority samples. The network architecture parameters of our SNNs with imbalanced R-STDP learning rule and the baseline SNNs for different datasets are shown in Table 3. The hyperparameters of our networks such as convolutional window size and initial learning rates are optimized using a grid search.

TABLE 3. Network Architecture Parameters of Our SNNs with Imbalanced R-STDP and Baseline SNNs for Different Datasets.

Dataset	Number of STDP convolutional layers	Number of imbalanced R-STDP convolutional layers	Number of neurons on convolutional layers (from bottom to up)
F-MNIST	1	1	12500, 2000
CIFAR-10	0	1	4500
CIFAR-100-household	1	1	150, 80
CIFAR-100-tree1	1	1	150, 80
CIFAR-100-tree2	1	1	150, 80
ISIC-2018	0	1	77760

C. EVALUATION METRICS

We use four evaluation metrics, sensitivity, specificity, G-mean and F_β-score to evaluate the classification performances of the algorithm we proposed and other methods. These metrics are more suitable than the overall accuracy when dealing with class imbalance problems [79]. Sensitivity, also called recall, measures the percentage of the positive group that is correctly predicted to be positive. It is not affected by class imbalance because it is only dependent on the positive group. Similarly, specificity measures the percentage of the negative group that was correctly predicted to be negative, and it is also not affected by imbalance. G-mean and F_β-score are the overall measures of network classification performance. The formulas of the evaluation metrics are as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{9}$$

$$Specificity = \frac{TN}{TN + FP} \tag{10}$$

$$G - mean = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}} \tag{11}$$

$$F_\beta - score = \frac{(1 + \beta^2) \times Precision \times Sensitivity}{\beta^2 \times Precision + Sensitivity} \tag{12}$$

where TP and TN are the number of positive and negative examples that are classified correctly, while FN and FP are the number of misclassified negative and positive examples respectively, and $Precision = TP/(TP + FP)$. F_β-score computes a weighted harmonic mean of precision and sensitivity. β is a parameter that controls the relative importance given to sensitivity over precision (typical values are 1/4, 1/2, 1, 2, 4). For F_1-score, the importance given to sensitivity and precision are balanced. For F_2-score, the value of β is 2.

TABLE 4. G-means of Different Methods on F-MNIST Dataset.

Dataset	Imbalance ratio	Baseline CNNs	CNNs with DQNimb [78]	CNNs with FL [62]	CNNs with MSFE [26]	Baseline SNNs	Our SNNs	SNNs with RUS	SNNs with ROS
F-MNIST	1	0.978	–	–	–	0.975	–	–	–
	4%	0.921	0.971	0.964	0.965	0.929	0.974	0.974	0.960
	2%	0.885	0.966	0.956	0.950	0.901	0.969	0.961	0.964
	1%	0.853	0.959	0.925	0.923	0.858	0.963	0.949	0.953

TABLE 5. G-means of Different Methods on CIFAR-10 Dataset.

Dataset	Imbalance ratio	Baseline CNNs	CNNs with DQNimb [78]	CNNs with FL [62]	CNNs with MSFE [26]	Baseline SNNs	Our SNNs	SNNs with RUS	SNNs with ROS
CIFAR-10	1	0.966	–	–	–	0.967	–	–	–
	10%	0.853	0.944	0.935	0.936	0.867	0.975	0.961	0.973
	5%	0.826	0.920	0.918	0.882	0.839	0.960	0.956	0.940
	2%	0.744	0.817	0.849	0.789	0.747	0.942	0.935	0.879

TABLE 6. G-means of Different Methods on CIFAR-100 Dataset.

Dataset	Imbalance ratio	Baseline CNNs	CNNs with DQNimb [78]	CNNs with FL [62]	CNNs with MSFE [26]	Baseline SNNs	Our SNNs	SNNs with RUS	SNNs with ROS
CIFAR-100-household	1	0.621	–	–	–	0.598	–	–	–
	20%	0.492	0.584	0.538	0.558	0.563	0.614	0.591	0.585
	10%	0.401	0.486	0.471	0.451	0.542	0.606	0.552	0.536
	5%	0.267	0.367	0.340	0.334	0.387	0.546	0.488	0.458
CIFAR-100-tree1	1	0.719	–	–	–	0.657	–	–	–
	20%	0.601	0.624	0.613	0.634	0.604	0.650	0.635	0.647
	10%	0.547	0.555	0.569	0.558	0.583	0.636	0.615	0.599
	5%	0.300	0.412	0.397	0.340	0.503	0.566	0.429	0.475
CIFAR-100-tree2	1	0.691	–	–	–	0.663	–	–	–
	20%	0.485	0.513	0.525	0.419	0.550	0.562	0.547	0.558
	10%	0.316	0.462	0.485	0.374	0.499	0.543	0.536	0.540
	5%	0	0.403	0.443	0.316	0.447	0.520	0.516	0.513

This value 2 gives twice weightage to sensitivity. In our experiments, we set the class label of the minority class as positive, and the class label of the majority class as negative.

D. RESULTS AND COMPARISONS

Our results on the deliberately imbalanced splits of the three benchmark datasets F-MNIST, CIFAR-10 and CIFAR-100 are reported in Table 4, 5, and 6. For these three benchmark datasets, the imbalanced reward coefficients for majority class samples of our SNNs are set as the value of their imbalance ratios. Each result value in the tables is the average value from three repeated experiments on the same dataset in order to make conclusive comparisons.

The experimental results on the imbalanced F-MNIST dataset are reported in Table 4. Our results show that compared with the baseline SNNs we achieve a gain of 0.045, 0.068 and 0.105 for the imbalance ratio of 4%, 2% and 1%, respectively. It can be seen that our imbalanced R-STDP learning rule can improve the classification performance of the SNNs on the imbalanced F-MNIST dataset, especially for high imbalance degree. Our SNNs also outperform the methods of RUS, ROS, and the methods based on CNNs for all the imbalance ratios. Moreover, we find that the performance of the baseline SNNs is better than that of the baseline CNNs for all the imbalance ratios.

Our experimental results on the imbalanced CIFAR-10 dataset are shown in the Table 5. It can be seen that compared with the results on the F-MNIST dataset, the performances of the two baseline networks drop more obviously with the increase of imbalance degree. This is because the CIFAR-10 dataset is more complex than the F-MNIST dataset. Compared with the baseline SNNs, our method can significantly improve the classification performance of the SNNs, with a gain of 0.108, 0.121 and 0.195 for the imbalance ratio of 10%, 5% and 2%, respectively, while the best boosts of the methods based on CNNs over their baseline CNNs are 0.091 (CNNs with DQNimb), 0.094 (CNNs with DQNimb) and 0.105 (CNNs with FL), respectively. Compared with the baseline networks, our SNNs have more obvious improvements than the methods based on CNNs. And for all the imbalance ratios, the performance of our SNNs is much better than that of the methods based on CNNs, and our SNNs also outperform the methods of RUS and ROS. The performances of RUS and ROS method are better than that of the methods based on CNNs on this imbalanced dataset. Note that for the highest imbalance degree of 2%, our G-mean value is obviously larger than those of other methods, which demonstrates the effectiveness of our proposed method on the classification of imbalanced data.

Our experiment results for the three binary-class imbalanced datasets extracted from the CIFAR-100 dataset are

reported in Table 6. The overall G-mean values of the listed methods seem to be not very high. This is because classification of CIFAR-100 is a more challenging task than CIFAR-10, and the number of training samples for the minority class is very small. For the different imbalanced ratios of the three datasets CIFAR-100-household, CIFAR-100-tree1 and CIFAR-100-tree2, our SNNs with imbalanced R-STDP learning rule significantly outperform the baseline SNNs and the methods based on CNNs. Compared with the baseline SNNs, the higher the imbalance degree, the more the improvement of the G-mean value of our method. Our performance is also better than those of RUS and ROS methods for all the three datasets. We also note that for these difficult datasets, the classification results of RUS and ROS methods are better than that of the methods based on CNNs.

We further evaluate the effectiveness of our algorithm on the medical imbalanced dataset, ISIC-2018 dataset which has an imbalance ratio of 12.6% in our experiments. The experiment results are summarized in Table 7. In order to mitigate the bias caused by the random selection of the data, each result in the table represents the average value from five experiments using five-fold cross validation. The G-mean, F_2-score, sensitivity and specificity are used as metrics to evaluate the classification performance of the different methods on this dataset. In medical diagnose domain, recognition of minority class is more important because they contain more useful information, therefore sensitivity should be given more importance [80], [81]. It represents the probability of identifying patients who do in fact have the disease. The higher the value of sensitivity, the less likely a diagnostic test returns false-positive results. We use F_2-score rather than F_1-score since it gives twice the weightage to sensitivity, thus identification of minority class is given more importance.

TABLE 7. Classification Performances of Different Methods on ISIC-2018 Dataset.

Method	G-mean	F_2-score	sensitivity	specificity
baseline SNNs	0.191	0.049	0.040	0.990
Our SNNs with λ=IR	0.906	0.873	0.857	0.946
Our SNNs with λ=IR/2	0.894	0.911	0.921	0.868
Baseline SNNs with ROS	0.606	0.475	0.437	0.840
Baseline SNNs with RUS	0.696	0.757	0.748	0.648
SNNs with STDP [38]	0.871	0.893	0.903	0.841
baseline CNNs	0.387	0.182	0.152	0.984
CNNs with DQNimb [78]	0.814	0.770	0.724	0.916
CNNs with FL [62]	0.783	0.731	0.676	0.907
CNNs with MSFE [26]	0.777	0.692	0.633	0.954

From Table 7, it can be seen that the class imbalance problem has great impact on the baseline SNNs which produce very low values of G-mean, F_2-score and sensitivity on this dataset. When the imbalanced reward coefficient for the majority class samples (NV) $\lambda = IR$, our SNNs can greatly improve the performance of the baseline SNNs, with the G-mean, F_2-score, and sensitivity increasing from 0.191 to 0.906, 0.049 to 0.873, and 0.040 to 0.857, respectively. Besides, although the methods based on CNNs, ROS and RUS can also improve the classification performance, they are not comparable to our methods. When we further increase the importance of minority class and take the imbalanced reward coefficient for the majority class samples as one-half of the imbalance ratio ($\lambda = IR/2$), we achieve a significant boost of the F_2-score and sensitivity over our SNNs with $\lambda = IR$, while maintaining acceptable G-mean and specificity. The F_2-scores and G-means of our SNNs with different reward coefficients for the NV are shown in Fig. 3. It can be seen that the F_2-score gets the maximum value when the reward coefficient of NV takes the value of $IR/3$. However, when the reward coefficient of NV is $IR/4$, because the reward given to NV is too small, the networks tend to recognize all NV samples as MEL during learning, so the G-mean value of the networks drop to 0. Therefore, for the imbalanced medical image datasets, by adjusting our proposed imbalanced reward coefficient to an appropriate value, our method can obtain high classification accuracy of the minority class which is very important for medical diagnosis, while keep good overall classification performance. We also compare with the classification results of STDP-based SNNs [38] which used convolutional SNNs with STDP learning rule and extra SVM classifier. And they used RUS method on this dataset to obtain balanced training datasets. Note that in the experiments of [38], they removed the images with serious noise while our experiments use all the images, and they used several image preprocessing steps while we only use one preprocessing of hair removal. The four metrics of our SNNs with $\lambda = IR/2$ which has only one learning layer are all better than this SNNs which contains three learning layers.

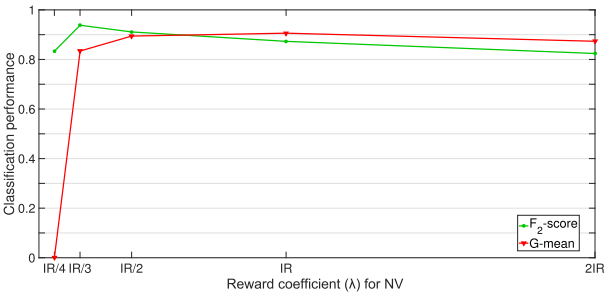

FIGURE 3. F_2-score and G-mean with different reward coefficients of NV on ISIC-2018 dataset.

Based on the results on the three imbalanced benchmark datasets and one imbalanced medical image dataset, it can be seen that comparing with the state-of-the-art methods for class imbalanced learning, our proposed imbalanced R-STDP learning rule for SNNs present the best performances on all of the tested datasets. Therefore, our method is quite promising to improve the performance of SNNs on imbalanced medical image classification problem. The improvement is very significant especially when there is a high imbalance degree in the training dataset. We also find that with the increase of data

imbalance degree, the performance decline of our SNNs with imbalanced R-STDP is much less than those of the CNNs for comparison. This is because our SNNs can learn better the features of the minority class than the CNNs when there is only a small number of training data. Moreover, the method has another advantage, by adjusting our imbalanced reward coefficient to an appropriate value, our SNNs can obtain high classification accuracy for a certain class. This is very useful for identify disease samples in medical diagnosis tasks.

E. ABLATION STUDY: EFFECT OF IMBALANCED R-STDP LEARNING RULE

In this section, we analyze why our proposed imbalanced R-STDP learning rule can improve the performance of SNNs on imbalanced datasets. In the learning progress of our SNNs and the baseline SNNs, the input weights of convolutional layers gradually convergence towards 0 or 1 after an adequate number of training epochs. Here we define the learning convergence of one convolutional layer as

$$C = \frac{\sum_f \sum_i W_{f,i} \times (1 - W_{f,i})}{n_W} \tag{13}$$

where, $W_{f,i}$ is the *ith* input synaptic weight of the *fth* feature of this convolutional layer, and n_W is the total number of input synaptic weights of all the features. Since these feature maps corresponding to these features correspond to the neurons of the decision layer through global pooling, they can also be assigned to different input categories. We denote the learning convergences of the minority class and the majority class of this layer by C_{min} and C_{maj} respectively. The smaller the value of C_{min} and C_{maj}, the better the features learned by this convolutional layer.

We make comparative studies of the learning process of the convolutional layer of our SNNs with imbalanced R-STDP learning rule and the baseline SNNs on the CIFAR-10 benchmark dataset. The C_{min} and C_{maj} values of our SNNs and those of the baseline SNNs over the 2000 training epochs are illustrated in Fig. 4. It can be seen that when the training data is balanced, both C_{min} and C_{maj} of the baseline SNNs keep changing towards 0 and then convergence to the minimum almost at the same time. However, when the training data is imbalanced with $IR = 10\%$, C_{maj} quickly convergences to 0, while C_{min} cannot convergence towards 0. After about 100 training epochs, its value no longer changes and stays at around 0.18. This indicates that the features of the minority class have not been well learned, while the learning of the majority class quickly finished. This shows that when the training data is imbalanced, learning of the baseline SNNs with balanced R-STDP learning rule tends to bias towards the majority class which has much more training samples than the minority class. Therefore, the baseline SNNs cannot effectively learn from imbalanced data features that discriminate the minority class from majority classes. Our proposed imbalanced R-STDP learning rule can be a good solution to this problem, as can be seen from the Fig. 4, both C_{min} and C_{maj} of our SNNs can convergence towards 0 at the same time. It indicates that our proposed imbalanced R-STDP learning rule can balance the effects of the majority class and the minority class on the learning of the networks, and consequently avoid the learning bias caused by the class imbalanced training data.

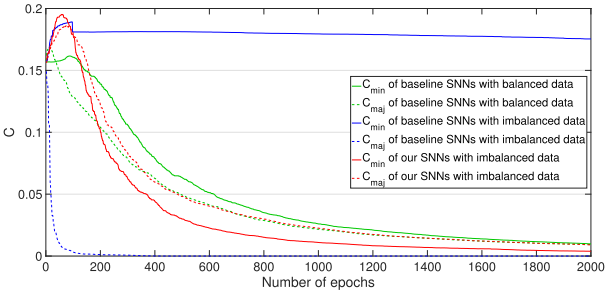

FIGURE 4. C values of baseline SNNs and SNNs with imbalabced R-STDP on balance data (CIFAR-10) and imbalanced data (CIFAR-10, IR=10%).

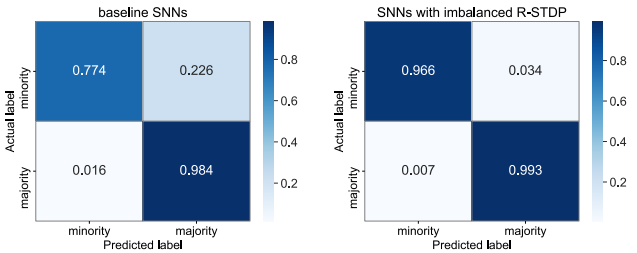

FIGURE 5. Confusion matrix comparisons of baseline SNNs and SNNs with imbalanced R-STDP on the CIFAR-10 dataset ($IR = 5\%$). (a) Confusion matrix of baseline SNNs. (b) Confusion matrix of SNNs with imbalanced R-STDP.

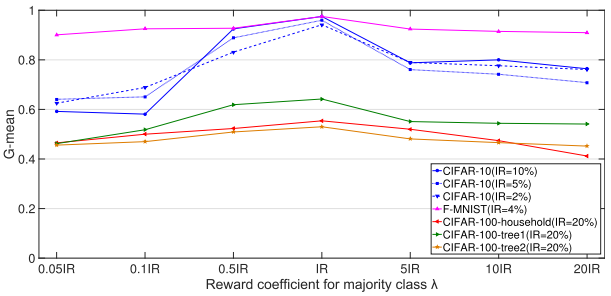

FIGURE 6. G-means for different reward coefficients of the majority class on the three benchmark datasets with different imbalance ratios.

We also compared the classification results of the baseline SNNs and our SNNs with imbalanced R-STDP rule on the CIFAIR-10 dataset with $IR = 5\%$ which is a very high imbalance degree. The confusion matrices of the two SNNs are represented in Fig. 5. It can be seen that our proposed imbalanced R-STDP rule significantly improves the classification rate of the minority class from 0.774 to 0.966, and the classification rate of the majority class is also improved from 0.984 to 0.993. It demonstrates that our proposed method can well improve the classification performance of SNNs on imbalanced datasets.

In addition, we study the effect of different values of reward coefficients of the majority class λ, on the G-mean of our SNNs with imbalanced R-STDP on the three benchmark datasets for different imbalance ratios. The results are shown in Fig. 6. It can be seen that for all these imbalanced datasets of different imbalance ratios, the G-mean of our SNNs reaches its maximum value when λ takes the value of IR. It indicates that our SNNs have the best overall classification performance for both classes when $\lambda = IR$.

V. CONCLUSION

Although DL methods have achieved remarkable success in class imbalanced medical image analysis tasks in recent years, their expensive computational costs and the need for a great amount of labeled data limit their application. SNNs with the advantages of low power consumption, hardware friendliness and can learning feature form few data then become a good candidate for medical image analysis. However, the class imbalanced problem of medical image will have a serious negative effect on the overall performance in SNNs. Thus, we proposed an imbalanced R-STDP learning rule for SNNs by introducing a reward coefficient, and this is the first work that modifies the learning algorithm of SNNs to solve the problem of class imbalance. The classification results show that the SNNs with imbalanced R-STDP learning rule present the best performances on all of the tested datasets comparing with the compared state-of-the-art methods. We also find that our SNNs can obtain high classification accuracy for a certain class by adjusting our imbalanced reward coefficient to an appropriate value. Moreover, our method does not alter the original data distribution, so it does not lose data information and has little effect on training times. The other advantage of using our method is that the reward coefficient can set the class-dependent rewards based on the data statistic of the training set, so the method is relatively easy to implement and integrate into existing models.

In the future, we will try to apply our method to solve the multi-class imbalance problem, which is more challenging than solving the binary class imbalance problem. We will also try to further improve our network architectures by using multiple imbalanced R-STDP convolutional layers, which is possible to further improve the recognition accuracy of the networks.

REFERENCES

[1] A. A. Agyeman and R. Ofori-Asenso, "Perspective: Does personalized medicine hold the future for medicine?" *J. Pharmacy Bioallied Sci.*, vol. 7, no. 3, pp. 239–244, Jul./Sep. 2015.

[2] J. D. Fauw, J. R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O'Donoghue, and D. Visentin, "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nature Med.*, vol. 24, no. 9, pp. 1342–1350, Aug. 2018.

[3] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," in *Classification in BioApps*. Cham, Switzerland: Springer, 2018, pp. 323–350.

[4] G. Litjens, T. Kooi, B. E. Bejnordi, M. Ghafoorian, A. A. A. Setio, F. Ciompi, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[5] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, May 2019.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.

[7] V. Mnih, K. Kavukcuoglu, M. G. Bellemare, D. Silver, A. A. Rusu, J. Veness, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[8] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, and B. Hu, "Liver fibrosis classification based on transfer learning and FCNet for ultrasound images," *IEEE Access*, vol. 5, pp. 5804–5810, 2017.

[9] P. Tschandl *et al.*, "Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: An open, Web-based, international, diagnostic study," *Lancet Oncol.*, vol. 20, no. 7, pp. 938–947, Jul. 2019.

[10] Z. Yan, Y. Zhan, Z. Peng, S. Liao, Y. Shinagawa, D. N. Metaxas, and X. S. Zhou, "Bodypart recognition using multi-stage deep learning," in *Information Processing in Medical Imaging*, vol. 24. Cham, Switzerland: Springer, 2015, pp. 449–461.

[11] H. R. Roth, C. T. Lee, H.-C. Shin, A. Seff, L. Kim, J. Yao, L. Lu, and R. M. Summers, "Anatomy-specific classification of medical images using deep convolutional nets," in *Proc. IEEE 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 101–104.

[12] B. Kong, S. Sun, X. Wang, Q. Song, and S. Zhang, "Invasive cancer detection utilizing compressed convolutional neural network and transfer learning," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018*. Cham, Switzerland: Springer, 2018, pp. 156–164.

[13] J. G. Nam, S. Park, E. J. Hwang, J. H. Lee, K.-N. Jin, K. Y. Lim, T. H. Vu, J. H. Sohn, S. Hwang, J. M. Goo, and C. M. Park, "Development and validation of deep learning–based automatic detection algorithm for malignant pulmonary nodules on chest radiographs," *Radiology*, vol. 290, no. 1, pp. 218–228, Jan. 2019.

[14] D. Nie, L. Wang, R. Trullo, J. Li, P. Yuan, J. Xia, and D. Shen, "Segmentation of craniomaxillofacial bony structures from MRI with a 3D deep-learning based cascade framework," in *Proc. Int. Workshop Mach. Learn. Med. Imag.* New York, NY, USA: Springer, 2017, pp. 266–273.

[15] J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein, M. Bendszus, and A. Biller, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, vol. 129, pp. 460–469, Apr. 2016.

[16] G. Wu, M. Kim, Q. Wang, B. C. Munsell, and D. Shen, "Scalable high-performance image registration framework by unsupervised deep feature representations learning," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1505–1516, Jul. 2016.

[17] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: A survey," *Mach. Vis. Appl.*, vol. 31, nos. 1–2, p. 8, Jan. 2020.

[18] F. Altaf, S. M. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: Concepts, methods, challenges, and future directions," *IEEE Access*, vol. 7, pp. 99540–99572, 2019.

[19] M. Rezaei, T. Uemura, J. Näppi, H. Yoshida, C. Lippert, and C. Meinel, "Generative synthetic adversarial network for internal bias correction and handling class imbalance problem in medical image diagnosis," *Proc. SPIE*, vol. 11314, pp. 82–89, Art. no. 113140E.

[20] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.

[21] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.

[22] N. Codella, V. Rotemberg, D. Gutman, P. Tschandl, M. E. Celebi, S. Dusza, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC)," 2019, *arXiv:1902.03368*. [Online]. Available: http://arxiv.org/abs/1902.03368

[23] M. Sahare and H. Gupta, "A review of multi-class classification for imbalanced data," *Int. J. Adv. Comput. Res.*, vol. 2, no. 3, p. 160, 2012.

[24] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, no. 1, p. 27, Mar. 2019.

[25] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing, "Learning from class-imbalanced data: Review of methods and applications," *Expert Syst. Appl.*, vol. 73, pp. 220–239, May 2017.
[26] S. Wang, W. Liu, J. Wu, L. Cao, Q. Meng, and P. J. Kennedy, "Training deep neural networks on imbalanced data sets," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 4368–4374.
[27] V. Garcia, J. S. Sanchez, R. A. Mollineda, R. Alejo, and J. M. Sotoca, "The class imbalance problem in pattern classification and learning," in *Proc. II Congreso Español de Informática*, Zaragoza, Spain, 2007, pp. 283–291.
[28] M. Hao, Y. Wang, and S. H. Bryant, "An efficient algorithm coupled with synthetic minority over-sampling technique to classify imbalanced PubChem BioAssay data," *Analytica Chim. Acta*, vol. 806, pp. 117–127, Jan. 2014.
[29] A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera, "Algorithm-level approaches," in *Learning from Imbalanced Data Sets*. Cham, Switzerland: Springer, 2018, pp. 123–146.
[30] Y. Park and M. Kellis, "Deep learning for regulatory genomics," *Nature Biotechnol.*, vol. 33, no. 8, pp. 825–826, Aug. 2015.
[31] S. Ghosh-Dastidar and H. Adeli, "Spiking neural networks," *Int. J. Neural Syst.*, vol. 19, no. 4, pp. 295–308, Aug. 2009.
[32] K. Roy, A. Jaiswal, and P. Panda, "Towards spike-based machine intelligence with neuromorphic computing," *Nature*, vol. 575, no. 7784, pp. 607–617, Nov. 2019.
[33] M. Mozafari, M. Ganjtabesh, A. Nowzari-Dalini, S. J. Thorpe, and T. Masquelier, "Bio-inspired digit recognition using reward-modulated spike-timing-dependent plasticity in deep convolutional networks," *Pattern Recognit.*, vol. 94, pp. 87–95, Oct. 2019.
[34] S. R. Kheradpisheh, M. Ganjtabesh, S. J. Thorpe, and T. Masquelier, "STDP-based spiking deep convolutional neural networks for object recognition," *Neural Netw.*, vol. 99, pp. 56–67, Mar. 2018.
[35] C. Lee, G. Srinivasan, P. Panda, and K. Roy, "Deep spiking convolutional neural network trained with unsupervised spike-timing-dependent plasticity," *IEEE Trans. Cognit. Develop. Syst.*, vol. 11, no. 3, pp. 384–394, Sep. 2019.
[36] N. Kasabov, V. Feigin, Z.-G. Hou, Y. Chen, L. Liang, R. Krishnamurthi, M. Othman, and P. Parmar, "Evolving spiking neural networks for personalised modelling, classification and prediction of spatio-temporal patterns with a case study on stroke," *Neurocomputing*, vol. 134, pp. 269–279, Jun. 2014.
[37] S. Ghosh-Dastidar and H. Adeli, "Improved spiking neural networks for EEG classification and epilepsy and seizure detection," *Integr. Comput.-Aided Eng.*, vol. 14, no. 3, pp. 187–212, May 2007.
[38] Q. Zhou, Y. Shi, Z. Xu, R. Qu, and G. Xu, "Classifying melanoma skin lesions using convolutional spiking neural networks with unsupervised stdp learning rule," *IEEE Access*, vol. 8, pp. 101309–101319, 2020.
[39] S. Singh, D. Gupta, R. S. Anand, and V. Kumar, "Nonsubsampled shearlet based CT and MR medical image fusion using biologically inspired spiking neural network," *Biomed. Signal Process. Control*, vol. 18, pp. 91–101, Apr. 2015.
[40] N. Caporale and Y. Dan, "Spike timing–dependent plasticity: A Hebbian learning rule," *Annu. Rev. Neurosci.*, vol. 31, pp. 25–46, Jul. 2008.
[41] T. Masquelier, R. Guyonneau, and S. J. Thorpe, "Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains," *PLoS ONE*, vol. 3, no. 1, p. e1377, Jan. 2008.
[42] T. Masquelier, "STDP allows Close-to-Optimal spatiotemporal spike pattern detection by single coincidence detector neurons," *Neuroscience*, vol. 389, pp. 133–140, Oct. 2018.
[43] R. V. Florian, "Reinforcement learning through modulation of Spike-Timing-Dependent synaptic plasticity," *Neural Comput.*, vol. 19, no. 6, pp. 1468–1502, Jun. 2007.
[44] R. Legenstein, D. Pecevski, and W. Maass, "A learning theory for reward-modulated Spike-Timing-Dependent plasticity with application to biofeedback," *PLoS Comput. Biol.*, vol. 4, no. 10, Oct. 2008, Art. no. e1000180.
[45] E. M. Izhikevich, "Solving the distal reward problem through linkage of STDP and dopamine signaling," *Cerebral Cortex*, vol. 17, no. 10, pp. 2443–2452, Oct. 2007.
[46] M. Mikaitis, G. Pineda García, J. C. Knight, and S. B. Furber, "Neuromodulated synaptic plasticity on the SpiNNaker neuromorphic system," *Frontiers Neurosci.*, vol. 12, p. 105, Feb. 2018.
[47] M. Mozafari, S. R. Kheradpisheh, T. Masquelier, A. Nowzari-Dalini, and M. Ganjtabesh, "First-spike-based visual categorization using reward-modulated STDP," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 6178–6190, Dec. 2018.
[48] J. C. Vasquez Tieck, P. Becker, J. Kaiser, I. Peric, M. Akl, D. Reichard, A. Roennau, and R. D. Dillmann, "Learning target reaching motions with a robotic arm using brain-inspired dopamine modulated STDP," in *Proc. IEEE 18th Int. Conf. Cognit. Informat. Cognit. Comput. (ICCI*CC)*, Jul. 2019, pp. 54–61.
[49] Z. Bing, C. Meschede, K. Huang, G. Chen, F. Rohrbein, M. Akl, and A. Knoll, "End to end learning of spiking neural network based on R-STDP for a lane keeping vehicle," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 1–8.
[50] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[51] E. Nasr-Esfahani, S. Samavi, N. Karimi, S. M. R. Soroushmehr, M. H. Jafari, K. Ward, and K. Najarian, "Melanoma detection by analysis of clinical images using convolutional neural network," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 1373–1376.
[52] A. Menegola, M. Fornaciali, R. Pires, F. V. Bittencourt, S. Avila, and E. Valle, "Knowledge transfer for melanoma screening with deep learning," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 297–300.
[53] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinge, "Skin lesion classification using hybrid deep neural networks," in *Proc. ICASSP-IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1229–1233.
[54] C. X. Ling and C. Li, "Data mining for direct marketing: Problems and solutions," *Knowl. Discovery Databases*, vol. 98, pp. 73–79, Aug. 1998.
[55] M. Kubat, R. C. Holte, and S. Matwin, "Machine learning for the detection of oil spills in satellite radar images," *Mach. Learn.*, vol. 30, nos. 2–3, pp. 195–215, 1998.
[56] N. Japkowicz, "The class imbalance problem: Significance and strategies," in *Proc. Int. Conf. Artif. Intell.*, vol. 56, 2000, pp. 111–117.
[57] J. Wang, M. Xu, H. Wang, and J. Zhang, "Classification of imbalanced data by using the SMOTE algorithm and locally linear embedding," in *Proc. 8th Int. Conf. Signal Process.*, vol. 3, 2006, pp. 1815–1818.
[58] R. Huang, J. Liang, F. Jiang, F. Zhou, N. Cheng, T. Wang, and B. Lei, "MelanomaNet: An effective network for melanoma detection," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 1613–1616.
[59] A. B. Qasim, I. Ezhov, S. Shit, A. Sekuboyina, O. Schoppe, J. C. Paetzold, F. Kofler, J. Lipkova, H. Li, and B. Menze, "Red-GAN: Attacking class imbalance via conditioned generation. Yet another medical imaging perspective," 2020, *arXiv:2004.10734*. [Online]. Available: http://arxiv.org/abs/2004.10734
[60] T. Iqbal and H. Ali, "Generative adversarial network for medical images (MI-GAN)," *J. Med. Syst.*, vol. 42, no. 11, p. 231, Oct. 2018.
[61] N. V. Chawla, N. Japkowicz, and A. Kotcz, "Editorial: Special issue on learning from imbalanced data sets," *ACM SIGKDD Explor. Newslett.*, vol. 6, no. 1, pp. 1–6, Jun. 2004.
[62] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.
[63] G. S. Tran, T. P. Nghiem, V. T. Nguyen, C. M. Luong, and J.-C. Burie, "Improving accuracy of lung nodule classification using deep learning with focal loss," *J. Healthcare Eng.*, vol. 2019, pp. 1–9, Feb. 2019.
[64] N. Thai-Nghe, Z. Gantner, and L. Schmidt-Thieme, "Cost-sensitive learning methods for imbalanced data," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2010, pp. 1–8.
[65] T. Li, Z. Han, B. Wei, Y. Zheng, Y. Hong, and J. Cong, "Robust screening of COVID-19 from chest X-ray via discriminative cost-sensitive learning," 2020, *arXiv:2004.12592*. [Online]. Available: http://arxiv.org/abs/2004.12592
[66] E. Donati, M. Payvand, N. Risi, R. Krause, and G. Indiveri, "Discrimination of EMG signals using a neuromorphic implementation of a spiking neural network," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 5, pp. 795–803, Oct. 2019.
[67] S. G. Wysoski, L. Benuskova, and N. Kasabov, "Fast and adaptive network of spiking neurons for multi-view visual pattern recognition," *Neurocomputing*, vol. 71, nos. 13–15, pp. 2563–2575, Aug. 2008.
[68] A. Tavanaei and A. Maida, "Bio-inspired multi-layer spiking neural network extracts discriminative features from speech signals," in *Neural Information Processing*. Cham, Switzerland: Springer, 2017, pp. 899–908.

[69] M. J. McMahon, O. S. Packer, and D. M. Dacey, "The classical receptive field surround of primate parasol ganglion cells is mediated primarily by a non-GABAergic pathway," *J. Neurosci.*, vol. 24, no. 15, pp. 3736–3745, Apr. 2004.
[70] H. Meinhardt and A. Gierer, "Pattern formation by local self-activation and lateral inhibition," *BioEssays*, vol. 22, no. 8, pp. 753–760, Jul. 2000.
[71] H. Mostafa and G. Cauwenberghs, "A learning framework for winner-take-all networks with stochastic synapses," *Neural Comput.*, vol. 30, no. 6, pp. 1542–1572, Jun. 2018.
[72] D. Scherer, A. Müller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in *Proc. Int. Conf. Artif. Neural Netw.* Berlin, Germany: Springer, 2010, pp. 92–101.
[73] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 411–426, Mar. 2007.
[74] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, MA, USA: MIT Press, 2016.
[75] M. E. Celebi, Q. Wen, H. Iyatomi, K. Shimizu, H. Zhou, and G. Schaefer, "A state-of-the-art survey on lesion border detection in dermoscopy images," *Dermoscopy Image Anal.*, vol. 10, pp. 97–129, Sep. 2015.
[76] T. Lee, V. Ng, R. Gallagher, A. Coldman, and D. McLean, "Dullrazor: A software approach to hair removal from images," *Comput. Biol. Med.*, vol. 27, no. 6, pp. 533–543, Nov. 1997.
[77] L. Talavera-Martínez, P. Bibiloni, and M. González-Hidalgo, "Comparative study of dermoscopic hair removal methods," in *Proc. ECCOMAS Thematic Conf. Comput. Vis. Med. Image Process.* Cham, Switzerland: Springer, 2019, pp. 12–21.
[78] E. Lin, Q. Chen, and X. Qi, "Deep reinforcement learning for imbalanced classification," *Applied Intelligence*, vol. 50, pp. 2488–2502, Mar. 2020.
[79] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 4, pp. 463–484, Jul. 2012.
[80] A. Dascalu and E. O. David, "Skin cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope," *EBioMedicine*, vol. 43, pp. 107–113, May 2019.
[81] D. Devarriya, C. Gulati, V. Mansharamani, A. Sakalle, and A. Bhardwaj, "Unbalanced breast cancer data classification using novel fitness functions in genetic programming," *Expert Syst. Appl.*, vol. 140, Feb. 2020, Art. no. 112866.

QIAN ZHOU received the Ph.D. degree in control theory and control engineering from the University of Nankai, Tianjin, China, in 2010. She is currently an Associate Professor with the Hebei University of Technology, Tianjin. Her current research interests include biological neural networks, spiking neural networks, STDP plasticity, and deep learning.

CONG REN received the B.S. degree in biological engineering from the Hebei University of Technology, Tianjin, China, where she is currently pursuing the M.S. degree in biomedical engineering. Her current research interests include spiking neural networks and deep learning.

SAIBING QI received the B.S. degree in electrical engineering from the Hebei University of Science and Technology, Hebei, China. He is currently pursuing the M.S. degree in biomedical engineering with the Hebei University of Technology, Tianjin, China. His current research interests include spiking neural networks and deep learning.