Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

3D CNN-based fingerprint anti-spoofing through optical coherence tomography

Yilong Zhang [a], Shichang Yu [a], Shiliang Pu [c], Yingyu Wang [c], Kanlei Wang [c], Haohao Sun [b], Haixia Wang [a,*]

[a] *College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, 310023, China*
[b] *College of Information Engineering, Zhejiang University of Technology, Hangzhou, 310023, China*
[c] *Hikvision Research Institute, Hangzhou, 310023, China*

ABSTRACT

Optical coherence tomography (OCT) is a noninvasive high-resolution imaging technology that can accurately acquire the internal characteristics of tissues within a few millimeters. Using OCT technology, the internal fingerprint structure, which is consistent with external fingerprints and sweat glands, can be collected, leading to high anti-spoofing capabilities. In this paper, an OCT fingerprint anti-spoofing method based on a 3D convolutional neural network (CNN) is proposed, considering the spatial continuity of 3D biometrics in fingertips. Experiments were conducted on self-built and public datasets to test the feasibility of the proposed anti-spoofing method. The anti-spoofing strategy using a 3D CNN achieved the best results compared with classic networks.

1. Introduction

Every individual possesses unique biological characteristics, such as fingerprints, palm prints, voice prints, gait, and iris patterns. The limitations of traditional identity verification can be overcome by utilizing biometrics for identity authentication, enhancing both the accuracy and convenience of identity recognition [1].

Among various biometric features, fingerprints have emerged as the most widely adopted and popular because of their uniqueness, invariance, and convenience. This attribute has positioned fingerprints as a commonly employed trait for authentication [2,3], with more than 40% of biometric authentication systems worldwide employing fingerprint identification [4]. Even if two individuals are identical twins [5], an identification system can accurately and reliably determine their identities using fingerprints. Currently, automatic fingerprint identification systems primarily assess external fingerprint images using detailed features obtained by imaging the surface of fingers. However, external fingerprints can be susceptible to artificial when using materials such as silica gel and capacitive glue. Furthermore, unauthorized individuals can easily acquire and provide various counterfeit fingerprint samples to gain unauthorized access [6,7]. Additionally, self-made artificial fingers have been used to attack automatic fingerprint identification systems [5], and a considerable number of artificial fingerprints have been misjudged by automatic fingerprint identification systems. Hence, the anti-spoofing ability of fingerprint recognition systems has attracted the attention of researchers in biometrics and related fields. Researchers have begun to explore methods that can distinguish artificial fingerprints in different directions.

In recent years, the use of neural networks for computer vision has produced astounding results in areas such as face pose estimation [8–10], click prediction [11,12], and face recognition [13,14]. Neural networks also perform well in the field of fingerprint anti-spoofing. A series of neural-network-based methods have been proposed. The characteristics of bonafide and artificial

[*] Corresponding author.
E-mail address: hxwang@zjut.edu.cn (H. Wang).

https://doi.org/10.1016/j.heliyon.2023.e20052
Received 5 June 2023; Received in revised form 8 September 2023; Accepted 9 September 2023
Available online 14 September 2023

2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

fingerprints, including the type of minutiae, ridge flow or orientation, ridge contour or shape, and ridge spacing or ridge density, are used to detect demonstration attacks of artificial samples using these neural-network-based methods. Park et al. [15] proposed a small fully convolutional neural network (CNN) that can be integrated into a fingerprint acquisition system. Anti-spoofing identification can be performed by classifying the collected images into bonafide fingerprints, artificial fingerprints, and background information. Wang et al. [16] cut the fingerprint image into blocks and input the local fingerprint image into a fully connected layer network for classification to determine the authenticity of the fingerprint. Uliyan et al. [17] used discriminative restricted Boltzmann machines to accurately recognize the fingerprints against fabricated materials used for spoofing. Maheswari et al. [18] proposed convolution neural network and dynamic differential annealing (CNN-DDA)-based spoofed fingerprint detection to analyze and evaluate fingerprint spoofing and forgery authentication systems. Kong et al. [19] proposed a novel method for handling noisy information: channel-wise feature denoising for fingerprint presentation attack detection (CFD-PAD). The aforementioned methods are all based on surface fingerprints for anti-spoofing.

Simultaneously, hardware-based live detection methods have been introduced into fingerprint anti-spoofing research. Baldisiserra et al. [20] integrated odor sensors into a fingerprint collection system to sense the odor difference between human skin and artificial samples and to judge the authenticity of the detected samples. Through biological research, Drahansky et al. [21] found that heart contraction causes changes in blood flow, thus causing changes in the contact volume of the finger on the fingerprint reader. The tiny changes were collected by the sensor, amplified, and transmitted back to the control center to conduct anti-spoofing research on the tested fingerprint. Venkata et al. [22] used sensors to obtain the absorbance of the finger and artificial finger films for light of different wavelengths to determine the blood oxygen content of the finger and realize automatic anti-spoofing of the finger. Yau et al. [23] developed a fingerprint tactile sensing system that utilized the conductive characteristics of a finger to identify the authenticity of the collected finger. To further enhance the anti-spoofing ability of fingerprint recognition systems, researchers have explored the use of advanced imaging technologies such as optical coherence tomography (OCT).

OCT is based on the principle of low-coherence interferometry [24] and has been widely used in the field of biomedical imaging for more than 30 years [25,26]. OCT is suitable for various biomedical applications, such as retinal nerve fiber layer [27], COVID-19 [28], and skin layer detection [29]. OCT can collect three-dimensional (3D) volumetric data from 0 to 3 mm below the surface of the finger skin. Based on the rich fingertip information collected by OCT systems, researchers have explored their anti-spoofing characteristics. Chugh et al. [30] extracted the features of a single slice and input them into the Inception-V3 network and compared the network prediction value with the threshold value to determine whether the fingerprint was bonafide or artificial. Liu et al. [31,32] proposed the one-class PAD (OCPAD) method, which only requires bonafide cross-sectional OCT-based fingerprint images for training and an efficient fingerprint anti-spoofing system with high accuracy and robustness using OCT to anti-counterfeit the fingerprints. Zhang et al. [33] combined one-class wavelet transforms to propose a frequency-domain OCT presentation attack detection method.

However, existing OCT-based fingerprint anti-spoofing methods suffer from significant limitations, pertaining to the early employment of conventional techniques as well as the more recent utilization of neural networks for fingerprint anti-spoofing. These approaches primarily focus on individual cross-sectional images (B-scans) extracted from the volume data acquired using the OCT system. Unfortunately, these approaches do not fully leverage the inherent 3D spatial continuity within the volume data during the anti-spoofing procedure, and fail to consider the correlation between B-scans.

Building on the limitations of existing OCT-based fingerprint anti-spoofing methods, our study seeks to bridge this gap by harnessing the inherent 3D spatial continuity present within the volume data acquired through OCT. As shown in Fig. 1 and (a) presents a series of B-scans of OCT volume data and Fig. 1 (b) is an example of the 3D volume data reconstructed from (a). It can be observed that 3D biometrics, such as sweat glands as well as the epidermis and dermis of the fingertips in OCT cross-sectional images, are continuous in 3D space. The biometrics in the volume data of the fingertips collected by OCT have 3D information and spatial continuity, which represent significant advantages over most current fingerprint collection devices. To leverage the spatial coherence present in OCT data within a 3D space, researchers have incorporated 3D CNNs across various domains of OCT data. They aimed to capitalize on intrinsic spatial relationships and enhance the effectiveness of their methodologies in their respective domains. For instance, Maetschke et al. [34] used 3D CNNs to differentiate between the OCT-captured images of healthy eyes and those afflicted by glaucoma. Their approach focused on direct analysis of the optic nerve papillae (ONH) to accurately categorize OCT-captured eye images. Similarly, Yang et al. [35] utilized 3D fully CNNs to partition and extract both internal and external patterns from OCT fingerprint data.

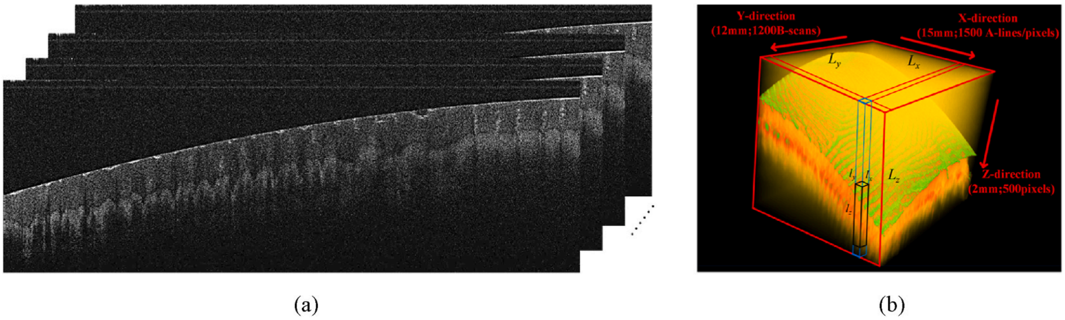

Fig. 1. (a) A series of B-scans of OCT volume data. (b) An example of OCT volume data [35].

Considering the exceptional capabilities of 3D CNNs in processing OCT 3D volume data, we explored its application in extracting continuous 3D spatial information from OCT volume data in the realm of fingerprint anti-spoofing research. To the best of our knowledge, this is the first study to use the spatial continuity of OCT 3D volume data for fingerprint anti-spoofing research using a 3D CNN.

This study proposes an anti-spoofing method that utilizes the 3D spatial continuity of OCT for the anti-spoofing field of OCT fingerprints. This method focuses on continuous information in the 3D space of fingerprints collected using OCT. We devised a methodology to extract small patches of the region of interest (ROI) from OCT 3D volume data. Furthermore, we crafted a 3D CNN in conjunction with a suite of anti-spoofing strategies, leveraging the extracted small patches to accurately determine the authenticity of the OCT volume data. As shown in Fig. 3 (c), the double-layer fingerprint film is locally very similar to the bonafide fingerprint, resulting in a high score in the certain area. We then designed a strategy to prevent the partial score of artificial fingerprints from being too high, which may cause the final anti-spoofing results to be incorrect. The anti-spoofing capability of this technique was tested on two OCT databases obtained using different OCT systems and different types of artificial materials. The results were satisfactory, demonstrating the robustness and security of the proposed method.

2. Methods

2.1. OCT system and data acquisition

Our research focused on OCT systems utilizing a central light source with a wavelength of 1310 nm [36]. The setup of the spectral-domain OCT system is illustrated in Fig. 2 (a). The OCT system is enclosed within a protective casing to enhance system stability and mitigate external influences. A schematic diagram of the system in the shell is shown in Fig. 2 (b), whose basic structure is that of a Michelson interferometer. The light from the broad light source is divided into reference and sample arms using a 50 × 50 fiber coupler. In the reference arm, the light is returned to the fiber through mirror reflection; while in the sample arm, the light is reflected off the detection tissue material. The two beams of returned light interfere with the fiber coupler, and a spectrometer with a charge-coupled device (CCD) is used to obtain the depth information of the fingertips (A-line). Finally, OCT volumetric data corresponding to the real finger scanning area, measuring 18 mm × 14 mm, are acquired. These data possess dimensions of 500 × 1800 × 1400 pixels owing to the scanning conducted by the X and Y scanners, as shown in Fig. 2(c) and (d) is the corresponding physical

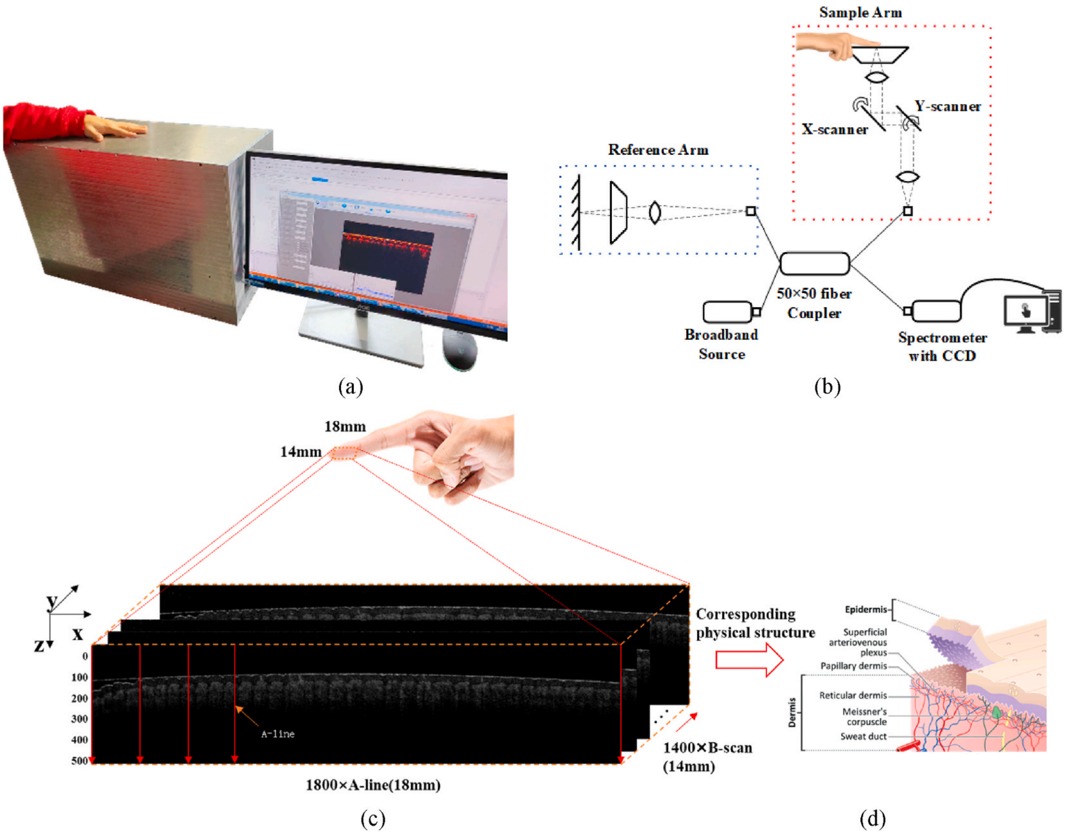

Fig. 2. The OCT system. (a) The experimental platform. (b) A schematic diagram. (c) A schematic diagram of the OCT system for scanning fingertips to obtain the 3D volumetric data. (d) The corresponding physical structure [37].

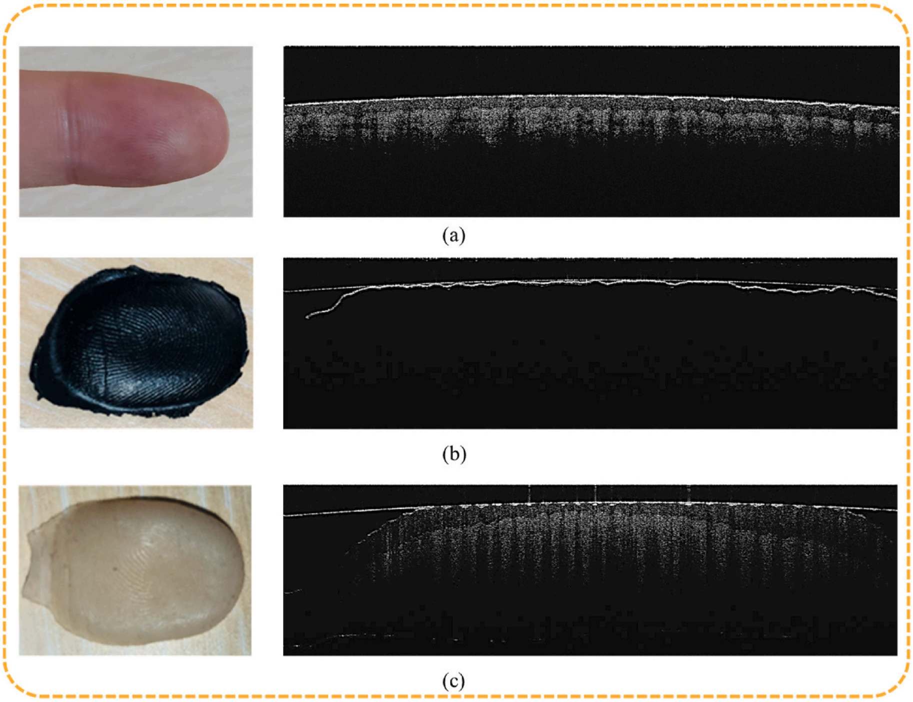

Fig. 3. (a) The bonafide fingerprint and the corresponding B-scan. (b) The artificial fingerprint with external pattern simulated (conductive silicone) and the corresponding B-scan. (c) The artificial fingerprint with the internal structure simulated (a double-layer fingerprint film made in clear silicone + flesh pigmented silicone) and the corresponding B-scan.

structure.

In this study, we developed 21 types of artificial fingerprint materials [38] that can be roughly divided into two categories: external pattern simulation (e.g., clear silicone and conductive silicone) and internal structure simulation (e.g., double-layer and ultrathin fingerprint films). Fig. 3 illustrates typical examples of bonafide and artificial B-scans. In contrast to Fig. 3 (b), (c) is more prone to inducing identification errors in automatic fingerprint identification systems. Consequently, anti-spoofing methods that focus solely on a single cross-sectional image (B-scan) from the volume data can result in misjudgment. Therefore, we introduced an OCT fingerprint anti-spoofing method based on a 3D CNN, enabling us to extract discriminating features more accurately from OCT volume data while preserving spatial continuity.

2.2. *Dataset description*

As shown in Table 1, our self-made OCT fingerprint database is denoted as the Zhejiang University of Technology External and Internal Fingerprint Database (ZJUT-EIFD) [38]. The training and testing datasets used in the proposed method were selected from this database. A total of 160 fingers from 20 individuals aged 22–55 years were used as bonafide samples. Each finger was sampled eight times. Finally, 1280 sets of bonafide OCT volume data were obtained.

For artificial fingers, 21 common materials were used, and each material consisted of five different artificial fingers. Each artificial finger was sampled five times. Finally, 525 sets of artificial OCT volume data were obtained.

Generally, the potential failure of the fingerprint anti-spoofing method can often be attributed to the presence of unknown artificial fingerprint materials. Therefore, the generalization ability of the model is the key to evaluating the anti-spoofing method. We divided the ZJUT-EIFD data into two parts: D_1 for the network performance experiment and D_2 for the anti-spoofing performance experiment. The details are presented in Table 1. In the D_1 *dataset*, we used 24 fingers of three people, resulting in a total of 72 bonafide volume data points for training. Simultaneously, D_1 was divided into three equal parts for three-cross validation. Illustration of some bonafide

Table 1
The data description of ZJUT-EIFD.

Type	Number of Volume Data	D_1	D_2
		Volume Data	Volume Data
Bonafide	20 × 8 × 8 = 1280	3 × 8 × 3 = 72	17 × 8 × 8 = 1088
Artificial	21 × 5 × 5 = 525	12 × 2 × 3 = 72	9 × 5 × 5 = 255

B-scans in D_1 are shown in Fig. 4(a). To balance the positive and negative sample sizes, 24 subjects using 12 materials were utilized, resulting in a total of 72 artificial volume data points for training. In the D_2 *dataset*, we used 1088 bonafide volume data points for the remaining 17 people and 225 artificial volume data points for the remaining nine artificial samples for testing, as shown in Fig. 4(c).

Fig. 4(b)–(d) shows some of the samples used for training and testing. Artificial samples composed of various materials can be classified into two categories: simulated external patterns and simulated internal structures. Unlike the traditional automatic fingerprint recognition system, which is easily deceived by artificial fingerprints of the simulated external patterns, OCT is deceived more by artificial fingerprints of the simulated internal structures because of its ability to collect images of 0–3 mm below the fingertips. Therefore, to better test the generalization performance, the artificial fingerprints used for training in D_1 were all simulated external pattern artificial fingerprints, while those used in D_2 for testing were simulated internal structure artificial fingerprints. The anti-spoofing performance was tested on D_2 using unknown samples.

2.3. The 3D convolutional neural network

ResNet [39,40], based on a 3D CNN combined with a convolutional block attention module (CBAM) [41], was used for fingerprint anti-spoofing collected by the OCT system. As shown in Fig. 5, the network used in this study maintains the network weight of key features, thus enhancing the network focus on the subcutaneous key organization information of fingerprints to achieve the authenticity identification of fingerprints. Most importantly, we used 3D convolution to extract 3D spatial information based on the spatial continuity of OCT fingerprint data for anti-spoofing, instead of using 2D convolution to extract the feature information of a single B-scan as in most studies.

2.4. The proposed method for OCT fingerprint anti-spoofing

Our proposed method utilizes a 3D CNN to extract anti-spoofing features from the volume data of fingertips collected using OCT. This was achieved by leveraging the 3D spatial continuity of the volume data to accurately distinguish bonafide fingerprints from artificial fingerprints. The main anti-spoofing procedures for OCT fingerprints are shown in Fig. 6. Original 3D volume data are collected using the OCT system, and eight ROI patches with a size of 128 × 128 × 128 pixels are randomly extracted. The identification results are then used for anti-spoofing. If a single patch is predicted to be an artificial fingerprint, the volume data are predicted to be artificial fingerprints. Meanwhile, if all patches are predicted to be bonafide fingerprints, the volume data are confirmed to be bonafide fingerprints.

2.4.1. Preprocessing

A significant number of black background regions were present in the OCT-collected fingerprint data. These regions have no impact on OCT fingerprint anti-spoofing and are therefore considered redundant information. To this end, we extracted the ROI from the OCT volume data, which were small patches. Small ROI patches are shown in Fig. 7(a)(b). Small patches can reduce the interference of redundant black background information while preserving important anti-spoofing information, such as the epidermis, sweat glands, and viable epidermis. This approach can effectively enhance the judgment accuracy while reducing the amount of data processing and training time required. In this section, we describe the process of extracting small ROI patches. This is shown in Algorithm 1.

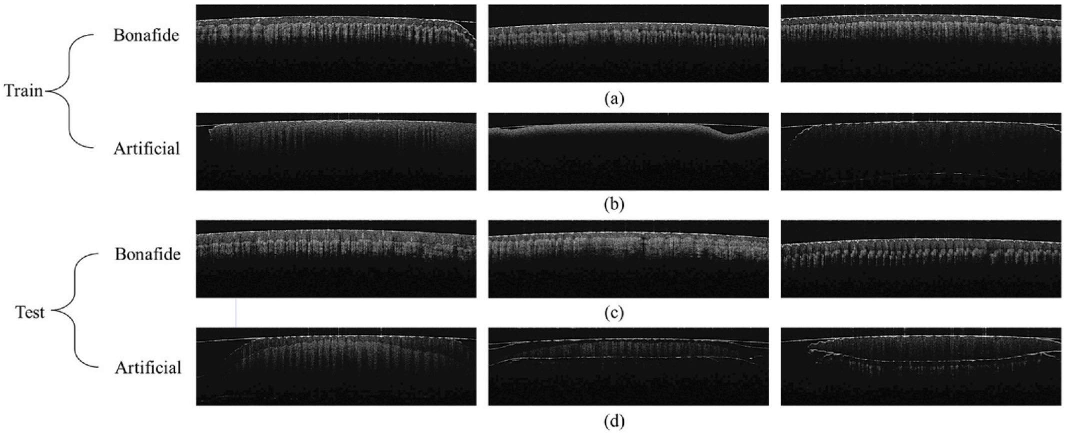

Fig. 4. (a) and (c) illustration of some bonafide B-scans collected from different human fingers for training and testing. (b) illustration of some artificial B-scans collected from artificial fingers with external pattern simulated for training. (d) illustration of some artificial B-scans collected from artificial fingers with internal structure simulated for testing.

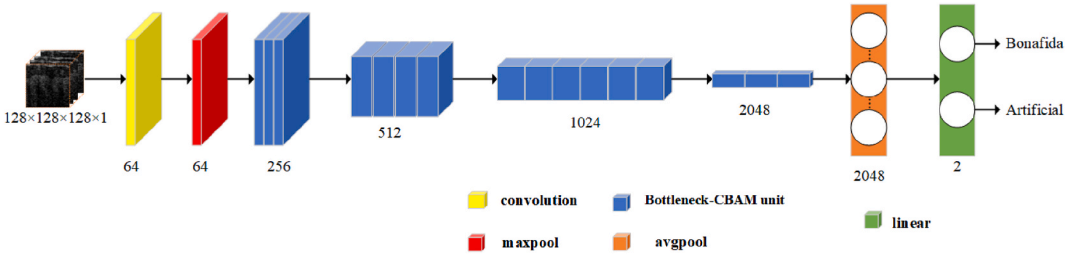

Fig. 5. Network structure diagram.

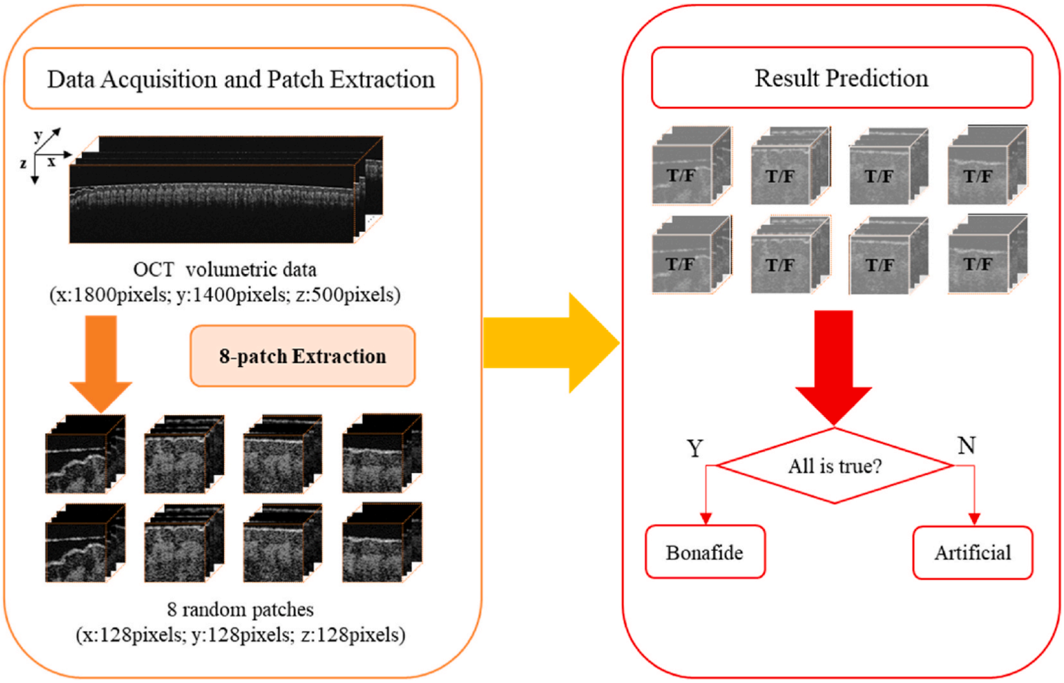
Data Acquisition and Patch Extraction
y
z
x
OCT volumetric data
(x:1800pixels; y:1400pixels; z:500pixels)
8-patch Extraction
8 random patches
(x:128pixels; y:128pixels; z:128pixels)
Result Prediction
T/F
All is true?
Y
N
Bonafide
Artificial

Fig. 6. The flowchart of the proposed anti-spoofing of OCT fingerprint algorithm.

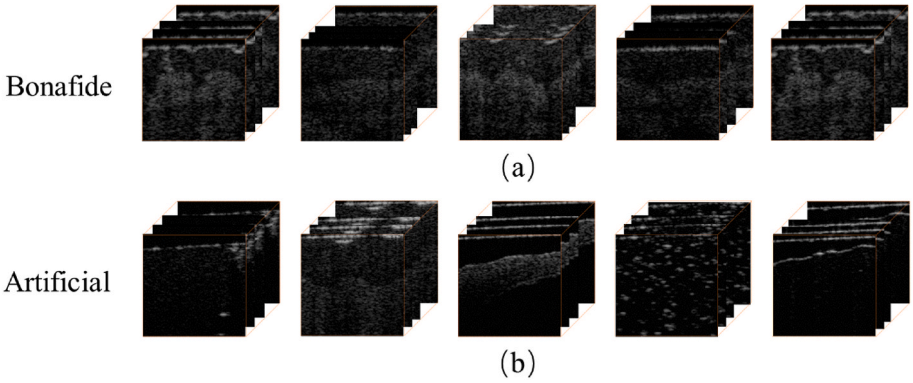

Fig. 7. (a) Illustration of bonafide patches using patch extraction method. (b) Illustration of artificial patches using patch extraction method. (To better demonstrate the operation of image enhancement).

Algorithm 1

Extract the ROI patches of OCT volume data

```
Input: The OCT volume data T with the size of D×L×W (where 'D' represents the Depth of the volume data, 'L' represents the Length of the volume data, and 'W' represents the Width of the volume data, which is the number of B-scans.)
Output: The ROI patches of the OCT volume data
Define constants: the arrays of random numbers A, B, the number of autocorrelation terms $N_{ct}$, the size of patches $D' \times L' \times W'$
for (i = 0, i < L, i++) do
    if (i in A) then
        for (j = $N_{ct}$, j < D, j++) do
            obtain sum: $X_S$ = SUM(T[i][j])
        end for
        obtain the max value of $X_S$: $V_{max}$ = max ($X_S$)
        obtain the index of the $V_{max}$: $H$ = index ($V_{max}$) + $N_{ct}$//$H$ as cover glass line
        if $H$ + D' > D then
            $H$ = D-D'-1//prevent value out of bounds
        end if
        for (m = 0, m < W, m++) do
            if (m in B) then
                obtain the ROI patches: Dataset = [T [i-L': i, H:H + D', j: j + W']]
            end if
        end for
    end if
end for
return Dataset
```

First, we summed the gray values of each row of pixels in the B-scan. It can be clearly found that the gray value of biological tissues in the image acquired by OCT was generally higher than the background area. In the OCT B-scans, the glass layer exhibits the largest pixel. Leveraging this characteristic, we accumulate the pixel values along the length direction to obtain the cumulative values along the depth direction. The maximum value within this accumulation approximates the position of the glass layer, which is denoted as the 'cover glass line' [31,38], where the selection range for x is [Nct, D). The value is not from the position of the 0th pixel because the top rows of the autocorrelation terms are caused by the OCT system itself, which interferes with the selection of the cover glass line. We selected the maximum value from the sum of the values, determined the position of the maximum value, and added the pixels of Nct that were originally deleted. The final position was the intercepted cover glass line. As the difference between the positions of the cover glass line of the adjacent B-scan was small, after obtaining the cover glass line of the B-scan, a small patch was cut into the previous B-scan.

The separation between the cover glass line and critical anti-spoofing data within the OCT images, such as internal fingerprints and sweat glands, did not exceed 100 pixels [42]. To save GPU computer power, we set the size of the cut patches to 128 pixels in length, width, and height. Fig. 7(a)(b) illustrates some bonafide and artificial patches by the patch extraction method.

The skin curvature of the edge area in the B-scan was relatively large. Simultaneously, the middle part of the B-scan, finger, and glass surface were closely fitted, and the subcutaneous information of the finger was complete and clear. Therefore, the method described in this study discarded the 300-pixel areas at each boundary in the x-direction. During the network training, to obtain more small patches, 36 feature patches were cut from each volume of data to ensure that the network had sufficient learning data, to ensure learning on different biological tissue parts, and to fully learn the forged information of different artificial materials. Then, we used eight patches that were extracted randomly in the anti-spoofing performance experiment and the robustness experiment to ensure the performance of the proposed method. We chose eight small ROI patches because they could cover the anti-spoofing information distribution of OCT volume data. The slider windows slid along the cover glass line in the x-direction, 128 B-scans were cut forward in the y-direction at the B-scan position of the cover glass line, and 128-pixel regions were cut downward in the z-direction according to the cover glass line.

2.4.2. Process of anti-spoofing using 3D CNN

Algorithm 2

Process of fingerprint anti-spoofing strategy using 3D CNN

```
Input: The OCT volume data T
Output: Fingerprint data authenticity identification results
V = Algorithm 1(T);
//Step 1. Divide the input OCT volume data T into patches randomly
for (i = 0, i < the length of V, i++) do
    R[i] = model(V[i]);
end for//Step 2. Input the local feature patch into the trained network model and return the prediction result of each patch
for (j = 0, j < the length of V, j++) do
    if (R[j] = = false) then
        $R_S$ = false;
        break;
    end if
end for//Step 3. Judge whether there is false sample in the recognition results
return Rs//Step 4. Output classification results
```

Fig. 6 shows the OCT fingerprint anti-spoofing process proposed in this study and the input is the OCT volume data collected by the OCT system.

Algorithm 2 outlines the specific anti-spoofing process used in the fingerprint anti-spoofing stage. We used OCT collection equipment to obtain fingertip data, from which we randomly selected eight ROI patches as input into our trained network model to generate the prediction scores and results for each patch. Specifically, if the prediction result for any of the eight ROI patches was artificial, the prediction result of the volume data was output as an artificial fingerprint, and the predicted results were then output.

3. Experiments and analysis

This section verifies the superiority of the proposed anti-spoofing method. The network performance experiment was used to assess the classification effect of each patch, while the anti-spoofing performance experiment was used to identify each OCT volume data point after the anti-spoofing process described in Section 2.4.1. Subsequently, we conducted cross-device experiments on the SZU database [43] to demonstrate the robustness of the proposed method by estimating the recognition error rate on this database. Finally, we used Grad-CAM [44] for visualization and analyzed whether our proposed method is based on the location of the OCT volume data for anti-spoofing.

The identification error rate (ERR) is a crucial evaluation metric for fingerprint anti-spoofing because of its quantifiable nature. A lower ERR indicates a higher accuracy in distinguishing genuine from counterfeit products and improving security and trustworthiness. Therefore, it was used as an evaluation metric to assess the efficiency of the proposed method. The ERR is defined as Eq. (1):

$$ERR = \frac{FP + FN}{TP + TN + FP + FN}, \tag{1}$$

where TP (True Positive) refers to the number of positive instances correctly predicted as positive, TN (True Negative) represents the number of negative instances correctly identified as negative, FP (False Positive) denotes the number of negative instances incorrectly classified as positive, and FN (False Negative) indicates the number of positive instances incorrectly predicted as negative. This metric is fundamental for evaluating the performance of binary classifiers and assessing their ability to correctly identify positive and negative samples.

3.1. Network performance experiment

To validate the effectiveness of the proposed method, we conducted a comparison with current mainstream supervised networks, including PreResNet [45], DenseNet [46], WideResNet [47], and R(2 + 1)D [48] Simultaneously, we replicated the state-of-the-art methods, DSResNet [13] and the ResNet used CDCN [14], which are widely used in the field of face recognition, which is similar to fingerprint anti-spoofing. We compared the experimental results to explore whether the algorithms developed for face recognition could deliver exceptional performance in the field of OCT fingerprint anti-spoofing. The proposed patch extraction method was configured for all the supervised networks to make the comparison more intuitive.

The D_1 dataset was used for the experiment. We adopted the three-fold cross-validation method for network training to verify and compare the ERRs of the aforementioned four different network models. The artificial materials in D_1 were divided into three equal groups, with two groups used for training and one group used for verification in each fold. The experimental results are listed in Table 2.

Observing the experimental results of the different network comparisons in Table 2, the experimental results presented in the table were obtained by combining the results from the three-fold cross-validation. From the results, compared to other supervised-based models, our method achieved optimal performance in ERR. With our method, the ERR was 3.57%, which was 0.36% lower than the second-best result. In this experiment, all 3D supervised learning networks achieved relatively good performance when tested on both bonafide and simulated artificial fingerprints, even with limited training data. In future experiments, we will test artificial samples of unknown materials and fingerprint volume data collected from different databases to verify the robustness and generalizability of our method for unknown artificial materials and different OCT acquisition systems. The results were obtained by averaging three-fold cross-validation.

Table 2
Performance comparison of different network (%).

	ERR
PreResNet [45]	4.29
DenseNet [46]	4.44
WideResNet [47]	4.57
R(2 + 1)D [48]	3.93
CDCNResNet [14]	4.36
DSResNet [13]	4,48
Proposed Method	**3.57**

3.2. Anti-spoofing performance experiment

In this section, we compare the proposed method with current mainstream supervised networks configured using the proposed patch extraction method and anti-spoofing process. In addition, we also compared the OCT fingerprint anti-spoofing methods with Chugh's method [30], which is designed for single B-scan anti-spoofing; to ensure a fair comparison, we used an equal number of B-scans for anti-spoofing when applying their approach. The final anti-spoofing score for the 3D volume data was obtained by averaging the individual anti-spoofing scores of all B-scans involved in the prediction process. For this experiment, we used the D_2 dataset, which contained 1088 bonafide volume data points and 255 artificial volume data points. The results of the different methods are listed in Table 3.

As described in Section 3.3, the anti-spoofing process yielded extremely low error rates for all supervised networks that identified artificial samples. In this experiment, CDCNResNet and DSResNet, did not demonstrate significant advantages over the other supervised learning methods. The proposed method was able to identify all the artificial samples in this experiment and achieved the best recognition results, with a recognition error rate of 0.92% on the bonafide fingerprint database, thus verifying its superior anti-spoofing performance.

Based on the experimental results, it is evident that the performance of the supervised networks that use 3D convolution is considerably superior to that of Chugh's method, which relies on 2D convolution. This is because OCT volume data possess spatial continuity, and 3D convolution can extract more anti-spoofing features than 2D convolution. Furthermore, the materials used to generate artificial fingerprints are highly complex, resulting in poor performance when identifying fingerprints. Additionally, compared with other supervised networks, the proposed approach demonstrated a better ability to proactively focus on internal structural differences with limited training sets, comprising only two individuals for bonafide fingerprints and eight materials with only simulated external patterns. Therefore, the proposed method can ensure the generalizability of artificial samples.

3.3. Runtime experiment

We measured the runtime of the CNN using a machine equipped with an Intel Xeon Gold 5218 R 2.10GHZ and an NVIDIA GeForce RTX 3090.

Regarding OCT fingerprint anti-spoofing, it is essential to consider the entire 3D volume data captured by the OCT system because focusing solely on a single B-scan cross-section would not yield meaningful results. For this experiment, all time performance calculations were based on the OCT volume data. The approach for evaluating the time performance involved predicting 20 sets of OCT volume data consecutively, recording the time overhead for each set, and subsequently computing the average value.

To better demonstrate the time overhead for each method, we present the prediction time for each method using the volume data in two segments. These two segments are the processing and prediction runtimes.

As shown in Table 4, for the 3D CNN-based supervised learning methods, the preprocessing methods used are the same, so their prediction runtime, total runtime and params are recorded in the table. Processing runtime is about 1.643 s. For prediction, excluding PreResNet, which boasts a simpler network structure resulting in notably lower prediction times compared to other methods, the remaining approaches did not exhibit substantial advantages or disadvantages in terms of time performance. Nevertheless, in terms of the total runtime, these methods did not demonstrate particularly noticeable advantages or disadvantages in terms of temporal performance. In terms of parameter count, DenseNet has significantly fewer parameters compared to other methods due to its densely connected architecture. WideResNet increases the parameter count by widening the network's width while keeping its depth unchanged. Meanwhile, DsResNet employs a dual-stream branching structure, resulting in roughly twice the parameter count of a single-branch structure. The proposed method slightly increases in parameter count due to the incorporation of attention modules.

To enhance the accuracy and persuasiveness of our experimental results, we computed the time overhead of Chugh's method [30]. During the data preprocessing phase, the time overhead was 368.162 s, whereas the network prediction phase took 214.364 s. These runtimes significantly exceeded those of the 3D CNN-based supervised learning method. The reason for the high processing time overhead was the Non-Local Means denoising of the B-scan. The extra time overhead added when using Non-Local Means denoising for a single B-scan might not be very obvious; however, in the ZJUT-EIFD [38] database, the OCT fingerprint volume data had 1400 B-scans, resulting in a significant increase in time overhead. This explains why the processing time overhead of Chugh's method was much higher than that of other methods. During the prediction runtime, their method used no more than 60 ROI patches to predict a single B-scan. As a result, when predicting the authenticity of the entire data volume, the number of ROI images used for prediction

Table 3
The Anti-spoofing performance of different methods (%).

Methods	ERR of Bonafide	ERR of Artificial
PreResNet [45]	3.89	0.44
DenseNet [46]	4.81	0.29
WideResNet [47]	1.41	0.59
R(2 + 1)D [48]	2.27	0.89
CDCNResNet [14]	3.52	0.74
DSResNet [13]	3,56	0.44
Chugh's method [30]	3.31	5.34
Proposed method	**0.92**	**0**

Table 4
Runtime of each method.

Method	Runtime of prediction (s)	Total Runtime (s)	Params (M)
PreResNet [45]	**0.125**	**1.768**	46.97
DenseNet [46]	0.215	1.858	**11.24**
WideResNet [47]	0.269	1.912	158.25
R(2 + 1)D [48]	0.213	1.856	47.02
CDCNResNet [14]	0.191	1.834	46.97
DSResNet [13]	0.195	1.838	93.11
Proposed method	0.217	1.86	49.50

exceeded 10,000. By contrast, the proposed method only needs to validate eight ROI patches, making the time overhead of Chugh's method in the network prediction phase significantly higher than that of other methods.

3.4. Cross-device research

The OCT data distribution can vary significantly because of the use of different OCT acquisition devices and methods. Consequently, robustness and device generalizability are major challenges facing OCT fingerprint anti-spoofing methods.

To assess the robustness and device generalizability of the proposed anti-spoofing method, it is necessary to evaluate its performance on fingerprints collected using different OCT devices. To this end, we used the SZU database [34], which contains 1807 bonafide volume data collected by an 840 nm spectral domain OCT system, with an OCT data volume of $500 \times 1500 \times 400$ pixels, which is different from our database and has large differences in image distribution. For our OCT system, we collected 1400 B-scans in volume data, which correspond to an actual finger length of 14 mm. Liu's database only has 400 B-scans for volume data, but corresponds to an actual finger length of 15 mm. Our dataset features a significantly higher scanning density in the B-scan direction than Liu's dataset. This allowed us to determine whether the proposed method was affected by imaging differences between different OCT devices.

The results presented in Table 5 demonstrate that the proposed fingerprint anti-spoofing method achieved the best experimental results. In addition to the proposed method, the performance of other methods on the SZU database significantly decreased. As shown in Fig. 2(a), the SZU database contains only 400 B-scans in the y-direction, while our database comprises 1400 B-scans in the same direction. The density of the data collected from fingertips in our database was higher than that in the SZU database. This is evident in the 3D space, where the 3D information of patches with the same size was notably different, thereby resulting in a significantly inferior performance of other networks on the SZU database. From Table 5, the two methods applied to face recognition did not show good robustness on this task. Furthermore, the proposed method achieved an error rate of 6.40% on the SZU database, which was 4.99% lower than the lowest result. This indicates that the proposed method can accurately distinguish bonafide fingerprints from those collected by unknown OCT devices. Additionally, the results demonstrate the generalizability and robustness of the proposed anti-spoofing method.

3.5. Visualization

As shown in Fig. 8, we used Grad-CAM [37] to draw heat maps of the network results to visualize the network learning characteristics of the results and more intuitively display the information features on which the network prediction labels depend.

Fig. 8 (a) shows a partial cross-sectional view of a bonafide fingerprint block. The network focus on the internal fingerprint information and sweat glands under the skin for bonafide fingerprints. The network pays more attention to this internal information than to the epidermal information near the glass layer. Simultaneously, along the y-direction (which is the direction of generating the B-scan), the network continues to focus on the internal structure of the bonafide fingerprints. From the perspective of the entire volume of data, the content on which the network focuses exhibits spatial continuity. For the partial cross-sectional images of the artificial fingerprints with the external pattern simulation blocks shown in Fig. 8 (b), the network focuses more on the artificial fingerprint information through the glass layer, which is completely different from the bonafide fingerprint. For the artificial patch with the simulated internal structure, as shown in Fig. 8 (c), the network also focuses on the internal artificial structure. However, the spatial continuity of artificial fingerprints with simulated internal structures cannot support the network in focusing on their internal structures from the perspective of the entire 3D volume data. Therefore, the network distinguishes artificial fingerprints with simulated internal structures from bonafide fingerprints.

Our method enhances the ability to extract internal features of OCT volume data and automatically identifies the differences in the internal structure of bonafide and fake fingerprints, thereby greatly improving the effect of fingerprint anti-spoofing.

4. Conclusion

This paper presented an OCT fingerprint anti-spoofing method based on a 3D CNN. Using a 3D CNN, this approach effectively focuses on key structural information and 3D spatial continuity from the volume data of the fingertips collected by OCT. Compared with existing methods, our proposed method demonstrated the best network performance and anti-spoofing capabilities, exhibiting high distinguishability for both bonafide and artificial fingerprint samples. Furthermore, we thoroughly explored the anti-spoofing

Table 5
Error rates for Cross-device research (%).

Methods	ERR
PreResNet [45]	11.03
DenseNet [46]	23.96
WideResNet [47]	12.11
R(2 + 1)D [48]	18.78
CDCNResNet [14]	18,58
DSResNet [13]	21.20
Proposed method	**6.40**

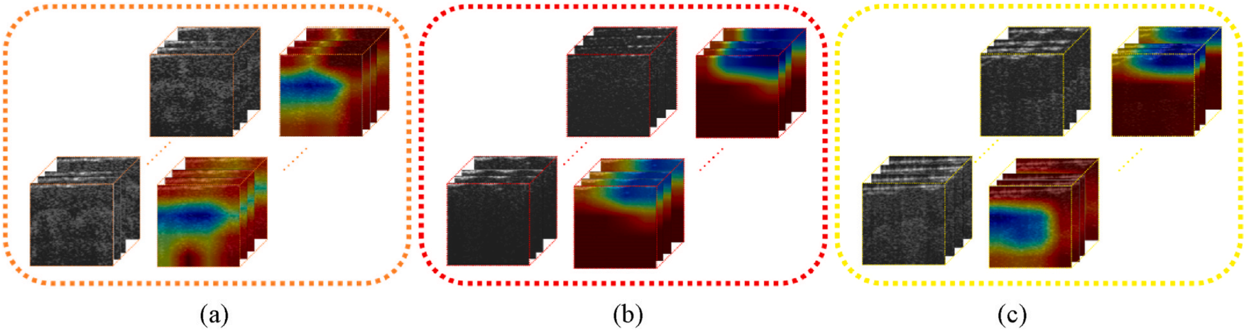

Fig. 8. Gard-CAM [37]: Visualization of the patches. (a) Heat maps of the cross-sectional images of the bonafide patches. (b) Heat maps of the cross-sectional images of artificial patches with simulated external patterns. (c) Heat maps of cross-sectional images of artificial patches with simulated internal structures (cold-tone areas are the focus of the neural networks).

performances of various OCT acquisition devices. The results highlighted the excellent generalizability and robustness of the proposed method for different OCT equipment setups, thus demonstrating its effectiveness in addressing security concerns related to fingerprint authentication.

Funding

This work was supported in part by the Natural Science Foundation of China under Grants 62276236 and 61976189 and the Zhejiang Provincial Leading Innovation and Entrepreneurship Team under Grant 2021R01002.

Data availability statement

Data associated with this study has been deposited at A portion of the data is shown in the: https://github.com/ZJUT-ERCISS-home/ZJUT-EIFD, if readers need to obtain data, please call for protocol and get the complete database via email hxwang@zjut.edu.cn.

Author contribution statement

Yilong Zhang: Conceived and designed the experiments; Wrote the paper. Shichang Yu: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper. Shiliang Pu; Yingyu Wang; Kanlei Wang: Contributed reagents, materials, analysis tools or data. Haohao Sun; Haixia Wang: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D. Bhattacharyya, R. Ranjan, F. Alisherov, M. Choi, Biometric authentication: a review, Int. J. u-e Serv. Sci. Technol. 2 (2009) 13–28.
[2] D. Maltoni, R. Cappelli, Advances in fingerprint modeling, Image Vis Comput. 27 (2009) 258–268.
[3] S.S. Ali, V.S. Baghel, I.I. Ganapathi, S. Prakash, Robust biometric authentication system with a secure user template, Image Vis Comput. 104 (2020), 104004.
[4] S. Meissner, R. Breithaupt, E. Koch, Defense of fake fingerprint attacks using a swept source laser optical coherence tomography setup, in: Frontiers in Ultrafast Optics: Biomed. Sci. Ind. Appl., 2013, pp. 49–52. XIII(SPIE.

[5] T. Matsumoto, H. Matsumoto, K. Yamada, S. Hoshino, Impact of Artificial" Gummy" Fingers on Fingerprint Systems, SPIE, 2002, pp. 275–289.

[6] S. Marcel, M.S. Nixon, J. Fierrez, N. Evans, Handbook of Biometric Anti-spoofing: Presentation Attack Detection, Springer, 2019.

[7] J.J. Engelsma, S.S. Arora, A.K. Jain, N.G. Paulter, Universal 3D wearable fingerprint targets: advancing fingerprint reader evaluations, IEEE Trans. Inf. Forensics Secur. 13 (2018) 1564–1578.

[8] C. Hong, J. Yu, J. Wan, D. Tao, M. Wang, Multimodal deep autoencoder for human pose recovery, IEEE Trans. Image Process. 24 (2015) 5659–5670.

[9] C. Hong, J. Yu, J. Zhang, X. Jin, K.-H. Lee, Multimodal face-pose estimation with multitask manifold deep learning, IEEE Trans. Ind. Inf. 15 (2018) 3952–3961.

[10] C. Hong, J. Yu, D. Tao, M. Wang, Image-based three-dimensional human pose recovery by multiview locality-sensitive sparse retrieval, IEEE Trans. Ind. Electron. 62 (2014) 3742–3751.

[11] J. Yu, M. Tan, H. Zhang, Y. Rui, D. Tao, Hierarchical deep click feature prediction for fine-grained image recognition, IEEE Trans. Pattern Anal. Mach. Intell. 44 (2019) 563–578.

[12] J. Yu, D. Tao, M. Wang, Y. Rui, Learning to rank using user clicks and visual features for image retrieval, IEEE Trans. Cybern. 45 (2014) 767–779.

[13] S. Chen, T. Yao, K. Zhang, Y. Chen, K. Sun, S. Ding, J. Li, F. Huang, R. Ji, A dual-stream framework for 3D mask face presentation attack detection, ICCV (2021) 834–841.

[14] Z. Yu, C. Zhao, Z. Wang, Y. Qin, Z. Su, X. Li, F. Zhou, G. Zhao, Searching central difference convolutional networks for face anti-spoofing, CVPR (2020) 5295–5305.

[15] E. Park, W. Kim, Q. Li, J. Kim, H. Kim, Fingerprint liveness detection using CNN features of random sample patches, BIOS (2016) 1–4.

[16] C. Wang, K. Li, Z. Wu, Q. Zhao, A DCNN based fingerprint liveness detection algorithm with voting strategy, CCBR (Springer (2015) 241–249.

[17] D.M. Uliyan, S. Sadeghi, H.A. Jalab, Anti-spoofing method for fingerprint recognition using patch based deep learning machine, Eng. Sci. Technol. Int. J. 23 (2020) 264–273.

[18] B.U. Maheswari, M. Rajakumar, J. Ramya, Dynamic differential annealing-based anti-spoofing model for fingerprint detection using CNN, Neural Comput. Appl. 34 (2022) 8617–8633.

[19] F. Liu, Z. Kong, H. Liu, W. Zhang, L. Shen, Fingerprint presentation attack detection by channel-wise feature denoising, IEEE Trans. Inf. Forensics Secur. 17 (2022) 2963–2976.

[20] D. Baldisserra, A. Franco, D. Maio, D. Maltoni, Fake fingerprint detection by odor analysis, ICB (Springer (2005) 265–272.

[21] M. Drahansky, R. Notzel, W. Funk, Liveness detection based on fine movements of the fingertip surface, WIA (2006) 42–47.

[22] P.V. Reddy, A. Kumar, S. Rahman, T.S. Mundra, A new antispoofing approach for biometric devices, IEEE Trans. Biomed. Circuits Syst. 2 (2008) 328–337.

[23] W.-Y. Yau, H.-L. Tran, E.-K. Teoh, Fake finger detection using an electrotactile display system, ICARCV (2008) 962–966.

[24] P.J. Rosenfeld, Chapter 3 - Optical Coherence Tomography, Retina, 2013.

[25] C.K. Hitzenberger, Optical coherence tomography in optics express, Opt Express 26 (2018) 24240–24259.

[26] D. Huang, E.A. Swanson, C.P. Lin, J.S. Schuman, W.G. Stinson, W. Chang, M.R. Hee, T. Flotte, K. Gregory, C.A. Puliafito, Optical coherence tomography, Science 254 (1991) 1178–1181.

[27] C. Bowd, R.N. Weinreb, J.M. Williams, L.M. Zangwill, The retinal nerve fiber layer thickness in ocular hypertensive, normal, and glaucomatous eyes with optical coherence tomography, Arch. Ophthalmol. 118 (2000) 22–26.

[28] G. Cennamo, M. Reibaldi, D. Montorio, L. D`Andrea, M. Fallico, M. Triassi, Optical coherence tomography angiography features in post-COVID-19 pneumonia patients: a pilot study, Am. J. Ophthalmol. 227 (2021) 182–190.

[29] A.G. Podoleanu, J.A. Rogers, D.A. Jackson, S. Dunne, Three dimensional OCT images from retina and skin, Opt Express 7 (2000) 292–298.

[30] T. Chugh, A.K. Jain, OCT Fingerprints: Resilience to Presentation Attacks, 2019 arXiv preprint arXiv:1908.00102.

[31] F. Liu, G. Liu, X. Wang, High-accurate and robust fingerprint anti-spoofing system using optical coherence tomography, Expert Syst. Appl. 130 (2019) 31–44.

[32] F. Liu, H. Liu, W. Zhang, G. Liu, L. Shen, One-class fingerprint presentation attack detection using auto-encoder network, IEEE Trans. Image Process. 30 (2021) 2394–2407.

[33] W. Zhang, H. Liu, F. Liu, Fingerprint Presentation Attack Detection by Learning in Frequency Domain, PRML, 2021, pp. 183–189.

[34] S. Maetschke, B. Antony, H. Ishikawa, G. Wollstein, J. Schuman, R. Garnavi, A feature agnostic approach for glaucoma detection in OCT volumes, PLoS One 14 (2019), e0219126.

[35] H. Wang, X. Yang, P. Chen, B. Ding, R. Liang, Y. Liu, Acquisition and extraction of surface and internal fingerprints from optical coherence tomography through 3D fully convolutional network, Optik 205 (2020), 164176.

[36] H. Sun, Y. Zhang, P. Chen, H. Wang, Z. Guo, Y.-H. He, R. Liang, Synchronous fingerprint acquisition system based on total internal reflection and optical coherence tomography, IEEE Trans. Instrum. Meas. 69 (2020) 8452–8465.

[37] Madhero88, Layers of the skin. https://en.wikipedia.org/wiki/File:Skin_layers.png.

[38] Haohao Sun, Peng Chen, Haixia Wang, Yipeng Liu, Ronghua Liang, A new approach in automated fingerprint presentation attack detection using optical coherence tomography, IEEE Trans. Inf. Forensics Secur. (2023).

[39] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, CVPR (2016) 770–778.

[40] K. Hara, H. Kataoka, Y. Satoh, Learning spatio-temporal features with 3d residual networks for action recognition, ICCV (2017) 3154–3160.

[41] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, Cbam: Convolutional Block Attention Module, ECCV, 2018, pp. 3–19.

[42] Y. Zhang, X. Li, H. Wang, R. Wang, P. Chen, R. Liang, Sweat gland extraction from optical coherence tomography using convolutional neural network, IEEE Trans. Instrum. Meas. (2022).

[43] F. Liu, C. Shen, H. Liu, G. Liu, Y. Liu, Z. Guo, L. Wang, A flexible touch-based fingerprint acquisition device and a benchmark database using optical coherence tomography, IEEE Trans. Instrum. Meas. 69 (2020) 6518–6529.

[44] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: visual explanations from deep networks via gradient-based localization, ICCV (2017) 618–626.

[45] K. He, X. Zhang, S. Ren, J. Sun, Identity Mappings in Deep Residual Networks, ECCV (Springer, 2016, pp. 630–645.

[46] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, CVPR (2017) 4700–4708.

[47] S. Zagoruyko, N. Komodakis, Wide residual networks (2016) arXiv preprint arXiv:1605.07146.

[48] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, M. Paluri, A closer look at spatiotemporal convolutions for action recognition, CVPR (2018) 6450–6459.