

Received April 10, 2020, accepted May 16, 2020, date of publication May 28, 2020, date of current version June 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2998098

Classifying Melanoma Skin Lesions Using Convolutional Spiking Neural Networks With Unsupervised STDP Learning Rule

QIAN ZHOU^{ID}, YAN SHI, ZHENGHUA XU, RUOWEI QU^{ID}, AND GUIZHI XU^{ID}

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: Zhenghua Xu (zhenghua.xu@hebut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61305077, Grant 61906063, Grant 51977060, and Grant 51737003, in part by the Natural Science Foundation of Tianjin, China, under Grant 19JCQNJC00400, in part by the 100 Talents Plan of Hebei Province under Grant E2019050017, and in part by the Yuanguang Scholar Fund of Hebei University of Technology, China.

ABSTRACT Deep learning methods have made some achievements in the automatic skin lesion recognition, but there are still some problems such as limited training samples, too complicated network structure, and expensive computational costs. Considering the inherent power-efficiency, biological plausibility and good image recognition performance of spiking neural networks (SNNs), in this paper we make malignant melanoma and benign melanocytic nevi skin lesions classification using convolutional SNNs with unsupervised spike-timing-dependent plasticity (STDP) learning rule. Efficient temporal coding, event driven learning rule and winner-take-all (WTA) mechanism together ensure sparse spike coding and efficient learning of our networks which achieve an average accuracy of 83.8%. We further propose to use feature selection to select more diagnostic features to improve the classification performance of our networks. Our SNNs with feature selection reach an average accuracy of 87.7%. Experimental results show that comparing to CNNs that need to be trained from scratch, our SNNs (with and without feature selection) not only achieve much better classification accuracies but also have much better runtime efficiency. Moreover, although the pretrained CNNs models can achieve similar running time, our proposed SNNs are more stable and easier to use than the pretrained CNNs because we do not need to try many pretrained models any more, and our SNNs also have much better classification accuracies than the pretrained CNNs. In addition, our networks have only three convolutional layers, and the complexity of the model and the parameters that need to be trained in the networks are greatly reduced. Our works show that STDP-based SNNs are very beneficial for the implementation of automated skin lesion classifiers on small portable devices.

INDEX TERMS Melanoma recognition, convolutional spiking neural networks, STDP, deep learning.

I. INTRODUCTION

Skin cancer is one of the most common worldwide malignancy [1]. It has been found that over the past three decades, the people diagnosed as skin cancer is more than those diagnosed as all other cancers combined [2]. Malignant melanoma is a kind of high-risk deadly skin cancer. Early detection of malignant lesions has great significance for helping the clinicians to improve the chances of survival [3]. Because of the visual similarities of some lesion types, correct diagnosis

is a challenging task for clinicians, and is largely dependent on the experience.

Machine learning (ML) methods have shown their advantages in detecting key features and patterns from complex datasets, thus are suitable to perform classification, prediction or estimation tasks [4]. In recent years there is growing trend on the application of ML methods as an aid to accurate and automated cancer diagnosis and detection [4], [5]. The application of ML techniques has significantly improved the accuracy of cancer prediction by 15%–20% over the past decades [5]. Deep learning, in particular CNNs supported by advanced computing technology and large datasets, has become one of the most popular and powerful machine learning

The associate editor coordinating the review of this manuscript and approving it for publication was Ganesh Naik^{ID}.

methods in image recognition and classification [6], and has been applied to the classification of skin lesions [7]–[9]. The CNNs can learn from training image set and automatically extract important features for classification. The prior knowledge and complicated image preprocessing which are very necessary in the image classification using traditional ML algorithms, are no longer greatly demanded. Some classifiers based on deep learning methods have shown to classify images of skin cancer with the performance comparable to the level of skilled dermatologists [8]. Thus CNNs have the potential to help develop dermatologist-level, computer-aided fast skin lesion classifiers.

However, at present there is still a lack of high-quality medical image datasets those can be used for training. This mainly refers to the lack of annotated/labeled data or lack of images for the abnormal classes [10]. CNNs with simple architectures are prone to overfitting on the small training datasets. Some researchers apply very deep CNNs architecture (for example, Resnet152, which has 152 layers [11]). Although this can improve the classification performance of the networks, it also adds more computing costs, which is major barrier in clinical applications [12]–[14]. More researchers tend to use pretrained CNNs to classify skin lesions [15]–[26], which avoids the problem of overfitting, but the network architectures are still very complex. And the pretrained CNNs use features learn from natural image datasets (such as ImageNet) without sufficiently considering the features of medical images [17], which limits the application of transfer learning in the field of medical image analysis. So unsupervised learning methods are needed in medical image analysis when labeled training data is scarce. Moreover, in order to facilitate the implementation of automated skin lesion classifiers, and to make the access convenient and cheap, for example through installation of apps on mobile devices, the computational costs of the CNNs-based skin lesion classifiers must be reduced.

SNNs have been emerged as an ideal biologically inspired neuromorphic-computing paradigm for realizing energy-efficient on-chip intelligence hardware [27], [28]. Like in the brain, information in SNNs is encoded not only by spike rates, but also by precise spike times and spike latencies of neurons. Furthermore, SNNs usually apply bio-inspired STDP as unsupervised local learning (synaptic weight modification) rule, which is crucial for brain learning. STDP is observed in different brain areas [29]–[32], in particular in the visual cortex [33]–[35]. Weight modification of this synaptic plasticity depends on the temporal order and time difference between presynaptic and postsynaptic spikes. Since individual spike events in the networks can be made sparse in time, learning in SNNs in principle is sparse and event-driven, leading to low computational consumption. Moreover, SNNs equipped with unsupervised STDP learning rule have the capability to learn the spatio-temporal patterns of input signals, especially in an online mode [36]–[38]. SNNs with multiple hidden layers can extract more complex features from input to obtain high classification performance [39], [40]. There are deep SNNs with

unsupervised learning rules whose performances are comparable to the traditional CNNs for small-scale image recognition tasks [41]–[43]. Recently SNNs have shown very good performance in the task of pattern recognition such as visual processing [41], [44], [45] and speech recognition [46], [47]. They have also been applied to predict strokes and seizures in the medical diagnosis based on electroencephalograms (EEG) classification [48], [49]. Kasabov *et al* [48] proposed a new spiking neural networks reservoir system for early prediction of occurrence of stroke on an individual basis. Their results showed that their method had obvious advantages in the accuracy and time of stroke event prediction compared with standard ML algorithms. Ghosh-Dastidar *et al* [49] proposed an efficient SNNs model for epilepsy and epileptic seizure detection using EEG. Their model reached a classification accuracy of 92.5%.

Although SNNs have shown good performance in natural image classification, to our knowledge, this is the first work that applying SNNs on medical image classification tasks. It is worth noting that medical image classification is more difficult than traditional natural image classification [50]. This is mainly because unlike natural images, the boundaries between different issues, or between normal tissues and lesion areas are usually not clear, and their texture differences are not obvious. Besides, some imaging principles used in medical images are different from those used in natural images. In addition, comparing to the most similar existing SNNs (i.e., Kheradpisheh *et al.* [41]), in order to improve the classification performance of the SNNs, we further propose to use feature selection operation to help pick out diagnostic features which may not be so prominent. In our model, input skin lesions images are encoded into a spiking train by a difference of gaussians (DOG) filter. Then, we use unsupervised STDP rule in combination with WTA and lateral inhibition mechanisms to extract prominent hierarchical features from the spiking trains of skin images in CNNs-architectures (with only a few convolutional and pooling layers). We use feature selection method to select diagnostic features from extracted features. Finally, a SVM classifier is used to classify the categories of the input skin lesions images based on the selected features.

In our networks, STDP learning rule and lateral inhibition mechanism work together to enable the networks to process input images with sparse but informative spikes. For the SNNs with unsupervised STDP learning rule, this kind of convolutional SNNs architectures provide so far the best classification performance in processing large scale natural images [41]. Our STDP-based SNNs can learn from the small training datasets quickly and effectively in an unsupervised manner and do not suffer from the overfitting challenge caused by small training set size [41], [51], [52]. Therefore, our networks can well solve the problem of lacking labeled data in the field of medical image analysis. Moreover, our networks have only a few convolutional layers and pooling layers, so there are a very small number of parameters need to be trained. Besides, efficient temporal coding, WTA and

lateral inhibition mechanism ensure the sparseness of the spike coding, these together enable the networks to require very little computational and memory consumption, which facilitate the deployment of skin lesions automatic diagnosis systems to smart phones and other portable devices, as well as cloud computations.

The classification results of our networks are compared to those of the art-of-date CNNs. The results show that our networks which have only a few hidden layers can learn features well from the small skin lesions training datasets, and after using feature selection they have better classification performance than CNNs trained from scratch and pretrained. Our networks take much less computational costs than traditional CNNs trained from scratch, and their running times are comparable to those of pretrained CNNs models. Our works evaluate the validity of SNNs with unsupervised STDP learning rule in classifying medical images and discuss the advantages of the SNNs over traditional deep neural networks.

II. RELATED WORKS

Since high-quality published skin cancer datasets are limited, using CNNs trained from scratch will make network models prone to overfitting. Therefore, more researchers tend to use pretrained CNNs to classify skin lesions [15]–[26]. Most such works are to pretrain CNNs on ImageNet [53] dataset, and then the weights of CNNs are fine-tuned according to the practical classification problems. The most common CNNs architectures used are VGG16 [18], [25], Resnet [22], [54]–[56], GoogleNet [16], [20], [57], and Alexnet [58]. Haenssle *et al.* [16] made the classification of melanoma and melanocytic nevi using a pretrained Inception v4 model, and they achieved an area under the receiver operating characteristic curve (AUC) of 0.95. Compared with a large international group of 58 dermatologists, their results were better than most dermatologists. Yu *et al.* [17] confirmed that very deep CNNs with effective training can be used to make skin lesions diagnosis, even with limited training data. They achieved an accuracy of 85.5% on the 2016 ISIC dataset, and won the first place at ISIC challenge 2016. Menegola *et al.* [18] used VGG16 networks pretrained on diabetic retinopathy dataset and on ImageNet respectively, and their results showed that the networks pretrained on ImageNet had better performance, and reached an AUC of 80.9% on the 2016 ISIC dataset. Codella *et al.* [19] combined deep learning, sparse coding and SVM for melanoma recognition, and they achieved an accuracy of 93.1% for the melanoma vs all non-melanoma lesions task, and an accuracy of 73.9% for the melanoma vs melanocytic nevi task. Mahbod *et al.* [21] proposed a method combining the classification results of three pretrained CNNs (Alexnet, VGG16 and Resnet18) to discriminate between three lesion classes (malignant melanoma, seborrheic keratosis and benign nevi). This method had been proven to be more effective than a single networks, and obtained an AUC of 0.838 for melanoma classification. In addition, to solve the problems

that handwork can not deal with the huge intra class variation of melanoma and there is high visual similarity between melanoma and non-melanoma, some researchers proposed to perform segmentation first, and then recognize melanoma according to the segmentation results [17], [59]–[62].

Although very deep CNNs or combined deep network models can improve the classification accuracy, the problem of hardware resources consumption will be more serious, which make it more difficult to deploy diagnostic systems on small portable devices. Besides, recent studies have shown that the success of transfer learning depends on the distance between the data used to pretrain CNNs and the data in the actual classification task [63]. Therefore, these CNNs pretrained on natural images may not well deal with the challenge of melanoma recognition [17], [64]. In recent years, SNNs have shown good performance in many image recognition tasks [40], [51], [65]–[69]. And due to their sparse, spiked-based communication framework, they have huge advantages in the computational costs and hardware friendliness [27], [70]–[74]. Kheradpisheh *et al.* [41] proposed a deep SNNs which use STDP learning rule to train multiple convolutional/pooling layers to extract the visual features of images and then send them to the SVM classifier. Their model reached an accuracy of 98.4% on MNIST datasets, and their results showed that the networks performs well on small training datasets. Qiao *et al.* [74] proposed a biologically plausible SNNs model for hardware implementation, which used a hardware-friendly STDP mechanism to achieve unsupervised learning. Their model achieved a recognition accuracy of 94.5% on the MNIST datasets, and significantly reduced hardware costs and power consumption.

Therefore, we try to use convolutional SNNs to make melanoma classification. Our convolutional SNNs use STDP-based unsupervised learning method which can learn features from small training datasets. Moreover, the networks use sparse coding method and only three convolutional layers, which make the networks run faster and consume less memory.

III. MATERIALS AND METHODS

A. MATERIALS

In order to evaluate the performance of our STDP-based convolutional SNNs on the skin lesion classification, we use the data from the International Skin Imaging Collaboration (ISIC) 2018 Challenge, which is an international effort to automatic skin lesion analysis towards melanoma detection, including lesion segmentation, dermoscopic feature extraction and lesion classification. The ISIC datasets are currently the largest public dermoscopic image collection of skin lesions in the world. The dataset we used is from ISIC 2018 challenge Task 3: Disease classification. Melanoma is the most deadly type of skin cancers, and is responsible for an overwhelming majority of skin cancer deaths. So, one of the most practical task for an AI skin cancer diagnostic model is to classify melanoma from non-melanoma. Since melanocytic nevi (one type of benign skin lesions) has high

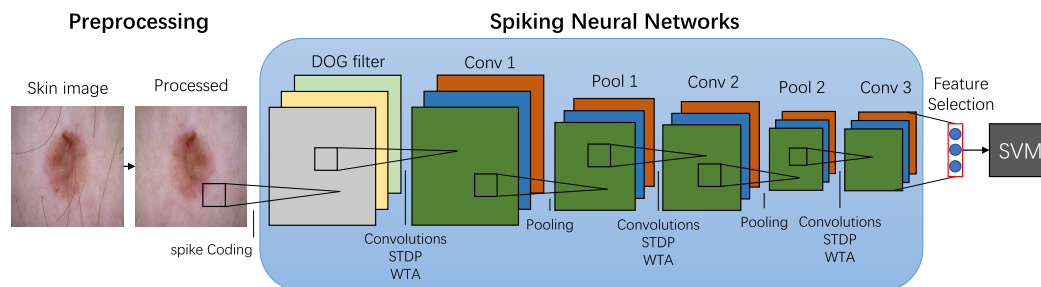


FIGURE 1. The workflow of the skin lesion classification using STDP-based convolutional SNNs.

similarity with melanoma in morphological appearance, and are easily confused with melanoma in clinical diagnosis, we adopt the melanocytic nevi images as the non-melanoma images in this work and try to use our proposed model to distinguish melanoma from melanocytic nevi.

B. OVERVIEW OF THE PROPOSED METHODS

Our method mainly consists of three parts, skin lesion image data preprocessing, feature extraction based on spiking neural networks, and skin lesions classification using SVM classifier as shown in Fig 1. Our SNNs are similar to the networks of Kheradpisheh *et al.* [41], which include a DOG encoding layer, three convolutional layers (Conv1, Conv2 and Conv3), and three pooling layers (Pool1, Pool2 and Pool3). The DOG filter is applied to convert the preprocessed skin images into spikes using the intensity-to-latency coding scheme. Convolutional layers and pooling layers are all consisting of integrate-and-fire (IF) neurons. Each convolutional layer learns features from its input by STDP learning rule, in combination with WTA weight updating strategy and lateral inhibition mechanism. Different to [41], in this work, we replace the final global maximum pooling layer (Pool3) with a feature selection layer in order to improve the classification performance of the SNNs. The outputs of feature selection layer are then used to make the skin lesion classification by SVM classifier.

C. DATA PREPROCESSING

Data preprocessing is used to remove the hairs or noises in the raw images. We use three preprocessing methods which are hair deleting, media filter and global contrast normalization. Many images in the ISIC dataset have a lot of hairs which affects the classification accuracy. In our work hair deleting is implemented by using an algorithm called DullRazor, which was proposed by Lee *et al.* [75] and used to preprocess skin images [76], [77]. The DullRazor can remove hairs effectively, and also additional noises by using Fast Median Filtering. The example results of hair deleting algorithm is shown in Fig 2. Then, a media filter with a window size of 7×7 is used to reduce small pores on the skin and light reflections or shines in the dermoscopic images. Finally, we use global contrast normalization to eliminate the effects of



FIGURE 2. Results of the algorithm of hair deleting.

different light conditions on the pixel values of the images. It works by subtracting the mean of the intensities in the image to each pixel.

D. THE PROPOSED CONVOLUTIONAL SPIKING NEURAL NETWORKS WITH STDP LEARNING RULE

1) INPUT SPIKE ENCODING

Efficient input spike coding can lead to fast and accurate responses of SNNs. A DoG filter is applied to encode each input skin image into discrete spikes. The output of the filter is spikes of the DOG cells which detect contrasts in the input image according to their receptive fields. The higher the contrast, the earlier the cell fires when its activation is above a certain threshold. That is, the firing time of a DoG cell is inversely proportional to its activation value. This intensity-to-latency temporal coding is shown to be effective to detect V1 like oriented edge features as well as complex visual features in higher cortical areas [78]. The output spikes of the DOG filter are grouped into some sequential time steps to be processed in the following convolutional layer. These sequential time steps can be efficiently propagated in parallel on the GPU. Therefore, each input two-dimensional image is encoded into a three-dimensional spiking train with time step as one dimension.

2) CONVOLUTIONAL LAYER

Each convolutional layer in the networks has several feature maps to learn different features determined by their input synaptic weights. Like CNNs, each visual feature obtained in one convolutional layer is a combination of several simpler features extracted from the previous layer. Each neuron receives input spikes from the neurons located in the same

convolutional window of all feature maps of the previous layer. Input synaptic weight sharing is applied to neurons belonging to the same feature map. At each time step, the membrane potential of an IF neuron is updated as follows:

$$V(t) = V(t - 1) + \sum_j W_j S_j(t - 1). \quad (1)$$

where $V(t)$ is the membrane potential of this neuron at time step t , W is the shared input synaptic weight matrix of the corresponding feature map. W_j is the synaptic weight between the j th presynaptic neuron in the previous layer and this neuron. S_j is the spiking train of the j th presynaptic neuron. When its membrane potential exceeds its threshold V_{thr} , the neuron fires a spike, and its value is reset: $V(t) = 0$ and $S(t) = 1$. Lateral inhibition mechanism is applied to the neurons of all convolutional layers. When a neuron belonging to one feature map fires, it inhibits neurons in that same location but belonging to other feature maps to fire. In addition, each neuron is allowed to fire only once. Each spike of a feature map indicates the detection of a particular feature at that location, and the earlier the spike, the more prominent the detected feature.

3) STDP LEARNING RULE

In the networks, the synaptic learning of each convolutional layer is unsupervised and done layer by layer. A simplified STDP rule in combination with a winner-takes-all (WTA) mechanism is used to update the input synaptic weights of neurons in convolutional layers. The modification of one synaptic weight is calculated as follows:

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

where ω_{ij} is the synaptic weight from the j th neuron in the input layer (pre-synaptic neuron) to the i th neuron in one convolutional layer (post-synaptic neuron). t_j and t_i are the firing times of the pre-synaptic neuron and the post-synaptic neuron, respectively. And a^+ and a^- are the two learning rate parameters of STDP. In a convolutional layer, neurons corresponding to the same feature map detect the same feature but at different locations. They compete with each other to update their shared input synaptic weights. The neuron that fires the earliest is the winner which is then modify their shared weights according to STDP rule, and the other neurons of the same feature map are prevented from doing weights updating. Moreover, in order to encourage different feature maps to learn different and prominent features, there is local lateral inhibition between the feature maps of one convolutional layer. That is, if one neuron is allowed to modify its input synaptic weights, it prevents other neurons at the same location and belonging to other feature maps to update their input synaptic weights. These biological mechanisms make the learning event-driven and the information processing sparse and effective in the networks.

4) POOLING LAYER

The first two pooling layers in the networks are local pooling. Each local pooling layer performs a nonlinear max pooling operation over its previous convolutional layer. Such an operation helps the networks to gain invariance and also to reduce the dimensionality of the input. A pooling neuron receives input spikes from the neurons in a pooling window located in the corresponding feature map of its previous convolutional layer. Each pooling neuron is allowed to fire at most once, and its input synaptic weights and threshold are all set to one. Due to the rank-order coding used in the DOG encoding layer, this max pooling operation can simply propagate the first spike in each local pooling window, which corresponds to the most prominent feature.

After the last convolutional layer, there can be a global pooling layer whose outputs are used to classify the input prototypes. This layer performs global maximum pooling over their corresponding feature maps in the last convolutional layer. That is, a global pooling neuron receives input spikes from all the neurons located in the corresponding feature map of its previous convolutional layer. The thresholds of the global pooling neurons are set to infinity. Therefore, the output of each global pooling neuron is the maximum neural membrane potential of its corresponding neural map and it is also the maximum membrane potential value of all the time steps of this neural map. So, there is only one output value for each feature map, which indicates the presence of that feature in the input image. These output values are used to classify the input prototypes by SVM classifier.

E. FEATURE SELECTION

The features extracted by the convolutional layers in the convolutional SNNs are to some extent redundant (including irrelevant features). In [41], Kheradpisheh *et al.* used a global maximum pooling to compress input information and remove the redundancy. After the global pooling, there is only one output value for each feature map, which represents the most prominent feature of this feature map. These output values are then used to classify the input prototypes by SVM classifier. However, when classifying images with very high similarity, it is very likely that the global maximum pooling might filter some diagnostic but not the most prominent features, which will affect the classification result of the classifier. Therefore, to improve classification accuracy we use univariate feature selection based on chi-square test to replace the global maximum pooling in order to select more diagnostic features, while reducing redundancy.

After the learning of all convolutional layers have finished, we use each training sample as the input of the SNNs, then the output of Conv3 (extracted features) is flattened and input to the feature selection layer. The chi-square test is used to measure the relationship coefficient between each extracted feature and input category, and these relationship coefficients are sorted by value. After that, an optimal feature percentage (counting from the feature with the biggest relationship

coefficient in the ranking) can be determined according to the best classification performance of SVM classifier. This optimal feature percentage is later used for feature selection in the test process. The single variable feature selection is implemented with SelectPercentile method in the scikit-learn toolbox [79].

IV. EXPERIMENTAL STUDY AND RESULTS

In this section, we evaluate the performances of our STDP-based convolutional SNNs with and without feature selection on the melanoma (MEL) vs melanocytic nevi (NV) dataset. We compare the melanoma classification results of our networks to those of art-of-date CNNs which had been used in the melanoma skin lesions classification before. And we analyze the effectiveness of feature selection method on improving the melanoma classification performance of the STDP-based convolutional SNNs.

A. DATASET

In our experiments, our task is to discriminate between melanoma and melanocytic nevi. Melanocytic nevi are non-malignant lesions characterized by atypical melanocytic hyperplasia in a lentiginous epidermal pattern, which are visually similar to melanoma and are very difficult for experts to distinguish. The ISIC 2018 dataset includes 1113 images of MEL and 6705 images of NV. After removing the images with serious noise, we use 1081 images of MEL and 6638 images of NV in the task. Considering the data of these two categories is imbalanced, we randomly select 1081 samples from the 6638 NV images, then we use four-fold cross validation on the 1081 images of the two categories. The samplings of NV images are performed 5 times (20 experiments in total).

B. EXPERIMENTAL SETTINGS

Our networks are implemented with SpykeTorch [80], an open-source high-speed simulation framework based on PyTorch. The computations of the framework are based on tensor, which is totally done by PyTorch functions, so that the networks can run efficiently on the GPU platform. CNNs used for comparison study are also implemented based on PyTorch. All these models use the same data preprocessing method as mentioned in this paper, and run on a computer server with an NVIDIA GeForce RTX 2080TI GPU (11GB GPU memory) and an Intel Core i7-8700 3.20GHz CPU.

In our experiments, we use a DOG filter with a size of 5×5 in the encoding layer of our networks. Each skin image is encoded into 15 spiking time packets (time steps) for processing. The first, second and third convolutional layers consist of 6, 20 and 20 feature maps with the kernel sizes of 5×5 , 17×17 and 5×5 , respectively. And their IF neurons have thresholds of 4, 28 and 6, respectively. The pooling window sizes of the first and second pooling layers are 7 and 5, and the strides are 6 and 5, respectively. The learning rates of all the three convolutional layers are set to $a^+ = 0.07$ and $a^- = -0.01$. The thresholds of the IF neurons, the sizes of the feature maps, and the learning rates a^+ and a^- are

determined according to the visualization of the learned features and the classification results of the SVM classifier. For STDP-based convolutional SNNs with feature selection, the optimal feature percentage used in the test is 8% determined in the feature selection phase. The SVM classifier uses a liner kernel and the penalty parameter $C=2$ (optimized by the grid search).

C. EVALUATION METRICS

The output values of the global maximum pool or that of feature selection are used to train the linear SVM classifier to classify input prototypes. In the skin lesion diagnosis, it is more important to correctly predict melanoma lesion which has high mortality than to incorrectly predict benign melanocytic nevi lesion [55], therefore in our experiments the performance of the SVM classifier is evaluated in terms of accuracy (AC), precision (PREC), sensitivity (SE), specificity (SP) and AUC to measure the classification performance of the networks. The criteria are defined as:

$$AC = \frac{N_{tp} + N_{tn}}{N_{tp} + N_{fp} + N_{fn} + N_{tn}}, \quad (3)$$

$$SE = \frac{N_{tp}}{N_{tp} + N_{fn}}, \quad (4)$$

$$SP = \frac{N_{tn}}{N_{tn} + N_{fp}}, \quad (5)$$

$$PREC = \frac{N_{tp}}{N_{tp} + N_{fp}}, \quad (6)$$

where N_{tp} is the number of true positives, N_{tn} is the number of true negatives, N_{fn} is the number of false negatives, and N_{fp} is the number of false positives.

D. MAIN RESULTS

1) QUALITATIVE EVALUATION OF FEATURES EXTRACTED BY STDP-BASED CONVOLUTIONAL SNNs

The features of input skin lesions images are learned in the first, second and third convolutional layers of the STDP-based spiking neural networks. As shown in Fig 3, the neuronal maps of the first convolutional layer work as filters to detect the edge features with different orientations over the images. The second convolutional layer learns visual features which are the combination of oriented edges. The third convolutional layer learns the more complex features which are related to category-specific prototypes. In all the learning processes of the three convolutional layers, STDP tends to learn those frequent and salient features, and the lateral inhibition mechanism helps the spiking neurons to learn different features.

2) QUANTITATIVE EVALUATION AND COMPARISON

The features extracted by the convolutional layers are then filtered by global maximum pooling layer or feature selection, and the final filtered outputs are used to train SVM classifiers to distinguish melanoma from melanocytic nevi. We compare the melanoma classification results of our proposed networks

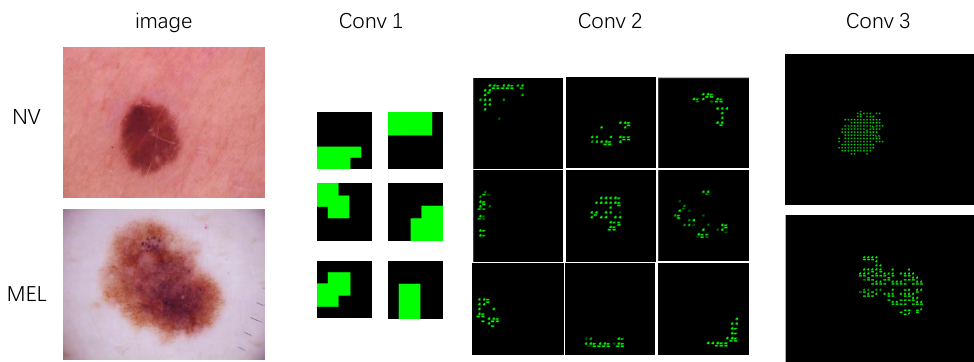


FIGURE 3. The visualization of features learned by different convolutional layers on the MEL/NV dataset.

TABLE 1. Classification performances of our SNNs and some other traditional CNNs on the MEL/NV dataset.

Model	AUC	AC	PREC	SE	SP	Running Time (min)	Number of weights
STDP-based convolutional SNNs	0.907	0.838	0.8	0.875	0.801	25	44,830
STDP-based convolutional SNNs + Feature selection	0.936	0.877	0.846	0.903	0.847	32	44,830
Resnet152	0.846	0.813	0.793	0.847	0.779	739	62,647,106
Pretrained Resnet152	0.876	0.816	0.808	0.872	0.794	21	25,909,250
Pretrained Resnet50	0.746	0.773	0.756	0.805	0.741	19	26,285,058
Pretrained Inception v3	0.885	0.854	0.83	0.89	0.817	25	12,657,475
Pretrained Alexnet	0.878	0.850	0.828	0.883	0.816	26	54,542,338
Pretrained VGG16	0.651	0.711	0.726	0.678	0.743	32	119,554,050
CNNs-three	0.54	0.577	0.584	0.535	0.619	7	44876

with other traditional CNNs on the MEL/NV dataset. These traditional CNNs include VGG16, Alexnet, Inception v3, Resnet50, Resnet152 which are all pretrained on ImageNet, and Resnet152 trained from scratch. All these CNNs are from the latest studies in the melanoma detection literatures. The pretrained CNNs are implemented by transferring the weights trained on ImageNet and then fine-tuning all the last fully connected layers using our MEL/NV dataset. In addition, we also train a CNNs with the same structure as our SNNs from scratch, which have the same number of convolutional and pooling layers (CNNs-three). All these networks use the same data preprocessing method as mentioned in this paper. The classification results of our SNNs and other traditional CNNs are listed in Table 1. Each result value in the table represents mean value from five sets of experiments using four-fold cross validation.

It can be seen from Table 1 that on the melanoma vs melanocytic nevi classification task, our SNNs without feature selection reach an average AUC of 0.907, an average AC of 0.838, an average PREC of 0.8, an average SE of 0.875, and an average SP of 0.801, and its AUC value is higher than all the 6 traditional CNNs. And after using feature selection our SNNs achieve the best AUC (0.936), AC (0.877), PREC (0.846), SE (0.903) and SP (0.847) among the presented networks. Moreover, our networks show the highest sensitivity (SE), which means that melanoma lesions are much easier to be correctly identified by using our SNNs, which is very important in the actual diagnosis of skin lesions. Therefore, compared with these listed traditional CNNs, our STDP-based convolutional SNNs provide better

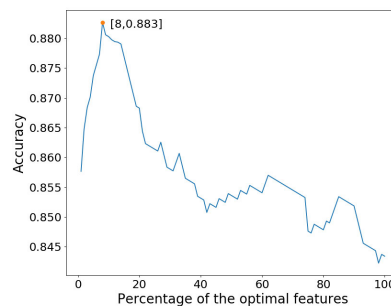


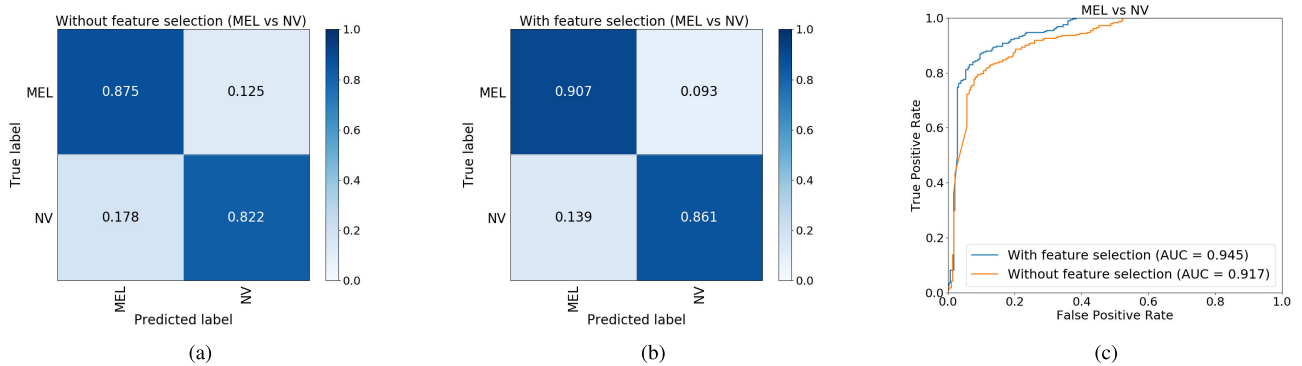
FIGURE 4. The accuracies of melanoma recognition when using different percentages of optimal features.

classification performance on the melanoma vs melanocytic nevi classification task.

Moreover, to compare the hardware resources consumption of these networks for the melanoma vs melanocytic nevi classification task, the running times and the weight numbers of these networks are measured and listed in Table 1. Running time and weight numbers can represent the computational and memory resource requirements of a network model during training [81]. It can be seen from Table 1 that CNNs-three have the least running time, but their classification performance is very poor, so shallow CNNs can not be used in this classification task. With the same network structure, CNNs-three run faster than our SNNs, this is because our SNNs have an additional time dimension, and therefore more data needs to be processed. By comparing the running times of Resnet152 and pretrained Resnet152, it can be seen that existing deep CNNs, if not pretrained, will consume dozens

TABLE 2. Classification performances of our SNNs and some other traditional CNNs on the remaining 5557 NV test images.

Model	AUC	AC	PREC	SE	SP
STDP-based convolutional SNNs	0.902	0.799	0.178	0.874	0.795
STDP-based convolutional SNNs + Feature selection	0.929	0.887	0.224	0.903	0.841
Resnet152	0.828	0.762	0.15	0.847	0.758
Pretrained Resnet152	0.862	0.785	0.168	0.873	0.781
Pretrained Resnet50	0.731	0.73	0.13	0.805	0.726
Pretrained Inception v3	0.861	0.798	0.179	0.89	0.794
Pretrained Alexnet	0.848	0.789	0.171	0.883	0.784
Pretrained VGG16	0.633	0.698	0.102	0.676	0.699
CNNs-three	0.526	0.567	0.059	0.534	0.568

**FIGURE 5.** Classification performance comparisons of our STDP-based convolutional SNNs with and without feature selection on the MEL/NV dataset. (a) Confusion matrix of our SNNs without feature selection. (b) Confusion matrix of our SNNs with feature selection. (c) ROC curves of our SNNs with and without feature selection when MEL is the positive sample.

of times more training time and more hardware resources. Therefore, pretrained CNNs are often used to reduce the consumption of computing resources. However, as shown in Table 1, we find that the five state-of-art pretrained CNNs have very different classification performance. Some work well in melanoma classification but some do not. Therefore, it proves that improper pretraining is likely to weaken the model's learning capability, so it is usually needed to try many pretrained models to find the most appropriate one, which is also time-consuming. The proposed STDP-based convolutional SNNs is generally a good choice. On one hand, as shown in Table 1, comparing to non-pretrained CNNs models (i.e., CNNs that need to be trained from scratch), our SNNs (with and without feature selection) not only achieve much better classification accuracies but also have much better runtime efficiency (tens of times better); on the other hand, although the pretrained CNNs models (i.e., the fine-tuned CNNs) can achieve similar running time, our proposed SNNs are more stable and easier to use than the pretrained CNNs because we do not need to try many pretrained models any more, and SNNs also have much better classification accuracies than the pretrained CNNs. In addition, our networks have only three convolutional layers, and there are no fully connected layers. The complexity of the model and the parameters that need to be trained in the network are greatly reduced, which is very beneficial for building a neuromorphic computing platform or deploying to small portable devices.

In order to further show the good performance of our proposed SNNs, we use the remaining 5557 NV images after random sampling and 281 MEL images from the four-fold

cross validation as the test sets, and make the MEL/NV classification using the models listed in Table 1. It should be noted that in the case of unbalanced datasets, the AC value and PREC value will be unreliable, but the AUC value can still measure the performance of the classifier well. The classification results are listed in Table 2. It can be seen that on the remaining 5557 NV images, our SNNs using feature selection still show the best AUC value, SE value and SP value, which shows the superiority of our model in the MEL/NV classification over other CNNs. And for each listed network model, the AUC and SP values decrease slightly, but there is no much difference between each of these three classification values in Table 1 and Table 2, which proves that the sampling method we used in the experiments is feasible.

E. ABLATION STUDY: EFFECT OF USING FEATURE SELECTION

In this section we analyze why our SNNs perform better when using feature selection. In our SNNs without feature selection, the neurons in the last pooling layer perform a global maximum pooling over their corresponding neuronal maps of the last convolutional layer. Only one output value is taken for each feature to represent the presence of this feature in the input image. But when the networks recognize images with high similarity, those rare but diagnostic information is easily filtered by the global maximum pooling operation, which affects the final classification performance of the SVM classifier. Therefore, we use the univariate feature selection algorithm to replace the global maximum pooling. Using

feature selection can retain enough optimal features while removing redundant features. Figure 4 shows the melanoma classification accuracies of our SNNs when using different percentages of optimal features during the testing phase. It can be seen that when 8% of the optimal features are selected and used in the classification, we obtain the best accuracy of 0.883 (the highest point in the figure) for the melanoma vs melanocytic nevi classification. Figure 5 shows the performance comparison of our SNNs with and without feature selection in a single experiment. Figure 5 (a) and (b) are the confusion matrices of our SNNs without and with feature selection, Figure 5 (c) shows the ROC curves of our SNNs without and with feature selection. We can see that on the MEL/NV dataset, the recognition rate of MEL increases from 0.875 (without feature selection) to 0.907 (with feature selection), the recognition rate of NV increases from 0.822 (without feature selection) to 0.861 (with feature selection), and the AUC increases from 0.917 (without feature selection) to 0.945 (with feature selection). It is obvious that the use of feature selection improves the classification performance of our SNNs on the MEL/NV dataset.

V. CONCLUSION

Although deep CNNs have achieved remarkable success in medical image analysis tasks in recent years, their expensive computational costs limit the application of CNNs in the real-time clinical practice. Besides, since labeled skin cancer images are limited, using CNNs trained from scratch will make networks prone to overfitting, and the CNNs pretrained on natural images are not very suitable for skin lesion image analysis. In this paper, we use a STDP-based convolutional SNNs to distinguish melanoma from melanocytic nevi, which is the first work that applying SNNs on medical image classification tasks. The classification results show that efficient temporal coding, event driven learning rule and winner-take-all (WTA) mechanism together ensure sparse spike coding and efficient learning of our networks which achieve an average accuracy of 83.8%. We also find that after the learning of all convolutional layers have finished, using feature selection can effectively help pick out diagnostic information from the features extracted by SNNs, and can improve the classification performance of our SNNs to an average accuracy of 87.7%. In addition, compared with listed traditional CNNs, our SNNs using feature selection show the best classification accuracy and AUC value, and more importantly, our networks have significantly lower computational costs than CNNs that need to be trained from scratch. So, our SNNs are very friendly to smartphones and portable devices with low computational power (CPUs or low-end GPUs) limitations [82]. And also our SNNs are more stable and easier to use than the pretrained CNNs. This is very helpful to facilitate the implementation of automated skin lesion classifiers and makes the access convenient and cheap. Moreover, our networks can be easily extended to other medical image analysis fields.

In the future, we will try to improve the spike coding in the input layer of the SNNs in order to better encoding

some visual information such as texture and color which the DOG filter may not deal well with [41]. We will also try to further improve our network architectures by adding feedback connections, which has been proven to play an important role in the visual pathway [83]. In addition, our future work will include multi-classification of skin cancer lesions with high performance, which is still challenging to STDP-based SNNs.

REFERENCES

- [1] A. K. Gupta, M. Bharadwaj, and R. Mehrotra, "Skin cancer concerns in people of color: Risk factors and prevention," *Asian Pacific J. Cancer Prevention*, vol. 17, no. 12, p. 5257, 2016.
- [2] R. S. Stern, "Prevalence of a history of skin cancer in 2007: Results of an incidence-based model," *Arch. Dermatol.*, vol. 146, no. 3, pp. 279–282, 2010.
- [3] J. Premaladha and K. S. Ravichandran, "Novel approaches for diagnosing melanoma skin lesions through supervised and deep learning algorithms," *J. Med. Syst.*, vol. 40, no. 4, p. 96, Apr. 2016.
- [4] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction," *Comput. Struct. Biotechnol. J.*, vol. 13, pp. 8–17, 2015.
- [5] J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and prognosis," *Cancer Inform.*, vol. 2, pp. 59–77, Jan. 2006.
- [6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [7] T. J. Brinker, A. Hekler, J. S. Utikal, N. Grabe, D. Schadendorf, J. Klode, C. Berking, T. Steeb, A. H. Enk, and C. von Kalle, "Skin cancer classification using convolutional neural networks: Systematic review," *J. Med. Internet Res.*, vol. 20, no. 10, Oct. 2018, Art. no. e11936.
- [8] Y. Fujisawa, S. Inoue, and Y. Nakamura, "The possibility of deep learning-based, computer-aided skin tumor classifiers," *Frontiers Med.*, vol. 6, p. 191, Aug. 2019.
- [9] M. Q. Khan, A. Hussain, S. U. Rehman, U. Khan, M. Maqsood, K. Mehmood, and M. A. Khan, "Classification of melanoma and nevus in digital images for diagnosis of skin cancer," *IEEE Access*, vol. 7, pp. 90132–90144, 2019.
- [10] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, 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.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [12] B. Kong, S. Sun, X. Wang, Q. Song, and S. Zhang, "Invasive cancer detection utilizing compressed convolutional neural network and transfer learning," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2018, pp. 156–164.
- [13] S. Wu, Z. Gao, Z. Liu, J. Luo, H. Zhang, and S. Li, "Direct reconstruction of ultrasound elastography using an end-to-end deep neural network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2018, pp. 374–382.
- [14] B. Kong, Y. Zhan, M. Shin, T. Denny, and S. Zhang, "Recognizing end-diastole and end-systole frames via deep temporal regression network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2016, pp. 264–272.
- [15] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors*, vol. 18, no. 2, p. 556, Feb. 2018.
- [16] H. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kallou, A. B. H. Hassen, L. Thomas, A. Enk, and L. Uhlmann, "Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Ann. Oncol.*, vol. 29, no. 8, pp. 1836–1842, 2018.
- [17] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Trans. Med. Imag.*, vol. 36, no. 4, pp. 994–1004, Apr. 2017.
- [18] 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.

- [19] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, and J. R. Smith, "Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images," in *Proc. Int. Workshop Mach. Learn. Med. Imag.* Berlin, Germany: Springer, 2015, pp. 118–126.
- [20] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115, 2017.
- [21] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Elling, "Skin lesion classification using hybrid deep neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1229–1233.
- [22] S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, "Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm," *J. Investigative Dermatol.*, vol. 138, no. 7, pp. 1529–1538, Jul. 2018.
- [23] L. Bi, J. Kim, E. Ahn, and D. Feng, "Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks," 2017, *arXiv:1703.04197*. [Online]. Available: <http://arxiv.org/abs/1703.04197>
- [24] J. Kawahara and G. Hamameh, "Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers," in *Proc. Int. Workshop Mach. Learn. Med. Imag.* Berlin, Germany: Springer, 2016, pp. 164–171.
- [25] X. Sun, J. Yang, M. Sun, and K. Wang, "A benchmark for automatic visual classification of clinical skin disease images," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2016, pp. 206–222.
- [26] A. Romero-Lopez, X. Giro-i-Nieto, J. Burdick, and O. Marques, "Skin lesion classification from dermoscopic images using deep learning techniques," in *Proc. 13th IASTED Int. Conf. Biomed. Eng. (BioMed)*, Feb. 2017, pp. 49–54.
- [27] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura, B. Brezzo, I. Vo, S. K. Esser, R. Appuswamy, B. Taba, A. Amir, M. D. Flickner, W. P. Risk, R. Manohar, and D. S. Modha, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, Aug. 2014.
- [28] M. Davies et al., "Loihi: A neuromorphic manycore processor with on-chip learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan. 2018.
- [29] W. Gerstner, R. Kempter, J. L. Van Hemmen, and H. Wagner, "A neuronal learning rule for sub-millisecond temporal coding," *Nature*, vol. 383, no. 6595, p. 76, 1996.
- [30] H. Markram, "Regulation of synaptic efficacy by coincidence of post-synaptic APs and EPSPs," *Science*, vol. 275, no. 5297, pp. 213–215, Jan. 1997.
- [31] G.-Q. Bi and M.-M. Poo, "Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type," *J. Neurosci.*, vol. 18, no. 24, pp. 10464–10472, 1998.
- [32] P. J. Sjöström, G. G. Turrigiano, and S. B. Nelson, "Rate, timing, and cooperativity jointly determine cortical synaptic plasticity," *Neuron*, vol. 32, no. 6, pp. 1149–1164, Dec. 2001.
- [33] C. D. Meliza and Y. Dan, "Receptive-field modification in rat visual cortex induced by paired visual stimulation and single-cell spiking," *Neuron*, vol. 49, no. 2, pp. 183–189, Jan. 2006.
- [34] S. Huang, C. Rozas, M. Trevino, J. Contreras, S. Yang, L. Song, T. Yoshioka, H.-K. Lee, and A. Kirkwood, "Associative hebbian synaptic plasticity in primate visual cortex," *J. Neurosci.*, vol. 34, no. 22, pp. 7575–7579, May 2014.
- [35] Y. Guo, W. Zhang, X. Chen, J. Fu, W. Cheng, D. Song, X. Qu, Z. Yang, and K. Zhao, "Timing-dependent LTP and LTD in mouse primary visual cortex following different visual deprivation models," *PLoS ONE*, vol. 12, no. 5, May 2017, Art. no. e0176603.
- [36] 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.
- [37] T. Masquelier and S. R. Kheradpisheh, "Optimal localist and distributed coding of spatiotemporal spike patterns through STDP and coincidence detection," *Frontiers Comput. Neurosci.*, vol. 12, p. 74, Sep. 2018.
- [38] N. Kasabov, K. Dhoble, N. Nuntalid, and G. Indiveri, "Dynamic evolving spiking neural networks for on-line spatio-and spectro-temporal pattern recognition," *Neural Netw.*, vol. 41, pp. 188–201, May 2013.
- [39] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, and A. Maida, "Deep learning in spiking neural networks," *Neural Netw.*, vol. 111, pp. 47–63, Mar. 2019.
- [40] C. Lee, P. Panda, G. Srinivasan, and K. Roy, "Training deep spiking convolutional neural networks with STDP-based unsupervised pre-training followed by supervised fine-tuning," *Frontiers Neurosci.*, vol. 12, p. 435, Aug. 2018.
- [41] 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.
- [42] C. Lee, G. Srinivasan, P. Panda, and K. Roy, "Deep spiking convolutional neural network trained with unsupervised spike-timing-dependent plasticity," *IEEE Trans. Cogn. Develop. Syst.*, vol. 11, no. 3, pp. 384–394, Sep. 2019.
- [43] J. C. Thiele, O. Bichler, and A. Dupret, "Event-based, timescale invariant unsupervised online deep learning with STDP," *Frontiers Comput. Neurosci.*, vol. 12, p. 46, Jun. 2018.
- [44] M.-J. Escobar, G. S. Masson, T. Vieville, and P. Kornprobst, "Action recognition using a bio-inspired feedforward spiking network," *Int. J. Comput. Vis.*, vol. 82, no. 3, p. 284, 2009.
- [45] D. Liu and S. Yue, "Fast unsupervised learning for visual pattern recognition using spike timing dependent plasticity," *Neurocomputing*, vol. 249, pp. 212–224, Aug. 2017.
- [46] A. Tavanaei and A. Maida, "Bio-inspired multi-layer spiking neural network extracts discriminative features from speech signals," in *Proc. Int. Conf. Neural Inf. Process.* Berlin, Germany: Springer, 2017, pp. 899–908.
- [47] M. Dong, X. Huang, and B. Xu, "Unsupervised speech recognition through spike-timing-dependent plasticity in a convolutional spiking neural network," *PLoS ONE*, vol. 13, no. 11, Nov. 2018, Art. no. e0204596.
- [48] 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.
- [49] 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.
- [50] J. Tian, "Deep learning in medical image analysis and its challenges," *ACTA Automatica Sinica*, vol. 44, no. 3, pp. 401–424, 2018.
- [51] P. Panda and K. Roy, "Unsupervised regenerative learning of hierarchical features in spiking deep networks for object recognition," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 299–306.
- [52] T. Masquelier and S. J. Thorpe, "Unsupervised learning of visual features through spike timing dependent plasticity," *PLoS Comput. Biol.*, vol. 3, no. 2, pp. 247–257, 2007.
- [53] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [54] F. Cicero, A. Oliveira, G. Botelho, and C. D. C. da Computacao, "Deep learning and convolutional neural networks in the aid of the classification of melanoma," in *Proc. SIBGRABI*, 2016, pp. 1–4.
- [55] J. Yap, W. Yolland, and P. Tschandl, "Multimodal skin lesion classification using deep learning," *Experim. Dermatol.*, vol. 27, no. 11, pp. 1261–1267, Nov. 2018.
- [56] J. Yang, F. Xie, H. Fan, Z. Jiang, and J. Liu, "Classification for dermoscopy images using convolutional neural networks based on region average pooling," *IEEE Access*, vol. 6, pp. 65130–65138, 2018.
- [57] 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.
- [58] T. Majtner, S. Yildirim-Yayilgan, and J. Y. Hardeberg, "Combining deep learning and hand-crafted features for skin lesion classification," in *Proc. 6th Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Dec. 2016, pp. 1–6.
- [59] T. Tommasi, E. La Torre, and B. Caputo, "Melanoma recognition using representative and discriminative kernel classifiers," in *Proc. Int. Workshop Comput. Vis. Approaches Med. Image Anal.* Berlin, Germany: Springer, 2006, pp. 1–12.
- [60] M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, and R. H. Moss, "A methodological approach to the classification of dermoscopy images," *Comput. Med. Imag. Graph.*, vol. 31, no. 6, pp. 362–373, Sep. 2007.
- [61] G. Schaefer, B. Krawczyk, M. E. Celebi, and H. Iyatomi, "An ensemble classification approach for melanoma diagnosis," *Memetic Comput.*, vol. 6, no. 4, pp. 233–240, Dec. 2014.
- [62] A. A. Adegun and S. Viriri, "Deep learning-based system for automatic melanoma detection," *IEEE Access*, vol. 8, pp. 7160–7172, 2020.
- [63] H. Azizpour, A. S. Razavian, J. Sullivan, A. Maki, and S. Carlsson, "From generic to specific deep representations for visual recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2015, pp. 36–45.

- [64] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
- [65] T. Liu, Z. Liu, F. Lin, Y. Jin, G. Quan, and W. Wen, "MT-spike: A multilayer time-based spiking neuromorphic architecture with temporal error backpropagation," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2017, pp. 450–457.
- [66] S. K. Esser, R. Appuswamy, P. Merolla, J. V. Arthur, and D. S. Modha, "Backpropagation for energy-efficient neuromorphic computing," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1117–1125.
- [67] B. Rueckauer, I.-A. Lungu, Y. Hu, M. Pfeiffer, and S.-C. Liu, "Conversion of continuous-valued deep networks to efficient event-driven networks for image classification," *Frontiers Neurosci.*, vol. 11, p. 682, Dec. 2017.
- [68] S. K. Essera, "Convolutional networks for fast energy-efficient neuromorphic computing," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 41, pp. 11441–11446, 2016.
- [69] E. Stomatias, D. Neil, F. Galluppi, M. Pfeiffer, S.-C. Liu, and S. Furber, "Scalable energy-efficient, low-latency implementations of trained spiking deep belief networks on SpiNNaker," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2015, pp. 1–8.
- [70] Z. Mainen and T. Sejnowski, "Reliability of spike timing in neocortical neurons," *Science*, vol. 268, no. 5216, pp. 1503–1506, Jun. 1995.
- [71] W. Bair and C. Koch, "Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey," *Neural Comput.*, vol. 8, no. 6, pp. 1185–1202, Aug. 1996.
- [72] R. Herikstad, J. Baker, J.-P. Lachaux, C. M. Gray, and S.-C. Yen, "Natural movies evoke spike trains with low spike time variability in cat primary visual cortex," *J. Neurosci.*, vol. 31, no. 44, pp. 15844–15860, 2011.
- [73] J.-S. Seo, B. Brezzo, Y. Liu, B. D. Parker, S. K. Esser, R. K. Montoye, B. Rajendran, J. A. Tierno, L. Chang, D. S. Modha, and D. J. Friedmann, "A 45nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Sep. 2011, pp. 1–4.
- [74] G. Qiao, S. G. Hu, J. J. Wang, C. M. Zhang, T. P. Chen, N. Ning, Q. Yu, and Y. Liu, "A neuromorphic-hardware oriented bio-plausible online-learning spiking neural network model," *IEEE Access*, vol. 7, pp. 71730–71740, 2019.
- [75] 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, 1997.
- [76] J. A. Jaleel, S. Salim, and R. B. Aswin, "Computer aided detection of skin cancer," in *Proc. Int. Conf. Circuits, Power Comput. Technol. (ICCPCT)*, Mar. 2013, pp. 1137–1142.
- [77] D. Ruiz, V. Berenguer, A. Soriano, and B. SaNchez, "A decision support system for the diagnosis of melanoma: A comparative approach," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 15217–15223, 2011.
- [78] G. Portelli, J. M. Barrett, G. Hilgen, T. Masquelier, A. Maccione, S. Di Marco, L. Berdondini, P. Kornprobst, and E. Sernagor, "Rank order coding: A retinal information decoding strategy revealed by large-scale multielectrode array retinal recordings," *Eneuro*, vol. 3, no. 3, pp. 1–18, 2016.
- [79] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, and M. Blondel, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
- [80] M. Mozafari, M. Ganjtabesh, A. Nowzari-Dalini, and T. Masquelier, "SpykeTorch: Efficient simulation of convolutional spiking neural networks with at most one spike per neuron," 2019, *arXiv:1903.02440*. [Online]. Available: <http://arxiv.org/abs/1903.02440>
- [81] G. Antipov, S.-A. Berrani, and J.-L. Dugelay, "Minimalistic CNN-based ensemble model for gender prediction from face images," *Pattern Recognit. Lett.*, vol. 70, pp. 59–65, Jan. 2016.
- [82] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5353–5360.
- [83] J. Tang and N. Suga, "Modulation of auditory processing by cortico-cortical feed-forward and feedback projections," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 21, pp. 7600–7605, May 2008.



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, China. Her current research interests include biological neural networks, spiking neural networks, STDP plasticity, and deep learning.



YAN SHI received the B.S. degree in biomedical engineering from the Hebei University of Technology, Tianjin, China, where he is currently pursuing the M.S. degree in biomedical engineering. His current research interests include spiking neural networks and deep learning.



ZHENGHUA XU received the M.Phil. degree in computer science from the University of Melbourne, Australia, in 2012, and the D.Phil. degree in computer science from the University of Oxford, U.K., in 2018.

From 2017 to 2018, he worked as a research associate at the Department of Computer Science, University of Oxford. He is currently a Professor with the Hebei University of Technology, China, and a awardee of "100 Talents Plan" of Hebei Province. He has published dozens of articles in top AI or database conferences, such as, AAAI, IJCAI, ICDE, EDBT, CIKM, and so on. His research focuses on topics within artificial intelligence and data mining, especially deep learning, medical artificial intelligence, health data mining, and computing vision.



RUOWEI QU received the M.S. degree from the School of Science, University of Science and Technology Beijing, China, in 2012. She is currently pursuing the Ph.D. degree with the School of Electrical Engineering, Hebei University of Technology, Tianjin, China. Her research interests include medical image processing, pattern recognition, and machine learning.



GUIZHI XU received the Ph.D. degree from the School of Electrical Engineering, Hebei University of Technology, Tianjin, China, in 2002. She is currently a Professor and a Ph.D. supervisor with the Hebei University of Technology. She has acted as the Principal Investigator (PI) of tens of national and provincial projects, including Natural Science Foundation of China. She has published more than 90 SCI or EI indexed academic papers and three books. She was awarded the Outstanding Contribution Award of Science and Technology in Hebei Province, the Second Prize of the Natural Science of Hebei Province, the Hebei Science and Technology Progress, and the Hebei Province Excellent Teaching Achievement, and the Third Prize of the Natural Science of Hebei Province.

• • •