IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 16, 2019, accepted June 26, 2019, date of publication July 8, 2019, date of current version July 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2927357

Slim-ResCNN: A Deep Residual Convolutional Neural Network for Fingerprint Liveness Detection

YONGLIANG ZHANG[1], **DAQIONG SHI**[1], **XIAOSI ZHAN**[3], **DI CAO**[1], **KEYI ZHU**[4], **AND ZHIWEI LI**[2]
[1]College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China
[2]Hangzhou Jing Lianwen Technology Company Ltd., Hangzhou 310014, China
[3]School of Science and Technology, Zhejiang International Studies University, Hangzhou 310023, China
[4]Glasgow College, University of Electronic Science and Technology, Chengdu 611731, China

Corresponding author: Yongliang Zhang (titanzhang@zjut.edu.cn)

This work was supported in part by the Public Welfare Technology Research Program of Zhejiang Province under Grant LGF18F030008 and in part by the Hangzhou Jinglianwen Technology Company Ltd.

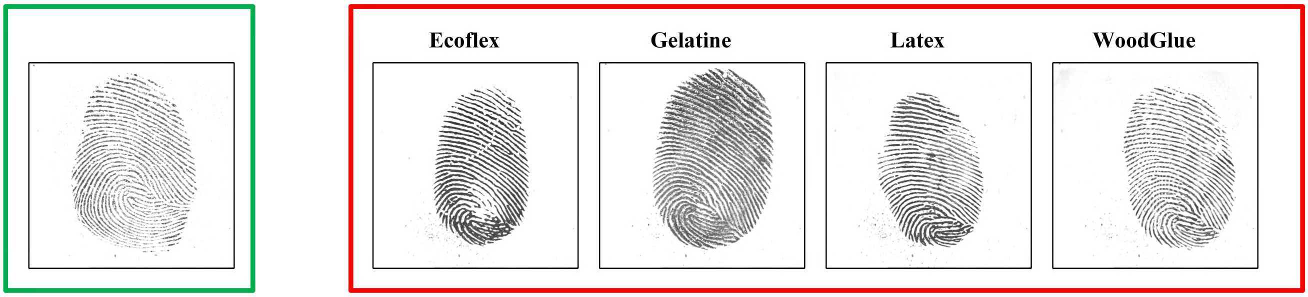
ABSTRACT Fingerprint liveness detection has gradually been regarded as a primary countermeasure for protecting the fingerprint recognition systems from spoof presentation attacks. The convolutional neural networks (CNNs) have shown impressive performance and great potential in advancing the state-of-the-art of fingerprint liveness detection. However, most existing CNNs-based fingerprint liveness methods have a few shortcomings: 1) the CNN structure used on natural images does not achieve good performance on fingerprint liveness detection, which neglects the inevitable differences between natural images and fingerprint images; or 2) a relative shallow architecture (typically several layers) has not paid attention to the capability of deep network for spoof fingerprint detection. Motivated by the compelling classification accuracy and desirable convergence behaviors of the deep residual network, this paper proposes a new CNN-based fingerprint liveness detection framework to discriminate between live fingerprints and fake ones. The proposed framework is a lightweight yet powerful network structure, called Slim-ResCNN, which consists of the stack of series of improved residual blocks. The improved residual blocks are specifically designed for fingerprint liveness detection without overfitting and less processing time. The proposed approach significantly improves the performance of fingerprint liveness detection on LivDet2013 and LivDet2015 datasets. Additionally, the Slim-ResCNN wins the first prize in the Fingerprint Liveness Detection Competition 2017, with an overall accuracy of 95.25%.

INDEX TERMS Fingerprint spoofing, presentation attacks, fingerprint liveness detection, center of gravity, Slim-ResCNN.

I. INTRODUCTION

Recently, the fingerprint recognition technology is extensively employed in border control applications and personal identification verification systems, owing to its high reliability, high generalization, and low cost. Meanwhile, the fingerprint recognition technology has shown some weaknesses related to the problem of security as the widespread use of personal verification systems based on fingerprint. Especially, the growing trend of mobile devices which use fingerprint for unlocking and payment has generated a challenging problem: spoofing attacks. Fingerprint-based authentication systems are directly attacked by these spoof attacks through artificial fingerprint replicas. Artificial fingerprint replicas, also called spoofs, or fake fingerprints, make possible to jeopardize the security of the fingerprint recognition systems. In order to enhance the security of these fingerprint recognition systems, fingerprint liveness detection is regarded as a primary countermeasure against fingerprint spoof attacks. Fingerprint liveness detection prevents direct attacks to scanners by analyzing images which are captured from live fingers or fake ones [1]. The fingerprint liveness detection has extremely become urgent to thwart spoof attacks at fingerprint authentication systems, especially

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



(a) Live fingerprint.

(b) Spoof fingerprints.

FIGURE 1. Fingerprint samples taken from the greenbit sensor of LivDet2015 dataset [4]: Live fingerprint and the corresponding spoof fingerprints (of the same finger) fabricated with different materials.

when artificial fingerprints are easily fabricated by commonly available materials, such as latex, gelatin, silicone, and play-doh [2], [3]. Figure 1 shows some samples of live and spoof fingerprints.

The various spoof attack detection approaches have been proposed to assess if the input fingerprint image is from a "live" entity or a "spoof" artefact. The fingerprint liveness detections can be broadly divided into hardware-based and software-based methods [5]. Hardware-based methods exploit additional hardware devices to detect whether the input fingerprint image is coming from alive user or artificial replicas by measuring additional life characteristics of fingerprint, such as temperature, blood pressure, odor, pulse oximetry and so on [6]. Although the additional hardware devices can distinguish between live and fake fingerprints precisely, they also make the fingerprint recognition systems more complex and expensive. Moreover, it is difficult to update these additional hardware devices when the attackers improve artificial replicas with new manufacturing technology and pass the fingerprint recognition systems successfully. Software-based methods, on the other hand, have gained an increasing attention, which uses image processing technology to extract features from the captured fingerprint images so as to identify the live and fake fingerprints without additional hardware devices. Compared with hardware-based methods, software-based methods make the fingerprint recognition systems less cost and more easily to update [7].

Software-based methods exploit dynamic behaviors (e.g., ridge distortion, perspiration) or static characteristics (e.g., textural characteristics, ridge frequencies, elastic properties of the skin) [8] which are extracted from the fingerprint images. The dynamic behaviors are obtained from a sequence of images, which is very time-consuming because the users need to collect fingerprint many times during the fingerprint registration phase. Compared with the dynamic behaviors, the static characteristics are more applicable because only one or a few images are used for fingerprint liveness detection instead of a sequence of images. In summary, the static characteristics-based approaches not only prevent spoofs from attacking fingerprint authentication systems but also don't complex the fingerprint authentication system in practical applications. Texture-based features extraction approaches are the most common in static characteristics-based approaches. Texture characteristics are different in continuity, clarity and ductility between live fingerprints and fake ones, therefore the texture features can be used to compute the liveness of fingerprints. Local binary pattern (LBP) histograms based on gradient, which extracts texture details by binary coding, are firstly used to capture textural details for fingerprint liveness detection [9]. Some modifications inspired by LBP, such as multi-scale local binary pattern [10] and uniform local binary pattern [11], achieve high classification accuracy on some standard databases. The local phase quantization (LPQ) descriptor [12] is acquired by short time Fourier transform (STFT) to discriminate the differences between live samples from fake ones due to the loss of information which may occur during the replica fabrication process. In literature [13], the weber local descriptor (WLD) is utilized to prevent spoof attacks on fingerprint sensors, where the input fingerprint images are represented by extracting two-dimensional histogram features from differential excitation and square bipartite. Gragnaniello *et al.* further [14] propose a new local contrast phase descriptor (LCPD), which combines gradient with local phase information together, achieving a commendable liveness detection accuracy. Inspired by weber local descriptor, Xia *et al.* [15] propose a new local descriptor named Weber local binary descriptor, which consists of the local binary differential excitation component that extracts intensity-variance features and the local binary gradient orientation component that extracts orientation features. The potential of feature fusion approaches is evaluated in the area of fingerprint liveness detection by analyzing different features and different methods for their aggregation, which shows that the feature fusion methods improve the accuracy of those methods based on individual feature [16].

Notably, most of software-based approaches using texture feature usually rely on expert knowledge to engineer handcraft features. However, it is very difficult to find out effective handcraft features to distinguish live fingerprints from fake ones. Moreover, the handcraft features are not easy to generalize due to the poor robustness of new materials and types of sensors [17]. In other words, the fingerprint liveness detection with handcraft features need to be redesigned provided that the attackers improve the technology of fabricating fake fingerprints or the fingerprint sensors change [18].

In terms of software-based methods, many deep learning based techniques have been employed to detect the liveness of an input fingerprint image recently. In contrast to approaches using handcraft features, there are a growing number of researches using deep learning to design robust and interpretable fingerprint liveness detection methods. Convolutional Neural Networks (CNNs), which have been widely used in computer vision, make outstanding performance in image classification [19], object detection [20] and many other tasks [21], attributing to the impressive ability of extracting local features. In literature [22], CNN is firstly introduced to determine whether an input fingerprint image is live or fake. However, it is difficult to optimize the feature extraction and classification simultaneously since they are designed into two separate parts.

CNNs have played a significant role in advancing the development of fingerprint liveness detection, which can get rid of or reduce the dependence on domain knowledge. Many CNNs-based approaches have been proposed for fingerprint liveness detection, such as MobileNet-v1 [17], VGG-19 [23], CaffeNet and GoogLeNet [24]. However, most the existing approaches using CNNs are transferring pre-trained CNN models instead of redesigning a new network structure oriented to fingerprint liveness detection. These approaches use fingerprint images to fine tune the CNN models pre-trained on natural images. The inevitable differences between fingerprint images and natural images make the parameters of pre-trained models on natural images not achieve good performance on the fingerprint liveness detection. Moreover, the fingerprint images are cut up randomly to fit the network's input size, causing partial fingerprint information loss. Most of the existing approaches pour attention into increasing classification performance without considering the speedup of training CNN-based models. Some researches select multiple patches of single fingerprint image for fingerprint liveness detection, which improves the performance at the cost of testing time. Chugh *et al.* [24] propose a CNN-based method, which adopts a voting strategy based on minutiae-centered multiple local patches, showing state-of-the-art average classification accuracy. Obviously, as minutiae's increase, so does running time. It is not a wise choice in practical application, leading to poor user experience.

In order to deal with the problems mentioned above, we propose a new approach for fingerprint liveness detection. Firstly, the foreground region of fingerprint image has been extracted to eliminate the interference of blank areas by the statistical histograms of rows and columns. The proposed approach utilizes the center of gravity [25] to select local patches from the foreground region because of the continuity of pixel distribution of fingerprint ridges and valleys, which further effectively avoids a blank area with no fingerprint information before feeding into the network. Furthermore, the extracted local patches are flipped horizontally and vertically, and rotated at four different angles for data augmentation. Inspired by residual networks (ResNets) [26], [27] showing compelling accuracy and desirable convergence behaviors, slim residual convolutional neural network, called Slim-ResCNN, is specially designed for fingerprint liveness detection. The proposed Slim-ResCNN framework is different from the original residual network in that only nine improved residual blocks are stacked into Slim-CNN and less convolutional kernels are employed, making less training time and improving classification performance for fingerprint spoof detection. The main contributions of this paper are enumerated as follows:

- Different from most existing random selection local block methods, the statistical histogram and center of gravity CoG) are used to remove the blank area and select local patches from the fingerprint image before into the network, which makes full use of image information. The local patches instead of the entire fingerprint sample are sent to the network, which can decrease execution time and adapt to different scale of fingerprint samples.
- We propose a new residual network architecture, called Slim-ResCNN, which is a relative simple and lightweight CNN structure specialized for fingerprint liveness detection. The new architecture consists of several improved residual blocks in which the dropout layer is added to each pair of convolutional kernels of original residual block to prevent overfitting. When the dimensions increase, the padding channel layer replaces the convolutional layer of original residual block by extra zero entries padded, without bringing in extra parameters.
- Experiments demonstrate that the proposed approach provides high classification accuracy for fingerprint liveness detection on LivDet2013 [28], LivDet2015 [4] and LivDet2017 [29] datasets. The excellent performance of Slim-ResCNN model is further confirmed in that the model wins the first place on the Fingerprint Liveness Detection Competition 2017, with an overall accuracy of 95.25%.

The paper is structured as follows. The proposed method is discussed in Section 2. Section 3 presents the experimental results and discussions. Finally, Section 4 concludes the paper and gives some prospective research directions.

II. PROPOSED METHOD

Motivated by the favorable performance of deep residual network, a relative simple and lightweight residual convolutional neural network structure, called Slim-ResCNN, is designed especially for fingerprint liveness detection. In this section, the foreground region is firstly extracted from entire fingerprint image by the statistical histograms of rows

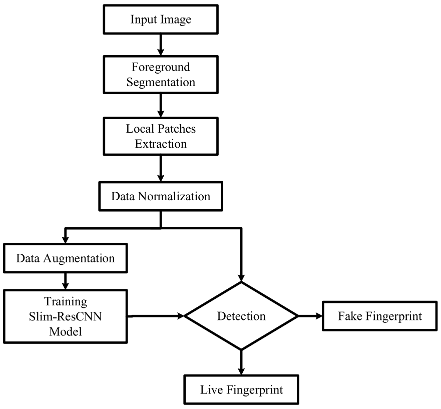


FIGURE 2. The flow chart of fingerprint spoof detection using the Slim-ResCNN structure.

and columns. Next, the local patches are segmented from the foreground region by center of gravity (CoG). During the training stage, the amount of local patches are augmented by flipping and rotating before feeding into Slim-ResCNN and training model. During the testing stage, the score of the local block of fingerprint obtained by the trained model can assess the liveness of the input fingerprint image to detect whether it's an ''alive'' entity or a ''spoof'' artefact. The flow chart of the proposed approach for fingerprint liveness detection is shown in Figure 2.

A. PATCH EXTRACTION

In order to build fingerprint liveness detection system, the local patches with fixed size need to feed into the network instead of using entire fingerprint images in that the local patches are crucial for CNN-based fingerprint liveness detection to reduce the time of training and decrease the parameter of model. A large quantity of local patches, which is extracted from the fingerprint images, not only works out the drawback of significant loss of discriminatory information which is caused by downsizing or resizing fingerprint images randomly, but also is adequate to train CNNs model from scratch without overfitting. Moreover, the suitable size of local patches can be adapted to both conventional sensors on the market and sensors on mobile devices.

The pretreatment of fingerprint images involves the foreground extraction and the patch segmentation. Firstly, the foreground region is extracted by the statistical histograms of rows and columns, which removes blank area without any information from the entire fingerprint image. Secondly, the patch segmentation utilizes CoG to select local patches from the foreground area, which reduces the network execution time and model parameters by reducing the size of input effectively.

1) FOREGROUND EXTRACTION

Considering the influence of effective fingerprint region on the performance of fingerprint liveness detection, the foreground extraction is first carried out. The blank area without any fingerprint information can be removed from the whole fingerprint image so that the size of fingerprint image will be decreased. Due to the continuity of the pixel distribution of fingerprint, the effective region is extracted by obtaining the rows and columns of foreground from the rows and columns statistical histograms of fingerprint image. The rows and columns statistical histograms on a sample, which is captured from CrossMatch sensor in LivDet2015, are shown in Figure 4. The image area that does not satisfy the condition is removed according to the threshold of the row and the column, then the foreground area of fingerprint image is obtained. The foreground region extracted from whole fingerprint image is shown in Figure 3.

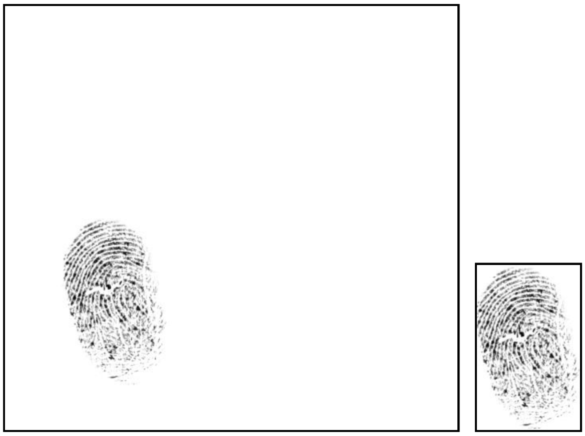
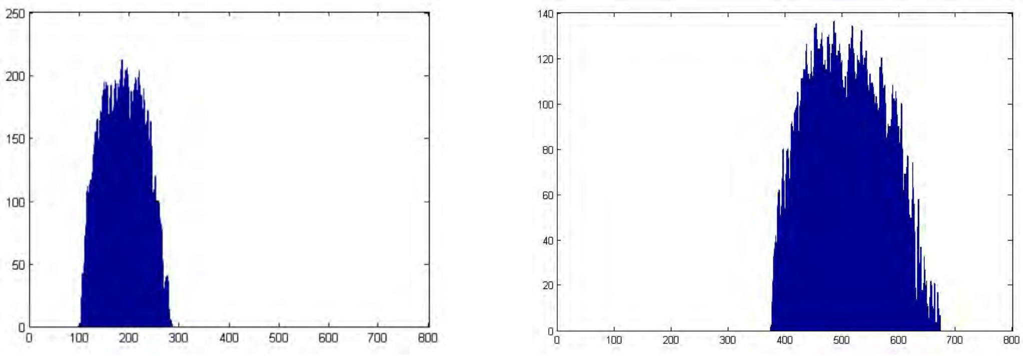


FIGURE 3. The foreground region (181 × 291) extracted from whole fingerprint image (800 × 800) captured from CrossMatch sensor in LivDet2015.

2) PATCH SEGMENTATION

The local patches are segmented based on CoG after the foreground extraction. To uniform the size of input and reduce the network execution time, the $w \times w$ ($w = 160$) patch is segmented from the foreground area of fingerprint. The local patches not only contain abundant fingerprint information but also ensure the effective implementation of the network training process.

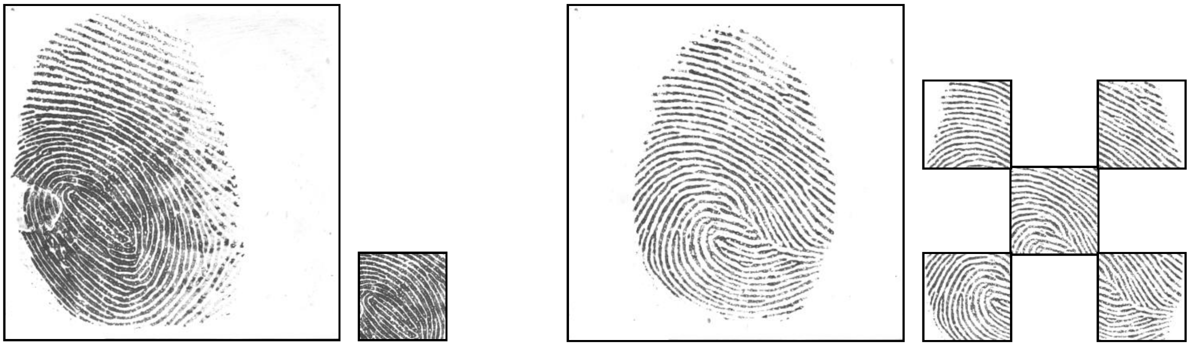
Data enhancement is employed to augment the number of samples in that it requires a large amount of data to train the Slim-ResCNN model from scratch, aiming at preventing overfitting on the small-scale fingerprint databases. In the testing stage, one $w \times w$ ($w = 160$) local patch is segmented based on the point of CoG from the foreground region, as shown in Figure 5 (a). Nevertheless, apart from the point of CoG in the training stage, another four points are selected separately at the top, bottom, right, and left of CoG according to a step size of 50 pixels. Next, some $w \times w$ patches are cut centered each of these points from the foreground region, which makes full use of the fingerprint image information, as shown in Figure 5 (b). Furthermore, some of these selected patches have little fingerprint information, which is not conducive to train network, hence they must be excluded from the training set. Color reversal and normalization are



(a) The columns histograms.

(b) The rows histograms.

FIGURE 4. The statistical histograms of fingerprint image by the columns (left) and rows (right).



(a) Single patch segmentation.

(b) Multiple patches segmentation.

FIGURE 5. Single local patch and multiple local patches extracted from the foreground of fingerprint image (the samples from LivDet2015 dataset acquired by Green Bit sensor).

performed on each $w \times w$ local patch. Next, the maximum closure of the binary patch is obtained. When the maximum closure area of local patch is more than 60% of the local patch area, it will be selected as one sample, and otherwise it will be excluded. Furthermore, the extracted local patches are flipped horizontally and vertically, and rotate at four different angles which are respectively 0°, 90°, 180° and 270°, to deal with the problem of insufficient fingerprint samples. The large amount available data, obtained by the pretreatment and data enhancement, is sufficient to train the Slim-ResCNN model from scratch to prevent over-fitting. The extracted local patches are randomly cropped into 112×112 in the training and testing stages for building model stability.

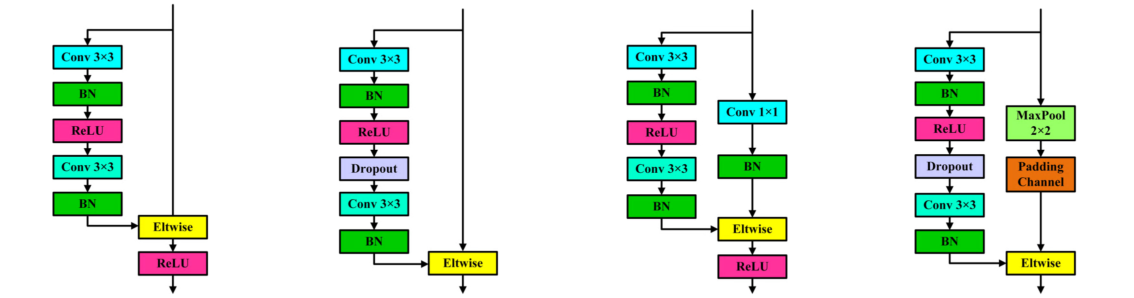
B. THE SLIM-RESCNN STRUCTURE

Inspired by the success of residual networks applied to some challenging object detection, localization and segmentation tasks, a relative simple and lightweight residual network, called Slim-ResCNN, is especially designed for fingerprint liveness detection in this paper. The Slim-ResCNN framework based on data-driven avoids tediously designing handcraft features.

1) THE IMPROVED RESIDUAL BLOCKS

Based on the original residual block_a (Figure 6 (a)) in original residual network [26], the improved residual block_b (Figure 6 (b)) firstly removes the activation function (ReLU) of the second convolutional kernel, which helps to maintain the representation power in the narrow layers [30]. As we want to study the effect of residual blocks, the bottleneck blocks, which is initially used to make blocks less computationally expensive to increase the number of layers, are not adopted. Secondly, the improved residual block_b is more effective where the convolutional kernels are broadened by twice and the dropout layer is inserted into each pair of convolutional layers [29]. In the forward propagation, the dropout layer makes the activation value of a certain neuron stop working with a certain probability p ($p = 0.5$), which makes the model more generalized for it does not rely too much on some local features.

When the dimensions increase, the improved residual block_d (Figure 6 (d)), where the 1×1 convolutional layer replace with the padding zero channel (called the padding channel layer), is employed instead of the original residual block_c (Figure 6 (c)). Neither are additional parameters introduced nor is the efficiency of reverse gradient



(a) Original residual block_a. (b) Improved residual block_b. (c) Original residual block_c. (d) Improved residual block_d.

FIGURE 6. The original residual blocks and the improved residual blocks.

flow reduced. In this paper, a detailed experimental study is conducted on the architecture of residual blocks to verify the effectiveness of the improvement. On the basis of the improved residual blocks, we propose a novel architecture which consist of nine improved residual blocks, called Silm-ResCNN, which is superior for fingerprint liveness detection.

2) THE OVERALL STRUCTURE OF SLIM-RESCNN

Following the principles of neural network structure design, a binary classification network structure is constructed for spoofs presentation attacks detection. The Slim-ResCNN consists of Conv1, Conv2, Conv3 (Conv3_1, Conv3_2), and Conv4 (Conv4_1, Conv4_2), followed by a global average pooling (Avg_Pool) and a final classification layer.

In order to adapt the neural network structure on small-scale datasets, the number of convolutional kernels is only broadened by twice, and the depth of the network is greatly compressed compared with original residual network. The number of network occupancy parameters is greatly reduced, therefore the network is called as Slim-ResCNN. The structure of Slim-ResCNN is illustrated in Table 1. The overall network structure is explained as follows:

TABLE 1. The overall structure of Slim-ResCNN consisting of improved residual blocks. Final classification layer is omitted for clearance. The network uses a ResNet block of type B(3,3).

Ground Name	Output Size	Block Type=B(3,3)
Conv1	112×112	$[3 \times 3/1, 32]$
Conv2	112×112	$\begin{bmatrix}3 \times 3/1, 32\\3 \times 3/1, 32\end{bmatrix} \times 3$
Conv3_1	56×56	$\begin{bmatrix}3 \times 3/2, 64\\3 \times 3/1, 64\end{bmatrix}$
Conv3_2	56×56	$\begin{bmatrix}3 \times 3/1, 64\\3 \times 3/1, 64\end{bmatrix} \times 2$
Conv4_1	28×28	$\begin{bmatrix}3 \times 3/2, 128\\3 \times 3/1, 128\end{bmatrix}$
Conv4_2	28×28	$\begin{bmatrix}3 \times 3/1, 128\\3 \times 3/1, 128\end{bmatrix} \times 2$
Avg_Pool	1×1	$[28 \times 28]$

(1) The Conv1 is responsible for connecting the input local patches and extract the initial features which are delivered to the followed residual block.

(2) The Conv1 is followed Conv2, Conv3 (Conv3_1, Conv3_2) and Conv4 (Conv4_1, Conv4_2). If the size of feature map is reduced by half, the convolutional kernel is doubled to preserve the time complexity per layer, as can be seen from Conv3_1 and Conv4_1.

(3) To reduce network model parameters, the global average pooling layer is used instead of the fully connected layer. The reduction of network parameters prevents overfitting and reduce the network computation cost.

(4) The Slim-ResCNN structure is trained on local patches using the cross-entropy loss function.

III. EXPERIMENTS

A. DATASETS

The efficiency of the proposed network is evaluated on three public datasets included LivDet2013 [28], LivDet2015 [4], and LivDet2017 [29]. LivDet is specially conducted for fingerprint presentation attack detection, which is held every two years from 2009. There are two methods to fabricate artificial fingerprints: the cooperative and the non-cooperative methods. LivDet2013 is composed of four datasets captured by four different fingerprint readers. Gelatine, latex, Ecoflex, modasil and wood glue are used to fabricate fake fingerprints with the non-cooperative method. In addition to Biometrika and ItalData, LivDet2013 involves two other datasets which are Swipe and CrossMatch. However, the Swipe sensor obtains the fingerprint images by swiping fingerprint from top to bottom in which the images are vastly different from existing data, hence these fingerprint images are excluded from the data for experimental analysis, and the LivDet2013 CrossMatch dataset is also eliminated because the original selection of subjects is an anomaly. The same spoof fingerprint materials are used in the training and testing sets of LivDet2013. LivDet2015 also contains four datasets, which are respectively Biometrika, Digital Persona, Green Bit and CrossMatch. It should be noted that the testing sets on LivDet2015 include fake fingerprints fabricated using new materials which includes liquid Ecoflex and RTV for Biometrika, Digital Persona, and Green Bit readers, and

TABLE 2. Summary of the liveness detection (LivDet) datasets utilized in this study (the training sets are a similar size but did not include the unknown materials which are bold).

Dataset	Sensor	Image Size	Resolution(dpi)	# of training(Live/spoof)/ # of testing(Live/spoof)	Cooperative Subject	Spoof Materials
LivDet2013	Biometrika Ita1data	315 × 372 640 × 480	569 500	(1000/1000)/(1000/1000) (1000/1000)/(1000/1000)	No No	Ecoflex, Gelatine, Latex, Modasil, WoodGlue
LivDet2015	Biometrika Digital Persona GreenBit	1000 × 1000 252 × 324 500 × 500	1000 500 500	(1000/1000)/(1000/1500) (1000/1000)/(1000/1500) (1000/1000)/(1000/1500)	Yes Yes Yes	Ecoflex, Gelatine, Latex, Liquid Ecoflex, RTV, WoodGlue
	CrossMatch	640 × 480	500	(1510/1473)/(1500/1448)	Yes	Body Double, Ecoflex, PlayDoh, **OOMOO**, **Gelatin**
LivDet2017	GreenBit Orcanthus Digital Persona	500 × 500 300 × n 252 × 324	500 500 500	(1000/1200)/(1700/2040) (1000/1200)/(1700/2676) (999/1199)/(1700/2028)	Yes Yes Yes	Body Double, Ecoflex, WoodGlue, **Gelatine**, **Latex**, **Liquid Ecoflex**

OOMMOO and gelatin for Crossmatch reader. However, only three datasets on LivDet2017, the fingerprint images captured from Digital Persona, Orcanthus, and GreenBit sensors. Furthermore, the materials used in the training sets (Wood Glue, Ecoflex, Body Double) are completely different with ones (Gelatine, Latex, Liquid Ecoflex) in the testing sets. Notably, LivDet2017 employs a new peculiarity where fake samples in the training sets are built by an operator and fake samples in the testing sets are built by two other persons for simulating a real scenario. The fingerprint images on every dataset are equally divided into the training set and the testing set. The training sets are often used to train the models, and then the detection ability of models is evaluated on the testing sets. In LivDet datasets, all live fingerprint images come from multiple acquisitions of all fingers of different subjects, while spoof fingerprint images are collected using cooperative method or non-cooperative method. The sizes of fingerprint images vary on different fingerprint scanners, ranging from 252 × 324 to 1000 × 1000 pixels, hence the fingerprint images need to be unified to the same size before feeding into the network. The LivDet datasets used in this experiment is outlined in Table 2.

B. EXPERIMENTAL ENVIRONMENT AND PERFORMANCE EVALUATION METRICS

We train our models using stochastic gradient descent with a batch size of 15 samples. The learning rate is initialized at 0.01 and reduced 20% per 200,000 iterations in the training stage. We train the network for roughly 600,000 iterations when the loss function became convergent. Our implementation is derived from the publicly available C++ Caffe toolbox. The GPU used in experiments is NVIDIA GeForce GTX 1080.

In all experiments of this paper, we follow the performance measurement stated by [4] and [23] to ensure consistency during experimental results comparison. The Average Classification Error (ACE) is the average of the error detection rate of live fingerprints ($F_{errlive}$) and the error detection rate of fake fingerprints ($F_{errfake}$). The ACE parameters adopted for the performance evaluation is defined as:

$$ACE = \frac{F_{errlive} + F_{errfake}}{2} \tag{1}$$

In addition to using generic ACE to evaluate the performance of fingerprint liveness detection, the rate of samples (live fingerprints and fake ones) correctly classified, called accuracy, is also a fundamental parameter of classification. The threshold for determining liveness of fingerprint is set to 0.5. The fingerprint image with liveness score over 0.5 is considered as "alive" entity, otherwise it is considered as "spoof" artefact. The following experimental results are calculated based on the threshold, except for special explanations.

C. EXPERIMENTAL RESULTS AND ANALYSIS

In this paper, several different types of experiments have been implemented to prevent presentation attacks bypassing the fingerprint authentication systems. The training samples of live and fake fingerprints need to be shuffled before feeding into the network. In all experiments, the model parameters initialization is performed the Gaussian distributions where the weights are randomly drawn with fixed mean and fixed standard deviation. It should be emphasized that the labels corresponding to fingerprints are necessary, where the label of live fingerprint is "1" and fake fingerprint is "0". The model is a two-category, which is either a live fingerprint or a spoof fingerprint. When a fingerprint sample enters network, the final output of model, called the predicted probability, is changed into the predicted label according to the threshold. If the predicted value is the same as the real label, then the prediction is correct, otherwise the prediction is wrong. In LivDet2013, LivDet2015 and LivDet2017 datasets, Slim-ResCNN models are trained for every training set respectively where the fingerprint images are captured from the same fingerprint scanner.

Firstly, this paper evaluates and compares the performance of three different kinds of Slim-ResCNNs which are modified version of Residual Network (ResNet), special for fingerprint liveness detection on LivDet2015 dataset. The three Slim-ResCNNs have the same depth, as described in Table 1, yet subtle difference on network structure. The impact of stride size of convolutional kernel in the first convolutional layer is evaluated by comparing the performance of two different CNN models trained on the same experimental conditions. Compared to one model with a stride of 2,

TABLE 3. The performance comparison on ACEs of different types of Slim-ResCNN structures in LivDet2015 dataset (%).

Datasets	padding channel layer stride=1	padding channel layer stride=2	1 × 1 convolutional layer stride=1
Biometrika	2.78	3.79	4.23
CrossMatch	2.82	3.51	3.65
Digital Persona	4.53	6.40	4.81
GreenBit	2.17	4.27	2.41
Average	3.07	4.49	3.52

TABLE 4. The performance comparison on ACEs of three Slim-ResCNN models with different numbers of improved residual block_bs in LivDet2015 dataset (%).

Datasets	4 improved residual block_bs	7 improved residual block_bs	10 improved residual block_bs
Biometrika	4.51	2.77	3.80
CrossMatch	3.12	2.82	4.10
Digital Persona	4.88	4.53	4.53
GreenBit	1.80	2.17	2.71
Average	3.58	3.07	3.80

the other model with a stride of 1 performs better on the four datasets of LivDet2015, as shown in Table 3. The comparison of experimental results indicates that the resolution of fingerprint image input sent into the network should not be decreased too early, which may throw away some inherent feature information permanently at the beginning of training. When changing the input or output channels, a linear projection by shortcut connections is performed to match dimensions. The linear projection can be 1 × 1 convolutional layer (Figure 6 (c)) or proposed padding channel layer (Figure 6 (d)). We evaluate the two types of linear projections by compared the performance of two CNN models with different linear projections, one CNN model using the 1 × 1 convolutional layer and the other one using the proposed padding channel layer. The model with the padding channel layer has brought about a consistence improvement in ACE, except on GreenBit sensor, which argues that the padding channel layer improves the performance of model effectively when the dimensions increase. The 1 × 1 convolutional layer brings in extra parameters, resulting in a certain degree of network overfitting. The experimental results on kinds of Slim-ResCNNs with different structures demonstrate that the Slim-ResCNN with padding channel and a stride with 1, is more effective than other two network structures. The comparison results of different improvements are outlined in Table 3.

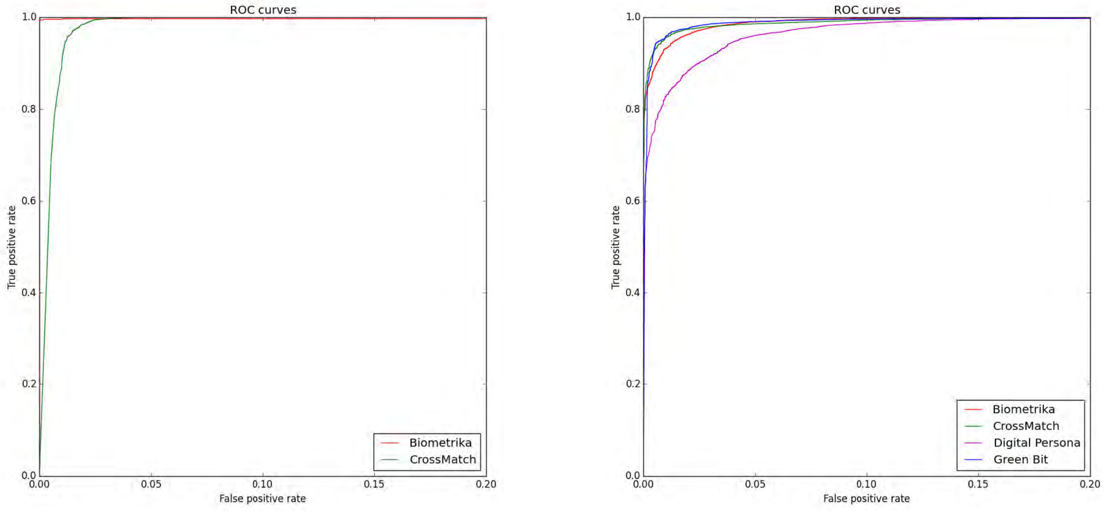
Secondly, the impact of network depth on the classification performance of real and fake fingerprints is evaluated on LivDet2015 dataset through three different depths Slim-ResCNN structures which are comprised of different numbers of improved residual blocks. The network structure, which employs the stride = 1 of the convolutional kernel on the first convolutional layer and uses the padding channel layer as a linear projection when the dimension increase, is chosen according to the first experimental results. Every Slim-ResCNN structure contains two improved residual block_ds (Figure 6 (d)) and k improved residual block_bs (Figure 6 (b)), where $k = 4$, 7 and 10. The comparison experimental results at different depths show that the Slim-ResCNN model with 4 improved block_bs performs worst among the three Slim-ResCNNs for all fingerprint sensors. The Slim-ResCNN model with 7 improved block_bs and one with 10 improved block_bs have some differences in the performance of ACE, as shown in Table 4. It indicates that the increase in network depth has a certain improvement in fingerprint liveness detection, however when the network depth reaches a certain level, the deeper network structure does not necessarily achieve the best classification performance on this experiments. Furthermore, when the network is deepened, it results in longer training time, more execution time, and larger model memory.

Thirdly, the effectiveness of the proposed method has been verified based on the first two sets of comparative experiments. The ACEs of proposed method are compared to the existing works on LivDet2013 and LivDet2015 datasets. Table 5 presents the comparisons between Silm-ResCNN and the most recent algorithms, including fPADnet [31], VGG-19 [23], Gram-128 model [32] and residual network (ResNet). Furthermore, the optimal threshold is selected when the best rate of samples correctly classified is obtained in the training set, which is applied to the testing set. The performance of optimal threshold differs according to sensors, but the performance is typically better on average ACE compared to the fixed threshold of 0.5. The experimental results show in the "Slim-ResCNN(thres)" column in Table 5, which is compared with the fixed threshold mentioned earlier. The overall comparison results is shown in Table 5. The ROC (Receiver Operating Characteristic) curves of the Slim-ResCNN model on LivDet2013 and LivDet2015 are shown in Figure 7. The false positive rate of ROC is zoomed between 0 and 20% so that the changing trend can be seen more clearly.

Experiments on LivDet2013 dataset use fake fingerprints fabricated with the same material for training and testing. However, the testing sets on LivDet2015 dataset consist of fake fingerprints made of unknown materials.

TABLE 5. The ACE comparisons of different methods on LivDet2013 and LivDet2015 datasets (%).

Datasets	sensor	fPADnet [31]	VGG-19 [23]	Gram-128 model [32]	ResNet	Slim-ResCNN	Slim-ResCNN (thres)
LivDet2013	Biometrika	0.9	1.8	0.85	4.09	0.47	0.47
	Italdata	1.3	0.4	1.25	17.16	5.21	3.01
LivDet2015	Biometrika	1.4	4.6	4.1	32.06	2.78	3.10
	CrossMatch	4.1	5.6	0.27	9.59	3.03	4.32
	Digital Persona	8.5	6.3	8.5	40.61	4.48	2.37
	GreenBit	0.3	1.9	1.35	14.74	2.14	2.64
average		4.1	3.4	2.72	19.71	3.07	2.65



(a) ROC curves of LivDet2013 dataset.

(b) ROC curves of LivDet2015 dataset.

FIGURE 7. The ROC curves of the proposed method on LivDet2013 and LivDet2015 datasets.

TABLE 6. The detailed Performance comparison between the proposed approach (bottom) and state-of-the-art (top) reported on LivDet2015 dataset.

	LivDet2015	Fcorrlive (%)	Fcorrfake (%)	Fcorrfake known (%)	Fcorrfake unknown (%)	Accuracy (%)
State-of-the-Art [4]	Biometrika	91.50	96.27	97.30	94.20	94.36
	CrossMatch	99.07	97.10	97.88	95.98	98.10
	Digital Persona	91.90	94.93	95.40	94.00	93.72
	GreenBit	96.50	94.67	95.70	92.60	95.40
	Average	94.74	95.74	96.57	94.20	95.40
Proposed Approach	LivDet2015	Fcorrlive (%)	Fcorrfake (%)	Fcorrfake known (%)	Fcorrfake unknown (%)	Accuracy (%)
	Biometrika	96.65	95.77	97.56	92.28	97.02
	CrossMatch	98.28	96.09	96.82	95.42	97.01
	Digital Persona	95.72	95.22	96.26	93.60	95.42
	GreenBit	97.78	97.35	97.58	96.89	97.81
	Average	96.86	96.11	97.01	94.55	96.82

Table 6 presents the more detailed performance comparisons on LivDet2015 dataset between the proposed approach and the state-of-the-art results reported at the Fingerprint Liveness Detection Competition 2015 [4]. Fcorrlive counts the percentage of correctly classified live fingerprints and Fcorrfake counts the percentage of fakes classified as such for all fake fingerprint images (known and unknown). Fcorrfake known and Fcorrfake unknown are the percentage correctly classified spoof image from known materials and unknown materials respectively. A comparison between Fcorrfake known and Fcorrfake unknown shows that the proposed method has a small decrease in classification accuracy when fake fingerprints are fabricated with unknown materials to test the model. The proposed approach achieves 96.82% overall accuracy over four datasets compared with 95.40% achieving by the champion of the LivDet2015 competition. The accuracy has been improved for most datasets of LivDet2015, except CrossMatch dataset.

Lastly, we have no opportunity to do experiments on the testing sets of LivDet2017 dataset which have not been opened so far. The fingerprint images of the training sets come from all fingers of 20 different subjects. In order to evaluate the performance of the trained models, every official training set is divided into the training set and the verification

TABLE 7. The accuracy of the LivDet2017 verification sets (%).

Datasets	Slim-ResCNN
Green Bit	99.52
Digital Personal	99.03
Orcanthus	98.43
Average	98.99

TABLE 8. Accuracy of the algorithms on the testing sets of the LivDet2017 dataset [%]. 1 = GreenBit, 2 = Digital persona, 3 = Orcanthus.

Algorithm	1	2	3	Overall
SLFD	93.58	94.33	93.14	93.68
JLW_A	95.08	94.09	93.52	94.23
JLW_B	**96.44**	**95.59**	**93.71**	**95.25**
OKIBrB20	84.97	83.31	84.00	84.09
OKIBrB30	92.49	89.33	90.64	90.82
ZYL_1	95.91	95.13	91.66	94.23
ZYL_2	96.26	94.73	93.17	94.72
SNOTA2017_1	95.03	91.26	91.58	92.62
SNOTA2017_2	94.04	86.72	86.74	89.17
ModuLAB	94.25	90.40	90.21	91.62
ganfp	95.67	93.66	94.16	94.50
PB_ LivDet_1	93.85	89.97	91.85	91.89
PB_ LivDet_2	92.86	90.43	92.60	91.96
hanulj	97.06	92.34	92.04	93.81
SpoofWit	93.66	88.82	89.97	90.82
LCPD	89.87	88.84	86.87	88.52
PDfV	92.86	93.31	N. A	N. A

set according to the number of people. The training set accounts for four copies and the verification set occupies one copy, which means the fingerprint images of 16 people are used for training the model and 4 people's fingerprint images are used to evaluate the classification performance of Slim-ResCNN model. The experimental results of the verification sets show that the proposed approach perform well, where the average accuracy of three testing sets achieve 98.99% on the verification sets, as shown in Table 7. Since there are not many results of LivDet2017's state-of-the-art performance in literature thus far, the results of the Fingerprint Liveness Detection Competition 2017 are used in this paper. Inspired by score fusion [33], the fusion of multiple local patches extracted from the fingerprint image is effective for fingerprint liveness detection. The multiple local patches and single local patch are utilized to detect the liveness of fingerprint image. The overall accuracy of single local patch achieves 94.23%. In contrast, multiple local patches averaging the score-level fusion of multiple patches achieve an overall classification accuracy of 95.25%, which improves 1% approximately compared to single local patch. It reveals that the multiple local patches are more efficient than single local patch. The outstanding performance of the proposed method using multiple local patches has been awarded the first place in the Fingerprint Liveness Detection2017 Competition. Table 8 summarizes the correct classification rates of each algorithm on every testing set of LivDet2017 dataset, in addition to the overall average accuracy. The complete results of the LivDet2017 competition have been reported at [29].

Since the fake fingerprints in the testing sets are fabricated using new materials that have not seen in the training sets, the accuracy of fingerprint liveness detection has a slight discrepancy between the testing sets and the verification sets. Based on the experiment results on three testing sets of LivDet2017 dataset, the good performance is observed for the proposed method even regarding the unknown material fake fingerprint in the testing sets.

D. EXPERIMENT PROCESSING TIME

The time performance of the proposed method is evaluated on PC with GeForce GTX 1080. It takes one second to process 50 local patches on average, thus the proposed algorithm can satisfy the real-time processing requirement in a real scenario.

IV. CONCLUSION

Spoof attacks with artificial replicas threat the safety of fingerprint recognition systems severely to a large extent, therefore, it is urgent to require effective countermeasures against spoof attacks. On the basic of analyzing the deficiency of the existing methods, we propose a new method based on statistical histogram and CoG to extract local patches, which prevents the blank area feeding into the network. Moreover, we design a new residual network structure specially for the fingerprint liveness detection, which compensates for the drawback of the pre-trained CNN models on natural images that neglect the inevitable differences between fingerprint images and natural images. In this paper, a relative simple and lightweight residual convolutional neural network, called Slim-ResCNN, consists of several improved residual blocks, which is designed to distinguish between live fingerprint and fake ones. The advantages of Slim-ResCNN is showed through the comparative experiments. Compared with other methods, the Slim-ResCNN structure is suitable for fingerprint liveness detection which wins the first place of LivDet2017 competition with an overall accuracy of 95.25%. In some extend, the Slim-ResCNN is robust to fake fingerprints with new materials. In the future, we will focus on researching the fingerprint liveness detection algorithm that is robust to new materials and different types of fingerprint sensors.

REFERENCES

[1] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Comput. Surv.*, vol. 47, no. 2, pp. 28:1–28:36, Jan. 2014.

[2] S. S. Arora, K. Cao, A. K. Jain, and N. G. Paulter, Jr., "Design and fabrication of 3D fingerprint targets," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 10, pp. 2284–2297, Oct. 2016.

[3] C. W. Schultz, J. X. H. Wong, and H.-Z. Yu, "Fabrication of 3D fingerprint phantoms via unconventional polycarbonate molding," *Sci. Rep.*, vol. 8, Jun. 2018, Art. no. 9613.

[4] V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay, and S. A. Schuckers, "LivDet 2015 fingerprint liveness detection competition 2015," in *Proc. IEEE 7th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Arlington, VA, USA, Sep. 2015, pp. 1–6.

[5] P. Coli, G. L. Marcialis, and F. Roli, "Vitality detection from fingerprint images: A critical survey," in *Proc. Int. Conf. Adv. Biometrics (ICB)*, S.-W. Lee and S. Z. Li, Eds. Berlin, Germany: Springer-Verlag, 2007, pp. 722–731.

[6] B. Tan and S. Schuckers, "Liveness detection for fingerprint scanners based on the statistics of wavelet signal processing," in *Proc. Conf. Comput. Vis. Pattern Recognit. Workshop (CVPRW)*, New York, NY, USA, 2006, p. 26.

[7] A. S. Abhyankar and S. C. Schuckers, "A wavelet-based approach to detecting liveness in fingerprint scanners," *Proc. SPIE*, vol. 5404, pp. 278–286, Aug. 2004.

[8] R. K. Dubey, J. Goh, and V. L. L. Thing, "Fingerprint liveness detection from single image using low-level features and shape analysis," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1461–1475, Jul. 2016.

[9] S. B. Nikam and S. Agarwal, "Texture and wavelet-based spoof fingerprint detection for fingerprint biometric systems," in *Proc. 1st Int. Conf. Emerg. Trends Eng. Technol.*, Nagpur, India, 2008, pp. 675–680.

[10] X. Jia, X. Yang, K. Cao, Y. Zang, N. Zhang, R. Dai, X. Zhu, and J. Tian, "Multi-scale local binary pattern with filters for spoof fingerprint detection," *Inf. Sci.*, vol. 268, pp. 91–102, Jun. 2014.

[11] Y. Jiang and X. Liu, "Uniform local binary pattern for fingerprint liveness detection in the Gaussian pyramid," *J. Elect. Comput. Eng.*, vol. 2018, Jan. 2018, Art. no. 1539298.

[12] L. Ghiani, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection by local phase quantization," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Tsukuba, Japan, 2012, pp. 537–540.

[13] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "Fingerprint liveness detection based on Weber local image descriptor," in *Proc. IEEE Workshop Biometric Meas. Syst. Secur. Med. Appl.*, Naples, Italy, Sep. 2013, pp. 46–50.

[14] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "Local contrast phase descriptor for fingerprint liveness detection," *Pattern Recognit.*, vol. 48, no. 4, pp. 1050–1058, 2015.

[15] Z. Xia, C. Yuan, R. Lv, X. Sun, N. N. Xiong, and Y.-Q. Shi, "A novel weber local binary descriptor for fingerprint liveness detection," *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published. doi: 10.1109/TSMC.2018.2874281.

[16] A. Toosi, A. Bottino, S. Cumani, P. Negri, and P. L. Sottile, "Feature fusion for fingerprint liveness detection: A comparative study," *IEEE Access*, vol. 5, pp. 23695–23709, 2017.

[17] E. Marasco, P. Wild, and B. Cukic, "Robust and interoperable fingerprint spoof detection via convolutional neural networks," in *Proc. IEEE Symp. Technol. Homeland Secur. (HST)*, Waltham, MA, USA, Aug. 2016, pp. 1–6.

[18] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcão, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 864–879, Apr. 2015.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[20] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 580–587.

[21] N. Zhang, M. Paluri, M. Ranzato, T. Darrell, and L. Bourdev, "PANDA: Pose aligned networks for deep attribute modeling," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 1637–1644.

[22] R. F. Nogueira, R. de Alencar Lotufo, and R. C. Machado, "Evaluating software-based fingerprint liveness detection using convolutional networks and local binary patterns," in *Proc. IEEE Workshop Biometric Meas. Syst. Secur. Med. Appl. (BIOMS)*, Rome, Italy, Oct. 2014, pp. 22–29.

[23] R. F. Nogueira, R. de Alencar Lotufo, and R. C. Machado, "Fingerprint liveness detection using convolutional neural networks," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 6, pp. 1206–1213, Jun. 2016.

[24] T. Chugh, K. Cao, and K. Anil Jain, "Fingerprint spoof buster: Use of minutiae-centered patches," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 9, pp. 2190–2202, Sep. 2018.

[25] Y. Li, J. Zhou, F. Huang, and L. Liu, "Sub-pixel extraction of laser stripe center using an improved gray-gravity method," *Sensors*, vol. 17, no. 4, p. 814, Apr. 2017.

[26] K. M. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[27] S. Zagoruyko and N. Komodakis, "Wide residual networks," in *Proc. BMVC*, 2016, pp. 87.1–87.12.

[28] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G. L. Marcialis, F. Roli, and S. Schuckcrs, "LivDet 2013 fingerprint liveness detection competition 2013," in *Proc. Int. Conf. Biometrics (ICB)*, Madrid, Spain, 2013, pp. 1–6.

[29] V. Mura, G. Orrù, R. Casula, A. Sibirium, G. Loi, P. Tuveri, L. Ghiani, and G. L. Marcialis, "LivDet 2017 fingerprint liveness detection competition 2017," in *Proc. Int. Conf. Biometrics (ICB)*, Gold Coast, QLD, Australia, 2018, pp. 297–302.

[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 4510–4520.

[31] T. H. B. Nguyen, X. Cui, V. Nguyen, and H. Kim, "fPADnet: Small and efficient convolutional neural network for presentation attack detection," *Sensors*, vol. 18, no. 8, p. 2532, 2018.

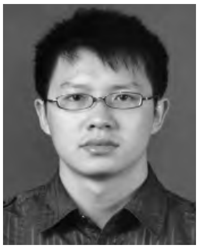
[32] E. Park, X. Cui, W. Kim, and H. Kim, "End-to-end fingerprints liveness detection using convolutional networks with gram module," 2018, *arXiv:1803.07830*. [Online]. Available: https://arxiv.org/abs/1803.07830

[33] C. Wang, K. Li, Z. H. Wu, and Q. Zhao, "A DCNN based fingerprint liveness detection," in *Proc. Chin. Conf. Biometric Recognit.* Cham, Switzerland: Springer, 2015, pp. 241–249.

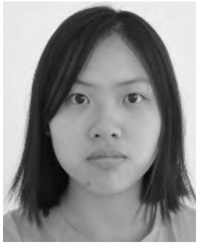
YONGLIANG ZHANG received the B.S. and M.S. degrees from Jilin University, China, in 2000 and 2003, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, China, in 2007. He is currently an Associate Professor with the College of Computer Science and Technology, Zhejiang University of Technology. His research interests include biometric identification, pattern recognition, and artificial intelligence.

DAQIONG SHI received the B.S. degree from Huzhou University, China, in 2017. She is currently pursuing the M.S. degree with the School of College of Computer Science and Technology, Zhejiang University of Technology. Her research interests include image processing and machine learning.

XIAOSI ZHAN received the B.S. and M.S. degrees from Jilin University, China, in 1998 and 2001, respectively, and the Ph.D. degree from Nanjing University, China, in 2004. He is currently a Professor with the School of Science and Technology, Zhejiang International Studies University. His research interests include image processing, pattern recognition, and biometric identification.



DI CAO received the B.S. and Ph.D. degrees from the University of Strathclyde, in 2008 and 2013, respectively. He is currently a Teacher with the College of Computer Science and Technology, Zhejiang University of Technology. His research interests include the Internet of Things and wireless sensor network.



KEYI ZHU was born in Hangzhou, China, in 1997. She is currently pursuing the B.E. degree in electrical and electronic engineering with the University of Electronic Science and Technology, China. Her research interests include image signal processing, computer vision, and machine learning.

ZHIWEI LI received the M.S. degree from the Center for Optics and Optoelectronics Research, College of Science, Zhejiang University of Technology, China, in 2016. He is currently an Algorithm Engineer with Hangzhou Jing Lianwen Technology Company Ltd. His research interests include biometric recognition, machine learning, and digital forensic.