# A Universal Anti-Spoofing Approach for Contactless Fingerprint Biometric Systems

Anonymous IJCB 2023 submission

## Abstract

*With the increasing integration of smartphones into our daily lives, fingerphotos are becoming a potential contactless authentication method. While it offers convenience, it is also more vulnerable to spoofing using various presentation attack instruments (PAI). The contactless fingerprint is an emerging biometric authentication but has not yet been heavily investigated for anti-spoofing. While existing anti-spoofing approaches demonstrated fair results, they have encountered challenges in terms of universality and scalability to detect any unseen/unknown spoofed samples. To address this issue, we propose a universal presentation attack detection method for contactless fingerprints, despite having limited knowledge of presentation attack samples. We generated synthetic contactless fingerprints using StyleGAN from live finger photos and integrating them to train a semi-supervised ResNet-18 model. A novel joint loss function, combining the Arcface and Center loss, is introduced with a regularization to balance between the two loss functions and minimize the variations within the live samples while enhancing the inter-class variations between the deepfake and live samples. We also conducted a comprehensive comparison of different regularizations' impact on the joint loss function for presentation attack detection (PAD) and explored the performance of a modified ResNet-18 architecture with different activation functions (i.e., leaky ReLU and RelU) in conjunction with Arcface and center loss. Finally, we evaluate the performance of the model using unseen types of spoof attacks and live data. Our proposed method achieves a Bona Fide Classification Error Rate (BPCER) of 0.12%, an Attack Presentation Classification Error Rate (APCER) of 0.63%, and an Average Classification Error Rate (ACER) of 0.37%.*

## 1. Introduction

Biometric systems have been used in wide range of applications such as law enforcement and forensics, individual identification, healthcare, and access control for smart phone and tablet which increase convenience in our daily life. Some of the traditional contact-based biometric systems such as fingerprints and palm prints required the physical touch of the user to a sensor, which then increases user concern about hygiene and public shared devices. Moreover, aside from hygiene-related issues, the elasticity of human skin can lead to shape and detail changes in captured touch base biometric when direct contact is made with scanner [14, 3]. Due to this concern and the recent advancement in sensors and cameras, contactless biometrics have gained major popularity for commercial use. Contactless biometric systems also offer high speed authentication/identification since it does not require physical contact. For instance, smartphones can be used to capture finger photos to be used for biometric authentication. However, obtaining high quality image to extract minutia of fingerprints is challenging compared to contact based sensors due to low contrast of ridge and valley patterns [16, 7]. Furthermore, although contactless biometrics provide a number of advantages as mentioned earlier, they are also more susceptible to deepfake and spoofing attack [5]. For instance, facial recognition systems are vulnerable to deepfake such as masking, or contactless fingerprint are susceptible to photopaper or synthetic.

The presence of various types of attacks, as well as the emergence of unpredictable attack techniques, highlights the need for generalization in presentation attack detection (PAD) systems to effectively detect unseen types of attacks [11]. Although face biometric anti-spoofing has been extensively studied in the literature, there is currently a lack of research regarding the study of presentation attack detection on contactless fingerprint biometric systems. Furthermore, the current research on PAD heavily relies on supervised learning techniques, where both genuine and spoofed samples are used during the model training – exhibits poor performance against unseen attacks. To address that limitation, we present a semi-supervised learning model that utilizes a residual network (ResNet18) with training of only live and synthetic spoofed samples. The main contributions of this paper are as follows:


Figure 1. Different spoofed samples from CLARKSON and COLFISPOOD dataset.

- We proposed a universal presentation attack detection mechanism for contactless fingerprints based on limited knowledge of presentation attack samples. To that end, we generated synthetic contactless fingerptints from live samples using StyleGAN to train a semi-supervised RestNet18. The model is trained on genuine data along with only synthetic spoof attacks.
- We introduced a novel joint loss function by combining the Arcface and Center Loss functions along with regularization to balance between two loss functions. By employing the joint loss function, we aim to minimize the variations within the live samples, while simultaneously enlarging the inter-class variations between deepfake and live samples.
- We conducted a comprehensive comparison of various regularizations' impact on the joint loss function for PAD. Additionally, we evaluated and demonstrated the results of using a modified ResNet-18 architecture with different activation functions, such as leaky ReLU and ReLU, in combination with Arcface and Center loss. Our findings shed light on the effectiveness of these combinations and their performance in the context of anti-spoofing.
- To evaluate our universal PAD, we have stressed out the proposed model with unseen spoofed samples. The model is tested under two public PAI data adopted from COLFISPOOF and CLARKSON. The BPCER and APCER are used as the standard metrics to demonstrate the effectiveness of our proposed technique. The results show that we were able to achieve 0.12% of BPCER, 0.63% APCER on various types of spoofed samples.
- We also conducted a comprehensive comparison with state-of-the-art techniques in terms of various scenarios such as evaluation metrics, number of subjects, and detection of unseen attacks under two public datasets. Our proposed work achieved a remarkable 99% improvement in BPCER at an APCER of 10% compared to the results in [6]. Furthermore, when compared to the findings in [13], our model exhibited substantial improvements of 69.58% for Photopaper, 1.55% for Playdoh, and 0.94% for the synthetic sample

The paper is structured as follows: In section 2 we provide an overview of the previous research on contactless fingerprint anti-spoofing detection. In section 3, we present our deep learning architecture used in our experiments. We will demonstrate the database, and experimental set up such as database, evaluation metrics in Section 4. We present and analyze the results obtained from our experiments, comparing them with several approaches found in the literature in Section 4. Finally, we conclude our paper in Section 5.

## 2. Related Work

In this section, we present a comprehensive overview of recent research in the field of contactless fingerprint anti-spoofing. We summarize key aspects of each study, including the employed methods (handcrafted features or deep learning-based approaches), the number of subjects involved, the type of database used (public or private), the variety of presentation attack instruments considered, the evaluation metrics utilized (e.g., BPCER and APCER), and the scalability of the proposed methods. Earlier work on PAD in contactless fingerprint system was based on handcrafted techniques. These approaches were proposed in [15, 17, 19] involved extracting feature vectors from fingerphotos such as local binary patterns (LBP), dense scale invariant feature transform (DSIFT), and locally uniform comparison image descriptor (LUCID), histogram of oriented gradients (HOG), binarized statistical image features (BSIF) and making decisions based on binary classification using support vector machine (SVM). The evaluation results of Stein et al. [15] reported a 77% success rate of attack detection with 37 subjects. Taneja et al. [17] also obtained 3.71% EER based on photo printed attacks. Wasnik et al. [19] also achieved D-EER of 4.43% based on 50 subjects using three types of attack samples. Previous works relying

| Author | Year | Method | Database | device | Spoof type | Metrics | Results |
|---|---|---|---|---|---|---|---|
| Tanej et al. [17] | 2016 | Hand crafted | IIITD:<br>class: 128<br>images: 5100 | iphone 5 | Print Attack<br>Photo Attack | EER<br>TAR<br>FAR | EER = 3.71% |
| Wasnik et al. [19] | 2018 | Hand crafted LBP,<br>BSIF, HOG, SVM | subjects: 50<br>images: 250<br>videos: 150 | IOS, iphone,<br>ipad | print artefact<br>electronic replay<br>elctronic dispaky | APCER<br>BPCER | BPCER = 1.8, 0, 0.66,<br>APCER = 10 |
| Fujito et al. [2] | 2018 | AlexNet | Live: 4096<br>spoofe sample: 8192 | ios, android,<br>windows | Print Attack<br>Photo Attack | HTER | HTER = 0.04% |
| Marasco et al. [9], [10] | 2022 | AlexNet DenseNet201,<br>ResNet18,DenseNet121,<br>ResNet34, MobileNEt-V2 | IIITD | Apple iphone 5,<br>Flash off,<br>8MP resolution | Print AttackPhoto Attack | DEER | D-EER_AlexNet = 2.14<br>D-EER_ResNet = 0.97% |
| Kolberg et al. [5] | 2023 | Not Reported | COLFISPOOF:<br>7200 spoof samples<br>72 different PAI | Not Reported | Knetosil, Mould glue,<br>latex, silly putty,<br>paper printout, s<br>chool glue,<br>dragonskin,<br>ecoflex, gelatin,<br>glue, modelling clau,<br>playdoh | Not Reported | Not Reported |
| Purnapatra et al. [13] | 2023 | DenseNet 121,<br>NASNet | 35 subjects with 12 devices<br>attack sample: 7548<br>synthetic: 10000 | ios, andrios | ecoflex, playdoh, wood glue,<br>synthetic, fingerphoto, latex | APCER<br>BPCER | APCER = 0.14%<br>BPCER = 0.18% |
| Hailin Li et al. [6] | 2023 | AlexNet, DenseNet201,<br>MobileNet-V2, NASNet,<br>ResNet50, GoogleNet,<br>EfficientNet-B0 and<br>Vision Transformers | 5886 bonafide and 4247<br>attack samples with 4 PAIs | ios, android | ecoflex, playdoh, wood glue,<br>fingerphoto, latex | APCER<br>BPCER | EER = 8.26% |

Table 1. Summary of state-of-the-art approaches for contactless fingerprint anti-spoofing. HOG– histogram of oriented gradients (HOG), SVM– support vector machine, LBP–local binary patterns, BSIF–binarized statistical image features, EER – equal error rate, TAR – true acceptance rate, FAR – false acceptance rate BPCER–bonafide presentation classification error rate, HTER – half total error rate, APCER–attack presentation classification error rate.

on handcrafted features were not robust against state-of-the-art presentation attack scenarios, including those generated by synthetic models from a wide range of adversarial techniques. Additionally, these studies often used a limited number of attack samples and failed to demonstrate their evaluation metrics, such as BPCER and APCER.

The next-generation techniques were deployed based on deep learning approaches in which the model was trained on both live and spoofed samples. Fujio et al. [2] were pioneers in exploring the use of deep neural networks, specifically "Alexnet," for contactless fingerprint anti-spoofing. They trained the model using a combination of bonafide samples and photo-printed attack samples. Their dataset included 9,192 spoofed samples and 4,096 genuine samples, and they achieved an impressive half-error rate of 0.04%. Marasco et al. [9] utilized ResNet and AlexNet architecture based on the IIITD Spoofed Finger photo Database. The database contains 2,048 print attacks, 6,144 photo attacks, and 4,096 live samples. They achieved a D-EER of 2.14% for AlexNet and .07% for ResNet, respectively. In subsequent research by Marasco et al. [10], there was a slight improvement in the results compared to the baseline approach. However, it should be noted that the ResNet architecture used in their model was trained on both live and spoofed images, which may not be practical or representative of real-world scenarios. Recently, Purnapatra et al.[13] proposed DenseNet-121 and NasNetMobile models with new public database which includes 35 subjects with 65,972 images which includes 29,204 live samples. Their model achieved 88.3% APCER and 0.48% BPCER. Furthermore, Hailin Li et al. [6] demonstrated PAD using various models such as AlexNet, DenseNet-201, MobileNet-V2, NASNet, RestNet50, and Vision transformer. They have integrated 5,886 live samples and 4,247 spoofed samples and obtained an EER of 8.6% on RestNest50. Nevertheless, while recent models demonstrated fair performance on spoofed images during training, their generalizability is limited, leading to poor performance when tested on unseen spoofed images.

## 3. Proposed Method

As illustrated in Figure 2 we applied supervised learning on two fingerprint classes, live and synthetic. In order to increase the resolution and quality of dataset, we applied enhanced super-resolution generative adversarial networks (ESRGAN) on dataset. Additionally, we created synthetic attacks by applying StyleGAN2 with adaptive discriminator augmentation (ADA) [4] on the live dataset. Furthermore, we employed a pretrained ResNet18 architecture, making some modifications (details of architecture is described in Section 3.2). To improve presentation attack detection success rate, we introduced a novel loss function which is a combination of the Arcface loss and Center loss. Our model successfully classified live data samples from unseen spoof type of attacks. In the rest of this section, we will provide a comprehensive explanation of our method and new joint loss function. This method is specifically designed to address contactless fingerprint anti-spoofing.

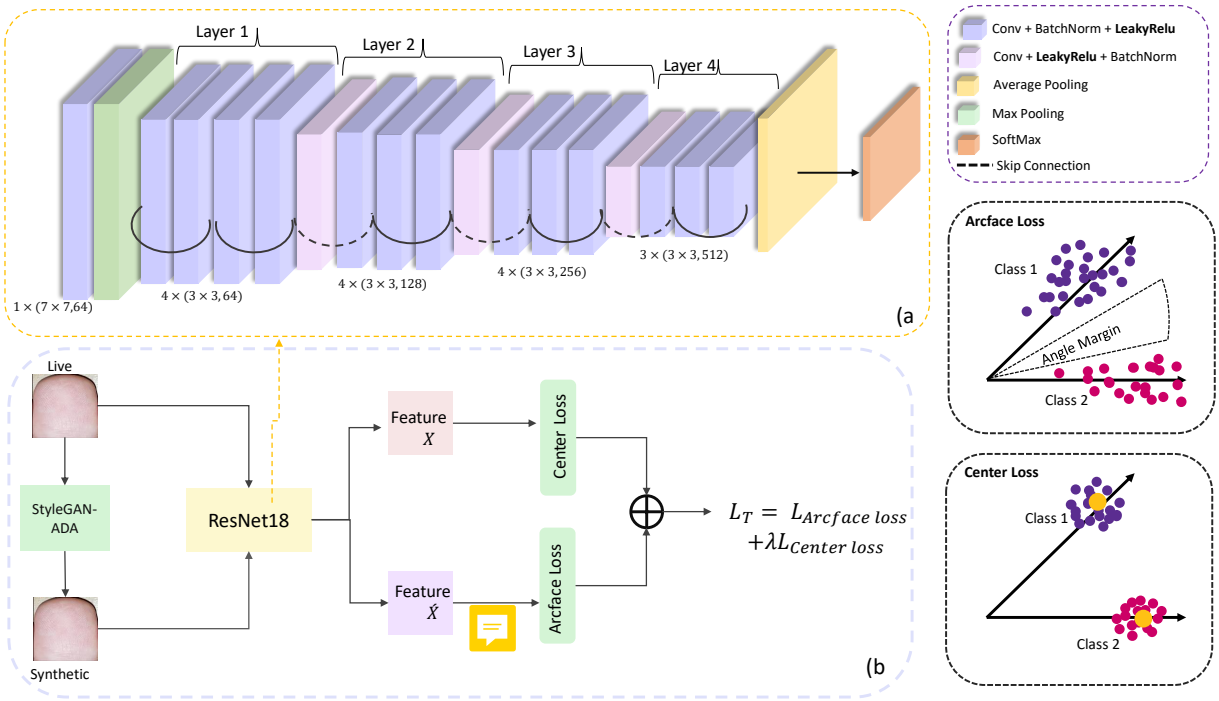


Figure 2. Resnet-18 with Leaky-Relu activation function and using a joint loss function a combination of the Arcface and center loss functions. The model is trained on live samples and synthetic spoofed fingerphotos and tested under various spoofed images such as photo paper, Ecoflex, playdoh, woodglue, and etc.

### 3.1. Joint Loss Function

In this project, we propose a novel approach for contactless fingerprint anti-spoofing using a joint loss function that combines the Arcface Loss [1] and Center Loss [20]. Since our approach focouses on universal deepfake detection, utilizing joint loss function will help us to obtain lower error detection rate. By employing the joint loss function, we aim to minimizing the variations within live samples, while simultaneously enlarging the inter-class variation between deep-fake and live samples. Equation 1 demonstrates how our proposed joint function has been calculated by combining Arcfac and center loss. We used the Arcface loss to map the input data (live and synthetic) to angular space and used center loss to minimizing the variations within each class and pulling together the embeddings of samples belonging to the same class in the feature space:

$$L_{joint} = -log\frac{e^{xcos(\theta_{y_i}+m)}}{e^{xcos(\theta_{y_i}+m)} + \sum_{J=1,J\neq y_i}^{N} e^{xcos\theta_{y_j}}} + \lambda\frac{1}{2}\sum_{i=1}^{m}\|x_i - c_{y_i}\|_2^2,$$

$$L_{joint} = L_a + \lambda L_c, \tag{1}$$

where $L_a$ is the Arcface loss, $L_c$ is the Center loss, and $\lambda$ is a regularization parameter which is used to balance the two loss functions. We achieve the best value for $\lambda$ during network training.

### 3.2. ResNet

We used ResNet-18 as a deep convolutional Neural Network architecture by applying some modifications to the architecture. We replaced the RELU activation function with Leaky-Relu. The ReLU activation function can encounter a problem called "Dead Neuron" [21]. Where neurons become inactive for negative inputs, resulting in unchanging weights during training. To tackle this problem, we replace the ReLU activation function with the Leaky-Relu activation function. Using Leaky-Relu addresses this issue by allowing a small, non-zero output for negative inputs, ensuring that neurons don't die out completely and giving an opportunity for weights to be updated during training [8]. In simpler terms, Leaky-ReLU prevents the dying neuron problem of ReLU by permitting some activity for negative inputs.

The ResNet-18 contains 5 layers and each layer contains convolutional layers, activation functions, and batch normalization. The input image is $1024 \times 1024$ with 3 channels (RGB) and fed to the ResNet-18 network. After the first convolutional the number of channels increases to 64, and in the last layer it would have 512 channels. In layers 2, 3, and 4 we have residual blocks in each of them. Each residual block contains two convolutional layers. After the last

layer, we have a fully connected layer which is the classification layer, is responsible for converting the learned high-dimensional features from the previous layers into a format suitable for making class predictions. The last block is softmax activation function and it is applied to the output of the last fully connected layer. It takes the single scores for each live and spoof class and converts them into a probability distribution. According to Figure 2.b, we use live and generated synthetic samples from the live dataset as an input to the modified ResNet-18 (See Figure 2-a). We then apply the Arcface and center loss on each feature vector and combine them to achieve a better classification.

| SPOOF | |
|---|---|
| API | NUMBER of IMAGES |
| ECOFLEX | 1248 |
| PHOTOPAPER | 1104 |
| PLAYDOH | 1700 |
| WOODGLUE | 272 |
| LIVE (26 subjects) | |
| LIVE | 5886 |

Table 2. Statistics of CLARKSON Dataset.

| **Spoof** | **Number of Images** |
|---|---|
| dragonskin | 1700 |
| ecoflex | 300 |
| gelafix | 100 |
| gelatin | 100 |
| glue | 200 |
| knetosil | 200 |
| latex | 100 |
| modelling-clay | 100 |
| mouldable-glue | 900 |
| paper-printout | 1200 |
| playdoh | 1700 |
| silly-putty | 600 |

Table 3. Statistics of the COLFISPOOF dataset.

# 4. Experimental Setup

## 4.1. Database

In our research, we made use of two publicly available databases: CLARKSON and COLFISPOOF. The CLARKSON database, introduced by Purnapatra et al. [13], consists of a variety of images. It includes 7,500 images of four-finger attacks, over 14,000 manually segmented images of single-fingertip attacks, and 10,000 synthetic fingertip images generated using deepfake techniques. These images were obtained from six different Presentation Attack Instruments (PAI) that cover three levels of difficulty. Furthermore, the CLARKSON database contains a total of 31,702 images of 26 subjects captured from live finger photos. Among these images, 2,150 were collected from the four-finger scenario, and 7,768 were collected from single fingertip scenarios. To assess the effectiveness and performance of each device, we evaluated the six different smartphones: iPhone X, iPhone 7, Samsung Galaxy S9, Google Pixel, Samsung Galaxy S6, and S7. For generating spoofed fingertip images, we used different smartphones, such as synthetic, Ecoflex PAI, Playdoh PAI, Wood Glue PAI, Finger Photo PAI, and Latex PAI.

In contrast, the COLFISPOOF database, introduced by Kolberg et al. [5], exclusively contains spoof images from various categories, including dragonskin, ecoflex, gelafix, gelatin, glue, knetosil, latex, modelling-clay, moduldable-glue, paper-printout, playdoh, and silly-putty. The statistics of the databases are presented in Table 2 and Table 3, showing the details of the CLARKSON and COLFISPOOF datasets, respectively.

| APCER% | | |
|---|---|---|
| API/METHOD | **ResNet-Leaky-Relu** | ResNet-Relu |
| ECOFLEX | **0** | 0 |
| PHOTOPAPER | **9.43** | 11.21 |
| PLAYDOH | **0** | 0 |
| WOODGLUE | **0** | 0 |
| SYNTHETIC | **0.15** | 0.15 |
| BPCER% | | |
| LIVE | **0.12** | 0.35 |

Table 4. Results on CLARKSON dataset (APCER%, BPCER%).

| APCER% | | |
|---|---|---|
| API/METHOD | **ResNet-LReLu** | ResNet-Relu |
| DRAGONSKIN | **0** | 0 |
| ECOFLEX | **0** | 0 |
| GELAFIX | **0** | 0 |
| GELATIN | **0** | 0 |
| GLUE | **0** | 0 |
| KNETOSIL | **0** | 0 |
| LATEX | **0** | 0 |
| MODELLING-CLAY | **0** | 0 |
| MODULABLE-GLUE | **0** | 0 |
| PAPER-PRINTOUT | **0** | 0 |
| PLAYDOH | **0** | 0 |
| SILLY-PUTTY | **0** | 0 |

Table 5. Results on COLFISPOOF dataset (APCER%).

It is important to note that the CLARKSON dataset used in our study contains a smaller number of live and spoofed samples compared to the one reported in the aforementioned reference. The original database also includes synthetic spoofed samples that were not investigated in our study. In order to generate synthetic samples from live fingers, we implement StyleGAN with Adaptive Discriminator

| APCER% | | |
|---|---|---|
| Metrics/METHOD | **ResNet-Leaky-Relu** | ResNet-Relu |
| APCER% | **0.63** | 0.75 |
| BPCER% | **0.12** | 0.35 |
| ACER% | **0.37** | 0.55 |

Table 6. Average of APCER, BPCER, and ACER on both COLFISPOOF and CLARKSON dataset.

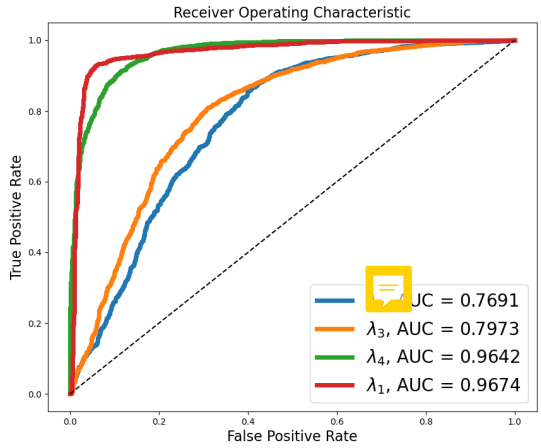


Figure 3. ROC curves for the ResNet architecture trained using live and synthetic samples from StyleGAN, varying the $\lambda$ regularization to find the optimal value.

Augmentation (ADA) [4] and generate 5,000 synthetic samples from 26 live subject from CLARKSON dataset. Also, in order to increase the resolutiuon of images, we apply Enhanced Super Resolution Generative Adversarial Networks (ESRGAN)[18] on both live and synthetic images.

Table 2 and Table 3 shows the detailed statistics of the CLARKSON and COLFISPOOF databases, respectively. Figure 1 shows different spoof types that we used in our study.

### 4.2. Implementation Details

We implemented our combined loss function on ResNet-18 with several modifications, including changing the activation function to Leaky ReLU. The model was trained for 20 epochs using the Adam optimizer with a learning rate of 0.001. Throughout the training process, we assessed the model's performance to determine the optimal value for the $\lambda$ in Equation 1, which balances the two loss functions. Figure 3 illustrates the ROC curve for different $\lambda$ values, and we obtained the best value based on the ROC curve of the validation set. In the Arcface loss function, we set the angle equal to $30^{\circ}$ and added a margin of 0.3 to the angular.

### 4.3. Metrics and Evaluation Protocols

During the validation phase of our study, we computed the True Positive Rate (TPR) and False Positive Rate (FPR) to determine the probability of correctly classifying live finger images as live and spoof finger images as live images. Additionally, we calculated the Bona Fide Presentation Classification Error Rate (BPCER), Attack Presentation Classification Error Rate (APCER), and Average Classification Error Rate (ACER). ACER is defined as the average of APCER and BPCER, and these metrics help evaluate the performance of our model in distinguishing between live and spoofed finger images.

We trained the RestNet model to determine the optimal $\lambda$ value (regularization parameter) used in the joint loss function (details in Section 1, Figure 3). To that end, both live samples and synthetic images generated from StyleGAN were used for binary classification. The goal was to minimize variations within live samples while enhancing the inter-class variation between deepfake and live samples. Finally, the evaluation of the proposed framework was based on unseen spoofed and live samples. In the testing phase, pretrained model was stressed out on testing datasets, which contains both spoof and live dataset (CLARKSON and CLFISPOOf), and the synthetic samples that we generated using the styleGAN-ADA. Subsequently, we computed APCER (for spoofed samples), and BPCER (for live dataset). By plotting the receiver operating characteristic (ROC) curve and calculating the area under the curve (AUC), we have evaluated the performance of our model. It is important to consider that we evaluate the performance of our model on unseen spoof types of attacks (expect synthetic type) and unseen live subjects.

### 4.4. Results

During the training and validation process, we determined the best value for $\lambda$. Based on Figure 3, we selected $\lambda_{[illegible]}$ it demonstrated the best performance during validation. We tested the trained model on both the CLARKSON and COLFISPOOF datasets, which included live and spoof data. As previously mentioned, we trained the model on a combination of live data (from the CLARKSON dataset) and synthetic data generated using StyleGAN-ADA [4] from the live dataset. Next, we evaluated the model's performance on unseen spoof attack types from both the CLARKSON and COLFISPOOF datasets, as well as on unseen live subjects. Tables 5 and 6 present the results showcasing the performance of our model on the respective datasets. To provide a better comparison, we trained the ResNet-18 model with both ReLU and LeakyReLU activation functions, utilizing the combined Arcface and Center loss functions. According to Figure 4, the ResNet-18 model with LeakyReLU activation function and the combined loss achieved the best performance among the other methods.

| Method | (%)APCER | | | | | (%)BPCER |
|---|---|---|---|---|---|---|
| | ECOFLEX | PHOTOPAPER | PLAYDOH | WOODGLUE | SYNTHETIC | LIVE |
| DenseNet-121 [13] | 0 | 88.03 | 0.14 | 0 | 0.13 | 0.18 |
| DenseNet-121 (keras) [13] | 0 | 79.01 | 1.55 | 0.94 | 0.79 | 3.64 |
| NasNetMobile | 0 | 82.15 | 0.71 | 5.96 | 4.12 | 9.04 |
| DenseNet-121 (grayscale) | 0.16 | 98.9 | 1.98 | 11 | 11.58 | 0.18 |
| ResNet-18/Relu (Combined Loss) | 0 | 11.21 | 0 | 0 | 0.15 | 0.35 |
| **Resnet-18/Leaky Relu (Combined Loss)** | **0** | **9.43** | **0** | **0** | **0.15** | **0.12** |
| ResNet-18/Relu (Arcface Loss) | 0 | 45.12 | 0 | 0 | 0.32 | 0.37 |
| ResNet-18/Relu (Center Loss) | 0 | 20.14 | 0 | 0 | 0.26 | 0.21 |

Table 7. Performance of different deep learning architectures across spoofed and non-spoofed samples, measured in terms of BPCER and APCER.

### 4.5. Discussions

Based on the results presented in Table 7, our proposed method, ResNet-18 with Leaky-ReLU and the joint loss function, achieved the best performance compared to other methods in classifying unseen spoof attacks and unseen live datasets. In comparison to DenseNet-121 (keras)[13], our model's error rate in classifying Photopaper improved by 69.58%, Playdoh improved by 1.55%, and for the synthetic samples improved by 0.94%. Additionally, we achieved an APCER of 0% for both Ecoflex and Woodglue samples. Furthermore, our model's performance on the live dataset improved by 3.52%. We also conducted experiments with ResNet-18 using the ReLU activation function and the joint loss function to extract features from images. However, we found that the performance of ResNet with Leaky ReLU was superior. This is because Leaky ReLU addresses the "dying neuron" problem, which occurs when ReLU-activated neurons in a neural network become stuck and stops learning during training [12, 8]. Additionally, we trained ResNet-18 with Leaky ReLU using center loss and Arcface loss independently. According to Table 7, the modified version of ResNet-18 with Arcface loss improved by 42.91% and 0.14% compared to DenseNet-121 in detecting photopaper and playdoh spoof attacks, respectively. However, the results improved even further when applying the Center loss function (refer to Table 7). ResNet-18 with the center loss function exhibited a 67.89% improvement in detecting photopaper compared to DenseNet-121 and a 0.14% improvement in detecting playdoh.

To achieve the best performance, we decided to combine these two loss functions, and as shown in Table 7, our proposed method achieved the best performance in detecting photopaper, which was the most challenging among the other types of spoof attacks. In addition, in our study, we compared our proposed work with the research conducted by Hailin Li et al. [6]. They considered four scenarios in their study, and their ResNet50 architecture was trained as follows: Case-1: Training with photopaper, playdoh, and woodglue, and testing with ecoflex. Case-2: Training with ecoflex, playdoh, and woodglue, and testing with photopaper. Case-3: Training with ecoflex, photopaper, and woodglue, and testing with playdoh. Case-4: Training with ecoflex, photopaper, and playdoh, and testing with woodglue. In our proposed approach, our model was exclusively exposed to synthetic and live samples during training and subsequently tested on all types of unseen spoofed attacks. This strategy allowed us to evaluate the robustness and generalizability of our model against various unseen spoofing scenarios. Based on the findings, in case two where the model is tested on photo-printed attacks, our model demonstrated a remarkable 99% improvement in BPCER at an APCER of 10%. This significant improvement highlights the effectiveness of our proposed model in detecting photo-printed attacks.

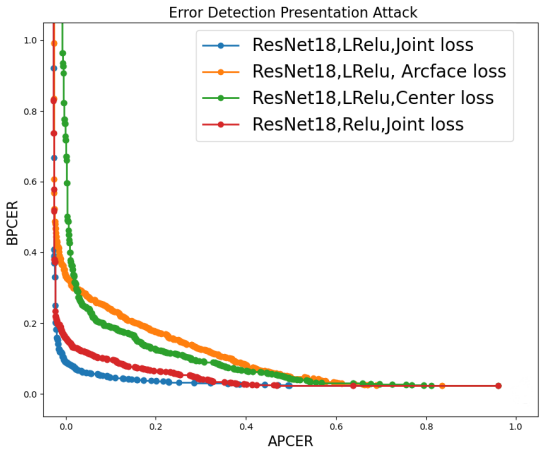


Figure 4. ROC curves illustrate the performance of different loss functions and activation functions using the proposed ResNet architecture.

## 5. Conclusion

The rising popularity of contactless fingerprint biometric systems has led to their potential replacement of conventional touch-based fingerprint recognition systems. However, these systems have some drawbacks, particularly their vulnerability to presentation attacks involving photo-

printed or paper printout spoof samples. Current research in presentation attack detection (PAD) predominantly relies on supervised learning techniques, utilizing both genuine and spoofed samples during training. Nevertheless, these methods often exhibit poor performance against unseen attacks, limiting their scalability. In this paper, we propose a novel approach to address this issue. We introduce a new loss function that combines the Arcface loss to minimize the intra-class variation and the Center Loss to maximize the intra-class variation. By finding the optimal value for a parameter called lambda ($\lambda$), we strike a balance between the two loss functions. Importantly, our deep learning approach is trained solely on genuine images and focuses on detecting a specific type of spoof attack (synthetic). During the testing phase, we evaluate our model using unseen spoofed samples. The proposed scheme demonstrates promising results, with an average BPCER (Bonafide Presentation Classification Error Rate) of 0.12% and an APCER (Attack Presentation Classification Error Rate) of 0.63% for presentation attacks involving various types of spoofed samples.

## 6. Acknowledgements

This project received partial support from the National Science Foundation through Grants No. xxx (blind for review purposes).

## References

[1] J. Deng, J. Guo, N. Xue, and S. Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4690–4699, 2019.

[2] M. Fujio, Y. Kaga, T. Murakami, T. Ohki, and K. Takahashi. Face/fingerphoto spoof detection under noisy conditions by using deep convolutional neural network. In *BIOSIGNALS*, pages 54–62, 2018.

[3] S. A. Grosz, J. J. Engelsma, E. Liu, and A. K. Jain. C2cl: Contact to contactless fingerprint matching. *IEEE Transactions on Information Forensics and Security*, 17:196–210, 2021.

[4] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila. Training generative adversarial networks with limited data. *Advances in neural information processing systems*, 33:12104–12114, 2020.

[5] J. Kolberg, J. Priesnitz, C. Rathgeb, and C. Busch. Colfispoof: A new database for contactless fingerprint presentation attack detection research. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 653–661, 2023.

[6] H. Li and R. Ramachandra. Deep features for contactless fingerprint presentation attack detection: Can they be generalized? *arXiv preprint arXiv:2307.01845*, 2023.

[7] C. Lin and A. Kumar. Matching contactless and contact-based conventional fingerprint images for biometrics identification. *IEEE Transactions on Image Processing*, 27(4):2008–2021, 2018.

[8] A. L. Maas, A. Y. Hannun, A. Y. Ng, et al. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3. Atlanta, GA, 2013.

[9] E. Marasco and A. Vurity. Fingerphoto presentation attack detection: Generalization in smartphones. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 4518–4523. IEEE, 2021.

[10] E. Marasco, A. Vurity, and A. Otham. Deep color spaces for fingerphoto presentation attack detection in mobile devices. In *International Conference on Computer Vision and Image Processing*, pages 351–362. Springer, 2021.

[11] O. Nikisins, A. Mohammadi, A. Anjos, and S. Marcel. On effectiveness of anomaly detection approaches against unseen presentation attacks in face anti-spoofing. In *2018 International Conference on Biometrics (ICB)*, pages 75–81. IEEE, 2018.

[12] L. Pa[illegible]D. Neagu, R. Ma, and F. Campean. Qrelu and m-qrelu: Two novel quantum activation functions to aid medical diagnostics. *arXiv preprint arXiv:2010.08031*, 2020.

[13] S. Purnapatra, C. Miller-Lynch, S. Miner, Y. Liu, K. Bahmani, S. Dey, and S. Schuckers. Presentation attack detection with advanced cnn models for noncontact-based fingerprint systems. pages 1–6, 2023.

[14] A. Ross, S. Dass, and A. Jain. A deformable model for fingerprint matching. *Pattern Recognition*, 38(1):95–103, 2005.

[15] C. Stein, V. Bouatou, and C. Busch. Video-based fingerphoto recognition with anti-spoofing techniques with smartphone cameras. In *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*, pages 1–12. IEEE, 2013.

[16] H. Tan and A. Kumar. Towards more accurate contactless fingerprint minutiae extraction and pose-invariant matching. *IEEE Transactions on Information Forensics and Security*, 15:3924–3937, 2020.

[17] A. Taneja, A. Tayal, A. Malhorta, A. Sankaran, M. Vatsa, and R. Singh. Fingerphoto spoofing in mobile devices: a preliminary study. In *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–7. IEEE, 2016.

[18] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops*, pages 0–0, 2018.

[19] P. Wasnik, R. Ramachandra, K. Raja, and C. Busch. Presentation attack detection for smartphone based fingerphoto recognition using second order local structures. In *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 241–246. IEEE, 2018.

[20] Y. Wen, K. Zhang, Z. Li, and Y. Qiao. A discriminative feature learning approach for deep face recognition. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*, pages 499–515. Springer, 2016.

[21] J. Xu, Z. Li, B. Du, M. Zhang, and J. Liu. Reluplex made more practical: Leaky relu. In *2020 IEEE Symposium on Computers and communications (ISCC)*, pages 1–7. IEEE, 2020.